



**HAL**  
open science

# Interphonologie et représentations orthographiques. Du rôle de l'écrit dans l'enseignement / apprentissage du français oral chez des étudiants japonais

Sylvain Detey

► **To cite this version:**

Sylvain Detey. Interphonologie et représentations orthographiques. Du rôle de l'écrit dans l'enseignement / apprentissage du français oral chez des étudiants japonais. Linguistique. Université Toulouse le Mirail - Toulouse II, 2005. Français. NNT : . tel-00458366

**HAL Id: tel-00458366**

**<https://theses.hal.science/tel-00458366>**

Submitted on 22 Feb 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Université de Toulouse-Le Mirail**

*Département des Sciences du Langage*

**INTERPHONOLOGIE**

**ET**

**REPRÉSENTATIONS ORTHOGRAPHIQUES.**

Du rôle de l'écrit dans l'enseignement/apprentissage du français oral  
chez des étudiants japonais.

Thèse en vue de l'obtention du grade de

**DOCTEUR DE L'UNIVERSITÉ TOULOUSE II**

*Discipline : SCIENCES DU LANGAGE*

**Sylvain DETEY**

Sous la co-direction de Messieurs les Professeurs  
**Jacques DURAND et Jean-Luc NESPOULOUS**

**2005**

Jury de soutenance :

Pr. Jacques Durand (*E.R.S.S., UMR 5610, UTM*) (co-directeur)  
Pr. Jean-Luc Nespoulous (*Laboratoire Jacques-Lordat, EA 1941, UTM*) (co-directeur)  
Pr. Rémy Porquier (*Université de Paris 10*) (rapporteur)  
Pr. Jean-Emile Gombert (*Université de Rennes 2*) (rapporteur)  
Pr. Lorraine Baqué (*Université Autonome de Barcelone*) (rapporteur)  
Pr. Michel Billières (*UTM*) (examineur)  
Dr. Laurence Labrune (*Université de Bordeaux 3*) (MC, HDR) (examineur)

Date et lieu de soutenance : 10 octobre 2005 à Toulouse (UTM).

Mention: Très honorable avec les félicitations du jury à l'unanimité.

## **Table des Matières**

<b>NOTES LIMINAIRES .....</b>	<b>- 14 -</b>
<b>INTRODUCTION GENERALE .....</b>	<b>- 16 -</b>
<b>PARTIE DIDACTIQUE I : INTRODUCTION.....</b>	<b>- 28 -</b>
TABLE DES MATIERES PARTIELLE .....	- 29 -
I. DIDACTIQUE DES LANGUES, PSYCHOLINGUISTIQUE ET PHONOLOGIE : UN TRIPTYQUE DISCIPLINAIRE POUR LA FORMATION A L'ENSEIGNEMENT DE LA PAROLE EN LANGUE ETRANGERE.....	- 30 -
1.1. <i>La didactique des langues et des cultures (DLC)</i> .....	- 30 -
1.2. <i>La phonologie</i> .....	- 41 -
1.3. <i>La psycholinguistique</i> .....	- 48 -
1.4. <i>Intégration et transdisciplinarité</i> .....	- 49 -
II. MULTIMEDIA ET MULTIMODALITE .....	- 51 -
II.1. <i>Domaine et historique</i> .....	- 51 -
II.2. <i>Multimodalité/ multicanalité : un atout pour l'apprentissage ?</i> .....	- 55 -
III. ENSEIGNEMENT DE LA PRONONCIATION .....	- 58 -
III.1. <i>Situation actuelle de l'enseignement de la prononciation du FLE au Japon</i> .....	- 58 -
III.1.1. Raisons d'ordre institutionnel .....	- 59 -
III.1.2. Raisons d'ordre socioculturel.....	- 60 -
III.1.3. Raisons d'ordre (psycho-)linguistique .....	- 62 -
III.1.4. Raisons d'ordre didactique.....	- 63 -
III.2. <i>Qu'est-ce que la « prononciation » ?</i> .....	- 64 -
III.2.1. Définitions .....	- 64 -
III.2.2. Interfaces : langue et communication .....	- 66 -
III.3. <i>Observations, pratiques et enquêtes au Japon</i> .....	- 67 -
III.3.1. Observations.....	- 67 -
III.3.1.1. De la perception auditive à la production écrite : erreurs de « phonétique » en production sous dictée.....	- 70 -
III.3.1.2. Perception auditive et production orale .....	- 74 -
III.3.1.3. Lecture et production orale.....	- 77 -
III.3.2. Pratiques .....	- 79 -
III.3.2.1. Matériel pédagogique .....	- 79 -
III.3.2.2. Formation des enseignants : de la phonétique articulatoire à la méthode verbo-tonale d'intégration phonétique .....	- 82 -
III.3.3. Enquêtes : représentations et motivations concernant la prononciation du français au Japon .....	- 92 -
III.3.3.1. Enquête 1 : Apprenants en atelier de correction phonétique en école privée (n=31) .....	- 93 -
III.3.3.2. Enquête 2 : Enseignants français (n = 32) et japonais (n = 21) de français (total=53).....	- 101 -
III.3.3.3. Enquête 3 : Apprenants de français en université (n = 273) .....	- 104 -
IV. TRANSITION.....	- 105 -



**PARTIE LINGUISTIQUE : SYSTEMES PHONETICO-PHONOLOGIQUES ET PHONOGRAPHEMIQUES FRANÇAIS ET JAPONAIS..... - 107 -**

TABLE DES MATIERES PARTIELLE .....	- 108 -
<b>I SYSTEMES PHONETICO-PHONOLOGIQUES .....</b>	<b>- 110 -</b>
<b><i>I.1. Le système français.....</i></b>	<b><i>- 112 -</i></b>
I.1.1. Norme francophone et projet « Phonologie du Français Contemporain ».....	- 113 -
I.1.2. Description classique du français standardisé .....	- 122 -
I.1.2.1 Aspects segmentaux.....	- 124 -
I.1.2.2. Aspects suprasegmentaux.....	- 128 -
I.1.2.2.1. Aspects phonotactiques .....	- 128 -
I.1.2.2.1.1. Distribution et opposition .....	- 128 -
I.1.2.2.1.2 Phénomènes de contacts .....	- 132 -
I.1.2.2.2. Aspects syllabiques : des représentations .....	- 133 -
I.1.2.2.2.1. Extramétricalité .....	- 137 -
I.1.2.2.2.2. Ambisyllabité.....	- 137 -
I.1.2.2.2.3. Sonorité.....	- 138 -
I.1.2.2.2.4. Structure interne de la syllabe .....	- 141 -
I.1.2.2.2.5. Syllabes primitives ou émergentes ?.....	- 142 -
I.1.2.2.3. Aspects syllabiques : des contraintes (OT et TCRS).....	- 144 -
I.1.2.2.3.1. La Théorie de l'Optimalité .....	- 146 -
I.1.2.2.3.2. La Théorie des Contraintes et des Stratégies de Réparation.....	- 151 -
I.1.2.2.3.3. OT versus TCRS .....	- 157 -
I.1.2.2.4. Aspects syllabiques : syllabes complexes .....	- 165 -
I.1.2.2.4.1. Syllabes complexes et groupes consonantiques .....	- 166 -
I.1.2.2.4.2. Syllabes complexes et schwa.....	- 174 -
I.1.2.2.4.3. Syllabes complexes, enchaînements et liaisons.....	- 181 -
I.1.2.2.5. Aspects prosodiques.....	- 185 -
I.1.2.2.5.1. L'intonation .....	- 186 -
I.1.2.2.5.2. L'accentuation .....	- 187 -
I.1.2.2.5.3. Le rythme.....	- 189 -
<b><i>I.2. Le système japonais.....</i></b>	<b><i>- 192 -</i></b>
I.2.1. Aspects segmentaux .....	- 196 -
I.2.1.1. Consonnes .....	- 196 -
I.2.1.2. Voyelles.....	- 199 -
I.2.2. Aspects suprasegmentaux .....	- 201 -
I.2.2.1. Aspects phonotactiques .....	- 201 -
I.2.2.2. Aspects syllabiques : syllabes et mores.....	- 202 -
I.2.2.2.1. More ou syllabe ?.....	- 202 -
I.2.2.2.2. Mores « spéciales ».....	- 205 -
I.2.2.3. Aspects syllabiques : emprunts étrangers.....	- 208 -
I.2.2.4. Aspects syllabiques : syllabes complexes et groupes consonantiques .....	- 215 -
I.2.2.5. Aspects syllabiques : syllabes complexes et voyelles dévoisées .....	- 216 -
<b>II SYSTEMES PHONOGRAPHEMIQUES .....</b>	<b>- 222 -</b>
<b><i>II.1. Système d'écriture, script et orthographe .....</i></b>	<b><i>- 222 -</i></b>
<b><i>II.2. Typologie des systèmes d'écriture et « profondeur orthographique » .....</i></b>	<b><i>- 229 -</i></b>
<b><i>II.3. Le système d'écriture français .....</i></b>	<b><i>- 232 -</i></b>
<b><i>II.4. Le système d'écriture japonais.....</i></b>	<b><i>- 238 -</i></b>
II.4.1. Un système pluriel.....	- 239 -
II.4.2. Les kana .....	- 242 -
II.4.3. Les romaji .....	- 248 -
II.4.4. Transcription des emprunts, katakana et alphabet .....	- 252 -
<b>III. PERSPECTIVE CONTRASTIVE .....</b>	<b>- 262 -</b>
<b><i>III.1. Contrastes phonético-phonologiques.....</i></b>	<b><i>- 263 -</i></b>
III.1.1. Aspects segmentaux.....	- 263 -

III.1.2. Aspects syllabiques.....	- 270 -
III.2. Graphie.....	- 272 -
<b>PARTIE PSYCHOLINGUISTIQUE : TRAITEMENTS ET RESSOURCES CHEZ L'APPRENANT DE LANGUE ETRANGERE.....</b>	<b>- 275 -</b>
TABLE DES MATIERES PARTIELLE.....	- 279 -
I. PERCEPTION ET APPRENTISSAGE DE LA PAROLE EN LANGUE ETRANGERE : DESCRIPTION GENERALE.....	- 282 -
I.1. <i>Système perceptivo-attentionnel et perception de la parole</i> .....	- 283 -
I.1.1. Perception de la parole: considérations générales.....	- 286 -
I.1.1.1. Processus bas-haut et haut-bas.....	- 287 -
I.1.1.2. Théories de la perception de la parole.....	- 292 -
I.1.1.3. De la perception à la compréhension.....	- 296 -
I.1.2. Perception de la parole et reconnaissance linguistique.....	- 300 -
I.1.2.1. Unités perceptives : niveau sublexical.....	- 301 -
I.1.2.2. Reconnaissance linguistique.....	- 310 -
I.1.2.2.1. Appariement, catégorisation et reconnaissance : considérations générales.....	- 310 -
I.1.2.2.2. Modèles de reconnaissance.....	- 318 -
I.1.2.2.2.1. Le modèle COHORTE.....	- 319 -
I.1.2.2.2.2. Le modèle TRACE.....	- 320 -
I.2. <i>Système mnésique et connaissances</i> .....	- 322 -
I.2.1. Système mnésique.....	- 323 -
I.2.1.1. Organisation de la mémoire : perspectives.....	- 324 -
I.2.1.1.1. Concepts généraux.....	- 324 -
I.2.1.1.2. Modélisation.....	- 327 -
I.2.1.1.2.1. Modèle Modal.....	- 327 -
I.2.1.1.2.2. Les Niveaux de Traitement.....	- 328 -
I.2.1.1.2.3. Composantes de la mémoire.....	- 330 -
I.2.1.1.2.4. Mémoire et apprentissage.....	- 332 -
I.2.1.1.2.5. Perspectives connexionnistes.....	- 333 -
I.2.1.1.3. Format et Temporalité.....	- 336 -
I.2.1.1.3.1. Format et Modalité.....	- 336 -
I.2.1.1.3.2. Temporalité.....	- 338 -
I.2.1.1.3.3. Parenthèse didactique.....	- 339 -
I.2.1.2. Mémoire de travail et boucle phonologique.....	- 341 -
I.2.1.2.1. Nature de la boucle phonologique.....	- 344 -
I.2.1.2.2. Fonction de la boucle phonologique.....	- 348 -
I.2.1.2.2.1. Acquisition de formes phonologiques nouvelles.....	- 348 -
I.2.1.2.2.2. Lecture.....	- 353 -
I.2.1.2.3. Parenthèse didactique.....	- 355 -
I.2.2. Connaissances linguistiques.....	- 357 -
I.2.2.1. Savoirs déclaratifs et savoir-faire procéduraux.....	- 358 -
I.2.2.2. Le lexique mental : primitives et émergentes.....	- 363 -
I.2.2.2.1. Primitives phonologiques : perspective symbolique.....	- 364 -
I.2.2.2.1.1. Production.....	- 365 -
I.2.2.2.1.2. Perception.....	- 370 -
I.2.2.2.2. Emergentisme et connexionnisme : perspective subsymbolique.....	- 371 -
I.2.2.2.3. Parenthèse didactique.....	- 375 -
II. PERCEPTION DE LA PAROLE ET MULTIMODALITE.....	- 380 -
II.1. <i>Influences intermodales</i> .....	- 382 -
II.2. <i>Théories et Modèles</i> .....	- 388 -
II.2.1. Le modèle flou (Massaro).....	- 388 -
II.2.2. La théorie de la perception pour le contrôle de l'action (Schwartz et al.).....	- 396 -
II. 3. <i>Perception, multimodalité et apprentissage de la parole en langue étrangère</i> .....	- 400 -

II.3.1. Rôle des indices visuels dans la perception auditive en langue étrangère.....	400 -
II.3.1.1. Influence positive des indices visuels .....	400 -
II.3.1.2. Influence moindre des indices visuels.....	402 -
II.3.2. Rôle de la bimodalité dans l'apprentissage de la parole en langue étrangère .....	404 -
II.3.2.1. Indices visuels oro-faciaux et catégories phonémiques .....	404 -
II.3.2.1.1. Indices visuels externes .....	405 -
II.3.2.1.2. Indices visuels internes.....	407 -
II.3.2.2. Indices visuels orthographiques et mémorisation.....	409 -
II.3.3. Conclusion.....	410 -
<b>III. PERCEPTION DE LA PAROLE ET ORTHOGRAPHE .....</b>	<b>412 -</b>
<b>III.1. Introduction .....</b>	<b>413 -</b>
<b>III.2. Perception de l'oral et de l'écrit : principes invariants.....</b>	<b>414 -</b>
III.2.1. Primauté de l'oral sur l'écrit .....	414 -
III.2.2. Traitement de l'information .....	416 -
III.2.3. Contexte et effets haut-bas .....	420 -
<b>III.3. Perception de l'écrit : spécificités et variation selon les langues.....</b>	<b>421 -</b>
III.3.1. L'hypothèse de la profondeur orthographique.....	423 -
III.3.1.1. Orthographe profondes vs. superficielles .....	423 -
III.3.1.2. Procédure d'assemblage (voie phonologique) vs. procédure d'adressage (voie lexicale) ....	424 -
III.3.1.3. Approche universelle : implication des deux voies dans les deux types d'orthographe .....	426 -
III.3.1.4. Facteurs déterminant l'implication de chaque voie .....	428 -
III.3.1.5. Conclusion .....	429 -
III.3.2. Taille des unités perceptives : lettre, graphème, corps, syllabe et mots .....	429 -
III.3.2.1. Lettre vs. Graphème.....	430 -
III.3.2.2. Graphème vs. Corps et Mots .....	431 -
III.3.2.3. Fréquence et taille des unités.....	433 -
III.3.3. Le cas du français .....	437 -
III.3.3.1. Le graphème .....	437 -
III.3.3.2. La syllabe.....	438 -
III.3.3.3. Fréquence, unités de différentes tailles et autres facteurs en lecture à voix haute.....	439 -
III.3.3.4. Conclusion .....	443 -
III.3.4. Le cas du japonais.....	443 -
III.3.4.1. Introduction .....	443 -
III.3.4.2. Les Kana .....	445 -
III.3.4.2.1. Distinction katakana/hiragana et familiarité scripturale .....	445 -
III.3.4.2.2. Kana et unités phonologiques.....	447 -
III.3.4.2.3. Fréquence et voie d'adressage pour les kana.....	447 -
III.3.4.2.4. Hémisphère gauche vs. hémisphère droit pour les kana .....	448 -
III.3.4.2.5. Conclusion .....	449 -
III.3.4.3. Les Romaji.....	449 -
III.3.4.4. Les Kanji.....	450 -
III.3.4.4.1. Implication de la voie phonologique pour les kanji.....	450 -
III.3.4.4.2. Hémisphère gauche vs. hémisphère droit pour les kanji.....	451 -
III.3.4.5. Conclusion .....	453 -
<b>III.4. Reconnaissance lexicale et lecture à voix haute : modèles.....</b>	<b>455 -</b>
III.4.1. Le Modèle Cohorte .....	455 -
III.4.2. Le Modèle à deux voies en cascades ( <i>Dual-Route Cascaded Model, DRC</i> ).....	456 -
III.4.3. Modèles Interactifs : connexionnisme et intégration.....	460 -
III.4.3.1. Les Modèles d'Activation Interactive.....	460 -
III.4.3.2. Le Modèle Flou de la Perception .....	461 -
III.4.3.3. Morphologie et Prosodie en reconnaissance visuelle .....	463 -
III.4.4. Conclusion.....	465 -
<b>III.5. Influence de l'orthographe en perception de la parole.....</b>	<b>466 -</b>
III.5.1. Interactions entre codes orthographiques et phonologiques .....	466 -
III.5.1.1. L'hypothèse de l'autonomie orthographique : les données neurolinguistiques .....	466 -
III.5.1.2. Interactivité orthographe/phonologie : au niveau sublexical .....	469 -
III.5.1.3. Interactivité orthographe/phonologie : au niveau lexical .....	471 -
III.5.1.4. Conclusion : de la perception à la production .....	473 -
III.5.2. La conscience phonologique .....	474 -

III.5.2.1. Conscience phonologique précoce vs. Conscience phonémique .....	475 -
III.5.2.2. Apprentissage du code écrit et développement de la conscience phonologique.....	476 -
III.5.2.2.1. La conscience phonémique.....	476 -
III.5.2.2.2. La conscience phonologique non-phonémique .....	478 -
III.5.2.2.3. Conclusion .....	479 -
<b>PROBLEMATIQUES : APPRENTISSAGE PHONOLOGIQUE MULTIMODAL .....</b>	<b>483 -</b>
TABLE DES MATIERES PARTIELLE .....	484 -
I. PROBLEMATIQUE D' APPRENTISSAGE PHONOLOGIQUE : L' INTERPHONOLOGIE .....	486 -
I. 1. Introduction : apprentissage phonétique ou phonologique ? .....	486 -
I. 2. Interphonologie : perspectives structuraliste et fonctionnaliste.....	494 -
I.2.1. Le crible phonologique.....	494 -
I.2.2. Facteurs sociolinguistiques .....	495 -
I.2.3. Du crible au transfert.....	498 -
I. 3. Interphonologie : perspectives générativistes .....	499 -
I.3.1. Perspectives représentationnelles : Marque, Transfert et Grammaire Universelle .....	499 -
I.3.1.1. La Marque .....	500 -
I.3.1.2. Le Transfert et la Grammaire Universelle.....	503 -
I.3.1.2.1. Le Transfert L1/L2.....	503 -
I.3.1.2.2. L'Accès à la Grammaire Universelle.....	505 -
I.3.2. Perspectives non-représentationnelles : optimalité et contraintes .....	507 -
I.3.2.1. Interphonologie et Théorie de l'Optimalité : une nouvelle approche du transfert.....	507 -
I.3.2.2. Exemple 1 : épenthèses et émergence du non-marqué chez des apprenants brésiliens d'anglais .....	509 -
I.3.2.3. Exemple 2 : épenthèses et variation chez des apprenants chinois d'anglais.....	512 -
I.3.2.4. Exemple 3 : reclassement de contraintes et émergence du non-marqué chez des apprenants zoulous d'anglais .....	513 -
I.3.2.5. Exemple 4 : coda et émergence du non-marqué chez des apprenants italiens d'anglais .....	515 -
I.3.2.6. Exemple 5 : substitutions segmentales chez des apprenants japonais d'anglais .....	516 -
I.3.2.7. Exemple 6 : groupes consonantiques et voyelles épenthétiques en interphonologie .....	520 -
I.3.3. Synthèse .....	525 -
I.3.3.1. Marque et Transfert dans la Théorie de l'Optimalité .....	525 -
I.3.3.2. Accès à UG, Transfert et Perception dans la Théorie de l'Optimalité .....	528 -
I.4. Interphonologie : perspectives exemplaristes.....	530 -
I.4.1. Des épisodes phonétiques aux catégories phonologiques.....	531 -
I.4.1.1. Encodage épisodique acoustique de la parole.....	531 -
I.4.1.2. Variation allophonique et acquisition des catégories phonémiques .....	535 -
I.4.2. Mémoire phonologique et émergentisme.....	538 -
II. INVESTIGATIONS PHONOLOGIQUES : L' INTERPHONOLOGIE DES APPRENANTS JAPONAIS DE FLE .-	540 -
II.1. Aspects segmentaux.....	541 -
II.1.1. Axe syntagmatique : la position du segment cible dans le mot phonologique influence-t-elle le taux de confusion segmentale ? .....	543 -
II.1.2. Axe paradigmatique : la nature du segment cible dans le mot phonologique influence-t-elle le taux de confusion segmentale ? .....	545 -
II.1.2.1. Segments consonantiques .....	545 -
II.1.2.2. Segment vocalique .....	547 -
II.1.2.3. Confusion phonétique ou phonologique ? .....	550 -
II.2. Aspects syllabiques.....	553 -
II.2.1. Axe syntagmatique : la position du groupe cible dans le mot phonologique influence-t-elle le taux de confusion syllabique ?.....	556 -
II.2.2. Axe paradigmatique : la nature du groupe cible dans le mot phonologique influence-t-elle le taux de confusion syllabique ?.....	557 -
II.2.3. Axe paradigmatique : la structure de surface du groupe cible dans le mot phonologique influence-t-elle le taux de confusion syllabique ? .....	568 -
II.2.4. Résumé.....	569 -

II.2.5. Aspects phonographémiques .....	571 -
<b>III. INVESTIGATION PSYCHOLINGUISTIQUE : ORTHOGRAPHE ET INTERPHONOLOGIE CHEZ LES APPRENANTS JAPONAIS DE FLE .....</b>	<b>576 -</b>
<i>III.1. Introduction : variables générales, tâches d'apprentissage et niveaux d'analyse.....</i>	<i>576 -</i>
<i>III.2. Nature des tâches cognitives.....</i>	<i>584 -</i>
III.2.1. Tâche de Lecture – unimodalité visuelle.....	584 -
III.2.2. Tâche de Répétition – unimodalité auditive.....	594 -
III.2.3. Tâche de Lecture/Répétition – bimodalité audiovisuelle .....	601 -
III.2.4. Tâche de Dictée Orthographique .....	604 -
<i>III.3. Influence de l'orthographe : niveau segmental et identification .....</i>	<i>605 -</i>
III.3.1. Traitement psycholinguistique « erroné » des unités phonético-phonologiques et orthographiques.....	606 -
III.3.2. Variation et alternance inter-tâches et intra-tâche .....	612 -
III.3.3. Schémas d'activation des unités phonologiques et orthographiques.....	616 -
III.3.4. Protocole quasi-expérimental.....	620 -
III.3.5. Substitutions inattendues et stratégies d'apprentissage .....	622 -
<i>III.4. Influence de l'orthographe : niveau syllabique et segmentation .....</i>	<i>624 -</i>
III.4.1. L'orthographe comme outil de segmentation de la parole.....	624 -
III.4.2. Le cas du japonais : contraintes phonologiques et/ou orthographiques ? .....	626 -
III.4.3. Protocole quasi-expérimental.....	634 -
<i>III.5. Résumé.....</i>	<i>636 -</i>
<b>PARTIE EXPERIMENTALE .....</b>	<b>639 -</b>
TABLE DES MATIERES PARTIELLE .....	641 -
<b>PARTIE EXPERIMENTALE I : PRE-TESTS.....</b>	<b>645 -</b>
I. INTRODUCTION .....	645 -
II. METHODOLOGIE.....	651 -
III. RESULTATS ET INTERPRETATION .....	666 -
IV. RESUME GENERAL ET DISCUSSION .....	679 -
V. ASPECTS METHODOLOGIQUES ET PERSPECTIVES .....	685 -
<b>PARTIE EXPERIMENTALE II : TESTS - PROTOCOLE.....</b>	<b>693 -</b>
TABLE DES MATIERES PARTIELLE .....	694 -
I. METHODOLOGIE .....	695 -
II. DESCRIPTIF GENERAL DE LA POPULATION .....	696 -
III. DESCRIPTIF GENERAL DES PROTOCOLES .....	704 -
IV. PREMIERE PARTIE : AXE SEGMENTAL (GROUPE 1) .....	710 -
IV.1. Test A (Identification) .....	710 -
IV.2. Test F (Identification et Discrimination).....	716 -
V. DEUXIEME PARTIE : AXE SYLLABIQUE (GROUPE 1 ET 2).....	721 -
V.1. Test C1 (Groupe 1 – Unimodalité Auditive).....	721 -
V.2. Test C2 (Groupe 2 – Unimodalité Auditive).....	728 -
V.3. Test D (Groupe 2 – Unimodalité Visuelle).....	732 -
V.4. Test E (Groupe 2 – Bimodalité Audio-Visuelle).....	735 -
V.5. Méta-test C2-D-E (Groupe 2 – Effet de la Modalité).....	743 -
VI. VARIABLES ET CONSIDERATIONS SUPPLEMENTAIRES .....	744 -

<b>PARTIE EXPERIMENTALE III : TESTS - RESULTATS</b> .....	<b>- 749 -</b>
TABLE DES MATIERES PARTIELLE .....	- 750 -
I. RESULTATS DES TESTS .....	- 753 -
<i>I.1. Considérations statistiques</i> .....	- 753 -
<i>I.2. Axe segmental</i> .....	- 758 -
TACHE A .....	- 758 -
A) Résultats descriptifs .....	- 758 -
A) 1) Axe syntagmatique .....	- 759 -
A) 2) Axe paradigmatique .....	- 762 -
B) Résultats inférentiels : analyses de variance (Test F) .....	- 767 -
C) Résultats complémentaires : choix de graphème existant/absent en L1 (41 sujets) .....	- 768 -
D) Résumé .....	- 771 -
TACHE F .....	- 773 -
A) Résultats descriptifs .....	- 773 -
A) 1) Confusion phonographémique (premier stimulus) .....	- 773 -
A) 1) 1. Axe syntagmatique .....	- 774 -
A) 1) 2. Axe paradigmatique .....	- 776 -
A) 1) 3. Variable liée à la tâche : nature du couple (identique/différent) .....	- 780 -
A) 2) Confusion phonologique (entre les deux stimuli auditifs) .....	- 781 -
A) 2) 1. Axe syntagmatique .....	- 782 -
A) 2) 2. Axe paradigmatique .....	- 784 -
A) 2) 3. Variables liées à la tâche : nature du couple (identique/différent) .....	- 788 -
B) Résultats inférentiels : analyses de variance (Test F) .....	- 789 -
D) Résumé .....	- 790 -
COMPARAISON TACHES A ET F.....	- 792 -
A) Résultats globaux .....	- 792 -
B) Position .....	- 792 -
C) Nature.....	- 794 -
D) Résultats inférentiels.....	- 797 -
E) Résumé.....	- 798 -
<i>I.3. Axe syllabique</i> .....	- 799 -
TACHE C1 .....	- 799 -
A) Résultats descriptifs .....	- 799 -
A) 1) Axe syntagmatique .....	- 800 -
A) 2) Axe paradigmatique.....	- 805 -
B) Résultats inférentiels : analyse de variance (Test F) : .....	- 816 -
C) Résumé.....	- 819 -
C) 1) Résumé détaillé .....	- 819 -
C) 2) Résumé essentiel .....	- 826 -
TACHE C2.....	- 827 -
A) Résultats descriptifs .....	- 827 -
A) 1) Axe syntagmatique .....	- 828 -
A) 2) Axe paradigmatique.....	- 831 -
B) Résultats inférentiels : analyse de variance (Test F).....	- 851 -
C) Résumé.....	- 854 -
C) 1) Résumé Détaillé .....	- 854 -
C) 2) Résumé Essentiel .....	- 862 -

COMPARAISON TACHES C1 ET C2 .....	- 863 -
TACHE D .....	- 871 -
A) Résultats descriptifs .....	- 871 -
A) 1) Axe syntagmatique .....	- 871 -
A) 2) Axe paradigmatic .....	- 873 -
A) 3) Variation entre sujets .....	- 879 -
B) Résumé Essentiel .....	- 880 -
COMPARAISON TACHES C2 ET D .....	- 881 -
TACHE E .....	- 890 -
CAS 1 - STIMULI AUDITIFS ET VISUELS CONVERGENTS .....	- 890 -
A) Résultats descriptifs .....	- 890 -
A) 1) Axe syntagmatique .....	- 891 -
A) 2) Axe paradigmatic .....	- 892 -
B) Résultats inférentiels .....	- 903 -
C) Résumé Essentiel .....	- 903 -
CAS 2 - STIMULI AUDITIFS ET VISUELS DIVERGENTS .....	- 905 -
A) Modalité dominante en tâche de comptage syllabique .....	- 905 -
B) Perception de la divergence intermodale .....	- 909 -
1) Résultats Globaux .....	- 909 -
2) Influence de la position .....	- 909 -
3) Influence de la nature .....	- 910 -
4) Influence de la structure/longueur .....	- 918 -
Comparaison des performances : Comptage syllabique vs. Perception de la divergence inter-stimuli .....	- 921 -
COMPARAISON DES TACHES C2 (AUDITIVE), D (VISUELLE) ET E (AUDIOVISUELLE) .....	- 924 -
1. Résultats globaux .....	- 924 -
2. Position .....	- 924 -
3. Famille Générale .....	- 926 -
4. Famille .....	- 927 -
5. Sous-Famille .....	- 928 -
6. Nature .....	- 929 -
7. Type de Sonorité .....	- 930 -
8. Ecart de Sonorité .....	- 931 -
<b>PARTIE EXPERIMENTALE IV : TESTS – INTERPRETATION .....</b>	<b>- 935 -</b>
TABLE DES MATIERES PARTIELLE .....	- 938 -
I. TACHE A : APPARIEMENT SEGMENTAL PHONEME - GRAPHEME (L/R ; B/V) .....	- 940 -
I.1. <i>Introduction</i> .....	- 940 -
I.1.1. Mode de traitement .....	- 940 -
I.1.2. Interphonologie : catégories phonétiques, phonologiques et graphémiques .....	- 943 -
I.2. <i>Interprétation linéaire des résultats</i> .....	- 945 -
I.2.1. La position .....	- 945 -
I.2.2. La nature de la consonne .....	- 947 -
I.2.3. La nature de la voyelle adjacente .....	- 949 -
I.2.4. Choix de graphèmes .....	- 953 -

II. TACHE F : DISCRIMINATION PHONOLOGIQUE ET APPARIEMENT SEGMENTAL PHONEME - GRAPHEME (L/R ; B/V).....	- 954 -
II.1. Introduction.....	- 954 -
II.2. Interprétation linéaire des résultats phonographémiques (premier stimulus).....	- 955 -
II.2.1. La position.....	- 955 -
II.2.2. La nature de la consonne.....	- 956 -
II.2.3. La nature de la voyelle adjacente.....	- 958 -
II.2.4. La nature du couple de stimuli (identiques/différents).....	- 960 -
II.3. Interprétation linéaire des résultats phonologiques.....	- 960 -
II.3.1. La position.....	- 960 -
II.3.2. La nature des consonnes.....	- 960 -
II.3.3. La nature de la voyelle adjacente.....	- 961 -
II.3.4. La nature du couple de stimuli (identiques/différents).....	- 962 -
III. TACHES C1 ET C2 : COMPTAGE SYLLABIQUE EN UNIMODALITE AUDITIVE.....	- 963 -
III.1. Tâche C1 - Epenthèses.....	- 964 -
III.1.1. Position.....	- 964 -
III.1.2. Nature.....	- 966 -
III.1.3. Sonorité.....	- 969 -
III.2. Tâche C1 – Effacements.....	- 969 -
III.2.1. Position.....	- 971 -
III.2.2. Nature.....	- 971 -
III.2.3. Sonorité.....	- 972 -
III.3. Tâche C1 – Comparaison Epenthèses/Effacements.....	- 973 -
III.4. Tâche C2 - Epenthèses.....	- 974 -
III.4.1. Position.....	- 974 -
III.4.2. Nature.....	- 974 -
III.4.3. Sonorité.....	- 978 -
III.5. Tâche C2 - Effacements.....	- 979 -
III.5.1. Position.....	- 979 -
III.5.2. Nature.....	- 979 -
III.6. Tâche C2 – Comparaison Epenthèses/Effacements.....	- 980 -
IV. TACHE D : COMPTAGE SYLLABIQUE EN UNIMODALITE ORTHOGRAPHIQUE.....	- 980 -
IV.1. Position.....	- 982 -
IV.2. Nature.....	- 983 -
V. TACHE E : COMPTAGE SYLLABIQUE EN BIMODALITE AUDIO-ORTHOGRAPHIQUE.....	- 983 -
V.1. Stimuli auditifs et orthographiques convergents.....	- 987 -
V.1.1. Position.....	- 987 -
V.1.2. Nature.....	- 988 -
V.1.3. Sonorité.....	- 989 -
V.2. Stimuli auditifs et orthographiques divergents.....	- 989 -
V.2.1. Modalité dominante en tâche de comptage syllabique.....	- 989 -
V.2.2. Discrimination intermodale de stimuli divergents.....	- 991 -
V.2.2.1. Position.....	- 991 -
V.2.2.2. Nature.....	- 992 -
V.2.2.3. Sonorité.....	- 994 -
V.2.2.4. Longueur/Structure des stimuli divergents.....	- 994 -
V.3. Comptage syllabique vs. Discrimination intermodale.....	- 995 -
VI. COMPARAISON TACHES C2, D ET E : UNIMODALITE AUDITIVE VS. UNIMODALITE ORTHOGRAPHIQUE VS. BIMODALITE AUDIO-ORTHOGRAPHIQUE.....	- 995 -
VII. SYNTHESE ET PERSPECTIVES.....	- 1003 -
VII.1. Synthèse.....	- 1004 -
VII.2. Implications méthodologiques.....	- 1011 -
VII.3. Perspectives.....	- 1013 -



VII.3.1. Phonétique-Phonologie.....	- 1013 -
VII.3.2. Psycholinguistique .....	- 1016 -
VII.3.3. Didactique et Apprentissage.....	- 1019 -

**PARTIE DIDACTIQUE II : IMPLICATIONS POUR L'ENSEIGNEMENT DE L'ORAL - 1021 -**

TABLE DES MATIERES PARTIELLE .....	- 1027 -
------------------------------------	----------

**I. IMPLICATIONS SPECIFIQUES..... - 1029 -**

*I.1. Axe segmental : tâches A (identification) et F (discrimination)..... - 1032 -*

I.1.1. Tâche A (identification)..... - 1034 -

I.1.2. Tâche F (discrimination)..... - 1035 -

I.1.3. Tâches A et F..... - 1038 -

*I.2 Axe syllabique : tâches C1, C2, D et E (comptage)..... - 1038 -*

**II. IMPLICATIONS GENERALES..... - 1041 -**

*II.1. Prononciation et modalité visuelle..... - 1042 -*

II.1.1. Fonctions de l'écrit pour l'apprentissage de l'oral en classe de langue : fixation et segmentation de la parole..... - 1044 -

II.1.2. Caractéristiques phonético-phonologiques et input visuel..... - 1049 -

II.1.2.1. Préliminaires..... - 1049 -

II.1.2.1.1. Systèmes d'écriture en L1 et homographes hétérophones L1/L2..... - 1051 -

II.1.2.1.2. Perception et mémorisation de l'oral..... - 1057 -

II.1.2.1.3. Correction phonétique..... - 1058 -

II.1.2.2. Information oro-faciale..... - 1060 -

II.1.2.3. Information gestuelle non-faciale..... - 1063 -

II.1.2.3.1. Axe segmental..... - 1064 -

II.1.2.3.2. Axe rythmique/syllabique..... - 1067 -

II.1.2.3.3. Axe accentuel/intonatif..... - 1068 -

II.1.2.3.4. Compréhension et mémorisation..... - 1068 -

II.1.2.4. Information picturale..... - 1070 -

II.1.2.5. Information graphique..... - 1080 -

II.1.2.5.1. Caractéristiques prosodiques..... - 1080 -

II.1.2.5.2. Segmentation de la parole..... - 1082 -

II.1.2.5.3. Code graphique et modifications orthographiques..... - 1085 -

II.1.2.5.3.1. Elaboration d'un code graphique pour l'apprentissage de l'oral..... - 1088 -

II.1.2.5.3.2. Le système Alfonic..... - 1089 -

II.1.2.5.3.3. Conscience phonémique et correspondances phonographémiques..... - 1091 -

II.1.2.5.3.4. L'Alphabet Phonétique International (API)..... - 1093 -

II.1.3. Tâches et utilisation du support visuo-graphique..... - 1096 -

II.1.3.1. Lecture vs. Transcription..... - 1097 -

II.1.3.2. Utilisation passive vs. Utilisation active..... - 1099 -

*II.2. Enseignement explicite ou implicite ?..... - 1102 -*

II.2.1. Apprentissage implicite ou explicite ?..... - 1105 -

II.2.2. L'enseignement explicite est-il bénéfique au développement des savoirs implicites ?..... - 1109 -

II.2.3. Apprentissage explicite et réflexion métalinguistique..... - 1116 -

II.2.4. Enseignement explicite et « prononciation »..... - 1125 -

**III. CONSIDERATIONS METHODOLOGIQUES : ATTENTION SUR LA FORME, OUTILS LINGUISTIQUES ET**

**NORME DIDACTIQUE..... - 1135 -**

*III.1. Attention sur la forme et approches didactiques..... - 1135 -*

III.1.1. Attention : niveau cognitif..... - 1140 -

III.1.2. Attention : niveau métacognitif..... - 1147 -

*III.2. Modèles linguistiques et corpus..... - 1150 -*

III.2.1. Modèles linguistiques et modèles didactiques..... - 1150 -

III.2.2. Corpus oraux..... - 1166 -

*III.3. Norme didactique..... - 1176 -*

III.3.1. Description vs. Prescription vs. Adaptation .....	- 1177 -
III.3.2. De la sociolinguistique à la psycholinguistique <i>pour la didactique</i> .....	- 1180 -
III.3.3. Le cas de la « prononciation » .....	- 1185 -
<b>III.4. Perspectives didactiques méthodologiques.....</b>	<b>- 1193 -</b>
III.4.1. Intégration méthodologique de l'articulation oral/écrit.....	- 1193 -
III.4.1.1. Module méthodologique.....	- 1200 -
III.4.1.2. Module explicite.....	- 1201 -
III.4.1.3. Module implicite .....	- 1202 -
III.4.2. Parcours d'apprentissage de l'oral.....	- 1203 -
III.4.2.1. Utilisation des « méthodes » de FLE.....	- 1206 -
III.4.2.2. Exercices, Lecture et API .....	- 1210 -
III.4.2.2.1. Les « exercices » .....	- 1211 -
III.4.2.2.2. La lecture à voix haute .....	- 1211 -
III.4.2.2.3. L'utilisation de l'Alphabet Phonétique International (API) .....	- 1212 -
III.4.2.3. Correction phonétique et correction phonographémique .....	- 1213 -
III.4.2.3.1. Utilisation du support visuo-graphique lors de la correction phonétique.....	- 1213 -
III.4.2.3.1.1. Identification/discrimination.....	- 1214 -
III.4.2.3.1.2. Segmentation .....	- 1217 -
III.4.2.3.2. Correction phonographémique en production (dictée) .....	- 1218 -
III.4.2.3.3. Correction phonographémique en perception (lecture à voix haute).....	- 1219 -
III.4.5. Ouvertures .....	- 1223 -
<b>CONCLUSION GENERALE.....</b>	<b>- 1227 -</b>
<b>BIBLIOGRAPHIE .....</b>	<b>- 1233 -</b>
<b>ANNEXES .....</b>	<b>- 1273 -</b>
I. MANUELS RELATIFS A LA PRONONCIATION DU FRANÇAIS .....	- 1273 -
II. PRE-TESTS.....	- 1281 -
II.1. Document traduit en japonais et distribué aux étudiants.....	- 1281 -
II.2. Notice de l'instructeur.....	- 1286 -
III. TESTS GENERAUX.....	- 1288 -
III.1. La notice de l'instructeur.....	- 1289 -
III.1.1. Tests A, F et C1 .....	- 1289 -
III.1.2. Tests C2, D et E .....	- 1292 -
III.2. Description des conditions de passation.....	- 1295 -
III.3. Documents de rendu des sujets .....	- 1296 -
III.3.1. Tests A, F et C1 (en français) .....	- 1296 -
III.3.2. Tests C2, D et E (en français) .....	- 1305 -

## **Notes liminaires**

Voici ci-dessous quelques indications relatives aux codes d'écriture utilisés dans cette thèse :

1) En ce qui concerne la problématique du genre en français dans notre texte, nous sommes conscients des problèmes que peuvent poser des appellations génériques masculines du type « les chercheurs », appellation qui recouvre également « les chercheuses ». Nous n'avons adopté aucun principe d'écriture à ce sujet, et tenons simplement à indiquer ici la prise en considération de la question.

2) Il en est de même pour les appellations « langue maternelle » et « langue première » d'une part, et « langue étrangère » et « langue seconde » d'autre part, notions que nous assimilerons souvent en L1/LM et L2/LE, n'apportant de précisions supplémentaires que lorsque la distinction semble pertinente ou nécessaire dans notre propos.

3) En ce qui concerne les symboles phonético-phonologiques utilisés, nous nous sommes essentiellement basés sur ceux de l'Alphabet Phonétique International. Les crochets [ ] renvoient au niveau phonétique, les barres obliques / / au niveau phonologique et les guillemets simples < > ou doubles « » au niveau graphémique. Cependant, par souci de lisibilité et de clarté dans notre propos, nous avons pu parfois utiliser d'autres codes idiosyncrasiques qui, en contexte, facilitent la lecture et la compréhension du texte. Une de ces idiosyncrasies concerne le phonème /r/ et ses différentes réalisations en français et en japonais : afin de clairement différencier l'unité dans les deux langues, nous avons utilisé, dans certaines parties de cette thèse, la majuscule *R* pour le français et la minuscule *r* pour le japonais.

4) Lorsque nous citons certaines références dont nous n'avons pu disposer en main propre, celles-ci sont indiquées en note de bas de page et non dans la bibliographie générale.

5) Seules certaines citations en anglais ont été traduites par nous en français, lorsque cela semblait opportun. Les autres citations en anglais ont été reproduites telles quelles, notamment afin d'éviter tout risque d'interprétation ou de traduction approximative, voire erronée, du texte original.

6) Enfin, le choix du « nous » comme « sujet écrivant » de cette thèse : nous sommes en effet conscients du fait que certains membres de la communauté académique préfèrent l'utilisation du « je » à celle du « nous » lorsque l'auteur est unique, comme c'est le cas ici. Néanmoins, notre préférence pour le « nous », si elle s'explique essentiellement en termes de « style personnel », se justifie également par le fait que, dans notre optique, un travail comme celui-ci, et, à vrai dire, tout travail de recherche, nous semble-t-il, ne peut être véritablement considéré comme la production originale d'un être unique et isolé : un tel travail se base en effet sur un ensemble de travaux antérieurs et actuels dont il s'inspire, et bénéficie également d'un encadrement et d'échanges humains qui alimentent ledit travail, au sein d'une communauté humaine, comprenant, dans notre cas, des chercheurs, des enseignants, des concepteurs, des apprenants et d'autres encore. Les idées formulées et les démarches adoptées sont, de même, bénéficiaires de ce que l'on nomme parfois « l'air du temps », dans le domaine de la recherche. Ainsi, l'utilisation du « nous », loin d'exprimer une quelconque prétention ou suffisance, qu'elle pourrait évoquer à certains, correspond au contraire pour nous à la modeste reconnaissance de la dimension relativement collective de tout effort de recherche, remettant ainsi à sa place l'*ego*, somme toute bien quelconque, du chercheur individuel.

## **Introduction Générale**

Alors que l'essor des sciences dites « cognitives » surprend par son ampleur depuis maintenant plusieurs décennies, éclairant chaque jour davantage le fonctionnement cognitif de l'être humain, la didactique des langues (ou la didactologie des langues et des cultures comme l'appelleront certains) en France ne semble toujours pas s'être suffisamment fondée scientifiquement. Tout en évitant une quelconque comparaison hâtive avec le domaine anglo-américain de *Second Language Acquisition* (SLA), comparaison qui soulèverait des problèmes d'ordre épistémologique et socio-politique, il semble difficile de nier que l'utilisation qui est y respectivement faite des outils théoriques et méthodologiques d'autres disciplines apparemment mieux établies (en particulier la psychologie cognitive et les sciences du langage) diffère de manière parfois importante, comme cela apparaît à la lecture des grandes revues des domaines concernés. Inévitablement pluridisciplinaire, la didactique des langues appartient au champ des sciences que l'on qualifie encore de nos jours d'« appliquées », lesquelles posent précisément un problème d'application et de transposition de connaissances issues de disciplines jugées plus fondamentales. Loin de constituer un ensemble de repères stables, ces dernières, de par leur caractère scientifique, se voient pourtant régulièrement bouleversées, éclatées et renouvelées, évolution scientifique indispensable d'un point de vue heuristique.

Tandis que la linguistique dite « appliquée », dans sa relation à la didactique des langues, se voit supplantée chaque jour davantage par ces sciences de la cognition, qui, en fait, engloberaient en grande partie les sciences du langage, il semblerait que ce que l'on serait tenté aujourd'hui de qualifier de lapalissade plus ou moins obsolète selon les écoles, à savoir que « la didactique des langues est aux mains des linguistes », est en passe de devenir caduque : après le « linguistique » et, dans une certaine mesure, le « culturel », c'est le « cognitif » qui, depuis plus d'une décennie à présent, semblerait dominer en partie l'actualité didactique. Du moins en surface. Car même s'il est vrai que les enseignants de langue s'intéressent enfin à « l'apprentissage », resté jusque-là

dans l'ombre de « l'enseignement », voilà bien des années que ce tournant a cédé la place à d'autres dans la communauté des chercheurs, et force est de constater qu'une fois de plus la transposition vers les terrains d'enseignement / apprentissage reste toujours difficile à négocier.

La position didactique que nous adoptons à ce jour pourra peut-être sembler quelque peu ambivalente : d'un côté nous rejetons la perspective applicationniste précédemment mentionnée, car nous considérons que seuls les didacticiens peuvent envisager de mettre en œuvre des paradigmes de recherche véritablement adaptés aux problématiques et aux finalités didactiques, reliant théories et concepts du domaine aux pratiques et aux outils des classes de langue. D'un autre côté, nous considérons que cette mise en œuvre exige de ces didacticiens une formation/information suffisante dans les domaines connexes auxquels se rattachent leurs problématiques d'étude, au premier rang desquels figure, selon nous, la psycholinguistique, lorsque l'on s'intéresse à l'apprentissage d'une langue étrangère, psycholinguistique qui présuppose une solide base linguistique. De la même manière, si nous considérons que les différents paramètres impliqués dans toute situation *réelle* d'enseignement / apprentissage œuvrent en synergie, de manière relativement holistique, il nous semble en revanche que la recherche dans le domaine ne peut porter ses fruits qu'à condition de faire preuve d'un certain discernement dans les tâches et les opérations impliquées dans de telles situations. Cette parcellisation provisoire des objets de recherche, incontournable dans l'état actuel de la démarche scientifique, ne peut cependant satisfaire le didacticien, qui doit, par la suite, réintégrer ces données à un ensemble didactique complexe, dont les finalités doivent être effectives. Dès lors, l'intégrité de la didactique des langues ne peut, nous semble-t-il, se constituer qu'en intégrant la pluridisciplinarité qu'elle invoque nécessairement, une intégration qui conduit, non plus à un quelconque domaine « appliqué » à l'enseignement / apprentissage des langues, mais plutôt à une approche « implicationniste » en didactique.

Ce passage de la « linguistique appliquée » à la « didactique impliquée » suppose, pour le didacticien, de formuler des problématiques d'enseignement / apprentissage ancrées

dans le quotidien des classes de langues, mais traitées en faisant usage des données, voire des démarches, des disciplines pertinentes, en vue d'en retirer des *implications* potentielles pour la théorisation didactique et surtout pour les démarches et les techniques d'enseignement, ces implications devant être soumises à validation empirique.

Cette position se trouve parfaitement reflétée dans l'architecture et la démarche adoptées dans notre travail de thèse : formulant en premier lieu des problématiques d'ordre didactique, ancrées dans le terrain et dans la théorie, nous nous plongeons ensuite dans les domaines linguistique et psycholinguistique pour en retirer certaines informations pertinentes, lesquelles nous permettent de préciser nos problématiques initiales, en vue de formuler des hypothèses que nous confrontons quasi-expérimentalement<sup>1</sup> au milieu écologique des classes de langue pour le public qui nous concerne. A partir de ces résultats empiriques, ainsi que de notre réflexion didactique, alimentée par les données de la linguistique et de la psycholinguistique, nous reprenons alors notre perspective didactique initiale pour tenter de retirer certaines implications, générales et spécifiques, à la fois « théoriques », mais aussi très « pratiques », pour l'enseignement des langues d'une part et pour le développement de la recherche en didactique d'autre part.

Comme nous l'avons affirmé, l'objet d'étude de la recherche en didactologie des langues et des cultures, à savoir, de manière très générale, l'enseignement / apprentissage des langues et des cultures, présente une complexité certaine, qu'il est difficile de traiter dans sa globalité. Nous sommes donc, dans le cadre d'un travail comme le nôtre, contraints d'effectuer des choix, quant aux aspects de cet objet que nous souhaitons plus finement analyser. Partant du principe que l'un des objectifs

---

<sup>1</sup> Nous utilisons l'expression « quasi-expérimental(e) » pour désigner une recherche qui implique une certaine manipulation artificielle des variables en jeu par le chercheur (dimension « expérimentale »), tout en préservant une certaine écologie à la situation, qui la distingue de situations expérimentales de laboratoire (d'où le préfixe « quasi »). Dans notre cas, par exemple, les tests (constructions artificielles à visée expérimentale) ont été passés par des étudiants dans leurs conditions d'apprentissage habituelles (lieux, horaires, encadrement) et leur ont été présentés comme de simples activités d'apprentissage, insérées dans leur routine pédagogique hebdomadaire.

ultimes de ce domaine de recherche est l'amélioration, en des termes variables selon les différents points de vue envisageables, de l'apprentissage *via* celle de l'enseignement, puis considérant que celle-ci ne peut s'opérer que grâce à une meilleure connaissance des processus d'apprentissage, sur lesquels peut bien évidemment influencer l'enseignement, et pour diverses raisons enfin, liées à nos intérêts personnels, à notre formation, initiale et continue, et à notre pratique de l'enseignement, nous avons donc décidé de nous concentrer sur certains aspects de cet objet, de la manière suivante :

	• Didactologie des langues et des cultures étrangères.
↓	• Didactique des langues étrangères.
↓	• Didactique du français langue étrangère (FLE).
↓	• Enseignement/apprentissage du FLE.
↓	• Enseignement/apprentissage de l'oral en FLE.
↓	• Utilisation de la multimodalité pour l'apprentissage de l'oral en en FLE.
↓	• Utilisation de l'oral et de l'écrit pour l'apprentissage de l'oral en FLE.
↓	• Utilisation de l'oral et de l'écrit pour l'apprentissage de la prononciation en FLE.
↓	• Utilisation de l'oral et de l'écrit pour l'apprentissage des phonèmes /r, l, b, v/ et des structures syllabiques à groupes biconsonantiques en français par des apprenants (faux-) débutants adultes japonais au Japon.

Cette « réduction » progressive, contrainte par les exigences de recherche propres à notre travail de thèse, n'implique en aucun cas, comme nous l'avons indiqué, une quelconque « déconnexion », très certainement illusoire, des autres dimensions de l'enseignement / apprentissage qui figurent dans les lignes supérieures de cette progression. Si cette orientation « haut - bas » représente bien la focalisation de notre étude, et ainsi les cinq premières parties de cette thèse, la dernière partie, en revanche, est orientée « bas - haut », puisque l'approche « implicationniste » que nous venons de décrire succinctement, si elle nous autorise à suivre la première orientation, nous contraint également à effectuer le parcours inverse pour réinvestir nos données



« partielles » dans une problématique didactique, à la fois plus générale, mais également, dans une certaine mesure, plus *réaliste*.

Au cœur de notre étude se trouve donc la problématique des interactions entre représentations visuelles et représentations auditives au cours de l'apprentissage multimodal (essentiellement visuel et auditif) de l'oral en langue étrangère. Nous nous intéressons plus particulièrement à l'apprentissage du système phonético-phonologique du français par des apprenants adultes japonais débutants et nous nous interrogeons sur l'influence potentielle des représentations phonographémiques du japonais et/ou graphémiques du français sur l'apprentissage des formes phonético-phonologique du FLE.

Au vu de la variété des différentes représentations visuelles sémiotiques existantes (graphiques et iconiques) et au vu de celles qui sont utilisées en didactique des langues, en particulier l'alphabet romain dans le cas du FLE, nous nous sommes demandés, étant donné ce que l'on sait désormais en psycholinguistique de la gestion cognitive des inputs visuels et auditifs dans l'apprentissage d'une langue étrangère, si ces représentations (orthographiques, phonétiques, iconiques, etc.), ou bien l'utilisation qui en est faite (combinaisons spatio-temporelles entre inputs auditifs et visuels, consignes attentionnelles explicites ou implicites, etc.), étaient les plus « efficaces » pour l'apprentissage de l'oral en langue étrangère. Cette question s'est posée à nous pour au moins deux raisons :

- D'une part, *le développement de l'Apprentissage des Langues Assisté par Ordinateur (ALAO)*. Etant donné le développement des outils informatiques dans le domaine de l'enseignement / apprentissage des langues (Chapelle, 2001), sous forme de didacticiel notamment, les concepteurs de cours de langues assistés par ordinateur peuvent désormais faire appels à des ergonomes et des infographistes pour produire de nouveaux formats de présentation de l'input linguistique et de nouveaux outils pédagogiques. Les environnements multimédia, et *a fortiori* hypermédia, donnent ainsi la possibilité de combiner des inputs visuels et auditifs,

d'une manière auparavant difficilement envisageable, étant donné le matériel didactique existant et disponible alors, c'est-à-dire, essentiellement, textes, cassettes audio, cassettes vidéo, images fixes (photos, dessins, schémas) et images mobiles (diapositives, etc.).

- D'autre part, *le développement des sciences cognitives*. La gestion cognitive de la bimodalité (auditive et visuelle) est aujourd'hui mieux connue, et différents modèles rendent compte des rapports entre représentations auditives et visuelles, à un niveau périphérique (perceptif) mais aussi plus profond (symbolique et/ou subsymbolique). Les didacticiens disposeraient donc, sinon de recommandations, du moins de pistes d'investigation relativement précises concernant l'organisation souhaitable du matériel visuel et sonore, afin d'améliorer la perception, la compréhension, la mémorisation et *in fine* l'apprentissage du matériel verbal en langue étrangère.

Or, si la multimodalité semble être un thème d'actualité, dans chacun des deux domaines mentionnés, il nous a paru essentiel de ne pas oublier que la bimodalité audiovisuelle commençait par la simple combinaison du texte et du son, renvoyant alors, pour l'apprentissage de la parole, à une problématique déjà quelque peu ancienne en didactique des langues, celle de l'articulation entre supports oraux et écrits pour l'enseignement / apprentissage de l'oral en langue étrangère, qui avait notamment été spécifiquement traitée par les concepteurs de la méthodologie *Structuro-Globale Audio-Visuelle* (SGAV) (Rivenc et Boudot, 1962 ; Rivenc, 2000).

Par ailleurs, lorsque l'on examine la littérature didactique relative à l'apprentissage de l'oral, et plus particulièrement de la « prononciation », en langue étrangère, on constate qu'il y a là un vaste domaine de recherche, préoccupé depuis très longtemps par les phénomènes d'« interférences » phonético-phonologiques entre langue première (L1) et langue étrangère (L2) durant l'apprentissage de celle-ci. Ce domaine semble récemment connaître un nouvel essor, en particulier *via* ce que l'on nomme « l'interphonologie » des apprenants, notamment en raison des nouvelles approches et

des nouveaux outils linguistiques disponibles (modèles à contraintes, analyses acoustiques, etc.). Parallèlement, dans le domaine de l'apprentissage de la lecture, et de l'écriture, en L1, comme en L2, il existe une abondante littérature, souvent ciblée sur les populations enfantines, et relevant généralement davantage de la psychologie développementale que de la didactique des langues étrangères.

Cependant, il nous semble que les études croisant les thèmes « interphonologie » et « multimodalité » sont à ce jour peu nombreuses, notamment lorsque ladite « multimodalité » correspond à une bimodalité audio-orthographique, négligeant ainsi l'impact potentiel du code écrit, en L1 et/ou en L2, sur la constitution de cette « interphonologie », alors que le texte écrit semble être encore aujourd'hui omniprésent dans les classes de langue, y compris d'« oral ». A notre connaissance, aucune étude transversale de la sorte, adoptant une approche à la fois phonologique, psycholinguistique et didactique, n'avait jusqu'à présent été réalisée pour des étudiants japonais apprenant le français.

C'est donc dans un projet de recherche relativement pluridisciplinaire que nous nous sommes lancés pour cette thèse, tentant d'intégrer certains des travaux les plus récents, du moins dans le domaine francophone, en psycholinguistique, en phonologie et en didactique des langues, en assurant ainsi une base théorique solide à des investigations empiriques conduites en milieu écologique dans une perspective résolument didactique. Une telle tentative n'est pas sans risque, nous exposant inmanquablement aux dangers de la pluridisciplinarité : notre travail sera sans doute jugé insuffisamment linguistique par les linguistes, insuffisamment psycholinguistique par les psycholinguistes et insuffisamment didactique par les didacticiens. Pourtant, en tant qu'enseignant de FLE, formé en didactique des langues et des cultures et co-dirigé, dans ce travail, par un linguiste, spécialiste de phonologie, de phonétique et de linguistique générale, et un neuropsycholinguiste, spécialiste de psycholinguistique, de neurosciences et de sciences cognitives, il nous semble que ce travail relève pleinement de la pluridisciplinarité inhérente à la didactique des langues, et c'est donc bien d'une thèse

de didactique des langues, et non de linguistique ou de psycholinguistique, qu'il s'agit, axée sur le rôle de l'écrit dans l'apprentissage de l'oral en langue étrangère.

Ce travail se présente en six parties :

- 1) Une partie didactique introductive posant le cadre de notre étude et la problématique générale ;
- 2) Une partie linguistique consacrée à la comparaison des systèmes phonétophonologiques et phonographémiques français et japonais ;
- 3) Une partie psycholinguistique décrivant l'appareil cognitif des apprenants de langue étrangère, particulièrement axée sur la perception de l'oral et de l'écrit ;
- 4) Une partie « problématique » permettant de formuler des hypothèses de recherche précises concernant le rapport entre « interphonologie » et « représentations graphémiques » ;
- 5) Une partie « expérimentale » présentant nos investigations quasi-expérimentales en milieu écologique, visant à tester nos hypothèses de recherche ;
- 6) Une partie didactique conclusive dans laquelle nous tentons de tirer les implications didactiques, spécifiques et générales, de notre étude.

Dans la partie didactique introductive, après avoir souligné la nécessité de faire appel à la phonologie, à la psycholinguistique et à la didactique des langues lorsque l'on souhaite se former à l'enseignement, et plus encore conduire des recherches sur l'apprentissage, de la parole en langue étrangère, nous abordons le thème de la multimodalité *via* le thème de l'Enseignement / Apprentissage des Langues Assisté par Ordinateur (E/ALAO). Nous introduisons ensuite celui de l'enseignement de la prononciation en L2, en nous concentrant sur le cas du FLE au Japon, à travers notamment des observations de classes, de manuels, de formation d'enseignants et des enquêtes réalisées sur le terrain. Cette première partie nous permet ainsi de poser notre problématique générale, le rapport entre « enseignement/apprentissage de l'oral » et « multimodalité », ainsi que le cadre didactique dans lequel nous l'inscrivons, à savoir celui de l'enseignement du français au Japon. Cependant, en vue de cerner plus

précisément ce que l'on entend par « oral » et « prononciation », il nous faut faire appel au domaine linguistique, et ainsi aborder notre deuxième partie.

Dans la partie linguistique, nous établissons dans un premier temps les systèmes phonético-phonologiques du français et du japonais, en introduisant certaines théories phonologiques utiles à la compréhension d'études récentes en interphonologie, et en nous concentrant sur les inventaires segmentaux et les structures syllabiques des deux langues en question. Nous présentons, dans un second temps, leur système phonographémique, en insistant sur les particularités du japonais, avant, finalement, de contraster les deux systèmes en présence, nous permettant ainsi de sélectionner deux aspects de la langue française qui posent des difficultés d'apprentissage aux sujets japonais, à savoir les oppositions phonémiques /r/ vs. /l/ et /b/ vs. /v/ d'une part, et les structures syllabiques biconsonantiques branchantes de type /CCV/ ou /VCC/ d'autre part. Afin de mieux établir les interactions potentielles entre systèmes phonético-phonologiques et phonographémiques en L1 et en L2, dans l'apprentissage de l'oral en L2, nous devons cependant disposer de données suffisantes concernant la perception et le traitement de l'oral, la perception et le traitement de l'écrit, et ainsi, plus généralement l'appareil cognitif de l'apprenant, ce qui nous conduit au domaine psycholinguistique, et donc à notre troisième partie.

Dans la partie psycholinguistique, nous proposons dans un premier temps un descriptif global des systèmes perceptivo-attentionnel et mnésique des apprenants de langue étrangère, en insistant d'une part sur les processus à l'œuvre en perception de la parole et sur sa variation selon les langues, et d'autre part sur le fonctionnement de la mémoire de travail, particulièrement sollicitée lors du traitement de matériel verbal inconnu. Nous nous penchons ensuite sur la dimension multimodale, et plus précisément audiovisuelle, de la perception de la parole, d'abord de manière générale, puis dans le cas d'apprenants de langue étrangère. De la modalité visuelle faciale, nous passons alors à la modalité visuelle orthographique, qui nécessite de présenter certaines données relatives à la perception de l'écrit, en particulier dans le cas du français et du japonais. A partir de ces données, nous examinons finalement les interactions possibles

entre représentations orales et écrites, et plus particulièrement l'influence potentielle des représentations orthographiques sur la perception et le traitement de la parole, que ce soit en termes d'unités ou de conscience phonologiques. A ce stade de notre étude, il est alors possible de synthétiser les données récoltées dans la littérature linguistique et psycholinguistique pour affiner notre problématique didactique initiale, ce qui nous amène à notre quatrième partie.

Dans la partie « Problématiques », nous recentrons en effet notre étude autour du concept d'« interphonologie », en présentant un ensemble de travaux à ce sujet, dans une approche tant structuraliste-fonctionnaliste, que générativiste ou exemplariste, et formulons nos premières hypothèses de recherche vis-à-vis des apprenants japonais de français, essentiellement en termes phonético-phonologique. A ces premières hypothèses, nous donnons ensuite une orientation davantage psycholinguistique en précisant, selon différentes tâches et différentes variables, le rôle potentiel des représentations orthographiques dans la constitution de ladite interphonologie, en particulier dans le cas de Japonais, sur les points linguistiques que nous avons sélectionnés. Les hypothèses de recherche ainsi formulées nous mènent à la présentation des grands traits de notre protocole quasi-expérimental, et donc à notre cinquième partie.

La partie « expérimentale » se divise en quatre temps : dans le premier, après avoir brièvement décrit la méthodologie adoptée, nous présentons des « pré-tests » (protocole, résultats et interprétations) que nous avons effectués auprès de 22 sujets dans une école de langue privée au Japon. Ces pré-tests devaient essentiellement nous permettre de calibrer les tests généraux dont nous décrivons le protocole dans un second temps et qui comportent six tâches : une tâche d'identification phonographémique segmentale, une tâche d'identification phonographémique et de discrimination phonologique segmentale, deux tâches de comptage syllabique en unimodalité auditive, une tâche de comptage syllabique en unimodalité orthographique

et une tâche de comptage syllabique en bimodalité audio-orthographique<sup>2</sup> synchrone, convergente et divergente. Les trois premières ont été passées par 120 étudiants universitaires et les trois autres par 62 d'entre eux, dans leur environnement didactique habituel. Les deux tâches segmentales concernaient les couples phonémiques /r, l/ et /b, v/ : il s'agissait pour la première de choisir la transcription orthographique d'un non-mot présenté auditivement entre deux choix proposés ne s'opposant que par un graphème (<R> vs. <L> et <B> vs. <V>); la deuxième comportait deux non-mots présentés auditivement de manière successive, les sujets devant alors choisir deux transcriptions orthographiques, identiques ou différentes, parmi quatre proposées. Trois variables indépendantes principales étaient contrôlées : la nature du phonème consonantique, sa position dans le non-mot et la nature de la voyelle adjacente. L'objectif principal de ces tâches était de mesurer le taux d'erreurs phonographémiques dans les deux tâches, en vue de soulever la question du niveau, phonétique, phonologique ou phonographémique auquel l'on devait assigner celles-ci. D'un point de vue didactique, il s'agissait de mettre au jour la dimension phonographémique de l'apprentissage des catégories phonético-phonologiques considérées. Les quatre tâches de comptage syllabique impliquaient, pour les étudiants, de compter le nombre de syllabes des non-mots présentés, contenant des groupes consonantiques /CC/, entourant un nombre d'ellipses identiques. Les deux premières tâches ne différaient que par la nature des groupes /CC/ et par la population testée. Les non-mots et les sujets testés étaient en revanche identiques dans les trois dernières tâches (unimodalité auditive, unimodalité orthographique et bimodalité synchrone audio-orthographique), ce qui nous a permis d'évaluer l'effet de la modalité de présentation sur les résultats de comptage syllabique, et donc sur le traitement des groupes /CC/. Les deux variables indépendantes principales pour chacune de ces tâches étaient la nature du groupe

---

<sup>2</sup> Une telle terminologie pourra surprendre le lecteur : on oppose en effet généralement « auditif » à « visuel » et « phonético-phonologique » à « orthographique ». Les raisons pour lesquelles nous avons décidé de substituer le terme « orthographique » à celui, plus général, de « visuel » dans certaines parties de la thèse sont les suivantes : 1) cela permet de préciser le type exact de stimulus visuel (qui, autrement, peut être de nature variée), 2) cela permet d'éviter toute confusion entre nos travaux d'une part et ceux, portant sur la perception audiovisuelle de la parole, dans lesquels les stimuli visuels correspondent à la vision du visage, en particulier de la bouche et des lèvres, du locuteur d'autre part (notamment les travaux de Massaro). Nous assumons donc la responsabilité de ce choix terminologique.

consonantique et sa position dans le non-mot. Dans un troisième temps, nous présentons les résultats de ces tests, qui sont ensuite interprétés dans un quatrième temps. Cette interprétation nous permet d'ouvrir certaines perspectives pour les trois domaines, phonétique-phonologie, psycholinguistique et didactique des langues, auxquels nous nous rattachons ici, et qui nous conduisent vers les implications didactiques de notre dernière partie.

Dans la partie didactique conclusive, nous effectuons un retour vers notre problématique centrale, en considérant, dans un premier temps, les implications spécifiques de nos tests pour l'enseignement des phonèmes /r, l, b, v/ ainsi que celui des groupes consonantiques aux apprenants japonais, avant d'élargir notre réflexion, en tirant parti de l'ensemble de notre étude, à des implications didactiques plus générales, notamment méthodologiques, pour l'utilisation de l'écrit dans l'enseignement de l'oral en langue étrangère. A cette fin, nous examinons deux grands axes thématiques : celui du rôle de la modalité visuelle dans l'apprentissage de la prononciation en langue étrangère, puis celui de la nature, que l'on peut globalement qualifier d'explicite ou d'implicite, de l'enseignement. Ces deux axes nous permettent alors de nourrir certaines considérations didactiques méthodologiques, ayant trait d'une part au rôle de « l'attention sur la forme » dans l'enseignement / apprentissage, d'autre part au rapport entre modèles linguistiques et modèles didactiques, dans lequel est incluse la notion de « corpus », et enfin à la conception de la « norme didactique », dans une perspective à la fois sociolinguistique et psycholinguistique, pour la didactique. Ces considérations débouchent, en fin de partie, sur des perspectives méthodologiques, incluant l'intégration de l'articulation oral/écrit dans un parcours d'apprentissage de l'oral, dans lequel doit être couplée à la correction phonétique une indispensable correction phonographémique.

A l'issue de ce parcours, ce sont bel et bien des ouvertures multiples, tant sur le plan de la recherche, linguistique, psycholinguistique ou, bien évidemment, didactique, que de la conception et des pratiques pédagogiques, qui viennent clore notre étude.



## **Partie Didactique I :**

### **Introduction**

## Table des Matières Partielle

I. DIDACTIQUE DES LANGUES, PSYCHOLINGUISTIQUE ET PHONOLOGIE : UN TRIPTYQUE DISCIPLINAIRE POUR LA FORMATION A L'ENSEIGNEMENT DE LA PAROLE EN LANGUE ETRANGERE.....	- 30 -
1.1. <i>La didactique des langues et des cultures (DLC)</i> .....	- 30 -
1.2. <i>La phonologie</i> .....	- 41 -
1.3. <i>La psycholinguistique</i> .....	- 48 -
1.4. <i>Intégration et transdisciplinarité</i> .....	- 49 -
II. MULTIMEDIA ET MULTIMODALITE .....	- 51 -
II.1. <i>Domaine et historique</i> .....	- 51 -
II.2. <i>Multimodalité/ multicanalité : un atout pour l'apprentissage ?</i> .....	- 55 -
III. ENSEIGNEMENT DE LA PRONONCIATION .....	- 58 -
III.1. <i>Situation actuelle de l'enseignement de la prononciation du FLE au Japon</i> .....	- 58 -
III.1.1. Raisons d'ordre institutionnel .....	- 59 -
III.1.2. Raisons d'ordre socioculturel.....	- 60 -
III.1.3. Raisons d'ordre (psycho-)linguistique .....	- 62 -
III.1.4. Raisons d'ordre didactique.....	- 63 -
III.2. <i>Qu'est-ce que la « prononciation » ?</i> .....	- 64 -
III.2.1. Définitions.....	- 64 -
III.2.2. Interfaces : langue et communication .....	- 66 -
III.3. <i>Observations, pratiques et enquêtes au Japon</i> .....	- 67 -
III.3.1. Observations.....	- 67 -
III.3.1.1. De la perception auditive à la production écrite : erreurs de « phonétique » en production sous dictée.....	- 70 -
III.3.1.2. Perception auditive et production orale .....	- 74 -
III.3.1.3. Lecture et production orale.....	- 77 -
III.3.2. Pratiques .....	- 79 -
III.3.2.1. Matériel pédagogique .....	- 79 -
III.3.2.2. Formation des enseignants : de la phonétique articulatoire à la méthode verbo-tonale d'intégration phonétique .....	- 82 -
III.3.3. Enquêtes : représentations et motivations concernant la prononciation du français au Japon .....	- 92 -
III.3.3.1. Enquête 1 : Apprenants en atelier de correction phonétique en école privée (n=31).....	- 93 -
III.3.3.2. Enquête 2 : Enseignants français (n = 32) et japonais (n = 21) de français (total=53).....	- 101 -
III.3.3.3. Enquête 3 : Apprenants de français en université (n = 273) .....	- 104 -
IV. TRANSITION.....	- 105 -

## **I. Didactique des langues, psycholinguistique et phonologie : un triptyque disciplinaire pour la formation à l'enseignement de la parole en langue étrangère**

Encadrer et guider l'apprentissage de la parole en langue étrangère, implique, pour le sujet enseignant, de disposer d'un ensemble d'informations suffisant pour prendre des décisions à même de faciliter ou d'améliorer cet apprentissage. Trois domaines de recherche sont alors particulièrement susceptibles de lui fournir ces informations : la didactique des langues et des cultures, la psycholinguistique et le couple phonétique/phonologie, lorsqu'il s'agit plus particulièrement des aspects phonéto-phonologiques de l'apprentissage, que ce soit en production ou en perception. Quel que soit le projet didactique entrepris (enseignement, formation, conception ou recherche), ces trois domaines sont ceux qu'il convient de connaître autant que possible en vue de pouvoir réaliser une intégration personnelle des données et des problématiques disponibles, pertinentes vis-à-vis dudit projet. Cela devient particulièrement crucial lorsque l'on envisage de recourir à la multimodalité pour l'enseignement / apprentissage de la parole en langue étrangère.

### **I.1. La didactique des langues et des cultures (DLC)**

Si l'on cherche à caractériser le champ particulier que semble couvrir la didactique des langues et des cultures (DLC), il est possible de parvenir à la perspective suivante (Detey, 2003) : la DLC apparaît comme un domaine doublement orienté, « théorique » et « pratique », dont les sources et les extensions franchissent les frontières disciplinaires traditionnelles afin de cerner une réalité dont l'étude ne peut être que clairement pluridisciplinaire<sup>3</sup>. Pluridisciplinaire, car la communication inter-individuelle *via* une langue particulière implique, pour simplifier, des individus, un temps et un espace, bref, plusieurs sphères énonciatives d'une redoutable complexité,

---

<sup>3</sup> Ou « transdisciplinaire » selon certains (Pothier, 2003, pp. 38-41), dont nous partageons en grande partie la perspective.

dont le nombre d'angles d'étude possibles se voit reflété par celui des champs disciplinaires qui se les sont appropriés. Pluridisciplinaire aussi, car « la compréhension et la production du langage nécessitent l'apport de plusieurs fonctions cérébrales fondamentales telles que la perception d'une structure (à la fois visuelle et auditive), l'analyse temporelle, la mémoire, et la production temporelle pour n'en nommer que quelques-unes » (Gordon et Weide, 1983, p. 45). Cette pluridisciplinarité relève donc notamment de celle qui anime les sciences dites « cognitives », des sciences dont les démarches et les données manquent encore, à notre connaissance, d'être suffisamment réinvesties en DLC, alors même qu'elles pourraient en constituer le socle scientifique. Les sciences cognitives « tentent à la fois de comprendre et d'expliquer les mécanismes de la connaissance, ainsi que d'améliorer les performances cognitives de l'homme », peut on lire dans l'introduction d'un ouvrage collectif qui leur est consacré (Gordon et Paugam-Moisy, 1997, p. 16). Si l'on considère le passage du monolinguisme au plurilinguisme comme une amélioration des « performances cognitives de l'homme », alors la DLC a bien sa place dans le vaste ensemble de recherches que constituent lesdites sciences.

Il faut cependant reconnaître l'extrême difficulté de la tâche : décompartmenter, puis intégrer les disciplines alors même qu'elles sont en plein développement, en vue de générer une interaction entre leurs apports respectifs pose non seulement des problèmes d'expertise dans les domaines considérés, mais aussi des problèmes d'adaptation de formes, de contenus, de terminologies et de perspectives. Lorsque l'on songe ensuite aux différents modèles et aux différentes théories en compétition au sein de chaque micro-champ (qu'il s'agisse de modèles mnésiques, phonologiques ou d'apprentissage par exemple), on ne peut que se résigner ou être en proie à la confusion, alors que les didacticiens, et à leur suite les enseignants, ont besoin de concepts opératoires et limités en nombre.

Néanmoins, il suffit d'examiner la *compétence de communication* (Moirand, 1982) que le sujet apprenant est aujourd'hui censé acquérir, pour admettre la nécessité de faire appel à plusieurs domaines de recherche, sans pour autant dénaturer la cohérence et la

cohésion propres du corps didactique. Le tableau que présente Intravaia à la fin de son ouvrage (2000, Annexe 1, p. 277) et que nous reproduisons ci-dessous, malgré sa présentation « en liste » qui le rend quelque peu trop englobant, résume bien, de notre point de vue, sans être exhaustif cependant<sup>4</sup>, la pluralité des problématiques sous-jacentes à l'apprentissage d'une langue étrangère (LE), ainsi que l'évolution de celles-ci à travers celle de la notion de compétence de communication. Nous sommes en effet passé d'une compétence que l'on pourrait grossièrement qualifier de « linguistique », à un conglomérat de micro-compétences qui constituent aujourd'hui ladite « compétence de communication »<sup>5</sup>, et qui expliquent le caractère nécessairement pluridisciplinaire de la DLC.

---

<sup>4</sup> La « marguerite de la didactique » de Pothier (2003, p. 37), présentant « le champ de la didactique et ses rapports avec les disciplines d'intervention » complète utilement ce tableau, en distinguant quatre grands domaines : les sciences du langage, les sciences psychologiques, les sciences de l'éducation et les sciences sociales, humaines et de la communication.

<sup>5</sup> Il convient ici de rappeler que la notion de *compétence linguistique* a été introduite par Chomsky, et que celle-ci a ensuite été incluse par Hymes dans celle, plus large, de *compétence de communication*.

<b>De la compétence linguistique à la compétence communicative</b>	<b>Micro-compétences</b>	<b>Champs théoriques</b>
	<p>audio-phonatoire</p> <p>phonologique sémantique lexicale morpho-syntaxique</p>	<p>Phonétique acoustique Phonétique articulatoire Phonétique perceptuelle Phonologie Sémantique Lexicologie Théories de la grammaire</p>
<p>supralinguistique paralinguistique</p>	<p>rythmico-mélodique phonostylistique</p>	<p>Phonétique suprasegmentale Phonostylistique</p>
<p>extralinguistique</p>	<p>kinésique proxémique</p>	<p>Kinésique Proxémique</p>
	<p>socio-culturelle</p>	<p>Anthropologie culturelle Sociolinguistique Ethnographie de la communication</p>
	<p>pragmatico-énonciative interactionnelle stylistique dialectique</p>	<p>Pragmalinguistique Ethnométhodologie Stylistique collective Dialectique comparée</p>
<p><i>compétence stratégique</i></p>	<p><i>didactique</i></p>	

La notion de compétence de communication a ainsi subi plusieurs changements au cours des dernières décennies, et le développement récent des Nouvelles Technologies de l'Information et de la Communication (NTIC) dans le domaine didactique augure de nouveaux ajouts : comme l'indique Chapelle (2001, p. 1), « Forward-looking members of the profession have suggested that the nature of communicative competence has changed in a world where communication occurs with computers and with other people through the use of computers ». Il conviendrait donc, selon certains auteurs (Jarvis, 2000), de rajouter également une « composante NTIC » à la compétence de communication.

Théorie et pratique, tels sont les deux pôles de la DLC, un domaine qui, d'une part vise à accroître les connaissances relatives aux phénomènes d'enseignement/apprentissage des langues, et, d'autre part, à investir ces savoirs dans une pratique de l'enseignement qui puisse faciliter cet apprentissage. Les questions qui se posent aux chercheurs et aux praticiens sont les suivantes :

- |                              |                              |
|------------------------------|------------------------------|
| 1. Qui apprend ?             | A. Qui enseigne ?            |
| 2. Quoi apprendre ?          | B. Quoi enseigner ?          |
| 3. Comment apprendre ?       | C. Comment enseigner ?       |
| 4. Pour (-) quoi apprendre ? | D. Pour (-) quoi enseigner ? |

Ces deux séries peuvent être condensées en trois questions fondamentales :

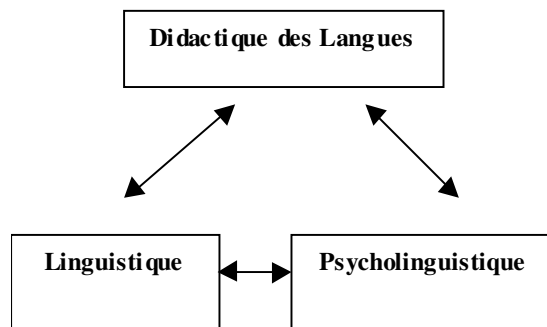
- 1) Qu'apprend-on (objet d'apprentissage/ comportement à apprendre) ?
- 2) Comment l'apprend-on (sujet apprenant/ mode d'apprentissage) ?
- 3) Comment peut-on améliorer cet apprentissage (mode d'enseignement/ procédures et outils) ?

Tandis que la première semblerait trouver réponse dans le champ de la linguistique, et la seconde dans celui de la psycholinguistique, il ne faudrait pas croire que le domaine de la didactique des langues serait celui auquel renverrait exclusivement la troisième, car même si c'est à celui-ci qu'incomberait la tâche d'y répondre, une telle répartition

est à bien des égards abusivement simplificatrice. Dans la perspective qui est la nôtre, il semble en effet nécessaire de hiérarchiser le questionnement : la réponse au « Comment apprendre » nécessite de disposer d'informations suffisantes sur le « Quoi apprendre », tandis que le « Comment enseigner » implique d'avoir répondu autant que possible aux deux premières questions.

Il ne faudrait cependant pas croire à une hiérarchisation linéaire, en termes de dominance conceptuelle et méthodologique stricte, qui nous renverrait aux anciennes querelles entre didactique des langues et linguistique appliquée (Béacco et Chevalier, 1988 ; Py, 1988 ; Vivès, 1988 ; Pothier, 2003 ; Cuq et *al.*, 2003). Il s'agit aujourd'hui bien moins de « dominance », conceptuelle ou méthodologique, que d'information et d'interactions entre les trois domaines concernés, puisque l'on sait, par exemple, que le mode d'enseignement peut influencer, dans une certaine mesure, le mode d'apprentissage et que l'objet effectivement appris résulte d'interactions entre l'objet à apprendre et les deux variables précédentes. Cette approche est d'ailleurs parfaitement illustrée par la multiplication de travaux que l'on pourrait facilement qualifier d'interdisciplinaires, combinant linguistique et psycholinguistique (par exemple : rôle de la perception en phonologie, Hume et Johnson, 2001), psycholinguistique et didactique des langues (par exemple : rôle des supports iconiques en compréhension orale, Merlet et Gaona'h, 1995) ou encore didactique des langues et linguistique (par exemple : apprentissage de la phonologie d'une langue étrangère, Peperkamp, 2002 ; Escudero et Boersma, 2004 ), l'objectif étant de bénéficier des apports respectifs des connaissances et des pratiques des acteurs des domaines en question.





Il serait bien évidemment grossier de traiter de didactique des langues sans mentionner le terme de « culture », étant donné le rôle que peuvent jouer certains facteurs contextuels non-linguistiques d'ordre socioculturel dans celle-ci. Mais puisqu'il convient, comme dans tout travail scientifique, d'effectuer des choix, notre travail se situe donc explicitement au centre de ce triangle disciplinaire, axé sur la langue et la parole.

Si dire que la didactique des langues est un domaine pluridisciplinaire est aujourd'hui devenu une banalité, la mise en pratique d'une telle affirmation n'est restée pas moins complexe et, encore à ce jour nous semble-t-il, inachevée. L'intégration, et non l'imposition, tant dans les données que dans les démarches, de la linguistique et de la psycholinguistique à la didactique des langues semble cependant devenir une nécessité, au fur et mesure de leurs développements respectifs, lesquels suivent l'évolution générale et rapide des sciences cognitives dans leur ensemble.

La didactique des langues se doit en effet de tirer parti des avancées effectuées dans les divers domaines la concernant : perception de l'oral, perception de l'écrit, fonctionnement du système mnésique, gestion de la multimodalité et organisation du lexique mental, pour ne citer que quelques-uns des grands axes de recherche, dans l'un des domaines centraux de ces sciences, à savoir la psycholinguistique, une psycholinguistique qui, comme l'a maintes fois affirmé Jean-Luc Nespoulous, ne peut

être menée avec rigueur qu'à condition de disposer d'une formation linguistique suffisamment solide (Nespoulous, 1994).

Or, ni les linguistes, ni les psycholinguistes n'ont pour vocation de réinvestir leurs savoirs et leurs découvertes dans le domaine didactique. C'est bien aux didacticiens que revient la tâche de s'informer, en vue, non « d'importer des produits préfabriqués », comme on a pu l'affirmer à propos de concepts et de théories linguistiques parfois malhabilement injectés dans le corps didactique, mais plutôt de mettre en œuvre leur propre paradigme, fonction des objectifs qu'ils se sont assignés. Car s'il nous semble indispensable de mettre à profit les diverses avancées des sciences cognitives, la pratique de l'enseignement du français langue étrangère (FLE), tant en milieu homoglotte qu'hétéroglotte, ne laisse aucune illusion quant aux écarts qui peuvent séparer les situations expérimentales de laboratoire des situations d'enseignement de terrain, à une époque où l'influence du contexte et des tâches apparaît comme déterminante dans les résultats obtenus. C'est précisément pour cela que la didactique des langues ne peut plus souffrir d'applicationnisme, mais bien au contraire adopter une approche implicationniste. Il ne s'agit cependant pas « d'impliquer la linguistique plutôt que de l'appliquer », comme cela a pu être proposé par Roulet (Pothier, 2003, p. 12), puisque nous pensons, comme Pothier (*ibid.*), que la polémique entre didactique et linguistique est « à la fois non productive et un peu dépassée ». Bien au contraire, *c'est aux didacticiens de s'impliquer* dans les domaines connexes, s'informant et se formant *en vue d'en retirer de possibles implications*, tant dans les données que dans les pratiques, pour leurs propres problématiques. Combiner et faire interagir ces informations aux expériences et observations d'enseignement/apprentissage de terrain, dans une perspective résolument didactique, telle est la démarche qui semble à ce jour la plus à même de faire progresser la didactique des langues sur le chemin qui mène à la réponse de notre troisième question fondamentale. C'est donc bien de formation qu'il s'agit, formation des futurs enseignants et des futurs chercheurs du domaine.

Connaître la nature et le fonctionnement des processus d'apprentissage en vue d'y adapter les outils et les modes d'enseignement, tel semble ainsi être l'un des enjeux

principaux de la didactique des langues de demain, une didactique qui dès lors ne peut être autre que foncièrement cognitive (Detey, Durand, Nespoulous, 2005). Cette didactique « cognitive » n'est donc certainement pas vouée à un « néo-applicationnisme », mais semble au contraire pouvoir adopter une approche que nous avons qualifiée d'implicationniste. Plus qu'une mise à l'épreuve du contexte, c'est bien un réinvestissement des données des sciences cognitives dans une démarche d'expérimentation de terrain, linguistiquement fondée (Nespoulous, 1994), psycholinguistiquement informée mais *foncièrement didactiquement orientée*, qui semble alors nécessaire.

Défendre l'intégrité et l'autonomie de la DLC, tout en tirant parti des données des champs connexes n'est donc pas une position incohérente, bien au contraire, comme l'illustre la perspective de Pothier (2003, p. 13) à propos de la linguistique :

*« La linguistique et la didactique n'ont pas exactement le même objet, ni les mêmes objectifs [...]. Il est clair cependant qu'une description fiable et approfondie des langues est sans conteste un besoin absolu pour l'enseignant, à charge pour lui de la comprendre et de l'utiliser à sa manière ».*

Il ne fait donc nul doute que la linguistique par exemple, pour ne nommer qu'elle, a toujours un rôle à jouer, dans une certaine mesure au moins, dans la formation des enseignants, des concepteurs de programmes et surtout des chercheurs-didacticiens. Encore faut-il préciser la linguistique dont il s'agit. Juff (2002, p. 87) présente la linguistique formelle<sup>6</sup> comme telle :

*« A theory of the capacity underlying the grammars of natural languages that meets the requirements of "explanatory adequacy" [...]. A formal grammar is an explicit description of a speaker's knowledge of his or her language(s) ; this*

---

<sup>6</sup> Et l'on se doit ici de faire référence aux travaux fondateurs de Chomsky, notamment vis-à-vis des notions d'adéquation (« adequacy ») observationnelle, descriptive et explicative (nous remercions J. Durand à ce propos).

*means that all the properties are specified fully and precisely as a system of operations on linguistic categories [...]. The grammar will achieve explanatory adequacy if linguistic expressions can be derived from the grammar and if the theory accounts for how the grammar might have arisen in the mind of the speaker ».*

Pour l'auteur, les deux grandes questions auxquelles la linguistique formelle essaie de répondre sont les suivantes : « What does it mean to say we "know" a language? » et « How does that knowledge arise in the mind of the speaker; that is, how is it acquired? ». Dans son appréciation du rôle que la linguistique formelle a à jouer dans le domaine de l'acquisition des langues secondes, l'auteur adopte la position de Lightbown et White auxquels il se réfère. Pour eux, les théories formelles sont utiles pour l'un des sous-domaines du champ : « the development and ultimate attainment of linguistic competence » (p. 88), la compétence linguistique étant entendue dans une acception chomskyenne, comme la connaissance implicite que possède chaque locuteur des propriétés abstraites de sa langue. Elle ne constitue donc qu'une sous-composante de ce qu'on appelle la compétence communicative, et ne prétend pas englober les aspects pragmatiques et socio-linguistiques, pourtant essentiels à l'utilisation de la langue.

Juff (2002, p. 88) évoque ainsi plusieurs points sur lesquels la linguistique formelle a contribué au développement du domaine de l'Acquisition des Langues Secondes (SLA) :

- Description détaillée des principes sous-tendant les langues humaines, fournissant ainsi des cadres théoriques, falsifiables, conduisant à un ensemble de, et nécessaire aux, expérimentations empiriques consécutives.
- Questionnement des chercheurs suscité par et à travers ces cadres théoriques.
- Développement méthodologique pour la recherche et développement de matériel pédagogique.

En se concentrant sur certains aspects morphosyntaxiques, dans une approche en principes et paramètres (par exemple le temps (*tense*) et l'ordre des mots (*word order*)), Juff montre que ces théories formelles dotent le chercheur en SLA d'une description détaillée de structures qui, apparemment différentes entre, et à l'intérieur des langues, se réduisent à des catégories (avec les opérations associées) communes similaires, caractérisées par des valeurs binaires (par exemple la catégorie fonctionnelle « TEMPS » (TENSE) à laquelle on attribue la caractéristique [+/- STRONG] selon les langues (« + » en français et allemand, « - » en anglais et chinois, par exemple)) : « The grammar is *learnable* based on pre-existing knowledge of the categories provided by UG and the binary values of the categories » (p. 92). Ces travaux permettent ainsi de poser certaines questions (par exemple : « Do strength values of [functional categories] transfer from the L1 to the L2 ? » ou « Is there an advantage for native speakers of a language with a strong or a weak functional category system ? » (pp. 92-93)) plus précises que d'autres (« Are second language learners able to acquire tense and agreement ? » (p. 93)).

De la même manière, les modèles phonologiques sont utiles pour aborder le problème de l'apprentissage des « nouveaux » sons : « Some languages employ certain segments as part of the system of contrasts (i.e., the segments are phonemes), whereas these same sounds in other languages are merely phonetic variants of an underlying phoneme (allophones) » (p. 99). C'est donc au moins pour le travail de fond que la linguistique formelle peut présenter un réel intérêt pour le chercheur-didacticien (p. 102) :

*« Without [formal linguistics], researchers cannot hope to ask sophisticated questions about what it means to know a second language. [...] While the results of this research often do not have direct pedagogical applications, the results can inform an understanding of the process of second language acquisition and provide a background for teachers to understand the progress or lack of progress learners make in the classroom ».*

C'est la raison pour laquelle il semble indispensable, lorsque l'on s'intéresse à l'enseignement/apprentissage de la parole en LE, de disposer d'une formation, au moins minimale, tant en linguistique qu'en psycholinguistique. Dans notre cas, c'est essentiellement à la phonétique et à la phonologie que nous devons ici recourir en ce qui concerne le domaine linguistique.

## **I.2. La phonologie**

En ce qui concerne la phonologie, on peut définir cette discipline (en première approximation) comme la composante linguistique qui traite des phonèmes et de leurs combinaisons possibles (Durand, 1990, p. 4). D'un point de vue relativement traditionnel, elle constitue l'un des sous-domaines de la linguistique formelle, définie par Archangeli (1997, p. 2) comme suit<sup>7</sup> :

*« L'étude, d'une part des universaux langagiers, c'est-à-dire l'étendue et le type des propriétés partagées, d'une manière ou d'une autre, par toutes les langues, et d'autre part de la variation langagière, c'est-à-dire l'étendue et le type des variations possibles entre les langues ».*

L'auteur rappelle également que les quatre branches de la linguistique formelle sont (p.4) :

- La phonologie : L'étude de la manière dont les sons se combinent entre eux pour former des morphèmes et des mots.
- La morphologie : L'étude de la manière dont les morphèmes se combinent entre eux pour former des mots.
- La syntaxe : L'étude de la manière dont les mots se combinent entre eux pour former des phrases.
- La sémantique : L'étude de la manière dont les significations (« meanings ») des sous-parties se combinent entre elles pour rendre la signification du tout.

---

<sup>7</sup> Notre traduction.

Bien évidemment, ces quatre niveaux d'analyse manquent de refléter ici les interactions incessantes qui ont lieu au cours des activités langagières *in vivo*. Néanmoins, c'est principalement sur la base de cette division que s'est établie la linguistique moderne.

Dans l'histoire de la phonologie (Durand et Laks, 2002b ; Durand et Lyche, 2000), on peut distinguer trois grandes périodes, qui s'articulent autour de l'un des ouvrages de référence du domaine, celui de Chomsky et Halle, *The Sound Pattern of English* (SPE):

- ⇒ Pré-SPE : la période structuraliste ;
- ⇒ SPE : la période générative-transformationnelle ;
- ⇒ Post-SPE : éclatement des approches et (ré-) émergence de nouvelles perspectives (contraintes et connexionnisme).

L'ouvrage de Durand (1990), permet d'obtenir un panorama à la fois diachronique et synchronique des concepts et problématiques essentiels de la phonologie : phonèmes, allophones, phones, phonotactique, distinction entre les niveaux phonétique et phonémique, paires minimales, distribution complémentaire, variation libre et similarité phonétique en font partie. L'une des questions centrales est la suivante : « What are the primitive building blocks of sound structure ? » (p. 12) Cela renvoie à la distinction entre phonèmes (tradition structuraliste) et traits (travaux de Jakobson, entre autres). Comme l'écrit Durand (1990, p. 12) : « a traditional answer to this question is that phonemes are the ultimate units of phonological structure, *i.e.* the elements that keep morphemes apart from one another ». Cependant, si l'on choisit d'adopter le concept de trait, il convient de bien distinguer les niveaux phonétique et phonologique : « it is crucial to establish a fundamental distinction between phonological and phonetic features. [...] Not every phonetic characteristics of a segment will be relevant to differentiate it from the other segments of the language » (p. 15). Néanmoins, la question du rapport entre les deux niveaux reste délicate et controversée. Dans le cadre de la théorie des traits distinctifs, un segment est donc défini comme « a set of simultaneous Distinctive Features (DFs) » (p. 37), ce qui peut être représenté de la

manière suivante pour le morphème anglais « at » (exemple tiré de Durand, 1990, p. 37) :

	æ	t
<b>syllabic</b>	+	-
<b>consonantal</b>	-	+
<b>nasal</b>	-	-
<b>coronal</b>	-	+
<b>sonorant</b>	+	-
...		

La représentation des traits est ici binaire, mais il convient de continuellement distinguer les niveaux (Durand, 1990, p. 40) :

*« The features will be assumed here to be binary at the **phonological** level. [...] Of course, the postulation of binariness at the classificatory (or contrastive) level does not mean that, **phonetically**, features have to be binary. For instance, nasality is not distinctive for vowels in English, but a vowel may exhibit a small amount of nasality either inherently or under the influence of a neighbouring nasal consonant. In this case, the features are functioning as scales ».*

Durand présente (p. 38) plusieurs arguments avancés par la phonologie générative en faveur de la reconnaissance des traits distinctifs, et non des phonèmes, comme primitives phonologiques (notamment leur caractère universel (et non spécifique à une langue particulière comme le sont les phonèmes) et leur autonomie phonologique qui peut ainsi les conduire à dépasser l'unité phonémique, par exemple dans le cas de l'harmonie vocalique ou de la diffusion tonale). Insistons bien sur le fait que nous nous situons ici à un niveau phonologique.



En ce qui concerne les aspects articulatoires et/ou acoustiques des traits, il faut indiquer que la dimension auditive a toujours eu ses défenseurs dans l'approche des DFs, notamment Jakobson, Fant et Halle (*Preliminaries (to Speech Analysis)*, 1952), pour lesquels, comme l'indique Durand (p. 61), « the emphasis was placed on the receptive side of the communication circuit ». Pour Jakobson en particulier, commente Durand (p. 62), « there is an imbalance between production and perception : invariance for them [Jakobson and Waugh] is to be found at the acoustic and not the articulatory level ». La primauté de la dimension auditive-acoustique est notamment illustrée par la spécification des voyelles dans l'espace vocalique (p. 64) pour lesquelles, d'un point de vue articulatoire, la variabilité est importante et attestée par de nombreuses études, ou bien encore par le « r rétroflexe » (Durand, p. 67) :

*« On the basis of X-rays and cineradiography, it can be shown that two very dissimilar articulations can be used to produce this /r/. [...] As Delattre (1967 : 22) argues, this is a striking « case of different articulatory patterns which can produce similar acoustic patterns or, to express it differently, of differences in articulation which do not result in significant differences in spectrograms ». He notes further that auditory perception matches the spectrographic evidence since listeners cannot detect whether the /r/ is bunched or retroflex ».*

Cependant, Durand (p. 68) indique que le problème de l'invariance remet en question cette approche, notamment à travers la théorie motrice de la perception pour laquelle les auditeurs décodent les schémas acoustiques sur la base des gestes articulatoires utilisés pour la production de parole. Durand conclut que (p. 69) :

*« From the point of view of phonology, the best position is that advocated by Delattre (1967) and Halle (1983). Distinctive features are neutral between production and perception. They must not be confused with their actual phonetic implementation. They are abstract, classificatory dimensions which provide a link between the articulatory and the acoustic dimension of speech. In*

*some case, the tie between a distinctive feature and its typical implementation is more direct on the articulatory side and in other cases on the acoustic side ».*

Depuis, même si le débat persiste, d'autres problématiques ont vu le jour, et la phonologie s'est par ailleurs de plus en plus diversifiée : on constate en effet une parcellisation accrue des sous-domaines traditionnels de la linguistique, en fonction des approches théoriques et méthodologiques (par exemple pour la phonologie : phonologie auto-segmentale, phonologie naturelle, phonologie de laboratoire, etc. (Durand et Laks, 2002a)), mais aussi des nouveaux outils (linguistique de corpus, modélisations connexionistes, par exemple).

Cependant, pour Goldsmith (1995), la plupart des travaux en phonologie actuelle touchent trois questions sous-jacentes essentielles (p. 1) :

- Celle de la *phonotactique*, c'est-à-dire l'ensemble des conditions de bonne formation des mots phonologiques dans une langue donnée ;
- Celle des *alternances* qui renvoie à la morphophonologie, puisqu'il s'agit des différentes réalisations d'un même morphème dans des contextes, phonologiques et/ou morphologiques différents ;
- Celle enfin de la *contrastivité*<sup>8</sup> qui se situe à la limite de la phonologie puisqu'elle s'interroge sur la valeur contrastive de différences phonétiques dans une langue donnée.

Pour tenter de répondre à ces questions, trois grandes catégories d'outils sont à la disposition des linguistes, qui sont les suivants (p. 2) : des représentations, des niveaux et des règles phonologiques. Goldsmith propose alors le tableau suivant (p. 3) qui, d'une certaine manière, synthétise les problématiques phonologiques de l'époque :

---

<sup>8</sup> J. Durand nous rappelle utilement (communication personnelle) que, dans la tradition française, Martinet opposait l'*opposition* (paradigmatique) au *contraste* (syntagmatique) et que le terme « contrastivité » emprunté à l'anglais renvoie ici à l'*opposition* française.

	<b>Phonotactics</b>	<b>Alternations</b>	<b>Contrasts</b>
<b>Representations</b>	Autosegmental and metrical structure Prosodic hierarchy	Feature geometry and limits on kinds of assimilation Prosodic morphology	Underspecification theory
<b>Levels</b>	Licensing Abstractness Structure-preservation	Issues of stratal organization	Organization of the lexicon
<b>Rules</b>	Metrical theory Harmonic rule application Optimality theory	Equation of phonotactics and alternations : strict cyclicity	Structure-building Versus Structure-changing operations

Comme on peut ainsi le voir, les travaux en phonologie ont acquis un degré de complexité et de technicité qui en rend l'approche peu aisée au bétien et qui peut parfois rendre difficile la démarche interdisciplinaire que nous avons tenté d'adopter ici. Nous ne prétendons ainsi certainement pas pouvoir comprendre et maîtriser tous les concepts et cadres théoriques qui y ont été développés, mais plutôt essayer d'en retirer des informations utiles et nécessaires à une approche rigoureuse de la problématique qui nous occupe.

On peut aujourd'hui trouver de nombreux chercheurs/-euses qui possèdent une double formation en phonologie et en psycholinguiste et qui travaillent sur des problématiques se situant aux nombreux croisements des deux disciplines (Escudero et Boersma, 2004). Nous partageons d'ailleurs la vision de Nespoulous qui, à plusieurs reprises, a souligné la nécessité, pour les aspirants psycholinguistes, de disposer d'une formation suffisante en linguistique, formation sans laquelle il est difficile de manipuler scientifiquement du matériau verbal dans le cadre d'études en psycholinguistique (Nespoulous, 1994, 2004). S. Peperkamp fait partie de ces chercheuses/-eurs et elle résume bien (2002), dans une perspective diachronique, la problématique des rapports entre phonologie et

psycholinguistique, puisque ses axes de recherche, par exemple celui du traitement et de l'acquisition phonologique, nécessitent les apports des deux disciplines. Comme elle le note, l'écart séparant la linguistique de la psycholinguistique, tant sur le plan des données empiriques que sur celui des méthodes de recherche (Peperkamp, 2002, p. 3), s'est creusé au cours des trente dernières années, alors qu'il n'en avait pas toujours été ainsi (pp. 4-5). En outre, en raison de la recherche d'un certain *réalisme psychologique* des structures linguistiques, certains, à la suite de Chomsky (cité par Peperkamp, p. 5), considèrent la linguistique comme l'une des branches de la psychologie, comme par exemple Moeschler et Auchlin qui écrivent (1997, p. 11) :

*« Le linguiste ne décrit pas les langues, mais la connaissance que les sujets parlants ont des langues. Dans cette perspective, la linguistique est une science qui appartient de plein droit à ce qu'on appelle la psychologie cognitive, à savoir le domaine de la psychologie qui s'intéresse aux facultés mentales à l'origine des comportements, des pensées et des manifestations langagières ».*

Il est donc temps, comme l'écrit Peperkamp, de « réintégrer la linguistique et la psycholinguistique », afin que chacun puisse bénéficier des résultats obtenus, par exemple, sur la perception/production de la parole d'une part, et sur les processus et structures phonologiques d'autre part. Le cas de l'acquisition phonologique en est représentatif, puisque « des modèles réalistes de l'acquisition phonologique ne peuvent être construits que par une intégration des approches des linguistes et des psycholinguistes » (Peperkamp, 2002, p. 7). Haunz (2002) suit la même ligne d'idée, en déplorant la négligence des chercheurs en phonologie des emprunts vis-à-vis des travaux sur la perception de la parole en langue seconde.

Dans la perspective didactique qui est la nôtre, si une formation linguistique suffisante semble indispensable, le domaine crucial sur lequel repose la DLC est bien celui de la psycholinguistique, puisque c'est lui qui traite de la perception, de la production et de l'apprentissage du matériau langagier.

### **I.3. La psycholinguistique**

La psycholinguistique pourrait à première vue sembler faire partie de ce que l'on appelle les *linguistiques à traits d'union*. Pourtant, comme cela a été indiqué plus haut, l'étude du langage (la partie « linguistique ») semble difficile à réaliser sans locuteurs, lesquels sont des êtres fondamentalement psychiques<sup>9</sup> (la partie « psycho- »). Cela peut nous amener à poser le problème de la distinction esprit/cerveau<sup>10</sup>, mais, surtout, cela nous laisse penser que si l'on souhaite établir le lien entre, d'un côté des codes sémiotiques abstraits, et de l'autre leur *mise en réel* (pour ne pas dire « mise en action ») la psycholinguistique apparaît comme une discipline non pas satellite mais bien centrale dans l'étude du langage. Elle apparaît d'ailleurs comme l'étape intermédiaire (et ainsi le lien indispensable) entre l'étude des structures linguistiques abstraites et celle du substrat biologique, en particulier cérébral (Nespoulous, 2004 ; Nespoulous et Leclercq, 1990), puisque, comme l'écrit Nespoulous (2004, p. 2) :

*« La linguistique spécifie les propriétés structurales de telle ou telle langue naturelle à chacun de ses niveaux d'organisation : phonologique, morphologique, syntaxique...Son objet est donc l'identification de l'« architecture structurale » des langues et, à travers ces dernières et leur diversité, de l'architecture structurale du langage dans l'espèce humaine.*

*La psycholinguistique, de son côté, s'assigne pour objectif de caractériser les niveaux de représentation et les processus cognitifs présidant au traitement par l'esprit humain – si possible en temps réel – des dites structures linguistiques, et ce en production comme en compréhension, à l'oral comme à l'écrit. Son objet est donc de spécifier l'« architecture fonctionnelle » sous-jacente du langage.*

---

<sup>9</sup> On pourrait sans doute nous rétorquer que ce sont plutôt des êtres fondamentalement sociaux, mais cela nous plongerait dans un débat dépassant notre cadre de travail.

<sup>10</sup> Il est d'ailleurs intéressant de noter qu'en japonais deux termes coexistent pour désigner la « psycholinguistique » : « There are actually two interrelated usages in Japanese, *Gengoshinrigaku* and *Shinrigengogaku*, which could be discriminated when translating into English. The first might be better translated as 'the psychology of language', since most of these scholars are psychologists and not as focused on linguistic knowledge as defined by the field of linguistics » (Kess et Miyamoto 1999, p. 5, note 12).

*La neuropsycholinguistique, pour sa part, tente, encore fort modestement, de « réconcilier l'esprit et le corps » (= le cerveau), dans un effort d'identification des structures cérébrales ou réseaux neuronaux mobilisé(e)s lors du traitement cognitif de telle ou telle composante de « l'architecture fonctionnelle » du langage dans telle ou telle activité langagière. Son objet est donc d'appréhender l'« architecture cérébrale », neuronale, du langage ».*

La psycholinguistique représente ainsi pour nous un sous-domaine incontournable des sciences cognitives relatives au langage, et ceci plus particulièrement lorsque l'on en vient au problème des apprentissages linguistiques. Que l'on traite d'apprentissage ou d'acquisition, pour reprendre la célèbre distinction de Krashen, c'est bien le domaine psycholinguistique qui est chargé de rendre compte des phénomènes d'apprentissage, *via* les multiples étapes perceptives, mnésique et autres que l'apprenant doit suivre.

Sur ce point, nous devons admettre que le domaine anglophone SLA semble avoir *a priori* réalisé une synthèse disciplinaire davantage réussie entre didactique des langues et psycholinguistique, alors que la traduction française de « SLA » oscille toujours avec hésitation entre « Didactique des langues » et « Psycholinguistique appliquée à l'enseignement/apprentissage des langues ». Si l'on peut regretter la connotation applicationniste de la dernière formulation, on ne peut que s'interroger sur les raisons de cette hésitation dans la traduction et soulever la nécessité d'intégrer systématiquement une dimension psycholinguistique à toute problématique d'apprentissage linguistique (et nous pourrions très facilement ajouter « et culturel »).

#### **I.4. Intégration et transdisciplinarité**

Intégrer les apports de la phonologie et de la psycholinguistique, afin de mieux comprendre le traitement de la parole en multimodalité par des sujets apprenants, de manière à pouvoir en aval adopter une démarche didactique (conception et enseignement) propice à l'acquisition du système phonético-phonologique du français,

tel est donc l'un des objectifs essentiels de notre travail. Cette approche transdisciplinaire<sup>11</sup>, à orientation triple, si elle s'efforce d'être novatrice et intégratrice, présente évidemment l'inconvénient de ne laisser suffisamment ni d'espace ni de temps pour un approfondissement parfaitement expert de chaque domaine. L'expression est bien connue : « on ne peut être spécialiste de tout... ». Il semble donc utile d'insister sur le fait que la présente étude relève du domaine de la didactique des langues étrangères. A l'image du parcours en trois étapes proposé par Nespoulous (2004), de la linguistique à la neuropsycholinguistique, nous avons donc adopté la perspective suivante dans notre étude :

**Didactique →            Linguistique →            Psycholinguistique →            Didactique**  
(Horizon didactique)            (Horizon didactique)

L'important, nous semble-t-il, est d'avoir initialement un horizon didactique clairement déterminé, ainsi que la formation plurielle appropriée, sans laquelle les questionnements se trouvent trop fractionnés, puis impossible à recomposer de manière cohérente et réaliste. Les deux « thèmes didactiques » qui ont constitué la base de ce travail sont les suivants : a) utilisation didactique de la multimodalité et b) enseignement/apprentissage de la « prononciation » du français.

---

<sup>11</sup> On pourrait discuter de la pertinence et de la signification du terme « transdisciplinaire », en l'opposant à « interdisciplinaire » et « pluridisciplinaire ». Si l'on considère que l'« interdisciplinaire » renvoie à une véritable collaboration et un questionnement partagé entre spécialistes de différentes disciplines, alors que le « pluridisciplinaire » ne renvoie qu'à un questionnement pluriel, en provenance de différentes disciplines, sur un même objet (par exemple), le « transdisciplinaire » s'en distinguerait alors par son caractère « transversal » : notre approche « transdisciplinaire » est issue d'une *unique* discipline (la didactique), mais la dépasse, puisque nous allons puiser (des informations, des démarches, des réflexions) dans d'autres disciplines, tout en gardant notre propre perspective disciplinaire. Si notre objet d'étude peut donc être considéré comme interdisciplinaire ou pluridisciplinaire, notre approche semble en revanche transdisciplinaire, en ce qu'elle nous conduit à partir d'une discipline et à en dépasser les frontières, tout en gardant une perspective unique (et non plurielle) sur l'objet en question, dépassant ainsi (d'où le préfixe « trans -») chacune des disciplines en question (les données et les démarches pour la première, les perspectives pour les suivantes). Il va de soi que l'on pourrait aussi bien réduire la question à un problème terminologique et que l'on pourrait facilement substituer interdisciplinaire à transdisciplinaire.

## **II. Multimédia et multimodalité**

Depuis une vingtaine d'années environ, les avancées technologiques en matière de traitement de l'information ont été considérables, au point de s'intégrer dans la plupart des secteurs de la vie quotidienne. Dans le domaine de l'éducation, notamment linguistique, l'introduction, encore récente, des outils dits « hypermédias » semble très prometteuse pour l'avenir, puisque ces outils permettent certaines manipulations ergonomiques, spatiales et/ou temporelles, de l'input visuel et auditif impraticables auparavant. Par « hypermédia » on entend habituellement la « conjonction du multimédia et d'un ensemble de liens hypertextes » (Bruillard et De la Passardière, 1998, p. 97)<sup>12</sup>.

C'est alors vis-à-vis de la problématique de la multimodalité que ces outils trouvent leur place dans notre réflexion, puisque l'un des atouts didactiques qu'ils affirment offrir au public, enseignant et apprenant, est précisément celui de la multimédiatisation de l'enseignement/apprentissage. Il convient donc de savoir dans quelle mesure l'agencement des différents médias constitue effectivement un atout pour l'apprentissage, dans notre cas celui de l'oral en FLE : de quelle manière la manipulation d'informations multimodales doit être conduite à l'aide de ces outils, si l'on souhaite optimiser l'apprentissage d'une langue étrangère, et plus particulièrement celui du système phonético-phonologique de cette langue ? Cette question n'est pas négligeable, car l'ordinateur est appelé à devenir, tôt ou tard, l'outil de formation par excellence, omniprésent et banalisé, de l'enseignement primaire au supérieur.

### **II.1. Domaine et historique**

Autour de ces outils d'apprentissage s'est construit un véritable domaine de recherche, celui de l'Enseignement/ Apprentissage des Langues Assisté par Ordinateur (E/ALAO), qui est l'équivalent du *Computer Assisted Language Learning* (CALL) de la sphère

---

<sup>12</sup> « La technique de l'hypertexte permet de mettre en rapport un premier texte, présent à l'écran, avec d'autres textes qui sont appelés et apparaissent alors à leur tour sur ce même écran » (Lancien, 1998, p. 20).



anglophone. Comme l'écrit Chapelle, durant les années 1990 la question « Doit-on utiliser l'ordinateur en classe de langue ? » s'est progressivement transformée en « Quelle est la meilleure manière d'utiliser l'ordinateur en classe de langue ? ». La question de son utilisation ou de sa non-utilisation semble donc ne plus se poser (Chapelle, 2001, p. 1). Il faut aussi noter que ce secteur est en constant développement, un développement accéléré, essentiellement conduit par l'innovation technique.

Les atouts de ces outils sont nombreux : comme l'indiquent Warschauer et Healey (1998, p. 4), « *the benefits of adding a computer to language instruction are many, and include :*

1. *Multimodal practice and feedback,*
2. *Individualization in a large class,*
3. *Pair and small group work on projects, either collaboratively or competitively,*
4. *The fun factor,*
5. *Variety in the resources available and learning styles used,*
6. *Exploratory learning with large amounts of language data,*
7. *Real-life skill-building in computer use ».*

Il alors est intéressant de noter ici que la « pratique multimodale » figure en tête de liste.

Dans un ouvrage de synthèse bien connu dans le domaine, Levy (1997) passe en revue une grande partie de la littérature disponible sur le domaine, avant de rendre compte d'une enquête menée auprès d'enseignants de langue sur leur approche de l'ALAO. L'enquête, qui date de 1990, abordait de nombreux aspects, du rôle de l'ordinateur (tuteur/outil) à la conception de l'enseignement en passant par la question de l'autonomie. Les réponses sont très variées et il est difficile d'en tirer des grandes lignes. Comme l'auteur le note lui-même : « When CALL is viewed as a whole, as a body of work, the diversity is remarkable » (p. 151), notamment en raison des différents cadres et sources conceptuels adoptés par les développeurs, différences qui naturellement entraînent un problème terminologique. L'ordinateur est un outil technique qui peut se décomposer en *hardware* et *software*. Le *hard* renvoie aux

techniques de réalisation et à leur progression : de plus en plus de miniaturisation et d'intégration de différents modules à l'intérieur d'un outil unique (webcams, vidéo-casque, gants de données, etc.), de plus en plus de numérisation, d'accélération du traitement des données, ainsi que d'amélioration globale de la qualité (finesse) et de la quantité (capacités de stockage) d'informations. Le *soft* renvoie à la « matière grise », à la conception des logiciels, tant du point de vue des contenus que des formes (ergonomie), qui progresse, elle aussi, tirant parti des résultats et des évaluations des produits précédents, malgré le manque de recul. Il faut aussi remarquer que l'ordinateur se décline de moins en moins au singulier : nous vivons à l'ère des réseaux : Internet ou Intranet, laboratoire multimédia, avec postes reliés entre eux ainsi qu'à un projecteur et à un poste central (celui de l'enseignant). Ces développements, techniques ou ergonomiques, concernent de très près les modalités de présentation et de distribution (par exemple temporelle) de l'information.

Si l'on suit Warschauer (1996), on peut distinguer trois phases dans l'histoire du CALL : le CALL *béavioriste*, le CALL *communicatif* et enfin le CALL *intégratif*. Naturellement, ces trois phases ne se sont pas enchaînées les unes à la suite des autres et les trois coexistent ainsi encore de nos jours à des degrés divers, correspondant à la fois à un certain niveau technologique mais aussi à une certaine approche pédagogique. Le CALL *béavioriste* (50s-60s) va naturellement de pair avec les modèles de conditionnement béavioristes en psychologie et avec la méthodologie audio-orale en ce qui concerne l'approche didactique. L'ordinateur servait alors de *tuteur* ou de *répétiteur* pour que l'apprenant puisse effectuer des batteries d'exercices structuraux (les fameux « drills »). Le passage de cette phase à la suivante a été déclenché, comme toujours, d'une part par l'évolution méthodologique (abandon du béaviorisme), et d'autre part par les avancées techniques (nouvelles possibilités techniques). Le CALL *communicatif* (70s-80s) pourrait se caractériser ainsi<sup>13</sup> :

---

<sup>13</sup> J. Underwood, *Linguistics, computers, and the language teacher : a communicative approach*, Rowley, MA, Newbury House, 1984, p. 52, cité par Warschauer, 1996, p. 2.

*« Focuses more on using forms rather than on the forms themselves ; teaches grammar implicitly rather than explicitly ; allows and encourages students to generate original utterances rather than just manipulate prefabricated language ; does not judge and evaluate everything the students nor reward them with congratulatory messages, lights or bells ; avoid telling students they are wrong and is flexible to a variety of students responses ; uses the target language exclusively and creates an environment in which using the target language feels natural, both on and off the screen ; will never try to do anything that a book can do just as well ».*

Avec le déclin de l'approche communicative et le développement de l'approche basée sur les tâches et les projets, sans oublier l'évolution technologique, on arrive au *CALL intégratif*, qui est basé sur deux avancées technologiques : l'environnement multimédia et l'Internet, le multimédia débouchant lui-même sur l'hypermédia. La distinction entre le communicatif et l'intégratif, si elle peut sembler quelque peu floue, semble tenir à la fois aux développements des ressources techniques et de leur utilisation, et aux changements, progressifs, dans les options méthodologiques didactiques. Parmi les avantages de l'hypermédia pour l'apprentissage des langues mentionnés par Warschauer (1996), nous pouvons retenir les suivants :

- Environnement perceptif plus « authentique » (audiovisuel).
- Intégration des compétences (lire, écrire, écouter et parler réunis par le multimédia).
- Individualisation et contrôle par l'étudiant de son propre apprentissage (d'où aussi une responsabilisation) : propre rythme d'apprentissage, mais aussi propre parcours d'apprentissage (allers-retours, écrémage, etc.).
- Enfin l'un des avantages majeurs de l'hypermédia est de pouvoir à la fois se concentrer sur le contenu sans pour autant sacrifier la forme, puisque grâce aux liens, l'apprenant, sans perdre de vue la leçon principale qui se trouve dans l'arrière-plan, a accès instantanément à de nombreuses explications grammaticales, des exercices, des paraphrases, des informations relatives à la prononciation, etc.

Dans les quatre cas, la question du choix de la (ou des) modalité(s) mise(s) en jeu se pose au concepteur et au chercheur-didacticien.

## **II.2. Multimodalité/ multicanalité : un atout pour l'apprentissage ?**

En ce qui concerne la variété des médias, Jarvis (2000) note une certaine évolution : « we are now seeing an integration of text, audio and video material all in one package, whereas in the past everything was essentially "text-based CALL" ». Cela présente des avantages certains (Develotte, 2000, pp. 23-26) :

*« Les caractéristiques constitutives du multimédia multiplient les possibilités d'adaptation aux différents types d'apprentissage. Ainsi la multicanalité permet de toucher, par exemple, aussi bien des apprenants de type « auditif » que ceux qui développent des stratégies plus « visuelles » dans leur démarche d'apprentissage [...]. Outre les effets de mise en page et l'existence de liens hypermédias, le fait que l'information fasse l'objet de codages différents (son, image, texte) engendre de potentialités de reformulation, de simplification, de transposition, potentialités susceptibles de faciliter l'activité de qui s'informe ou apprend ».*

Mais l'intégration de ces médias nécessite de disposer d'un certain équipement (Portine, 1996) :

*« Intégrer le support multimédia à la méthodologie suppose que la classe soit dotée à la fois de moyens dits classiques (tableau, écran pour la projection de films et de diapositives) et de moyens techniques multimédias sous une double forme : « tablettes à cristaux liquides » (ou data show) pour le travail collectif en multimédia, postes de formation multimédias individuels sur chaque table d'étudiant ».*

La question de la multimodalité concerne ainsi à la fois le contenu, mais aussi, sinon surtout, la forme et donc *l'ergonomie* des outils (Tricot et Rouet, 1998). Les

concepteurs/chercheurs s'interrogent donc sur les meilleures combinaisons entre canaux qui permettent d'optimiser l'apprentissage. Ainsi, Lancien (1998, pp. 25-26) à propos de la multicanalité écrit-il :

*« On s'attachera, lors de l'évaluation de la multicanalité dans un produit multimédia, à examiner : la place accordée aux différents canaux ou médias ; la pertinence du recours à un canal plutôt qu'à un autre ; la nature, la pertinence et la richesse des liens entre les médias ; les rapports de sens entre les trois médias ».*

Il ne faut en effet pas ignorer les limites de la multimodalisation, limites contre lesquelles les (neuro-)psycholinguistes nous mettent en garde (Nespoulous, 2000, p. 14) :

*« Si trop peu de stimulations nuisent clairement au développement de l'être humain (cf. les enfants « sauvages »), il convient de veiller à ce que trop de stimulations ne viennent saturer les capacités de traitement de la cognition humaine. [...]. Pour gérer convenablement une pluralité d'informations qui lui sont livrées « en temps réel », l'être humain n'a qu'une solution [...] en perception/compréhension : il doit partiellement négliger une des informations pour concentrer davantage d'attention sur l'autre information qu'il reçoit de manière synchrone. La littérature scientifique en psychologie cognitive foisonne d'exemples dans lesquels le sujet, soumis à une « double tâche », est confronté à un « trade-off » (ou « conflit procédural » dans nos propres termes) qui le conduit à négliger X pour bien appréhender Y ».*

Nous prendrons comme exemple la question du rôle de l'image dans la compréhension orale (Merlet, 1998 ; Bertin, 2001). Comme l'indique Merlet dans une sous-partie intitulée « Discussion générale : les conditions de l'efficacité d'une présentation multimodale », il faut comprendre que la présence de l'image n'est pas systématiquement avantageuse et que « l'efficacité de l'illustration semble en fait

dépendre d'un certain nombre de facteurs ayant trait notamment aux caractéristiques du matériel présenté, mais aussi aux caractéristiques individuelles des apprenants » (Merlet, 1998, pp. 153-154). Insérer des documents imagés dans une méthode de langue ne signifie pas en exploiter les effets, loin s'en faut. Plusieurs questions doivent être posées si l'on souhaite utiliser à bon escient « l'image » en cours de langue. Quel type d'image (analyse de type sémiotique et format), quels rapports (explicatif, illustratif, etc.) entretient-elle avec les autres éléments (verbaux ou pas, visuels ou auditifs) (Develotte, 2000, p. 27), comment est-elle agencée vis-à-vis de ces derniers (d'un point de vue spatial et temporel), quel rôle souhaite-t-on lui faire jouer dans l'apprentissage (aide à la compréhension, indice mnésique, etc.) et enfin quel rôle ne doit-elle pas jouer (parasite attentionnel, surcharge cognitive, source d'ambiguïté)? Ainsi Bertin, lors de la conception d'un CD-ROM d'apprentissage de l'anglais du Transport International et de la Logistique, a décidé de remplacer une photographie par un dessin stylisé pour « une appréhension plus simple et plus rapide du message de la part du message dévolue à l'image » (Bertin, 2001, p. 132). Le choix des canaux de transmission, son, texte, image, ainsi que leur intégration n'est par ailleurs par neutre du point de vue de la méthodologie didactique adoptée. Comme l'écrit Portine (1996) :

*« Concevoir le multimédia comme l'agencement de ces trois canaux de transmission, c'est opérer sur deux types de données de la perception : celles qui permettent de communiquer à l'écrit, celles qui permettent de communiquer à l'oral. Or, de ce point de vue, les méthodologies ont toujours appartenu à deux paradigmes : celui des méthodologies centrées sur la compréhension-production de l'écrit et celui des méthodologies centrées sur la compréhension-production de l'oral ».*

Ce dernier point soulève une question essentielle : si l'avènement des outils multimédia semble certainement avoir donné la possibilité d'augmenter considérablement la quantité de matériel sonore (verbal et non-verbal) et pictographique (images fixes ou mobiles) en classe, ainsi que de les manipuler et combiner facilement, cela ne semble en revanche avoir que peu donné aux didacticiens l'idée de réexaminer la

problématique des combinaisons texte-son (écrit/oral) pour l'apprentissage de l'oral. Pourtant, c'est bien à un questionnement *méthodologique* que devraient conduire ces avancées *techniques*. Il convient en effet de rappeler ce que certains pourraient considérer comme une vérité de Lapalisse : la multimodalité commence avec la bimodalité, et la bimodalité commence avec les formats auditifs et orthographiques. La présente étude vise donc à contribuer à ce travail d'analyse méthodologique en amont, préalable à toute implémentation technique multimodale.

### **III. Enseignement de la prononciation**

Ayant rapidement présenté le thème didactique de la multimodalité, il nous faut à présent préciser le public et le contexte didactique auxquels nous nous intéressons pour l'enseignement/apprentissage de la « prononciation », et plus généralement de l'oral en français.

#### **III.1. Situation actuelle de l'enseignement de la prononciation du FLE au Japon**

Au Japon, la question de l'enseignement de la prononciation du français et la pratique de la phonétique corrective n'est pas neuve (Cortès, 1970) : chaque année peut-on trouver des articles ou des communications portant sur le sujet. Ainsi lors du congrès d'automne 2003 de la Société Japonaise de Langue et Littérature Françaises, dans la catégorie « Enseignement », figurait la communication suivante « Découpe syllabique et rythme en Français : application pédagogique » (Azra, J.-L., 2003). Lors de celui de la Société Japonaise de Didactique du Français, tenu à la même époque, on trouve « La Méthode d'Accord et La Méthode d'Alternance ; Deux nouvelles méthodes d'enseignement de la prononciation » (Wada, Y.). L'année précédente, lors des Journées Pédagogiques annuelles de l'Université Dokkyo en novembre 2002, deux ateliers étaient consacrés à la méthode Verbo-Tonale d'intégration phonétique. Par ailleurs, et contrairement à ce que certains pourraient penser, les enseignants s'efforcent d'y accorder du temps, afin de pratiquer quelques exercices et mettre en œuvre certaines techniques, à commencer par la simple répétition (ce qui n'est pas

rien). En outre, alors que les Japonais ont parfois la réputation d'avoir une « mauvaise prononciation » en langue étrangère (LE), il semblerait que les nouvelles générations d'étudiants, de plus en plus exposées aux LE, notamment grâce aux NTIC, soient capables d'intégrer de nouvelles unités phonético-phonologiques plus facilement que leurs prédécesseurs, ce qui est à relier à la situation socio-phonologique du Japon, avec, par exemple, l'apparition récente de « nouvelles consonnes » (Labrune, 2001a) issues de « l'invasion » des emprunts étrangers, en particuliers d'origine anglo-américaine. Pourtant, malgré les efforts des enseignants et l'évolution des sujets apprenants, les difficultés d'apprentissage persistent, et il est important d'en comprendre les raisons. Selon nous, celles-ci peuvent être rangées essentiellement dans quatre catégories : institutionnelles, socioculturelles, (psycho-)linguistiques et didactiques.

### **III.1.1. Raisons d'ordre institutionnel**

Dans la situation socioéconomique japonaise du début du 21<sup>ème</sup> siècle, l'apprentissage du français, et plus généralement des langues étrangères autres que l'anglais, n'est pas perçu comme une priorité. L'anglais, qui est enseigné durant les études secondaires, se pose en effet dans ce pays comme langue internationale majeure, tant académique que professionnelle. La crise économique du début des années 1990, ainsi que certains développements internes aux pays, ont entraîné un nombre important de réformes du système de l'enseignement supérieur, parmi lesquelles une réévaluation de l'importance des langues étrangères : le français a perdu l'avantage dont il disposait alors, en termes de crédits et de moyens, vis-à-vis d'autres langues étrangères comme l'italien ou le coréen, et les secondes langues étrangères, jusqu'alors obligatoires, sont devenues facultatives. Dans le même temps, un projet de diminution du nombre d'universités a vu le jour, conduisant à la fermeture de certaines, à la fusion d'autres, et en tous cas à la diminution du nombre de classes et d'enseignants. Le nombre d'heures d'enseignement a été ainsi drastiquement réduit, et la fusion des classes a conduit à une augmentation des effectifs par classe, variant entre une trentaine et une cinquantaine—



conditions bien difficiles pour des classes dites de « conversation »<sup>14</sup>. Etant donné les méthodologies en vigueur dans la plupart des universités (en général peu axées sur l'oral), et la nécessité de respecter le programme imposé, il ne reste que très peu de temps et de moyens (laboratoire de langue par exemple, lequel nécessite également un certain volume horaire) à consacrer à la qualité phonético-phonologique des productions orales des étudiants. Le manque de temps et de moyens constitue donc une raison institutionnelle quasi-insurmontable au problème de la prononciation, puisque l'exposition sonore et l'entraînement audio-phonatoire sont largement insuffisants.

### **III.1.2. Raisons d'ordre socioculturel**

Il faut en effet considérer la place occupée par l'expression personnelle (verbale et non-verbale) et plus particulièrement l'expression orale dans les modes de communication (Baylon et Mignot, 1999, pp. 127-138 et pp. 141-148) standard des Japonais. Les travaux consacrés à la communication interculturelle sont innombrables et l'on peut mentionner les travaux pionniers de E.T. Hall par exemple. Dans cette perspective, le Japon a toujours attiré de nombreux chercheurs, que ce soit en raison de leur longue isolation nationale ou de leur poids économique durant la deuxième moitié du 20<sup>ème</sup> siècle. De ces travaux ressortent alors certaines images, plus ou moins stéréotypées, qui circulent à grande échelle. On parle ainsi de « culture du silence », de « communication implicite » ou « minimale », dans laquelle le « non-dit » joue une part considérable et où seules les expressions « subtiles », verbales ou non-verbales, sont appréciées. Il existe par ailleurs des modes de communication bien différents selon l'âge, le rang social (le concept de « classe sociale » au Japon ne saurait être transposé si facilement) et surtout le sexe. On caractérise alors les sujets japonais comme plutôt introvertis (en particulier pour les femmes), et on les oppose au « type latin », qualifié d'extraverti (en particulier pour les hommes). Ce sont là des stéréotypes sociaux à présent considérés comme « standards », qui se répercutent bien évidemment dans le contexte didactique.

---

<sup>14</sup> A ce sujet voir les travaux de Azra et ses collaborateurs, à propos des classes dites de « conversation » avec de grands effectifs.

Il est en effet courant, lorsque l'on évoque les différences de paysages didactiques et d'habitudes d'apprentissage à travers le monde, de relever certains clichés : telle population d'apprenants sera dite plutôt animée et bruyante, tandis que d'autres seront qualifiés de studieux et silencieux ; certains seront portés sur l'expression orale, tandis que d'autres seront férus de descriptions grammaticales. Il s'agit là essentiellement de stéréotypes sociaux, stéréotypes qui circulent au sein du corps enseignant comme ailleurs et qui peuvent par la suite influencer les styles d'enseignement. Parmi ceux qui n'ont jamais vécu au Japon, nombreux sont ceux qui auraient tendance à catégoriser les étudiants japonais dans le deuxième groupe : travailleurs, calmes, discrets, (trop) silencieux, réfléchis, pouvant être très compétents dans des tâches d'expression écrite et plus encore lorsqu'il s'agit de se confronter à la panoplie des exercices grammaticaux, mais en revanche peu enclins à animer une discussion polémique à l'oral. Voilà l'image que beaucoup de ces étudiants peuvent donner lorsqu'ils se trouvent en France au sein de classes internationales et multilingues. Ces clichés sont bien connus des chercheurs qui travaillent sur l'interculturel d'une part, et les styles et stratégies d'apprentissage d'autre part. Ceux qui enseignent au Japon, quant à eux, jettent un regard nuancé sur cette image : il faut y faire la part des choses. N'étant pas spécialiste du domaine interculturel, nous ne pouvons ici apprécier en détail les tenants et les aboutissants de ces analyses. Néanmoins, il faut signaler que cette culture de « la non-expression » joue un rôle crucial dans la mise en œuvre ou l'absence d'entraînement audio-phonatoire pour les étudiants japonais.

Il faut ensuite considérer l'organisation et surtout la perception du système éducatif supérieur par les étudiants japonais : pour beaucoup d'entre eux, l'université constitue, dans une certaine mesure, une étape de « repos » et de « libération », entre une course à l'excellence effrénée qu'ils ont dû mener durant le secondaire (avec des cours du soir dans des « Juku », écoles privées complémentaires), en vue des concours d'entrée aux universités, et une entrée dans le monde du travail souvent difficile pour les jeunes diplômés. L'université ne serait donc pas systématiquement un lieu d'apprentissage intensif et la plupart des universités, souvent privées, accorderaient alors, sans sélectivité interne trop rigoureuse, leur diplôme, au terme de quatre années d'études, à

la plupart des étudiants (*dans une certaine mesure*, bien sûr), ce qui peut aussi s'expliquer par le coût exorbitant (d'un point de vue français au moins) de la scolarité annuelle. On peut ainsi schématiser *très grossièrement* la situation : en France, il est assez facile d'entrer à l'université, mais chaque année constitue un véritable défi auquel les étudiants doivent répondre pour pouvoir passer en année supérieure, avec une grande importance accordée à la composition du cursus et aux résultats obtenus pour la suite de leur parcours professionnel ; au Japon, en revanche, les étudiants ont travaillé très dur durant toutes leurs études primaires et secondaires, afin de pouvoir réussir les concours d'entrée aux universités les plus prestigieuses, concours extrêmement sélectifs ; une fois l'entrée à l'université acquise, les examens semblent *relativement* faciles et seul compte le prestige de l'université, ainsi que l'obtention d'une qualification finale, ce qui rappelle, *d'une certaine manière*, le système des classes préparatoires et des Grandes Ecoles en France. En ce qui concerne la composition du cursus, elle semble avoir moins d'importance sur le marché du travail, étant donné l'uniformité (*relative*) des parcours des étudiants : quatre ans d'étude, ni plus ni moins, pour tous les étudiants « undergraduate » et une recherche d'emploi institutionnalisée (« job-fairs ») et ritualisée (« job-hunting ») qui ne nécessite que peu d'initiative personnelle. Les curriculum vitæ se présentent tous sous un modèle standard (type « fascicule administratif ») et la formation est dispensée par la suite de manière systématique au sein de l'entreprise ou de l'organisation d'embauche. Conclusion : nul besoin de se former avant, puisque l'on est formé ensuite. Tout cela ne favorise pas, du point de vue de la motivation, l'investissement personnel des étudiants dans leur étude de la langue. Rappelons tout de même qu'il ne s'agit là que de généralités, sans aucun doute *très abusives*, mais qui nous permettent de souligner certaines carences, notamment d'ordre motivationnel.

### **III.1.3. Raisons d'ordre (psycho-)linguistique**

Les facteurs de nature (psycho-)linguistique à l'origine des « problèmes de prononciation » des étudiants japonais constitue, en grande partie, le cœur de la présente étude. Nous ne développerons pas ainsi ce point ici, mais nous pouvons d'ores et déjà indiquer que les différences linguistiques et psycholinguistiques (système

phonético-phonologique, comportements perceptifs, organisation du lexique phonologique, système scripturaux et conscience phonologique, etc.) entre le français et le japonais, qui appartiennent à des familles linguistiques différentes (*Pour la Science*, 1997) jouent bien évidemment un rôle déterminant dans les difficultés rencontrées au cours de l'apprentissage.

#### **III.1.4. Raisons d'ordre didactique**

D'un point de vue didactique, il convient de souligner trois aspects de la situation des étudiants japonais :

- D'abord, les objectifs d'apprentissage (largement déterminés par les objectifs d'enseignement en contexte universitaire) ne sont pas favorables au développement des compétences audio-phonatoires. Il ne s'agit bien souvent pour les étudiants que de réussir leurs examens de fin d'années : examens sur support écrit, de type QCM et exercices de transformations grammaticales.
- Par ailleurs, l'environnement exogène des étudiants japonais au Japon est peu propice à la conversation exolingue comme le souligne Ishikawa (2003), ce qui n'encourage ainsi pas les étudiants dans leur étude de la langue orale.
- Enfin, en ce qui concerne la didactique générale et les modes d'enseignement et d'apprentissage au Japon, il est habituel de mentionner que ceux-ci sont essentiellement axés sur la lecture et l'expression écrite, une culture éducative de l'écrit qui reflète peut-être à la fois le mode de communication sociale extra-scolaire d'une part, et d'autre part la nature de l'un des codes scripturaux en vigueur au Japon, celui des « kanji ». L'apprentissage des kanji est en effet un apprentissage à vie, qui nécessite une réactualisation constante (il est facile d'oublier certains kanji, même pour des sujets japonais, en raison de leur faible fréquence d'occurrence), et la connaissance des kanji, tout comme, dans une certaine mesure, la connaissance de l'orthographe française en France, constitue un facteur de distinction sociale qui peut expliquer en partie l'attachement porté à leur étude.

Le contexte dans lequel se trouve notre public d'apprenants étant précisé, nous pouvons à présent détailler ce que nous avons jusqu'à présent nommé « prononciation ».

### **III.2. Qu'est-ce que la « prononciation » ?**

Dans le corps enseignant, entre autres, il existe malheureusement une fâcheuse tendance à manquer de précision dans les termes que l'on utilise, et ce manque de précision entraîne parfois un flou conceptuel qui ne facilite en rien la recherche de solutions concrètes aux problèmes posés. Il nous semble donc important de clarifier ce que l'on entend par le terme « prononciation » et de délimiter son domaine d'application ou de pertinence.

#### **III.2.1. Définitions**

Dans le *Larousse de Poche 2001*, on trouve les entrées suivantes :

Prononciation : Action, manière de prononcer. (p.629)

Prononcer : 1. Articuler : *prononcer les mots*.

2. Dire, débiter : *prononcer un discours*.

3. Déclarer avec autorité : *prononcer un arrêt*. (p. 629)

On voit bien ici que *Prononcer*, dans un de ses sens, semble synonyme de *Articuler*.

Articuler : 1. Faire entendre distinctement des sons, les syllabes des mots, à l'aide des organes de la parole. (p. 49)

Apparaît alors ici le pendant audio-acoustique de la prononciation, puisqu'il s'agit de « faire entendre distinctement ». La structure « faire entendre » implique bien un « entendant » : celui qui articule, ou bien une autre personne, son interlocuteur. Si l'on ouvre l'ouvrage de M. et P. Léon (1997), intitulé *La prononciation du français*, on ne trouve pas véritablement de « définition », mais voici ce que l'on peut lire dans l'avant-propos (p. 7) :

*« Ce petit livre [...] présente élément par élément les mécanismes de la prononciation [...]. On a trop fait de la phonétique une science aride où les données rébarbatives de la physiologie et de l'acoustique étaient présentées souvent de manière trop savante [...]. Les divers aspects de l'étude de la prononciation sont ici passés en revue, montrant comment les sons sont formés, codés linguistiquement, quel usage nous en faisons [...]. [L]a prononciation est un champ d'études passionnant pour quiconque s'intéresse à la langue en tant que reflet de l'humain sous toutes ses faces, individuelle, sociale, régionale, aussi bien que communicative ».*

A travers ces deux présentations, la « prononciation » se transforme ainsi d'une simple « manière de faire entendre » à un « champ d'étude » très varié, dont les différents aspects se reflètent dans le sommaire de l'ouvrage :

1. L'oral et l'écrit
2. Classement articulatoire
3. Lois phonétiques
4. Problèmes d'orthoépie : E caduc et liaison
5. Aspects fonctionnels
6. Aperçus phonologiques
7. Aspects psycho-physiologiques
8. Aspects acoustiques et auditifs
9. L'accentuation, les pauses et le rythme
10. L'intonation
11. Les différents types de variation

Vouloir résoudre des « problèmes de prononciation » implique donc de cerner avec une relative précision la nature, l'origine et l'extension des dits problèmes. En parlant d'extension, on pose alors la question des effets de ces problèmes : la prononciation ne constitue que l'un des aspects de tout processus langagier oral, et l'on doit prendre en compte les relations entre les différents aspects, afin de déterminer la véritable origine des problèmes (de quel niveau (psycho)linguistique relèvent-ils ?). La démarche est

identique à celle de l'aphasiologiste qui cherche à évaluer le déterminisme sous-jacent aux manifestations pathologiques de surface d'un patient lésé : s'agit-il d'un trouble profond ou superficiel ? Central ou périphérique (Nespoulous, 2000) ? Ceci nous conduit donc aux interfaces entre les différents aspects des processus langagiers oraux.

### **III.2.2. Interfaces : langue et communication**

Sans rentrer en détail dans l'analyse linguistique (aspects structurels) ou psycholinguistique (aspects fonctionnels) des questions abordées ici, puisque nous les traitons plus loin, il nous semble cependant nécessaire d'évoquer les enjeux *didactiques* de la prononciation, car, de ce point de vue, la « prononciation » semble bien souvent marginalisée, réduite, voire amputée de ses implications pour l'apprentissage général de la langue, dans l'esprit de nombreux enseignants.

Du point de vue du système linguistique, la prononciation renvoie au niveau d'analyse phonologique, lequel reflète une dimension centrale de la langue, étroitement imbriquée aux autres, en particulier au lexique (formes phonologiques) et à la morphologie (allomorphes). La prononciation n'est donc pas marginale, mais au contraire essentielle, en situation d'interlocution, en production ou en compréhension, pour assurer la grammaticalité et la clarté sémantique du discours, produit ou reçu.

Du point de vue de la communication orale, la qualité de la « prononciation » va être déterminante sur la réussite ou non de l'interlocution. On aborde ici la dimension pragmatique de la langue, en considérant les facteurs non-linguistiques (situationnels) et les actes de paroles (fonctions illocutoires et effets perlocutoires) sous-tendant les énoncés (Baylon et Mignot, 1999, pp. 97-107). Il nous faut alors considérer ici les schémas de la communication, tels que ceux de Jakobson ou de Kerbrat-Orecchioni par exemple, mais en les adaptant à la situation particulière des apprenants de langue étrangère, laquelle modifie sensiblement en particulier la forme et l'importance des stratégies de communication (stratégies palliatives, d'évitement, etc.).

Toujours dans une perspective « communicativiste », il est important de prendre en considération la dimension identitaire de la « prononciation ». Celle-ci renvoie en effet à ce que l'on appelle « l'accent » d'une personne : « accent francophone » lorsqu'il s'agit de marques sociolinguistiques à l'intérieur de l'espace francophone (ou plutôt hexagonal pour la majorité des français de la métropole), « accent étranger » lorsque la personne en est linguistiquement en dehors. Or, qui dit « accent étranger » dit différence, dit regard de l'autre et dit sensibilité culturelle, voire susceptibilité culturelle ou individuelle. Certaines personnes, pour diverses raisons, désirent alors minimiser leur « accent » et se rapprocher d'une certaine norme phonétique, tandis que d'autres, au contraire, soit souhaitent le conserver comme marque d'appartenance à une certaine communauté, soit n'y attachent aucune importance. Ce qui apparaît alors, c'est que, contrairement peut-être à d'autres aspects de la langue à apprendre, la prononciation touche de près au système émotionnel des individus, un fait qu'il convient de prendre très sérieusement en compte dans un cadre didactique, étant donné l'influence du système émotionnel sur la motivation et ainsi l'apprentissage des sujets.

### **III.3. Observations, pratiques et enquêtes au Japon**

Afin d'ancrer nos problématiques dans une approche « de terrain », nous présentons ici certaines données empiriques issues de notre fréquentation dudit « terrain ».

#### **III.3.1. Observations**

Si l'expatriation (même provisoire) présente de nombreux inconvénients, notamment en ce qui concerne l'environnement scientifique et l'accès aux ressources académiques et didactiques, elle nous semble cependant être une étape indispensable pour tous ceux qui souhaitent s'inscrire dans la problématique de l'Apprentissage des Langues, du Français Langue Etrangère et de la Francophonie. Pour les chercheurs-didacticiens en particulier, le terrain d'enseignement exogène constitue un champ d'investigation riche et formateur, sans l'expérience duquel les travaux produits risquent parfois de manquer soit de recul, soit d'amplitude, soit de réalisme. Idéalement, il faudrait pouvoir



enseigner dans deux ou trois pays assez différents, tant d'un point de vue linguistique que culturel, pour pouvoir disposer de repères et de références suffisamment variés.

Notre expérience de l'enseignement du FLE au Japon en contexte privé et universitaire (après l'avoir enseigné en Grande-Bretagne en contexte universitaire et brièvement en France dans les deux types de contexte), nous a ainsi permis d'être confrontés à des populations d'apprenants très variées (sexe, âge, catégorie socio-professionnelle, origine géographique, compétences linguistiques, etc.) qui constituent l'une des réalités de la Francophonie à travers le monde : ce sont eux, notamment, qui font vivre le français à l'étranger.

Nos réflexions et nos travaux académiques ont donc été nourris *empiriquement* par des *observations* et des *interventions* didactiques quotidiennes *in vivo* pendant plusieurs années. Nous espérons nous épargner ainsi, au moins en partie, d'une part les risques de produire du « travail de papier », et d'autre part les accusations potentielles de théorisation ou d'abstraction exagérées coupées du réel... Vouloir traiter d'apprentissage des langues implique en effet de se confronter, de manière suffisante, à des situations réelles d'enseignement/apprentissage. Il va néanmoins de soi qu'il n'est pas aisé de « tout mener de front » : l'énergie et le temps accordés à l'enseignement ne le sont pas à la réflexion et à l'analyse académique, et vice-versa. L'équilibre est donc toujours délicat à trouver, et surtout à maintenir.

Dans notre expérience d'enseignement en contexte privé japonais, grâce à nos ateliers de phonétique corrective (3 classes d'une heure par semaine) (durant lesquels occasionnellement nous entraînions nos étudiants à la lecture à voix haute), à nos classes audiovisuelles (3 classes de deux heures par semaine) (au cours desquelles nous passions au minimum 30 minutes par classe en laboratoire de langue) et enfin grâce à nos classes de dictée (qui incluaient des tâches de lecture à voix haute) (différents niveaux, deux heures par semaine), nous avons pu, en particulier au niveau débutant, observer les erreurs et les confusions de nos étudiants, tant à l'oral, qu'à l'écrit, ainsi qu'aux interfaces entre les deux (tâches de dictée et de lecture à voix haute). Bien sûr,

ces observations informelles ont également été conduites durant les autres classes (avec manuels et/ou documents annexes). En moyenne, nous avons donc pu disposer de 22 heures d'enseignement hebdomadaires, réparties entre 7 et 10 classes différentes (entre 1H et 3H par séance). Le nombre des étudiants par classe variait entre 10 et 30, avec une moyenne approximative de 15. Les profils étaient très variés : étudiant(e)s universitaires, femmes au foyer, employé(e)s, retraité(e)s, lycéen(ne)s, etc., avec des écarts d'âge parfois très importants, tout comme leur niveau et éducation linguistique. La grande majorité était de langue première japonaise, mais certains étaient d'origine coréenne ou chinoise.

Avant de présenter ci-dessous quelques observations notées au cours de nos enseignements, il est important d'insister sur la nature du contexte : les étudiants universitaires et les apprenants d'écoles privées (même s'il se recoupent parfois) diffèrent non seulement du point de vue de leur profil (l'école privée demandant un effort d'adaptation du cours et de l'enseignant beaucoup plus important), mais aussi, sinon surtout du point de vue de leur motivation (l'université demandant alors un effort de la part de l'enseignant beaucoup plus grand). Les motivations étant sensiblement différentes, le contrat d'apprentissage devient lui aussi alors différent. Or, pour la phonétique corrective par exemple, ce contrat est très important dans la dynamique du groupe-classe : l'évaluation et la correction, notamment, doivent être acceptées par l'apprenant. Ce contrat est particulièrement sensible au contexte institutionnel (en contexte privé, les relations hiérarchiques et les transactions ne sont souvent pas les mêmes qu'en université), ainsi qu'au contexte socioculturel (les rapports de hiérarchie sociale au Japon sont assez complexes (sexe, âge, fonctions, etc.) et ils sont surtout encore très en vigueur dans la société actuelle (contrairement à d'autres sociétés), entretenus à travers la langue japonaise dont les différents registres sociolinguistiques en sont les marques mêmes). Par ailleurs, l'hétérogénéité des apprenants conduit à une diversification considérable de l'origine des difficultés audio-phonatoires (éducation

linguistique, déficience auditive ou phonatoire, etc.)<sup>15</sup>. Dans un tel contexte, la nécessité d'une personnalisation relative de l'enseignement apparaît clairement.

Voici donc ci-dessous quelques observations informelles relatives à la question de la prononciation, à travers ses manifestations orales ou écrites :

### III.3.1.1. De la perception auditive à la production écrite : erreurs de « phonétique » en production sous dictée

Dans l'institution privée dans laquelle nous avons enseigné, pour certains examens bi-trimestriels, les sujets sont fournis aux enseignants par l'institution, accompagnés d'un barème pour la notation. Ces examens comportent généralement une épreuve de dictée. Pour ces dictées, le barème propose trois catégories d'erreurs : « Usage », « Phonétique » et « Grammaire » (par ordre croissant de « gravité »). On pourrait d'ores et déjà s'étonner de la présence d'une catégorie « phonétique » dans un barème de correction d'un document écrit. La « phonétique » (ou « prononciation ») ne concernerait donc pas seulement « l'expression orale », mais également la perception auditive, la compréhension orale, l'expression écrite et l'orthographe. Bref, il ne s'agit donc pas d'une compétence triviale, et cela est encore plus marqué par le fait que, selon ce barème, une « erreur de phonétique » vaut deux fois plus (2 points-faute) qu'une « erreur d'usage » (1 point-faute), mais pas deux fois moins qu'une « erreur de grammaire » (qui vaut 3 points-faute). On pourrait bien sûr facilement critiquer ce barème et ce mode d'évaluation non-formatif, mais cela n'enlève rien à l'existence et à l'utilisation réelles de nombreuses grilles de ce genre à travers le monde. En tous cas, il a le mérite de souligner l'importance des aspects phonétiques dans la réalisation d'une épreuve dont le produit fini est écrit. Dans le cas de nos étudiants japonais, ce mérite

---

<sup>15</sup> Ainsi cet apprenant âgé d'environ 70 ans, qui, en tâche de répétition, substituait systématiquement le phonème /t/ au phonème /s/ en attaque syllabique, malgré nos efforts et les stratégies de lecture labio-faciale qu'il déployait sensiblement (et qui devaient certainement le conduire également en erreur). Nous avons dû nous résigner à transiter par la reconnaissance lexicale en français, voire en japonais, pour obtenir, isoler, puis réinsérer dans l'énoncé initial, la syllabe débutant par le phonème /s/. Ces cas de déficience auditive étaient cependant exceptionnels.

est nécessaire, car nous avons été très surpris par la quantité d'erreurs qui relevaient précisément de la catégorie « Phonétique » du barème, laquelle permet d'éclairer, *en partie* au moins, les deux exemples suivants :

Item dicté	Transcriptions
Dans cette œuvre, on retrouve la présence de l'auteur.	-Dans cet arbre, on retrouve la présence de retour. - Dans cette arbre, on retrouve la présence de l'hôtel <sup>16</sup> .
Nous écoutons	- Nous éctons
Ecoutant	- Ectant

Nous avons donc pu percevoir les problèmes phonético-phonologiques (« de prononciation ») des apprenants à travers leurs textes écrits. Si cela est vrai pour les épreuves de dictée, il en va de même pour des exercices écrits de type « Expression écrite, libre ou guidée » ou bien de simples exercices grammaticaux. Les productions écrites des apprenants semblent donc souffrir de leurs « problèmes de prononciation ».

Nous avons pu constater des problèmes identiques avec des étudiants qui avaient pourtant suivi des classes audiovisuelles, très axées sur l'entraînement audio-phonatoire. En tâche de dictée, nous avons ainsi trouvé « robe jaune » écrit "robe joune", ou bien « de la chambre » écrit "de la chamble". Est-ce un problème d'encodage phonographémique ou de perception auditive ? Nous avons été surpris de constater que des problèmes de la sorte persistaient même au plus haut niveau d'étude de l'établissement. Ainsi, en classe supérieure de niveau 2, en tâche de dictée, le mot « laboratoire » a été écrit « lavoratoire ». Nous avons également été surpris par les stratégies adoptées en tâche de dictée face à des groupes triconsonantiques : ainsi le mot « meurtre » a été écrit "motre" par un étudiant et "meurter" par un autre étudiant.

---

<sup>16</sup> Notons qu'il s'agissait d'un énoncé décontextualisé et que les énoncés précédents ne faisaient mention ni d'arbre ni d'hôtel.

Nous présentons ci-dessous quelques résultats issus d'une épreuve de dictée lors de l'examen final d'une classe basée sur le manuel *Champion 1* (21 étudiants, décembre 2003, les chiffres entre parenthèses indiquent le nombre d'occurrences):

<b>Original</b>	<b>Ecrit par les étudiants</b>
Janvier	Junvier / jonvier / janvie / jenbiér / jemvier
Médecin	Médechine / medecine (2) / médecine (2)/ médecine / médcin
Du	De (2)
De	Du (4)
L'a soignée	La soignait (2) / la soigner / la soigne / la soignée
Retourner	Retrouner (4) / retorner (3)/ retrener (2) / rortorner / rentere
S'est inscrite	S'est inscritee (3) / inscrit / inscritee / se inscrite / se insclit
Mois	Moin (2) / moins
Arrêter	Aller (2) / aller tout /
Sport	Suport (2)
Veut	Peut
Jambe	Jamb (2) / jeumb / jumb
Chez	Shez
Se reposer	Se réposer / reporze / opposer / reproser / repouse / reporser
Faisait	Fasait / fusée / fesait (2) / faiez / fusé
Fait	Faite / fair
Gymnastique	Gimenastique / gymenastique / gymunastique / gymnasetique
Doit	Doire
Quelques	Quelleques
D'arrêter	D'arrtêr
Semaines	Semain
Aussi	Assi
Pendant	Pondant (2)
Maintenant	Mantenon

L'analyse de ces erreurs permettrait certainement de mettre au jour certaines « erreurs de prononciation » caractéristiques des apprenants japonais de FLE, quoique l'on doive distinguer les niveaux phonétique (en perception) et phonographémique (en production) :

- Janvier → jonvier : mauvaise maîtrise des voyelles nasales (pas de voyelles nasales en japonais).
- Du → de : mauvaise maîtrise de l'opposition entre le /y/ et le /ø/ (une seule voyelle en japonais : /ɯ/).
- Gymnastique → gymunastique : mauvaise maîtrise de la syllabation du français (voyelle épenthétique).

Cependant, d'autres facteurs (sémantiques, morphosyntaxiques) interviennent dans une telle tâche, et ces exemples ne sont donc présentés qu'à titre illustratif. Notons que nous avons également trouvé « robe bleue » écrit "robe blue", une « erreur » qui renvoie à un autre facteur : celui des ressemblances lexicales. L'anglais étant très largement répandu, sinon pratiqué, parmi les Japonais (surtout des dernières générations) on peut en effet se demander s'il s'agit d'un problème de perception auditive ou bien d'un calque lexical de l'anglais vers le français (cette stratégie de calque, surtout lexical, anglais/français nous semble très fréquente chez nos apprenants, en particulier parmi les jeunes générations).

Enfin, on ne peut également manquer de citer le cas, que l'on pourrait presque qualifier de « contre-exemple », de cette étudiante qui, tout en s'obstinant à prononcer le /R/ français comme un /r/ japonais en tâche de répétition ou de production orale semi-guidée, nous a surpris en se révélant étonnamment compétente en tâche de dictée, une épreuve qu'elle accomplissait régulièrement avec très peu de substitutions du type <r>/<l>. Il faut néanmoins signaler que cette tâche de dictée suivait l'étude et la présentation du texte, qui était ensuite dicté : la distinction entre encodage auditif et

encodage orthographique pourrait ainsi sembler particulièrement nette dans le cas de cette étudiante, qui a le mérite de souligner la complexité de la problématique.

### III.3.1.2. Perception auditive et production orale

Parmi les nombreuses interrogations qu'ont suscité chez nous les réactions de nos apprenants, l'exemple suivant fait partie des plus marquants : dans une classe de faux-débutants, dont la durée hebdomadaire était de 3h chaque mercredi matin, basée sur un manuel de FLE de confection française (*Champion 1*), après plus de 6 mois d'étude du français, lors de l'écoute d'une dictée (ex. 5, p. 84), sur 18 étudiants (très variés), 12 d'entre eux ont rapporté entendre « la port » (même après plusieurs réécoutes et mise en focus dans la discussion), alors que seulement le tiers (c'est-à-dire 6 étudiants) et nous-mêmes entendions « le port ». Comment expliquer cela ? La classe était-elle divisée en 12 « malentendants » et 6 « normo-entendants » ? S'agissait-il d'un problème d'orientation *top-down* (attente lexicale ou autre) ? Ou bien d'un problème de catégorisation auditive ?

Voici ci-dessous quelques exemples supplémentaires, *a priori* surprenants, liés à la prononciation des étudiants, avec ou sans « correction phonétique » :

- Une étudiante prononce le mot « réfléchir » en substituant alternativement [si] ou [çi] à [ʃi], mais elle prononce « messieurs-dames » en substituant [ʃi] à [si]<sup>17</sup>. Alors que le premier exemple laisserait penser que cette étudiante est incapable de prononcer [ʃi] (ce qui est surprenant, puisque la syllabe initiale du nom japonais « shibuya » lui est quasiment identique), mais qu'elle est capable de prononcer [si] (ce qui est également surprenant puisque la séquence /si/ est illicite en japonais), le deuxième exemple (beaucoup plus « canonique » dans le domaine des erreurs de prononciation chez les apprenants japonais de français) révèle qu'il n'en est rien : il s'agit bien plus d'instabilité et de variation.

---

<sup>17</sup> Dans le cadre d'une leçon du manuel de FLE *Initial 1*, p. 96.

- Au cours d'un atelier de phonétique corrective, les étudiants doivent répéter le mot « fichier » et plusieurs d'entre eux s'obstinent à répéter [si] ou [çi] au lieu de [ʃi]. Cependant, en appliquant une technique de segmentation/concaténation du modèle à répéter (« fich---ier » → « fich—ier » → « fich-ier » → « fichier »), les apprenants parviennent à une production « acceptable ». Nous avons appliqué cette technique en tenant compte des contraintes phonotactiques du japonais qui interdit les séquences /si/. Cette technique est bien connue des praticiens de la phonétique corrective, voire même de phonéticiens de renoms (par exemple Wells (2000, p. 125), qui, pour des apprenants japonais d'anglais, propose le « découpage » suivant « step up » → « ste pup », en vue de les entraîner à la production de coda<sup>18</sup>).
- De la même manière, en tâche de répétition, et après lecture de la forme écrite, les apprenants devaient répéter la séquence « la fenêtre » en omettant le « e muet » (« la fnêtre »). Malgré plusieurs essais, les étudiants ne parvenaient pas à « effacer » le « e muet ». Nous avons donc appliqué la même technique de segmentation/concaténation (« laf---nêtre » → « lafnêtre » → « la fnêtre »), et sommes parvenus à obtenir de bons résultats. Ces deux exemples soulignent l'importance des contraintes phonotactiques (et donc syllabiques (Archangeli, 1997)) sur la production des apprenants, et ainsi l'intérêt pour l'enseignant de les connaître, afin d'adopter rapidement une stratégie de résolution adaptée.
- Un autre exemple d'adaptation de stratégie de résolution est le suivant : le mot « héros » prononcé par un étudiant avec une fricative uvulaire non-voisée [χ] au lieu d'une fricative uvulaire voisée [ʁ] ou d'une approximante uvulaire [ʁ̥].

---

<sup>18</sup> Wells (*ibid.*) insiste néanmoins sur la nature transitoire de la technique, essentiellement adaptée aux apprenants débutants, pour deux raisons principales: « first, because native speakers do not actually pronounce final consonants in exactly the same way as initial ones: a *great ape* [grɛɪt əɪp] is phonetically distinct from a *grey tape* [grɛɪ tɛɪp] [...]; and secondly, more importantly, because in real life most word-final consonants are not followed by a vowel-sound at the beginning of the next word ».



Malgré plusieurs répétitions, ainsi que l'utilisation de la technique de segmentation/concaténation appliquée selon la syllabation française « canonique » hér/ros, les résultats sont négatifs. Cependant, lorsque cette même technique est appliquée selon un découpage non-standard (hér/os)<sup>19</sup>, l'apprenant parvient à de bons résultats. Un tel exemple souligne que la syllabation « canonique » française n'est pas toujours la plus adaptée à la résolution de problèmes didactiques, en l'occurrence en raison de l'opposition attaque/coda pour le segment /r/, la dernière semblant plus « facile » à traiter par les apprenants dans une telle tâche, ce qui peut s'expliquer de plusieurs manières (allophonie en français et phonotactique en japonais du segment /r/).

- En ce qui concerne la variation des performances intra-individuelle d'une part et les phénomènes de coarticulation d'autre part, nous avons également été surpris, en tâche de production libre ou de répétition, d'entendre nos étudiants prononcer des réalisations phonétiques du phonème /R/ tout à fait différentes au cours d'un énoncé, voire du même énoncé qu'ils sont en train de répéter (un [R] dévoisé, un [r] américain, un [r] japonais, un [l] « clair », un [l] « sombre », une « jota espagnole » (fricative vélaire non-voisée [x]), etc.), avec, par exemple, le mot « préparé » prononcé [plepaʁe]. A ce propos, dans une perspective didactique, il faut noter que ces réalisations phonétiques, variant en fonction du contexte (segmental et suprasegmental), ne présentent pas toutes, aux oreilles du locuteur natif, le même degré de similarité auditive avec la réalisation d'un locuteur natif. Le degré d'allophonie vis-à-vis de la réalisation native varie ainsi en fonction du contexte. Si ces observations étaient systématisées, on pourrait alors peut-être soutenir que, selon la position dans le mot phonologique, il importerait d'entraîner davantage les apprenants sur certains allophones et certaines combinaisons phonotactiques, afin de rationaliser la pratique enseignante dans une perspective communicative

---

<sup>19</sup> Et avec l'aide d'une gestuelle adaptée.

(atteindre une réalisation allophonique « acceptable ») et non normative (réalisation « native », elle-même tout à fait idéale).

### III.3.1.3. Lecture et production orale

Ce troisième type d'exemple est particulièrement important pour nous, puisqu'il touche au cœur même de notre étude. On peut en effet s'étonner de constater que certains apprenants, capables d'effectuer, de manière relativement systématique, certaines oppositions phonémiques en production orale (en tâche de répétition ou de production libre), multiplient en revanche les « erreurs » lorsqu'ils doivent effectuer une tâche de lecture à voix haute. Ainsi, nous constatons très régulièrement des « substitutions » de ce type chez nos apprenants japonais, en particulier entre les phonèmes /R/ et /l/ et /b/ et /v/. Doit-on considérer de telles erreurs comme des « erreurs de prononciation » ? Doit-on, ou plutôt peut-on, alors appliquer des techniques de « correction phonétique », qui sont généralement sollicitées lorsque les apprenants sont engagés dans des tâches de production orale libre ou sous répétition ? Que doit-on penser de cette étudiante qui, en laboratoire de langue, écoute la séquence « tu t'arrêtes » tout en lisant la transcription orthographique, puis la répète/lit à voix haute en prononçant la dernière syllabe [rɛt] au lieu de [ʁɛt] ? Après avoir alors nous-même répété la séquence deux fois pour attirer son attention sur la forme phonétique de la séquence, l'étudiante parvient, après deux essais erronés, à une production « correcte ». On pourrait en conclure que, malgré le stimulus auditif qui accompagnait la lecture de la séquence, la lecture à voix haute/répétition (bimodalité) conduit à une production erronée, alors que cette étudiante est capable de répéter correctement la séquence (unimodalité auditive).

Le cas de la lecture à voix haute oblige ainsi le didacticien à approfondir la notion d'« erreur de prononciation », qui est souvent attribuée à un problème de « perception auditive » et qui associe ainsi systématiquement la « prononciation » au domaine phonétique, en négligeant quelque peu le domaine phonologique. A moins d'adopter une perspective exemplariste stricte (Coleman, 2002), le lien entre perception de stimuli orthographiques et production orale semble fortement impliquer la phonologie

d'une part, et la phonographémie d'autre part. Cette question nous conduit alors à la problématique centrale de notre étude, à savoir l'influence possible de l'écrit sur l'apprentissage du système phonético-phonologique étranger. L'existence d'une telle influence est particulièrement envisageable lorsque l'on écoute plusieurs de nos apprenants japonais, qui s'obstinent, en tâche de production orale libre ou de lecture à voix haute, à prononcer [e] les « e muets » finaux des formes verbales du type « il mange » ou « il regarde », entraînant, chez l'auditeur, une confusion avec les formes verbales correspondantes à l'imparfait : « il mangeait » ou « il regardait ».

A travers les exemples précédents, ainsi que de nombreuses autres observations de classe, il est possible de commencer à distinguer ce que l'on peut qualifier de « problèmes classiques » de « prononciation » des apprenants japonais de français. Au niveau segmental, il est habituel de mentionner principalement les couples B/V, R/L, EU/OU et U/OU. Il existe également d'autres zones de difficultés, mais celles-ci sont cependant moins fréquemment abordées, car elles ne peuvent être traitées en termes simplistes d'existence ou d'absence d'unités phonémiques : il s'agit en effet de problèmes qui peuvent être considérés soit comme phonotactiques (opposition et neutralisation), soit comme syllabiques (l'approche syllabique pouvant, dans une certaine mesure, rendre compte des faits phonotactiques ; il ne s'agit que d'une différence d'analyse (Archangeli, 1997)). A titre d'exemple, nous pouvons mentionner l'opposition française /s/ vs. /ʃ/ : en japonais il existe *au moins* trois types de réalisations phonétiques correspondant à ces unités ([s, ø, ʃ])<sup>20</sup>, alors qu'il n'y en a que deux en français ([s, ʃ]). Etant donné le système syllabique/moraïque du japonais, la nature de la voyelle qui suit le segment joue un rôle essentiel dans la détermination de la réalisation allophonique. Un énoncé comme « ces chaussures » pose ainsi de nombreux problèmes aux apprenants. Par ailleurs, d'autres aspects du système de traitement phonético-phonologique rentrent également en jeu dans la réalisation d'une tâche de répétition ou de lecture à voix haute (strate lexicale, reconnaissance lexicale,

---

<sup>20</sup> A ce sujet, voir notamment (Catford, 1994, p. 96).

etc.), permettant d'expliquer pourquoi certains étudiants sont capables de répéter correctement « assis » [asi], mais ne peuvent répéter « cinéma » et prononcent plutôt [ʃinema] ou [çinema], malgré une mise en focus attentionnel explicite.

### **III.3.2. Pratiques**

Une fois ces observations effectuées et les difficultés constatées, les enseignants s'interrogent alors sur les techniques de remédiation praticables : quelles sont les interventions didactiques possibles ? Beaucoup d'enseignants de français locaux ou d'enseignants-chercheurs se sont bien sûr déjà penchés sur la question, à commencer par Cortès (1970) lors du premier colloque franco-japonais sur l'enseignement du français au Japon. Il suffit de parcourir les différents numéros des dix dernières années de la revue *Enseignement du français au Japon*, publication de la Société Japonaise de Didactique du Français (Tokyo), pour constater que tous les ans des efforts et des propositions sont faits par les enseignants pour essayer de résoudre les problèmes (par exemple, parmi les articles *en français*, Richardot, 1988 ; Guimbretière, 1994 ; Abrial, 1995). Ces efforts peuvent aussi déboucher sur la publication de manuels spécialisés pour l'enseignement de la prononciation du français aux apprenants japonais).

Deux pistes principales sont alors à explorer : d'une part celle du matériel pédagogique existant, et d'autre part celle de la formation des enseignants de langue, en phonétique (corrective ou non) et en phonologie.

#### **III.3.2.1. Matériel pédagogique**

S'il existe sur le marché plusieurs ouvrages spécialisés dans « l'enseignement/apprentissage de la prononciation française », que ceux-ci soient de facture française et « universalistes »<sup>21</sup> ou<sup>22</sup> bien de facture franco-japonaise et

---

<sup>21</sup> Par exemple: Charliac, L. et Motron, A.-C., *Phonétique progressive du français*, Paris, CLE international, 1998.

<sup>22</sup> Certains auteurs travaillent sur les deux types d'ouvrage, comme par exemple E. Guimbretière :

davantage « ciblés » vers un public japonais<sup>23</sup>, l'utilisation de ce matériel présuppose généralement une certaine connaissance, même minimale, du domaine de la part de l'enseignant, et surtout un intérêt particulier accordé à cet aspect de l'apprentissage. L'utilisation de tels supports exige en outre un certain investissement temporel, dont ne disposent malheureusement pas bon nombre d'enseignants, contraints de respecter une progression linguistique généralement basée sur une méthode d'apprentissage plus « générale ».

C'est la raison pour laquelle nous avons décidé de nous pencher sur ces méthodes « générales » d'apprentissage de la langue française, à partir d'un niveau débutant, en examinant la manière dont étaient traités certains aspects du pan phonético-phonologique de l'enseignement/apprentissage du français. Ne pouvant examiner ici en profondeur l'ensemble des activités proposées, nous avons choisi deux points particuliers : d'une part le couple phonémique R/L, puisqu'il s'agit d'une opposition délicate à réaliser pour les étudiants japonais, et d'autre part la question de la syllabation et des groupes consonantiques, elle aussi posant problème à ces étudiants. Cinq grandes questions ont guidé notre lecture de ces manuels<sup>24</sup> :

- 1) La prononciation est-elle spécifiquement abordée (de manière régulière) ? ;
- 2) Le couple R/L est-il spécifiquement abordé ? ;
- 3) La syllabe est-elle spécifiquement abordée (autrement que par le rythme) ? ;
- 4) Les groupes consonantiques sont-ils spécifiquement abordés (autrement que par le cas du « e muet ») ?
- 5) Le rapport son / graphie est-il spécifiquement abordé (de manière régulière) ?

Les neuf manuels que nous avons examinés sont les suivants, classés par ordre de parution :

---

\*Franco-français : Kaneman-Pougatch, M. et Pédoya-Guimbretière, E., *Plaisir des Sons*, Paris, Hatier, 1989.

\*Franco-japonais : Anan, F. et Guimbretière, F., *Komyunikeishon no tame no furansugo hatsuon hou / Plaisir de communiquer: règles et exercices de phonétique pour bien communiquer*, Tokyo, Surugadai-Shuppansha, 1995.

<sup>23</sup> Par exemple: Sahara, T. et Pagès, C., *Hajimete no furansugo*, Tokyo, Natsume, 1999.

<sup>24</sup> Dont nous ne présentons ici que la synthèse.

- **CC** : *Café Crème 1*, Kaneman-Pougatch, M., Beacco di Giura, M., Trévisi, S. et Jennepin, D., Paris, Hachette, 1997.
- **RE** : *Reflets 1*, Capelle, G. et Gidon, N., Paris, Hachette, 1999.
- **AC** : *Accord 1*, Berger, D. et Spicacci, N., Paris, Didier, 2000.
- **IN** : *Initial 1*, Poisson-Quinton, S. et Sala, M., Paris CLE International, 2000.
- **CH** : *Champion 1*, Monnerie-Goarin, A., Siréjols, E. et Kempf, M.-C., Paris, CLE International, 2001.
- **ES** : *Escales 1*, Blanc, J., Cartier, J.-M. et Lederlin, P., Paris, CLE International/HER, 2001.
- **ST** : *Studio 100*, Lavenne, C., Bérard, E., Breton, G., Canier, Y. et Tagliante, C., Paris, Didier, 2001.
- **CA** : *Campus 1*, Girard, J. et Pécheur, J., Paris, CLE International/VUEF, 2002.
- **TA** : *Taxi 1*, Capelle, G. et Ménand, R., Paris, Hachette, 2003.

Les résultats de cette lecture peuvent alors être résumés de la manière suivante :

	<b>1</b> (Prononciation)	<b>2</b> (R/L)	<b>3</b> (Syllabe)	<b>4</b> (Groupes CC)	<b>5</b> (Son/Ecrit)
<b>CC</b>	Oui	Oui	Non	Non	Non
<b>RE</b>	Oui	Non	Plus ou moins	Non	Oui
<b>AC</b>	Oui	Non	Non	Non	Oui
<b>IN</b>	Oui	Plus ou moins	Plus ou moins	Non	Plus ou moins
<b>CH</b>	Oui	Oui	Non	Non	Oui
<b>ES</b>	Non	Non	Non	Non	Non
<b>ST</b>	Oui	Non	Non	Non	Oui
<b>CA</b>	Oui	Oui	Non	Oui	Plus ou moins
<b>TA</b>	Oui	Plus ou moins	Oui	Oui	Plus ou moins

Cette évaluation, très approximative<sup>25</sup> et parfois peut-être hâtive, ne vise qu'à illustrer les insuffisances potentielles des manuels de ce type sur les points qui nous préoccupent, mais surtout à souligner *leur manque d'adaptation* aux difficultés spécifiques du public japonais dans le domaine de la prononciation. Certaines activités semblent ainsi manquantes, tandis que d'autres semblent peu utiles à ces apprenants, puisqu'elles portent sur des points ne posant aucune difficulté aux apprenants japonais. Si l'enseignant souhaite pallier ces insuffisances, il lui faut alors disposer d'une formation suffisante en phonologie, en phonétique et en phonétique corrective. Or, cette formation ne semble pas toujours relever de l'évidence.

### III.3.2.2. Formation des enseignants : de la phonétique articulatoire à la méthode verbo-tonale d'intégration phonétique

En ce qui concerne la formation des enseignants de langue, on peut dire que la place qui est accordée à la phonétique (corrective) est généralement très limitée, cette discipline s'étant vue ainsi qualifiée de « parent pauvre » de la didactique des langues : beaucoup d'enseignants se sentent dès lors parfois désemparés face à des problèmes auxquels ils n'ont pas été préparés (Champagne-Muzarc et Bourdages, 1993). Etant donné la tentative de définition présentée plus haut de la « prononciation », on peut considérer deux aspects de la question (même si une telle division peut sembler arbitraire) :

- D'une part des aspects strictement linguistiques (système).
- D'autre part des aspects davantage énonciatifs (parole et communication).

Nous nous concentrons ici sur les premiers aspects, sans rentrer dans les débats relatifs d'une part aux théories (plutôt auditives ou plutôt motrices) de la perception de la

---

<sup>25</sup> Par exemple, alors que certains manuels ne proposent que quelques activités relatives aux correspondances entre l'oral et l'écrit (*Studio 100*), d'autres (*Champion 1*) le font de manière beaucoup plus systématique. De la même manière nous avons indiqué « Oui » à la question 4) pour le manuel *Campus 1*, mais celui-ci ne propose en fait qu'une activité pour les groupes /vR/ et /bR/. Il ne s'agit donc ici inévitablement que d'approximations.

parole, et d'autre part aux mérites d'une approche articulatoire ou auditive de la phonétique corrective. Pour les enseignants de français langue étrangère, la formation standard, très variable selon les centres de formation, est celle d'une initiation à la phonétique et à la phonologie du français, plus ou moins longue et plus ou moins détaillée. Maîtriser les concepts descriptifs de base de ces deux domaines associés nous semble en effet être une nécessité, ne serait-ce que pour pouvoir gérer les activités proposées dans les manuels de FLE. En général, lorsque dans les formations figurent des programmes de correction phonétique, il s'agit surtout de phonétique articulatoire, selon laquelle il faut, pour pouvoir « bien prononcer », savoir précisément et explicitement comment articuler. Parallèlement, un autre type de formation, relativement répandu dans le domaine du FLE, est celle issue de la méthode verbo-tonale (V-T) d'intégration phonétique, que nous présentons ci-dessous.

Les praticiens de cette méthode sont nombreux, et on peut ainsi en trouver à Toulouse, Besançon, ou Mons-Hainaut par exemple. La méthode V-T prend soin de mettre l'accent sur l'aspect perceptif et non articulatoire, de la pratique audio-phonatoire : les erreurs des apprenants ne seraient pas dues—généralement—à des incapacités articulatoires, mais à une mauvaise perception des modèles à reproduire. Naturellement les correspondances entre plans articulatoire et acoustique ne sont pas niées, mais le principe de base de la V-T repose sur l'affirmation suivante : pour bien reproduire un son, il faut d'abord l'avoir bien perçu. Cette méthode, élaborée principalement par Guberina (Rivenc, 2000), repose essentiellement sur le concept de *crible phonologique* (Troubetzkoy, 1976).

Il va de soi que dans de nombreux cas, ce sont des incapacités articulatoires ou des dysfonctionnements pathologiques qui sont à la source d'une mauvaise (re)production de sons. Néanmoins, ce sont les difficultés d'ordre perceptif--qui restent majoritaires--qui sont envisagées par la méthode verbo-tonale d'intégration phonétique (Renard, 1979 ; Borrell, 1991 ; Dubray et Kramer, 1994 ; Billières, 1999 ; Intravaia, 2000).



Vuletic (cité par Borrell, 1982, p. 11) écrit que « l'oreille normale en contact avec des sons étrangers se comporte comme une oreille pathologique ». Bien que son appareil auditif ne soit pas déficient, l'élève qui apprend une langue étrangère est « sourd » à certains sons de cette langue. Ce fondement du système verbo-tonal permet, à ce niveau, d'appréhender l'apprenant de langue étrangère comme un sujet pathologique. C'est pourquoi la méthode verbo-tonale, utilisée à l'origine dans le domaine de la rééducation des sourds, a été appliquée à celui de l'acquisition des langues étrangères, dans le cadre de l'élaboration de la méthodologie SGAV entre Zagreb et Saint-Cloud (Rivenc, 2000, 2003). Pour Guberina, pouvoir reproduire correctement les sons d'une langue exige au préalable de les percevoir correctement. La mauvaise perception d'un son étranger (et donc sa mauvaise reproduction) résulte de son interprétation au niveau cérébral en fonction du « crible phonologique » de la langue maternelle. A ce propos, Borrell (1982, p. 12) parle du « rôle structurant » du cerveau : « la perception de la parole se réalise en synthèse sur la base de stimuli acoustiques ». Il est fait ici allusion à la théorie motrice de la perception de la parole, selon laquelle la perception verbale auditive est effectuée grâce au couplage entre le stimulus externe et un stimulus généré de façon interne. L'auditeur génère inconsciemment ce signal interne par les mêmes mécanismes neuronaux qu'il utilise pour contrôler son canal vocal lors de la production d'énoncés (Lieberman et Blumstein, 1988, p. 147 ; Renard, 1979, p. 32). « Le conditionnement va consister à faire admettre une nouvelle structuration. Il s'agira de faire entendre de nouveaux sons en évitant qu'ils soient confondus avec d'autres sons connus [...]. Lorsqu'il sera entendu et reconnu, cela signifiera que notre cerveau lui a fait une place en restructurant ce qui existait. Le but [étant] de faire accepter cet élément nouveau comme quelque chose d'original » (Borrell, 1982, p. 15).

De manière très schématique, on présente ainsi aux futurs enseignants de langue des exemples comme celui-ci : le cas d'un étudiant espagnol qui, dans sa langue maternelle, ne dispose que des phonèmes /i/ et /u/, et non du phonème /y/. Alors, quand on lui dit « Tu vas où ? », celui-ci « entendrait » « Tou vas où ? » car /y/ n'existe pas dans son système et le phonème /u/ serait celui qui lui ressemblerait le plus en espagnol (sans

parler ici de phonotactique ou de ressemblances lexicales). Il s'agira donc de lui faire intégrer de nouveaux phonèmes dans son système phonologique, de nouveaux intonèmes et de nouveaux schémas rythmiques. Pour ce faire, les praticiens de la V-T adoptent une démarche particulière en caractérisant tout son essentiellement par deux paramètres : le *timbre* et la *tension*. D'un point de vue acoustique, un son est en effet caractérisé par certains paramètres physiques : intensité, fréquence, durée (ensemble de vibrations, vibrations qui se caractérisent par certains paramètres). La combinaison de ces fréquences va donner le timbre d'un son, plus particulièrement des voyelles. Ce sont ces timbres qui vont distinguer les voyelles entre elles. Un autre paramètre, valable plutôt pour les consonnes, est la tension, qu'on pourrait définir comme l'énergie (musculaire) impliquée dans la production des sons. Certaines consonnes vont nécessiter plus d'énergie que d'autres et seront alors qualifiées de tendues, les autres de relâchées<sup>26</sup>. C'est principalement sur ces deux facteurs que les praticiens de la V-T s'appuient pour la correction phonétique, car ceux sont eux qui seraient mal perçus dans la plupart des cas. Il est cependant utile de remarquer (Borrell, 1997, p. 54) « qu'un nombre important de problèmes perceptifs sont dus en réalité aux distributions des unités phonético-phonologiques dans les deux langues ». Par « distribution », Borrell entend « l'ensemble des contextes et positions dans lesquels une unité ou un complexe d'unités peuvent apparaître ».

Voici ci-dessous une « fiche technique abrégée » du praticien de la V-T, que peuvent se constituer tous les enseignants de langue qui y ont été formés<sup>27</sup>. Notons néanmoins qu'il ne s'agit que d'une lecture possible (la nôtre), et que les ouvrages de références de la V-T peuvent conduire à différentes approches et différents particularismes didactiques. Il ne s'agit donc que d'une « fiche » de présentation très globale, qui

---

<sup>26</sup> Notons cependant que le trait de « tension » a été critiqué à maintes reprises, dans des approches tant fonctionnaliste (Borrell, 1998) que générativiste (Durand, 2005).

<sup>27</sup> Ce sous-chapitre est inspiré essentiellement de Renard (1979), Billières (1988, 1999, 2005), Borrell (1982, 1991, 1993, 1996a, 1996b, 1997), Intraia (2000), ainsi que des notes des séminaires de Maîtrise FLE de M. Billières et A. Borrell, 2000-2001, UTM.

simplifie considérablement l'ensemble des principes et considérations de ses concepteurs et acteurs.

## Fiche simplifiée du Verbo-Tonaliste

La communication est un phénomène global : il faut proposer des situations de communication réelle, en travaillant sur des énoncés complets et non des sons isolés. De plus, les phonèmes sont tributaires de l'intonation et du rythme, et la gestualité (micro et macro), tant de l'enseignant que de l'apprenant, doit être sollicitée, notamment vis-à-vis des problèmes de « tension ». La tâche principale d'entraînement audio-phonatoire de la « méthode » est une tâche de répétition : pour « bien prononcer » un « son », il faut l'avoir « bien perçu ».

### **Principes :**

Motiver l'apprenant ; Pas d'intellectualisation ; Priorité aux éléments prosodiques ; Faire preuve de patience ; Point de départ : la « faute » de l'apprenant. Modèle à répéter : ralentir le rythme, et parfois articuler excessivement, sans pourtant faire de découpage syllabique, de manière à mieux faire percevoir la séquence, puis accélérer progressivement en vue de retrouver un rythme « normal ».

### **Procédure :**

- 1) « Partir » de la faute de l'apprenant et établir un diagnostic, c'est-à-dire évaluer ce qui a été mal ou imparfaitement perçu par l'apprenant et l'identifier (ex : « trop tendu »).
- 2) Proposer un nouveau modèle dans lequel l'élément mal perçu va être « optimisé » pour que le système perceptif de l'apprenant le prenne enfin correctement en compte.

### **Le diagnostic :**

Les erreurs sont généralement surtout dues au timbre (pour les voyelles) et à la tension (pour les consonnes). Il y a parfois aussi d'autres problèmes (distribution, rythme, etc.) mais ce ne sont pas les plus importants dans un premier temps. Pour pouvoir réaliser ce diagnostic, il faut parfaitement connaître le classement des « sons » en fonctions de leurs caractéristiques de timbre et de tension (car il faut réagir immédiatement en classe). Donnons deux exemples :

Voyelle : Un apprenant espagnol perçoit /u/ au lieu de /y/ car il a favorisé les fréquences « graves » (/u/ plus « sombre » que /y/). Il faudra donc faire davantage apparaître les fréquences « claires ».

Consonne : Un apprenant allemand qui dit /s/ au lieu de /z/ aura perçu/répété avec trop de tension (les sourdes sont plus « tendues » que les sonores). Il faudra qu'il prononce de manière moins « tendue ».

### **Les trois procédés fondamentaux :**

- 1) Position optimale : il s'agit de la position du phonème problématique dans la courbe intonative ou bien dans le mot. Par exemple, placer le phonème /y/ en fin d'énoncé interrogatif en pic intonatif favorise la perception des fréquences « claires » pour l'apprenant espagnol (« Il est au sud ? »).
- 2) Combinatoire : il s'agit de l'environnement consonantique et/ou vocalique du segment « défectueux ». Par exemple, les consonnes « claires » /s/ et /d/ dans le mot « sud » favorisent la perception du /y/ pour l'apprenant espagnol.
- 3) Prononciation nuancée : l'enseignant propose une variation du signal initial (« déformer » le modèle acoustique) de manière à mieux faire percevoir ce qui a été mal perçu, ce qui implique une « déformation » en « proportion inverse » : si le « son » a été perçu trop « clair », le modèle « déformé » doit être plus « sombre » que le modèle initial. Pour prendre un exemple « grossier », on peut considérer le cas d'un apprenant espagnol qui répète « lé fé » au lieu de « le feu » : l'enseignant lui propose alors un modèle du type « lo fo ».

Ces trois procédés peuvent en outre être combinés de manière à fournir aux apprenants un modèle « optimal ».

### **Pour corriger le rythme et l'intonation :**

- 1) Utilisation de suites logatomique. Trois types : « mumming sound », « dadada », la « bouche pleine ».
  - 2) Découpage régressif (pour l'intonation) ou progressif (pour le rythme).
  - 3) Utilisation de la gestuelle (scansion de la main, décompte syllabique de la main, tracé intonatif de la main ; le buste relâché ou tendu).
-

Nous ne détaillerons pas davantage ici les exemples ou le détail des techniques, qui sont abondamment commentés et développés dans les ouvrages de référence (en particulier Renard (1979) et Intravaia (2000)).

Comme cela apparaît donc ici, la méthode V-T, tout en s'appuyant sur les données de la phonétique acoustique et articulatoire (Renard, 1979), est un ensemble de principes et de techniques d'ordre *didactique*, avec une terminologie et des concepts qui lui sont, dans une certaine mesure, propres, destinée à être utilisée par des enseignants de langue préalablement formés, et non des spécialistes de phonétique ou de phonologie *a priori*. C'est précisément cela qui lui a assuré un succès certain au cours des quarante dernières années, sous sa forme originale ou à travers d'autres approches, notamment au Japon.

Sans mentionner ici tous les avantages de la V-T, on doit cependant en souligner certains, que nous résumons ci-dessous :

- Son intégration à une méthodologie d'apprentissage général (SGAV), qui lui permet de ne pas être « déconnecté » de l'apprentissage lexical et grammatical, tout en pouvant s'insérer dans d'autres approches didactiques (approches « communicatives »).
- Son insistance sur la perception auditive permet de se détacher d'une conception trop strictement articulatoire de la « prononciation ».
- Son caractère « différencié » : l'intérêt de la phonétique corrective individuelle est qu'elle permet de répondre à des problèmes différents : si certains étudiants n'ont pas de problème avec le /R/ alors ceux-ci n'ont pas envie de perdre leur temps sur ce point en classe. Il se pose donc une certaine nécessité d'individualisation. Ainsi dans l'une de nos classes, certains étudiants présentaient des problèmes tout à fait différents :
  - Etudiant 1 : [u] prononcé de manière trop « claire » (tel que cela est défini dans la Verbo-Tonale), ressemblant ainsi à un [y] ;
  - Etudiant 2 : [t] prononcé [t<sup>h</sup>] aspiré (influence de l'anglais ?) ;

- Étudiant 3) : difficultés avec les couples B/V ; R/L et EU/U.

Dans cet exemple, les deux premiers sembleraient avoir des problèmes d'ordre plutôt phonétique (allophones, qualité phonétique), tandis que le dernier semblerait ne pas encore avoir acquis certaines oppositions phonologiques essentielles du français.

- Son utilisabilité et son adaptabilité : une fois formés, les enseignants disposent d'un ensemble de techniques qu'ils peuvent utiliser à tout moment dans leurs classes, sans exigence temporelle ou matérielle particulière. Elle permet par ailleurs de s'adapter immédiatement aux productions des apprenants, quelle que soit leur(s) langue(s) première(s).

Pratiquant nous-mêmes quasi-quotidiennement la phonétique corrective, nous avons ainsi pu observer les atouts et les limitations de la V-T. Ni la version auditive, ni la version articulatoire de la phonétique corrective ne nous semble pouvoir être rejetée catégoriquement : il s'agit de s'adapter au cas par cas. Ainsi, en français, pour les étudiants japonais, certains problèmes nous semblent pouvoir être mieux traités en termes auditifs (notamment pour les voyelles), alors que d'autres problèmes, en particulier le /R/ ou le /v/ par exemple nous semblent pouvoir bénéficier d'une approche (complémentaire ou principale) en termes articulatoires. C'est la raison pour laquelle il semble nécessaire d'être formé aux deux approches, afin d'adapter son enseignement au public et au cadre.

C'est précisément cette pratique de la V-T qui nous a conduits à nous interroger sur certaines de ses limitations, que nous mentionnons brièvement ci-dessous :

- Correction phonétique ou apprentissage phonético-phonologique : la V-T est une méthode de « correction » (ou « d'intégration ») phonétique. Coupée de son cadre méthodologique d'origine (SGAV), le pendant quelque peu « négatif » de son adaptabilité, que nous avons précédemment présentée comme un atout, est

qu'elle ne propose pas véritablement de démarche planifiée d'enseignement/apprentissage du système phonético-phonologique de la langue à apprendre, notamment vis-à-vis des aspects les plus phonologiques (par exemple la syllabation) de ce système.

- L'intellectualisation : le rejet de « l'intellectualisation » (Renard, 1979) de l'apprentissage se trouve quelque peu en porte-à-faux vis-à-vis des nombreuses études mettant à jour l'intérêt, pour l'apprentissage linguistique, de développer les capacités métalinguistiques des apprenants. L'existence d'une activité métalinguistique, souvent comparatiste, chez les apprenants adultes est en effet aujourd'hui indiscutable, et nous avons nous-mêmes constaté dans nos classes que la mise en focus attentionnel explicite de certaines difficultés, *via* une certaine « intellectualisation », pouvait conduire à des « prises de conscience », suite auxquelles les apprenants amélioreraient notablement leurs productions.
- La modalité visuelle : si la V-T insiste sur l'utilisation de la gestualité faciale et corporelle (auxquelles nous recourons abondamment avec succès dans nos propres classes), on peut s'interroger sur le manque de formalisation ou de systématisation de telles procédures, en particulier lorsque l'on réalise que ce qui relève de la gestualité chez l'enseignant constitue bien souvent un ensemble de stimuli visuels (bucco-faciaux mais aussi corporels) dynamiques pour les apprenants. La modalité visuelle semble donc, sinon négligée, du moins manquer de considération dans les principes de la V-T, alors que la perception du langage est aujourd'hui considérée comme fondamentalement audiovisuelle (Massaro, 2001). Il s'agit donc ici davantage d'un « manquement » que d'un écueil.
- Orthographe et méthodologie : ce point rejoint celui que nous avons évoqué en premier lieu. Alors que la V-T était intégrée à la méthodologie SGAV, qui différait l'introduction de l'écrit en classe de langue (Rivenc et Boudot, 1962),



de manière à permettre l'apprentissage du système phonético-phonologique du français sans « interférence de l'écrit », la méthode V-T en elle-même n'accorde explicitement aucune attention à l'apprentissage du système *phonographémique* de la langue à apprendre, et néglige de ce fait un facteur considérable dans celui du système phonético-phonologique. Les séances de correction phonétique voient ainsi parfois leurs résultats réduits à néant en raison d'un manque de cohérence méthodologique dans l'enseignement, usant du support écrit sans se soucier de l'influence que celui-ci peut exercer sur l'acquisition du système sonore de la langue à apprendre.

- Transfert inter-tâche des compétences : à la suite de ce qui vient d'être présenté, on peut en effet s'interroger sur la portée d'une correction phonétique sporadique qui, si elle ne fait pas l'objet d'une attention particulière de la part des apprenants, peut conduire à de bons résultats pendant la correction (selon le principe du psittacisme), mais ne débouche sur aucun apprentissage à long-, voire à moyen-, terme, une fois encore en raison du manque de cohérence méthodologique forte, entre les différentes activités d'apprentissage proposées. Les exemples de production en tâches de répétition et de lecture à voix haute que nous avons présentés précédemment illustrent bien cette préoccupation.

Ainsi, tout en reconnaissant les mérites d'une méthode qui a « fait ses preuves » et que nous utilisons nous-mêmes, il semble important de continuer à tenter d'en améliorer les principes et les techniques, en soulignant en particulier deux points : le rapport à l'orthographe d'une part, et le rapport à l'apprentissage phonologique d'autre part.

### **III.3.3. Enquêtes : représentations et motivations concernant la prononciation du français au Japon**

Afin de mieux connaître la situation de l'enseignement/apprentissage de la « prononciation » du français au Japon, et plus particulièrement les facteurs d'ordre

*subjectif* (représentations et motivations, tant des apprenants que des enseignants), nous avons effectué trois enquêtes au Japon<sup>28</sup>, à l'aide de questionnaires portant sur le thème de l'apprentissage du français et de sa prononciation. La première a été conduite auprès d'apprenants adultes (n = 31) suivant des ateliers de correction phonétique en école privée ; la seconde a été conduite auprès d'enseignants français (n = 32) et japonais (n = 21) de français, tant en milieu universitaire qu'en école de langue privée ; la troisième a été conduite auprès d'étudiants universitaires japonais de français (n = 273). Nous ne présentons ci-dessous que les parties de ces enquêtes les plus pertinentes pour notre étude.

### III.3.3.1. Enquête 1 : Apprenants en atelier de correction phonétique en école privée (n=31)

La première enquête a été effectuée auprès d'apprenants de français de profils variés dans une école de langue privée. Ces apprenants avaient suivi les ateliers de correction phonétique que nous proposons, basés sur la méthode Verbo-Tonale de correction phonétique, et qui incluaient également des tâches de lecture à voix haute. Il s'agissait donc d'étudiants manifestant un certain intérêt pour la « prononciation » : soit intéressés par le sujet et souhaitant se perfectionner, soit peu au fait du sujet mais conscients de leurs lacunes en la matière. L'effectif de ces ateliers était limité à 10 étudiants, et chaque atelier durait 50 minutes. Les questionnaires ont été remplis par les étudiants à la fin de la session. Nous avons effectué ces enquêtes durant deux sessions :

- Session 1 (50 minutes/semaine pendant 12 semaines) (14 questionnaires).
- Session 2 (50 minutes/semaine pendant 6 semaines) (17 questionnaires).

Nous avons pu ainsi récolter 31 questionnaires. Les questions avaient été traduites en japonais, puisque la grande majorité des étudiants était de niveau faux-débutant. Nous

---

<sup>28</sup> Une partie des enquêtes 2 (enseignants japonais) et 3 (étudiants universitaires japonais) a été conçue et conduite par N. Akashi. Nous remercions l'ensemble des collègues et étudiants qui y ont pris part. Une partie des résultats de ces travaux communs ont été présentés dans Detey et Akashi (2003).

présentons et commentons ci-dessous les résultats qui nous paraissent les plus pertinents.

<b>Question</b>	<b>Pas de réponse</b>	<b>Non</b>	<b>Un peu</b>	<b>Assez</b>	<b>Oui</b>	<b>Très/Beaucoup</b>
Vous avez amélioré votre prononciation en classe ?	0	0	32%	35%	30%	3%
Vous avez amélioré votre capacité d'écoute en classe ?	6%	0	46%	6%	36%	6%
Vous avez amélioré votre lecture à voix haute en classe ?	0	3%	39%	32%	26%	0

Par-delà la traditionnelle « modestie » culturelle nippone, deux points méritent d'être relevés : a) le lien entre prononciation et « écoute », qui a été perçu par environ 42% d'entre eux, alors qu'il ne l'a pas du tout été par les 6% qui n'ont pas répondu à la question, et b) le lien entre prononciation et « lecture » qui a été bien moins perçu que le premier. De tels résultats pourraient suggérer que l'enseignant devrait explicitement sensibiliser les apprenants à ces deux liens.

<b>Dans la classe, vous aimez/ vous pensez que c'est important (de) :</b>	<b>Pas de réponse</b>	<b>Non</b>	<b>Un peu</b>	<b>Oui</b>	<b>Beaucoup</b>
Répéter ?	9%	0	0	42%	49%
Lire à voix haute ?	9%	3%	0	35%	53%
Que la pièce soit un peu sombre ?	6%	0	24%	58%	12%
Que le professeur s'occupe de vous personnellement ?	6%	0	3%	27%	64%
L'utilisation du corps ?	6%	0	24%	35%	35%
L'utilisation du rythme ?	3%	0	0	27%	70%

L'utilisation de la mélodie ?	16%	0	3%	20%	61%
Que le professeur se déplace et bouge ?	12%	0	3%	46%	39%
Que le professeur fasse des gestes avec ses bras et ses mains ?	9%	0	3%	39%	49%
Que le professeur vous montre sa bouche en prononçant ?	6%	0	6%	26%	62%
Que le professeur explique ce qu'il fait (par exemple : « Ici vous avez un problème avec le "V" », « Faites attention au rythme », etc...) ?	18%	3%	12%	32%	35%
Que le professeur imite une mauvaise prononciation (par exemple : en exagérant un défaut d'étudiant) pour montrer la différence ?	10%	0	12%	49%	29%

De ce tableau, on peut retenir :

- Que l'activité de répétition n'est pas perçue comme « rébarbative » par ces apprenants, contrairement à de nombreux enseignants ;
- Que l'importance de la lecture à voix haute n'est pas perçue par tous les apprenants ;
- Que, pour des apprenants japonais en particulier, l'obscurité de la salle est appréciée, probablement en tant que facteur de désinhibition « psycho-affective » ;
- Que la personnalisation de la correction phonétique est globalement appréciée par les apprenants ;
- Que l'utilisation du corps et de la gestualité est perçue positivement ;
- Que les apprenants sont sensibles à la dimension prosodique ;
- Que les apprenants font usage des indices visuels labio-faciaux ;
- Qu'une approche explicitement métalinguistique de la correction n'est pas appréciée par tous ;
- Que la répétition de la production erronée des apprenants peut être utile aux apprenants.

Lorsque l'on demande aux apprenants de classer les cinq éléments les plus importants, d'après eux, pour la correction phonétique, voici ce que l'on obtient :

<b>Classement :</b>	<b>Eléments les plus importants d'après les apprenants</b>
1	Répéter
2	Lire à voix haute / Que le professeur s'occupe personnellement de l'apprenant
3	L'utilisation du rythme
4	L'utilisation de la mélodie
5	Que le professeur montre sa bouche en prononçant

Ici encore apparaît l'importance de la prosodie et des indices visuels labio-faciaux. D'autres questions ont alors été posées aux étudiants :

<b>Question</b>	<b>Réponses</b>				
La prononciation, cela concerne	Seulement la bouche 0		L'oreille 0	Les deux 100%	
Quand vous prononcez un mot ou une phrase, est-ce que vous avez une image dans la tête ? <sup>29</sup>	Pas de réponse 6%	L'articulation de la bouche 26%	La phrase écrite 29%	Des règles de prononciation 35%	Une autre image 26%

Si les apprenants semblent tous être conscients du lien entre production articulaire et perception auditive<sup>30</sup>, les réponses sont en revanche beaucoup plus variées en ce qui concerne la « visualisation » associée à la répétition ou à la production d'un énoncé oral. Alors que le taux de représentation orthographique est assez élevé, on peut s'étonner du taux premier des « règles de prononciation » : de quelles règles s'agit-il ?

<sup>29</sup> Les étudiants pouvaient choisir plusieurs réponses, d'où les pourcentages.

<sup>30</sup> Ce qui doit certainement être rattaché au fait que l'enseignant, dans le cadre de la présentation de la méthode, a fortement insisté sur ce lien.

Lorsque l'on interroge les apprenants sur leurs principales stratégies de recherche de la prononciation d'un mot, voici le classement que l'on obtient :

<b>Pour connaître la prononciation d'un mot.....</b>	<b>Classement</b>
Vous cherchez dans le dictionnaire (transcription)	1
Vous comparez avec d'autres mots français que vous connaissez	2
Vous prononcez comme une autre langue que vous connaissez (anglais, etc.)	3
Vous prononcez comme dans votre langue maternelle (japonais, etc.)	4 -5
Autre :.....	6

Le dictionnaire constituerait donc une source d'information indispensable : on peut alors se demander si le dictionnaire consulté utilise l'alphabet phonétique international et si les apprenants maîtrisent ce dernier, sans quoi l'influence du code écrit pourrait ici s'avérer importante. Il faut néanmoins noter que deux personnes ont choisi « Autre » comme première stratégie, et que les étudiants qui ont également indiqué « Autre » ont ajouté les commentaires suivants (traduits du japonais) :

- Etudiant 1) : « Si la phrase est présentée visuellement (écrite), j'utilise des correspondances graphies - phonies que j'ai apprises ; mais je vérifie toujours quand même dans le dictionnaire ; de plus si la personne qui prononce est en face, alors je regarde soigneusement sa bouche » ;
- Etudiant 2) : « Je compare la transcription phonétique avec l'alphabet ou le mot écrit et je l'apprends » ;
- Etudiant 3) : « Je prononce exactement comme je l'entends » ;
- Etudiant 4) : « Je visualise la forme de la bouche et de la langue ».

Lorsqu'on leur demande s'ils aimeraient avoir davantage de cours de « prononciation » et de « lecture à voix haute », les réponses sont évidemment très positives. Quand on leur demande de préciser le *type* de cours qu'ils aimeraient avoir, les réponses sont les suivantes :

<b>Vous préféreriez avoir.....</b>	<b>Classement</b>
Des cours pratiques (répétition et correction)	1 (51%)
Pratiques et théoriques	2 (41%)
Un peu théorique (explications et exemples)	3 (8%)

Si les cours « pratiques » semblent préférés par ces étudiants, une grande proportion d'entre eux souhaiterait y associer une approche quelque peu « théorique ». Il faut cependant également noter que certains étudiants ont choisi le cours « Théorique » en premier (1 choix), d'autres ont choisi « Pratique et Théorie » en dernier (1 choix), et d'autres encore ont choisi le cours « Pratique » en dernier (2 choix).

En ce qui concerne le rôle du support écrit dans la prononciation, voici l'appréciation des étudiants :

<b>Question</b>	<b>Réponses</b>		
C'est plus facile de répéter une phrase....	Pas de réponse	Prononcée et écrite	Prononcée seulement
	10%	58%	32%

Alors que 10% d'entre eux sembleraient ne pas être conscients de la différence entre formes orales et écrites, 58% d'entre eux considèrent le support écrit comme une aide pour la répétition d'un énoncé, tandis que 32% d'entre eux estiment que ce support peut entraver la tâche. Ces trois nombres nous semblent révélateurs du manque de conscience de la majorité des étudiants de la distinction entre forme orale et forme écrite, ainsi, sinon surtout que du rôle précis de l'écrit dans cette impression de « facilitation ». Un tel écart se retrouve également dans la conscience du lien entre « prononciation » et « grammaire », c'est-à-dire essentiellement entre phonétique/phonologie et morphosyntaxe, puisque environ 42% d'entre eux ne perçoivent pas ce lien :

Pour vous, la prononciation c'est lié à la grammaire ?	Pas de réponse	Oui	Non
	3%	58%	39%

En ce qui concerne le lien entre « prononciation » et « lecture », ce manque de conscience apparaît également dans les pourcentages suivants, puisqu'on pourrait penser que 42% des apprenants supposent lire en français sans faire appel à la « prononciation » :

Quand vous lisez en français, vous articulez silencieusement avec la bouche ?	Pas de réponse	Oui	Non
	0	58%	42%

Les pourcentages suivants semblent confirmer notre impression d'un manque de conscience des différences entre l'oral et l'écrit en français, puisque 20% pensent que les deux sont identiques, et 67% pensent que « bien prononcer » suffit à « bien lire à voix haute ».

	Pas de réponse	Oui	Non
Le français écrit et oral, c'est identique ?	3%	20%	67%
Quand on sait prononcer, c'est facile de lire à voix haute ?	3%	67%	30%
Quand on sait lire à voix haute, c'est facile de prononcer ?	5%	63%	32%

Lorsque l'on demande enfin à ces étudiants de classer les « sons » les plus difficiles pour eux en français, voici ce que l'on obtient :



<b>Classement des sons difficiles</b>	<b>Rang</b>
R	1
B et U	2
V	3
L	4
EU	5
AN	6
ON	7
IN	8
OU	9

Il est alors intéressant de constater que :

- Le « son » B est en deuxième position, avant le son V, alors qu'il fait partie de l'inventaire segmental du japonais, contrairement à V ;
- Le « son » L est en quatrième position, alors qu'il ressemble fortement au /r/ japonais, dont il est un allophone ;
- Le « son » OU est en dernière position alors qu'il entraîne de nombreuses erreurs phonétiques, en perception et en production, notamment vis-à-vis du EU mais aussi du U.

Il est très intéressant, pour le didacticien, de comparer ainsi les représentations métalinguistiques de l'apprenant à celles de l'enseignant, formé ou non en phonétique, car une telle comparaison, en cas d'écart, comme ici, encouragerait à mieux informer/former les apprenants sur les phénomènes en jeu, et à sensibiliser l'enseignant aux préoccupations, et donc aux objectifs d'apprentissage, de ses étudiants.

Il faut cependant noter, pour conclure cette sous-partie, que les participants à cette enquête, de par leur démarche même de participation à une classe de correction phonétique, présentaient un intérêt relativement particulier pour cette dimension de

l'apprentissage de la langue. Ils étaient donc déjà très attentifs à leur prononciation et ils avaient en outre généralement surmonté la plupart des difficultés rencontrées par les débutants. Ces résultats, touchant à la subjectivité des apprenants, sont donc à prendre pour ce qu'ils sont, et ne peuvent être naïvement généralisés.

III.3.3.2. Enquête 2 : Enseignants français (n = 32) et japonais (n = 21) de français (total=53)

Cette enquête<sup>31</sup> a été conduite auprès d'enseignants de profils et de parcours très différents, enseignant tant en école de langue qu'en université. Notre objectif était d'aborder à la fois la question de la formation des enseignants en didactique du FLE, et plus particulièrement en phonétique corrective, et celle des représentations et des pratiques concernant l'enseignement de la prononciation du français aux Japonais. En voici les résultats les plus remarquables :

Enseignants		Français (32)	Japonais (21)
Avez-vous obtenu des diplômes d'appellation FLE ?	Aucun	59%	76%
	Licence	0%	5%
	Maîtrise	28%	5%
	DEA	6%	14%
	Doctorat	3%	5%
	Sans réponse	3%	0%
Souhaiteriez-vous faire des études (complémentaires) dans le domaine du FLE ?	Oui	56%	71%
	Non	40%	24%
	Sans Réponse	3%	5%

---

<sup>31</sup> Plusieurs réponses ont parfois été indiquées, d'où la variation dans ces pourcentages, qui doivent être considérés comme des indicateurs et non comme des « valeurs de vérité ».

Le travail sur l'oral est-il suffisant au Japon, selon vous ?	Suffisant	12,5%	14%
	Il faut absolument l'augmenter	69%	43%
	Il faudrait l'augmenter, mais ce serait difficile	15,5%	33%
	Sans réponse	3%	10%
Pour vous, la prononciation en français c'est important	Oui	97%	90%
	Non	3%	10%
Pour vous, la prononciation, cela concerne surtout	La bouche et l'articulation	6%	5%
	L'écoute et la perception sonore	12%	10%
	Les deux	81%	81%
	Sans réponse	0%	5%
Avez-vous eu une formation en phonétique générale et/ou phonologie ?	Oui	47%	47,5%
	Non	53%	47,5%
	Sans réponse	0%	5%
Avez-vous eu une formation à l'enseignement de la « prononciation » ?	Oui	25%	52, 5%
	Non	75%	47, 5%
Donnez-vous des cours de correction phonétique ?	Oui	25%	10%
	Non	75%	90%
Avez-vous un intérêt dans l'enseignement de la prononciation ?	Oui	90%	76%
	Non	10%	24%
Dans vos classes, faites-vous attention à la prononciation des étudiants ?	Oui	94%	38%
	Un peu	6%	57%
	Rarement	0%	5%
	Jamais	0%	0%

C'est plus facile de répéter une phrase	Prononcée et écrite	53%	48%
	Prononcée seulement	35%	24%
	Pas de différence	9%	19%
	Pas de réponse	3%	9%
Entre des cours pratiques (répétition) et un peu théoriques (explications et exemples), quels sont les plus importants ?	Pratique	37, 5%	9, 5%
	Un peu théorique	6%	19, 5%
	Les deux	56, 5%	67%
	Sans réponse	0%	5%
Faites-vous lire vos étudiants à voix haute ?	Oui	81%	86%
	Parfois	9%	14%
	Rarement	6%	0%
	Jamais	3%	0%
Pensez-vous qu'on peut améliorer sa prononciation en s'entraînant ?	Oui	97%	100%
	Non	3%	0%
En langue étrangère, si on peut déjà communiquer, même si on prononce mal, cela suffit ?	Oui	44%	43%
	Non	50%	52%
	Sans réponse	6%	5%
L'entraînement pour avoir une meilleure prononciation peut-il stimuler la motivation des étudiants pour l'apprentissage de la langue ?	Très possible	59%	33,5%
	Un peu	31%	33,5%
	Peu possible	10%	18%
	Absolument pas	0%	5%
	Sans réponse	0%	10%

Suite aux commentaires que nous avons apportés aux résultats de l'enquête précédente, il ne semble pas indispensable pour nous d'interpréter plus avant ces résultats, qui, dans une certaine mesure, « parlent d'eux-mêmes ». Rappelons en effet que ces enquêtes ne visaient qu'à donner un panorama général de la situation de l'enseignement/apprentissage de la prononciation du français au Japon. Pour achever ce panorama, nous avons donc effectué une enquête du même type auprès d'étudiants très variés dans plusieurs universités japonaises, apprenant généralement le français

comme deuxième ou comme troisième langue étrangère, en incluant également des questions plus générales relatives à leur apprentissage du français.

III.3.3.3. Enquête 3 : Apprenants de français en université (n = 273)

Voudriez-vous travailler à l'étranger dans le futur ?	Oui : 52% Non : 47% Sans réponse : 1%
Quelles sont vos motivations pour apprendre le français ?	Le français va servir pour mon travail : 10% Je m'intéresse à la langue française : 39% Je m'intéresse à la culture française : 38% Je n'avais pas d'autre choix : 8% Autres raisons : 12% Sans réponse : 1%
Pour vous, la prononciation en langue étrangère, c'est important ou pas ?	Oui : 89% Non : 10% Sans réponse : 1%
Pour vous, la prononciation, cela concerne surtout	La bouche et l'articulation : 21% L'écoute et la perception sonore : 18% Les deux : 61%
Quand vous apprenez une langue étrangère, faites-vous des efforts pour bien prononcer ?	Oui : 64% Non : 31% Sans réponse : 5%
Voudriez-vous suivre un cours spécial de prononciation ? <sup>32</sup>	Oui : 60% Non : 40%
Vous voudriez qu'on corrige votre prononciation ou non ?	Oui : 56% Non : 44%

<sup>32</sup> Pour les deux dernières questions, seule la moitié environ des questionnaires (130/273) a pu être comptabilisée, ce qui est non négligeable.

Ici encore, notre objectif ne sera pas d'interpréter ces pourcentages, qui sont naturellement fortement tributaires des imperfections méthodologiques de l'enquête et du questionnaire. Cependant, ils permettent à la fois d'obtenir le panorama que nous évoquions, mais également de réfléchir sur les questions soulevées par cette enquête (comme par exemple les résultats des deux dernières questions).

Ayant ainsi achevé notre descriptif général de la situation de l'enseignement de la prononciation du FLE au Japon, nous pouvons à présent revenir à notre problématique initiale : la perspective didactique de notre étude étant clairement posée, nous pouvons aborder le volet linguistique de celle-ci, en l'occurrence phonético-phonologique, en établissant puis comparant les deux systèmes phonético-phonologiques et phonographémiques en présence, à savoir celui du français et celui du japonais.

#### **IV. Transition**

Dans cette partie, nous avons exposé la perspective transversale dans laquelle nous avons voulu placer nos travaux, avec une visée, non pas applicationniste, mais bien intégrationniste, destinée à alimenter l'approche des problèmes d'enseignement/apprentissage de la parole en langue étrangère. Cette transversalité se pose :

- o D'une part au niveau des domaines concernés : phonologie, psycholinguistique et didactique des langues et des cultures ;
- o D'autre part au niveau de l'orientation des travaux : orientation *théorique*, si l'on songe aux modèles phonologiques et psycholinguistiques utilisés ; orientation *empirique*, si l'on songe à l'ancrage expérimental de cette étude ; orientation *pratique* enfin, si l'on songe à la perspective et au cadre didactiques constamment évoqués au cours de nos recherches ;
- o Enfin, au niveau de certaines problématiques propres à chaque domaine, que nous avons essayé de combiner de manière cohérente : développements en interphonologie et en phonologie des emprunts dans le cadre des modèles à

contraintes ; développements de modèles mnésiques et de modèles de lecture ;  
développement des apprentissages multimédiatisés.

Il nous a donc semblé important de maintenir ce cadre et cette perspective constamment présents dans nos travaux, ce qui explique certains choix, certaines carences potentielles, mais aussi certaines ouvertures. Nous pouvons donc à présent aborder le premier volet de notre étude à proprement parler : le volet linguistique, essentiellement descriptif.

**Partie Linguistique :**  
**Systemes phonético-phonologiques et**  
**phonographémiques français et japonais**

Puisque nous nous intéressons à l'apprentissage de la phonologie du français par des sujets adultes japonais, il convient, dans un premier temps, d'établir, avant de les comparer, puis de les contraster, les systèmes phonético-phonologiques, d'une part, et phonographémiques, d'autre part, du français et du japonais.



## Table des Matières Partielle

I SYSTEMES PHONETICO-PHONOLOGIQUES .....	- 110 -
<i>I.1. Le système français.....</i>	<i>- 112 -</i>
I.1.1. Norme francophone et projet « Phonologie du Français Contemporain ».....	- 113 -
I.1.2. Description classique du français standardisé .....	- 122 -
I.1.2.1 Aspects segmentaux.....	- 124 -
I.1.2.2. Aspects suprasegmentaux.....	- 128 -
I.1.2.2.1. Aspects phonotactiques .....	- 128 -
I.1.2.2.1.1. Distribution et opposition .....	- 128 -
I.1.2.2.1.2 Phénomènes de contacts .....	- 132 -
I.1.2.2.2. Aspects syllabiques : des représentations .....	- 133 -
I.1.2.2.2.1. Extramétricalité .....	- 137 -
I.1.2.2.2.2. Ambisyllabicit�.....	- 137 -
I.1.2.2.2.3. Sonorit�.....	- 138 -
I.1.2.2.2.4. Structure interne de la syllabe .....	- 141 -
I.1.2.2.2.5. Syllabes primitives ou �mergentes ?.....	- 142 -
I.1.2.2.3. Aspects syllabiques : des contraintes (OT et TCRS).....	- 144 -
I.1.2.2.3.1. La Th�orie de l'Optimalit� .....	- 146 -
I.1.2.2.3.2. La Th�orie des Contraintes et des Strat�gies de R�paration.....	- 151 -
I.1.2.2.3.3. OT versus TCRS .....	- 157 -
I.1.2.2.4. Aspects syllabiques : syllabes complexes .....	- 165 -
I.1.2.2.4.1. Syllabes complexes et groupes consonantiques .....	- 166 -
I.1.2.2.4.2. Syllabes complexes et schwa.....	- 174 -
I.1.2.2.4.3. Syllabes complexes, enchaînements et liaisons.....	- 181 -
I.1.2.2.5. Aspects prosodiques.....	- 185 -
I.1.2.2.5.1. L'intonation .....	- 186 -
I.1.2.2.5.2. L'accentuation .....	- 187 -
I.1.2.2.5.3. Le rythme.....	- 189 -
<i>I.2. Le syst�me japonais.....</i>	<i>- 192 -</i>
I.2.1. Aspects segmentaux .....	- 196 -
I.2.1.1. Consonnes .....	- 196 -
I.2.1.2. Voyelles .....	- 199 -
I.2.2. Aspects suprasegmentaux .....	- 201 -
I.2.2.1. Aspects phonotactiques .....	- 201 -
I.2.2.2. Aspects syllabiques : syllabes et mores.....	- 202 -

I.2.2.2.1. More ou syllabe ? .....	202 -
I.2.2.2.2. Mores « spéciales » .....	205 -
I.2.2.3. Aspects syllabiques : emprunts étrangers.....	208 -
I.2.2.4. Aspects syllabiques : syllabes complexes et groupes consonantiques .....	215 -
I.2.2.5. Aspects syllabiques : syllabes complexes et voyelles dévoisées .....	216 -
II SYSTEMES PHONOGRAPHEMIQUES .....	222 -
II.1. <i>Système d'écriture, script et orthographe</i> .....	222 -
II.2. <i>Typologie des systèmes d'écriture et « profondeur orthographique »</i> .....	229 -
II.3. <i>Le système d'écriture français</i> .....	232 -
II.4. <i>Le système d'écriture japonais</i> .....	238 -
II.4.1. Un système pluriel.....	239 -
II.4.2. Les kana .....	242 -
II.4.3. Les romaji .....	248 -
II.4.4. Transcription des emprunts, katakana et alphabet .....	252 -
III. PERSPECTIVE CONTRASTIVE .....	262 -
III.1. <i>Contrastes phonético-phonologiques</i> .....	263 -
III.1.1. Aspects segmentaux.....	263 -
III.1.2. Aspects syllabiques.....	270 -
III.2. <i>Graphie</i> .....	272 -

## **I Systèmes phonético-phonologiques**

De manière générale, lorsque l'on décide d'analyser le système structuro-fonctionnel que met en œuvre tout locuteur natif d'une langue donnée, afin de produire oralement une certaine suite d'énoncés à visée communicative, on doit distinguer plusieurs niveaux d'analyse, qui une fois recombinaisonnés, permettent de disposer d'une description dudit système. Classiquement, les linguistes distinguent quatre niveaux d'analyse : phonologique, morphologique, syntaxique et sémantique. Bien que la tendance actuelle soit de privilégier les interfaces et les interactions entre ces différents niveaux, tendance à laquelle nous souscrivons pleinement, il nous faut néanmoins, dans cette première partie, nous concentrer exclusivement, voire artificiellement, sur le premier niveau, avant de revenir aux interconnexions existant entre ces différents plans d'analyse. Il faut cependant noter que le terme « niveau » peut parfois être source d'ambiguïté, et il faut donc clairement définir les niveaux auxquels on se réfère. Comme le note Goldsmith (1995, p. 5) :

*« On the one hand, the traditional notion of a level derives from having a particular set of tools (syntactic categories, morphological categories, discourse categories, etc.) for analyzing each aspect of an utterance ; levels of this sort (syntactic, morphological, discourse, etc.) could, in principle at least, be said to hold « simultaneously » of an utterance. On the other hand, derivational analysis of phonology posits an underlying and a derived representation, and these distinct and apparently incompatible representations are also referred to as belonging to different levels ».*

La phonologie est habituellement couplée à une autre discipline, la phonétique, qui constitue, d'une certaine manière, le pendant concret et implémenté de la première. La frontière entre les deux domaines est souvent floue, et les deux se renvoient l'un à l'autre en permanence. Les relations entre ces deux domaines généraux peuvent être

illustrées dans un premier temps par les deux tableaux synthétiques<sup>33</sup> suivants que proposent Moeschler et Auchlin (1997, pp. 40-41) :

« Sons par lesquels se manifeste le langage »	Segmental	Phonétique	Acoustique
			Auditive
			Articulatoire
		Phonologie structurale	
	Suprasegmental	Analyse- synthèse de la parole	
		Phonosyntaxe	
		Phonologie intonative	
		Phonostylistique	

Tableau 1 : étude des sons du langage, domaines et disciplines

<b>Approches centrées sur :</b>	<b>Segmental</b>	<b>Suprasegmental</b>
La substance (« phonetic »)	Sons	Mélodies
Les traits formels pertinents (« phonemic »)	Phonèmes	Intonèmes

Tableau 2 : approches étiques [phonétiques] et émiques [phonémiques]

<sup>33</sup> Au demeurant discutables, puisque, par exemple, l'« analyse-synthèse de la parole » dans le tableau 1 relève autant du domaine segmental que suprasegmental.

Décrire la phonologie d'une langue, et plus globalement décrire une langue, est une démarche aux implications fortes. Une telle tentative semble en effet présupposer que ce que l'on souhaite décrire est stable, appréhensible, descriptible, et que l'on dispose des outils adéquats pour opérer une telle description. Il convient également de bien distinguer ce qui relève de la description de ce qui relève de l'analyse, une distinction tout à fait pertinente pour la problématique de l'enseignement/apprentissage. Deux termes, en particulier, permettent de synthétiser la problématique : norme et variation. Il ne faudrait alors pas croire que la phonologie ne s'intéresse qu'au système et à la norme, et la phonétique à la parole et à la variation : au contraire, la langue elle-même est sujette à variation et il convient donc de s'interroger sur ce que l'on peut, veut ou doit décrire, et plus encore lorsqu'il s'agit de l'enseigner.

### **I.1. Le système français**

Quel français décrire et quel français enseigner ? La question aurait pu sembler outrageuse du temps de Vaugelas et elle pourrait sans doute encore surprendre de nos jours. Pourtant, la variation a été l'un des thèmes clés des sciences du langage depuis déjà plusieurs décennies, à commencer, notamment, vis-à-vis de la didactique du français, par la vaste enquête conduisant à l'établissement du *Français Fondamental*, en ce qui concerne les aspects lexicaux et grammaticaux de la langue française dans son usage réel (Gougenheim, Michéa, Rivenc et Sauvageot, 1956). Comme l'indique Laks (1996, p. 64), du point de vue de la linguistique variationniste :

*« La langue n'est plus conçue comme un ensemble homogène, mais fondamentalement hétérogène et instable, tant du point de vue de la compétence du locuteur que du point de vue de l'objet social qui circule dans la communauté [...] comme le proposait déjà Saussure, la description conceptuelle de la langue, d'un locuteur donné ou d'une communauté linguistique concrète n'est qu'une moyenne, elle relève de la construction statistique ».*

En ce qui concerne le domaine de la prononciation (Blanchet, 2002), on doit bien évidemment citer les travaux de Martinet, puis ceux de Walter (1974). Or, dans son ouvrage de 1974, Carton ouvre justement la troisième partie du livre, consacrée à la question de la norme en synchronie, par un double point d'interrogation : « Comment prononcer le français contemporain ? » (pp. 195-224) et « Une norme du français ? ». Cette question perd encore plus de son illusoire trivialité lorsque l'on élargit la « donne hexagonale » à l'ensemble de la communauté francophone mondiale.

### **I.1.1. Norme francophone et projet « Phonologie du Français Contemporain »**

Dans son ouvrage, Carton (1974) s'interroge d'abord sur les critères qui permettraient d'établir une telle norme : régional (en référence à la prononciation tourangelle à laquelle des circonstances purement historiques et politiques ont conféré un prestige bien connu) ou socioculturel (par référence à un « bon usage » quasi-mythique de certaines classes de gens cultivées, non moins quasi-mythiques, comme l'auteur le suggère lui-même, p. 197). C'est dans la distinction entre modèle et usages, ou bien entre compétence et performance, au sens chomskyen, que Carton trouve le moyen, dans une orientation qu'il qualifie de fonctionnelle (*op. cit.*, p. 195), de définir un certain standard, largement diffusé par les médias, qui, s'il n'est pas pratiqué par tous en production, peut en revanche être compris de tous.

Ce n'est donc pas une norme phonétique du français à laquelle il parvient, mais à ce qu'il nomme des « registres de langue comportant divers degrés d'exigences orthoépiques » (Carton, 1974, p. 201) et qu'il présente ainsi :

POPULAIRE (non vulgaire)	Moins élaboré - Moins figé	Nombre minimal de traits caractérisants
FAMILIER (mais correct)	↑	↑
COURANT (moyen)	↓	↓
SOIGNE (soutenu, littéraire)	Plus élaboré - Plus figé	Nombre maximal de traits caractérisants

Les « traits caractérisants » qu'il mentionne sont explicités par la suite, et présentés dans les termes suivants (p. 202) :

- Nombre de liaisons facultatives ;
- Résistance à la neutralisation des voyelles inaccentuées (ex. le premier *o* dans *monotone*) ;
- Refus des harmonisations vocaliques (ex. fermeture de *ai* dans *raidir*) ;
- Nombre d'amuïssements de *e* muets ;
- Refus de l'assimilation (Israël/ Izraël)
- Tension articulaire (netteté des timbres, des articulations, etc.) ;
- Relative égalité syllabique ;
- Refus des traits d'identification régionale (dans certains cas) ;
- Maintien des groupes consonantiques (quatre/ quate)

Carton écrit alors : « Bien prononcer, ce n'est pas du snobisme ! C'est d'abord être intelligible et clair. Que de bafouilleurs ! Combien peu savent lire à haute voix ! » (*op. cit.*, p. 204). Cette conception d'une « bonne prononciation » renvoie au concept d'« orthoépie », présenté comme tel dans le *Dictionnaire de Didactique du Français Langue Etrangère et Seconde* (DDFLES) (Cuq, 2003, pp. 183-184) :

« L'orthoépie ou **phonétique normative** définit la norme de la prononciation des sons d'une langue et de leur enchaînement dans la parole. Elle [...] suppose une connaissance du système graphophonétique de la langue. Elle suppose également une définition de la norme de prononciation dans la langue : l'orthoépie recherche la mise en œuvre des formes phoniques à l'intérieur de certaines marges d'acceptabilité telles que l'utilisateur de la langue sera considéré comme n'ayant pas d'accent particulier, qu'on ne parviendra à le situer ni socialement, ni géographiquement ».

Il est particulièrement intéressant pour nous de noter que cette définition, offerte dans un ouvrage de didactique du français, établit un lien marqué entre « norme de prononciation » d'une part et « connaissance du système graphophonétique » d'autre part. Il semblerait, d'après cette définition, que « pour bien prononcer », il soit nécessaire de savoir « bien lire à voix haute ». Si l'on rapporte cela à l'apprentissage de la langue maternelle, un tel présupposé semblerait quelque peu surprenant, étant donné que la capacité de parole (capacité *acquise*) précède celle de lecture (capacité *apprise*) chez les enfants. Cependant, si l'on rapporte cela à l'apprentissage d'une langue étrangère par des adultes lettrés, le lien entre apprentissage du système phonético-phonologique de cette langue et apprentissage de son « système graphophonétique » semble alors clairement mis en évidence par une telle citation, tout en renvoyant, en partie, à la problématique centrale de notre réflexion. Mais loin de clarifier la question, la définition précédente pose un problème essentiel : comment peut-on apprendre à « bien prononcer » en langue étrangère si l'on ne connaît pas (encore) le système graphophonétique de cette langue, système dont l'apprentissage nécessite justement, en partie au moins, de connaître le système phonético-phonologique que l'on cherche à apprendre. La circularité du raisonnement, pour évidente qu'elle soit, permet de souligner l'importance de cette question.

Notons par ailleurs que le terme « graphophonétique » semble renvoyer à la réalisation *phonétique* des *graphèmes*, alors que le terme *phonographémique*, plus fréquent, nous semble-t-il, dans une certaine littérature psycholinguistique, semble davantage



renvoyer aux correspondances entre *phonèmes* et *graphèmes*. Or, alors que le domaine phonétique est généralement celui de la variation, c'est bien le domaine phonémique (et ainsi phonologique) qui est, dans une certaine mesure, celui de la norme, norme que transcrit, à des degrés divers, l'écrit.

La question du rapport entre norme écrite et norme orale d'une part, et norme et variation phonético-phonologique d'autre part, ne peut donc se résoudre en quelques paragraphes, alors que celles-ci continuent d'être pleinement des thèmes d'actualité dans les domaines concernés. L'influence de l'écrit sur la norme phonologique en particulier, puisque celle-ci est fortement liée à notre problématique, est particulièrement soulignée par Laks (2003b, p. 97 et p. 99) :

*« On voit ainsi apparaître très clairement depuis les années 1920 un effet rétroactif de la forme orthographique sur la forme phonique, effet pour la première fois mis en évidence par Buben [...]. La cristallisation progressive de ce que Morin [...] a justement baptisé Français de Référence ne s'est en effet pas faite sur l'oral mais sur la normation progressive de l'oralisation de l'écrit ».*

Pour Borrell et Billières (1989), l'évolution de la norme phonétique en français contemporain se caractérise par trois étapes qu'ils décrivent ainsi:

1. Le français parisien cultivé
2. Le français standard
3. Le français standardisé

Ces trois normes ne se sont évidemment pas succédées les unes aux autres de manière brutale, mais se sont plutôt chevauchées et concurrencées. Elles sont présentées de manière synthétique de la manière suivante (Billières et Borrell, 1990, p. 28) :

<b>Modèle de référence</b>	<b>Extension géographique</b>	<b>Extension sociale</b>	<b>Influence du modèle</b>
<b>« Français Parisien Cultivé »</b>	Paris uniquement	Haute bourgeoisie uniquement	Jusqu'aux années 60 et au-delà
<b>« Français Standard »</b>	De Paris vers la province	Radio, télévision, journalistes et politiques	A partir des années 60 et de nos jours
<b>« Français Standardisé »</b>	Hexagone	Classes moyennes plutôt favorisées	A partir des années 80

Dans leur article de 1990, les auteurs s'interrogent sur la norme et la variation linguistique, plus particulièrement phonético-phonologique, utilisées dans plusieurs méthodes de français langue étrangère (*Voix et Images de France*, 1960 ; *De Vive Voix*, 1972 ; *C'est le Printemps 1*, 1976 ; *Intercodes 1*, 1978 ; *Cartes sur Table 1*, 1981 ; *Archipel*, 1982 ; *Sans Frontières 1*, 1982 ; *Cartes sur Tables 2*, 1983 ; *Sans Frontières 2*, 1984). Ils indiquent (1990, p. 28) :

*« Pour l'enseignant de FLE, connaître la norme présente un avantage et un inconvénient. Un avantage parce que c'est ce qui doit être enseigné. Un inconvénient dans le cas où le professeur a lui-même un accent l'éloignant de la norme ».*

Nous ne nous arrêterons pas ici sur les implications et les fondements sociaux, pourtant cruciaux, de la norme linguistique. Borrell et Billière notent que se conformer à la norme « est une garantie d'intercompréhension à l'intérieur de la communauté francophone, s'en éloigner entraîne des incorrections qui seront très souvent jugées négativement » (1989, p. 45). Cette norme est « le produit de la pression exercée à un certain moment par un groupe social dominant, plus ou moins étendu » (*op. cit.*, p. 48) et les auteurs reprennent la distinction faite par Houdebine entre normes « objectives »

(établies à partir d'études et d'enquêtes) et « subjectives » (établies à partir du sentiment linguistique des locuteurs) (*op. cit.*, p. 47).

Cependant, comme le démontrent très clairement les travaux sur la variation de l'oral, le tableau présenté plus haut sur les trois normes est nettement réductionniste, voire simpliste, puisque (Laks, 2003b, pp. 97-98) :

*« Dès qu'il interroge les usages attestés, le linguiste se trouve en effet confronté à une variation qui déploie ses figures dans chacune des dimensions du phénomène langagier. [...] La variation ne constitue pas un phénomène externe, marginal ou surimposé à une langue, par ailleurs identifiable comme stable, homogène et parfaitement circonscrite. [...] Les variations diachroniques, générationnelles, géographiques, stylistiques, sociales, sexuelles se superposent et s'interpénètrent de telle sorte que la notion de langue ne constitue pas un objet défini (Saussure) ».*

Dans le DDFLES (Cuq, 2003), l'article « Norme » propose trois principaux types de norme, en prenant soin de rappeler au lecteur « la distinction transversale entre ce qui est du domaine du "normal" (norme = ensemble d'usages) et ce qui est du domaine du "normatif" (norme = ensemble de règles) » (p. 178) :

- La norme objective : « l'ensemble des faits répétitifs et collectifs n'appartenant pas nécessairement au système ou à la structure de la langue [...]. Cette norme peut également être descriptive ».
- La norme prescriptive : elle « définit des modèles et [...] hiérarchise des usages identifiés en description, d'où une fusion parfois gênante [...] entre norme objective et norme prescriptive ».
- La norme subjective : « elle se situe sur le terrain des attitudes et des représentations et consiste à accorder des valeurs esthétiques, affectives ou morales aux formes ».

Bien qu'une telle définition soit discutable, en particulier celle relative à la « norme objective », qui semble consister en une opération de « moyennage », ne permettant donc pas de refléter le système intégré par un locuteur potentiel de la langue, il semble surtout important pour nous d'insister ici sur le fait que ces trois types de normes sont bien souvent abusivement amalgamés par un certain nombre d'enseignants de FLE, peu au fait de la question. Par ailleurs, lorsque l'on consulte l'article « Correct » du DDFLES (Cuq, 2003, pp. 57-58), le lecteur est précisément mis en garde contre l'idéalisation de ses propres productions :

*« La distinction entre le plus et le moins correct s'appuie [...] sur des jugements de valeur qui réfèrent à une norme subjective susceptible de fortes variations. [...] Les productions langagières orales révèlent [...] des formes [...] décalées par rapport aux structures normatives [...] par exemple "il y en a qu'on peut s'en servir" énoncé par un président d'université dans une réunion de travail. [...] Soumises au jugement de ces mêmes locuteurs, [ces "incorrections"] seraient qualifiées d'incorrectes. C'est donc bien à une norme subjective, idéalisée par rapport aux pratiques langagières, que renvoie la notion de correct : est correct ce que l'on croit conforme à l'usage ou au bon usage ».*

Blanchet (2002, p. 47) souligne ainsi le privilège qui a longtemps été accordé à « la » norme dans le domaine du FLE :

*« La sensibilisation à la diversité du français, d'abord de France puis de l'ensemble de l'espace francophone, et, mieux, sa prise en compte dans la conceptualisation de l'objet d'enseignement [...] ne sont apparues que tardivement dans les méthodes de FLE ».*

Le rapport entre norme, variation et enseignement du FLE semble donc être pleinement un sujet d'actualité, puisque certains tentent d'intégrer les apports de la phonétique et

de la phonologie, dans une perspective variationniste, pour alimenter la réflexion à ce sujet et proposer d'enseigner une norme pédagogique « suppléée par une exposition à la richesse de la variation » (Girard et Lyche, 2003, p. 12).

L'une des linguistes les plus célèbres à avoir travaillé sur la norme linguistique en français est H. Walter : comme elle l'écrit (2001), la langue française « est perpétuellement en mouvement ». L'auteur observe ainsi, depuis le début du XX<sup>ème</sup> siècle (p. 8) :

- « 1. qu'en raison de l'adaptation à la vie moderne et à la mondialisation, c'est le lexique qui a connu une évolution spectaculaire [...] ;*
- 2. que la phonologie se dirige lentement mais sûrement, au moins en France, vers un système vocalique qui privilégie les oppositions de timbre et délaisse les oppositions de longueur, et vers un système nasal réduit à trois unités ;*
- 3. que la grammaire connaît des mouvements contradictoires, qui se traduisent d'un côté par l'élimination de certaines formes verbales et, en contrepoint, [...] par une liberté syntaxique de plus en plus généralisée ».*

Il faut noter que Walter a conduit l'une des dernières grandes enquêtes phonologiques portant sur le français, dont les résultats ont servi de références jusqu'à aujourd'hui, avec, il y a quelques années, le lancement de l'enquête *Phonologie du Français Contemporain* (PFC), dirigé par J. Durand, B. Laks et C. Lyche (Durand et Lyche, 2003). La nécessité de telles enquêtes (Laks, 2003a), essentiellement pour la linguistique mais également pour la didactique, apparaît clairement à travers les recommandations de Martinet d'une part, et Labov d'autre part, cités par Blanchet (pp. 41-42) :

- « L'introspection n'est pas un procédé recommandable dans la recherche scientifique » (Martinet)
- « Les linguistes ne peuvent plus désormais continuer à produire à la fois la théorie et les faits » (Labov).

Le projet PFC figure ainsi parmi les grandes enquêtes phonologiques récentes, visant à constituer un corpus oral francophone de grande ampleur. Comme l'écrit Laks dans une présentation succincte qu'il donne du projet (2001, p. 106), « la prononciation du français contemporain dans son usage réel est encore fort peu et fort mal connue ». Il indique que le projet n'est pas sociolinguistique, mais purement de phonologie descriptive, en vue de « recueillir et de décrire la diversité des usages et des systèmes phonologiques du français contemporain » (*ibid.*). C'est donc précisément à une remise en question de l'existence de « la norme » que peut conduire ce projet, puisque, comme l'indique Laks (2003b, p. 97 et p. 99) :

*« En dehors de l'observation directe des pratiques, ce que l'on obtient par introspection, recollection des données standards ou interrogation des traités d'orthoépie n'est autre que le nivellement normatif produit par la tradition grammairienne, c'est-à-dire l'instance de l'écrit dans l'oral. [...] A partir de l'analyse critique des bases normatives et orthoépiques sur lesquelles s'est construite la phonologie actuelle du français, et sur celles de la seule introspection il est apparu nécessaire d'en revenir à l'oral tel qu'il est attesté dans l'usage ordinaire ».*

Parmi les facteurs pris en compte dans la variation, on peut mentionner :

- La géographie (zones urbaine, suburbaine et rurale)
- L'âge
- Le sexe
- La catégorie socio-professionnelle

Pour la géographie, une cinquantaine de points d'enquête est prévue, avec une dizaine d'interviews au minimum à chaque fois. Ces points se situent en France (Lille, Béthune, Lyon, Villeurbanne, Paris-banlieue est, Paris-banlieue ouest, Paris-centre, Orléans, L'Orléanais, Rennes, l'île d'Yeu, Cherbourg, le Cotentin, le Perche, Caen, le Périgord,

Bordeaux, la Corrèze rurale, Toulouse-ville, Toulouse-banlieue, Montpellier, Pézenas, Grenoble, Saint-Etienne, Strasbourg, l'Alsace rurale, Metz, la Lorraine rurale, Rouen, le pays basque, Brives la Gaillarde, Nice, la Provence rurale, Marseille, Dijon), mais aussi hors de l'Hexagone (Belgique, Canada, Louisiane, Afrique francophone). A terme, il devrait donc fournir 500 heures de français oral, transcrit orthographiquement et numérisé, avec une attention particulière portée sur des phénomènes comme la liaison ou bien le comportement du « e » muet, ainsi que sur les inventaires vocaliques. Les outils informatiques permettent en outre d'aligner facilement des fichiers « textes » et « sons ».

Comme on le voit à travers ces enquêtes, le système phonético-phonologique du français n'est pas *un*. Cette délicate question de la variation, si elle apparaît clairement au niveau inter-individuel, se pose également au niveau intra-individuel, conduisant Borrell (1991) à poser la question dans les termes suivants : chaque locuteur français possède-t-il *un* ou *des* système(s) phonologique(s) ? Cette question nous entraîne vers celle de la compétence linguistique des locuteurs et ses diverses manifestations en performance, fonctions des tâches à réaliser.

### **I.1.2. Description classique du français standardisé**

Vouloir décrire le système phonologique du français implique de recourir à certains concepts et à une certaine terminologie, essentiellement issus de la phonologie et de la phonétique générales. L'histoire de la phonologie, intimement liée à celle de la phonétique, est clairement présentée de manière synthétique par Durand et Lyche (2000), ou bien Durand et Laks (2002b). Ce que l'on appelle parfois les *nouvelles phonologies*, tout comme de nombreux travaux récents très techniques et spécialisés en sciences du langage, manquent souvent de diffusion adéquate, et les concepts de base du domaine restent dans la plupart des manuels universitaires destinés aux étudiants de lettres, de langues, de psychologie ou bien même de linguistique, ceux de phonème, d'allophone, de phonotactique et de morphème par exemple. Ainsi, lorsque l'on cherche à obtenir une présentation du système phonético-phonologique du français,

l'on obtient plusieurs présentations qui se recourent et dont nous nous servirons ici dans un premier temps.

Puisqu'il est difficile, dans notre champ d'étude, de trouver une forme sans substance, il convient, lorsque l'on traite des aspects phonético-phonologique d'une langue, de partir de la réalité articulo-motrice et acoustique des continuums sonores qui servent de substance phonique aux énoncés linguistiques. Deux sous-domaines sont alors concernés : la phonétique articulatoire et la phonétique acoustique. Cependant, ce qui nous intéresse particulièrement ici, ce sont les formes que prendra cette substance phonique, formes qui vont se composer et s'assembler de manière à former un système d'unités, qu'elles soient distinctives et/ou significatives, pour reprendre les concepts désormais classiques de première et de deuxième articulation forgés par Martinet. Lorsque l'on dépasse le niveau de la substance phonique, et que l'on opère à des niveaux d'analyse plus abstraits, dans lesquels des phénomènes de catégorisation spécifiquement linguistiques entrent en jeu, alors c'est le domaine de la phonologie qui, en partie au moins, prend le relais. La question est alors : comment segmenter et analyser le flux sonore de la parole en unités pertinentes, qui permettent, tant en production qu'en compréhension, d'organiser du sens ? Par ailleurs, il n'est pas exclu que plusieurs types d'unités soient pertinents, en fonction des tâches à accomplir et des langues en question (trait *vs.* phonèmes *vs.* syllabes. *vs.* mores). Certains auteurs ont ainsi avancé que les locuteurs disposaient d'un inventaire syllabique phonétique en production orale (Levelt, 1989 ; Levelt et Wheeldon, 1994 ; De Bot, 1992). A ce propos, il convient de noter la remarque de Katz et Frost (1992) qui insistent sur la nécessaire variation entre les langues. Alors que le japonais présente un nombre de syllabes limité (moins de 113 d'après les auteurs (p. 68)) qui rend possible la conception d'un syllabaire (les kana), il n'en est pas de même pour l'anglais par exemple (p. 69) :

*« English is said to have at least 8000 syllables in its phonology, compared to fewer than 1300 for Chinese [...]. Eight-thousand is far too large a number for*



*an effective syllabary. For English, an alphabet, representing phonemes, is more efficient for learning to read and write ».*

Une fois ces unités identifiées, il n'est pas exclu que l'on puisse en proposer différentes descriptions, valides les unes autant que les autres (modèles linéaires vs. non-linéaires par exemple). Par ailleurs, une fois déterminé le support représentationnel (les dites unités pertinentes), il convient d'observer les processus qui s'y appliquent, variables en fonctions du contexte et des situations linguistiques. Commençons donc par la question des unités pertinentes. Dans la tradition des modèles linéaires phonématiques, il est habituel de distinguer, d'une part des aspects segmentaux, et d'autre part des aspects suprasegmentaux.

#### 1.1.2.1 Aspects segmentaux

Tout en tenant compte des réserves et des problèmes soulevés précédemment concernant la question de la norme phonético-phonologique en français, la plupart des manuels universitaires s'accordent pour dire que le français comporte entre 17 et 18 phonèmes consonantiques, et entre 16 et 10 phonèmes vocaliques. Il faut aussi mentionner les 3 semi-consonnes (ou semi-voyelles) du français, qui n'ont pas systématiquement de valeur phonologique (Léon, 1992, p. 70).

Du point de vue de la phonétique articulatoire, il est habituel de proposer, pour les consonnes, le tableau classificateur suivant, basé sur une description articulatoire relativement statique (Mode et Lieu d'articulation) de la réalisation phonétique des phonèmes :

		<i>Lieu d'articulation</i>								
		Bilabiales	Labio-dentales	Apico-dentales	Alvéolaires	Apico-palatales	Pré-Palatales (Médio)	Vélaires (Dorso) —	Uvulaires	
<b>Mode d'articulation</b>	<b>Plosives</b>	<b>Non voisées</b>	p		t				k	
		<b>Voisées</b>	b		d				g	
		<b>Nasales</b>	m		n			ɲ	ŋ	
	<b>Fricatives</b>	<b>Non voisées</b>		f		s	ʃ			
		<b>Voisées</b>		v		z	ʒ			
		<b>Liquides latérales</b>				l				
		<b>Liquides vibrantes</b>								R
	<b>Semi-consonnes</b>	<b>Non labiales</b>						j		
		<b>Labiales</b>						ɥ	w	

Les consonnes françaises se caractérisent notamment par trois aspects :

- *L'anticipation vocalique* (les organes phonateurs préparent la production de la voyelle qui suit, en particulier les lèvres et la langue),
- *La tension articuloire* (situation inverse : lorsque la consonne est finale, la voyelle précédente détermine, en partie, la position des organes durant l'articulation finale).

- *La détente finale* (les consonnes se prononcent complètement, c'est-à-dire qu'il y a apparition d'une détente finale pour les occlusives et les nasales en finale, et un maintien de la sonorité pour les consonnes sonores en finale).

Pour les voyelles, un tableau possible est le suivant :

	Voyelles antérieures				Voyelles postérieures			
	Ecartées		Arrondies		Ecartées		Arrondies	
	Orales	Nasales	Orales	Nasales	Orales	Nasales	Orales	Nasales
<b>Très fermées</b>	i		y				u	
<b>Fermées</b>	e		ø				o	õ
<b>Moyennes</b>								
<b>Ouvertes</b>	ɛ	ẽ	œ	œ̃			ɔ	
<b>Très ouvertes</b>	a				ɑ	ã		

Il faut aussi indiquer que l'on regroupe parfois les 16 voyelles du français en archiphonèmes, qui sont au nombre de 10 : /i E A O u y OE E ã õ/. Quant au « schwa », symbolisé par le signe ə et dont la réalisation typique est celle d'une voyelle antérieure moyenne arrondie, nous ne l'incluons pas ici, car il constitue un « phonème » à part dans le système, sur lequel nous revenons plus loin.

Les voyelles françaises se caractérisent notamment par :

- Leur *tension articulaire* élevée (tension des muscles phonateurs et maintien des muscles au cours de l'émission qui empêche toute diptongaison).
- Leur *antériorité* (la plupart des voyelles impliquent l'avancement de la masse de la langue dans la cavité buccale).

- Leur *labialisation* (en général, les lèvres jouent un rôle important dans l'articulation des voyelles).

Il faut alors ici prendre soin de maintenir la distinction entre les traits *articulatoires* (voisement, nasalité, lieu d'articulation, etc.), qui relèvent de la phonétique articulatoire, et les traits *distinctifs* correspondants, qui relèvent de la phonologie et qui permettent d'opposer un phonème à un autre (Durand, 1990, 2000a, 2005b). Ces traits distinctifs peuvent être présentés de manière articulatoire (comme ici) ou bien acoustique, avec des traits comme aigu/grave par exemple. Ces traits sont en effet, depuis plusieurs décennies, considérés comme les unités de base des représentations phonologiques, et même s'ils sont envisagés en tant qu'entités psycholinguistiques, ils sont généralement définis en termes de réalisations acoustiques et articulatoires (Clements and Hume, 1995, p. 245). L'émergence des phonologies non-linéaires correspond à une certaine autonomisation des traits vis-à-vis des segments, autonomisation, issue de la phonologie autosegmentale, qui, *via* les classes de traits, a ensuite conduit à un modèle général de l'organisation hiérarchique des traits, souvent représentée en graphe arborescent. Cette organisation donne lieu à deux assertions formulées ainsi par Clements et Hume (1995, p. 251) :

- « Feature values are arrayed on separate tiers, where they may enter into nonlinear (nonbijective) relations with one another ».
- « Features are at the same time organized into hierarchical arrays, in which each constituent may function as a single unit in phonological rule ».

Outre son intérêt phonologique, la caractérisation infra-segmentale en termes de traits permet au didacticien s'intéressant à l'apprentissage phonético-phonologique en langue étrangère d'avoir une approche plus fine de certains phénomènes, comme cela apparaît dans les travaux de plusieurs chercheurs (Archibald, 2000).

Après cette brève présentation classique des phonèmes du français, nous nous intéressons à présent aux phénomènes qui « dépassent » le phonème : il s'agit des aspects suprasegmentaux de l'analyse linguistique.

#### I.1.2.2. Aspects suprasegmentaux

Lorsque l'on s'intéresse aux phénomènes phonético-phonologiques qui dépassent le domaine du segment (qu'il s'agisse de certains rapports entre eux ou bien de phénomènes dont la durée excède celle d'un segment isolé), on en vient aux aspects suprasegmentaux de la parole, laquelle est fondamentalement continue, plutôt que discrète. Nous inclurons ici sous cette étiquette tous les phénomènes qui impliquent plus d'une unité phonémique, même si l'appellation « suprasegmental » renvoie souvent principalement à des aspects strictement prosodiques.

##### I.1.2.2.1. Aspects phonotactiques

Si les sons de la parole ne sont jamais isolés, ils ne présentent également quasiment jamais de comportement tout à fait erratique en ce qui concerne la régularité de leur apparition sur l'axe syntagmatique. Des schémas sont identifiables : il s'agit de leur distribution, parfois également appelée « phonotactique ».

###### I.1.2.2.1.1. Distribution et opposition

A un niveau phonologique, certaines restrictions s'appliquent en effet à la distribution des unités pertinentes, de manière plus ou moins systématique. Parmi ces restrictions, certaines entraînent des comportements d'alternance ou d'opposition entre allophones ou phonèmes, comme l'indique Mohanan (1995, p. 67, note 2) :

*« "Distribution" is the relation between a unit and the environment in which it occurs : the statement "[h] cannot occur at the end of a word in English" identifies a pattern of distribution. "Alternation" refers to the relation between two corresponding units in two related forms : the statement " the nucleus [ai] in divine corresponds to the nucleus [i] in divinity " identifies a pattern of*

*alternation. It is important to distinguish the classification of patterns of distribution and alternation from the classification of formal devices that capture the patterns into rules and constraints. Either rules or constraints can be used for the statement of both distribution and alternation ».*

Nous pouvons, dans un premier temps, simplement considérer la distribution d'un phonème comme l'ensemble des positions qu'il peut occuper au sein d'un mot<sup>34</sup> : initiale— médiale — finale. P. et M. Léon (1976, pp. 44-45) présentent ainsi, en première approximation, la distribution des phonèmes du français :

Voyelles: Chaque voyelle a une distribution complète.

	<b>Initiale</b>	<b>Médiale</b>	<b>Finale</b>
V	+	+	+

Exception: le [ə] caduc ne se trouve jamais à l'initiale et rarement en finale (ex : « Donnez-le. »)

Consonnes : Chaque consonne a une distribution complète.

	<b>Initiale</b>	<b>Médiale</b>	<b>Finale</b>
C	+	+	+

Exception : le [ŋ] ne se trouve jamais à l'initiale ni en médiale.

---

<sup>34</sup> Pour une définition des notions de « mot » et de « mot phonologique » voir (Dell, 1995).

Semi-consonnes :

	<b>Init.</b>	<b>Méd.</b>	<b>Fin.</b>
<b>j</b>	+	+	+
<b>ɥ</b>	+	+	0
<b>w</b>	+	+	0

Cependant l'analyse de la distribution telle qu'elle est envisagée ci-dessus est bien évidemment insuffisante. Elle ne nous donne aucun renseignement sur les contextes phonémiques, sinon phoniques, précis dans lesquels chaque unité de la langue peut ou, et c'est ce qui nous intéresse, ne peut pas apparaître. Par ailleurs, d'un point de vue phonologique, la distribution phonémique d'une langue donnée peut très simplement résulter de la structure syllabique de cette langue (Archangeli, 1997). Nous revenons plus bas sur cette question lorsque nous traitons de la structure syllabique du français.

En ce qui concerne les voyelles, il existe, pour certaines d'entre elles dites « à double timbre », des cas d'oppositions particulières, correspondant à ce que l'on appelle la *distribution complémentaire* (Léon, 1992, p. 85), laquelle est liée aux fonctionnements accentuel et syllabique du français.

- En syllabe inaccentuée (syllabe faible) : il se produit un phénomène qui conduit les voyelles à devenir moyennes. Il ne s'agit cependant pas de réduction vocalique totale qui irait jusqu'à une transformation en voyelle neutre de type « schwa », En français, en effet, même en position atone, les voyelles et les syllabes gardent leur identité. Selon Léon, cité par Borrell (1998, p. 59), « Le français est une langue à articulation tendue par rapport à des langues comme l'anglais où les syllabes inaccentuées se réduisent souvent à une voyelle moyenne indistincte proche du E caduc français ». Borrell nuance cependant l'affirmation selon laquelle le français

est une langue à prononciation tendue et l'anglais langue à prononciation relâchée en s'exprimant en termes de tendance et non d'absolu : « le français actuel manifeste une évolution qui a été dans le sens de plus de tension et l'anglais dans le sens de moins de tension ».

- En syllabe accentuée (syllabe forte) : il faut considérer le caractère ouvert ou fermé de la syllabe. Léon (1992, p. 85) en propose les définitions phonétiques suivantes :
  - Syllabe ouverte : syllabe se terminant par l'ouverture du canal buccal, c'est-à-dire par une voyelle prononcée.
  - Syllabe fermée : syllabe se terminant par une fermeture du canal buccal, c'est-à-dire une consonne prononcée.

Notons alors que le français privilégie la syllabation ouverte (plus de 80% des syllabes sont ouvertes), même si la chute des schwas augmente le nombre de syllabes fermées. Wioland (1991, p. 56) indique les pourcentages suivants, repris ici seulement pour les trois structures les plus fréquentes :

<b>Syllabes ouvertes</b>		<b>Syllabes fermées</b>	
CV	55,5 % (« vie »)	CVC	13,5 % (« toute »)
CCV	14 % (« pris »)	CCVC	2,5 % (« prête »)
V	10 % (« un »)	CVCC	1,5 % (« merde »)

La loi de distribution complémentaire se présente alors ainsi : « dans une syllabe accentuée fermée, la voyelle est ouverte et dans une syllabe accentuée ouverte, la voyelle est fermée » (Léon, 1992, p. 85), même s'il y a plusieurs exceptions à cette loi générale (*op. cit.*, p. 86).



	<b>E</b>	<b>EU</b>	<b>O</b>
<i>Syllabe ouverte</i>	E fermé [e]	EU fermé [ø]	O fermé [o]
<b>Syllabe fermée</b>	E ouvert [ɛ]	EU ouvert [œ]	O ouvert [ɔ]

Notons également pour la suite qu'il ne faut pas confondre *syllabe faible/forte* avec *syllabe légère/lourde* (Mohan, 1995, p. 68, note 14) :

- Syllabe légère : syllabe ouverte (sans coda) avec une voyelle brève.
- Syllabe lourde : syllabe avec, soit une coda, soit une voyelle longue, soit une diphtongue.

#### 1.1.2.2.1.2 Phénomènes de contacts

Comme nous l'avons dit, les phonèmes n'apparaissent jamais seuls sur l'axe de la chaîne parlée. Des phénomènes de contact se produisent, entraînant des modifications phonétiques et ainsi des variantes combinatoires. Parmi celles-ci, on peut citer : l'antériorisation / la postériorisation, la palatalisation, le voisement / le dévoisement, la nasalisation, la labialisation, le relâchement / la tension (Léon, 1992, pp. 70-71 et pp. 91-92). Il s'agit de ce que l'on appelle les phénomènes *d'assimilation*.

Dans le cadre phonétique de la méthode Verbo-Tonale, Renard (1979, p. 86) présente un tableau résumant les phénomènes de modification associés à l'assimilation de la manière suivante :

Influence des voyelles sur les consonnes qui les précèdent

Voyelles Antérieures (/iye/)	Palatalisation
Voyelles Postérieures (/uo/)	Vélarisation
Voyelles Labiales (/yu/)	Labialisation

Influence des consonnes sur les voyelles

labiales, vélares, laryngées, /R/	Assombrissement du timbre
nasales, sonores	Assombrissement
alvéolaires, post-alvéolaires	Eclaircissement
constrictives sourdes	Fermeture
occlusives sourdes	Ouverture
j	Fermeture
ʒ	Arrondissement

Influence des consonnes entre elles (Assimilation régressive)

Sonore précédant sourde	Assombrissement de la sonore
Sourde précédant sonore	Eclaircissement de la sonore

L'assimilation *régressive* est très fréquente en français. Il existe également des cas d'assimilation *progressive*, comme dans le mot « peuple » où le /l/ perd sa sonorité au contact de /p/. Pour les voyelles, il faut également noter les cas d'harmonisation vocalique (Léon, 1992, p. 85) et de diffusion de la nasalité.

1.1.2.2.2. Aspects syllabiques : des représentations

Le phonème a longtemps été, et est toujours pour certains, l'unité phonologique descriptive minimale. Cependant, la description « classique » phonématique présentée plus haut a, depuis plusieurs décennies, été mise en concurrence avec d'autres modèles

descriptifs/explicatifs, d'une certaine manière plus abstraits, et ainsi davantage phonologiques. Plusieurs éléments ont en effet conduit à remettre en question le statut de primitive phonologique du phonème, notamment sa linéarité supposée (Brandão de Carvalho et Kawaguchi, 2002). C'est en particulier avec la phonologie autosegmentale de Goldsmith que la phonologie multilinéaire a pris son essor, pour finalement déboucher sur une représentation non-linéaire à plusieurs niveaux (« tiers »), dont un baptisé *syllabique* (Lacharité et Paradis, 1993, p. 131) :

<b>Syllabic tier</b>		<b>Metrical tier</b>	
<b>TIMING TIER</b>			
<b>Segmental tier</b>	<b>Autosegmental tier</b>	<b>Tonal tier</b>	

Le niveau syllabique a encore été nommé « squelette syllabique », et ce squelette peut comporter des places « vides », c'est-à-dire sans association de matériel segmental (le segment étant défini comme un ensemble de traits distinctifs). Ainsi, tandis que, au niveau infra-segmental, le *trait* est devenu l'unité phonologique de base, au niveau supra-segmental, c'est la syllabe qui constitue, dans une certaine mesure, la première unité minimale. Les représentations multilinéaires permettent alors de distinguer deux aspects fondamentaux des items linguistiques : comme l'explique Encrevé, cité par Brandão de Carvalho et Kawaguchi (2002, p. 4), il y a « deux réalités linguistiques différentes : les places que comporte un item lexical et les unités phonétiques [...] qui les occupent ». Il peut donc arriver qu'un segment soit « flottant », c'est-à-dire sans association à une position syllabique, ou encore qu'il soit associé à deux positions syllabiques différentes.

La question du statut phonologique (primitif ou émergent) de la *syllabe* est alors devenue un sujet de discussion polémique parmi les spécialistes du domaine. Comme l'illustre Durand (1990, p. 198) à propos du mazatèque et du polonais, il est en effet

apparu que la syllabe doit être reconnue comme une *unité* à part entière, permettant d'aborder les spécifications phonotactiques de la phonologie des langues de manière beaucoup plus économique qu'en termes de positions segmentales. En effet, comme l'indique Archangeli (1997, p. 10), si l'on considère que les mots sont composés de syllabes, alors la distribution des consonnes et des voyelles dans une langue donnée découle des structures syllabiques de cette langue, et ne nécessite donc pas de description du type de celle que nous avons donnée plus haut, même si une connaissance des structures morphologiques est également nécessaire.

Définir précisément la syllabe, soit comme primitive phonologique, soit comme objet émergent, n'est pas chose aisée, d'autant que cela implique de se référer à d'autres aspects du matériau phonologique (Blevins, 1995, p. 207) :

*« Just as the feet of metrical theory supply rhythmic organization to phonological strings, syllables can be viewed as the structural units providing melodic organization to such strings [...]. The syllable then is the phonological unit which organizes segmental melodies in terms of sonority [(where the sonority of a sound is roughly defined as its loudness relative to other sounds produced with the same input energy (i.e., with the same length, stress, pitch, velocity of airflow, muscular tension, etc.))]; syllabic segments are equivalent to sonority peaks within these organizational units ».*

Dans son chapitre sur la place de la syllabe dans les théories phonologiques<sup>35</sup>, Blevins offre un bon aperçu de la problématique syllabique, qui s'articule comme suit (p. 206) :

1. La reconnaissance de la syllabe comme constituant phonologique ;
2. Le rôle de ce constituant dans l'organisation des segments en termes de sonorité ;
3. La structure interne de la syllabe (notamment avec les questions de sonorité et de poids) ;

---

<sup>35</sup> Nous nous contentons d'aborder ici la syllabe d'un point de vue strictement phonologique. Nous l'examinons ensuite du point de vue de la psycholinguistique, dans le cadre de modèles de perception et de production langagière.

#### 4. Le processus de syllabification.

En ce qui concerne le premier point, l'un des arguments en sa faveur est l'intuition des locuteurs natifs (Blevins, 1995, pp. 209-210), généralement capables de définir avec précision le nombre et les frontières des syllabes. Cette référence à l'intuition du locuteur natif, qui évoque le jugement de grammaticalité<sup>36</sup>, nous renvoie à la frontière entre argument phonologique et argument psycholinguistique : « If phonology is in part the study of mental representations of sound structure, then such intuitions support the view of the syllable as a plausible phonological constituent » (*ibid.*). Blevins considère ainsi les syllabes comme des « substantive linguistic universals » (*ibid.*). Elle note que certaines langues peuvent ne pas devoir systématiquement faire appel à de tels constituants, mais cela n'enlève rien, selon elle, à leur universalité (note 12, p. 236) : « though a language may lack positive phonological evidence for syllables, this is not inconsistent with the claim that syllables exist in all languages ». Il semble que ce point de discussion pourrait largement profiter de l'intégration des résultats des études psycholinguistiques sur la réalité cognitive de la syllabe en tant qu'unité de traitement, en réception ou en production, lesquels rappellent la nécessité de maintenir une distinction claire entre entité linguistique réelle et outil conceptuel de description des ces entités ou des processus qui s'y appliquent.

Blevins (*op. cit.*, p. 210) présente ainsi la hiérarchie prosodique universelle dans laquelle la syllabe aurait une position fixe :

Phonological phrase  
Prosodic word  
Foot  
Syllable  
Syllable terminals

Suivant Archibald (2000), cette hiérarchie peut être développée comme suit :

---

<sup>36</sup> A ce sujet, voir le chapitre « Norme » de Ducrot et Todorov (1972, pp. 162-169), et aussi Azra, 1999.

Phonological phrase

Prosodic word

Foot

Syllables

Moras

Segments

Features

Trois concepts sont alors essentiels pour traiter de la syllabe : celui d'extramétricalité, celui d'ambisyllabité et celui de sonorité.

I.1.2.2.2.1. Extramétricalité

Le concept d'*extramétricalité* évoque la possibilité pour certains segments sous-jacents de ne pas être systématiquement intégrés à une structure syllabique (Durand, 1990, p. 211), et a été élaboré au vu de certains phénomènes morphophonologiques, au cours desquels l'échelle de sonorité n'est plus respectée. Durand (1990, p. 213) en donne une illustration en analysant le phénomène de liaison en français du midi (tel qu'il le définit), avec l'exemple de la séquence « petit écrou » : étant donné les formes dérivées (« petitesse ») et flexionnelles (« petite »), il est en effet apparu que la forme sous-jacente de mots comme « petit » doivent inclure une consonne latente finale (/t/ dans le cas de « petit »). Traiter la liaison en termes de consonne extramétrique (qui, dans l'exemple de Durand, est rattachée à l'attaque de droite si elle est inoccupée (« écrou ») permet notamment de bien distinguer *consonnes latentes* (qui n'apparaissent que si elles sont suivies d'une voyelle (attaque suivante inoccupée)) et *consonnes fixes* qui apparaissent dans tous les cas.

I.1.2.2.2.2. Ambisyllabité

Le concept d'*ambisyllabité*, lié à la syllabification, peut être pertinent dans notre étude pour des cas comme celui-ci : « monsieur le juge ». Lorsque le schwa est effacé

dans le déterminant « le », on obtient : « Monsieur l'juge », c'est-à-dire [møsjøɫʒyʒ].

La question de la syllabification du /l/ se pose alors : /mø.sjøl.ʒyʒ/ ou /mø.sjøl.ʒyʒ/ ? Cette question permet de ne pas simplement qualifier le groupe /ɫʒ/ de groupe consonantique secondaire en français (c'est-à-dire résultant de la chute d'un schwa) : il peut s'agir d'un groupe secondaire tautosyllabique, d'un groupe secondaire hétérosyllabique, ou bien d'un groupe secondaire ambisyllabique. Il s'agit là du problème de la syllabification en français, qui est, par exemple, clairement et succinctement présenté par Laks (1995, p. 54) : pour tout locuteur francophone natif, la longueur syllabique d'un énoncé, c'est-à-dire le nombre de syllabes dont il est composé, ne présente, en général, aucune ambiguïté, contrairement aux frontières syllabiques, c'est-à-dire les points de discontinuité syllabique qui sont parfois difficiles à déterminer avec précision et sans hésitation. Le concept d'ambisyllabité est alors invoqué, par exemple pour des groupes comme /str/ ou /sm/ en position interne. Laks l'illustre avec le mot « catastrophe » :

- /ka-tas-trɔ-fə/ (canonique)
- /ka-ta-strɔ-fə/ (groupe tautosyllabique)
- /ka-tas-strɔ-fə/ (/s/ ambisyllabique)

Laks justifie l'emploi du terme « canonique » pour la première segmentation en indiquant qu'il s'agit d'une syllabification jugée possible par tous les francophones, soutenue par la tradition orthoépique normative et encouragée par l'apprentissage de la lecture, en particulier des vers.

#### 1.1.2.2.2.3. Sonorité

Un autre concept essentiel pour l'étude de la syllabe est celui de *sonorité* (Morelli, 1999) : la définition de celle-ci est contestée mais s'appuie généralement sur une base phonétique (intensité acoustique et saillance perceptive en particulier, même si d'autres facteurs phonétiques sont parfois invoqués). D'un point de vue phonologique c'est essentiellement le caractère primitif (trait distinctif) ou émergent (propriété dérivable)

de la sonorité des segments qui est en jeu. De ces différentes approches découlent plusieurs échelles de sonorité, ainsi qu'un principe *non absolu* la « Sonority Sequencing Generalization » (SSG), dont une formulation possible est la suivante (Blevins, 1995, p. 210) :

*« Between any member of a syllable and the syllable peak, a sonority rise or plateau must occur ».*

Cela correspond à ce que l'on appelle aussi le « Sonority Sequencing Principle » (SSP) (Morelli, 1998). Si la SSG semble parfois violée dans certaines langues, en particulier par les groupes « /s/ + Plosive », il semblerait pour certains (Blevins, 1995, p. 236, note 18) que cela soit en réalité dû à l'implication d'éléments extraprosodiques, aux frontières du domaine de syllabification. D'autres, comme Morelli (1999), soutiennent que l'organisation des groupes obstruants, au moins, n'est, en fait, pas contrainte par des principes de sonorité, et qu'il n'y a donc pas existence, ou nécessité de postuler, de cas exceptionnels.

L'une des questions à propos de la hiérarchie de sonorité est bien celle de l'existence d'une échelle universelle unique ou au contraire d'échelles spécifiques à chaque langue. Plusieurs auteurs ont ainsi proposé des échelles différentes. Morelli (1999) présente les échelles de :

Clements : Obstruantes < Nasales < Liquides < Glides

Butt : Obstruantes Non-voisées < Obstruantes Voisées < Nasales < Liquides < Glides < Voyelles

Selkirk : p, t, k < b, d, g < f, θ < v, z, ð < s < m, n < l < r

Durand (1990, p. 210), en se basant sur les travaux de Selkirk sur plusieurs langues, propose la hiérarchie suivante :

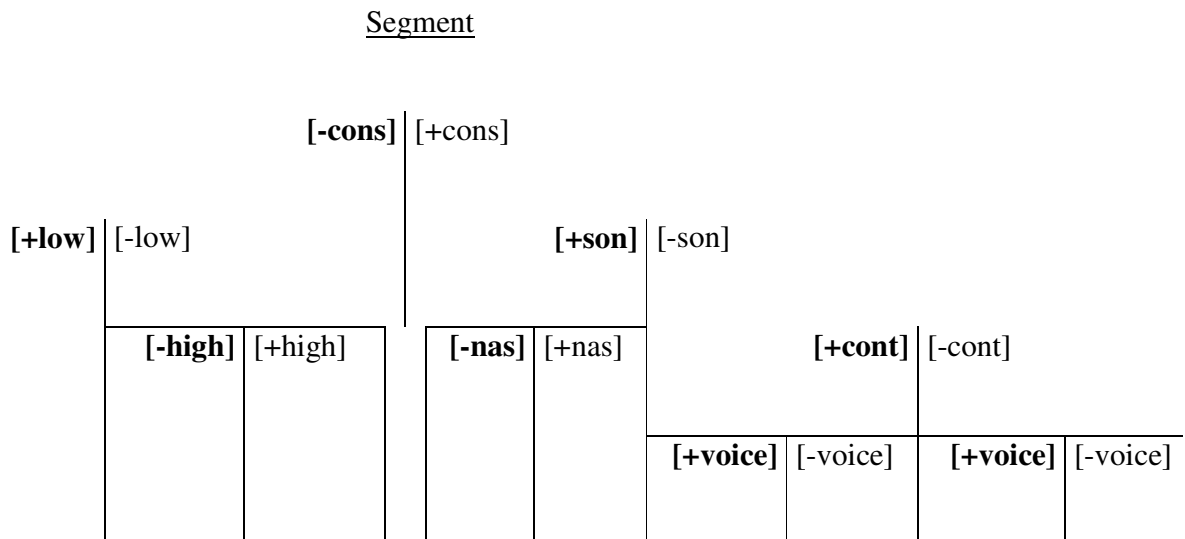


Durand : a > e, o, > i, u > r-sounds > laterals > nasals > voiced fricatives > voiceless fricatives > voiced stop > voiceless stop

Blevins (1995, p. 211) cite celle de Ladefoged pour l'anglais, qui se conforme à la plupart des autres échelles :

Ladefoged : a > æ > ε > ɪ > u > i > l > n > m > z > v > s > ʃ > d > t > k.

Blevins propose ensuite une version universelle de l'échelle de sonorité en termes de traits distinctifs, que nous adaptons ici :



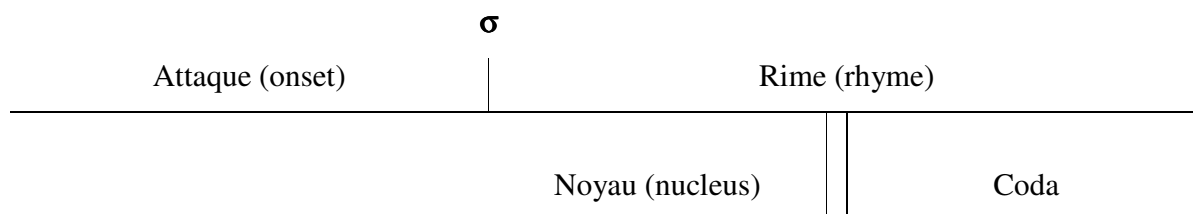
Pour chaque nœud (non-représentés ici), la branche de gauche (figurée en caractère gras) est plus sonore que celle de droite. Pour Morelli (1999), il est en effet nécessaire de regrouper fricatives et plosives du point de vue de la sonorité, sous peine de conduire à des prédictions typologiques erronées.

Le concept de sonorité a donc été fréquemment invoqué au cours de l'histoire de la phonologie pour rendre compte de l'organisation des segments au sein de la syllabe, ainsi qu'entre syllabes, que ce soit *via* la SSG, le *Principe de Distance de Sonorité*

*Minimale* (« The Minimum Sonority Distance Principle ») qui pose que la combinaison des segments est fonction de leur distance relative sur l'échelle de sonorité, ou encore la *Loi de Contact Syllabique* (« The Syllable Contact Law ») qui pose que le contact préféré entre deux syllabes adjacentes est celui où le segment final de la première consonne est moins sonore que le premier segment de la syllabe suivante (Morelli, 1999).

#### 1.1.2.2.4. Structure interne de la syllabe

En ce qui concerne la structure interne des syllabes, Blevins présente plusieurs arguments (1995, pp. 212-216) en faveur d'une structure binaire branchante avec rime :



L'un de ces arguments est notamment lié au concept de poids syllabique, que Blevins présente ainsi (p. 215)<sup>37</sup> :

	Légère	Lourde	Plus lourde
<b>Type 1</b>	Rime simple		Rime branchante
<b>Type 2</b>	Noyau simple		Noyau branchant
<b>Type 3</b>	Rime simple	Rime branchante	Noyau branchant

<sup>37</sup> A la lecture de ce tableau, on pourrait se demander s'il ne faudrait pas plutôt, pour la catégorie « Plus lourde » du Type 3 : Noyau branchant + Rime branchante. Il faut alors considérer ce qu'indique Blevins en note de fin de contribution (note 24, p. 237) : « [...] reflects the generalization that heavy syllables are more sonorous than light syllables. Given syllable-internal sonority profiles, a branching nucleus will always be more sonorous than (or equally as sonorous as) a nonbranching nucleus + coda sequence, resulting in the above stated generalization ».

La syllabe serait donc composée d'une attaque et d'une rime, elle-même composée d'un noyau et d'une coda. D'autres théories existent néanmoins à ce sujet, en raison notamment de la variation à travers les langues : Archangeli propose ainsi une liste des propriétés typiques des structures syllabiques à travers les langues, en soulignant leur caractère non-absolu (1997, p. 7) :

Syllables begin with a consonant	ONSET
Syllables have one vowel	PEAK
Syllables end with a vowel	NOCODA
Syllables have at most one consonant at an edge	*COMPLEX
Syllables are composed of consonants and vowels	ONSET & PEAK

Si l'on examine le français à la lumière de cette liste, nous pouvons voir que certaines tendances sont toujours respectées (PEAK), tandis que d'autres le sont moins.

#### 1.1.2.2.2.5. Syllabes primitives ou émergentes ?

Une fois établie l'existence des syllabes phonologiques, Blevins s'interroge sur leur caractère primitif (présentes dans le lexique) ou émergent (générées durant les processus de dérivation phonologique), en favorisant la seconde hypothèse. Celle-ci souligne que les représentations sous-jacentes et intermédiaires ne constituent pas forcément des syllabes bien formées dans une langue donnée, ce qui conduit notamment à l'effacement (« erasure ») de segments non syllabifiés (p. 223), en particulier les segments consonantiques. Cependant, si l'effacement et la modification, en termes de traits, constituent des modes de traitement de ces segments non-syllabifiables, une autre stratégie est celle de l'épenthèse vocalique (p. 224), qui permettrait de préserver des segments a priori non syllabifiables. Cette stratégie peut être abordée de deux points de vue : soit en considérant l'épenthèse comme un processus post-syllabique (« rule-based approach ») soit en la considérant comme faisant intégralement partie du processus de syllabification (« templatic approach »).

La directionnalité de la syllabification dans une langue donnée (gauche-droite ou droite-gauche) a alors été invoquée comme facteur de détermination des sites épenthétiques, mais cette proposition n'est pas partagée par tous. La polémique à ce sujet est clairement illustrée par cette citation de Brandão de Carvalho (2002, p. 79) :

« [...] We are assigning a representational basis to syllables. [...] this runs counter to the claims of output-based approaches where syllables emerge from smaller units ; in this respect, our proposals are clearly opposite to non-representational views such as those defended by Coleman [2002] ».

Certaines études adoptent en effet une perspective émergentiste non-représentationnelle (Coleman, 2002), notamment certains travaux que l'on place généralement au sein des approches dites « connexionnistes »<sup>38</sup>. En ce qui nous concerne, nous pouvons par exemple citer l'étude de Laks (1995) sur la syllabification du français, placée dans un modèle qu'il qualifie d'extension de la phonologie autosegmentale, *le Modèle Dynamique Linéaire* (« Dynamic Linear Model »), un modèle subsymbolique en ce qu'il propose « an intermediate level of modelling between pure symbolic, conceptual accounts, and non symbolic dynamic implementations » (1995, p. 53). Dans ce cadre subsymbolique, les syllabes ne sont pas des primitives phonologiques, mais des produits dérivés (« by-products ») de l'intégration linéaire des segments, lesquels interagissent de manière continue, donnant à l'universalité supposée de l'organisation syllabique « a materialistic, quasi-phonetic grounding » (p. 75). Dans cette approche, le concept de sonorité prend un autre tour, puisque, selon Laks, chaque segment posséderait *deux* types de sonorité (p. 56) :

- Une sonorité *inhérente*, liée aux propriétés articulatoires et acoustiques du segment ;
- Une sonorité *dérivée*, liée au contexte phonotactique.

---

<sup>38</sup> Il convient cependant de ne pas confondre l'interprétation du terme « émergent(iste) » en tant que « dérivé par règles », comme chez Blevins, et celle que l'on trouve dans les travaux connexionnistes. Néanmoins, dans le cas qui nous occupe, il ne nous semble pas absolument erroné d'établir un lien, aussi ténu soit-il, entre ces deux interprétations, en les opposant à une perspective primitiviste, pour la première et représentationnelle pour la seconde.

Selon l'approche proposée par Laks, le passage de l'un à l'autre correspondrait précisément au processus de syllabification, chaque segment étant pris dans les interactions avec ses deux segments adjacents. Contrairement aux approches symboliques qui, pour résoudre la question de la syllabification, posent de manière axiomatique une échelle de sonorité et un algorithme de syllabification, le modèle présenté par Laks, reposant sur les valeurs de sonorité et les poids de connectivité, considère que la hiérarchie de sonorité s'établit de manière contextuelle, l'échelle de sonorité et l'algorithme de syllabification ne faisant alors plus qu'un. La confrontation avec, et l'adaptation à, un large nombre de données relève ainsi d'une certaine thèse de l'apprentissage, se passant de ressources symboliques et de principes posés *a priori* (*op. cit.*, p. 75) :

« *What we have here is a kind of non-symbolic inferential process for, if at the descriptive level phonological concepts and labels are necessary, the behaviour itself does not need them* ».

Cependant, si l'opposition entre *primitivisme* et *émergentisme* touche à une question essentielle, une autre opposition (toute relative) qui a émergé ces quinze dernières années et qui a particulièrement affecté l'approche des phénomènes phonologiques, est celle entre *représentations* et *contraintes* phonologiques.

#### 1.1.2.2.3. Aspects syllabiques : des contraintes (OT et TCRS)

Depuis plusieurs années à présent (Durand et Lyche, 2000, pp. 228-233), la phonologie a vu se multiplier les travaux conduits dans des cadres théoriques de moins en moins axés sur la problématique représentationnelle et davantage sur celle des *contraintes*<sup>39</sup>. De cette manière, les contraintes (encore appelées « conditions », « filtres » ou même « principes » et « paramètres ») s'opposent à des « règles » arbitraires (LaCharité et

---

<sup>39</sup> Comme l'illustre l'intitulé d'un colloque international de phonologie organisé en 2003 à l'université de Toulouse-Le Mirail : « Des représentations aux contraintes ».

Paradis, 1993, p. 127). Durand et Lyche, à propos de la tendance en phonologie à vouloir se dégager des « règles », citent (p. 229) Kaye, Lowenstamm et Vergnaud<sup>40</sup> :

*« Pris ensemble, les principes et les ensembles de paramètres particuliers à une langue donnent une caractérisation complète du système phonologique de cette langue. Dans ce modèle, un système phonologique ne contient aucune composante de règle ».*

L'une des différences majeures entre les deux approches peut être présentée ainsi : les modèles à contraintes sont orientés vers *l'output*, alors que les règles sont orientées vers *l'input* : « Rules are input oriented. They take an input and tell you what changes to make : whatever results is the output » (Hammond, 1997, p. 39). Il est intéressant de souligner ici le lien établi par Goldsmith avec les questions de phonotactique (« in what ways can the items of phonology be put together to make a well-formed word » (1995, p. 3)), puisqu'il prend soin de noter en bas de page (note 14, p. 22) que, dans le cadre multilinéaire des théories basées sur des contraintes comme celle développée par Paradis et La charité, « the term "constraint" is equivalent to phonotactic ». En se référant à la phonologie harmonique, il indique alors que (1995, p. 7) :

*« The well-formedness of a representation is a scalar (indeed, a quantifiable) notion, not simply a matter of yes and no ».*

L'un des atouts des modèles à contraintes, est que ces derniers permettent d'éviter des problèmes de duplication de règles, comme il en existe dans des approches génératives standard qui ont nécessité l'addition d'une règle uniquement pour rendre compte de l'adaptation (Jacobs et Gussenhoven, 2000, p. 193). Deux grandes théories sont alors à considérer : la *Théorie de l'Optimalité* (OT) (Kager, 1999) et la *Théorie des Contraintes et des Stratégies de Réparation* (TCRS) (LaCharité et Paradis, 1993).

---

<sup>40</sup> « The internal structure of phonological representation : a theory of charm and government », *Phonology yearbook*, n°2, 1985.

#### I.1.2.2.3.1. La Théorie de l'Optimalité

Dans le cadre générativiste, l'un des développements majeurs des années 1990 a été la lancée de la Théorie de l'Optimalité par Prince et Smolensky en 1991, à tel point que certain(e)s linguistes comme Archangeli écrivent : « Optimality theory [...] is THE Linguistic Theory of the 1990s » (Archangeli, 1997, p. 1). OT fait partie de ces nouveaux cadres théoriques qui adoptent une approche en principes et paramètres, rebaptisées *contraintes* (de bonne formation), lesquelles sont privilégiées par rapport à des représentations de type linéaire ou non-linéaire, et qui tendent à se distinguer considérablement des approches dérivationnelles sérielles. Comme l'écrit Archangeli (1997, p. 27) : « Examining the nature of underlying representations and of rules has been core to linguistic research : OT changes the focus of this research ».

Dans OT, les contraintes caractérisent les universaux langagiers, tandis que les violations de ces contraintes permettent de caractériser les schémas, la variation et le degré de marque (Archangeli, 1997, p. 10). Ces contraintes peuvent être d'ordre phonologique, morphologique ou syntaxique. Nous reprenons ici en grande partie la présentation de OT donnée par Archangeli (1997). Dans ce cadre théorique, la grammaire universelle (UG) est considérée comme un ensemble de contraintes violables, tandis que les grammaires des langues spécifiques se définissent dans les classements particuliers, propres à chaque langue, de ces contraintes.

Le modèle formel de OT se présente comme suit (Archangeli, 1997, pp. 14-17 ; l'exemple est tiré du Yawelmani) :

<b>Input</b>	/xat-en/	<p>1. <b>GEN</b> : Pour un input donné, le Générateur (GEN) crée un ensemble de candidats d'outputs potentiels</p> <p>2. <b>EVAL</b> : De cet ensemble, l'Évaluateur (EVAL) sélectionne l'output optimal pour cet input-ci</p> <p>3. <b>CON</b> : EVAL utilise le classement spécifique des contraintes à partir de l'ensemble universel des contraintes (CON)</p>
	<b>GEN</b>	
<b>Ensemble des candidats</b>	xa.te.n/ xa.te.ni./ xa.ten/ ne.tax/ etc	
	<b>EVAL</b>  (contraintes)	
<b>Output optimal</b>	[xa.ten.]	

**Input → GEN → EVAL → Output**



La seule contrainte portant sur l'input est que celui-ci soit linguistiquement bien formé. Il en va de même pour les candidats proposés par GEN. A chaque fois, il s'agit d'objets linguistiques tirés du vocabulaire universel de UG.

Il existe aussi des *familles* ou des *classes* de contraintes : les deux principales sont celles de Fidélité (FAITHFULNESS) (F) et celle de Marque (MARKEDNESS) (M).

Parmi les contraintes typiques de ces deux familles, on peut citer :

<b>Contrainte de Fidélité</b>	<b>Contrainte de Marque</b>
<b>MAX</b> : No deletion	<b>ONSET</b> : Syllables have onsets
<b>DEP</b> : No epenthesis	* <b>CODA</b> : Syllables don't have codas
<b>IDENT(F)</b> : No change in feature specifications	* <b>CODA/VOI</b> : Codas are voiceless
	* <b>CLASH</b> : No adjacent stressed syllables

De manière générale, ces deux familles de contraintes sont en conflit permanent, un conflit qui peut être résolu à travers le classement des contraintes :

- Si les contraintes M sont haut placées ( $M \gg F$ ), alors les structures de surface non-marquées sont privilégiées.
- Si les contraintes F sont haut placées ( $F \gg M$ ), alors la fidélité entre l'input et l'output (I/O) est privilégiée.

En ce qui concerne EVAL et CON, on peut noter les propriétés suivantes (Archangeli, 1997, p. 15) :

- I. Les contraintes dans CON sont violables
- II. Les contraintes sont classées
- III. EVAL sélectionne le candidat qui satisfait le mieux les contraintes classées
  - A. La violation d'une contrainte de rang inférieur peut être tolérée si cela est nécessaire à la satisfaction d'une contrainte de rang supérieur.
  - B. Lorsque deux contraintes de rang supérieur se trouvent toutes deux soit satisfaites soit violées, alors c'est une contrainte de rang inférieur qui détermine le degré de satisfaction du candidat.

Notons que la propriété III B. fait partie de celles qui font apparaître la problématique de la *marque* dans le modèle : dans le cas où les contraintes de rang supérieur sont violées, « unmarked patterns that are encoded in low-ranked constraints occasionally emerge despite those patterns not being observed throughout the language » (Archangeli, 1997, p. 16). Dans le cas de la phonologie des emprunts notamment (une langue *emprunte* des items lexicaux à une autre en les *adaptant* à son système phonologique), l'une des problématiques centrales est en effet celle de l'émergence du non-marqué : contrairement à ce à quoi l'on pourrait s'attendre, les modifications subies ne sont pas explicables par les contraintes à l'œuvre dans le système empruntant, voire même les contredisent. Les chercheurs voient dans ces schémas émergents des propriétés de la Grammaire Universelle. Rappelons en effet que l'analyse des mots d'emprunts dans le cadre de OT est globalement la suivante : « the borrowing language's constraint hierarchy will evaluate the candidates produced by GEN, taking as input the source language's output of the borrowed word » (Archangeli, 1997, p. 30). Archangeli donne l'exemple des Hawaïens qui adaptent le mot anglais « velvet » à leur propre langue, en ajoutant notamment une voyelle supplémentaire (« weleweka »), car le mot anglais dans sa forme originale viole les contraintes syllabiques du hawaïen.

Archangeli résume ainsi le fonctionnement de OT, en indiquant comment cette théorie permet d'aborder l'ensemble de la problématique linguistique (*op. cit.*, pp. 16-17) :

<b>Fonctionnement de OT</b>	<b>Problématiques</b>
<p>I. La <i>grammaire universelle</i> comprend :</p> <p>Un alphabet linguistique :</p> <p style="padding-left: 20px;">A. Un ensemble de contraintes CON</p> <p style="padding-left: 20px;">B. Deux fonctions, GEN et EVAL</p> <p>II. La <i>grammaire d'une langue particulière</i> comprend :</p> <p style="padding-left: 20px;">A. Des formes basiques pour les morphèmes (à partir desquelles l'input est construit)</p> <p style="padding-left: 20px;">B. Un classement pour les contraintes de CON</p> <p>III. Pour chaque <i>input</i> :</p> <p style="padding-left: 20px;">A. GEN crée un ensemble de candidats d'output potentiels</p> <p style="padding-left: 20px;">B. EVAL sélectionne le meilleur candidat de cet ensemble</p>	<p>I. La <i>variation langagière</i> se caractérise par les classements différents du même ensemble de contraintes</p> <p>II. Les <i>schémas spécifiques</i> dérivent de ces classements spécifiques</p> <p>III. Les <i>universaux</i> sont représentés par les contraintes (universelles mais violables)</p> <p>IV. La <i>marque</i> est inhérente au modèle :</p> <p style="padding-left: 20px;">A. Chaque contrainte est une expression de marque</p> <p style="padding-left: 20px;">B. Les aspects spécifiques de la marque proviennent des classements</p>

Dans ce cadre, il est alors possible d'établir une typologie des systèmes syllabiques, en utilisant trois contraintes majeures (ONSET, NOCODA et FAITHFULNESS), comme le fait Hammond (1997, p. 36) qui propose le tableau suivant :

	<i>Rankings</i>	<i>Types</i>
A	Faithfulness >> Onset, NoCoda	(O)V(C)
B	Onset, NoCoda >> Faithfulness	OV
C	Onset >> Faithfulness >> NoCoda	OV(C)
D	NoCoda >> Faithfulness >> Onset	(O)V

D'après ce tableau, on pourrait dans un premier temps considérer que le français correspondrait au type A, tandis que le japonais correspondrait au type D.

#### 1.1.2.2.3.2. La Théorie des Contraintes et des Stratégies de Réparation

Comme nous l'avons déjà indiqué, l'autre grande théorie basée sur la notion de « contrainte » est celle initialement élaborée par C. Paradis, à savoir la Théorie des Contraintes et des Stratégies de Réparation (TCRS). La brève description que nous en donnons ci-dessous est essentiellement issue de Paradis (1996) et Paradis et Béland (2002). TCRS s'est notamment développée sur la base de données issues de la *phonologie des emprunts*, emprunts que Paradis (1996, p. 5) définit de la manière suivante :

*« An individual word, or compound functioning as a single word, from L2 that phonologically conforms to (at least) the outermost peripheral phonological constraints of the target language [...], and that is incorporated into the discourse of L1 ».*

A travers le lexique d'emprunt il est donc possible de déterminer en partie le classement des contraintes à l'œuvre dans le système empruntant. Dans TCRS, les contraintes proviennent de principes universels (contraintes universelles) et de paramètres réglés négativement. Les violations de contraintes sont alors traitées par des *Stratégies de Réparation* définies comme telles (Paradis et Béland, 2002, p. 198) :

**Repair Strategy** : « *A universal, non contextual phonological operation that is triggered by the violation of a phonological constraint, and which inserts or deletes content or structure to ensure conformity to the violated constraint* ».

Ces réparations s'appliquent selon le *Principe de Minimalité* (Paradis, 1996, p. 4) :

**Minimality Principle** : « *Repairs (a) apply at the lowest phonological level to which the violated constraint refers and (b) involve as few strategies (steps) as possible* ».

Ce « niveau phonologique » évoqué par Paradis renvoie à une échelle, sur laquelle repose également la *Convention de Précédence*. Cette convention est nécessaire lorsqu'il y a un conflit de contraintes, et elle est définie par Paradis et Béland (2002, p. 200) comme suit :

**Precedence Convention** : « *In a situation involving two or more violated constraints, priority is given to that constraint referring to the highest level of the Phonological Level Hierarchy* ».

L'échelle à laquelle renvoient tant le Principe de Minimalité que la Convention de Précédence (la Hiérarchie du Niveau Phonologique, PLH) se présente ainsi (*ibid.*) :

**Phonological Level Hierarchy** : *metrical level > syllabic level > skeletal level > root node > feature.*

Cependant, contrairement à ce qui est postulé dans OT, les épenthèses et les effacements ne bénéficient pas d'un statut identique dans TCRS, en raison des données des auteurs sur l'adaptation : dans celles-ci, l'insertion (96, 7%) est beaucoup plus fréquente que l'effacement (3, 3%) (Paradis et Béland, 2002, p. 193). Cela est attribué au *Principe de Préservation* (p. 194) :

**Preservation Principle** : « *Segmental information is maximally preserved within the limits of constraint conflicts [(i.e., within the limits of the Threshold principle (Paradis, 1996, p. 3)]* ».

Les structures et/ou contenus sont donc préservés autant que possible. Cependant, cela s'inscrit dans certaines limites, comme cela est notamment exprimé dans le *Principe de Seuil* (Paradis et Béland, 2002, p. 194):

**Threshold Principle** : « *a) all languages have a tolerance threshold to segment preservation, b) the tolerance threshold of a language sets the limit on the procedural weight (i.e. the number of steps) a repair is allowed within a constraint domain [this threshold is set at two steps (or two repairs) within a given constraint domain (Paradis, 1996, p. 3)]* ».

Le *domaine de contrainte* est défini comme tel par Paradis (1996, p. 11) :

**Constraint Domain** : « A "constraint domain" is the "scope" of a constraint, i.e. the material which is horizontally and vertically scanned by a constraint before structure-building rules are authorised to apply ».

La notion de poids procédural renvoie au nombre d'opérations psycholinguistiques nécessaires au traitement de matériel langagier. Les réparations dues au Principe de Préservation sont donc limitées à un certain domaine de contrainte. Il apparaît alors que les effacements consonantiques sont souvent dus à l'enchâssement de deux malformations, généralement l'une segmentale et l'autre syllabique.

Comme l'indiquent les auteurs, alors que la syllabe CV correspond à la structure universellement non-marquée, les contextes syllabiques qui posent souvent problèmes pour l'adaptation (non-syllabifiable en langue empruntante) de mots français sont les six structures marquées, illustrées de la sorte (Paradis et Béland, 2002, p. 196) :

\*codas : « patte » (coda)

\*branching onsets : « place » (attaque branchante)

\*branching codas : « contre » (coda branchante)

\*hiatus : « mosaïque » (hiatus)

\*diphthongs : « boisson » (diphthongue)

\*word intial empty onsets : « ambassade » (attaque vide en initiale de mot)

Les auteurs prennent par ailleurs soin de distinguer « forme » et « emprunt », la forme étant la réalisation concrète, et variable, tant d'un point inter-individuel que intra-

individuel, de l'emprunt, qui, lui, est une entité abstraite (Paradis et Béland, 2002, p. 194, note 1).

Ils proposent un exemple d'adaptation du mot français « biscuit » en Fula (langue africaine) (2002, p. 199). Etant donné qu'en fula les attaques branchantes et les codas branchantes sont illicites d'une part, et que les glides labio-coronaux sont également illicites, la séquence /kɥ/ présente deux malformations :

- Une attaque CC<sup>41</sup> non syllabifiable (\*kɥ) ;
- Un segment illicite (\*ɥ).

La séquence /skɥ/ n'étant pas syllabifiable (puisque le segment /s/ se situe en coda de la syllabe précédente, le segment /k/ ne peut occuper cette position), l'adaptation devrait requérir trois étapes :

- 1) L'insertion d'un noyau (structure)
- 2) Le remplissage de ce noyau épenthétique par un segment (contenu)
- 3) L'adaptation du segment malformé \*/ɥ/ en un segment acceptable /w/

Il s'agirait donc d'une réparation en trois étapes : /Cɥ/ → /Cuw/

Cependant, en fula, une telle réparation violerait le *Seuil de Tolérance* qui est limité à deux, d'où la déletion consonantique observée dans les données (/biski/), au détriment de l'insertion (\*/biskuwi/).

---

<sup>41</sup> Le glide /ɥ/ ayant ici un statut consonantique.



Dans le cas des attaques branchantes ainsi que des codas branchantes, on voit donc que l'adaptation en une syllabe de type CV nécessite au moins deux étapes (p. 209) :

- 1) L'insertion d'un noyau (création d'une nouvelle syllabe)
- 2) Le remplissage segmental du noyau ainsi inséré

Ces contextes entraînent ainsi une augmentation du nombre de syllabes dans le mot, raison pour laquelle les auteurs les appellent des « lengthening marked syllabic contexts » (p. 212).

Paradis (1996, p. 5) propose ainsi une vue schématique de l'organisation générale de TCRS :

<b>TCRS</b>	
<p><b>Constraint Violation</b></p> <p style="text-align: center;">↓</p> <p><b>Repair</b></p> <p>(insertion/deletion)</p>	<p>Governed by the</p> <ul style="list-style-type: none"> <li>• <b>Preservation Principle</b> (limited by the <i>Threshold Principle</i>)</li> <li>• <b>Minimality Principle</b> (based on the <i>PLH</i>)</li> <li>• <b>Precedence Convention</b> (based on the <i>PLH</i>)</li> </ul>

Deux hypothèses relatives à l'examen des malformations potentielles sont alors proposées par Paradis (1996, p. 10) :

**Vertical Scanning Hypothesis** : « *Scanning for syllabic and segmental malformations proceeds in parallel, i.e. when a cluster is examined for syllabic purposes, segmental constraints are also activated, and the cluster is then scanned for segmental purposes as well* ».

**Horizontal Scanning Hypothesis** : « *When a consonant within a cluster is unsyllabifiable, it is the whole consonant cluster which is identified as problematic for the syllabic constraint* ».

Le cadre TCRS permet de noter des différences de stratégies entre l'adaptation linguistique d'une part et la paraphasie (pathologie linguistique) d'autre part : alors que l'insertion vocalique est favorisée pour l'adaptation, c'est l'effacement consonantique qui est préféré dans la paraphasie, ce qui s'explique par son moindre coût en termes d'opérations phonologiques (Paradis et Béland, 2002).

En ce qui concerne la phonologie des emprunts, Paradis (1996) a souligné certaines différences entre TCRS et OT concernant certains principes apparemment similaires : le *Principe de Préservation* pour TCRS et ceux de *Fidélité* (« Parse » et « Fill ») pour OT. En effet, alors que les Principes de Fidélité prédisent une équivalence statistique entre les processus de délétion (*Parse*) et d'insertion (*Fill*) à travers les langues, cela n'est pas vérifié empiriquement dans les quatre corpora de Paradis, qui montrent un net avantage pour l'insertion ou la correspondance, au détriment de la délétion. Cela soulève la question des différences entre ces deux grandes théories, que nous rappelons ci-dessous.

#### 1.1.2.2.3.3. OT versus TCRS

Dans leur article de 1993, LaCharité et Paradis ont comparé trois modèles phonologiques majeurs basés sur la notion de « contrainte », la phonologie déclarative (DP), la théorie de l'optimalité (OT) et la théorie des contraintes et des stratégies de réparation (TCRS), en prêtant une attention toute particulière aux questions suivantes (p. 128) :

- Qu'est-ce qu'une contrainte ?
- Les contraintes peuvent-elles rentrer en conflit ?
- Comment sont résolus ces conflits ?

- Les contraintes peuvent-elles être violées ?
- Ces violations sont-elles réparées ?
- Comment rend-on compte des exceptions de surface ?

Pour notre part, nous essayerons ici de synthétiser les grandes lignes de cet article, en ne retenant principalement que les descriptions de OT et TCRS.

Dans OT, les contraintes sont plus passives que dans TCRS, où elles sont au contraire déterminantes dans la construction de l'output, avec des stratégies de réparation qui visent à assurer la conformité de cet output avec lesdites contraintes, qu'elles soient universelles ou spécifiques (p. 134). Ces contraintes phonologiques constituent donc des restrictions sur la forme phonologique possible de l'output (de surface ou dérivé). Les auteurs résument la nature et les fonctions des contraintes dans le tableau suivant (p. 135) :

Optimality Theory	TCRS
<ul style="list-style-type: none"> <li>• Universal</li> <li>• Positive, negative or implicational</li> <li>• Cannot construct, destroy or trigger any processes</li> <li>• Selects the output form</li> </ul>	<ul style="list-style-type: none"> <li>• Universal and language-specific (formally linked to UG [universal grammar])</li> <li>• Positive, negative or implicational</li> <li>• Can trigger constructive or destructive processes</li> <li>• Determines the output form</li> </ul>

En ce qui concerne la définition des contraintes, dans TCRS, voici ce que proposent LaCharité et Paradis (1993, p. 137) :

*« Constraints delimit universal or language-specific generalizations that account for the inventory, distribution and combinations of phonological elements and structures in a language, as well as phonological alternations.*

*Thus, like [Declarative Phonology], TCRS considers that language differ from one another by virtue of their constraint repertoires ».*

A la question, « les contraintes peuvent-elles rentrer en conflit ? », OT et TCRS répondent oui (p. 138) :

OT	TCRS
Two constraints can have incompatible requirements (one constraint favours selection of candidate A, another favours candidate B)	Two constraints can have incompatible requirements (blocking a change violates constraint A, not blocking it violates constraint B)

Un exemple de conflit dans le cadre de OT est celui de McCarthy cité par LaCharité et Paradis (p. 139) concernant un [r] épenthétique que l'on trouve en particulier à Boston, mais qui a un comportement inattendu. McCarthy considère que ce comportement est dû au conflit entre « Final-C » d'une part, qui nécessite une coda pour les syllabes se terminant par une voyelle courte, et « Coda-Cond » d'autre part, qui indique que ce [r] peut être en position de coda seulement s'il est ambisyllabique (c'est-à-dire lié simultanément à l'attaque suivante), ce qui n'est possible que lorsque l'attaque suivante appartient au même mot phonologique que cette coda [r].

Dans TCRS, un conflit de contraintes peut être déclenché par un élément ou une structure qui se montre problématique vis-à-vis de deux contraintes différentes ou plus. LaCharité et Paradis donnent l'exemple du mot français « cuivre » [kɥivʁ] qui, lorsqu'il est emprunté en fula, pose plusieurs problèmes, puisque le fula, d'une part ne dispose ni du segment [ɥ] ni du segment [v], et d'autre part interdit les attaques et codas branchantes, ainsi que les diphtongues ascendantes. Cependant, comme chacune des contraintes segmentales porte sur un élément différent du mot, et que la contrainte

syllabique est, de toutes façons, prioritaire en raison du mécanisme baptisé Hiérarchie du Niveau Phonologique (PLH) dans TCRS, il n'y a pas de conflits (p. 140).

En revanche, il existe un second type de conflit dans TCRS : les situations dans lesquelles une contrainte est violée lorsque l'on applique une réparation, tandis qu'une autre contrainte est violée lorsque l'on ne l'applique pas. Les auteurs prennent l'exemple des nasales flottantes en français, qui violent la contrainte de *licence* (« licensing ») en n'étant pas attaché à une unité de temps, mais qui violent la contrainte négative sur les voyelles nasales tendues en français si elles sont rattachées à la voyelle précédente par diffusion régressive de la nasalité [+ nasal] (p. 140).

Pour résoudre ces conflits de contraintes, OT et TCRS utilisent des mécanismes différents (p. 141) :

OT	TCRS
Settled by a language-specific prioritization of constraints	Averted or settled by the Phonological Level Hierarchy

Pour OT, c'est bien la hiérarchie spécifique à chaque langue des contraintes qui assure cette résolution, grâce à la prédominance des contraintes de rang supérieur (qui doivent être respectées) sur celles de rang inférieur (qui peuvent être violées).

Pour TCRS, il y a également prédominance d'un type de contrainte sur un autre. Cependant, contrairement à OT qui considère le classement comme spécifique à chaque langue, TCRS considère l'existence d'une hiérarchie indépendante et universelle dans le cas de conflits : la Hiérarchie du Niveau Phonologique (PLH). Celle-ci reflète une organisation phonologique générique, puisqu'elle se présente comme suit (p. 142) :

**Metrical level > Syllable > Skeleton > Root node > Class node > Terminal feature**

La PLH a par la suite été affinée de manière à se présenter ainsi (p. 142, note 12) :

**Metrical > Syllable > Timing slot > Segment > Non-terminal feature > Terminal feature**

En cas de conflit, les contraintes de niveau supérieur sont prédominantes par rapport à celles de niveau inférieur.

Dans TCRS comme dans OT, les contraintes sont violables. Mais, alors que les « réparations » sont possibles dans TCRS, elles ne le sont pas dans OT, qui se contente d'accepter le candidat le mieux formé, même si certaines violations persistent. Pour TCRS en revanche une représentation bien formée doit satisfaire toutes les contraintes actives de la langue, ordonnée selon la PLH. Ces réparations sont dirigées par certains principes universaux, dont *le Principe de Préservation* : « Preserve as much of the input as possible, according to the constraints of the language » (p. 146). Les raisons d'être de ce principe sont présentées dans LaCharité et Paradis (1993). Nous retiendrons pour l'instant que celui-ci, d'une part rend compte de la prédominance générale des opérations d'insertion / liaison sur celles de délétion / déliaison, et d'autre part conduit au *Principe de Minimalité* qui en est la version « opérationnelle » (p. 147) : « a repair must apply at the lowest phonological level to which the violated constraint it preserves refers ».

Pour rendre compte des exceptions de surface (p. 148), OT utilise la supériorité des contraintes de rang supérieur, reléguant ainsi les contraintes de rang inférieur. Mais comme l'indiquent LaCharité et Paradis, il convient de mentionner le rôle des strates (ou niveaux) lexicales : les contraintes sont certes spécifiques aux langues, mais elles le sont aussi aux strates lexicales au sein de chaque langue, un point qui est commun à OT et TCRS. Les domaines de contraintes sont donc déterminants, puisque chaque strate semblerait disposer de son propre classement de contraintes. Cela permet de tolérer des formes intermédiaires dont le degré de bonne formation s'établit en

référence à un certain domaine phonologique. LaCharité et Paradis (1993, p. 149) proposent donc le tableau suivant pour résumer l'utilisation des outils spécifiques et universaux dans chacun des deux cadres théoriques :

	<b>OT</b>	<b>TCRS</b>
<b>Language-specific tools</b>	<ul style="list-style-type: none"> <li>• Constraint ranking</li> <li>• Constraint-ranking domains</li> </ul>	<ul style="list-style-type: none"> <li>• Constraint domains</li> <li>• Parameter settings</li> </ul>
<b>Universal tools</b>	<ul style="list-style-type: none"> <li>• Constraints</li> <li>• GEN</li> </ul>	<ul style="list-style-type: none"> <li>• Repair strategies</li> <li>• Constraints (parameters and principles)</li> <li>• Constraint ranking</li> </ul>

A propos des domaines de contraintes, Paradis souligne sur ce point la supériorité de TCRS sur OT en montrant que (1996, p. 23) :

*« Even though derivational constraints increase the complexity of a framework (they require extra machinery such as the notion of « constraint domain »), they are necessary from the point of view of explanatory adequacy ».*

C'est également cette question qu'abordent LaCharité et Paradis (2000), montrant que, si le passage d'une approche dérivationnelle à une approche en termes de contraintes, comme cela a été réalisé dans le cadre OT, a été réussi, des affirmations comme celle de Itô et ses collègues (« there are no rules or repair strategies, and no serial derivation ») qu'elles citent (p. 211), seraient exagérées (LaCharité et Paradis, 2000, p.213) :

*« We see this as a fundamental paradox in OT: it repudiates the notions of being procedural or derivational, but it nonetheless relies on the operations that defined rules to transform an input form into an output surface form ».*

C'est, en effet, ce qui caractérise précisément les règles phonologiques, à savoir la mise en relation des représentations sous-jacentes et des représentations de surface. Lacharité et Paradis considèrent donc que, dans OT, GEN est utilisé précisément pour accomplir ce qui était auparavant pris en charge par des règles : la mise en relation d'un niveau à un autre *via* l'insertion et l'effacement de structures et de contenus (p. 218).

Alors que certains placent parfois la différence essentielle entre OT et TCRS dans l'utilisation que cette dernière fait de « règles », Lacharité et Paradis soulignent donc la nécessité de clarifier le terme de « règle » (2000, pp. 229-230) :

*« The rules [TCRS] resort[s] to, though, are universal and rely on a violated (universal or universally-based) constraint to provide context, and their power is limited to insertion/deletion of content/structure. TCRS terms such operations « repairs », but OT considers them rules, which is often misleading. [...] Surely, the idea that GEN is able to insert or delete any amount of content or structure casts doubt on the ideas that GEN constitutes a single operation. [...] OT cannot disavow the use of repair while depending on the operations that define repair ».*

Les auteurs en concluent que la vraie différence entre OT et TCRS ne se trouve pas, par exemple, dans l'opposition « règles vs. contraintes », mais plutôt dans la perspective avec laquelle sont abordées les relations entre règles et contraintes phonologiques.

Ainsi, la présentation des deux grands modèles phonologiques basés sur des contraintes que nous avons effectuée dans les précédents paragraphes se justifie au moins de deux manières :



- Représentations métalinguistiques : comme nous l'avons indiqué plus haut, la phonétique et la phonologie ont beaucoup évolué depuis les travaux structuralistes et fonctionnalistes. Pourtant, ces avancées ne semblent pas, ou peu, affecter les représentations métalinguistiques des didacticiens, et à plus forte raison des enseignants de FLE, qui doivent pourtant aider les apprenants à résoudre leurs problèmes de « prononciation », entre autres. Si les *applications* didactiques de tels modèles peuvent ne pas sembler évidentes, il va de soi que de telles applications ne pourront voir le jour sans une certaine formation/information desdits didacticiens. De plus, par-delà les éventuelles « applications », ce qui importe, nous semble-t-il, c'est que les représentations métalinguistiques des didacticiens et enseignants concernant la chaîne parlée puissent *évoluer*, puisque celles-ci conditionnent naturellement, en partie au moins, leurs principes et leurs pratiques d'enseignement. Loin de devoir suivre les *modèles linguistiques* les plus récents, il s'agit au contraire d'être suffisamment informé de manière à pouvoir faire évoluer ses propres *modèles didactiques*. Cela s'illustre bien dans la problématique de la syllabe.
- Interphonologie, phonologie des emprunts et syllabes : comme cela est apparu au cours de la présentation des modèles, ceux-ci se sont montrés particulièrement productifs dans le cadre de la phonologie des emprunts, notamment à propos des questions de structure syllabique. Or, l'interphonologie (la phonologie des apprenants d'une langue étrangère) est étroitement liée à la phonologie des emprunts, et de nombreuses études récentes dans le domaine de l'apprentissage de la phonologie d'une langue étrangère recourent précisément à ces modèles basés sur des contraintes. Il était donc nécessaire de proposer une description suffisamment détaillée de ces théories, en vue de permettre la lecture des études en interphonologie que nous présentons plus loin dans notre travail et qui ont inspiré notre propre étude.

Cette présentation générale étant effectuée, nous pouvons à présent revenir sur la structure syllabique du français, et plus particulièrement sur les syllabes dites « complexes ».

#### 1.1.2.2.4. Aspects syllabiques : syllabes complexes

La langue française comporte plusieurs types de syllabes : outre les syllabes simples CV et V, on trouve en effet des syllabes beaucoup plus complexes, dont la fréquence d'occurrence est présentée comme telle par Wioland à partir d'un corpus de 87000 syllabes (1991, p. 55, liste 3) :

<b>Fréquence d'occurrence des structures syllabiques toutes positions confondues</b>			
<b>Rang</b>	<b>Structures syllabiques</b>	<b>Pourcentages</b>	<b>Pourcentages cumulés</b>
1	CV	55,5	--
2	CCV	14,0	69,5
3	CVC	13,5	83,0
4	V	10,0	93,0
5	CCVC	2,5	95,5
6	CVCC	1,5	97,0
7	VC	1,3	98,3
8	CCCV	1,0	99,3
9	CCVCC	0,3	99,6
10	CCCVC	0,2	99,8
11	VCC	0,1	99,9
Le reste		0,1	

Si le type « universel » CV est le plus fréquent, on voit qu'en termes de syllabes complexes, ce sont les structures CCV (attaque branchante biconsonnantique) et CVC (coda monoconsonnantique) qui apparaissent le plus fréquemment. Ces deux types de structure nous conduisent à nous pencher :

- D'une part sur les groupes branchants biconsonantiques (en position initiale CCV, mais pas seulement) ;
- D'autre part sur les phénomènes liés au schwa (C@CV → CCV : effacement du schwa donnant lieu en surface à un groupe complexe hétérosyllabique) et à la liaison.

#### 1.1.2.2.4.1. Syllabes complexes et groupes consonantiques

D'un point de vue phonétique, P. et M. Léon (1976, p. 45) proposent un premier recensement de la distribution des groupes consonantiques les plus fréquents en français. Ceux-ci sont du type Consonne + Liquide, c'est-à-dire C + /R/ ou C + /l/. Dans la terminologie de Léon, ces groupes peuvent être :

<b><u>Primaires</u></b> :	C'est-à-dire en <i>langue</i> .	<u>Par exemple</u> : « draps » (/dʁ/) ou « grand » (/ɡʁ/).
<b><u>Secondaires</u></b> :	C'est-à-dire en <i>parole</i> , résultant de la chute d'un « [ə] caduc ».	<u>Par exemple</u> : « j(e) réponds » (/ʒʁ/) ou « le p(e)tit » (/pt/).

Leur distribution est donnée dans le tableau ci-dessous, le signe « + » indiquant l'existence, le signe « 0 » indiquant l'absence et le signe « 2<sup>nd</sup> » indiquant la nature « secondaire » des groupes, les autres étant primaires.

D'un point de vue phonologique, les groupes primaires correspondent à des groupes tautosyllabiques, tandis que les autres correspondent à des groupes hétérosyllabiques<sup>42</sup>.

---

<sup>42</sup> Pour une liste d'exemples en français contenant ces groupes, voir Léon (1976, p. 45). Si certains groupes semblent absents de ce tableau (comme [zʁ], alors qu'il apparaît en tant que groupe « secondaire », par exemple dans « bais'ra »), il faut sans doute le relier à la définition du terme « position » (dans le mot lexical, le mot phonologique, dans la syllabe).

	<b>Position</b>	
	<b>Initiale</b>	<b>Finale</b>
pR	+	+
tR	+	+
kR	+	+
fR	+	+
sR	+ 2 <sup>nd</sup>	0
ʃR	0	0
bR	+	+
dR	+	+
gR	+	+
vR	+	+
zR	0	0
ʒR	+ 2 <sup>nd</sup>	0
lR	+ 2 <sup>nd</sup>	0
nR	+ 2 <sup>nd</sup>	0
mR	+ 2 <sup>nd</sup>	0
ŋR	0	0
pl	+	+
tl	+ 2 <sup>nd</sup>	0
kl	+	+
fl	+	+
sl	+ 2 <sup>nd</sup>	0
ʃl	0	0
bl	+	+
dl	+ 2 <sup>nd</sup>	0
gl	+	+

vɫ	+	0
zɫ	0	0
ʒɫ	+ 2nd	0
lɫ	+ 2nd	0
nɫ	+ 2nd	0
mɫ	+ 2nd	0
ŋɫ	0	0

A propos de l'émergence des groupes « secondaires » via l'effacement du schwa, Carton note (1974, p. 206) :

*« Dans "nous v'nons", la probabilité de disparition du e est plus grande que dans "ses gu'nilles", bien que les deux e soient à l'initiale de mot. C'est sans doute une question de distribution : le groupe secondaire [vn] est plus fréquent que le groupe secondaire [gn]. Dans "vous ne s'rez" ou "vous ne f'rez", nous n'hésitons pas à supprimer le deuxième e muet car les groupes résultants [nsR] et [nfR] sont habituels en français »<sup>43</sup>.*

Les groupes secondaires nous entraînent donc vers le problème du schwa, une question complexe mais centrale dans la phonologie du français, sur laquelle nous revenons plus bas. Nous nous concentrerons ici essentiellement sur les groupes « primaires », en nous basant sur l'analyse de Dell (1995).

Le tableau proposé ci-dessus par Léon semble en effet insuffisant d'un point de vue phonologique. Dell (1995), en revanche, présente un aperçu de la distribution des consonnes en français, en rapport avec les syllabes phonologiques. Ces dernières, selon Dell, ont surtout joué un rôle dans l'étude des alternances phonologiques et moins dans

---

<sup>43</sup> Notons néanmoins qu'en dehors de ces effacements de tels groupes sont assez rares, pour ne pas dire inexistantes.

celle des restrictions phonotactiques. C'est donc dans cette perspective qu'il aborde la question, en considérant les syllabes comme composées de Attaque + Rime, avec Rime = Noyau + Coda. Les deux conclusions principales qu'il dégage pour le français sont les suivantes (p. 5) :

- ❑ Les rimes ne peuvent se terminer que par une consonne au maximum (il donne l'exemple de la première syllabe /ɔp/ du mot « obstruer »)
- ❑ Les groupes en fin de mot se composent de Coda+ Attaque (par exemple « mordre » et « compact », dans lesquels /dʀ et /t/ sont des attaques)

Dans son article, Dell considère les groupes consonantiques à un niveau de représentation nommé « L » qui n'est ni phonologique, ni phonétique, mais un niveau intermédiaire, auquel les schwas « post-toniques » auraient été effacés. Il considère ainsi les consonnes sous-jacentes suivantes, en excluant les glides: /p, t, k, b, d, g, f, s, ʃ, v, z, ʒ, m, n, ɲ, ʀ, l, ɲ/. L'étiquette « OBLI » représente l'ensemble des consonnes appartenant à l'ensemble suivant, c'est-à-dire toute séquence Obstruante + Liquide, excepté celles composées de deux segments coronaux et ceux qui commencent par un segment coronal fricatif : /p<sub>l</sub>, p<sub>ʀ</sub>, b<sub>l</sub>, b<sub>ʀ</sub>, f<sub>l</sub>, f<sub>ʀ</sub>, v<sub>l</sub>, v<sub>ʀ</sub>, t<sub>ʀ</sub>, d<sub>ʀ</sub>, k<sub>l</sub>, k<sub>ʀ</sub>, g<sub>l</sub>, g<sub>ʀ</sub>/.

Il situe le niveau L par rapport à un cadre d'analyse qu'il a globalement proposé dans son ouvrage de 1985. Dans celui-ci, l'analyse du schwa le conduit à poser un certain nombre de règles phonologiques :

- ❑ E-FINb (1985, p. 235) : déletion des schwas en fin de mot.
- ❑ VCE1 (et INI, qui rend compte du comportement du schwa après une consonne post-pausale) (1985, p. 229) : traitement du schwa après une consonne à l'initiale d'un mot.
- ❑ E-FUT (1985, p. 232) : déletion facultative du schwa devant le R du futur simple de l'indicatif.

- EPEN (1985, p. 235, 246) : insertion facultative du schwa entre un groupe consonantique en final de mot et la consonne initiale du mot suivant.
- VCE2 (1985, p. 200) : délétion obligatoire du schwa après une séquence VC appartenant à au même mot.

Ainsi au niveau L, les représentations phonético-phonologiques ont déjà subi E-FINb (au niveau L : "grandes" → /gRād/ et "partent" → /pārt/), tandis qu'elles n'ont encore subi ni VCE1 ("mesure" → /mēzyR/), ni E-FUT ("gardera" → /gārdəRā/), ni EPEN ("strict minimum" → /stRict#mi...), ni VCE2 ("projeter" → /pRɔʒətə/), lesquelles sont toutes postérieures à E-FINb.

Il indique par ailleurs qu'il existe de nombreuses séquences /VCCV/ dont on ne sait si, au niveau L, il convient de les représenter comme /VCəCV/ ou /VCCV/. Dans ce cas, Dell adopte la représentation la plus courte, adaptée à l'analyse, c'est-à-dire celle sans schwa. Il donne ainsi l'exemple du mot "matelas", auquel il attribue, au niveau L, un groupe consonantique interne /tɫ/.

Dans son analyse, l'auteur choisit de désigner par le terme « mot », ce qui reste dans un mot lorsque l'on en a retiré la couche externe de préfixation (par exemple : "postscolaire" est composé de deux « mots », avec le préfixe externe "post"). Ces préfixes externes constituent alors eux-mêmes des « mots » dans l'acception juste mentionnée. Cette clarification est nécessaire pour rendre compte du fait que certains groupes consonantiques de trois consonnes ou plus ne sont admissibles que dans une situation de préfixation externe et non à l'intérieur d'un mot. Il n'est ainsi pas toujours exact de dire que ces groupes consonantiques à l'intérieur d'un mot peuvent être analysés comme une consonne suivie d'un groupe bi-consonantique admissible à l'initiale d'un mot (1995, p. 9).

Voici ainsi les groupes relevés par Dell dans ses sources (certains groupes sont manquants) :

<b>Initial (1)</b>	<b>Médial (2)</b>	<b>Final (3)</b> <b>= (coda) + onset</b>
a) Toute consonne, excepté /ɲ/ b) Tout groupe OBLI c) sp, st, sk d) spl, spR, stR, skl, skR	a) Toute consonne b) La plupart des groupes CC, excepté /ɲC/ et les groupes de deux obstruantes identiques c) spl, spR, stR, skl, skR d) /l/ ou /R/ suivi de OBLI e) Rsp, Rst, RstR, lst, lstR f) ptR, pst, pstR, psk g) ktR, kst, kstR, ksk, kskR, kskl, ksp, kspR, kspl h) Rbs, Rbt, Rks, Rkt, Rkn, Rts, Rdz, lpt i) mst, mps, nst, sʃp	a) Toute consonne simple b) Groupes OBLI, excepté /vɪ/ <sup>44</sup> c) /l/ ou /R/ suivi de toute consonne, excepté /lR, ll, RR, lZ, lɲ/ d) sp, st, sk e) pt, kt, ft, ps, ks f) ts, tʃ, dʒ g) nt, nd, ns, nʃ, ms h) tm, sm, km, gm, dn, gn, mn i) tl, dl j) /l/ ou /R/ suivi de OBLI k) spl, spR, stR, skl l) ptR, ktR, kst, kstR m) lpt, Rts, Rtʃ, Rks, Rst n) ʃtR, nks
e) pn, ps, pf, pt f) kn, km, kv, ks, kt g) tl, tm, ts, tʃ h) sl, sm, sn, sv, sf i) ʃl, ʃR, ʃn, ʃv, ʃpR, psk j) gn, gz k) dz, dʒ l) zl, zv, zb, zgR m) mn, ft		

Dell note que les séries (1) a) à d) contiennent des attaques bien formées, alors que les séries (1) e) à m) sont beaucoup plus exceptionnelles, issues de mots rares ou empruntés.

<sup>44</sup> « (an accidental gap) » (Dell, 1995, p. 18, note 31).



En ce qui concerne la syllabification des groupes intervocaliques, l'auteur établit certaines correspondances entre la position des groupes consonantiques à l'intérieur des mots, et celles des mêmes groupes à l'intérieur des syllabes (1995, p. 6).

De ses analyses, il ressort que (pp. 14-16) :

- a) Une consonne prévoicative est tautosyllabique avec la voyelle suivante.
- b) Dans un groupe OBLI, les deux consonnes sont tautosyllabiques.
- c) Une consonne postvoicative est tautosyllabique avec la voyelle précédente, à condition qu'aucun conflit n'apparaisse avec les deux principes précédents a) et b).
- d) En français, une coda contient une consonne au plus.

Dell note cependant que ce n'est pas parce que tout groupe médial de trois consonnes ou plus peut être analysé comme une séquence coda - attaque que toute séquence de coda + attaque admissible constitue un groupe médial admissible, et il avance trois restrictions (p. 17) :

- Une attaque complexe de trois consonnes ou plus ne doit pas contenir une obstruante voisée.
- Une attaque complexe qui suit une plosive doit commencer par une coronale.
- Une attaque complexe de trois consonnes ou plus ne peut pas être précédée par une plosive coronale.

Concernant les groupes finaux, contrairement à Plénat, auquel il fait référence et pour lequel l'opposition présence *vs.* absence de *e* graphique est significative pour la phonologie du français standard (p. 18), Dell n'opère pas de distinction entre les mots se terminant par un *e* graphique (par exemple : "asile", "pacte") et les autres (par exemple : "avril", "tact"). En généralisant, Dell considère que (p. 19) :

- e) L'ensemble des groupes en finale de « mots » est un sous-groupe de celui des groupes médiaux.

Afin de rendre compte de ceci, et en accord avec le principe posé précédemment, selon lequel en français une coda contient une consonne au plus, Dell propose de considérer que les groupes finaux comprennent une coda suivi d'une attaque. Il donne les exemples suivants (p. 19) : /gar.d/ ("garde"), /maR.bR/ ("marbre"), /dɔg.m/ ("dogme"), /Ris.k/ ("risque"). Dans ces exemples, la deuxième syllabe est considérée comme « dégénérée », avec un noyau vide (/gaR.dØ/), ne devenant pleine que lorsqu'un schwa est associé au noyau vide : "garde-les" [gaRdøle]. Suivant cette ligne de raisonnement, le mot "livre" peut être syllabifié de deux manières (p. 20) : /liv.RØ/ ou /li.vRØ/, et seule la deuxième est cohérente avec l'analyse précédente.

Parmi tous les groupes mentionnés jusqu'ici, il nous faut souligner la relative originalité des groupes du type Sibilante + Consonne, spécialement /s/+C, qui ont reçu un traitement particulier par certains phonologues. L'une des principales raisons pour lesquelles ces groupes ont un comportement particulier (Archibald, 2000, pp. 9-10) est qu'ils violent l'échelle de sonorité (*Sonority Sequencing Generalization*, Blevins, 1995, p. 211) laquelle est respectée par la plupart des groupes consonantiques à travers les langues et qui exprime la tendance selon laquelle le noyau d'une syllabe en est l'élément le plus sonore et que la sonorité tend à décroître aux extrémités de la syllabe. Or dans le cas des groupes /s/+C, certains violent ce principe (par exemple : « stopper », puisque [s] est plus sonore que [t]) mais d'autres non (par exemple : « sniffer », puisque [s] est moins sonore que [n]).

Certains chercheurs ont invoqué le trait d'extrasyllabité, soutenant que le [s] n'appartenait pas vraiment à la syllabe. D'autres suggèrent d'analyser les groupes sp-, st- et sk- comme des unités uniques, en utilisant le « syllabic tier », conduisant à des analyses des mots anglais « strip » et « next » respectivement comme C<sub>1</sub>CVC et CVCC<sub>2</sub> (avec C<sub>1</sub> = st- et C<sub>2</sub> = -st) (Durand, 1990, p. 217). En effet, en attaque syllabique, les groupes « s + Plosive » constituent des « plateaux de sonorité » ou des « renversements de sonorité », selon l'échelle de sonorité choisie. Pourtant malgré leur violation du SSP, Morelli (1999, p. 155) fait remarquer que ces groupes sont assez

fréquents à travers les langues, et que pour réconcilier ces deux faits, les phonologues ont accordé un statut particulier à ces groupes, d'une manière ou d'une autre (hétérosyllabité ou monosegmentalité par exemple), afin de leur conférer une « immunité » contre le SSP. Cette position n'est pas partagée par tous (Morelli, 1999 ; Fleischhacker, 2000).

Après nous être particulièrement penchés sur les groupes tautosyllabiques, il convient d'examiner le statut des groupes hétérosyllabiques, en particulier ceux liés au schwa en français.

#### 1.1.2.2.4.2. Syllabes complexes et schwa

Afin d'introduire la notion de « schwa », on peut prendre comme point de départ ce qu'écrit Grammont en guise d'ouverture de son chapitre sur le « e caduc » (1954, p. 115) :

*« Les changements les plus nombreux et les plus importants qui apparaissent dans la forme des syllabes et des mots sont déterminés par le traitement de l'e inaccentué (oe). [...] Il est nommé ici caduc, c'est-à-dire susceptible de tomber. Lorsqu'il tombe, il n'est pas diminué et réduit, mais supprimé complètement ; lorsqu'il subsiste, sa prononciation est aussi pleine que celle de n'importe quelle voyelle inaccentuée ».*

Les raisons de la chute ou du maintien de ce « e » ont été, et continuent d'être, un fond d'investigation phonologique abondant dans la littérature consacrée à la phonologie du français (Eychenne, 2003a), que ce soit par exemple dans une approche plutôt « traditionnelle » (Grammont, 1954), générativiste-transformationnelle (Dell, 1985) ou dans le cadre générativiste de OT (Noske, 1996), entre autres approches possibles. L'intérêt du schwa pour notre travail repose sur le lien entre « effacement » du schwa et apparition de groupes consonantiques de surface, qui peuvent poser problème aux apprenants japonais de français. Il ne s'agit donc ici que de donner quelques éléments utiles à la réflexion sur ce sujet.

Il convient en premier lieu de bien distinguer *schwa phonétique* et *schwa phonologique*, comme le rappelle Jetchev (2003) :

- D'un point de vue phonétique, « schwa est le terme qui désigne la voyelle qui correspond à la position dite « neutre » des organes d'articulation : ni antérieure ni postérieure, ni fermée ni ouverte, réalisée sans arrondissement des lèvres. [...] Il se trouve que cette réalisation phonétique est souvent le son « par défaut » [...] dans les positions d'affaiblissement (syllabes inaccentuées, syllabes post-toniques, cas de neutralisation d'oppositions vocaliques) » (p. 102).
- D'un point de vue phonologique, en français, c'est cette voyelle, qui correspond à la prononciation de /œ/ dans les syllabes inaccentuées<sup>45</sup>, qui permet une syllabification correcte de certains mots, notamment les emprunts étrangers.
- D'un point de vue orthographique, cette voyelle est généralement transcrite par la lettre « e », mais l'identification des authentiques « e instables », c'est-à-dire ceux qui correspondent à des alternances « réalisation phonétique – zéro phonétique », par opposition aux « e » toujours réalisés phonétiquement, pose souvent problème.

Globalement, les alternances Schwa/Zéro apparaissent dans deux grands contextes (Côté et Morrison, 2004) :

1. A l'intérieur de morphèmes (ex. : "cerise" - [s(ə)riz])
2. En position de joncture, en particulier de frontières de mots ("valse rapide" - [vals(ə)rapid]) et de clitiques ("ce bateau" - [s(ə)baʔo]).

Comme l'indique Côté et Morrison (2004), les premiers sont considérés comme des schwas sous-jacents, c'est-à-dire assumant une fonction contrastive (ex. : "p(e)louse" vs. "blouse"). Les seconds, en revanche, ne sont jamais contrastifs, et leur présence est

---

<sup>45</sup> J. Durand (communication personnelle) attire cependant notre attention sur le fait que dans certains cas (position faible de pied), il pourrait y avoir un véritable schwa phonétique (par exemple dans « autre » ou « sable »).

déterminée par des facteurs segmentaux et prosodiques. Les auteurs insistent en particulier sur le fait que tous les schwas de joncture sont épenthétiques et non lexicaux.

On peut encore illustrer la question du schwa à l'aide des exemples donnés par Eychenne (2003b), qui, en basant ses travaux sur ceux de Dell, insiste sur la nécessité de disposer d'un schwa lexical (et pas un schwa « épenthétique ») pour rendre compte desdits exemples :

Jacques secoue [ʒaksæku]	Henri secoue [ãrisku]
Jacques skie [ʒakski]	Henri skie [ãriski]
secret [sækrɛ] ou [skrɛ]	scribe [skrib(ə)] mais pas *[sækrɪb(ə)] scrabble [skrabl(ə)] mais pas *[sækrabl(ə)]

Il faut par ailleurs noter que « la chute du muet » n'est pas le seul effacement vocalique que l'on trouve en français : la séquence « s'tu veux » présente un « effacement » de la voyelle [i] dans « si », qui conduit à l'apparition du groupe secondaire /st/. Mais ces effacements sont, dans une certaine mesure, secondaires par rapport à ceux du « e muet ».

L'approche dite « traditionnelle » du schwa repose sur la « loi des trois consonnes », énoncée comme telle par Grammont (1954, p. 115) :

*« La règle générale est qu'il se prononce seulement lorsqu'il est nécessaire pour éviter la rencontre de trois consonnes. [...] Son maintien ou sa chute dépend essentiellement de ce qui précède ».*

Qualifiée d'« absolue » par Grammont dans certains contextes (pour les consonnes occlusives et en position interne), d'autres, en revanche, conduisent à des traitements particuliers. Pour ce qui nous intéresse, Grammont mentionne en particulier le « s » (p. 120) et les liquides (p. 123) :

*« Quand un s se trouve entre deux consonnes à l'intérieur ou au début d'un mot, les trois consonnes se prononcent sans aucune insertion d'e, parce que l's a assez de son par lui-même pour n'avoir pas besoin de s'appuyer sur une voyelle et peut servir de soutien aux consonnes qui l'entourent. [...] Si la dernière consonne est une liquide (r ou l), on se borne à la chuchoter ou souffler ; elle est par conséquent dépourvue de vibrations glottales et de voix ».*

Grammont note cependant que « la prononciation dont on vient de lire les règles est celle de la conversation familière et courante, mais non pas la prononciation populaire » et que « diverses causes peuvent maintenir un *e* dont l'application mécanique de la règle entraînerait la chute, et que le nombre des *e* que l'on prononce augmente à mesure que la diction se ralentit ou s'élève » (p. 126).

Cette première approche de la question a naturellement servi de support aux travaux ultérieurs, notamment ceux de Dell (1985), sur lesquels s'appuie à son tour Noske (1996). Ce dernier aborde en effet la question de l'alternance Schwa/ Zéro en français dans le cadre de OT. Ce cadre permet, selon l'auteur, de simplifier le problème en passant d'une description de l'environnement en termes d'input à une description en termes d'output, étant donné les trois principes définitoires majeurs de OT, qu'il évoque au début de son article :

- Les processus phonologiques sont dirigés par l'output ;
- Les contraintes sont classées et peuvent être violées ;
- Il n'y a pas véritablement de dérivation, mais plutôt une simple adéquation input / output, au moins à l'intérieur de chaque strate phonologique.

Noske rappelle au lecteur, qu'une approche générative classique (type SPE) comme celle de Dell de l'alternance Schwa/ Zéro en français, nécessite au moins une dizaine de règles pour rendre compte des phénomènes, tandis que la loi des trois consonnes de Grammont présente de très nombreuses exceptions, en raison de la non prise en compte de la structure syllabique (note 2). En général, indique Noske, cinq types d'alternance sur six sont traités comme des cas de déletion du schwa, tandis que le sixième est traité comme un cas d'insertion.

Il présente d'abord les quatre cas suivants :

- I. Schwa in prevocalic position : deletion obligatory (ex. : "l'ours")
- II. Schwa in postvocalic position : deletion obligatory (ex. : le féminin ("jolie"))
- III. Schwa in a « two-sided open syllable » (VC\_\_CV) (ex. : "tu devenais") :
  - A. Deletion of a single schwa : optional
  - B. Deletion of both schwa : forbidden
- IV. Schwa in more complex consonantal environments
  - A. Deletion optional (par exemple : VC(\_\_)CCV)
  - B. Deletion forbidden (par exemple : CC\_\_CCV)

Il rappelle alors la « Loi des trois consonnes », utilisée pour rendre compte du cas III en particulier. Il qualifie celle-ci de contrainte sur l'output, puisque c'est à ce dernier qu'elle interdit de comporter une séquence de trois consonnes et il explique la constitution de cette « loi » en considérant la structure syllabique du français qui, en position médiale au moins, interdit des séquences finales du type [dv] et initiales du type [vn]. L'exemple auquel il recourt, « tu devenais », est déjà proposé par Dell, dont la démarche dérivationnelle, qualifiée de « input-driven » par Noske, est présentée comme laborieuse, avec des règles parfois presque identiques et donc peu économes :

- VCE1  $\emptyset \rightarrow \emptyset / V\#_1C\_$  (optionnel)
- VCE2  $\emptyset \rightarrow \emptyset / VC\_$  (obligatoire)

D'après Noske, lorsque l'on considère les conditions de syllabification du français, les quatre cas indiqués plus haut peuvent se ramener, à trois « core-types » de déletion du schwa français :

- I. Déletion du schwa prévocalique, obligatoire ;
- II. Déletion du schwa postvocalique, obligatoire ;
- III. Déletion du schwa partout ailleurs, qui s'effectue de manière facultative tant que la forme résultante puisse être syllabifiée.

A ces trois cas de déletion viennent s'ajouter ceux qui concernent les syllabes situées aux frontières des groupes rythmiques<sup>46</sup> : en français, davantage de consonnes sont permises en positions initiale et finale qu'en position médiale, donnant lieu à l'apparition d'attaques et de codas particulièrement marquées par les groupes consonantiques résultants. Noske donne les exemples suivants :

<b>Initial</b>	<b>Final</b>
<ul style="list-style-type: none"> <li>• Revenez [RƏVƏne] - [RVƏne] demain</li> <li>• Te fais [tƏfe] - [tfe] pas de bile</li> </ul>	<ul style="list-style-type: none"> <li>• Je vois l'autre [ʒƏvwalotR(Ə)]</li> <li>• Voilà mon oncle [vwalamõnõkl(Ə)]</li> </ul>

Ces situations de frontières conduisent à un dernier cas d'alternance schwa/zéro, cette fois-ci traitée comme une insertion, et non plus une déletion, de schwa : il s'agit des environnements consonantiques issus de la concaténation de deux mots phonologiques dont le premier présente une coda complexe. Noske donne les exemples suivants :

- Un contact fugitif [əkõtakt(Ə)fyʒitif]
- L'index boursier [lɛdɛks(Ə)bursje]
- Un film doublé [ɛfilm(Ə)dublɛ]

---

<sup>46</sup> Pour un traitement de l'interaction entre accentuation et syncope dans le cadre OT (déletion de voyelles, en particulier en parole rapide, par exemple en anglais : « opera » → « opra ») voir Hammond (1997, pp. 33-58).



Noske indique que les groupes [kt], [ks] et [lm] constituent des codas internes illicites en français, ce qui implique de les analyser en tant qu'attaques de syllabes suivantes, dont le noyau vide est rempli par la voyelle par défaut, le schwa. En ce qui concerne la variabilité de l'épenthèse, Noske en rend compte en invoquant le domaine de syllabification : lorsque « un contact » est considéré pour la syllabification, l'épenthèse n'a pas lieu puisque [kt] est licite en position finale ; en revanche lorsque « un contact fugitif » est pris comme domaine, alors il y a épenthèse.

Comme on le voit, la question du schwa est une question complexe. L'important est de garder à l'esprit que le schwa est d'abord fondamentalement une alternance voyelle/zéro, et que celle du « e muet » est la plus fréquente en français (Eychenne, 2003a, p. 31). Malgré la variabilité de ses réalisations phonétiques, Eychenne propose quatre caractéristiques formelles (illustrées par des exemples) partagées par toutes ses réalisations (2003a, pp. 32-33), dont nous ne présentons ici que les trois premières, qui sont les plus établies :

- Critère 1 (Alternance schwa/zéro) : « Ce phénomène est très souvent le premier, parfois le seul, critère considéré ». Ex. : "la s'maine" vs. "une semaine".
- Critère 2 (Effet de syllabe fermée) : « La présence d'un schwa dans une "syllabe" a pour effet d'abaisser une voyelle moyenne /o, e, ø/ en [ɔ, ε, œ] respectivement. [...] Tout se passe, descriptivement, comme si la consonne d'"attaque" du schwa jouait en fait le rôle de "coda" de la syllabe précédente ». Ex. : "bonnement" (trois syllabes, voyelle initiale lâche) vs. "bonnet" (deux syllabes, voyelle initiale tendue) vs. "bol" (une syllabe, voyelle lâche)<sup>47</sup>.

---

<sup>47</sup> Il faudrait cependant prendre en compte la variation sociolinguistique dans la validité de ce critère, en particulier lorsque l'on considère le « français du midi ».

- Critère 3 « Ce phénomène est souvent corrélé aux critères 1 et 2 lorsque ce (Morphologie) : dernier joue un rôle ». Ex. : le mot "affablement" possède un schwa (et non une voyelle stable) car il est formé sur l'adjectif "affable" qui en possède un<sup>48</sup>.

Ainsi, il apparaît que « l'effacement » du schwa peut conduire, en particulier dans certaines variétés linguistiques du nord de la France, à l'apparition de groupes consonantiques de surface dans les contextes présentés plus haut, groupes qui peuvent s'avérer problématiques pour les apprenants japonais, mais peut-être moins que les groupes « primaires ».

#### 1.1.2.2.4.3. Syllabes complexes, enchaînements et liaisons

Alors que la question du schwa était liée à la question des syllabes complexes *via* les groupes consonantiques « secondaires », celles de la liaison et de l'enchaînement lui sont liées *via* les codas consonantiques : "petit(e) amie" pourrait en effet *a priori* être syllabifié "pe-ti-ta-mie" ou "pe-tit-a-mie". Alors que le premier cas permettrait un traitement « non-problématique » par les apprenants japonais, le second est moins évident, en raison de la présence d'une coda. La question se pose avec plus de force encore pour les codas complexes, comme dans le syntagme "votr(e) ami" syllabifié en "vo-tra-mi", dans lequel la coda complexe (phonétique) se transforme en attaque complexe une fois le syntagme syllabifié. Dans les deux cas, le rapport entre graphie et phonie d'une part, et entre syntagmes non-syllabifiés et syntagmes syllabifiés d'autre part, présente une apparente opacité qui peut rendre ardu le processus de syllabification pour des apprenants de français, en l'occurrence japonais. Ainsi, d'un point de vue phonographémique se pose la question des rapports entre syllabation (ouverte/fermée) et graphies consonantiques muettes en finale de mot. Wioland (1991, pp. 58-63) donne la liste des 20 principales graphies finales qui *peuvent* être muettes (mais ne le sont pas toujours) :

---

<sup>48</sup> Nous simplifions ici l'argumentation de l'auteur.

<b>b</b>	Plomb
<b>c</b>	tabac ; blanc ; porc
<b>ch</b>	Almanach
<b>ct</b>	Respect
<b>d</b>	chaud ; crapaud
<b>ds</b>	Poids
<b>f</b>	nerf ; cerf
<b>g</b>	sang ; étang
<b>gt</b>	vingt
<b>l</b>	soûl ; fusil

<b>ls</b>	pouls
<b>p</b>	trop ; drap
<b>ps</b>	temps
<b>pt</b>	prompt
<b>r</b>	danger ; dernier
<b>rs</b>	volontiers
<b>s</b>	bas ; frais ; après
<b>t</b>	état ; salut
<b>x</b>	choix ; deux
<b>z</b>	assez

C'est la raison pour laquelle, au-delà d'un motif purement descriptif, la question de l'enchaînement et de la liaison, à laquelle est liée celle du schwa, comme cela apparaît dans notre exemple, présente certaines difficultés pour les apprenants de français, et nous oblige ainsi à disposer de certaines informations à son sujet.

Les phénomènes d'enchaînement et de liaison sont des phénomènes de jointure, que l'on définit classiquement de la manière suivante :

- Enchaînement : lorsqu'un mot se termine par une consonne et que le mot suivant commence par une voyelle, la consonne finale tend à prendre appui sur la voyelle qui la suit et qui forme le noyau syllabique. La consonne est liée à la voyelle qui la suit. Ex. : "sac à main" → "sa-cà-main".
- Liaison : la liaison s'applique à des consonnes finales non prononcées lorsque le mot auquel elles appartiennent est isolé, mais prononcées lorsqu'elles précèdent une voyelle à l'intérieur d'un même groupe rythmique. Ex : "ils" et "arrivent" vs. "Ils arrivent".

Pour Wioland (1991, p. 88) :

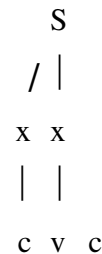
*« L'élision, l'enchaînement et la liaison, à l'exception des h disjonctifs et des faits de jointure entre unités rythmiques successives, ne sont que des conséquences de cette habitude fondamentale de la structuration syllabique française qui, dans le cadre de l'unité rythmique, tend à réunir en une seule syllabe toute consonne avec la voyelle qui suit, même si à l'écrit elles sont séparées par un blanc ».*

Avec le développement de la phonologie non-linéaire, le traitement de la liaison (et de l'enchaînement) s'est en effet de plus en plus tourné vers le domaine syllabique pour rendre compte des phénomènes observés, et cette approche est clairement présentée par Tranel (1995a, 1995b), dont nous nous inspirons largement dans les lignes qui suivent. Comme le rappelle l'auteur, deux types de consonnes finales peuvent être distingués dans la phonologie du français : les consonnes *fixes*, qui sont toujours réalisées phonétiquement, et les consonnes *latentes*, dont la réalisation dépend du contexte. Le premier cas est illustré par l'adjectif « honnête », dans « un honnête citoyen ; honnêteté ; un honnête homme ; une honnête personne », tandis que le deuxième cas est illustré par l'adjectif « petit », dans « un petit citoyen ; petitesse ; un petit homme ; une petite personne » (Tranel, 1995b, p. 131). Il faut alors se rappeler la distinction entre position syllabique sur le squelette phonologique et contenu segmental : « si la consonne se prononce c'est qu'elle dispose d'une "place" à laquelle elle pourra "s'associer", sinon, c'est qu'il n'y en a pas, la consonne restant alors "flottante". [...] [Il y a] à la fois des segments flottants, qui ne s'associent à aucune position, et des places "vides", auxquelles ne s'associe aucun segment » (Brandão de Carvalho et Kawaguchi, 2002, p. 5). Si l'on souscrit à la thèse selon laquelle les consonnes latentes sont des segments flottants, deux approches du phénomène de liaison peuvent alors être envisagées : l'une considère ces consonnes comme des codas, tandis que l'autre les considère comme des attaques de syllabes « sans tête » (« empty-headed syllables », Tranel, 1995b, p. 133). La première semble cependant la plus reconnue, et s'illustre de la manière suivante :

Consonnes Finales Fixes



Consonnes Finales Latentes



L'une des questions essentielles est alors celle de la « syllabification » de la consonne latente, qui nécessite, pour cela, une position sur le squelette phonologique. Plusieurs auteurs soutiennent que les mots à initiale vocalique présentent en fait une attaque vide, qui permettrait à cette consonne latente d'être rattachée au squelette en bénéficiant de cette position vide. Cependant une telle représentation, qui considère la liaison comme une syllabification en position d'attaque, peut poser problème lorsque l'on considère le phénomène d'enchaînement (car on trouve alors une position supplémentaire sur le squelette, celle de la coda du premier mot avant la resyllabification de celle-ci en attaque du mot suivant), en particulier dans le cas des liaisons sans enchaînements illustrées comme tel par Tranel (1995a, p. 806 ; 1995b, p. 141) :

« J'avais un rêve » :

a) [.ʒa.və.ẽ.rɛv.] (pas de liaison)

b) [.ʒa.və.zẽ.rɛv.] (liaison avec enchaînement)

c) [.ʒa.vəz.ẽ.rɛv.] (liaison sans enchaînement)

La liaison sans enchaînement révèle les cas où la consonne latente peut apparaître aussi bien en position de coda que d'attaque dans des contextes de liaison. De plus, certains contextes de liaison ne débouchent pas systématiquement sur la réalisation de telles liaisons, comme dans « un court entracte » *vs.* « de courts entractes ». La variation sociolinguistique à ce sujet constitue justement l'un des points centraux du projet PFC

(Durand, Laks et Lyche, 2003), qui devrait, à terme, fournir un panorama *réaliste* de la question. Au vu de ces « écueils », Tranel, parmi d'autres, préfère distinguer la liaison de la syllabification, en considérant la première comme une insertion de position sur le squelette phonologique, la seconde s'opérant ensuite. Plusieurs modèles de la liaison, basés sur la notion de « flottement », ont ainsi été proposés, se distinguant notamment sur le niveau par rapport auquel cette consonne est flottante : squeletal, syllabique ou bien encore aux deux. Dans tous les cas, l'opposition entre consonnes fixes et latentes semble à présent bien établie, Tranel penchant pour une représentation dans laquelle les premières sont ancrées au squelette phonologique, et les secondes sont sans position sur ce squelette (1995b, p. 165).

En dépit de ce foisonnement théorique phonologique, il semble important, d'un point de vue didactique, que les enseignants de français puissent se dégager de l'attachement quelque peu naïf des phénomènes de liaison à la dimension écrite de la langue, et disposer de perspectives différentes sur la question, en vue d'en gérer l'enseignement à leurs étudiants vis-à-vis de ses dimensions écrite, phonétique, phonologique et morphologique, tout en abordant le problème des « consonnes muettes » à l'écrit.

#### 1.1.2.2.5. Aspects prosodiques

Bien que notre propre étude soit concentrée sur les problèmes segmentaux et syllabiques rencontrés par les apprenants japonais de français langue étrangère, il va de soi que la dimension prosodique des langues en présence ne peut être abusivement ignorée, puisque de celle-ci peuvent fortement dépendre certains des phénomènes que nous étudions. Il convient donc d'en proposer ici une description succincte.

De nombreux auteurs utilisent les termes « traits suprasegmentaux », « traits prosodiques » ou simplement « prosodie » sans distinction (Catford, 1994, p. 172). Lacheret-Dujour et Beaugendre (1999, p. 13)<sup>49</sup> écrivent que :

---

<sup>49</sup> Lacheret-Dujour, A. et Beaugendre, F., *La prosodie du français*, Paris : CNRS Editions, 1999. Cités par Billières, 2001-2002.

*« Le mot "prosodie" est un terme générique qui renvoie à deux mécanismes suprasegmentaux distincts : le premier est de nature lexicale, le second relève de la phonologie postlexicale, puisqu'il dépend de contraintes syntactico-sémantiques et rythmiques globales. Alors que l'accentuation correspond a priori à un processus lexical – les morphèmes lexicaux étant générateurs d'accents—l'intonation relève de la prosodie postlexicale, étant donnée qu'elle utilise les paramètres prosodiques dans une fonction phrastique et énonciative ».*

En effet, même si le domaine suprasegmental couvre plusieurs phénomènes et domaines liés (accentuation, intonation, tonalité, rythmicité, pauses, syllabicité, etc.), il est d'usage de distinguer d'une part des phénomènes que l'on peut globalement qualifier d'intonatifs et d'autres d'accentuels (Carton, 1974, p. 89). La présentation que nous en offrons plus bas est très succincte et peu détaillée, et nous n'approfondirons pas les différences de perspective entre phonologues et phonéticiens sur les phénomènes évoqués.

#### 1.1.2.2.5.1. L'intonation

D'un point de vue phonétique, Léon (1992, p. 119) définit l'intonation comme la « structuration mélodique des énoncés », la mélodie du discours correspondant à l'évolution de la fréquence fondamentale des cordes vocales (*ibid.* p. 43). Lorsque celle-ci est absente, « la perception mélodique s'effectue également à l'aide de la différence de fréquence entre deux harmoniques consécutifs » (p. 43). L'intonation peut ainsi être définie comme la courbe mélodique de la voix, liée au mouvement du fondamental ou à l'écart des harmoniques. Ce mouvement mélodique est souvent exprimé en termes de montées et de descentes de la voix pendant la production d'un énoncé. Celles-ci correspondent essentiellement à des variations de hauteur et de vibration des cordes vocales.

D'un point de vue linguistique (par opposition aux plans physiologique et acoustique), Carton (1974, p. 92) mentionne les traits (distinctifs) qui permettent de décrire les systèmes intonatifs<sup>50</sup> : la hauteur, la durée et l'intensité, la courbe mélodique et le niveau. Delattre a été l'un des pionniers du domaine en dégagant « les 10 intonations de base du français »<sup>51</sup> (Continuation mineure ; Continuation majeure ; Finalité ; Interrogation ; Question (oui/non) ; Parenthèse haute (écho) ; Implication ; Parenthèse basse ; Exclamation ; Commandement). Depuis, de nombreux chercheurs ont proposé des modèles bien plus affinés. Néanmoins, on retrouve à travers ces modèles certaines grandes fonctions de l'intonation, comme par exemple les fonctions démarcative, syntaxique ou encore expressive.

#### 1.1.2.2.5.2. L'accentuation

D'un point de vue essentiellement physiologique, Catford (1994, p. 226) propose la définition suivante du mot *stress*, équivalent anglais du terme « accent » en français, dans ce contexte : « the energy expended by (usually) the pulmonic initiator in pushing air up from the lungs against the loads imposed by phonatory and articulatory strictures ». D'un point de vue davantage linguistique, Léon (1992, p. 107), qui préfère le terme *d'accentuation* à celui *d'accent*, indique que « l'accentuation est une prééminence d'énergie articulatoire qui se manifeste par une augmentation physique de *longueur*, *d'intensité*, et éventuellement un changement de *fréquences* en passant de syllabe inaccentuée à accentuée et/ou au cours de l'évolution de la syllabe accentuée. Au plan de la perception, on parlera de paramètres de *durée*, *d'intensité* et de *hauteur* ». Cette « mise en valeur » (Carton, 1974, p. 98) d'un élément de la chaîne parlée peut en effet être opérée de différentes manières : comme l'écrit Carton (1974, p. 99) : « il est faux de dire que l'accent du français se marque par la durée, l'intonation par la hauteur. L'auditeur ne compte pas sur la reconnaissance de quelques facteurs spécifiques pour repérer les accents : il associe un grand nombre d'indices polyvalents, dont le sémantisme n'est pas le moindre ». Il convient de distinguer deux types

---

<sup>50</sup> Ce point a donné lieu à de nombreux débats (Léon, 1992, p. 12).

<sup>51</sup> Delattre, P., « Les dix intonations de base du français », *French Review*, 41/3, pp. 326-339, 1966.



d'accentuation : l'une, dite linguistique, est relativement fixe et relève de la langue ; l'autre, dite libre, fluctue en fonction des locuteurs et relève de la parole.

En français, cette mise en valeur accentuelle porte sur la syllabe finale de chaque groupe rythmique des énoncés. Les groupes rythmiques coïncident avec les groupes de sens et s'articulent en blocs phoniques. Les groupes de souffle, délimités par la reprise de respiration, contiennent plusieurs groupes rythmiques. Le groupe rythmique est ainsi constitué par une suite de syllabes dont la dernière est porteuse de l'accent primaire. Tout mot prononcé isolément est toujours accentué sur la dernière voyelle prononcée. En ce qui concerne le type de syllabes accentuées en français, Wioland (1991, p. 57, liste 4) donne les pourcentages suivants, repris ici pour les cinq structures les plus fréquentes :

<b>Rang</b>	<b>Fréquence d'occurrence des structures syllabiques en position accentuée</b>	
1	CV	41 %
2	CVC	36 %
3	CCV	8,5 %
4	CCVC	7 %
5	CVCC	6 %

Dans une suite de mots, l'accent perd son individualité au profit du groupe. Wioland (1991, p. 43) parle d'accent rythmique et indique que « la prononciation de ces syllabes accentuées s'accompagne donc d'une énergie articulatoire et non d'une intensité acoustique ». De manière générale, une position d'attaque pour les consonnes est une position forte, tandis qu'une position de coda est faible (relâchement de l'énergie articulatoire) (p. 57). L'accentuation linguistique en français se caractérise donc par : une absence d'énergie acoustique (mais présence d'énergie musculaire) ; un glissando vocalique ; un allongement du fait de l'énergie articulatoire (p. 44).

Le phénomène d'accentuation en français est donc souvent résumé ainsi :

- L'accent primaire (linguistique) est un accent de durée et de groupe. Il est fixe et porte sur la syllabe terminale du groupe rythmique. Il a une fonction démarcative et non distinctive, contrairement à des langues comme l'anglais ou, dans une certaine mesure, le japonais.
- L'accent secondaire (linguistique) est aussi un accent de durée. Il se trouve à l'intérieur de certains groupes rythmiques, à gauche de l'accent primaire. Il est flottant, séparé par au moins une syllabe de l'accent primaire, et apparaît en général lorsque trois syllabes au moins avant l'accent primaire ne sont pas accentuées.
- L'accent libre est un accent d'intensité. Il est laissé à la liberté du sujet qui peut ou non y avoir recours. C'est un accent affectif, émotionnel. Il porte sur la première syllabe du mot ou la deuxième s'il commence par une voyelle.

#### 1.1.2.2.5.3. Le rythme

Le rythme de la parole, fixé par la répétition des accents, est fondamentalement déterminé par la répartition du discours en groupes, modulé éventuellement par des accents secondaires qui dépendent de facteurs variés (sens et type de communication, personnalité et milieu du locuteur). Comme l'écrit Léon (1992, p. 110) : « le *rythme* est marqué par la perception du retour d'une proéminence accentuelle ». Parmi les facteurs qui influent sur ce retour, et donc sur la périodicité accentuelle, figurent d'une part le nombre de syllabes à l'intérieur du groupe, et d'autre part (mais les deux sont liés) leur durée. Pour des locuteurs français, les syllabes inaccentuées dans un énoncé semblent avoir la même durée. Toutes les syllabes, accentuées ou pas, ont tendance à apparaître à intervalles de temps réguliers, ce qui implique que le temps entre les syllabes accentuées sera plus ou moins long selon le nombre de syllabes inaccentuées entre elles. On parle alors dans la littérature anglophone de *syllable-timed rhythm*, par opposition à des langues caractérisées par un *stress-timed rhythm*, comme l'anglais, et on qualifie le rythme français d'isosyllabique (Carton, 1974, p. 125). Certains nuancent cependant cette affirmation, puisque, comme l'écrit Wioland (1991, p. 38) :

*« Le mode croissant qui caractérise le français ne doit pas laisser croire à l'apprenant que le rythme du français serait isosyllabique. L'impression d'égalité syllabique ne peut concerner que les syllabes qui précèdent la dernière syllabe prononcée d'une unité rythmique, la dernière étant par position plus longue ; de plus, lorsque la différence du nombre de syllabes entre unités successive est grande, l'égalité n'est pas identique pour les suites de syllabes concernées d'une unité à l'autre. [...] Contrairement à la tendance naturelle de tout apprenant, le débit est plus lent pour l'unité qui contient moins de syllabes prononcées et plus rapide pour celle qui en contient plus ».*

On distingue ainsi traditionnellement trois grandes classes rythmiques : syllabique (par exemple en français), accentuelle (par exemple en anglais) et moraïque (par exemple en japonais). Cependant, certaines études récentes en psycholinguistique, dans une approche phonético-acoustique du rythme, basée sur des expériences perceptives, conduisent à envisager d'autres classes rythmiques (par exemple dans le cas du polonais), tout en confirmant l'existence des classes mentionnées (Raimus, Dupoux et Mehler, 2003).

Ainsi, en guise de conclusion de cette brève et incomplète présentation du système phonético-phonologique du français, essentiellement envisagé, pour certaines parties, dans une tradition fonctionnaliste à orientation davantage phonétique que phonologique, on peut reprendre les caractéristiques fondamentales du français telles que résumées par Carton (1974, p. 132) :

Antériorité	1. Prédominance des voyelles d'avant
Mode croissant	<ol style="list-style-type: none"> <li>1. Progressivité</li> <li>2. Assimilation progressive</li> <li>3. Anticipation vocalique</li> <li>4. Nasalité nette</li> <li>5. Syllabation ouverte</li> <li>6. Enchaînement</li> </ol>
Tension	<ol style="list-style-type: none"> <li>1. Non-diphtongaison</li> <li>2. Synchronisme</li> <li>3. Labialité</li> <li>4. Fermeture</li> <li>5. Palatalisation</li> </ol>
Egalité rythmique	<ol style="list-style-type: none"> <li>1. Importance du groupe accentuel</li> <li>2. Tendance à l'isosyllabisme</li> </ol>

## **I.2. Le système japonais**

Comme pour le français, la norme linguistique au Japon n'est pas une, puisque le japonais, outre une inévitable évolution diachronique, présente une variation socio-géographique considérable. Le japonais au singulier n'existe donc pas. Cependant, notre domaine d'investigation étant axé sur la région de Tokyo, c'est cette variété de japonais, considérée comme « standard », que nous nous attachons ci-dessous à décrire dans ses grands traits.

Lorsque l'on aborde l'étude de la phonologie du japonais, il importe d'avoir une certaine représentation de la manière dont est organisé son lexique mental, et plus particulièrement phonologique. Cette importance apparaît dès que l'on souhaite décrire l'inventaire segmental du japonais, puisque comme l'écrit Labrune (2001a, p. 93) :

*« L'un des traits les plus remarquables de la phonologie du japonais réside dans le fait qu'on ne peut procéder à un inventaire des consonnes et de leurs possibilités phonotactiques sans tenir compte de la strate lexicale [...] à laquelle appartiennent les mots dans lesquels ces sons apparaissent ».*

D'un point de vue psycholinguistique, la question du lexique mental (Dubois, 1990) reste toujours un sujet de controverse, vis-à-vis de son organisation interne (mode de représentation de l'information sémantique, morphosyntaxique et phonologique), mais aussi vis-à-vis des modes d'accès à ce lexique. Cependant, en ce qui concerne son organisation, d'un point de vue linguistique au moins, l'une des caractéristiques majeures du japonais est sa structure stratifiée, avec quatre grandes classes morphémiques :

- La strate *Yamato* (« yamato-kotoba »), parfois qualifiée de « native »,
- La strate *Sino-japonaise* (« kan-go »),
- La strate *Etrangère* (« gairai-go »),

- La strate *Mimétique/ Onomatopéique* (« gitai-/gisei-go »)<sup>52</sup>.

Itô et Mester (1995, 1999) l'illustrent de la manière suivante, en présentant quatre morphèmes occupant approximativement le même champ sémantique, mais en indiquant bien que les éléments indiqués ne sont en aucun cas synonymes et qu'ils sont utilisés dans des contextes tout à fait différents :

	Classe Morphémique			
Glose	Yamato	Sino-Japonaise	Etrangère	Mimétique
« Chien »	inu	-ken-	doggu	waN-waN

La classe « étrangère » est indiscutablement non-négligeable dans le lexique japonais, étant donné que les emprunts sont de plus en plus nombreux au Japon. Comme l'écrit Crystal (1987/2003, p. 350) :

*« Japan is the only monolingual country to make frequent use of foreign languages (primarily English and French) in its commercials. The viewer usually does not understand them, but the connotations of prestige associated with these languages are enough to warrant their use ».*

En ce qui concerne les emprunts européens, ceux-ci peuvent globalement être divisés en cinq catégories (Taylor et Taylor, 1995, p. 289) :

- Les mots qui représentent des objets et des concepts européens. Par exemple : "banana" (« banana » : banane)<sup>53</sup>, "arufuabetto" (« alphabet » : alphabet).
- Les mots qui représentent des objets et des concepts qui possèdent des équivalents dans la classe lexicale Yamato. Par exemple : "risuto" (« list » : liste), "ruutsu" (« root » : racine).

<sup>52</sup> Que Labrunne (2001a, p. 94) place au sein de la strate Yamato.

<sup>53</sup> Bien qu'aujourd'hui l'« objet » « banane » ne soit pas particulièrement plus « européen » que « japonais » : cela renvoie donc à son introduction historique au Japon.

- Les mots européens tronqués. Par exemple : "masukomi" (« mass communication » : communication de masse), "waapuro" (« word processor » : traitement de texte).
- Les mots européens dont le sens a été quelque peu modifié. Par exemple : "haikara" (« high collar » → « modish » : à la mode), "waishatsu" (« white shirt » → « dress shirt of any color » : une chemise (« formelle ») de couleur quelconque), "abekku" (avec → un couple).
- Les néologismes formés à partir de mots européens. Par exemple : "ooeru" (« OL » pour « office lady » : employée de bureau), "oorudomisu" (« old miss » pour « spinster » : une vieille fille).

Or, cette stratification peut jouer un rôle considérable dans l'approche de la phonologie du japonais. En effet, dans des cadres théoriques à base de contraintes comme OT, si certaines contraintes s'appliquent à l'ensemble de ces strates lexicales, d'autres ne s'appliquent qu'à certaines d'entre elles. De plus, ces strates ne sont pas toujours elles-mêmes homogènes, puisque, dans la classe « Etrangère », le degré de nativisation des mots d'origine étrangère varie, notamment en fonction de son ancienneté, mais aussi de la fréquence d'usage. Itô et Mester (1995) distinguent ainsi les mots étrangers assimilés (« Foreign ») et non-assimilés (« Alien »). Fukazawa et ses collègues (1998, Note 1) indiquent aussi que certains mots d'emprunt, qui étymologiquement appartiennent à la strate des mots étrangers, sont traités comme s'ils appartenaient à la strate native, s'ils ont été assimilés phonologiquement.

Pour rendre compte du comportement des contraintes au sein de ce lexique stratifié, Itô et Mester ont donc proposé un modèle de type Centre-Périphérie, dans lequel chaque strate est incluse dans la strate supérieure ( $Lex^0 \subset Lex^1 \subset Lex^2 \subset \dots \subset Lex^{Max}$ ) de la plus native (Yamato) à la plus étrangère (mots étrangers non-assimilés). La strate minimale correspondrait à la strate native. Plus la strate est éloignée du cœur, plus les contraintes phonologiques natives sont violées par les items lexicaux qui en relèvent. Les auteurs se refusent cependant à parler de « cophonologies », étant donné qu'il ne

s'agit, pour eux, que d'une seule grammaire et d'une seule phonologie (1999, p. 17, note 43). Ils proposent ensuite de rendre compte des différences phonologiques entre ces strates simplement à l'aide de variation de classement des contraintes de Fidélité, qui sont indexées sur les strates.

Bien sûr ce modèle est critiquable : Fukazawa et ses collègues (1998), par exemple, proposent de diviser le lexique en sous-ensembles possédant chacun leur classement propre, plutôt que de permettre un reclassement des contraintes générales en fonction des strates, comme le proposent Itô et Mester. Cette analyse, en contraintes de fidélité multiples, permettrait selon eux de mieux rendre compte des données empiriques, notamment dans le cas de mots hybrides composés de mots appartenant à plus d'une strate lexicale, et pour lesquels les contraintes de chaque strate impliquée devraient être satisfaites simultanément. Ils illustrent ainsi leur argumentaire à l'aide des exemples suivants (p. 5) :

- [tombo-kenkiuuka] « a dragonfly-researcher » (un spécialiste des libellules)  
([tombo]<sub>Yamato</sub> –[kenkiuuka]<sub>Sino-Japanese</sub>)
- [toNkatsu-domburi] « bowl of rice with pork cutlet » (bol de riz avec côtelettes de porc)  
([toN]<sub>Sino-Japanese</sub> –[katsu]<sub>Foreign</sub> –[domburi]<sub>Yamato</sub>)
- [supootsukappa] « a sport raincoat » (un imperméable de sport)  
([supootsu]<sub>Foreign</sub> –[kappa]<sub>Yamato (assimilé)</sub>)
- [tSiimu-tiitSiNgu] « team-teaching » (enseignement en équipe)  
([tSiimu]<sub>Foreign</sub> –[tiitSiNgu]<sub>Alien</sub>)

Ils concluent en ne considérant la structure Centre- Périphérie du lexique que comme une tendance résultant d'un processus d'assimilation lexicale, et non comme une



propriété inhérente de la grammaire, et les débats persistent encore quant à la possibilité d'une version « stratifiée » de OT.

Cette présentation générale du lexique phonologique japonais étant faite, et tout en gardant à l'esprit l'importance de sa stratification vis-à-vis des contraintes phonologiques pouvant s'y exercer, nous pouvons à présent aborder l'inventaire segmental du japonais dit « standard ».

### **I.2.1. Aspects segmentaux**

Nous pouvons dans un premier temps suivre la distinction classique entre consonnes et voyelles, même si la structure moraïque du japonais, que traduit son système d'écriture, dont une partie est parfois qualifiée de « syllabaire », rend, d'une certaine manière, moins pertinente une telle présentation phonémique, par rapport à celle du français, qui possède un système d'écriture alphabétique, plus à même d'assurer, chez ses locuteurs, une certaine réalité psychologique à de tels constituants.

#### I.2.1.1. Consonnes

Labrune (2001a, p. 94) propose l'inventaire consonantique suivant :

	<b>labiales</b>	<b>coronales</b>	<b>palatales</b>	<b>vélaires</b>	<b>glottale</b>
<b>occlusives</b>	<b>p b</b>	<b>t d</b>		<b>k g</b>	
<b>fricatives</b>	<b>{f} {v}</b>	<b>s z</b>	<b>(ʃ) (ʒ)</b>		<b>h</b>
<b>affriquées</b>		<b>(ts)</b>	<b>(tʃ)</b>		
<b>nasales</b>	<b>m</b>	<b>n</b>		<b>(ŋ)</b>	
<b>glissantes</b>	<b>w</b>		<b>j</b>		
<b>Liquide</b>		<b>r</b>			

A cet inventaire, Labrune ajoute les deux « contoïdes » /N/ et /Q/, et elle indique que les consonnes entre parenthèses peuvent être omises de l'inventaire (pour des raisons

qu'elle développe), tandis que les consonnes entre accolades sont des « phonèmes apparus récemment en japonais, sous l'influence des langues occidentales » (p. 94, note 14). /N/ et /Q/ se distinguent des autres phonèmes en ce que /N/ correspond à une consonne nasale sous-spécifiée sur le lieu d'articulation, c'est-à-dire qu'elle présente une grande variété de réalisations allophoniques en fonction du contexte (notamment [n], [m] ou [ŋ]), tandis que /Q/ est également sous-spécifiée puisqu'elle représente la première partie d'une obstruante géminée, et peut donc se réaliser [p] devant [p], [k] devant [k], etc. Dans les deux cas (/N/ et /Q/), l'articulation est déterminée par assimilation régressive.

Il convient à présent de souligner certains traits du japonais qui peuvent avoir une incidence particulière pour notre problématique.

Notons d'abord, suivant Labrune, que le système consonantique japonais se caractérise par la grande variété allophonique de certains de ces phonèmes : par exemple /t/ et /d/, qui sont généralement affriquées ([tʃ, ts, ʒ, ʒ̥, dʒ, dʒ̥, z, dz]) devant /i/ et /u/. A propos des chuintantes, celles-ci sont souvent considérées, non comme des consonnes simples, mais comme des consonnes palatalisées du type /Cy/<sup>54</sup> (le signe /y/ correspondant dans ce cas au symbole /j/ de l'Alphabet Phonétique International). Comme l'indiquent Akamatsu (2001) et Labrune (2001a, p. 99), toutes les consonnes japonaises possèdent en effet un équivalent palatalisé /Cy/. Enfin, à propos des obstruantes voisées /b, d, g, z/, Labrune (2001a, p. 98) présente plusieurs éléments qui font d'elles des consonnes très marquées, dans la classe Yamato au moins.

En ce qui concerne le /r/, il s'agit d'une consonne de type « flap » apico-alvéolaire notée [ɾ] (réalisation prototypique), mais on trouve d'autres réalisations telles que [l] ou [r]. D'un point de vue distributionnel, /r/ est absent en position initiale de mot pour

---

<sup>54</sup> Labrune (2001a, p. 98) fait remarquer, à ce propos, que le système de transcription officiel « si, ti, sya, tya, zya » rend bien mieux compte de la réalité phonologique des segments considérés (qui se retrouve dans la graphie kana) que le système de transcription Hepburn.

les strates Yamato et Mimétique. En revanche, c'est la consonne la plus fréquente en position non-initiale. /r/ semble donc être la plus instable des consonnes : en début de syllabe, elle subit parfois une syncope, mais apparaît parfois en épenthèse de manière inattendue (Labrune, à paraître). Certains parlent de « tap » (Kawakami) (en français « battue »), ou de « flick » (Catford) (cités par Vance, 1987, p. 27). De plus, dans la présentation que Akamatsu (2001, p. 138) donne du système consonantique japonais, on voit figurer, aux côtés de cette « battue » apico-alvéolaire, une latérale apico-alvéolaire /l/<sup>55</sup>. L'auteur indique que celle-ci apparaît après la contoïde N (par exemple dans « aNrakushi » (= euthanasie), mais que ce /l/ est différent du /l/ français. Par ailleurs, comme l'indique Labrune (2003, p. 11), suivant Itô et Mester, le /r/ japonais est la consonne non-marquée du système japonais, sous-spécifiée, comme en atteste la variété de ses réalisations entre autres indices<sup>56</sup>.

En ce qui concerne les fricatives, lorsque l'on examine le tableau d'Akamatsu (2001, p. 138), on remarque, d'une part un couple {/ɸ/ (bilabiale fricative) ; /h/} et d'autre part un triplet {/ç/ (dorso-palatale) (par exemple : « hito » (= personne)) ; /ç̣/ (laminodorsopalatale) (par exemple : « shumi » (= passe temps)) ; /s/ (lamino-alvéolaire) (par exemple : « kasa » (= parapluie))}. L'auteur note que /s/ ne peut pas se trouver devant /i/, contrairement au français.

Il est par ailleurs remarquable que Labrune (2001a, p. 100) consacre tout un sous-paragraphe à ce qu'elle nomme les « nouvelles consonnes » (et qui correspondent à des séquences Consonne(s) + Voyelle inédites), qui ne se rencontrent que dans des emprunts récents :

- wi, we, wo
- kwa, kwi, kwe, kwo
- gwa, gwi, gwe, gwo

---

<sup>55</sup> Sans mentionner les versions palatalisées.

<sup>56</sup> Voir Labrune (2003, p. 12) pour d'autres arguments.

- sye, zye, tsye
- tsa, tse, tso
- ti, di, tu, du
- tyu, dyu
- fa, fi, fe, fo
- va, vi, vu, ve, vo
- ye

A propos de ces « nouvelles consonnes », Labrune mentionne en particulier le /v/ (prononcé /β/) qui « représente la seule introduction récente d'un segment véritablement inconnu en japonais auparavant, mais [qui] est loin de s'être imposée encore » (2001a, p. 101).

#### 1.2.1.2. Voyelles

Le système vocalique japonais se compose de cinq voyelles orales /i, e, a, o, u/, le /u/ pouvant être précisé phonétiquement comme [ɯ], plus centralisé que le /u/ français par exemple, et surtout non arrondi. Ce [ɯ] peut lui-même être encore plus centralisé ([ɯ̠])<sup>57</sup> lorsqu'il suit les consonnes [s], [ts] ou [(d)z] (Akamatsu, 2001, p. 31). Le [e] japonais se situerait entre le [e] et le [ɛ] français, tout comme le [o] japonais, à peine arrondi, situé entre le [o], tout à fait arrondi, et le [ɔ] français, ou le [a] japonais, situé lui aussi entre le [a] et le [ɑ] français. Akamatsu (2001, p. 29) en propose la description phonétique suivante :

---

<sup>57</sup> Le signe diacritique [ ̠ ] indique la centralisation.

[i] (très fermée antérieure)

[u] (très fermée postérieure)

[e] (fermée antérieure)

[o] (fermée postérieure)

[a]

(ouverte

centrale)

Parmi ces cinq voyelles, qui possèdent toutes un équivalent de longueur double et avec lesquelles elles s'opposent, les voyelles /i/ et /u/ sont mentionnées par Labrune (2001a, p. 101) comme remarquables, de par leur utilisation épenthétique dans les emprunts et de par leur effet sur les consonnes environnantes (en particulier affrication). Keating et Huffman (1984, p. 197, fig.2) proposent une description acoustique du système vocalique moyen du japonais standard, basée sur l'analyse de sept locuteurs :

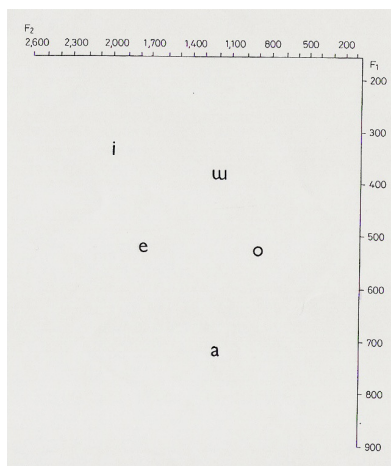


Figure 1 (p. 197) : Valeurs moyennes pour les sept locuteurs

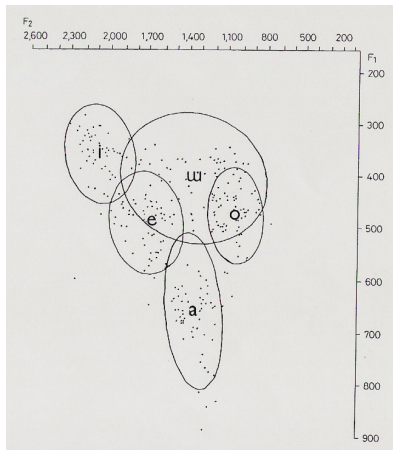


Figure 2 (p. 205) : Valeurs pour le locuteur n° 6

### **I.2.2. Aspects suprasegmentaux**

Nous regroupons ici les aspects phonotactiques, syllabiques et prosodiques.

#### **I.2.2.1. Aspects phonotactiques**

Parmi les aspects suprasegmentaux du japonais, on peut évidemment mentionner, comme nous l'avons fait pour le français, la dimension phonotactique : comme nous l'avons vu précédemment dans l'inventaire segmental, certaines combinaisons sont illégales, au moins dans la strate Yamato (par exemple : \*/si/). Ces contraintes sont facilement appréhendables à travers les deux « syllabaires » kana du japonais, qui en sont les témoins fidèles. Dans une perspective phonotactique, on doit également noter qu'il existe en japonais, à l'instar des phénomènes de contact tels que l'assimilation en français, un phénomène nommé « rendaku » (« voisement de liaison ») (Vance, 1987, p. 133 ; Labrune 2001a, p. 109) que l'on trouve dans des mots composés par concaténation (Labrune donne l'exemple de « amai » (sucré/doux) + « sake » (alcool de riz) → « amazake » (alcool de riz sucré/doux)) et qui consiste en un voisement consonantique, celui de l'obstruante initiale de la seconde partie du mot. Il vaut pour les consonnes [k, s, t, h, ʃ, tʃ].

Par exemple : « kabe » (mur) + « kami » (papier) → « kabegami » (papier peint) ;  
« hana » (fleur) + « hi » (feu) → « hanabi » (feux d'artifice)<sup>58</sup>.

Cependant, ce qui nous intéresse plus particulièrement ici, ce sont les aspects « syllabiques » de la phonologie du japonais.

### 1.2.2.2. Aspects syllabiques : syllabes et mores

Lorsque l'on cherche à décrire la phonologie du japonais, on peut être surpris, de prime abord, de ne pouvoir opérer « naïvement » cette description dans les termes phonémiques et syllabiques traditionnels de la linguistique « occidentale ». En effet, comme l'indique Labrune (2001b, p. 111) :

*« Pour les locuteurs japonais, l'objet « syllabe » tel que peut l'appréhender la linguistique occidentale ne possède aucune réalité intuitive [...] ce terme renvoyant tantôt à la more, tantôt à une unité supérieure à celle-ci de type syllabe lourde, tantôt aux deux à la fois ».*

#### 1.2.2.2.1. More ou syllabe ?

Le terme de « more » est en effet parfois utilisé dans la phonologie occidentale pour désigner une unité de « poids », les syllabes dites « lourdes » comportant alors deux mores (Broselow, 1995, p. 188). Mais comme le rappellent Otake, Hatano, Cutler et Mehler (1993, p. 260),

*« Les concepts tels que phonème, more, syllabe et unité accentuelle ne sont pas les propriétés de langues particulières ; ce sont des abstractions phonologiques grâce auxquelles toute langue peut être décrite. Cependant, de la même*

---

<sup>58</sup> Il faut expliquer ici que, en japonais, le /h/ est lié au /b/ (et au /p/), comme cela apparaît dans le système d'écriture kana : les séquences « ba » et « pa », par exemple, sont formées sur la base du kana « ha », auquel on adjoint un signe diacritique orthographique, l'un marquant le voisement (pour « ba »), l'autre marquant l'absence de voisement (pour « pa »).

*manière que les contraintes de versification ou encore les relations entre phonologie et orthographe varient selon les communautés linguistiques, le degré d'implication de ces abstractions phonologiques dans les processus phonologiques ainsi que leur importance dans la reconnaissance de la parole varient selon les langues ».*

Labrune (2001a, p. 90, note 6) indique ainsi que trois termes sont actuellement utilisés par les linguistes japonais : « haku » (terme d'origine musicale signifiant « battement, coup frappé »), « onsetsu » (« groupe de sons », terme pouvant signifier tout aussi bien syllabe que more) et « moura » (more), ce dernier étant généralement inconnu des non-spécialistes (Akamatsu, 1997, pp. 160-165). Pour Akamatsu (1997), dans une perspective plutôt phonétique, le terme de more devrait d'ailleurs être abandonné au profit de celui d'unité moraïque, puisque la more est fondamentalement, pour l'auteur, une unité de temps. C'est précisément ce système moraïque qui est transcrit par les « kana » japonais, que l'on nomme ainsi parfois abusivement « syllabaires ».

Généralement, le japonais est donc caractérisé, d'un point de vue rythmique par son découpage en mores, lequel est retranscrit dans l'écriture kana, notamment la deuxième partie d'une voyelle longue, la more nasale ou encore la première partie d'une consonne géminée, comme le souligne Labrune (2001a, p. 90), qui rappelle que le terme more est issu du terme « mozi » signifiant littéralement « lettre ». D'un point de vue prosodique, le pied est le second constituant majeur, canoniquement bimoraïque. Labrune (2001a, p. 91) donne plusieurs exemples confirmant l'importance du pied dans l'organisation métrique des énoncés et sa composition en mores, et non en syllabes (le pied ne respectant pas toujours l'intégrité des syllabes, en particulier lourdes).

On peut illustrer, de manière très didactique, la notion de more à l'aide des exemples suivants : le mot « kimono » peut-être segmenté en trois syllabes : « ki-mo-no ». Dans ce cas, chaque syllabe correspond à une more, et cette segmentation sera identique pour des locuteurs français et japonais. En revanche, dans le mot « Kōbe », qui comporterait



deux syllabes pour un locuteur français (kō-be), la voyelle /o/ est en fait une voyelle longue (« ō ») qui compte deux mores distinctes : « ko-u ». Le nom « Kōbe » serait donc composé de deux syllabes pour un Français, mais de trois mores pour un Japonais : « ko-u-be ». Cette voyelle longue « ō » peut être composée de deux mores différentes selon les unités lexicales : soit « o-u » (comme dans l'exemple précédent), soit « o-o ». Ainsi, le mot japonais « ōkī » (= « grand ») sera décomposé en quatre mores : « o-o-ki-i ». La more est donc fondamentalement une unité de temps, qui correspond traditionnellement à l'unité de segmentation naturelle des mots en japonais, à la différence du français qui s'attache, lui, à la syllabe.

Comme l'indiquent Otake et ses collègues (1993, p. 270), qui se sont penchés sur les caractéristiques phonologiques desquelles dépendent les procédures de segmentation des mots parlés, dans le cas du japonais et du français :

*« Les auditeurs japonais ne segmentent pas naturellement la parole syllabe par syllabe ; ils la segmentent naturellement more par more ».*

Dans de très nombreux cas, la more et la syllabe se superposent, mais pas dans tous les cas. Ce système moraïque fait que la structure syllabique du japonais peut en apparence sembler assez simple, de type CV ou V. Si le cas des « voyelles longues », comme dans « Kōbe » peut être appréhendé en tant que syllabe « lourde », d'autres exemples permettent de saisir la spécificité de la more en japonais, en particulier dans le cas de l'unité moraïque nasale N, et l'unité moraïque non-nasale (consonne géminée) Q :

- N:**
- « dengon » (message) → « de-N-go-N »
  - « hinkon » (pauvreté) → « hi-N-ko-N »
  - « kanshin-suru » (admirer) → « ka-N-shi-N-su-ru »

- Q:**
- « sekkan » (savon) → « se-Q-ka-N » ;
  - « gakka » (leçon) → « ga-Q-ka »
  - « kappatsuna » (actif) → « ka-Q-pa-tsu-na »

Wells (2000, p. 119) donne également l'exemple du mot d'emprunt « football » prononcé [ɸʊ-t-to-bo-o-ɾʊ] et qui serait composé de 6 mores (fu-Q-to-bo-o-ru), mais seulement de 4 syllabes (fuQ-to-boo-ru).

L'articulation entre more et syllabe, en particulier en japonais, constitue donc toujours un point de discussion (Kubozono, 1989 ; Otake, Hatano, Cutler et Mehler, 1993 ; Broselow, 1995, p. 202), notamment entre une vision hiérarchisée et une vision indépendante de ces deux types d'unités (Labrune, 2001a, p. 92) :

<b>Hierarchie Prosodique</b>	<b>Indépendance Prosodique</b>
La syllabe est moraïque : l'attaque est dépourvue de poids moraïque, un noyau dure une more, et un élément final V ou C dure une more également.	Syllabes et mores sont directement et indépendamment associées aux segments.

#### 1.2.2.2. Mores « spéciales »

Outre les deux types classiques de syllabes japonaises (légère/monomoraïque et lourde/bimoraïque), Labrune (2001a, p. 92) mentionne également des syllabes extra-lourdes trimoraïques de type (C)(y)VVC (par exemple « siin » (= consonne)). De plus, comme nous l'avons déjà indiqué, tandis que les mores se présentent généralement sous une forme de type CV ou V, quatre d'entre elles présentent certaines particularités (Labrune, 2001a, p. 103) :

- La contoïde /N/ (nasale générique sous-spécifiée ; articulation déterminée par assimilation régressive),

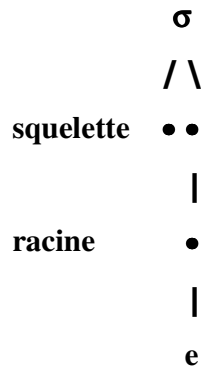
- La contoïde /Q/ (première partie d'une obstruante géminée sous-spécifiée ; articulation déterminée par assimilation régressive),
- La vocoïde /R/ (allongement vocalique sous-spécifié ; articulation déterminée par assimilation progressive),
- La vocoïde /J/ (palatale).

En se basant notamment sur plusieurs faits accentuels en japonais, Labrune (2001b) réexamine l'articulation entre mores (normales et spéciales) et syllabes (légères et lourdes), afin de donner un nouvel éclairage à la structure de la syllabe japonaise. Elle propose de considérer qu'en japonais (2001b, p. 120) :

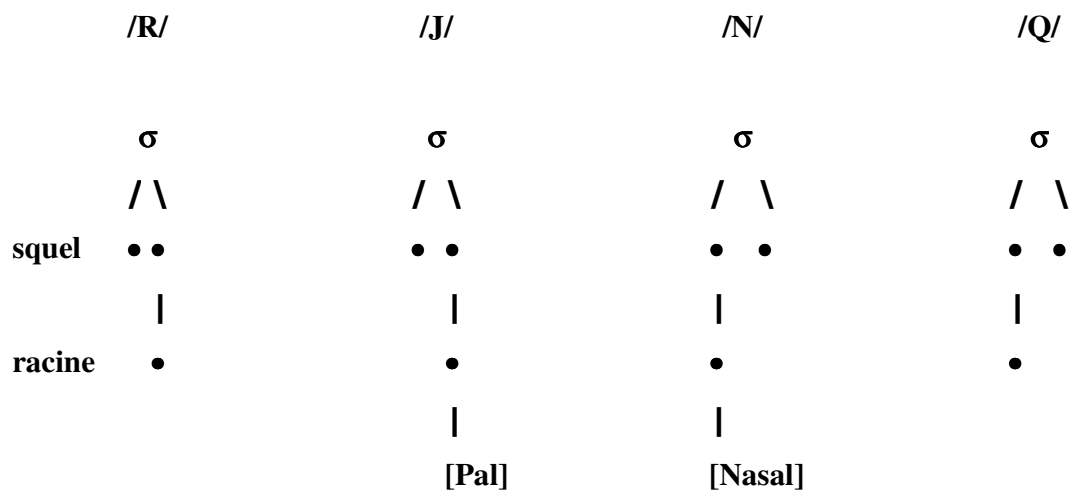
*« [Il n'y a que] deux types de syllabes : des syllabes de structure CV ou CyV, et des syllabes dont l'une des positions est vide : C• ou •V, que l'on peut qualifier de « syllabes déficientes ». Phonologiquement, il n'existe pas de syllabe à structure ternaire. Cette analyse rejoint celle de la tradition japonaise qui ne reconnaît pas les syllabes lourdes ».*

Cette analyse en syllabes exclusivement « légères » permet de neutraliser l'opposition more/syllabe : « toute syllabe vaut une more, et toute more correspond à une syllabe » (*op. cit.*, p. 118), les codas étant donc totalement exclues de la structure syllabique japonaise. Labrune illustre alors les implications de cette analyse pour les syllabes sans attaques, pour les mores « spéciales » et pour les syllabes à voyelle épenthétique ou dévoisée (2001b, pp. 118-119), à l'aide des graphes et des commentaires que nous reproduisons ci-dessous :

Syllabes à attaque vide : position d'attaque vide (exemple : /e/).

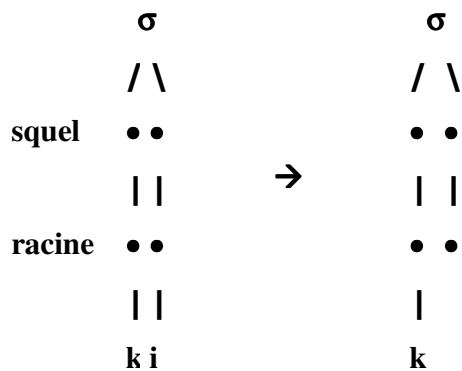


Mores spéciales (longueur vocalique /R/ ; palatale /J/, nasale /N/, géminée /Q/) : syllabes comportant une position squelettale vide avec une autre position dont le nœud racine est soit non spécifié, soit spécifié de manière minimale. La spécification se fait par assimilation régressive ou progressive.

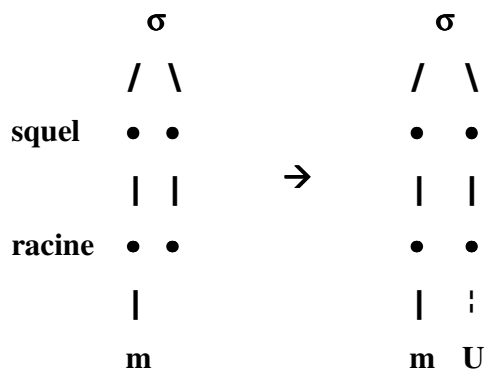


Syllabes à voyelle dévoisée : désassociation du matériel segmental, mais non structurel

(/ki/ → /k•/)



Syllabes à voyelle épenthétique : consonne suivie d'un noyau vide à un certain niveau de représentation (/m•/ → /mu/ dans /amusuterudamu/)



Labrune conclut donc (2001b, p. 122) que la syllabe lourde de la théorie moraïque n'existe pas, les deux types syllabiques licites étant le type « plein » (CV ou CyV) et le type « déficient » (C• ou •V) à position vide.

1.2.2.3. Aspects syllabiques : emprunts étrangers

La phonologie des emprunts, et plus généralement la grammaire des emprunts, est riche d'intérêt pour le chercheur puisqu'elle permet d'observer les transformations,

dans notre cas phonologiques, subies par les mots étrangers pour donner lieu aux formes d'emprunts. A travers ces transformations, on peut apprécier les contraintes qui sont à l'œuvre dans le système empruntant, le rôle de ces contraintes pouvant varier selon les strates lexicales. La phonologie des emprunts peut en outre mettre au jour certains faits occultés par le système phonologique « natif », comme dans le cas du japonais : « loan-word phonology reveals that coda possibilities are more extensive than evidenced by the native vocabulary » (Blevins, 1995, p. 228). Dans la perspective de l'apprentissage d'une langue seconde, ces phénomènes sont également importants, puisque, d'une certaine manière, et surtout du point de vue de l'apprenant débutant, la frontière entre emprunt (ou adaptation) lexical(e) et apprentissage lexical peut parfois être floue. Les contraintes phonologiques qui jouent lors de l'emprunt d'un mot étranger, peuvent également jouer sur la forme phonologique d'un mot étranger « appris », non pas par un groupe linguistique, mais par un individu apprenant.

Shinohara (1997) a ainsi consacré sa thèse à l'analyse phonologique de l'adaptation japonaise de mots étrangers, et plus particulièrement français, dans le cadre de OT. Le Japon se présente en effet comme un domaine riche d'intérêt pour la phonologie des emprunts, puisque quiconque va au Japon est frappé par le nombre impressionnant de mots ou d'expressions d'origine étrangère qui sont intégrés à la vie quotidienne. Certains le sont comme noms propres, pour des publicités, des slogans ou des établissements, soit comme noms communs pour des objets ou des comportements de la vie quotidienne par exemple. Il est d'ailleurs intéressant d'observer les différences entre le répertoire phonique traditionnel japonais (qui se rapporte au syllabaire traditionnel) et celui que propose le *Tuttle new dictionary of loanwords in Japanese* (Kamiya, 1994), dans lequel sont inclus les phones (« sounds » dans le texte original) « nouveaux », utilisés dans les emprunts, signalés par un astérisque dans l'ouvrage de Kamiya (pp. xi-xiv). Ces différences soulignent une fois de plus l'importance des strates lexicales dans l'organisation de la phonologie du japonais. Shinohara (1997, p. 8) donne un premier exemple d'adaptation tout à fait illustratif avec le mot français « gras » qui présente une attaque complexe CCV, attaque qui, du fait d'une certaine

contrainte, est illicite dans le lexique japonais. Cette contrainte \*CCV déclenche l'apparition d'un /u/ épenthétique, débouchant sur la forme d'emprunt /gura/, dont la séquence syllabique est licite en japonais. Comme le rappellent Itô et Mester (1999, p. 14, note 37), les voyelles épenthétiques en japonais sont principalement :

- /u/,
- /o/ après les plosives coronales (par exemple /t, d/),
- /i/ après /k/ dans certains emprunts plus anciens : "sutoraiki" (« labor strike » : grève) (emprunt ancien) vs. "sutoraiku" (« a strike in a baseball game » : une « prise » (« bonne balle ») en baseball) (emprunt plus récent).

Par ailleurs, Dupoux et ses collègues (1999b) notent que la majorité des cas d'emprunts faisant usage de la voyelle épenthétique /i/ correspond à des groupes consonantiques de type Plosive Sourde + Fricative Sourde comme dans « textile » → /tekisutaro/.

Rappelons que, dans OT, les contraintes caractérisent les universaux langagiers, tandis que les violations de ces contraintes permettent de caractériser les schémas, la variation et le degré de marque (Archangeli, 1997, p. 10). La grammaire universelle (UG) y est considérée comme un ensemble de contraintes violables, tandis que les grammaires des langues spécifiques se définissent dans les classements particuliers, propres à chaque langue, de ces contraintes. Ainsi, lorsque l'on considère le lexique phonologique japonais, on peut considérer que les formes existantes résultent d'un certain classement de contraintes, qui n'est bien évidemment pas le même que celui du lexique phonologique français.

Il est alors intéressant de noter que, dans ses analyses, Shinohara relève deux types de formes adaptées (1997, p.9) :

- Le premier type correspond à la grammaire du japonais ;
- Le second présente des régularités qui, tout en se conformant à cette grammaire, n'apparaissent que de manière limitée dans le lexique japonais.

Ce dernier type soulève la problématique de la marque, comme l'indique Shinohara qui distingue ainsi trois catégories de données (1997, p. 10) :

*« 1) Des schémas conformes aux régularités des autres types du lexique, qui confirment la grammaire de L1 ; 2) Des schémas inexistant (ou limités) dans les formes du lexique non étranger, qui dévoilent la grammaire de L1 (émergence de cas non marqués) ; 3) Des schémas inexistant dans les formes du lexique non étranger, qui dévoilent la Grammaire Universelle (émergence de grammaires non marquées) ».*

En ce qui concerne les strates lexicales, malgré les critiques de Fukazawa et ses collègues (1998), le modèle Centre-Périphérie de Itô et Mester (1995) constitue une base d'analyse utile à l'investigation, sur laquelle s'est appuyée Shinohara pour son étude, notamment pour distinguer l'adaptation de l'emprunt. Par « Adaptation », elle entend (2004, p. 2) : le processus par lequel des locuteurs natifs de L1 ajustent les mots étrangers (L2) de manière à ce que les formes résultantes soient acceptables en tant que séquence de sons en L1. Shinohara prend par ailleurs bien garde de distinguer « emprunt » et « adaptation » (1997, p. 10), le premier correspondant généralement à des formes déjà assimilées et elle souligne que son travail porte bien sur l'adaptation des mots étrangers, adaptation qui aurait lieu dans la partie périphérique du lexique (1997, p. 53). Shinohara (2004, p. 3) présente le tableau synthétique suivant (duquel nous avons omis le niveau accentuel) en rappelant, pour le niveau syllabique, les seules structures licites en japonais.



<b>Processus d'adaptation</b>		
<b>Inputs</b>	Formes françaises	<u>Exemples</u> « calmer », « bac » /kalme/, /bak/
<b>Niveau segmental</b>	Correspondance segmentale	karme, bak
<b>Niveau syllabique</b>	Insertion de voyelle épenthétique <sup>59</sup> Allongement syllabique préfinal	karume, bakku
<b>Outputs</b>	Formes adaptées	/ka 'rume/, /ba 'kku/

La structure syllabique du japonais qu'elle adopte dans son analyse est la suivante :

- Syllabes légères :
- (C) (j) V (par exemple : /ha/ (dent))
- Syllabes lourdes :
- (C) (j) V V (par exemple : /dai/ (sujet))
  - (C) (j) V N (par exemple : /teN/ (ciel))
  - (C) (j) V Q (par exemple : /haQ.ki.n/ (platine))

Avec : N= nasale moraique et Q = obstruante moraique ; la première moitié d'une obstruante géminée.

Examinant les facteurs à l'œuvre dans l'adaptation, certains chercheurs ont reproché à ceux qui travaillaient en phonologie des emprunts d'avoir négligé le rôle de la perception dans leur analyse et plusieurs travaux ont ainsi adopté une approche qualifiée de « perceptive » (Côté, 2000). Comme l'écrit Kenstowicz (2003, p. 1), « In adapting a foreign word, the speaker is often faced with choices as to which information of the source to preserve and which to sacrifice ». Pour lui et d'autres, la

---

<sup>59</sup>Voir Shinohara (1997, p. 84 et suivantes) pour le détail de l'adaptation des groupes consonantiques, dans une approche en termes de mores.

saillance et la similarité auditives seraient des facteurs déterminants dans ce choix. Par exemple, dans le cas de l'adaptation fidjienne de mots anglais, Kenstowicz montre que l'une des propriétés les plus saillantes des mots anglais (l'accent principal) est fidèlement copiée dans la forme adaptée. Des ajustements sont alors effectués sur le rythme et la longueur pour correspondre au système métrique fidjien. Le rôle de la saillance y est également important pour l'adaptation des groupes consonantiques.

Archibald (2000, p. 15), au contraire, rejette l'interprétation selon laquelle les sujets ne peuvent utiliser que les indices (« cues ») acoustiques qui sont saillants en L2 et pas les autres, puisque la saillance, selon lui, est dérivée de représentations mentales et pas seulement de la chaîne acoustique :

*« The distinction between [l] et [r] is salient to an English speaker because that is a contrastive feature for English speakers. It is not a salient acoustic feature for Japanese subjects because the sound pair is not contrastive. However, other sounds (like clicks) may well be notable in an acoustic string because they are so unlike any L1 sound of a particular subject ».*

Contrairement à Jacobs et Gussenhoven (2000) qui considèrent que tous les locuteurs possèdent un vocabulaire phonologique universel et, ainsi, que la perception ne joue aucun rôle dans l'adaptation, d'autres chercheurs, comme Silverman (cité par Haunz (2002) et par Jacobs et Gussenhoven (2000)), considèrent que l'adaptation segmentale a précisément lieu lors de l'étape perceptive.

Dans tous les cas, ces deux positions pourraient sembler ne laisser que peu de place à la phonotactique des deux langues dans le processus d'adaptation, ce qui n'est pas soutenu par l'ensemble des chercheurs (Dupoux et *alii.*, 1999). Haunz a ainsi travaillé sur l'adaptation anglaise de mots russes originaux présentant des groupes consonantiques en position initiale. Sa remarque concernant les travaux de Dupoux et

ses collègues nous semble particulièrement importante pour notre propre étude (2002, pp. 5-6)<sup>60</sup> :

*« The methodological issue with Dupoux et al's study is that Japanese has a wealth of loanwords, mainly from English, in which consonant clusters are always adapted in the same way, that is by epenthesis of a vowel. [...] This means that that any Japanese speaker is aware of this convention of adaptation, and this knowledge might play a dominant role in this automatic assumption of a vowel between two consonants by Japanese listeners ».*

Ses expériences la conduisent à conclure que (p. 8) : « native phonotactics is, if not a perceptual filter selecting only sequences that are well-formed in L1, certainly a factor in perception ».

Deux positions semblent donc actuellement se confronter en phonologie des emprunts : l'une, « perceptive/phonétique », soutient que l'adaptation des mots étrangers s'effectue lors de la perception *via* une transformation phonétique minimale, établie par assimilation perceptive entre l'input acoustique en L2 et la catégorie phonétique en L1 la plus « proche », cette « proximité » étant définie soit en termes acoustiques, soit en termes de gestes articulatoires « fins » (« fine-grained ») (Peperkamp et Dupoux, 2003 ; Vendelin et Peperkamp, 2004, 2005) ; l'autre, « phonologique », correspond à une analyse « classique » en phonologie des emprunts, selon laquelle l'adaptation a lieu en production selon les règles et/ou contraintes de la L1 : la représentation sous-jacente, qui serait relativement conforme à l'input de la L2, subirait ainsi des transformations phonologiques pour résulter en une représentation de surface conforme à la grammaire en L1 (Jacobs et Gussenhoven, 2000 ; Smith, 2005). Or, plusieurs études se sont précisément basées sur le japonais pour soutenir tant l'approche « perceptive » (Vendelin et Peperkamp, 2004, à propos de l'adaptation de la coda [n] en

---

<sup>60</sup> Même s'il faudrait nuancer l'affirmation selon laquelle les groupes consonantiques sont « toujours » adaptés de la même manière (Shinohara, 1997; 2001 ; Smith, 2005).

fin de mot, tantôt en more nasale lorsque l'input est énoncé par un locuteur anglais, tantôt en nasale géminée suivie d'une voyelle épenthétique lorsqu'il l'est par un locuteur français) que « phonologique » (Smith, 2005, à propos des « doublets » en japonais, c'est-à-dire des mots d'emprunts ayant deux formes adaptées, l'une avec épenthèse et l'autre avec effacement dans le cas des codas et des groupes consonantiques), et il semblerait donc que les deux approches doivent être prises en considération, voire combinées, pour rendre compte de l'ensemble des faits.

#### 1.2.2.4. Aspects syllabiques : syllabes complexes et groupes consonantiques

Comme nous l'avons indiqué, les groupes consonantiques, tout comme les codas, semblent tout à fait illégaux en japonais. Cependant, il faut bien distinguer les groupes hétérosyllabiques des groupes tautosyllabiques, puisque, comme l'écrit Lombardi (1998, p. 1) : « the only possible clusters [in Japanese] are voiceless geminate and homorganic NC ». Citant Itô et les exemples qu'elle donne, elle indique qu'en japonais les seules consonnes possibles en position de coda sont celles qui sont liées par le lieu (« Place ») à la consonne en position d'attaque suivante, c'est-à-dire les groupes homorganiques, de type Nasale + Obstruante, ainsi que certains types de géminées. Itô et Mester (1993, p. 202) donnent également des exemples du même genre, que nous rassemblons ci-dessous :

#### Groupes NC assimilés sur le lieu (homorganiques)

- "kam.pai" (santé)
- "sam.po" (ballade)
- "dom.buri" (bol)
- "mon.dai" (problème)
- "han.tai" (opposé)
- "haŋ.ko" (sceau)
- "daŋ.go" (confiserie à base de pâte de riz)

### Géminées

- "kip.pu" (ticket)
- "gak.kō" (école)
- "kap.pa" (être légendaire)
- "tos.sa" (impulsivement)
- "kit.te" (timbre)
- "on.na" (femme)

Cela conduit donc à une contrainte de lieu sur la coda (« Coda Condition on Place »). Par ailleurs, les seules géminées possibles sont les obstruantes non-voisées et les nasales (les obstruantes non-voisées étant notablement moins marquées que les obstruantes voisées) (Lombardi, 1998). Lombardi indique aussi que les groupes homorganiques Nasale + Consonne (généralement Plosive) sont des groupes non-marqués. Si les groupes consonantiques tautosyllabiques sont donc bel et bien illégaux en japonais, il apparaît que l'existence de certaines codas permette néanmoins la contiguïté de certaines consonnes hétérosyllabiques.

#### 1.2.2.5. Aspects syllabiques : syllabes complexes et voyelles dévoisées

De la même manière qu'il est inexact, à un niveau phonologique, d'affirmer que deux consonnes ne peuvent se suivre dans un mot japonais, puisqu'il faut distinguer les cas des géminées et des groupes homorganiques NC, une telle affirmation est également inexacte à un niveau phonétique, puisque l'une des caractéristiques du japonais est de présenter, dans certains contextes, un dévoisement vocalique. Comme l'indique Vance (1987), le japonais est une langue qui présente la particularité de posséder des voyelles non-voisées (« voiceless »). Pour simplifier, on dit parfois que les voyelles japonaises /i/ et /u/ sont dévoisées en particulier entre les consonnes sourdes /k, s, t, p, h/ (le mot [suki] ressemblant alors à [ski] et [ʃita] ressemblant à [ʃta]) et, pour la voyelle /u/ en position finale de mot suivant une consonne sourde (la copule /desu/ étant alors prononcé [des]). Les voyelles non-voisées sont traitées la plupart du temps comme des

allophones des voyelles voisées qui leur correspondent, lorsqu'elles sont traitées dans un cadre phonémique. Vance (1987, p. 48) cite Fromkin et Rodman qui écrivent « In Japanese, high vowels are devoiced when preceded and followed by voiceless obstruents », avant d'indiquer que la situation est plus complexe que ce dont peut rendre compte une telle règle de simple assimilation.

En effet, les voyelles non hautes perdent également parfois leur voisement, mais le dévoisement de voyelles japonaises varie considérablement en fonction de facteurs linguistiques (position syllabique) et extra-linguistiques (locuteurs et situations de communication). Il existe cependant certaines tendances, que l'on peut résumer en disant (Han, cité par Vance, 1987, pp. 49-50) que pour un tempo « normal » (se pose alors le problème de la norme), c'est-à-dire durant une conversation polie avec des locuteurs natifs, les voyelles /a e o/ ne sont pas dévoisées alors que /i u/ le sont, et que le dévoisement s'applique en général à des syllabes non accentuées.

Il faut en effet souligner que le dévoisement est, bien évidemment, étroitement lié au schéma accentuel, d'autant que le japonais présente un type d'accent particulier (« pitch accent »), qui permet de distinguer des mots autrement apparemment identiques comme « hashi » (pont) et « hashi » (baguettes). L'intonation joue également un rôle : lorsque la voyelle susceptible d'être dévoisée figure dans la dernière syllabe d'un énoncé associé à un patron intonatif montant, le dévoisement opère moins systématiquement : par exemple « Nani ka arimasu ka ? » vs. « Nani ka arimasu » (= est-ce que vous avez quelque chose) ? Ceci peut notamment apparaître dans des énoncés qui ne se distinguent que par le contour intonatif : « Ashita ikimasu ? » (= vous y allez demain ?) vs. « Ashita ikimasu » (= j'y vais demain)<sup>61</sup>. De plus, les chances de dévoisement sont plus ou moins grandes selon l'environnement consonantique non voisé. Il semblerait (Vance, 1987, p. 51) que les possibilités de

---

<sup>61</sup> Niveau de langue familier; la forme moins familière serait : « Ashita ikimasu ka ? ».

dévoisement de la voyelle en question augmentent selon la classification suivante des consonnes qui la précèdent :

Fricatives (en particulier /s/ et /ʃ/) >> Affriquées >> Occlusives.

La question centrale qui nous occupe dans notre examen contrastif entre les systèmes phonético-phonologiques français et japonais est celle de l'opposition entre dévoisement et effacement : les voyelles situées dans un environnement dévoisant sont-elles simplement dévoisées ou bien effacées ? Vance (1987, pp. 52-53) cite plusieurs auteurs dont l'approche est différente : Bloch, dans sa démarche de phonémisation du japonais transcrit phonémiquement la plupart des séquences CVC comme /CC/. La seule exception notable est celle de la more obstruante /Q/ à la fin de la séquence : Bloch présente la voyelle précédente comme dévoisée et non effacée, comme par exemple dans « *suppai* » (aigre), transcrit phonétiquement par Vance comme [sɯ̥pːpai] qui, dans le système de Bloch donnerait /suppai/. Han ne partage pas cette vue, en particulier vis-à-vis du système phonotactique du japonais, qui, alors, comprendrait de nombreux groupes consonantiques. Kawakami considère que les voyelles en question sont parfois dévoisées (lorsqu'elles sont suivies par une consonne identique à la précédente ou par /Q/) et parfois effacées (en particulier après une fricative ou une affriquée), même si les phénomènes de coarticulation du [i] et du [ɯ] persistent sur la consonne précédente en cas d'effacement. Vance, quant à lui, considère que, dans un tempo rapide de conversation normale entre locuteurs natifs, certaines voyelles dévoisées disparaissent et que certaines combinaisons phonétiques, jugées phonologiquement impossibles hors contexte, apparaissent. Il en conclut que, même si pour des tempos rapides le phénomène d'effacement des voyelles dévoisées produit certains groupes consonantiques, au niveau phonologique auquel l'intuition des locuteurs natifs s'applique, un traitement phonémique du système japonais interdit ce genre de phénomène et inclut ainsi les voyelles considérées dans tous les cas.

A propos du dévoisement vocalique, Labrune (2001a, p. 102) indique que /i/ et /u/ se dévoisent soit lorsqu'elles sont précédées d'une consonne sourde et suivies d'une autre consonne sourde, soit lorsqu'elles précèdent une pause et non accentuées. Cependant, lorsque la voyelle est initialement accentuée, le dévoisement peut déclencher le déplacement de l'accent vers la more suivante. Comme le fait remarquer Labrune (2001a, p. 103) la présence des allophones [tʃ] et [ts] dans certains mots sont révélateurs de la présence de voyelles sous-jacentes (/i/ et /u/) apparemment effacées, puisque ce sont elles qui déclenchent cette affrication.

A propos du dévoisement des voyelles /i/ et /u/, Keating et Huffman (1984, p. 194) considèrent, quant à elles, comme « inadéquate » la formulation suivante : les voyelles hautes, entre consonnes sourdes ou entre une consonne sourde et une pause, sont dévoisées ou non-voisées, en particulier en ton bas (« low tone »). Citant Han, le dévoisement serait à relier à la longueur des voyelles, les voyelles courtes étant les plus sujettes au dévoisement. Cependant, pour Beckman, (cité par Keating et Huffman), les voyelles en question seraient bien effacées et non simplement dévoisées.

Comme on le voit à travers ces différents points de vue, la question semble donc ne pas encore avoir débouché sur un consensus général. Néanmoins, dans le cas d'apprenants japonais de français, ce phénomène peut conduire à la réalisation *phonétique* de groupes consonantiques, jugés pourtant difficiles pour ces sujets. D'un point de vue didactique, une telle tendance peut être envisagée de deux manières : positive en production, puisque les apprenants semblent à même d'imiter le modèle phonétique qui leur a été fourni ; moins positive dans d'autres tâches, puisque la représentation phonologique, activée en tâche de lecture ou d'écriture par exemple, peut différer de la forme phonétique précédente.

Enfin, en guise de conclusion de cette brève présentation d'une partie de la phonologie du japonais, il faut reconnaître que l'on ne peut traiter de la structure syllabique de cette langue, sans en évoquer, même brièvement, les caractéristiques accentuelles,



comme cela apparaît à travers le cas du dévoisement syllabique. L'accent japonais, qui constitue depuis longtemps un important thème de recherche dans le domaine, est défini par Labrune (2001a, p. 106) comme :

*« Un accent de hauteur distinctif, marqué phonétiquement par le passage d'une more haute (H) à une more basse (B). La dernière more du mot portant le ton H devant une more basse est considérée comme la more accentuée du mot. Toutes les mores placées devant elle, s'il y en a, sauf la première du mot, sont articulées avec un ton haut ».*

D'un point de vue phonétique, l'accent japonais est désigné en anglais par le terme « pitch accent », le « pitch » étant le corrélat perceptif de la fréquence fondamentale, déterminée par les vibrations des cordes vocales (Vance, 1987, pp. 77-78) : « The accent patterns of isolated Japanese words can be represented by specifying which moras carry high pitch and which moras carry low pitch ».

Le lexique japonais comporte donc des mots accentués et d'autres qui ne le sont pas. L'accent en japonais est réglé par deux principes généraux, celui de la consécuitivité (pas de tons bas entre des tons hauts dans un mot) et celui de la dissimilation initiale (la hauteur des deux premières mores du mot est toujours différente (HB, ou BH), sauf dans les cas de mores spéciales (HH)) (Labrune, 2001a, p. 106). D'un point de vue phonologique, le japonais est donc une langue à accent, « et pas une langue à tons comme le chinois, où chaque syllabe possède un ton propre, ne pouvant être déduit de celui des syllabes voisines » (*ibid*, p. 107). Il semble cependant exister une importante variation sociolinguistique (notamment due à l'âge et au dévoisement vocalique) (Labrune, 2001a, p. 109) dans l'accentuation, certains mots possédant ainsi plusieurs schémas accentuels.

Ainsi, après avoir présenté les systèmes phonético-phonologiques du français et du japonais, nous pouvons, avant d'effectuer une mise en contraste, introduire leur système phonographémique, et donc leur système d'écriture.

## **II Systèmes phonographémiques**

Puisque l'un des objectifs de notre étude est d'examiner les interactions possibles entre représentations orthographiques et phonologiques dans l'apprentissage de l'oral en L2, il nous faut, à ce stade, présenter les systèmes phonographémiques français et japonais. Par « système phonographémique », nous entendons le système d'associations existant, pour une langue donnée, entre les unités du système phonético-phonologique de cette langue et les unités graphiques qui les transcrivent sous une forme écrite normée, ce système étant bidirectionnel (oral → écrit / écrit → oral).

### **II.1. Système d'écriture, script et orthographe**

Ce que l'on nomme parfois « la primauté de l'oral sur l'écrit » se justifie en premier lieu par le fait que « toute langue est parlée avant d'être écrite » (Léon, 1993, p. 29). Katz et Frost (1992, p. 67) insistent eux aussi sur ce point : « Although writing systems are, in general terms, systems for communication, what they actually communicate is the spoken language — as opposed to communicating nonverbal ideas and meanings ». L'articulation entre oral, écrit et faculté de langage est clairement discutée par Durand (2000b), qui rappelle, en début d'article (p. 41), la position, à ce sujet, de deux grands linguistes qui ont inspiré l'ensemble de la linguistique moderne, à savoir F. de Saussure (dans son *Cours de linguistique générale*<sup>62</sup>) et L. Bloomfield (dans son ouvrage *Language*<sup>63</sup>) :

---

<sup>62</sup> 1916, Paris, Payot, Edition Tullio de Mauro, 1972, p. 45.

<sup>63</sup> 1933, Londres, Georges Allen & Unwin, 13<sup>ème</sup> impression 1976, p. 21.

<b>Saussure</b>	<b>Bloomfield</b>
<p>« Langue et écriture sont deux systèmes de signes distincts : l'unique raison d'être du second est de représenter le premier ; l'objet linguistique n'est pas défini par la combinaison du mot écrit et du mot parlé ; ce dernier constitue à lui seul son objet. Mais le mot écrit se mêle si intimement au mot parlé dont il est l'image, qu'il finit par usurper son rôle principal ; on en vient à donner autant et plus d'importance à la représentation du signe vocal qu'à ce signe lui-même. C'est comme si l'on croyait que pour connaître quelqu'un, il vaut mieux regarder sa photographie que son visage ».</p>	<p>« Writing is not language but merely a way of recording language by means of visible marks ».</p>

Durand expose alors les arguments (phylogénèse, ontogénèse et nature des systèmes d'écriture) qui ont conduit à cette « domination » de l'oral dans la linguistique post-saussurienne, « domination » due non pas à une quelconque « supériorité » de l'oral sur l'écrit, mais à son caractère plus « naturel » (p. 44). Analysant ensuite les critiques portées contre ces arguments par des auteurs comme Derrida ou Kristeva, et tout en reconnaissant, d'une part que « l'actualisation du langage dans le canal vocal et auditif n'est pas un trait définitoire de la faculté de langage » (p. 46), notamment vis-à-vis des langues signées utilisées par les mal-entendants, et d'autre part qu'il est « faux de voir les systèmes d'écriture comme de simples tentatives de codage de la langue orale au sens strict du terme » (p. 54), notamment vis-à-vis des « écritures » non-alphabétiques, en particulier picturales, Durand insiste cependant sur la nécessité de maintenir une distinction entre le « linguistique », le « paralinguistique » et l' « extralinguistique ». Il conclue en indiquant que, « bien que la faculté de langage ne soit pas en droit lié à une

modalité particulière, il fallait distinguer son inscription naturelle dans la substance orale ou signée de son inscription dans l'écriture au sens ordinaire du terme » (p. 56).

Il semblerait dès lors que la prépondérance accordée en linguistique aux formes orales des langues naturelles, pour des motifs tout à fait justifiables, ait relégué le rôle des formes écrites dans la langue au second plan, alors que celui est pourtant loin d'être inexistant. Pour Ducrot et Todorov (1972, p. 435) : « presque partout (partout en Occident, sous la domination de l'écriture phonétique), le langage *parlé* s'est trouvé privilégié comme constituant le langage par excellence, dont le langage *écrit* ne serait qu'une image redoublée, une reproduction auxiliaire ou un instrument commode – *signifiant de signifiant* ». Or, comme le note Jaffré (2003, p. 38), l'écrit constitue un objet d'étude linguistique à part entière :

*« La plupart des linguistes reprennent la position inspirée de Saussure pour qui l'écriture n'est encore qu'un système de signes second, une sorte de "vêtement" de la langue orale, et finalement un objet relevant plutôt de l'anthropologie. Pour la linguistique de l'écrit, au contraire, l'oral et l'écrit sont à mettre d'emblée sur le même plan. L'écriture est un système sémiologique ni plus ni moins complexe que celui de l'oralité, qui interagit avec lui et doit remplir des fonctions spécifiques. Le fait qu'il ait fallu attendre 2001 pour trouver un chapitre sur l'écriture dans le Handbook of Linguistics (Daniels, 2001) témoigne de ce long désaccord épistémologique ».*

C'est également à ce « phonocentrisme » que s'attaque Coulmas (1989) dans sa préface, soulignant lui aussi le rôle de l'écrit dans l'évolution globale de la langue :

*« This book is one response to the disregard of writing prevalent in modern linguistics. [...] It was through my concern for non-alphabeticallly written languages that I began to see more clearly that languages are affected in various ways by the writing systems used for their visual materialization, and hence that writing should not only be accorded a proper place in linguistics as*

*an object in its own right, but also be reckoned with as a factor of language change ».*

Dans le cas du français, une telle influence a notamment été envisagée sous l'appellation « effet Buben » (Chevrot et Malderez, 1999), comme le note Durand (2000b, p. 55) : « On sait aussi que dans des langues comme le français l'écriture a une influence [...] sur les usages oraux ». Catach en donne l'illustration suivante (1995, p. 67) :

*« Notre langue, qui ne prononçait il y a deux siècles pratiquement aucun groupe de consonnes, si ce n'est avec les liquides (r et l), se surcharge actuellement, sous les actions combinées des emprunts, de l'école, de l'orthographe, etc., de nombreux groupes internes et finaux (cheptel, dompter, sangsue, etc.) ».*

Ces prémisses étant posées, nous pouvons à présent nous concentrer sur les fonctions de l'écrit. Coulmas distingue quatre fonctions essentielles de l'écriture (1989, pp. 11-13) :

- **La fonction mnémonique :** *« History becomes possible thanks to the mnemonic function of writing, as well as the accumulation of knowledge ».*
- **La fonction de distanciation :** *« Writing [...] enables communication over any distance in space or time. [...] Writing is a distancing medium not only with respect to sender and receiver, but also as regards the sender and the message ».*

- **La fonction de réification :** *« More obviously than its oral counterpart, which disappears as soon as it materializes, the written message assumes the qualities of an object. [...] The spoken word is ephemeral and spontaneous [...]. In writing, on the other hand, words become stable and tangible ».*
- **La fonction de contrôle social :** *« The notions of law and right, of standard and correctness, are closely linked with writing ».*

Dans la perspective qui est la nôtre, ce sont bien les première et troisième fonctions qui retiennent notre attention, puisque ces deux fonctions sont généralement jugées positives en classe de langue, l'écrit servant de support, mnésique et perceptif, indispensable à l'apprentissage de la langue étrangère. Comme l'indique Mondada et Dubois (1995, p. 295) : « L'écriture permet de disposer et de fixer dans des relations spatiales le flux temporel des mots du discours oral ». Permettre une telle disposition dans l'espace (fonction de réification) et une telle fixation (fonction mnémotechnique) sont deux des atouts majeurs du support écrit. Sans un tel support, on pourrait imaginer que les difficultés posées aux apprenants lors de l'apprentissage d'une langue étrangère seraient bien plus nombreuses qu'elles ne le sont déjà. Pourtant, malgré ces deux arguments de poids, on ne peut manquer de s'interroger sur les possibles « effets secondaires » d'un tel support sur l'apprentissage des formes orales que celui-ci est censé transcrire.

Afin de clarifier l'usage des termes « système d'écriture » (« writing system »), « script » (« script ») et « orthographe » (« orthography »), Coulmas donne les exemples suivants (1989, p. 37-38) : « The term writing system will be used to differentiate systems depicting linguistic units of different structural levels. These units are words, syllables and phonemes ». Ainsi, l'anglais et le néerlandais partagent le

même système d'écriture : le système alphabétique. Ils partagent également le même script : l'alphabet romain. « Scripts are thus graphical instanciations of writing systems ». Le système alphabétique est ainsi réalisé à travers différents scripts : l'alphabet romain, grec ou russe par exemple. En ce qui concerne l'orthographe, l'auteur la définit en présentant, à titre d'exemple, certaines différences entre l'anglais britannique et américain, ou l'allemand standard et l'allemand suisse. Ainsi du *système d'écriture* à *l'orthographe*, via le *script*, on passe du plus général au plus spécifique.

Taylor et Taylor (1995), quant à eux, distinguent également les termes « système d'écriture » et « script » de la manière suivante (p. 10) :

*« "Writing system" may refer to a type of system such as alphabet, syllabary, and logography [logo=word ; graph=written symbol]. Or it may refer to a set of different scripts used to represent one language, in expressions such as the "Japanese writing system", which consists of a set of logographs and two forms of a syllabary. "Script" may then refer to an individual form of some type of writing system, such as the Roman alphabet, the Greek alphabet, or the Cyrillic alphabet (used in Russia) under the writing system "alphabet". The term script may refer also to a style of writing, such as the "clerical" and "cursive" script styles of Chinese characters ».*



La distinction entre les trois termes est synthétisée par Coulmas dans le tableau suivant (1989, pp. 38-39) :

<b>Proto-écriture</b>		Picture writing Idea writing
<b>Système d'écriture</b>	« Every writing system makes a selection of the linguistic units to be graphically represented (not language specific) ».	Word writing Morpheme writing Syllable writing Phoneme writing Phonetic writing
<b>Script</b>	« Every script makes a specific selection of the possibilities of a given system in accordance with the structural conditions of a given language ».	Chinese script Arabic script Greek script
<b>Orthographe</b>	« Every orthography makes a specific selection of the possibilities of a script for writing a particular language in a uniform and standardized way ».	Chinese/Taiwanese orthography Standard German/Swiss German orthography

Cependant, comme il le note (p. 39), la sélection opérée par le script pose le problème de l'existence de ce qui est sélectionné :

*« The objects which are « selected » are not necessarily given. To some extent, they are also created. Writing, in other words, not only maps, but also imposes structure. Part of the general problem of defining linguistic categories and units is that often no proper attention is paid to the important differences between spoken language and written language ».*

La notion même de « mot » est difficile à cerner : d'après l'auteur, on pourrait envisager un mot orthographique, un mot phonétique et un mot sémantique (p. 40). Coulmas cite ainsi certaines sources (Goody) selon lesquelles certaines langues d'Afrique de l'Ouest sans tradition d'écriture n'auraient pas de mot pour « mot » (langues LoDagaa et Gonja).

Coulmas en arrive donc à la conclusion suivante (p. 40) :

*« (1) The independent existence of the units supposedly depicted by a given writing system cannot be taken for granted, and (2) that in mapping language, writing systems, scripts and orthographies are as much based on analytic perception as they form them and the resulting conceptualizations of the structural units of language ».*

C'est précisément dans la deuxième partie de cette conclusion que ressort l'influence active de l'écrit sur la perception et la conceptualisation des unités linguistiques, deux étapes qui sont essentielles dans l'apprentissage d'une langue étrangère.

## **II.2. Typologie des systèmes d'écriture et « profondeur orthographique »**

Taylor et Taylor (1995, p. 9) distinguent *en premier lieu* (cette dichotomie étant bien souvent abusivement utilisée) deux types de système d'écriture : ceux qui représentent

essentiellement des unités de sens, et ceux qui représentent essentiellement des unités de son. Ils l'illustrent avec l'exemple du dessin d'un parapluie sur un paquet<sup>64</sup> :

*« A drawing of an umbrella on a package could stand for various messages, such as "This package contains umbrellas", "The content of this package should not be exposed to rain", and simply "Keep dry". Some cultures [...] may never have seen an umbrella ».*

Citant DeFrancis, Katz et Frost distinguent, quant à eux, trois catégories de systèmes d'écriture, fonctions des rapports entretenus entre le script (ensemble de symboles) et la langue transcrite, rapports qui correspondent à la notion d'orthographe (1992, p. 68) :

- Type logographique-phonétique.
- Type syllabique.
- Type alphabétique.

A l'intérieur de cette dernière catégorie, on peut également trouver une grande variété d'orthographe alphabétiques. Katz et Frost (1992) opposent ainsi l'hébreu au serbo-croate, dont le principe est le suivant (p. 69) : « spell a word like it sounds and speak it the way it is spelled ». A l'origine de cette variété peuvent se trouver plusieurs raisons, mais ce sont deux motifs essentiels purement linguistiques sur lesquels insistent les auteurs (pp. 70-71) :

- D'une part le degré de complexité de la langue orale, en ce qui concerne les relations entre phonologie et morphologie ;
- D'autre part la possibilité, pour une langue phonologiquement complexe, de représenter soit l'invariance morphologique (ce qui débouche sur une orthographe « profonde »), soit l'invariance graphème-phonème (ce qui débouche sur une orthographe « superficielle »).

---

<sup>64</sup>A propos de la « communication iconique », voir (Yazdani et Barker, 2000).

Ce degré de transparence des correspondances graphèmes-phonèmes correspond à la « profondeur » de l'orthographe, laquelle permet de répartir les différentes orthographes alphabétiques en deux grandes catégories (Katz et Frost, 1992, p. 71) :

- **Superficielle** : Si les correspondances lettres-phonèmes sont uniques (« completely and consistently »).
- **Profonde** : Si les relations lettre-phonèmes sont plurielles (« substantially »), c'est-à-dire que : « some letters have more than one sound and some phonemes can be written in more than one way or are not represented in the orthography ».

Ces distinctions d'ordre linguistique ont également été testées dans le domaine neurolinguistique : ainsi, en ce qui concerne les écritures dites « phonétiques », certains chercheurs ont pu, grâce à des cas cliniques, mettre en évidence le fait que « des aires fonctionnelles différentes en HG [hémisphère gauche] peuvent être impliquées dans le traitement de différentes écritures phonétiques, selon le degré de correspondance son-symbole » (Galloway, 1983, p. 83). En ce qui concerne les écritures dites « idéographiques »<sup>65</sup>, il semblerait que ces dernières, sous certaines conditions, soient mieux traitées par l'hémisphère droit, alors que les écritures phonétiques le seraient par l'hémisphère gauche. L'exemple du japonais est intéressant, puisqu'il comporte les deux types de code. On aurait ainsi une participation supérieure de l'HG dans le traitement des kanas, et une implication accrue de l'HD dans celui des kanjis, des résultats expérimentaux confirmés par les données cliniques (Galloway, 1983, p. 83)<sup>66</sup>. Vis-à-vis de l'influence de l'écrit sur l'apprentissage de la L2, en considérant ici un

---

<sup>65</sup> Une appellation à présent nettement abandonnée par les spécialistes du domaine, certains d'entre eux qualifiant par exemple de « syllabo-morpho-phonétique » l'écriture chinoise, pour insister sur l'information phonétique qu'elle contient dans la très grande majorité des cas, en dépit des changements diachroniques, et pour se couper ainsi de cette « tradition » de l'idéographie.

<sup>66</sup> Cependant, certaines données plus récentes (Kess et Miyamoto, 1999) indiquent qu'une telle dichotomie est abusivement simplificatrice.

type d'apprentissage institutionnel « traditionnel », axé sur la lecture et l'écriture au détriment de la production et de la compréhension orale, il est intéressant de noter les observations de Galloway (1983, p. 100), qui indique qu'il existe « d'une part une relation entre la disponibilité des indices pragmatiques et de l'expérience interactionnelle en L2 et l'absence de différence asymétrique entre L1 et L2, et d'autre part une relation entre l'absence d'expérience pragmatique contextualisée en L2 et une implication plus forte de l'HG en L2 qu'en L1, [...] [cette dernière pouvant disparaître] à mesure que s'accroît [la] compétence et [l']habileté pragmatique en L2 ».

Cependant, si les systèmes orthographiques sont habituellement classés en trois types (alphabétique (symboles pour phonèmes), syllabique (symboles pour syllabes) et logographique (symboles pour morphèmes ou mots)), Kess et Miyamoto (1999, p. 8) soulignent que, au sein d'un même système, les choses ne sont pas toujours aussi tranchées, ce qu'ils illustrent dans le cas de l'anglais : « Graphemic units are often tied to an intermediate morphophonemic level before they are related to sound ». On sait également par ailleurs que, en chinois (système généralement qualifié de « logographique »), l'accès à l'information phonologique est souvent possible, voire nécessaire, en procédure de reconnaissance visuelle.

Ces considérations générales étant établies, nous pouvons à présent aborder les cas spécifiques du français et du japonais. Il ne s'agira pas pour nous de développer toute la complexité de ces systèmes, mais plus modestement, et comme pour la partie phonologique, de mettre en relief certains points pertinents pour notre étude.

### **II.3. Le système d'écriture français**

Le français ne possède qu'un système d'écriture, le système alphabétique, et qu'un type de script, l'alphabet romain, composé de 26 lettres : a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z<sup>67</sup>. L'orthographe française peut être qualifiée de

---

<sup>67</sup> Il faut néanmoins mentionner la différence entre caractères majuscules et minuscules d'une part, l'utilisation de diacritiques, tels les accents, d'autre part, et enfin celle des chiffres « arabes ».

« profonde », en ce que les correspondances de phonème à graphème sont loin d'être univoques.

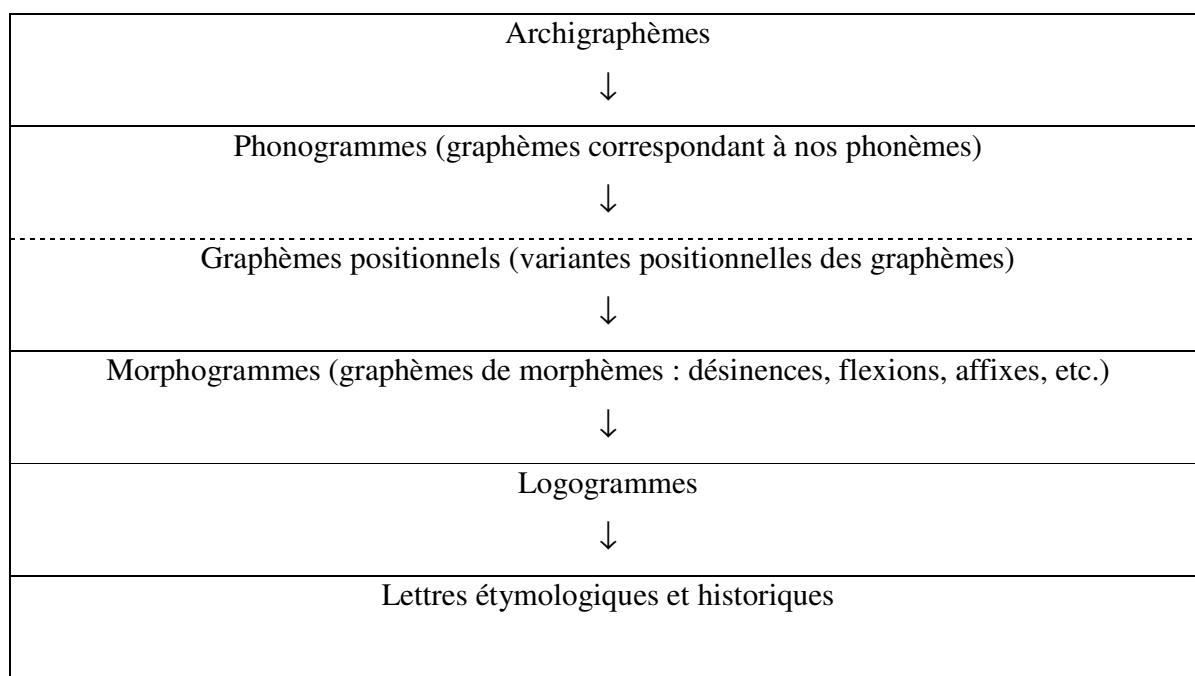
Cependant, alors que le français est bien connu pour sa complexité et ses ambiguïtés dans le codage son→lettres, l'orthographe française est également considérée comme assez systématique dans la conversion lettres→ son, une observation confirmée par l'analyse statistique du système grapho-phonologique français effectuée par Lange et Content (1999, p. 438). Les associations graphèmes-phonèmes semblent en effet assez prévisibles, avec 103 graphèmes, 172 associations et une probabilité d'association moyenne assez élevée (60%). Plus de 40% des associations sont tout à fait régulières et non ambiguës. Lorsque de multiples prononciations existent pour un graphème donné (en moyenne 1,7 prononciations par graphème), les prononciations alternatives ont des probabilités d'association assez faibles (moins de 15%).

Cinq principes peuvent être évoqués pour rendre compte du système phonographémique du français (recensés par Gak, cité par Zesiger et Partz, 1994, p. 419) :

- Le principe phonético-graphique ;
- Le principe morphologique (<d> : grand/grande/grandir) ;
- Le principe étymologique (<h> : homme/ latin *homo*) ;
- Le principe historique (graphies traditionnelles conservées) ;
- Le principe de différenciation (homonymes : poids, pois, poix).

L'une des caractéristiques de l'orthographe française, qui semble parfois la rendre « opaque », tient évidemment à ses nombreuses et rapides mutations graphiques au cours des siècles. Ainsi, si l'on compare la première édition du dictionnaire de l'Académie française, avec celle de 1935, sur 17168 mots dépouillés, on note 17750 modifications graphiques et ainsi, depuis le XVIème siècle, 8807 *mots* (soit 51,3 % du corpus total) qui ont changé de forme (Catach, 1978/1995, pp. 44-45).

Pourtant, malgré cet apparent foisonnement de formes et de variations, Catach et son équipe HESO (Histoire et Structure des Orthographe et Systèmes d'écritures) du CNRS, pour lesquels « l'orthographe n'est pas seulement un code ou une institution sociale, mais un ensemble complexe de signes linguistiques » (*op. cit.*, p. 52), sont parvenus à une conception systémique de l'orthographe française, en tant que « plurisystème ». En se basant sur la terminologie de Martinet (phonème<sup>68</sup>, morphème et lexème), Catach (1978/1995) propose de représenter la structure de notre orthographe par une série de cercles concentriques, qui se présente, du cœur à la périphérie, comme suit :



Cherchant à définir le « graphème », Catach insiste sur la nécessité de prendre en compte l'ensemble du système pour parvenir à une telle définition. Tout en considérant les « phonogrammes (transcription des sons) » comme les fondamentaux du système, Catach distingue des graphèmes « centraux » et « secondaires » (p. 57) :

---

<sup>68</sup> Catach (*op. cit.*, p. 53, note 1) note cependant que: « Au phonème correspond non pas une unité seulement distinctive, mais aussi en partie *significative*, le graphème. Cette différence tient au fait que le système graphique est relativement *autonome* par rapport à la langue, tout comme le morse ».

« On aboutit ainsi à une définition du graphème comme **forme maximale stable** [...] du phonème ou du morphème, en opposition, d'une part aux **variantes combinatoires** (par rapport au groupe de mots, au mot, à la syllabe, à l'entourage immédiat), d'autre part aux sous-graphèmes et aux exceptions ».

En vue de définir les graphèmes de base, cinq critères sont retenus (*op. cit.*, p. 58) :

- 1) La fréquence de l'unité : Par exemple, le graphème « ch » vs. « sch » (beaucoup moins fréquent).
- 2) La cohésion et la stabilité : Par exemple, le graphème « eau » vs. « ean ». (pour les graphèmes complexes)
- 3) La « signifiante » ou pertinence phonologique : Par exemple, les lettres doubles ou grecques, comme « rh », sont peu significatives phonologiquement.
- 4) La « rentabilité » linguistique : Par exemple les rapports morphologiques sériels.
- 5) La créativité linguistique : Par exemple, certains suffixes disparaissent (« -eux » dans « boueux » au profit de « -eurs »), tandis que d'autres apparaissent (« -ing » dans « parking »).

Selon ces critères, Catach parvient à 36 graphèmes vocaliques, 30 graphèmes consonantiques et 6 graphèmes de semi-voyelles (*op. cit.*, p. 59). Parmi tous ces graphèmes, 45 d'entre eux présentent un haut degré de fréquence, de stabilité et de pertinence, et peuvent être considérés comme les graphèmes de base du français. Cependant, une analyse encore plus fine permet de dégager les 33 archigraphèmes du français, qui en seraient le noyau graphémique (*op. cit.*, p. 61) :



A	E	I	O	U	EU	OU
AN		IN	ON	UN		
		ILL				
		Y				
			OI			
			OIN			
P.B – T.D – C.G – F.V – S.Z – X – CH.J – L.R – M.N. - GN						

D'après ce schéma, Catach conclut que le « plurisystème » orthographique français est (p. 64) :

*« Une orthographe **phonologique** (à plus de 80%), dans une certaine mesure morphologique (avec une concentration extrême des "morphogrammes grammaticaux", bien moindre des "morphogrammes lexicaux"), et enfin en partie distinctive (les distinctions des homophones grammaticaux étant de loin plus informatives que celles des homophones lexicaux) ».*

La définition que propose Catach du graphème est donc résumée ainsi (*op. cit.*, p. 118) :

*« **Graphème** : la plus petite unité distinctive et/ou significative de la chaîne écrite, composée d'une lettre, d'un groupe de lettres (digramme, trigramme), d'une lettre accentuée ou pourvue d'un signe auxiliaire, ayant une référence phonique et/ou sémique dans la chaîne parlée. Ex. : "p, ou, r, ch, a, ss, e, r", dans "purchasser". On peut classer les graphèmes en trois catégories :*

- *Les **phonogrammes**, ou graphèmes chargés de transcrire les sons ; leur usage est réglé par les lois de position ; ex. : g dans gare, gu dans gué ;*
- *Les **morphogrammes**, surtout situés, pour les renforcer, aux jointures des mots, maintenus graphiquement identiques qu'ils soient prononcés ou non ; ex. :*

*marques de féminin/masculin, singulier/pluriel, suffixes/préfixes, radicaux/dérivés, etc. ;*

- *Les **logogrammes**, ou "figures de mots", dans lesquels, à la limite, la "graphie" ne fait qu'un avec le mot, dont on ne peut la dissocier (ex. : sept, thym, pouls, coing, etc.) ».*

L'archigraphème, quant à lui, est défini ainsi (p. 119) :

*« **Archigraphème** : graphème fondamental, représentant d'un ensemble de graphèmes correspondant au même phonème ou au même archiphonème (ex. O pour o, ô, au, eau, etc.). Certains archigraphèmes sont des "idiotismes graphiques" (X, OI, OIN) ».*

Un fait particulièrement intéressant pour notre problématique est lié à l'évolution, orale et écrite, du français. En effet, tout en notant les différents types de modifications possibles, Catach indique que certaines d'entre elles sont d'ordre purement graphique (*op. cit.*, p. 67) :

*« Beaucoup de consonnes ou de groupes de consonnes qui n'avaient jamais été prononcées auparavant le deviennent aujourd'hui, sous l'influence de l'orthographe. Le phénomène a été étudié par Vl. Buben<sup>69</sup> en 1935 [...]. Il n'a pas cessé depuis. Le r, le l sont les plus touchés : les meilleures exemples en sont d'une part les infinitifs des deuxième et troisième groupes (mourir, devoir, dont le r ne se prononçait généralement pas au XVIIIème siècle), d'autre part les mots en -il (grésil, mil, gril, etc.) où la consonne est très généralement prononcée aujourd'hui. Le besoin de distinguer entre eux les monosyllabes les fait aujourd'hui prononcer dans certains cas : nez et net, sot et soc, rot et roc, art et arc, dos et dot, etc. [...] Notre langue, qui ne prononçait il y a deux siècles pratiquement aucun groupe de consonnes, si ce n'est avec les liquides (r*

---

<sup>69</sup> *Influence de l'orthographe sur la prononciation du français moderne*, Bratislava et Paris, Droz, 1935.

*et l), se surcharge actuellement, sous les actions combinées des emprunts, de l'école, de l'orthographe, etc., de nombreux groupes internes et finaux (cheptel, dompter, sangsue, etc.) ».*

Cela rejoint ce qu'écrit Carton (1974, p. 200) : « Vers 1900, la seule prononciation admise pour *cerfs*, *ours*, au pluriel, était : [sɛ :ʀ], [u :ʀ], et on devait dire [ɛgza], [almana] pour *exact*, *almanach*. Sous l'influence de l'orthographe, les jeunes ont de plus en plus souvent prononcé les consonnes finales ». Ce phénomène, désigné sous le nom d'« effet Buben » (Chevrot et Malderez, 1999), est particulièrement marquant pour nous, puisque l'une des questions sous-jacentes à notre problématique pourrait être formulée ainsi :

Existe-t-il un « effet Buben » dans l'apprentissage d'une langue étrangère, cet effet pouvant être envisagé en termes intra-linguistiques (influence de l'orthographe en L2 sur la prononciation en L2) ou, plus encore, en termes inter-linguistiques (influence de l'orthographe en L1 sur la prononciation en L2) ?

Il convient donc dans un premier temps, pour tenter de répondre à cette question, d'établir le système orthographique de la L1, en l'occurrence celui du japonais.

#### **II.4. Le système d'écriture japonais**

L'une des caractéristiques majeures du système d'écriture japonais est certainement sa pluralité. Outre sa pluralité « traditionnelle », la situation sociolinguistique urbaine du Japon a également conduit à une évolution rapide du lexique orthographique, et dans une certaine mesure phonologique, japonais, principalement due à l'utilisation considérable des emprunts étrangers, et plus particulièrement anglo-américains, dans la vie quotidienne.

### **II.4.1. Un système pluriel**

Le système d'écriture japonais est traditionnellement triple, voire actuellement quadruple puisque « with the increasing use of the Latin alphabet-base *romaji*, [Japanese] has now four script types » (Kess et Miyamoto, 1999, p. 9).

Trois systèmes coexistent donc dans l'écriture japonaise :

<ul style="list-style-type: none"> <li>• <u>Morphophonique</u> :</li> </ul>	<p>Les <i>kanjis</i>, caractères d'origine chinoise, parfois malhabilement (Matsunaga, 1994) nommés "idéogrammes".</p>	<p><u>Par exemple</u> : 漢字 (kan – ji = kanji)</p>	
<ul style="list-style-type: none"> <li>• <u>Moraïque</u> :</li> </ul>	<p>Les <i>kana</i>, caractères japonais, séparés en deux types de scripts, que l'on appelle parfois « syllabaires ».</p>	<ul style="list-style-type: none"> <li>• Les <i>hiragana</i></li> </ul>	<p><u>Par exemple</u> : かんじ (ka – N – ji = kanji)</p>
		<ul style="list-style-type: none"> <li>• Les <i>katakana</i></li> </ul>	<p><u>Par exemple</u> : カンジ (ka – N – ji = kanji)</p>

<ul style="list-style-type: none"> <li>• <u>Alphabétique :</u></li> </ul>	<p>Les <i>romaji</i>, caractères issus de l'alphabet romain : a, i, u, e, o pour les voyelles ; b, c<sup>70</sup>, d, f, g, h, j, k, m, n, p, r, s, t, z pour les consonnes et w, y pour les semi-consonnes. Cela correspond à 22 lettres sur 26 (puisque l, q, v et x n'en font pas partie).</p>	<p><u>Par exemple :</u> kanji (ka – N – ji)</p>
---	---	---

Cette pluralité s'explique à travers l'organisation, historique et fonctionnelle, du lexique japonais, qui comporte, comme nous l'avons décrit dans la partie phonologique, quatre grandes classes morphémiques. A chacune d'entre elle est globalement associé un type d'écriture : les kanji à la strate Sino-japonaise, les hiragana à la strate Yamato et les katakana aux strates Etrangère et Mimétique. Les romaji, quant à eux, sont utilisés pour transcrire les mots japonais en alphabet romain. Cependant, comme le notent Kess et Miyamoto (1999, p. 14), « there is no strict one-to-one correspondence between type of vocabulary item and script type » et un tel découpage constitue une simplification grossière des modes d'utilisation réelle de ces différents types d'écriture. Nous revenons sur ce point dans la partie psycholinguistique de notre étude.

Alors que les kanji, qui représentent une grande majorité des morphèmes lexicaux des énoncés, ont reçu une attention considérable de la part des chercheurs, notamment occidentaux, en raison de leurs différences vis-à-vis des systèmes alphabétiques, nous sommes, quant à nous, davantage préoccupés par le rôle des kana et des romaji dans la perspective didactique qui est la nôtre, puisque ce sont eux qui peuvent être utilisés par les apprenants pour éventuellement transcrire des items étrangers, et qui peuvent, dans

---

<sup>70</sup> Dans le graphème « ch » ou « tch », selon le système de translittération adopté.

le cas des romaji, se superposer en partie à l'alphabet romain utilisé pour le français. C'est donc sur eux que sera concentrée notre présentation.

Les textes japonais sont pour l'essentiel des mélanges de kanji et de kana (« kanji/kana majiribun » = phrases mixtes kanji/kana), et ils peuvent être écrits dans deux modes directionnels différents (Kess et Miyamoto, 1999, p. 163) :

<ul style="list-style-type: none"> <li>• <u>Traditionnel</u> (« tategaki ») :</li> </ul>	Mode vertical, avec les caractères placés de haut en bas et les colonnes de droite à gauche (origine chinoise).	<div style="text-align: right;">             まく米は日              す食を、本              。&gt;よおで           </div> Nihon de wa, okome wo yoku tabemasu ( = Au Japon, on mange beaucoup de riz).
<ul style="list-style-type: none"> <li>• <u>Importé</u> (« yokogaki ») :</li> </ul>	Mode horizontal, avec les caractères placés de gauche à droite et les lignes de haut en bas (influence occidentale).	日本では、お米をよく食べます。 Nihon de wa, okome wo yoku tabemasu ( = Au Japon, on mange beaucoup de riz).

Un troisième mode existe : horizontal et de droite à gauche. Ce mode est très peu utilisé, actuellement réservé à certains textes traditionnels, notamment religieux.

Il faut souligner que de nombreux magazines que l'on trouve quotidiennement au Japon, qu'ils soient importés comme « *Elle Japon* » ou simplement de production locale japonaise, mélangent les modes, en ce que la lecture du magazine (ordre des pages) s'effectue de droite à gauche, tandis que les caractères horizontaux se lisent de gauche à droite. Le mélange des genres semble ainsi être la règle dans le Japon urbain d'aujourd'hui, comme en témoigne cette inscription que nous avons observée sur une affiche publicitaire à Tokyo en juillet 2004 : le mot « shopping » y était écrit ainsi : « (shi) (petit yo) (petit tsu) PPING » (« しよつPPING »), c'est-à-dire un mélange entre

hiragana (prononciation japonaise, avec gémination de la consonne) et alphabet romain (mot étranger et non adapté, conservant le digraphe "PP").

#### **II.4.2. Les kana**

Contrairement au français (orthographe profonde), le japonais présente une transparence élevée pour le système kana. Kess et Miyamoto (1999, p. 89) indiquent en effet que :

*« The grapheme-to-mora ratio is extremely high for kana in both syllabary systems, and the idealized one-to-one principle in orthographic correlations is close to being realized in Japanese orthography. [...] [Exceptions are] the hiragana symbol « ん » and « ん » [which] are read as /ha/ and /he/ normally, but when they function as case-marking particles after nominals, they take on an archaic reading of /wa/ and /e/ ».*

Les caractères kana se sont développés historiquement à partir des kanji, selon l'évolution suivante (Taylor et Taylor, 1995, p. 308) : kanji → kanji en tant que signes phonétiques → formes simplifiées des kanji → kana. Ces derniers caractères sont organisés sous forme de tableaux, dans un ordre déterminé (« Gojuonzu » = le tableau des 50 sons), que nous présentons ci-dessous :

Syllabaire Hiragana										
	A		I		U		E		O	
.	あ	a	い	i	う	u	え	e	お	o
K	か	ka	き	ki	く	ku	け	ke	こ	ko
S	さ	sa	し	shi	す	su	せ	se	そ	so
T	た	ta	ち	chi	つ	tsu	て	te	と	to
N	な	na	に	ni	ぬ	nu	ね	ne	の	no
H	は	ha	ひ	hi	ふ	fu	へ	he	ほ	ho
M	ま	ma	み	mi	む	mu	め	me	も	mo
Y	や	ya			ゆ	yu			よ	yo
R	ら	ra	り	ri	る	ru	れ	re	ろ	ro
W	わ	wa							を	wo
N	ん	n								

Syllabaire Hiragana Dérivé										
	A		I		U		E		O	
G	が	ga	ぎ	gi	ぐ	gu	げ	ge	ご	go
Z	ざ	za	じ	ji	ず	zu	ぜ	ze	ぞ	zo
D	だ	da	ぢ	ji	づ	zu	で	de	ど	do
B	ば	ba	び	bi	ぶ	bu	べ	be	ぼ	bo
P	ぱ	pa	ぴ	pi	ぷ	pu	ぺ	pe	ぽ	po



Syllabaire Katakana										
	A		I		U		E		O	
.	ア	a	イ	i	ウ	u	エ	e	オ	o
K	カ	ka	キ	ki	ク	ku	ケ	ke	コ	ko
S	サ	sa	シ	shi	ス	su	セ	se	ソ	so
T	タ	ta	チ	chi	ツ	tsu	テ	te	ト	to
N	ナ	na	ニ	ni	ヌ	nu	ネ	ne	ノ	no
H	ハ	ha	ヒ	hi	フ	fu	ヘ	he	ホ	ho
M	マ	ma	ミ	mi	ム	mu	メ	me	モ	mo
Y	ヤ	ya			ユ	yu			ヨ	yo
R	ラ	ra	リ	ri	ル	ru	レ	re	ロ	ro
W	ワ	wa							ヲ	wo
N	ン	n								

Syllabaire Katakana Dérivé										
	A		I		U		E		O	
G	ガ	ga	ギ	gi	グ	gu	ゲ	ge	ゴ	go
Z	ザ	za	ジ	ji	ズ	zu	ゼ	ze	ゾ	zo
D	ダ	da	ヂ	ji	ヅ	zu	デ	de	ド	do
B	バ	ba	ビ	bi	ブ	bu	ベ	be	ボ	bo
P	パ	pa	ピ	pi	プ	pu	ペ	pe	ポ	po

Chaque script kana contient ainsi 46 symboles de base auxquels viennent s'ajouter 25 symboles avec des signes diacritiques, ce qui élève le tout à 71 symboles. A cela, il faut également additionner 33 syllabes de type « palatales » issues de la combinaison « Consonne + Semi-voyelle /j/ + Voyelle », ce qui fait un grand total de 104 kana. La structure interne du syllabaire, si elle s'inspire de critères phonétiques, s'inspire aussi de critères morphophonémiques, en particulier en ce qui concerne le voisement et les signes diacritiques qui le signalent : les couples sonores/sourdes reflètent en effet une alternance morphologique connue sous le nom de « rendaku » dont un exemple, déjà cité précédemment, est le suivant : « te » (main) + « kami » (papier) → « tegami » (lettre). Par ailleurs l'organisation spatiale du syllabaire (lignes et colonnes) n'est pas triviale, puisque c'est ainsi qu'il reste en mémoire chez les locuteurs, qui associent ainsi k/g, t/d et h/b/p par exemple (Kess et Miyamoto, 1999, p. 86).

Tandis que les hiragana sont utilisés pour certains morphèmes lexicaux, grammaticaux et surtout pour « the grammatical scaffolding of Japanese sentences » (Kess et Miyamoto, 1999, p. 31), les katakana sont généralement utilisés pour (Taylor et Taylor, 1995, p. 310) :

- Les mots d'emprunts européens et chinois ;
- Les onomatopées ;
- Les noms de certains animaux et végétaux rares (par exemple : « kame » = tortue de terre) ;
- Les noms de substances chimiques peu fréquents ;
- Les représentations des prononciations japonaise (type « On ») et chinoise (type « Kun ») des kanji<sup>71</sup>.

---

<sup>71</sup> La plupart des kanji japonais sont en effet utilisés tant en Chine, dont ils sont originaires, qu'au Japon, mais chacun d'entre eux possède généralement au moins deux « prononciations » : l'une « chinoise », l'autre « japonaise ». Une telle situation semble parfois permettre aux touristes japonais se rendant en Chine de pouvoir décrypter, en partie au moins, certaines indications écrites « en chinois » (voire de pouvoir communiquer, de manière minimale, par écrit, comme nous l'a rapporté une de nos étudiantes), tout en étant incapables d'utiliser la langue orale du pays.

En effet, alors que le japonais présente de nombreux homophones dont la différence sémantique ressort efficacement de manière visuelle grâce aux kanji distincts qui les représentent (Kess et Miyamoto, 1999, p. 31), l'apprentissage des kanji est un apprentissage à vie pour les japonophones, même natifs, qui, tout au long de leur existence en tant que lecteurs, peuvent rencontrer des kanji jusque là inconnus, ou bien oubliés. C'est une des raisons (il y en a d'autres) pour lesquelles certains kanji sont parfois surmontés, dans les journaux notamment, mais pas uniquement, d'une transcription baptisée « furigana » (= kana d'annotation). Comme le notent Taylor et Taylor (1995, p. 311), les furigana servent à indiquer la prononciation de kanji dont la prononciation est inhabituelle ou non-familiale. En écriture horizontale, les furigana sont généralement placés au-dessus du kanji ; en écriture verticale, ils le sont à sa droite. Il existe par ailleurs d'autres utilisations occasionnelles des furigana, que mentionnent Taylor et Taylor (1995, pp. 312-313) à travers les exemples suivants :

- 1) « *A foreign word, [like] the French word "sentimentalisme", is annotated with Furigana to give its pronunciation* » ;
- 2) « *A case when the furigana annotate Kanji meaning "shachoo" ("company president") and "shain" ("company workers"), but their sounds given in Furigana are Japanese transcriptions of the English words "one man" and "gentlemen", which neither sound nor mean the same as the Kanji. Conceivably the pair "one man" and "company president" and the pair "gentlemen" and "company workers" have related meanings. Its author [...] makes a questionable claim that a reader views Furigana and its Kanji word simultaneously, thus obtaining extra meaning* » ;
- 3) « *A case when a kanji word is accompanied by Furigana that represent the sound of its English translation. Perhaps the concept in question, "transcript", is not familiar either as a kanji word or as an English loan word, but the two together provide securely the intended meaning* » ;
- 4) « *Cases when the meanings of words in phonetic scripts can be ambiguous or hard to recognize until they are annotated with kanji, i.e., with Furigani. [In our examples] Furigani annotates a foreign loan word in Katakana "peepaa" [...] and a Sino-Japanese word "izen" ("as before") in Hiragana* ».

Outre leur valeur descriptive, ces derniers exemples soulignent un trait assez original du système d'écriture japonais, celui du double codage écrit : l'exemple 1), en particulier, renvoie à une pratique relativement courante dans les manuels de FLE de production japonaise, qui est d'annoter en katakana les énoncés français, en vue de rendre leur prononciation plus accessible. Ce type d'annotation peut également fréquemment être observé en classe de français au Japon, placé par les étudiants sous la forme orthographique alphabétique des mots nouveaux qui leur sont présentés.

Enfin, ce qui semble essentiel pour nous, c'est le lien particulièrement fort entre unités phonologiques moraïques (qui sont traditionnellement les unités phonologiques intuitivement minimales des locuteurs japonais) et unités graphémiques kana. Ainsi, en réponse à la question que nous lui avons posée, un sujet japonais non-linguiste nous a indiqué que, de son point de vue, si l'on demandait à un Japonais combien de syllabes possédait tel énoncé en japonais, le sujet (non-spécialiste) effectuerait très probablement ce comptage en termes de caractères kana. Il semblerait en aller de même pour les kanji, puisque, face à des kanji, et à la même question, le sujet aurait de fortes chances d'effectuer en premier lieu une transcription de ce kanji en kana (*cf.* les furigana) et compter alors de nouveau le nombre de syllabes en termes de kana. Ainsi, lorsque nous avons présenté la suite de kanji signifiant « mariage » (en japonais "kek'kon", formé de deux kanji) à ce sujet et que nous lui avons demandé combien de syllabes possédait ce mot, le sujet était quelque peu confus, car en hiragana, la consonne géminée se traduirait par un « petit » « tsu » :

<b>Kanji</b>	<b>結</b>		<b>婚</b>	
<b>Hiragana</b>	<b>け</b>	<b>っ</b>	<b>こ</b>	<b>ん</b>
<b>Romaji</b>	« ke »	« petit » « tsu »	« ko »	« n »

C'est alors ici qu'intervient la notion de more (Otake, Hatano, Cutler et Mehler, 1993). En effet, comme l'indique Akamatsu (1997, p. 165) une unité moraïque (qu'il convient

de distinguer d'une more, c'est-à-dire une unité temporelle) correspond à un caractère kana (simple ou composé). Cette dépendance vis-à-vis des « syllabaires » kana, et les difficultés rencontrées lors du traitement des consonnes géminées (qui constitue l'un des cas où unité moraïque et syllabe ne se recoupent pas), par exemple, pourrait donc avoir des conséquences sur la lecture et la conscience phonologique en L2<sup>72</sup>.

### **II.4.3. Les romaji**

Il convient dans un premier temps de bien distinguer le système romaji (dont la fonction officielle est de transcrire les mots japonais en alphabet romain, et dont l'inventaire ne comporte que 22 lettres sur 26, <l>, <q>, <v> et <x> en étant absentes) de l'alphabet romain, qui est généralement appris dans le cadre de l'apprentissage scolaire de l'anglais. Au Japon, l'alphabet latin est appris à l'école durant les classes d'anglais, si bien que tout le monde (ou presque) sait épeler en alphabet latin (avec une prononciation cependant propre aux sujets japonais et fortement influencée par l'anglais, comme par exemple : <b> → [bi:]<sup>73</sup>). Une telle situation conduit bien souvent à certaines confusions entre les deux systèmes, romaji et alphabet latin, confusion alimentée par le développement actuel d'une certaine « pluriscripturalité »

---

<sup>72</sup> Otake, Hatano, Cutler et Mehler (1993, p. 261) notent néanmoins :

*« It should be noted that syllables may also be quite tangible constructs to Japanese language users. The kana orthographies are often thought of as syllabaries, and indeed, of all morae only N and Q cannot function as separate syllables. The Chinese characters of the kanji orthography, moreover, when given a Chinese reading, more often represent single syllables. Thus although we have argued above that orthographic influences may render the mora a salient unit to literate Japanese, it is also true that orthographic influences may render the syllable equally salient ».*

<sup>73</sup> Lorsque l'on demande à certains Japonais « d'épeler » les romajis, leur réaction peut ainsi être assez confuse, étant donné qu'il y a chevauchement entre, d'un côté une tendance à épeler en unités syllabiques l'alphabet latin appris en classe d'anglais, et de l'autre la tendance à « épeler » en unités moraïques le système romaji. Par exemple, pour le mot japonais « ringo » (= pomme), la réaction d'un sujet a été approximativement la suivante : "r" → "aaru" ; "i" → "ai" ; "n" → "enu" ; "g" → "jii" ; "o" → "oo". De la même manière, le mot japonais « itsu » (= quand) est « épelé » de manière différente en hiragana et en romaji chez ce locuteur particulier (on ne peut évidemment pas généraliser ici) : Hiragana: "i-tsu" / Romaji: "ai-tii-esu-yu". « Epeler » est donc une tâche relativement ambiguë en japonais, puisque, dans le système romaji, le mot « itsu » devrait être épelé « i-tsu », comme en kana. Il y a donc une certaine interférence entre l'apprentissage de l'alphabet latin utilisé en anglais et celui du système romaji, interférence qui apparaît clairement dans une telle tâche, où certains sujets confondent, en partie, les deux systèmes.

liée aux emprunts étrangers, en particulier dans le domaine de la communication de masse en milieu urbain. Une telle situation pourrait en outre laisser croire aux enseignants de français que les Japonais savent « lire » le français écrit, ce qui est bien loin de la réalité.

En ce qui concerne les romaji en tant que système de translitération, des discussions se poursuivent encore à propos du choix du meilleur système à adopter. Taylor et Taylor (1995, p. 316) indiquent ainsi :

*« Transcribing Japanese in Romaji should be straightforward because of the simple Japanese sound system. Yet there are over dozen different Romaji systems, which depend partly on which European language – e.g., Portuguese, Dutch, German, English – underlies a system, and partly on how the system deals with the several tricky Japanese sounds ».*

Le plus répandu est le système Hepburn, datant de 1885, qui correspond parfaitement aux codes phonographémiques de la langue anglaise, un facteur qui a pu influencer son extension. Mais deux autres systèmes, notamment, existent aussi, le plus récent étant le « Style Cabinet » (« Kunreishiki ») officiellement adopté en 1937, puis en 1954. Kess et Miyamoto (1999, p. 116) résument les principales différences entre ces trois systèmes dans le tableau que nous reproduisons ci-dessous :

<b>Hepburn</b>	<b>Kunreishiki</b>	<b>Nihonshiki</b>
shi	si	si
sha	sya	sya
shu	syu	syu
sho	syo	syo
ji	zi	zi
ja	zya	zya
ju	zyu	zyu
jo	zyo	zyo
chi	ti	ti
cha	tya	tya
chu	tyu	tyu
cho	tyo	tyo
ji	zi	di
ja	zya	dya
ju	zyu	dyu
jo	zyo	dyo
tsu	tu	tu
zu	zu	du
fu	hu	hu
o	o	wo

Nous pouvons tout de suite faire quelques remarques à propos des implications de telles différences sur l'apprentissage du français. Ainsi, en ce qui concerne l'opposition « shi/si » : l'utilisation du « si » graphique pour transcrire un [ɕi] qui, bien souvent pour une oreille française ressemble davantage à un [ʃi] qu'à un [si], risque ensuite de poser problème lors de la lecture à voix haute de la séquence « si » en français (qui est bien [si] ou [zi], et non [ɕi] ou [ʃi]). Si, d'un point de vue phonologique la séquence

/si/ est illégale en japonais, au moins dans la strate Yamato, le système Hepburn a au moins le mérite de « préserver » la séquence graphémique « si », tandis que les deux autres systèmes associent cette séquence à une séquence phonologique différente de ce qu'elle est en français par exemple. De plus, pour les sujets qui connaissent les trois systèmes de translittération, l'opposition graphique « shi/si » semble neutralisée : il ne s'agirait alors que d'allographes, contrairement au français dans lequel ces deux séquences ne sont en aucun cas homophones. Pour de tels sujets, on pourrait se demander dans quelle mesure cette connaissance du système romaji interfère dans la perception et la production (en tâche de répétition ou de lecture à voix haute) de mots comme « cinéma » ou comme « situation ». On pourrait faire des remarques similaires à propos des oppositions « chi/ti », « ji/zi/di », « tsu/tu » ou encore « fu/hu » entre ces trois systèmes. Le choix d'un système de translittération n'est donc pas sans importance pour l'enseignement/ apprentissage des langues étrangères, puisqu'il revient à assigner à une séquence phonologique une séquence graphique partagée, en partie au moins, par d'autres communautés linguistiques, et ce faisant, à associer, de manière indirecte, ces unités graphiques à certaines valeurs phonologiques en L1. Il n'est donc pas étonnant que, face à un texte en langue étrangère utilisant ces mêmes séquences graphiques, les automatismes de reconnaissance conduisent initialement à l'activation des valeurs phonologiques qui leur correspondent en L1. D'où l'importance du choix de ce système (puisque en français, par exemple, « tu » et « tsu » ne peuvent pas être assimilés, au risque de perturber fortement la réussite de la tâche à accomplir par les sujets).

Taylor et Taylor (1995, p. 318) donnent un exemple d'instructions officielles sur l'usage des romaji :

1. a) *For a long vowel, put a ^ or overbar over it, as in Rōmaji ;*
- b) *But for a long /i/ write ii ;*
- c) *In capital letters, a vowel letter can be doubled for a long vowel, as in ROOMAJI.*



2. *For the nasal final N, which may sound like -m, -n, -ng in different contexts, use one letter n, as in pinpon (« ping-pong »). (But m is often used, as in shimbun « newspaper »).*
3. *To separate the nasal N from the following vowel or y, use the apostrophe ', as in Man'yōshū.*
4. *For a doubled (geminate) consonant, double its letter, as in gakkō.*
5. *Capitalize the first letter of a sentence or of a proper noun.*
6. *For special sounds, use any Rōmaji.*

Si le système d'écriture japonais est donc traditionnellement multiple, cette « pluriscripturalité » s'est vue renforcée ces dernières décennies par l'utilisation massive des emprunts étrangers, et plus particulièrement par la coexistence de leur forme orthographique adaptée en katakana et de leur forme orthographique d'origine en alphabet latin.

#### **II.4.4. Transcription des emprunts, katakana et alphabet**

La quantité d'emprunts étrangers en japonais a augmenté de manière exponentielle ces dernières années, avec une convention officielle de transcription de ces mots en katakana promulguée en 1991 (Kess et Miyamoto, 1999, p. 97), enrichissant considérablement le lexique japonais. Kess et Miyamoto notent en outre que l'usage des katakana est également en évolution (1999, p. 30) :

*« The Katakana syllabary is [...] commonly declared as the appropriate transliteration medium for loan words coming into Japanese from other languages. However it sees a good deal of use in modern printed Japanese as a kind of visual italics ».*

Il est alors important de noter que la transcription en katakana des mots étrangers implique une adaptation phonologique, comme le soulignent Taylor et Taylor (1995, p. 314) :

« About 10% of the words in the Japanese vocabulary are borrowed from foreign languages, mostly European, and especially from English. In order to represent the sounds of foreign words, there are more Katakana signs than there are Hiragana signs. Despite the extra signs, the Japanese rendition system of foreign words is anything but faithful because [of] the Japanese sound system ».

Il faut également noter que certains mots en katakana peuvent s'écrire de différentes manières, par exemple (Kess et Miyamoto, 1999, p. 94) :

- バイオリン (« baiorin ») et ヴァイオリン (« vaiorin ») pour « violin » (= violon)<sup>74</sup>.
- クラスメイト (« kurasu meito ») et クラスメート (« kurasu meeto ») pour « class mate » (= camarade de classe).
- アイデア (« aideaa ») et アイディア (« aidiaa ») pour « idea » (= idée).

Si l'on examine le *Tuttle New Dictionary of Loanwords in Japanese* (Kamiya, 1994), on pourrait, dans une certaine mesure, adopter une démarche identique à celle que l'on trouve en phonologie des emprunts, dans un cadre OT ou TCRS par exemple, en essayant de découvrir les contraintes qui pèsent sur le système graphique japonais, et comment des formes graphiques étrangères, mal formées du point de vue du système japonais, sont ainsi modifiées (« réparées ») (même si ces modifications ne sont, en théorie, que le reflet de leurs modifications phonologiques).

L'existence même d'un dictionnaire de la sorte souligne l'importance, grandissante, de cette strate lexicale (*garaigo* en japonais) dans l'ensemble du lexique japonais. Environ 4000 mots sont ainsi présentés, avec des variations de forme mais aussi de sens et d'usage vis-à-vis des formes originales : par exemple le mot japonais *sâbisu*, issu de l'anglais *service*, signifie « gratuit » en japonais, tandis que le mot japonais *hansuto*

---

<sup>74</sup> La deuxième syllabe de notre propre prénom, « Sylvain », a ainsi été transcrite de manière variable selon les scripteurs japonais que nous avons rencontrés : soit バン (« -ban ») soit ヴァン (« -van »).

résulte de la contraction des deux unités composant le lexème anglais *hunger strike* (« grève de la fin »). Les origines de ces mots sont très variées (portugais, néerlandais, français, allemand, etc.), mais c'est l'anglais qui domine, en particulier depuis la seconde guerre mondiale.

Dans l'introduction au dictionnaire, l'auteur propose un inventaire des sons japonais, à l'aide de leur transcription en katakana et en alphabet, que nous reproduisons ci-dessous. En katakana, la combinaison de deux unités graphémiques se fait à l'aide d'une réduction de taille des caractères : par exemple pour « kiya », la première partie (« ki ») est de taille normale, alors que la seconde (« ya ») figurant à sa gauche est de taille réduite. Il est intéressant de noter l'écart séparant la présentation traditionnelle du système phonético-phonologique japonais (plus réduit) de celle-ci : l'auteur indique d'ailleurs par une astérisque certains sons « utilisés seulement pour les emprunts ». On pourrait donc considérer qu'il s'agit ici du système de la strate lexicale « étrangère » maximale.

A		I		U		E		O	
ア	A	イ	I	ウ	u	エ	e	オ	o
カ	ka	キ	ki	ク	ku	ケ	ke	コ	ko
クア	kwa*	クイ	kwi*			クエ	kwe*	クオ	kwo*
ガ	ga	ギ	Gi	グ	gu	ゲ	ge	ゴ	go
グア	gwa*								
サ	sa	シ	shi	ス	su	セ	se	ソ	so
ザ	za	ジ	Ji	ズ	zu	ゼ	ze	ゾ	zo
タ	ta	チ	chi	ツ	tsu	テ	te	ト	to
ツア	tsa*	テイ	ti*	トウ	tu*	ツイ	tse*	ツオ	tso*
		ツイ	tsi*	テユ	tyu*				
ダ	da	ダイ	di*	ドウ	du*	デ	de	ド	do
				デュ	dyu*				
ナ	na	ニ	ni	ヌ	nu	ネ	ne	ノ	no
ハ	ha	ヒ	hi	フ	fu	ヘ	he	ホ	ho
バ	ba	ビ	bi	ブ	bu	ベ	be	ボ	bo
ヴァ	va*	ヴィ	vi*	ヴ	vu*	ヴェ	ve*	ヴォ	vo*
パ	pa	ピ	pi	プ	pu	ペ	pe	ポ	po
ファ	fa*	フィ	fi*			フェ	fe*	フォ	fo*
マ	ma	ミ	mi	ム	mu	メ	me	モ	mo
ヤ	ya			ユ	yu	イエ	ye*	ヨ	yo
ラ	ra	リ	ri	ル	ru	レ	re	ロ	ro
ワ	wa	ウィ	wi*			ウェ	we*	ウォ	wo
ン	n								
キャ	kya			キュ	kyu			キョ	kyo
ギャ	gya			ギョ	gyu			ギョ	gyo
シャ	sha			シュ	shu	シェ	she*	ショ	sho
ジャ	ja			ジュ	ju	ジェ	je*	ジョ	jo

チャ	cha		チュ	chu	チ	che*	チョ	cho
ニャ	nya		ニユ	nyu			ニョ	nyo
ヒャ	hya		ヒユ	hyu			ヒョ	hyo
			フユ	fyu*				
ビャ	bya		ビユ	byu			ビョ	byo
			ヴユ	vyu*				
ピャ	pya		ピユ	pyu			ピョ	pyo
ミャ	mya		ミユ	myu			ミョ	myo
リャ	rya		リユ	ryu			リョ	ryo

Il faut aussi noter que, dans la partie « Ecrire les emprunts en katakana » du dictionnaire, les codes de transcription sont particulièrement souples, puisque chaque « son » mentionné peut être transcrit de deux ou trois manières différentes :

« Sons »	Transcriptions possibles
She / Je	シエ ou セ (She) / ジエ ou ゼ (Je)
Ti / Di	テイ, チ, ou テ (Ti) / ジ, デ, ou デイ (Di)
Fa / Fi / Fe / Fo	ハ, ファ, ou ファ (Fa) / ヒ, フイ, ou フイ (Fi) / ヘ, フェ, ou フェ (Fe) / ホ, フォ, ou フォ (Fo)
Dyu/Ju	ジュ ou テュ
Ye	イエ ou イエ
Wi / We / Wo	ウイ ou ウイ (Wi) / ウエ ou ウエ (We) / ウオ ou ウオ (Wo)
Kwa / Kwi / Kwe / Kwo	カ, クア, ou クア (Kwa) / キ, クイ, ou クイ (Kwi) / ケ, クエ, ou クエ (Kwe) / コ, クオ, ou クオ (Kwo)
Gwa	グア, グア, ou ガ
Tsi	チ ou ツイ
Tu / Du	ツ, トウ, ou ト (Tu) / ズ, ドウ, ou ド (Du)
Va / Vi / Vu / Ve / Vo	バ ou ヴァ (Va) / ビ ou ヴイ (Vi) / ブ ou ヴ (Vu) / ベ ou ヴエ (Ve) / ボ ou ヴォ (Vo)
Tyu	チュ ou テュ

Fyu	ヒユ ou フユ
Vyu	ビユ ou ヴユ
N	ン ou ム
Son associé à la lettre X	キサ (Kisa) ou クサ (Kusa), キシ (Kishi) ou クシ (Kushi), キス (Kisu) ou クス (Kusu), キソ (Kiso) ou クソ (Kuso)

Ce relevé d'emprunts se pose ainsi, d'une certaine manière, en complément du travail de Shinohara (1997), puisque celle-ci s'est précisément attachée à ne pas travailler sur les emprunts déjà établis, qui, en raison de la variabilité de leur médium d'origine, phonético-phonologique, sociologique ou orthographique, ne permettent pas d'établir les régularités recherchées et conduisent à traiter certaines données comme des « exceptions ». L'influence possible de l'orthographe sur la forme phonologique des emprunts est ainsi signalée par Shinohara (1997, p. 12) :

*« La deuxième voyelle /ee/ dans l'emprunt /messezi/ <message> (anglais) /mesɪdʒ/ « message » correspond à la voyelle /ɪ/ dans le mot d'origine. Nous pensons que cette forme a été introduite par un locuteur japonais qui a fait l'analogie avec l'adaptation de la voyelle /eɪ/ orthographiée également a dans d'autres mots comme age /eɪdʒ/ ».*

Pour notre part, étant donné l'intérêt que nous portons dans notre étude au traitement des groupes de consonnes par les sujets japonais, en particulier ceux de type OBLI, c'est-à-dire incluant une liquide (/l/ ou /r/), nous avons procédé à un relevé quasi-systématique des mots d'emprunts de ce dictionnaire possédant un tel groupe à l'initiale (dans le mot d'origine), que nous présentons dans le tableau ci-dessous. Nous avons limité le nombre d'exemples maximal à cinq. Dans ce tableau, le « L » tient lieu de liquide (/l/ ou /r/), puisque le dictionnaire ne dispose que d'une seule entrée, celle du « R » (et donc pas d'entrée « L »). Certains mots d'origine sont français, espagnols ou allemands, mais la plupart sont anglais. Nous n'avons inclus que les groupes /CC/ qui possédaient un représentant dans le dictionnaire.

<b>Groupe CC en position initiale</b>	<b>Voyelle suivante en Japonais</b>	<b>Mot d'origine</b>	<b>Forme adaptée</b>
PL	V=A	Platina	Purachina
	V=E	Play	Puree
	V=I	Prima	Purima
	V=O	Professional	Purofesshonaru
BL	V=A	Bravo	Buraboo
	V=E	Brakes	Bureeki
	V=I	Briefing	Buriifingu
	V=O	Brooch	Buroochi
	V=U	Blue	Buruu
FL	V=A	Fry/ Fly	Furai
	V=E	Flexibility	Furekishibiritii
	V=I	Free dial	Furii daiaru
	V=O	Floor	Furoa
	V=U	Flute	Furutoo
TL	V=A	Travail	Torabaiyu/Torabaayu
	V=E	Trade	Toreedo
	V=I	Trigger	Torigaa
	V=O	Trophy	Torofii
DL	V=A	Draft	Dorafuto
	V=E	Dressing	Doresshing
	V=I	Drip	Dorippu
	V=O	Drawn game	Doron geemu

KL	V=A V=E V=I V=O V=U	Club Credit Clear Cloak <b>!Croquette !</b> Crew	Kurabu Kurejitto Kuria Kurooku <b>!Korokke !</b> Kuruu
GL	V=A V=E V=I V=O V=U	Gravure Grey Green Globalization Groupie	Gurabia Guree Guriin Guroobarizeeshon Guruupii
SP	V=A V=E V=I V=O V=U	Spaghetti Spectacle Speech Sponge Spoon	Supagetti Supekutakuru Supiichi Suponji Supuun
ST	V=A V=E V=I V=O	Star Steady Steam Stockings	Sutaa Sutedii Sutiimu/suchiimu Sutokkingu
SK	V=A V=E V=I V=O V=U	Scarf Schedule Skim milk Score Scoop	Sukaafu Sukejuuru Sukimu miruku Sukoa Sukuupu
SL	V=A V=E V=I V=O	Slide Slender Slim Slow	Suraido Surendaa Surimu Suroo
SN	V=A	Snack	Sunakku



PT	V=I	Petit bourgeois	Puchiburu
SM	V=A	Smash	Sumasshu
	V=O	Smog	Sumoggu
	V=U	Smooth	Sumuusu/Sumuuzu

On ne peut alors manquer de remarquer que l'adaptation des mots étrangers présente des régularités qui semblent, comme le souligne Haunz (2002, p. 5), à présent constituer des conventions d'adaptation que connaissent, plus ou moins bien, tous les locuteurs japonais :

*« Japanese has a wealth of loanwords, mainly from English, in which consonant clusters are always adapted in the same way, that is by epenthesis of a vowel. [...] This means that any Japanese speaker is aware of this convention of adaptation, and this knowledge might play a dominant role in this automatic assumption of a vowel between two consonants by Japanese listeners ».*

Kamiya (1994), dans la partie introductive de son dictionnaire, propose ainsi plusieurs consignes pour lire et écrire les mots d'emprunts en katakana, dont ce que l'on peut appeler la « règle d'épenthèse ». Ces consignes constituent donc certaines conventions d'adaptation, de plus en plus connues des locuteurs japonais. On peut également noter que les ordinateurs locaux sont équipés de logiciels permettant la transcription instantanée de mots écrits en romaji en caractères katakana. Il semble donc y avoir une habitude de conversion des séquences alphabétiques illicites en formes phonologiques licites *via* le script katakana, comme par exemples les noms « PARCO » (coda syllabique simple), « MacDonald » (codas syllabiques simple et complexe, ainsi que segment /l/) ou « General Nezu » (coda syllabique simple, ainsi que segments /r/ et /l/) (il s'agit de noms de magasins ou de résidence), dont la forme orthographique, puisqu'elle est *effectivement* utilisée, semble donc licite, dans une certaine mesure, en japonais contemporain. La situation phonographémique du Japon urbain actuel semble

donc assez complexe : les romaji servent à transcrire les mots japonais, mais l'alphabet latin est également disponible, notamment pour transcrire les mots étrangers.

Ce qui est plus frappant encore, c'est la coexistence des formes orthographiques d'origine<sup>75</sup> en alphabet latin et des formes adaptées en katakana. Ainsi, dans le métro de Tokyo, on peut voir deux inscriptions placées l'une sous l'autre et renvoyant au même référent (une carte de métro) : « Metro Card » / « メトロカード » (= ME - TO - RO - KA - A - DO). Un autre exemple similaire est celui de cette marque de viennoiserie : « Le chocolat » / « ル・ ショ コラ » (= RU - SHI-yo - KO - RA).

On pourrait donc dire que, dans une certaine mesure, les Japonais ont l'habitude de lire les caractères alphabétiques latins selon la structure moraique du japonais, en fonction d'un code socialement normé de conversion, tant au niveau syllabique que segmental. Il semblerait donc que se développe actuellement au Japon un système « pluriscriptural » mais toujours « monophonologique »<sup>76</sup>, dans lequel les formes orthographiques étrangères sont intégrées, mais pas leurs formes phonologiques. Une telle situation, si elle peut sembler positive pour l'apprentissage des langues étrangères en raison de la familiarisation accrue des Japonais avec le lexique orthographique étranger, s'avère en réalité tout à fait néfaste pour l'intégration de la phonologie étrangère, puisque les formes orthographiques alphabétiques sont systématiquement, et préalablement à l'apprentissage, ramenées à des formes phonologiques conformes à la phonologie du japonais. Dans ce cas, l'écrit, s'il n'est pas explicitement traité

---

<sup>75</sup> Sauf pour ce que l'on nomme parfois de manière humoristique le « Japanese Engrish », défini comme l'ensemble des erreurs humoristiques commises par les Japonais, en anglais, dans le domaine commercial. En particulier, on parle de « Japanese Engrish » lorsque l'on trouve des mots anglais, utilisés dans la vie quotidienne, dont les formes orthographiques contiennent certaines substitutions graphémiques, en particulier entre les graphèmes <r> et <l>, des substitutions *attestées* qui font sourire les locuteurs anglophones. A titre d'exemple on peut mentionner : « Eric Crapton » (« crap » = « nul ») au lieu de « Eric Clapton », « Erection » au lieu de « Election » et le classique « Lice » (= poux) au lieu de « Rice » (= riz).

<sup>76</sup> Même si, comme cela apparaît dans le dictionnaire des emprunts de Kanya (1994), de nouveaux « sons » (en particulier le [v]) et surtout de nouvelles combinaisons phonotactiques s'introduisent très progressivement dans le répertoire « phonique » japonais, témoignant une nouvelle fois de l'influence que peut avoir, dans notre cas *via* les emprunts, l'orthographe sur la langue orale, produisant ainsi ce que l'on pourrait nommer un « effet Buben » en japonais.

didactiquement, peut constituer un obstacle à l'apprentissage du système phonologique de la langue étrangère, en raison de ces codes de conversion automatique.

Cette présentation est évidemment quelque peu simplificatrice et sujette à discussion (Kess et Miyamoto, 1999, p. 110). Mais ce qui importe ici pour nous c'est :

- 1) Le caractère quasi-univoque des relations entre unités moraiques phonologiques et unités graphémiques kana. Par exemple, en caractères kana, le mot « kimono » serait transcrit à l'aide de trois signes : きもの (= ki-mo-no). Le lien entre unités phonologiques de segmentation et caractères kana (de type CV) est donc particulièrement fort, à travers une orthographe que l'on pourrait facilement qualifier de « superficielle ».
- 2) L'utilisation quasi-systématique des caractères katakana pour transcrire les mots étrangers (qui sont alors adaptés pour se conformer à la structure phonologique et graphémique du japonais, très souvent de type CV, comme par exemple le mot « métro » transcrit avec trois caractères kana : メトロ (= me-to-ro)).
- 3) La pluralité des types d'écriture utilisés en synchronie vis-à-vis d'un système phonologique unique.

Nous pouvons à présent comparer et contraster les systèmes phonologiques et phonographémiques français et japonais sur les points que nous avons sélectionnés.

### **III. Perspective contrastive**

D'un point de vue linguistique, la comparaison des deux systèmes phonologiques en présence, français et japonais, vise essentiellement à en identifier les similarités et les divergences. Plusieurs questions se posent néanmoins lorsque l'on entreprend une telle démarche, la première étant : que compare-t-on exactement ? Si l'on admet comparer, non pas tant les systèmes linguistiques eux-mêmes, que les systèmes conceptuels et sémiotiques de description et/ou de modélisation desdits systèmes, on doit alors s'assurer de la comparabilité de ces méta-systèmes, tout en affinant les critères et/ou les modes de comparaison (par exemple dans le cas de l'opposition syllabe/more). A

quel niveau d'analyse se situe-t-on (en particulier dans le cadre des phonologies non-linéaires) ? Quels sont les modèles choisis (théoriques et descriptifs) ? Quelles sont les frontières des domaines de comparaison (par exemple syllabe vs mot phonologique) ? Peut-on contraster les systèmes phonologiques sans se soucier du domaine morphologique ? Peut-on se limiter aux domaines syllabique et segmental en négligeant les niveaux prosodiques supérieurs ?

Du cadre structuraliste au cadre OT, via le fonctionnalisme et le générativisme, la démarche contrastive a, elle aussi, dû suivre l'évolution des modèles, distinguant principes universels et paramètres spécifiques, représentations et relations, règles et contraintes. Dans le cadre de notre étude, nous nous limiterons ici, dans un premier temps, à une mise en contraste sélective, essentiellement phonémique, de certaines parties des systèmes segmentaux et syllabiques français et japonais, avant de modifier progressivement par la suite notre cadre théorique vers celui des modèles à contraintes pour traiter ultérieurement des questions d'apprentissage phonologique.

### **III.1. Contrastes phonético-phonologiques**

Nous examinons ici, en effectuant une certaine sélection des unités et des phénomènes, deux niveaux d'analyse descriptive : le niveau segmental (phonémique) et le niveau syllabique.

#### **III.1.1. Aspects segmentaux**

En ce qui concerne le niveau segmental, il faut noter qu'une comparaison strictement phonétique serait beaucoup plus fine et nous conduirait à d'autres développements, en particulier au vu de la grande variété allophonique que l'on rencontre dans le système japonais. Cependant, il est parfois délicat de séparer de manière tranchée ce qui relève de la description phonétique et ce qui relève de la description phonologique, d'autant que l'une sert généralement de fondement, conceptuel ou terminologique, à l'autre, en particulier dans un cadre phonémique, structuraliste ou fonctionnaliste, mais également

dans les approches en traits. La mise en contraste que nous proposons ci-dessous<sup>77</sup>, et qui n'inclut par ailleurs aucune indication phonotactique, en partie parce que nous traitons les aspects syllabiques ultérieurement, reflète donc cette étroite imbrication.

---

<sup>77</sup> Les caractéristiques contrastives entre les deux langues sont signalées par des caractères "gras" dans les descriptions données dans le tableau.

Segments Consonantiques Français	Segments Consonantiques Japonais
<p style="text-align: center;"><b><u>L</u></b> (API=[l])</p> <p><i>Phonétique</i> : consonne constrictive apico-alvéolaire (latérale/sonnante/liquide) (l'air s'échappe par les côtés).</p> <p><i>Phonologique</i> : [+sonorante] [+coronale].</p> <p><i>Statut phonémique</i> : <b>phonème</b>.</p>	<p style="text-align: center;"><b><u>L</u></b> (API=[l])</p> <p><i>Phonétique</i> : consonne constrictive apico-alvéolaire latérale (latérale/sonnante/liquide) (l'air s'échappe par les côtés)<sup>78</sup>.</p> <p><i>Phonologique</i> : [+sonorante] [+coronale].</p> <p><i>Statut phonémique</i> : <b>allophone</b> du /r/.</p>
<p style="text-align: center;"><b><u>R</u></b> (API=[ʀ])</p> <p><i>Phonétique</i> : consonne <b>constrictive</b> (vibrante/sonnante/liquide) <b>dorso-uvulaire</b>. (/R/ est réalisé par contact ou rapprochement de la luette contre la partie postérieure du dos de la langue, avec un léger battement [ʀ] ou sans battement [ʁ]).</p> <p>Articulation <b>maintenable</b>.</p> <p><i>Phonologique</i> : [+sonorante] [+<b>continuante</b>] [-<b>coronale</b>].</p> <p><i>Statut phonémique</i> : phonème.</p>	<p style="text-align: center;"><b><u>r</u></b> (API=[ɾ])</p> <p><i>Phonétique</i> : consonne de type <b>flap apico-alvéolaire</b> (liquide) noté [ɾ] (réalisation prototypique).</p> <p>Articulation <b>non-maintenable</b>.</p> <p><i>Phonologique</i> : [+sonorante] [-<b>continuante</b>]<sup>79</sup> [+<b>coronale</b>] ; consonne non-marquée du système japonais, sous-spécifiée, comme en atteste la variété de ses réalisations entre autres indices<sup>80</sup>.</p> <p><i>Statut phonémique</i> : phonème.</p>

<sup>78</sup> Akamatsu (2001) indique cependant que ce [l], qui apparaît généralement après la réalisation de /N/ (par exemple dans « aNrakushi » (= euthanasie)), serait différent du [l] français.

<sup>79</sup> Pour le caractère [+/-cont] des flaps, voir Durand, 1990, p. 52.

<sup>80</sup> Labrune, 2003, pp. 11-12.

	<p style="text-align: center;"><b><u>H</u></b> (API=[h])</p> <p><i>Phonétique</i> : consonne constrictive <b>glottale</b> non voisée.</p> <p><i>Phonologique</i> : [+obstruante] [+continuante] [-sibilante] [-voisée].</p> <p><i>Statut phonémique</i> : phonème<sup>81</sup>.</p>
<p style="text-align: center;"><b><u>B</u></b> (API=[b])</p> <p><i>Phonétique</i> : consonne occlusive bilabiale orale voisée.</p> <p><i>Phonologique</i> : [+obstruante] [+plosive] [-coronale] [+antérieure] [+voisée].</p> <p><i>Statut phonémique</i> : phonème.</p>	<p style="text-align: center;"><b><u>B</u></b> (API=[b])</p> <p><i>Phonétique</i> : consonne occlusive bilabiale orale voisée.</p> <p><i>Phonologique</i> : [+obstruante] [+plosive] [-coronale] [+antérieure] [+voisée].</p> <p><i>Statut phonémique</i> : phonème.</p>
<p style="text-align: center;"><b><u>V</u></b> (API=[v])</p> <p><i>Phonétique</i> : consonne constrictive <b>labiodentale</b> voisée.</p> <p><i>Phonologique</i> : [+obstruante] [-sibilante] [+voisée].</p> <p><i>Statut phonémique</i> : <b>phonème</b>.</p>	<p style="text-align: center;"><b><u>β</u></b></p> <p><i>Phonétique</i> : consonne constrictive <b>bilabiale</b> voisée.</p> <p><i>Phonologique</i> : [+obstruante] [-sibilante] [+voisée].</p> <p><i>Statut phonémique</i> : <b>allophone</b> du /B/<sup>82</sup>.</p>

<sup>81</sup> Akamatsu (2001, p. 97) note qu'un allophone fréquent du /h/ est la fricative dorso-uvulaire non voisée [χ], qui, nous le notons, est elle-même un allophone occasionnel du /R/ français.

<sup>82</sup> Essentiellement en contexte intervocalique (Akamatsu, 2001, p. 75).

<p style="text-align: center;"><b><u>F</u></b> (API=[f])</p> <p><i>Phonétique</i> : consonne constrictive <b>labiodentale</b> non voisée.</p> <p><i>Phonologique</i> : [+obstruante] [-sibilante] [-voisée].</p> <p><i>Statut phonémique</i> : <b>phonème.</b></p>	<p style="text-align: center;"><b><u>ϕ</u></b></p> <p><i>Phonétique</i> : consonne constrictive <b>bilabiale</b> non voisée.</p> <p><i>Phonologique</i> : [+obstruante] [-sibilante] [-voisée].</p> <p><i>Statut phonémique</i> : <b>allophone</b> du /H/.</p>
<p style="text-align: center;"><b><u>S</u></b> (API=[s])</p> <p><i>Phonétique</i> : consonne constrictive <b>prédorso-alvéolaire</b> (ou lamino-alvéolaire) non voisée.</p> <p><i>Phonologique</i> : [+obstruante] [+coronale] [+sibilante] [+antérieure] [-voisée].</p> <p><i>Statut phonémique</i> : phonème.</p>	<p style="text-align: center;"><b><u>ʃ</u></b> (API=[s])</p> <p><i>Phonétique</i> : consonne constrictive <b>prédorso-alvéolaire</b> (ou lamino-alvéolaire) non voisée.</p> <p><i>Phonologique</i> : [+obstruante] [+coronale] [+sibilante] [+antérieure] [-voisée].</p> <p><i>Statut phonémique</i> : phonème.</p>
<p style="text-align: center;"><b><u>ʃ</u></b></p> <p><i>Phonétique</i> : consonne constrictive <b>prédorso-prépalatale</b> (ou apico-palatoalvéolaire) non voisée.</p> <p><i>Phonologique</i> : [+obstruante] [-coronale] [+sibilante] [-antérieure] [-voisée].</p> <p><i>Statut phonémique</i> : <b>phonème.</b></p>	<p style="text-align: center;"><b><u>ʃ</u></b></p> <p><i>Phonétique</i> : consonne constrictive <b>prédorso-dorsopalatale</b> (ou laminodorsopalatale) non voisée.</p> <p><i>Phonologique</i> : [+obstruante] [-coronale] [+sibilante] [+antérieure] [-voisée].</p> <p><i>Statut phonémique</i> : <b>allophone</b> du /S/.</p>



	<p style="text-align: center;">ç</p> <p><u>Phonétique</u> : consonne constrictive <b>dorso-palatale</b> non voisée.</p> <p><u>Phonologique</u> : [+obstruante] [-coronale] [+sibilante] [-antérieure] [-voisée].</p> <p><u>Statut phonémique</u> : <b>allophone</b> du /H/.</p>
--	---

Segments Vocaliques Français	Segment Vocalique Japonais
<p style="text-align: center;"><u><b>OU</b></u> (API=[u])</p> <p><i>Phonétique</i> : voyelle orale postérieure <b>arrondi</b> très fermée.</p> <p><i>Phonologique</i> : [-antérieure] [+haute] [+<b>arrondi</b>].</p> <p><i>Statut phonémique</i> : phonème.</p>	<p style="text-align: center;"><u><b>OU</b></u> (API=[ɯ])</p> <p><i>Phonétique</i> : voyelle orale <b>postérieure non-arrondi</b> (non-labialisée) très fermée<sup>83</sup>.</p> <p><i>Phonologique</i> : [-antérieure] [+haute] [-<b>arrondi</b>].</p> <p><i>Statut phonémique</i> : phonème.</p>
<p style="text-align: center;"><u><b>Y</b></u> (API=[y])</p> <p><i>Phonétique</i> : voyelle orale <b>antérieure arrondi</b> très fermée.</p> <p><i>Phonologique</i> : [+antérieure] [+haute] [+<b>arrondi</b>].</p> <p><i>Statut phonémique</i> : phonème.</p>	
<p style="text-align: center;"><u><b>EU</b></u> (API= [ø])</p> <p><i>Phonétique</i> : voyelle orale <b>antérieure arrondi fermée</b>.</p> <p><i>Phonologique</i> : [+antérieure] [-haute] [-basse] [+<b>arrondi</b>].</p> <p><i>Statut phonémique</i> : phonème.</p>	

<sup>83</sup> Rappelons que, d'après Akamatsu (2001, p. 31), le [ɯ] après [s, ts, (d)z] tend à se centraliser [ɰ], et ainsi à devenir voyelle neutre, particulièrement sujette au dévoisement, ce qui peut notamment conduire à une confusion avec la voyelle neutre centrale française « schwa » [ə].

A propos de l'espace vocalique phonétique, Keating et Huffman (1984, p. 193) notent que les systèmes vocaliques de différentes langues peuvent être distingués de plusieurs manières :

- Différences segmentales *quantitatives* (nombres différents de segments et peu de variation allophonique par segment).
- Différences segmentales *qualitatives* (nombres de segments quasi-équivalents et grande variation allophonique pour chaque segment).
- Différences segmentales *phonotactiques* (distributions différentes des segments et peu de variation allophonique par segment).
- Combinaison de ces trois types de différences.

Dans un tableau comme celui-ci, l'absence d'indications phonotactiques détaillées (autres que celles sous-entendues par certains allophones) pourrait être jugée criante (par exemple l'impossibilité pour le /s/ japonais dans la strate Yamato de se trouver avant /i/, contrairement au français), mais cette absence s'explique par sa visée représentationnelle (en termes d'existence / de non-existence et de statut phonémique) : les problèmes de contraintes phonotactique sont donc mis à l'écart pour le moment. De la même manière, tous les phénomènes liés à la coarticulation, comme par exemple le dévoisement, et ainsi les allophones qui en résultent, appartiennent soit au domaine strictement phonétique, soit au domaine phonotactique, et n'apparaissent donc pas ici.

### **III.1.2. Aspects syllabiques**

Sans négliger la discussion sur la dichotomie more/syllabe que nous avons présentée précédemment, on peut établir la comparaison suivante :

Structure syllabique du français	Structure syllabique du japonais
(C)(C)(C)V(C)(C)(C)	<p><u>Syllabes légères</u> : (C) (j) V</p> <p><u>Syllabes lourdes</u> : (C) (j) V V  (C) (j) V N  (C) (j) V Q</p> <p style="text-align: right;">OU<sup>84</sup></p> <p><u>Syllabes pleines</u> : C(j)V</p> <p><u>Syllabes déficientes</u> : C• ou •V</p>

Ainsi, on peut comparer les deux langues de la manière suivante (les glides sont ici exclues de la catégorie consonantique) :

	V	V :	CV	CVC	VC	CCV	CCVC	CVCC	VCC	CCVCC
<b>Français</b>	Oui	Non	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
<b>Japonais</b>	Oui	Oui	Oui	Limité	Limité	Non	Non	Non	Non	Non

Ou encore en termes de variation paramétrique dans le type de syllabe, selon les paramètres proposés par Blevins pour plusieurs langues (1995, p. 219) :

	Noyau complexe	Attaque obligatoire	Attaque complexe	Coda	Coda complexe
<b>Français</b>	Non	Non	Oui	Oui	Oui
<b>Japonais</b>	Oui	Non	Non	Limité	Non

Comme elle le note (1995, p. 218): « In second language acquisition, speakers have little difficulty in shifting from a « yes » to a « no » value for a given parameter, but do show difficulty in switching from a « no » value to a « yes » ».

---

<sup>84</sup> Labrune, 2001b.

On pourrait enfin comparer et contraster les classements, spécifiques, des contraintes, universelles, pertinentes dans notre étude, pour le français, le japonais, et, à terme, pour l'interlangue des sujets japonais apprenant le français, (suivant par exemple la démarche de Grijzenhout et Van Rooy (2001, p. 16) pour des sujets Zoulous apprenant l'anglais), mais cela serait prématuré, puisque ce sont précisément, en partie du moins, ces classements, ou certains aspects de ces classements, que notre étude vise à mettre à jour.

### **III.2. Graphie**

Nous ne développerons pas ici la complexité de chaque système phonographémique : nous nous contentons d'effectuer une mise en contraste des codes *graphiques* correspondant aux points phonologiques qui nous préoccupent. D'un point de vue segmental, seuls les romaji japonais et l'alphabet latin nous permettent d'effectuer une comparaison, à un niveau strictement graphique : il s'agit en effet d'examiner dans un premier temps, simplement en termes d'existence / de non-existence, les systèmes graphiques des deux langues, de manière à savoir quels sont les graphes qui y sont disponibles. Nous ne traiterons ici que du répertoire romaji « traditionnel », c'est-à-dire n'incluant pas les « nouvelles consonnes » par exemple. Nous laisserons pour l'instant de côté également les correspondances phonographémiques, plus complexes.

<b>Français</b>	<b>Japonais</b>
B/b	B/b
<b>V/v</b>	<b>B/b</b>
F/f	F/f
R/r	R/r
<b>L/l</b>	<b>R/r</b>
<b>OU/ou</b>	<b>U/u</b>
<b>EU/eu</b>	<b>U/u</b>
U/u	U/u

On voit ici qu'à chaque graphe français correspond, de manière non systématiquement isomorphe, un graphe japonais, sauf pour les digraphes français « OU » et « EU » qui correspondent à un monographe en japonais. Sont théoriquement absents du répertoire romaji « traditionnel » les graphes « V », « L », « OU » et « EU ».

Ainsi, après avoir présenté les deux systèmes phonético-phonologiques et phonographémiques en présence, et avant de formuler certaines problématiques d'apprentissage linguistique issues de la mise en contraste de ces deux systèmes, il nous faut à présent caractériser la manière dont ces derniers sont sollicités, en particulier en perception, au cours de l'apprentissage. Cela nous conduit à la partie psycholinguistique de notre étude.



## **Partie Psycholinguistique :** **Traitements et Ressources** **chez l'Apprenant de Langue Etrangère**

Rappelons que la population dont nous nous occupons ici correspond à une population d'étudiants universitaires, d'adultes en formation continue, voire de lycéens. Il s'agit d'une population linguistiquement assez homogène, plutôt monolingue, scolarisée, de niveau débutant ou intermédiaire en français langue étrangère, et se trouvant en situation d'apprentissage institutionnel en milieu hétéroglotte, c'est-à-dire un milieu linguistique autre que celui de la langue à apprendre (Cuq, 2003, p. 121). C'est une population dont le développement cognitif, social et linguistique en L1 est globalement censé avoir atteint un certain palier de stabilisation (par opposition à une population enfantine, dite « précoce » dans le domaine de la didactique du FLE).

Si l'on souhaite connaître la manière dont ces apprenants perçoivent, mémorisent, et ainsi apprennent le système phonético-phonologique du français, et plus généralement leur mode de traitement de l'oral en français, en vue d'adopter des démarches didactiques qui y soient adaptées, il faut nous plonger dans le domaine psycholinguistique, tant vis-à-vis des « universaux » cognitifs, valables pour tous les apprenants, quelle que soit leur origine linguistique et culturelle, que des spécificités pouvant caractériser les apprenants japonais de français qui nous préoccupent ici. Trois grands thèmes peuvent alors être abordés, eu égard à notre problématique globale : (I) la perception et l'apprentissage de la parole en langue étrangère en général ; (II) le rapport entre perception de la parole et multimodalité ; (III) le rapport entre perception de la parole et orthographe.

Le premier thème (I) se décompose en deux sous-parties :

- L'une a trait au système perceptivo-attentionnel (I.1.), et concerne d'une part la perception de la parole (I.1.1.), avec les processus bas-haut et haut-bas (I.1.1.1.), les



théories de la perception de la parole (I.1.1.2.) et le rapport entre perception et compréhension (I.1.1.3.), et d'autre part la reconnaissance de la parole (I.1.2.), avec les unités perceptives impliquées (I.1.2.1.), les processus de catégorisation et d'appariement (I.1.2.2.) et les modèles de reconnaissance de la parole (I.1.2.3.) ;

- L'autre a trait à la mémoire et aux connaissances du sujet (I.2.), et concerne d'une part le système mnésique (I.2.1.), avec son organisation générale et sa modélisation (I.2.1.1.), ainsi que celles, plus spécifiques, de la mémoire de travail et de la boucle phonologique (I.2.1.2.), et d'autre part les connaissances linguistiques du sujet (I.2.2.), avec la distinction entre savoirs déclaratifs et savoirs procéduraux (I.2.2.1.) et l'organisation du lexique mental (I.2.2.2.).

Le deuxième thème (II) vise à établir ce que l'on sait des influences intermodales dans la perception de la parole (II.1.), en se référant à certaines théories et modèles en vigueur (II.2.), notamment le « modèle flou » de Massaro (II.2.1.) et la « théorie de la perception pour le contrôle de l'action » de Schwartz, Abry et leurs collègues (II.2.2.), avant d'aborder plus spécifiquement le rôle de la multimodalité dans la perception de la parole en situation d'apprentissage d'une langue étrangère (II.3.). Ce rôle est traité ici sous deux angles : celui des indices visuels dans la perception auditive en L2 (II.3.1.), et celle de la bimodalité audiovisuelle dans l'apprentissage de la parole en L2 (II.3.2.), que les indices visuels soient oro-faciaux pour l'apprentissage de catégories phonémiques (II.3.2.1.), ou qu'ils soient orthographiques pour la mémorisation (II.3.2.2.). La question de la bimodalité audio-orthographique, qui est celle qui nous occupe principalement dans notre étude, permet d'aborder le troisième thème de cette partie, celui du rapport entre la perception de la parole et celle de l'écrit.

Ce troisième thème (III) se divise en cinq parties :

- La première et la deuxième visent à établir quelques généralités sur le rapport entre perception de la parole et de l'écrit (III.1. et III.2.), que ce soit dans leurs différences, en particulier la primauté de l'oral sur l'écrit (III.2.1.), ou dans leurs

similitudes, en termes de traitement de l'information bas-haut (III.2.2.) ou haut-bas (III.2.3.).

- La troisième permet ensuite de préciser les spécificités de la perception de l'écrit et sa variation en fonction des langues (III.3) : après avoir examiné certains concepts généraux relatifs à ce type de perception, liés à l'hypothèse de la profondeur orthographique (III.3.1.) et à la taille des unités perceptives en lecture (III.3.2.), nous traitons le cas du français (III.3.3.), *via* les notions de « graphème » (III.3.3.1.), de « syllabe » (III.3.3.2.) et de « fréquence » (III.3.3.3.), et celui du japonais (III.3.4.), dans sa pluralité scripturale particulière (III.3.4.1.), *via* les caractères « kana » (III.3.4.2.), « romaji » (III.3.4.3.) et « kanji » (III.3.4.4.).
- Dans la quatrième partie, nous nous référons aux modèles de reconnaissance lexicale et de lecture à voix haute (III.4.), de type « cohorte » (III.4.1.), à « deux voies » (III.4.2.) ou interactif (III.4.3.), afin de finaliser le panorama que nous donnons de l'activité de lecture, considérée sous l'angle de la reconnaissance lexicale et sublexicale (III.4.4.).
- La cinquième partie reprend alors notre problématique générale, à savoir celle de l'influence possible des représentations orthographiques sur la perception et le traitement de la parole (III.5.). Deux grands types d'influence sont envisagés : d'abord une interaction entre codes phonologiques et orthographiques (III.5.1.), malgré certaines données neurolinguistiques (III.5.1.1.), tant au niveau sublexical (III.5.1.2.) que lexical (III.5.1.3.); puis une influence de l'écrit sur le développement de la conscience phonologique (III.5.2.), et plus précisément de la conscience phonémique (III.5.2.1.). L'interaction entre l'apprentissage du code écrit et le développement de la conscience phonologique rejoint alors notre problématique d'ensemble (III.5.2.2.).

Au terme de cet examen de l'appareil cognitif de l'apprenant de langue étrangère, axé sur la perception de la parole, auditive ou audiovisuelle, sur la perception de l'écrit et sur leurs interactions éventuelles, il nous est alors possible de rassembler les données présentées dans la partie linguistique de notre étude avec celles de la partie

psycholinguistique, pour tenter de formuler plus précisément nos problématiques relatives à l'apprentissage multimodal de la parole en langue étrangère, et plus spécifiquement celui du système phonético-phonologique du français, en bimodalité audio-orthographique, par des apprenants adultes japonais.

## Table des Matières Partielle

I. PERCEPTION ET APPRENTISSAGE DE LA PAROLE EN LANGUE ETRANGERE : DESCRIPTION GENERALE.....	- 282 -
.....	
<i>I.1. Système perceptivo-attentionnel et perception de la parole</i> .....	- 283 -
I.1.1. Perception de la parole: considérations générales.....	- 286 -
I.1.1.1. Processus bas-haut et haut-bas.....	- 287 -
I.1.1.2. Théories de la perception de la parole.....	- 292 -
I.1.1.3. De la perception à la compréhension.....	- 296 -
I.1.2. Perception de la parole et reconnaissance linguistique.....	- 300 -
I.1.2.1. Unités perceptives : niveau sublexical.....	- 301 -
I.1.2.2. Reconnaissance linguistique.....	- 310 -
I.1.2.2.1. Appariement, catégorisation et reconnaissance : considérations générales.....	- 310 -
I.1.2.2.2. Modèles de reconnaissance.....	- 318 -
I.1.2.2.2.1. Le modèle COHORTE.....	- 319 -
I.1.2.2.2.2. Le modèle TRACE.....	- 320 -
<i>I.2. Système mnésique et connaissances</i> .....	- 322 -
I.2.1. Système mnésique.....	- 323 -
I.2.1.1. Organisation de la mémoire : perspectives.....	- 324 -
I.2.1.1.1. Concepts généraux.....	- 324 -
I.2.1.1.2. Modélisation.....	- 327 -
I.2.1.1.2.1. Modèle Modal.....	- 327 -
I.2.1.1.2.2. Les Niveaux de Traitement.....	- 328 -
I.2.1.1.2.3. Composantes de la mémoire.....	- 330 -
I.2.1.1.2.4. Mémoire et apprentissage.....	- 332 -
I.2.1.1.2.5. Perspectives connexionnistes.....	- 333 -
I.2.1.1.3. Format et Temporalité.....	- 336 -
I.2.1.1.3.1. Format et Modalité.....	- 336 -
I.2.1.1.3.2. Temporalité.....	- 338 -
I.2.1.1.3.3. Parenthèse didactique.....	- 339 -
I.2.1.2. Mémoire de travail et boucle phonologique.....	- 341 -
I.2.1.2.1. Nature de la boucle phonologique.....	- 344 -
I.2.1.2.2. Fonction de la boucle phonologique.....	- 348 -
I.2.1.2.2.1. Acquisition de formes phonologiques nouvelles.....	- 348 -
I.2.1.2.2.2. Lecture.....	- 353 -
I.2.1.2.3. Parenthèse didactique.....	- 355 -
I.2.2. Connaissances linguistiques.....	- 357 -
I.2.2.1. Savoirs déclaratifs et savoir-faire procéduraux.....	- 358 -
I.2.2.2. Le lexique mental : primitives et émergentes.....	- 363 -
I.2.2.2.1. Primitives phonologiques : perspective symbolique.....	- 364 -
I.2.2.2.1.1. Production.....	- 365 -
I.2.2.2.1.2. Perception.....	- 370 -
I.2.2.2.2. Emergentisme et connexionnisme : perspective subsymbolique.....	- 371 -
I.2.2.2.3. Parenthèse didactique.....	- 375 -
II. PERCEPTION DE LA PAROLE ET MULTIMODALITE.....	- 380 -
<i>II.1. Influences intermodales</i> .....	- 382 -
<i>II.2. Théories et Modèles</i> .....	- 388 -
II.2.1. Le modèle flou (Massaro).....	- 388 -
II.2.2. La théorie de la perception pour le contrôle de l'action (Schwartz et al.).....	- 396 -

II. 3. Perception, multimodalité et apprentissage de la parole en langue étrangère.....	400 -
II.3.1. Rôle des indices visuels dans la perception auditive en langue étrangère.....	400 -
II.3.1.1. Influence positive des indices visuels .....	400 -
II.3.1.2. Influence moindre des indices visuels.....	402 -
II.3.2. Rôle de la bimodalité dans l'apprentissage de la parole en langue étrangère .....	404 -
II.3.2.1. Indices visuels oro-faciaux et catégories phonémiques .....	404 -
II.3.2.1.1. Indices visuels externes .....	405 -
II.3.2.1.2. Indices visuels internes.....	407 -
II.3.2.2. Indices visuels orthographiques et mémorisation.....	409 -
II.3.3. Conclusion.....	410 -
III. PERCEPTION DE LA PAROLE ET ORTHOGRAPHE .....	412 -
III.1. Introduction .....	413 -
III.2. Perception de l'oral et de l'écrit : principes invariants.....	414 -
III.2.1. Primauté de l'oral sur l'écrit .....	414 -
III.2.2. Traitement de l'information .....	416 -
III.2.3. Contexte et effets haut-bas .....	420 -
III.3. Perception de l'écrit : spécificités et variation selon les langues .....	421 -
III.3.1. L'hypothèse de la profondeur orthographique.....	423 -
III.3.1.1. Orthographes profondes vs. superficielles .....	423 -
III.3.1.2. Procédure d'assemblage (voie phonologique) vs. procédure d'adressage (voie lexicale) ....	424 -
III.3.1.3. Approche universelle : implication des deux voies dans les deux types d'orthographe .....	426 -
III.3.1.4. Facteurs déterminant l'implication de chaque voie .....	428 -
III.3.1.5. Conclusion.....	429 -
III.3.2. Taille des unités perceptives : lettre, graphème, corps, syllabe et mots .....	429 -
III.3.2.1. Lettre vs. Graphème.....	430 -
III.3.2.2. Graphème vs. Corps et Mots .....	431 -
III.3.2.3. Fréquence et taille des unités.....	433 -
III.3.3. Le cas du français .....	437 -
III.3.3.1. Le graphème .....	437 -
III.3.3.2. La syllabe.....	438 -
III.3.3.3. Fréquence, unités de différentes tailles et autres facteurs en lecture à voix haute.....	439 -
III.3.3.4. Conclusion .....	443 -
III.3.4. Le cas du japonais.....	443 -
III.3.4.1. Introduction .....	443 -
III.3.4.2. Les Kana .....	445 -
III.3.4.2.1. Distinction katakana/hiragana et familiarité scripturale .....	445 -
III.3.4.2.2. Kana et unités phonologiques.....	447 -
III.3.4.2.3. Fréquence et voie d'adressage pour les kana.....	447 -
III.3.4.2.4. Hémisphère gauche vs. hémisphère droit pour les kana .....	448 -
III.3.4.2.5. Conclusion .....	449 -
III.3.4.3. Les Romaji.....	449 -
III.3.4.4. Les Kanji.....	450 -
III.3.4.4.1. Implication de la voie phonologique pour les kanji.....	450 -
III.3.4.4.2. Hémisphère gauche vs. hémisphère droit pour les kanji.....	451 -
III.3.4.5. Conclusion.....	453 -
III.4. Reconnaissance lexicale et lecture à voix haute : modèles.....	455 -
III.4.1. Le Modèle Cohorte .....	455 -
III.4.2. Le Modèle à deux voies en cascades ( <i>Dual-Route Cascaded Model, DRC</i> ) .....	456 -
III.4.3. Modèles Interactifs : connexionnisme et intégration.....	460 -
III.4.3.1. Les Modèles d'Activation Interactive.....	460 -
III.4.3.2. Le Modèle Flou de la Perception .....	461 -
III.4.3.3. Morphologie et Prosodie en reconnaissance visuelle .....	463 -
III.4.4. Conclusion.....	465 -

<i>III.5. Influence de l'orthographe en perception de la parole.....</i>	<i>- 466 -</i>
III.5.1. Interactions entre codes orthographiques et phonologiques .....	- 466 -
III.5.1.1. L'hypothèse de l'autonomie orthographique : les données neurolinguistiques .....	- 466 -
III.5.1.2. Interactivité orthographe/phonologie : au niveau sublexical .....	- 469 -
III.5.1.3. Interactivité orthographe/phonologie : au niveau lexical .....	- 471 -
III.5.1.4. Conclusion : de la perception à la production .....	- 473 -
III.5.2. La conscience phonologique .....	- 474 -
III.5.2.1. Conscience phonologique précoce vs. Conscience phonémique .....	- 475 -
III.5.2.2. Apprentissage du code écrit et développement de la conscience phonologique.....	- 476 -
III.5.2.2.1. La conscience phonémique.....	- 476 -
III.5.2.2.2. La conscience phonologique non-phonémique .....	- 478 -
III.5.2.2.3. Conclusion .....	- 479 -

## **I. Perception et apprentissage de la parole en langue**

### **étrangère : description générale**

Si l'on souhaite établir un premier descriptif, partiel et non-fonctionnel, des caractéristiques psycholinguistiques d'un apprenant adulte d'une langue étrangère, voici ce que l'on peut obtenir (Detey, 2003b) :

- Deux hémisphères cérébraux,
- Une main dominante,
- Un appareil perceptif,
- Un appareil mnésique,
- Un appareil bucco-phonatoire,
- Une langue « première » (L1),
- Une compétence de communication en L1 :
  - Linguistique,
  - Discursive,
  - Référentielle,
  - Socio-interactionnelle.

Cette dernière se présente sous des aspects :

- Verbaux,
- Vocaux,
- Non-verbaux.

Elle est en outre liée :

- Aux compétences (méta)cognitives :
  - « Intelligence », raisonnements et réflexions (associations, inférences, attributions),
  - Savoirs/ savoir-faire,
- A un ensemble d'expériences individuelles et sociales.

Ces expériences débouchent sur un ensemble de:

- Croyances,
- Comportements,
- Catégories cognitives.

Cet ensemble détermine une :

- Matrice communicationnelle.

Deux sous-ensembles retiennent ici notre attention : les systèmes perceptivo-attentionnel et mnésique.

### **I.1. Système perceptivo-attentionnel et perception de la parole**

Le système perceptivo-attentionnel comprend, comme son nom l'indique, deux composantes étroitement associées : la perception d'une part, l'attention d'autre part. Si l'expression « perception subliminale » tend parfois à évoquer une « perception-sans-attention », une « perception inconsciente » ou encore une « perception sous le seuil de conscience », il convient cependant de bien distinguer, dès à présent, la *perception* de l'*apprentissage*.

De nombreux effets dit « de perception subliminale » (en particulier certains effets d'amorce, issus de la présentation préalable fulgurante d'un stimulus, rendant plus rapide le traitement ultérieur par le sujet d'un autre stimulus qui lui est lié d'une manière ou d'une autre) se mesurent en effet à travers certaines tâches, qui mettent à jour les effets d'une exposition à des stimuli non-consciemment perçus par les sujets, exposition qui conduirait à ce que l'on pourrait appeler un certain type d'apprentissage. Même si certains travaux récents en neuropsycholinguistique (Dehaene, Naccache, Le Clec'h, Koechlin, Mueller, Dehaene-Lambertz, Van de Moortele et Le Bihan, 1998), faisant appel à l'imagerie cérébrale, semblent pleinement confirmer l'hypothèse d'une perception subliminale, c'est-à-dire inconsciente, des mots en présentation visuelle, la



perception d'un événement par un sujet est étroitement liée à l'attention que celui-ci y prête, d'où l'association entre les deux sous-systèmes. Par ailleurs, dans le domaine de l'apprentissage des langues, il est par ailleurs admis que l'idée d'apprendre une langue de manière subliminale ne repose sur aucun fondement (« ce que l'on appelle l'apprentissage subliminal des langues n'existe pas »<sup>85</sup>, Van Patten, 2002, p. 108, se référant aux travaux de Schmidt), en particulier chez les auteurs qui ont travaillé sur le rôle de l'attention<sup>86</sup> dans l'acquisition des langues étrangères.

La perception, selon Bonnet (2003, p. 3) :

*« Désigne l'ensemble des mécanismes et des processus par lesquels l'organisme prend connaissance du monde et de son environnement sur la base des informations élaborées par ses sens. [...] [Elle] recouvre en effet l'ensemble des fonctions par lesquelles l'organisme « impose » une signification aux données sensorielles ».*

Cela rejoint pleinement la définition introductrice de la perception de la parole par Massaro (2001, p. 14870) :

*« Il est très courant d'avoir l'impression que les locuteurs de langues étrangères parlent bien plus rapidement que ceux de notre propre langue et qu'ils ne font pas de pause entre les mots et entre les phrases. Notre propre langue, en revanche, est perçue à un rythme normal, avec des pauses nettes entre les mots et entre les phrases. En fait, toutes les langues sont parlées à un rythme approximativement identique. Ainsi, nous définissons la perception de la parole comme le processus par lequel nous imposons une expérience*

---

<sup>85</sup> « There is no such thing as subliminal learning of language » (notre traduction).

<sup>86</sup> En anglais « noticing ».

*perceptive signifiante à un signal linguistique oral (parole) autrement dénué de toute signification »<sup>87</sup>.*

Cette tendance semble caractériser le système cognitif humain dans son ensemble, puisqu'on la retrouve, par exemple, dans le cas de la perception visuelle des formes, à travers les effets de contours illusoires et d'ombrage (Thorpe, 1994, p. 155).

Il convient, lorsque l'on traite de *perception* de distinguer *proprioception* d'une part (c'est-à-dire la perception des événements internes à l'organisme), et *extéroception* d'autre part (c'est-à-dire la perception des événements externes à l'organisme). Nous nous intéressons ici spécifiquement à la perception de la parole, en considérant celle-ci comme manifestation linguistique orale, émise par un locuteur lorsque celui-ci produit un ou des énoncés oraux. Nous nous concentrons dans un premier temps sur le canal auditif par lequel sont transmis ces énoncés, avant de prendre ensuite en considération l'autre canal essentiel de transmission, à savoir le canal visuel.

En ce qui concerne l'ensemble des processus à l'œuvre dans ce qu'il est convenu d'appeler *l'attention*, «on peut distinguer des processus qui opèrent directement sur l'information et des procédures qui vont moduler, arranger et faire aboutir ces différentes opérations mentales en fonction de la situation et des buts à atteindre» (Siéroff, 1994, p. 147). En fonction de la répartition et du degré d'attention d'un apprenant de langue étrangère, l'activité perceptive, et donc, par la suite, métalinguistique, de celui-ci va varier, puisque «l'attention règle la perception dans le cadre d'une organisation prévisionnelle» (Hamon, 1998, p. 24). Attention et perception sont donc étroitement liées : ainsi, «l'acquisition d'automatismes « complexes » correspondrait à la soustraction graduelle de la demande attentionnelle à mesure que la

---

<sup>87</sup> « It is commonplace to have the impression that foreign languages are spoken much more rapidly than our own, and without silent periods between the words and sentences. Our own language, however is perceived at a normal pace [...] with clear periods of silence between the words and sentences. In fact, languages are spoken at approximately the same rate [...]. Thus we define speech perception as the process of imposing a meaningful perceptual experience on an otherwise meaningless speech input » (notre traduction).

situation devient plus routinière » (Siéroff, 1994, p. 131). On oppose ainsi fréquemment un traitement sélectif conscient à un traitement automatique (Hamon et Juan de Mendoza, 1998, p. 77).

### **I.1.1. Perception de la parole: considérations générales**

Dès que l'on traite de perception de la parole (Nguyen, 2005), il faut distinguer « l'*audition*, qui relève de la sensibilité de l'oreille à entendre et la *perception*, qui procède d'une activité mentale de reconnaissance. [...] La perception concerne également l'interprétation de la réalité physique des sons » (Léon, 1992, pp. 41-42). Quatre paramètres physiques interviennent alors : l'intensité, la durée, la hauteur et le timbre. La perception de la parole se distingue en outre de la perception auditive de sons non-linguistiques : Celsis, Doyon, Boulanouar, et Nespoulous (1997) se sont en effet servis des techniques d'imagerie cérébrale fonctionnelle afin de déterminer « si la parole [était] perçue différemment des autres signaux auditifs et si cette différence [avait] des fondements cérébraux » (p. 85). Les résultats semblent indiquer que (pp. 92-93) :

*« Le contexte, par un effet de type « top-down », peut moduler le traitement d'un stimulus auditif, et plus particulièrement qu'un phonème présentant un caractère ambigu serait traité différemment selon ce contexte, peut-être de manière prédominante par l'hémisphère gauche en contexte de langue. Ces observations, sans démontrer l'existence d'un mode perceptif spécialement dédié à la parole, indiquent que des réseaux neuronaux différents seraient à l'œuvre pour traiter un son lorsqu'il a ou n'a pas valeur de phonème ».*

Les effets « haut-bas » mentionnés par les chercheurs jouent en effet un rôle non-négligeable dans la perception de la parole, effets jusqu'alors négligés dans une conception classique strictement ascendante de l'acte perceptif.

#### 1.1.1.1. Processus bas-haut et haut-bas

Le rôle des processus haut-bas, c'est-à-dire ceux mettant en jeu des informations connues du sujet mais non-explicitement présentes dans les stimuli cibles, dans la perception de la parole peut être envisagé de plusieurs manières, selon le niveau d'analyse auquel l'on se place (phonétique, phonologique, morphologique, etc.) et selon la perspective adoptée (en particulier lexicale ou sublexicale). La restauration d'informations manquantes (par exemple lorsqu'un phonème est absent ou masqué dans une unité lexicale) ou l'identification précise d'informations présentes (identification d'un phonème) peut ainsi être modulée de trois manières : l'une est liée à la connaissance du système linguistique (en particulier phonologique) de la langue du sujet et ainsi de son système phonotactique (connaissances statistiques), la seconde renvoie aux connaissances morphologiques (reconnaissance lexicale ou morphémique); la dernière enfin correspond à la connaissance du contexte (en particulier linguistique) dans lequel se trouve le stimulus cible.

- Connaissance du système phonologique (inventaire phonémique, structure syllabique, etc.) : la reconnaissance des mots isolés dépend en effet des propriétés phonologiques de la langue du sujet, puisqu'elles déterminent en partie au moins les procédures de segmentation du mot, sans doute même davantage que les propriétés acoustiques du stimulus (Ségui, 2003, p. 229). Ségui donne l'exemple du phonème /l/ dans le mot « balance » : tandis que les locuteurs français identifient les syllabes « ba-lance », les locuteurs anglais sont confrontés à l'ambisyllabité du /l/ dans le même mot dans leur langue. Si les premiers semblent ainsi faire usage des indices syllabiques, les seconds en revanche reposent davantage sur des indices d'ordre accentuel. A cela viennent s'ajouter des connaissances d'ordre statistique (phonotactiques).
- Reconnaissance lexicale : dans le cas de la restauration phonémique, lorsqu'un phonème est masqué par un autre signal acoustique dans un mot : « le système perceptif semble [...] être en mesure de compenser l'information sonore manquante

de manière telle que l'auditeur est victime d'une véritable illusion perceptive. [...] Les décisions prises par l'auditeur à propos des propriétés phonétiques du stimulus sont modulées par les informations provenant du lexique » (Ségui, 2003, p. 220). La fréquence lexicale constitue à cet égard un facteur non-négligeable. De tels phénomènes sont cependant dépendants de la focalisation de l'attention du sujet, ce qui souligne de nouveau l'importance des facteurs attentionnels dans le traitement perceptif de la parole. Il convient de clairement distinguer l'effet des connaissances pré-lexicales (comme les connaissances phonotactiques) des effets lexicaux haut-bas (connaissances lexicales) : l'étude de Dupoux, Pallier, Kakehi et Mehler (2001) à propos de la perception de voyelles illusoires au sein de groupes consonantiques illégaux par des sujets japonais l'illustre parfaitement, puisqu'elle leur permet de rendre compte de ce phénomène *via* des processus pré-lexicaux (connaissance phonotactique) et non *via* une influence lexicale de type haut-bas.

- Influence du contexte : il a en effet été montré que la catégorisation phonétique des stimuli effectuée par les sujets peut être influencée par le contexte lexical dans lequel ils se trouvent. Ségui (2003, p. 221) donne l'exemple d'un stimuli ambigu /p/ vs. /b/ sur la dimension du VOT<sup>88</sup>, qui est perçu tantôt /p/ tantôt /b/ selon qu'il se trouve respectivement dans le contexte « -alace » ou dans le contexte « -ateau ».

Plusieurs modèles de perception de la parole (par exemple le modèle TRACE), font ainsi la part belle au contexte et aux informations de niveaux supérieurs qui peuvent moduler le niveau d'activation des unités lexicales. Dans ces modèles interactifs, « un mot peut être "activé" par le seul contexte préalable, c'est-à-dire avant toute information sensorielle sur le mot lui-même » (Ségui, 2003, p. 230).

Ce qui, d'un point de vue linguistique, est vrai au niveau phonético-phonologique l'est également pour les autres : en tâche de compréhension orale, les prises d'informations

---

<sup>88</sup> Voice Onset Time (Durée d'Attaque Vocale): durée séparant le relâchement des articulateurs et le début des vibrations des cordes vocales (point de vue articulatoire).

s'effectuent à différents niveaux (phonétique, syntaxique, sémantique, contextuel). Il faut en effet se rappeler que « la notion d'*information* est polysémique. Au sens technique, ce terme fait référence à ce qui réduit de manière mesurable l'incertitude » (Bonnet, 2003, p. 3). Cette réduction de l'incertitude se réalise donc à travers des mécanismes complexes et « seule une conception reposant sur un fonctionnement interactif entre des opérations de différents niveaux, et sur une propagation bidirectionnelle de l'information (des informations linguistiques aux connaissances préalables de l'auditeur et inversement) permet de rendre compte de la formidable puissance du système de perception et d'identification humain » (Merlet, 1998, p. 143). La perception d'un mot est donc à la fois plus globale et plus analytique qu'elle peut paraître : « lors de la reconnaissance d'un mot, le système de traitement prend en considération non seulement ses propriétés intrinsèques mais également son contexte phonologique d'occurrence » (Ségué, 1997, p. 21).

Les effets d'amorce, d'attente et d'anticipation illustrent bien l'influence des connaissances du sujet sur la perception des stimuli linguistiques, et ils révèlent la nature non exclusivement linéaire et séquentielle du traitement. En ce qui concerne la perception de la parole (Abry et Schwartz, 1997, p. 112) :

*« Plus les attentes de l'auditeur (sa connaissance de la langue, du domaine...) sont pauvres, plus l'information contenue dans le signal doit être riche pour permettre une situation de communication réussie, d'où la nécessité pour le locuteur d' « hyperarticuler ». Au contraire, pour un auditeur averti, le locuteur peut diminuer le débit d'information présent dans le signal (en « hypoarticulant ») sans gêner la réussite de la communication » : c'est ce qui conduit un chercheur comme Lindblom à utiliser la notion d' « invariance adaptative ».*

Ces « attentes » peuvent alors conduire à des « illusions » perceptives : ainsi Dupoux, Kakehi, Hirose, Pallier et Mehler (1999) ont comparé, à travers quatre études, les comportements perceptifs d'auditeurs japonais et français, auxquels ont été présentées

des listes de non-mots. Il apparaît que les propriétés phonotactiques du japonais conduisent les auditeurs japonais à percevoir des syllabes « illusoires » au sein de groupes consonantiques. Dans un stimulus du type VCCV (ex : *ebzo*), les participants japonais ont indiqué percevoir une voyelle [u] entre les deux consonnes, montrant ainsi certaines difficultés à distinguer entre des stimuli /VCCV/ et /VCuCV/, contrairement aux participants français. Ces derniers ont en revanche des difficultés à distinguer des stimuli ne différant que par la longueur de la voyelle (*ebuzo* vs *ebuuzo*), une distinction qui est pertinente en japonais mais pas en français. Ces études conduisent les chercheurs à souligner la nécessité de prendre en compte les phénomènes d'assimilation d'origine phonotactique dans les modèles de perception de la parole, comme ils l'indiquent dans le résumé de leur article :

*« Il est apparu que les propriétés phonotactiques du japonais (un ensemble réduit de types syllabiques) conduisent les auditeurs japonais à percevoir des voyelles « illusoires » à l'intérieur de groupes consonantiques dans des stimuli du type voyelle-consonne-consonne-voyelle (VCCV). [...] Les participants japonais ont eu du mal à distinguer les stimuli de type VCCV et ceux de type VCuCV »<sup>89</sup>.*

En approfondissant cette étude, *via* la manipulation du voisinage lexical des non-mots utilisés dans leurs tests (par exemple *sokdo*, *sokudo*, *sokado*, le deuxième étant le seul mot japonais du triplet, et le premier présentant une séquence consonantique illégale en japonais), Dupoux, Fushimi, Kakehi et Mehler (1999) ont montré que cette « illusion » n'était pas due à une influence d'ordre lexical, mais bien pré-lexicale, c'est-à-dire avant l'activation de toute connaissance lexicale, soulignant ainsi la précocité de l'influence des connaissances phonotactiques sur la perception. Il ne s'agirait donc pas d'un effet haut-bas d'ordre lexical, mais bien sub-lexical. L'importance de cet effet a

---

<sup>89</sup> « We found that the phonotactic properties of Japanese (a reduced set of syllable types) induce Japanese listeners to perceive « illusory » vowels inside consonant clusters in vowel-consonant-consonant-vowel (VCCV) stimuli. [...] Japanese participants had trouble discriminating between VCCV and VCuCV stimuli » (notre traduction).

été particulièrement illustrée par les expériences de Hirose et Dupoux (2004), qui montrent que ces « illusions perceptives » peuvent même conduire le sujet à ignorer la présence d'informations acoustiques, en l'occurrence celles de la voyelle [i], placées entre deux consonnes.

De manière relativement similaire, Hallé, Ségui, Frauenfelder, et Meunier (1998) se sont penchés sur l'assimilation perceptive des groupes consonantiques français illégaux par des locuteurs français. La question des processus haut-bas rejoint alors, *via* le concept d'*assimilation perceptive*, la problématique du *crible phonologique*: « le système de sons dans la langue première agit comme un filtre phonologique intégrant, parfois "de force", certains phonèmes étrangers au sein de ses propres catégories phonémiques, à condition qu'il y ait une similarité suffisante »<sup>90</sup> (p. 592). Plusieurs chercheurs ont ainsi suivi cette piste et élaboré des modèles d'*assimilation perceptive*. « Tout auditeur a naturellement tendance à interpréter tout énoncé oral comme étant en adéquation avec l'organisation de sa langue première, ainsi qu'avec ses représentations linguistiques fonctionnelles »<sup>91</sup> (p. 592). Au sein de cette organisation figurent, outre un inventaire phonémique, des informations phonotactiques, lesquelles ont été moins étudiées que le premier dans ce domaine. Or, dans cette perspective : « Les groupes illégaux mais prononçables devraient tendre à être assimilés à des groupes légaux, sous réserve de similarité suffisante, de la même manière que les singletons étrangers ont tendance à être assimilés à des phonèmes de la langue première »<sup>92</sup> (p. 593).

Les auteurs se sont donc penchés sur les groupes consonantiques en français. En particulier, les groupes /dl/ et /tl/ ne sont pas légaux en position initiale, quoique prononçables. Comment vont alors être perçus les groupes /dl/ par des locuteurs

---

<sup>90</sup> « The native language sound system acts as a phonological filter molding some nonnative phonemes into native phonemic categories provided that there is sufficient similarity » (notre traduction).

<sup>91</sup> « Listeners are biased towards interpreting any spoken utterance as being consistent with the organization of their native language and with its functional linguistic representations » (notre traduction).

<sup>92</sup> « Illegal but pronounceable clusters should tend to be assimilated to clusters that are legal, whenever there is sufficient similarity, just as nonnative singleton phones tend to be assimilated to native phonemes » (notre traduction).



français ? Les chercheurs ont donc proposé une tâche de transcription orthographique libre aux sujets. Les résultats sont clairs : les groupes ont été respectivement transcrits /gl/ et /kl/ (qui sont des groupes autorisés en français) dans environ 85% des cas, ce qui montre que la perception du phonème /l/ est plus stable que celle de la consonne plosive initiale. Les tests ultérieurs ont confirmé l'existence d'une assimilation perceptive des groupes illégaux /tl/ et /dl/ en groupes légaux, /kl/ et /gl/, que les auteurs qualifient de « contextuelle », car cette confusion n'est pas due à une mauvaise perception du premier segment dental, mais est induite par la combinaison des deux segments formant l'attaque syllabique. Ils notent cependant qu'une généralisation aux autres groupes consonantiques française n'est pas justifiée, les groupes /tl/ et /dl/ ayant par ailleurs un statut particulier dans la distribution des groupes OBLI en français, puisqu'ils en sont absents (ce que Dell (1995) appelle un « accidental gap »). Après une discussion sur l'interprétation et les implications de ces résultats au regard de modèles psycholinguistique et phonologique, les auteurs concluent : « L'implication générale est que les auditeurs ont naturellement tendance à percevoir des événements et des objets davantage concordants avec leurs représentations internes qu'avec les stimuli réels »<sup>93</sup> (p. 606). Les effets haut-bas en perception de la parole semblent donc aujourd'hui solidement établis.

#### 1.1.1.2. Théories de la perception de la parole

Dans un premier temps, il est possible d'aborder la perception de la parole de manière strictement ascendante, dans un cadre classique de traitement de l'information (Ségui, 2003, p. 206). Cinq étapes relativement séquentielles peuvent alors être distinguées :

1. Niveau périphérique (propriété acoustique du signal et activité du nerf auditif) ;
2. Niveau auditif (encodage en mémoire à court terme et analyse acoustique) ;
3. Niveau phonétique (mise en relation des indices acoustiques et des propriétés phonétiques) ;

---

<sup>93</sup> « The general implication is that listeners are biased towards perceiving events and objects that fit with their internal representations rather than with the actual stimuli » (notre traduction).

4. Niveau phonologique (conversion des informations phonétiques en entités phonologiques: « c'est à ce niveau que le système perceptif « filtrerait » les variations phonétiques n'ayant pas de valeur distinctive » (Ségui, *ibid.*) ;
5. Niveau supérieur (accès lexical, reconnaissance et intégration des informations).

Il faut cependant noter que cette présentation repose en grande partie, à partir d'un certain niveau au moins, sur la notion d'atome phonologique, alors que la nature des primitives phonologiques est toujours débattue (trait phonologique, phonème, syllabe, approche non-symbolique). Par ailleurs, de manière générale en psycholinguistique, les modèles cognitifs de traitement de l'information sont aujourd'hui soit de type interactif (c'est-à-dire non strictement ascendant, mais combinant les apports informatifs de l'environnement et du sujet), soit de type connexionniste (Thorpe, 1994). Dans ces derniers, ce que l'on appelle le « paradigme subsymbolique » met en œuvre une « stratégie mixte, à la fois ascendante et descendante » qui aspire, par un ensemble de principes intermédiaires d'interprétation, à relier le niveau du substrat neuronal au niveau conceptuel dans une relation d'interprétabilité mutuelle (Laks, 1996, p. 60). La bidirectionnalité est donc de mise, pour les uns comme pour les autres. Par ailleurs, on connaît aujourd'hui de manière détaillée les phénomènes de filtrage perceptif auditif (Berger-Vachon et Collet, 1997, p. 29) :

*« Il est maintenant admis que la mécanique cochléaire n'est pas un simple transmetteur d'informations vers le cortex. La cochlée reçoit des contrôles venant des niveaux bulbaires et corticaux qui lui permettent de choisir, de sélectionner, les sons qui seront transmis vers les étages supérieurs ».*

Ségui (2003, pp. 217-218) propose un résumé de deux des théories de la parole les plus établies : la théorie motrice de la perception de la parole d'une part (qui renvoie à une *approche articulatoire*), les théories d'analyse par synthèse d'autre part (qui renvoient à une *approche acoustique*).

Dans la théorie motrice de la perception de la parole, la perception de sons linguistiques revient à les percevoir en tant que produits de gestes articulatoires. Perception et production des sons y sont donc étroitement liées, résolvant ainsi le problème de l'invariance phonémique par leur étroite association (l'absence d'invariants acoustiques est compensée par la présence d'invariants articulatoires). Cela correspond à la version « révisée » de la théorie (Liberman et Mattingly, 1985), dans laquelle l'information phonétique est perçue par un « module » spécialisé dans la détection des cibles gestuelles articulatoires (« intended gestures ») du locuteur, gestes articulatoires qui constituent la base des catégories phonétiques. Ce module se trouve par ailleurs en compétition avec un module de perception auditive générale. Les auteurs définissent un geste articulatoire comme suit (Liberman et Mattingly, 1985, p. 21) :

*« A phonetic gesture [...] is a class of movements by one or more articulators that results in a particular, linguistically significant deformation, over time, of the vocal-tract configuration. The linguistic function of the gesture is clear enough : phonetic contrasts ».*

La théorie se compose donc de deux parties (p. 10) : « gestures are the objects of perception and [...] perception of these gestures depends on a specialized module ». A l'appui de celle-ci vient un ensemble de travaux relatifs, notamment, à la coarticulation, qui conduisent Liberman et Mattingly à affirmer que (p. 12):

*« Putting together all the generalizations about the multiplicity and variety of acoustic cues, [...] there is simply no way to define a phonetic category in purely acoustic terms ».*

Les effets de coarticulation peuvent en effet conduire à des écarts entre segmentation phonétique et acoustique, ainsi qu'à des situations où, dans différents contextes, des sons différents peuvent aboutir au même percept d'une part, et des sons identiques à des percepts différents d'autre part.

Dans les théories d'analyse par synthèse (il en existe plusieurs versions), la perception s'effectue *via* l'ajustement progressif de patrons phonatoires générés par le sujet avec le signal acoustique, un ajustement qui résulte dans la transformation du signal acoustique en information phonémique : « l'auditeur s'efforce de produire à un niveau abstrait de représentation les configurations articulatoires correspondant à la séquence de phonèmes véhiculée par le signal acoustique » (Ségui, 2003, p. 218). Très proche de la théorie motrice, elle s'en distingue cependant en ce que la génération de patrons phonatoires correspond davantage à un processus de vérification opéré par le système de production en vue de l'identification phonémique des informations acoustiques ; il n'y a donc pas de lien « direct » entre signal et gestes articulatoires comme dans la théorie motrice.

Ces deux cadres théoriques adoptent néanmoins la même approche de la perception, à savoir celle de la perception catégorielle. La conception catégorielle de la perception constitue en effet la base de nombreux travaux dans le domaine (Harnad, 1987). Ségui (2003, p. 208) donne l'exemple de la distinction entre consonnes occlusives sourdes et sonores (/b/ vs. /p/, /d/ vs. /t/ et /g/ vs. /k/), qui tient, d'un certain point de vue, à la différence du temps d'attaque vocale (*Voice Onset Time*), paramètre acoustique qui, rappelons-le, d'un point de vue articulatoire, correspond à la durée de temps séparant le relâchement des articulateurs et le début des vibrations des cordes vocales. La spécificité linguistique de la perception catégorielle est un sujet par ailleurs encore débattu : tandis que certains travaux indiquent que « la perception catégorielle des phonèmes ne dépend pas de processus spécifiques à la parole et à l'homme » (Celsis, Doyon, Boulanouar, et Nespoulous 1997, p. 86), d'autres, en revanche, « sont en faveur de l'existence d'un mode de perception phonétique de la parole distinct du mode de perception auditif général » (*op. cit.*, p. 87).

Trois types d'approche ont été avancées pour rendre compte des phénomènes de perception catégorielle : l'approche motrice, qui considère que la nature catégorielle de la perception résulte de la nature catégorielle des gestes articulatoires, l'approche

auditive, qui suppose qu'elle relève d'une sensibilité naturelle du système auditif, sans référence à un module spécifique au langage, et enfin une approche basée sur l'apprentissage, qui estime que le recours aux catégories est déterminé, non par des facteurs strictement articulatoires ou auditifs, mais par l'apprentissage. Ces trois approches ont été comparées par Rosen et Howell (1987), une comparaison qui, loin de privilégier l'une ou l'autre, insiste sur la nécessité d'intégrer les trois en vue de rendre compte de l'ensemble des phénomènes: « the safest remark to make is that none of the three theories of CP [(Categorical Perception)] will serve to explain all the findings on its own. Some interplay among them is necessary » (p. 152).

#### 1.1.1.3. De la perception à la compréhension

Lorsque l'on traite de perception linguistique, particulièrement en situation d'apprentissage, il est rare que cette perception ne constitue pas l'étape préliminaire d'un processus de compréhension d'ordre sémantique. Or, de la même manière que le traitement perceptif de l'information intègre des informations non présentes dans le signal, la compréhension d'un énoncé, ou bien d'un discours, met en oeuvre un ensemble de savoirs pour traiter *l'input* linguistique. Klein propose de regrouper ces savoirs en quatre catégories (1986, p. 63) :

- Connaissances générales à propos de la nature du langage humain et de la communication verbale ;
- Connaissances spécifiques de la structure de sa langue première, et éventuellement d'autres langues qu'il connaît ;
- Connaissances de la langue-cible dont il dispose, ou dont il pense disposer ;
- Toutes sortes de savoirs non linguistiques.<sup>94</sup>

---

<sup>94</sup> « General knowledge about the nature of human languages and of verbal communication ; specific knowledge of the structure of his first language, or possibly of any other language he happens to know ; knowledge of the target language which he has or presumes to have ; all manner of non-linguistic knowledge » (notre traduction).

Loin d'être un sujet-récepteur passif, l'énonciataire joue ainsi un rôle essentiel dans l'acte de compréhension, car comme l'écrit Moirand (1982, p. 130), « comprendre c'est produire de la signification ». L'article de Oléron (1981, p. 110), bien que déjà ancien, nous semble clairement exposer la problématique, en précisant en outre que « la compréhension du langage comporte à la fois des aspects linguistiques, puisqu'il s'agit du langage, et des aspects cognitifs, puisque la compréhension est une démarche intellectuelle qui peut se manifester – et a largement l'occasion de s'exercer - sur des matières non verbales (situations, évènements, comportements de personne...) ». Pour l'auteur, la compréhension est constituée par l'interprétation d'indices, ces indices étant intégrés avec ce à quoi ils renvoient ou ce qu'ils évoquent, le contexte jouant naturellement un rôle de tout premier ordre. « Rentrent dans le contexte les éléments qui relèvent de la pragmatique où interviennent la connaissance du locuteur, celle des règles sociales imposant des formes indirectes et parfois déguisées aux affirmations, demandes, ordres, et sans laquelle la signification exacte de ceux-ci ne serait pas perçue » (p. 111). En mentionnant l'intervention des connaissances du locuteur et donc l'activité du sujet dans les démarches qui l'amènent à comprendre les énoncés, Oléron insiste sur « l'aspect constructif de la compréhension, en soulignant les aspects inférentiels de l'exploitation des indices et de l'intégration » (*ibid.*). Il affirme cependant que « le sujet n'exerce ses activités que dans le cadre des savoirs qu'il a acquis et dont il dispose sur l'univers des objets et des personnes » (*ibid.*). Notons qu'à l'heure actuelle, il est de plus en plus courant, lorsqu'on évoque les « savoirs » du locuteur de se référer à la notion de « schéma », ce que fait Oléron en retrouvant dans les processus de compréhension « l'articulation entre le connu (les schémas ou scénarios) et le nouveau ([les] données) et l'appel à la cohérence (schémas et scénarios sont structurés en fonction de la nature des choses ou des règles usuelles d'exposition) » (*ibid.*).

Dans une perspective didactique, Gremmo et Holec, traitant de la compréhension orale, (1990, p. 30), introduisent les deux grands types de modèles non connexionnistes de compréhension orale que les recherches en psycholinguistique portant sur la construction de sens opérée par le participant-auditeur d'un discours ont permis

d'élaborer : d'un côté les modèles sémasiologiques (bas-haut), de l'autre les modèles onomasiologiques (haut-bas). Cependant, comme nous l'avons dit, le modèle désormais le plus communément reconnu est un modèle interactif conjuguant des processus bas-haut (*bottom-up*) de décodage des signes, et haut-bas (*top-down*) d'anticipation et d'inférence. La construction de la signification dans une tâche de compréhension implique donc une interaction entre l'interprétation contrainte par les données (*data driven*) et celle contrainte par notre « savoir dormant » (*knowledge driven*). C'est précisément cette interaction qui permet de pouvoir comprendre un message sans passer par un traitement obligatoire de tous les niveaux (phonétique, syntaxique, etc.) dudit message.

Dans le modèle onomasiologique, la priorité est accordée aux informations apportées par l'auditeur, « à l'opération créative de préconstruction de la signification du message par l'auditeur » (Gremmo et Holec, 1990, p. 33). Comme Klein, les auteurs proposent ainsi une liste des « connaissances grâce auxquelles l'auditeur anticipe la signification du message » (p. 33) :

- Connaissances sociolinguistiques sur la situation de communication ;
- Connaissances socio-psychologiques sur le producteur du message ;
- Connaissances discursives sur le type de discours concerné ;
- Connaissances linguistiques sur le code utilisé ;
- Connaissances référentielles sur la thématique invoquée ;
- Connaissances culturelles sur la communauté à laquelle appartient le producteur du message.

Qu'il s'agisse de texte oral ou écrit, il faut donc, comme l'écrit Carrell (1990, p. 16) à propos des textes écrits :

*« Garder à l'esprit que [l'] interaction [entre énoncé et sujet-interprétant] implique tous les types de connaissances que le lecteur [(sujet-interprétant)] utilise dans une tâche de lecture [(compréhension)]-connaissances et*

*croyances sur le monde (schémas de contenu), connaissance de différents types de textes, de leur organisation et de leur structure typiques (schémas formels), tout autant que les connaissances linguistiques (schémas linguistiques), y compris des connaissances lexicales, syntaxiques, sémantiques et pragmatiques ».*

Carrell rapporte les résultats de nombreuses recherches qui montrent que les schémas, tant de contenu que formels, affectent significativement la compréhension textuelle.

Ces informations sont particulièrement déterminantes en ce qui concerne les dimensions implicite et pragmatique du langage. Comprendre un acte de langage non littéral implique, d'une part « l'intégration de ce qui est explicitement dit », et d'autre part « la capacité de dépasser cette signification littérale pour saisir l'intention du locuteur dans le contexte donné » (Champagne, 1999, p. 241). Dans certains modèles, auxquels se réfère Champagne, cette compréhension implique « un traitement simultané du sens littéral et du sens non littéral, et non pas un traitement séquentiel ». Néanmoins, d'autres études citées par Champagne semblent montrer que l'analyse de l'interprétation littérale n'est pas nécessairement effectuée par les sujets. En revanche, ils « seraient attentifs au sens conventionnel des énoncés, étant donné un contexte situationnel. Il existerait une interaction entre la conventionnalité et le type de contexte social dans lequel elle apparaît. La compréhension des énoncés non littéraux semblerait donc automatique » (p. 242). Nous pourrions presque considérer, en raison de routines contextuelles et langagières, qui nous ramènent aux « schémas » mentionnés par Oléron, qu'un acte de langage non littéral conventionnel change de statut, passant d'un acte indirect à un acte direct, de même qu'une métaphore morte ou figée perd quasiment son statut figuratif. L'importance des processus descendants apparaît donc nettement dans le traitement des actes de langage non littéraux et de leur dimension implicite, puisque ce traitement repose en grande partie sur les connaissances préalables du sujet et les capacités d'inférences qui leur sont liées.



Ainsi, nous avons vu à quel point, en tâche de compréhension orale, écrite ou visuo-spatiale, le traitement des stimuli met en œuvre l'utilisation des différents types de savoirs du sujet. Dans une perspective didactique, ce constat invite à la prise en compte de ces savoirs afin, soit d'améliorer l'acquisition d'une compétence de compréhension, soit de contrecarrer certains effets de ces connaissances, potentiellement nuisibles à l'apprentissage. En raison des phénomènes d'attente et d'anticipation, ainsi que du traitement de la dimension implicite des énoncés, ce constat impose également de prêter une attention toute particulière au rapport entre les structures langagières visées, qu'elles soient fonctionnelles ou notionnelles, et les contextes dans lesquels elles sont ou ne sont pas insérées.

Quel que soit le cadre adopté, ce qui importe ici, dans une perspective globale, c'est de définitivement bannir la conception strictement ascendante de l'intégration perceptive pour souligner le rôle actif du sujet en tâche de perception : que celle-ci soit visuelle, auditive ou qu'elle corresponde à une tâche de compréhension linguistique, les connaissances latentes du sujet imposent une certaine orientation interprétative au traitement des stimuli, dans une série de processus *haut-bas* : « pour certains auteurs, un effet "descendant" se définit comme la modification du traitement effectué sous l'influence de l'expérience (ou des connaissances acquises) » (Thorpe, 1994, p. 159). La conception strictement ascendante de l'acte perceptif, sous une forme ou sous une autre, est malheureusement encore la plus répandue parmi les profanes en sciences cognitives, et parmi eux une large majorité des enseignants de langue.

### **I.1.2. Perception de la parole et reconnaissance linguistique**

Comme le rappelle Ségui (2003, p. 205), trois problèmes principaux émergent dans l'étude de la perception de la parole:

1. Continuité du signal acoustique *vs.* Caractère discret de la perception ;
2. Variation du signal acoustique *vs.* Invariance des phonèmes ;
3. Non-linéarité de la distribution des indices acoustiques *vs.* Séquentialité de la perception des sons.

Le caractère discret de la perception, qui renvoie à la segmentation de la chaîne acoustique par l'auditeur, pose alors le problème des frontières et de la nature des ces éléments discrets, qui constituent les unités perceptives.

#### 1.1.2.1. Unités perceptives : niveau sublexical

Les travaux de Dupoux et de ses collègues (1999) sur la perception de voyelles « illusoires » entre deux consonnes par des sujets Japonais peuvent être reliés à la problématique évoquée ci-dessus : celle de la segmentation de la chaîne parlée en l'absence d'indices acoustiques invariants (notamment en raison des effets de coarticulation), problème qui, à son tour, renvoie à celui des unités perceptives de base. Il faut en effet s'interroger sur les unités qui sont perçues et traitées en perception de la parole. Deux questions essentielles, quoique générales, nous semblent alors devoir être formulées :

- Ces unités sont-elles universelles ou spécifiques aux langues ?
- Correspondent-elles aux unités descriptives formelles élaborées par les linguistes ?

Car, comme l'écrivent Otake, Hatano, Cutler et Mehler (1993, p. 260) :

*« Concepts such as phoneme, mora, syllable, and stress unit are not properties of particular languages ; they are phonological constructs in terms of which any language can be described. However, just as there are differences across language communities in the way verse forms are typically constrained, and in the way phonology may be reflected in orthography, so are there differences across languages in the degree to which various phonological constructs play a role in phonological processes within the language [...] [and] in the importance of the same phonological constructs in speech recognition processes ».*

Plusieurs études se sont ainsi attachées à évaluer la validité de telles unités, que ce soit en termes d'universalité linguistique ou de réalisme psychologique, comme par

exemple celle de Nesterenko (2002) qui, à travers une tâche de comptage syllabique par des auditeurs français, pour des stimuli français et russes, s'est penchée sur le statut cognitif de la syllabe en français. Car, si le phonème a longtemps constitué l'atome phonologique minimal en linguistique et en psycholinguistique, d'autres unités, de niveaux inférieur ou supérieur, semblent à présent pris en compte dans ces domaines, particulièrement dans celui de la perception de la parole. Plusieurs travaux semblent en effet indiquer que la syllabe constitue une unité de segmentation du langage à un niveau intermédiaire entre phonème et unité lexicale : en tâche de détection de phonème, la syllabe serait ainsi le socle à partir duquel le sujet peut effectuer une analyse phonémique. De plus, la syllabe serait aussi l'unité donnant accès aux représentations lexicales (Ségui, 2003, p. 222). Le « rôle de la syllabe dans la perception de la parole » est ainsi l'intitulé de la thèse de doctorat de C. Pallier (1994), dans laquelle il explore la question du format des représentations perceptives en tâche d'identification/de reconnaissance : formats classiques (suite de phonèmes, matrices de traits distinctifs) et plus récents (valeurs continues des traits ou des phonèmes, approche acoustique non-linguistique, syllabes). Suite à ses expériences, Pallier en arrive à considérer, s'inspirant de certains auteurs, que la perception de la parole s'effectue à l'aide d'un parseur phonologique, élaborant de manière continue une représentation linguistique dans laquelle l'organisation des traits distinctifs serait basée sur les unités syllabiques. La syllabe jouerait ainsi un rôle de « code correcteur » de l'interprétation phonétique du signal (p. 120).

Ce questionnement, relatif aux unités pertinentes en perception de la parole, accompagne naturellement celui des théories phonologiques : des *Preliminaries to Speech Analysis* de Jakobson, Fante et Halle, à la Théorie de l'Optimalité telle que présentée par Prince et Smolensky, par exemple (Boë et Schwartz, 1997 ; Durand et Laks, 2002b ; Durand et Lyche, 2000), les modèles phonologiques on en effet beaucoup évolué, et ces nouveaux cadres, comme par exemple celui des phonologies non linéaires, se trouvent déjà mis à profit dans des domaines tels que l'aphasiologie (Nespoulous et Moreau, 1998 ; Valdois, 1990 ; Valdois et Nespoulous, 1994), conduisant certains à affirmer que « le phonème est devenu de plus en plus marginalisé

dans les théories phonologiques »<sup>95</sup> (Harris, 1998, p. 91). Même si le phonème garde une certaine pertinence, il ne peut plus être abordé comme il l'était auparavant. Cela est parfaitement illustré par Brandão de Carvalho et Kawaguchi (2002) qui retracent l'évolution du concept de phonème, et plus particulièrement sa dimension « linéaire », depuis la phonologie structurale de l'Ecole de Prague jusqu'à la phonologie multilinéaire, et qui, arguments diachroniques à l'appui, parviennent à montrer que le phonème n'est précisément pas une unité linéaire, comme en témoignent tant certains phénomènes d'assimilations que de nombreuses expériences en phonétique perceptive (par exemple la perception de séquences /Cr/ dans lesquelles les allophones de /r/ en français sont manipulés conduisant les sujets à interpréter différemment la consonne initiale /C/ en fonction desdits allophones (opposition /k/ vs. /g/) : « dans les groupes /Cr/, la "place" fondamentale du trait pertinent de voix n'est pas le premier segment, mais le second » (p. 2)). Cette évolution a notamment pu se réaliser grâce à la distinction entre *position*, qui renvoie au squelette (lequel « encode le temps phonologique » (p. 6)), et *contenu*, qui renvoie au segment (qui est « la mise en forme de la substance phonique en traits pertinents » (p. 6)). Cette distinction permet en particulier de bien rendre compte du phénomène de liaison en français, également lié à la question du « e muet », grâce aux notions de positions vides et de segments flottants. Cela permet ainsi d'écarter, selon les auteurs, toute influence de la graphie sur la liaison (p. 6). Un exemple tiré du japonais est celui de la réalisation en [çi] de la syllabe /hi/. Tandis qu'une interprétation phonologique classique conduira à poser que « le phonème /h/ du japonais se réalise [ç] devant /i/, formulation qui sauvegarde la linéarité fondamentale des constituants de la syllabe /hi/ » (p. 2) (la linéarité phonologique étant alors préservée, malgré son abolition phonétique partielle), une interprétation multilinéaire conduira à affirmer que « la palatalité de /h/ appartient à une réalisation contextuelle de /i/. [...] il n'y a plus dès lors de séquence /h+i/. » (p. 16).

---

<sup>95</sup> « The phoneme has become increasingly marginalized in phonological theory » (notre traduction).

Les concepts de *trait* et de *syllabe*, notamment, ont ainsi pris une place de plus en plus importante dans les domaines concernés (Durand et Laks, 2002b, p. 15) :

*« L'un des résultats à présent acceptés par de nombreux phonologues est la multilinéarité (l'idée que les représentations sont organisées autour d'axes qui ne coïncident pas) et l'existence de constituants suprasegmentaux comme les syllabes. [...] La syllabe est considérée par certains auteurs comme l'unité fondamentale de laquelle les segments et les traits sous-segmentaux émergent ontogénétiquement et qui assure le lien entre un système de bas niveau (« physically driven system ») et des unités de plus haut niveau »<sup>96</sup>.*

La syllabe est donc souvent considérée comme l'unité d'interface stockée entre les niveaux prélexical et lexical de traitement (Ségui, 2003), comme cela est illustré en français dans une tâche de détection (« monitoring task ») souvent citée dans la littérature : la cible /pa/ est détectée plus rapidement dans le mot /pa.las/ que dans le mot /pal.mje/, tandis que la cible /pal/ est détectée plus rapidement dans le second que dans le premier (Ségui et Ferrand, 2002, p. 152). L'étude conduite par Kolinsky, Morais et Cluytens (1995) sur la perception de mots illusoires issus de la conjonction de deux stimuli différents présentés en écoute dichotique (par exemple : /bitō - kōzu/ perçu comme /biʒu/) soutient également l'hypothèse selon laquelle la syllabe jouerait un rôle majeur en tant qu'unité de segmentation en français, puisque la perception de ces mots illusoires, en tâche d'identification et de détection, semble dépendre crucialement des relations d'ordre syllabique entre les deux stimuli. Les chercheurs adoptent cependant une position neutre quant à l'agencement temporel du processus perceptif (p. 36) :

---

<sup>96</sup> « One of the results now taken for granted by many phonologists is multilinearity (the idea that representations are organised along tiers that do not coincide) and the existence of suprasegmental constituents such as the syllables. [...] The syllable is seen by some contributors as the fundamental unit from which segments and subsegmental features emerge ontogenetically and which provides the link between a physically-driven system and higher-level units » (notre traduction).

*« We do not intend to claim that the perception of syllables precedes the perception of their constituents. The reverse temporal processing order seems just as plausible. From this point of view, both syllabic effects would reflect a stage of processing in which the internal properties of the syllable are no longer represented ».*

A un niveau d'analyse encore plus fin, Dumay, Frauenfelder et Content (2002), à travers des tâches de repérage de mots (« word-spotting »), se sont penchés sur l'influence respective des attaques (« onsets ») et codas (« offsets ») sur la segmentation lexicale en français : les sujets devaient détecter des mots de type CVC enchâssés, à l'initiale ou en finale, dans des non-mots bisyllabiques dont la structure syllabique était tantôt alignée avec celle du mot-cible (CVC.CVC), tantôt en décalage (CV.CCVC). Les résultats semblent indiquer que les auditeurs français s'appuient sur les attaques syllabiques pour repérer le début des mots.

Cependant, d'un point de vue psycholinguistique, il semblerait que la fonction attribuée à la syllabe, notamment en perception, varie selon les langues, tout comme les unités de segmentation varient selon la langue maternelle de l'auditeur (Segui et Ferrand, 2002). En effet, tandis que l'existence de cet effet syllabique, mis à jour par Mehler et ses collègues, a été confirmée à plusieurs reprises chez des auditeurs francophones natifs<sup>97</sup>, quelle que soit l'appartenance linguistique des stimuli (stimuli français, anglais, japonais), et qu'un effet similaire a pu, sous certaines conditions, être mis en évidence chez des auditeurs catalans, espagnols et portugais (*op. cit.* p. 152), les langues germaniques (par opposition aux langues romanes précédemment citées) comme l'anglais et le néerlandais semblent présenter un usage différent de la syllabe. En anglais notamment, les frontières syllabiques sont, comparativement, fréquemment moins nettes qu'en français, avec de nombreuses consonnes ambisyllabiques, comme le [L] dans le mot « balance » : [BA[L]ANCE] et l'on trouve donc des résultats

---

<sup>97</sup> Bien que certaines études consacrées à ce thème ne soient pas parvenues à obtenir le même effet (Content, Meunier, Kearns et Frauenfelder, 2001).

contradictoires quant à l'existence ou non d'un effet syllabique similaire en anglais (*op. cit.*, p. 158). Il se pourrait que la syllabe intervienne en tant qu'unité de traitement, mais plus tardivement qu'en français. En néerlandais aussi, les résultats sont contradictoires, cette langue étant, comme l'anglais, une langue à rythme accentuel, à la différence du français, qui est à rythme syllabique. La nature des processus de segmentation rythmique (syllabique pour le français, accentuelle pour l'anglais, moraïque pour le japonais)<sup>98</sup> semblerait en effet déterminante dans le choix des unités perceptives de segmentation (*op. cit.*, p. 165), et plusieurs travaux en production de la parole semblent permettre d'établir un parallèle entre les deux processus à ce sujet. L'étude de Dupoux, Pallier, Sebastian et Melher (1997) sur la perception de l'accent<sup>99</sup> lexical par des sujets espagnols et par des sujets français (l'accent étant distinctif en espagnol, pas en français) conforte cette idée, puisque, notamment à travers deux expériences de détection de différences accentuelles et une d'ignorance de celles-ci, les chercheurs montrent que, les français sont « sourds » à l'accent lexical, tandis que les sujets espagnols le traitent automatiquement en vue d'identifier l'item lexical qui le porte, mettant à jour l'importance des facteurs prosodiques dans la perception de la parole, pour certaines langues en particulier, facteurs généralement peu intégrés dans les modèles de perception.

Dans le cas du japonais, nous pouvons citer deux études en particulier qui se sont penchées sur les statuts respectifs de la syllabe et de la more en japonais. Kubozono (1989), à travers l'étude de plusieurs phénomènes linguistiques, dont différents types de lapsus (substitutions et transpositions), démontre dans un premier temps la réalité

---

<sup>98</sup> Ce découpage linguistique traditionnel des classes rythmiques a cependant été réexaminé par Ramus et ses collaborateurs (Ramus, Dupoux et Mehler, 2003), qui, dans une approche psycholinguistique quantitative, empiriquement basée sur des expériences perceptives, en soulignent les faiblesses : tandis que certaines langues, comme l'anglais ou l'italien, semblent se conformer au découpage traditionnel en deux ou trois classes, le Polonais semble nécessiter la constitution d'une nouvelle classe. L'approche phonético-acoustique quantitative des chercheurs, axée sur la durée des intervalles vocaliques et consonnantiques, a été testée à l'aide de stimuli de synthèse, assurant que les sujets ne puissent disposer que d'indices rythmiques, quoique ceux-ci ne fassent appel qu'aux variations de durée et non d'intensité.

<sup>99</sup> En tant que simple prééminence phonétique/perceptive, indiquent les chercheurs (p. 407, note 1).

psychologique de la more en production de la parole en japonais, en particulier *via* l'importance des frontières moraïques, car comme il l'écrit (p. 258) :

*« Because all syllable boundaries are mora boundaries by definition, it is a priori, impossible to find any data in favor of the syllable that are against the mora, while it is possible to find evidence for the mora that argues against the syllable ».*

En ce qui concerne la constitution de la syllabe en japonais, Kubozono indique que, contrairement à des résultats similaires en anglais (p. 267) :

*« Speech error data in Japanese provide no substantial evidence for the existence of the rhyme as a syllable unit, thus reinforcing the view [...] that the traditional analysis of the syllable into onset and rhyme does not represent a universal syllable structure, as often implied in the literature ».*

Si la pertinence de la syllabe en japonais n'est donc pas rejetée, l'étude de Kubozono le conduit à insister sur l'importance de la more et des frontières moraïque en japonais, et à minimiser celle des frontières et des notions de rime syllabiques.

D'avantage orientés vers la perception de la parole, Otake et *al.* (1993) se sont penchés sur les caractéristiques phonologiques desquelles dépendent les procédures de segmentation des mots parlés, dans le cas du japonais et du français. Tandis que le rythme français est dit syllabique, conduisant ses locuteurs à adopter la syllabe comme unité de segmentation, le rythme japonais est considéré comme étant moraïque. La difficulté, en terme d'investigation psycholinguistique, dans le cas du japonais, réside en ce que très souvent, la more et la syllabe se superposent. La question abordée par les chercheurs était donc de vérifier si, en accord avec les données sur le français et l'anglais quant au lien entre caractéristiques rythmiques et unités de segmentation, les auditeurs japonais adoptent effectivement la more comme unité. La conclusion est nette : « Japanese listeners do not naturally segment speech syllable by syllable ; they



do naturally segment it mora by mora » (p. 270). De manière plus générale, au vu des résultats antérieurs, les chercheurs indiquent que : « the way listeners accomplish speech segmentation is influenced by the way their native language organizes its rhythm » (p. 276). En ce qui concerne la possibilité de développer plusieurs types de procédure, les chercheurs, suite à leurs travaux antérieurs, répondent par la négative, en se référant à l'anglais (p. 277) :

*« English-speaking monolinguals cannot exploit syllabic rhythm to aid segmentation, even when they are listening to speech in French, in which syllabic segmentation would be more efficient. More remarkably yet, maximally competent French-English bilinguals appear to have only one such segmentation procedure available to them [...] – either the syllabic segmentation [...], or the stress-based segmentation procedure ».*

Il existe ainsi toute une série de travaux ne soutenant pas l'hypothèse du statut universel d'unité perceptive primaire de la syllabe en perception de la parole. Pallier (1994, chap. VII), dans ses propres expériences, a ainsi obtenu de nombreux résultats syllabiques, mais aussi « anti-syllabiques », une contradiction qu'il propose de résoudre à travers un modèle élaboré en tenant compte du poids des processus décisionnels dans les tâches expérimentales, puisque « le choix d'une représentation plutôt qu'une autre est sous le contrôle du sujet, mais typiquement on suppose que celui-ci choisit la mieux adaptée à la tâche » (p. 118). Cette contradiction apparente, qui s'exprime à travers des résultats favorisant tantôt la syllabe, tantôt le phonème (ou le trait), serait due, d'après Pallier, à l'hypothèse tacite selon laquelle « ces unités [existent] à des niveaux de représentation distincts, et que l'un de ces niveaux [doit] [...] nécessairement se trouver avant les autres » (p. 124). Afin de résoudre ce problème, Pallier adopte au contraire l'hypothèse d'un format unique de représentation, qui sous-tend tant la perception que la production de la parole. La « clé » du modèle qu'il propose se formule ainsi (p. 124) :

*« Les effets "syllabiques" [...] s'interprètent très bien en supposant que notre système perceptif fournit une représentation du signal structurée syllabiquement. Ils n'impliquent pas (mais ne rejettent pas non plus) l'existence de détecteurs de syllabes, ni même d'une représentation purement syllabique où la parole serait représentée par une chaîne ».*

Dans son modèle, la représentation phonologique, unique, du signal s'élabore progressivement dans le temps, et si la syllabe ne constitue pas un « atome », puisqu'elle se construit progressivement, elle joue cependant un rôle organisationnel important.

En ce qui concerne la nature de la syllabe en production de la parole (soit des unités (« chunks ») dans lesquels le contenu segmental est spécifié, soit des schémas sans spécification segmentale), les données expérimentales semblent compatibles avec la notion de schéma, même si les résultats de nombreuses expériences passées, qui ne permettraient pas d'en distinguer les deux aspects, pourraient soutenir l'une ou l'autre interprétation (Ségui et Ferrand, 2002, p. 161). Cela conduit à un « modèle mixte » de la syllabe (p. 165) (la variation dans les résultats étant liée à la nature de la tâche expérimentale) dans lequel deux niveaux coexistent, celui des unités (« chunks ») et celui des schémas, tous deux reliés au niveau lexical.

La question des unités perceptives ayant été ainsi balisée, nous pouvons à présent aborder une étape ultérieure dans le traitement de la parole, à savoir celle de la reconnaissance. Nous devons bien sûr signaler qu'une telle présentation séquentielle ne fait pas l'unanimité : la question reste encore ouverte de savoir s'il faut toujours distinguer la « perception » de la « reconnaissance », suivant un modèle séquentiel à deux étapes qui rend traditionnellement bien compte de la différence entre agnosie aperceptive et agnosie associative (Peretz, 1994, p. 206 ; voir également le schéma simplifié du modèle « logogène » de Morton présenté par Kremin, 1994, p. 379) ou bien s'il convient de suivre une modélisation connexionniste « qui conçoit la

perception comme faisant partie intrinsèque du processus de reconnaissance » (Peretz, 1994, p. 206).

### I.1.2.2. Reconnaissance linguistique

Avant de traiter plus avant de reconnaissance phonético-phonologique, sublexicale ou lexicale, il importe de présenter quelques notions générales relatives au processus de reconnaissance, tout en rappelant la perspective didactique qui est, en grande partie, la nôtre.

#### I.1.2.2.1. Appariement, catégorisation et reconnaissance : considérations générales

Si nous examinons de manière globale la nature de l'interaction entre les traitements bas-haut et haut-bas, nous en venons au phénomène général d'appariement cognitif, et ainsi de reconnaissance. La problématique de l'appariement (« matching ») n'est pas neuve et on la retrouve sous diverses formes dans différentes propositions théoriques relatives, d'une part, à la perception visuelle (Bruyer, 1994, p. 188) et auditive (Perez, 1994, p. 206) et, d'autre part, à la compréhension linguistique et non linguistique. Il ne s'agit finalement, en simplifiant de manière exagérée, que de l'appariement, plus ou moins réussi, entre les stimuli externes et les connaissances préalables du sujet, appariement qui donne lieu, ou pas, à un phénomène de reconnaissance et donc d'identification, dans des zones qui, d'un point de vue cérébral, sont de mieux en mieux identifiées : « different brain structures that may serve to maintain a record of experience in a particular sensory domain, a record that serves as an internal standard for judgement » (Wilson, 1987, p. 413). On comprend donc que le processus d'appariement soit au cœur de l'interaction entre processus ascendants et descendants.

Si l'on peut appréhender le processus d'appariement selon la modalité (visuelle, auditive), selon le type de représentation (iconique, verbal), ainsi que selon les combinaisons et déroulements dans le temps (séquences et enchaînements) des stimuli

utilisés, il est utile, pour analyser, sous un certain angle, la manière dont il s'effectue, de faire notamment appel à la notion de perception catégorielle (Harnad, 1987 ; Handschin, 1994 ; Klose, 1994) et aux nombreux travaux relatifs, de près ou de loin, à la catégorisation cognitive (Dubois, 1990, 1991a ; Klir, 1996, pp. 133-136, par exemple) avec les termes-clés qui conviennent (prototypes, typicalité, etc.).

Le lien entre *appariement* et *catégorisation* apparaît clairement à la lecture de ce qu'écrit Kleiber (1991, p. 109) : « c'est un principe d'appariement qui est à la base de l'opération de catégorisation et non une vérification de CNS [(Conditions Nécessaires et Suffisantes)] : « un prototype est une instance typique d'une catégorie et les autres éléments sont assimilés à cette catégorie sur la base de leur ressemblance perçue avec le prototype » (R. W. Langacker, 1987, p. 371) ». Catégoriser présuppose donc d'effectuer un certain appariement entre faisceaux de traits perceptifs et catégories cognitives, tandis que l'appariement s'effectue à l'aide de catégories : items et prototypes, selon un gradient de typicalité. Il s'agit donc d'une activité constructive, qui s'effectue en permanence et qui permet de discrétiser le flux apparemment continu des stimuli en provenance de notre environnement : « l'activité cognitive individuelle, au niveau psychologique, est elle aussi une activité constante de catégorisation, et non pas une simple identification et reconnaissance d'objets pré-existants » (Mondada et Dubois, 1995, p. 287).

Cette reconstruction permanente s'agence autour de schémas et de figures prototypiques : « schémas et prototypes résultent de la capacité de notre système nerveux à détecter les corrélations entre traits, servant à abstraire des catégories, à détecter les écarts à ces catégories et, de façon plus générale, à anticiper les variations de l'environnement » (G. Tiberghien, 1994, p. 268). Ainsi, « le système cognitif construirait à l'aide de prototypes des invariants psychologiques qui donneraient une stabilité aux interprétations que les humains font du monde » (Mondada et Dubois, 1995, p. 292). Il est cependant important de noter que la plupart de ces catégories ne s'excluent pas strictement et qu'il existe des zones de chevauchement et d'incertitude : « La plupart des catégories que nous rencontrons et utilisons quotidiennement ont des

frontières vagues »<sup>100</sup> (Klir, 1996, p. 136). C'est d'ailleurs ce qui rend la théorie des ensembles flous si utile, puisque (*ibid.*) :

*« Les méthodes floues nous permettent de décrire la transition d'une adéquation parfaite à une incompatibilité totale de manière graduelle, plutôt qu'abrupte. C'est-à-dire que l'adéquation pour chaque trait est exprimée en termes de degré, et ces degrés sont correctement agrégés à l'aide opérations de moyennage (éventuellement avec des coefficients) pour obtenir une adéquation d'ensemble. La classification est alors basée sur le plus haut degré d'adéquation globale »<sup>101</sup>.*

Ainsi, dans le domaine phonético-phonologique (Repp et Liberman, 1987), la présentation graphique du système phonologique vocalique de telle ou telle langue en fonction des deux premiers formants (analyse acoustique des fréquences) correspond à la présentation de valeurs moyennes, chaque voyelle phonologique se situant à l'intérieur de son champ de dispersion phonétique (Renard, 1979, p. 22).

Dans une approche catégorielle de la reconnaissance, il faut alors distinguer *reconnaissance catégorielle* (c'est la catégorie à laquelle appartient l'item qui est reconnue, en anglais : « type ») et *reconnaissance individuelle* (c'est l'exemplaire individuel qu'il s'agit de reconnaître, en anglais : « token »), et aussi considérer les deux cadres théoriques actuellement dominants : celui qui considère la reconnaissance comme l'appariement entre représentation épisodique de l'élément perçu et représentation stockée en mémoire, et celui, d'obédience connexionniste, dans lequel la perception de l'objet conduit le système à tendre vers un état qui correspond, ou non, à celui auquel a été précédemment associé l'item correspondant (Bruyer, 1994, p. 182 et p. 197). Nous pourrions aussi aborder le concept d'appariement à travers la suggestion

---

<sup>100</sup> « Most categories we commonly encounter and employ have vague boundaries » (notre traduction).

<sup>101</sup> « Fuzzy methods allow us to describe the transition from a perfect match to a total mismatch as gradual rather than abrupt. That is, the match in each feature is expressed in terms of a degree, and these degrees are properly aggregated by some averaging operation (possibly weighted) to obtain an overall match. The classification is then based on the highest overall degree match » (notre traduction):

de Changeux et Dehaene (cités par Laks, 1996, p. 53) qui, de notre point de vue, y renvoie :

*« Le schéma traitement de l'information/input-output doit être abandonné au profit d'un « générateur » interne de variations, qui anticipe et teste continuellement le monde extérieur en référence à sa propre représentation invariante du monde ».*

Cependant, le concept d'appariement n'est pas purement et simplement un concept associationniste, appariant états internes et externes. Il y a en général tout un travail de traitement et de reconstruction qui est à l'œuvre. C'est ce que souligne aussi Dubois lorsqu'elle écrit (1991b, p. 23) :

*« La perception elle-même est (trop) souvent envisagée comme activité contemplative, un processus essentiellement « bottom-up » sous la dépendance des seuls filtres sensoriels, ou a contrario comme un « simple » appariement à des représentations préalables (« pattern matching »). [...] Des résultats de recherches actuelles [...] introduisent l'idée de sélectivité des activités perceptives humaines, en liaison avec les finalités ».*

L'importance des finalités de l'activité du sujet dans la constitution des catégories perceptives est aussi reprise par Mazet (1991, p. 100) qui écrit :

*« Il ne s'agit pas de minimiser la portée des catégorisations perceptives, mais au contraire d'en extraire leur caractère éminemment fonctionnel, en les insérant dans une approche plus globale considérant la finalité des activités. [...] Nous avons commencé de montrer que les fondements des structures catégorielles n'étaient plus assurés par les seules propriétés « intrinsèques » du monde (les corrélats d'attributs perceptifs), mais par les régularités des comportements en œuvre dans une activité déterminée ».*

Citant alors Dubois, Mazet poursuit : « ce n'est plus la perception comme processus « reflétant » des structures du monde « réel », mais les activités finalisées des sujets (ou opérateurs) qui doivent être pris en compte comme principe organisateur des catégories » (*ibid.*). Perception et action sont donc ici étroitement associées<sup>102</sup>, mettant de nouveau en avant l'activité du sujet comme facteur de détermination perceptive. Dans le cadre de la didactique des langues, la perception d'énoncés en langue étrangère se voit donc contrainte, non seulement par leurs attributs perceptifs (aspects structurels de la langue étrangère), mais aussi par le caractère plus ou moins routinier des situations dans lesquelles se trouvent les apprenants et des tâches que ces derniers sont censés accomplir.

Il est également possible d'aborder cette question d'un point de vue neuropsycholinguistique, en considérant notamment les propositions de Goldberg et Costa (1981). Ces chercheurs distinguent deux types de processus cognitifs : « Ceux qui ont recours à des codes préexistants et ceux qui n'y ont pas recours »<sup>103</sup> (p. 152). Confronté à un ensemble de stimuli de type sémiotique, chaque individu dispose en effet de connaissances lui permettant ou non d'appréhender ces stimuli en se référant, comme nous l'avons vu, aux dits codes/savoirs préexistants. Dans le cas de l'apprenant de langue étrangère, l'intégration de stimuli verbaux en L2 pose problème, puisque les apprenants ne disposent, initialement, que de leur compétence langagière en L1, laquelle renvoie à un certain « code préexistant ». Or les études en neuropsycholinguistique semblent indiquer que, de manière générale, « l'hémisphère gauche a une plus grande disposition à utiliser l'information précédemment apprise, tandis que l'hémisphère droit a tendance à appréhender chaque tâche comme une expérience nouvelle »<sup>104</sup> (Goldberg et Costa, 1981, p. 153). Considérant toute tâche à accomplir en langue étrangère, en début d'apprentissage, comme une tâche nouvelle,

---

<sup>102</sup> Il serait possible d'établir ici un lien, aussi ténu soit-il, avec, dans le domaine phonético-phonologique, la théorie motrice de la perception de la parole.

<sup>103</sup> « Those which draw on preexisting codes and those which do not » (notre traduction).

<sup>104</sup> « The left hemisphere has a greater facility for the utilization of previously learned information, while the right hemisphere tends to approach every task as a novel experience » (notre traduction).

avec un input peu familier, on peut envisager les situations d'enseignement/apprentissage à travers le découpage opéré par Goldberg et Costa (p. 154) :

- « *Orientation dans une tâche nouvelle, lorsqu'aucun système descriptif n'apparaît de manière immédiate, mais la tâche est cependant finalement reconnue comme se rapportant à un système préexistant. L'hypothèse prédit que l'hémisphère droit est essentiel pour l'orientation initiale, et que l'hémisphère gauche joue un rôle supérieur dans l'utilisation d'un système descriptif déjà établi, une fois ce dernier identifié ;*
- *Orientation dans une nouvelle tâche pour laquelle aucun des systèmes descriptifs disponibles ne peut être utilisé de manière réussie. Dans ce cas, un nouveau système descriptif, pertinent pour la tâche cognitive, doit être « assemblé ». L'hypothèse prédit que la participation de l'hémisphère droit est essentielle tant pour l'étape d'orientation que pour celle de l'assemblage. Une fois qu'un système approprié a été assemblé, l'hémisphère gauche prend, à terme, la direction de son utilisation »<sup>105</sup>.*

C'est donc bien ici le « concept de nouveauté de la tâche par rapport aux codes représentationnels préexistant dans le répertoire cognitif d'un sujet donné »<sup>106</sup> (Goldberg et Costa, 1981, p. 145) qui est mis en avant.

---

<sup>105</sup> « Orientation in a novel task, when no descriptive system is immediately apparent, but the task is ultimately recognized as relevant to an existing one. The hypothesis predicts that the right hemisphere is essential for the initial orientation, and that the left hemisphere is superior in the utilization of an existing descriptive system once it is identified. Orientation in a novel task to which none of the available descriptive systems can be successfully applied. In this case, a new descriptive system which is relevant to the cognitive task must be « assembled ». The hypothesis predicts that right-hemispherical participation is essential for both the orientation and assembly stages. Once an appropriate system has been assembled, the left hemisphere should achieve a leading role in its utilization » (notre traduction).

<sup>106</sup> « Invokes the concept of task novelty with respect to the representational codes preexisting in a given subject's cognitive repertoire » (notre traduction).



La conclusion d'une étude sur la gestion des inférences chez les cérébro-lésés droits (CLD) semble en parfait accord avec les propositions de Goldberg et Costa (Duchêne, 1999, p. 275) :

*« En situation de communication, les CLD ne cherchent pas à comprendre les propos du locuteur mais bien plutôt à s'approprier le message en le réajustant éventuellement, pour qu'il soit « en phase » avec leur propre représentation du monde, qui elle est première et toute puissante ».*

Comme l'indique l'auteur (p. 274) :

*« Les patients CLD manifestent ici la toute puissance de leurs schémas prédéterminés d'aménagement du territoire. Le poids des mots ne suffit pas à les détourner des sentiers battus. Une force d'inertie les maintient dans leur direction initiale et classe au rang d'impertinence toute interprétation innovante ».*

Ceci confirme pour nous l'idée générale selon laquelle l'hémisphère gauche, doté de ses connexions avec un hémisphère droit intègre, rigidifie et routinise les comportements interprétatifs, en ramenant les stimuli externes aux matrices préexistantes, quitte à les déformer dans un mécanisme d'appariement forcé. Il nous semble normal que cela apparaisse aussi clairement dans le traitement de la signification non littérale, puisque ce dernier repose sur des processus d'inférence fortement contraints par les connaissances préalables du sujet.

D'un point de vue didactique, ces hypothèses d'ordre neuropsycholinguistique nous semblent devoir être prises en compte lors de l'élaboration de l'input, et surtout de sa présentation aux apprenants, puisqu'elles constituent une réinterprétation du couple traditionnel analytique / global attaché, de manière souvent hâtive, aux hémisphères gauche et droit du cerveau humain. En fonction des systèmes descriptifs possédés par, et ainsi disponibles pour, les apprenants, le mode de présentation de l'input devrait être

pensé en vue de faciliter son traitement dominant par l'hémisphère approprié. Comme l'écrivent Gordon et Weide (1983, p. 50) :

*« Les deux hémisphères traitent les stimuli linguistiques selon leurs modes de traitement cognitif inhérents. Si la tâche favorise une stratégie de traitement séquentielle, analytique et phonétique, l'hémisphère gauche domine. L'hémisphère droit domine si la tâche favorise soit l'identification de configurations, soit la mémoire ».*

Soulignons en outre, de manière plus générale, le renouvellement constant des cadres catégoriels qui structurent nos savoirs et nos actions, celui-ci étant notamment dû à notre activité discursive (Vignaux, 1991, p. 310) :

*« Le processus de la schématisation discursive est donc indissociable d'une catégorisation permanente—réitérée de discours à discours—des objets du monde dans une sorte de « relecture » incessante de leurs propriétés, et dans une non moins incessante réaffectation de leurs "indexations" à l'intérieur de catégories de faits, d'événements, de situations ou de "mondes" ».*

Ces cadres sont donc évolutifs, et aussi, d'une certaine manière, en « auto-perpétuation ». C'est d'ailleurs ainsi que certains chercheurs appellent ce qui constitue à leurs yeux un mécanisme fondamental d'apprentissage implicite, le « mécanisme d'autoperpétuation d'encodage », mécanisme « par lequel nous avons tendance, lorsque nous sommes confrontés à des stimuli ambigus, à leur imposer des catégories préexistantes » (Meulemans, 1998, p. 115).

En se plaçant, par exemple, dans le cadre d'une classe d'enseignement de la physique, au cours de laquelle les élèves doivent acquérir de nouveaux concepts, Collet note « la présence d'une activité métalinguistique de reformulation dans toutes les catégories cognitives » (2000, p. 31), qui s'exprime sous trois formes : « recherche de synonymes

ou parasyonymes, la paraphrase conduisant à la décomposition de termes complexes en termes constituants, et la construction de phrases explorant conjointement les propriétés physiques ou fonctionnelles des objets et les propriétés proprement linguistiques des termes du modèle » (Collet, 2000, p. 32). On pourrait considérer ceci comme un processus de déconstruction/recomposition visant à « traduire » « l'inconnu »/l'ambigu en « connu »/catégories préexistantes, de manière à ce que l'appariement soit réussi et qu'il y ait intégration dans l'ancien système de connaissances du sujet, dès lors forcément recomposé.

D'un point de vue didactique, pour l'apprenant de langue étrangère, ces processus de *catégorisation* et *d'appariement*, s'ils peuvent parfois se révéler utiles (cas de transferts positifs), s'avèrent la plupart du temps néfastes (cas de transferts négatifs), en raison du degré de similarité établi par le sujet entre les items présentés et les éléments prototypiques de ses catégories préexistantes, qu'elles soient phonético-phonologiques ou lexico-sémantiques, par exemple.

#### 1.1.2.2.2. Modèles de reconnaissance

La plupart des travaux ayant porté sur la reconnaissance, ou l'identification, de mots parlés isolés, c'est sur ceux-ci que nous nous concentrerons. L'identification d'un mot peut dans un premier temps être envisagé comme « une opération d'appariement entre un code sensoriel issu de l'analyse du signal et une entrée du lexique mental » (Ségui, 1997, p. 17). En ce qui concerne les différentes voies d'accès à l'item lexical recherché, ainsi que le rôle des stimuli visuels et auditifs, plusieurs propositions ont été avancées (Bruyer, 1994, p. 196 ; Valdois et Nespoulous, 1994, p. 372). Mais comme l'indique Ségui (2003, p. 224), les deux questions essentielles à propos de l'identification des mots isolés sont les suivantes :

1. Quelle est la nature des informations utilisées dans cette tâche par le sujet ?
2. Quel est le rôle des connaissances lexicales du sujet dans le processus d'identification (facilitation et accélération) ?

Dans le domaine de la reconnaissance des mots parlés, deux modèles en particulier tentent de rendre compte des phénomènes à l'œuvre, dont Ségui résume comme suit les grandes lignes (Ségui, 2003, p. 225 ; 1997, p. 18).

#### I.1.2.2.2.1. Le modèle COHORTE

Le modèle *COHORTE* proposé par Marslen-Wilson et ses collègues suppose une activation de candidats lexicaux possibles (la cohorte du mot cible), suivie de la sélection du candidat correspondant au stimulus présenté, grâce à un processus d'élimination progressive des candidats incompatibles avec le stimulus. L'hypothèse centrale de ce modèle est la suivante : l'identification d'un mot aboutit lorsque l'un des candidats lexicaux finit par être le seul compatible avec la séquence sonore perçue. Le point à partir duquel un mot peut être identifié en théorie est appelé *point d'identification*. Ségui (2003, p. 225) donne l'exemple du mot « vocabulaire » pour lequel le point d'identification correspond au phonème associé à la lettre « u », puisque aucun autre mot français ne possède la séquence initiale « vocabu ». Ainsi, l'ensemble des unités lexicales partageant la séquence sonore initiale du mot cible est appelé la *cohorte* de ce mot. Ce processus est modifié lorsque le mot cible se trouve dans un contexte phrastique, puisque le contexte peut accélérer le processus d'élimination des candidats *via* certaines contraintes syntaxiques ou interprétatives : « le point d'identification d'un mot varie donc selon le contexte de sa réalisation » (Ségui, 2003, p. 231). Cependant, l'influence du contexte ne peut se réaliser qu'une fois la cohorte activée : la primauté reste donc à l'information sensorielle, qui domine l'information contextuelle. Hors-contexte, ce modèle confère donc néanmoins une importance certaine aux premiers sons du mot en question, ce qui en fait un modèle largement orienté. Cette orientation ne permet alors pas de rendre facilement compte de l'identification des mots à partir de leur partie finale (Ségui, 2003, p. 226). D'autres modèles, moins orientés, sont alors requis, comme le modèle TRACE. Les quatre principes sous-jacents au modèle COHORTE sont donc résumés comme tels par McClelland et Elman (1986, p. 98) :

1. Le modèle utilise le premier son du mot, afin de déterminer la cohorte initiale du mot ;

2. Une fois l'ensemble des candidats établi, leur élimination s'effectue progressivement, au fur et à mesure de la comparaison avec les phonèmes suivants. Dans certains cas, des contraintes sémantiques peuvent également être à l'œuvre.
3. La reconnaissance du mot s'accomplit dès qu'il ne reste plus qu'un candidat dans la cohorte.
4. La reconnaissance lexicale ne peut influencer l'identification des phonèmes dans un mot qu'une fois celui-ci reconnu.

#### 1.1.2.2.2. Le modèle TRACE

Le modèle TRACE de McClelland et Elman ressemble partiellement à ce modèle, puisqu'il s'en est inspiré en cherchant à en combler les faiblesses (McClelland et Elman, 1986, p. 98), mais il propose un processus de compétition entre candidats, plutôt que de sélection, avec un phénomène de feed-back lexical de type haut-bas. L'une des différences essentielles entre les deux modèles repose dans la nature interactive du second, avec différents niveaux d'unités : traits, phonèmes et mots. Toutes ces unités sont interconnectées par des liens, soit d'activation, soit d'inhibition. Ainsi lorsque l'unité-phonème /s/ est activée (au delà d'un certain seuil), celle-ci active les unités-mots qui la contiennent et inhibe les autres unités-phonèmes. Dans ce modèle, l'identification d'un mot aboutit lorsque l'unité-mot qui lui correspond atteint un certain seuil d'excitation. C'est cette interactivité qui permet au modèle de rendre compte d'influences lexicales de type haut-bas.

Si ce modèle n'est pas à l'abri de certaines faiblesses, ses performances peuvent cependant être comparées aux performances humaines sur les points suivants, entre autres (McClelland et Elman, 1986, p. 117) :

- Utilisation d'informations qui se chevauchent dans le flux linguistique ;
- Tendance à la perception catégorielle des phonèmes ;
- Intégration d'informations lexicales haut-bas pour l'identification des phonèmes ;
- Connaissances apparente des régularités phonotactiques influant sur l'identification phonémique ;

- Adaptation à l'influence de la coarticulation sur les propriétés acoustiques du phonème.

L'interactivité est ainsi l'un des concepts-clés de l'approche tenue par ses défenseurs, comme cela est illustré dès le premier paragraphe de McClelland et Elman (1986, p. 58) à propos de la perception du phonème /g/ dans la phrase « She received a valuable gift » :

*« There are a large number of cues in this sentence to the identity of this phoneme [...], acoustic cues [...], the other phonemes in the same word [...], the semantic and syntactic context [...]. [A]ll of these different sources of information are used in recognizing words and the phonemes they contain ».*

Or, il est à présent bien établi que c'est précisément l'interactivité entre les éléments des différents niveaux d'analyse et l'imbrication étroite entre les différents contextes qui permet de rendre compte de nombreux effets (Lambert, 1997, p. 157) :

*« L'appréhension dynamique du traitement de la parole, et tout particulièrement la notion d'interaction entre les divers niveaux de représentations (sublexicaux et lexicaux), peut donc rendre compte du fait qu'un sujet peut accéder à la représentation lexicale adéquate même si l'information correspondant aux traits ou aux phonèmes de ce mot a été incorrectement analysée. Cette interaction n'est pas limitée entre unités traits, phonèmes, mots. De nombreux travaux en psychologie expérimentale et en neuropsychologie attestent d'effets sémantiques, syntaxiques ou pragmatiques sur la reconnaissance des mots et sur la perception phonémique ».*

Dans ce sens, la vision strictement modulaire de la « grammaire » intériorisée des locuteurs, qui serait constituée de composants syntaxiques, phonologiques, etc., indépendants semble se trouver également remise en question par la théorie de

l'optimalité (OT). Dans celle-ci, l'approche est basée sur des contraintes qui peuvent, pour une langue donnée être classées au sein d'une même hiérarchie interne, qu'elles soient syntaxiques, morphologiques, phonologiques, phonétiques ou sémantiques, ce qui implique de possibles interactions entre lesdits modules. Ainsi, certaines contraintes syntaxiques pourraient être violées afin de satisfaire d'autres contraintes phonologiques par exemple (Archangeli, 1997, p. 30).

Néanmoins, malgré des résultats favorables aux modèles interactifs, Ségui (2003, p. 231) souligne que la question de la nature modulaire ou interactive des procédures d'accès lexical reste encore largement ouverte, car si certains effets de type contextuels sont constatés, il reste difficile de déterminer si ceux-ci se situent à un niveau pré-lexical (perceptif) ou post-lexical (ce qui correspond à un biais de réponse chez les sujets).

D'autres modèles ont ainsi été élaborés, certains mettant l'accent sur les relations de voisinage acoustico-phonétique globales entre le stimulus et l'ensemble des unités lexicales qui lui ressemblent. Dans cette approche, tant la fréquence d'usage dans la langue (qui est liée à la disponibilité lexicale) que le nombre des mots phonétiquement voisins du stimulus semblent jouer un rôle essentiel (Ségui, 2003, p. 228).

Enfin, il faut souligner que bien d'autres facteurs interviennent dans la perception de la parole, au-delà de la reconnaissance lexicale, en particulier les facteurs prosodiques qui sont liés aux propriétés syntaxiques et sémantiques des énoncés et qui permettent la construction en temps réel d'une représentation intégrée des énoncés (Ségui, 2003, p. 233).

## **I.2. Système mnésique et connaissances**

Si nous avons jusqu'à présent assez librement utilisé le terme *reconnaissance*, il va de soi que celui-ci présuppose l'existence d'une certaine *connaissance*. L'analyse de l'une ne peut donc être que difficilement conduite sans celle de l'autre. Or, du point de vue des sciences cognitives, au moins, la connaissance humaine repose fondamentalement

sur le système mnésique, au sens large, dont sont dotés les individus et sans lequel celle-ci ne serait qu'instantanée et fugace.

*Apprendre* signifiant acquérir des connaissances, qu'elles soient procédurales (en anglais *to know how*) ou déclaratives (en anglais *to know that*), l'apprentissage implique essentiellement, sinon exclusivement (tout dépend de la définition des notions employées ici), un ensemble de processus de mémorisation. Boogards, par exemple, considère qu'« apprendre une langue étrangère, c'est stocker des informations dans la mémoire sémantique ; dans ce stockage interviennent deux phénomènes : d'une part, la création de relations entre des concepts (déjà existants ou nouvellement créés) et des éléments de la langue à apprendre ; d'autre part, la mise en place de séquences de nœuds en vue de la création d'automatismes » (1988, p. 30). Quelle que soit l'approche que l'on adopte, l'apprentissage d'une langue étrangère est donc foncièrement dépendant de l'organisation et du fonctionnement du système mnésique.

Dans une perspective didactique, *faire-mémoriser*, *faire-apprendre*, *faire-intégrer*, sont, en définitive, trois expressions similaires qui renvoient toutes à l'un des objectifs centraux de l'enseignement. L'une des questions-clés qui harcèlent les enseignants et les méthodologues est alors la suivante : comment favoriser la mémorisation des éléments de la L2 par les apprenants ? Il est par exemple établi que « la récupération d'une information en mémoire est d'autant meilleure que cette information a été codée sémantiquement » (Van der Linden, 1994, p. 289). Mais répondre plus en détail à cette question implique de connaître, autant que possible, le fonctionnement de l'appareil mnésique humain, particulièrement dans sa relation au langage.

### **I.2.1. Système mnésique**

Etant donné l'étendue du champ d'étude relatif à la mémoire, nous nous concentrerons ici, après un bref rappel des conceptions actuelles en psychologie cognitive, sur l'un de ses aspects les plus pertinents pour notre étude, à savoir la mémoire de travail, et son implication dans l'apprentissage de nouvelles formes phonologiques.



### 1.2.1.1. Organisation de la mémoire : perspectives

Nous présentons d'abord quelques concepts généraux relatifs à l'étude de la mémoire, avant d'examiner plus en détail certaines tentatives de modélisation de son système.

#### 1.2.1.1.1. Concepts généraux

En raison de nombreuses avancées en sciences cognitives durant ces dernières décennies, les conceptions de la mémoire ont beaucoup évolué, si bien qu'elle se trouve aujourd'hui au cœur même de toute problématique d'apprentissage, puisque « la » mémoire humaine est désormais considérée à la fois comme connaissance des événements du passé, mais aussi comme cadre cognitif déterminant notre « présent perceptif » et nos nouvelles acquisitions : « Les propriétés du système de mémoire déterminent largement, sinon entièrement, les capacités générales du système cognitif », écrit à ce propos G. Tiberghien (1997, p. 140).

Deux précisions essentielles doivent être fournies lorsque l'on aborde le fonctionnement mnésique : *le type d'information* qui doit être mémorisé, et *la tâche*, dans laquelle va être impliquée la mémoire. En ce qui concerne la première, il convient en outre de distinguer *la nature* de cette information (verbale ou non, concrète ou abstraite, très ou peu imageable, etc.) de *la modalité sensorielle* (auditive, visuelle ou tactile, unique ou multiple) par laquelle le sujet y accède. Nous pouvons donc donner quelques exemples d'ordre linguistique très généraux :

⇒ des éléments externes :

- des règles, ou relations, entre forme et fonction ;
- des items lexicaux ;
- des chaînes graphiques ;
- des chaînes sonores ;
- des images ;
- des « connaissances » d'ordre sémantique ;

⇒ des conduites internes :

- stratégies (palliatives, attentionnelles ou autres) ;
- conduites (articulatoires, gestuelles ou autres).

En ce qui concerne la tâche, celle-ci constitue le socle sur lequel sont conduites toutes les recherches en psychologie de la mémoire, la variation des résultats en étant fortement tributaire. Il existe ainsi deux types majeurs de méthode d'étude de la mémoire, issue d'un même paradigme, schématisé par G. Tiberghien, (1997, p. 142) qui suit le déroulement suivant :

- Phase d'encodage et de stockage pendant laquelle une information d'origine perceptive est transformée en une représentation mentale plus ou moins stable et plus ou moins fortement associée à d'autres représentations ;
- Phase de réactivation de ces représentations mentales ;
- Phase de rétention qui sépare les deux phases précédentes.

Les méthodes *directes* placent l'individu dans une situation où il doit se souvenir intentionnellement d'un épisode antérieur de son existence, tandis que les méthodes *indirectes* demandent au sujet d'effectuer une tâche cognitive ou cognitivo-motrice particulière, qui peut être facilitée si elle a déjà été réalisée antérieurement. L'économie au ré-apprentissage (en anglais *savings*) et les effets d'amorçage (en anglais *priming*), sémantique ou de répétition, sont les deux techniques principalement employées dans ces dernières.

Si apprendre consiste, partiellement au moins, à mémoriser, c'est, à terme, pour pouvoir *reconnaître* ou *rappeler* les éléments stockés. La différence entre les deux tâches est ainsi résumée par G. Tiberghien (1997, p. 142) : « dans le rappel un contexte particulier constitue une clé d'accès à une *cible* alors que c'est l'inverse dans la reconnaissance, une cible sert de clé d'accès à un contexte spécifique ». Baddeley (1986, p. 51) indique également que la récupération en mémoire (en anglais *retrieval*) recouvre deux processus séparés : l'un semble impliquer un accès automatique à la

trace mnésique appropriée aux indices perceptifs disponibles, l'autre correspond davantage à une stratégie consciente de récupération (ce qui a parfois été nommé *recollection* en anglais) utilisant un ensemble d'indices potentiellement utiles qui sont ensuite évalués. Baddeley illustre ce dernier par la récupération du nom d'une personne *via* les lettres de l'alphabet ou, dans le cas d'un homme, le nom de son épouse.

En ce qui concerne le contexte, il convient de distinguer avec Baddeley (cité par Van der Linden, 1994, p. 290) :

- Le « contexte indépendant » : c'est-à-dire les éléments « qui restent à la périphérie de l'attention durant l'expérience d'apprentissage, qui n'influencent pas l'interprétation de l'événement-cible » ;
- Le « contexte interactif » : c'est-à-dire « les informations qui constituent l'arrière-plan de l'information-cible » et qui affectent sa signification.

Il apparaît alors que le rappel dépend davantage du contexte indépendant que la reconnaissance, alors que cette dernière est particulièrement tributaire du contexte interactif, qui détermine les caractéristiques de l'information à encoder (Van der Linden, 1994, p. 291).

Plusieurs dichotomies méthodologiques telles que celles qui précèdent sont ainsi bien établies, et à l'entrée « Mémoire » du glossaire d'un ouvrage collectif consacré aux Sciences Cognitives (Gordon et Paugam-Moisy, 1997, p. 254) on trouve :

*« Ensemble d'informations codées au sein d'un système cognitif (ou informatique, génétique, immunitaire, etc.) et structurées de façon à permettre leur stockage et leur rappel. On distingue les processus de maintien des représentations à court terme (limité en capacité et en temps), et les processus de stockage à long terme. Les spécialistes de la mémoire introduisent des distinctions entre mémoire associative, de travail, tampon, déclarative, procédurale, etc. Dans les modèles connexionnistes la mémoire est codée dans*

*les poids des connexions, sous forme distribuée, par opposition aux systèmes informatiques de stockage par adresse. Les états mémorisés sont les états des neurones ».*

Ainsi, comme cela apparaît dans cette brève présentation, les modèles fonctionnels récents de la mémoire, établis ou non à partir de cas de dissociations observés en pathologie du langage (G. Tiberghien, 1994, p. 273), peuvent être regroupés selon leur classe : *computo-symboliques* ou *connexionnistes*.

#### 1.2.1.1.2. Modélisation

Nous rappelons brièvement ici les grands traits des principaux modèles mnésiques.

##### 1.2.1.1.2.1. Modèle Modal

Comme l'écrit Baddeley (1986), il y avait, à la fin des années 60, une pléthore de modèles de la mémoire, distinguant mémoire à court-terme et à long-terme. Les modèles de mémoire à court-terme partageaient cependant tous essentiellement les mêmes caractéristiques, principalement issues des travaux de Atkinson et Shiffrin, eux-mêmes basés sur ceux de Broadbent (Baddeley, 1986, p. 16). Il s'agissait alors du modèle dit *modal*, dont les trois composants étaient les suivants :

- 1) Un ensemble d'unités-tampons périphériques de stockage sensoriel (« a bank of sensory buffer stores »), capables de stocker temporairement des informations de diverses modalités (une modalité par unité).
- 2) Une unité à capacité limitée de stockage à court terme (*short-term store, STS*), alimentée par les unités-tampons précédentes, et capable de fournir et de récupérer de l'information du troisième composant.
- 3) Une unité à grande capacité de stockage à long-terme (*long-term store, LTS*).

L'information ne pouvait donc être stockée en mémoire à long-terme qu'après avoir été traitée en mémoire à court-terme. La caractéristique centrale de ce modèle résidait en effet dans le second composant (STS), dont était dépendant le troisième (LTS)

(l'information devait transiter et être maintenue dans le STS avant de pouvoir être transférée dans le LTS, indiquant ainsi que l'apprentissage dépendait, de manière cruciale, de la durée de maintien dans le STS)<sup>107</sup>. Loin d'être une unité passive, le STS était censé être responsable de l'encodage mnésique de l'information, quelle qu'en soit la manière (par exemple établir des liens associatifs) (Baddeley, 1986, p. 16).

Cependant, le modèle ne semblait pas pouvoir rendre compte des données expérimentales et cliniques de plus en plus nombreuses, en particulier celles relatives au rapport entre STS et apprentissage à long-terme : plusieurs études avaient en effet remis en cause la relation de dépendance entre maintien de l'information en STS et apprentissage, à long-terme, de celle-ci. Ainsi, le maintien d'un item dans le STS ne semblait pas systématiquement déboucher sur un apprentissage à long-terme (Baddeley, 1986, p. 19) :

#### 1.2.1.1.2.2. Les Niveaux de Traitement

Le cadre théorique des Niveaux de Traitement (NT) (du niveau sensoriel périphérique jusqu'à un niveau sémantique profond), qui a succédé au modèle modal, suggérait qu'à chaque niveau de traitement une trace mnésique était laissée, trace dont la durabilité augmentait avec la profondeur. Dans ce cadre, deux types de processus de répétition (*rehearsal*) étaient envisagés :

- 1) Répétition de *maintenance* (à l'intérieur d'un niveau unique) (*maintenance rehearsal*) ;
- 2) Répétition *élaborative* (chaque traitement augmente la profondeur de l'encodage) (*elaborative rehearsal*).

L'une des hypothèses centrales de ce cadre était que l'apprentissage à long-terme ne pouvait s'effectuer que grâce au processus de répétition élaborative. Cependant

---

<sup>107</sup> D'autres ont également proposé que ce n'était pas tant de la durée de maintien en STS, que du nombre d'entrées de l'item informatif concerné dans celui-ci que dépendait la probabilité de son stockage en LTS.

l'hypothèse des NT s'est trouvée remise en question par plusieurs résultats (Baddeley, 1986, pp. 24-25), notamment :

- 1) Le rapport entre répétition de maintenance et apprentissage à long-terme : il semblerait que ce type de répétition ait une influence, même minime, sur la rétention à long-terme d'information.
- 2) Le rapport entre traitement de surface et apprentissage durable : il semblerait que l'encodage d'un item selon ses traits superficiels (« superficial features ») puisse donner lieu à un apprentissage durable, contrairement à l'hypothèse selon laquelle un tel encodage ne laisserait qu'une trace mnésique qui se dissiperait très rapidement.
- 3) La notion de continuum de niveaux de profondeur : le cadre NT proposait en effet qu'à chaque niveau d'encodage correspondait un degré de mémorabilité. Si cela a pu se vérifier entre domaines différents (par exemple, un mot était moins bien mémorisé lorsqu'il était traité respectivement de manière visuelle, phonologique ou sémantique), de telles différences ne semblent pas avoir été constatées au sein d'un même domaine. Comme l'indique Baddeley (1986, p. 26), cela suggère que le concept de profondeur, qui implique une certaine continuité, est moins à même de rendre compte des données que celui de domaine de traitement, qui relève du discret.
- 4) Le caractère séquentiel linéaire du traitement : contrairement à l'approche classique de traitement linéaire du cadre NT, il est depuis apparu que la perception était hautement interactive (traitement parallèle et simultané de plusieurs dimensions d'un même stimulus, ainsi qu'intégration des processus bas-haut et haut-bas).

Le cadre NT a donc lui aussi dû se confronter à ses faiblesses explicatives, donnant lieu à de nouvelles hypothèses et de nouveaux développements, particulièrement axés sur la distinction entre mémoire à long-terme et mémoire à court-terme, celle-ci conduisant les chercheurs à élaborer et à modéliser le concept de mémoire de travail.

### I.2.1.1.2.3. Composantes de la mémoire

Si ces modèles ont ainsi connu de nombreuses évolutions, ils constituent toutefois la base des conceptions actuelles du système mnésique, et plusieurs concepts opératoires ont pu être dégagés et reconnus au fil des décennies. Ainsi l'observation de cas de doubles dissociations a contribué à remettre en cause le modèle « modal », dans lequel les mémoires à court terme et à long terme étaient systématiquement hiérarchiquement organisées, l'information devant nécessairement transiter par la mémoire à court terme. De même, la conception unitaire de cette dernière a été remise en cause par des études en neuropsychologie montrant une préservation des informations de type visuo-spatiale tout en présentant un déficit dans la rétention de matériel verbal (Van der Linden, 1994, p. 282). Ainsi, dans un ouvrage notamment destiné à un public étudiant, Kekenbosch (1994) présente le système mnémonique comme pouvant être subdivisé en trois composants, dans lesquels l'on peut retrouver, du moins en partie, les fondements du modèle modal présenté plus haut : le *registre d'information sensorielle*, la *mémoire à court terme* et la *mémoire à long terme*.

- Dans le premier (durée de maintien de l'information très brève) se trouvent la *mémoire iconique*, qui concerne le stockage visuel, et la *mémoire échoïque*, qui concerne le stockage auditif.
- Dans la seconde, se trouvent la *mémoire de travail* et la *mémoire tampon* (interface entre la précédente et le registre sensoriel).

La mémoire iconique semble impliquer : 1) une persistance visuelle (fonction de l'intensité lumineuse) et 2) une persistance informationnelle, qui serait composée d'une représentation analogue visuelle et d'une représentation d'identité non visuelle (G. Tiberghien, 1994, p. 259). En ce qui concerne la représentation d'identité, celle-ci renvoie aux phénomènes de catégorisation et aux représentations préexistantes. L'image ou la chaîne graphique sera mieux mémorisée si elle peut être facilement identifiée, et l'on sait que « la sensibilité au contraste du système visuel humain est

fonction de la fréquence spatiale [...] [et] que les composantes de basse fréquence sont détectées plus rapidement que celles de haute fréquence » (Bruyer, 1994, p. 184).

Au sein de la mémoire permanente, on distingue la *mémoire épisodique* (mémoire des expériences personnelles individuelles contextualisées) de la *mémoire sémantique* (connaissances généralisées décontextualisées mais interconnectées et possiblement hiérarchisées). Ce que partagent les représentations de ces deux dernières est la possibilité d'être verbalisées (caractère explicite), et elles s'opposent en ceci aux représentations procédurales de tous ordres (caractère implicite). Les différentes composantes du système mnésique pourraient ainsi être présentées hiérarchiquement comme suit (adapté de Fabbro, 1999, p. 97) :

MEMOIRE				
De travail	Permanente			
	<i>Explicite</i>		<i>Implicite</i>	
	Sémantique	Episodique	Conditionnement	Procédurale

Tandis que la mémoire explicite renvoie à un rappel volontaire et conscient qui peut être rapporté ouvertement, la mémoire implicite renvoie à des circonstances dans lesquelles une tâche ou un stimulus influence une performance future même lorsque le sujet n'est pas conscient d'avoir été exposé à un tel stimulus ou d'avoir effectué une telle tâche (Mesulam, 1998, p. 1026, note 7). D'un point de vue neurolinguistique, Mesulam suppose qu'il n'y a pas de différence fondamentale d'encodage entre les deux types de mémoire (p. 1027) :

*« In implicit memory, the information remains in the form of isolated fragments, mostly within unimodal and heteromodal association areas ; in explicit memory*



*it becomes incorporated into a coherent context through the binding function of limbic nodes ».*

Et comme il l'ajoute (p. 1029) :

*« The process of explicit memory represents one of the most fundamental means of incorporating sensory information into cognition. It allows each individual to construct a unique record of experience and knowledge based on events of personal significance ».*

#### I.2.1.1.2.4. Mémoire et apprentissage

Il est également possible d'envisager la modélisation du système mnésique *via* le rapport entre mémorisation et apprentissage, rapport qui a constitué le cœur de nombreux travaux en psychologie de la mémoire. Les différentes approches dans ce domaine sont représentées de manière très synthétique dans ce tableau récapitulatif des « relations entre mécanismes d'apprentissage et formalismes de description des unités symboliques élémentaires et de leur organisation en mémoire » proposé par G. Tiberghien (1994, p. 270) :

<b>Organisation des unités élémentaires en mémoire</b>	<b>Mécanisme d'apprentissage</b>
S – R [(Stimulus-Réponse)]	<ul style="list-style-type: none"> <li>▪ Association</li> <li>▪ Répétition</li> <li>▪ Modification de la force des associations</li> <li>▪ Généralisation</li> <li>▪ Différenciation</li> </ul>
Structure de liste	<ul style="list-style-type: none"> <li>▪ Création de propositions</li> <li>▪ Suppression de propositions (macro-structure)</li> <li>▪ Modification de la hiérarchie des propositions</li> </ul>

Réseau sémantique	<ul style="list-style-type: none"> <li>▪ Création de nœuds</li> <li>▪ Création d'associations étiquetées</li> <li>▪ Convergence et « fan effect »</li> <li>▪ Re-structuration</li> </ul>
Système de production	<ul style="list-style-type: none"> <li>▪ Création de règles</li> <li>▪ Modification des conditions (discrimination et généralisation)</li> <li>▪ Modification des actions (idem)</li> <li>▪ Compactage de règles (automatisation)</li> <li>▪ Décompactage de règles (analyse, prise de conscience)</li> </ul>
Schéma, Script	<ul style="list-style-type: none"> <li>▪ Extraction de propriétés</li> <li>▪ Catégorisation</li> <li>▪ Instanciation</li> <li>▪ Détection de nouveauté</li> <li>▪ Hiérarchisation des schémas</li> <li>▪ Emboîtement de schémas</li> </ul>

Comme cela apparaît nettement dans ce tableau, plusieurs modélisations de l'appareil mnésique humain, tant du point de vue du traitement que du stockage de l'information, ont donc été avancées durant les cinquante dernières années. Parmi elles, il en est une qui mérite une mention particulière, puisque, au cours des dernières décennies, c'est par elle que de nombreux chercheurs ont abordé la problématique en question, dans une approche relativement singulière du point de vue de la tradition symbolique de récupération par adressage (pour laquelle la mémoire était métaphoriquement associée à la figure de la « bibliothèque »): il s'agit de la modélisation connexionniste.

#### 1.2.1.1.2.5. Perspectives connexionnistes

Comme l'indiquent McClelland et Rumelhart (1986, p. 171), l'un des dilemmes posés aux théories de la mémoire était celui du choix entre représentations générales et

spécifiques de l'information. En effet, d'un côté la mémoire humaine semble reposer sur la formation de généralisations issues des détails d'expériences spécifiques, généralisations que l'on retrouve sous forme de concepts abstraits ou encore de prototypes (la représentation d'exemplaires typiques). D'un autre côté, néanmoins, de nombreuses études ont montré que la mémoire était particulièrement sensible aux spécificités de chaque expérience passée : ainsi en tâche de catégorisation, d'identification perceptive ou de prononciation, la congruence entre stimuli spécifiques d'entraînement et de test semble jouer un rôle non négligeable dans la performance (p. 171). En réponse à ce dilemme, certains modèles, comme le modèle d'activation interactive de perception des mots de McClelland, considèrent que les « règles linguistiques » apparentes émergent d'une conjugaison (*conspiracy*) d'activations partielles de traces mnésiques individuelles. Mais comme ils le notent (p. 172), cette approche, qui prétend se passer de règles abstraites, nécessite néanmoins un certain niveau d'abstraction, puisque chaque occurrence perceptive d'un même mot doit être associée au même détecteur de mots (ce qui présuppose, d'une certaine manière, l'existence de similarités entre ces occurrences et ainsi la constitution d'une certaine classe, plus abstraite). Si enfin, afin de pallier cette difficulté, l'on adopte un point de vue strictement exemplariste, en considérant que la représentation de chaque mot, ou objet de représentation, est issue d'une énumération spécifique d'expériences, c'est essentiellement un problème de capacité de stockage qui se pose aux théoriciens.

A ce dilemme, McClelland et Rumelhart tentent d'apporter une solution qui serait capable de rendre compte à la fois des données suggérant que nous stockons des représentations abstraites et de celles qui suggèrent que la mémoire consiste en une énumération d'expériences spécifiques. Il s'agit d'un modèle distribué, qui se propose de modéliser la microstructure de la mémoire. L'une de ses limitations, qu'il partage avec les modèles d'énumération, est qu'il repose néanmoins sur un ensemble fixe de primitives représentationnelles (p. 173). Ce modèle, comme d'autres modèles distribués, suppose que le système mnésique consiste en un réseau d'unités fortement interconnectées qui présentent des valeurs d'activations et qui communiquent avec les autres à l'aide de signaux, modulés par des poids de connexions. En ce qui concerne la

valeur « primitive » de ces unités, les auteurs notent : « Sometimes we may think of the units as corresponding to particular representational primitives, but they need not » (pp. 173-174). Il est également important de noter que, dans ce modèle, la notion de *module* de traitement n'est pas exclue (p. 174) :

*« We assume that the units are organized into modules. Each module receives inputs from other modules ; the units within the module are richly interconnected with each other ; and they send outputs to other modules ».*

Chaque module combine par ailleurs différentes sources d'information, que celles-ci soit plutôt abstraites (conceptuelles, haut-bas) ou plutôt périphériques (sensorielles, bas-haut). Les auteurs établissent alors certaines correspondances entre leur modèle et plusieurs concepts mnésiques standards (p. 176) :

- ◆ Etat mental : un état mental correspond à un schéma d'activation des unités ; le traitement de l'information est le processus d'évolution dans le temps des états mentaux.
- ◆ Trace mnésique : les traces mnésiques correspondent aux changements de poids de connexions entre unités.
- ◆ Récupération : La récupération en mémoire (*retrieval*) correspond à une réinstanciation (*reinstatement*) partielle d'un état mental passé.

Ce modèle permettrait ainsi, d'une part d'extraire ce qui semblerait être le prototype ou la tendance centrale d'un ensemble de schémas, et d'autre part de faire coexister représentations spécifiques et connaissances prototypiques dans le même ensemble de connexions (p. 182).

Les architectures connexionnistes, en donnant un nouvel éclairage à la question, ont ainsi le mérite d'inviter les didacticiens à réévaluer leur conception de la mémoire et de la « mémorisation », une conception sur laquelle repose depuis des décennies tout un ensemble de pratiques et de considérations théoriques dans le domaine de la didactique

des langues. Le résumé qu'en propose G. Tiberghien (1997, p. 151) nous semble bien en donner la teneur :

*« Ce sont les codages des épisodes vécus par le système qui sont stockés en mémoire et le processus d'accessibilité à la mémoire résulterait toujours d'une interaction entre un contexte de récupération et un état du système de mémoire récapitulant son historique ; ce processus engendrerait un nouvel état reflétant la synergie des composantes d'encodage et de récupération dans l'acte de remémoration. La mémoire est ici fondamentalement de nature épisodique : l'accès à un état de mémoire est toujours contextualisé, l'abstraction résulte de l'interaction entre l'historique du système à mémoire et les conditions instantanées de la récupération. En d'autres termes, les représentations sémantiques ne sont pas stockées dans la mémoire mais elles sont la résultante de cette interaction : le sens n'est pas stocké en mémoire, mais il émerge des règles de fonctionnement épisodique de la mémoire ».*

#### I.2.1.1.3. Format et Temporalité

Nous considérons ici quelques questions transversales relatives à l'encodage mnésique de l'information, dans une perspective toutefois essentiellement computo-symbolique.

##### I.2.1.1.3.1. Format et Modalité

En ce qui concerne le format des représentations mnésiques, la question de l'unicité *versus* la pluralité des représentations sémantiques (stocks sémantiques verbal, visuel, tactile, etc.), liées aux canaux perceptifs, est toujours débattue par les chercheurs, certains se fondant sur « l'observation de déficits de dénomination limités à une modalité sensorielle de présentation » (Kremin, 1994, p. 383). Ainsi voit-on apparaître certains modèles morpholexicaux, établis sur la base d'études en neurolinguistique cognitive, contenant « four modality-specific lexical systems, a semantic system, and lexical processing routes » (De Bleser et Cholewa, 1998, p. 233) et permettant

notamment, au sein du système morpholexical, de différencier inflexions régulières et irrégulières.

Il convient cependant, sur cette question, de bien distinguer les codes perceptifs (registre sensoriel) des codes de stockage (format représentationnel en mémoire à long-terme). Les différentes théories en présence (double codage, verbal et imagé (Paivio), *versus* codage sémantique abstrait et amodal de nature propositionnelle (Anderson)) ont ainsi conduit certains chercheurs à considérer la question en fonction du niveau de « profondeur » des représentations (Kosslyn) (G. Tiberghien, 1994, p. 266). Se pose ici le problème du « formatage » des inputs auditifs et visuels en « représentations profondes » codées en mémoire : codage amodal, modal, ou encore distribué. La question à l'origine du débat peut être formulée ainsi (*op.cit.*, p. 265) : « le format des représentations en mémoire est-il unique ou, au contraire, est-il affecté par les propriétés des inputs sensoriels ? ». Il est établi, par exemple, que « la mémoire imagée [s'avère] nettement supérieure à la mémoire verbale » (*ibid.*), et « les données cliniques donnent des raisons de croire que le traitement sémantique trouve son origine dans le traitement sémantique visuel »<sup>108</sup> (Visch-Brink et Bastiaanse, 1998, p. 17)<sup>109</sup>. L'imageabilité des mots semblerait également être un facteur important dans leur mémorabilité (Baddeley, 1986, p. 121). Du point de vue des codes perceptifs, Baddeley note néanmoins, en référence aux travaux de Broadbent, et à propos d'une tâche de rappel, que (1986, p. 43) :

*« There is indeed good evidence to suggest that auditory presentation does give rise to a particularly durable memory trace which resists disruption by visual material, whereas material that is presented visually appears to be readily disrupted by subsequent auditory material ».*

---

<sup>108</sup> « The patient data give reason to believe that semantic processing should originate from visual semantic processing ».

<sup>109</sup> Hypothèse de G. Goldenberg, « Aphasic patients' knowledge about the visual appearance of objects », *Aphasiology*, 9, 50-56, 1995.

Il est ainsi établi que la modalité de présentation d'un item peut influencer la probabilité de rappel de celui-ci : une présentation auditive favoriserait, dans une certaine mesure, le rappel, en particulier dans la portion de récence de la liste (globalement, les derniers items de la liste)<sup>110</sup> (Baddeley, 1986, p. 86). Il semblerait par ailleurs qu'il existe une asymétrie entre l'interférence produite par la présentation de matériel visuel sur le rappel de matériel auditif et l'interférence inverse (1986, p. 87). Kekenbosch (1994, pp. 22-23) présente également des résultats qui montrent que pour le codage en Mémoire à Court Terme « il n'y aurait pas d'effet cumulatif du traitement des présentations visuelles, sonores et articulatoires » et que « le codage auditif et/ou articulatoire [y] domine le codage visuel ».

#### 1.2.1.1.3.2. Temporalité

A la question du format des représentations mnésiques, vient s'ajouter celle de la dimension temporelle de leur encodage en mémoire. Depuis quelques décennies, les chercheurs en neurosciences ont en effet accordé de plus en plus d'importance à ce que Peschanski appelle la « quatrième dimension du cerveau », à savoir la dimension temporelle, qui, à l'échelle cellulaire, s'exprime dans les phénomènes de neuroplasticité et de Potentialisation à Long Terme (*Long Term Potentialization, L. T. P.*) (Peschanski, 1993, pp. 49-50 ; Fabbro, 1999, pp. 90-91) et qui, à l'échelle cérébrale, implique des phénomènes de latéralisation fonctionnelle (Joanette et Roch Lecours, 1984, pp. 28-29). La question du codage temporel (Thorpe, 1994, pp. 162-163) de l'information dans les structures cérébrales (synchronie ou asynchronie entre les aires sollicitées ; neuroplasticité) est une question intimement liée aux processus de mémorisation et d'apprentissage. Bullier nous apprend par exemple que (1997, pp. 84-86) :

*« L'association entre neurones activés par le même stimulus fait intervenir le couplage temporel de leur réponse neuronale. [...] La synchronisation entre*

---

<sup>110</sup> L'effet de récence serait dû, d'après Baddeley, à une opération de stockage passive et automatique, dans laquelle l'information serait maintenue par le système sans que le sujet n'ait à la répéter activement (Baddeley, 1986, p. 145).

*neurones situés dans des aires corticales différentes constitue une bonne réponse au problème du groupement fonctionnel entre neurones situés dans des aires corticales différentes ».*

#### I.2.1.1.3.3. Parenthèse didactique

D'un point de vue didactique, ces données et hypothèses, concernant tant le format que la temporalité, permettent d'alimenter les interrogations quant à l'utilisation pédagogique des représentations auditives et visuelles dans l'enseignement d'une langue étrangère.

En ce qui concerne le format, plusieurs questions doivent être posées si l'on souhaite utiliser à bon escient tant « l'image » que « l'écrit » en cours de langue. Par exemple, en ce qui concerne la première : de quel type d'image s'agit-il (analyse de type sémiotique et format) ? Quel(s) rapport(s) (explicatif, illustratif, etc.) entretient-elle avec les autres éléments (verbaux ou non, visuels ou auditifs) de l'input didactique (Develotte, 2000, p. 27) ? Comment est-elle agencée vis-à-vis de ces derniers (d'un point de vue spatial et temporel) (Jamet, 2001) Quel rôle souhaite-t-on lui faire jouer dans l'apprentissage (aide à la compréhension, indice mnésique, etc.) (Merlet, 1998) ? Et enfin quel rôle ne doit-elle pas jouer (parasite attentionnel, surcharge cognitive, source d'ambiguïté) ? Ainsi Bertin, lors de la conception d'un CD-ROM d'apprentissage de l'anglais du Transport International et de la Logistique, a décidé de remplacer une photographie par un dessin stylisé pour « une appréhension plus simple et plus rapide du message de la part du message dévolue à l'image » (Bertin, 2001, p. 132). Il peut donc s'avérer utile, voire essentiel ici, de disposer de modèles de référence concernant la manière dont est traitée l'information visuelle.

Par ailleurs, toujours d'un point de vue didactique, présenter un stimulus auditif et un stimulus visuel pour favoriser l'apprentissage grâce à la multimodalité semble donc ne plus suffire : il faut aussi penser leur agencement temporel. Tout en étant conscient de



la variation des phénomènes entre échelle humaine et échelle cellulaire, le didacticien peut ainsi s'interroger sur les combinaisons temporelles optimales<sup>111</sup> entre les différentes modalités sollicitées (*inter-modales*), même si la problématique de la combinaison temporelle son/image en didactique des langues n'est bien sûr pas récente : elle avait déjà été examinée dans les années soixante (puis réexaminée ensuite) par les concepteurs de la problématique S.G.A.V. (Rivenc, 2000, Rivenc-Chiclet, 1989). Il peut ensuite s'interroger sur la répartition temporelle (durée et fréquence) des stimuli de même modalité (*intra-modale*), touchant ici, de près ou de loin, à ce qu'on appelle la « progression » en didactique des langues et donc à la distinction entre novices et experts. Notons que la notion didactique de « progression en spirale » se voit, d'une certaine manière, justifiée par les effets d'amorçage perceptif (*priming*) qui renvoient à « une facilitation de l'identification perceptive d'un stimulus due à une rencontre préalable avec ce stimulus »<sup>112</sup>, (Meulemans, 1998, p. 17), effets qui ont également été constatés à un niveau conceptuel. Mémorisation et temporalité sont ainsi étroitement liées en didactique des langues, s'inscrivant dans une perspective diachronique qui conduit à la notion de *progression*, et, avec elle, au couple *novice/expert*. Cette distinction peut être abordée sous un angle psycholinguistique, en s'appuyant par exemple sur les travaux de Anderson (1993) et en considérant alors la progression comme « le transfert des savoirs [(déclaratifs)] dans la mémoire procédurale et ainsi un accès direct aux productions procéduralisées »<sup>113</sup> (Towell et Hawkins, 1994, p. 220). Mais elle peut également être abordée sous un angle neuropsycholinguistique en considérant le processus de latéralisation hémisphérique de droite à gauche (Gordon et Weide, 1983, p.49-50), même si les prises de position diffèrent sur ce sujet : d'après Galloway (1983, p. 103), « il n'y a aucune indication que

---

<sup>111</sup> Peschanski (1993, p. 34) rapporte ainsi certains travaux montrant que « l'amélioration de la performance [...] dépend donc d'une meilleure synchronisation entre les neurones corticaux », la réponse étant différente non du point de vue du neurone individuel, mais du groupe de neurones.

<sup>112</sup> Ce que l'on pourrait relier, à l'échelle cellulaire, au fait que la L. T. P. est envisagée comme « une augmentation de la réponse post-synaptique à la stimulation d'une voie afférente excitatrice lorsque cette voie a, auparavant, subi une stimulation répétée de brève durée et de forte fréquence » (Peschanski, 1993, pp. 49-50).

<sup>113</sup> « The transfer of knowledge into the proceduralised memory and subsequent direct access to the proceduralised productions » (notre traduction).

l'HD [hémisphère droit] soit plus impliqué en L2 durant les premières étapes de [son] appropriation ». Pour Goldberg et Costa, en revanche (1981, pp. 144-173), « the gradient of the left-right hemispherical performance reflects the gradient of « trainedness » of the subjects for the tasks involved » (1981, p. 163), la supériorité de l'hémisphère gauche s'affirmant à mesure que les compétences nécessaires à la réalisation de la tâche sont acquises et deviennent routinières. Ceci est compatible avec les études psycholinguistiques de Kail (1997, p. 214) relatives aux *indices* (Kail, 1990) et à leur *coût de traitement* en tâches de compréhension dans le cadre du Modèle de Compétition, lesquelles indiquent qu'un « indice est d'autant plus valide pour l'attribution des fonctions et de traitement moins coûteux qu'il est plus local et [...] qu'avec le développement, les traitements locaux<sup>114</sup> sont privilégiés aux dépens des traitements topologiques ou globaux<sup>115</sup> ».

Ainsi, même si « aucune des théories actuellement en présence n'est réellement satisfaisante » (G. Tiberghien, 1997, p. 149), certains aspects du fonctionnement mnésique semblent de mieux en mieux connus, notamment ceux ayant trait à la mémoire de travail.

#### 1.2.1.2. Mémoire de travail et boucle phonologique

La nécessité d'intégrer un système de stockage temporaire et de manipulation de l'information durant l'accomplissement de diverses tâches cognitives, telles que la compréhension ou l'apprentissage, a été mise à jour depuis longtemps (comme par exemple le « STS », voir *supra.*). Le concept général de mémoire de travail a ainsi pu donner lieu à divers modèles spécifiques (Baddeley, 1986, p. 34), mais le modèle actuellement le plus reconnu en neuropsychologie (Van der Linden, 1994, p. 283) est celui de Baddeley (1986), dont la validité semble notamment appuyée par des études menées en imagerie cérébrale (Démonet, 1997, p. 121 ; Thierry, 2001). Dans celui-ci,

---

<sup>114</sup> Renvoyant de manière tendancielle à l'hémisphère gauche.

<sup>115</sup> Renvoyant de manière tendancielle à l'hémisphère droit.

la « mémoire de travail » (spécifique) vient remplacer la mémoire à court terme. Les trois composants essentiels de ce modèle sont les suivants :

- **Une boucle articulatoire** (ultérieurement nommée **phonologique**)<sup>116</sup>, composée d'un stock phonologique (dans lequel l'information est stockée temporairement et disparaît après une ou deux secondes si elle n'est pas « rafraîchie ») et d'un processus de récapitulation articulatoire. Cette boucle permet le stockage temporaire de l'information verbale. Sa fonction est double : lorsque l'information est présentée auditivement, celle-ci est directement reçue par le stock phonologique et est maintenue, à l'aide du processus de récapitulation articulatoire, sous forme de codes phonologiques. Lorsque l'information est présentée visuellement, celle-ci est d'abord convertie en un code phonologique *via* le processus de récapitulation articulatoire avant de pouvoir y être stockée (Van der Linden, 1994, figure 1, p. 283).
- **Un bloc-notes visuo-spatial** qui participe à la création et à la gestion des images mentales. De nombreuses études en neuropsychologie sembleraient en effet justifier une séparation des processus verbaux et visuels (Baddeley, 1986, p. 109). Notons qu'un « processus de récapitulation active des images mentales (analogue au processus de récapitulation articulatoire) pourrait être sous-tendu par le système impliqué dans le contrôle des mouvements des yeux » (Van der Linden, 1994, p. 284). Ainsi, certaines études indiquent que l'activité d'imagerie visuo-spatiale serait davantage perturbée par l'acte de lire que par l'acte d'écouter (Baddeley, 1986, p. 115).
- **Un système de supervision**, qui gère la répartition des ressources attentionnelles entre les deux sous-systèmes précédents (Tiberghien, 1994, p. 263) ainsi que les procédures de gestion (planification et contrôle). Van Der Linden nous rappelle que

---

<sup>116</sup> Le chapitre 5 (pp. 75-107) de Baddeley (1986) s'intitule "The articulatory loop", tandis que l'article de Baddeley, Gathercole et Papagno (1998) s'intitule "The phonological loop as a language learning device".

cet administrateur central, dans le modèle de Baddeley, est considéré comme système attentionnel de contrôle et que le modèle de contrôle attentionnel proposé par Norman et Shallice permet de mieux en comprendre le fonctionnement (Van der Linden, 1994, p. 284). Dans ce dernier modèle, il y aurait une hiérarchie attentionnelle, dont l'agencement et l'activation seraient fonction des difficultés rencontrées par le sujet :

- Peu de contrôle, activation de routines (de schémas d'action).
- Processus semi-automatiques de résolution de conflits (« Contention Scheduling Mechanism ») lorsqu'il y a conflit entre les routines.
- Système attentionnel de supervision (« Supervisory Attentional System », SAS) lorsque la sélection d'un schéma n'est pas suffisante.

Comme le rappelle Baddeley (1986, p. 228), le SAS a une capacité limitée, et est sollicité dans les circonstances suivantes :

- a) « Tasks that involve planning or decision making.
- b) Situations in which the automatic processes appear to be running into difficulties and some form of trouble-shooting is necessary.
- c) Where novel or poorly learned sequences of acts are involved.
- d) Where the situation is judged to be dangerous or technically difficult.
- e) Where some strong habitual response or temptation is involved ».

Le modèle de contrôle de Norman et Shallice comporterait donc deux niveaux : l'un, horizontal, permettant d'éviter la production de réponses incompatibles, l'autre, vertical correspondant au SAS. Baddeley semble, dans une certaine mesure, juger appropriée l'identification, même partielle du SAS et du système de supervision de son propre modèle : « After 20 years of trying to understand [...], I was delighted to find that the Norman and Shallice model provided a simple and plausible interpretation » (1986, p. 228).

#### 1.2.1.2.1. Nature de la boucle phonologique

La boucle phonologique se compose de deux unités : un stock phonologique et un processus de répétition (*rehearsal*) subvocalique. Le stock contient le matériel verbal sous sa forme phonologique, pouvant correspondre aux phonèmes ou bien à des traits sous-phonétiques (*subphonetic*) (Gathercole, 1998). La durée de stockage peut être prolongée *via* le processus de répétition. L'information peut donc accéder au stock *via* deux routes : une route directe impliquant l'input auditif ; une route indirecte pour les stimuli non-auditifs. La boucle phonologique permet donc de stocker des séquences phonologiques externes ainsi que des séquences générées de manière interne. Cette boucle est cohérente avec l'observation de plusieurs effets :

- *Effet de similarité phonologique* : deux items phonologiquement proches sont moins bien rappelés en rappel sériel immédiat (Baddeley, 1986, p. 76).
- *Effet de longueur* : dans le même type de tâche, un mot qui est long à prononcer va gêner le rappel du mot précédent en raison du processus de récapitulation (Baddeley, 1986, p. 78).
- *Effet de suppression articulaire* (ou d'articulation concurrente), dans certaines conditions : l'effet de similarité phonologique est supprimé lorsque les items sont présentés visuellement et lorsque les sujets sont empêchés de subvocaliser (*subvocal rehearsal*) par une tâche de suppression articulaire, comme par exemple répéter un item non-pertinent (ex. : « le ») tout au long de la présentation<sup>117</sup> (Baddeley, 1986, p. 78).

Les deux principaux effets reflètent les deux composants de la boucle : tandis que l'effet de longueur est lié à la répétition (« longer words take longer to say »), l'effet de similarité est lié au stock phonologique, maintenu et rafraîchi par la boucle (Baddeley, 1986, p. 84).

---

<sup>117</sup> Baddeley note cependant (1986, p. 82) que lorsque les éléments à mémoriser sont présentés auditivement, l'effet de similarité phonologique résiste à la suppression articulaire, une observation qui a motivé l'évolution du modèle initial de la boucle articulaire.

En ce qui concerne la nature « phonologique » de cette boucle, Baddeley indique (1986, p. 75) :

*« I shall use the term "phonological" in a purely neutral sense as meaning speech-based. The term "articulatory" will be used when the coding is assumed to be based on the subject's speech production system, while the term "acoustic" will be used when the coding is assumed to be based on the subject's speech perception processes ».*

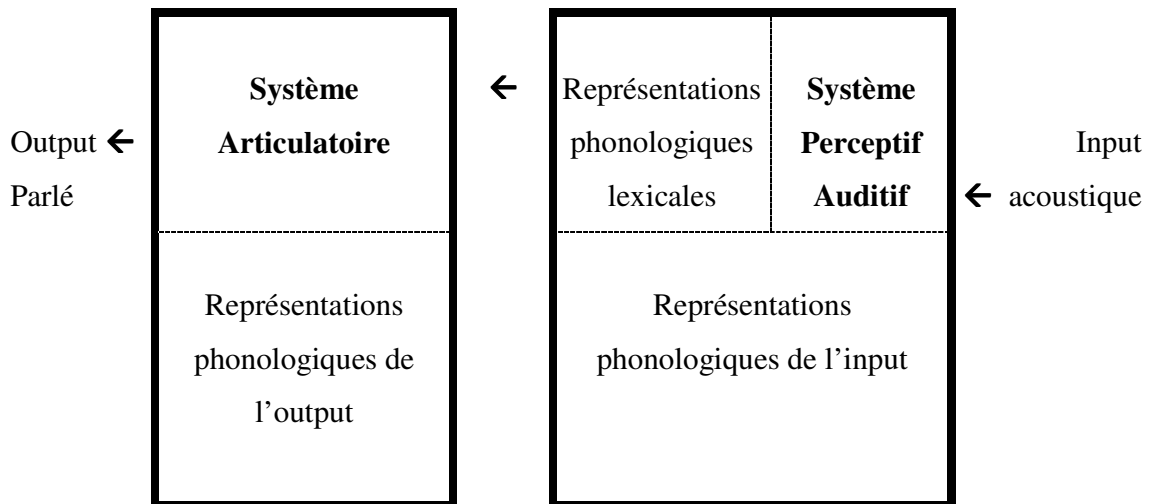
Cette question a cependant été abordée plus en détail quelques années plus tard par Gupta et MacWhinney, qui posent la question suivante : « Is the phonological loop articulatory or auditory ? » (1993), c'est-à-dire, quelle est la nature du processus qui « rafraîchit » les représentations emmagasinées ? En effet, alors que le paradigme du rappel sériel immédiat (tâche dans laquelle on présente au sujet une liste d'items verbaux indépendants que celui-ci est censé rappeler dans leur ordre de présentation, selon diverses modalités de présentation et de rendus) avait constitué la base de nombreux travaux sur la mémoire de travail, son rapport à la nature des processus phonologiques en jeu n'avait jusqu'alors été que peu examiné.

Comme l'indiquent les auteurs (1993, p. 1, note 1) :

*« It appears well established that this process cannot be articulatory in the sense of involving the speech musculature ; rather, "articulatory" connotes that the process is somehow involved in phonological production planning [...] and it is in this sense that we will use the term "articulatory" or "output" ».*

Cela rejoint ce qu'écrit Baddeley : « the loop and its rehearsal processes are operating at a much deeper level than might at first seem likely, apparently relying on central speech control codes which appear to be able to function in the absence of peripheral feedback » (1986, p. 107). Cela expliquerait les échecs relatifs des études électromyographiques de la musculature des organes articulatoires portant sur la boucle phonologique.

Gupta et MacWhinney tentent donc de proposer une interprétation davantage phonologique de la boucle, en se basant sur un modèle initial basique de traitement de mots isolés oraux (que nous adaptons ici) (1993, p. 3):



A travers ce schéma minimaliste et grandement sous-spécifié, comme les notent les auteurs, peuvent être abordées les tâches de répétition, de production et d'apprentissage, de mots et de non-mots isolés. La question essentielle est alors de savoir dans quelle mesure l'information, concernant les « produits » issus du traitement phonologique lié à l'output, est disponible pour le traitement phonologique lié à l'input, l'interprétation standard du processus de récapitulation phonologique étant de considérer que les représentations phonologiques de l'input sont « rafraîchies » *via* le processus lié à l'output.

Or, selon les auteurs, cette interprétation, basée sur l'output, n'est pas la seule envisageable : le processus de récapitulation pourrait ne se baser que sur l'input. Se basant sur l'effet d'articulation concurrente, les auteurs examinent alors les deux perspectives tentant de rendre compte dudit effet : la première, « articulatoire », standard ; l'autre, « auditive », qui impliquerait un certain « écho mental » des stimuli

auditifs, via une « boucle d'écho mental » des « images auditives ». Cette dernière pourrait en effet rendre compte de l'effet en question si l'on considère que celui-ci est dû à des interférences auditives, créées par l'articulation concurrente du sujet lui-même. Cette interprétation ne nécessiterait donc aucune connectivité entre l'input et l'output.

Chacune des deux interprétations semble à même de pouvoir rendre compte des effets de compétition des ressources cognitives, telles que représentés par l'effet d'articulation concurrente, lequel implique :

- 1) L'utilisation de ressources générales en vue d'accomplir une tâche motrice.
- 2) La création d'interférences auditives.
- 3) L'utilisation de ressources articulatoires (Gupta et MacWhinney, 1993).

Cependant, comme l'indiquent les auteurs, ces interprétations diffèrent quant au rôle qu'elles accordent à une compétition des ressources articulatoires : postulée par la version articulatoire, une telle compétition est en revanche absente de la version auditive. Les chercheurs ont donc réalisé un protocole expérimental dans lequel un groupe de sujets devait accomplir une tâche de suppression articulatoire (répéter le mot « the ») durant la présentation des items à mémoriser (tâche articulatoire), tandis qu'un autre groupe devait taper du doigt tout en entendant la répétition du mot « the » au cours de la présentation (tâche non-articulatoire). Les résultats semblent soutenir l'interprétation articulatoire standard basée sur l'output et impliquant une connexion forte entre l'output et l'input. Notons que Baddeley avait cependant déjà évoqué la question en considérant la possibilité de deux systèmes de stockage phonologique : l'un, la « voix interne » (*inner voice*), était supposé correspondre à la boucle articulatoire, l'autre, l'« oreille interne » (*inner ear*) était censé impliquer une forme d'image acoustique, indépendamment de tout processus articulatoire (1986, p. 86).

Le succès de la boucle phonologique a par ailleurs conduit certains à tenter d'en proposer une modélisation de type connexionniste. Le modèle de Burgess et Hitch (1992), par exemple conduit à de nombreux résultats satisfaisants relativement comparables à ceux trouvés en performance humaine : empan mnésique, effet de



similarité phonémique, effet de longueur du mot, effet de primauté (les premiers items sur la courbe de rappel sériel sont mieux rappelés que ceux qui se trouvent dans la partie centrale), types d'erreurs de rappel (erreurs d'ordre, d'items et omissions) et, dans une certaine mesure, effet de suppression articulaire. Bien sûr, le modèle présente également quelques faiblesses, notamment l'incapacité à produire un effet de récence (en performance humaine, les derniers items sur la courbe de rappel sériel sont en effet mieux rappelés que ceux qui se trouvent dans la partie centrale) et l'absence de variation de l'effet de suppression articulaire selon la modalité, contrairement à la performance humaine (Burgess et Hitch, 1992, pp. 452-453).

#### 1.2.1.2.2. Fonction de la boucle phonologique

Nous présentons ici la fonction essentielle, à ce jour, de la boucle phonologique, à savoir l'acquisition de nouvelles formes phonologiques, avant d'évoquer rapidement la question de la lecture.

##### 1.2.1.2.2.1. Acquisition de formes phonologiques nouvelles

Dans une perspective développementale, il apparaît que l'augmentation de l'empan mnésique chez l'enfant est due à l'augmentation de la vitesse et de l'efficacité du processus de répétition (Gathercole, 1998). Un rythme plus rapide de répétition permet en effet de contenir davantage de matériel dans le stock et de le maintenir sans perte. Cette augmentation est constatée jusqu'à l'âge de 12 ans, à partir duquel le rythme d'articulation correspond à un rythme adulte.

En ce qui concerne la fonction de la boucle phonologique, des études en neuropsychologie ont montré que celle-ci ne jouait pas un rôle crucial pour la production ou la compréhension langagière. C'est alors en étudiant des populations d'enfants, plutôt que d'adultes, et en se concentrant sur la capacité à apprendre de nouvelles informations linguistiques, plutôt que sur les performances langagières, que les fonctions de la boucle phonologique sont apparues plus clairement (Gathercole, 1998). Ainsi, il apparaît que les compétences de la mémoire phonologique chez les

enfants sont étroitement liées à leur capacité à apprendre du nouveau matériel phonologique.

Il semblerait cependant que la relation entre la mémoire à court-terme phonologique et l'apprentissage à long terme soit restreinte à l'acquisition des formes phonologiques, et non sémantiques, des nouveaux mots. Gathercole (1998) avance l'hypothèse selon laquelle un stockage temporaire adéquat de la forme phonologique d'un nouveau mot constitue la première étape cruciale vers la construction d'une représentation stable à long-terme de sa structure. Le lien entre boucle phonologique et acquisition lexicale semble clairement mis en évidence par de nombreuses études : « the capacity of the phonological loop to hold telephone numbers and to support serial recall of lists of unrelated words is fortunate for experimental psychologists, but secondary » (*ibid.*). Le rôle de la boucle phonologique dans l'apprentissage de vocabulaire nouveau, notamment en langue étrangère, a ainsi été particulièrement étudié au cours des quinze dernières années.

Par exemple, dans une étude longitudinale sur l'apprentissage de l'anglais par des enfants finlandais, à travers trois tâches (répétition orale de non-mots parlés (création de représentations phonologiques), copie retardée de non-mots présentés visuellement (pas de modalité orale, mais un recodage articulatoire) et analyse métalinguistique (pas de manipulation phonologique, mais traitement de matériel verbal)), Service (1992) obtient des résultats qui soutiennent l'importance de la boucle phonologique dans l'apprentissage du vocabulaire étranger. Ses trois tâches, conçues comme des indicateurs de capacité à apprendre du vocabulaire étranger, ont conduit au classement suivant : la répétition est un indicateur légèrement meilleur que la copie, elle-même bien meilleure que la tâche métalinguistique de comparaison sémantico-syntaxique (p. 42). Ces résultats indiquent donc que la capacité à représenter du matériel phonologique non-familier dans la mémoire de travail sous-tend l'acquisition de vocabulaire nouveau en langue étrangère. La capacité de répétition de non-mots semble donc être un très bon indicateur pour les enfants de leur capacité à apprendre un vocabulaire étranger.

Un autre exemple est l'étude de Papagno, Valentine et Baddeley (1991), qui ont tenté de vérifier l'hypothèse selon laquelle le stockage phonologique à court-terme jouait un rôle important dans l'acquisition du vocabulaire natif par de jeunes enfants, en utilisant la technique de la suppression articulatoire avec des enfants anglais et italiens. Leurs résultats se divisent en deux catégories : ceux qui soutiennent l'idée selon laquelle la suppression articulatoire parasite l'acquisition de vocabulaire étranger, tant en modalité auditive (puisque la suppression empêche les sujets de subvocaliser) qu'en modalité visuelle (puisque la suppression interfère avec le recodage phonologique, forçant les sujets à s'appuyer sur des ressources lexicales et sémantiques), et ceux qui n'étaient pas attendus par les chercheurs. Ces derniers suggèrent en effet que, pour l'apprentissage de mots étrangers, lorsque des associations sémantiques sont possibles (notamment entre mots « cognates ») entre le mot à apprendre et le lexique de la langue première des sujets (comme dans le cas de l'italien et de l'anglais), les sujets utilisent ces associations, transformant la tâche phonologique en tâche sémantique (p. 344).

Baddeley, Gathercole et Papagno (1998), passent à leur tour en revue diverses études (population adulte, infantile, pathologique) qui montrent que la boucle phonologique joue un rôle crucial dans l'apprentissage des formes phonologiques de nouveaux mots. Il semblerait en effet que l'apprentissage à long-terme de la structure sonore de mots nouveaux, phonologiquement non-familiers, dépende de la disponibilité de représentations sonores adéquates dans la boucle phonologique (p. 163) :

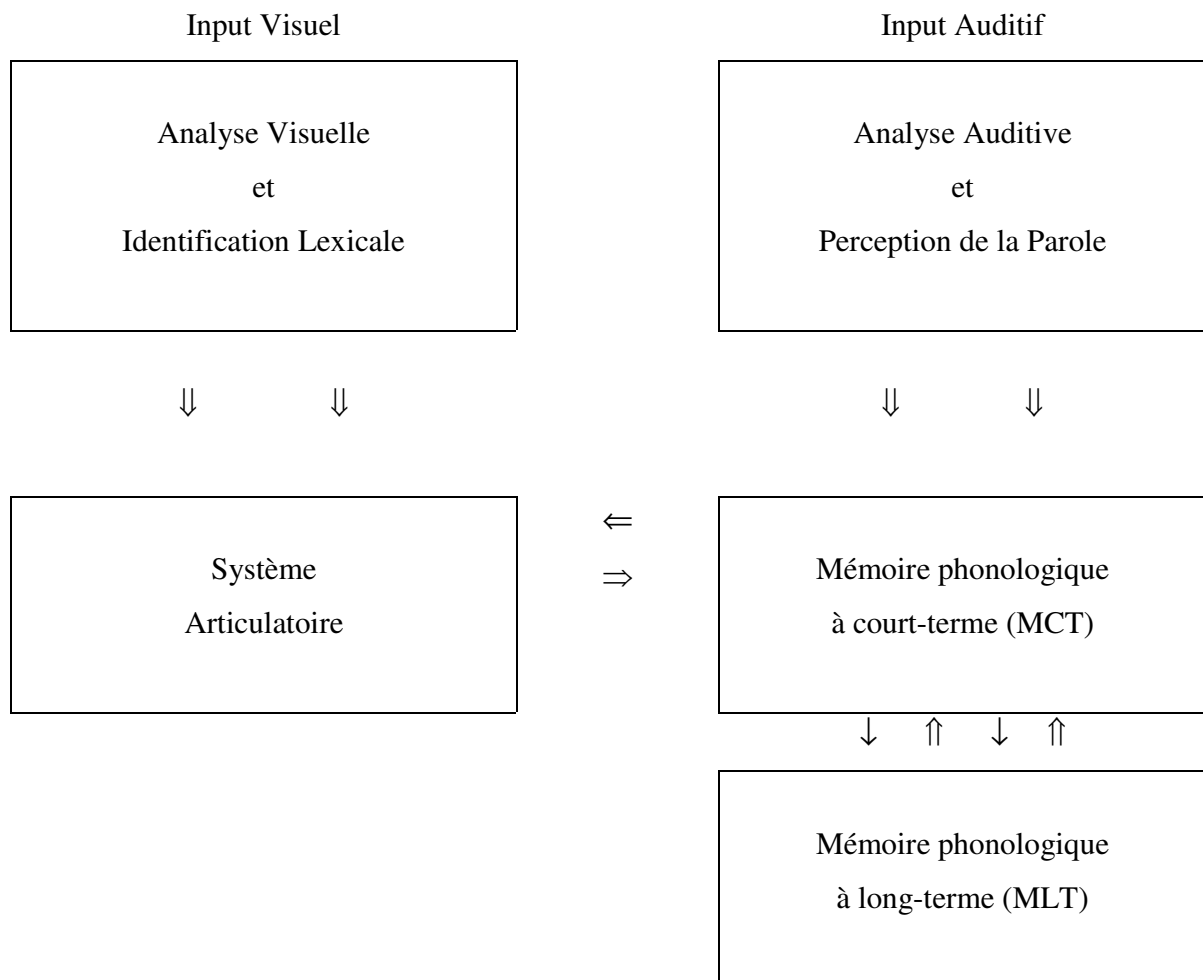
*« Thus the phonological loop appears to provide a critical input to the construction of the more permanent phonological structures that are stored in the mental lexicon. Learning of associations that require the production of familiar lexical items, on the other hand, is achieved typically either without any reliance on the phonological loop or with reduced loop support and is presumably mediated instead by the use of existing knowledge of the native language ».*

Les auteurs mentionnent également certaines études récentes qui se sont penchées sur le rôle de la boucle phonologique dans l'apprentissage de formes morphosyntaxiques. Les résultats seraient en accord avec l'idée selon laquelle l'apprentissage des items lexicaux et de leurs propriétés morphologiques serait sous-tendu par un mécanisme unique : il n'y aurait pas de distinction fonctionnelle entre la manière dont un mot et, par exemple, celle dont la morphologie inflectionnelle serait appris (p. 167).

A la question « Quel aspect de la boucle phonologique est essentielle à cette fonction d'apprentissage ? », les auteurs répondent : le *stock phonologique*, même si la répétition pourrait être importante pour maintenir la qualité de la représentation. La répétition de non-mots fournirait ainsi, non une mesure de la répétition (*rehearsal*) phonologique, mais du stock phonologique (p. 168). Il apparaît par ailleurs que la boucle phonologique est capable d'utiliser des connaissances préalablement acquises (p. 169) :

*« Long-term knowledge is used to "fill-in" representations in the phonological loop that are incomplete, as a consequence of either decay or interference, by using a process [...] whereby partial traces resulting from familiar words or highly wordlike novel words have a greater likelihood of being correctly reinstated than nonwords with unusual sound patterns such as words in a foreign language ».*

Le schéma suivant propose une vue simplifiée de la relation entre stockage à court-terme et apprentissage à long-terme (p. 170) :



Les flèches simples entre la MCT et la MLT indiquent que la modification de la MLT est lente et progressive. Les auteurs concluent donc : « the primary function of the phonological loop is the processing of novel speech input » (p. 170). Cette boucle permet ainsi de stocker des schémas sonores non-familiers, tandis que se bâtit une représentation plus permanente en mémoire à long-terme.

#### 1.2.1.2.2.2. Lecture

En ce qui concerne l'implication de la boucle phonologique dans la lecture, Baddeley notait déjà en 1979 qu'il fallait bien distinguer les lecteurs *novices*, qui pouvaient utiliser la boucle, des lecteurs *experts*, qui n'en faisaient que peu usage en lecture. Les expériences qu'il présente indiquent en effet que les sujets peuvent lire et comprendre des énoncés écrits sans utiliser la boucle phonologique, alors nommée articulatoire. Les conditions dans lesquelles la boucle serait requise sembleraient être les suivantes :

- Lorsque des jugements de similarité phonologique sont nécessaires.
- Lorsque la rétention mnésique du passage à lire est nécessaire.
- Lorsque l'ordre des mots est essentiel à la compréhension.
- Peut-être aussi lorsque le rythme de présentation du texte dépasse celui de traitement sémantique du texte.

A propos du lien entre boucle phonologique et lecture, Gathercole (1998) commence par souligner, elle aussi, que la lecture experte (*skilled reading*) ne fait que très peu, voire pas, appel à la mémoire à court terme. La mémoire phonologique n'intervient donc pas (ou peu) pour des lecteurs experts. Plusieurs hypothèses ont été avancées concernant les liens certains entre le développement de la lecture et la boucle phonologique, notamment le rôle de la boucle dans l'apprentissage des associations sons-lettre et dans le stockage de segments générés par une stratégie de recodage phonologique préalable à la fusion (*blending*). Il semblerait d'ailleurs d'après Baddeley (1986, p. 218) que « specific reading disability is associated with impaired functioning of the articulatory loop ». Dans ce domaine, le centre d'intérêt dominant a été le lien entre capacité de lecture et conscience phonologique : afin de devenir un lecteur compétent, l'enfant doit développer une conscience métalinguistique de la structure de la langue parlée. Cependant, comme l'ont souligné les travaux de Bruxelles (Bertelson, Morais, Content et al.), le développement de cette conscience est-elle même fortement stimulée par l'apprentissage de la lecture. Les liens entre le système de stockage à court-terme et l'apprentissage de la lecture ont donc été peu étudiés.

En ce qui concerne la nature des processus mnésiques à l'œuvre en tâche de compréhension textuelle, d'autres modèles de la mémoire de travail ont été proposés, notamment celui de Ericsson et Kintsch (1995) qui, étant donné la taille du stockage informatif requis dans une telle tâche, propose de dépasser le cadre de la mémoire à court-terme pour permettre un stockage en mémoire à long-terme. Dans ce cadre, les produits du traitement informatif seraient stockés dans la MLT mais directement accessibles à l'aide d'indices de rappel en MCT : il s'agit de la mémoire de travail à long-terme (MT-LT). Selon les chercheurs, dans le cas de la lecture d'un texte, la séquence de représentations intermédiaires générées lors du traitement et les étapes de stockage provisoires qui y sont associées, d'un point de vue mnésique, se présente ainsi :

1. Icône rétinotopique : mise en œuvre des photorécepteurs permettant la perception d'un mot isolé.
2. Mémoire visuelle spatiotopique : le texte est représenté comme une structure spatiale stable, grâce à l'intégration de l'information rétinienne issue des fixations oculaires successives.
3. Mémoire visuelle réatopique : intégration des informations relatives à la configuration des traits et des schémas visuels du segment textuel perceptible.
4. Mémoire-tampon conceptuelle : à ce niveau, les mots et les objets sont perçus et compris.
5. Boucle articulatoire : l'information visuelle textuelle est transformée en forme acoustique.
6. Le calepin visuo-spatial : la fonction est la même que la précédente, mais cette fois-ci pour le traitement des informations visuo-spatiales.
7. La mémoire de travail : l'information concernant le segment textuel en cours de traitement est stockée sous une forme rapidement accessible.

Selon ce modèle de mémoire de travail en compréhension textuelle, lors de la lecture experte de textes correctement écrits, les représentations mentales des phrases successives sont donc générées dans la mémoire de travail à court-terme. Certains

éléments de celles-ci sont liés d'une part aux représentations précédemment construites (mémoire épisodique textuelle) et déjà stockées en mémoire à long-terme, et d'autre part aux connaissances du lecteur. Ces connexions résultent en une mémoire de travail à long-terme qui permet un accès direct aux informations pertinentes de ces structures à partir des indices présents en mémoire à court-terme.

Cependant, ce modèle est particulièrement destiné à rendre compte de performances cognitives complexes (compréhension textuelle, calcul mental), et la ML-LT n'est pas une capacité généralisable à toute activité cognitive : elle est acquise dans certains domaines particuliers afin de répondre à certaines demandes particulières en termes de stockage et de récupération mnésique dans le contexte d'activités spécialisées spécifiques. Nous ne nous étendrons donc pas plus sur ce sujet, puisque nous nous intéressons davantage à l'acquisition de nouvelles formes phonologiques qu'à la compréhension textuelle.

#### 1.2.1.2.3. Parenthèse didactique

D'un point de vue didactique, l'organisation fonctionnelle de la boucle phonologique permet d'éclairer, d'une part l'importance de l'activité de « répétition » pour la mémorisation du matériel verbal, et d'autre part les possibles interactions entre code écrit et prononciation, puisque l'information visuelle (orthographique) est d'abord recodée phonologiquement, selon les codes préexistants naturellement. Si elle suggère de renouveler (au sens de « rafraîchir ») la présentation du matériel pour entretenir la persistance de l'information, elle implique également de laisser une marge temporelle suffisante entre la présentation de deux séquences pour que la récapitulation puisse avoir lieu. Si la présentation visuelle du matériel verbal semble offrir l'avantage de pouvoir laisser le sujet gérer sa récapitulation avant de passer aux fragments suivants, il convient d'exploiter ce support visuel en tenant compte du processus de recodage lorsque le matériel est orthographique, mais également, sinon surtout, de bien planifier la répartition temporelle du matériel présenté, que ce soit en termes de fréquence d'apparition des mêmes séquences, ou en termes de pauses.



Ces remarques relatives aux aspects temporels de la boucle phonologique d'un point de vue didactique peuvent également être formulées vis-à-vis du rôle du calepin visuo-spatial en tâche de lecture, en considérant la fréquence d'exposition et les pauses. Par ailleurs, les habitudes de lecture étant étroitement liées aux mouvements oculaires, il conviendrait de prendre en considération le public auquel est présenté le matériel, de façons à ne pas entraver le processus de récapitulation. Modifier les habitudes de lecture est une autre tâche, et il faut donc savoir les séparer dans un premier temps, si l'on souhaite faciliter leur accomplissement progressif.

En ce qui concerne la lecture, la distinction entre expert et novice soulignée par Baddeley et Gathercole semble tout à fait pertinente dans le cas d'apprenants de langue étrangère. La prise en compte de la boucle phonologique dans l'approche méthodologique adoptée en pédagogie de la lecture en langue étrangère (à laquelle est intimement lié, comme on l'oublie trop souvent, l'apprentissage du système phonético-phonologique de la langue en question, pour les apprenants débutants au moins) semble donc essentielle.

Le processus de récapitulation indique ainsi à quel point le sujet peut être actif dans le processus de mémorisation. Cette activité peut considérablement influencer sur la manière dont sont encodées les informations (G. Tiberghien, 1994, p. 267) :

*« La mémoire dépend de façon critique de l'activité déployée lors de l'encodage. Lorsqu'on est exposé à une information perceptive, celle-ci est interprétée et transformée afin d'être intégrée à la mémoire permanente. Cette « élaboration » intègre à la trace mnésique des informations contextuelles, mais y applique aussi une activité d'inférence qui peut modifier, éventuellement, l'information initiale, l'appauvrir ou l'enrichir. Cette élaboration dépend évidemment de l'expérience antérieure et de la motivation cognitive et affective des conduites ».*

Il convient alors, d'une part de *laisser le temps* à l'apprenant d'être actif, et d'autre part d'orienter son attention, par exemple en activant le « SAS » de Norman et Shallice, sur les caractéristiques du matériel qui seraient fortement susceptibles d'être involontairement modifiées lors de leur encodage mnésique, afin de prévenir de telles modifications, modifications que l'on pourrait lier à l'activation de « routines » préexistantes.

### **I.2.2. Connaissances linguistiques**

Comme cela est apparu précédemment lors de l'examen du système mnésique, la question du format de stockage cognitif de l'information est sujette à débat. Ce débat s'articule en particulier autour de deux oppositions : la première, entre format *propositionnel* et *non-propositionnel* des connaissances ; la deuxième, entre représentations *symboliques* et *subsymboliques*. Les différentes conceptions du savoir linguistique que peuvent avoir les spécialistes sont assez bien illustrées par cette citation de Laks (1996, p. 99) :

*« Pour un structuraliste, la langue est le sujet de deux formes radicalement différentes et disjointes de savoir : un savoir explicite et propositionnel qui est le propre de l'observateur extérieur qui objectivise et idéalise les productions des locuteurs afin de dégager le système des régularités linguistiquement pertinentes, et un savoir psychologique propre au locuteur dont la linguistique n'a, par principe, rien à dire. Pour un connexionniste réductionniste, connaître une langue c'est, pour un organisme, se trouver dans un état tel que l'activité neuronale qui y a lieu présente une forme de régularité et de stéréotypie reproductible et donc interprétable par soi et par autrui, sans pour autant négliger les facteurs sociaux et culturels qui bornent et contraignent cette régularité ».*

Si la tendance post-structuraliste a été d'établir une relation d'identité de plus en plus forte entre savoir propositionnel « extérieur » et savoir « intérieur », notamment en vue de conférer un certain réalisme cognitif aux représentations et règles postulées par les

linguistes, de nombreux chercheurs ont également exploré d'autres hypothèses relatives aux représentations cognitives. Comme l'indiquent Bechtel et Abrahamsen (pp. 157-162), plusieurs propositions ont en effet été avancées en sciences cognitives pour défendre l'hypothèse d'un format non-propositionnel des connaissances, quoique symbolique. Cela s'est notamment illustré par les travaux sur l'imagerie mentale (Denis, 1994, p. 9) :

*« Par représentation, nous désignons cette forme de l'activité humaine qui consiste à produire des **symboles** ayant pour caractéristique **de tenir lieu** d'autres entités. [...] L'image est une modalité de représentation mentale qui a pour caractéristique de conserver l'information perceptive sous une forme qui possède un degré élevé de similitude structurale avec la perception ».*

Comme nous l'avons déjà indiqué, l'opposition entre codage amodal et codage modal de l'information a alimenté de nombreux travaux. Il semblerait néanmoins que le connexionnisme constitue « une remise en question plus profonde de la conception selon laquelle les connaissances doivent être représentées propositionnellement » (Bechtel et Abrahamsen, 1991, p. 163). A la base de cette remise en question, se trouve notamment la distinction entre savoirs déclaratifs et savoir-faire procéduraux.

#### *1.2.2.1. Savoirs déclaratifs et savoir-faire procéduraux*

En regard du mode cognitif de stockage et de traitement de l'information, tous les individus se trouvent fédérés par l'existence de deux types de connaissances : celles que l'on peut qualifier de *déclaratives*, et les autres de *procédurales*. L'un des chercheurs à avoir développé cette distinction est G. Ryle, qui, en 1949, distingue *savoir faire* et *savoir propositionnel* (Bechtel et Abrahamsen, 1991, p. 163). Cette distinction peut par exemple s'illustrer dans l'opposition entre, *savoir que Paris est la capitale de la France*, et *savoir jouer du saxophone*. Dans le premier cas, il est possible de retrouver en mémoire, soit la proposition « Paris est la capitale de la France », soit d'autres propositions qui permettent d'accéder ensuite à la première par déduction.

Cela n'est en revanche pas possible dans le deuxième cas. L'idée sous-jacente à la thèse de Ryle est la suivante : « la connaissance théorique exprimée de manière propositionnelle (*savoir que*) n'est pas primitive, mais dépend du fait que l'on *sache* accomplir certaines activités » (Bechtel et Abrahamsen, 1991, p. 165). Les auteurs citent alors Ryle : « La pratique intelligente n'est pas l'enfance de la théorie. Au contraire, la théorie est une pratique parmi d'autres et est elle-même effectuée de manière intelligente ou stupide » (Bechtel et Abrahamsen, *ibid.*).

Cette distinction s'illustre de manière plus nette encore en aphasiologie. Nespoulous et Dordain (1994), par exemple, présentent le cas d'un patient agrammatique (« =production orale "télégraphique" avec structures syntaxiques à complexité réduite et fréquente omission des morphèmes grammaticaux », p. 267) qui, tout en présentant d'importants troubles de la production linguistique, est tout à fait capable d'effectuer des jugements de grammaticalité corrects : cette situation révèle ainsi clairement « une importante dissociation entre son "savoir" sur la langue – qui apparaît absolument intact – et son "faire" qui est massivement perturbé » (p. 271). Les dissociations présentées par les chercheurs soutiennent donc la pertinence de la distinction entre, d'une part "déclaratif" et "procédural" (« ce dernier étant de beaucoup le plus perturbé chez l'aphasique qui doit traiter en "déclaratif" ce qu'il ne peut plus traiter en "procédural" automatique », p. 276), d'autre part "traitement linguistique" et "capacités métalinguistiques", et finalement "traitement automatique" et "traitement contrôlé" (p. 276).

Le couple *savoir propositionnel / savoir faire* correspond ainsi à ce que l'on nomme, en sciences cognitives, connaissances déclaratives et procédurales, un couple qui a notoirement été élaboré et intégré par J. R. Anderson dans son modèle d'acquisition des habiletés cognitives (1983). Globalement, dans son modèle *Adaptive Control of Thought* « *star* » (A.C.T.\*), Anderson postule, que les connaissances du sujet sont initialement déclaratives (sous forme propositionnelle, verbalisables), avant d'être transformées en connaissances procédurales par paliers :

1. Interprétation : Interprétation des connaissances déclaratives en règles de production du type « Si-Alors ».
2. Compilation : Compilation des règles par réalisation répétée de la tâche, qui s'effectue en deux temps :
  - 2.1. Procéduralisation : elle permet de ne plus transiter par la mémoire de travail ;
  - 2.2. Composition : elle permet à l'exécution de gagner en rapidité.

Dans une autre perspective, certains chercheurs (Dreyfus et Dreyfus) ont proposé une version encore davantage procéduralisée de la connaissance, impliquant cinq stades de développement du savoir-faire (Bechtel et Abrahamsen, 1991, pp. 168-169) :

1. Novice : application de règles précises mais acontextuelles :
2. Débutant avancé : début de prise en compte du contexte, de reconnaissance et de modification des règles selon le contexte :
3. Exécutant compétent : construction de buts facilitant la coordination des règles et des faits, responsabilisation :
4. Exécutant accompli : utilisation de ressemblances holistiques, utilisation de souvenirs pertinents :
5. Expert : réponse intuitive (« l'expert ne manifeste plus l'aspect délibératif de la performance compétente »).

Une telle approche est commentée comme telle par Laks (1996, p. 101-102) :

*« La compétence experte [...] n'est donc plus caractérisée comme un savoir explicitement représenté sous forme déclarative et comme un ensemble de règles formelles s'y appliquant [...], mais au contraire, comme la dissolution progressive des règles au profit du stockage organisé de milliers de cas spécifiques. L'expert peut ainsi violer telle règle, sans même le savoir, pour la raison que son comportement n'est pas le produit d'une délibération explicite quant à l'applicabilité des règles, mais découle d'une perception directe et sémantiquement très riche de l'environnement. [...] Pour l'expert, ce savoir*

*procédural se tient en dessous du niveau de conscience et s'avère très difficilement exprimable dans les termes d'un discours déclaratif ».*

Ainsi, tandis que dans la conception « classique », l'acquisition du savoir est conçue comme « un processus progressif d'abstraction et de généralisation » (Laks, 1996, p. 101), dans la conception procédurale, telle que présentée par Dreyfus et Dreyfus, « l'acquisition d'une expertise semble consister à partir de règles très générales, et par là peu efficaces, à s'adapter progressivement aux contextes et aux expériences, jusqu'à dissoudre les maximes de départ dans un savoir pratique très sensible au contexte et qui ne s'exprime plus de façon déclarative » (*ibid.*).

Il apparaît cependant ici qu'au couple déclaratif/procédural est étroitement lié celui d'explicite/implicite. Le rapport entre explicite/implicite et déclaratif/procédural est présenté ainsi par Laks (1996, p. 103), qui distingue explicite structural et explicite procédural :

*« Il est nécessaire de distinguer le caractère explicite ou implicite de la représentation elle-même de celui des règles et procédures qui s'y appliquent. [...] Est explicite, non ce qui est physiquement et réellement représenté, mais ce que l'on peut mobiliser de façon immédiate ».*

Ainsi la discussion que conduit Laks à propos de l'opposition entre modélisations représentationnaliste et subsymbolique de la connaissance le conduit à la conclusion suivante (p. 103) :

*« Toutes deux voient la compétence comme inconsciente, mais la première la modélise comme un ensemble déclaratif de propositions structurellement explicites, alors que la seconde la modélise comme un savoir certes procéduralement explicite mais non directement représenté dans l'esprit ».*

Il faut néanmoins noter que les résultats des études concernant la distinction entre apprentissage (et ainsi connaissances) explicite et implicite ne sont pas convergents : il existe plusieurs points de vue en compétition concernant cette distinction (Meulemans, 1998, pp. 20-21), sa valeur heuristique et surtout les relations entretenues entre les deux types d'activité et de connaissances. De plus le contenu même des expressions « savoirs déclaratifs » et « savoirs procéduraux » varie en fonction des auteurs, qui les associent, de façon variable, aux couples savoirs/savoir-faire, savoirs conscients/savoirs inconscients, contrôle/automatisme, apprentissage explicite/apprentissage implicite. Pour Towell et Hawkins, par exemple, nous devons nous garder de confondre « savoirs déclaratifs » et « savoirs conscients », dans le cadre du modèle de Anderson au moins (1994, p. 252). Il convient donc de bien s'accorder sur le sens que l'on attribue à ces termes.

L'une des caractéristiques apparemment constantes qui permettraient de les différencier est la capacité ou l'incapacité du sujet à verbaliser ces savoirs. Des savoirs qui ne pourront pas être verbalisés seront qualifiés de *procéduraux*. Les autres seront qualifiés de *déclaratifs* (Meulemans, 1998, p. 20 ; G. Tiberghien, 1994, p. 268). Il faudra également se garder de confondre *verbalisation de savoirs* et *description verbale de comportements à la source desquels se trouvent ces savoirs*. Précisons aussi que la verbalisation peut ici englober tout autre médium, visuo-spatial en particulier, que celui de la parole : un geste, une illustration ou un schéma accompagné de flèches peuvent très facilement être codés linguistiquement et aboutir à une suite d'énoncés.

Il semble cependant également nécessaire d'insister sur l'autre critère déterminant, évoqué par Laks : celui de l'accessibilité immédiate. Dans les deux conceptions du savoir qu'il examine, « la connaissance est donc inconsciente, mais alors que dans la première, règles et représentations sont explicitement représentées et manipulées, dans la seconde, représentations et processus restent totalement implicites et ne sont pas manipulés en tant que tels » (Laks, 1996, p. 102). Les deux approches se distinguent donc, non par « le caractère conscient ou non des représentations et des processus postulés mais [par] leur caractère explicite ou implicite » (*ibid.*). S'il convient de

distinguer représentations explicites et implicites (structural), il faut aussi distinguer « le caractère explicite ou implicite [...] des règles et des procédures qui s'y appliquent [(procédural)]. [...] C'est le critère de rapidité d'accès et de disponibilité de l'information qui importe » (*ibid.*, p. 103).

Cela nous conduit au troisième couple notionnel associé aux deux premiers, à savoir le couple *automatique/contrôlé* : en fonction du degré d'attention et d'énergie qu'un processus cognitif implique pour un sujet donné (en tâche de traitement linguistique par exemple), ce processus peut en effet être qualifié d'*automatique* ou de *contrôlé*, une dichotomie établie par Shiffrin et Schneider en 1977<sup>118</sup> et qu'on retrouve constamment, sous des aspects et des dénominations divers, au sein des Sciences Cognitives. Cette dichotomie s'illustre notamment par les cas de dissociation automatico-volontaire, des cas qui sont fréquemment rencontrés en aphasiologie (Nespoulous et Dordain, 1994). Or, si le *procédural* relève, dans une certaine mesure seulement, de l'*automatique*, il n'en est pas de même pour le *déclaratif*.

Ainsi des structures (connaissances déclaratives/procédurale) aux processus (automatiques/contrôlés), deux grands types de conceptions de la connaissance émergent, que l'on ne peut simplement qualifier d'explicite ou d'implicite, au risque de les dénaturer. Ces deux approches se retrouvent dans la conception du lexique mental, qui se trouve à la base de nos connaissances linguistiques, et plus particulièrement du lexique mental phonologique qui nous intéresse au premier plan et dans lequel sont censées être encodées les représentations phonologiques des locuteurs.

#### 1.2.2.2. Le lexique mental : primitives et émergentes

Comme nous l'avons déjà évoqué à propos du stockage de l'information en mémoire, l'étude du format des représentations cognitives est un sujet d'actualité, comme en témoigne l'un des programmes de recherche actuels du Laboratoire de Psychologie

---

<sup>118</sup> « Controlled and automatic human information processing », *Psychological Review*, vol. 84, pp. 127-190.



Cognitive de l'Université de Provence (France) intitulé « Structure et fonctionnement du lexique mental ». Comme l'indique son directeur, J. Grainger, dans la présentation du programme, celui-ci vise à « spécifier l'architecture fonctionnelle du lexique mental chez le lecteur adulte, ainsi que les différents types de codage impliqués dans la compréhension et la production des mots parlés et écrits ».

Il faut en effet souligner que l'importance accordée au lexique mental tient, en partie, à sa position de « carrefour » cognitif linguistique : impliqué tant en perception qu'en production, tant en modalité auditive que visuelle, il est censé contenir l'ensemble des informations (sémantiques, morphosyntaxiques, phonologiques) propres à chaque item lexical, même si certains ont pu postuler l'existence de plusieurs systèmes lexicaux, selon la modalité et la tâche, étant donné, notamment, l'asymétrie entre modèles de compréhension et de production linguistique. Ici encore deux conceptions s'opposent, au moins au partie : une conception symbolique, et une conception émergentiste.

#### 1.2.2.2.1. Primitives phonologiques : perspective symbolique

Dans ce qu'il est aujourd'hui convenu d'appeler la conception « classique » des représentations mentales, le lexique phonologique se compose de primitives (phonèmes ou traits par exemple) à partir desquelles se construisent les formes phonologiques des représentations lexicales, lesquelles seraient stockées en mémoire et récupérables par une procédure d'adressage. Des règles déclarativement explicites s'appliqueraient à ces représentations. Si le développement de la multilinéarité en phonologie a contribué à faire progresser la conception de ces représentations, la perspective multilinéaire n'en reste pas moins symbolique, un symbolisme qui se trouve à la base même de la phonologie depuis ses origines et du développement des sciences cognitives au cours de la seconde moitié du vingtième siècle. Si le phonème, puis le trait ont longtemps fait, et font toujours, figure d'unités primitives du lexique phonologique, l'importance accordée à la syllabe ces dernières décennies ont conduit certains à l'inclure dans leur modèle de compétence linguistique (Levelt et Wheeldon, 1994). D'un point de vue phonologique, certains auteurs accordent en effet un statut représentationnel à la

syllabe, en opposition à la conception selon laquelle la syllabe émergerait d'unités primitives plus réduites. C'est le cas de Brandão de Carvalho (2002) qui, en refusant un statut symbolique à l'aspiration et au voisement, accorde un tel statut à la syllabe : « by assuming that both VOT-values and length contrasts are segmental effects of onset and nucleus weight, defined as articulatory tension versus perceptual sonority, we are assigning a representational basis to syllables » (p. 79). Rappelons néanmoins qu'il est nécessaire de distinguer syllabes phonétiques et syllabes phonologiques, ainsi que de bien spécifier le niveau d'analyse auquel l'on se situe.

#### 1.2.2.2.1.1. Production

Dans son modèle de production monolingue (appliqué aux sujets bilingues par DeBot (1992)), Levelt (1989) propose de distinguer quatre modules séquentiellement organisés (p. 27) :

1. Un *Conceptualiseur*, qui génère un message préverbal (information conceptuelle).
2. Un *Formulateur*, qui se divise en deux parties : a) Un *Encodeur Grammatical* et b) Un *Encodeur Phonologique*. Ces deux encodeurs sont reliés au lexique mental, lequel se compose de *lemmes* (information sémantico-syntaxique) et de *lexèmes* (information phonologique). Plus précisément, le lemme contient : a) l'information sémantique (sens, rôles thématiques<sup>119</sup>) et b) l'information syntaxique (catégorie grammaticale, sous-catégorisation<sup>120</sup>). Quant au lexème, celui-ci correspond à la forme phonologique sous-jacente des mots, c'est-à-dire, l'information sur leur forme segmentale et métrique. Le lemme est donc nécessaire à l'encodage grammatical, tandis que le lexème est nécessaire à l'encodage phonologique. Chaque item du lexique mental comporte ainsi quatre composants (sémantique, syntaxique, morphologique, phonologique) (p. 188) répartis entre le lemme et le lexème. L'encodeur grammatical accède aux lemmes et établit les relations

---

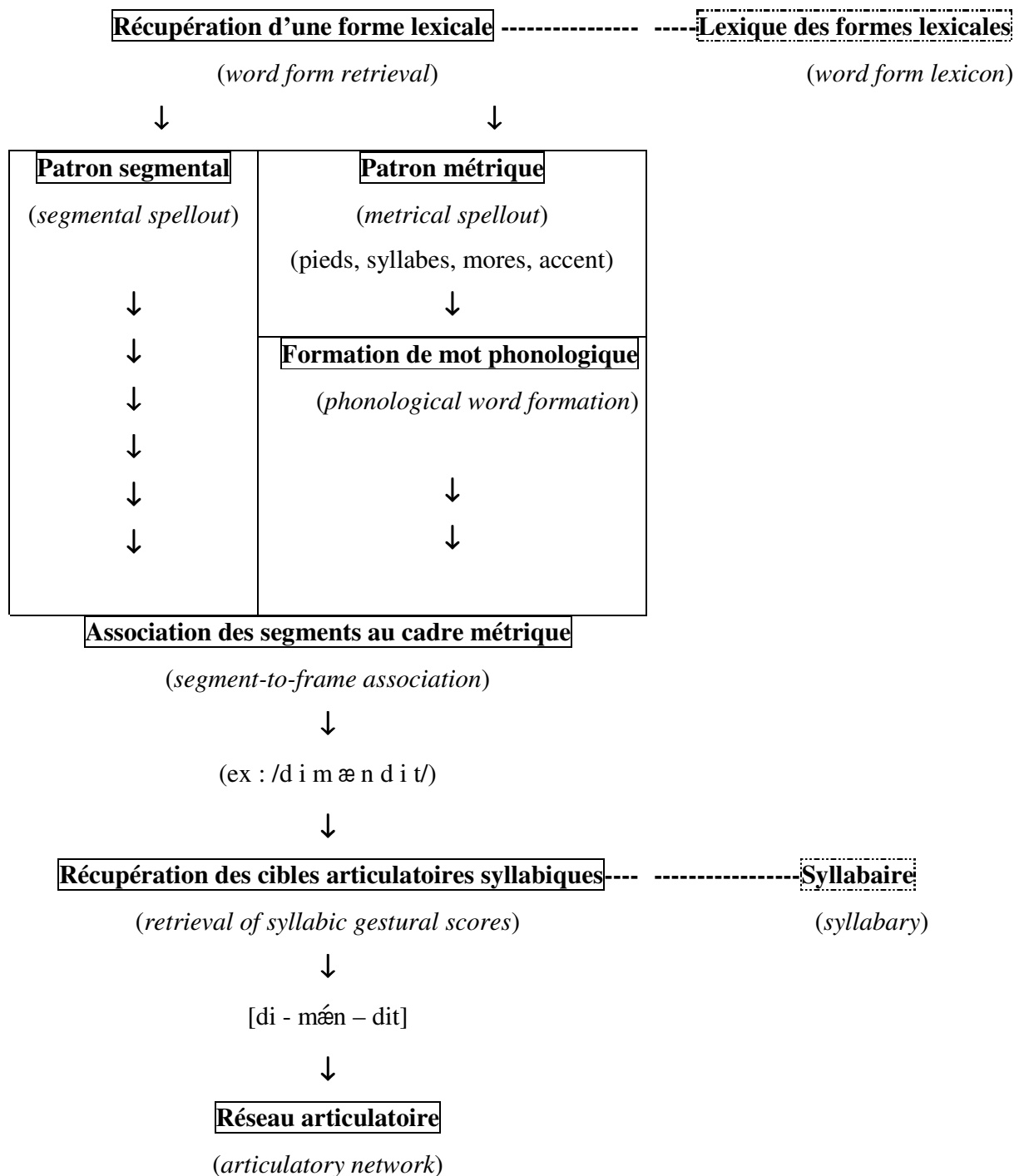
<sup>119</sup> Le rôle thématique correspond au rapport sémantique entre les groupe verbaux et nominaux d'une phrase, le plus connu étant le couple agent/thème, l'agent « effectuant l'action » et le thème (ou patient) « subissant l'action ».

<sup>120</sup> Par exemple transitif/intransitif pour un verbe, impliquant une relation sujet/objet pour le premier.

grammaticales reflétant le message préverbal, produisant ainsi une « structure de surface ». C'est sur la base de cette structure que l'encodeur phonologique crée un plan phonétique interne.

3. Un *Articulateur*, qui traduit ce plan phonétique en instructions neuromusculaires. L'étape d'encodage phonologique précède donc celle d'encodage phonétique, qui consiste à accéder aux gestes articulatoires adéquats.
4. Le *Système de Compréhension* de la parole, qui permet au locuteur d'exercer un auto-contrôle sur sa production interne et externe.

Prenant appui sur ce modèle, Levelt et Wheeldon (1994) ont proposé que les locuteurs pouvaient avoir accès à un syllabaire phonétique mental, qui serait un mécanisme de traduction d'une représentation phonologique abstraite d'un énoncé en une représentation phonétique dépendante du contexte, suffisamment détaillée pour guider l'articulation (p. 240). Cette proposition s'oppose ainsi aux conceptions traditionnelles selon lesquelles les formes phonétiques des syllabes seraient recréées lors de chaque production, car « most syllables that a speaker uses are highly overlearned articulatory gestures » (p. 246). Le cadre avancé par les auteurs repose sur l'idée que la syllabification d'un énoncé ne se réalise que lors de l'étape de production de mots phonologiques (impliquant souvent plusieurs items lexicaux), et non de mots lexicaux hors contexte, ce qu'ils nomment « connected speech » (p. 241). Pour eux, la syllabification des mots phonologiques n'est pas une « resyllabification » (qui présupposerait une syllabification lexicale préalable) (p. 244). Ce cadre de production, et ainsi d'encodage phonologique, se présente ainsi (p. 242, partiellement illustré ici par le mot phonologique anglais « demand it ») :



Concernant l'accès au syllabaire, Levelt et Wheeldon précisent d'abord ce que sont les « gestural scores », avec l'exemple du mot « apple » qui nécessite une fermeture des lèvres (p. 245) :

*« The gestural score only specifies that the lips should be closed, but not how it should be done. [...] Gestural scores are abstract. They specify the tasks to be performed, not the motor patterns to be executed ».*

Le syllabaire proposé est un ensemble fini de paires consistant, d'une part en une spécification phonologique syllabique (input), et de l'autre en une cible articulatoire syllabique (*syllabic gestural score*) (output, et input pour l'encodage phonétique final, c'est-à-dire la réalisation motrice du geste). La taille du syllabaire varierait en fonction de la langue (quelques centaines en chinois ou japonais et plusieurs milliers en anglais ou en néerlandais). Levelt (2001, p. 13465) indique que l'encodage phonétique implique également la concaténation « douce » (« smooth ») des routines syllabiques ainsi récupérées.

L'encodage phonologique présenterait donc deux étapes, successives et indépendantes (p. 247) :

- 1) Récupération du lexème,
- 2) Récupération de la cible articulatoire syllabique.

Les auteurs considèrent alors que les segments phonétiques ne sont que des propriétés d'un geste articulatoire syllabique (attaque, noyau et coda), et que ces gestes syllabiques sont stockés mentalement, du moins ceux correspondant aux syllabes utilisées régulièrement (p. 246). En effet, comme ils le notent, « An important aspect of form retrieval [...] is that it is frequency sensitive » (p. 243). La manière dont ce modèle traite le rapport segment/syllabe s'accorde ainsi notamment bien avec les théories phonologiques de la sous-spécification (p. 264) :

*« There is no need to complete the specifications of successive segments in a word if one condition is met. It is that each phonological syllable arising in the process of segment-to-frame association [...] corresponds to one and only one gestural score in the syllabary. [...] Even if a syllable's segments are underspecified, their combination can still be unique ».*

Par ailleurs, certains aspects du modèle semblent également concordants avec certaines théories récentes (MacNeilage, 1998) relatives à l'évolution du langage et de la production de la parole, qui insistent sur la séparabilité entre cadres syllabiques (*frame*) et contenu segmental (*content*), une approche en accord avec plusieurs travaux phonologiques (Brandão de Carvalho, 2002, p. 79 ; Rousset et Vallée, 2002). Déjà avancée par d'autres auteurs dans les années 80 (Shattuck-Hufnagel), c'est sur cette conception que reposent en effet les travaux neurolinguistiques de MacNeillage et ses collègues, considérant, comme il l'indique dans son résumé (MacNeilage, 1998), que « syllabic "Frames" and segmental "Content" elements are separately controlled in speech production process ». La syllabe est alors considérée comme le « réceptacle » des segments au cours de l'organisation motrice, une hypothèse soutenue par certaines observations sur les « erreurs linguistiques ».

Après avoir ainsi décrit leur cadre théorique, Levelt et Wheeldon présentent certains résultats expérimentaux penchant en leur faveur, dont les plus importants sont résumés comme tels (p. 263) :

- 1) « Syllable frequency affects naming latency in bisyllabic words »,
- 2) « The effect is independent of word frequency »,
- 3) « The effect is due to the frequency of the word's ultimate syllable »,
- 4) « Second syllable complexity does not affect naming latency, and hence cannot be the cause of the frequency effect ».

De ces résultats il ressort : 1) que les mots finissant par des syllabes fréquentes sont nommés plus rapidement que ceux finissant par une syllabe de basse fréquence, 2) que les fréquences syllabique et lexicale ont un effet additif (dans la tâche considérée).

Par rapport à ce modèle initial, Levelt (2001, p. 13465) précise que l'accès au lemme implique la récupération des codes phonologiques morphémiques : lorsqu'ils sont marqués par le pluriel, de nombreux lemmes ont en effet un code multimorphémique. Dans ce cas, un code phonologique est récupéré pour chaque morphème, comme dans l'exemple du mot anglais « horses » proposé par Levelt, qui se compose de deux morphèmes: <horse> et <iz>. L'encodage de la forme sonore (à bien distinguer de la sélection lexicale qui le précède et qui résulte dans la sélection du lemme approprié) se présente donc globalement en trois étapes :

1. Récupération des codes phonologiques morphémiques → codes phonologiques.
2. Prosodification, syllabification → mot phonologique.
3. Encodage phonétique → gestes articulatoires (« articulatory score »).

L'existence du syllabaire postulé par Levelt est cependant loin d'être reconnue par tous les chercheurs, mais cette hypothèse constitue une tentative intéressante d'intégration de la syllabe dans un modèle de production linguistique.

#### 1.2.2.2.1.2. Perception

En ce qui concerne le plan perceptif, le titre, passablement argumentatif, de la contribution de Pallier (2000), « Word recognition : do we need phonological representations ? », renvoie précisément à la problématique du lexique mental, comme l'indique la question qu'il pose dans son résumé :

*« Under what format(s) are spoken words memorized by the brain ? Are word forms stored as abstract phonological representations ? Or rather, are they stored as detailed acoustic-phonetic representations ? ».*

Tandis que de nombreux chercheurs, linguistes ou psycholinguistes se sont depuis longtemps attachés à établir la réalité psychologique de telle ou telle représentation phonologique, d'autres ont envisagé l'existence de représentations autres, en particulier de représentations acoustiques associées à chaque mot, dans une perspective exemplariste et non abstractionniste. Certains ont également considéré la possibilité de distinguer entre production et perception de la parole : les représentations acoustiques (issues de l'input) ne seraient pertinentes qu'en perception, alors directement superposées (en anglais *mapping*) aux représentations lexicales. Pallier présente ainsi une série d'expériences visant à tester une telle hypothèse, à l'aide de sujets catalans et espagnols (vocabulaire identique, mais systèmes phonémiques différents) d'une part, et en référence à une population de sujets japonais (Dupoux et *al.*, 1999) d'autre part. Les résultats le confortent dans l'hypothèse qu'il avait développée auparavant (Pallier, 1994) : le système perceptif semble bâtir une représentation phonologique propre à la langue première des sujets, dans laquelle la structure syllabique jouerait un rôle de code correcteur perceptif pour des stimuli ambigus ou incomplets phonétiquement. Si ces résultats semblent aller dans le sens des approches abstractionnistes, Pallier note cependant que les modèles acoustiques pourraient également, sous certaines conditions, en rendre compte. Comme il l'indique en conclusion, les deux approches pourraient en effet être réconciliées, notamment *via* la notion de « représentations phonologiques gradées » : « and then, one must admit, the distinction between acoustic and phonological models starts to blur... ».

#### 1.2.2.2. Emergentisme et connexionnisme : perspective subsymbolique

Depuis une vingtaine d'années environ, la conception des représentations cognitives s'est trouvée de plus en plus sujette à de nouveaux débats, notamment sous la pression des modélisations de type connexionniste. Dans la conception connexionniste de la sémantique des représentations mentales, le concept d'*émergence* occupe une position centrale, car « la valeur sémantique est d'abord contextuelle et le contenu « intrinsèque » émerge de la récurrence systématique des états subsymboliques » (Laks, 1996, p. 78). La perspective connexionniste est donc bien une perspective de



construction du sens (par opposition aux primitives sémantiques conceptuelles), ascendante et contextuelle. Notons que, d'un point de vue didactique, celle-ci semblerait justifier ainsi en partie, d'une part les approches grammaticales inductives en classe de langue, et d'autre part ce que les didacticiens nomment la "grammaire en contexte": « l'usage et l'expérience, par leurs régularités et parce que la récurrence des événements fait apparaître des structures statistiques de dépendance, permettent la construction (ascendante) de catégories et de concepts abstraits » (Laks, 1996, p. 79). Si telle est la position adoptée vis-à-vis des représentations sémantiques et grammaticales, il en est de même pour les représentations phonologiques (Laks, 1996, p. 79) :

*« Dans le cadre que nous défendons, la notion de phonème, d'atome ou de caractéristique phonologiquement pertinente peut être conçue comme le résultat d'un processus de prototypage des régularités perceptives et motrices. [...] Les catégories de la phonétique universelle [...] émergent [...] par catégorisation du flux informationnel, tant en production qu'en réception ».*

Nous retrouvons là l'idée sous-jacente à la thèse de Ryle, que développent ainsi Bechtel et Abrahamsen : « Si le connexionnisme cesse de se servir de chaînes de symboles dans la modélisation, peut-il encore expliquer la capacité humaine d'affirmer des propositions et d'étayer leur vérité par des arguments ? Si c'est le cas, *le savoir propositionnel* pourrait être compris comme un cas particulier du *savoir-faire* » (Bechtel et Abrahamsen, 1991, p. 167).

Laks résume bien, nous semble-t-il, l'opposition entre les deux conceptions dominantes (extrêmes) du lexique mental :

- Modélisation représentationnaliste (de type lexicographique) :

*« Le lexique mental est alors conçu comme la mémorisation d'une longue liste d'items, un dictionnaire, spécifiant pour chaque entrée représentation phonologique, catégorie*

*syntactique d'appartenance, propriétés sémantiques et contraintes syntagmatiques liées à l'usage de cet item. Ces représentations lexicales [...] sont par nature acontextuelles* » (p. 103).

- Modélisation subsymbolique (de type réseau récurrent) :

*« Les mots, au sens d'entrée dans une liste, n'existent pas parce qu'il n'y a, à proprement parler, pas de remémoration d'un savoir acontextuel stocké indépendamment. Les mots sont toujours réactivés dans un contexte particulier à partir des traces mnésiques que constituent les connexions pondérées par l'expérience. En tant qu'états mentaux ainsi réactivés, ils correspondent à des indices interprétatifs adéquats à l'analyse d'une situation donnée [...]. S'ils ont une existence indépendante (i.e. lexicale), il ne faut y voir qu'un effet secondaire de leur récurrence au même titre qu'un prototype ne correspond qu'à la partie invariant de toutes ses actualisations. [...] Encore faut-il ajouter que la capacité pour un locuteur à reconnaître des unités discrètes appelées mots est fortement tributaire d'un apprentissage explicite, celui de l'écriture et de la lecture »* (pp. 103-104). Une telle approche soutient donc *« l'existence d'états mentaux qui sont procéduralement explicites sans être physiquement représentés »* (p. 104).

Cette approche non-représentationnaliste est notamment illustrée par la contribution de Coleman, « *Phonetic representations in the mental lexicon* » (2002). Le point de vue de l'auteur est annoncé dès la première page, puisqu'il se propose de défendre les arguments suivants (p. 96) :

- 1) « *Our mental representations of the forms of words are essentially phonetic, rather than symbolic-phonological* »,
- 2) « *Phonological competence includes and makes use of statistical properties of the time-course of phonetic representations* »,
- 3) « *A combination of phonetic, statistical, and semantic knowledge is sufficient to explain many aspects of phonological structure* ».

Il s'agit donc d'une conception « épisodique » des représentations lexicales, motivée par la question suivante (p. 97) :

*« If we entertain the existence of non-abstract, superficial phonetic representations [...], are there any phonological phenomena which cannot be described or formulated using phonetic representations, and which require us also to employ more abstract, phonological representations ? ».*

Afin de répondre à cette question, et après avoir passé en revue plusieurs types de travaux soutenant l'existence d'unités phonologiques abstraites (syllabes, pieds, composants syllabiques, dans une certaine mesure traits, et de manière beaucoup moins assertive, phonèmes), Coleman examine d'autres données (psycholinguistiques, sociolinguistiques, développementales) qui le conduisent à proposer une version alternative aux modèles phonologiques symboliques traditionnels, lesquels semblent négliger certaines informations pourtant apparemment présentes dans le lexique mental : « Taking these studies together, it is difficult to escape the conclusion that very specific pronunciation details, such as speaker-specific detail, and word frequency, are encoded in the lexicon » (p. 125). Il semblerait en effet que la connaissance des formes lexicales inclue :

- 1) « Knowledge of fine phonetic details not normally reckoned among the inventory of features »,
- 2) « Knowledge of word frequencies, and frequencies of part of words »,
- 3) « Knowledge of the statistical patterns of phonetic variation and its correlations with linguistic and non-linguistic factors, such as structural context, word frequency, communicative situation, social identity, and other factors » (p. 125).

L'approche « néo-empiriste » défendue par Coleman, visant à mieux incorporer ces données dans la modélisation du lexique mental, et à établir ainsi une théorie phonologique cognitivement réaliste essentiellement orientée « bas-haut », repose sur trois idées principales :

- 1) « Word forms are stored as memories of psychophysical (auditory and articulatory) experience (not abstract structures of distinctive features) ».
- 2) « Phonological constituents are statistical regularities over those psychophysical spaces ».
- 3) « Some phonological abilities follow from the manner of storage (e.g. the discrimination of contrastive words); other phonological properties must be computed on-line (e.g. phoneme monitoring) » (p. 126).

Cette approche s'accorde ainsi très bien avec certaines modélisations connexionnistes, lesquelles permettent un fort degré d'interconnectivité entre « espaces psychophysiques », résolvant par-là même le problème posé par Coleman (p. 128) :

*« Auditory representations alone are not sufficient for encoding phonological knowledge. In order to encode the semantically contrastive function of differences between the sounds of different words, the auditory space must be associated with semantic spaces [...]. The phonetic space may be quasi-continuous, but in places it is effectively carved up by virtue of these associations to discrete semantic categories ».*

L'auteur évoque également d'autres associations nécessaires à la constitution d'un modèle phonologique cognitif (auditif/articulatoire, sémantique/articulatoire, auditif/visuel-orthographique, etc.), qui renvoient à une conception modale des représentations lexicales : « Rather than seeking for a single kind of word-form representation, we are not surprised by the wealth of evidence for a variety of many different kinds of modality-specific representations, and various paths of association between them » (p. 130).

#### 1.2.2.2.3. Parenthèse didactique

Ayant mis à jour les débats actuels relatifs à la nature et au format du lexique mental phonologique parmi les spécialistes de la cognition, nous souhaitons ici reprendre

brièvement la perspective didactique qui est en partie la nôtre, et adopter un point de vue tout à fait généraliste. De manière globale, les représentations en réseau semblent connaître à l'heure actuelle un engouement certain dans plusieurs domaines de la société. Elles constituent, comme l'ont été en leur temps les représentations non-linéaires en phonologie par exemple, des tentatives d'ouvrir de nouveaux chemins d'accès, de repousser les limites conceptuelles, spatiales, mais aussi temporelles, avec l'activation de processus parallèles et rétroactifs. Les réalisations techniques (simulations informatiques, Internet), les échanges mondiaux (communication et circulation), ainsi que l'émergence des neurosciences et la mise en valeur des architectures fonctionnelles cérébrales ont par ailleurs contribué au développement de cette tendance. Comme l'écrivait Rumelhart dans une préface en 1991 « the last several years have witnessed a remarkable growth in interest in the study of brain-style computation » (Ramsey, Stich et Rumelhart, 1991, ix) et cet intérêt s'est concentré dans ce qu'on peut nommer le connexionnisme, attirant tout autant des chercheurs en neurosciences, des ingénieurs, des spécialistes en intelligence artificielle, des mathématiciens, des psychologues ou encore des philosophes, chacun y trouvant un intérêt commun ou spécifique à son domaine. Il existe cependant différents modèles, et différentes théories connexionnistes. Nous ne nous attachons donc pas à l'un ou à l'autre, mais plutôt aux perspectives globales qui en découlent, en considérant « le connexionnisme en tant qu'*idée* de ce qu'est l'esprit » (Bechtel et Abrahamsen, 1991, p. 280).

Le connexionnisme, en tant que tradition intellectuelle, fournit un cadre scientifique et ainsi un certain paradigme : « un ensemble d'hypothèse, de constructions, de techniques et d'objectifs qui guident la recherche » (Bechtel et Abrahamsen, 1991, p. 272, citant T. S. Kuhn, *La structure des révolutions scientifiques*, Paris, Flammarion, traduit par L. Meyer, 1983). Bechtel et Abrahamsen (1991, pp. 287) présentent ainsi de manière condensée quelques propriétés du connexionnisme :

- 1) Une manière différente d'aborder les règles et les représentations, qui est moins fragile que l'approche symbolique traditionnelle.
- 2) L'extraction de traits adaptée au moyen d'unités cachées.

- 3) L'implémentation de contraintes souples multiples.
- 4) La résistance au dommage et autres caractéristiques liées à la plausibilité neuronale.
- 5) « La forme d'apprentissage la plus patente est l'ajustement des poids des connexions » (p. 288).

Une autre caractéristique essentielle des représentations en réseaux tient au caractère distribué de l'information (Laks, 1996, p. 109) :

*« Dans un réseau de neurones, la connaissance est entièrement distribuée sur l'architecture du réseau et l'état stationnaire atteint résulte de l'interaction d'un très grand nombre de processeurs interconnectés fonctionnant en parallèle. Lorsqu'une partie du réseau est détériorée, détruite ou coupée de l'ensemble, cela ne grève les performances du système que dans la proportion où la partie affectée y contribuait. On a une dégradation progressive [...] [qui] est liée au comportement holistique d'un système où chaque unité participe à la création ou la re-création d'un schéma, mais où aucune ne le représente vraiment. ».*

Cependant, tous les chercheurs ne s'accordent pas sur la définition ou la valeur qui doit être accordée au concept de « distribution » : Van Gelder (1991), par exemple, critique l'une des définitions principales de la distribution telle que définie par Hinton, McClelland et Rumelhart : « Each entity is represented by a pattern of activity distributed over many computing elements, and each computing element is involved in representing many different entities ». Pour Van Gelder, cette définition est trop limitée, car il y a au moins deux types de schémas : « the patterns of activity themselves, and the patterns of connectivity that mediate their transformation » (1991, p. 36). Il critique également l'imprécision que lui donnent certains auteurs : « Distributed systems or representations are often described as *holistic*. This is an extraordinarily vague term, and usually contributes nothing to our understanding of the phenomenon in question »

(p. 52). A la question « what is distribution ? » Van Gelder répond finalement, en prenant certaines précautions et en constatant que le thème de la superposition<sup>121</sup> est celui qui est le plus commun à toutes ses caractérisations, que : « distribution is the superposition of representings, and distributed representations are those which belong to schemes defined around a core method of generating superposed representations. [...] A distributed representation must also be more than just a « pattern » over some extended area, for virtually any representation will count as a pattern of some kind » (*ibid.*, pp. 53-54). Par ailleurs, il convient de noter que le concept de distribution n'est pas systématiquement invoqué par l'ensemble des chercheurs travaillant dans un cadre connexionniste, comme cela apparaît dans le titre de Bowers (2002), « Challenging the widespread assumption that connectionism and distributed representations go hand-in-hand ».

Ayant ainsi bien trop brièvement survolé la problématique non-représentationnaliste, il convient de s'interroger sur la valeur, pragmatique, de celle-ci pour les praticiens de la didactique des langues. On peut dans un premier temps affirmer, sans rentrer dans la moindre considération technique, que les didacticiens et enseignants de langue ont tous une certaine *représentation* de la manière dont fonctionne le cerveau, et plus encore l'être humain, représentations forgées sous une « pression informationnelle », issue de leur éducation, de leur culture générale, de l'environnement médiatique et des magazines de vulgarisation scientifique. Or le modèle qui prédominait jusqu'au début des années 80 était lié à la métaphore de l'ordinateur « à la Von Neumann », traitant les processus cognitifs « as rule governed symbol manipulation » (Ramsey, Stich et Rumelhart, 1991, xi). On pourrait ainsi considérer, d'un point de vue épistémologiquement *naïf*, qu'après avoir continuellement « fractionné » l'objet d'étude pour mieux en étudier les parties dans une approche structuraliste/modulariste

---

<sup>121</sup> Pour la différence entre « superposed » et « superimposed », voir Van Gelder, 1991, p. 43 : « When two representations are effectively superposed they become a single new item representing both the original contents ».

issue d'une tradition facilement qualifiable de cartésienne<sup>122</sup>, l'idée connexionniste pourrait être envisagée comme une reconnexion de toutes ces parties, dans un retour au « tout ». En ce qui concerne la distinction entre *structure* et *fonction*, Bechtel et Abrahamsen notent : « Il est important de se rappeler qu'une entité qui est considérée comme une structure du point de vue d'une discipline donnée peut elle-même être considérée comme le produit fonctionnel d'autres structures dans une autre. la distinction structure-fonction est donc relative » (1991, p. 303). Bien entendu, les querelles entre modularistes et connexionnistes « pure laine » perdent leur intensité si l'on adopte une position médiane en considérant que leur différend de surface ne provient en essence que d'une différence de perspective et surtout d'échelle. Les modèles envisagés sont alors « mixtes » ou « interactifs », essentiellement connexionnistes et fonctionnellement modulaires. Cependant, l'influence de ces représentations sur les pratiques et les constructions théoriques didactiques ne sont pas négligeables, bien au contraire, comme cela apparaît nettement d'un point de vue diachronique, lorsque l'on examine l'évolution des relations entre didactique des langues, linguistique et psychologie cognitive. La question que nous souhaitons soulever est alors la suivante : quel serait l'impact, en didactique des langues, d'un changement majeur des représentations (méta-)cognitives des enseignants et didacticiens, passant d'une perspective symbolique à une perspective subsymbolique ? Qu'en serait-il, par exemple, si, suivant la tendance, nous tentions de donner une représentation « en réseaux » aux modèles de compétence de communication ? Que peuvent apporter les perspectives connexionnistes à la didactique des langues, un domaine tout jeune, dans lequel les sciences cognitives non-connexionnistes elles-mêmes sont encore finalement assez peu sollicitées ? C'est notamment d'une question de formation initiale des futurs enseignants et futurs chercheurs qu'il s'agit. La problématique de la « grammaire », par exemple, peut facilement prendre un nouveau tour à la lecture de ce qu'écrit Laks (1996, p. 100) : « ce qui se laisse décrire comme un savoir n'est pas la cause mais la conséquence de la régularité des activités du substrat

---

<sup>122</sup> Cf. le second principe de Descartes, situé dans la seconde partie du *Discours de la méthode pour bien conduire sa raison, et chercher la vérité dans les sciences* : « diviser chacune des difficultés que j'examinerois, en autant de parcelles qu'il se pourroit, et qu'il seroit requis pour les mieux résoudre ».



physique. Les principes et paramètres qui expriment cette compétence sont donc des produits, qui peuvent éventuellement contraindre rétroactivement les productions, et non en eux-mêmes des principes de production. ». Si l'on considère la problématique des catégories (sémantiques ou grammaticales), la flexibilité des systèmes en réseaux est également à rapprocher de l'évolution des modèles de catégorisation (Mondada et Dubois, 1995, pp. 278-279) :

*« Les catégories sont généralement instables, variables et flexibles. [...] Depuis la théorie des prototypes de Rosch, les approches de la catégorisation ont évolué du cadre philosophique classique traitant de la formation des concepts vers une perspective plus écologique qui considère que l'organisation de la connaissance humaine est motivée par des buts adaptatifs. Une telle conceptualisation [...] introduit des frontières floues entre catégories plutôt que sur des décisions tranchées d'appartenance catégorielle ».*

On remarquera néanmoins que la problématique connexionniste a déjà commencé à pénétrer « en force » les recherches sur l'acquisition des langues étrangères, essentiellement dans le domaine anglophone de « Second Language Acquisition », et notamment chez les chercheurs se situant dans un cadre variationniste et/ou sociolinguistique (Tarone, 2002). Signalons en particulier les travaux de N. Ellis (1998, 2002) qui, à notre connaissance (voir aussi Broeder et Plunkett, 1994), fait partie des chercheurs qui ont le plus exploré ces pistes de recherche en SLA, mettant à jour différents axes d'études pour les didacticiens dans cette perspective.

## **II. Perception de la parole et multimodalité**

De manière générale en situation de perception, l'information sensorielle externe est saisie de différentes manières par l'organisme, qui intègre ensuite ces informations de sources multiples à des niveaux cognitifs supérieurs. L'architecture neuro-cognitive de ces processus est d'ailleurs de mieux en mieux connue (Mesulam, 1998) : des aires sensorielles primaires vers les aires associatives unimodales, l'information est ensuite

transférée vers les aires associatives hétéromodales, puis vers les niveaux supérieurs. Il semblerait d'ailleurs, comme l'indiquent Driver et Spence (2000), que la question de la convergence et de l'intégration multimodale soit, en neuropsychologie, un sujet d'actualité (voir par exemple Calvert, Campbell et Brammer, 2000), après plusieurs décennies consacrées à l'identification de modules spécialisés et distincts les uns des autres. Un tel regain d'intérêt est dû à la mise en avant du fait que « understanding the interplay between components in an extended network is as important as fractionating that network into its component parts » (Driver et Spence, 2000, p. 731). Les travaux de Massaro sur l'intégration des informations multimodales constituent à cet égard une version particulière d'une idée plus générale, résumée comme telle par Driver et Spence (2000, p. 732) : « crossmodal integration must be produced by convergence of separate information sources, with suitable weighting of each as they are combined ». Cela rend compte du fait que la dominance d'une modalité par rapport à une autre puisse varier en fonction des circonstances. Les études récentes dans le domaine, faisant appel à l'imagerie cérébrale fonctionnelle, suggèrent en outre que, si en termes d'information afférente, les aires cérébrales dites unimodales, le sont en effet, celles-ci subiraient néanmoins une influence de type haut-bas en provenance des aires multimodales Driver et Spence (2000, p. 734) :

*« Back projections from multimodal convergence areas could result in responses to the primary modality in « unimodal » brain areas being modulated by stimulation in a second modality ».*

Cette idée est également celle suivie par Shimojo et Shams (2001), qui passent en revue différentes études mettant à jour d'importantes interactions entre les modalités sensorielles, et réexaminent ainsi d'un point de vue critique le « dogme » de la modularité (p. 508) :

*« Interaction between modalities is the rule as opposed to the exception in brain function, and brain imaging and recording studies provide evidence*

*against modularity and for interaction in areas traditionally thought to be unimodal ».*

Parmi ces arguments, Shimojo et Shams distinguent ceux ayant trait :

- A la plasticité cérébrale : au cours des premières étapes du développement, les modalités sensorielles ne seraient pas aussi indépendantes et intrinsèquement distinctes qu'on le pensait autrefois : « sensory deprivation in one modality starting from an early period of life causes the cortical area that is normally devoted to that modality to be used by some other modality » (p. 505).
- Aux interactions entre modalités : influence de la vision sur la perception auditive (effet McGurk ; ventriloquisme), influence du son sur la perception visuelle (dimension temporelle ; ambiguïté spatio-motrice)

En ce qui concerne la direction de l'influence intermodale, Shimojo et Shams émettent l'hypothèse que celle-ci est en partie déterminée par la nature des stimuli perçus (p. 508) :

*« The modality that carries a signal which is more discontinuous (and hence more salient) becomes the influential or modulating modality. [...] The transient/discontinuous signals possess a special status in determining the direction of cross-modal interactions, with the transient stimulus strongly influencing the multimodal perception, regardless of the modality in which it occurs ».*

## **II.1. Influences intermodales**

Si le canal auditif semble être, pour des normo-entendants, le médium privilégié de la perception de la parole, celle-ci n'y est certainement pas limitée, recourant tant au canal visuel qu'au canal tactile. Comme l'écrit Massaro (2001, p. 14872) :

*« Speech perception has traditionally been viewed as a unimodal process, but in fact appears to be a prototypical case of multimodal perception. This is best seen in face-to-face communication ».*

Les influences des modalités entre elles en perception de la parole constituent donc pleinement un sujet d'actualité (Fowler et Dekle 1991 ; Massaro, 1998). Dans le cas de la parole, le deuxième canal qui mérite une mention particulière est le canal visuel. De l'œil au cerveau, on passe du champ de ce que l'on appelle la vision de bas-niveau (*low-vision*), qui consiste à récupérer de l'information sur les caractéristiques physiques de l'environnement, telles que l'orientation, la profondeur, les frontières entre objets, le mouvement dans l'espace, etc. (Yuille et Ullman, 1990) à celui des représentations issues de ces premières étapes de traitement (vision de haut-niveau, *high-level vision*) (Biederman, 1990) et, par la suite, à celui de l'imagerie mentale (Kosslyn, 1990 ; Denis, 1994). Il convient évidemment de ne pas confondre l'objet, la perception de cet objet et l'image mentale qui en découle. Cependant on peut établir des relations de similitude, fonctionnelle et structurelle, entre eux (Denis, 1994), comme cela se justifie d'un point de vue neuro-cognitif : « the neural substrates for mental imagery appear to include the same areas that would have supported the corresponding acts of perception if the imagined scene were actually unfolding in the external world » (Mesulam, 1998, p. 1034). La reconnaissance visuelle implique plusieurs étapes de traitement, et l'on peut par exemple adopter l'analyse de Riddoch et Humphreys (cités par Bruyer, 1994, p. 199) qui considèrent six opérations :

1. Détection des composantes de formes, lignes, bords, etc., ainsi que leur localisation et orientation, sous différentes échelles d'analyse ;
2. Mise en relation de ces éléments pour créer des ensembles et dissocier la figure du fond ;
3. Repérer l'orientation de l'objet entier et pouvoir le reconnaître sous différents points de vue ;
4. Accès en mémoire à une représentation structurale de l'objet ;
5. Accès en mémoire à une représentation de la fonction de l'objet et des objets auxquels il est souvent associé ;

6. Comparaison entre les informations perceptives (1, 2, 3) et celles récupérées en mémoire (4, 5).

La pertinence, pour l'étude de la perception de la parole, des données sur la perception visuelle n'est cependant pas neuve, puisque l'un des fondements du domaine est le célèbre « effet McGurk », mis à jour dans les années 70 : la perception auditive d'un [ba] couplée avec la perception visuelle de mouvements articulatoires correspondant à un [ga] donne l'illusion perceptive d'un [da]. On distingue ainsi aujourd'hui des mouvements audibles et visibles, audibles non visibles et non audibles mais bien visibles (Abry et Schwartz, 1997, p. 107). Or, l'effet McGurk illustre parfaitement le fait que la question de la multimodalité perceptive doit être reliée à la problématique, beaucoup plus générale, de la concurrence catégorielle : « La variation et la concurrence catégorielle émergent notamment lorsqu'une scène est vue de différentes perspectives, qui impliquent différentes catégorisations de la situation, des acteurs et des événements » (Mondada et Dubois, 1995, p. 280).

Notamment issue des travaux sur la réhabilitation des malentendants et sur la lecture labiale, l'étude de l'influence des représentations visuelles, en particulier manuo-faciales, sur la perception de la parole occupe ainsi aujourd'hui de nombreux chercheurs à travers le monde, motivés par le développement des outils informatiques (par exemple Massaro et Light, 1997 ; Attina, Beautemps et Cathiard, 2002 ou encore Muller et Prelaz, 1994). On peut ainsi mentionner les travaux de l'Institut de la Communication Parlée (IPC) de Grenoble (France) qui y consacre, sous diverses formes, plusieurs projets de recherche : par exemple l'association de représentations gestuelles à une tête parlante audiovisuelle virtuelle déjà existante à l'IPC, dispositif destiné à la réhabilitation des déficients auditifs. Comme l'indique son responsable, D. Beautemps, dans la présentation du projet, « la bimodalité de la parole audiovisuelle se situe au cœur même du mécanisme de la communication parlée. [...] Les objectifs de la lecture labiale [...] sont de retrouver la réalité motrice de la parole, c'est-à-dire les mouvements des articulateurs, non pas simplement à partir de leur conséquences audibles [...] mais aussi à partir d'autres voies sensorielles telles que la vision ». Les

chiffres qu'il donne parlent d'eux-mêmes : « la lecture labiale seule permet d'appréhender 40 à 60% des phonèmes d'une langue donnée ». Le projet en question combine alors deux techniques : la *Réalité Augmentée*, qui permet d'afficher certains des articulateurs non visibles sur un locuteur humain naturel, et *le Langage Parlé Complété* (équivalent français du *Cued Speech*), qui permet de saisir à l'aide de la main et des doigts, des informations non-visibles (voisement, nasalité, mouvements de la langue), complétant ainsi la lecture labiale.

Si la lecture labiale semble ainsi constituer un outil indispensable aux malentendants, elle l'est également pour les sujets normaux en situation de parole bruitée : il s'agit des trois rôles attribués à la lecture labiale par Summerfield, l'un des chercheurs majeurs du domaine dans les années 80, cité par Grantz et Seitz (2000, pp. 1197-1198), qui se sont, quant à eux, attachés à tester, avec un succès relatif, la partie « temporelle » du troisième :

1. La lecture labiale fournit une information (segmentale et suprasegmentale) qui est *redondante* par rapport aux indices acoustiques (information acoustique également présente).
2. La lecture labiale fournit une information qui est *complémentaire* par rapport aux indices acoustiques (information acoustique absente, en raison du bruit ou d'interférences dans le signal).
3. La coïncidence spatiale et temporelle du signal visuel et acoustique permet de diriger l'analyse acoustique vers l'information pertinente plutôt que vers le bruit environnant.

D'un point de vue neurolinguistique, l'influence de la lecture labiale a été testée par certains chercheurs à l'aide des techniques d'imagerie cérébrale fonctionnelle (Calvert, Bullmore, Brammer, Campbell, Williams, McGuire, Woodruff, Iversen et David, 1997). Il apparaît alors que ces indices visuels suffisent pour activer le cortex auditif d'un locuteur normo-entendant, même en l'absence de stimuli auditifs. Lorsque les stimuli visuels faciaux sont non-linguistiques (avec la bouche fermée), le cortex auditif n'est pas activé, mais il l'est néanmoins lorsqu'il s'agit de stimuli visuels pseudo-

linguistiques (mouvements ressemblant à des mouvements linguistiques, mais sans signification). Cela conduit les auteurs à conclure que la lecture labiale silencieuse, qui active donc les mêmes zones que celles engagées dans la perception de la parole, module la perception de la parole à un niveau pré-lexical (p. 595).

En ce qui concerne l'intégration temporelle des informations auditives et visuelles dans une situation perceptive conduisant à un effet « McGurk », Munhall et Tohkura (1998) ont souligné la complexité de cette intégration (p. 537) :

*« Information from the visual and auditory modalities is not synchronous and does not unfold continuously at the same rate. Thus it seems unlikely that the perceptual system uses temporal coincidence or any tight cross-modal timing as the basis of audiovisual integration ».*

Cette conclusion semblerait soutenue par d'autres travaux jouant sur la synchronie des stimuli : par exemple, malgré un délai de 180 millisecondes du signal acoustique par rapport au signal visuel, l'effet McGurk n'est pas perturbé. Cependant, cette résistance à la désynchronisation est limitée, en particulier lorsque le signal acoustique précède le signal visuel. L'effet semble en effet davantage perturbé lorsque le signal auditif est premier, un phénomène expliqué par certains comme étant dû au caractère naturellement premier des mouvements articulatoires par rapport au son, mais que Munhall et Tohkura préfèrent expliquer par une question de seuil (p. 537) :

*« The limitation is not time per se but rather the information processing of the auditory information ; visual information might not influence perceptual categorization if the auditory information has reached a criterion threshold. The more natural temporal precedence of visual information may have led to a priming role for the visual signal in normal audiovisual perception ».*

Si l'influence de certaines fonctionnalités d'ordre visuel sur la perception de la parole semble donc à présent bien établie, l'influence de la modalité auditive sur la perception

visuelle a en revanche beaucoup moins été étudiée. Bien qu'il ne s'agisse pas de perception linguistique, on peut ainsi mentionner certains travaux récents, qui, une fois de plus, soulignent l'importance des interactions possibles entre modalités perceptives lors de l'intégration de l'information.

Ainsi, Shimojo et ses collaborateurs ont récemment mis en évidence une capacité de modulation de la perception visuelle par des stimuli auditifs (Shams, Kamitani, Shimojo, 2000) : lorsqu'un flash lumineux unique est accompagné de multiples bips auditifs, ce flash est illusoirement perçu comme une série de flashes multiples. Cet effet persiste tant que les deux types de stimuli se situent dans un intervalle de 100 millisecondes, et il a également été testé d'un point de vue neuropsychologique en utilisant la technique des potentiels évoqués (ERP) (Shams, Kamitani, Thompson et Shimojo, 2001). Les chercheurs ont comparé les potentiels visuellement évoqués en présence et en absence de stimuli auditifs, et il est apparu que les potentiels correspondant aux flashes illusoires étaient qualitativement similaires à ceux correspondant à un flash réel. Cet effet vient s'ajouter aux autres effets étudiés jusqu'alors : l'influence de stimuli auditifs sur la perception de l'intensité visuelle (influence quantitative et non qualitative) d'une part et sur la perception de la direction d'un mouvement visuel ambigu d'autre part. Les travaux de Shimojo et de ses collaborateurs indiquent donc que le son peut altérer qualitativement la perception visuelle, même lorsque le stimulus visuel n'est pas ambigu. Ces travaux ont ensuite été étendus, en considérant d'une part la dimension temporelle (résolution) de la perception visuelle, qui semblerait pouvoir être améliorée ou dégradée selon les stimuli auditifs associés (séquence et délais), et d'autre part la perception du mouvement de deux objets orientés l'un vers l'autre, continu ou de rebond, selon le synchronisme ou l'asynchronisme des stimuli auditifs (Watanabe et Shimojo, 2001). Au vu de l'ensemble de ces résultats, l'hypothèse des auteurs est la suivante : « audition affects vision particularly when there are sharp transient signals which are adjacent to a particular visual event in time » (Shimojo, Scheier, Nijhawan, Shams, Kamitani et Watanabe, 2001, p. 65). L'ensemble de ces travaux souligne donc la nécessité



d'analyser en profondeur les relations entre les différents supports impliqués en tâche de perception, auditive, visuelle et audiovisuelle.

## **II.2. Théories et Modèles**

Nous présentons ici deux grands types de travaux relatifs à l'intégration d'informations multimodales en perception de la parole :

- Dans le domaine anglophone, ceux, fondateurs dans une certaine mesure, de Massaro, conduisant au « modèle flou » de perception de la parole ;
- Dans le domaine francophone, ceux, davantage liés à la phonologie, de Schwartz, Abry et leurs collaborateurs, relatifs à la théorie de « la perception pour le contrôle de l'action ».

### **II.2.1. Le modèle flou (Massaro)**

Comme l'explique Massaro (1987), le Modèle Flou (*Fuzzy Logical Model of Perception, FLMP*) a notamment été élaboré au vu des limitations du modèle de perception catégorielle (CP). Passant en revue plusieurs arguments soutenant sa position, en particulier dans le domaine du traitement de l'information auditive, il indique : « Although most of our actions reflect a categorical partitioning of the environment, perception of the instances within a category is not necessarily categorical » (p. 254). Il oppose ainsi perception *catégorielle* et perception *continue* de la manière suivante (p. 255) :

*« Continuous perception refers to a relatively continuous relationship between changes in a stimulus and changes in the perceptual experience of that stimulus. In contrast, categorical perception refers to a mode of perception in which changes along a stimulus continuum are not perceived continuously, but in a discrete manner ».*

La base des travaux sur la CP repose sur un ensemble d'expériences conduisant à établir une correspondance entre identification et discrimination : les stimuli ne

peuvent être discriminés qu'à condition d'être identifiés comme différents. Or, pour Massaro, il est nécessaire de distinguer la perception catégorielle des résultats de perception catégorielle (« categorical perception » vs. « categorical-perception results ») (p. 257), car les processus de décision peuvent transformer de l'information sensorielle continue en résultats généralement considérés comme reflétant une perception catégorielle (p. 257) :

*« "categorical perception" is defined in terms of whether the sensory system makes available continuous or categorical information, whereas "categorical-perception results" refer to the accuracy with which identification predicts discrimination. There are alternative explanations of CP results that do not depend on CP ».*

Parmi ces alternatives, il mentionne les limitations de la mémoire auditive en tâche de discrimination (« the failure to discriminate two sounds may simply reflect a limitation in short-term auditory memory for the sounds » (p. 257)), la présence de dimensions qualitatives (« a qualitative dimension might be responsible for the apparent categorical perception of a quantitative dimension » (p. 258)) et l'influence du contexte (« contextual variables can always serve to encourage CP results, even though perception is continuous » (p. 258)).

Il poursuit par ailleurs cette critique vis-à-vis de certains travaux plus récents, comme la *Théorie de l'Aimant Perceptif* élaborée par Kuhl, qui postule que la capacité de discrimination doit être moins bonne au voisinage du cœur de la catégorie, puisque celui-ci est davantage homogène que sa périphérie, en raison de « l'effet attracteur » du prototype (Boë et Schwartz, 1997 ; Iverson, Kuhl, Akahane-Yamada, Diesch, Tohkura, Kettermann et Siebert, 2003). A ce propos, Massaro indique donc (2001, pp. 14871-14872) :

*« Too often, behavioral scientists allow their phenomenal impressions to spill over into their theoretical constructs. [...] Although communication forces us to*

*partition the inputs into discrete categories for understanding, this property in no way implies that speech perception is categorical. [...] Although CP has been discredited, it is often reinvented under new guises. Most recently, the perceptual-magnet effect (PME) has had a tremendous impact [...]. The critical idea is that the discriminability of a speech segment is inversely related to its category goodness. Ideal instances of a category are supposedly very difficult to distinguish from one another relative to poor instances of the category ».*

Cela renvoie, selon Massaro, à une reformulation de la CP, en ce que la discrimination est supposée être plus précise lorsqu'elle est inter-catégorielle que lorsqu'elle est intra-catégorielle. Dans le cadre PME, de manière similaire au cadre CP dans lequel la discrimination doit être directement prédite par la performance d'identification, la discrimination doit être prédite par une mesure d'appartenance à la catégorie (« category goodness »).

Ces limitations conduisent donc à se tourner vers d'autres approches. Si certains défenseurs de la CP accordent aux frontières catégorielles une certaine flexibilité, Massaro considère que cette flexibilité s'accorde davantage avec une interprétation continue que catégorielle de l'information sensorielle. Dans ce sens, une nouvelle approche a été d'utiliser des jugements perceptifs continus plutôt que catégoriels, une approche qui respecte la distinction entre les deux étapes de traitement de l'information que sont la perception sensorielle (« sensory stage ») et la décision perceptive (« decision stage ») et qui a permis de montrer que de l'information continue est bien disponible pour les auditeurs (Massaro, 1987, p. 271). La question est alors de déterminer comment cette information continue est évaluée et intégrée dans la perception de la parole. Le FLMP vise précisément à répondre à cette question.

Si une approche binaire en termes de Tout-ou-Rien convient bien au traitement d'informations discrètes, le traitement de l'information continue nécessite une base logique quelque peu différente : il s'agit de la logique floue qui, au lieu de disposer de deux valeurs (vrai/faux), propose une série continue de valeurs de vérité (plus ou moins

vrai) (Klir, 1996). En effet, face à de l'information continue, le sujet percevant doit assigner aux sources d'information des valeurs qui peuvent être comparées entre elles, avant de pouvoir prendre une décision. C'est ce que permet d'effectuer la logique floue : à chaque source d'information est assignée une valeur sur un continuum de vérité, valeur basée sur le degré de soutien à une certaine catégorie qu'apporte la source (Massaro, 1987, p. 271).

D'après le FLMP, la reconnaissance est effectuée en trois étapes (p. 272-273) :

- ◆ Evaluation des traits : l'information est traitée par le système sensoriel, conduisant à la dérivation de traits indépendants auxquels est assignée une valeur de vérité.
- ◆ Intégration des traits : les traits sont comparés avec les unités perceptives prototypiques afin de déterminer le degré de réalisation de chaque prototype dans le son perçu. De cette manière, un trait informatif a d'autant plus de poids que le trait suivant est ambigu : l'indice le plus informatif a le plus de poids sur le jugement.
- ◆ Classification des schémas : les prototypes potentiels sont alors évalués en regard de leur « mérite ».

Ce modèle a par la suite été affiné (Massaro, 2001), mais les trois étapes essentielles restent les suivantes :

1. Evaluation,
2. Intégration,
3. Décision.

Ces processus font usage de prototypes stockés en mémoire à long terme et, tout en se suivant dans le temps, se chevauchent en partie (l'un peut opérer sans attendre la fin du précédent). Les trois principes de ce modèle sont les suivants (2001, p. 14872) :

- ◆ Chaque source d'information est évaluée afin de déterminer le degré continu auquel cette source spécifie différentes alternatives,
- ◆ Les sources d'informations sont évaluées indépendamment les unes des autres,

- ◆ Les sources sont intégrées afin de fournir un soutien globalement continu à chaque alternative,
- ◆ L'identification perceptive et l'identification découlent du degré de soutien relatif à chaque alternative.

Ces processus semblent universels, quelle que soit l'information à traiter (Chen et Massaro, à paraître) :

*« Information [...] available for the evaluation operation naturally varies across individuals or groups because of differences in the prototypes related to the perceiver's native language and linguistic experience. However, the information-processing instantiated by the evaluation, integration, and decision operations appears to be universal across individuals and languages »*

Massaro (2001, p. 14873) donne l'exemple d'une tâche d'identification syllabique pour les syllabes /ba/, /va/, /a/ et /da/, en uni- et bimodalité (vision faciale). Lorsque la présentation bimodale est « consistant » (deux sources convergentes), le rapport d'identification correcte est le suivant pour les syllabes /ba/ et /va/ : Bi (bimodal) > Vi (visuel) > Au (auditif). Pour les syllabes /a/ et /da/, le classement est le suivant : Bi > Au > Vi. Dans tous les cas, la performance en bimodalité est supérieure à celle en unimodalité, sauf lorsque les deux sources sont « inconsistant ». Les résultats indiquent ainsi une influence considérable des deux modalités sur la performance, avec une dominance globale de la modalité auditive sur la modalité visuelle.

En ce qui concerne les influences contextuelles haut-bas, Massaro (2001, p. 14874) soulève l'importante question de savoir si les informations haut-bas et bas-haut sont traitées de la même manière. D'après son analyse et les résultats d'études conduites dans le cadre du FLMP, il semblerait que celles-ci soient effectivement traitées de la même manière, le sujet percevant traitant et intégrant ces différentes sources d'information comme différentes sources d'information bas-haut. D'ailleurs, dans son

analyse de l'effet McGurk, Massaro (1998) remarque un fait intéressant : croisant les syllabes auditives et visuelles /ba/, /da/ et /ga/ (unimodalité et bimodalité), il apparaît, comme cela était prévisible, qu'il y a de nombreuses confusions entre les syllabes visibles /da/ et /ga/, étant donné que ces syllabes appartiennent à la même catégorie de visèmes. Ce qui est plus surprenant en revanche, c'est que, en unimodalité visuelle, les participants ont identifié les syllabes /da/ et /ga/ deux fois plus comme /da/ que comme /ga/. Un tel fait avait jusqu'alors été ignoré car les sujets n'avaient jamais été testés en unimodalité visuelle (p. 22) :

*« This result offers a new explanation of why an auditory /ba/ paired with a visual /ga/ produces the response /da/. Apparently, people are biased to report /d/ over /g/ because /d/ occurs much more often than /g/ in spoken language [in English] ».*

La fréquence d'occurrence des items, et ainsi la connaissance phonotactique des sujets, semble donc influencer l'identification de syllabes visuelles.

En ce qui concerne l'étape d'intégration, Massaro et ses collègues (Massaro, Cohen et Smeele, 1996) se sont également penchés sur la dimension temporelle du processus, en considérant les situations de décalage, en temps, mais également en contenu, entre informations visuelles et auditives. Afin de tester la perception asynchrone des stimuli, convergents et divergents, les chercheurs ont donc fait varier l'asynchronisme de l'attaque du stimulus (« Stimulus Onset Asynchrony », *SOA*) pour les syllabes /ba/, /va/, /ða/ et /da/. Ils rappellent d'abord que la vision du visage du locuteur peut améliorer l'intelligibilité du message et que les indices visuels sont particulièrement informatifs en ce qui concerne le lieu d'articulation des consonnes, mais également pour leur mode d'articulation, ainsi que pour l'identification des voyelles. La vision peut donc fortement influencer la perception de la parole, au point de conduire parfois à des illusions perceptives.

Dans leurs expériences, les sujets devaient regarder le visage du locuteur, écouter sa parole et identifier *ce qu'ils entendaient*. Lorsque les stimuli étaient divergents (/ba/ et /da/), les résultats obtenus étaient les suivants :

	<b>Modalité visuelle</b>	<b>Modalité auditive</b>	<b>Identification</b>
<b>Synchrone</b>	/ba/	/da/	/bda/ (52%)
<b>Auditif avant Visuel</b>	/ba/	/da/	/bda/ (31%)
<b>Visuel avant Auditif</b>	/ba/	/da/	/bda/ (69%)

Les chercheurs notent alors que l'articulation d'un /ba/ est visuellement similaire à celle d'un /bda/, et que l'augmentation du taux d'identification du groupe /bda/ semble être due au traitement d'une modalité (ici visuelle) avant l'autre. En revanche, l'articulation d'un /da/ est assez différente de celle d'un /dba/ et lorsqu'un /da/ visuel précède un /ba/ auditif, seuls 10% des sujets ont identifié un groupe /dba/, un taux bien différent des 69% précédents. Deux facteurs semblent donc être à l'œuvre dans l'identification des groupes : d'une part le temps d'arrivée des informations, mais également le degré de concordance entre celles-ci et l'articulation d'un groupe.

Suite à leurs expériences, les chercheurs concluent que le système perceptif semble capable d'intégrer des informations asynchrones, sans doute grâce à un système de stockage sensoriel, qui permet l'extension du processus de traitement des informations une fois les stimuli disparus, dans un intervalle variant globalement entre 150 et 250 ms : « The independent evaluation and multiplicative integration of audible and visible speech is very robust across small changes in temporal occurrence » (Massaro, 1998, p. 23).

Par ailleurs, toujours en relation avec la dimension temporelle, il semblerait que le caractère dynamique de l'information visuelle ne soit pas essentiel à la reconnaissance de syllabes visuelles : les expériences conduites par Massaro à l'aide de son système de tête parlante virtuelle (BALDI), lui ont en effet permis de réduire le nombre de

variations faciales, remplaçant une information dynamique par une information statique, sans que le pourcentage de reconnaissance ne soit significativement affecté, dans une certaine limite du moins (Massaro, 1998, p. 24).

Alors que la majorité des expériences conduites l'ont été en langue anglaise (stimuli et sujets), mis à part quelques rares études incluant le japonais, l'espagnol et le néerlandais, Chen et Massaro (à paraître) se sont penchés sur le cas du mandarin, comparant ainsi les performances de locuteurs américains et mandarins, afin d'évaluer la portée, potentiellement universelle, des principes sous-tendant le FLMP. Les auteurs rappellent d'abord que, contrairement aux résultats obtenus précédemment par Massaro, certains chercheurs (Sekiyama et ses collaborateurs) ont mis à jour une variation inter-langue dans l'importance de l'effet McGurk : l'influence visuelle semblait bien plus grande chez les locuteurs américains que chez les locuteurs japonais et chinois. L'idée présentée par les chercheurs était que les Japonais s'appuyaient davantage sur l'information auditive, l'information visuelle n'étant sollicitée que lorsque la première n'était pas complètement identifiée. Ainsi, les Japonais séparent les informations auditives et visuelles tant que la première serait suffisamment informative. Les explications avancées étaient de deux ordres : culturellement, les Japonais (et les Chinois) auraient tendance à éviter de regarder directement le visage de leur locuteur, (attitude considérée comme impolie) ; linguistiquement, la dépendance vis-à-vis de l'information auditive serait à relier aux propriétés tonales des langues en question.

Afin de réconcilier ces résultats avec les leurs, Chen et Massaro suggèrent de distinguer l'information du traitement de l'information : « Different language speakers may be influenced by visual and/or auditory sources of information in different quantities, but the quality or the type of audio-visual integration may not necessarily be different across languages ». Ils mentionnent également le rôle des consignes et des tâches expérimentales, qui peuvent fortement influencer l'orientation (plutôt visuelle ou plutôt auditive) attentionnelle des sujets. Chen et Massaro « do not believe that participants switch among different types of audiovisual processing depending on the physical properties of the stimulus and whether it is native or foreign ». Les



expériences qu'ils décrivent visaient précisément à approfondir la question, puisque le chinois mandarin présente les caractéristiques, culturelles (regard non-direct) et linguistiques (langue tonale), évoquées plus haut. Les résultats soutiennent leurs hypothèses : même si les réponses des locuteurs mandarins diffèrent de celles des locuteurs américains, tous sont significativement influencés par les stimuli auditifs et visuels, d'une manière apparemment parfaitement décrite par le FLMP. Le modèle de Massaro semblerait donc offrir un cadre solide d'analyse du phénomène de perception audiovisuelle de la parole.

### **II.2.2. La théorie de la perception pour le contrôle de l'action (Schwartz et al.)**

Issu de plusieurs travaux menés au sein de l'Institut de la Communication Parlée (ICP) de Grenoble (France), notamment sur la communication audiovisuelle, Schwartz, Abry, Boë, Vallée et Cathiard présentent un cadre théorique visant à intégrer les relations entre perception, contrôle de l'action et phonologie (Schwartz, Abry, Boë, Vallée et Cathiard, 1999), ces relations étant, selon eux, au cœur de l'étude de la parole. La perception de la parole dans un tel cadre est définie comme suit (Schwartz, Abry, Boë et Cathiard, 2002, p. 255) : l'ensemble des processus (auditifs, visuels, voire tactiles) qui permettent au locuteur, au niveau segmental, de récupérer et de spécifier l'ordre temporel et les cibles des gestes articulatoires, et qui fournissent un ensemble de représentations pour le contrôle de ses propres actions d'une part, et pour assurer le suivi des actions d'autrui d'autre part. Leur approche est donc centrée sur la co-structuration des systèmes perceptifs et actionnels en relation avec la phonologie (p. 255).

Cette approche se distingue donc ainsi des théories purement *auditives*, dans lesquelles l'interprétation du signal, par un lien théorique direct entre sons et phonèmes, semble indépendante des gestes articulatoires, ainsi que des théories *motrices* (en particulier « La » Théorie Motrice de Liberman et Mattingly), pour lesquelles la perception n'est que le reflet de l'action, étant donné le lien supposé direct entre sons et gestes articulatoires. Leur théorie se centre donc sur des représentations intrinsèquement sensori-motrices qui ne sont ni de purs schémas sensoriels, ni de purs objets moteurs

inférés, mais plutôt des percepts multimodaux contraints par les régularités de l'action (p. 271). L'architecture qu'ils ont proposée pour l'élaboration de telles représentations repose sur deux notions essentielles à la production de gestes articulatoires : l'agencement temporel (« timing ») et les cibles (« targets »). L'idée générale est donc que tout locuteur perçoit les sons émis en accord avec sa propre capacité à reproduire ces sons (p. 273).

Dans cette théorie, la place faite au caractère multimodal de la parole n'est pas minimale : « A basic component of our theory [...] is that gesture recovery and specification involves multisensorial inputs » (p. 278). Le recouvrement des gestes articulatoires est ainsi considéré dans sa dimension multimodale, comme l'indique l'intitulé d'une de leur sous-partie, « gesture recovery/shaping is multimodal » (pp. 262-270) :

*« The need for supplementary information when the auditory specification is incomplete opens another and very important field, the human ability to follow speech gestures through sensorial modalities other than audition, and eventually fuse multisensorial inputs for this aim. [...] this ability is not a marginal 'side effect' in speech perception, but on the contrary a core component of speech perception, in relationship with action control and phonology ».*

Ils listent ensuite plusieurs arguments justifiant une telle affirmation :

1. La possibilité de suivre, partiellement, les gestes articulatoires lorsque le signal auditif est manquant. Ceci est remarquablement développé chez les malentendants, particulièrement *via* le système de lecture labiale (les auteurs indiquent que, en moyenne, entre 40% et 60% des phonèmes et entre 10% et 20% des mots peuvent être saisis par la seule lecture labiale, bien que ces taux soient fortement variables selon les individus), mais également par le canal tactile, notamment grâce à la méthode Tadoma (conçue à l'origine pour les malentendants non-voyants).

2. L'amélioration de l'intelligibilité d'un énoncé dans des conditions de bruit auditif grâce à la lecture labiale. De nombreuses études indiquent en effet que dans ces conditions le taux de perception correcte, tant des mots que des syllabes ou même des traits phonétiques, est plus élevé en présentation bimodale audiovisuelle que monomodale, auditive ou visuelle.
3. L'utilité de l'input visuel même lorsque l'input auditif n'est pas altéré. En particulier, les auteurs citent Davis, C. et Kim, J (1998, « Repeating and remembering foreign language words : does seeing help ? », *Proceedings of Audio-Visual speech Processing 98*, Sydney, 121-5) pour soutenir leur troisième point : « repetition is more efficient when the visual input is present [...] if it is uttered in a foreign language ».
4. L'utilisation de l'input visuel lors du recouvrement des gestes articulatoires, même lorsqu'il ne concorde pas avec l'input auditif. Cela est notoirement illustré par l'effet dit de McGurk : un [ga] visuel couplé à un [ba] auditif est perçu [da] ou [ɔ̃a], tandis qu'un [ba] visuel couplé à un [ga] auditif est perçu comme [bga]. Il semblerait par ailleurs que l'information tactile soit utilisée de manière similaire à l'information visuelle dans ce type de manipulation entre information auditive et information tactile (p. 264).
5. L'implication de l'input visuel dans la perception quotidienne de la parole, et cela dès l'acquisition du langage par les jeunes enfants.
6. L'acquisition quasi-complète d'un module phonologique par les enfants sourds grâce à la vision des gestes articulatoires, en particulier dans le cadre de la méthode de la parole indicée (« cued speech method ») qui fournit des informations visuelles labiales pour les gestes articulatoires visibles et manuelles pour ceux qui ne le sont pas.

Plusieurs modèles de perception multimodale ont déjà été proposés, mais seul un certain type de modèle, selon les auteurs, permet de rendre compte de toutes ces données : il s'agit de considérer que le format des représentations internes issue des deux sources perceptives est amodal, permettant ainsi une comparaison des deux informations, comparaison qui vise au recouvrement des caractéristiques des gestes articulatoires (processus de recodage moteur (« motor recoding, (MR) », p. 266)). L'accès aux représentations phonologiques serait ainsi assuré par ce recodage (p. 267) :

*« In summary, an MR model in which the audio-visual interaction would be organized towards the recovery of speech gestures, and regularized and constrained by their characteristics, seems to provide the best framework for dealing with available audio-visual speech perception data [...]. This model could also include a tactile input ».*

En ce qui concerne le format des représentations visuelles des gestes articulatoires (p. 268), de nombreux travaux ont porté sur l'arrondissement labial et la capacité des sujets à suivre visuellement ce mouvement, et ainsi à anticiper sur ses conséquences acoustiques : à propos du couple [i, y] en français, et du passage de la première voyelle à la seconde (arrondissement labial), il semblerait ainsi que le geste anticipatoire d'arrondissement est bien visible, mais non audible. Cela démontre l'existence d'un certain désynchronisme naturel entre les lèvres et les sons. Des expériences ultérieures ont montré (p. 269) que l'audition n'est pas la modalité dominante pour la perception de l'arrondissement (couple [i, y]), alors qu'elle l'est pour la perception de la hauteur (couple [i, a]).

En ce qui concerne le rôle de la dynamique des mouvements, il semblerait que la perception du mouvement des lèvres dans son ensemble (plutôt que ses simples extrémités statiques) soit profitable à l'identification de l'arrondissement. Cependant, cet effet disparaît lorsque l'énonciateur est présenté de profil et non plus de face (p. 270) :

*« Static and dynamic profile views provided exactly the same performances as the best front view condition, that is, the dynamic one. This confirms [...] that profile is not poorer than face for rounding visual perception ».*

Certains chercheurs (en particulier Cathiard) ont ainsi défendu l'idée selon laquelle le composant essentiel des représentations visuelles des gestes de la parole serait la forme tridimensionnelle, et non la dynamique du mouvement.

Ainsi, si le rôle de la modalité visuelle dans la perception de la parole par des sujets natifs semble aujourd'hui indiscuté, il convient de s'interroger à présent sur la valeur des indices visuels pour la perception et l'apprentissage de la parole en langue étrangère.

### **II. 3. Perception, multimodalité et apprentissage de la parole en langue étrangère**

Deux axes sont considérés ici : d'abord la question spécifique du rôle des indices visuels (labio-faciaux) dans la perception des catégories phonético-phonologiques en langue étrangère, puis celle, plus générale, de la bimodalité (audio-faciale ou audio-orthographique) dans l'apprentissage de la parole en langue étrangère.

#### **II.3.1. Rôle des indices visuels dans la perception auditive en langue étrangère**

Dans le cas d'apprenants de langue étrangère, plusieurs études ont également abordé la question de l'intégration des indices audiovisuels en perception de la parole.

##### **II.3.1.1. Influence positive des indices visuels**

Hardison (1999), qui y a en partie consacré sa thèse de doctorat, s'est intéressée au rôle des indices labio-faciaux dans l'identification de phonèmes par des apprenants avancés d'anglais langue étrangère de langues premières différentes (japonais, coréen, espagnol et malais), ainsi que par des locuteurs anglais natifs. Usant d'un protocole comparable à celui mis en place par McGurk et ses collègues (syllabes CV présentées de manière

audiovisuelle, les deux modalités étant tantôt concordantes, tantôt discordantes), son objectif était d'évaluer la contribution relative des deux sources d'information à l'identification d'un percept unique. Les consonnes cibles se divisaient en deux groupes, selon leur discernabilité articulo-visuelle : /p, f, w/ présentaient une information labiale visuellement saillante, contrairement à /r, t, k/. La voyelle adjacente était /ɑ/ : comme le rappelle Hardison (1999, p. 224), il a été montré que l'effet McGurk variait selon la voyelle, étant donné que la saillance visuelle de l'articulation consonantique variait en fonction de la voyelle qui la suivait, les voyelles arrondies étant celles qui masquaient le plus l'information visuelle consonantique. En particulier, à propos des phonèmes anglais américains /ɹ/ et /l/, qui constituent généralement des catégories visuelles distinctes (Hardison, 2003, p. 497), Hardison indique :

*« The visual intelligibility of a consonant is reduced in contexts with rounded vowels, especially /u/ [...] and in clusters with bilabial stops and /f/ versus clusters with velar stops and singleton positions ».*

Parmi les réponses obtenues pour les sujets japonais (qui nous intéressent particulièrement) dans le cas du /r/ et du /w/, nous pouvons mentionner les suivantes (Hardison, 1999) :

	<b>Indice visuel</b>		
<b>Indice auditif</b>	<b>aucun</b>	<b>w</b>	<b>r</b>
<b>w</b>	/w/ 97% /l/ 2%	/w/ 100%	/w/ 100%
<b>r</b>	/r/ 65% /l/ 35%	/r/ 85% /l/ 15%	/r/ 87% /l/ 13%

Hardison rappelle en effet que les difficultés rencontrées par les apprenants japonais pour distinguer les phonèmes anglais /r/ et /l/ ont été abondamment commentées et étudiées dans la littérature phonético-phonologique (la difficulté étant plus grande dans

les groupes consonantiques prévoicaliques, et moindre en position de singleton post-vocalique en fin de mot) (p. 229). Elle note également que certaines études ont montré que les Japonais percevaient une catégorie /w/ entre le /r/ et le /l/, et que, en termes de similarité articulatoire phonétique, le phonème anglais /w/ était plus proche d'un phonème japonais que ne l'étaient les phonèmes anglais américains /r/ et /l/. En termes de visèmes (c'est-à-dire des sons visuellement contrastifs), certains auteurs ont d'ailleurs rassemblé les phonèmes anglais /w/ et /r/ dans une unique catégorie (p. 221).

De ses expériences, Hardison conclut que, pour les non-natifs, l'influence d'un indice visuel dépend de sa valeur informative, de l'intelligibilité de l'indice auditif, de l'évaluation de la similarité entre les deux indices, ainsi que de la langue première des sujets. Cependant le recours aux indices labiaux dans une telle tâche semble globalement positif.

#### II.3.1.2. Influence moindre des indices visuels

Il faut cependant noter que d'autres travaux menés sur le même thème sont moins conclusifs : Ortega-Llebaria, Faulkner et Hazan (2001) se sont penchés sur le cas de sujets espagnols apprenant l'anglais et ont conduit des tests sur l'identification de consonnes et de voyelles en anglais britannique. Les résultats obtenus en condition audiovisuelle ne semblent pas supérieurs à ceux obtenus en condition auditive pour les consonnes ayant un statut phonémique différent en espagnol (allophones) et en anglais (phonèmes). Cependant, l'addition d'indices visuels a globalement amélioré la perception des consonnes, tant pour les sujets espagnols que pour les sujets anglais natifs, en particulier vis-à-vis de certaines caractéristiques acoustiques, ce que les auteurs expliquent de la manière suivante (p. 49) :

*« Spanish speakers of English may have used the visual cues to place/manner as allophonic features, not as distinctive cues to a phonemic distinction. [...] Visual features, like auditory features, can have different weights when cueing phonemic and allophonic distinctions ».*

En particulier (p. 48) dans le contexte syllabique VC, dans lequel l'identification auditive de la consonne est la plus difficile, les deux groupes (espagnols et natifs) ont nettement bénéficié de l'apport des indices visuels.

Ces résultats mitigés ont été en partie confirmés par l'étude de Hazan, Sennema et Faulkner (2002) qui se sont précisément intéressés au rôle des indices visuels dans la perception des contrastes /b/-/v/ (contraste allophonique en espagnol, mais phonémique en anglais) et /p/-/b/ (contraste phonémique en espagnol et anglais) par des sujets espagnols apprenant l'anglais. Outre leur statut différent en espagnol et en anglais, ces deux contrastes se distinguent également en termes de discernabilité visuelle, puisque le contraste /b/-/v/ (visèmes différents) est beaucoup plus marqué visuellement que ne l'est celui entre /p/ et /b/. Cependant les deux contrastes (de lieu et de mode pour /b/-/v/, et de voisement pour /p/-/b/) conduisent généralement à des difficultés perceptives pour les apprenants espagnols.

Globalement, les résultats n'indiquent pas de supériorité pour la condition audiovisuelle. Cependant, la variabilité des performances a conduit les chercheurs à distinguer deux niveaux d'expertise : pour le contraste /b/-/v/ les indices visuels sembleraient en effet avoir profité aux apprenants avancés, c'est-à-dire ceux qui avaient de bons résultats en condition unimodale auditive et visuelle, contrairement aux autres (p. 1688) :

*« In the process of L2 category acquisition, sensitivity to visual as well as to acoustic cues is acquired, but this sensitivity may not be present at early stages of the process, for certain contrasts at least ».*

Au vu des différences avec les résultats obtenus par Hardison (1999, 2003), Sennema, Hazan et Faulkner (2003) se sont penchés sur le cas de la perception du contraste /r/-/l/ en anglais britannique par des apprenants japonais. Les résultats se sont cependant une fois de plus révélés négatifs (p. 135) :



*« Overall identification rates in Audio and A[udio]V[isual] conditions did not differ significantly and few individual listeners showed evidence of AV benefit ».*

Si les auteurs proposent d'expliquer cette différence de résultats par une possible différence « d'informativité visuelle » (« visual informativeness ») entre les contrastes /r/-/l/ en anglais *américain* (plus marqué) et *britannique* (moins marqué), il faut également relever la possible différence d'expertise entre les sujets de Hardison (avancés) et ceux de Sennema, Hazan et Faulkner (de niveau moyen (« lower to lower intermediate level »)), qui rejoindrait alors les conclusions de Hazan, Sennema et Faulkner (2002). Cependant, il semblerait que même les sujets ayant obtenu de bons résultats en unimodalité visuelle (« visually aware ») n'aient pas obtenu d'amélioration en condition audiovisuelle.

#### Conclusion :

Plusieurs facteurs sembleraient donc influencer les résultats obtenus, et d'autres travaux semblent ainsi nécessaires avant de pouvoir tirer des conclusions plus nettes de ces travaux. En particulier, outre le niveau d'expertise des apprenants, la valeur distinctive (phonémique ou allophonique) des traits visuels entre les deux langues semble devoir être prise en considération.

### **II.3.2. Rôle de la bimodalité dans l'apprentissage de la parole en langue étrangère**

Deux types de travaux sont présentés ici : les uns relatifs à l'influence des indices faciaux dans la perception et l'apprentissage de nouvelles catégories phonémiques, les autres relatifs à l'influence des indices orthographiques sur la perception et la compréhension de la parole en langue étrangère.

#### II.3.2.1. Indices visuels oro-faciaux et catégories phonémiques

Grâce aux outils informatiques, plusieurs types d'indices visuels oro-faciaux peuvent aujourd'hui être utilisés par les apprenants de langue étrangère. Alors que les indices

les plus courants sont de type labio-faciaux (vision externe des articulateurs), d'autres permettent, à l'aide de représentations en trois dimensions, d'inclure également la position de la langue (vision interne des articulateurs).

#### II.3.2.1.1. Indices visuels externes

Les études de Hardison (1999, 2000, 2003) d'une part, et celles de Hazan et ses collaborateurs d'autre part (2001, 2002, 2003), présentent, pour l'apprentissage, une situation quelque peu mitigée, analogue à celle présentée pour la perception.

A la question « Is it possible to enhance the information value of visual cues to speech perception for nonnatives ? » Hardison (1999, p. 272) semble répondre plutôt positivement, et l'intérêt d'un entraînement audiovisuel, plutôt qu'auditif, pour les apprenants japonais entre autres, semble clair d'après les résultats de Hardison (2000, p. 313)<sup>123</sup> :

*« Auditory-visual perceptual training was more successful than auditory only providing significantly greater improvement in identification accuracy for such sounds as /ɪ/, /l/, /f/, /θ/. [...] Visual input contributed the most to the bimodal percept for the more difficult phonetic environments (i.e., initial singleton and clusters for the Japanese [...]). »*

De plus, outre l'identification phonémique, l'information labiale semble également utile à la reconnaissance lexicale, et ainsi à la compréhension orale (Hardison, 2000, p. 313) :

*« On can conclude that talker's lip movements do facilitate L2 learners' comprehension of spoken English. By increasing the information value of input*

---

<sup>123</sup> Nous remercions ici D.M. Hardison pour son aide dans l'obtention de ces informations.

*from both modalities, training reduces the set of potential word candidates to the identification process. »*

La supériorité, en termes d'amélioration de précision perceptive, de l'entraînement perceptif audiovisuel (les deux modalités étant entraînées simultanément) comparé à l'entraînement auditif semble donc avoir été clairement mise en évidence par plusieurs travaux de Hardison (1999, 2000, 2003), notamment dans le cas du /ɹ/ et du /l/ pour des apprenants japonais (et coréens). De plus, il semblerait que cet entraînement ait conduit à un certain transfert des compétences (Hardison, 2000, p. 313) :

*« Tests of generalization revealed successful transfer of training to novel stimuli and a new talker. Perceptual training also transferred to production improvement without explicit instruction or attention in training sessions to production ».*

Ainsi, si un entraînement perceptif uniquement auditif conduit à une amélioration des performances, un entraînement audiovisuel comme celui proposé par Hardison semble accroître cette amélioration de manière tout à fait significative (2003, p. 514), dans le cas de sujets japonais et coréens s'entraînant à identifier et produire les phonèmes anglais américains /r/ et /l/. Cependant, plusieurs facteurs influencent ces résultats, notamment la nature de la voyelle adjacente, la position de la consonne dans le mot, les caractéristiques phonétiques du locuteur ainsi que le système phonologique de la langue première des sujets.

Contrairement aux résultats obtenus par Hardison, Sennema, Hazan et Faulkner (2003) n'obtiennent pas un tel avantage pour la condition audiovisuelle, chez des apprenants japonais, entraînés à distinguer les phonèmes /r/ et /l/ en anglais britanniques. L'entraînement perceptif, qu'il soit uniquement auditif ou audiovisuel, conduit en effet à une amélioration des résultats, mais la différence entre les deux types d'entraînement ne semble pas significative. Cependant, les sujets entraînés audiovisuellement ont, en

revanche, davantage amélioré leur compétence de discrimination visuelle (modalité uniquement visuelle), que les autres sujets, ce qui est non-négligeable (p. 138) :

*« Although the training has not led to the development of an L2 A[udio]V[isual] benefit, a positive effect on listener's sensitivity to visual information could be established. The ability to make use of visual cues may be important for the everyday use in a foreign language environment ».*

Cependant, les résultats obtenus dans des tests ultérieurs (dans le cadre du projet « Acoustic and visual enhancement of speech for computer-based auditory training » conduit par Hazan, Faulkner et Campbell), concernant la distinction /b-/v/ seraient nettement avantageux pour les sujets entraînés audiovisuellement, ce qui montre que l'efficacité des indices visuels dans un tel entraînement varierait en fonction de la visibilité du contraste entre les deux sons. Une fois encore, plusieurs facteurs semblent donc intervenir dans l'influence que peuvent avoir les indices visuels sur l'acquisition de nouvelles catégories phonémiques.

#### II.3.2.1.2. Indices visuels internes

Massaro et Light (2003) se sont eux aussi intéressés à la perception et à l'apprentissage des segments /r/ et /l/ en anglais par des locuteurs japonais, pour lesquels une telle opposition phonético-phonologique est toujours source de difficulté. Leur étude visait à évaluer l'efficacité d'un système informatique « BALDI », destiné à faciliter l'apprentissage de certains contrastes phonétiques en langue étrangère, en comparant deux types d'enseignement :

- 1) L'un permettait aux apprenants de visualiser les processus articulatoires intra-buccaux (position et mouvements des articulateurs buccaux) à l'aide d'une simulation faciale rendue partiellement transparente ;
- 2) L'autre ne fournissait aux apprenants que la vision externe de cette même simulation faciale non transparente (pas de visualisation des articulateurs internes).

Comme l'indiquent les auteurs, l'apprentissage du couple /l/ - /r/ pour des locuteurs japonais apprenant l'anglais a été le sujet de nombreuses études. Il a été montré qu'un entraînement purement auditif pouvait aider à améliorer la capacité de discrimination des sujets, laquelle était encore meilleure lorsque les sujets disposaient d'une vision frontale du visage du locuteur - moniteur, cette présentation bimodale étant particulièrement nette pour des items plus complexes. L'étude de Massaro et Light a donc visé à tester l'hypothèse selon laquelle les sujets, en présentation bimodale, bénéficieraient, en perception et en production, de la visualisation des mouvements des articulateurs buccaux internes. Cette visualisation était par ailleurs renforcée par un système de modifications de couleurs.

Les résultats n'ont cependant pas confirmé l'hypothèse initiale des chercheurs : l'entraînement à l'aide de ce système semble bien avoir amélioré les performances des apprenants, mais aucune différence significative entre les deux types d'input n'a été mise en évidence (p. 4) :

*« The training method with a view of the visible articulation [...] was no more effective than a frontal view of the tutor ».*

D'un point de vue méthodologique, il faut noter que le test de généralisation (p. 2) consistait en la lecture à voix haute de listes de paires minimales contrastant /l/ et /r/ et que l'évaluation des productions des apprenants durant le pré-test était évalué informatiquement, tandis que pour le post-test l'évaluation des productions était humaine (« because the voice recognition system was not always accurate » (p. 3)).

Il est par ailleurs intéressant de noter que Hazan, Faulkner et Campbell, dans le cadre du projet « Acoustic and visual enhancement of speech for computer-based auditory training », se sont également servis du système « BALDI », ce qui leur a permis de comparer deux types d'entraînement : l'un en vision faciale (externe) naturelle (locuteur humain), l'autre en vision faciale (externe) artificielle (tête virtuelle « BALDI »). La distinction ne reposait donc pas dans la visibilité des articulateurs

internes, mais simplement dans la nature artificielle ou naturelle de l'instructeur : les résultats obtenus semblent alors meilleurs pour le groupe ayant suivi l'instructeur humain.

#### Conclusion :

Le nombre d'études rigoureuses portant sur ce thème étant encore limité, il serait abusif de vouloir déjà conclure. Cependant, tout en tenant compte de la multiplicité des facteurs, il semblerait que, dans certaines conditions, l'intégration des indices visuels puisse faciliter l'apprentissage de certains contrastes phonémiques.

#### II.3.2.2. Indices visuels orthographiques et mémorisation

Dans ses travaux sur l'effet McGurk, Massaro (1998) a comparé la perception de stimuli auditifs couplés, d'une part avec la vision faciale du locuteur, et d'autre part avec la vision de la forme orthographique des stimuli. Le résultat est que les stimuli visuels, tant faciaux qu'orthographiques, influencent la perception auditive dans une situation de perception bimodale.

Cependant, un autre aspect de l'influence de l'orthographe sur l'apprentissage de la parole est celui relatif à la mémorisation lexicale. Bird et Williams (2002) se sont ainsi penchés sur l'effet de la modalité (unimodalité auditive, orthographique ou audio-orthographique) sur l'apprentissage lexical, à travers des tests mnésiques explicites et implicites, pour des sujets anglais natifs, mais également des sujets non-natifs de niveau avancé. Ce travail est basé sur la thèse de doctorat de Bird, dont l'un des objectifs était d'apporter des éléments de réponse à la question suivante : le sous-titrage, dans la même langue, de documents vidéos utilisés en tant que matériel d'enseignement en langue étrangère est-il bénéfique à l'apprentissage de la langue en question ou pas ? En effet, tandis que certaines études ont mis en évidence un rôle positif de tels sous-titrages pour la compréhension globale et lexicale, certains critiques affirment qu'un tel dispositif détourne l'attention des étudiants du canal auditif, crée une dépendance vis-à-vis du texte et influence l'apprentissage des formes lexicales

orales. Le résultat global de cette étude est qu'un input bimodal simultané peut améliorer certains aspects implicites et explicites de l'apprentissage de formes lexicales orales sans coût apparent (p. 525) :

*« The present results suggest, therefore, that phonological information derived from text and sound both contribute to the improvements in processing of spoken words ».*

Il est important de souligner que cette étude se concentre essentiellement sur la rétention lexicale, et non sur l'apprentissage du système phonético-phonologique de la langue à apprendre (p. 527) :

*« The evidence presented here, and in other repetition priming studies, suggests that orthographic information can, under certain circumstances, have a significant facilitatory impact on long-term implicit and explicit learning of spoken word forms ».*

### **II.3.3. Conclusion**

Ainsi, l'ensemble de ces travaux permet de montrer que la perception de la parole est fondamentalement multimodale, et que certains stimuli visuels peuvent fortement contraindre et orienter la perception et l'interprétation des stimuli auditifs, et ainsi l'apprentissage d'un nouveau système phonético-phonologique chez des sujets apprenants. Le manque d'homogénéité dans les résultats actuels souligne précisément le besoin d'approfondir la question de la multimodalité dans le cadre de l'apprentissage d'une langue étrangère, alors qu'une grande partie des efforts de recherche s'était jusqu'alors concentré sur la perception unimodale, essentiellement auditive.

D'un point de vue neurolinguistique, il est intéressant de s'attarder un instant ici sur les propositions de Goldberg et Costa, dont l'essentiel tient en une distinction fondamentale résumée au début de leur article « Hemisphere Differences in the Acquisition and Use of Descriptive Systemes » (Goldberg et Costa, 1981, p. 144) :

*« L'hémisphère droit a une plus grande capacité à accomplir une intégration intermodale et à traiter des stimuli nouveaux ; l'hémisphère gauche est davantage capable d'effectuer des traitements moteurs et unimodaux, ainsi que l'emmagasinage de codes compacts. Au cours du processus d'acquisition d'un nouveau système descriptif, l'hémisphère droit joue un rôle crucial dans les premières étapes de l'acquisition, alors que l'hémisphère gauche est davantage capable d'utiliser des codes bien routinisés. Ceci conduit à un transfert droite-gauche de supériorité hémisphérique comme une fonction de compétence accrue vis-à-vis d'un type particulier de traitement »<sup>124</sup>.*

Ainsi, l'hémisphère droit aurait une plus grande capacité à gérer d'une part la complexité informationnelle, et d'autre part la pluralité des modes de représentations à l'intérieur d'une tâche cognitive unique. En revanche, l'hémisphère gauche serait prédominant dans les tâches qui nécessitent « une fixation sur un seul mode de représentation ou d'exécution »<sup>125</sup> (Goldberg et Costa, 1981, p. 148). Ceci est à relier à la différence de schémas de connectivité entre les deux hémisphères : davantage inter-régional pour l'hémisphère droit et intra-régional pour le gauche. Par ailleurs, les mesures neuroanatomiques « suggèrent que l'asymétrie hémisphérique dans le cortex postérieur n'est pas « modality specific » comme cela a précédemment été avancé (auditif *versus* visuel), mais plutôt que cette asymétrie est telle que les représentations distinctes attachées à une modalité spécifique sont plus saillantes dans l'hémisphère gauche »<sup>126</sup> (Goldberg et Costa, 1981, p. 146) ; « l'accent est mis sur une intégration interne pour différentes aires corticales attachées à une modalité spécifique dans

---

<sup>124</sup> « The right hemisphere has a greater ability to perform intermodal integration and to process novel stimuli; the left hemisphere is more capable of unimodal and motor processing as well as the storage of compact codes. In the process of acquisition of a new descriptive system, the right hemisphere plays a critical role in initial stages of acquisition, whereas the left hemisphere is superior at utilizing well-routinized codes. This leads to a right-to-left shift of hemisphere superiority as a function of increased competence with respect to a particular type of processing » (notre traduction).

<sup>125</sup> « Fixation upon a single mode of representation or execution » (notre traduction).

<sup>126</sup> « Suggest that hemispherical asymmetry in the posterior cortex is not modality specific as has been previously implied (auditory versus visual), but rather is such that distinct modality-specific representations are more prominent in the left hemisphere » (notre traduction).



l'hémisphère gauche ; et sur une intégration externe (intermodale) entre des aires associatives dans l'hémisphère droit »<sup>127</sup> (*op. cit.*, p. 166).

Ceci pourrait, à terme, avoir des implications didactiques quant aux modes de présentation souhaitables de l'input et des tâches à accomplir. Si l'hémisphère droit est en effet plus à même de « traiter la multimodalité », pour être bref, et étant donné sa propension à traiter les tâches ou à appréhender les systèmes jusque-là inconnus, il faudrait alors tenter de restreindre, et développer, l'utilisation de la multimodalité dans ce cadre-ci uniquement.

Or, en termes de multimodalité, si l'étude des stimuli visuels oro-faciaux et manuels semblent avoir dominé le champ au cours des dernières décennies, celle du rôle des représentations orthographiques semble un sujet d'actualité, comme en témoigne, pour la production de la parole, l'intitulé d'une communication de F.-X. Alario « Has orthography an effect on speech production ? »<sup>128</sup> (2004). Dans la perspective qui est la nôtre, il convient alors de s'intéresser à l'influence potentielle du code écrit sur la perception et l'apprentissage de la phonologie d'une langue étrangère, ce qui nécessite d'examiner en premier lieu la manière dont est perçu et traité le langage écrit.

### **III. Perception de la parole et orthographe**

Après nous être penchés sur la perception de la parole, tant en modalité auditive qu'audiovisuelle (essentiellement oro-faciale), il nous faut à présent rassembler certaines données sur la perception de l'écrit, puisque, d'une certaine manière, c'est précisément à l'influence possible de celle-ci sur la précédente qu'une grande partie de notre étude est consacrée. Comme nous l'avons vu précédemment, la perception du langage peut s'effectuer *via* divers canaux, impliquant différents types de

---

<sup>127</sup> « The emphasis is on modality-specific cortical areas and integration within them in the left hemisphere; and on associative areas and intermodal integration in the right hemisphere » (notre traduction).

<sup>128</sup> Journée internationale de psycholinguistique cognitive en l'honneur de Juan Segui, Université Paris 5, 6 novembre 2004.

représentations cognitives. L'écrit se pose comme le second principal médium de communication linguistique, et les relations entre la parole et l'écrit, tant en perception qu'en production, constituent un champ d'investigation considérable, liant en premier lieu orthographe, phonologie et morphologie.

### **III.1. Introduction**

Otake, Hatano, Cutler et Mehler (1993), dans leur étude sur les procédures de segmentation de la parole en japonais, envisagent la possibilité que les procédures *a priori* phonologiques mises en œuvre par les sujets dans leur protocole aient été influencées par le code orthographique correspondant aux unités phonologiques évoquées dans une telle tâche (p. 274) :

*« It is appropriate to consider whether the close relationship between Japanese phonology and (one type of) orthography may have played a role in our findings. The mora is properly a phonological construct [...] [but it is] also an orthographically real construct to our Japanese subjects, in that kana orthographies directly encode mora structure. The possibility exists, therefore, that our subjects may have produced their moraic response patterns via an orthographic representation rather than a phonological representation. This would require first that subjects formed an orthographic representation of the specified target, in terms of the mora-based kana orthography, and second that this representation could be directly matched against an orthographic representation of each incoming word, again in kana ».*

Une telle remarque soulève donc la possibilité qu'interviennent, dans une tâche de traitement de la parole, les représentations orthographiques qui lui sont associées. Or, lorsque l'on s'intéresse à l'influence d'une L1 sur l'apprentissage du système phonético-phonologique d'une L2, s'il est d'usage de prendre en considération les « interférences » possibles entre systèmes phonético-phonologiques en présence, il est en revanche beaucoup plus rare, à notre connaissance, de se pencher sur l'influence de

son système phonographémique et ainsi de ses représentations orthographiques sur ledit apprentissage.

C'est précisément à l'influence possible des représentations orthographiques sur le traitement de la parole qu'est consacrée la présente partie. Dans un premier temps (III.2.), nous présentons certaines caractéristiques communes à la perception de la parole et de l'écrit en termes de traitement de l'information (information ascendante et descendante). Puis (III.3.), nous nous penchons sur les spécificités de la perception de l'écrit (qu'il s'agisse de reconnaissance visuelle lexicale ou de lecture de mots inconnus ou de non-mots), en soulignant la variation interlangagière des processus : cette variation semble liée tant à la « profondeur orthographique » des systèmes phonographémiques, qu'aux unités perceptives sollicitées. Nous détaillons alors le cas particulier du japonais, qui présente un système d'écriture triple, voire quadruple. Nous présentons ensuite (III.4.) les principaux modèles de reconnaissance visuelle lexicale et de lecture à voix haute, avant d'examiner (III.5.) de manière plus détaillée la manière dont le code orthographique peut influencer le traitement de la parole. Cette influence peut être appréhendée à un niveau sublexicale ou lexical en ce qui concerne les unités perceptives, mais également à travers la formation de la conscience phonologique des sujets.

### **III.2. Perception de l'oral et de l'écrit : principes invariants**

Avant de traiter des similitudes entre les deux types de perception, il nous faut rappeler certaines différences fondamentales entre l'oral et l'écrit.

#### **III.2.1. Primauté de l'oral sur l'écrit**

En ce qui concerne la primauté de l'oral sur l'écrit, d'un point de vue tant ontogénétique que phylogénétique, celle-ci s'explique *via* un certain nombre de différences que nous rappelle Liberman (1992). Parmi celles-ci, on peut signaler que, contrairement à l'écrit :

- 1) La parole est universelle (chaque communauté possède un langage parlé).

- 2) La parole est première, du point de vue de l'histoire collective, mais aussi individuelle.
- 3) La parole doit être apprise, mais ne nécessite pas d'enseignement.

Le caractère plus « naturel » de la parole par rapport à l'écrit est justifié par Liberman dans le cadre de la théorie motrice de la parole de la manière suivante : tandis que certains auteurs définissent les éléments phonétiques (en tant qu'unités de la parole) comme des sons (ce que Liberman nomme « conventional view », p. 169), les tenants de la théorie motrice les considèrent comme des gestes (« unconventional view », p. 172), lesquels permettraient en particulier une meilleure approche de la coarticulation. Dans une telle perspective, il n'y aurait alors aucunement besoin d'une transformation cognitive d'une représentation auditive initiale, puisqu'une telle représentation n'existerait tout simplement pas (p. 173). La différence entre la parole et la forme écrite du langage s'illustre alors ainsi dans le cas de la lecture (p. 174) :

*« In reading, [...] the primary perceptual representations are [...] inherently visual, not linguistic. Thus these representations are, at best, arbitrary symbols for the natural units of language, hence unsuited to any natural language process until and unless they have been translated into linguistic forms. On the other hand, the representations that are evoked by the sounds of speech are immediately linguistic in kind, having been made so by the automatic processes of the phonetic module ».*

Liberman poursuit en indiquant que, dans une telle perspective, le langage est considéré comme n'étant ni auditif ni visuel (p. 174) :

*« If it seems auditory, that is only because the appropriate stimulus is commonly acoustic [...]. But optical stimuli will, under some conditions, evoke equally convincing phonetic percepts, provided [...] they specify the same articulatory movements (hence, phonetic gestures) that the sounds of speech evoke ».*

Cette vue « non-conventionnelle » considère donc les représentations motrices et perceptives primaires comme spécifiquement phonétiques, contrairement aux représentations écrites, moins « naturelles ». Quelle que soit la position adoptée, la primauté de l'oral sur l'écrit semble, dans tous les cas, indiscutable, malgré les critiques qui ont pu lui être adressées et qui sont soigneusement examinées par Durand (2000b). Ceci étant, la perception de l'oral et celle de l'écrit sont régies par des principes généraux communs, abordables en termes de traitement de l'information.

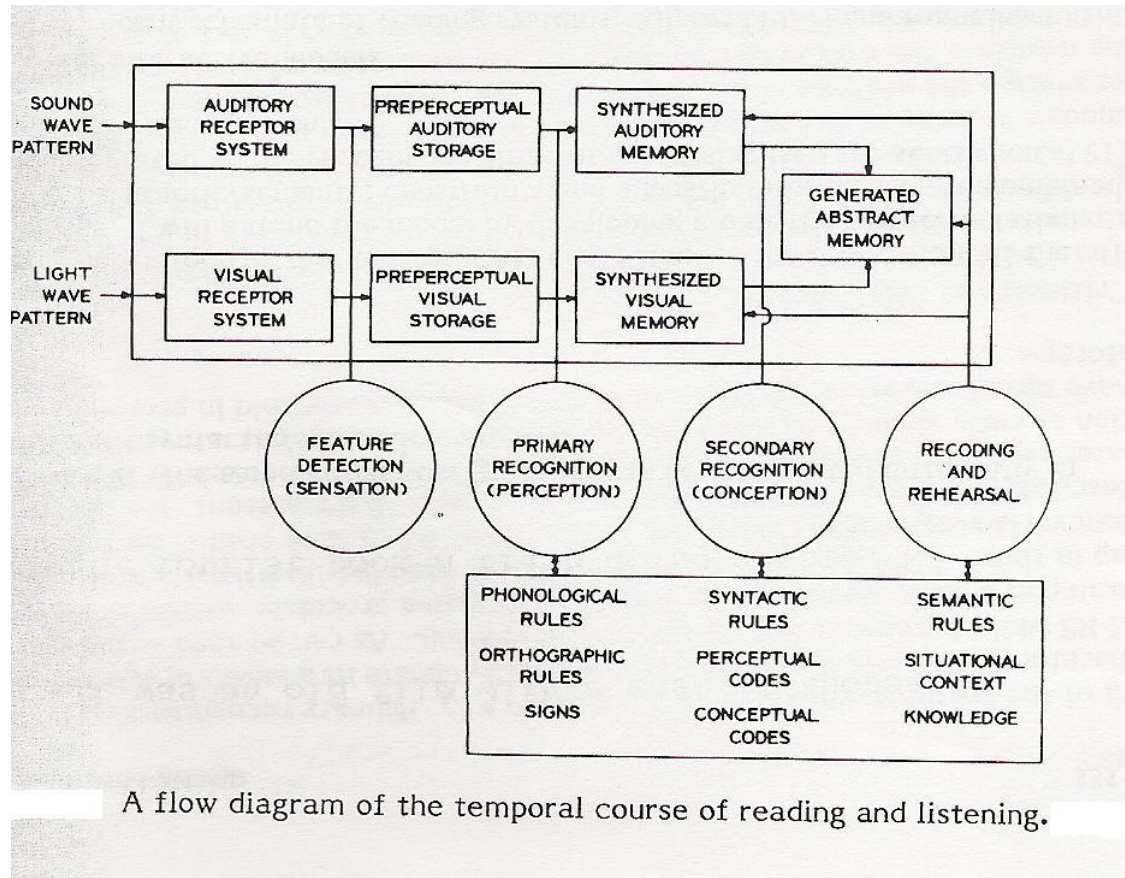
### **III.2.2. Traitement de l'information**

L'étude des relations entre la perception de l'oral et celle de l'écrit n'est pas récente. En 1979, alors que les approches interactives et parallèles étaient en cours de développement, l'approche la plus répandue était encore celle des modèles « bas-haut » séquentiels, même si l'influence du contexte et des informations de haut-niveau n'était pas ignorée ou négligée (Allport, 1979), comme dans le modèle décrit par Massaro (1979).

Massaro se place alors dans un cadre de traitement essentiellement séquentiel de l'information : dans celui-ci, le passage de la forme au sens se déroule en plusieurs étapes, dans lesquelles entrent en jeu des modules de *stockage*, qui définissent le type d'information disponible, et des modules de *traitements* qui spécifient les procédures opérant sur l'information stockée dans le module correspondant. Dans une telle perspective, la perception de l'oral et celle de l'écrit partagent des structures et des processus analogues. Le modèle qu'il propose distingue quatre modules de traitement :

1. Détection de traits,
2. Reconnaissance primaire,
3. Reconnaissance secondaire,
4. Recodage en répétition (« rehearsal-recoding »).

La schématisation qu'il en donne est la suivante :



Détection des traits : En ce qui concerne la détection des traits, il convient de distinguer traits phonétiques (essentiellement continus) et phonémiques (essentiellement discrets). Dans le modèle décrit par Massaro, les traits acoustiques sont considérés comme continus (p. 337) :

*« The listener will be able to hear the degree of presence or absence of a particular feature, even though his judgment in a forced choice task will be discrete. [...] [We] have described acoustic features as fuzzy predicates which may be more or less true rather than only absolutely true or false ».*

De la même manière que pour la parole, les traits visuels sont détectés en termes flous. Les ensembles de traits sont généralement déterminés par la présence ou l'absence de lignes horizontales, verticales ou obliques, de courbes, d'intersections, d'angles, etc. Ces traits ont généralement été étudiés en rapport avec les lettres majuscules, tandis que l'étude des lettres minuscules a plutôt conduit à la notion d'*enveloppe* de la lettre, c'est-à-dire la forme globale de la lettre (par exemple circulaire pour la lettre « e », mais triangulaire pour la lettre « v ». Ici encore la perception de l'enveloppe doit être considérée en termes flous.

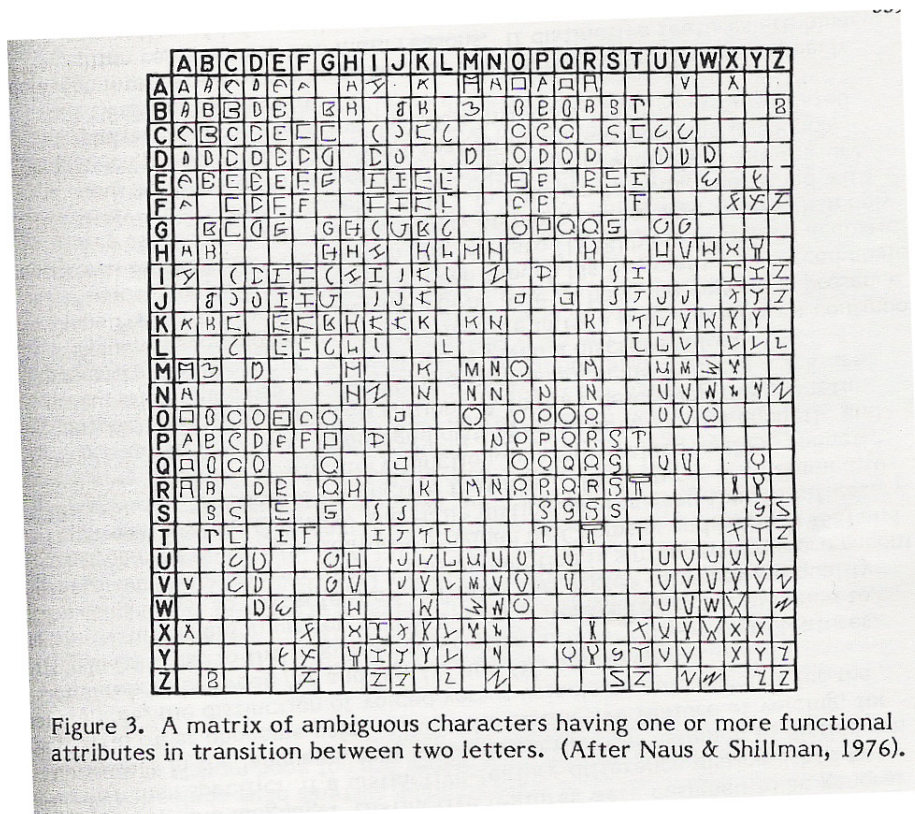


Figure 3. A matrix of ambiguous characters having one or more functional attributes in transition between two letters. (After Naus & Shillman, 1976).

Reconnaissance primaire : En ce qui concerne la reconnaissance primaire, il s'agit du passage de l'intégration des traits en un percept stocké en mémoire. La reconnaissance primaire permet donc de trouver la meilleure association entre le résultat de l'intégration des traits et la description des unités perceptives dans la mémoire à long terme. La question des unités perceptives se pose alors. Tandis que pour la parole, la

syllabe semble la meilleure candidate, car moins invariante acoustiquement que le phonème, pour l'écrit, les lettres imprimées sont parfaitement définies de manière invariante quant à leurs caractéristiques visuelles, ce qui en fait des candidats tout à fait acceptables (Massaro évoque cependant les polémiques au sujet des unités perceptives tant à l'oral qu'à l'écrit).

Reconnaissance secondaire: En ce qui concerne la reconnaissance secondaire, elle permet de transformer les percepts ainsi intégrés en formes signifiantes *via* l'appariement entre l'information perceptive et le meilleur candidat dans le lexique mental. De plus, l'information contenue pour chaque mot dans le lexique mental est de double nature : perceptive, mais également contextuelle (sémantique et syntaxique) (p. 345) :

*« Both low-level perceptual information and high-level contextual information contribute to speech processing (and reading). What remains to be discovered is exactly how these two sources of information are combined in the abstraction of the meaning of the message ».*

Recodage en répétition: Le dernier module, celui de la répétition et du recodage, est destiné à maintenir et construire les structures sémantico-syntaxiques. Les informations issues des canaux auditif et visuel sont toutes deux stockées dans la même structure abstraite, qui correspond à la mémoire de travail. Quant au code à l'œuvre dans cette structure, il semblerait qu'il soit dépendant de la langue en question.

Une telle approche permet ainsi à Massaro de conclure (p. 350) :

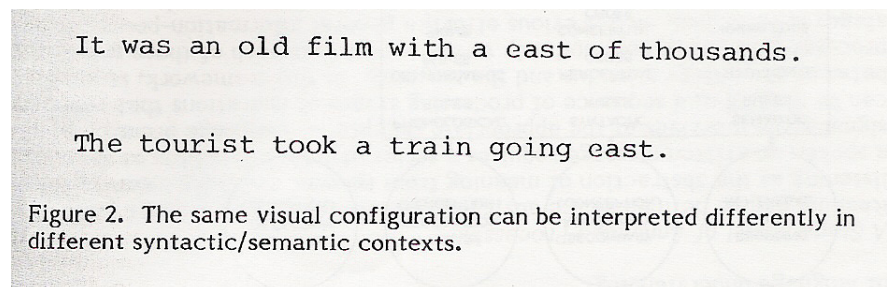
*« Although the surface structures of written text and speech present unique questions to each skill, the apparent similarities in deep structure offer the hope of a single framework for understanding both reading and listening ».*



### **III.2.3. Contexte et effets haut-bas**

Parmi les nombreuses similitudes entre les deux types de perception, il est nécessaire d'insister sur l'importance des effets d'anticipation et d'attente (Carrell, 1990), qui sont intimement liés au contexte (contexte socio-interactionnel de production du discours, contexte morphosyntaxique, lexical et phonologique) dans lequel apparaît l'item traité : c'est le contexte qui permet d'activer un ensemble de potentialités linguistiques et d'inhiber celles qui ne sont pas pertinentes. A titre purement illustratif, nous insérons ici deux exemples, l'un donné par Massaro (1979, p. 335), l'autre, apparemment courant<sup>129</sup>, qui nous a été fourni par hasard par un collègue non-spécialiste. Chaque exemple illustre des aspects différents de l'importance des effets haut-bas (pré-lexicaux ou lexicaux) :

#### **Exemple 1 :**



#### **Exemple 2 :**

*Sleon une édtue de l'Uvinertisé de Cmabrigde, l'odrrre des ltteers dnas les mtos n'a pas d'ipmrotncae, la suele coshe ipmrotnate est que la pmeirère et la drenèire soit à la bnnoe pclae. Le rsete peut êrte dans un dsérorde ttoal et vuos puoevz tujoruos lrie snas porlbème.*

Tandis que la forme d'une lettre est altérée dans le premier exemple, ce sont les positions des lettres qui le sont dans le deuxième. Ce dernier peut d'ailleurs constituer

---

<sup>129</sup> Nous remercions S.L. pour l'illustration qu'il nous a ainsi fournie. Nous ne disposons pas des références exactes s'il y en a.

une illustration intéressante des travaux conduits par Grainger et ses collaborateurs sur le codage de la position des lettres : ceux-ci montrent que la position est codée de manière relative pour les lettres internes au mot (droite/gauche de telle ou telle lettre), mais de façon plus nette pour les extrémités (première/dernière lettre du mot). Dans les deux exemples cependant, il est possible de comprendre aisément chacun des textes. De manière plus générale, l'influence haut-bas (dans laquelle il est nécessaire de distinguer plusieurs niveaux), n'est d'ailleurs négligée dans aucun domaine, même dans l'étude de la vision (générale) de bas-niveau (Yuille et Ullman, 1990, p. 9) :

*« Too little information arrives at the retina to recover unambiguously the full three-dimensional world. [...] different physical events can give rise to the same pattern of light on the retina. In this sense, vision is ill-posed and it needs to make assumptions about the world to interpret the data ».*

Néanmoins, même s'il existe de nombreuses similitudes entre la perception de la parole orale et la perception de l'écrit, Ségui (2003, p. 223) note que :

*« Les caractéristiques du stimulus correspondant à la réalisation du mot parlé diffèrent assez radicalement de celles du mot écrit. Ceci concerne non seulement l'extrême variabilité des réalisations des phonèmes selon un grand nombre de paramètres mais également la durée intrinsèque du signal ainsi que son organisation séquentielle ».*

La perception de l'écrit, au sein de laquelle on peut distinguer la reconnaissance visuelle morpholexicale de la lecture de mots inconnus ou de non-mots, présente donc certaines caractéristiques qui lui sont propres et que nous présentons ci-dessous.

### **III.3. Perception de l'écrit : spécificités et variation selon les langues**

A un niveau périphérique, il importe au préalable de disposer de quelques notions relatives au système oculaire humain (Kess et Miyamoto, 1999, pp. 154-155), puisque l'étude des mouvements oculaires dans différentes tâches liées à la lecture ou à la

reconnaissance visuelle constitue une technique d'investigation assez répandue dans le domaine psycholinguistique concerné (Levy-Schoen et O'Reagan, 1979 ; McConkie, 1979 ; Rayner, 1979 ; Shebilske, W. L. et Reid, 1979 ; Matsunaga, 1994).

En tâches de lecture, le mouvement oculaire de base se compose de séries de *fixations* et de *saccades*. Les fixations correspondent à des périodes durant lesquelles les yeux restent relativement immobiles, entre 150 et 350 millisecondes. Leur fonction est triple :

- 1) Permettre la transmission du stimulus visuel.
- 2) Fournir une période de calme ouverte à toute interférence de stimuli autres.
- 3) Permettre l'intégration des informations dans les niveaux supérieurs (compréhension des informations). Elles dépendent de facteurs comme la longueur lexicale, les traits syntaxiques, la prévisibilité en fonction du contexte, et surtout la fréquence lexicale.

A tout moment lors de la fixation des yeux du lecteur, on peut distinguer, du centre à la périphérie du champ oculaire :

- La vision fovéale : elle correspond au focus central et s'occupe des détails, impliquant ainsi une grande acuité visuelle.
- La vision parafovéale : moins précise, elle est responsable du contrôle des mouvements saccadés oculaires d'un point de fixation à un autre.
- La vision périphérique : elle n'intervient que très peu dans la lecture et fournit des informations générales dans le traitement global d'images ou de textes, comme la longueur des phrases, le format des textes, etc.

En ce qui concerne les saccades, les points de fixation ne sont pas choisis au hasard mais, grâce à la vision parafovéale, les yeux cherchent constamment le point de vision optimal, c'est-à-dire celui qui va permettre une reconnaissance la plus rapide. En situation réelle de lecture, le jeu entre fixations et saccades est complété par celui de mouvements régressifs, nécessaires à la compréhension et à l'identification de mots non-compris. Cela est particulièrement vrai pour les mots de basse fréquence. Ces

mouvements régressifs sont ensuite suivis de longs mouvements saccadés jusqu'au point où la régression avait commencé.

### **III.3.1. L'hypothèse de la profondeur orthographique**

Cette hypothèse, qui a conduit à de nombreux débats, se trouve, en partie au moins, à la base de deux oppositions : l'une entre orthographes profondes et orthographes superficielles, l'autre entre voie d'assemblage et voie d'adressage. Les travaux les plus récents indiquent cependant qu'une telle binarité ne rend pas compte de la réalité des processus à l'œuvre en lecture, quelle que soit l'orthographe en question.

#### *III.3.1.1. Orthographes profondes vs. superficielles*

Lorsque l'on examine les différents systèmes orthographiques à travers le monde, il est possible de qualifier certains de *superficiels* (pour lesquels les correspondances phonème-graphème sont essentiellement univoques, associant une lettre unique à un son unique ; on peut citer, avec certaines réserves, le serbo-croate ou l'italien) et d'autres de *profonds* (pour lesquels ces relations sont beaucoup plus complexes ; on peut citer, avec certaines réserves, le français ou l'hébreu). La prononciation d'un mot écrit « superficiel » est donc beaucoup plus prévisible que celle d'un mot « profond ». L'apprentissage de la lecture et les procédures à l'œuvre en reconnaissance lexicale visuelle varieraient alors selon le type d'orthographe.

Le titre d'un chapitre de Katz et Frost (1992) soulève de nombreuses questions, et d'une certaine manière, résume bien le débat : « The reading process is different for different orthographies ». Il s'agit là de l'hypothèse de la profondeur orthographique (*Orthography Depth Hypothesis*, ODH) qui suppose que (p. 71) :

- Les orthographes *superficielles* sont plus à même d'impliquer une voie phonologique pour la reconnaissance des mots (voie d'assemblage/phonologique/sublexicale) ;

- Les orthographe*s profondes* sont plus à même de conduire le lecteur à reconnaître les mots en se référant à la structure visuelle-orthographique de leur morphologie (voie d'adressage/visuelle/lexicale).

Il convient alors de préciser les fonctions de ces deux « voies ».

III.3.1.2. Procédure d'assemblage (voie phonologique) vs. procédure d'adressage (voie lexicale)

Les deux procédures peuvent être représentées comme suit, associant préférentiellement les mots non familiers à la voie d'assemblage et les mots familiers à la voie d'adressage (Zesiger et Partz, 1994, p. 422) :

<b>Procédure d'assemblage</b>	<b>Procédure d'adressage</b>
Mot non familier écrit ↓	Mot familier écrit ↓
Analyse visuelle graphique ↓	Analyse visuelle graphique ↓
Segmentation graphémique ↓	Lexique orthographique d'entrée ↓
Conversion segments graphémiques/segments phonémiques ↓	Système sémantique ↓
Assemblage phonémique ↓	Lexique phonologique de sortie ↓
Mémoire tampon phonologique ↓	Mémoire tampon phonologique ↓
Production orale	Production Orale

Il existe une version forte et une version faible de l'hypothèse de la profondeur orthographique : la première soutient la possibilité d'un assemblage pur en tâche de dénomination ou de décision lexicale. Mais Katz et Frost soulignent les insuffisances de cette version, en donnant l'exemple de la représentation accentuelle en serbo-croate, laquelle est absente de la représentation orthographique, celle-ci étant pourtant l'une des plus superficielles. Cette version forte a été critiquée par de nombreux auteurs, qui soulignent au contraire le caractère universel de l'opération d'oralisation des mots écrits, en dépit de la variété des orthographe : il s'agit de « l'hypothèse universelle » (Besner et Smith, 1992). Katz et Frost soutiennent quant à eux l'hypothèse faible de l'ODH (1992, p. 72) :

*« The phonology needed for the pronunciation of printed words comes not only from prelexical letter-phonology correspondences but also from stored lexical phonology, that is to say, from memory. The latter is the result of a visual-orthographic addressing of lexicon, i.e., a search process that matches the spelling of a whole word or morpheme with its stored phonology »*

Les auteurs regroupent ainsi différentes études apparemment non-convergentes en soulignant la nature hétérogène des processus impliqués et rejetant le principe d'exclusivité de l'une ou de l'autre voie. C'est donc de degré de mise en œuvre de chaque voie qu'il s'agit, quelle que soit l'orthographe. Ce degré est néanmoins précisément fonction de la profondeur de l'orthographe en question. La version faible de l'ODH contient en effet les deux propositions suivantes (p. 74) :

- (1) « Both orthographic information and prelexically assembled phonological information are used for lexical access ».
- (2) « The degree to which one kind of information predominates is a function of the structural relationship between orthography and the lexical entry ».

L'approche de Katz et Frost est donc compatible avec les modèles à deux voies d'une part (qui considèrent la reconnaissance lexicale visuelle comme une procédure à deux voies distinctes) et connexionnistes d'autre part (dans lesquels le degré d'interactivité entre les différentes unités impliquées est beaucoup plus élevé) (p. 81). Katz et Frost notent toutefois qu'ils se limitent dans leur étude aux problèmes de dénomination et de décision lexicale, et qu'ils n'abordent pas les questions postlexicales telles que la compréhension textuelle (1992, p. 73).

### III.3.1.3. Approche universelle : implication des deux voies dans les deux types d'orthographe

Cette approche mixte, dans laquelle chacune des deux voies joue un rôle dans la reconnaissance des mots, est généralement de plus en plus acceptée par l'ensemble de la communauté scientifique (Massaro et Cohen, 1994) : la question ne concerne plus tant l'implication absolue et exclusive de l'une ou de l'autre voie (Carello, Turvey et Lukatela démontrent l'existence de l'implication prélexicale de la phonologie pour trois systèmes aussi différents que l'anglais le serbo-croate et l'hébreu (1992, p. 222) tandis que Besner et Smith apportent de nouveaux arguments soutenant l'implication du système lexical dans la lecture de mots « superficiels », en particulier *via* le facteur de fréquence lexicale (1992, pp. 50-53)), mais bien les facteurs déterminant le degré d'implication de chacune des voies (Katz et Frost, 1992, p. 82). Comme le démontrent Besner et Smith (1992) à travers le perse, l'italien, le néerlandais, l'espagnol et le serbo-croate, la présence ou l'absence d'effets de fréquence lexicale et d'amorce (contextuelle) ne suffit plus à établir l'implication exclusive de l'une ou l'autre voie dans l'un ou l'autre type d'orthographe. Leur conclusion est donc la suivante (p. 60) :

*« Overall the data are most compatible with the hypothesis that the addressed and the assembled routines are available in virtually all orthographies [...]. However, we assume that the assembled route makes more of a contribution to word recognition in reading shallow scripts, because the reliable relationship between spelling and sound supports faster computation ».*

Si la voie *lexicale* semble ainsi également jouer un rôle dans les orthographe*s superficielles*, l'implication de la *phonologie* dans la lecture d'orthographe*s profondes* semble à présent tout autant établie, non seulement dans des langues alphabétiques comme l'allemand et le français, mais également dans des langues dites logographiques comme le chinois. Perfetti, Zhan et Berent (1992) montrent en effet que la reconnaissance lexicale visuelle en anglais et en chinois suit un principe « universel » : en dépit de divergences attendues (phonologie « pré-lexicale » (« early prelexical ») en anglais, mais phonologie « quasi-pré-lexicale » (« still-early lexical ») en chinois), ce n'est pas tant l'implication ou non du système phonologique qui diffère, que le temps d'activation de ce dernier (p. 245). Les auteurs concluent donc à l'universalité de l'implication de la phonologie en lecture : « The acquisition of visually based spelling representations may (or may not) reduce the role phonology plays in recognizing words, but it does not entirely eliminate it » (p. 245). Cela rejoint les conclusions de Matsunaga (1994) qui y a consacré sa thèse de doctorat : l'idée selon laquelle les caractères kanji, tant en chinois qu'en japonais, seraient des symboles de type iconique représentant, non pas des sons mais du sens, et que les lecteurs pourraient obtenir ce sens sans faire appel au son semble tout à fait non fondée. C'est la raison pour laquelle les termes « pictographes », « idéographes », « logographes » ou « morphographes » doivent être abandonnés au profit de « morphosyllabique » pour le chinois et « morphophonique » ou « morphonique » pour le chinois et le japonais (p. 39) : « kanji do not represent only meanings or trigger only meanings when fluent readers read them ; instead, kanji were found to represent sounds in a significant way and trigger sounds, if not primarily, then simultaneously with meanings » (p. 114). Ses résultats semblent ainsi bien s'intégrer aux modèles à deux voies de la parole, la voie lexicale étant utilisée pour les mots ou caractères familiers et la voie phonologique pour les autres types de mots, même si certains phénomènes en japonais et en chinois restent à éclaircir, en particulier les ellipses lexicales (« skipping words ») en lecture de phrases ou de textes : « phonological coding is evidenced to occur in English, Chinese, or Japanese, at all levels of processing, except when familiar words or characters are



identified in isolation, and possibly when words in sentences and texts are skipped » (p. 118).

#### III.3.1.4. Facteurs déterminant l'implication de chaque voie

L'idée selon laquelle le système de traitement en lecture fait usage de plusieurs ressources (phonologique et lexicale) et est influencé par plusieurs facteurs (régularité, fréquence, familiarité, etc.), se retrouve chez Seidenberg (1992, pp. 109-110) :

*« There may be important individual differences in the division of labor within a given orthography [...]. [...] the orthographic depth hypothesis is simply a generalization of this observation to the case of orthographies [...]. There may be important differences between orthographies [...] owing to variation in properties such as transparency of orthographic-phonological correspondences, the number of orthographic vs. phonological word forms, and other factors. The processing style of a given reader will then depend on interactions among these factors ».*

Le facteur de fréquence, qui apparaît ici, a notamment été mis en avant par certains chercheurs, comme Seidenberg, qui ont présenté la fréquence d'occurrence des mots comme l'un des facteurs essentiels déterminant le degré d'implication de chacune des deux voies dans le processus de reconnaissance lexicale. Les mots les plus fréquents favoriseraient une stratégie visuo-orthographique, alors que les moins fréquents déclencheraient préférentiellement la voie phonologique. Certains résultats semblent même montrer que la nature de l'orthographe perd toute importance lorsque les mots en question sont très fréquents, le facteur de fréquence devenant alors prédominant (Katz et Frost, 1992, p. 78). Seidenberg (1992) se situe dans un cadre connexionniste, lequel peut cependant également être considéré comme « dual » dans le sens où sont possibles des connexions « directes » entre orthographe et sémantique et « indirectes » entre orthographe, phonologie et sémantique (Seidenberg, 1992, p. 92). Ce cadre permet notamment d'éviter l'hypothèse d'un traitement lexical séparé pour les mots à

prononciation irrégulière (p. 97), et il permet également de prendre en compte le degré de régularité (« consistency ») des mots (*via* la pondération des connexions). Dans cette perspective, la reconnaissance d'un mot implique l'activation, à divers degrés, tant des sources phonologiques qu'orthographiques. Il s'agit donc d'un système « multi-routes » (p. 113), contraint par différents facteurs, dont la profondeur orthographique.

### III.3.1.5. Conclusion

La question de l'implication des deux voies mentionnées plus haut en tâche de lecture semble ainsi s'orienter, de manière très relative, vers un certain consensus. Cependant, si ce degré d'implication constitue un facteur de distinction entre les langues, un autre facteur essentiel est celui des unités perceptives sollicitées en tâche de lecture : ceci est particulièrement important lors du traitement de non-mots et de mots inconnus, qui sollicite ainsi, à un niveau sublexical, la voie d'adressage. Car si la distinction linguistique entre graphèmes, lettres, et corps (l'équivalent de la rime phonologique) semble universelle, la nature des unités psycholinguistiques effectivement utilisées pourrait varier selon les langues, comme cela apparaît en perception auditive (par exemple : phonèmes *vs.* syllabes). En effet, comme le souligne Mattingly (1992, p. 16), il est important de distinguer « the linguistic unit used for the actual processing of an utterance by writer and reader, and the linguistic unit to which the various graphemic units correspond ».

### III.3.2. Taille des unités perceptives : lettre, graphème, corps, syllabe et mots

De la même manière que l'implication des deux « voies » en lecture semble varier selon les langues, les unités perceptives fonctionnelles sollicitées pour une telle opération semblent, non seulement variables en fonction de la langue en question, mais également multiples en fonction des conditions de lecture.

### III.3.2.1. Lettre vs. Graphème

Comme le rappellent Ray, Ziegler et Jacobs (2000), les graphèmes sont généralement définis comme la représentation écrite des phonèmes. Les graphèmes sont donc des « ponts fonctionnels » minimaux dans l'appariement entre orthographe et phonologie. Il existe des graphèmes simples (une lettre) et complexes (multi-lettres). En ce qui concerne le rapport entre graphème (et donc phonème) et lettre, Rey, Jacobs, Schmidt-Weigand et Ziegler (1998) se sont interrogés sur un possible effet phonémique en reconnaissance visuelle. Dans des langues à orthographe alphabétique comme le français ou l'anglais, il est en effet fréquent de lire des mots dont le nombre de lettres est supérieur à celui des phonèmes de ce mot, en raison de l'existence de graphèmes multi-lettres, ce qui est illustré par le mot anglais BEACH, qui comporte trois phonèmes et cinq lettres (B-EA-CH). Les auteurs se sont penchés sur le rôle des graphèmes/phonèmes dans la reconnaissance visuelle lexicale en tâche d'identification. En manipulant des mots anglais monosyllabiques de cinq lettres et de faible fréquence, il est apparu que les temps d'identification étaient plus longs pour les mots avec peu de phonèmes que pour les mots avec davantage de phonèmes. Cet effet a été également mis en évidence pour des mots français de même type, mais seulement pour les mots de basse fréquence, pas pour ceux de fréquence élevée. Cet « effet phonémique » conduit les auteurs à considérer que ces éléments subsyllabiques, qui constituent des unités orthographiques fonctionnelles, jouent un rôle essentiel dans la reconnaissance visuelle lexicale.

Cela rejoint l'étude de Ray, Ziegler et Jacobs (2000) dont l'hypothèse a été de considérer les graphèmes comme des unités perceptives du système de lecture. Les auteurs ont conduit des expériences en français et en anglais, montrant qu'il est plus facile de détecter un graphème correspondant à une lettre qu'à plusieurs. Ces résultats soutiennent donc l'hypothèse des auteurs selon laquelle les graphèmes sont traités comme des unités perceptives dans des systèmes d'écriture alphabétique comme l'anglais ou le français.

### III.3.2.2. Graphème vs. Corps et Mots

Cependant, les travaux de Grainger et ses collaborateurs sur la comparaison des traitements de mots similaires (orthographe, prononciation et signification similaires) en allemand et en anglais par des lecteurs adultes monolingues de chaque langue (ex : zoo/Zoo) ont permis de montrer que les unités traitées préférentiellement par les lecteurs allemands sont de petites taille (graphème), alors que les lecteurs anglais présentent une préférence pour les unités de grande taille (rimes ou mots). Ces résultats sont expliqués par l'irrégularité orthographique des petites unités en anglais, qui conduisent les lecteurs experts anglais à traiter des unités plus larges.

De tels résultats concordent avec ceux de Ziegler, Perry, Jacobs et Braun (2001), qui ont cherché à tester l'hypothèse suivante :

*« It is hypothesized that written languages differ in the preferred grain size of units that emerge during reading acquisition. Smaller units (graphemes, phonemes) are thought to play a dominant role in relatively consistent orthographies (e.g., German) whereas larger units (bodies, rhymes) are thought to be more important in relatively inconsistent orthographies (e.g., English) ».*

Les résultats indiquent que « the English participants exhibited stronger body-rhyme effects, [and] the German participants exhibited a stronger length effect for words and non-words »<sup>130</sup>. Cela amène les auteurs à conclure que des items identiques ont été traités de manière différente dans différentes orthographe (p. 379) :

*« Orthographic consistency determines not only the relative contribution of orthographic versus phonological codes within a given orthography, but also the preferred grain size of units that are likely to be functional during reading ».*

---

<sup>130</sup> Le « corps » (« body ») est l'équivalent orthographique de la rime phonologique.

De tels résultats nécessitent, comme nous l'avons vu précédemment, une révision de la version forte de l'Hypothèse de la Profondeur Orthographique : « We no longer believe that the difference (between shallow and deep orthography) is one of whether or not phonology is routinely involved in visual word recognition » (Lukatela and Turvey, cités par Ziegler et al., p. 379). Ziegler et ses collaborateurs partagent cette opinion (p. 379) :

*« Orthographic consistency may affect no so much the relative contribution of phonology (i.e. the specific mix between orthographic and phonological pathways), but rather the very nature of the phonological processes themselves. [...] What seems to differ between orthographies is not the amount of phonological recoding required but rather the nature of the phonological recoding process ».*

Dans des langues comme l'espagnol ou l'allemand, les stratégies de de décodage graphème-phonème en lecture semblent favorisées car l'appariement graphème-phonème est relativement non-ambigu. C'est donc le grain le plus fin qui est utilisé (d'où un effet de longueur lexicale important). En revanche, pour des langues dont l'orthographe est moins régulière (« consistant »), comme l'anglais, les enfants qui apprennent à lire utilisent une variété de stratégies, complémentaires de celle de conversion graphèmes-phonèmes, comme la reconnaissance de groupes de lettres pour les rimes et des essais de reconnaissance lexicale globale (whole-word), ce qui les conduit à utiliser un grain plus gros (d'où un effet de longueur lexicale réduit). Ainsi, dans des langues comme l'anglais, les représentations lexicales résultantes « should be shaped to represent orthography and phonology at multiple levels » (p. 380). En revanche, dans des langues à orthographe plus régulière (« consistant »), « phonological recoding should be based mainly on the processing of small grain sizes, such as graphemes and phonemes, and the resulting lexical representations should show less variability » (p. 380).

L'hypothèse de travail des chercheurs peut donc être formulée de la manière suivante : « consistency of an orthography should have a measurable effect on the grain size of units that are likely to play a role during reading and reading development » (p. 380). Dans leur protocole, ils ont utilisé des mots « cognates »<sup>131</sup> anglais-allemand, deux langues qui ont des orthographe et des phonologies similaires mais qui diffèrent énormément dans la régularité (« consistency ») des correspondances orthographe-phonologie. Ils donnent ainsi l'exemple des mots « ball », « park » et « hand » qui existent dans les deux langues, pour lesquels le graphème « a » se prononce de manière identique en allemand, mais de trois manières différentes en anglais. C'est ce qui conduit les lecteurs anglais à faire appel à des unités plus larges, de manière à réduire l'irrégularité (« the inconsistency ») au niveau graphémique. Ils notent cependant que « finding evidence for a preferential use of one particular grain size does not mean that processing at other grain sizes does not occur » (p. 383). Leurs résultats indiquent que c'est donc peut-être dans la taille des unités de traitement dominantes, dans le nombre de niveaux de « grains » de différentes tailles et dans la capacité du lecteur de passer de l'un à l'autre de ces niveaux que se trouve l'une des différences essentielles entre orthographe différentes, plutôt que dans les quantités de traitement orthographique, phonologique et sémantique, qui, elles, pourraient au contraire être similaires.

### III.3.2.3. Fréquence et taille des unités

Toujours dans cette problématique, Perry et Ziegler (2004) se sont penchés sur ce que l'on nomme en anglais « spelling », c'est-à-dire, la capacité d'épeler un mot ou un non-mot, et donc d'associer une forme orthographique à une forme sonore. La plupart des modèles associés à cette performance reposent sur deux procédures :

- L'une, lexicale, consiste à accéder à la représentation lexicale en mémoire,
- L'autre, sublexicale, a recours à un ensemble de règles phonographémiques.

---

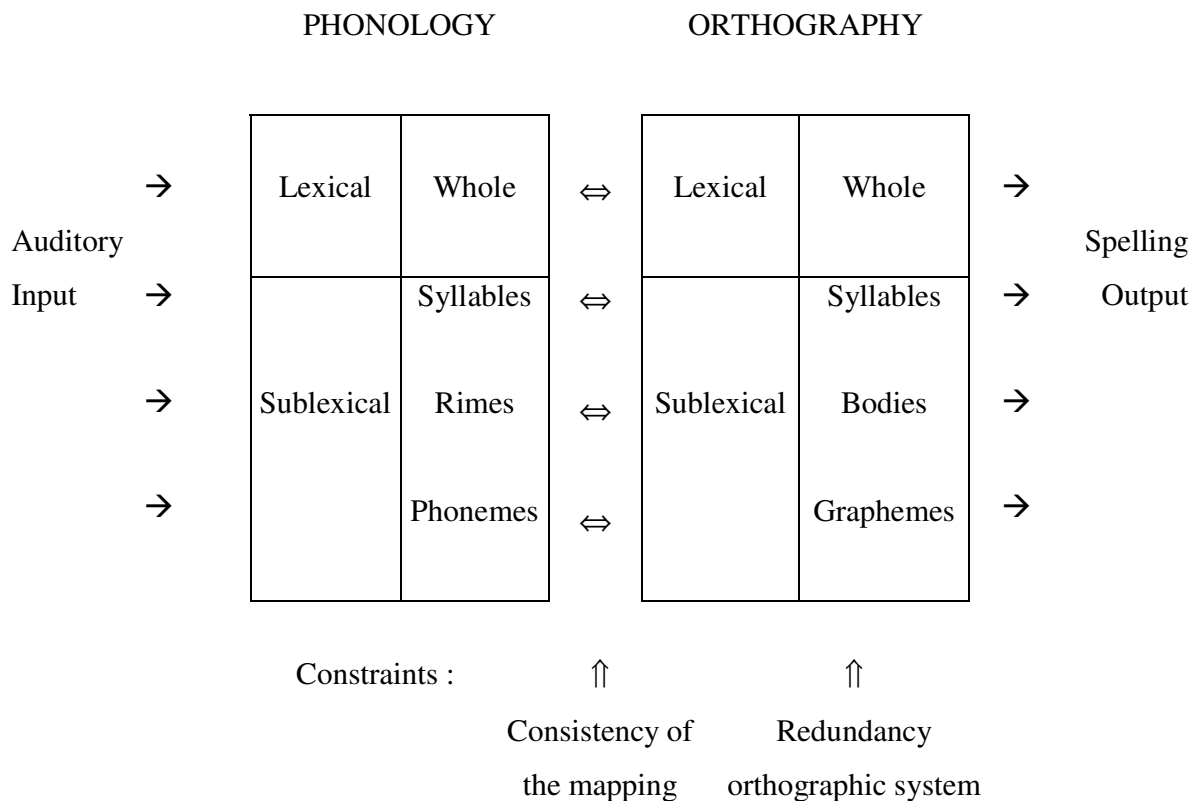
<sup>131</sup> « Cognates are words that look similar, sound similar, and have the same meaning in two languages » (p. 380).

Les auteurs se sont concentrés sur une tâche de dictée (différente d'une tâche de dénomination à partir d'images par exemple) et pour des items monosyllabiques. Dans leur étude, ils ont émis l'hypothèse que les stratégies sublexicales pouvaient être multiples, et non réduite aux correspondances phonèmes-graphèmes, suivant ainsi plusieurs études suggérant que d'autres unités sublexicales, comme les rimes et les groupes CV, pouvaient jouer un rôle important dans une telle tâche.

Ils avaient en effet par ailleurs déjà mis en évidence une variation de la taille des relations phonographémiques activées en fonction de la tâche (Perry, Ziegler et Coltheart, 2002) : alors que dans une tâche de dictée (« spelling ») de non-mots, les sujets semblaient davantage reposer sur des relations de type phonème-graphème, dans une tâche de conscience orthographique (« orthographic awareness »), dans laquelle les sujets devaient émettre un jugement de fréquence sur l'occurrence de relations subsyllabiques, les sujets semblaient pouvoir recourir à des relations de taille plus grande. Les auteurs observaient ainsi une dissociation entre les tâches de « dictée » et celles de « conscience orthographique » vis-à-vis de la taille des relations phonographémiques utilisées.

Il s'agissait donc pour eux ici (Perry et Ziegler, 2004) d'évaluer plus précisément les rôles respectifs des correspondances de petite taille (phonème-graphème) et de grande taille (rime-corps) dans une tâche de dictée (« spelling »).

Cela est illustré par le schéma qu'ils présentent (p. 327) :



Un deuxième point d'investigation concerne le rôle de l'information purement orthographique sublexicale dans l'accès aux représentations lexicales en mémoire : il s'agit de savoir si, dans une telle procédure, les sujets utilisent la redondance orthographique d'une langue (le fait que certaines lettres apparaissent plus souvent avec certaines qu'avec d'autres). Dans le premier cas, ce sont donc les régularités entre phonologie et orthographe qui sont à l'étude, tandis que le deuxième cas concerne les régularités au sein du système orthographique. La redondance orthographique est évaluée à l'aide du voisinage de corps (« BodyN »), c'est-à-dire le nombre de mots qui partagent un schéma orthographique correspondant au corps (toutes les lettres après le groupes consonantique initial, par exemple le mot *bleu* a pour voisinage de corps *feu*, *jeu*, *peu*, etc.).



Les expériences ont été conduites à l'aide de sujets natifs de langue anglaise. Deux utilisaient des mots, les trois autres des non-mots. A la base de leur première expérience se trouvait l'idée suivante : lorsque le sujet doit choisir l'orthographe d'un mot en présentation auditive, lequel possède un ou plusieurs homophones, son choix va particulièrement être influencé par la fréquence d'occurrence de chaque alternative (« people will be more likely to use a high-frequency alternative than a low-frequency alternative », p. 329). Ainsi le mot anglais « might » sera préféré au mot « mite », moins fréquent. Les auteurs notent (p. 329, note 2) que le choix de l'orthographe d'un homophone en milieu écologique est évidemment principalement déterminé par la signification appropriée au contexte, et que les conditions de laboratoire qui sont les leurs permettent précisément d'éviter de telles influences sémantiques, en présentant des mots hors-contextes.

Les conclusions auxquelles conduisent ces résultats sont les suivantes : lorsqu'ils doivent épeler tant des mots que des non-mots, les sujets ont tendance à exploiter autant l'information orthographique que la régularité des correspondances phonographémiques. En ce qui concerne la taille des unités sollicitées, il semblerait que les sujets reposent davantage sur des relations de l'ordre des phonèmes-graphèmes que des rimes-corps.

L'interprétation de ces résultats s'accorde alors plus facilement, selon les auteurs, avec les modèles à deux voies (Ziegler, Perry et Coltheart, 2000) qu'avec certaines modélisations connexionnistes (Zorzi, Houghton et Butterworth, 1998) : l'effet du voisinage de corps en particulier soutient l'existence d'une représentation lexicale du corps orthographique sensible à la fréquence : « the idea here is that words with common orthographic body representations are easier to spell than those with uncommon orthographic body representations » (p. 347). D'autres études ont par ailleurs également fait état de la pertinence d'un tel niveau de représentation, mettant en jeu le corps orthographique en lecture. En revanche, comme le souligne les auteurs, alors que le modèle de Zorzi et de ses collègues présente un effet facilitant pour le

voisinage de corps et la régularité phonographémique, tant pour les mots que pour les non-mots, les résultats de Perry et Ziegler indiquent au contraire une double dissociation : effet de voisinage mais pas de régularité phonographémique de rime (« rime consistency ») pour les mots, et le contraire pour les non-mots.

Les conclusions des chercheurs sont donc les suivantes (p. 348) :

- Les effets de fréquence orthographique (au niveau lexical et sublexical) peuvent être distingués des effets de correspondance phonographémique (de différentes tailles).
- Les correspondances de petite taille sont préférées à celles de grande taille, même si l'existence et l'utilisation de représentations de corps orthographique semblent devoir être prises en considération.

### **III.3.3 Le cas du français**

Trois notions clés doivent être évoquées dans le cas du français : celles de graphème, syllabe et fréquence.

#### **III.3.3.1. Le graphème**

Comme nous l'avons vu, le graphème semble être une unité fonctionnelle en français (Ray, Ziegler et Jacobs, 2000), du moins pour les mots de basse fréquence (Rey, Jacobs, Schmidt-Weigand et Ziegler, 1998), ce qui rejoint les conclusions de Lange et Content (1998). Ces auteurs se sont en effet également interrogés sur la nature (lettre ou graphème) des unités mises en jeu en français. Leur travail est basé sur une étude de Rastle et Coltheart qui montrait que, pour des non-mots anglais, ceux-ci étaient lus plus rapidement lorsque les nombres de graphème et de lettre étaient équivalents, que lorsqu'il y avait plus de lettres que de graphèmes. Ces résultats étaient expliqués dans le cadre du « Modèle à deux voies en cascades » (modèle « DRC »), qui propose un mécanisme de traitement sériel lettre par lettre. Il semble alors que les multi-lettres ralentissent la performance car la première unité phonémique activée (pour la première lettre) inhibe l'activation du graphème correspondant au multi-lettre, qui est activé plus tard. Lange et Content ont donc tenté de répéter l'expérience pour des non-mots

français. Ils observent le même phénomène, mais pas de manière aussi significative. Les deux résultats peuvent cependant concorder, d'après les auteurs, lorsqu'est prise en compte l'influence de la fréquence des graphèmes, fréquence qui n'avait pas été contrôlée dans l'expérience originale.

### III.3.3.2. La syllabe

Cependant, si le graphème semble être une unité fonctionnelle de lecture en français, Ferrand et New (2003) ont également mis à jour un effet syllabique important en tâche de reconnaissance visuelle et de dénomination lexicale dans cette langue. D'après leurs résultats, il semblerait que les mots français, au moins ceux de basse fréquence, nécessitent, pour les deux tâches considérées, une décomposition syllabique. Ces résultats peuvent être facilement interprétés dans le cadre de modèles considérant la lecture sous l'angle de deux procédures : l'une, globale, traiterait en parallèle la chaîne de lettres (pas d'effet syllabique) et serait particulièrement adaptée aux mots de fréquence élevée ; l'autre, plus analytique, traiterait la chaîne de lettres de manière séquentielle (fort effet syllabique) et serait particulièrement sollicitée pour les mots de basse fréquence. Comme l'indiquent Ferrand et New, le modèle DRC pourrait également rendre compte des résultats obtenus, à conditions de dépasser l'unité phonémique et d'y inclure des unités syllabiques. Leur conclusion, pour le français, est donc la suivante (p. 180) :

*« Overall, our results suggest that reading involves a global and parallel procedure for common words and an analytic and sequential procedure for low-frequency words (in both naming and visual recognition) and for nonwords (in naming only), and that the syllable constitutes an important unit in reading in French »<sup>132</sup>.*

---

<sup>132</sup> Brand, Rey et Peereman (2003), dans un protocole de dénomination lexicale en français, ne parviennent cependant pas à obtenir l'effet d'amorce syllabique qui avait été mis en évidence dans les travaux antérieurs de Ferrand, mitigeant ainsi les conclusions que l'on peut tirer de l'effet d'amorce syllabique en reconnaissance lexicale visuelle en français.

Il semblerait donc que la dimension des unités fonctionnelles sollicitées en lecture (équivalent orthographique du phonème vs. équivalent orthographique de la syllabe) varie non seulement en fonction de la profondeur de l'orthographe (relativement profonde dans le cas du français), mais également en fonction de la fréquence lexicale.

### III.3.3.3. Fréquence, unités de différentes tailles et autres facteurs en lecture à voix haute

Pour le français, les facteurs influençant la lecture à voix haute ont notamment été étudiés par Lange et Content (1999). Ceux-ci ont effectué une analyse statistique du système grapho-phonologique du français, suivi d'une expérience (tâche de lecture à voix haute de non-mot) destiné à tester leurs prédictions. Il apparaît que, pour cette tâche, la performance est influencée à la fois par la fréquence des graphèmes (la contrepartie orthographique des phonèmes) et par leur entropie (voir ci-dessous).

Leurs résultats impliqueraient (p. 437) que les sujets disposent d'une connaissance gradée (« graded knowledge ») des correspondances graphème-phonème, et s'accorderaient ainsi mieux avec une modélisation de type connexionniste, laquelle considère que le système de traitement saisit l'ensemble des régularités grapho-phonologiques, à travers un ensemble d'associations pondérées de manière différentielle entre unités de différentes tailles pour le codage tant des lettres que des phonèmes : « In this framework, the rule-governed behavior is explained by the gradual encoding of the statistical structure that governs the mapping between orthography and phonology » (p. 441). Ils s'accorderaient en revanche moins avec une approche symbolique comme les modèles à deux voies, qui, outre un index de règles de correspondances graphèmes-phonèmes, postulent la mise en œuvre d'une connaissance lexicale capable de traiter les déviations et les ambiguïtés. En effet, dans le cas de pseudo-mots, ce modèle postule un processus analytique de conversion grapho-phonémique à base de règles, qui ne laisserait aucune place aux prononciations alternatives. Lange et Content citent néanmoins Rastle et Coltheart (1999) qui écrivent :

*« One refinement of dual-route modeling that goes beyond DRC in its current form is the idea that different GPC rules might have different strengths, with the strength of the correspondence being a function of, for example, the proportion of words in which the correspondence occurs ».*

Les variables issues de l'analyse statistique de Lange et Content sont les suivantes (p. 438) :

<u>La fréquence graphémique</u>	Le nombre d'occurrences du graphème, indépendamment de sa valeur phonologique.
<u>Le nombre de prononciations alternatives</u>	Pour chaque graphème.
<u>L'entropie graphémique</u>	Cette mesure est liée à la probabilité de distribution d'un ensemble de phonèmes pour un graphème donné, reflétant ainsi le degré de prévisibilité de la prononciation du graphème (par exemple, l'entropie est minimale pour le graphème <J> qui est invariablement associé au phonème /ʒ/, mais elle est maximale lorsque la prononciation est tout à fait incertaine).
<u>La probabilité d'association phonème-graphème</u>	La fréquence d'association graphème-phonème divisée par la fréquence graphémique totale.
<u>Le rang de dominance d'association</u>	Pour une association graphème-phonème donnée, son classement parmi les différentes relations possibles entre un graphème et les différents phonèmes possibles, classement ordonné par probabilité décroissante.

A la suite de ce travail, Lange s'est plus spécialement penchée sur le cas des associations phonémiques (et ainsi des « prononciations ») multiples des graphèmes. Dans son article, Lange (2002) prend comme point de départ la position sous-jacente au modèle DRC selon laquelle le système de conversion de l'écrit au son est basé exclusivement sur des règles de correspondance graphème-phonème simples (« one-to-one, all-or-none ») : chaque graphème est converti en son associé phonémique le plus fréquent, sans modulation possible liée à la force de l'association phonographémique (règle binaire, activée, ou non) ou aux ambiguïtés de prononciation pour les graphèmes possédant plusieurs prononciations. Ce choix des concepteurs du DRC semble avoir été guidé par un souci de parcimonie, soutenu par l'absence de données relatives à cette question.

Or, Lange présente précisément les résultats d'études montrant que les lecteurs évoquent effectivement les multiples prononciations d'un graphème en tâche de reconnaissance lexicale visuelle, ce qui conduit à remettre en question la position du DRC sur ce point. Par exemple, certains résultats indiquent que le graphème G peut dans certaines conditions conduire à une compétition entre les codes phonémiques /ʒ/ et /g/ durant la conversion graphème-phonème. Lange passe ainsi en revue plusieurs travaux soutenant son point de vue, notamment ses propres travaux basés sur l'analyse statistiques des correspondances graphème-phonème en français et en anglais qui montrent une certaine sensibilité des lecteurs à la force de ces correspondances.

Cela apparaît notamment à travers leur sensibilité à ce que Lange et Content ont nommé *l'entropie moyenne de prononciation des graphèmes*. Cette variable d'entropie correspond à la mesure de l'incertitude de la prononciation de chaque graphème dans la séquence, de manière telle que cela reflète à la fois la probabilité de l'association régulière et celle des prononciations alternatives du graphème. Ainsi, pour des non-mots français, les graphèmes avec une forte valeur d'entropie (incertitude de prononciation faible) étaient lus plus rapidement que ceux avec une faible entropie.

Ainsi, dans une tâche de détection de lettres dans laquelle les sujets devaient indiquer si une lettre était présente ou non dans un mot rapidement présenté, alors que BEUDI et GEUDI sont tous deux des voisins du mot JEUDI, davantage d'erreurs ont été commises pour GEUDI (homophone par règle), indiquant l'existence d'une activation précoce de codes phonologiques sublexicaux qui influencent le processus de décision. Dans la comparaison des mots BONGOUR (homophone par activations multiples) et BONDOUR qui sont similaires du point de vue du voisinage orthographique de BONJOUR, les différences de performance constatées par Lange sont ainsi attribuées au fait que le premier active partiellement la prononciation /ʒ/ contextuellement irrégulière au cours de la conversion sublexicale. Or d'après le modèle DRC, le G dans BONGOUR ne devrait pas activer /ʒ/. Cela conforte donc la position de Lange selon laquelle la représentation de la force d'association graphème-phonème doit être incluse dans le système : si de multiples prononciations sont activées, c'est la plus forte et la plus fréquente qui est favorisée. Une telle approche s'accorde en revanche tout à fait avec des modèles interactifs comme celui de Norris (1994), dans lesquelles tous les associés phonémiques sont partiellement activés, activation qui est modulée par la force de l'association phonème-graphème dans la langue en question.

Cela rejoint également les travaux de Grainger, Nguyen Van Kang et Ségui (2001) sur les amorces inter-modales pour des homophones hétérographes : alors qu'un homophone hétérographe active simultanément toutes les représentations phonologiques associées, ainsi que les représentations sémantiques correspondantes, un homophone hétérographe déclenche, dans un premier temps, l'activation de l'ensemble des formes orthographiques compatibles, ainsi que celle des représentations orthographiques correspondantes. Dans ce cas, la fréquence jouerait un rôle crucial dans la sélection du candidat, *via* un processus d'inhibition.

Cette nouvelle dimension s'ajoute à celle mise en évidence par les travaux relatifs aux unités effectivement traitées (graphème-phonème *vs.* corps-rime) qui soulignent la

nécessité d'introduire plusieurs niveaux de représentations dans des modèles comme le DRC. Les résultats de Lange confirment également l'importance des bas-niveaux (graphème-phonème et lettre-phonème) dans le processus de conversion.

#### III.3.3.4. Conclusion

Ainsi, alors que l'orthographe française peut être qualifiée de « profonde », il semblerait que plusieurs facteurs influencent la tâche de lecture à voix haute de mots ou de non-mots en français, et que la taille des unités perceptives dépende précisément de ces facteurs (fréquence lexicale et sublexicale, phonologique, orthographique et phonographémique). Les interactions à l'œuvre entre ces facteurs soulignent ainsi l'interactivité existant entre les codes phonologique et orthographique dans les tâches considérées. Il semblerait, en tous cas, que l'unité « lettre » en français (cas de graphème simple) ne soit pas l'unique unité perceptive, et que dans certaines conditions, celle-ci soit négligée au profit d'unités de plus grande taille comme les graphèmes multi-lettres, les corps ou même l'équivalent orthographique des syllabes.

#### III.3.4. Le cas du japonais

Dans la perspective qui est la nôtre, il nous faut à présent examiner le cas du japonais, qui présente une complexité particulière, dû à son système d'écriture pluriel.

##### III.3.4.1. Introduction

Le système d'écriture du japonais étant triple, voire actuellement quadruple (kanji, katakana, hiragana et romaji), il constitue un champ d'étude intéressant pour ceux qui s'intéressent à la reconnaissance visuelle de mots. Ces différents types de script sont, de plus, combinés entre eux dans une orthographe « unique » dans les textes normaux de la vie quotidienne (Kess et Miyamoto, p. 191). On ne peut donc simplement qualifier le système japonais de « profond » ou de « superficiel », puisqu'il combine ces différentes caractéristiques.



Il apparaît aujourd'hui que les représentations « classiques » des processus de traitement de l'écrit en japonais (distinctions binaires entre kana et kanji, en termes d'hémisphères cérébraux mais aussi de voie d'accès lexical) doivent être revus, au vu de la multiplicité des facteurs influençant les résultats. Une telle révision bénéficie en particulier de bases de données psycholinguistiques lexicales, relatives, notamment, à la fréquence lexicale, à la familiarité scripturale et à la modalité perceptive, qui ont récemment commencé à être constituées en japonais (Kondo, Amano et Mazuka, 1996).

Ainsi, Katz et Frost (1992, p.79) rapportent certains résultats qui, grâce à un paradigme de variation graphique entre katakana et hiragana, laisseraient penser que les lecteurs japonais adoptent constamment une stratégie visuo-orthographique quelle que soit la profondeur du système scriptural : kanji (profonds) ou kana (superficiels). De la même manière, alors que le script kana est dit « superficiel », son traitement n'implique pas toujours la « voie d'assemblage », puisque plusieurs études ont mis à jour des effets de familiarité. La voie d'adressage jouerait donc un rôle dans la lecture des kana (Besner et Smith, 1992). De plus, le système graphémique adopté est lié à la strate lexicale de l'item en question dans le lexique japonais (Ito et Mester, 1999).

Ainsi, comme nous l'avons déjà indiqué, le réexamen de l'hypothèse de la profondeur orthographique et les études récentes sur la nature des unités fonctionnelles en perception de l'écrit soulignent la complexité des processus à l'œuvre. Cette complexité se retrouve donc en japonais.

Un ouvrage récent de description et de synthèse des études consacrées au traitement psycholinguistique du système d'écriture japonais est celui de Kess et Miyamoto (1999), dont la particularité est de puiser, non pas tant dans la littérature « occidentale », bien connue, mais plutôt dans la littérature japonaise, qui l'est encore aujourd'hui insuffisamment, alors que, comme ils l'écrivent en introduction, « Japanese scholarship into the nature of the mental lexicon represents one of the most important Japanese contributions to the field of cognitive science » (p. 1). Dans les lignes qui suivent, nous puisons donc abondamment dans cet ouvrage.

### III.3.4.2. Les Kana

Quatre aspects nous intéressent ici : la distinction katakana/hiragana, le lien entre unités phonologiques et unités kana, l'influence de la fréquence sur leur traitement psycholinguistique et enfin les bases cérébrales de leur traitement neurolinguistique.

#### III.3.4.2.1. Distinction katakana/hiragana et familiarité scripturale

Comme nous l'avons vu les kana se divisent en deux scripts : les katakana et les hiragana. La différence entre les deux est essentiellement fonctionnelle, étant donné la structure stratifiée du lexique japonais. Kess et Miyamoto (p. 88) citent plusieurs études contradictoires qui conduisent à soutenir la notion de vocabulaires distincts, bien que se chevauchant parfois, entre hiragana et katakana. Néanmoins, le facteur clé semble bien être celui de la familiarité : « if the form is appearing in katakana, it must be a new word and thus categorically more familiar in the katakana script than in the hiragana script » (p. 89). Il faut en effet souligner que le type de script a son importance, en tâche de lecture par exemple, étant attaché à un certain sous-ensemble lexical. Il faut donc clairement distinguer les niveaux auxquels l'on se situe : « At some level, the two script are seen as alternate paths to the same lexical address. But a common finding in the lexical access literature is that word familiarity, in addition to word frequency, is closely related to the ease of word perception » (p. 99).

Les deux formats scripturaux ne constituent donc pas deux représentations lexicales tout à fait distinctes, mais ne sont pas non plus complètement identiques. Le choix du script est ainsi parfois sujet à variation : certains mots peuvent être écrits en katakana dans certains contextes et en hiragana dans d'autres (ex : le mot « unagi » = anguille), d'autres varient entre kanji et katakana (ex : « gan » = cancer) et d'autres enfin entre hiragana, katakana et kanji (ex : « furo » = bath). Cette variation est cependant assez limitée. Il semblerait par ailleurs (p.100) qu'il y ait une certaine dépendance entre la familiarité et la modalité de présentation (visuelle vs. orale), le choix de l'orthographe

la plus familière (type de script) variant selon la modalité de présentation. Une certaine partie du lexique japonais serait donc « modality dependent ».

La familiarité du script ne peut donc s'exprimer en termes binaires : Kess et Miyamoto (p. 104) citent les travaux de Ukita, Sugishima, Minagawa, Inoue et Kashu<sup>133</sup> qui ont établi 10 catégories pour décrire le degré de fréquence interactive entre les trois types de script :

1. Mots souvent écrits en kanji mais rarement en katakana ou hiragana,
2. Mots souvent écrits en hiragana mais rarement en kanji ou katakana,
3. Mots souvent écrits en katakana mais rarement en kanji ou hiragana,
4. Mots souvent écrits en kanji,
5. Mots souvent écrits en hiragana,
6. Mots souvent écrits en katakana,
7. Mots souvent écrits en kanji et hiragana,
8. Mots souvent écrits en kanji et katakana,
9. Mots souvent écrits en hiragana et katakana,
10. Mots souvent écrits dans chacun des types scripturaux.

Il est par ailleurs important de souligner les différences générationnelles entre sujets étudiants universitaires et sujets d'âge moyen : « it seems that younger age groups exhibit a greater subjective variety of script types than older age groups, with the boundary between kanji and hiragana not as sharply defined in absolute terms » (p. 105).

Cependant, exclure l'exclusivité scripturale ne signifie pas exclure la familiarité scripturale, puisque le type scriptural semblerait en effet fonctionner comme un attribut mnésique (« a memory attribute ») (p. 107), comme cela a été montré dans des tâches

---

<sup>133</sup> Nihongo no hyoki keitai ni kansuru shinrigakuteki kenkyu [Psychological research on orthographic forms in Japanese], Psychological Monograph N°25, Tokyo, Japanese Psychological Association, 1996.

de reconnaissance orthographique. Ainsi, en japonais, le facteur de *familiarité lexicale* possède une autre dimension, celle de la *familiarité scripturale* (p. 204):

*« In some cases, the theory [of word recognition in Japanese] must reckon with script familiarity in the exclusive assignment of script for certain words. In other cases, it must deal with statistical approximations of script type assignments for other words, for script type assignments in these cases are not defined in binary terms ».*

Il faut donc noter que, à la fréquence lexicale et à la familiarité lexicale, vient s'ajouter le fait que les deux types de script présentent parfois certains chevauchements (p. 199).

#### III.3.4.2.2. Kana et unités phonologiques

En ce qui concerne la nature moraique ou phonémique des unités attachées aux kana, Kess et Miyamoto (p. 90) citent une étude indiquant:

*« Although the mora is the smallest unit of orthographic representation, Japanese subjects were obviously sensitive to phonemic segments in processing these kana "words", suggesting that phonemic segments are relevant to phonological processing of kana representations ».*

Une fois encore, il semblerait donc que le traitement des unités orthographiques soit davantage pluriel (phonème, mora, syllabe) qu'unique.

#### III.3.4.2.3. Fréquence et voie d'adressage pour les kana

En ce qui concerne la voie (d'assemblage ou d'adressage) sollicitée en tâche de lecture, il serait illusoire de croire que seuls les kanji utilisent la voie d'adressage : les kana aussi, en particulier les hiragana. Le critère clé ici est celui de la fréquence d'occurrence : « low frequency katakana words require phonemic processing while

high frequency katakana words allow direct access to their meaning » (p. 96). Ces résultats proviennent d'études croisant les deux variables « type de script » (kanji, kana) (renvoyant à la familiarité, tel mot étant normalement écrit avec tel script) et fréquence d'occurrence (basse/élevée).

Toujours à propos des kana et de la différence entre les voies d'assemblage et d'adressage, il semblerait que la fréquence d'une configuration de kana, en termes de formes de mot fréquentes, joue un rôle important dans la reconnaissance lexicale. Ainsi, « the type of orthography per se does not exclusively control processing time ; familiarity with the frequency of the orthographical shape is a highly influential factor in lexical access » (p. 103). Il faut également noter que la variation individuelle constatée dans certaines études conduit les auteurs à affirmer :

*« These results point to individual strategies favoring the possible involvement of either direct access via the orthographic shape or access via the phonological assembly route, as well as the differential weighting of factors such as frequency, familiarity, and conventionality determining which one of the two routes will be taken » (p. 103).*

#### III.3.4.2.4. Hémisphère gauche vs. hémisphère droit pour les kana

D'un point de vue neurolinguistique, il semblerait que le traitement de mots en kana qui nécessite un encodage phonémique implique l'hémisphère gauche, et que l'hémisphère droit soit séparé de ces fonctions d'encodage. Cependant, « the assumption that the right hemisphere is dissociated with the processing of kana words does not necessarily mean that this hemisphere does not contribute to the processing of individual kana characters » (p. 172). Malgré cela, au vu des résultats obtenus jusqu'à présent à ce sujet, « it would appear that the left hemisphere undertakes the linguistic analysis that kana symbols require in their usual appearance in language materials. The evidence for anything other than this conclusion seems to be weak » (p. 173).

### III.3.4.2.5. Conclusion

En conclusion de leur partie consacrée aux kana, Kess et Miyamoto indiquent que, même si les deux syllabaires partagent de nombreux attributs en tant que types scripturaux, « they also differ in the way in which they are tied to the words which they represent, and not in the simple way they have been traditionally viewed as carriers of native vs. borrowed words » (p. 110). Du point de vue du traitement de l'information, on ne peut donc pas affirmer que ces deux types de syllabaire ne constituent que des variantes représentationnelles des mêmes structures syllabiques, même si l'on trouve des effets d'amorce « inter-script » et même si, du point de vue de la transcription, les hiragana et les katakana partagent les mêmes points de référence syllabiques, pouvant transcrire les mêmes mots selon les mêmes principes. Pour des lecteurs japonais natifs, les deux systèmes de kana se distinguent en effet au moins de trois façons essentielles (p. 205) :

- « *Differential frequencies with individual words* »;
- « *Katakana words receive the longest eye-fixations, invoke shorter saccades and invite regressive movements more than kanji and hiragana (from eye-movements studies)* »;
- « *Katakana themselves may be more difficult to process because of their minimalist angular shapes, the fact that they often overlap in whole or part, and the fact that they give rise to perceptual confusion in a way that the more distinctive, cursive hiragana do not* ».

### III.3.4.3. Les Romaji

A propos du système des romaji, Kess et Miyamoto indiquent que, bien qu'il ne soit pas reconnu formellement comme faisant partie du système orthographique traditionnel, il est omniprésent (p. 111). Il tire son nom, romaji, du fait qu'il utilise 22 des 26 lettres de l'alphabet romain ou latin. Par ailleurs, les numéraux arabes sont majoritairement utilisés dans les écritures horizontales (*yokogaki*), tandis que les numéraux en kanji restent majoritaires dans les écritures verticales (*tategaki*).

L'utilisation des romaji, en particulier dans le domaine des médias et de la publicité, est en augmentation exponentielle, d'après certains auteurs, ce qui les conduit à perdre leur caractère exotique : « words from English, French, German, and other languages that only a few years ago were written in katakana are now often imported and written in the original roman alphabet » (p. 112). En ce qui concerne leur apprentissage par les écoliers japonais, la carte des romaji est introduite dès la quatrième année (« Grade 4 »), ce qui n'est pas au goût de tous, comme le soulignent Kess et Miyamoto (p. 114) :

*« Some educators see its introduction as a disadvantage in English language studies, arguing that romaji hinders rather than facilitates the acquisition of English vocabulary because the gap between English source words (table, computer) and Japanese borrowed renditions (teeburu, kompyuutaa) is simply too large to be anything but confusing ».*

#### III.3.4.4. Les Kanji

Le traitement cognitif des kanji a été l'objet de nombreuses études, notamment en neuropsycholinguistique. Cependant, les résultats les plus récents obligent à réexaminer, ou du moins à nuancer, certaines conclusions antérieures.

##### III.3.4.4.1. Implication de la voie phonologique pour les kanji

A propos du traitement psycholinguistique des kanji en reconnaissance visuelle, l'implication de la phonologie dans une telle tâche est à présent bien documentée (Matsunaga, 1994), contrairement aux « croyances » traditionnelles selon lesquelles les kanji permettaient, de manière systématique, un accès au sens sans évocation phonologique. La question restante à propos de l'activation de l'information phonologique est cependant la suivante (Kess et Miyamoto, 1999, p. 57) : « Does it take place at the pre-lexical stage before word recognition has been achieved ? Or at the post-lexical once word recognition has been achieved ? Or is it at the *at-lexical* event as an integral part of the word recognition event ? ». Dans tous les cas,

l'argument selon lequel « Japanese language users are somewhat dependent upon the graphic image of the word that is stored in their minds, and knowing the kanji shape for the word is then an important factor in their comprehending what was just said » doit être passé à la loupe afin de ne pas mélanger, comme le soulignent Kess et Miyamoto (p. 87), « cultural beliefs » et mesures psycholinguistiques.

#### III.3.4.4.2. Hémisphère gauche vs. hémisphère droit pour les kanji

Comme pour les kana, il y a de nombreux facteurs à l'œuvre dans le traitement des kanji : « the observed right hemisphere superiority of single kanji may be due to other factors, such as kanji stimuli being interpreted as visuospatial stimuli. Then too, the presentation of single kanji stimuli does not invariably lead to right hemisphere superiority processing » (p. 175). En ce qui concerne les kanji composés en chinois, excepté une étude citée par Kess, aucune supériorité droite n'a été montrée à ce jour (p. 177) :

*« If anything, kanji are generally processed as linguistic elements, and as a result, require the involvement of the left hemisphere. [...] The dichotomous view of kana and kanji processing by script type and hemisphere is too simplistic a view to explain the data which arise from various experimental and clinical enquiries ».*

En particulier, le traitement des kanji est sensible à l'interaction entre le type de stimuli et le type de tâche : Kess et Miyamoto mentionnent la préférence droite pour les numéraux en kanji et la préférence gauche pour les numéraux arabes, une préférence qui disparaît en tâche de calcul mental.

*« The simple view that kanji are processed directly as « sight words » and kana are processed through phonological recoding is simply not the case either. If anything, both logographic and syllabary scripts give evidence of both types of*



*informational features being accessed, depending upon the task, the frequency and familiarity of the individual word, and the linguistic context » (p. 110).*

Toujours à propos de la latéralité hémisphérique chez les Japonais, les auteurs citent certains travaux de Paradis, Hagiwara et Hildebrandt<sup>134</sup> (1985) qui concluent : « Single Chinese hanzi or Japanese kanji can produce the same laterality pattern as English words, given the right conditions of presentation » (p. 168).

Ainsi, de nombreuses études ont été conduites depuis l'établissement de la première dichotomie hémisphérique chez les Japonais (« left hemisphere-phonological processing of kana / right hemisphere-visual processing of kanji »), et toutes ces études conduisent à une révision de la version « binaire » (« the "dichotomous view" ») du traitement des kana et des kanji (p. 170). Ce ne serait donc pas tant la nature du stimulus en lui-même qui orienterait la latéralité que l'interaction entre le stimulus et la tâche cognitive demandée (p. 12).

De plus, outre les variables physiques propres aux manipulations expérimentales, il y a aussi des variables qualitatives comme les parties du discours, la distinction abstrait/concret ou encore la familiarité des kanji stimuli qui peuvent avoir un effet sur la latéralisation. Kess et Miyamoto concluent (p. 186) : « while the configurational aspects of kanji may be processed both by the right and left hemisphere, the phonemic and semantic aspects of kanji processing are predominantly handled by the left hemisphere ».

Toujours d'un point de vue neurolinguistique, Nakamura et Kouider (2003), au vu des données récentes en neuro-imagerie fonctionnelle concernant la production écrite en

---

<sup>134</sup> Paradis, M., Hagiwara, H., Hildebrandt, N. (1985), *Neurolinguistic Aspects of the Japanese Writing System*, Orlando, FL : Academic Press.

japonais, concluent également que celle-ci implique des structures cérébrales communes, quel que soit le type de script : la seule différence entre l'écriture des kana et celle des kanji serait que celle des derniers nécessite la récupération d'informations visuo-spatiales concernant des caractères, une récupération qui constitue une opération cognitive supplémentaire.

### III.3.4.5. Conclusion

C'est donc pour une certaine universalité des processus que penchent Kess et Miyamoto (p. 193) :

*« Our review of a wide range of studies suggests that the same basic principles of lexical access operate across the three types of orthographic activity. That is, orthographic configurations can be used to jump directly to phonological form, to meaning, or to both, depending upon the processing task at hand and their advantageous or necessary interaction ».*

On ne peut donc plus aujourd'hui continuer à penser que « the cognitive considerations in processing Japanese orthography are unique, with kana processed by the left hemisphere and kanji processed by the right hemisphere ». Des facteurs comme la fréquence semblent être cruciaux dans la médiation phonologique, ou son absence, en reconnaissance visuelle, les mots à fréquence élevée étant alors indépendants de la profondeur de l'orthographe (p. 197) :

*« If this is true, then differences between orthographies will reside more in how the lower items are represented phonologically and the way in which phonology plays out its role in reading. [...] The prototypical example of this must be the function words. [...] Kana are not always superior to kanji in*

*accessing the phonological code, and kanji do not always invoke rapid and automatic access to the semantic code » (p. 196).*

Ainsi, dans certaines conditions, un accès direct est possible pour les kana, tandis que l'information phonétique peut être évoquée, lorsque celle-ci est disponible, dans certains cas de traitement des kanji. Lorsque l'on compare le système orthographique japonais avec un système alphabétique comme celui de l'anglais ou du français, il est donc important de se rappeler que (p. 197) :

*« Both the phonological and the orthographic pathways can interact as the processor seeks to identify and activate the mental representation of the words. In fact, a realistic interactive explanation should postulate that all information is being accessed, namely, phonological, orthographic (or graphemic), kinetic (or motoric), and semantic information ; lexical access procedures utilize what they can and what they need from the sources until successful word recognition takes place ».*

Le japonais se présente donc comme un système particulièrement complexe, dans lequel interviennent de nombreux facteurs, certains étant universels (par exemple la fréquence lexicale), d'autres spécifiques (par exemple la familiarité scripturale). Dans notre perceptive, il convient de retenir :

- La complexité du lien entre strate lexicale et script (mots étrangers écrits en katakana et/ou alphabet romain ; mots japonais transcrits en romaji, et lien entre alphabet romain et romaji) ;
- L'influence de la fréquence sur les scripts « superficiels » kana et romaji ;
- La nature essentiellement moraique des unités orthographiques minimales en japonais (en particulier pour les scripts kana et romaji).

### **III.4. Reconnaissance lexicale et lecture à voix haute : modèles**

Nous présentons ici les deux types de modèle actuellement dominants dans le domaine de la reconnaissance lexicale visuelle, les modèles à deux voies et les modèles interactifs. Ces deux types de modèle reposant tous deux, dans une certaine mesure au moins, sur un principe d'activation et d'inhibition des unités pertinentes et non pertinentes, nous décrivons en premier lieu le modèle « Cohorte » de Johnson (1992), qui permet d'introduire le concept de « cohorte », et ainsi ceux de « cascade » et de « voisinage ». Nous concluons cette partie en évoquant brièvement la place accordée à la morphologie et de la prosodie dans de tels modèles.

#### **III.4.1. Le Modèle Cohorte**

En ce qui concerne la reconnaissance visuelle de mots, Johnson (1992) présente un modèle de type Cohorte, qu'il a élaboré en s'inspirant des travaux de Marslen-Wilson dans le domaine de la reconnaissance de mots parlés. Ce modèle a été constitué en tenant compte de trois facteurs (p. 147) :

- « Words conceal their letters » (« word-priority effect »).
- « Lexical processing is facilitated by increasing word frequency ».
- « Lexical processing is influenced [...] by the number of words that are visually similar to the target ».

Ce modèle comporte deux composants essentiels :

1. L'un concerne l'encodage visuel de l'information orthographique ;
2. L'autre concerne la relation entre l'information issue de cet encodage et l'activation ou la désactivation des représentations lexico-sémantiques.

En outre, trois types de représentations sont employés : « feature codes », « letter codes » et « wickelgraphs ». Les wickelgraphs visent à saisir le contexte orthographique dans lequel se trouve la lettre cible, et ne constituent donc pas l'encodage d'un simple trigramme. Johnson donne l'exemple du mot anglais « THE »

et des trois wickelgraphes correspondants : #Th, tHe, hE# (avec # représentant une frontière de mot). Il écrit (p. 148) :

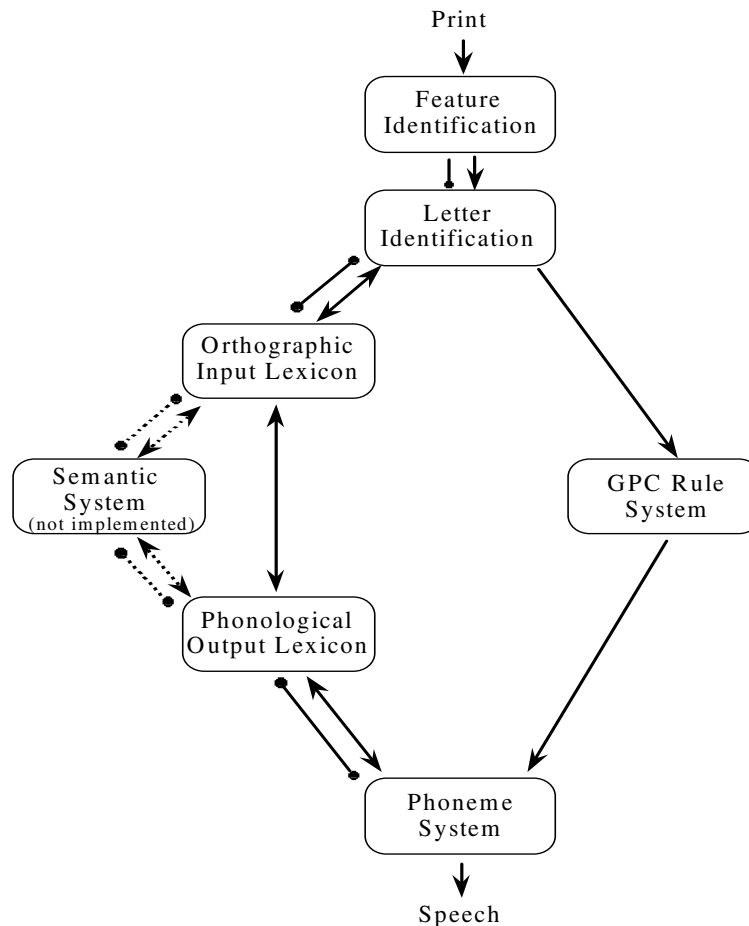
*« The wickelgraph is an encoding of the identity of the critical letter (the H in tHe), plus an opaque representation of the context letters, with the latter portion of the encoding being used as a means of defining the orthographic context in which the wickelgraph's identity can appear ».*

Ces trois types de représentation correspondent aux trois étapes de traitement de l'encodage orthographique, lequel relève de deux modules : l'un pour les traits simples, le second pour les relations structurelles du type inclusion d'unité inférieure (traits ; lettres) à supérieure (lettres, groupes de lettres). Dès que le premier wickelgraphe est encodé, l'ensemble des entrées lexicales auxquelles il correspond, *via* un schéma orthographique abstrait (p. 152), est activé (la cohorte initiale). Le schéma d'activation se modifie graduellement, au fur et à mesure de l'encodage des autres wickelgraphes : « the cohort is resolved when all nontarget members of the cohort have lost activation » (p. 151). L'activation des entrées lexicales, et leur maintien dans la cohorte, est notamment fonction de leur fréquence d'usage (« the words' cultural frequencies » (p. 155)). Il peut aussi arriver que la conversion sous forme phonétique d'un mot dont l'orthographe, mais pas la prononciation, est inconnue ou incertaine, déclenche l'émergence d'une cohorte phonologique cette fois-ci.

#### **III.4.2. Le Modèle à deux voies en cascades (Dual-Route Cascaded Model, DRC)**

Parmi l'ensemble des modèles dits « à deux voies », le modèle DRC de reconnaissance visuelle lexicale et de lecture à voix haute est sans doute le plus reconnu.

Il se présente ainsi (Ziegler, Perry et Coltheart, 2000, p. 416) :



Le modèle présente donc deux « voies » ou « routes » :

- 1) Route lexicale (voie d'adressage) : deux principes sont à l'œuvre. Le premier est celui d'une activation interactive bidirectionnelle entre différents niveaux (axe vertical) : activation excitatrice entre unités compatibles et activation inhibitrice entre unités incompatibles. Le second est celui d'une inhibition au sein d'un même niveau (axe horizontal). Les lexiques phonologiques et orthographiques sont reliés par des connexions uniquement excitatrices : un mot orthographique ne peut

inhiber tous les mots phonologiques incompatibles et vice-versa. Ce mécanisme est assuré par les inhibitions latérales au sein d'un même niveau.

- 2) Route non-lexicale (voie d'assemblage): les lettres ou groupes de lettres sont converties en phonèmes uniques *via* un processus sériel de conversion phonographémique (*Grapheme-Phoneme Correspondences*, GPC).

Paap, Noel et Johansen (1992) soulignent que la voie d'assemblage est davantage *contrôlée* et nécessite davantage de ressources attentionnelles, alors que la voie d'adressage serait plus *automatique*, sans pour autant être incontournable (p. 312).

Ce modèle, initialement conçu pour l'anglais, a également été appliqué au français et à l'allemand. Deux notions sont ici importantes (Ziegler et *al.*, 2000, p. 415) :

- ◆ Le principe de constance morphémique : quels que soient les changements de prononciation induits par le contexte, l'orthographe (« spelling ») des morphèmes reste constante (par exemple, en anglais, *Heal vs. Health*)
- ◆ Irrégularité orthographique : Pour toute orthographe alphabétique, le degré d'irrégularité de celle-ci peut être établi, une fois un ensemble de règles de correspondances phonographémiques (GPC) spécifié pour la langue en question. Le degré d'irrégularité correspond alors au pourcentage de mots pour lesquels les règles de prononciation sont en désaccord avec la prononciation lexicale. Dans le cas de l'allemand, elle est assez faible et provient généralement des emprunts, des noms propres et des termes géographiques.

Dans le modèle original (anglais) trois types de règles coexistent (*op. cit.*, p. 418) :

- ◆ Règles sur lettre simple : elles s'appliquent lorsqu'une lettre unique correspond à un phonème unique.

- ◆ Règles sur groupes de lettres : elles s'appliquent lorsque plusieurs lettres correspondent à un phonème. Le nombre de ces règles va donc varier selon le nombre de graphèmes multi-lettres que comporte la langue en question. Ainsi l'anglais en possède 146, tandis que l'allemand en possède 48.
- ◆ Règles sensibles au contexte : elles s'appliquent lorsque les lettres précédentes ou suivantes déterminent de manière consistante la prononciation d'un graphème particulier.

Toutes ces règles sont du type graphème-phonème et l'unité phonémique est l'unité maximale. Elles sont par ailleurs liées à la position dans le mot : initiale, médiale ou finale.

Ce type de modèle est défendu par de nombreux chercheurs, comparant la théorie de la double voie à une « course de chevaux », qui peut parfois aboutir à une compétition et interférence entre les réponses obtenues par chacune des deux voies (Paap, Noel et Johansen, 1992, p. 298), compétition dans laquelle interviennent les facteurs de fréquence et de régularité (p. 302). Ces chercheurs sont par ailleurs favorables à une vision « souple » de ces modèles (p. 307) :

*« Dual-process theory need not embrace rules based on all-or-none GPC's [Grapheme-Phoneme Correspondences] [...] and a more flexible OPC [Orthographic to Phonological Conversion] rulebase is immune to many of the shortcomings of a GPC system. We are perfectly willing to [...] contemplate dual-route models that have no rules at all ».*

Or, c'est précisément la « rigidité » relative du modèle et son approche basée sur des « règles », qui ont conduit à l'élaboration de modèles davantage « interactifs », certains d'entre eux pouvant, dans une certaine mesure, également être considérés comme étant « à double voie » (Zorzi, Houghton et Butterworth, 1998).



### **III.4.3. Modèles Interactifs : connexionnisme et intégration**

Si l'appellation « interactif » est généralement essentiellement attribuée aux modèles des approches connexionnistes (Norris, 1994 ; Zorzi, Houghton et Butterworth, 1998 ; Kroll et Dijsktra, 2000), nous y associons également ici le Modèle Flou de la Perception (Massaro et Cohen, 1994), étant donné le poids qui est accordé à l'ensemble des informations disponibles au sujet et au traitement parallèle qu'il est supposé leur appliquer. Nous qualifions alors ce modèle d'« intégratif ». Ces deux types de modèle sont destinés à rendre compte de la lecture tant des mots que des non-mots.

#### **III.4.3.1. Les Modèles d'Activation Interactive**

Les modèles dits « interactifs » de reconnaissance visuelle, sont fondamentalement tributaires des travaux de McClelland et Rumelhart, qui, en 1981, ont élaboré le Modèle d'Activation Interactive. Ce dernier repose sur les mêmes fondements, connexionnistes, que le modèle TRACE de perception de la parole (McClelland et Elman, 1986). Dans ce modèle, trois niveaux sont distingués : celui des traits, celui des lettres et celui des mots (McClelland, 1986). Ces trois niveaux sont reliés dans un réseau de type connexionniste (activation/ inhibition/ pondération).

Kroll et Dijsktra (2000) ont adapté ce modèle au cas des sujets bilingues (une perspective intéressante pour nos sujets apprenants de langue étrangère) : dans leur modèle, les auteurs proposent un lexique unique intégrant les deux langues, et ils ajoutent aux trois niveaux originaux un quatrième niveau (supérieur) dont les unités permettent « d'étiqueter » les mots pour indiquer la langue à laquelle ils appartiennent (p. 304). Le fonctionnement d'un tel modèle, face à une séquence de lettres, est alors le suivant : 1) Activation parallèle de mots possibles (le « voisinage ») dans les deux langues, 2) Compétition entre candidats activés et suppression des activations (*via* des inhibitions latérales) jusqu'à ce que l'un d'eux dépasse son seuil d'activation et soit reconnu. Dans le cas des bilingues non équilibrés (comme peuvent l'être des apprenants de langue étrangère), ce modèle permet de rendre compte de l'influence de la langue dominante sur la seconde (notamment en termes de fréquence et ainsi en

termes de seuil ou de mode d'activation). Le quatrième niveau (celui de l'appartenance linguistique) permet de rendre compte de l'effet du contexte, en exerçant une influence haut-bas inhibitrice sur les mots de la langue non pertinente. Kroll et Dijsktra notent cependant que leur modèle n'aborde pas le cas des bilingues dont les langues ne partagent pas le même « alphabet » (p. 321).

Un autre modèle connexionniste d'activation interactive pour la lecture est celui de Norris (1994) : comme l'indique son nom, le *Modèle à Multiples Niveaux* consiste en un système d'activation interactive combinant plusieurs niveaux de correspondances phonographémiques : phonèmes, rimes et mots. L'idée n'est pas neuve, puisque ce modèle s'inspire notamment des travaux de Shallice et de ses collaborateurs, qui avaient proposé sept niveaux de correspondances sublexicales (dont le morphème, la syllabe et la rime), allant ainsi au-delà des règles de correspondances phonèmes-graphèmes traditionnelles des modèles à double voie. Le principe essentiel du modèle est celui d'une compétition entre correspondances phonographémiques à plusieurs niveaux. Cette multiplicité fait que le modèle peut également être considéré comme un modèle à double voie (p. 1229).

#### III.4.3.2. Le Modèle Flou de la Perception

Dans le cadre du *Modèle Flou de Perception* (FLMP), Massaro et Cohen (1994) insistent sur la multiplicité des sources d'information utilisées en tâche de lecture (reconnaissance lexicale et sublexicale) : notamment la connaissance de la structure orthographique, celle des correspondances phonographémiques et la fréquence lexicale. Pour les auteurs, la lecture ne constitue qu'un cas particulier de reconnaissance perceptive (« perceptual recognition »), traitée dans le modèle *via* trois étapes (évaluation, intégration, décision) : « Written language [...] is just one of many situations in which context and stimulus information are exploited in pattern recognition » (p. 1124).

Ce modèle se distingue d'autres, comme le DRC, en ce qu'il suppose que les sources d'information sont traitées simultanément, et que l'ensemble des contraintes (orthographiques, phonologiques, syntaxiques, sémantiques et pragmatiques) influencent le traitement lexical et sublexical : « For example, "the boy down fell the hill" is gramatically incorrect, perhaps making it more difficult to read the word "fell" » (p. 1109). Cependant, en ce qui concerne l'influence du contexte sur la reconnaissance des lettres, le FLMP se distingue notamment du *Modèle d'Activation Interactive* (IAM) de McClelland et Rumelhart : alors que dans le IAM, le contexte modifie la représentation de la lettre cible, le FLMP considère que le contexte lexical fonctionne comme une source d'information supplémentaire (p. 1111), traitée simultanément (contrairement au IAM qui suppose que le contexte intervient ultérieurement). Pour Massaro et Cohen, l'effet de supériorité lexicale (une lettre est plus facilement reconnue dans un mot que dans un non-mot) serait simplement dû au fait qu'il est plus facile de reconnaître un mot qu'un non-mot, et non pas, comme le suppose le IAM, au fait que l'activation de la lettre serait, d'une certaine manière, modifiée selon que celle-ci se trouve dans un mot ou un non-mot (p. 1111).

En ce qui concerne le voisinage orthographique (et ainsi les « cohortes » lexicales), Massaro et Cohen adoptent une position critique vis-à-vis de l'importance qui lui a été accordée, défini comme « a mismatch on one letter between the target word and other words of the same length » (p. 1120). Une telle mesure ne prend en effet pas en compte d'autres facteurs important de la structure orthographique : a) la position des lettres discordantes, b) les voisins lexicaux de longueur différente, c) la similarité (visuelle) des lettres discordantes. L'influence du voisinage orthographique sur la lecture doit donc être intégrée à celles des autres variables impliquées.

Ainsi le FLMP se présente comme (p. 1126) :

- Un modèle « indépendant » (« independence model ») : au niveau des lettres le traitement sensoriel n'est pas influencé par le contexte (contrairement au modèle

« non-indépendant » IAM), ce dernier étant en revanche considéré comme une source d'information supplémentaire.

- Un modèle intégratif : il prend en effet en compte l'ensemble des facteurs influant sur la performance de lecture : fréquence lexicale, fréquence sublexicale et correspondances phonographémiques.
- Un modèle mixte vis-à-vis des « deux voies » (lexicale et phonologique) utilisables en lecture : plutôt qu'une « course de chevaux » (Paap, Noel et Johansen, 1992) à vainqueur unique, Massaro et Cohen considèrent que l'information orthographique et les correspondances phonographémiques contribuent simultanément et en parallèle à l'accès lexical, lequel s'effectuerait, dans l'approche du FLMP, de manière continue et progressive, grâce à l'accumulation d'informations de sources multiples.

#### III.4.3.3. Morphologie et Prosodie en reconnaissance visuelle

Deux points nous ont semblé manquer dans la présentation précédente : il s'agit de la morphologie et de la prosodie, que nous nous contentons ici d'évoquer.

##### A) Et la morphologie ?

Si de nombreux modèles de reconnaissance visuelle lexicale ont considéré que le traitement lexical impliquait des informations orthographiques, phonologiques et sémantiques, l'information morphologique, en revanche, avait pendant longtemps été relativement négligée, alors que celle-ci se trouve précisément à l'intersection des trois précédentes (Frost et Grainger, 2000). Or, lorsque l'on s'intéresse à la reconnaissance de mots, il faut précisément examiner le type de mot à reconnaître : en particulier d'un point de vue morphologique, s'agit-il d'un mot « racine », dérivé ou infléchi, préfixé ou suffixé ? L'information morphologique doit en effet être intégrée aux modèles de reconnaissance, étant donné que son importance dans le processus de reconnaissance des mots est de plus en plus reconnue (Beauvillain et Ségui, 1992).

Différents modèles et approches existent à ce sujet (décomposition morphologique sublexicale, modèle duaux, ou modèle de représentation morphologique supralexical), et le rapport entre modalité et morphologie est justement au cœur d'un projet de recherche en cours (2003-2007), animé par J. Grainger, D. Sandra et K. Diependale et intitulé « Are mental representations of language morphology modality specific or not ? ». Rastle, Davis et New (à paraître), par exemple, ont conduit certains tests, dont les résultats suggèrent qu'il existerait un certain niveau de représentation auquel des mots apparemment complexes seraient décomposés en fonction de leurs propriétés morpho-orthographiques.

Si la question peut s'avérer moins pertinente dans le cas de la reconnaissance en langue étrangère pour des apprenants non experts que pour des sujets natifs, il importe néanmoins de mentionner l'existence du facteur morphologique.

#### B) Et la prosodie ?

On ne peut manquer de remarquer l'absence globale des phénomènes prosodiques dans les travaux sur la reconnaissance visuelle des mots. Il serait en effet surprenant de croire que les unités lexicales sont totalement dépourvues de caractéristiques prosodiques, et cette question se pose particulièrement lorsque sont traitées des langues tonales ou à accent lexical. En lecture de textes ou d'énoncés, on peut en effet se demander : Comment déterminer le patron intonatif à assigner aux énoncés que l'on est en train de lire ? Ajoute-t-on l'enveloppe intonative une fois le « sens » de l'énoncé compris (post-lexical/post-sémantique) ? A quel moment la courbe intonative assignée à telle partie du discours en tâche de lecture à voix haute, en particulier en langue étrangère est-elle déterminée ?

Kondo et Mazuka (1996) par exemple se sont interrogés sur la manière dont était générée en temps réel la prosodie d'une phrase lue à voix haute en japonais, *via* la mesure de l'écart entre positions oculaires et production orale, en variant la complexité syntaxique des phrases. Il apparaît, étant donné le peu de caractères visualisés à l'avance, que les lecteurs génèrent les patrons prosodiques non pas sur la base d'une

analyse syntaxique globale, mais plutôt locale, ce qui entraîne des erreurs et des « réparations » lorsque ces productions prosodiques locales contredisent le sens de la phrase en question. On peut également citer la thèse de troisième cycle de P. Nicolas intitulée « La structuration prosodique du texte lu en français » (1995), mais il s'agit d'un travail de description et de modélisation davantage ancré dans le domaine phonétique-prosodique que dans celui, psycholinguistique, de la reconnaissance visuelle. On ne peut donc que constater la faible quantité d'information disponible à ce sujet.

#### **III.4.4. Conclusion**

A travers ces derniers modèles interactifs, mais également à travers l'évolution des modèles à double voie « classiques », il apparaît que la question de *l'interactivité* entre codes phonologiques et orthographiques à différents niveaux est aujourd'hui devenue incontournable. Si le code orthographique dépend bien évidemment de la phonologie de la langue en question, et si l'activation des codes phonologiques, à divers degrés, semble inévitable en tâche de lecture, il n'en reste pas moins que les représentations orthographiques ont une existence cognitive à part entière. Comme nous l'avons vu, l'information disponible en tâche de lecture est au moins triple : phonologique, phonographémique, mais également purement orthographique.

Une fois acquises, il serait ainsi possible que ces représentations orthographiques puissent à leur tour influencer les représentations phonologiques, ou du moins interagir avec elles, lors du traitement de l'écrit, mais également lors du traitement de la parole, puisque chaque unité linguistique possède alors une représentation phonologique et une représentation orthographique. Il convient donc à présent de s'interroger plus en détail sur l'influence que peuvent exercer les représentations orthographiques sur la perception et le traitement de la parole, en premier lieu pour des sujets natifs monolingues, puis pour des sujets apprenants de langue étrangère.

### **III.5. Influence de l'orthographe en perception de la parole**

Nous envisageons cette influence potentielle sous deux angles : celui des activations et inhibitions possibles entre unités phonologiques et orthographiques à un niveau lexical et sublexical, et celui du développement de la conscience phonologique.

#### **III.5.1. Interactions entre codes orthographiques et phonologiques**

Dans une perspective fonctionnellement modulariste, les systèmes de traitement orthographique et de traitement phonologique peuvent être considérés de manière relativement indépendante l'un de l'autre. Cependant, de nombreuses études font également état d'une dépendance forte entre les deux systèmes, tant au niveau sublexical qu'au niveau lexical, que ce soit en tâche de reconnaissance visuelle ou auditive.

##### **III.5.1.1. L'hypothèse de l'autonomie orthographique : les données neurolinguistiques**

Dans un cadre neuropsycholinguistique, certaines données issues de la pathologie du langage invitent en effet, *via* les phénomènes de « double dissociation » ou de « dissociation simple » (Nespoulous, 2004), à distinguer les systèmes de traitement de la parole de ceux de l'écrit, et à l'intérieur de ceux-ci à distinguer encore des sous-modules liés à la performance à accomplir (production/perception, orale/écrite). Par exemple, en ce qui concerne les sujets globalement qualifiés de « dyslexiques » (présentant des troubles de la lecture), Zesiger et Partz (1994, p. 426), en se référant au modèle à deux voies de la lecture, illustrent des cas de distinction entre altération de la procédure d'assemblage et altération de la procédure d'adressage. En ce qui concerne les sujets globalement qualifiés d' « aphasiques », Nespoulous et Dordain mentionnent les cas de patients (1994, p. 266) :

- Qui peuvent comprendre une phrase qu'ils ne peuvent produire (→ distinction compréhension/production),
- Qui peuvent répéter une phrase qu'ils ne peuvent comprendre (→ distinction répétition/compréhension),

- Qui peuvent énoncer oralement une phrase qu'ils ne peuvent écrire (→ distinction production orale/écrite),
- Qui ne peuvent lire/relire un énoncé qu'ils ont eux-mêmes écrit (→ distinction production/compréhension écrite),
- Qui répètent mieux qu'ils ne lisent à voix haute, et vice-versa pour d'autres patients (cas de double dissociation) (→ distinction répétition/lecture à voix haute).

Ces dissociations entre traitement de l'oral et de l'écrit, en aphasiologie, sont essentiellement constatées chez des patients souffrant de lésions dans le cortex primaire (moteur, auditif, visuel), lequel constitue « la porte d'entrée et de sortie du langage ». Il ne s'agit donc pas de perturbations linguistiques « profondes », puisque la capacité linguistique abstraite de ces sujets reste intacte. Les quatre principaux cas sont les suivants (Nespoulous, 2002) :

- Anarthrie pure : trouble spécifique de la production orale (au plan phonétique/articulatoire, mais pas au plan phonémique), en l'absence de tout autre trouble (compréhension orale, production/compréhension écrite intactes).
- Agraphie pure : trouble spécifique de la production écrite (mais capacité de dactylographie et capacité de « jouer au scrabble » intactes), en l'absence de tout autre trouble (compréhension écrite, production/compréhension orale intactes).
- Surdit  verbale pure : trouble spécifique de la compréhension orale (incapacité à décoder les messages auditifs : les sujets réagissent comme si on leur parlait en langue étrangère. Mais, par exemple, ils reconnaissent un miaulement de chat ou le caractère interrogatif d'une intonation) ; production orale quasi-normale et production/compréhension écrite intactes.
- Cécité verbale pure (alexie sans aggraphie) : trouble spécifique de la compréhension écrite (perte de la capacité de lire). Le texte apparaît aux sujets comme un texte



écrit en langue étrangère. Ceux-ci peuvent cependant reconnaître leur propre écriture (niveau graphique), mais ne peuvent déchiffrer ce qu'ils ont écrit (niveau graphémique). Ils peuvent également reconnaître une lettre qui leur est tracée dans la paume de la main (modalité tactile). C'est donc bien la modalité visuelle qui est perturbée, la modalité auditive étant, quant à elle, intacte.

Cependant, en ce qui concerne les lésions plus profondes, par exemple dans le cortex secondaire dans le cas d'une aphasie de Broca (troubles massifs de la production, mais pas ou peu de troubles en compréhension), les patients présentent un certain parallélisme (relatif) dans leurs symptômes entre l'oral et l'écrit : anomie (manque du mot), agrammatisme (style télégraphique), troubles phonologiques/orthographiques, troubles phonétique/graphiques. Toutefois, ce parallélisme n'est pas parfait, puisque l'on trouve par exemple des sujets chez qui l'agrammatisme est plus marqué à l'oral qu'à l'écrit.

Kouider et Dupoux (2001) ont également mis à jour une certaine autonomie des systèmes à l'œuvre en reconnaissance lexicale visuelle et orale. Ils rappellent dans un premier temps que les tentatives de réponse à la question « Is the reading processing system dependent from the auditory system, or has it acquired the status of an autonomous processing module ? » ne font pas encore l'unanimité (p. 36) : tandis que certaines études neuropsychologiques démontrent l'indépendance des deux systèmes (*l'hypothèse de l'autonomie orthographique*, selon laquelle la médiation phonologique entre l'orthographe et le sens n'est qu'optionnelle), d'autres études en psycholinguistique rapportent de nombreuses interactions entre les systèmes phonologiques et orthographiques dans de nombreuses tâches, suggérant que la voie phonologique est impliquée dans des tâches pouvant en principe être effectuées sans elle. En ce qui concerne l'influence de l'orthographe sur la réalisation de tâches purement phonologiques, certaines études (détection de rimes ou de phonèmes) montrent que le code orthographique y joue un rôle non-négligeable. Plusieurs modèles psycholinguistiques comportent ainsi des connexions directes et automatiques entre codes phonologiques et orthographiques, que celles-ci soient prélexicales ou lexicales.

Réconcilier ces deux perspectives implique, selon Kouider et Dupoux (p. 37) de soigneusement distinguer les deux étapes impliquées dans la plupart des tâches psycholinguistiques : celle du traitement automatique des stimuli et celle de la prise de décision relative à la tâche en question. Les auteurs se sont donc concentrés sur les relations orthographe/phonologie au niveau lexical, en distinguant deux types de conditions : traitement conscient et inconscient du stimulus visuel d'amorce. Cela devait leur permettre de savoir si le lexique phonologique d'entrée était obligatoirement et automatiquement activé lors de la présentation de mots écrits, ou si cette activation n'était effective qu'à travers des stratégies contrôlées par le système de contrôle central des sujets (« central executive ») (p. 37).

Les résultats qu'ils obtiennent à travers une tâche d'amorçage intermodal pour des sujets français (un mot oral précédé de son équivalent orthographique, de manière « masquée ») semblent montrer que :

- Pour une amorce intra-modale visuelle (visuel-visuel) : un effet d'amorce est observé, que l'amorce soit consciemment perçue ou non.
- Pour une amorce intermodale (visuel-oral) : l'effet d'amorce n'est observé que lorsque l'amorce est consciemment perçue.

La conclusion des auteurs est la suivante : « the systems underlying written and spoken word processing are, respectively, autonomous, and connected only under conscious conditions » (p. 35). Il n'y aurait donc pas d'activation automatique et obligatoire du lexique phonologique lors de la présentation d'un mot écrit lorsque le traitement est inconscient.

### III.5.1.2. Interactivité orthographe/phonologie : au niveau sublexical

Malgré les études précédemment mentionnées, de nombreux travaux psycholinguistiques ont conduit à souligner la forte connectivité existant entre les deux codes, activés tant en perception auditive que visuelle. Ainsi, Dijkstra, Frauenfelder et

Schreuder (1993) ont conduit certains tests de reconnaissance impliquant un partage attentionnel, afin de mettre en évidence les activations ou inhibitions mutuelles possibles entre graphèmes et phonèmes en tâche de détection bimodale. L'intérêt de la bimodalité ici est qu'elle permet d'étudier non seulement l'influence des phonèmes sur les graphèmes, comme dans les travaux consacrés à la reconnaissance visuelle lexicale, mais également l'inverse.

L'objectif de ces chercheurs était, à l'aide de ces techniques de présentation bimodale, d'étudier les relations entre le traitement de stimuli linguistiques présentés visuellement et celui de stimuli linguistiques présentés auditivement, en particulier au niveau sublexical (graphème/phonème). Comme ils l'indiquent dans leur introduction, peu d'études se sont penchées sur les activations de phonème à graphème, la majorité de la littérature étant consacrée aux activations de graphème à phonème, à un niveau lexical (mots ou non-mots), sans détailler les aspects temporels des processus d'activation intermodale. Dans leurs tests, certaines cibles étaient identiques (« congruent ») en bimodalité (par exemple A visuel et /a/ auditif), tandis que d'autres étaient différentes (« incongruent ») (par exemple U visuel et /a/ auditif). La dimension temporelle a également été prise en compte dans ces tests via un contrôle du délai de présentation des stimuli bimodaux en asynchronie.

D'un point de vue théorique, les auteurs se réfèrent à deux classes de modèles pouvant rendre compte du traitement de stimuli non-linguistiques en tâche de détection bimodale : les modèles à activations séparées, qui supposent deux processus de détection séparés pour chaque modalité mais parallèles et synchrones, et ceux à co-activation, qui supposent combinées les activations de chaque canal, l'activation augmentant graduellement dans le temps. Grâce à cette tâche de détection, trois aspects des relations graphèmes-phonèmes ont pu être abordés :

- L'existence d'effets de facilitation intermodale ;
- La taille et la directionnalité de ces effets ;
- L'existence d'effets d'inhibition intermodale.

Une activation entre graphèmes et phonèmes, dans les deux sens, a été mise à jour dans les trois expériences. D'un point de vue temporel, l'effet de facilitation a été obtenu dans les trois conditions suivantes : visuel – 100 ms – auditif ; auditif – 100 ms – visuel ; auditif et visuel simultanés. Les chercheurs considèrent d'après leurs résultats que ces effets sont bien de nature représentationnelle, sublexicale, et qu'ils sont automatiques. Cependant, si l'existence d'effets de facilitation intermodale semble clairement montrée, il n'en va pas de même pour les effets d'inhibition intermodale, alors qu'ils sont pourtant postulés en unimodalité dans certains modèles, tels que le modèle TRACE par exemple.

La conclusion de cette étude est donc la suivante : « the existence of automatic bidirectional activation and the absence of cross-modal inhibition between graphemes and phonemes » (p. 948). Cela soutiendrait l'hypothèse selon laquelle les graphèmes joueraient un rôle dans la reconnaissance auditive de mots et les phonèmes dans la reconnaissance visuelle de mots, ce qui irait dans le sens de certains modèles duaux de reconnaissance visuelle. Comme le notent les auteurs (p. 948) :

*« An interesting consequence of this viewpoint would be that the recognition of an auditory target word would be influenced not only by the set of similar words within the same modality (its auditory cohort), but also by the visual counterpart of that set of word candidates ».*

### III.5.1.3. Interactivité orthographe/phonologie : au niveau lexical

Certains travaux de Grainger et de ses collaborateurs soulignent également l'étroite imbrication des deux codes, orthographique et phonologique, à un niveau lexical. Leurs études sur les homophones-hétérographes (ex : /fwa/ → <fois>/<foie>) révèlent ainsi l'existence d'une interférence « phonologique » (mots homophones plus coûteux que mots non-homophones), comme il y a une interférence « orthographique » (liée aux voisins orthographiques de fréquence élevée). Pour Grainger, Nguyen Van Kang et

Ségui (2001), il semble donc aujourd'hui justifié de considérer que les codes orthographiques sont automatiquement activés lors de la présentation auditive d'un mot. Ainsi, d'après ces travaux, il semblerait :

- Qu'il y ait un effet inhibiteur du nombre de voisins phonologiques (densité du voisinage phonologique) en reconnaissance visuelle.
- Qu'il y ait un effet facilitateur du nombre de voisins orthographiques (densité du voisinage orthographique) en reconnaissance auditive.

Cela serait dû à une compétition lexicale entre mots phonologiquement similaires (« effet de voisinage phonologique »), alors que l'effet de voisinage orthographique reflèterait la régularité (« consistency ») des appariements sublexicaux entre phonologie et orthographe. Ces résultats suggèrent une fois de plus que l'information orthographique influence la reconnaissance lexicale auditive (Ziegler, Muneaux et Grainger, 2003).

L'étude de Chéreau, Hallé et Ségui (1999), conduite en français, semble aller dans le même sens, puisque les auteurs, à l'aide de stimuli dont les formes phonologiques (de surface) d'une part, et orthographiques-morphophonémiques ou seulement orthographiques d'autre part, étaient divergentes, montrent que le traitement de mots parlés est influencé par le code (abstrait) orthographique et/ou morphophonémique dans le premier cas, et, dans une moindre mesure, par le code orthographique dans le second cas. Un exemple de stimulus est le mot « absurde » dans lequel la lettre « b » est prononcée /p/, pour le premier cas (préfixe « ab- »), et le mot « asile », dans lequel la lettre « s » est prononcée /z/ pour le second cas. Dans chaque cas, la forme phonologique de surface et le script sont « conflictuels ». Leur conclusion est la suivante :

*« Orthography interacts with phonetic perception in spoken words [...] [and] both the visual and auditory processing of words engage orthographic and phonetic representations that interact in both directions ».*

Il est intéressant de noter que, si cette dernière étude souligne l'influence de l'orthographe sur le traitement des formes phonologiques « de surface » (en particulier phonémique), l'étude de Ventura, Kolinsky, Brito-Mendès et Morais (2001) sur des locuteurs portugais souligne, quant à elle, l'influence de l'orthographe sur les représentations syllabiques, représentations que l'on pourrait qualifier de « profondes ». Il semblerait donc que l'influence de l'orthographe sur le traitement de la parole puisse s'exercer à plusieurs niveaux de représentation phonologique.

#### III.5.1.4. Conclusion : de la perception à la production

En guise de conclusion de cette sous-partie sur les interactions entre codes phonologiques et orthographiques dans le cadre de la perception de l'écrit et de la parole, en particulier celles témoignant de l'influence des seconds sur le traitement des premiers, il semble intéressant d'évoquer une question plus rarement abordée dans la littérature, celle de l'influence possible de l'orthographe sur la *production* de la parole. Damian et Bowers (2003) se sont ainsi demandé si l'orthographe pouvait, dans une certaine mesure, contraindre la production de mots isolés. A l'aide de tests d'amorce pour des sujets anglais, les auteurs ont ainsi comparé le traitement de mots partageant le même segment phonologique initial (par exemple /k/), mais épelés différemment (par exemple « coffee »-« kennel »). Les résultats semblent soutenir leur hypothèse initiale<sup>135</sup> (p. 127) :

*« The study thus provides evidence for interactions between orthography and phonology in speech production, consistent with the few studies that have observed similar results in perception (and even fewer in production) ».*

Tout en admettant, au vu des données neurolinguistiques disponibles, une modularité fonctionnelle cérébrale pouvant être perturbée et conduire à des dissociations entre

---

<sup>135</sup> Bien qu'une étude ultérieure, basée sur celle de Damian et Bowers, ne soit pas parvenue à la même conclusion (Alario et Ziegler, 2004).

traitement de l'écrit et de l'oral, il semblerait néanmoins que, pour des sujets adultes à l'état « stable », les codes phonologiques et orthographiques soient étroitement connectés. Quelle que soit la modalité de présentation (auditive ou visuelle) des stimuli, les représentations activées (sublexicales ou lexicales) sembleraient donc relever des deux types de codes, à des degrés divers selon la modalité et la tâche. Les représentations orthographiques peuvent donc exercer une certaine influence sur la perception (mais également la production) de la parole, et, *a fortiori*, sur son apprentissage.

Le peu d'information disponible à ce sujet s'explique aisément : comme nous l'avons indiqué plus haut, tant d'un point de vue individuel que collectif, la parole est systématiquement première par rapport à l'écrit (du moins pour des sujets « non-déficients »). Il va alors de soi que, dans de telles conditions, l'on ne puisse évaluer à proprement parler l'influence de l'orthographe sur l'acquisition de la parole en langue maternelle. Le domaine de l'acquisition des langues secondes est dès lors à cet égard domaine privilégié, puisqu'il pourrait, dans une certaine mesure, offrir la possibilité d'évaluer celle-ci en langue étrangère. Dans le domaine de l'acquisition de la langue première cependant, malgré les limitations évoquées, il est une notion qui a particulièrement attiré l'attention des chercheurs s'intéressant à l'apprentissage de la lecture. Il s'agit de la notion de *conscience phonologique*. C'est précisément sur celle-ci que repose un deuxième argument majeur en faveur de l'hypothèse selon laquelle le traitement de la parole, et ainsi son apprentissage, peut être influencé par l'orthographe.

### **III.5.2. La conscience phonologique**

La notion de « conscience » semble, aujourd'hui encore, délicate à manipuler, traitée tant en sciences cognitives, qu'en sciences du langage ou en philosophie. D'un point de vue neurocognitif, Mesulam (1998, p. 1042) indique :

*« Human consciousness is not a special faculty occupying a specific site of the brain but an integrative manifestation of many [...] systems. [...] The accessibility of an event to explicit consciousness and introspective commentary*

*is the by-product of a special type of cortical activity and not an automatic consequence of sensory input ».*

Dans la perspective qui est la nôtre, la notion de « conscience (méta-)linguistique » renvoie en partie au débat articulé autour des couples déclaratif/procéduraux, explicite/implicite, automatique/contrôlé (Laks, 1996). Dans le domaine de la lecture en particulier, la notion de « conscience phonologique » a joué un rôle non négligeable dans l'étude des relations entre perception de la parole et perception de l'écrit.

### III.5.2.1. Conscience phonologique précoce vs. Conscience phonémique

Comme l'explique Bentin (1992), si les enfants sont capables dès leur plus jeune âge de faire preuve de discrimination phonétique et phonologique très fine, ils sont en revanche peu capables de réfléchir consciemment sur le système phonologique de leur langue ou d'en manipuler intentionnellement certains traits. C'est-à-dire que les jeunes enfants ne disposent pas d'une capacité métalinguistique de manipulation d'éléments sublexicaux, capacité nommée *conscience phonologique* (p. 193). Cependant, il existe différentes formes de conscience phonologique, celle-ci se révélant être une compétence métalinguistique hétérogène (p. 194). Les tests par lesquels cette compétence est évaluée diffèrent ainsi par la nature de la tâche (isolation de phonèmes, segmentation phonémique, comptage phonémique, etc.) mais aussi par la taille des unités auxquels ils renvoient (phonème, attaque syllabique, rime, syllabe, etc.). Bentin distingue ainsi trois dimensions caractéristiques de ces tests (p. 194) :

1. Opération requise (détection, isolation ou manipulation du segment phonologique).
2. Mode d'évaluation de la conscience phonologique (indirect ou explicite).
3. Taille du segment phonologique en question (syllabique, subsyllabique ou phonémique).

Au vu des performances des enfants dans différentes tâches, telles que rapportées dans la littérature concernée (bonnes performances pour les unités syllabiques et subsyllabiques, moins bonnes pour les unités phonémiques), Bentin suggère de



considérer une différence qualitative entre une conscience phonologique phonémique (singleton) et une autre multiphonémique. Pour d'autres auteurs il ne s'agit pas tant d'une différence qualitative que d'un continuum, en considérant la conscience phonologique comme une capacité en développement. Cependant, en se tenant à la caractérisation en trois dimensions des tests d'évaluation, Bentin conclue qu'il n'y a que deux formes de conscience phonologique (p. 198) :

1. *Conscience phonémique* (« phonemic awareness ») : nécessite une connaissance explicite du segment phonémique et est démontrée par la capacité d'isoler et de manipuler des phonèmes simples ;
2. *Conscience phonologique précoce* (« early phonological awareness ») : se signale indirectement dans le repérage de ressemblances et de dissemblances entre les mots sur la base d'unités subsyllabiques.

#### III.5.2.2. Apprentissage du code écrit et développement de la conscience phonologique

Une fois établie la nature de la conscience phonologique, il convient de s'interroger sur les facteurs qui influent sur son développement (p. 199). Deux points de vue existent à ce sujet : certains considèrent l'apprentissage de la lecture comme un facteur déterminant, tandis que d'autres estiment que la conscience phonologique se développe préalablement à l'apprentissage de la lecture. Bentin réconcilie ces deux perspectives en estimant qu'ils ne traitent pas de la même forme de conscience phonologique : en particulier les derniers se réfèrent à la conscience phonologique précoce (celle qui traite d'unités subsyllabiques et qui est testée de manière indirecte), tandis que les premiers se réfèrent à la conscience phonémique.

##### III.5.2.2.1. La conscience phonémique

Plusieurs études, en particulier certaines conduites auprès de sujets illettrés, montrent en effet que l'apprentissage de la lecture d'une orthographe alphabétique permet le développement de la conscience phonémique : contrairement à la parole, où les phonèmes subissent les effets de la coarticulation, l'écrit permet aux phonèmes d'être clairement définis par des segments orthographiques, lettres ou graphèmes. C'est

également ce que souligne Liberman (1992), dans la perspective de la théorie motrice de la parole qui est la sienne, puisqu'il insiste sur les contraintes exercées par la coarticulation dans le développement de la conscience phonologique (p. 175) :

*« Experience with speech is normally not sufficient to make one consciously aware of the phonological structure of its words. [...] Coarticulation [...] destroys any simple correspondences between the acoustic segments and the phonological segments they convey ».*

Trois arguments sont alors cités par Liberman (p. 176) :

- 1) « Preliterate and illiterate people [...] lack such phonological awareness ».
- 2) « The amount of awareness they do have predicts their success in learning to read ».
- 3) « Teaching phonological awareness makes success in reading more likely ».

Morais, Content, Cary, Mehler et Segui (1989), en référence à leurs travaux antérieurs, rappellent ainsi que l'apprentissage de la lecture et de l'écriture dans un système alphabétique entraîne le développement de la capacité à analyser la parole intentionnellement en unités phonémiques. Ce développement peut ainsi conduire à de nouvelles stratégies de traitement : certaines études sur les stratégies d'écoutes montrent ainsi que les sujets lettrés semblent prêter davantage attention aux phonèmes individuels, alors que les illettrés adoptent une stratégie plus globale (p. 58).

Il importe de souligner que la nature, alphabétique ou non-alphabétique, de l'orthographe ainsi apprise a son importance, comme le révèle l'étude de Him Cheung et Chen (2004) qui ont comparé les performances de deux types de locuteurs cantonnais : le premier groupe ne connaissait que le système « logographique » des kanji, tandis que le deuxième groupe maîtrisait tant les kanji que le système alphabétique Pinyin. Les résultats n'indiquent d'analyse de niveau phonémique que pour le deuxième groupe, rejoignant ainsi les conclusions précédentes : l'apprentissage d'une orthographe alphabétique influence fortement la constitution de la conscience phonologique (phonémique), ainsi que celles des représentations utilisées dans la

communication orale quotidienne. Certaines études montrent ainsi que les lecteurs à orthographe non-alphabétiques, comme ceux de Hong Kong, ont une conscience phonologique (phonémique) réduite pouvant entraîner des difficultés à traiter des non-mots ou des mots nouveaux (Holm et Dodd, 1996).

#### III.5.2.2.2. La conscience phonologique non-phonémique

Plusieurs travaux montrent cependant que la conscience phonologique précoce, ou du moins la conscience phonologique non-phonémique (attaque, rime, syllabe), peut se développer sans instruction explicite et indépendamment de l'apprentissage de la lecture. Morais, Content, Cary, Mehler et Segui (1989), par exemple, ont évalué les performances de sujets portugais adultes n'ayant pas été scolarisés dans leur enfance, dans une tâche de détection syllabique auditive (CV ou CVC). Les sujets étaient divisés en deux groupes : 20 étaient totalement illétrés ; 20 suivaient des cours d'apprentissage de la lecture pour adulte (ex-illettrés). Le point de départ de cette étude était celle menée précédemment par Mehler et ses collègues auprès de sujets français lettrés, mettant à jour un important « effet syllabique » dans une telle tâche : la cible était détectée plus rapidement lorsque celle-ci correspondait à la syllabe initiale du mot porteur (par exemple la cible « pa » dans le mot « palace », plutôt que dans le mot « palmier », et la cible « pal » dans le mot « palmier », plutôt que dans le mot « palace »). Cet effet n'apparaissait cependant pas pour des sujets anglais, une observation notamment liée au fait que l'anglais présente davantage de consonnes ambisyllabiques (comme le /l/ dans « palace ») que le français. L'hypothèse de Morais et de ses collègues était que si le développement de la procédure de syllabification ne dépendait que des caractéristiques (phonologiques) de la langue des sujets (et absolument pas de l'alphabétisation), alors les réponses des sujets illettrés devaient présenter un effet syllabique comparable à celui des ex-illettrés.

Si les résultats indiquent que les illettrés ont moins bien réussi cette tâche que l'autre groupe, un effet syllabique important a cependant été mis à jour pour les deux groupes : tant pour les illettrés que pour les ex-illettrés, les résultats étaient meilleurs

lorsque la cible (CV ou CVC) correspondait à la syllabe initiale (CV ou CVC) du mot porteur. Cet apparent « effet syllabique », comparable à celui obtenu en français pour des sujets lettrés, semble donc suggérer que la procédure de syllabification en traitement de la parole serait indépendante de toute instruction formelle. Ainsi, tandis que certaines unités ou procédures sembleraient naturellement acquises, d'autres dépendraient davantage de la présence ou l'absence d'une instruction formelle. Il faut d'ailleurs noter que l'apprentissage de la lecture ne constitue pas l'unique mode d'instruction de compétence métaphonologique : Bentin cite par exemple une méthode d'utilisation d'indices visuels de représentation des phonèmes (p. 202).

### III.5.2.2.3. Conclusion

Comme le souligne Mattingly (1992, p. 14), il convient de distinguer l'accès (« access ») de la conscience (« awareness »), ce qui nous renvoie à la discussion menée par Laks (1996) sur la connaissance procédurale et déclarative. Or la conscience linguistique doit en grand partie être acquise, une acquisition essentiellement stimulée par l'apprentissage de la lecture et de l'écriture. « Once it has become established, the writing system itself shapes the linguistic awareness, and even the phonology » (p. 25).

Les récents travaux neuropsycholinguistiques de Castro-Caldas, Petersson, Reis, Stone-Elander et Ingvar (1998) apportent de nouveaux résultats soutenant un tel point de vue. Ces derniers ont eu recours à l'imagerie cérébrale fonctionnelle (PET) afin de comparer les performances de sujets lettrés et illettrés en tâche de répétition de mots et de pseudo-mots. Pour les mots, les performances des deux groupes et les régions cérébrales activées étaient similaires. En revanche, pour les non-mots, les sujets illettrés ont eu davantage de difficultés, et les structures cérébrales activées étaient différentes de celles des lettrés : « Only the literate group had a trained system for phonological attention/awareness driving the organization and production of motor sequences not previously learned » (p. 1060). D'après ces résultats, il semblerait (p. 1054):

- 1) Que l'apprentissage du code écrit interagisse avec le traitement du langage oral :  
« The emergence of the necessary phonological processing capacity for correct pseudoword repetition is sculptured by the learning and knowledge of reading and writing ».
  
- 1) Que cet apprentissage influence l'organisation fonctionnelle du cerveau humain adulte : « learning to read and write modifies the phonological system by adding a visuographic dimension, i.e. that of graphemic-phonemic matching ».

La conclusion de Bentin (1992) est donc claire : le lien entre apprentissage de la lecture et conscience phonologique (phonémique) est indiscutable. Déchiffrer une orthographe alphabétique nécessite d'être capable d'isoler et de manipuler des phonèmes simples dans la parole, marquée par la coarticulation. L'exposition à l'alphabet déclenche donc le développement de cette capacité. Cependant il semble que ce développement ne peut se faire qu'à condition que la conscience phonologique précoce soit elle-même suffisamment développée. (p. 206). Il est intéressant de noter à ce propos, comme le fait Gathercole (1998), que les mesures de la mémoire phonologique ne sont pas sensibles au milieu socioculturel de l'enfant, alors que celles de la conscience phonologique sont fortement influencées par le statut socio-économique.

Ainsi, si l'on ne peut nier l'existence d'un certain type de conscience phonologique relativement indépendant de tout apprentissage du code écrit, il semblerait que le lien entre conscience phonémique et apprentissage d'une orthographe alphabétique soit à présent indiscutable. Ce lien pourrait donc jouer un rôle important en tâche de segmentation de la parole d'une part, et d'identification phonémique d'autre part.

Dans le cas d'apprenants adultes de langue étrangère, ceux-ci possèdent généralement au moins un système phonético-phonologique et un système d'écriture, ceux de leur langue première. Quatre (ou six) grands systèmes « se rencontrent » alors au cours de leur apprentissage :

<b>Langue Première</b>	<b>Langue Etrangère</b>
Système phonético-phonologique	Système phonético-phonologique
Système phonographémique/ orthographique <sup>136</sup>	Système phonographémique/ orthographique

L'évaluation de l'influence du système orthographique de la langue première sur l'acquisition du système phonético-phonologique de la langue étrangère apparaît difficile à accomplir, étant donné la multiplicité des facteurs en jeu et la complexité des relations qui les unissent. Néanmoins, il est possible de retirer quelques implications générales des questions évoquées précédemment (unités perceptives et conscience phonologique) pour les apprenants japonais de français dans certaines tâches fondamentales d'apprentissage, celles-ci impliquant en premier lieu une ou plusieurs tâches perceptives. Deux niveaux linguistiques retiennent en particulier notre attention : le niveau segmental (phonémique) et le niveau syllabique. Nous pouvons ainsi, en nous concentrant sur ces deux niveaux, formuler nos problématiques d'apprentissage, et à leur suite, d'enseignement.

---

<sup>136</sup> La distinction entre les deux pouvant varier selon les langues et les contextes, nous utilisons tantôt l'un tantôt l'autre terme selon les circonstances et la pertinence du terme. Il va de soi que, dans un certain contexte et à un certain niveau d'analyse, ils constituent deux systèmes clairement distincts.



## **Problématiques :**

### **Apprentissage Phonologique Multimodal**

La rencontre entre les termes « phonologie », strictement linguistique, et « apprentissage », à orientation psycholinguistique, élargit considérablement le champ de notre problématique initiale, hors des frontières des descriptions ou des modélisations linguistiques. Cependant, le domaine de l'apprentissage de la phonologie des langues étrangères a vu de nombreux développements récents, traités dans un cadre aussi exclusivement linguistique que possible. Avant d'aborder la question sous un angle psycholinguistique, il convient donc de l'aborder, autant que se peut, dans une perspective strictement phonologique.

Cette quatrième partie se présente donc comme suit : dans un premier temps (I) nous présentons la problématique générale de l'interphonologie dans une perspective essentiellement phonologique : d'abord structuraliste-fonctionnaliste (I.2.), puis générative (I.3.) et finalement exemplariste (I.4.). Nous nous concentrons ensuite (II) sur le cas spécifique des apprenants japonais de français, vis-à-vis de certains aspects segmentaux (II.1.), puis syllabiques (II.2.), de l'apprentissage phonologique. Dans un deuxième temps (III), nous adoptons une perspective davantage psycholinguistique, en considérant le rôle potentiel des représentations orthographiques dans le développement de l'interphonologie. Après avoir présenté les variables générales qui y sont impliquées (III.1.) ainsi que les tâches didactiques et cognitives considérées (III.2.), nous examinons le rôle de l'orthographe à un niveau segmental (III.3.) et syllabique (III.4.), résumant enfin (III.5.) les problématiques psycholinguistiques que notre protocole quasi-expérimental vise à traiter.



## Table des Matières Partielle

I. PROBLEMATIQUE D' APPRENTISSAGE PHONOLOGIQUE : L' INTERPHONOLOGIE .....	- 486 -
<i>I. 1. Introduction : apprentissage phonétique ou phonologique ?</i> .....	- 486 -
<i>I. 2. Interphonologie : perspectives structuraliste et fonctionnaliste</i> .....	- 494 -
I.2.1. Le crible phonologique .....	- 494 -
I.2.2. Facteurs sociolinguistiques .....	- 495 -
I.2.3. Du crible au transfert .....	- 498 -
<i>I. 3. Interphonologie : perspectives générativistes</i> .....	- 499 -
I.3.1. Perspectives représentationnelles : Marque, Transfert et Grammaire Universelle .....	- 499 -
I.3.1.1. La Marque .....	- 500 -
I.3.1.2. Le Transfert et la Grammaire Universelle .....	- 503 -
I.3.1.2.1. Le Transfert L1/L2 .....	- 503 -
I.3.1.2.2. L' Accès à la Grammaire Universelle .....	- 505 -
I.3.2. Perspectives non-représentationnelles : optimalité et contraintes .....	- 507 -
I.3.2.1. Interphonologie et Théorie de l' Optimalité : une nouvelle approche du transfert .....	- 507 -
I.3.2.2. Exemple 1 : épenthèses et émergence du non-marqué chez des apprenants brésiliens d' anglais .....	- 509 -
I.3.2.3. Exemple 2 : épenthèses et variation chez des apprenants chinois d' anglais .....	- 512 -
I.3.2.4. Exemple 3 : reclassement de contraintes et émergence du non-marqué chez des apprenants zoulous d' anglais .....	- 513 -
I.3.2.5. Exemple 4 : coda et émergence du non-marqué chez des apprenants italiens d' anglais .....	- 515 -
I.3.2.6. Exemple 5 : substitutions segmentales chez des apprenants japonais d' anglais .....	- 516 -
I.3.2.7. Exemple 6 : groupes consonantiques et voyelles épenthétiques en interphonologie .....	- 520 -
I.3.3. Synthèse .....	- 525 -
I.3.3.1. Marque et Transfert dans la Théorie de l' Optimalité .....	- 525 -
I.3.3.2. Accès à UG, Transfert et Perception dans la Théorie de l' Optimalité .....	- 528 -
<i>I.4. Interphonologie : perspectives exemplaristes</i> .....	- 530 -
I.4.1. Des épisodes phonétiques aux catégories phonologiques .....	- 531 -
I.4.1.1. Encodage épisodique acoustique de la parole .....	- 531 -
I.4.1.2. Variation allophonique et acquisition des catégories phonémiques .....	- 535 -
I.4.2. Mémoire phonologique et émergentisme .....	- 538 -
II. INVESTIGATIONS PHONOLOGIQUES : L' INTERPHONOLOGIE DES APPRENANTS JAPONAIS DE FLE ..	- 540 -
<i>II.1. Aspects segmentaux</i> .....	- 541 -
II.1.1. Axe syntagmatique : la position du segment cible dans le mot phonologique influence-t-elle le taux de confusion segmentale ? .....	- 543 -
II.1.2. Axe paradigmatique : la nature du segment cible dans le mot phonologique influence-t-elle le taux de confusion segmentale ? .....	- 545 -
II.1.2.1. Segments consonantiques .....	- 545 -
II.1.2.2. Segment vocalique .....	- 547 -
II.1.2.3. Confusion phonétique ou phonologique ? .....	- 550 -
<i>II.2. Aspects syllabiques</i> .....	- 553 -
II.2.1. Axe syntagmatique : la position du groupe cible dans le mot phonologique influence-t-elle le taux de confusion syllabique ? .....	- 556 -
II.2.2. Axe paradigmatique : la nature du groupe cible dans le mot phonologique influence-t-elle le taux de confusion syllabique ? .....	- 557 -
II.2.3. Axe paradigmatique : la structure de surface du groupe cible dans le mot phonologique influence-t-elle le taux de confusion syllabique ? .....	- 568 -

II.2.4. Résumé.....	- 569 -
II.2.5. Aspects phonographémiques .....	- 571 -
<b>III. INVESTIGATION PSYCHOLINGUISTIQUE : ORTHOGRAPHE ET INTERPHONOLOGIE CHEZ LES APPRENANTS JAPONAIS DE FLE .....</b>	<b>- 576 -</b>
<i>III.1. Introduction : variables générales, tâches d'apprentissage et niveaux d'analyse.....</i>	<i>- 576 -</i>
<i>III.2. Nature des tâches cognitives.....</i>	<i>- 584 -</i>
III.2.1. Tâche de Lecture – unimodalité visuelle.....	- 584 -
III.2.2. Tâche de Répétition – unimodalité auditive.....	- 594 -
III.2.3. Tâche de Lecture/Répétition – bimodalité audiovisuelle .....	- 601 -
III.2.4. Tâche de Dictée Orthographique .....	- 604 -
<i>III.3. Influence de l'orthographe : niveau segmental et identification .....</i>	<i>- 605 -</i>
III.3.1. Traitement psycholinguistique « erroné » des unités phonético-phonologiques et orthographiques.....	- 606 -
III.3.2. Variation et alternance inter-tâches et intra-tâche .....	- 612 -
III.3.3. Schémas d'activation des unités phonologiques et orthographiques.....	- 616 -
III.3.4. Protocole quasi-expérimental.....	- 620 -
III.3.5. Substitutions inattendues et stratégies d'apprentissage .....	- 622 -
<i>III.4. Influence de l'orthographe : niveau syllabique et segmentation .....</i>	<i>- 624 -</i>
III.4.1. L'orthographe comme outil de segmentation de la parole .....	- 624 -
III.4.2. Le cas du japonais : contraintes phonologiques et/ou orthographiques ? .....	- 626 -
III.4.3. Protocole quasi-expérimental.....	- 634 -
<i>III.5. Résumé.....</i>	<i>- 636 -</i>

## **I. Problématique d'apprentissage phonologique :**

### **l'interphonologie**

Le concept d'interphonologie n'est pas récent. Même si le terme n'apparaît pas dans le travail de Broselow en 1983, qui traite plutôt de « second language phonology », ni dans celui de Tarone en 1987 qui parle de « interlanguage phonology », c'est pourtant bien d'interphonologie qu'il s'agit. L'interphonologie renvoie au pendant phonologique de l'interlangue, un système phonologique provisoire et instable, qui peut néanmoins à un moment donné, être décrit et caractérisé, en dépit d'une variation inter- et intra-individuelle inévitable (Broselow, 1983 ; Tarone, 1987 ; Lombardi, 2000 ; Grijzenhout et Van Rooy, 2001 ; Monahan, 2001 ; Keys, 2002 ; Lin, 2003 ; Escudero et Boersma, 2004). L'apprentissage d'un système phonologique étranger s'effectue donc de manière graduelle, à travers les développements de ce système « interphonologique », dont l'étude, notamment revitalisée par la Théorie de l'Optimalité, occupe aujourd'hui de nombreux chercheurs, comme en témoigne le récent article de Eckman (2004), qui y propose un « état des lieux » (« state of the art ») de la question.

### **I. 1. Introduction : apprentissage phonétique ou phonologique ?**

Lorsque l'on se penche, dans le domaine psycholinguistique, sur un modèle de production, ou bien encore de compréhension linguistique, la partie inférieure du modèle est généralement dédiée à la dimension phonétique des processus (encodage ou dé-/recodage). On ne peut donc se pencher sur les problèmes dits de « prononciation » ou bien encore plus généralement sur les problèmes de traitement de la parole au cours de l'apprentissage d'une langue étrangère, sans se pencher sur l'interface entre niveaux phonétique et phonologique, puisque si certains problèmes relèvent de l'un ou de l'autre niveau, d'autres semblent difficiles à assigner à l'un ou à l'autre exclusivement. Comme l'indiquent Leather et James (1991, p. 331) à propos de l'acquisition de la parole en langue étrangère : « rigid distinctions between (psycho-)linguistic and

phonetic levels of investigation have come to seem unfruitful if not actually untenable ».

En didactique des langues, et plus particulièrement du FLE, l'enseignement ou la correction de la « prononciation », en tant que composante de « l'oral », se trouve généralement identifié à un enseignement « phonétique », voire de « phonétique corrective », et le domaine phonologique se trouve ainsi, sinon négligé, du moins peu explicitement abordé, laissant presque croire que les difficultés des apprenants seraient d'ordre strictement phonétique, et non phonologique, même si les aspects syllabiques et phonotactiques, par exemple, de l'apprentissage, lorsqu'ils sont évoqués, soulignent l'impossibilité, toute relative, de se limiter à un niveau phonétique. L'apprentissage de la phonologie semblerait donc négligé, dissimulé sous ses aspects phonétiques.

Dans le domaine de la pathologie du langage en revanche, la distinction entre niveaux phonétique et phonémique s'est toujours révélée fructueuse dans l'analyse des « erreurs », de nature pathologique, produites par des patients aphasiques par exemple (Valdois, 1990 ; Valdois et Nespoulous, 1994). Ainsi, chez les adultes, on distingue par exemple des cas de (Nespoulous, 2002) :

- Dysarthrie<sup>137</sup> : troubles phonétiques, mais pas phonémiques ;
- Anarthrie pure<sup>138</sup> : troubles phonétiques, mais pas phonémique ;
- Aphasie de conduction ou de Wernicke<sup>139</sup> : troubles phonémiques, mais pas phonétiques.

Il est vrai que, si l'on se place dans la dichotomie saussurienne Langue/Parole, et que l'on souhaite opposer le « système » au « discours », on peut en effet vouloir

---

<sup>137</sup> Problèmes d'exécution des gestes moteurs de la parole, observables, par exemple, dans le contexte de la Maladie de Parkinson.

<sup>138</sup> Problèmes de planification de la parole, en l'absence d'autres perturbation linguistiques, consécutifs à une lésion cérébrale focale cortico-sous-corticale.

<sup>139</sup> Perturbations des processus de sélection et/ou de combinaison des phonèmes de la langue en l'absence de tout trouble de la motricité de la parole.

s'interroger davantage sur les réalisations phonétiques des apprenants, que sur leur compétence phonologique. Néanmoins, il nous semble nécessaire de nous demander si, en fonction de l'approche adoptée (segmentale, syllabique ou métrique par exemple), les problèmes rencontrés relèvent davantage de l'un ou l'autre pôle.

Dans une perspective segmentale, par exemple, si l'on considère l'unité abstraite U, associée par exemple à l'une des voyelles dites cardinales, il semblerait facile de lui assigner une valeur phonémique tant en français, qu'en japonais. Pourquoi, alors, de nombreux étudiants donnent-ils l'impression de « mal prononcer » le U, comme nos étudiants qui fréquemment prononcent le mot « écouter » comme [ekøte], [ekte] ou bien [ekɥte]<sup>140</sup> ? On pourrait alors penser que la raison se situe simplement au niveau phonétique, niveau qui renvoie au concept d'allophonie, lequel à son tour renvoie au problème de la variation phonétique. Ainsi, les symboles de l'A.P.I. pour le U français et japonais sont différents : [u] français et [ɯ] japonais. Pourtant, lorsque l'on considère les unités au sein de leur système (vocalique dans le cas du U par exemple), on peut s'interroger sur la nature purement phonétique ou phonético-phonologique du problème. D'un point de vue phonologique, on pourrait considérer que la raison ne provient pas tant d'une déficience propre à l'unité phonémique qu'au système d'unités distinctives auquel elle appartient. Keating et Huffman (1984), en présentant le système de variation des voyelles japonaises, nous indiquent combien différent les zones de dispersion des voyelles japonaises et françaises, avec des zones de chevauchement rendant bien compte de certaines confusions commises par les apprenants. Cette version acoustique renvoie évidemment à son pendant articutoire (Lieberman et Blumstein, 1988), comme nous l'indiquent les tenants de la théorie motrice de la parole (Lieberman et Mattingly, 1985). Résoudre la question nous demande ainsi de nous pencher sur les modèles de production et de perception de la parole, puisque, si d'un point de vue phonétique le [ɯ] japonais présente une zone de dispersion pouvant recouvrir celles du [y], du [u] et du [ø] français, l'objectif, en termes essentiellement

---

<sup>140</sup> Nous négligeons ici les questions suprasegmentales, notamment accentuelles.

psychoacoustiques, pour les apprenants, serait d'établir de nouvelles frontières intra-catégorielles (Harnad, 1987a ; Schwartz, Abry, Boë et Cathiard, 2002<sup>141</sup>). Nous le constatons fréquemment dans nos classes, lorsque, face à une tâche de répétition des trois phonèmes français mentionnés, contextualisés ou non, plusieurs de nos étudiants nous indiquent ne pas « entendre de différences » entre les trois stimuli. Cela exige donc aussi de ne pas oublier que les réalisations phonétiques de tel ou tel phonème ne se font précisément pas sans phonème, et donc sans une certaine valeur phonologique, plus abstraite mais pourtant bien effective.

Cet aspect de l'apprentissage en pose un autre : selon quels critères juger l'acceptabilité des productions allophoniques des apprenants ? Comment décider si un allophone de type [u] est un allophone du /ɯ/ japonais ou du /u/ français ? Comme le notent Flege et Hillenbrand (1987, p. 199, note 1):

*« The accuracy with which a learner produces the sounds of a foreign language can be objectively assessed in a variety of ways:*

- (1) through the use of rating scale judgements by native speakers of the target language,*
- (2) by calculating the frequency with which L2 phones are correctly identified, and*
- (3) through acoustic analyses ».*

Le choix de l'évaluation semble donc renvoyer, en partie au moins, aux objectifs et aux cibles de l'apprentissage : s'agit-il d'abord d'établir un nouveau système d'opposition phonémique (entre /y/ et /u/ et /ø/ par exemple), ou bien de privilégier l'identité phonétique de chaque unité du nouveau système ? Cette question renvoie ici encore à l'articulation entre traits phonologiques et phonétiques, les premiers étant généralement traités de manière binaire {0, 1}, tandis que les seconds sont généralement décrits à l'aide d'une échelle scalaire, par exemple [0, 1] (Durand, 1990, p. 40). Mais comme le

---

<sup>141</sup> Voir notamment pp. 260-261 à propos du recouvrement articulatoire-auditif pour le [u].

rappellent Boë et Schwartz (1997), la relation entre substance (phonétique) et forme (phonologique) est une relation complexe, étant donné « la non-biunivocité entre codes symboliques et acoustiques » (pp. 52-53) : « les *frontières* de segments n'existent ni au niveau articulatoire, ni au niveau acoustique, ni au niveau perceptif ». Odlin (1989, p. 113) souligne ainsi la nécessité d'intégrer les deux dimensions, phonétique et phonémique, à l'apprentissage, puisque, comme il l'écrit « two languages frequently have sounds which may seem identical but which in fact are acoustically different », ce qui, dans notre cas, nous renvoie aux contrastes entre /u/ français et japonais. Cela est illustré par l'étude de Flege et Hillenbrand (1987), qui s'intéressent précisément à la question de l'apprentissage phonétique, à partir d'un exemple donné par Valdman, qu'ils citent (p. 177):

*« French and English /s/ differ with regard to place of articulation. The former is a dental, and the latter is an alveolar. The partial similarities he perceives in the acoustic signal of French /s/ will lead an English speaker to respond with the alveolar rather than the dental sound ».*

Leur étude visait à déterminer si les apprenants de L2 produisaient de nouveaux « phones » avec davantage de « précision » (« accuracy ») que les « phones » de L2 qui avaient une contrepartie en L1 (p. 177). Les résultats semblent soutenir cette hypothèse, par exemple dans le cas de sujets américains apprenant le français: « The most important finding regarding vowel production was that the American talkers matched the French talkers in producing /y/ but not /u/ » (p. 195). Il semble par ailleurs que l'expérience permette aux apprenants adultes de faire preuve de davantage de précision en production (p. 197), même si la notion de « précision » (« accuracy ») reste toujours difficile à définir et évaluer. Les résultats présentés soulignent donc que les apprenants adultes ne recourent pas toujours aux phones de leur langue première pour prononcer les mots étrangers.

Ainsi, à propos des erreurs commises par les apprenants, Odlin (1989, p. 116) cite la taxonomie de Moulton, qui distingue des erreurs :

1. Phonémiques : Inventaires phonémiques différents.
2. Phonétiques : Equivalence phonémique, mais non phonétique.
3. Allophoniques : Equivalence phonémique et phonétique, mais non allophonique, c'est-à-dire pas pour tous les allophones.
4. Distributionnelles : Equivalence d'unités mais pas de distribution.

Wells (2000), qui donne plusieurs exemples détaillés « d'interférences » *phonétiques* entre le japonais et l'anglais, distingue en outre deux types d'erreurs allophoniques (p. 121) :

- a) Celles qui résultent de l'incapacité à acquérir des règles allophoniques appropriées en L2 et non en L1 ;
- b) Celles qui résultent du « transfert » en L2 de règles allophoniques pertinentes seulement en L1.

Il convient donc de bien identifier les pôles dont relèvent les « erreurs » des apprenants :

Perception		Production	
Phonologique	Phonétique	Phonologique	Phonétique

On peut par exemple considérer que certains apprenants ont achevé leur apprentissage phonologique, mais pas phonétique (par exemple lorsqu'ils sont capables de discrimination phonologique en perception mais pas en production (phonétique)), tandis que d'autres semblent avoir acquis les compétences phonétiques requises, alors que, sur le plan phonologique, l'apprentissage reste encore à accomplir : par exemple lorsque certains apprenants prononcent [pti] (exemple de groupe hétérosyllabique) ou [klavje] (exemple de groupe tautosyllabique), donnant l'impression de maîtriser la syllabation du français, mais comptent respectivement deux/ trois syllabes au lieu d'une/ de deux en tâche de comptage syllabique. La structure phonétique [CCV]



correspond en fait pour eux à une structure phonologique /CVCV/, dissimulée par la rapidité du débit et d'autres phénomènes phonétiques. A ce propos, Blevins (1995, 232) s'interroge sur les discordances (« mismatches ») entre représentations syllabiques phonologiques et phonétiques, en particulier en ce qui concerne les profils de sonorité : en effet, les réalisations phonétiques sont typiquement sous-produites (« undershoot ») conduisant à des décalages entre pics de sonorité phonologiques et phonétiques. Blevins cite le cas de langues dans lesquelles des voyelles réduites non accentuées au niveau phonologique sont effacées, soit de manière facultative, soit lorsque celles-ci se trouvent entre deux consonnes adjacentes identiques, en production langagière rapide. Dans ce cas, un pic de sonorité présent au niveau phonologique est absent au niveau phonétique : la représentation phonologique est inchangée et la perte de la voyelle ne résulte que de la sous-production (« undershoot ») de la cible phonétique. Le japonais fait notamment partie de ces langues, tandis que certaines langues présentent aussi la situation inverse (1995, p. 233).

Dans le domaine segmental, la question de l'articulation entre apprentissage phonétique et phonologique peut également être illustrée par l'étude de Colantoni et Steele (2005), qui, à propos de l'apprentissage du phonème /R/ par des apprenants anglophones en fonction du contexte, parviennent à la conclusion suivante, formulée ainsi dans leur résumé :

*« L2 learners' acquisition of French /r/ does not proceed in an all-or-none manner. Rather, learners first master frication. Subsequently, voicing appears gradually, first in those contexts where it is articulatorily favoured, i.e. intervocalically ».*

Clairement, comme l'écrit Durand (1990, p. 41), « the issue of the mapping between phonology and phonetics is an extremely controversial one ». Celle-ci a, en outre, été ravivée ces dernières années par une certaine réinsertion du rôle de la perception de la parole dans les problématiques phonologiques, mêmes si les perspectives ne sont pas unanimes sur le sujet. Certains chercheurs se préoccupent en effet du rôle de la

perception de la parole dans la constitution des systèmes phonologiques, en évoquant notamment le rôle de contraintes de distinctivités dans la formation des contrastes phonologiques, par exemple dans le cas de la dispersion vocalique ou encore des neutralisations contextuelles (Flemming, 2005). Cependant, alors que certains considèrent le rôle de la perception de la parole (en termes d'audition et de reconnaissance auditive) en phonologie comme un facteur extra-linguistique, qui ne constituerait que l'un des facteurs externes influant sur la constitution et les changements des systèmes phonologiques (ce que Hume et Johnson (2001) nomment « forces externes » (« external forces ») et qui comprend, outre la perception, la production linguistique (en termes de coordination et d'aérodynamique), la généralisation (en termes de catégories cognitives) et la conformité (en termes de communication sociale)), d'autres (Escudero, 2004) estiment que la perception des sons fait entièrement partie de la connaissance linguistique de la structure sonore d'une langue donnée. Contrairement à certains modèles « universalistes » qui distinguent des niveaux acoustique, phonétique (avec des catégories phonétiques universelles) et phonologique, Escudero soutient, sur la base de données empiriques, que la perception de la parole est pleinement dépendante de la langue des locuteurs, et ainsi à intégrer à leur compétence linguistique (2004, pp. 3-4) :

*« The mapping from the signal to discrete categories may not be universal and automatic [...], but [...] may depend on the way in which listeners have learned to perceive the speech sounds of their particular language environment. [...] Speech perception research has shown that the way in which listeners integrate acoustic information depends on their particular language environment because the attention paid to acoustic cues that signal sound contrasts varies in adult speakers of different languages ».*

A l'aide du modèle de grammaire perceptive de Boersma, formalisé dans le cadre de la Théorie de l'Optimalité (Escudero et Boersma, 2004), Escudero considère donc que la perception de la parole dépend systématiquement des propriétés acoustico-phonétiques

des langues particulières, ce qui la conduit à l'inclure dans la compétence phonologique, procédurale et représentationnelle, des locuteurs.

Il apparaît ainsi que la conception phonétique ou phonologique de l'apprentissage repose également fondamentalement sur le positionnement théorique plus général que l'on adopte face, d'une part à la perception de la parole (notamment, approche plutôt articulatoire *vs.* plutôt acoustique), et d'autre part à la nature des représentations cognitives (approche symbolique *vs.* approche émergentiste). Ce débat peut également trouver un écho en phonologie des emprunts *via* l'opposition entre approche « perceptive/phonétique » (Peperkamp et Dupoux, 2003) et approche « phonologique » (Smith, 2005) de l'adaptation des mots étrangers.

Ainsi, si la question ne se résout pas en quelques paragraphes, nous devons cependant effectuer certains choix. En particulier, il importe de signaler que nous tenons à placer au cœur de notre étude la question de l'apprentissage phonologique, essentiel lorsque l'on aborde la dimension syllabique et dépassant ainsi le simple thème de la « prononciation », tout en étant inévitablement reliée à celle d'un apprentissage phonétique, antérieur, ultérieur ou parallèle, stable et systématique.

## **I. 2. Interphonologie : perspectives structuraliste et fonctionnaliste**

Plusieurs notions et facteurs doivent ici être mentionnés.

### **I.2.1. Le crible phonologique**

Dans un cadre structuraliste ou fonctionnaliste (Blanchet, 2002, p. 48), l'apprentissage de la phonologie d'une langue a longtemps été, et est toujours, associé au concept de *crible phonologique*, élaboré par E. Polivanov puis développé par N.S. Troubetzkoy. Ce concept a été repris jusqu'à nos jours par la littérature qui s'intéresse à la phonologie comparée des langues, et il peut être envisagé à travers quelques lignes, maintes fois citées, écrites par Troubetzkoy (1976, 54) :

*« Le système phonologique d'une langue est semblable à un crible à travers lequel passe tout ce qui est dit. Seules restent dans le crible les marques phoniques pertinentes pour individualiser les phonèmes. Tout le reste tombe dans un autre crible où restent les marques phoniques ayant une valeur d'appel ; plus bas se trouve encore un crible où sont triés les traits phoniques caractérisant l'expression du sujet parlant ».*

Dans une perspective développementale, la mise en place du crible phonologique (Billières, 1988) chez l'enfant correspondrait à une période de réduction fonctionnelle autour du 12<sup>ème</sup> mois, après la période du babil, durant laquelle toutes les possibilités phonatoires sont explorées (Ségui, 2003, p. 216). Dès lors, les sons de la parole s'inscrivent dans un continuum sonore que l'apprenant entend, découpe (plus ou moins consciemment et correctement) et interprète à travers son/ ses crible(s) phonologique(s) et/ou phonétique(s)<sup>142</sup> (Billières, 1988, p. 7):

*« Les sonorités de la langue [étrangère] étudiée sont [...] perçues sur la base du système de référence phonologique constitué par la langue maternelle. Certaines sont donc mal entendues, mal interprétées et par conséquent mal reproduites ».*

Le concept de *crible phonologique* a ainsi servi, et sert toujours, de support à la majeure partie des méthodes actuelles dites de « correction phonétique », et en premier lieu la méthode Verbo-Tonale d'intégration phonétique (Renard, 1979).

### **I.2.2. Facteurs sociolinguistiques**

Blanchet (2002) note cependant que, dans le cadre fonctionnaliste fondé par Martinet, au cœur duquel se trouve la notion de « double articulation », les phonèmes se présentent comme des « unité[s] immatérielle[s] qui recouvre[nt] diverses réalisations phonétiques parfois matériellement très différentes, mais qui partagent la même

---

<sup>142</sup> Pour une discussion sur le ou les crible(s) phonologique(s) et/ou phonétique(s), voir Borrell, 1991.

fonction distinctive » (Blanchet, 2002, p. 38). La variété et l'hétérogénéité y sont donc privilégiées comme cela apparaît dans les propos de Builles, cité par Blanchet (2002, p. 45) conduisant à une approche essentiellement sociolinguistique : « Il existe plusieurs systèmes phonologiques au sein d'une même langue ». En ce qui concerne l'enseignement de la prononciation, le fonctionnalisme semble donc rejeter l'approche de « perfectionnisme glottocentré » (Blanchet, 2002, p. 46) qui impose aux apprenants de se conformer « à la perfection » à la prononciation d'un locuteur monolingue natif idéal. Beebe (1987, p. 173) insiste également, d'une manière quelque peu différente, sur ce point :

*« The fourth and final myth about interlanguage phonology [...] concerns sociolinguistic variations. There seems to be a tacit assumption that interlanguage is not subject to sociolinguistic variation [...]. This popular assumption persists despite explicit evidence to the contrary »<sup>143</sup>.*

Dans une perspective fonctionnelle, qui s'intéresse aux effets communicatifs des différents types de simplifications interlangagières, d'autres facteurs, notamment issus de la sociolinguistique, ont ainsi été pris en considération dans l'analyse des productions des apprenants. Lin (2003), par exemple, s'est intéressée à l'influence de quatre facteurs sociolinguistiques (niveau de langue, sexe du locuteur, degré

---

<sup>143</sup> Nous pourrions l'illustrer de manière anecdotique en mentionnant le cas de ces deux étudiantes japonaises qui, dans une classe de correction phonétique de niveau intermédiaire, ont surpris leur enseignant, précisément en raison de cette variation sociolinguistique, et plus précisément, dans leur cas, géolinguistique, variation à laquelle celui-ci ne s'attendait pas. Les faits peuvent être résumés de la manière suivante : en tâche de répétition d'énoncés, une étudiante persistait à produire une intonation globalement montante alors que l'enseignant attendait une intonation descendante. Ne souhaitant pas trop insister, afin d'éviter tout sentiment d'embarras, l'enseignant « passa » alors à l'apprenante suivante et fut surpris de constater la même erreur, alors qu'aucun des autres étudiants n'avaient présenté cette « déficience », jugée par ailleurs surprenamment « grossière » par l'enseignant. Au vu de la surprise de l'enseignant, la deuxième étudiante s'était alors exclamée, de manière humoristique et en souriant, de manière à dissimuler son embarras, qu'elles étaient toutes deux originaires de la région du Kansai (et non du Kanto, région de Tokyo où se tenait la classe), la première ajoutant « de Osaka » (l'« accent » ou plutôt le dialecte d'Osaka étant bien connu à travers le Japon, notamment pour ses variations intonatives), ce qui a provoqué l'hilarité dans la salle. Les praticiens doivent donc garder à l'esprit le facteur sociolinguistique dans le travail de correction phonétique, afin d'y adapter leurs attentes.

d'expertise linguistique du locuteur et type d'interlocuteur) sur la production de groupes consonantiques en position initiale par des sujets chinois apprenants l'anglais. Dans une perspective fonctionnelle, ce type de productions conduit à deux catégories « d'erreurs », qui doivent être distinguées :

- 1) Les *épenthèses*, qui sont considérées comme les productions les « moins erronées », préservant les formes sous-jacentes, assurant une meilleure intelligibilité et ainsi une bonne communication ;
- 2) Les *effacements*, qui sont considérés comme « destructifs » et mettant en péril la communication.

D'un point de vue méthodologique, si l'on peut apprécier la quantité de la population testée (40 sujets, divisés en deux groupes), on doit cependant souligner d'un regard critique le recours au support orthographique pour trois des tâches principales : lecture de paires minimales, lecture de liste de mot, lecture de phrases. L'apparente négligence de l'influence du facteur orthographique dans les études en interphonologie constitue en effet un point de discussion sur lequel nous reviendrons plus bas. Tarone notait déjà (1987, p. 73) à propos d'une étude conduite par Nemser en 1971 que les résultats obtenus jusqu'alors dans ce domaine étaient sévèrement réduits par les limites méthodologiques des approches adoptées, puisque les résultats dépendaient sensiblement des tâches expérimentales (transcription de syllabes sans signification, traduction de L1 en L2, etc.) :

*« The method used to gather data on interlanguage phonology influenced the nature of the data gathered ».*

Si cela semble inévitable, d'un point de vue psycholinguistique en particulier, il convient alors précisément de ne pas ignorer ces facteurs dans l'analyse qui est faite des données.

Les résultats de son étude conduisent Lin à insister sur la prise en compte du *type* d'erreurs commises (épenthèses vs effacements), plutôt que celle du simple *taux* d'erreurs. En effet, seule une analyse basée sur le *type* d'erreurs semble permettre de mettre pleinement à jour l'influence des facteurs à l'étude. Les hypothèses initiales de Lin, basées sur les études déjà existantes dans le domaine, seraient alors confirmées :

- 1) Différence de niveau de langue : le rapport épenthèses/effacements augmente avec le niveau de langue (plus la langue est formelle, plus le rapport augmente) ;
- 2) Différence d'expertise linguistique : le rapport épenthèses/effacements augmente avec le degré d'expertise des locuteurs (plus les apprenants sont avancés, plus le rapport augmente) ;
- 3) Différence de sexe : le rapport épenthèses/effacements est plus élevé pour les femmes que pour les hommes ;
- 4) Différence liée à l'interlocuteur : chez les femmes, le rapport épenthèses/effacements est plus élevé lorsque l'interlocuteur est un(e) co-apprenant(e) que lorsque l'interlocuteur est un(e) professeur(e) natif (-ve).

### **I.2.3. Du crible au transfert**

Dans le domaine de l'apprentissage des langues, le concept de *crible phonologique* a progressivement été complété par d'autres, plus généraux, notamment ceux de *transfert* (Selinker, 1983, Odlin, 1989) et d'*interférence*. Le phénomène général d'*interférence* phonético-phonologique désigne le phénomène par lequel la langue maternelle des sujets interfère avec les efforts qu'ils font pour fonctionner en langue cible (Wells, 2000). Wells, par exemple, mentionne ainsi plusieurs types de difficultés rencontrées par des sujets japonais apprenant l'anglais : a) phonémiques, b) allophoniques, c) phonotactiques et d) suprasegmentales. Le phénomène de transfert, quant à lui, constitue un cas particulier d'interférence, et peut être positif ou négatif.

Tarone (1987, p. 71) donne un exemple de transfert, et des différentes formes qu'il peut prendre selon la comparaison entre les deux systèmes phonologiques en présence :

<b>Transfert</b>	<b>L1</b>	<b>L2</b>
Transfert positif	/t/	/t/
Transfert négatif convergent	/f/	/f/
	/v/	--
Transfert négatif divergent	/l/	/l/
	--	/r/

Le transfert négatif divergent correspond bien évidemment à la situation la plus difficile pour les apprenants. Il est cependant important de noter que le transfert, qu'il soit positif ou négatif, n'est pas un processus systématique, mais bien variable : « positive and negative transfer do not work invariably but variably » (Dickerson, L., citée par Tarone, 1987, p. 76).

### **I. 3. Interphonologie : perspectives générativistes**

Deux types de perspectives doivent ici être distingués : les unes représentationnelles, les autres non représentationnelles.

#### **I.3.1. Perspectives représentationnelles : Marque, Transfert et Grammaire**

##### **Universelle**

Parallèlement aux développements des théories linguistiques, les faiblesses, notamment prédictives, des approches contrastives classiques et les limites des descriptions phonémiques traditionnelles ont vu l'émergence, dans le cadre générativiste, de nouvelles problématiques, notamment vis-à-vis du concept de *transfert*, même si les applications didactiques possibles étaient, et sont peut-être toujours, beaucoup moins évidentes que celles des approches antérieures. Comme l'écrivent Dupoux et ses collègues à propos des inventaires phonémiques et de l'influence de la langue maternelle (1999) :



« *Human languages also differ in the rules that govern what sequences of phonemes are allowed in an utterance. For instance, some languages (e.g., French or English) allow complex clusters of consonants, whereas others (e.g., Japanese) disallow them. One may expect that language-specific influences may be demonstrated beyond phonemic categorization* ».

De plus, si les transferts phonémiques ne suffisent plus à rendre compte des phénomènes observés, le concept même de *transfert* voit ses limites atteintes : en particulier, le concept de transfert de la langue maternelle s'est vu parfois opposé, parfois associé, à celui d'*Accès à la Grammaire Universelle* (U.G.), telle qu'elle est postulée par les générativistes. De nouvelles études ont ainsi vu le jour, dans le domaine phonologique, mais aussi syntaxique (Towell et Hawkins, 1994), entre autres. Comme l'indique Odlin (1989, p. 120) les recherches conduites ces dernières années dans le domaine ont mis à jour certains facteurs phonético-phonologiques pouvant influencer la prononciation en langue étrangère, jusqu'alors négligés dans les analyses contrastives classiques. Parmi ces facteurs figure par exemple la fréquence d'occurrence des phonèmes à travers les langues : un « son » plus rare qu'un autre en L2 serait plus difficile à acquérir, même si les deux sont absents du système de la L1. Cela renvoie au concept de *marque*, qui vaut pour les unités segmentales, mais aussi pour certaines règles phonologiques, comme le dévoisement consonantique final.

### 1.3.1.1. La Marque

Introduit à l'origine par Troubetzkoy, le concept linguistique de *Marque* est présenté comme suit par Archangeli (1997, p. 2), pour qui la linguistique est l'étude d'une part des universaux langagiers<sup>144</sup> et d'autre part de la variation langagière<sup>145</sup> :

---

<sup>144</sup> « The range and type of properties shared in some way by all languages ».

<sup>145</sup> « The range and type of variation possible between languages ».

*« The term **markedness** is used to refer to this continuum [(between language-universal and language-specific properties)], with completely unmarked properties being those found in virtually all languages and extremely marked properties being found quite rarely ».*

Dans le domaine de l'acquisition des langues secondes, le concept de marque a notamment été popularisé par Eckman (1977/1987), qui, dans ses travaux, tentait de pallier certaines insuffisances des analyses contrastives originales, basées sur l'Hypothèse de l'Analyse Contrastive (« the Contrastive Analysis Hypothesis », CAH) de Lado. La CAH se présentait alors ainsi (Lado, 1957/1983, pp. 1-2, cité par Eckman 1987, p. 56) :

*« In the comparison between native and foreign language lies the key to the ease or difficulty in foreign language learning [...]. We assume that the student who comes in contact with a foreign language will find some features of it quite easy and others extremely difficult. Those elements that are similar to his native language will be easy for him and those elements that are different will be difficult ».*

Dans son étude, Eckman (1977/1987) aborde précisément l'une des faiblesses fondamentales de la CAH, à savoir son incapacité à prévoir la difficulté relative des types d'erreurs au sein d'une interlangue. Cela le conduit à poser l'Hypothèse de la Différence de Marque (« the Markedness Differential Hypothesis », MD) (p. 61) :

*« The areas of difficulty that a language learner will have can be predicted on the basis of a systematic comparison of the grammars of the native language, the target language, and the markedness relations stated in universal grammar, such that :*

- (a) *Those areas of the target language will differ from the native and are more marked than the native language will be more difficult.*
- (b) *The relative degree of difficulty of the areas of the target language which are more marked than the native language will correspond to the relative degree of markedness.*
- (c) *Those areas of the target language which are different from the native language, but are not more marked than the native language will not be difficult. »*

Après avoir rappelé les formes *forte* (qui pense pouvoir prédire toutes les erreurs des apprenants sur la base de la comparaison des descriptions des deux systèmes en présence) et *faible* (qui ne prétend pas avoir de valeur prédictive, mais qui sert de principe directeur à l'analyse des erreurs commises par les apprenants) de la CAH (p. 56), Eckman prend ainsi comme point de départ la forme forte de cette hypothèse, afin de proposer une alternative plus à même de rendre compte des données empiriques, illustrées dans son article par l'apprentissage de l'anglais par des apprenants de langue allemande. Sa proposition repose donc sur la notion de « degré de difficulté » qui correspond à celle de « marque typologique », où la marque est définie ainsi (p. 60) :

*« A phenomenon A in some language is more marked than B if the presence of A in a language implies the presence of B ; but the presence of B does not imply the presence of A ».*

Ainsi, l'un des fondements de la MDH est d'affirmer que, étant donné un certain nombre de différences entre les langues, et ainsi un certain nombre de difficultés potentielles pour les apprenants de langue étrangère, seules certaines de ces différences constitueront des zones de difficultés, tandis que d'autres non (p. 66).

### 1.3.1.2. Le Transfert et la Grammaire Universelle

Comme nous l'avons déjà indiqué, plusieurs facteurs semblent aujourd'hui influencer la constitution de l'interphonologie des apprenants (Keys, 2002). Distinguer l'influence respective de chacun de ces facteurs reste néanmoins un travail délicat.

#### 1.3.1.2.1. Le Transfert L1/L2

En 1987, les deux questions centrales pour Tarone, relatives à l'interphonologie, étaient les suivantes (1987, p. 71) :

- The nature of the process which shape interlanguage phonology.
- The phenomenon of fossilization of interlanguage phonology.

Si les explications du deuxième phénomène semblaient alors peu conclusives (Tarone semble insister sur les facteurs plutôt affectifs et culturels que physiologiques), les processus à l'œuvre dans la constitution de l'interphonologie sont résumés comme tels (Tarone, 1987, p. 79) :

- Negative transfer from Native Language
- First language acquisition process
- Overgeneralization
- Approximation
- Avoidance

Beebe (1987) ajoute à la catégorie des *approximations* celle des *composés* (« composites »), qu'elle a relevée dans l'étude dont elle traite (p. 171) :

*« Composites may be considered akin to approximation, but they are termed differently here to emphasize the distinction that they involve sequential production of two variants, not just coloring of one. In some cases, composites appear to be a sequential merger of an L1 and an L2 sound ».*

Elle l'illustre par les productions de sujets chinois et japonais apprenant l'anglais :

<b>Production des sujets chinois</b>	<b>Production des sujets japonais</b>	<b>Cible en anglais</b>
sθ θs	θs	/s/
rl	rl wl	/l/
rl lr	--	/r/

Par ailleurs, les contraintes suivantes semblent également à l'œuvre (Tarone, 1987, p. 79) :

- The inherent difficulty of certain Target Language sounds and phonological contexts
- The tendency of the articulators to rest position
- The tendency of the articulators to a CV pattern
- The tendency to avoid extremes of pitch variation
- Emotional and social constraints

Ce qui apparaît alors ici, c'est que le concept de *Transfert*, spécifique aux langues en présence, tel qu'il avait été abordé à l'origine, semble insuffisant pour rendre compte des phénomènes observés. D'autres facteurs, universaux, semblent devoir être pris en considération et c'est alors qu'intervient la question de l'accès à la Grammaire Universelle, qui renvoie à ce que Tarone nomme « First language acquisition process » dans la liste précédemment citée.

### I.3.1.2.2. L'Accès à la Grammaire Universelle

Dans un cadre générativiste représentationnel, l'acquisition d'un nouveau système phonologique est en effet tributaire à la fois du système phonologique L1 et des propriétés de la Grammaire Universelle (UG) :

$$UG = Principes \text{ (universels et innés) } + Paramètres \text{ (spécifiques et acquis).}$$

Pour Archibald (2000, p. 2), l'apprenant effectue certaines opérations de mise en place (« mapping ») des informations perceptives (générales ou linguistiques) extraites de l'input dans le cadre des structures phonologiques fournies d'une part par UG (*Accès à UG*) et d'autre part par la L1 (*Transfert*). L'acquisition phonologique peut en effet être abordée comme suit : UG fournit des structures sous-spécifiées auxquelles doivent ensuite être assignées des valeurs issues des informations environnementales.

Dans le cas d'une deuxième langue, c'est précisément la question des rôles respectifs (Plein ou Partiel) de ces deux cadres (*valeurs de L1 (Transfert) / valeurs par défaut modifiables (UG)*) qui a donné lieu à plusieurs hypothèses. Celles-ci sont résumées comme suit par Archibald (p. 3) :

<b>Full Transfer/Partial Access</b>	L'input en L2 est mis en place dans le cadre de L1, lequel est sous-tendu par UG, débouchant sur la constitution d'une interlangue stable.
<b>NoTransfer/Full Access</b>	L'input en L2 est mis en place dans le cadre de UG, sans relation avec L1, débouchant sur la constitution de L2
<b>Full Transfer/Full Access</b>	L'input en L2 est mis en place dans le cadre de L1, mais UG intervient tout au long du processus débouchant sur la constitution d'une interlangue stable.

<b>Partial Transfer/Full Access</b>	L'input en L2 est mis en place en partie dans le cadre de L1 et en partie dans le cadre de UG, avec UG intervenant tout au long du processus débouchant sur la constitution d'une interlangue stable.
-------------------------------------	---

Les deux questions essentielles que ces modèles permettent d'aborder, selon Archibald, sont les suivantes :

- 1) Les apprenants adultes d'une seconde langue peuvent-ils remodifier la valeur des paramètres fixés en L1 ?
- 2) Peuvent-ils déclencher l'apparition de nouvelles structures qui n'étaient pas dans leur L1 ?

En se penchant sur l'hypothèse de Full Transfer/Full Access, aux plans segmental, syllabique, moraique et métrique, Archibald répond positivement à la première question, mais reste incertain quant à la deuxième (p. 15).

A chacun de ces plans doit en effet correspondre un apprentissage. Au niveau infrasegmental par exemple, la question se pose de savoir si de nouveaux traits peuvent être acquis ou non, en se basant sur un modèle de géométrie des traits (p. 6). Archibald donne l'exemple de locuteurs japonais, incapables de discriminer /l/ et /r/ en anglais, étant donné l'architecture des traits de leur L1, et il présente l'argument de Brown, selon lequel les locuteurs ne sont capables de percevoir un contraste absent de leur langue première qu'à condition que le trait qui distingue ces deux segments soit présent dans la géométrie des traits de la langue première. Dans cet exemple, /l/ et /r/ sont distingués par le trait [coronal]. Selon Brown, des locuteurs du mandarin vont être capables de percevoir ce contraste, puisqu'en mandarin il y a opposition entre phonèmes coronaux (par exemple pour distinguer des segments alvéolaires et post-alvéolaires). Cela n'est pas le cas du japonais qui ne possède pas de nœud (« node ») coronal. Cependant, d'autres exemples, comme l'aspiration ([spread glottis]) en thaï

pour des apprenants anglais, montrent que l'acquisition de traits absents en L1 est possible.

### **I.3.2. Perspectives non-représentationnelles : optimalité et contraintes**

Dans un cadre générativiste non-représentationnel, l'émergence de la *Théorie de l'Optimalité* (OT) et de la *Théorie des Contraintes et des Stratégies de Réparation* (TCRS) dans les années 1990 ont conduit à de nouveaux développements et de nouvelles approches dans ce qu'il est convenu d'appeler le domaine de l'interphonologie. Les concepts de *transfert* et de *marque* en particulier y prennent un nouveau tour. Nous présentons dans un premier temps plusieurs exemples d'études dans ce domaine, illustrant tant l'approche que les problématiques.

#### **I.3.2.1. Interphonologie et Théorie de l'Optimalité : une nouvelle approche du transfert**

L'étude de Monahan, « Evidence of transference and emergence in the interlanguage » (2001) nous semble être un excellent exemple d'étude originale dans le domaine de l'acquisition des langues seconde, et plus particulièrement de l'interphonologie. Faisant appel à un cadre linguistique théorique récent, la théorie de l'optimalité, il aborde la question du transfert phonético-phonologique de la L1 vers la L2, non plus en termes classiquement phonémiques, ou bien même en termes de traits, mais plutôt en termes métriques, puisqu'il s'intéresse aux structures syllabiques des langues en question et cherche à observer l'influence du système phonologique de la L1 sur celui du système interphonologique L1-L2 en termes de contraintes et de classement de ces contraintes. D'une certaine manière, cette approche (interlangue, syllabe, OT) représente la version « moderne » des études comparatives ou contrastives classiques des années 50 (interférence, phonème, structuralisme). C'est d'ailleurs bien ce qu'annonce Archangeli (Archangeli, 1997, p. 31): une nouvelle approche de l'analyse contrastive,



non plus en termes de traits ou de phonèmes, mais en terme de hiérarchisation des contraintes dans le cadre de OT.<sup>146</sup>

De nombreux chercheurs ont ainsi conduit des études portant sur l'interphonologie en utilisant la théorie de l'optimalité. Monahan cite Broselow et ses collègues<sup>147</sup> qui se sont penchés sur la production d'obstruantes en position de coda par des locuteurs natifs chinois (mandarin) apprenant l'anglais. Comme l'indique le titre de leur article, « The emergence of the unmarked in second language phonology », le résultat le plus intéressant concerne la préférence des apprenants pour des formes non-marquées qui ne relèvent ni de l'anglais ni du mandarin. En particulier, l'étude montre que, à l'aide d'épenthèses et de délétions, les formes dissyllabiques (moins marquées) sont favorisées, tout comme les obstruantes non-voisées (moins marquées que les voisées). Monahan indique (p. 17) que, d'après Broselow et ses collègues, ces formes proviendraient du reclassement de contraintes qui ne sont pas dominantes en mandarin mais qui le deviennent dans l'interlangue en raison de la richesse de la base pour la langue-cible (ici l'anglais).

Cela correspond, dans un cadre théorique différent, à ce qu'avait déjà suggéré Broselow en 1983 (p. 279) :

*« I have argued that language transfer does play a significant role in second language acquisition : certain systematic errors can be directly attributed to the use by language learners of a phonological rule in the production of second language forms. [...] I have noted one sort of error which is not attributable to*

---

<sup>146</sup> Notons donc dès à présent que si ce type d'approche est encore non majoritaire dans les domaines purement linguistiques, il est sans aucun doute totalement minoritaire, pour ne pas dire absent, des préoccupations de la linguistique appliquée, tout au moins dans le domaine francophone, et plus particulièrement de la didactique des langues et des cultures. Certaines analyses contrastives à visées didactiques devraient ainsi bénéficier des nouvelles approches « linguistiques » qui sont aujourd'hui disponibles.

<sup>147</sup> Broselow, E., Su-I C. et Wang, C., « The emergence of the unmarked in second language phonology », *Studies in Second Language Acquisition*, 20 : 261-280, 1998.

*language transfer and have suggested that universal principles of phonological patterning also play a role in accounting for the errors of language learners ».*

Tarone (1987, p. 70) avait souligné le même point :

*« Transfer is only a part – and often a small part—of the influence on interlanguage phonology ».*

### *1.3.2.2. Exemple 1: épenthèses et émergence du non-marqué chez des apprenants brésiliens d'anglais*

Dans son étude, Monahan vise à déterminer si des locuteurs brésiliens natifs, parlant le portugais brésilien (BP) et apprenant l'anglais, transfèrent le classement des contraintes syllabiques de leur L1 (BP) vers celui de leur L2 (anglais). Monahan envisage ainsi le transfert, en début d'article, de la manière suivante (p. 3):

*« Second language learners have a tendency to transfer phonological constraints and pronunciations from their first language (L1) into their second language (L2) ».*

Il a donc fait lire à voix haute par ses sujets brésiliens des phrases anglaises contenant certaines constructions phonologiques licites en anglais mais absentes en BP. Il a ensuite analysé phonétiquement et phonologiquement les données récoltées, puis établi le classement des contraintes syllabiques en BP, en L2 anglais et en L1 anglais, avant de les comparer. En BP, comme en japonais, il souligne la nécessité de bien distinguer les différentes strates du lexique (en particulier celles des emprunts), puisque dans sa description de la structure syllabique du BP, Monahan prend soin de noter, par exemple, que : « [dl] does not occur [in onset position] except across morpheme boundaries in borrowed lexical items, which means that it is not syllabified in the onset » (p. 8). Le phénomène d'épenthèse se trouve par exemple chez les locuteurs portugais (p. 8) :

*« In Portuguese, non-permissible onset clusters appear in the input forms (underlying forms and orthographically), in such cases as \*pn pneu « tyre », \*gn gnomo « gnomo », and \*ps psicologia « psychology ». In these examples we find epenthesis, whereby [i] is inserted to break up impermissible cluster. For example, pneu is [pi-néw] and gnomo is [gi-nómu] [...]. This shows that when native speakers of Portuguese are confronted with non-native like clusters they use epenthesis in order to retain the ideal CV syllable structure »*

Cependant, dans l'étude de Monahan, les processus d'épenthèse sont bien moins nombreux que ce à quoi l'on pouvait s'attendre : à la question « est-ce que les locuteurs natifs de BP ont recours à l'épenthèse afin que leur phrases en anglais contiennent des syllabes bien formées en BP », les résultats de son analyse le conduisent à répondre négativement (Monahan, 2001, p. 21).

De cette étude, il ressort trois résultats principaux (p. 4):

- Il y a eu un transfert de contraintes significatif dans le classement de l'interlangue en ce qui concerne la coda (ce qui entraîne des processus comme des assimilations régressives de nasalité par exemple) ;
- Il y a eu ce que certains appellent « l'émergence du non-marqué » : des contraintes qui conduisent à l'émergence de structures non-marquées dans l'interlangue, mais absente tant de la L1 (BP) que de la L2 (anglais) ;
- Il n'y a pas eu (contrairement à l'hypothèse de départ) d'épenthèse lorsque les sujets on fait face à des marges syllabiques complexes, comme il y en a en anglais.

D'un point de vue méthodologique, on doit signaler le problème de la lecture à voix haute comme source de données phonético-phonologiques et du rapport entre graphisme et phonie. Ce point méthodologique est également à soulever, en partie, dans l'étude de Shinohara (1997). Notons également le nombre de participants : 5 seulement. L'enquêteur leur a demandé de lire des phrases en anglais avec trois rythmes différents : soigné (lent), familier (normal) et rapide, tandis que la cible (X)

était contextualisée et entourée de deux voyelles : « I will say X again » (p. 18). Les mots cibles ont été choisis parce qu'ils contenaient un groupe consonantique ou bien un segment absent en BP.

D'un point de vue formel, Monahan fait appel à un certain nombre de contraintes, toutes issues de OT, certaines générales, d'autres plus spécifiques :

<b>Contraintes</b>	
<b>DEP-IO :</b>	Output segments must have input correspondents (no epenthesis) (Faithfulness)
<b>MAX-IO :</b>	Input segments must have output correspondents (no deletion) (Faithfulness)
<b>*ó[Obs+[-Appx] :</b>	Obstruents plus non- approximants (obstruents and nasals) are not permissible in onsets
<b>NoCODA (*C]ó) :</b>	Syllables are open
<b>*ComplexCODA (*CC]ó) :</b>	Codas must be simple
<b>ONSET (ONS) :</b>	Onsets are obligatory
<b>*Vnasal :</b>	Vowels must not be nasal
<b>*VoralN :</b>	Before a tautosyllabic nasal, vowels must not be oral
<b>MAX-IO (Nasal) :</b>	If the feature [+nasal] is in the input form then it must also be in the output form
<b>IDENT-IO (Nasal) :</b>	Corresponding segments in the input and output must have identical values for [nasal]
<b>MAX (Lateral) :</b>	If the feature [+lateral] appears in the input form then it must also appear in the output form
<b>CODA-COND :</b>	Only [s] and [r] are permissible in the syllable coda
<b>MAX-IO (Obs) :</b>	Every obstruent (oral stop, fricative, affricate) in the input must be retained in the output

Si l'on voulait suivre la même démarche que celle de Monahan (p. 27), il nous faudrait :

1. Etablir le classement des contraintes phonotactiques syllabiques du japonais et du français ;
2. Examiner le classement de ces contraintes dans l'interlangue de nos sujets à travers les données récoltées, afin de comparer les trois classements.

### I.3.2.3. Exemple 2 : épenthèses et variation chez des apprenants chinois d'anglais

Une autre étude en interphonologie dans le cadre OT est celle de Radwanska-Williams et Yam (2001) sur l'acquisition des plosives anglaises en coda syllabique par des apprenants chinois, locuteurs de chinois cantonais et mandarin. Ici encore les phénomènes observés se ramènent essentiellement à des processus d'épenthèses vocaliques et de délétions consonantiques, qui sont dues aux interactions entre transferts de la L1 et facteurs universaux de type développemental. Dans leur étude, 12 sujets de niveaux avancé ont été choisis pour enregistrer leur prononciation de mots anglais mono-, bi- tri- et tétrasyllabiques contenant des singletons consonantiques /plosive/ ainsi que des codas branchantes du type /nasale + plosive/ en finale de mot. Trois types de formes déviantes ont été notés (p. 36) : 1) schwa épenthétique après plosive créant une nouvelle syllabe, 2) effacement consonantique (une seule consonne dans les codas branchantes), 3) plosives non-relâchées. L'analyse théorique effectuée à partir de ces données met en œuvre cinq contraintes syllabiques, classées différemment pour les Mandarins et les Cantonais (Syllable Structure Constraint Ranking, SSCR) :

SSCR Mandarins : \*Complex >> \*Obstruent coda >> Fill (V), Parse (C)

SSCR Cantonais : \*Complex >> \*Released Obstruent coda >> Fill (V) >> Parse (C)

L'une des principales conclusions de cette étude est que les contraintes de fidélité Fill (V) et Parse (C) sont haut placées dans le classement des contraintes en anglais, alors qu'elles ne le sont pas dans le système des apprenants chinois, reflétant là le système

phonologique de leur langue première. On peut apprécier dans cette étude la mise en relief de la variabilité inter et intra-individuelle dans la prononciation de mots stimuli.

*1.3.2.4. Exemple 3 : reclassement de contraintes et émergence du non-marqué chez des apprenants zoulous d'anglais*

Grijzenhout et Van Rooy (2001) se sont penchés sur le cas de locuteurs natifs de langue zoulou apprenant l'anglais, en comparant leur apprentissage à celui de l'anglais langue première vis-à-vis de ce qui est appelé dans la littérature « l'émergence du non marqué ». S'inscrivant dans le cadre OT, ils adhèrent à l'hypothèse selon laquelle l'apprentissage d'une langue première implique le déclassement graduel des contraintes de Marque (initialement haut placées dans le classement) au profit de contraintes de Fidélité. Leur étude vise à montrer que, lorsque certaines contraintes de marque ne nécessitent pas de déclassement en langue première, celles-ci restent haut classées dans la grammaire du locuteur et sont responsables du phénomène d'émergence du non marqué durant l'apprentissage d'une seconde langue.

Les auteurs se sont concentrés sur l'absence d'attaque complexe et sur le phénomène de dévoisement de coda. Comme ils le notent (p. 5), il est en effet troublant de constater que, en ce qui concerne le dévoisement des obstruantes en coda (correspondant au cas non marqué), s'il n'est pas étonnant d'observer ce phénomène chez des locuteurs néerlandais ou allemands par exemple (langues ne tolérant pas d'obstruantes voisées en coda) qui apprennent l'anglais (langue tolérant le voisement des obstruantes en coda), on ne peut en revanche parler de « transfert » de la langue première vers l'interlangue pour des locuteurs de Zoulou ou de Mandarin puisque ces langues ne présentent pas de trace de règle de dévoisement, tandis que les locuteurs natifs de ces langues semblent pourtant suivre ce cas « non-marqué » dans leur apprentissage de l'anglais. Les auteurs visent donc à établir et comparer les classements des contraintes pertinentes en anglais, en zoulou et dans l'interlangue des sujets. Dans leur analyse des phénomènes syllabiques observés chez les locuteurs zoulous, les auteurs font notamment appel aux contraintes suivantes (p. 12) :

<b>Contraintes</b>	
<b>NOCODA :</b>	Syllables must be open
<b>NOCOMPLEXONSET :</b>	Consonant clusters do not occur in onset position
<b>MAXCONS :</b>	All consonants in the input must have a correspondent in the output (no deletion)
<b>DEPVOWEL :</b>	All vowels in the output must have a correspondent in the input (no epenthesis)

Leur étude les conduit à dire que la différence entre l'acquisition d'une première et d'une deuxième langue provient du classement initial des contraintes :

- Pour la langue première, le classement est issu d'un classement universel dans lequel la Marque est plus importante que la Fidélité.
- Pour la seconde langue, c'est le classement de la langue première qui est pris comme base de départ.

La prononciation en langue étrangère est donc bien tributaire non seulement de la langue première (spécificité du classement de contraintes), mais aussi de facteurs universaux (universalité des contraintes). Ils soulignent par ailleurs avec justesse que le processus d'intégration est différent entre des sujets qui *adaptent* des mots étrangers et des sujets qui *apprennent* la langue étrangère : les premiers ne cherchent pas à se conformer à la phonologie de la langue étrangère, mais plutôt à les adapter à leur propre système de strates lexicales ; les seconds en revanche essaient de respecter autant que possible les formes étrangères. Cela est tout à fait pertinent pour notre étude, et il faut donc garder à l'esprit que, entre adaptation/emprunt et apprentissage, les objectifs sont bien différents et l'attention, entre autres ressources cognitives, qui y est allouée également.

I.3.2.5. Exemple 4 : coda et émergence du non-marqué chez des apprenants italiens d'anglais

Une autre étude encore est celle de Villafaña (2000) sur des sujets anglais apprenant l'italien : se concentrant sur la production de coda, elle se base sur le *Principe de Dispersion* de Clements qu'elle cite (« sonority rises maximally towards the peak and falls minimally towards the end » p. 35) pour indiquer que, de manière générale à travers les langues, les codas sonorantes sont préférées aux codas obstruantes et donc moins marquées que celles-ci. Comparant les structures syllabiques anglaises et italiennes, Villafaña cite le *Principe de Coda Harmonique* de Nagy et Napoli (« Harmonic Coda Principle », p. 35) selon lequel plus une consonne est sonore (en termes de sonorité), moins elle est marquée en position de coda, qu'elle reformule dans le cadre OT en « NoObstCoda », une contrainte plus active en italien qu'en anglais.

Ses résultats indiquent une préférence pour les codas sonorantes, ce qui est surprenant puisque la contrainte à l'œuvre n'est pas active en anglais. Rejetant d'une part l'idée de transfert (puisque en L1 il n'y a pas de préférence pour les sonorantes), et d'autre part l'idée de mise en œuvre de connaissances de la langue cible (L2) (étant donné le niveau d'apprentissage des sujets, ainsi que le dispositif méthodologique suscitant la production de codas), Villafaña suggère donc que ces résultats constituent un cas « d'émergence du non marqué » : même si la contrainte NoObstCoda n'est pas active en anglais, elle est cependant présente dans la grammaire du locuteur et refait surface lors de l'apprentissage de la langue étrangère débouchant sur une interlangue provisoire.

Dans cette perspective, chaque locuteur possède, dans sa grammaire, des contraintes négatives, certaines étant dominées par d'autres en langue maternelle. Les fluctuations propres aux classements interlangiers permettent la ré-émergence de ces contraintes jusqu'alors sous-classées. Ainsi, des trois raisons qui pourraient expliquer pourquoi ces locuteurs présentent une certaine préférence de sonorité, pourtant absente en langue



maternelle, à savoir le transfert, la connaissance de la langue cible et l'accès à UG, c'est bien la troisième qui semble être plausible dans la démonstration de Villafaña.

D'un point de vue méthodologique, malgré l'échantillon réduit de sujets, on peut apprécier l'intérêt du dispositif qui, de par le caractère non-explicite des consignes et par la technique du jeu langagier favorise la dimension écologique des tests, tout en se préservant des facteurs orthographiques puisque les mots sélectionnés étaient lus non pas par les sujets mais par l'enquêteur. Cette étude œuvre en faveur de l'hypothèse selon laquelle les apprenants adultes auraient accès à UG, selon une théorie de contraintes négatives universelles innées qui sont accessibles en permanence aux apprenants adultes.

#### *1.3.2.6. Exemple 5 : substitutions segmentales chez des apprenants japonais d'anglais*

A propos des substitutions segmentales en langue étrangère, on peut se pencher sur les travaux de Lombardi (2000), qui s'est intéressée aux substitutions des segments interdentaires anglais ([θ, ð]), analysées à l'aide de contraintes de Manière (« Mode ») dans le cadre de OT. Comme elle l'indique, ce qu'on appelle « effet de transfert » dans le domaine, de L1 à L2, ne permet pas de rendre compte de tous les effets décelés en interphonologie, notamment lorsque certains segments en L2 n'existent pas en L1 : c'est par exemple le cas des interdentaires en anglais qui sont relativement systématiquement remplacées par tel ou tel segment selon la L1, par exemple [s] pour des sujets germanophones et [t] pour des sujets russophones. (p. 1).

Dans le cas des substitutions phonémiques, le choix du substitut issu du transfert a d'abord été envisagé comme conditionné par un facteur de « proximité ». Mais c'est bien cette « proximité » qui est délicate à définir. Outre les effets de transfert, des effets développementaux et universaux ont été considérés : les premiers sont supposés refléter les similarités entre acquisition de L1 (ordre d'acquisition des phonèmes par l'enfant, par exemple plosives avant fricatives) et acquisition de L2, tandis que les

seconds reflètent la préférence pour les structures non marquées (par exemple structure CV, même lorsque les syllabes fermées sont licites en L1 et L2).

Cependant, Lombardi soutient que ces études ne sont pas satisfaisantes, et elle propose donc une analyse dans le cadre de OT, qui permet de montrer que quel que soit le type d'effets, effets universels ou effets de transferts « non-évident », selon les L1, ceux-ci résultent toujours du classement de contraintes en L1. Ainsi, dans le cas des interdentes, Lombardi indique que (pp. 2-3) :

- Si le mode original est respecté (fidélité), alors le substitut est [s]. Cela se produit dans des langues où les contraintes portant sur le mode en L1 sont explicitement classées (cas de « transfert », par exemple pour le japonais ou pour l'allemand).
- Si le résultat est un segment moins marqué, alors le substitut est [t]. En effet, étant donnée que les plosives sont acquises plus tôt que les fricatives par les enfants d'une part, et que toutes les langues possèdent des plosives mais non des fricatives d'autre part, il apparaît que les plosives sont moins marquées que les fricatives du point de vue du Mode. Cela se produit donc dans les langues où les contraintes de mode ne sont classées que par défaut (le cas du non-marqué, « universel », par exemple pour le thaï ou pour le russe).

Lombardi pose donc le classement suivant, relatif au degré de marque des plosives et des fricatives (p. 4) :

\*[cont] >> \*[stop]

Dans d'autres cas, le schéma de substitution est plus complexe, par exemple pour le néerlandais (p. 15) où le [t] est substitué en position d'attaque syllabique, mais où le [s] est substitué en position de coda syllabique.

Son étude la conduit finalement (p. 19) à poser que, dans le cas des interdentes, les substitutions par plosives résultent d'effets universaux, propres au classement par défaut de UG qui n'a pas été modifié en L1 lors de l'acquisition de la langue maternelle. En revanche, les substitutions par fricatives correspondent à un effet de transfert explicite, le classement initial des contraintes ayant été modifié lors de l'acquisition de la L1. Pour en rendre compte, Lombardi présente une certaine formulation des contraintes de fidélité sur le mode : la grammaire initiale (en L1), dans laquelle les contraintes de marques dominant les contraintes de fidélité, doit contenir une contrainte unitaire *IdentManner* avec fidélité pour les plosives, les continuantes, et les stridentes. Par la suite, le rapport de dominance change entre marque et fidélité, et la contrainte *IdentManner* peut alors (mais ce n'est pas systématique) être divisée en trois contraintes indépendantes propres à chaque trait.

Ainsi, si la grammaire en L1 ne nécessite pas de division de *IdentManner*, on obtient la situation suivante, conduisant à une substitution par plosive (p.7) :

/θ/	*θ	<b>IdentManner</b>	*Cont	*Stop
θ	* !		*	
+ t		*		*
s		*	* !	

Si, au contraire, la grammaire en L1 nécessite une division de *IdentManner* en *IdentCont* et *IdentStop*, alors on obtient la situation suivante, conduisant à une substitution par fricative (p. 9) :

/θ/	*θ	<b>IdentCont</b>	<b>IdentStop</b>
θ	* !		
t		*	*
+ s			

Seule une analyse de la sorte permettrait au mode fricatif d'être remplacé par le mode plosif, moins marqué.

Lombardi se base notamment sur le cas du japonais (pp. 11-13) en se référant à l'analyse de l'affrication de /t/ en /ts/ devant /u/ par Ito et Mester, qui soutiennent que, dans la strate périphérique du lexique phonologique : \*TU >> FaithFeat >> \*TS. Cela pose problème lorsque l'on prend en compte [su] comme candidat optimal possible. La solution, selon Lombardi, réside dans la division de FaithFeat en IdentStop et IdentCont. Il en va de même dans la strate centrale (Core), dans laquelle la proposition initiale de Ito et Mester (\*TU >> \*TS >> FaithFeat), pose un problème identique.

Lombardi propose alors les classements suivants :

<b>Strate</b>	<b>/tu/</b>	<b>*TU</b>	<b>IdentStop</b>	<b>IdentCont</b>	<b>*TS</b>
<b>Périphérique</b>	<b>tu</b>	<b>* !</b>			
	<b>+ tsu</b>			*	*
	<b>su</b>		<b>* !</b>	*	

<b>Strate</b>	<b>/tu/</b>	<b>*TU</b>	<b>IdentStop</b>	<b>*TS</b>	<b>IdentCont</b>
<b>Centrale</b>	<b>tu</b>	<b>* !</b>			
	<b>+tsu</b>			*	*
	<b>su</b>		<b>* !</b>		*

Ainsi, le japonais est une langue L1 qui nécessite de ses locuteurs un classement séparé de IdentStop et IdentCont, ce qui s'accorde avec la substitution des interdentes en fricatives.

I.3.2.7. Exemple 6 : groupes consonantiques et voyelles épenthétiques en interphonologie

En ce qui concerne le traitement des groupes consonantiques, on peut se pencher sur les travaux de Fleischhacker (2000), qui indique que les épenthèses vocaliques figurent parmi les principales stratégies de traitement des groupes consonantiques lorsque ceux-ci sont illicites en langue empruntant, dans le cas de l'adaptation des mots d'emprunts, ou en langue maternelle, dans le cas des interphonologies. La localisation de la voyelle épenthétique dépend principalement de la nature du groupe consonantique, le schéma général étant le suivant :

- Insertion préconsonatique (« prothèse ») pour les groupes Sibilantes + Stop
- Insertion intraconsonantique (« anaptyxe ») pour les groupes Obstruantes + Sonorantes.

Rappelons à ce propos qu'en français :

- Les sibilantes [+ sib] sont : /s, z, ʃ, ʒ/.
- Les plosives [- cont] sont : /p, b, t, d, k, g/.
- Les obstruantes [- son] sont : les plosives [- cont] et les fricatives [+ cont] (/f, v/ et les sibilantes).
- Les sonorantes [+ son] sont : les liquides (/l, R/), les glides (/w, ɥ, j/) et les nasales (/m, n, ɲ, ŋ/).

Fleischhacker a principalement travaillé sur les groupes consonantiques en position initiale, en se basant sur des données issues de productions interlangières d'apprenants de langue étrangère, ainsi que d'adaptation de mots étrangers. Elle est ainsi parvenue à établir une typologie, non-systématique, des stratégies épenthétiques mises en œuvre pour la résolution des groupes consonantiques, selon différentes langues (p. 10). Cette typologie présente un continuum de groupes consonantiques (Sibilantes+Plosives (ST), Sibilantes+Sonorantes (SR), Plosives+Sonorantes (TR)) partiellement agencés selon un profil de sonorité décroissante pour les sonorantes. Pour chaque langue est indiqué le

site épenthétique principal en fonction des groupes, ainsi que l'existence d'asymétries ou non (p. 9). Cette typologie permettrait ainsi certaines prédictions quant au type d'épenthèse possible et à sa localisation, selon les langues et les groupes consonantiques.

L'étude de Fleischhacker se base sur les travaux de Broselow et ceux de Singh, qu'elle remet partiellement en question en adoptant une approche « perceptive » (« perceptual »), approche qui insiste sur le rôle des facteurs perceptifs dans les processus phonologiques. Comme elle l'explique, ces deux chercheurs établissent en effet une distinction binaire entre les groupes Obstruantes + Sonorantes (OR) d'une part et Sibilantes + Plosives (ST) d'autre part : Broselow considère l'anaptyxe comme la stratégie épenthétique de résolution des groupes consonantiques par défaut, en attribuant aux groupes Sibilantes + Plosives un statut quasi-monosegmental exceptionnel (« structure-based approach », p. 12), tandis que Singh présente la localisation des sites épenthétiques comme fonction du profil de sonorité desdits groupes (« sonority-based approach », p. 11). Or, la dichotomie ST-OR, telle que présentée par des chercheurs comme Singh ou Broselow, ne permettrait pas, selon Fleischhacker, de rendre compte de toutes les schémas épenthétiques à travers les langues, ce qu'elle illustre en mentionnant les langues Kazakh et Farsi (p. 13). En ce qui concerne le japonais, ainsi que le coréen, Fleischhacker note que pour les groupes ST à l'initiale, c'est l'anaptyxe qui est néanmoins favorisée (« across-the-board anaptyxis », pp. 7-8), puisqu'une prothèse dans ce cas résulterait en une structure illicite d'un point de vue phonotactique/syllabique.

Fleischhacker, dans son étude, suppose que, pour les langues présentant des asymétries entre anaptyxes et prothèses, la localisation est déterminée en fonction de la similarité auditive (« perceptive ») que présente chacun des deux schémas épenthétiques entre input (avant épenthèse) et output (après épenthèse). Elle s'appuie sur deux expériences dont les résultats indiquent que les anaptyxes favorisent la similarité auditive pour les groupes Obstruantes + Sonorantes, tandis que les prothèses sont préférées pour les groupes Sibilantes + Plosives. Dans ces expériences, des mots anglais comportant des

groupes ST, OR et STR en initiale étaient modifiés, puis présentés à des sujets anglophones qui devaient exprimer un jugement de préférence concernant la similarité auditive des items modifiés avec les mots initiaux. L'analyse qu'elle propose de ses résultats se situe dans le cadre de la théorie de l'Optimalité : elle s'attache en particulier, à étendre l'une des contraintes de correspondance, DEP-V, en lui assignant une sensibilité au contexte, ce qu'elle formule ainsi (p. 30):

**DEP-V/X\_Y** : a vowel present in the output context X\_Y has a correspondent in the input context X\_Y

Cela lui permet de poser la hiérarchie de contraintes, suivante (p. 30) :

**Partial DEP-V/X\_Y scale** : DEP-V/S\_T >> DEP-V/S\_N >> DEP-V/S\_L >> DEP-V/T\_R

(avec : S=voiceless sibilant ; T=stop ; N=nasal ; L=liquid ; R=sonorant).

C'est également dans cette approche perceptive que se situent les travaux de Côté (2001), qui s'est basée sur des données en français parisien, du Québec et du pays Basque. Celle-ci s'est penchée sur les processus d'épenthèses et de délétion vocalique et consonantique en fonction de l'environnement consonantique. Pour elle, ces processus sont fonctions, pour un environnement consonantique donné, de la qualité et de la quantité des indices auditifs (« auditory cue ») qui lui sont associés. Elle se démarque ainsi du *Principe de Licence Prosodique*, qui insiste sur la nécessaire syllabification de tous les segments. Elle établit de nouvelles généralisations à base empirique, notamment le fait que, statistiquement, les plosives ont de plus fortes chances d'être effacées, de déclencher des épenthèses ou de bloquer l'effacement vocalique que les autres consonnes.

Côté soutient que cette approche permet de simplifier considérablement l'appareil théorique nécessaire à l'analyse de tels processus, en mettant au premier plan la notion de saillance perceptive, qui permet, entre autres de se passer de mécanismes

exceptionnels comme l'extra-syllabité aux frontières de domaines. La notion de saillance perceptive renvoie au principe selon lequel une consonne est effacée, déclenche une épenthèse ou bloque un effacement vocalique lorsque les indices qui permettent à l'auditeur de repérer leur présence sont diminués. Tandis que l'effacement consonantique permet d'enlever le segment déficient, l'épenthèse et le blocage d'un effacement vocalique permettent d'ajouter de tels indices.

Côté propose ainsi, dans le cadre de la théorie de l'optimalité, plusieurs contraintes de marques, motivées d'un point de vue perceptif :

- 1) C deletion and V insertion/retention only or preferentially affect Cs not adjacent to a V.
- 2) C deletion and V insertion/retention affect stops more than other types of Cs.
- 3) C deletion and V insertion/retention preferentially target stops followed by a [-cont] C.
- 4) C deletion and V insertion/retention preferentially target Cs similar to adjacent segments.

En ce qui concerne la nature des voyelles épenthétiques, Lombardi (2002) a proposé que, de manière générale, les voyelles antérieures sont plus marquées que les voyelles postérieures, et que les voyelles arrondies sont marquées. Le choix de la voyelle épenthétique selon les langues dépend du classement relatif de marque des voyelles basses et non basses, la voyelle épenthétique étant, normalement, la moins marquée du système vocalique. En ce qui concerne le japonais (p. 17), elle note que, contrairement à ce qui est fréquemment affirmé :

*« Epenthesis of [u] is sometimes claimed to exist but it is unclear whether there are any cases that don't involve additional contextual conditioning. Epenthesis in Japanese loans is not a case if the high back vowel is unrounded. In addition, Japanese uses different vowels in different vocabulary strata and contextual concerns seem to be relevant : the use of [u] and [o] in loans instead of the [i]*



*of the native strata allows all consonants in loan words to be realized without phonological modification ».*

Une autre approche concernant les groupes /s/ + Plosive est celle de Morelli (1999) qui a essentiellement travaillé sur les groupes consonantiques en attaque syllabique (p. 30), et plus particulièrement les groupes obstruants, qu'elle définit pour son étude comme des séquences tautosyllabiques de plosives (S) et fricatives (F) (p. 32). Travaillant sur une trentaine de langues dans une démarche typologique, elle montre que les différents groupes se positionnent les uns par rapport aux autres selon différents degrés de marque, et que, pour les dimensions pertinentes, en particulier celle du trait [continuatif] (Mode), les groupes /s/ + Plosive sont les moins marqués. Morelli (1998, p. 4) établit ainsi les relations de marque suivantes :

$$\text{ET} \left| \begin{array}{l} \text{FS} > \text{FF} \\ \text{FS} > \text{SF} > \text{SS} \end{array} \right.$$

Les groupes FS sont les moins marqués, tandis que les groupes SS sont les plus marqués. En ce qui concerne la dimension de lieu (« Place »), ce sont également les moins marqués puisque la fricative en position de pré-obstruante est une coronale (1999, p. 154). Elle soutient alors que les attaques /s/ + Plosive ne sont finalement que des attaques tautosyllabiques régulières, et non des groupes hétérosyllabiques ou monosegmentaux, montrant que les indices en faveur de telles analyses peuvent être traités de manière indépendante.

Dans cette analyse, les groupes /s/ + Plosive sont donc les moins marqués, au lieu d'être les plus marqués comme c'est le cas dans les analyses antérieures (statut exceptionnel). Selon elle, la phonotactique des groupes obstruants peut être plus facilement saisie en faisant explicitement référence aux faits phonétiques, le schéma d'occurrence le plus fréquent étant Fricative Coronale + Plosive. Morelli conclue donc

(1999, p. 193) que les groupes avec obstruantes sont particuliers en ce qu'ils ne sont pas contraints par des principes de sonorité, ce qui lui permet d'affirmer que la particularité des groupes /s/ + Plosive provient de leur caractère non-marqué pour les dimensions relatives aux groupes obstruants, et non de leur caractère marqué vis-à-vis des groupes de base (« core clusters »). Les « core clusters » sont ceux qui respectent le SSP, c'est-à-dire dont la sonorité augmente en direction du pic syllabique (Morelli, 1998, p. 1).

Il faut souligner que, dans le cadre OT, le degré de marque correspond à celui d'« harmonie » : une forme est marquée par rapport à une certaine contrainte C si elle viole celle-ci, et reçoit alors une marque. Etablir une hiérarchie de marque entre formes revient alors à comparer l'harmonie relative de ces formes par rapport à telle ou telle contrainte ou système de contrainte (Morelli, 1998, p. 6). Cela explique l'absence de relations entre les groupes FF d'une part, et SF et SS d'autre part dans la hiérarchie de Morelli (1998, p.4), puisque, d'après elle, s'ils sont équivalents en termes de marque (puisque'ils impliquent tous deux l'existence des groupes FS, moins marqués), ils ne le sont pas en termes d'harmonie car les marques qu'ils reçoivent ne sont pas identiques, étant donné que les contraintes auxquelles elles se réfèrent sont différentes<sup>148</sup>.

### **I.3.3. Synthèse**

Nous pouvons à présent tenter de synthétiser les différentes perspectives évoquées.

#### **I.3.3.1. Marque et Transfert dans la Théorie de l'Optimalité**

Uffmann (2003) propose une synthèse originale de la dichotomie *Transfert* vs. *Marque* dans le domaine de l'acquisition des langues secondes, dichotomie qu'il parvient à réduire *via* la théorie de l'Optimalité. Nous nous en inspirons donc largement ci-dessous.

---

<sup>148</sup> Pour le détail des contraintes, voir Morelli, 1998, pp. 7-8.

Comme il le rappelle, alors que pour Lado (1957) le transfert (de règles ou d'unités) explique tout, Eckman (1977/1987) souligne l'importance des effets de marque, qui conduit à l'émergence du non-marqué dans l'interlangue. Le débat a donc porté par la suite sur le rôle à attribuer à chacun de ces deux grands pôles, et à éclaircir la nature et la fonction des effets de marque.

Dans le cadre de la théorie de l'optimalité, l'acquisition d'une langue première résulte d'un certain déclassement de contrainte (« constraint demotion ») à partir d'une strate initiale, en fonction de la grammaire à acquérir. En particulier les contraintes de Fidélité (F) doivent être en fin de classement, alors que les contraintes de Marque (M) doivent être haut classées. Il s'agit du Principe de Faible Fidélité (« low faithfulness ») qui indique que, si M et F ne peuvent être classées en fonction des données disponibles, alors F sera déclassée. L'hypothèse « nulle » de l'apprenant est donc :  $M \gg F$ .

Dans ce même cadre, l'acquisition d'une langue seconde se distingue de la première en ce que « la strate initiale » correspond au classement (spécifique) des contraintes (universelles) de la L1. Il s'agit donc alors d'un *re*-classement de ces contraintes en fonction de leur classement en L2, au cours duquel s'applique également le Principe de Faible Fidélité. Ce principe fait que 1) le déclassement de F est une stratégie spontanée, alors que 2) le déclassement de M est provoqué par les données de la L2. Dans un premier temps, le classement en L1 est transféré dans l'interlangue, puis on peut distinguer plusieurs étapes dans ce re-classement, avant de parvenir au classement de la L2. Les effets de transfert résulteraient alors d'un reclassement imparfait.

Dans ce cas, certains phénomènes, comme par exemple le dévoisement des codas, peuvent constituer à la fois un cas de transfert (codas dévoisées en L1) et un effet de marque (les codas dévoisées sont moins marquées que les codas voisées). Ainsi OT permettrait d'unifier les deux approches (Uffmann, 2003, p. 7) :

Si en L1  $M \gg F$  et si dans l'interlangue  $M \gg F$ , alors il y a convergence des effets de transfert et de marque.

Une telle approche permet, selon Uffmann, de rendre de compte de certains phénomènes jusque là difficilement explicables. Il prend ainsi l'exemple de certaines L1, comme le japonais, qui, n'ayant pas de coda, n'ont également pas de règle de dévoisement pour les codas. Cela renvoie en partie à ce que l'on nomme les « classements cachés » : outre le classement apparent en L1, il existe aussi chez les sujets un classement « caché » pour des contraintes qui ne sont a priori pas pertinentes en L1, mais qui peuvent le devenir lors de l'apprentissage de la L2.

Ainsi, en japonais, le classement serait le suivant :

\*CODA/VOI >> \*CODA >> FAITH

En revanche en français, il serait le suivant :

FAITH >> \*CODA/VOI >> \*CODA

D'après les données du français (langue cible) pour l'apprenant japonais, il y a des codas en français, donc il y aurait déclassement de \*CODA, conduisant, pour l'interlangue à :

\*CODA/VOI >> FAITH >> CODA.

Ainsi, alors qu'il n'y a apparemment pas de « dévoisement de coda » en japonais (puisque'il n'y a pas de coda), dans ce contexte d'apprentissage la contrainte haut placée \*CODA/VOI peut devenir active, conduisant à l'émergence du non-marqué (dévoisement final), en raison du déclassement de \*CODA.

Ainsi, grâce au cadre OT, Uffmann parvient à unifier *marque* et *transfert*, en ramenant la première au second (2003, p. 15) :

*« Markedness effects are due to transfer. They either result from the transfer of the L1 constraint ranking or from minimal constraint demotion (F demotion) within the transferred grammar. Markedness [...] is not an extralinguistic*

*variable [...]. The scientific focus should thus return to transfer and to basic operations performed on a transferred constraint ranking. The notion of an independent linguistic variable "markedness" is fundamentally misguided. »*

### I.3.3.2. Accès à UG, Transfert et Perception dans la Théorie de l'Optimalité

Travaillant sur la perception et l'acquisition de voyelles anglaises par des apprenants espagnols, et sur la base du concept de *crible phonologique* de Polivanov, Escudero et Boersma (2004) obtiennent des résultats qui, dans le cadre de la Théorie de l'Optimalité soutiennent deux hypothèses précédemment évoquées :

- Celle du *Transfert Plein* (« Full Transfer »), selon laquelle le système L1 est entièrement transféré dans l'interlangue comme base de développement (ce qui rejoint l'hypothèse de Uffmann, 2003).
- Celle de *l'Accès Plein* (« Full Access »), selon laquelle l'apprenant a par la suite pleinement accès aux principes de UG, comme lors de l'acquisition de sa L1.

Dans leur étude, Escudero et Boersma montrent qu'il est possible d'étendre ces résultats à la perception de la parole : le savoir à l'aide duquel le sujet établirait les correspondances entre les traits auditifs continus de la parole et les structures phonologiques discrètes, telles que segments et syllabes, pourrait ainsi être modélisé en termes de classement de contraintes. Ils estiment en effet que derrière le processus de perception se trouve une grammaire perceptive formelle, qui détermine un output optimal (une catégorie phonologique) sur la base d'un input donné (un événement auditif) (p. 11). Ils formalisent alors cette grammaire dans une version probabiliste du cadre de OT : d'après eux, les auditeurs créent des catégories (ex. /ɪ/ et /p/) sur la base d'informations distributionnelles, puis utilisent ces catégories pour créer des formes phonologiques dans leur lexique (ex. |ʃɪp|, ainsi que des contraintes de correspondance (« mapping constraints ») dans leur grammaire de perception (par exemple à propos de la longueur vocalique : « 74 millisecondes n'est pas /ɪ/ »), avant d'optimiser graduellement leur classement de contraintes en fonction des données

(reconnaissance dans le lexique) (p. 14). C'est donc bien essentiellement d'une grammaire de perception qu'il s'agit, mais qui s'intègre dans un modèle de production/perception phonologique plus général (note 4, p. 16). Ainsi, dans la perception de sons étrangers, l'utilisation des catégories de la L1 constitue un cas de transfert de *structure*, alors que l'utilisation des « correspondances » (« mappings ») constitue un cas de transfert de *processus*.

En conséquence, l'hypothèse du Transfert Plein est soutenue par les trois types de transfert abordés par les chercheurs (p. 23) :

- Transfert d'éléments lexicaux : réutilisation des catégories de la L1 dans l'état initial de l'interlangue. D'après les auteurs, cela est compatible avec les travaux de Flege (*Speech Learning Model*), qui considère qu'un apprenant ne dispose que d'un espace phonologique pour la L1 et la L2, dans lequel, à l'état initial, ne figurent que les catégories de la L1.
- Transfert de la grammaire : réutilisation des correspondances perceptives dans l'état initial de l'interlangue. D'après les auteurs, cela est compatible avec les travaux de Best (*Perceptual Assimilation Model*), qui considère que « high-level linguistic experiences determines what listeners do when confronted with the cross-linguistic task of having to classify foreign language sounds with L1 categories ».
- Transfert de « trous » dans la grammaire : réutilisation des « dimensions vierges » (« blank slates ») de la L1 dans l'état initial de la L2. Comme ils l'indiquent à propos des sujets espagnols, ceux-ci n'ont jamais eu à créer de catégories discrètes sur le continuum de durée vocalique. La perception d'un tel continuum est donc pour eux dans un état « vierge » (« blank slate »). Cela s'accorde avec l'idée que de *nouveaux sons* (c'est à dire des sons dans de nouvelles aires de l'espace phonétique) soient relativement faciles à acquérir.

En ce qui concerne l'hypothèse de l'Accès Plein à UG, celle-ci se trouve soutenue par deux types d'accès (p. 24) :

- Accès à UG et « dimensions vierges »: nouvelle catégorisation des « dimensions vierges ». Face à un état vierge comme la durée vocalique en espagnol, les sujets adoptent un apprentissage distributionnel similaire à celui adopté lors de l'acquisition de leur L1. Cela s'accompagne donc du transfert des « trous » dans la grammaire.
- Accès à UG et reclassement : reclassement des contraintes dans l'interlangue. Les sujets ont la possibilité de reclasser graduellement les contraintes de leur interlangue, comme c'est le cas dans l'acquisition de la L1.

Ainsi, à travers les études récentes de Uffmann (2003) d'une part, et celle de Escudero et Boersma (2004) d'autre part, il apparaît que le cadre OT permet de combiner : 1) la marque et le transfert, 2) le transfert et l'accès à UG et 3) la perception et la phonologie. Dans ce cadre, l'apprentissage de la phonologie d'une langue étrangère semble donc se dérouler en deux étapes : transfert plein, puis reclassement progressif des contraintes en fonction, d'une part des données disponibles dans la langue cible, mais également en fonction de contraintes universelles.

#### **I.4. Interphonologie : perspectives exemplaristes**

Jusqu'à présent, nous avons essentiellement abordé la question de la perception et de l'apprentissage de la phonologie dans une perspective cognitivement symbolique, que celle-ci soit représentationnelle ou non. L'approche, en partie « perceptive » et en partie « distributionnelle » de Escudero et Boersma (2004) ont cependant pu évoquer une autre approche de la question : il s'agit de l'approche exemplariste.

#### **I.4.1. Des épisodes phonétiques aux catégories phonologiques**

Comme cela est bien présenté par Coleman (2002), certains chercheurs considèrent en effet que les représentations formelles dans le lexique mental sont essentiellement phonétiques, plutôt que symboliques-phonologiques. Une telle version phonéto-acoustique des représentations renvoie alors à une perspective exemplariste/ épisodique (Kirchner, 2004) :

*« A speech processing system in which words are represented as cohorts of exemplars – detailed memories of the sound signal (and articulatory plan) for previous tokens with the same semantic labels. »*

De nouvelles techniques représentationnelles sembleraient en outre permettre aux modèles exemplaristes de se passer, dans une certaine mesure, d'unités phoniques ou phonologiques sublexicales (Kirchner, 2004) :

*« Rather, the (quasi-)phonemic organization of speech can be shown to emerge from the relatively stable correlations between articulatory gestures and perceptual cues over the lexicon of exemplars ».*

##### **I.4.1.1. Encodage épisodique acoustique de la parole**

A l'origine d'une telle approche, dans le domaine de l'apprentissage d'une L2, se trouvent plusieurs observations (Fisher, Church et Chambers, 2004), dont celles notées par Hardison (2000, p. 314) :

*« The significant effects produced by stimulus and talker variability in perceptual learning, and results of studies with native speakers of English demonstrating an interdependence in the processing of words and voices [...] suggest that the memory encoding of speech involves storage of detailed individual episodes or traces. [...] This contrasts with the traditional abstractionist view that perceptual idiosyncrasies of words are normalized,*



*with input matched against stored prototypical representations such as phonemes ».*

L'importance de l'encodage de ces « détails » représentationnels et contextuels est justifiée ainsi par Fisher, Church et Chambers (2004, p. 3) :

*« Detailed and context-sensitive representations of spoken words are needed, both to learn the native-language phonetics and phonology, and to learn to identify and compensate for variations in dialect, accent and social register ».*

Cependant, pour Hardison, le fait que l'encodage de la parole implique également celui de la variabilité contextuelle et des caractéristiques du locuteur (sexe, dialecte, etc.), à conditions que ces détails aient été remarqués (« attended details »), n'est pas en opposition avec le concept d'une représentation phonémique abstraite. La différence doit être abordée en termes de temporalité : de son point de vue, les catégories abstraites sont des agrégats d'exemplaires individuels activés simultanément au moment de leur récupération en mémoire, et qui peuvent donc être dérivés des informations épisodiques au moment de leur rappel, mais pas immédiatement abstraits au moment de l'encodage de la parole, tel que cela est postulé dans certaines théories phonologiques classiques. Les phonèmes ne sont donc pas « perçus », mais assemblés ultérieurement.

Dans une telle perspective, les facteurs attentionnels et affectifs jouent donc un rôle non-négligeable dans la perception (Hardison, 2000, p. 319) :

*« Recall that the central premise [...] is that all **attended** perceptual details are stored. In other words, the features preserved in traces represent those which drew the observer's attention and would thus be said to hold some significance for the task. [...] Focused attention to a stimulus sustains activation and facilitates its learning and memory. Thus the affective valence (pleasant-*

*unpleasant) and task significance of sensory stimuli influence further processing ».*

D'un point de vue didactique, une telle conception encourage la présentation répétée de stimuli, et ainsi d'exemplaires, variés, de manière à permettre aux sujets de concentrer leur attention sur les propriétés essentielles (auditives et/ou visuelles) des stimuli, même si un certain équilibre doit être trouvé. Elle cite ainsi (p. 321) Leather<sup>149</sup>, qui écrit :

*« Too much or too little variability at too early a stage may prevent the learner from discovering with sufficient accuracy the prototypical forms that exemplars expound ».*

Cette attention portée à la perception de « bas-niveau » se traduit également chez certains par une approche acoustique des stratégies perceptives des sujets apprenants, que l'on retrouve dans plusieurs cadres théoriques (Escudero et Boersma, 2004). Ainsi, lorsque l'on s'intéresse à la perception des consonnes anglaises américaines [l] et [ɹ] par des sujets japonais, on peut en effet proposer une caractérisation acoustique de ces consonnes et ainsi des stratégies perceptives des sujets. Comme le note Hardison (2003, p. 496), le [ɹ] se distingue en particulier de [l] et de [w] par un formant F3 de basse fréquence. Les Japonais présentent alors de bonnes performances de discrimination en ce qui concerne ce formant sur des sons isolés non-linguistiques. En revanche, lorsque celui-ci est combiné avec les formants F1 et F2 pour produire les phonèmes /l/ et /r/ en contexte, les performances sont moins bonnes, ce qui serait dû au fait, d'après certaines expériences, que les Japonais se concentrent sur le formant F2 pour essayer de distinguer les deux phonèmes.

---

<sup>149</sup> Leather, J., « Perceptual and productive learning of Chinese lexical tone by Dutch and English speakers », in Leather, J. et James, A. (Eds.), *New Sounds 90*, Amsterdam, University of Amsterdam, 1990, pp. 72-97.

C'est ce que présentent Iverson, Kuhl, Akahane-Yamada, Diesch, Tohkura, Kettermann et Siebert (2003) qui se sont intéressés à la perception des phonèmes anglais /r/ et /l/ par des sujets adultes japonais, allemands et américains. Leur hypothèse était que les processus perceptifs de bas-niveau, modifiés par l'expérience linguistique précoce des sujets dans leur langue première, pouvaient conduire à des interférences avec la formation et l'adaptabilité de représentations linguistiques de haut-niveau. D'après leurs résultats, alors que les sujets américains étaient sensibles aux différences dans le formant F3 qui distingue /r/ et /l/, les sujets japonais étaient plus sensibles aux variations du formant F2. Si l'insensibilité des Japonais aux variations de F3 était déjà connue, cette étude démontre en revanche que ceux-ci ne sont pas totalement incapables de discerner les deux phonèmes (p. B53) :

*« Japanese adults are more sensitive to variation along a dimension, F2, that is mostly irrelevant to the English /r/-/l/ categorization, and can have high within-category sensitivity to F3 differences, particularly for /l/. Japanese adults thus have a distorted perceptual space for these phonemes, but not a total lack of perceptual sensitivity (i.e. the critical acoustic variation is not filtered out) ».*

En s'appuyant sur des indices acoustiques perceptivement saillants, mais non robustes pour la catégorisation, comme F2, les Japonais pourraient former des représentations catégorielles erronées de /r/ et de /l/. Les auteurs rejoignent alors Hardison sur l'intérêt, pour l'apprentissage, d'un entraînement perceptif à haute variabilité (p. B54) :

*« Training procedures involving smaller stimulus sets are easier to learn, but do not readily generalize to new stimuli [...]. Training with larger stimulus sets may generalize better, because the variability provides information about which cues are more robust and trains individuals to ignore irrelevant variation ».*

#### 1.4.1.2. Variation allophonique et acquisition des catégories phonémiques

Un autre versant de la dimension phonétique de la perception des catégories phonémiques concerne la perception des allophones des phonèmes. Peperkamp, Pettinato et Dupoux (2003) se sont ainsi penchés sur la perception des contrastes allophoniques par des sujets adultes. Comme ils le rappellent, dans la perspective de Kuhl et ses collègues, l'acquisition des catégories phonémiques par les enfants s'effectuerait *via* une analyse statistique de l'espace acoustique (Boersma, Escudero et Hayes, 2003) et la formation de prototypes. Etant donné que les allophones sont réalisés comme des prototypes acoustiques distincts, les enfants anglais ou coréens, par exemple, établiraient ainsi trois catégories distinctes [t], [d] et [t<sup>h</sup>], sans distinguer les contrastes phonémiques et allophoniques (p. 650). Les adultes, en revanche, sembleraient avoir acquis, à un moment donné, la distinction entre les deux types de contrastes. Les auteurs explorent alors l'hypothèse que cette dernière acquisition s'effectue sur la base d'une analyse distributionnelle. Ils présentent donc deux expériences en français :

- La première concerne la simple perception des deux types de contraste : d'une part la distinction phonémique /m/ - /n/, et d'autre part la distinction allophonique [ɞ] (fricative uvulaire voisée) – [χ] (fricative uvulaire non-voisée, que l'on trouve après ou avant une consonne non-voisée), tant en contexte que de manière isolée. Les résultats indiquent que, hors-contexte, la distinction allophonique est clairement perçue, alors que, en contexte, elle l'est beaucoup moins. L'interprétation favorisée par les chercheurs, au vu de leurs résultats est la suivante : les allophones seraient représentés comme une unique catégorie phonologique, leur discrimination étant alors similaire à celle de contrastes non-natifs, comme le /r/ et le /l/ pour les sujets japonais, l'isolation hors-contexte permettant d'accéder plus facilement aux corrélats phonético-acoustiques du contraste.

- La seconde expérience vise à déterminer si les sujets sont capables de construire des catégories phonémiques abstraites sur la base d'informations distributionnelles, en utilisant un continuum entre [ʁ] et [χ] et en contrôlant leur distribution statistique de manière artificielle : une distribution bimodale (de nombreuses valeurs aux extrêmes, mais peu au milieu du continuum) conduirait à l'existence de deux catégories, alors qu'une distribution unimodale (peu de valeurs aux extrêmes, mais de nombreuses valeurs au milieu du continuum) conduirait à une unique catégorie. Les résultats indiquent que, outre le mécanisme d'extraction de prototype local, l'apprentissage statistique prend en compte le contexte de réalisation, étant sensible, en particulier, à la distribution complémentaire. Les résultats de cette étude ne sont cependant pas suffisamment conclusifs.

Mais, comme le notent Chambers, Onishi et Fisher (2004), dans une perspective similaire :

*« A bimodal distribution of phonetic values does not always correspond to a phonemic contrast for speakers of that language. The key to determining whether a phonetic difference is contrastive or allophonic is whether values on the relevant phonetic dimension are bimodally distributed in the same phonological environment ».*

Les auteurs ont donc élaboré un protocole relativement similaire à celui de Peperkamp, Pettinato et Dupoux (2003), en présentant un continuum phonétique entre deux fricatives uvulaires (l'une voisée, notée « gh » ; l'autre dévoisée, notée « x ») produites par un locuteur natif arabophone à des sujets adultes anglais. Outre cette variation phonétique, les auteurs ont distingué deux positions phonologiques syllabiques, attaque et coda, conduisant à des stimuli du type : « ghik », « xik », « bigh » et « bix ». L'hypothèse des auteurs était la suivante :

*« If subjects separately kept track of the phonetic distributions in different phonological environment (onset and coda), then for either onset or coda position, subjects who received the bimodal training in that syllable position should respond « different » more often than those who received unimodal training in the same position ».*

Leurs résultats semblent confirmer cette hypothèse, indiquant que les adultes sont capables d'acquérir et de maintenir des distributions séparées des valeurs phonétiques pour différents contextes phonologiques, et ceci pour des contrastes non-natifs. Un tel résultat soutient l'idée que l'apprentissage distributionnel joue un rôle important dans le développement phonologique (Chambers, Onishi et Fisher, 2004), une idée également soutenue par Peperkamp et Dupoux (2004). Contrairement à l'idée selon laquelle les enfants acquerraient leurs catégories phonémiques *via* des oppositions sémantiques entre paires minimales, ce qui nécessiterait un lexique assez large, Peperkamp et Dupoux (2004) supposent que ceux-ci construisent ces catégories de manière pré-lexicale à l'aide d'informations distributionnelles. Comparant ainsi les deux approches à l'aide de tâches d'apprentissage artificiel auprès de sujets adulte français, leur conclusion est la suivante :

*« Based on information regarding the distribution of segments, learners can create abstract phoneme categories that encompass different segments ; lexical knowledge may be helpful but is not required, and the presence of minimal pair is irrelevant ».*

L'importance de l'information distributionnelle est également soulignée par Boersma, Escudero et Hayes (2003) qui l'intègrent dans leur modèle d'apprentissage des catégories phonologiques, dans le cadre de la théorie de l'Optimalité. Celui-ci, qui concerne l'émergence des catégories phonologiques chez l'enfant, comporte deux étapes. La première est dirigée par l'input auditif (« auditory-driven learning »), et conduit, notamment *via* l'effet d'attracteur perceptif de Kuhl, à la création de catégories phonétiques. La seconde est dirigée par le lexique (« lexically-driven

learning ») et, grâce au développement de représentations abstraites et à l'optimisation de la perception, conduit à la transformation de ces catégories phonétiques en catégories phonologiques. Le modèle qu'ils proposent comporte donc un mécanisme sous-jacent aux deux types d'apprentissage ainsi que l'articulation entre les deux, *via* un processus d'apprentissage perceptif graduel.

Tous ces travaux semblent donc relativement convergents, en insistant sur le rôle des informations statistiques dans le passage de l'acquisition de catégories phonétiques à des catégories phonologiques plus abstraites. Si l'acquisition de telles catégories par des enfants diffère de celle d'adultes apprenants de langue étrangère, il n'en reste pas moins que l'information distributionnelle et la distinction entre catégories phonétiques et phonologiques jouent un rôle non négligeable dans le développement de l'interphonologie de ces derniers. Une telle perspective permet par ailleurs de relier, dans une certaine mesure, les approches exemplariste et abstractionniste de l'apprentissage, et d'articuler ainsi l'apprentissage phonétique et l'apprentissage phonologique.

#### **I.4.2. Mémoire phonologique et émergentisme**

Dans le domaine de l'acquisition des langues secondes (en anglais « second language acquisition », SLA), l'importance de l'information statistique, liée à la fréquence et à la distribution des items concernés dans l'input linguistique, dans l'apprentissage est également un thème d'actualité.

N. Ellis, par exemple, considère que l'acquisition d'une langue revient en fait essentiellement à un apprentissage de séquences (1996) : pour le vocabulaire, celui des unités phonologiques et leur phonotactique ; pour le discours, celui des unités lexicales et leurs collocations. Les connaissances grammaticales abstraites résulteraient alors de l'analyse de ces séquences, stockées en mémoire à long-terme et servant de base de données pour l'analyse. Les différences entre apprenants de langue étrangère reposeraient alors sur les différences de capacité à opérer cette rétention de séquences, soulignant ainsi le rôle de la mémoire phonologique à court-terme : les individus

diffèrent en effet dans leur capacité à répéter les séquences phonologiques (mémoire phonologique à court-terme) et ces différences déterminent la facilité avec laquelle les apprenants peuvent acquérir le vocabulaire et la grammaire en langue étrangère (p. 92), ce qui renvoie aux travaux de Baddeley et ses collègues.

C'est donc une perspective exemplariste/émergentiste (MacWhinney, 2001) qu'adopte N. Ellis (2002), qui considère que les régularités de la langue émergent de l'expérience en tant que catégories et schémas prototypiques. Au cours de ce processus, l'influence de la fréquence n'est pas neutre, comme cela est illustré dans un numéro spécial de la revue *Studies in Second Language Acquisition*, consacré au thème « Frequency Effects in Language Processing and Acquisition » (2002), sur les plans phonologique, orthographique, morphosyntaxique et lexical. Passant en revue de très nombreuses études psycholinguistiques relatives aux effets de fréquence dans le traitement et l'acquisition du langage, N. Ellis (2002, p. 166) justifie ainsi sa position :

*« These psycholinguistic demonstrations that frequency sensitivity pervades all aspects of language processing have profound implications for theories of language acquisition – namely, that language learning is exemplar based. [...] [T]he knowledge underlying fluent use of language is not grammar in the sense of abstract rules or structures but a huge collection of memories of previously experienced utterances. [...] Linguistic regularities emerge as central tendencies in the conspiracy of the database of memories of utterances. »*

Dans une telle perspective, une grande partie du travail d'apprentissage linguistique revient à graduellement renforcer les associations entre éléments co-occurents de cette langue, afin de fournir une connaissance probabiliste nécessaire à une utilisation performante de la langue (p. 173). L'apprentissage implicite semble donc au cœur de l'apprentissage linguistique, même si N. Ellis prend soin par la suite d'apporter certaines précisions, relatives, notamment, au rôle de l'attention et de l'identification consciente d'une part, et à celui de l'instruction explicite d'autre part (p. 173) :



*« Because the conscious experiences of language do not revolve around counting, to the extent that language processing is based on frequency and probabilistic knowledge, language learning is implicit learning. »*

Cependant Ellis se défend de soutenir un retour à un certain néo-comportementalisme : certains pourraient en effet facilement assimiler ses positions à celles se trouvant à l'origine de la méthode Audio-orale (« Audio-Lingual Method ») de Lado, en établissant un lien abusivement simplificateur entre fréquence et régularité d'une part, et répétition « simpliste » d'autre part. Comme le note N. Ellis (2002, p. 177) :

*« Language learning involves much more than mindless repetition. **Mindful** repetition in an engaging communicative context by motivated learners is somewhat closer to the mark. »*

Ainsi, les grandes perspectives théoriques relatives au développement de l'interphonologie que nous avons examinées jusqu'à présent, nous ont progressivement fait passer du domaine strictement phonologique à celui, psycholinguistique, de l'acquisition des langues secondes. Avant de prolonger cette extension en considérant le rôle de l'orthographe dans ledit développement, nous pouvons à présent formuler certaines problématiques précises dans les cas d'apprenants japonais de français, en adoptant pour l'instant une perspective essentiellement phonético-phonologique.

## **II. Investigations phonologiques : l'interphonologie des apprenants japonais de FLE**

D'un point de vue strictement linguistique, il apparaît ainsi que les locuteurs japonais apprenant la phonologie du français vont être confrontés à plusieurs problèmes, résultant essentiellement des divergences entre les systèmes phonologiques et phonographémiques français et japonais. Dans le cadre d'une étude comme la nôtre, il

serait trop ambitieux de vouloir traiter l'ensemble des phénomènes en jeu. Nous avons donc sélectionné deux pistes d'investigation :

- L'une, *spécifique*, au niveau *segmental*, concerne le traitement des unités phonématiques françaises /b/ et /v/ d'une part, /l/ et /R/ d'autre part, et, dans une certaine mesure, /EU/ et /OU/.
- L'autre, de portée plus *générale*, au niveau *syllabique*, concerne le traitement des structures syllabiques complexes françaises biconsonantiques.

Ce choix se justifie, d'une part par la comparaison théorique des systèmes phonologiques en présence, et d'autre part par l'observation empirique des productions des sujets :

Niveau	Portée	Comparaison Théorique		Observations empiriques
		Français	Japonais	
Segmental	Spécifique	/b/	/b/	Substitutions /v/ ← → /b/
		/v/	Pas d'« équivalent »	
		/l/	/r/	Substitutions /R/ ← → /l/
		/R/	Pas d'« équivalent »	
		/OU/	/U/	Substitutions /EU/ ← → /OU/
		/EU/	Pas d'« équivalent »	
Syllabique	Générale	(C)(C)V(C)(C)	(C)(j)V(N) ou (C)(j)V(Q)	Epenthèses vocaliques et/ou effacement consonantique

### **II.1. Aspects segmentaux**

D'un point de vue *phonético-phonologique*, nous avons décidé de récolter certaines informations quantitatives sur les substitutions, phonémiques ou allophoniques, des segments concernés, préalables à notre analyse des phénomènes en jeu. Ces informations correspondent à deux questions assez générales, qui s'articulent sur les axes classiques syntagmatique et paradigmaticque :

(A) *la position du segment cible dans le mot phonologique influence-t-elle le taux de confusion segmentale ?*

(B) *la nature du segment cible dans le mot phonologique influence-t-elle le taux de confusion segmentale ?*

Concernant les productions observées, il convient de distinguer les productions licites (1), les substitutions d'ordre phonético-phonologique et/ou phonographémique (2) et finalement les substitutions qui relèvent de stratégies de production ou d'apprentissage (3), étant donné la situation d'apprentissage des sujets.

(1)	<ul style="list-style-type: none"> <li>• /b/ réalisé [b] (plosive bilabiale) ;</li> <li>• /l/ réalisé [l] (latérale apico-alvéolaire, allophone du /r/ japonais).</li> </ul>
(2)	<ul style="list-style-type: none"> <li>• /b/ réalisé [β] (fricative bilabiale, allophone du /b/ japonais, en général en contexte intervocalique, Akamatsu, 2001, p.75) ;</li> <li>• /v/ réalisé [b] (plosive bilabiale) ou [β] (fricative bilabiale) ;</li> <li>• /l/ réalisé [ɾ] (tap apico-alvéolaire, réalisation prototypique du /r/ japonais) ou [d] (plosive alvéolaire) (allophones du /r/ japonais, Labrune, 2003, p. 2) ;</li> <li>• /R/ réalisé [ɾ] (tap apico-alvéolaire, réalisation prototypique du /r/ japonais) ou [l] (latérale apico-alvéolaire) ;</li> <li>• /OU/ réalisé [ɯ] ;</li> <li>• /EU/ réalisé [ɯ] ou [o]<sup>150</sup>.</li> </ul>
(3)	<ul style="list-style-type: none"> <li>• /b/ réalisé [v] (fricative labio-dentale) ;</li> <li>• /l/ réalisé [R] (liquide uvulaire) ;</li> <li>• /OU/ réalisé [∅].</li> </ul>

<sup>150</sup> A ce propos, notons la transcription en hiragana de notre propre nom (prononcé [døte] en français) sur les documents officiels japonais, que nous donnons ici en romaji: DO-u-TE. On voit alors bien ici que le /EU/ est transcrit, puis réalisé, comme un combinaison entre le [o] et le [ɯ].

Nous ne nous intéressons, dans cette partie, qu'aux cas (2), en particulier aux réalisations de /v/, /R/ et, dans une moindre mesure, /EU/. Les cas (3) seront traités ultérieurement.

### **II.1.1. Axe syntagmatique : la position du segment cible dans le mot phonologique influence-t-elle le taux de confusion segmentale ?**

D'un point de vue syntagmatique, la notion de *position* que nous évoquons ici, correspondant à la position temporelle (« temporal slot ») sur le squelette phonologique, est importante, puisque dans une architecture non-linéaire des représentations phonologiques, elle est à relier à d'autres plans de représentation, en particulier au plan syllabique et au plan accentuel. Par ailleurs la distribution des segments sur l'axe syntagmatique renvoie à des problèmes de contraintes phonotactiques et de domaines phonologiques.

POSITION				
Domaine	Mot Phonologique (#-#)		Syllabe ((C)V(C))	
<b>Phonotactique</b>	<b>Frontière</b> (#CV(C) ou (C)VC#)	<b>Interne</b> ((C)VCV(C))	<b>Attaque</b> (CV(C))	<b>Coda</b> ((C)VC)
<b>Accentuel</b>	<b>Syllabe faible</b> (#CV(C))	<b>Syllabe forte</b> ((C)´V(C)#)		

Le terme de *confusion*, quant à lui, a été choisi eu égard au protocole de récolte des données, afin de ne pas préjuger des phénomènes à l'œuvre. Il nous permet, dans un premier temps, de rester relativement neutre, d'une part vis-à-vis des niveaux d'analyse (phonétique ou phonologique), et d'autre part, dans une perspective psycholinguistique, vis-à-vis des pôles perception/production.

Cette « confusion » peut ensuite être analysée de plusieurs manières :

- Dans un cadre phonémique : la confusion impliquerait, sur la base du concept de « crible phonologique », d'une part l'appel à la notion de « similarité », à base essentiellement phonétique/acoustique, et d'autre part celui au concept de « transfert ».
- Dans un cadre infra-segmental : la confusion pourrait être interprétée en terme de substitutions ou de modifications de traits. Les segments à l'étude peuvent en effet être, partiellement au moins, opposés de la manière suivante :
  - /b/ : [-cont] vs /v/ : [+cont]
  - /l/ : [+ cor] vs /R/ : [- cor]
  - /OU/ : [- ant] [+ haut] vs /EU/ : [+ ant] [- haut]
- Dans le cadre OT : la confusion pourrait correspondre au résultat de la violation de plusieurs contraintes, notamment de la contrainte Ident-F (« every segment in the output has identical values for a feature F as the corresponding segment in the input ») et refléter ainsi des différences de classement de contraintes entre le français, le japonais et l'interlangue des sujets.
- Dans le cadre TCRS : la confusion pourrait alors être interprétée en termes de « réparation », ou d'adaptation, selon le schéma : mauvaise formation/violation de contrainte(s) → réparation(s).

L'établissement de certaines corrélations entre *position* et *confusion* nous conduirait dans un premier temps, d'une part à apporter des informations descriptives sur le système interlangagier des sujets, et d'autre part à proposer, sur une base empirique quasi-expérimentale, un classement des contraintes pertinentes, d'ordre phonotactique, propre à l'interlangue des sujets japonais apprenant le français. A titre d'exemple, on pourrait imaginer que les taux de confusion pour les segments /l/ et /R/ soient beaucoup

plus élevés en coda qu'en attaque syllabique. Dans ce cas, on pourrait proposer le classement suivant : \*Coda [Liqu] >> \*Onset[ Liqu].

## **II.1.2. Axe paradigmatique : la nature du segment cible dans le mot phonologique influence-t-elle le taux de confusion segmentale ?**

### *II.1.2.1. Segments consonantiques*

D'un point de vue paradigmatique, on peut en effet s'interroger sur le caractère plus ou moins « problématique » des segments à l'étude (/v, R/), ce qui revient, ici encore, à établir un classement de contraintes pour l'interlangue des sujets : certains segments (par exemple /R/, Sonorante) sont-ils plus illicites, c'est-à-dire plus *marqués*, que d'autres (par exemple /v/, Obstruante) dans l'interlangue des sujets ? Cela correspondrait au classement suivant : \*/R/ >> \*/v/. Si oui, l'explication réside-elle dans le domaine universel (par exemple un segment plus marqué que l'autre sur une échelle de marque universelle, relevant d'un classement initial par défaut de certaines contraintes de marque) ou spécifique (classement spécifique au système interphonologique de certaines contraintes de marque) ? S'il s'agit du domaine universel, le classement correspond-il à celui du japonais (L1) ou bien correspond-il à un cas d' « émergence du non-marqué » ?

A propos des substitutions consonantiques en langue étrangère, on peut initialement s'inspirer de la démarche de Lombardi (2000). Bien que notre propre étude diffère de celle de Lombardi, les réponses à nos questions peuvent se trouver en partie éclairées par l'analyse qu'elle propose, notamment à propos du japonais, puisque dans le cas du /v/ (fricative labio-dentale) on peut considérer que, en L1, les deux substituts possibles sont [b] (plosive bilabiale) et [β] (fricative bilabiale), tandis que dans le cas du /R/ (fricative/approximante liquide vibrante uvulaire), les deux substituts possibles sont [r] (tap apico-alvéolaire) et [l] (fricative/approximante liquide latérale apico-alvéolaire). Nous pouvons alors poser comme elle certaines contraintes (\* /v/ ; \* /R/) et garder la même hiérarchie universelle de marque :

\*[β, l] >> \*[b, r]

Si l'on adapte les tableaux qu'elle propose pour le japonais aux segments qui nous intéressent, alors :

– Dans le cas de la fricative labiodentale /v/ :

/v/	*/v/	IdentStop	IdentCont
<b>V</b>	* !		
<b>B</b>			* !
+ β			

– Dans le cas de la liquide uvulaire /R/

/R/	*R	IdentStop	IdentCont
<b>R</b>	* !		
r			* !
+ l			

Il est alors intéressant de constater que nos observations empiriques informelles ne correspondent pas systématiquement à ces prédictions théoriques, puisque, bien souvent, le [v] est remplacé par la plosive [b] et le [R] par le tap [r]. Bien sûr, si l'on distingue les plans phonémique et allophonique, et si l'on considère que /b/ et /β/ d'une part, et /r/ et /l/ d'autre part, ne correspondent en fait qu'à deux entités abstraites uniques, la question perd de sa pertinence, du point de vue que nous avons adopté jusqu'à présent. Cependant, s'il s'agit bien de réalisations allophoniques en japonais, on ne peut *a priori* en dire autant pour l'interlangue des sujets : en ce qui concerne le /v/, s'il devait être systématiquement remplacé par le /β/, alors on pourrait attribuer au

/β/ un statut phonémique propre dans le système interlangagier des apprenants, en opposition notable au /b/. En ce qui concerne le /R/, s'il devait être systématiquement remplacé par le /l/, alors le système interlangagier serait réduit par rapport au système français, puisque le /l/ y possède déjà un statut phonémique propre. Dans ce cas, on pourrait effectivement peut-être considérer que les réalisations [l] et [r] ne sont que des allophones d'une même unité phonémique, même si on peut douter d'une telle interprétation puisque, en japonais, [l] n'est qu'un allophone secondaire par rapport au [r] qui est la réalisation prototypique du /r/ japonais. C'est donc précisément ce problème qui nous a conduit plus bas à intégrer certains aspects psycholinguistiques complémentaires à notre étude, en particulier les aspects phonographémiques, afin d'éclairer ce dont une analyse strictement phonologique semble parfois bien en peine de rendre compte lorsqu'il s'agit d'apprentissage de la phonologie d'une langue étrangère ou seconde.

#### II.1.2.2. Segment vocalique

En ce qui concerne le segment vocalique /EU/, la question diffère légèrement étant donné notre protocole d'investigation qui implique au moins deux variables combinées : le segment consonantique (/b, v, R, l) et le segment vocalique (OU, EU). L'approche du /EU/ passe donc par une perspective phonotactique/syllabique, puisqu'il s'agit de déterminer si certains couples (par exemple /b+EU/ et /v+EU/) sont plus illicites/marqués que d'autres (par exemple respectivement /b+OU/ et /v+OU/). Cette approche indirecte devrait nous permettre de proposer un classement de contraintes relatives à la fois aux combinaisons segmentales et aux segments /EU/ et /OU/. Plusieurs classements pourraient être envisageables, par exemple :

- (a) \*Cont+/EU/ >> \*Plos+/EU/ >> \*Cont+/OU/ >> \*Plos+/OU/
- (b) \*Cont+/EU/ >> \*Cont+/OU/ >> \*Plos+/EU/ >> \*Plos+/OU/



Les plosives étant moins marquées que les fricatives (Lombardi, 2000), on pourrait en effet se demander si le classement de contraintes relatives aux segments consonantiques, en attaque ou en coda syllabique, prime sur celui qui concerne les segments vocaliques, ou inversement.

Dans le cadre TCRS, l'analyse peut notamment s'effectuer via celle du poids procédural des réparations (Paradis et Béland, 2002, p. 194), c'est-à-dire le nombre d'opérations de réparation nécessaires, fonction du nombre de malformations (une (C ou V) ou deux (C et V)) et le domaine auquel elles s'appliquent (segmental ou syllabique). En ce qui concerne les réparations relatives aux codas, par exemple, deux types sont possibles : délétion consonantique (1 réparation), épenthèse vocalique (2 réparations : insertion nucléique et remplissage segmental). Dans ce cas, nous obtenons le tableau prédictif suivant :

Nombre de malformations						
Domaine	Nul	Un	Deux	Total		
Segmental	BOU LOU ----- BA BI LA LI	VOU/OUV BEU/EUB ROU/OUR LEU/EUL ----- VA/AV VI/IV RA/AR RI/IR	VEU/EUV REU/EUR	Segments	Nombre	Code
				BA BI LA LI	<b>NUL</b>	<b>00</b>
Syllabique	BOU LOU	AV IV	Aucun	BOU LOU		
	BA BI LA LI	OUV EUV AB IB		VA VI RA RI VOU BEU ROU LEU	<b>UN (segment)</b>	<b>01</b>
	VOU BEU ROU LEU VA VI RA RI VEU REU	OUB EUB AL IL OUL EUL AR IR OUR EUR		AB IB OUB AL IL OUL	<b>UN (syllabe)</b>	<b>10</b>
				VEU REU	<b>DEUX (segment)</b>	<b>02</b>
				AV IV OUV AR IR OUR	<b>DEUX (segment et syllabe)</b>	<b>11</b>
				EUV EUR	<b>TROIS (2 segments et 1 syllabe)</b>	<b>12</b>

Si, dans le cas des adaptations segmentales qui nous concernent, le nombre de réparation correspond au nombre de malformations, ce n'est pas le cas dans le domaine syllabique :

- Pour les cas (10) : étant donné le *Principe de Préservation*, l'insertion (poids=2) devrait primer sur l'effacement (poids=1) ;
- Pour les cas (11) : le segment nécessitant adaptation, l'effacement (poids=1) devrait primer sur l'insertion (poids=2), étant donné que celle-ci nécessiterait l'adaptation du segment consonantique en attaque ( $2+1=3$ ) ;
- Pour les cas (12) : comme pour les cas (11), l'effacement devrait primer sur l'insertion, pour des raisons similaires.

Lorsque certains stimuli présentent des malformations segmentales et syllabiques, étant donné la *Convention de Précédence*, les contraintes syllabiques s'appliqueront en premier. Si l'on considère que, dans le cas des codas, illicites en japonais, les sujets adopteront une stratégie de délétion, plus économe que l'épenthèse en nombre d'opérations, alors les séquences /EUV/ et /EUR/ ne nécessiteront en fait plus que deux opérations : effacement consonantique (domaine syllabique) et adaptation vocalique (domaine segmental). On pourrait donc s'attendre à ce qu'il n'y ait pas de différence de traitement entre les stimuli de types (11) et (12).

### II.1.2.3. Confusion phonétique ou phonologique ?

L'un des points d'achoppement de notre problématique, nous semble-t-il, est toujours celui de l'écart entre statut *phonétique* et statut *phonologique* des segments considérés, en français, en japonais, et dans l'interlangue des sujets. Les récents travaux en « phonologie perceptive » (Fleischhacker, 2000 ; Côté, 2000 ; Escudero, 2004), s'ils ont le mérite de réintégrer la perception (et ainsi la psycholinguistique) dans les analyses phonologiques, ne facilitent pas la clarification de ce point fondamental, toujours débattu. Comme nous l'avons vu, certains soutiennent même que les

représentations formelles dans le lexique mental sont essentiellement phonétiques, plutôt que symboliques-phonologiques (Coleman, 2002).

Si l'on reprend la taxonomie des erreurs segmentales de Moulton, cité par Odlin (1989, p. 116), on peut considérer que les segments français sont généralement traités de la manière suivante :

- R → erreur phonémique
- l → erreur phonétique
- b → erreur allophonique
- v → erreur phonémique
- OU → erreur phonétique
- EU → erreur phonémique

Si l'on prend l'exemple du /R/ français et du /r/ japonais, une approche laissant davantage place aux considérations perceptives, et ainsi phonétiques, modifie en effet le « grain » de l'analyse, puisque, d'après l'*Alphabet Phonétique International* (API) de 1989 (reproduit dans Catford, 1994, Appendice), on a :

[ʁ] = fricative uvulaire ; [R] = trill uvulaire ; [l] = latérale approximante.

Catford (1994, p. 70) note d'ailleurs qu'en français le « /r/ » est souvent une fricative uvulaire ou une approximante uvulaire (la fricative laisse place à l'approximante lorsque le canal articulaire s'ouvre légèrement davantage (Catford, 1994, p. 99)). Dans les deux cas, il s'agit, pour ces deux uvulaires, d'articulations maintenables (« maintainable stricture types »). En revanche les taps, comme le /r/ japonais, sont essentiellement des articulations momentanées et donc non maintenables. La principale différence entre la fricative et le trill uvulaires est présentée comme telle par Catford (1994, p. 99) :

*« Whereas the tongue back is relatively flat or convex for [the fricative], a deep longitudinal groove is formed in the back of the tongue for the trill [...], and the uvula vibrates in the groove ».*

Nous avons d'ailleurs observé que, chez les apprenants japonais de français, la production de l'approximante est plus difficile à obtenir que celle de la fricative et l'ordre d'acquisition que nous avons généralement constaté correspond globalement souvent à celui-ci : Tap apico-alvéolaire (japonais) [ɾ] → Fricative uvulaire non-voisée [χ] → Trill uvulaire [ʀ] → Fricative uvulaire voisée [ʁ] → Approximante uvulaire [ʁ̥]. Pourrait-on y voir l'évolution d'un apprentissage phonologique à un apprentissage phonétique ? Ou l'inverse ? Cette évolution ressemble fortement à celle constatée par Colantoni et Steele (2005), dont les apprenants anglophones, visant également l'apprentissage du segment français /R/, maîtrisaient d'abord la « frication », puis graduellement le voisement, en premier lieu en contexte intervocalique, où il est favorisé articulatoirement.

Ainsi, lorsque l'on reprend la divergence que nous avons constatée à propos des segments /b/ et /R/ entre prédiction théorique et observations empiriques, on peut se demander si cette divergence ne se réduit finalement pas à un simple écart d'analyse entre niveaux phonétique et phonologique : la valeur de la distinction b/β et r/l est-elle vraiment une distinction phonologique, ou bien une simple distinction phonétique ? D'un point de vue infra-segmental, cette question renvoie à la distinction entre traits phonétiques et traits phonologiques (Durand, 2000a). Cela implique également de définir précisément le système phonologique auquel on se réfère : français, japonais ou interlangagier.

<b>Unité</b>	<b>Japonais</b>	<b>Interlangue</b>	<b>Français</b>
<b>b</b>	phonème (réalisation prototypique)	?	phonème (réalisation prototypique)
<b>v</b>	Nul	?	phonème (réalisation prototypique)
<b>β</b>	allophone	?	nul
<b>R</b>	Nul	?	phonème (réalisation prototypique)
<b>l</b>	allophone	?	phonème (réalisation prototypique)
<b>r</b>	phonème (réalisation prototypique)	?	nul

Comme on peut le voir, si la question clé de l'apprentissage phonologique en termes de contraintes est celle du reclassement des contraintes, une approche représentationnelle plus classique nécessite de déterminer le statut, phonémique ou allophonique, pertinent ou non pertinent, des (nouvelles) unités constituant le système interlangagier.

Dans tous les cas, comme nous l'avons déjà indiqué, il semble que, dans le cas de sujets japonais apprenant le français, une analyse strictement phonologique, ou phonético-phonologique, ne suffise pas, même lorsqu'elle prend en compte certaines contraintes d'ordre phonotactique, en ce qu'elle néglige certains facteurs d'ordre psycholinguistique, en particulier celui des représentations phonographémiques. C'est précisément l'une des problématiques centrales de la présente étude.

## **II.2. Aspects syllabiques**

D'un point de vue strictement *phonologique*, c'est sur la question du traitement syllabique, et celle des stratégies épenthétiques qui l'accompagnent, que nous nous sommes penchés, car de ce traitement peuvent dépendre, en partie au moins, certains

aspects rythmiques, intonatifs et accentuels des énoncés produits en français, voire même certains aspects segmentaux, *via* le lien entre contraintes syllabiques et phonotactiques. L'apprentissage de la structure syllabique du français, et plus particulièrement des syllabes complexes biconsonantiques peut être abordé de plusieurs manières.

D'un point de vue *représentationnel*, on peut s'interroger, entre autres, sur la nature de la syllabe en français, et sur l'influence que celle-ci peut avoir sur l'apprentissage de la syllabation du français. On peut par exemple supposer que le traitement d'une syllabe de type CCV et celui d'une syllabe de type CəCV avec effacement ne poseront pas les mêmes difficultés. A ce propos, nous suivrons dans un premier temps l'analyse de Dell (1995) sur les groupes consonantiques et les syllabes phonologiques en français. On peut également se demander si, en termes de syllabification, les syllabes de type SCV ne correspondent pas en fait à des syllabes de type S-CV, avec un premier segment extra-syllabique, ce qui, ici encore, pourrait modifier l'analyse de leur traitement par les sujets japonais.

D'un point de vue *non-représentationnel*, on peut d'une part s'interroger sur le reclassement des contraintes syllabiques à l'œuvre en japonais, et d'autre part sur les facteurs linguistiques déterminant les stratégies épenthétiques, comme cela est illustré par exemple dans la distinction entre les groupes Obstruantes + Sonorantes d'une part et Sibilantes + Plosives d'autre part. Notons que, si l'on peut puiser dans les travaux en phonologie des emprunts, nous partageons cependant pleinement le point de vue de Grijzenhout et Van Rooy (2001, p. 6) lorsqu'ils indiquent :

*« The difference between loan-word phonology and second language phonology is that loan words are adjusted to the first language grammar, whereas it is the grammar that has to be adjusted when learning a second language. Under OT, the task of the second language learner is to establish the constraint ranking of the target language ».*

En fonction de ce classement, confronté à un mot comme /psa.mo.te/ (3 syllabes) par exemple, le sujet pourrait en effet :

– <u>Effacer l'une des consonnes du groupe :</u>	→ /sa.mo.te/ (3 syllabes) ou → /pa.mo.te/ (3 syllabes)
– <u>Insérer une voyelle épenthétique entre les consonnes :</u>	→ /pu.sa.mo.te/ (4 syllabes)
– <u>Insérer une voyelle avant les consonnes :</u>	→ /ep.sa.mo.te/ (4 syllabes) (avec violation de *Coda)
– <u>Etre fidèle à l'input :</u>	→ /psa.mo.te/ (3 syllabes)

Pour des mots comme /motesafl/, les candidats seraient [mo.te.sa.flu.], [mo.te.saf.lu], [mo.tes.af.lu.], etc. Il conviendrait alors de chercher le classement correspondant à l'interlangue des sujets, de manière à rendre compte des données récoltées, comme dans le tableau hypothétique illustratif suivant :

/psamote/	*Complex	*Coda	Fill (V)	Parse (C)
<b>psa.mo.te</b>	* !			
<b>+ pu.sa.mo.te</b>			*	
<b>up.sa.mo.te</b> 151		* !	*	
<b>sa.mo.te</b>				*
<b>pa.mo.te</b>				*

<sup>151</sup> Notons que la forme /up.sa.mo.te/ est liée à la contrainte CONTIGUITY-IO qui interdit les épenthèses médiales ou l'effacement de segments : on la trouve par exemple en espagnol où elle est placée très haut dans le classement.



Nous avons donc décidé, dans un premier temps, de récolter certaines informations quantitatives sur les épenthèses vocaliques opérées au cours du traitement des structures complexes trisyllabiques biconsonantiques françaises, préalables à notre analyse des phénomènes en jeu, à l'aide d'un protocole de comptage syllabique. Ces informations correspondent à trois questions assez générales, qui s'articulent sur les axes classiques syntagmatique et paradigmaticque :

- (A) *La position du groupe cible dans le mot phonologique influence-t-elle le taux de confusion syllabique ?*
- (B) *La nature du groupe cible dans le mot phonologique influence-t-elle le taux de confusion syllabique ?*
- (C) *La structure de surface du groupe cible dans le mot phonologique influence-t-elle le taux de confusion syllabique ?*

Il est alors important de préciser le niveau d'analyse auquel sont abordés ces groupes, en particulier leur caractère tauto- ou hétéro-syllabique (voir Dell, 1995, pour la sélection des groupes).

### **II.2.1. Axe syntagmatique : la position du groupe cible dans le mot phonologique influence-t-elle le taux de confusion syllabique ?**

En termes phonologiques, cette question revient à se demander si l'apparition d'épenthèse(s) vocalique(s) produite(s) par des sujets japonais lors du traitement de groupes biconsonantiques français est plus probable lorsque le groupe CC est :

- *Attaque branchante de la syllabe initiale du mot phonologique (syllabe non accentuée)*  
(#CCV) ?
- *Attaque branchante de la syllabe finale du mot phonologique (syllabe accentuée)*  
(CCV#) ?

- Coda branchante de la syllabe *finale* du mot phonologique (syllabe *accentuée*)  
 (VCC#) ?

A cette fin, les mots phonologiques utilisés dans notre étude ont été composés de la manière suivante :

- #C(V)CVCVCV# (tri- ou tétra-syllabiques)
- #CVCVC(V)CV# (tri- ou tétra-syllabiques)
- #CVCVCVC(V)C(V)# (tri- ou penta-syllabiques)

### **II.2.2. Axe paradigmatique : la nature du groupe cible dans le mot phonologique influence-t-elle le taux de confusion syllabique ?**

En termes phonologiques, cette question revient à se demander si la nature des segments constitutifs des groupes biconsonantiques français influence de manière significative l'apparition d'épenthèses lors de leur traitement par des sujets japonais.

Ces groupes peuvent être classés en termes de :

- Couples phonématiques (36 couples)
- Traits distinctifs (4 familles générales, 10 familles et 12 sous-familles)
- Profils de sonorité (3 types de profil et 7 types d'écart)

Si, tout en tenant compte des débats sur la question de la sonorité (Morelli, 1999), l'on adopte la hiérarchie de sonorité proposée par Durand (1990, p. 210), on peut alors suivre l'échelle scalaire suivante :

Segment	a	e,o	i,u	r- sounds	laterals	nasals	voiced fricatives	voiceless fricatives	voiced stop	voiceless stop
Valeur	1	2	3	4	5	6	7	8	9	10

On obtient alors le classement suivant, issu des groupes présentés par Dell (1995) :

<b>Sujets 1 (120)</b>						
CC	N°	Traits distinctifs <sup>152</sup>		Sous-Famille	Sonorité	
		Segment 1	Segment 2		Seg. 1	Seg. 2
pn	1	OBS/plos/-cor/+ant	SON/nas/+cor	Stop + Nas	10	6
ps	2	OBS/plos/-cor/+ant	OBS/+sib/+ant	Stop + Sib	10	8
pf	3	OBS/plos/-cor/+ant	OBS/-sib/-voiced	Stop + Cont	10	8
pt	4	OBS/plos/-cor/+ant	OBS/plos	Stop + Stop	10	10
kn	5	OBS/plos/-cor/-ant	SON/nas/+cor	Stop + Nas	10	6
km	6	OBS/plos/-cor/-ant	SON/nas/-cor	Stop + Nas	10	6
kv	7	OBS/plos/-cor/-ant	OBS/-sib/+voiced	Stop + Cont	10	7
ks	8	OBS/plos/-cor/-ant	OBS/+sib/+ant	Stop + Sib	10	8
kt	9	OBS/plos/-cor/-ant	OBS/plos	Stop + Stop	10	10
tl	10	OBS/plos/+cor	SON/liqu/+lat	Stop + Lat	10	5
tm	11	OBS/plos/+cor	SON/nas/-cor	Stop + Nas	10	6
ts	12	OBS/plos/+cor	OBS/+sib/+ant	Stop + Sib	10	8
tʃ	13	OBS/plos/+cor	OBS/+sib/-ant	Stop + Sib	10	8
ft	36	OBS/-sib/-voiced	OBS/plos/+cor/-voiced	Cont+Stop	8	9

<sup>152</sup> Pour une présentation détaillée des traits distinctifs, voir Durand, 1990, ch. 2 ; Clements et Hume, 1995 ; Durand, 2005a. Dans ce codage, nous avons distingué les segments français /l/ et /R/ via une opposition [+/-lat] (lat = latérale) ou [+/-rhot] (rhot = rhotique). Cependant, nous aurions pu les distinguer de manière plus rigoureuse phonologiquement à travers l'opposition [+/-cor] (cor = coronale), comme nous l'avons fait précédemment.

<b>Sujets 2 (62)</b>						
<b>CC</b>	<b>N°</b>	<b>Traits distinctifs</b>		<b>Sous-Famille</b>	<b>Sonorité</b>	
pl	14	OBS/plos/-cor/+ant/-voiced	SON/liqu/+lat	Stop+Lat	10	5
pR	15	OBS/plos/-cor/+ant/-voiced	SON/liqu/-lat	Stop+Rhot	10	4
bl	16	OBS/plos/-cor/+ant/+voiced	SON/liqu/+lat	Stop+Lat	9	5
bR	17	OBS/plos/-cor/+ant/+voiced	SON/liqu/-lat	Stop+Rhot	9	4
fl	18	OBS/-sib/-voiced	SON/liqu/+lat	Cont+Lat	8	5
fR	19	OBS/-sib/-voiced	SON/liqu/-lat	Cont+Rhot	8	4
vl	20	OBS/-sib/+voiced	SON/liqu/+lat	Cont+Lat	7	5
vR	21	OBS/-sib/+voiced	SON/liqu/-lat	Cont+Rhot	7	4
tR	22	OBS/plos/+cor/-voiced	SON/liqu/-lat	Stop+Rhot	10	4
dR	23	OBS/plos/+cor/+voiced	SON/liqu/-lat	Stop+Rhot	9	4
kl	24	OBS/plos/-cor/-ant/-voiced	SON/liqu/+lat	Stop+Lat	10	5
kR	25	OBS/plos/-cor/-ant/-voiced	SON/liqu/-lat	Stop+Rhot	10	4
gl	26	OBS/plos/-cor/-ant/+voiced	SON/liqu/+lat	Stop+Lat	9	5
gR	27	OBS/plos/-cor/-ant/+voiced	SON/liqu/-lat	Stop+Rhot	9	4
sp	28	OBS/+sib/+ant	OBS/plos/-cor/+ant/-voiced	Sib+Stop	8	10
st	29	OBS/+sib/+ant	OBS/plos/+cor/-voiced	Sib+Stop	8	10
sk	30	OBS/+sib/+ant	OBS/plos/-cor/-ant/-voiced	Sib+Stop	8	10
sl	31	OBS/+sib/+ant	SON/liqu/+lat	Sib+Lat	8	5
sm	32	OBS/+sib/+ant	SON/nas/-cor	Sib+Nas	8	6
sn	33	OBS/+sib/+ant	SON/nas/+cor	Sib+Nas	8	6
sv	34	OBS/+sib/+ant	OBS/-sib/+voiced	Sib+Cont	8	7
sf	35	OBS/+sib/+ant	OBS/-sib/-voiced	Sib+Cont	8	8

Les familles se présentent ainsi :

Familles générales		Familles		Sous-Familles		Couples
<b>FG1</b>	Obs+Liqu	<b>F1</b>	Obs+Lat	<b>SF1</b>	Stop+Lat	10-14-16-24-26
				<b>SF2</b>	Cont+Lat	18-20
		<b>F2</b>	Obs+Rhot	<b>SF3</b>	Stop+Rhot	15-17-22-23-25-27
				<b>SF4</b>	Cont+Rhot	19-21
<b>FG2</b>	Obs+NonLiqu	<b>F3</b>	Obs+Nas	<b>SF5</b>	Obs+Nas	1-5-6-11
		<b>F4</b>	Obs+Cont	<b>SF6</b>	Obs+Cont	3-7
		<b>F5</b>	Obs+Sib	<b>SF7</b>	Obs+Sib	2-8-12-13
		<b>F6</b>	Obs+Stop	<b>SF8</b>	Obs+Stop	4-9-36
<b>FG3</b>	Sib+Stop	<b>F7</b>	Sib+Stop	<b>SF9</b>	Sib+Stop	28-29-30
<b>FG4</b>	Sib+NonStop	<b>F8</b>	Sib+Nas	<b>SF10</b>	Sib+Nas	32-33
		<b>F9</b>	Sib+Cont	<b>SF11</b>	Sib+Cont	34-35
		<b>F10</b>	Sib+Lat	<b>SF12</b>	Sib+Lat	31

Les profils de sonorité se présentent ainsi (M=Montant ; P=plateau ; D=Descendant) :

Type		Profil général	Couples
<b>P</b>	Plateau	10/10	9-4
<b>M</b>	Montant	10/8	2-3-8-12-13
<b>M</b>	Montant	10/7	7
<b>M</b>	Montant	10/6	1-5-6-11
<b>M</b>	Montant	10/5	10-14-24
<b>M</b>	Montant	10/4	15-22-25
<b>M</b>	Montant	9/5	16-26
<b>M</b>	Montant	9/4	17-23-27
<b>D (Violation de la SSG)</b>	Descendant	8/10	28-29-30
<b>D (Violation de la SSG)</b>	Descendant	8/9	36

<b>P</b>	Plateau	8/8	35
<b>M</b>	Montant	8/7	34
<b>M</b>	Montant	8/6	32-33
<b>M</b>	Montant	8/5	18-31
<b>M</b>	Montant	8/4	19
<b>M</b>	Montant	7/5	20
<b>M</b>	Montant	7/4	21

Enfin, voici le classement établi à partir des écarts de sonorité entre les segments composant les groupes :

<b>Ecart</b>	<b>Couples</b>
<b>0</b>	9-35-4 (3 couples)
<b>1</b>	34-36 (2 couples)
<b>2</b>	2-3-8-12-13-20-28-29-30-32-33 (11 couples)
<b>3</b>	7-18-21-31 (4 couples)
<b>4</b>	1-5-6-11-16-19-26-(7 couples)
<b>5</b>	10-14-17-23-24-27 (6 couples)
<b>6</b>	15-22-25 (3 couples)

Par ailleurs, dans la perspective de TCRS, on peut s'interroger sur le comportement du taux de confusion syllabique lorsque le groupe consonantique pose un double ou triple problème : syllabique (structure illicite en japonais) d'une part et segmental (nature absente en japonais) d'autre part. Dans un premier temps, on peut compter le nombre de malformations apparentes (à ne pas confondre avec le nombre de réparations nécessaires) des groupes :

Nombre de malformations	Un (syllabique)  (21 groupes)		Deux (syllabique et segmentale)  (13 groupes)		Trois (une syllabique et deux segmentales)  (2 groupes)	
	<b>Groupes</b>	pn	1	pf	3	fR
	ps	2	kv	7	vR	21
	pt	4	pR	15		
	kn	5	bR	17		
	km	6	fl	18		
	ks	8	vl	20		
	kt	9	tR	22		
	tl	10	dR	23		
	tm	11	kR	25		
	ts	12	gR	27		
	tʃ	13	sv	34		
	pl	14	sf	35		
	bl	16	ft	36		
	kl	24				
	gl	26				
	sp	28				
	st	29				
	sk	30				
	sl	31				
	sm	32				
	sn	33				

Evidemment, cette classification simplifie quelques peu les problèmes, puisqu'elle ne tient pas compte de certaines contraintes supplémentaires, en particulier phonotactiques

(par exemple, nous considérons les sibilantes /s/ et /ʃ/ comme non-problématiques d'un point de vue segmental, alors qu'en fait elles sont soumises à certaines contraintes distributionnelles).

Comme le rappelle Paradis (1996, p. 12) les contraintes syllabiques précèdent les contraintes segmentales, et le traitement des groupes CC non-syllabifiables, dans le domaine CC, nécessite au moins deux réparations : insertion nucléique et remplissage segmental. En termes de stratégies, le choix des réparations peut varier en fonction de la position syllabique des groupes :

\*En attaque : soit (a) il y a effacement du segment consonantique le plus marqué du point de vue du profil de sonorité (Paradis, 1996, p. 14) et adaptation éventuelle du segment restant (1 ou 2 opérations), soit (b) il y a insertion nucléique, remplissage segmental et adaptation d'un ou de deux segments (2, 3 ou 4 opérations). D'après le *Principe de Préservation*, quand l'adaptation segmentale n'est pas nécessaire, l'épenthèse devrait prévaloir. Dans le cas contraire, ce devrait être l'effacement.

\*En coda : soit (a) il y a effacement du segment consonantique le plus marqué du point de vue du profil de sonorité, puis insertion nucléique et remplissage segmental, avec éventuellement adaptation du segment restant (3 ou 4 opérations), soit (b) effacement des deux segments consonantiques (deux opérations mais qui en fait se divisent en (1) et (1), le domaine de contrainte se modifiant au fur et à mesure des réparations), soit (c) insertion nucléique et remplissage segmental avec éventuellement adaptation du segment restant (2 ou 3 opérations), réparation qui doit ensuite être répétée sur la coda restante. Mais dans ce cas le domaine n'est plus le même, puisque, comme l'indique Paradis (1996, p. 23), dans un cadre phonologique-dérivationnel et non dans un cadre OT, « a constraint domain disappears as soon as the violated constraint is complied with » (donc 4, 5 ou 6 opérations, mais qui en fait se divisent en (2 ou 3) et (2 ou 3)). Selon la nature des segments constitutifs des groupes (illicites ou non en japonais), ce sont donc les stratégies (b) ou (c) qui devraient prévaloir.



En détaillant, les groupes, on peut donc tenter de prévoir la stratégie adoptée. Pour cela, il nous faut reprendre l'échelle de sonorité proposée par Durand (1990, p. 210), afin d'indiquer quels segments, parmi les candidats à l'effacement, sont les plus marqués du point de vue de la sonorité, et donc seront effacés :

<b>Segment</b>	<b>a</b>	<b>e,o</b>	<b>i,u</b>	<b>r- sounds</b>	<b>laterals</b>	<b>nasals</b>	<b>voiced fricatives</b>	<b>voiceless fricatives</b>	<b>voiced stop</b>	<b>voiceless stop</b>
<b>Valeur</b>	1	2	3	4	5	6	7	8	9	10

**En attaque :**

En attaque, si le segment moins marqué ne nécessite pas d'adaptation, alors le poids procédural se réduit à un.

<b>Poids Procédural</b>	<b>Attaque</b>					
	<b>Un</b> (réparation syllabique = effacement consonantique + réparation segmentale = adaptation <i>a priori</i> MAIS NON <i>a fortiori</i> ). *(C1*(C2)V) → C1V  (10 groupes)		<b>Deux</b> (réparation syllabique = effacement consonantique + réparation segmentale = adaptation <i>a priori</i> ET <i>a fortiori</i> ). *(C1*(C2)V) → *(C2)V → C3V  (5 groupes)		<b>Deux</b> (réparation syllabique = épenthèse vocalique) (pas d'adaptation segmentale <i>a priori</i> ).  *(C1C2V) → C1V.C2V  (21 groupes)	
<b>Groupe</b>	pf	3	sv	34	pn	1
	kv	7	sf	35	ps	2
	pR	15	ft	36	pt	4
	bR	17	fR	19	kn	5
	fl	18	vR	21	km	6
	vl	20			ks	8
	tR	22			kt	9
	dR	23			tl	10
	kR	25			tm	11
	gR	27			ts	12
					tʃ	13
					pl	14
					bl	16
					kl	24
					gl	26
					sp	28
					st	29
					sk	30
					sl	31
					sm	32
					sn	33

**En Coda :**

En coda, si aucune réparation segmentale n'est nécessaire, alors, étant donné le *Principe de Préservation*, l'épenthèse devrait prévaloir : deux opérations pour la première épenthèse sur le premier domaine, puis deux opérations pour la deuxième épenthèse pour le deuxième domaine. En revanche, dans les cas où au moins l'un des segments en position de coda nécessiterait adaptation s'il se trouvait en position d'attaque, alors l'effacement devrait prévaloir, car une stratégie d'insertion, quelle qu'elle soit, entraînerait une réparation en trois opérations, donc au-delà de la limite imposée par le *Principe de Seuil*.

		<b>Coda</b>		
<b>Poids Procédural</b>	<b>Un</b>		<b>Deux</b>	
	(réparation syllabique en deux temps = deux effacements consonantiques) (au moins un segment illicite). C*(V*)(C1*)(C2)) → C*(V*)(C1) → CV  (15 groupes)		(réparation syllabique en deux temps = deux épenthèses vocaliques) (pas de segment illicite). C*(VC1C2) → CV.C1*(VC2) → CV.C1V.C2V  (21 groupes)	
<b>Groupe</b>	pf kv pR bR fl fR vl vR tR dR kR gR sv sf ft	3 7 15 17 18 19 20 21 22 23 25 27 34 35 36	pn ps pt kn km ks kt tl tm ts tʃ pl bl kl gl sp st sk sl sm sn	1 2 4 5 6 8 9 10 11 12 13 14 16 24 26 28 29 30 31 32 33

Il s'agit donc de savoir si certains groupes, caractérisés par leur nature phonématique, leur famille infra-segmentale ou leur profil de sonorité, vont déclencher davantage d'épenthèses que d'autres et ainsi se présenter comme plus marqués pour les sujets

concernés. Cela devrait également nous permettre de confirmer ou d'infirmer les prédictions relatives aux stratégies de réparations présentées ci-dessus.

### **II.2.3. Axe paradigmatique : la structure de surface du groupe cible dans le mot phonologique influence-t-elle le taux de confusion syllabique ?**

En termes phonologiques, la question que nous posons est la suivante : si l'on suppose que les groupes de type CC en *input* conduisent à la sélection de groupes CvépC en *output* (avec vép=voyelle épenthétique), alors les groupes de surface CvépC sont-ils traités de la même manière que les groupes CC ou pas ?

Ainsi, parmi les mots phonologiques présentés dans notre protocole figurent des mots pour lesquels les groupes CC ont été « artificiellement séparés » par l'insertion de voyelles épenthétiques propres au système théorique d'épenthèse du japonais. Ces voyelles épenthétiques ont été insérées de manière relativement semblable à, par exemple, celle utilisée par Fleischhacker dans ses expériences (2000, p. 19). Si l'on considère que les formes de surface [CC] et [CvépC] correspondent à la même forme sous-jacente /CvépC/ pour les sujets japonais, alors les résultats de nos épreuves, en termes de comptage syllabique, devraient être les mêmes pour les deux types de mots phonologiques. L'intérêt d'une telle question est notamment lié à la nature des voyelles neutres et/ou épenthétiques dans les deux langues : principalement le schwa en français vs. le /ɯ/ en japonais (généralement non arrondie (« unrounded ») ; à ce propos voir l'analyse de Lombardi, 2002).

Si l'on constate l'apparition de voyelles épenthétiques dans des groupes CC, par exemple [abusutorakuto] pour « abstract », on constate aussi la réduction (le dévoisement ?), pour ne pas dire l'effacement, de voyelles censées être pleines en français et qui correspondent aux voyelles neutres/épenthétiques du japonais (principalement le /ɯ/), par exemple [ekte] pour « écoutez ». Cela nous renvoie au problème de discordances (« mismatches ») entre représentations syllabiques phonologiques et phonétiques que mentionne Blevins (1995, p. 233), avec d'un côté

l'effacement phonétique de voyelles phonologiques, et de l'autre une augmentation du nombre de pics de sonorité phonétiques par rapport au nombre initial de syllabes phonologiques. Ainsi, Labrune (2001b, p. 117), en mentionnant les formes d'emprunts « senegarū » et « amusuterudamū » (/ɯ/ final = voyelle épenthétique) d'une part, et le dévoisement vocalique dans le mot « nagasaki-si » (premier [i̥] = voyelle dévoisée) d'autre part, souligne la nécessité de bien distinguer syllabes phonétiques et phonologiques, puisque « on ne peut nier que des syllabes fermées comme *sak* [(dans ce cas issue du dévoisement vocalique du premier [i] dans le mot « nagasaki-si »)] aient une réalité phonétique » (*op. cit.*, note 10).

Nous avons donc décidé de vérifier, d'une part que le traitement des groupes CC correspondait à nos prévisions (c'est-à-dire épenthèses, conduisant à une équivalence de traitement entre groupes CC et CvépC), et d'autre part que celui des groupes CvépC était stable (qu'il n'y ait pas de réduction/effacement significatifs de vép). Si les résultats ne correspondaient pas à nos attentes, il serait intéressant de pouvoir identifier le phénomène majoritaire (épenthèse ou délétion de vép).

#### **II.2.4. Résumé**

Ainsi, nous pouvons synthétiser nos problématiques phonologiques de la manière suivante :

Domaine	Segmental (phonético-phonologique)		Syllabique (phonologique)		
	Syntagmatique	Paradigmatique	Syntagmatique	Paradigmatique	
Aspects	Position	Nature	Position	Nature	Structure
Questions	A	B	C	D	E
Expérience	(F)	(F)	(C1 et C2)	(C1 et C2)	(C1 et C2)
<b>A</b>	<i>La position du segment cible dans le mot phonologique influence-t-elle le taux de confusion segmentale ?</i>				
<b>B</b>	<i>La nature du segment cible dans le mot phonologique influence-t-elle le taux de confusion segmentale ?</i>				
<b>C</b>	<i>La position du groupe cible dans le mot phonologique influence-t-elle le taux de confusion syllabique ?</i>				
<b>D</b>	<i>La nature du groupe cible dans le mot phonologique influence-t-elle le taux de confusion syllabique ?</i>				
<b>E</b>	<i>La structure de surface du groupe cible dans le mot phonologique influence-t-elle le taux de confusion syllabique ?</i>				

Les données collectées devraient nous permettre de tenter de répondre à ces questions, étape préalable à toute démarche explicative. A partir de ces données statistiques, nous devrions ensuite être en mesure, dans le cadre OT, de proposer certains classements de contraintes, segmentale et syllabique, propres à l'interlangue des sujets japonais apprenant le français.

Nous sommes bien conscients des limites méthodologiques, d'un point de vue phonologique, de notre protocole, en particulier en ce qui concerne les tâches de comptage syllabique : comme le montre Paradis (1996, pp. 21-22), dans une perspective différente cependant, tant le comptage systématique des changements entre input et output, que le comptage de ces changements à l'intérieur des domaines de contraintes, se montre problématique. Ceci étant signalé, nous examinerons les implications et les limites de notre protocole lors du descriptif exact de celui-ci.

### **II.2.5. Aspects phonographémiques**

Au cours de la mise en place de nos problématiques phonologiques, nous avons été conduits à nous interroger sur l'existence de facteurs non strictement phonologiques devant être pris en compte lors de l'analyse de l'apprentissage de la phonologie d'une langue étrangère. Parmi ceux-ci, nous avons mentionné les facteurs phonographémiques, c'est-à-dire ceux qui renvoient aux relations entre représentations phonético-phonologiques et formes graphémiques/graphiques et donc à l'articulation entre représentations orales et écrites, entre perception phonético-phonologique et reconnaissance visuelle de graphèmes, entre parole et écriture.

Car si nous nous intéressons bien à l'apprentissage d'un système phonologique étranger, et non à celui du système phonographémique qui lui correspond, on ne peut cependant ignorer, dans une perspective *réaliste* (Detey, 2003c), la présence et les interactions des deux systèmes (quatre si l'on tient compte des systèmes de la L1) au cours de l'apprentissage phonologique, sauf dans des cas extrêmement rares (dont nous n'avons d'ailleurs pas connaissance, mais que l'on peut facilement envisager, notamment dans certaines sociétés plurilingues) d'apprentissage strictement oral, dans des sociétés et pour des systèmes linguistiques dépourvus de système phonographémique.

Dans le cas qui nous occupe, de sujets japonais apprenant le français, ce facteur est tout à fait pertinent, comme nous l'avons vu plus haut en contrastant leur(s) système(s) phonographémique(s) respectif(s). Si l'on considère que l'apprentissage phonologique



est étroitement imbriqué à celui du système phonographémique, l'un servant de support ou de médium à l'autre, on ne peut passer sous silence les éventuelles influences qui s'exercent entre eux. Cette « négligence » est malheureusement parfois constatée, dans une certaine mesure, dans les protocoles d'enquête des études en interphonologie (Monahan, 2001) ou en phonologie des emprunts (Shinohara, 1997). Cependant, l'orthographe en tant que variable cognitivement non négligeable semble de plus en plus prise en compte, notamment en phonologie des emprunts (Peperkamp et Dupoux, 2003 ; Smith, 2005), comme l'indiquent Vendelin et Peperkamp (2005) dans leur résumé :

*« Loanword adaptations are traditionally analyzed as phonologically minimal changes that are computed by the production grammar of the borrowing grammar. Several authors, however, have proposed that some loanword adaptations take place in perception and are based on phonetic distance [...]. Peperkamp and Dupoux (2003) argued that all loanword adaptations result from perceptual assimilation and are based on the phonetic closeness of the surface form of a foreign word to the surface form of possible native words. [...] We examine a possible confounding factor. That is the impact of perception might be obscured by the role of orthography. In fact, a great deal of loanword adaptations are arguably based on written input, and often some standardization has taken place as to how the L2 graphemes are pronounced in the L1. [...] The results [...] suggest that subjects are sensitive to fine phonetic detail when they have to base their adaptations on perception only, and tend to use learned grapheme-to-phoneme correspondence rules when an orthographic representation is available. [...] We conclude that loanword adaptations made on the basis of the written input are heavily influenced by the L1 speakers knowledge of the L2 orthography. Thus, the loanword adaptation is best studied in on-line adaptation of oral words ».*

Le facteur orthographique semble ainsi à présent de moins en moins négligé dans les protocoles d'investigation sur la perception de la parole, comme l'illustre celui de

Escudero et Boersma (2004, p. 6) sur la perception de voyelles anglaises par des apprenants espagnols :

*« For both L1 and L2 perception, we used pictures rather than written words, in order to avoid orthographic effects ».*

Un tel effet avait par ailleurs déjà été envisagé par Shinohara, qui s'était servie, de manière non exclusive, d'un input orthographique lors de la constitution de son corpus sur l'adaptation de mots français et anglais en japonais (Shinohara, 1997, p. 12) :

*« La deuxième voyelle /ee/ dans l'emprunt /messeezi/ <message> (anglais) /mesɪdʒ/ « message » correspond à la voyelle /ɪ/ dans le mot d'origine. Nous pensons que cette forme a été introduite par un locuteur japonais qui a fait l'analogie avec l'adaptation de la voyelle /eɪ/ orthographiée également a dans d'autres mots comme age /eɪdʒ/ ».*

Ainsi, après avoir développé la question du « Quoi » apprendre en termes strictement phonético-phonologiques, c'est à présent celui du « Comment » apprendre qui nous occupe, nous conduisant à des questions d'ordre non plus seulement phonologique mais aussi phonographémique. En effet, si, en général, l'apprentissage de la phonologie d'une langue se fait essentiellement *via* le canal auditif, et donc phonétique, il est depuis longtemps clairement établi que la perception auditive ne repose pas uniquement sur l'appareil auditif. La théorie motrice de la parole par exemple, ou bien encore la théorie de la perception-pour-le-contrôle-de-l'action (Schwartz, Abry, Boë et Cathiard, 2002), insistent sur les étroites relations entre perception et production dans la boucle audio-phonatoire, ce qui, d'une certaine manière, ramène alors à l'une des caractéristiques de la perception phonético-phonologique : son caractère multimodal et plus particulièrement bimodal, auditif et visuel (lecture labiale par exemple), sans négliger cependant la dimension tactile.

On ne peut alors tenter de rendre compte des réalisations de /b/ en [β], de /v/ en [b] ou en [β], de /l/ en [r], de /R/ en [r], de /OU/ en [u] et de /EU/ en [u] ou en [o] en ignorant ces aspects psycholinguistiques de l'apprentissage de la phonologie. Comme l'écrit Odlin (1989, p. 112):

*« Pronunciation often shows other influences besides cross-linguistic ones, [...] developmental and other factors [...] as well as transfer involving writing systems ».*

Il cite l'exemple du /r/ dans le mot anglais « route » prononcé par des apprenants français, influencés tant par l'orthographe que par la forme lexicale de même origine (« cognate words ») (1989, p. 114). Il est à noter que, dans son chapitre sur les transferts phonético-phonologique, il en consacre précisément l'avant-dernier sous-chapitre aux systèmes d'écriture, indiquant que les similarités et les différences entre les systèmes d'écriture en présence pouvaient déboucher sur des transferts positifs ou négatifs. Dans notre cas, nous pouvons tenter de schématiser dans un premier temps les problèmes émergents de la manière suivante :

Segments L2 (français)		Système phonologique L1 (japonais)	Système graphémique L1 (japonais)
<i>Phono</i>	<i>Graph</i>		
/b/	« b »	Licite	Licite
/v/	« v »	Illicite → /b/	Illicite → « b »
/l/	« l »	Licite mais → /r/	Illicite → « r » → /r/
/R/	« r »	Illicite → /r/	Licite mais → /r/
/EU/	« e » ou « eu »	Illicite → /u /	Licite mais → /e/ ou /e+/u/
/OU/	« ou »	Licite	Licite mais → /o+/u/

Structure syllabique L2 (français)		Système phonologique L1 (japonais)	Système graphémique L1 (japonais)
<i>Phono</i>	<i>Graph</i>		
/CCV/	« CCV »	Illicite → /CVCV/	Illicite → « CVCV »
/CVCV/ <sup>153</sup>	« CCV »	Licite	Illicite → « CVCV »
/CCV/	« CVCV » <sup>154</sup>	Illicite → /CVCV/	Licite

Bien sûr ce schéma ne rend pas compte de la complexité des systèmes phonographémiques français et japonais (homophones, homographes, strates lexicales, emprunts récents), mais il permet néanmoins d'envisager certaines questions générales, relatives au jeu entre phonèmes, allophones, graphèmes et allographes (par exemple dans le cas de la syllabe française /li/ perçue [li], considéré comme allophone de la syllabe japonaise /ri/, puis transcrit « RI » selon le système graphique japonais, renforçant ainsi, par un effet de boucle, l'identification erronée des phonèmes /R/ et /l/ et *a fortiori* de leurs graphèmes correspondants « R » et « L »).

S'il semble donc nécessaire de se pencher sur le pendant phonographémique de nos problématiques phonologiques, tant au plan segmental qu'au plan syllabique, cela nous conduit à sortir du champ strictement linguistique de la phonologie pour entrer pleinement dans celui de la psycholinguistique.

---

<sup>153</sup> Cas rare, à aborder dans une perspective sociolinguistique.

<sup>154</sup> Ce cas de figure évoque le cas du « schwa/ e muet » en français (« petit » prononcé [pti]), même s'il convient de ne pas confondre structure syllabique profonde /CəCV/ et de surface /CCV/.

### **III. Investigation psycholinguistique : orthographe et interphonologie chez les apprenants japonais de FLE**

Comme l'écrivent Kess et Miyamoto (1999, p. 11) :

*« Western linguists have been largely guided by Saussure's (1916) suggestion that spoken language is the first and foremost province of true language study, and that written language is but a pale and uninformative imitation of 'real' language ».*

Cependant, comme nous l'avons vu, l'apprentissage de la parole en langue étrangère implique plusieurs variables, dépassant largement le champ strict de la linguistique « classique ».

#### **III.1. Introduction : variables générales, tâches d'apprentissage et niveaux d'analyse**

A propos des variables impliquées dans l'apprentissage de la parole en langue étrangère, Leather et James (1991) mentionnent notamment l'âge, la motivation, la pression sociale (intégration et distance sociales), la personnalité, le sexe ou encore les capacités audiophonatoires. D'un point de vue psycholinguistique, l'apprentissage de la phonologie du français par des sujets adultes japonais implique en effet de nombreux facteurs, linguistiques, cognitifs et contextuels, que l'on peut diviser en deux catégories générales, individuelles et collectives.

	<b>Collectives</b>	<b>Individuelles</b>
<b>Linguistiques</b>	<ul style="list-style-type: none"> <li>• Langue première (phonologie)</li> <li>• Langue première (orthographe)</li> <li>• Adaptation des emprunts</li> </ul>	<ul style="list-style-type: none"> <li>• Variation sociolinguistique en langue première</li> <li>• Biographie langagière / Capital langagier</li> <li>• Traits audio-phonatoires</li> </ul>
<b>Cognitives</b>	<ul style="list-style-type: none"> <li>• Système perceptivo-attentionnel</li> <li>• Système mnésique</li> <li>• Processus d'apprentissage</li> </ul>	<ul style="list-style-type: none"> <li>• Contenus de savoirs déclaratifs et procéduraux</li> <li>• « Style » cognitif</li> <li>• Stratégies d'apprentissage</li> </ul>
<b>Contextuelles</b>	<ul style="list-style-type: none"> <li>• Cadre</li> <li>• Tâches</li> </ul>	<ul style="list-style-type: none"> <li>• Conditions psycho-physiologiques</li> <li>• Motivation</li> </ul>

Dans la première partie de notre étude, nous nous sommes concentrés sur les variables linguistiques, essentiellement collectives, en adoptant une perspective strictement linguistique, phonético-phonologique. Cela nous a permis d'examiner d'une part le contenu dudit apprentissage (représentations phonologiques) et d'autre part certains processus relatifs au traitement et à l'apprentissage de ce contenu. Il est cependant apparu qu'une telle perspective pouvait parfois s'avérer limitative, et ainsi ne pas suffisamment rendre compte des processus d'apprentissage dans leur ensemble. Nous avons en particulier souligné, lors de l'examen de plusieurs études en interphonologie, que les protocoles d'enquête semblaient négliger le facteur orthographique, et plus généralement le rôle des représentations visuo-graphiques (de surface ou profondes), dans cet apprentissage, ainsi que dans les méthodes de récolte des données. De manière schématique, la perspective que nous avons adoptée dans cette première partie nous a permis de définir précisément le *Quoi* apprendre, mais ne nous a que partiellement renseignés sur le *Comment* apprendre.

Or, nous avons vu par la suite que la perception et l'apprentissage de la parole en langue étrangère étaient foncièrement multimodaux, en particulier audiovisuels. Si l'on décide de se concentrer sur une population de sujets adultes sachant lire, écrire et ne souffrant d'aucune déficience perceptive ou pathologique, l'apprentissage d'une langue étrangère et de son système phonologique s'opère en effet généralement à l'aide de plusieurs canaux perceptifs, en particulier visuel et auditif. Ces canaux permettent de recevoir plusieurs types de signaux, sémiotiques ou non, et les productions des apprenants, quelle qu'en soit la nature, présentent également ce caractère multimodal. En ce qui concerne les représentations profondes, les « savoirs », la valeur du concept de multimodalité varie en fonction du paradigme théorique adopté (en particulier symbolique *vs.* non symbolique), ainsi que du cadre disciplinaire dans lequel il est placé (linguistique, psycholinguistique ou neuropsycholinguistique). Dans tous les cas, il s'agit d'une question toujours débattue par les chercheurs, qui rend délicate l'utilisation du terme « représentation(s) » et qui s'articule en particulier sur deux axes généraux :

	<b>Modal</b> (ex : double codage, imagé et propositionnel)	
<b>Computo-symbolique</b> (primitivisme)	<i>Versus</i>	<b>Sub-symbolique</b> (émergentisme)
	<b>Amodal</b> (propositionnel)	

Si l'on considère que l'apprentissage de la phonologie d'une langue étrangère se fait essentiellement en bimodalité, c'est-à-dire à l'aide de stimuli auditifs et visuels qui conduisent à l'activation de processus et de représentations attachés aux deux modalités, nous pouvons alors nous interroger sur le rôle joué par la modalité visuelle dans l'apprentissage du système phonologique, des niveaux périphériques au niveaux cognitifs plus profonds<sup>155</sup>. D'un point de vue périphérique, nous devons préciser la nature des stimuli visuels qui nous intéressent : il s'agit de celles les plus fréquemment utilisées en situation d'apprentissage linguistique, à savoir les « représentations » dites « orthographiques ». Si elles sont sans doute les mieux connues et celles qui ont été le plus abondamment étudié en psycholinguistique, elles n'en perdent pas moins leur intérêt étant donné la complexité et la multiplicité des systèmes de correspondances entre systèmes orthographiques et systèmes phonético-phonologiques à travers les langues. Notre champ d'investigation peut alors, dans un premier temps, se présenter ainsi :

<b>Linguistique</b>		<b>Cognitif</b>		<b>Contextuel</b>
<b>Phonologie</b>	<b>Orthographe</b>	<b>Perception</b>	<b>Apprentissage</b>	<b>Tâche</b>
Français	Français	Français	Mémoire	Conditions et Actions
Japonais	Japonais	Japonais		

A chacun de ces grands champs correspond une abondante littérature scientifique, parfois susceptible d'établir des connexions avec les champs adjacents, mais généralement limitée à son cadre d'origine : interphonologie et phonologie des emprunts pour le domaine de la phonologie, lecture et reconnaissance visuelle pour le domaine de l'orthographe et de la perception visuelle, perception catégorielle et multimodale pour la perception de la parole, mémoire de travail, boucle phonologique

<sup>155</sup> Il apparaît en effet que l'on ne peut pas, dans l'approche qui est la nôtre, traiter d'apprentissage phonologique en négligeant la question des modalités perceptives, celles-ci pouvant influencer celui-là. Il est, à cet égard, significatif que l'un des chapitres d'un manuel universitaire de phonologie à paraître bientôt soit intitulé « The visual element in phonological perception and learning » (Hardison, D. M., in Pennington, M. C. (éd.), *Palgrave Guide to Phonology*, éditions Macmillan, à paraître).



et lexicale mentale pour l'encodage mnésique, théories de l'apprentissage, transfert et grammaire universelle pour l'apprentissage linguistique et finalement tous les effets et variables propres aux travaux psycholinguistiques en ce qui concerne les tâches, expérimentales ou non.

Les études tentant d'intégrer ces différents champs sont cependant rares, pour des raisons évidentes d'un point de vue scientifique. Néanmoins, étant donné notre problématique, il nous a semblé important de tenter de rassembler et d'ordonner ces différents champs, afin de parvenir à une vision suffisamment large des processus à l'œuvre, ce qui nous permet de poser certaines questions très générales :

<b>1) Phonologie</b>	Le système phonologique japonais influence-t-il la perception et/ou l'apprentissage du système phonologique français ?
<b>2) Orthographe</b>	Le système orthographique/phonographémique japonais influence-t-il la perception et/ou l'apprentissage du système phonologique français ?
<b>3) Perception</b>	La perception de la parole et la reconnaissance visuelle en japonais (unités et processus) influencent-elles la perception et/ou l'apprentissage du système phonologique français ?
<b>4) Apprentissage</b>	Quel rôle la mémoire de travail, plus particulièrement la boucle phonologique, joue-t-elle dans l'influence probable exercée par le système orthographique du japonais sur la perception et/ou l'apprentissage du système phonologique français ?
<b>5) Tâche</b>	Certaines variables expérimentales (complexité de la tâche, fréquence d'occurrence en japonais des stimuli auditifs utilisés, régularité en japonais des stimuli visuels utilisés, etc.) et descriptives (sexe, degré d'expertise en français, etc.) ont-elle une influence significative sur les résultats obtenus ?

Afin d'aborder ces questions, nous avons sélectionné deux axes dans la phonologie du français, que nous rappelons ici :

- L'un, *spécifique*, au niveau *segmental*, concerne les unités phonématiques françaises /b/ et /v/ d'une part, /l/ et /R/ d'autre part, et, dans une certaine mesure, /EU/ et /OU/.
- L'autre, de portée plus *générale*, au niveau *syllabique*, concerne les structures syllabiques complexes françaises biconsonantiques.

Ayant traité la première question dans la première partie de notre étude, nous nous penchons à présent sur les quatre autres questions, étant entendu que les trois dernières sont ici pour nous pleinement subordonnées à la deuxième, laquelle constitue le centre de préoccupation majeur de notre travail, envisagée sous l'angle d'un possible transfert intermodal.

Peu d'études, à notre connaissance, se sont penchées sur l'influence de l'orthographe en L1 sur la perception de la parole en L2, la plupart étant consacrée à la reconnaissance lexicale visuelle (Chikamatsu, 1996), à la lecture (Akamatsu, 2003) ou à l'influence de l'orthographe en L2 sur la perception de la parole en L2 (Steele, 2005). On peut néanmoins citer celle de Byung-jin (2004) conduite auprès d'apprenants coréens d'anglais à propos des oppositions de voisement entre /p/ et /b/, en position initiale et finale : il s'agissait alors d'évaluer la perception du voisement à travers le choix d'une transcription orthographique. Un groupe disposait de transcriptions en orthographe alphabétique anglaise, tandis que l'autre groupe disposait de transcriptions en orthographe coréenne *Hangul*, qui est une orthographe assez « superficielle ». Il semblerait alors que les confusions ont été plus nombreuses pour le second groupe, suggérant une influence de l'orthographe sur la perception du contraste.

Si l'on considère que l'apprentissage de la phonologie du français s'effectue via une série de tâches cognitives à accomplir, l'influence du système orthographique japonais sur celui-ci peut être envisagée différemment selon la nature de la tâche :

- Influence d'origine *externe* lorsqu'il s'agit d'une tâche de lecture (à voix haute ou non), puisque les stimuli visuels français activent, directement si les graphes ou si les formes lexicales visuelles sont partagées par les deux langues, ou indirectement dans le cas contraire, les graphèmes japonais. Cette situation s'apparente à celles de *reconnaissance*.
- Influence d'origine *interne* lorsqu'il s'agit d'une tâche de perception phonéto-phonologique (unimodale ou bimodale avec vision faciale ; tâche de reconnaissance, de discrimination, voire de dictée), puisque même si aucun stimulus orthographique n'est présenté, certaines représentations internes d'ordre orthographique, peuvent être activées, automatiquement ou pas. Cette situation s'apparente davantage à celles de *rappel*.

Par ailleurs, cette influence peut s'exprimer *via* des *unités* de représentation (en particulier au niveau segmental), mais également *via* des *processus* de représentations, en particulier celui de segmentation et ainsi de discrétisation d'ensembles linguistiques perceptivement continus (en particulier au niveau métrique / syllabique).

On peut schématiser dans un premier temps la situation ainsi :

	<b>Externe</b>	<b>Interne</b>
<b>Unités</b>	(pseudo)homographes	(pseudo)homophones
<b>Processus</b>	Segmentation/Syllabification	Segmentation/Syllabification

Il convient également de distinguer deux niveaux d'analyse dans notre étude :

- L'un, général, concerne les unités et les processus mobilisés en tâche de perception bimodale, quelles que soient l'origine linguistique des sujets et des stimuli utilisés ;
- L'autre, spécifique, implique les particularismes propres aux locuteurs et aux langues en présence, dans notre cas ceux du japonais et du français.

Si le premier niveau nous permet de nous référer à certains modèles (modèles à double voie de reconnaissance visuelle ; modèle de mémoire de travail), le deuxième niveau nous oblige à adapter ces modèles généraux à notre situation particulière, tant au niveau des unités qu'à celui des processus : mores, segmentation moraïque, système tri-, tétra-, voir penta-scriptural et strates lexicales pour le japonais ; syllabes, segmentation syllabique et système alphabétique opaque pour le français. L'apprentissage du français par des locuteurs espagnols (système alphabétique avec orthographe relativement transparente), mandarins (système mono- ou bi-scriptural) ou anglais (système de segmentation métrique accentuel non-syllabique) nous conduirait ainsi à d'autres analyses.

<b>Niveau</b>	
<b>Général</b>	<b>Spécifique</b>
Modèle de reconnaissance visuelle	Systèmes orthographiques français et japonais
Modèle de perception de la parole	Unités et processus perceptifs français et japonais
Modèle de mémoire de travail	Processus d'apprentissage vs d'adaptation

A cette distinction, vient s'ajouter celle entre le niveau lexical / morphémique (reconnaissance de mots ou de morphèmes) et le niveau sub-lexical / graphémique (activation de graphèmes). Celle-ci se superpose en partie à la division entre les deux axes phonologiques de notre étude : syllabique pour le niveau lexical (segmentation de mot phonologique) et segmental pour le niveau sub-lexical (appariement phonème-graphème).

<b>Niveau</b>	<b>Tâche</b>
Sub-lexical	Appariement phonème – graphème
Lexical	Comptage syllabique

### **III.2. Nature des tâches cognitives**

Si jusqu'à présent nous nous sommes efforcés de mettre en parallèle les deux types de reconnaissance évoqués (auditive et visuelle) en vue de procéder à leur intégration, il convient cependant de respecter leurs spécificités, qui peuvent nous aider à mieux appréhender les schémas d'activation.

#### **III.2.1. Tâche de Lecture – unimodalité visuelle**

Cette tâche implique :

- Les deux pôles perception/production ;
- En perception : reconnaissance inter-modale  
(visuel/orthographique → auditif/phonologique)

Nous devons cependant bien préciser ce que nous entendons par tâche de lecture (en tant qu'oralisation de signes écrits, et non en tant qu'activité de compréhension textuelle) et quels en sont les aspects qui importent pour nous.

a) Lecture de mot vs. Lecture d'énoncés : nous nous intéressons essentiellement ici à la lecture de mots, en ce que la lecture d'énoncés implique une participation accrue des informations morphophonologiques (liaisons, flexions verbales, etc.) et syntaxico-sémantiques (prosodie). La lecture d'énoncés constitue donc une complexification de la première tâche, alors que nous nous intéressons d'abord au domaine segmental et syllabique.

b) Lecture de mots/morphèmes inconnus vs. Lecture de mots/morphèmes connus : nous nous intéressons essentiellement ici à la lecture de mots/morphèmes inconnus, étant donné que ce sont eux qui nécessitent la plus grande implication de la voie d'assemblage. La lecture de mots connus renvoie en effet aux processus de reconnaissance lexicale. Dans le cas d'apprenants en début d'apprentissage, la lecture de mots « transparents » ou d'emprunts, comme, pour le japonais, les mots « tomate » ou « football », active automatiquement les formes phonologiques lexicales existant

dans leur langue première via la voie d'adressage. La lecture de tels mots, si elle n'est certes pas bénéfique du point de l'apprentissage des formes phonologiques lexicales, ne concerne finalement que de manière limitée la constitution de l'interphonologie des apprenants, qui, dans une certaine mesure, semble davantage liée à l'acquisition des inventaires segmentaux, syllabiques (et ainsi phonotactiques) et phonographémiques de la langue à apprendre. Nous nous intéressons donc davantage au traitement des unités sublexicales, à leur reconnaissance éventuelle et ainsi à la lecture des mots inconnus.

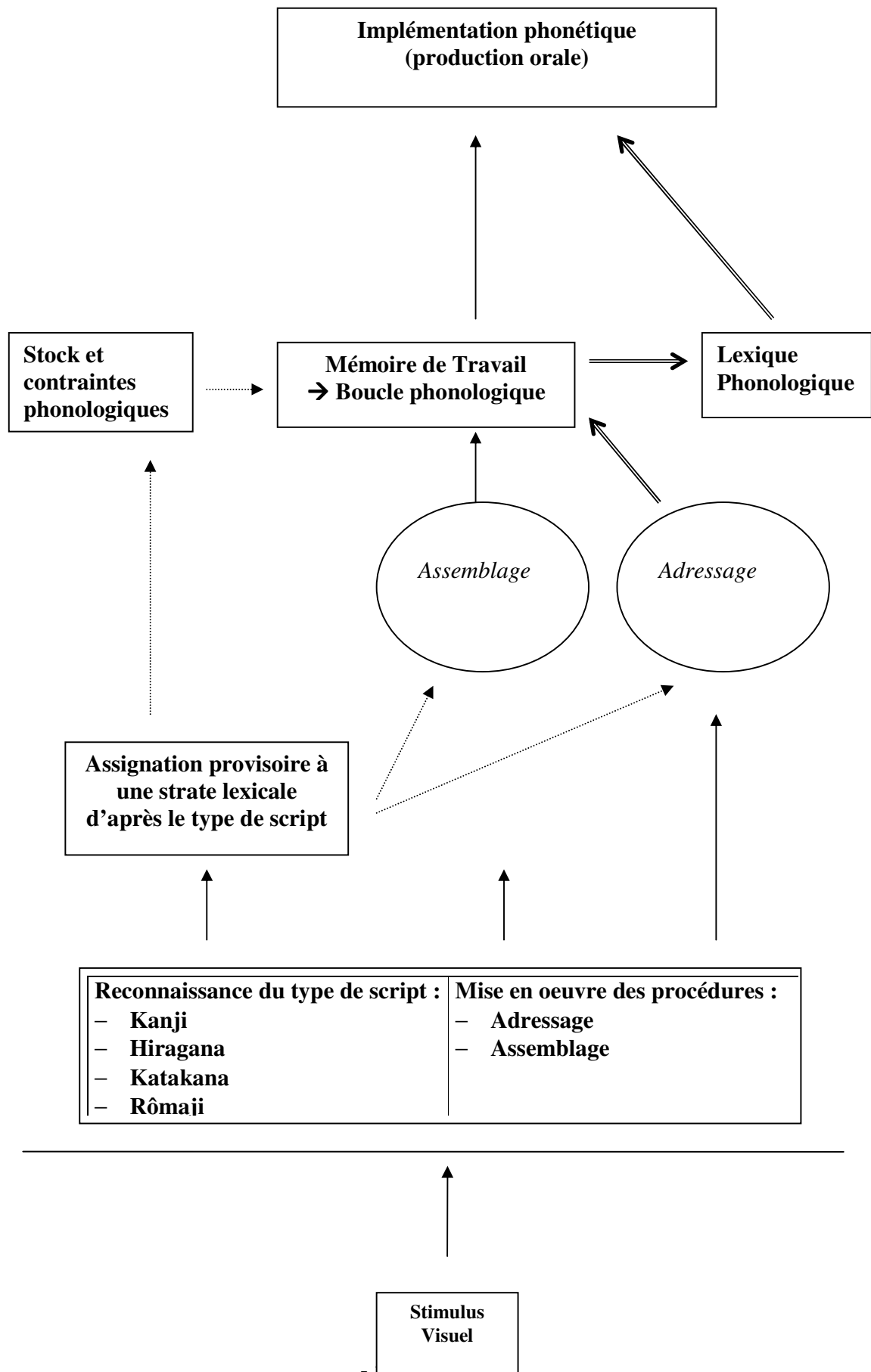
c) Lecture à voix haute vs. Lecture silencieuse : cette distinction pourrait apparemment sembler négligeable, en ce que l'on sait que même la lecture silencieuse implique l'activation de représentations phonologiques, et s'accompagne de toutes façons d'une certaine articulation subvocalique, comme cela apparaît nettement dans le modèle de mémoire de travail de Baddeley. Nous nous intéressons donc aux deux types de lecture. Mais, dans la perspective didactique qui est la nôtre, il convient néanmoins de souligner que les deux ne peuvent être considérées comme équivalentes : leur équivalence s'arrête en effet au niveau de l'encodage phonologique, puisque la lecture silencieuse ne conduit à pas l'activation complète du plan phonétique articulato-moteur. Si la différence peut donc sembler relativement négligeable du point de vue de la syllabation phonologique, il n'en est pas de même en ce qui concerne le domaine segmental, puisque l'absence de vocalisation effective supprime en effet pour les apprenants, d'une part une certaine partie de l'apprentissage moteur des organes articulatoires (mise en place complète des gestes articulatoires), d'autre part, du point de vue de la perception auditive de leurs propres productions (autorégulation), certaines oppositions relatives au voisement (puisque la lecture silencieuse est, en général, non-voisée), et surtout la dimension allophonique de certains phonèmes relative à leur distribution (notamment dans le cas du /R/ français). Il se pourrait donc que l'influence de la lecture silencieuse sur l'acquisition de la phonologie du français par les apprenants diffère en partie de celle de la lecture « à voix haute ».

Dans le cas du japonais, avec ses quatre codes orthographiques et la structure stratifiée de son lexique phonologique, si l'on souhaite intégrer la problématique des voies de

reconnaissance des mots (assemblage/adressage) à la boucle phonologique du modèle mnésique Baddeley, on peut, dans un premier temps esquisser les premières étapes du processus de lecture à voix haute de mots phonologiques de la manière suivante<sup>156</sup> :

---

<sup>156</sup> Il ne s'agit que d'une esquisse; il ne s'agit *pas* d'un schéma psycholinguistiquement fonctionnel.





Les flèches en pointillés correspondent à des « influences », les flèches pleines à des activations. Parmi ces activations, la flèche simple correspond à la voie d'assemblage, tandis que la flèche double correspond à la voie d'adressage, ces deux voies étant représentées par les ellipses.

Les étapes sont donc les suivantes (certaines sérielles, d'autres parallèles) :

<b>Etape 1)</b>		Perception visuelle du stimulus.
<b>Etape 2)</b>	<b>2)a)</b>	Assignation provisoire à une strate lexicale d'après le type de script : Kanji, Hiragana, Katakana, Rômaji.
	<b>2)b)</b>	Mise en œuvre parallèle des processus de reconnaissance, influencés par la strate assignée (2)a) : Adressage/Assemblage.
<b>Etape 3)</b>	<b>3)a)</b>	<i>Si l'adressage est premier</i> : alors il n'y a pas systématiquement de recodage phonologique dans la boucle phonologique (expertise des lecteurs). Le mot est reconnu, avec les contraintes associées, soit dans le lexique japonais pour les apprenants débutants, soit dans le lexique interphonologique pour les apprenants plus avancés (l'analyse dépend alors ici du modèle lexical adopté pour les apprenants).
	<b>3)b)</b>	<i>Si l'assemblage est premier</i> : alors il y a eu recodage phonologique via la boucle phonologique. Nous considérons pour l'instant que le stock phonologique utilisé (contraintes incluses) correspond à celui de la strate lexicale assignée provisoirement initialement. Les contraintes pesant sur le recodage sont donc celles de cette strate.

Cependant, un tel schéma reste tout à fait approximatif, et l'intégration, en particulier temporelle, des voies de lecture (adressage et assemblage) au modèle de mémoire de travail de Baddeley reste ici imparfaite.

Dans le cas qui nous occupe, les apprenants japonais sont confrontés à des mots inconnus écrits en alphabet latin. Trois possibilités d'identification existent alors en début de lecture :

- 1) Mot étranger/ Alphabet : identification d'un mot étranger écrit en alphabet latin (activation des procédures de conversion des emprunts et assignation à la strate lexicale étrangère (contraintes phonologiques « souples », variables selon les individus et les générations)).
- 2) Mot étranger/ Rômaji : identification d'un mot étranger écrit en romaji (activation du code katakana habituellement utilisé pour transcrire les mots étrangers et assignation à la strate lexicale étrangère (contraintes phonologiques « souples », variables selon les individus et les générations)).
- 3) Mot japonais/ Rômaji : identification d'un mot japonais écrit en romaji (activation du code kanji/hiragana habituellement utilisé pour transcrire les mots japonais et assignation à la strate lexicale « native » ou « sino-japonaise » (contraintes phonologiques « strictes »)).

Etant donné le contexte d'apprentissage des sujets, il semble raisonnable de privilégier la première, ou les deux premières, de ces trois possibilités. Dans ce cas, la tâche de lecture de mots inconnus implique, du pôle perceptif au pôle productif :

- La segmentation du mot en unités perceptives (orthographiques)
- La procédure d'identification et de conversion des segments (orthographe → phonologie) via la mémoire de travail
- La réalisation phonétique assemblée (mot) des unités phonologiques identifiées

Prenons deux exemples, l'un d'ordre segmental, l'autre d'ordre syllabique.

**Exemple 1 (segmental) : le non-mot « RAPIKO » (problème segmental : /R/)**

- **Segmentation** : « RA-PI-KO » (processus de segmentation « moraique » en japonais (Otake et al., 1993), identique, dans ce cas, à la segmentation syllabique en français pour des mots de basse fréquence (Ferrand et New, 2003)).

- **Identification et conversion** :

Deux facteurs interviennent ici :

- **Le degré d'expertise en français des apprenants** :

Celui-ci influence deux variables :

\**L'entropie graphémique* (Lange et Content, 1999) : plus le degré d'expertise est élevé, plus l'entropie graphémique est élevée, dans le cas qui nous occupe. Cependant, comme nous l'avons signalé, dans le cas d'un étudiant japonais universitaire apprenant le français, la lecture d'une syllabe comme « RA » comporte deux types d'entropie :

° L'une est *scripturale* et spécifique au japonais : c'est-à-dire que pour une syllabe orthographique donnée, celle-ci peut être associée à un script unique (le script rômaji pour un sujet débutant complet, avec une entropie scripturale minimale) ou à plusieurs scripts (scripts rômaji et alphabet latin, avec une entropie scripturale plus élevée). Cette entropie peut alors influencer la taille des unités à convertir : more orthographique en rômaji (« RA ») ou graphèmes en alphabet latin (« R-A »).

°L'autre est *graphémique* et correspond à celle mise en évidence par Lange et Content pour le français : plus l'apprenant est expert, plus la force d'association entre la syllabe orthographique « RA » et la syllabe phonologique française augmente, plus l'entropie graphémique (ou plutôt « syllabique ») augmente, puisque cette syllabe possède alors deux prononciations alternatives, l'une « japonaise » et l'autre « française ». Il apparaît alors que, dans la perspective qui nous intéresse, les deux types d'entropie sont étroitement liés : l'augmentation de l'entropie scripturale semble en effet faciliter celle de l'entropie graphémique, puisqu'une conversion en graphèmes simples (« R-A ») implique une prononciation « française » de la syllabe orthographique « RA » (car le graphème simple « R » n'existe pas en rōmaji). Une telle approche s'intègre tout à fait, dans une perspective sublexicale, au modèle interactif de Kroll et Dijkstra (2000), en considérant les apprenants comme des bilingues non équilibrés : pour des débutants, la langue première domine alors la langue seconde, mais le contexte (apprentissage du français) peut influencer l'interprétation des stimuli, et ainsi la conversion « française » de la syllabe orthographique « RA ».

Dans tous les cas, plus le sujet est expert, plus la force d'association entre la syllabe orthographique « RA » et la syllabe phonologique française /RA/ devrait être élevée, à condition que l'apprentissage de ces correspondances et celui du système phonético-phonologique du français aient lieu.

\**La conscience phonémique* (Bentin, 1992) : plus le degré d'expertise est élevé, plus la conscience phonémique est élevée, dans le cas qui nous occupe. Cela rejoint ce que nous avons appelé « l'entropie scripturale » : plus l'apprenant japonais est expert, plus sa capacité de segmentation phonémique (graphèmes simples) augmente, alors qu'elle est quasi-nulle chez un monolingue japonais. Or les graphèmes simples n'ont de correspondance qu'en français (ou d'autres

langues à système d'écriture alphabétique que connaît le sujet, comme l'anglais par exemple). Face à un mot inconnu comme « RAPIKO », alors que pour un sujet monolingue japonais la première unité minimale est « RA », un sujet apprenant expert est capable de segmenter cette unité en graphèmes français « R-A », ce qui favorise la conversion phonémique française.

Cependant, le cas du japonais présente une complexité particulière : alors que l'apprentissage d'une orthographe alphabétique est censé favoriser le développement de la conscience phonémique/ graphémique, l'apprentissage de l'alphabet latin, notamment lors de l'apprentissage de l'anglais par les écoliers japonais, entre, d'une certaine manière, en concurrence avec l'apprentissage, antérieur, des caractères rōmaji. En effet, alors que le système rōmaji utilise 22 des 26 lettres de l'alphabet, il est cependant utilisé pour transcrire un système moraique : l'unité minimale est bien « ra », et non « r » puis « a ». Ce chevauchement entre alphabet latin et rōmaji pourrait ainsi rendre compte des difficultés d'acquisition de la conscience phonémique/ graphémique par les Japonais.

- **Le degré d'attention des apprenants** : l'influence de l'attention sur l'identification et la conversion des unités segmentées, tout en interagissant avec celle du degré d'expertise des sujets, se ramène à celle de la nature, automatique ou contrôlée, des processus à l'œuvre. Pour des sujets novices, une faible attention conduira à l'activation de processus automatiques, et ainsi à la conversion la plus fréquente des graphèmes, c'est-à-dire la conversion « japonaise ». Plus l'expertise du sujet augmente, plus la demande attentionnelle, pour une conversion « française » devrait diminuer, jusqu'à ce que le choix de la conversion (« française ou japonaise ») ne dépende plus que du contexte, et soit automatisée, dans le cas de sujets bilingues équilibrés, selon le modèle de Kroll et Dijkstra (2000).

- **Réalisation phonétique assemblée** : une fois la conversion des unités effectuées, reste la question de la re-syllabification des unités et de leur réalisation phonétique. La distinction entre les plans phonologique et phonétique se pose alors, renvoyant à que nous avons déjà évoquée à propos de l'apprentissage, en particulier dans des modèles comme celui de Levelt et Wheeldon (1994). Les problèmes pouvant apparaître à ce niveau-là dépendent en partie des modèles adoptés, et ne concernent plus que très peu la problématique des rapports entre orthographe et phonologie.

**Exemple 2 (syllabique) : le non-mot « PSIKOMA » (problème syllabique : /ps/)**

- **Segmentation** : alors qu'un locuteur français proposerait la segmentation « PSI-KO-MA », la séquence « PSI » pose problème au locuteur japonais. Une telle séquence, inexistante en japonais, peut cependant être traitée grâce à la convention d'adaptation des mots étrangers partagée par les locuteurs japonais (Haunz, 2002). Cette convention, liée aux contraintes phonotactiques du japonais, intervient alors au cours du recodage phonologique en mémoire de travail de la séquence orthographique, puisque que, comme l'indiquent Gathercole, Frankish, Pickering et Peaker (1999), la mémoire à court-terme est sensible aux influences phonotactiques. Dans ce cas, la frontière entre les étapes de segmentation et de conversion est difficile à cerner : la segmentation orthographique serait initialement « PSI-KO-MA », mais la conversion phonologique donnerait lieu à une segmentation phonologique /PU-SI-KO-MA/, avec l'insertion d'une voyelle épenthétique /U/.
- **Identification et conversion** : comme nous l'avons vu, il y aurait donc ici une correspondance écrit-oral du type : « PSI » → /PUSI/ (adaptation des emprunts). L'apprentissage des correspondances écrit-oral est donc étroitement lié à celui de la structure syllabique du français. Les facteurs *d'entropie graphémique* et de *conscience phonémique* jouent également un rôle dans cet apprentissage, puisque la séquence « PSI » posséderait deux « prononciations » (/PUSI/ et /PSI/) et que la

capacité de segmentation phonémique (/P-S-I/) augmenterait avec la conscience phonémique.

- **Réalisation phonétique assemblée** : cette étape est identique à celle de l'exemple précédent.

En ce qui concerne l'influence de l'orthographe sur l'apprentissage du système phonologique français pour les apprenants d'un point de vue didactique dans une telle tâche, il apparaît ainsi que, en l'absence d'un apprentissage simultané des correspondances phonographémiques, la tâche de lecture peut « renforcer » celles existant en langue première, et ainsi, par rétroaction, faire obstacle à l'apprentissage des catégories phonémiques et des structures syllabiques françaises, de par l'activation du système phonologique japonais (et non français ou interphonologique) en tâche de lecture de séquences françaises. Par conséquent, dans le cas qui nous occupe, l'apprentissage de la phonologie du français semble donc devoir être effectué en relation avec celui des correspondances phonographémiques françaises.

### **III.2.2. Tâche de Répétition – unimodalité auditive**

Cette tâche implique :

- Les deux pôles perception/production ;
- En perception : reconnaissance intra-modale (auditif/phonologique → auditif/phonologique).

Cependant, il convient de distinguer deux types de répétition :

- **Répétition « périphérique »** : Phonétique → (Phonologique →) Phonétique
- **Répétition « profonde »** : Phonétique → Phonologique → Phonétique

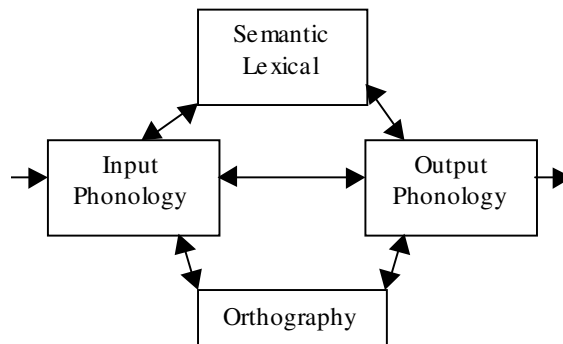
Il apparaît en effet, notamment à travers les données de la neurolinguistique, qu'il est possible de répéter une phrase que l'on ne comprend pas (il s'agit alors de ce que l'on pourrait qualifier de « psittacisme ») (Nespoulous et Dordain, 1994) et que certains

aphasiques souffrent de troubles phonétiques mais non phonologiques et vice-versa (Nespoulous, 2002). Il est également possible de répéter une séquence de « sons » inconnus, pouvant être illicite dans notre propre langue s'il s'agissait d'une séquence linguistique, et l'on sait que les zones cérébrales activées en perception auditive diffèrent selon la perception de la nature, linguistique ou non, des stimuli (Celsis, Doyon, Boulanouar et Nespoulous, 1997).

On pourrait considérer que, dans le cadre d'une étude sur l'apprentissage, seul le type « profond » est pertinent, puisqu'il semble être le seul à mettre en œuvre une véritable procédure de reconnaissance. Cependant, le temps où les modèles, psycholinguistiques entre autres, ne pouvaient se dégager du carcan binaire du tiers exclu est révolu : sans remettre en cause l'intérêt incontournable et la validité des approches modulaires qui, en particulier sur la base de phénomènes tels que les (doubles) dissociations, ont permis des avancées considérables, il apparaît aujourd'hui que ces modules, sans perdre ni leur modularité ni leur autonomie relative, ne sont pas aussi rigides que certains auraient pu le croire : la variabilité des phénomènes d'une part et les multiples influences qui s'y exercent d'autre part, font le jeu des approches, sinon neuromimétiques, du moins connexionnistes, dont certaines reconnaissent, à une certaine échelle, la modularité de l'architecture cognitive. La cognition semble ainsi de plus en plus ne pas fonctionner sur un mode binaire de Tout-Ou-Rien, mais bien sur un mode scalaire, voire statistique, comme cela a pu être formalisé dans les théories de logiques polyvalentes, et plus particulièrement dans la théorie des ensembles flous, sur lequel se basent précisément les travaux de Massaro en perception de la parole. Il serait donc déraisonnable d'affirmer que la répétition « périphérique » n'active *aucunement* le domaine phonologique : il serait plus prudent de dire que cette activation est optionnelle, et qu'elle doit être caractérisée de manière scalaire (fonction des seuils d'activation et des poids de connexion) plutôt que binaire (activé ou pas). Nous nous intéressons donc aux deux types de répétition.



Castro-Caldas et ses collègues (1998) proposent de schématiser ainsi les processus à l'œuvre en tâche de répétition verbale, une représentation de type parallèle adaptée des travaux de McClelland et de Seidenberg :



Deux cas sont à distinguer :

- Répétition de Mots : le traitement est orienté vers le système lexicosémantique dans un mode de traitement phonologique implicite.
- Répétition de Non-Mots : le traitement est orienté dans un mode de traitement phonologique explicite.

Dans le cas de nos apprenants japonais de niveau débutant, nous nous intéressons essentiellement au deuxième cas, puisque si le mot français est perçu, assimilé et identifié à un mot japonais (par exemple un mot d'emprunt, comme le mot « football »), alors la forme phonologique lexicale japonaise est activée, ce qui s'explique très bien dans un modèle de type Cohorte ou encore dans le modèle d'activation interactive de Kroll et Dijsktra (2000).

D'un point de vue didactique, on doit alors noter que la répétition de mots d'emprunts ou de mots « cognates », pour des apprenants débutants, peut s'avérer à la fois avantageuse et néfaste pour le développement de l'interphonologie : avantageuse, car l'input est « compréhensible », permettant aux apprenants de se concentrer sur sa forme phonético-phonologique ; néfaste, car si l'attention des apprenants n'est pas

précisément portée sur cette forme en langue étrangère, alors de tels mots sont traités automatiquement via le système lexicosémantique dans un mode de traitement phonologique implicite, perpétuant ainsi, à un niveau lexical et sublexical, l'équivalence entre système phonologique de L1 et l'interphonologie.

Ce qui, en revanche, est particulièrement remarquable dans le schéma de Castro-Caldas *et al.* (1998) est le lien entre « input phonologique », « orthographe » et « output phonologique ». Comme nous l'avons vu, il semblerait en effet que, dans certaines conditions, les représentations orthographiques des sujets puissent influencer la tâche de répétition. Reprenons alors les deux exemples précédemment utilisés.

**Exemple 1 (segmental) : le non-mot /RAPIKO/ (problème segmental : /R/)**

Face à un stimulus auditif comme [RAPIKO], présenté comme un mot français (ce qui exclut a priori pour des sujets débutants une quelconque reconnaissance lexicale), il faut d'abord noter que la répétition problématique du phonème /R/ est liée à celle de la syllabe /RA/, puisque la segmentation naturelle d'un tel non-mot sera moraique (et ici syllabique) pour un apprenant japonais novice : [RA-PI-KO] et non [R-A-P-I-K-O]. L'identification et la répétition de la première syllabe/more [RA] peuvent alors emprunter « plusieurs voies » :

- Voie phonétique : l'apprenant perçoit en attaque de la syllabe une catégorie phonétique [R] n'existant pas en japonais et s'efforce de la reproduire par approximation audioarticulatoire. Une telle « voie » pourrait favoriser, sur la base d'exemplaires phonétiques, l'émergence d'une nouvelle catégorie phonologique.
- Voie phonologique : l'apprenant identifie de la syllabe /RA/ comme une syllabe japonaise. D'un point de vue acoustique, la syllabe japonaise [HA] semblerait ressembler le plus à la syllabe française [RA], mis à part la différence de voisement en position initiale prévocale, qui est l'une des plus notables (l'allophone dévoisé du /R/ français est en effet également allophone du /H/ japonais).

Pourtant, c'est généralement (sur la base d'observations empiriques) la syllabe japonaise /ra/ qui est identifiée à la syllabe /RA/. Si le voisement constitue certainement un facteur phonologique non-négligeable conduisant à l'assimilation perceptive de [RA] et [ra], le schéma présenté plus haut souligne une autre influence possible : celle des représentations orthographiques. Comme nous l'avons vu plus haut, en particulier dans le cadre du modèle de Massaro, l'identification phonétophonologique fait appel à l'ensemble des informations disponibles aux sujets : acoustiques et visuelles pour le bas-niveau, contextuelles, phonotactiques et autres pour le haut-niveau.

Si l'on considère que le voisement constitue un indice inhibiteur pour la syllabe japonaise /HA/, on peut alors s'interroger sur les facteurs qui conduisent à l'identification de /ra/ et non à la perception d'une catégorie phonétique inconnue. La réponse pourrait se trouver, nous semble-t-il, dans le domaine phonographémique. Partant d'un principe général d'économie cognitive, il apparaît que « confronté à l'inconnu », l'assimilation à du « connu » et ainsi la mise en œuvre de procédures automatiques est moins coûteuse (d'un point de vue articulatoire et attentionnel notamment), que le travail d'approximation audioarticulatoire précédemment évoqué. L'importance des facteurs attentionnels et motivationnels est ici évidente. D'un point de vue phonologique, on peut bien sûr considérer que le /R/ français et le /r/ japonais sont toutes deux des liquides, ce qui constituerait un facteur de similarité non-négligeable.

Cependant, dans un contexte *réaliste* d'apprentissage du français pour la population qui nous occupe, il faut savoir que, sauf cas exceptionnel, les apprenants japonais sont conscients très tôt de l'existence d'un /R/ français tout à fait différent du /r/ japonais (à travers le premier cours de français, ou encore à travers les médias). Cependant, la similarité *orthographique* des deux phonèmes, conduit généralement très tôt les apprenants à identifier la syllabe [RA] comme l'équivalent phonétique de la syllabe orthographique « RA », laquelle est associée à la syllabe phonologique /ra/. La

catégorie phonétique [R] (en perception) serait alors graduellement comprise dans la catégorie phonologique /r/ de leur interphonologie, dont l'implémentation phonétique prototypique (en production) automatique (et donc peu coûteuse) pour les Japonais est [r]. Il se pourrait donc que les étudiants qui répètent [ra] au lieu de [RA] soient précisément influencés par leurs connaissances phonographémiques. Une telle interprétation n'exclut pas l'influence de facteurs purement phonologiques, mais elle souligne la possibilité que les facteurs phonographémiques peuvent prendre part à ce processus d'identification erronée.

Par ailleurs, l'acquisition du /R/ français est généralement liée à celle de l'opposition entre /R/ et /l/ en français. Or, le [l] étant un allophone du /r/ japonais, il est très souvent transcrit « r ». La transcription du /l/ en « r » conduit ainsi à une relative équivalence phonologique du /l/ et du /R/ interphonologique, via le graphème « r » : [l] et [R] constituent alors des allophones d'une unique catégorie phonologique (/r/ japonais ?), tous deux transcrits « r ». L'opposition phonético-phonologique entre /l/ et /R/ en français se trouverait alors neutralisée dans le système interphonologique en raison d'une équivalence orthographique supposée ("l" serait alors un allographe du graphème « r »).

On voit ainsi que les difficultés rencontrées par les Japonais face au /R/ français ne sont pas exactement identiques à celles qu'ils rencontrent face au /r/ anglais américain, par exemple, phonétiquement beaucoup plus proche du /r/ japonais que le /R/ français. Mais dans ce cas également, nous suspectons que les connaissances phonographémiques des locuteurs jouent un rôle important, si l'on considère que, à l'instar du « Transfert Plein » initiale du système phonologique L1 dans celui de l'interphonologie, il existe un « Transfert Plein » similaire en ce qui concerne le système phonographémique. On pourrait alors considérer que, d'une certaine manière et dans un premier temps au moins, la reconnaissance phonographémique peut constituer un obstacle à la répétition.

**Exemple 2 (syllabique) : le non-mot « PSIKOMA » (problème syllabique : /ps/)**

Face à un stimulus auditif comme [PSIKOMA], présenté comme un mot français (ce qui exclut a priori pour des sujets débutants une quelconque reconnaissance lexicale), plusieurs voies sont à nouveau possibles :

- Voie strictement acoustique : l'apprenant perçoit une séquence de « sons » qu'il essaye de reproduire par approximations audioarticulatoires. Une telle démarche favoriserait un certain apprentissage *phonétique* (l'apprenant est capable de percevoir et de reproduire une séquence sonore).
- Voie phonético-phonologique : au cours de l'identification des séquences phonologiques, sur une base acoustique, il est possible que l'apprenant procède à un effacement consonantique : /SIKOMA/ ou /PIKOMA/ par exemple. /SIKOMA/ présente alors un problème nouveau problème phonotactique, puisque la séquence /SI/ n'est pas légale en japonais. Mais ce qui importe ici, c'est de souligner que, de par la nature transitoire de l'input acoustique, il est possible que les apprenants japonais n'accordent d'attention qu'à l'un ou l'autre segment consonantique, en fonction de leurs caractéristiques acoustiques, étant donné que la structure phonologique /CCV/ est illégale en japonais : la nature phonétiquement transitoire du stimulus combinée aux « attentes » phonologiques des auditeurs japonais peuvent ainsi conduire à un effacement consonantique.

Cependant, si les apprenants ne peuvent acoustiquement négliger l'un ou l'autre des segments consonantiques, une autre stratégie d'encodage est de considérer que la séquence phonétique [PSI] correspond à la séquence phonologique /PUSI/, ce qui se justifie d'une part par la nature de la voyelle /U/ en japonais (voyelle neutre en japonais et sujette à dévoisement en japonais) et d'autre part par la convention d'adaptation phonologique des mots d'emprunts (insertion d'une voyelle épenthétique entre deux consonnes). D'un point de vue didactique, il est donc possible que les apprenants effectuent un apprentissage *phonétique* (dévoisement

de la voyelle [U] donnant l'impression d'une séquence consonantique), sans effectuer d'apprentissage *phonologique* (la structure phonologique restant /CVCV/ et non /CCV/)

En ce qui concerne l'influence de l'orthographe dans un tel exemple, on doit noter ici que la structure moraique de l'écriture japonaise (en particulier en katakana, qui est le script utilisé pour les mots étrangers) semble être un obstacle majeur à l'acquisition de la structure syllabique du français, puisque, si une représentation orthographique est activée chez les auditeurs lors de l'écoute du stimulus, c'est précisément une représentation en katakana (comme en témoignent les « notes » prises en katakana par les étudiants japonais de français pour transcrire les mots inconnus entendus en classe), voire en rōmaji, laquelle ne tolère, à quelques exceptions près, aucune structure consonantique complexe de type /CC/. Outre les contraintes phonotactiques du japonais, le système d'écriture japonais, qui les reflète en partie de par la structure moraique des syllabaires kana, constitue donc un obstacle supplémentaire à l'acquisition de la syllabation du français, puisqu'il ne permet aucune représentation orthographique appropriée des séquences française. Cela rejoint alors en partie les conclusions de Byung-jing (2004).

Ceci étant clarifié, il nous faut à présent considérer la tâche la plus complexe, celle de lecture/répétition bimodale.

### **III.2.3. Tâche de Lecture/Répétition – bimodalité audiovisuelle**

Si l'on examine une tâche complexe de perception bimodale comme celle-ci, quelle « étiquette » doit-on lui assigner : « répétition avec support visuel » ou « lecture à voix haute avec support auditif » ? Ainsi comment doit comprendre l'affirmation de Schwartz, Abry, Boë, et Cathiard (2002, p. 264) qui se réfèrent aux travaux de Davis et Kim (1998)<sup>157</sup> : « Repetition is more efficient when the visual input is present [...] if it

---

<sup>157</sup> Davis, C. et Kim, J (1998), « Repeating and remembering foreign language words : does seeing help ? », *Proceedings of Audio-Visual speech Processing 98*, Sydney, 121-5.

is uttered in a foreign language ». De quelle *efficacité* s'agit-il ? Le titre de la communication de Davis et Kim (« Repeating and remembering foreign language words ») est sans doute éclairant : il semblerait bien s'agir de mémorisation. Dire que le support écrit facilite la répétition est une affirmation un peu rapide : facilitation mnésique (ressource) peut-être ; facilitation de traitement (processus) pas forcément dans tous les cas.

On constate par ailleurs certaines « doubles dissociations » entre les tâches de lecture à voix haute et de répétition (Nespoulous et Dordain, 1994) : certains patients répètent mieux qu'ils ne lisent à voix haute et vice-versa. L'une des distinctions sans doute essentielles entre les deux modalités réside dans la nécessité d'activer ou non des unités « profondes » : tandis qu'en répétition auditive unimodale, les sujets peuvent théoriquement traiter les stimuli de manière strictement acoustique / phonique et non-linguistique, ce qu'on pourrait qualifier de psittacisme et qui, d'un point de vue (neuro-)psycholinguistique correspondrait à une activation de zones non-linguistiques ou en tous cas « périphériques », la tâche de lecture à voix haute implique, en raison de la « permanence » relative du stimulus visuel, à relier à la dimension attentionnelle, l'identification d'unités « profondes » qui sont associées aux stimuli. Cela revient à considérer que, chez des sujets non-déficients, les représentations orthographiques, malgré la variation *graphique*, constituent en quelque sorte l'équivalent visuel des « prototypes » phonologiques : contrairement au stimuli acoustiques, qui peuvent être associés à différentes valeurs phonétiques selon le contexte et les individus, un stimulus orthographique (graphème) serait inévitablement associé à l'exemplaire typique, pour un auditeur/locuteur donné, d'une catégorie phonologique.

Si l'on adopte un modèle interactif de l'architecture cognitive sous forme de réseaux d'activations, modèle notamment justifié par les nombreux travaux récents en (neuro-)psychologie portant sur la perception bimodale et indiquant l'existence d'effets d'intégration ou de modulation des modalités perceptives entre elles, il nous faut donc considérer deux situations en tâche de perception bimodale :

- 1) Dominance auditive : tâche de répétition influencée par les stimuli orthographiques ;
- 2) Dominance visuelle : tâche de lecture à voix haute influencée par les stimuli phonético-phonologiques.

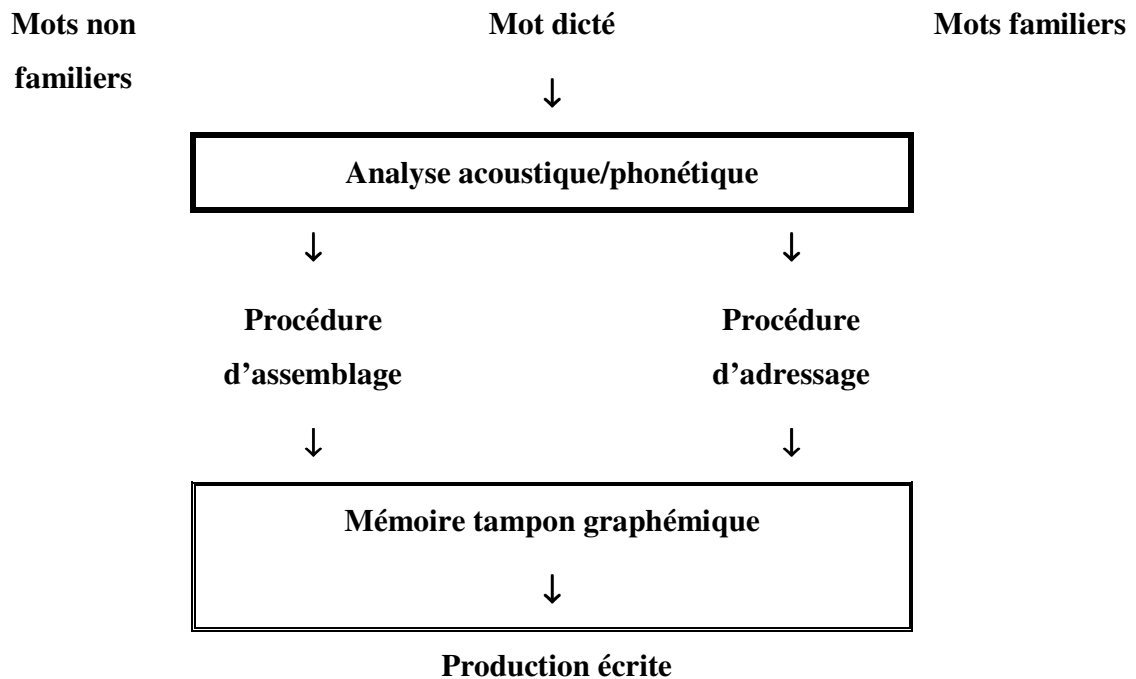
Bien sûr, cela n'exclut pas l'existence de tâches relativement équilibrées. Les questions qui se posent sont alors les suivantes : quels sont les facteurs déterminant la voie dominante ? Des facteurs individuels ? Collectifs ? Contextuels ? Identifier ces facteurs et en déterminer l'influence font donc partie du travail d'établissement des schémas d'activation, nécessaire à une tentative d'explication des comportements observés dans notre étude.

D'un point de vue didactique, si l'objectif est l'acquisition de catégories phonologiques (à ne pas confondre avec celles des relations phonographémiques) grâce à une tâche de répétition, l'intérêt ou le danger de fournir un support orthographique dans une telle tâche doit donc être examiné à travers la fonction que l'on souhaite faire jouer à ce support.



### **III.2.4. Tâche de Dictée Orthographique**

Zesiger et Partz proposent le schéma suivant pour aborder la tâche de génération de séquences de lettre (1994, p. 423) :



La tâche de dictée constitue donc une complexification de la tâche de répétition auditive unimodale. Cependant, alors que la répétition offre la possibilité d'un traitement « superficiel », la tâche de dictée oblige l'apprenant à identifier des graphèmes, et ainsi à effectuer des choix phonologiques catégoriels (par exemple entre les couples (/R/ ; « R ») et (/L/ ; « L »)).

D'un point de vue didactique, on pourrait considérer que, à travers le « feedback » de l'enseignant, relatif aux correspondances phonographémiques, une telle tâche permettrait aux apprenants, dans certains cas, d'acquérir de nouvelles catégories phonétiques, préalables aux catégories phonologiques. Par exemple, dans le cas du couple français R/L, on pourrait considérer que, dans un premier temps, pour les

raisons expliquées plus haut, les catégories phonétiques [R] et [L] soient associées à une unique catégorie phonologique /r/, transcrite « R ». L'introduction d'un nouveau symbole orthographique « L » (qui n'appartient pas au système rômaji), *plus facilement perceptible* qu'une nouvelle catégorie phonétique, pourrait alors permettre l'émergence d'une frontière intra-catégorielle phonétique entre [R] et [L]. Mais comme nous l'avons vu, la multiplicité des influences réciproques entre codes orthographiques et phonétio-phonologiques en L1 et en L2 doit être soigneusement prise en compte, avant de concevoir de tels processus isolés.

A travers cette dernière tâche, il apparaît ainsi que l'influence de l'orthographe sur l'acquisition du système phonologique étranger peut être *a priori* positive ou négative selon les tâches et les contextes. Plutôt négative en perception, elle pourrait en revanche être positive en production, à condition de disposer de retour correctif approprié.

### **III.3. Influence de l'orthographe : niveau segmental et identification**

Si l'on prend comme point de départ les conclusions de Dijkstra, Frauenfelder et Schreuder (1993), il y aurait, dans les tâches concernées, activations automatiques bidirectionnelles entre graphèmes et phonèmes, ce qui soutiendrait l'hypothèse selon laquelle les graphèmes joueraient un rôle dans la reconnaissance auditive de mots et les phonèmes dans la reconnaissance visuelle de mots, allant ainsi dans le sens de certains modèles à double voie de reconnaissance visuelle.

Dans notre cas, si l'on met de côté, temporairement au moins, la complexité des lexiques phonographémiques français et japonais ((pseudo)homophones, (pseudo)homographes, mots irréguliers, strates lexicales, emprunts récents, etc.), et si l'on se situe bien à un niveau sublexical, on peut proposer le schéma suivant comme base de description, de manière très simplifiée, des processus à l'œuvre en tâche de reconnaissance visuelle et auditive pour nos sujets :

<b>Reconnaissance</b>			
<b>Visuelle</b>		<b>Auditive</b>	
<b>Stimuli français</b>	<b>Traitement japonais</b>	<b>Stimuli français</b>	<b>Traitement japonais</b>
« b »	Licite → « b » → /b/	[b]	Licite → « b »
« v »	Illicite → « b » → /b/	[v]	Illicite → /b/ → « b »
« l »	Illicite → « r » → /r/	[l]	Licite mais → /r/ → « r »
« r »	Licite mais → /r/	[R]	Illicite → /r/ → « r »
« e » ou « eu »	Licite mais → /e/ ou /e+/ɯ/	[EU]	Illicite → /ɯ / → « u »
« ou »	Licite mais → /o+/ɯ/	[OU]	Licite → « u »

Les flèches représentent les activations, présentées ici de manière unidirectionnelle.

A vrai dire, une telle présentation est trompeuse, puisque le système orthographique japonais est bien moraique : face au mot « bato », c'est bien le graphèmes « ba » qui est activé, et non un hypothétique graphème « b » associé au graphème « a ». Cependant, cette première approche nous permet de réexaminer les substitutions observées empiriquement qui ont motivé notre démarche.

### **III.3.1. Traitement psycholinguistique « erroné » des unités phonéto-phonologiques et orthographiques**

Le tableau ci-dessous correspond à une première tentative d'explication des phénomènes observés ; nous verrons par la suite qu'il n'est pas totalement satisfaisant.

<p align="center"><b>Reconnaissance / Perception Visuelle</b> (Lecture à voix haute)</p>	<p align="center"><b>Reconnaissance / Perception Auditive</b> (Répétition)</p>
<p align="center"><b><u>« b » réalisé [β]</u></b></p> <p><u>Origine phonético-phonologique</u> : « b » active /b/ qui dans certains contextes est réalisé [β] (transfert phonotactique L1 / L2).</p> <p><b>Substitution allophonique : /b/ → [β]</b></p>	<p align="center"><b><u>/b/ réalisé [β]</u></b></p> <p><u>Origine phonético-phonologique</u> : [b] active /b/ qui dans certains contextes est réalisé [β] (transfert phonotactique L1 / L2).</p> <p><b>Substitution allophonique : /b/ → [β]</b></p>
<p align="center"><b><u>« v » réalisé [b] (réalisation principale)</u></b></p> <p><u>Origine phonographémique</u> : « v » active /b/ qui active « b » qui est réalisé [b] dans tous les cas.</p> <p><b>Substitution graphème-phonème : « v » → /b/</b></p> <p>Il faut en effet se rappeler que le graphème « v » n'existe pas en « romaji » : la valeur phonologique qui lui est généralement assignée dans la strate lexicale des emprunts correspond à celle du romaji « b ».</p>	<p align="center"><b><u>/v/ réalisé [b] (réalisation principale)</u></b></p> <p>A priori, deux explications seraient envisageables :</p> <p>a) <u>Origine phonético-phonologique</u> : [v] est perçu comme [β] qui active /b/ et est réalisé [b] sous l'effet de contraintes phonotactiques dans le contexte où /b/ → [b].</p> <p><b>Substitution perceptive : [v] → [β]</b></p> <p>b) <u>Origine phonographémique</u> : /v/ active /b/ qui active « b » qui est réalisé [b] dans tous les cas.</p> <p>En effet, il faut se rappeler que le « romaji » « b » se « prononce » [b] (c'est sa valeur prototypique : [β] est un allophone de /b/ mais le « romaji » « b » NE SE prononce PAS (canoniquement) [β]).</p>

	<p><b>Substitution phonème-graphème : /v/ → « b »</b></p> <p>Dans les contextes phonologiques où /b/ pourrait / devrait être réalisé [β], si l'explication a) était privilégiée, il y aurait <b>conflit</b> entre d'un côté certaines contraintes phonotactiques (qui s'appliqueraient par exemple en tâche de répétition) et de l'autres certaines contraintes relatives aux codes de conversion graphèmes-phonèmes (en lecture à voix haute).</p> <p>On voit que dans ce cas, seule l'explication b) permettrait de rendre compte de la réalisation de /v/ en [b] dans tous les contextes (contextes /b/ → [b] et /b/ → [β]).</p> <p>C'est donc bien d'une substitution phonème-graphème qu'il s'agirait.</p>
<p><b><u>« v » réalisé [β] (réalisation occasionnelle)</u></b></p> <p><u>Origine phonographémique</u> : « v » active /b/, qui est réalisé [β] sous l'effet de contraintes phonotactiques dans le contexte où /b/ → [β].</p> <p><b>Substitution graphème-phonème : « v » → /b/</b></p> <p>Il faut en effet se rappeler que le graphème</p>	<p><b><u>/v/ réalisé [β] (réalisation occasionnelle)</u></b></p> <p><u>Origine phonético-phonologique</u> : [v] est perçu comme [β] qui active /b/ et est réalisé [β] sous l'effet de contraintes phonotactiques.</p> <p><b>Substitution perceptive : [v] → [β]</b></p>

<p>« v » n'existe pas en « rômaji » : la valeur phonologique qui lui est généralement assignée dans la strate lexicale des emprunts correspond à celle du rômaji « b ».</p> <p>Une autre explication envisageable, quoique sans fondement <i>a priori</i>, serait d'imaginer que, dans le système interlangagier des apprenants, le graphe « v » correspondrait au phonème [β] qui aurait acquis une valeur phonémique propre /β/. Le système phonographémique interlangagier serait alors le suivant :</p> <p>« b » → /b/ → [b]          « v » → /β/ → [β]</p> <p>On pourrait alors considérer que « l'erreur » provient d'une substitution « phonémique » /v/ → /β/.</p>	
<p style="text-align: center;"><b><u>« l » réalisé [r]</u></b></p> <p><u>Origine phonographémique</u> : « l » active /r/ qui active « r » qui est réalisé [r] dans tous les cas.</p> <p><b>Substitution graphème-phonème</b> : « l » → /r/</p> <p>Il faut en effet se rappeler que le graphème « l » n'existe pas en « rômaji » : la valeur</p>	<p style="text-align: center;"><b><u>/l/ réalisé [r]</u></b></p> <p>A priori, deux explications seraient envisageables :</p> <p>a) <u>Origine phonético-phonologique</u> : [l] est perçu comme [r] qui active /r/ et est réalisé [r] sous l'effet de contraintes phonotactiques dans le contexte où /r/ → [r].</p>

<p>phonologique qui lui est généralement assignée dans la strate lexicale des emprunts correspond à celle du romaji « r ».</p>	<p><b>Substitution perceptive : [l] → [r]</b></p> <p>b) <u>Origine phonographémique</u> : /l/ active /r/ qui active « r » qui est réalisé [r] dans tous les cas.</p> <p>En effet, il faut se rappeler que le « romaji » « r » se « prononce » [r] (c'est sa valeur prototypique : [l] est un allophone de /r/ mais le « romaji » « r » NE SE prononce PAS (canoniquement) [l]).</p> <p><b>Substitution phonème-graphème : /l/ → « r »</b></p> <p>Dans les contextes phonologiques où /r/ pourrait / devrait être réalisé [l], si l'explication a) était privilégiée, il y aurait <b>conflit</b> entre d'un côté certaines contraintes phonotactiques (qui s'appliqueraient par exemple en tâche de répétition) et de l'autres certaines contraintes relatives aux codes de conversion graphème-phonème (en lecture à voix haute).</p> <p>On voit que dans ce cas, seule l'explication b) permettrait de rendre compte de la réalisation de /l/ en [r] dans tous les contextes (contextes /r/ → [r] et /r/ → [l]).</p> <p>C'est donc bien d'une substitution phonème-graphème qu'il s'agirait.</p>
--	---

<u>« R » réalisé [r] (réalisation principale)</u>	<u>/R/ réalisé [r] (réalisation principale)</u>
<p><u>Origine phonographémique:</u> « R » active /r/ qui active « r » qui est réalisé [r] dans tous les cas. En effet il faut se rappeler que [r] est la réalisation prototypique du rômaji « r ».</p> <p><b>Substitution graphème-phonème : « R » → /r/</b></p>	<p><u>Origine phonético-phonologique et phonographémique :</u> /R/ active /r/ qui active « r » qui est réalisé [r] dans tous les cas.</p> <p><b>Substitution « phonémique » : /R/ → /r/</b></p> <p>Cette explication ne semble cependant pas satisfaisante, car on peut se demander quels sont les éléments / indices justifiant l'activation directe de /r/ par /R/ :</p> <ol style="list-style-type: none"> <li>1) Il ne peut s'agir d'une substitution perceptive, puisque [R] et [r] sont difficilement identifiables l'un à l'autre, [h] étant la réalisation d'un phonème japonais la plus proche du [R] français ;</li> <li>2) L'autre source d'information serait distributionnelle (statistique) et non phonétique (perceptive), mais ici encore se poserait le problème de la variation allophonique en japonais : si /R/ active /r/, alors les réalisations de /r/ devraient varier entre [r] et [l] en fonction du contexte phonético-phonologique. Dans le cas qui nous occupe ici (/R/ réalisé [r]), une telle variation n'apparaît pas, et cette absence d'alternance ne peut s'expliquer que par l'activation du graphème « r ».</li> </ol>



<b><u>« R » réalisé [l] (réalisation occasionnelle)</u></b>	<b><u>/R/ réalisé [l] (réalisation occasionnelle)</u></b>
<p><u>Origine phonographémique:</u> « R » active /r/ qui est réalisé [l] dans le contexte où /r/ → [l].</p> <p><b>Substitution graphème-phonème : « R » → /r/</b></p>	<p><u>Origine phonético-phonologique :</u> /R/ active /r/ qui est réalisé [l] dans le contexte où /r/ → [l].</p> <p><b>Substitution « phonémique » : /R/ → /r/</b></p> <p>Ici encore se pose le problème de l'activation directe de /r/ par /R/.</p> <p>Ce problème soulève un certain degré de complexité négligé jusqu'ici.</p>

Ce premier examen phonographémique des substitutions observées nous a permis d'avancer quelques propositions d'explications, simplistes en apparence au moins. Cependant, les derniers points abordés, la réalisation de /R/ en [r] et [l], ont particulièrement mis à jour les limites, au moins partielles, de nos premières propositions. Plusieurs points doivent être discutés, afin de distinguer ce qui relève des pôles phonético-phonologique et phonographémiques:

### **III.3.2. Variation et alternance inter-tâches et intra-tâche**

L'une des pistes à suivre est alors celle de la question de l'alternance allophonique/phonémique dans une approche systémique, et non discrète, des substitutions observées : il convient en effet de considérer le système de substitutions des sujets, *intra-tâche* (lignes dans le tableau précédent) et *inter-tâches* (colonnes dans le tableau précédent), plutôt que d'examiner les substitutions de manière isolée. Par exemple, dans le cas du /l/ auditif (intra-tâche) : s'il n'y a pas alternance intra-tâche, c'est-à-dire, si le /l/ est systématiquement réalisé [r] quel que soit le contexte phonético-phonologique, alors nous avons vu que, *a priori*, seule l'explication phonographémique semblait à même de rendre compte de ce comportement. En

revanche, s'il y a alternance intra-tâche (réalisation en [r] dans les contextes où en japonais /r/ → [r] et réalisation en [l] dans les contextes où en japonais /r/ → [l]), alors c'est bien l'explication phonético-phonologique qui semble devoir être privilégiée.

A cette question d'alternance ou de non-alternance intra-tâche, s'ajoute celle de l'alternance inter-tâches, qui peut donner un éclairage nouveau à la première. Si l'on prend l'exemple du (« v » ; /v/) et si l'on reprend nos premières propositions d'explication, voici ce que l'on obtient :

Alternance	OUI	NON
<b>Intra-tâche</b>	/v/ réalisé [b] ET /v/ réalisé [β] DONC Substitution perceptive ([v] → [β]) ET Substitution perceptive ([v] → [β]) DONC Substitution perceptive (une voie : phonétique-phonologique)	Il s'agit des quatre possibilités présentées dans le tableau précédent.
<b>Inter-tâches</b>	Deux possibilités : <u>1) « v » réalisé [b] ET /v/ réalisé [β] (et donc aussi [b] selon le contexte) :</u> DONC Substitution graphème-phonème (« v » → /b/) ET Substitution perceptive ([v] → [β]) DONC Pas de conclusion unique (deux voies : grapho-phonémique et phonétique-phonologique ?)	Deux possibilités : <u>1) « v » réalisé [b] ET /v/ réalisé [b] :</u> DONC Substitution graphème-phonème (« v » → /b/) ET Substitution phonème-graphème (/v/ → « b ») DONC Substitution phonographémique bidirectionnelle (une voie : bidirectionnelle graphème/phonème)

	<p><u>2) « v » réalisé [β] (et donc aussi [b] selon le contexte) ET /v/ réalisé [b] :</u></p> <p>DONC</p> <p>Substitution graphème-phonème          (« v » → /b/)</p> <p>ET</p> <p>Substitution phonème-graphème          (/v/ → « b »)</p> <p>DONC</p> <p>Substitution phonographémique bidirectionnelle          (une voie : bidirectionnelle graphème/phonème)</p>	<p><u>2) « v » réalisé [β] (et donc aussi [b] selon le contexte) ET /v/ réalisé [β] (et donc aussi [b] selon le contexte) :</u></p> <p>DONC</p> <p>Substitution graphème-phonème          (« v » → /b/)</p> <p>ET</p> <p>Substitution perceptive          ([v] → [β])</p> <p>DONC</p> <p>Pas de conclusion unique          (deux voies : grapho-phonémique et phonétique-phonologique ?)</p>
--	---	--

Derrière cet examen, se trouve une hypothèse forte : celle selon laquelle la représentation orthographique (par exemple « v ») ne possède qu'une valeur phonétique (réalisation [b]) tandis que la représentation phonologique en possède deux (réalisations [b] et [β]). Selon la présence ou l'absence d'alternance allophonique, on pourrait donc déterminer la nature (orthographique ou phonologique) de la représentation « dominante », c'est-à-dire celle qui conduit à une réalisation phonétique. Il apparaît alors, que cette première présentation réductrice et qu'il nous faut détailler l'organisation interne de ces représentations supposées, de type (« R » ; /R/ ; [R]).

Avant d'aborder ce point ci-dessous, notons qu'une dernière possibilité d'explication, sans doute abusivement simplificatrice, serait de considérer que la variation allophonique est attribuable à des « ratés » phonétiques et tout à fait négligeable d'un

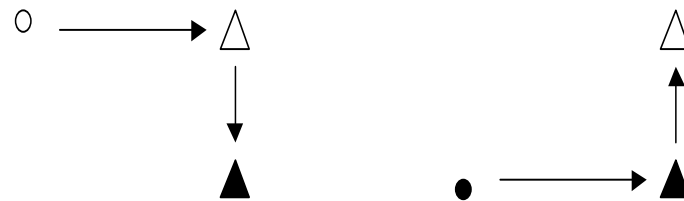
point de vue phonologique (dans ce cas on ne prendrait en compte que les substitutions (« v » ; /v/) → [b], (« l » ; /l/) → [r] et (« R » ; /R/) → [r]). Cela simplifierait certainement le problème, mais ne permettrait peut-être pas pour autant de rendre compte des observations empiriques.

### **III.3.3. Schémas d'activation des unités phonologiques et orthographiques**

Nous avons jusqu'à présent abordé les substitutions observées, sans poser, de manière explicite, de modèle d'organisation sous-jacente des modes d'activation et de représentation des unités supposées, alors que nos propositions semblent précisément se référer implicitement à un modèle présupposé de cette organisation. En effet, nous avons mentionné les termes « activation », « unités », « voies », « phonologiques/phonémiques », « orthographiques/graphémiques » et « réalisations phonétiques/allophoniques » sans cadrage précis, alors qu'un tel cadrage semble nécessaire à une approche plus fine des phénomènes observés. La base de ce cadrage repose notamment sur la conception du lexique mental (phonologique *vs.* phonétique ; symbolique *vs.* subsymbolique) et de son organisation. Si l'on maintient la distinction entre niveaux phonétique et phonologique (sans développer davantage ici la problématique), et si l'on accorde aux « représentations/unités orthographiques » un statut cognitif relativement autonome, les relations entre ces trois domaines doivent être précisées, si l'on souhaite résoudre les difficultés rencontrées jusqu'ici. Plusieurs possibilités sont alors envisageables, articulées sur l'axe linéaire/ parallèle.

Cas 1 : Activation linéaire des unités phonémiques et graphémiques.

1.a. En unimodalité



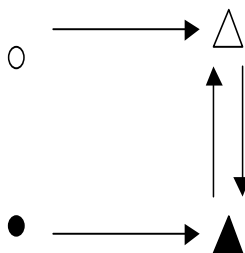
Activation  
Phonème-Phonème-Graphème

et/ou

Activation  
Graphème-Graphème-Phonème

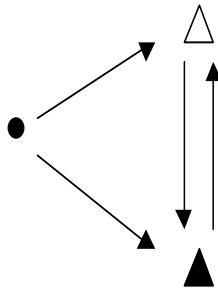
Les stimuli sont représentés par les ellipses, les unités internes par les triangles, les activations par les flèches. La couleur de remplissage noire est attribuée au domaine orthographique, la couleur blanche au domaine phonologique.

1.b. En bimodalité

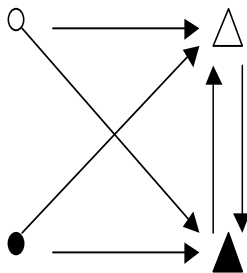


## Cas 2 : Activation parallèle des unités phonémiques et graphémiques.

### 2.a. En unimodalité :



### 2.b. En bimodalité

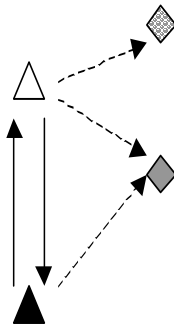


Comme nous l'avons dit, dans les tâches qui nous occupent ici, il faut également situer le plan phonétique (répétition et lecture à voix haute), lequel correspond à la deuxième étape des tâches mentionnées :

- 1) Reconnaissance (auditive et/ou visuelle) → Perception
- 2) Vocalisation (phonique) → Production

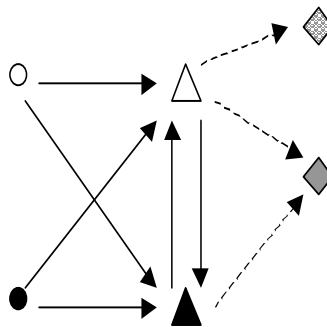
Pour cela, il nous faut prendre en compte l'hypothèse forte que nous avons posée, selon laquelle les représentations phonologiques peuvent être réalisées différemment selon le contexte phonético-phonologique (allophones), tandis que les représentations

graphémiques qui nous occupent ici (« b » et « r »)<sup>158</sup> se réalisent systématiquement de manière unique.



Les flèches en pointillés correspondent aux réalisations phonétiques, la couleur grise est attribuée au plan phonétique, les losanges correspondent aux productions des apprenants et le motif intérieur correspond aux allophones phonétiques.

Si l'on combine à présent les versants perceptifs et productifs de notre approche, voici ce que l'on obtient pour le modèle d'activation parallèle :



---

<sup>158</sup> Il en serait tout autrement au niveau lexical, avec la prise en compte des caractéristiques accentuelles par exemple, ou bien encore vis-à-vis des caractères kanjis, typiquement caractérisés par leurs séries homophoniques.



Etant donné un tel réseau d'activations, que le modèle soit linéaire ou parallèle, et étant donné que, d'après Dijkstra, Frauenfelder et Schreuder (1993) il serait difficile à ce jour d'affirmer que, parallèlement aux activations bidirectionnelles graphèmes-phonèmes, il y aurait un réseau d'inhibitions identiques, la résolution de nos interrogations sembleraient devoir passer par deux termes clés : celui de « poids de connexion » (activation plus ou moins dominante selon plusieurs facteurs) et celui de « temporalité » (activations synchrones ou asynchrones, en compétition ou en complémentarité). Ainsi pour un stimulus R (« R » ou /R/) plusieurs résultats sont envisageables :

- R → /R/
- R → « R »
- R → (« R », /R/)

### **III.3.4. Protocole quasi-expérimental**

Les questions soulevées jusqu'à présent nous ont ainsi permis de préciser notre problématique générale. C'est sur cette base que nous avons élaboré nos protocoles quasi-expérimentaux, destinés à nous fournir des informations sur certains aspects du traitement des relations phonographémiques françaises par les apprenants japonais.

Deux tâches ont été sélectionnées : l'une, simple, d'appariement phonographémique ; l'autre, complexe, de discrimination phonologique et d'appariement phonographémique. Tandis que la première vise à nous renseigner sur le traitement des relations phonographémiques par les sujets, la deuxième vise à nous renseigner sur l'influence hypothétique que peut avoir le support orthographique sur une tâche de discrimination phonologique, considérée comme tâche d'apprentissage, en terme de complexité : en effet, il ne s'agissait pas tant ici de comparaison inter-tâches *avec* et *sans* support orthographique, mais plutôt de comparer les taux de confusion phonologique et de confusion phonographémique. Cette question comporte des implications tant didactiques (apprentissage de la phonologie d'une part, des relations

phonographémiques d'autre part) que méthodologiques (constitution de corpora phonologiques avec ou sans support orthographique). Une fois les taux de confusion phonographémique et phonologique établis pour chaque sujet, cela nous conduit aux questions fines suivantes :

### **1) Position**

La position du segment consonantique cible dans le mot phonologique influence-t-elle le taux de confusion phonographémique ?

Tâche simple : appariement

Tâche complexe : appariement et discrimination

### **2) Nature**

La nature du segment consonantique cible dans le mot phonologique influence-t-elle le taux de confusion phonographémique ?

Tâche simple : appariement

Tâche complexe : appariement et discrimination

### **3) Environnement**

La nature du segment vocalique adjacent au segment cible dans le mot phonologique influence-t-elle le taux de confusion phonographémique ?

Tâche simple : appariement

Tâche complexe : appariement et discrimination

### **4) Influence L1**

Les confusions phonographémiques correspondent-elles à une préférence pour les graphes disponibles en japonais ?

Tâche simple : appariement

Tâche complexe : appariement et discrimination

## **5) Effets supplémentaires (questions liées à la tâche complexe)**

### **5.1. Effet de différenciation :**

La nature des couples de mots phonologiques (identiques / différents) influence-t-elle le taux de confusion phonographémique ?

### **5.2. Effet de complexité :**

Le taux de confusion phonographémique est-il supérieur en tâche simple ou complexe ?

Ce sont précisément ces questions auxquelles notre étude devrait nous permettre de répondre.

## **III.3.5. Substitutions inattendues et stratégies d'apprentissage**

Ayant présenté les substitutions segmentales liées à un traitement psycholinguistique « erroné » des unités phonético-phonologiques et orthographiques françaises, *a priori* en raison des divergences entre les deux systèmes L1 et L2, il nous faut à présent, rapidement au moins, évoquer les substitutions d'origine *a priori* « strictement psycholinguistique », dans le sens où elles seraient liées aux stratégies d'apprentissage et à la constitution des systèmes interlangagiers des apprenants.

- /b/ ou « b » réalisé [v]
- /l/ ou « l » réalisé [R]
- /OU/ ou « ou » réalisé [∅]

Ce type de substitution peut être fréquemment observé chez des apprenants de niveau faux - débutant ou intermédiaire, et il serait facile de les assigner à ce que, dans le domaine didactique, on a coutume d'appeler la tendance à l'hypercorrection ou à la surgénéralisation. Deux cas peuvent se présenter : A) substitutions occasionnelles, *a priori* aléatoires ; B) substitutions systématiques, pendant une courte période de temps au moins. Les cas n°1 semblent bien évidemment être les plus fréquents. A première

vue, ces substitutions sont tout à fait déconcertantes, et il serait assez délicat de les traiter de manière strictement linguistique. D'un point de vue psycholinguistique en revanche, on peut les replacer dans le cadre de l'apprentissage des langues étrangères, en considérant plus particulièrement le système linguistique instable qu'est l'interlangue provisoire des apprenants, ainsi que leurs stratégies d'apprentissage. Dans les deux cas, on peut considérer que les apprenants ont 1) pris conscience de l'existence des unités (/v/ ; « v »), (/R/ ; « R ») et (/EU/ ; « EU ») et 2) qu'ils sont parvenus à acquérir les gestes articulatoires leur permettant de donner une implémentation phonétique adéquate à ces unités. Ce qui semble encore leur faire défaut cependant, c'est la stabilité du système d'opposition phonémique : après s'être efforcés d'établir une frontière intra-catégorielle dans ce qui pour eux n'était auparavant qu'une unique catégorie (assimilation catégorielle de /b/ et /v/ par exemple), les sujets ne parviennent pas encore à stabiliser et à fixer ladite frontière. On pourrait penser que celle-ci ne s'est donc pas encore établie en tant que frontière inter-catégorielle. Nous pouvons avancer deux types d'explications générales pour les deux cas évoqués, que ce soit en reconnaissance auditive ou visuelle :

Cas A) : Lorsque les substitutions sont occasionnelles, on pourrait penser que, à ce niveau, les productions des apprenants ne sont pas encore « procéduralisées », mais bien sous la dépendance de savoirs déclaratifs qui s'expriment à travers une unité de « contrôle » de leur production (ce qu'on pourrait identifier au « monitor » de Krashen pour un modèle d'apprentissage, ou au « monitoring » de Levelt pour un modèle de production). Ces substitutions en sont précisément les indicateurs, puisqu'elles pourraient indiquer un manque de contrôle (en particulier lorsque la tâche se montre assez complexe), ou bien au contraire un contrôle exagéré et erroné (en particulier lorsque la tâche est assez simple). Le terme adéquat serait alors celui d'*hypercorrection*.

Cas B) : Lorsque les substitutions sont quasi-systématiques (cas plus rares), on peut penser que, au lieu d'établir une frontière intra-catégorielle, les apprenants ont substitué une catégorie (nouvelle) à une autre (ancienne) sans établir d'opposition.

Cette fixation erronée des unités pourrait être caractérisée par le terme de *surgénéralisation*.

Ces types de substitutions, relativement marginaux au niveau débutant, ne constituent cependant pas le sujet principal de notre étude, et nous nous contenterons donc de ces courtes ébauches d'explication pour l'instant.

Jusqu'à présent, nous nous sommes principalement concentrés sur la dimension *phonémique* des *unités* phonographémiques sublexicales. L'influence des représentations orthographiques sur l'apprentissage de l'inventaire segmental de la phonologie d'une langue étrangère a par ailleurs déjà été commenté dans la littérature, tant psycholinguistique que didactique (Steele, 2005). En revanche, peu d'études, à notre connaissance, ont été conduites, dans cette perspective, sur la dimension *syllabique* desdites unités et sur les *processus de segmentation* des mots phonologiques. C'est ce que nous abordons à présent.

#### **III.4. Influence de l'orthographe : niveau syllabique et segmentation**

Traiter de substitutions phonémiques ou graphémiques ne doit pas faire oublier que la parole est essentiellement continue, et que le mode de segmentation de la parole n'est pas « universel » : il se pourrait alors que, outre les caractéristiques phonologiques des systèmes langagiers en présence, leurs caractéristiques orthographiques exercent également une influence sur ces opérations de segmentation.

##### **III.4.1. L'orthographe comme outil de segmentation de la parole**

S'il est parfois nécessaire, utile ou préférable d'aborder la parole en la décomposant, de manière à pouvoir mieux en examiner l'architecture et le fonctionnement, cette discrétisation, linéaire ou non, totale ou partielle, n'enlève rien au caractère intrinsèquement continu de la parole. Des modèles tels que celui de Massaro (2001) remettent d'ailleurs au goût du jour cette problématique, succédant à une conception peut-être exagérément atomiste des unités et des processus perceptifs. Ce qui sous-tend, néanmoins, l'articulation discret/continu peut facilement renvoyer à une autre, tout

autant fondamentale, celle entre les dimensions spatiale et temporelle : on a alors vite fait de dire que les écrits (discrets dans l'espace mais continus dans le temps) « restent », tandis que les paroles ((relativement) continues dans l'espace mais discrètes dans le temps) « s'envolent ». La double spécialisation hémisphérique toute relative du cerveau humain vise notamment à gérer cette articulation entre le continu et le discret, une articulation que l'on retrouve dans l'axe langagier perception/production : tandis que les modèles de perception/compréhension peuvent s'accommoder relativement bien de stimuli continus, il n'en va pas de même pour les modèles de production qui présentent une hiérarchie temporelle passablement linéaire, malgré les phénomènes de chevauchement, de rétroaction, de parallélisme ou de distribution.

Gérer le continu constitue donc évidemment une source de difficulté pour les apprenants de langue étrangère, car le continu implique une disparition immédiate des stimuli que l'on tente d'appréhender. Pour pallier cette disparition, l'être humain est notamment doté d'un système mnésique, dont l'un des composants, la mémoire de travail, permettrait précisément de gérer cette fulgurance. Mais là encore, il y a des limites : 7, plus ou moins 2, c'est le chiffre « magique » mis en lumière par Miller et ses collaborateurs pour le nombre maximal d'items pouvant être stockés en mémoire de travail. Depuis, bien sûr, les études se sont affinées. Ces limites ont cependant certaines implications pour nos sujets : la discrétisation du continu semble indispensable à son traitement et plus encore à son apprentissage. Cette discrétisation peut être envisagée à travers deux autres termes : celui de *segmentation*, et à sa suite celui de comptage, et celui de *fixation*. La segmentation vise en effet, généralement, la fixation des unités résultantes : cela s'illustre immédiatement à travers l'existence des divers codes orthographiques de transcription de la parole orale. Une fois cette segmentation effectuée, il est alors possible de se pencher sur les unités ainsi isolées, ce qui nous renvoie à la partie précédente pour les unités segmentales phonographémiques sublexicales. Nous pourrions faire de même pour les unités syllabiques. Cependant, si c'est bien au domaine syllabique que cette sous-partie est consacrée, ce n'est pas seulement de manière statique (unités syllabiques), mais aussi dans une perspective dynamique (processus de syllabification).

Si la plupart des études dans le domaine de la reconnaissance visuelle linguistique se concentrent en effet sur des unités, la plupart segmentales, peu en revanche, à notre connaissance, se sont intéressées, *dans la perspective qui est la nôtre*, aux interactions entre représentations phonologiques et orthographiques dans le cadre des processus de segmentation/syllabification de mots. Ainsi, s'il apparaît que les représentations orthographiques en langue première peuvent influencer sur la perception des segments (phonologiques et orthographiques) en langue étrangère, nous nous demandons s'il n'existe pas de phénomènes similaires au niveau de la segmentation/syllabification : il semble en effet de plus en plus établi que, en ce qui concerne la perception du langage parlé, « les procédures de segmentation peuvent varier selon les langues, et, d'autre part, quand le sujet est confronté à des séquences n'appartenant pas à sa langue il applique à celles-ci les procédures d'analyse adaptées aux propriétés de sa propre langue » (Ségui, 1997, p. 16). Plus précisément, « les travaux actuels suggèrent que les procédures de segmentation du langage parlé sont fortement contraintes par l'organisation rythmique des langues » (Ségui, 1997, p. 17). Suite à cette affirmation de Ségui, il faut bien se garder de confondre segmentation rythmique et segmentation syllabique/syllabification. Mais on peut néanmoins se demander si les représentations orthographiques (en langue première, en langue étrangère, ou communes, en partie, aux deux) n'influencent pas la segmentation (découpage en unités discrètes), la syllabification (voir Laks, 1995), et ainsi en partie le « comptage » d'unités à l'intérieur de mots phonologiques. Cela renvoie donc à la distinction entre lettres, graphèmes et kana dans le cas du japonais.

#### **III.4.2. Le cas du japonais : contraintes phonologiques et/ou orthographiques ?**

Nous avons vu l'importance, dans le cadre strictement linguistique de OT, des contraintes phonologiques dites de « fidélité », ou bien encore dans le cadre TCRS du *Principe de Préservation*. Dans le cas qui nous intéresse, à savoir celui du traitement des syllabes complexes biconsonantiques françaises par des sujets japonais, nous pouvons nous demander si la structure orthographique de mots présentés visuellement peut influencer sur le traitement de la structure phonologique syllabique, en particulier

lorsque l'une ou l'autre forme semble ne pas respecter les contraintes linguistiques de la langue première des sujets. Il serait en particulier intéressant de savoir comment sont gérés les cas où les deux formes semblent ne pas converger, conduisant ainsi à d'hypothétiques conflits entre contraintes phonologiques et contraintes orthographiques.

Le tableau suivant permet d'illustrer, de manière très simplifiée, cette problématique :

Structure syllabique L2 (français)		Système phonologique L1 (japonais)	Système graphémique L1 (japonais)
<i>Phono</i>	<i>Graph</i>		
/CCV/	« CCV »	Illicite → /CVCV/	Illicite → « CVCV »
/CVCV/	« CCV » <sup>159</sup>	Licite	Illicite → « CVCV »
/CCV/	« CVCV » <sup>160</sup>	Illicite → /CVCV/	Licite

Ce premier tableau appelle cependant une remarque importante, relative, d'une part aux strates lexicales et aux codes scripturaux, et d'autre part aux modes de conversion/ d'adaptation interscripturale. Il faut en effet se rappeler que l'une des caractéristiques du système orthographique japonais réside dans sa pluralité, et que cette pluralité se retrouve au niveau de l'organisation des contraintes phono-lexicales. On en trouve de nombreux exemples dans les « marques » de produits ou d'établissements : par exemple le grand magasin « PARCO » (« CVCCV ») dont la transcription alphabétique (et non en romaji) coexiste sur les affiches avec la transcription en katakana « PA-RU-KO » (« CVCVCV »). On voit donc ici qu'une structure orthographique « CCV » semble licite dans le code alphabétique étranger, mais pas dans le code katakana japonais. Un exemple similaire est celui du grand magasin « Printemps » à Ginza (Tokyo), transcrit « PU-RA-N-TA-N ». On notera en passant le croisement suivant, même si la variation est aujourd'hui de mise dans l'usage qui est fait de ces deux codes :

<sup>159</sup> Cas plutôt rare en français, à aborder dans une perspective sociolinguistique.

<sup>160</sup> Ce cas de figure évoque, en partie seulement, le cas du « schwa/ e muet » en français : « petit » prononcé [pti].



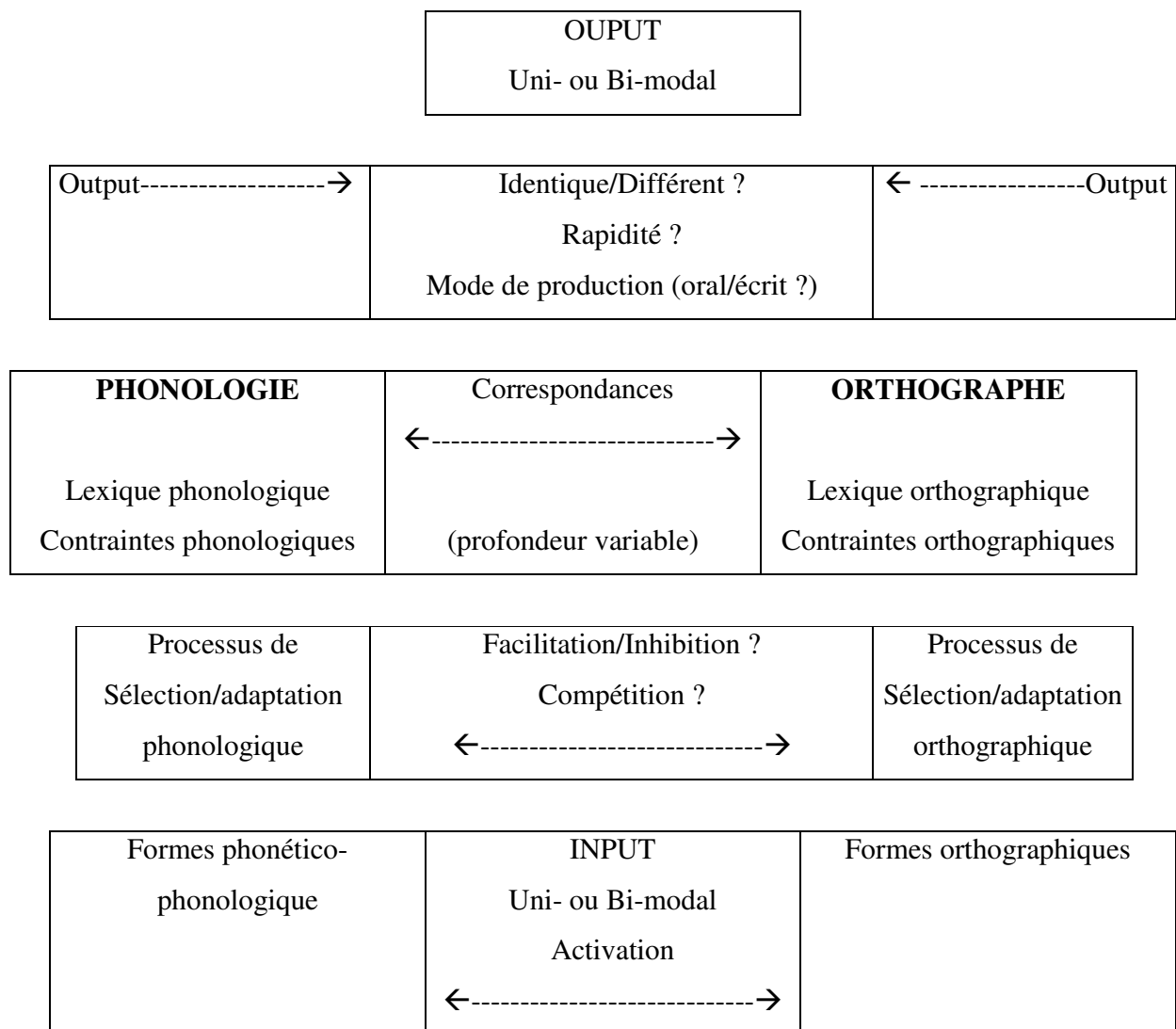
<b>Script</b>	<b>Katakana</b>	<b>Romaji</b>
<b>Origine graphique</b>	japonaise	Etrangère
<b>Origine lexicale</b>	étrangère	Japonaise

En revanche, cette situation confirme un point important de notre réflexion : il existe des codes de conversion/ d'adaptation relativement standards et établis dans la communauté japonaise des formes orthographiques étrangères en formes scripturales katakana.

Cependant, il reste à savoir comment établir le lien entre procédures d'assemblage et d'adressage d'une part, et strates lexicales d'autre part, en particulier lorsque l'on considère le cas des non-mots. Car si l'on considère que tel script est généralement associé à telle strate, on pourrait alors penser que le système romaji (qui est celui en partie partagé avec le français) peut être associé à la strate des mots étrangers et permettre ainsi des combinaisons illicites dans les strates « natives ». Néanmoins, on peut également s'interroger sur la manière dont le gradient d'« étrangeté » ("foreignness") au sein de la strate lexicale peut se traduire au niveau scriptural. C'est peut-être bien à partir de cette gradience que l'on pourrait rendre compte de l'influence du système scriptural commun au français et au japonais. Dans le cadre de OT par exemple, il conviendrait donc d'examiner si, dans le processus d'adaptation, la perception de l'input n'entraînerait pas, dans la perspective de modèles de type Cohorte, une compétition de contraintes, entre contraintes phonologiques et contraintes orthographiques, selon la modalité de présentation (uni- ou bi-modale), étant donné qu'il y a activation de deux types de représentation (orthographique et phonologique).

On pourrait ainsi se demander si, à l'instar du système de contraintes phonologiques, il ne serait pas possible de formaliser un système de contraintes orthographiques, qui pourrait éventuellement permettre de rendre compte de certaines interactions entre les

deux systèmes, et que l'on pourrait esquisser dans un premier temps de la manière suivante (à lire de bas en haut) :



Deux remarques importantes sont à noter à propos de cette esquisse :

- D'abord, la symétrie apparente entre phonologie et orthographe dans ce schéma est bien évidemment trompeuse : le système scriptural est un artéfact social, second par rapport au système phonologique qu'il est censé représenter. Néanmoins, l'influence du script sur la parole, des graphèmes sur les phonèmes, est attestée et on ne peut donc nier, ni cette influence, ni l'existence de représentations mentales de type visuo-graphique (Denis, 1994), parallèlement à celles de type phonologique.

- Ensuite trois facteurs vont influencer considérablement sur ce schéma :
  - La différence de scripts entre les deux systèmes linguistiques en présence, L1 et L2.
  - La profondeur du système scriptural dans le système L1, vis-à-vis de la phonologie du même système.
  - Le format de présentation (input) (uni- ou bi-modal) et de restitution (output) (uni- ou bi-modal) du matériel linguistique.

Même si le développement d'une telle piste de recherche demanderait plus ample réflexion, celle-ci permet néanmoins de souligner que, si l'adaptation phonologique a reçu beaucoup d'attention dans le passé (Shinohara, 1997), l'adaptation orthographique n'est cependant pas une question triviale, bien que l'on puisse sans doute la considérer comme un produit dérivé de la première, ne nécessitant que peu d'attention de la part des chercheurs. Si la question peut effectivement sembler limitée d'un point de vue strictement linguistique, elle l'est cependant beaucoup moins d'un point de vue psycholinguistique, puisque, d'une part cela concerne les codes de conversion graphème - phonème (reconnaissance visuelle, activation unidirectionnelle) et d'autre part cela peut conduire à l'existence hypothétique d'une influence graphème-phonème dans le processus d'adaptation phonologique (activation bidirectionnelle). Si l'on considère que ces modes d'adaptation sont largement automatisés chez les sujets japonais, on peut supposer que ces automatismes vont constituer des obstacles à leur apprentissage de la phonologie du français, en particulier si celui-ci se fait à l'aide d'un support orthographique.

Nous avons donc élaboré un protocole nous permettant de répondre aux deux questions principales évoquées jusqu'ici :

- La segmentation phonologique syllabique varie-t-elle en fonction du mode de présentation (auditif vs orthographique vs audio-orthographique) ?
- Comment cette segmentation est-elle effectuée lorsque, en présentation bimodale, il y a divergence entre les formes auditives et les formes visuelles ?

Pour ce faire, nous nous sommes basés sur une tâche de comptage syllabique de non-mots, tâche qu'il convient d'examiner à présent. Que doit accomplir le sujet en tâche de comptage syllabique sur des non-mots ? « Compter » implique deux processus : la segmentation et le comptage. La question se pose de savoir si la segmentation et le comptage s'effectuent de manière simultanée (parallèle) ou consécutive (linéaire), et si le comptage nécessite une (sub-) vocalisation numérique cardinale (qui pourrait interférer avec la (sub-)vocalisation du non-mot nécessaire à son maintien en mémoire de travail) ou pas. Les deux questions sont en fait sans doute liées : le protocole que nous avons élaboré, en proposant trois groupes d'ellipses adjacentes (3, 4 et 5 ellipses) comme format de rendu, visait à proposer un support de comptage visuel non-linguistique permettant de se passer de numérisation cardinale : en tapant du doigt par exemple, les sujets devaient pouvoir directement associer chaque item résultant de la segmentation à une ellipse, de manière à effectuer une tâche de segmentation pure, sans interférence numérique. Il apparaît ici qu'il nous est indispensable de nous référer à un modèle de mémoire de travail, modèle que nous empruntons à Baddeley (1986).

La question qui se pose alors est la suivante: en unimodalité auditive, en supposant qu'il y ait activation du niveau phonologique, la segmentation est-elle effectuée de manière purement phonologique, ou bien certaines représentations d'ordre graphémique sont-elles activées qui pourraient influencer la tâche de segmentation ? Etant donné ce que nous avons dit plus haut, il apparaît maintenant évident qu'il y a de fortes chances pour que les représentations phonologiques ne soient pas les seules activées.

A cela, il faut ajouter deux remarques: d'abord, d'un point de vue général, et malgré le support orthographique de notre protocole, nous savons bien que « l'écrit » constitue l'outil de segmentation par excellence de l'oral (continu) dont nous disposons, outil de segmentation, mais aussi de fixation sur l'espace des unités segmentées dans le temps. La question est alors : quel « écrit », c'est-à-dire quels outils graphiques/graphémiques vont être sollicités ? On peut supposer, pour nos sujets débutants en français, prompts à

utiliser les caractères kana en classe de français pour transcrire les mots entendus inconnus, que les unités activées vont être celles du « syllabaire » japonais, katakana ou romaji (en particulier au vu de l' « opacité », relative, du système orthographique français). En retour, ces outils de segmentation risquent alors fortement d'imposer des contraintes sur la segmentation phonologique des non-mots entendus, étant donné la structure des syllabes japonaises telles que transcrites par le « syllabaire ».

On pourrait en outre considérer que l'influence de la graphie sur la perception de l'organisation phonologique des énoncés (syllabes *vs.* mores par exemple) provient du fait que les unités graphiques sont des unités discontinues (lettres/kana), contrairement, en partie, à la parole, qui, elle, est continue. Même si, en japonais, les blancs graphiques sont bien moins nombreux qu'en français, les unités graphiques sont tout de même sous-tendues par des opérations de segmentation. L'apprentissage de la lecture/écriture conditionnerait donc, dans une certaine mesure au moins l'apprentissage de la segmentation de la parole, et, à sa suite, l'organisation prosodique, syllabique/moraïque et ainsi phonotactique des énoncés. Bien sûr, la réalité est sans doute plus complexe, le système orthographique étant lui-même, en théorie, censé n'être qu'une transcription du système parolier. Cependant, comme cela est parfaitement illustré dans la langue française via les réformes de l'orthographe, mais aussi dans la langue japonaise via l'apparition de « nouvelles consonnes » (par exemple la transcription relativement récente en katakana « de la lettre "v" »), le système orthographique peut à son tour exercer une influence sur le système parolier d'une communauté linguistique donnée. Si cela est bien connu sous le terme d' « effet Buben » pour le développement de la langue première des sujets, les travaux relatifs à l'apprentissage des langues étrangères dans cette perspective sont en revanche bien moins nombreux.

Dans le cas qui nous occupe, ce point est renforcé par la « culture de l'écrit » propre aux Japonais : les enseignants de français au Japon sont habitués à la question « Comment ça s'écrit ? » posée par leurs étudiants dès qu'une forme orale leur échappe. Cette question est la même lorsqu'il s'agit d'identifier un morphème japonais, dont

l'indétermination peut parfois être résolue par la visualisation du kanji qui lui correspond. Les correspondances étroites entre « syllabaire » japonais phonologique et système scriptural kana y participent également fortement.

### **III.4.3. Protocole quasi-expérimental**

Les questions fines auxquelles nos deux interrogations générales nous conduisent, et auxquelles notre protocole est censé nous aider à répondre, une fois les taux de confusion syllabiques établis, sont alors les suivantes :

#### **I) En présentation bimodale audio-orthographique :**

(Tâche complexe de comptage syllabique phonologique et discrimination syllabique phonographémique).

##### **I).1. Lorsque les stimuli auditifs et orthographiques sont identiques**

###### **(/CCV/)**

I).1.1. Position : La position du groupe cible dans le mot bimodal influence-t-elle le taux de confusion syllabique ?

I).1.2. Nature : La nature du groupe cible dans le mot bimodal influence-t-elle le taux de confusion syllabique ?

##### **I).2. Lorsque les stimuli auditifs et orthographiques sont différents (/CCV/-/CVCV/ et vice-versa)**

I).2.1. Position : La position du groupe cible dans le mot phonologique influence-t-elle le taux de discrimination entre les deux stimuli ?

I).2.2. Nature : La nature du groupe cible dans le mot phonologique influence-t-elle le taux de discrimination entre les deux stimuli ?

I).2.3. Longueur : Le rapport entre modalité et structure syllabique (légale ou illégale) des stimuli influence-t-il le taux de discrimination entre les deux stimuli ?

I).2.4. Lorsqu'il y a discrimination entre les deux stimuli : Le taux de discrimination correct est-il supérieur au taux de discrimination erronée ?

I).2.5. Lorsqu'il n'y a pas discrimination entre les deux stimuli : Quel mot (phonologique ou graphémique), et donc quelle modalité, semble privilégié pour le comptage syllabique ?

**I).3. Effet de complexité :** Trouve-t-on un écart significatif entre les taux de confusion syllabique (comptage) et phonographémique (discrimination) pouvant indiquer un conflit procédural d'ordre attentionnel ?

## **II) En présentation unimodale visuelle**

II).1. Position : La position du groupe cible dans le mot graphique influence-t-elle le taux de confusion syllabique ?

II).2. Nature : La nature du groupe cible dans le mot graphique influence-t-elle le taux de confusion syllabique ?

## **III) Comparaison des modalités de présentation**

III).1. Modalité : Le mode de présentation des mots (auditif / orthographique / audio-orthographique) influence-t-il le taux de confusion syllabique ?

III).2. Position : Cette influence varie-t-elle en fonction de la position des groupes cibles ?

III).3. Nature : Cette influence varie-t-elle en fonction de la nature des groupes cibles ?

Les questions relatives au traitement en présentation unimodale auditive ont déjà été formulées dans la partie phonologique de cette étude.



### **III.5. Résumé**

Nous pouvons donc à présent synthétiser l'essentiel de nos problématiques psycholinguistiques<sup>161</sup> :

<b>Domaine</b>	<b>Segmental</b>	<b>Syllabique</b>
<b>Expériences</b>	<b>A et F</b>	<b>C2, D et E</b>
<b>A et F</b>	<p>La position du segment consonantique cible dans le mot phonologique influence-t-elle le taux de confusion phonographémique ?</p> <p><u>Tâche simple</u> : appariement</p> <p><u>Tâche complexe</u> : appariement et discrimination</p>	
<b>A et F</b>	<p>La nature du segment consonantique cible dans le mot phonologique influence-t-elle le taux de confusion phonographémique ?</p> <p><u>Tâche simple</u> : appariement</p> <p><u>Tâche complexe</u> : appariement et discrimination</p>	
<b>A et F</b>	<p>La nature du segment vocalique adjacent au segment cible dans le mot phonologique influence-t-elle le taux de confusion phonographémique ?</p> <p><u>Tâche simple</u> : appariement</p> <p><u>Tâche complexe</u> : appariement et discrimination</p>	
<b>F (et A)</b>	<p><u>Questions liées à la tâche complexe</u> :</p> <p><u>Effet de différenciation</u> :</p> <p>La nature des couples de mots phonologiques (identiques / différents) influence-t-elle le taux de confusion phonographémique ?</p> <p><u>Effet de complexité</u> :</p> <p>Le taux de confusion phonographémique est-il supérieur en tâche simple ou complexe ?</p>	

---

<sup>161</sup> La nomenclature des « expériences » est tout à fait arbitraire.

<p><b>E</b></p>	<p><u>Tâche complexe</u> : comptage syllabique phonologique et discrimination syllabique phonographémique</p> <p><u>Cas 1</u> : longueur syllabique identique</p> <p>*La position du groupe cible dans le mot bimodal influence-t-elle le taux de confusion syllabique ?</p> <p>*La nature du groupe cible dans le mot bimodal influence-t-elle le taux de confusion syllabique ?</p> <p><u>Cas 2</u> : longueur syllabique différente</p> <p>*La position du groupe cible dans le mot phonologique influence-t-elle le taux de discrimination phonographémique ?</p> <p>*La nature du groupe cible dans le mot phonologique influence-t-elle le taux de discrimination phonographémique?</p> <p><u>2-A</u> : Lorsqu'il y a discrimination : le taux de discrimination phonographémique correct est-il supérieur au taux de discrimination phonographémique erronée ?</p> <p><u>2-B</u> : Lorsqu'il n'y a pas discrimination : quel mot (phonologique ou graphique) et donc quelle modalité est privilégiée pour le comptage syllabique ?</p> <p><u>Effet de complexité</u> :</p> <p>Trouve-t-on un écart entre les taux de confusion syllabique (comptage) et phonographémique (discrimination) pouvant indiquer un conflit procédural d'ordre attentionnel ?</p>
<p><b>D</b></p>	<p>La position du groupe cible dans le mot graphique influence-t-elle le taux de confusion syllabique ?</p>
<p><b>D</b></p>	<p>La nature du groupe cible dans le mot graphique influence-t-elle le taux de confusion syllabique ?</p>

<b>C2-D-E</b>	Le mode de présentation des mots (auditif / orthographique / audio-orthographique) influence-t-il le taux de confusion syllabique ? L'influence (potentielle) de la position et de la nature des groupes cibles varie-t-elle en fonction de la modalité ?
---------------	--

Nous pouvons donc maintenant présenter la partie quasi-expérimentale de notre étude.

## **Partie Expérimentale**



## **Table des Matières Partielle**

<b>PARTIE EXPERIMENTALE I : PRE-TESTS.....</b>	<b>- 645 -</b>
I. INTRODUCTION.....	- 645 -
II. METHODOLOGIE.....	- 651 -
III. RESULTATS ET INTERPRETATION .....	- 666 -
IV. RESUME GENERAL ET DISCUSSION .....	- 679 -
V. ASPECTS METHODOLOGIQUES ET PERSPECTIVES .....	- 685 -
<b>PARTIE EXPERIMENTALE II : TESTS - PROTOCOLE.....</b>	<b>- 693 -</b>
TABLE DES MATIERES PARTIELLE .....	- 694 -
I. METHODOLOGIE .....	- 695 -
II. DESCRIPTIF GENERAL DE LA POPULATION .....	- 696 -
III. DESCRIPTIF GENERAL DES PROTOCOLES .....	- 704 -
IV. PREMIERE PARTIE : AXE SEGMENTAL (GROUPE 1) .....	- 710 -
IV.1. <i>Test A (Identification)</i> .....	- 710 -
IV.2. <i>Test F (Identification et Discrimination)</i> .....	- 716 -
V. DEUXIEME PARTIE : AXE SYLLABIQUE (GROUPE 1 ET 2).....	- 721 -
V.1. <i>Test C1 (Groupe 1 – Unimodalité Auditive)</i> .....	- 721 -
V.2. <i>Test C2 (Groupe 2 – Unimodalité Auditive)</i> .....	- 728 -
V.3. <i>Test D (Groupe 2 – Unimodalité Visuelle)</i> .....	- 732 -
V.4. <i>Test E (Groupe 2 – Bimodalité Audio-Visuelle)</i> .....	- 735 -
V.5. <i>Méta-test C2-D-E (Groupe 2 – Effet de la Modalité)</i> .....	- 743 -
VI. VARIABLES ET CONSIDERATIONS SUPPLEMENTAIRES .....	- 744 -
<b>PARTIE EXPERIMENTALE III : TESTS - RESULTATS.....</b>	<b>- 749 -</b>
TABLE DES MATIERES PARTIELLE .....	- 750 -
I. RESULTATS DES TESTS .....	- 753 -
I.1. <i>Considérations statistiques</i> .....	- 753 -
I.2. <i>Axe segmental</i> .....	- 758 -

TACHE A .....	- 758 -
A) Résultats descriptifs .....	- 758 -
A) 1) Axe syntagmatique .....	- 759 -
A) 2) Axe paradigmatique .....	- 762 -
B) Résultats inférentiels : analyses de variance (Test F) .....	- 767 -
C) Résultats complémentaires : choix de graphème existant/absent en L1 (41 sujets) .....	- 768 -
D) Résumé .....	- 771 -
 TACHE F .....	 - 773 -
A) Résultats descriptifs .....	- 773 -
A) 1) Confusion phonographémique (premier stimulus) .....	- 773 -
A) 1) 1. Axe syntagmatique .....	- 774 -
A) 1) 2. Axe paradigmatique .....	- 776 -
A) 1) 3. Variable liée à la tâche : nature du couple (identique/différent) .....	- 780 -
A) 2) Confusion phonologique (entre les deux stimuli auditifs) .....	- 781 -
A) 2) 1. Axe syntagmatique .....	- 782 -
A) 2) 2. Axe paradigmatique .....	- 784 -
A) 2) 3. Variables liées à la tâche : nature du couple (identique/différent) .....	- 788 -
B) Résultats inférentiels : analyses de variance (Test F) .....	- 789 -
D) Résumé .....	- 790 -
 COMPARAISON TACHES A ET F .....	 - 792 -
A) Résultats globaux .....	- 792 -
B) Position .....	- 792 -
C) Nature .....	- 794 -
D) Résultats inférentiels .....	- 797 -
E) Résumé .....	- 798 -
 <i>1.3. Axe syllabique</i> .....	 - 799 -
 TACHE C1 .....	 - 799 -
A) Résultats descriptifs .....	- 799 -
A) 1) Axe syntagmatique .....	- 800 -
A) 2) Axe paradigmatique .....	- 805 -
B) Résultats inférentiels : analyse de variance (Test F) : .....	- 816 -
C) Résumé .....	- 819 -
C) 1) Résumé détaillé .....	- 819 -
C) 2) Résumé essentiel .....	- 826 -
 TACHE C2 .....	 - 827 -
A) Résultats descriptifs .....	- 827 -
A) 1) Axe syntagmatique .....	- 828 -
A) 2) Axe paradigmatique .....	- 831 -
B) Résultats inférentiels : analyse de variance (Test F) .....	- 851 -
C) Résumé .....	- 854 -
C) 1) Résumé Détaillé .....	- 854 -
C) 2) Résumé Essentiel .....	- 862 -
 COMPARAISON TACHES C1 ET C2 .....	 - 863 -
 TACHE D .....	 - 871 -
A) Résultats descriptifs .....	- 871 -
A) 1) Axe syntagmatique .....	- 871 -
A) 2) Axe paradigmatique .....	- 873 -
A) 3) Variation entre sujets .....	- 879 -
B) Résumé Essentiel .....	- 880 -
 COMPARAISON TACHES C2 ET D .....	 - 881 -

TACHE E.....	- 890 -
CAS 1 - STIMULI AUDITIFS ET VISUELS CONVERGENTS.....	- 890 -
A) Résultats descriptifs.....	- 890 -
A) 1) Axe syntagmatique.....	- 891 -
A) 2) Axe paradigmatique.....	- 892 -
B) Résultats inférentiels.....	- 903 -
C) Résumé Essentiel.....	- 903 -
CAS 2 - STIMULI AUDITIFS ET VISUELS DIVERGENTS.....	- 905 -
A) Modalité dominante en tâche de comptage syllabique.....	- 905 -
B) Perception de la divergence intermodale.....	- 909 -
1) Résultats Globaux.....	- 909 -
2) Influence de la position.....	- 909 -
3) Influence de la nature.....	- 910 -
4) Influence de la structure/longueur.....	- 918 -
Comparaison des performances : Comptage syllabique vs. Perception de la divergence inter-stimuli.....	- 921 -
COMPARAISON DES TACHES C2 (AUDITIVE), D (VISUELLE) ET E (AUDIOVISUELLE).....	- 924 -
1. Résultats globaux.....	- 924 -
2. Position.....	- 924 -
3. Famille Générale.....	- 926 -
4. Famille.....	- 927 -
5. Sous-Famille.....	- 928 -
6. Nature.....	- 929 -
7. Type de Sonorité.....	- 930 -
8. Ecart de Sonorité.....	- 931 -
<b>PARTIE EXPERIMENTALE IV : TESTS – INTERPRETATION.....</b>	<b>- 935 -</b>
TABLE DES MATIERES PARTIELLE.....	- 938 -
I. TACHE A : APPARIEMENT SEGMENTAL PHONEME - GRAPHEME (L/R ; B/V).....	- 940 -
I.1. Introduction.....	- 940 -
I.1.1. Mode de traitement.....	- 940 -
I.1.2. Interphonologie : catégories phonétiques, phonologiques et graphémiques.....	- 943 -
I.2. Interprétation linéaire des résultats.....	- 945 -
I.2.1. La position.....	- 945 -
I.2.2. La nature de la consonne.....	- 947 -
I.2.3. La nature de la voyelle adjacente.....	- 949 -
I.2.4. Choix de graphèmes.....	- 953 -
II. TACHE F : DISCRIMINATION PHONOLOGIQUE ET APPARIEMENT SEGMENTAL PHONEME - GRAPHEME (L/R ; B/V).....	- 954 -
II.1. Introduction.....	- 954 -
II.2. Interprétation linéaire des résultats phonographémiques (premier stimulus).....	- 955 -
II.2.1. La position.....	- 955 -
II.2.2. La nature de la consonne.....	- 956 -
II.2.3. La nature de la voyelle adjacente.....	- 958 -
II.2.4. La nature du couple de stimuli (identiques/différents).....	- 960 -
II.3. Interprétation linéaire des résultats phonologiques.....	- 960 -
II.3.1. La position.....	- 960 -
II.3.2. La nature des consonnes.....	- 960 -
II.3.3. La nature de la voyelle adjacente.....	- 961 -
II.3.4. La nature du couple de stimuli (identiques/différents).....	- 962 -



III. TACHES C1 ET C2 : COMPTAGE SYLLABIQUE EN UNIMODALITE AUDITIVE .....	- 963 -
III.1. <i>Tâche C1 - Epenthèses</i> .....	- 964 -
III.1.1. Position .....	- 964 -
III.1.2. Nature .....	- 966 -
III.1.3. Sonorité.....	- 969 -
III.2. <i>Tâche C1 – Effacements</i> .....	- 969 -
III.2.1. Position .....	- 971 -
III.2.2. Nature .....	- 971 -
III.2.3. Sonorité.....	- 972 -
III.3. <i>Tâche C1 – Comparaison Epenthèses/Effacements</i> .....	- 973 -
III.4. <i>Tâche C2 - Epenthèses</i> .....	- 974 -
III.4.1. Position .....	- 974 -
III.4.2. Nature .....	- 974 -
III.4.3. Sonorité.....	- 978 -
III.5. <i>Tâche C2 - Effacements</i> .....	- 979 -
III.5.1. Position .....	- 979 -
III.5.2. Nature .....	- 979 -
III.6. <i>Tâche C2 – Comparaison Epenthèses/Effacements</i> .....	- 980 -
IV. TACHE D : COMPTAGE SYLLABIQUE EN UNIMODALITE ORTHOGRAPHIQUE .....	- 980 -
IV.1. <i>Position</i> .....	- 982 -
IV.2. <i>Nature</i> .....	- 983 -
V. TACHE E : COMPTAGE SYLLABIQUE EN BIMODALITE AUDIO-ORTHOGRAPHIQUE .....	- 983 -
V.1. <i>Stimuli auditifs et orthographiques convergents</i> .....	- 987 -
V.1.1. Position .....	- 987 -
V.1.2. Nature .....	- 988 -
V.1.3. Sonorité.....	- 989 -
V.2. <i>Stimuli auditifs et orthographiques divergents</i> .....	- 989 -
V.2.1. Modalité dominante en tâche de comptage syllabique.....	- 989 -
V.2.2. Discrimination intermodale de stimuli divergents .....	- 991 -
V.2.2.1. Position .....	- 991 -
V.2.2.2. Nature.....	- 992 -
V.2.2.3. Sonorité.....	- 994 -
V.2.2.4. Longueur/Structure des stimuli divergents.....	- 994 -
V.3. <i>Comptage syllabique vs. Discrimination intermodale</i> .....	- 995 -
VI. COMPARAISON TACHES C2, D ET E : UNIMODALITE AUDITIVE VS. UNIMODALITE ORTHOGRAPHIQUE VS. BIMODALITE AUDIO-ORTHOGRAPHIQUE .....	- 995 -
VII. SYNTHESE ET PERSPECTIVES .....	- 1003 -
VII.1. <i>Synthèse</i> .....	- 1004 -
VII.2. <i>Implications méthodologiques</i> .....	- 1011 -
VII.3. <i>Perspectives</i> .....	- 1013 -
VII.3.1. Phonétique-Phonologie.....	- 1013 -
VII.3.2. Psycholinguistique .....	- 1016 -
VII.3.3. Didactique et Apprentissage.....	- 1019 -

## **Partie Expérimentale I :**

### **Pré-tests**

#### **I. Introduction**

Rappelons que le travail que nous menons se situe dans une perspective de psycholinguistique appliquée à l'enseignement/ apprentissage de la phonologie d'une langue étrangère. Nous cherchons à déterminer l'existence ou l'absence d'une influence des représentations orthographiques des systèmes langagiers en présence sur cet apprentissage. En particulier, lorsque les codes orthographiques sont partagés entre la L1 et la LE (avec des cas d'homographie interlinguale, totale ou partielle), il est possible que les tâches de lecture, massivement présentes tout au long de l'apprentissage d'une langue étrangère, notamment au début, puissent ne pas être favorables à l'apprentissage du système phonologique, en particulier lorsqu'il n'y a pas eu d'instruction spécifique préalable, car les codes graphiques ont tendance à être recodés avec le système phonologique de la L1, associant ainsi à des unités lexicales ou sublexicales en LE une forme phonologique propre à la L1. L'idée n'est pas neuve en didactique des langues : la méthodologie SGAV en particulier préconisait d'attendre, pour introduire le support écrit en classe, une vingtaine d'heures, de manière à « fixer le nouveau système phonologique » au préalable.

Notre hypothèse de recherche peut donc se formuler ainsi : le système phonographémique de la L1, sous certaines conditions, influence l'acquisition du système phonologique de la L2. Cette influence peut s'exprimer *structurellement* (système orthographique en partie, au moins, partagé par les deux langues) ou *procéduralement* (habitudes de lectures liées au système phonographémique 1 appliquées au système phonographémique 2). Afin de pouvoir évaluer cette influence, il convient donc au préalable préciser les relations suivantes :

- Phonologie 1 / Phonologie 2

- Orthographe 1 / Orthographe 2
- Phonologie 1 / Orthographe 1
- Phonologie 2 / Orthographe 2

Il faut ensuite préciser dans quelle tâche didactique/psycholinguistique cette influence peut se manifester :

- Lecture à voix haute
- Répétition
- Dictée
- Production orale spontanée

Chacune de ces tâches peut cependant être décomposée en sous-tâches, ou du moins être analysées à différents niveaux, linguistiques et/ou psycholinguistiques, précisément selon le matériel linguistique employé. C'est ici qu'il convient de préciser, dans notre cas, les conditions évoquées plus haut :

- Age : adulte
- L1 : japonais
- L2 : français
- Phénomènes linguistiques : phonologie française segmentale (/l, R, v, b/) et syllabique (groupes biconsonantiques /CC/).
- Phénomènes psycholinguistiques : perception de ces phénomènes en code oral et en code écrit, puis apprentissage de ce sous-système en perception et en production.

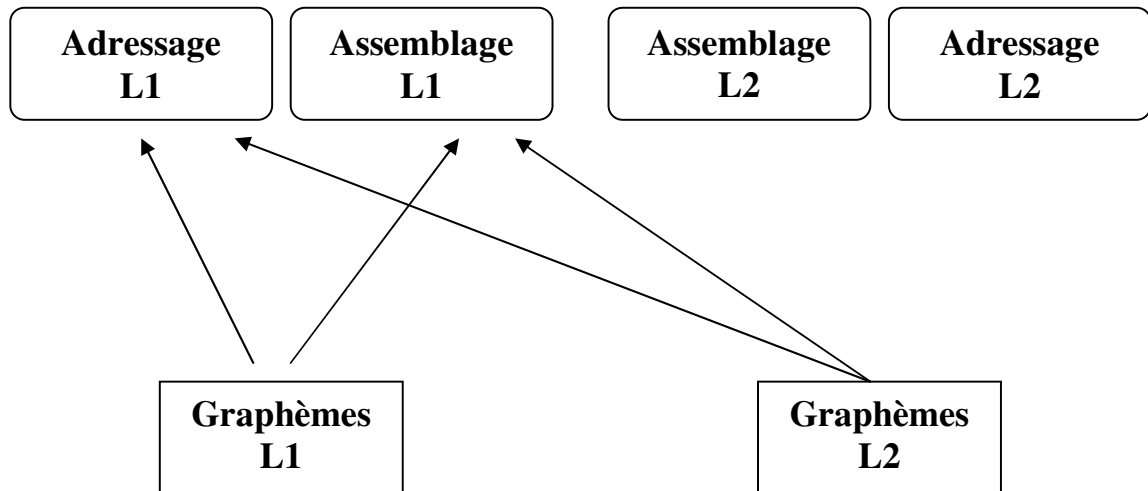
Un certain réseau de contraintes (au sens général) commence alors à se dessiner pour les apprenants japonais, notamment :

- A) Les habitudes de lecture et les codes d'écriture en japonais.
- B) Le chevauchement partiel du code orthographique romaji avec le code orthographique français.

- C) La profondeur de l'orthographe française (lettres muettes, homophones, etc.).
- D) Les différences de codes phonologiques japonais/français.
- E) Les différences de traitement entre matériel oral et écrit (gestion temporelle en particulier).

En ce qui concerne A) : les habitudes de lecture sont effectivement différentes, puisqu'elles impliquent par exemple une orientation différente (droite-gauche ; haut-bas ; gauche-droite), des scripts différents (et de profondeurs différentes) selon les strates lexicales (kanji, katakana, hiragana, romaji), l'oralisation de tous les caractères (pas de « caractères muets » : même si les relations phonographémiques japonaises présentent une certaine complexité, la profondeur des systèmes japonais et français est bien différente), une absence globale d'espace (blancs) entre les mots écrits, une ponctuation partiellement différente, et bien sûr un degré d'implication des voies d'assemblage (phonologique, sublexicale) et d'adressage (visuo-orthographique, lexicale) différentes.

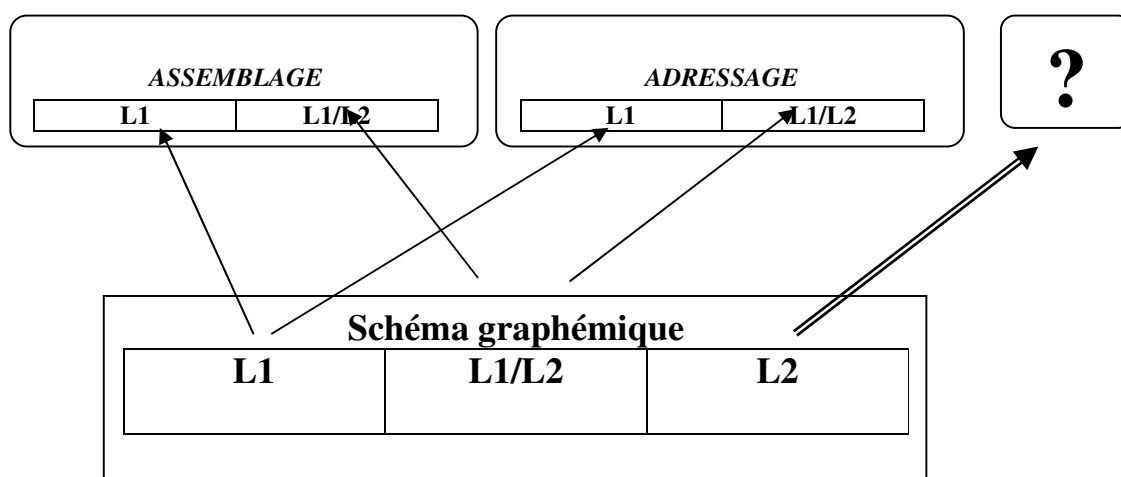
En ce qui concerne B) : les systèmes de transcription qui permettent de passer des kanji et des kana vers une forme alphabétique, en particulier le système romaji, établissent ainsi une zone de chevauchement graphique entre les deux systèmes orthographiques. Le transfert négatif le plus évident, *a priori*, serait de considérer les graphèmes étrangers comme ceux de la L1, ce qui nous conduit au schéma suivant :



Assemblage Graph2/Phon1	Adressage Graph2/Phon1
<p>En ce qui concerne la procédure d'assemblage Graph2/Phon1, même si l'on peut facilement en imaginer le déroulement, plusieurs problèmes se posent :</p> <ul style="list-style-type: none"> <li>- D'un point de vue segmental, que se passe-t-il lorsqu'un graphème est inconnu en L1 ?</li> <li>- D'un point de vue syllabique, que se passe-t-il lorsqu'une combinaison de graphèmes est inconnue ou illicite en L1 ?</li> <li>- D'un point de vue prosodique, la pauvreté informative du code écrit implique une implication forte des processus top-down pour l'encodage prosodique en lecture à voix haute par exemple. Ce vide apparaît</li> </ul>	<p>On pourrait <i>a priori</i> penser qu'une telle opération d'adressage ne peut avoir lieu, puisque les mots de L2 ne sont pas ceux de L1. Cependant il existe plusieurs possibilités :</p> <ul style="list-style-type: none"> <li>- Les emprunts ;</li> <li>- Les mots « cognates » ou homographes interlingaux ;</li> <li>- La ressemblance visuo-orthographique pour des mots autres.</li> </ul>

<p>nettement dans le cas de l'accentuation par exemple. Néanmoins, le problème est identique en ce qui concerne l'interprétation des signes de ponctuation et donc la restitution rythmique et intonative du discours.</p>	
--	--

Nous pouvons donc proposer une autre version du schéma précédent :



En ce qui concerne C) : l'absence de connaissance du système orthographique français peut conduire les sujets à établir des relations phonographémiques et développer des formes phonologiques erronées. Il faut prendre en compte ici le stade d'apprentissage global de la langue française des apprenants, puisque certaines unités orthographiques correspondent à des unités morphologiques qui ne sont pas décomposables (par exemple la marque verbale de la troisième personne du pluriel « -ent »). Les « lettres muettes », par exemple, peuvent être prononcées par les sujets novices ; les di-/trigraphes peuvent être di-triphtongués ou bien décomposés en unités plus réduites, etc.

En ce qui concerne D) : les différences de codes phonologiques conduisent à des problèmes para- et post-orthographiques.

En ce qui concerne E): les différences de traitement de l'oral et de l'écrit, en particulier au niveau des contraintes temporelles (par exemple : fixations oculaires et retour en arrière à l'écrit *vs.* impossibilité à l'oral ; absence relative de schéma rythmique et prosodique à l'écrit *vs.* présence à l'oral), peuvent conduire à des apprentissages différents et rendent non-triviale la question du transfert des compétences de l'un à l'autre canal.

Si l'on souhaite mettre en évidence une influence possible du système orthographique en L1 sur l'apprentissage du système phonologique en L2, il nous faut également expliciter la ou les théories de l'apprentissage linguistique qui sous-tendent notre approche. Si l'on se place dans une perspective dichotomique entre apprentissage implicite et explicite, nous nous intéressons ici pour l'instant à la dimension implicite de l'apprentissage, voire à ce que, en didactique des langues, on peut appeler « l'apprentissage par l'action ». En effet, ce qui nous intéresse ici, c'est d'observer le comportement des apprenants précisément lorsque la question du transfert phonographémique est négligée par l'enseignant ou par le corps méthodologique auquel renvoie le mode d'instruction. Notre hypothèse est qu'une telle négligence peut s'avérer néfaste pour l'apprentissage, alors qu'un traitement raisonné de la question peut au contraire le faciliter. Dans la perspective d'un mode d'instruction favorisant l'apprentissage implicite, on peut considérer que la fréquence d'exposition au matériel verbal français est un facteur déterminant pour l'apprentissage. Le rapport entre fréquence d'exposition aux formes orales et fréquence d'exposition aux formes écrites semble alors pertinent dans notre étude. Enfin, si l'on considère que la perception (au sens large) de matériel verbal en L2 constitue le préalable indispensable à tout apprentissage, alors l'étude de l'influence du système phonographémique en L2 sur la perception constitue une première étape vers celle de l'influence de ce même système sur l'apprentissage.

Afin de confronter cette hypothèse générale à la réalité des phénomènes d'apprentissage, et de l'illustrer de manière rigoureuse et précise, nous avons adopté

une démarche quasi-expérimentale portant sur deux points précis de la phonologie du français, l'un segmental, l'autre syllabique. Ayant posé certaines hypothèses relatives à l'apprentissage de la phonologie du français par des sujets japonais à partir de modèles phonologiques (structure syllabique, inventaire segmental, OT, TCRS) et de modèles psycholinguistiques (perception de la parole, mémoire de travail, reconnaissance visuelle des mots), nous avons monté un plan d'expérience nous permettant de manipuler certaines variables, d'où le qualificatif d' « expérimental ». Le préfixe « quasi » se justifie étant donné les conditions « écologiques » de passation des tests, qui comportent un nombre important de variables parasites. L'objectif était d'associer autant que possible la rigueur des tests expérimentaux au naturel de situations d'apprentissage réelles, ces dernières étant indispensables, à notre sens, à toute validation et généralisation didactique, voire psycholinguistique, le passage du laboratoire de psychologie expérimentale à la salle de classe de langue étant toujours délicat à effectuer sans hésitation. Ce protocole expérimental n'a évidemment pas pour ambition de valider notre hypothèse générale de recherche, mais plutôt d'apporter certains éléments de réponse à nos interrogations.

## **II. Méthodologie**

Avant de nous lancer dans la mise en œuvre de tests définitifs, il nous a semblé utile<sup>162</sup> d'effectuer certains prétests préliminaires sur une population réduite de sujets, afin de calibrer et d'affiner les tests ultérieurs. Nous avons donc effectué certains prétests exploratoires, qui nous ont certes permis d'obtenir quelques résultats, mais qui nous ont surtout permis de réaménager notre protocole pour des tests plus étendus.

Concernant les aspects méthodologiques de notre démarche, nous nous sommes référés dans un premier temps à Keller (1976). Dans son *Introduction aux systèmes psycholinguistiques*, l'auteur décrit une procédure expérimentale standard de la manière schématique suivante (élaboration et exécution du projet de recherche) :

---

<sup>162</sup> Nous tenons notamment à remercier ici J.-L. Nespoulous, J. Durand et J.-L. Azra.



1. Sélection d'un phénomène (doit refléter le fonctionnement linguistique) ;
2. Observations informelles (fréquence et importance du phénomène) ;
3. Transcription et structuration des observations informelles (essayer de découvrir un schéma, une règle de fonctionnement interne dans les données recueillies) ;
4. Hypothèse de recherche (très générale, en termes de relation causale probable (cause--variable indépendante ; effet--variable dépendante)) ;
5. Hypothèse de travail (vérifier un ou plusieurs aspects ou implications de l'hypothèse de recherche ; c'est l'opérationnalisation de cette dernière hypothèse, ainsi ramenée à une situation particulière) ;
6. Les modalités de l'expérience (taille de l'échantillon, lieu et temps, mode d'observation, mode de transcription, et classification des données établies) ;
7. Exécution de l'expérience (tenir compte de toutes les variables : âge, durée de l'apprentissage, classe sociale, etc.) ;
8. Analyse des données ;
9. Interprétation des résultats (en regard de l'hypothèse de travail) ;
10. Intégration des résultats (en regard de l'hypothèse de recherche, avec reformulation possible de celle-ci, suite aux observations).

Même si, depuis 1976, des avancées méthodologiques considérables ont été effectuées dans le domaine, qu'il s'agisse des outils ou bien des procédures, les grandes lignes tracées par Keller restent toujours valables. En ce qui concerne la présentation de la recherche, toujours d'après Keller, il faut :

- 1) Situer l'expérience par rapport à son cadre théorique ;
- 2) Présenter les hypothèses de recherche et de travail ;
- 3) Communiquer juste assez de détails méthodologiques concernant l'expérience pour une duplication éventuelle ;
- 4) Suffisamment de résultats pour que le lecteur puisse décider si l'hypothèse de recherche est appuyée ou infirmée.

C'est donc principalement de cette démarche que nous nous sommes inspirés, ainsi que de celles présentées dans certaines références (par exemple Dupoux et al., 1997 ; Monahan, 2001 ; Loonis, 2002) et dans les grandes revues des domaines concernés.

### Phénomène sélectionné

En situation d'apprentissage de la phonologie d'une langue étrangère, on constate que l'input et l'output ne sont pas identiques. L'input subit des modifications lors de son passage en intake/output, et les apprenants ont des difficultés à intégrer fidèlement le système phonologique étranger. Cela débouche sur une prononciation non-native et sur un certain degré d'incapacité à accomplir certaines tâches (en production ou en compréhension) vis-à-vis des standards natifs. Il convient de distinguer deux étapes, qui chacune renvoie à une problématique phonologique particulière : la perception renvoie, dans une certaine mesure, à la phonologie des emprunts ; l'apprentissage renvoie à l'interphonologie. La perception *pour* l'apprentissage renvoie aux deux. Dans notre cas, nous nous sommes concentrés sur les groupes biconsonantiques tautosyllabiques et sur les segments /l, R, v, b/.

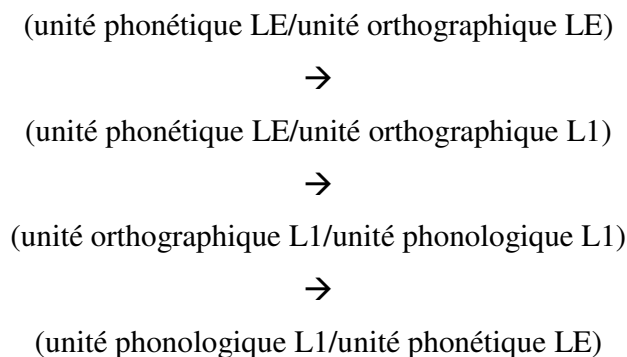
### Observations informelles, transcriptions et structuration

Dans le cas de sujets japonais adultes faux-débutants apprenant le français, nous pouvons très fréquemment observer des « difficultés » dans nos classes, qui conduisent soit à des épenthèses vocaliques dans le cas des groupes consonantiques, soit à des substitutions apparemment segmentales dans le cas des segments concernés.

### Hypothèse de recherche

Nous émettons l'hypothèse générale que, si des sujets adultes débutants apprennent une langue étrangère utilisant, en partie au moins, le même code orthographique que celui de leur langue première, alors, étant donné les processus à l'œuvre dans la boucle phonologique d'une part, et dans la reconnaissance visuelle de signes graphémiques d'autre part, le système phonographémique de la L1, ou certaines parties de ce système, va influencer sur l'apprentissage de la phonologie de la LE. Cette influence pourrait être soit positive soit négative pour cet apprentissage, une orientation qui ne peut être

déterminée qu'en fonction des tâches et des contextes précis d'enseignement / apprentissage, mais qui apparaîtrait *a priori* généralement négative, précisément en raison des écarts entre les deux systèmes phonographémiques en présence. Dans le cadre des études en interphonologie basées sur des modèles à contraintes (par exemple Grijzenhout et Van Rooy, 2001), on pourrait considérer que cette influence correspond à un obstacle au reclassement des contraintes de la L1, au vu du schéma général illustratif suivant, qui évoque un « effet Buben » du code orthographique de la L1 sur le système interphonologique en cours d'apprentissage :



Ce schéma général nécessite d'être détaillé en fonction des tâches et des unités linguistiques.

#### Hypothèse de travail

Dans le cas de non-mots, pour le public qui nous occupe, qu'il s'agisse de tâche de répétition ou de lecture à voix haute, les quatre possibilités suivantes de traitement sont à considérer :

- 1) Ressemblance du stimulus avec un mot japonais et pas/peu d'apprentissage préalable : mise en œuvre d'une procédure d'adressage conduisant à l'activation de la forme phonologique du mot japonais ;

- 2) Ressemblance du stimulus avec un mot japonais et apprentissage préalable : phénomène de compétition entre une procédure d'adressage en L1 et une procédure d'assemblage en LE.
- 3) Pas de ressemblance avec un mot japonais et pas/peu d'apprentissage préalable : mise en œuvre d'une procédure d'assemblage à l'aide de la boucle phonologique et du stock phonologique de la L1 conduisant à l'activation d'une forme phonologique adaptée (strate lexicale des emprunts) ;
- 4) Pas de ressemblance avec un mot japonais et apprentissage préalable : comme 2) mais avec utilisation du stock phonologique de la LE, conduisant à l'activation d'une forme phonologique française.

Nous supposons que les non-mots ne ressemblent, pour les apprenants au moins, à aucun mot français connu d'eux, excluant ainsi *a priori* la procédure d'adressage en LE. En tâche de dictée, vient s'ajouter la phase ultérieure de rappel ou d'assemblage de la forme graphique

Les hypothèses de travail relatives à chaque épreuve des prétests, conçus à des fins exploratoires, seront détaillées au fur et à mesure.

### Méthode :

Nous présentons ici la méthode suivie.

### *Sujets*

Il s'agit ici d'étudiants adultes en école de langue privée. Ils présentent donc des profils très variés, et leur motivation est généralement plus importante que celle de la majorité des étudiants universitaires. Quatre assistants, ainsi qu'un enseignant, japonais ont également pris part aux tests. Nous avons établi une première distinction entre les étudiants (A) d'une part, et les assistants-enseignants (B) d'autre part.

A) Etudiants :

<u>Nombre :</u>	17
<u>Sexe :</u>	9 femmes, 8 hommes.
<u>Age :</u>	6 sujets âgés entre 17-24 ans et 11 sujets âgés de plus de 25 ans
<u>Langue maternelle :</u>	japonais
<u>Discipline majeure :</u>	français : 5 ; Autre : 12
<u>Etude du français :</u>	0-6 mois : 6 sujets ; 6 mois-1 an : 1 sujet ; 1-2 ans : 5 sujets ; 2-4 ans : 3 sujets ; + de 4 ans : 2 sujets

Nous distinguons ici dans un premier temps les étudiants qui ont entre 0 et 6 mois d'étude du français (débutants) (A-A), et les autres (avancés) (A-B).

B) Assistants - enseignants

<u>Nombre :</u>	5
<u>Sexe :</u>	4 femmes, 1 homme.
<u>Age :</u>	1 sujet âgé entre 17-24 ans et 4 sujets âgés de plus de 25 ans.
<u>Langue maternelle :</u>	japonais
<u>Discipline majeure :</u>	français : 5
<u>Etude du français :</u>	0-6 mois : 0 ; 6 mois-1 an : 0 ; 1-2 ans : 1 sujet ; 2-4 ans : 2 sujets ; + de 4 ans : 2 sujets

***Epreuves***

Genre de tâche

Nous avons sélectionné 11 tâches, qui se répartissent en trois groupes, en fonction de la modalité des stimuli :

- Auditive ;
- Visuelle ;
- Audiovisuelle.

Les tâches sont de deux types :

- Discrimination entre deux stimuli (identiques ou différents) ;
- Choix de transcription (appariement stimuli/réponses).

Les tâches de discrimination concernent le versant perceptif du traitement linguistique, tandis que les tâches d'appariement impliquent, d'une certaine manière, le versant productif.

Les tâches ont donc été regroupées en 5 classes :

I Discrimination auditive,

II Choix de transcription à partir de stimuli auditifs (appariement auditif → orthographe/ symbole),

III Discrimination visuelle,

IV Choix de transcription à partir de stimuli visuels (appariement orthographe → symbole),

V Choix de transcription à partir de stimuli audiovisuels (appariement auditif - orthographe → orthographe/ symbole).

<b>I Présentation auditive</b> (6 tâches)	<b>I.1. Discrimination</b> (3 tâches)	I.1.1. Segmentale I.1.2. Syllabique I.1.3. Syllabique-Q
	<b>I.2. Choix de transcription</b> (3 tâches)	I.2.1. Segmentale I.2.2. Syllabique-Q et D-CC I.2.3. Syllabique-Q-Symbole
<b>II Présentation visuelle</b> (3 tâches)	<b>II.1. Discrimination</b> (1 tâche)	Segmentale
	<b>II.2. Choix de transcription</b> (2 tâches)	II.2.1. Syllabique-Q-Symbole II.2.2. Syllabique-D-CC
<b>III Présentation audiovisuelle</b> (2 tâches)	<b>Choix de transcription</b> (2 tâches)	III.1. Syllabique-Q et D-CC III.2. Syllabique-Q-Symbole

Note :

Q = Quantité (nombre de syllabes perçues),

D = Découpage (frontières syllabiques),

CC = présentation orthographique des réponses possibles ("CC" vs "CvC"),

Symbole = présentation non-orthographique des réponses possibles (ellipses : "OO" vs "OOO").

Les sujets devaient à chaque fois entourer la réponse de leur choix parmi plusieurs qui leur étaient proposées.

### ***Modalités d'expérience***

Les prétests ont été conduits de manière collective, dans un environnement familier aux apprenants, avec des instructions écrites en japonais et lues à voix haute par l'une de leurs professeurs habituels<sup>163</sup>. Celle-ci a pu clarifier certains points à l'oral, le cas échéant, pour assurer la complétion des documents. Lorsque l'on utilise des non-mots lors de tâches expérimentales, afin de minimiser le nombre de facteurs non-phonético-phonologiques, on peut reprocher à ce type d'expérience son manque de validité écologique. Rappelons que, à terme, dans une perspective didactique, l'environnement naturel que nous envisageons est celui des classes de langue. Puisqu'il s'agit d'apprentissage de langue étrangère, la critique portée perd de sa force, puisque, initialement, les lexies en langue étrangère pourraient facilement être considérées comme des non-mots par les apprenants débutants. En ce qui concerne les conditions de passation en revanche, s'il est vrai que des tests en laboratoire peuvent assurer un plus grand contrôle des différentes variables et ainsi une plus grande fiabilité des résultats, ces conditions sont loin d'être « écologiques » : l'environnement de la classe, avec tous les facteurs parasites, et la familiarité de cet environnement pour les apprenants, le sont bien davantage. Pour nous, qui tenons à prendre en considération la réalité des situations d'enseignement / apprentissage, afin d'assurer une certaine plausibilité à la valeur appliquée de notre travail, les conditions de passation de nos tests ne sont pas lacunaires, bien au contraire. Si nous reconnaissons la nécessité d'une approche rigoureuse des phénomènes à l'étude, et si nous apprécions à leur juste valeur la qualité des travaux en milieu contrôlé, c'est leur validité écologique qui assure à nos conditions expérimentales un certain intérêt et une certaine ouverture. Nous nous situons ainsi à mi-chemin entre étude de terrain et étude expérimentale, c'est-à-dire donc une étude quasi-expérimentale (Keller, 1976).

Nous reproduisons en annexe un exemplaire de la version française du document distribué aux étudiants, avec les consignes qui leur ont été données.

---

<sup>163</sup> Nous tenons à vivement remercier ici la collègue, K. K., qui nous a permis d'effectuer ces prétests auprès de ses étudiants.



Voici ci-dessous une description détaillée des conditions de passation des épreuves :

<ul style="list-style-type: none"> <li>⇒ <u>Date</u> :</li> <li>⇒ <u>Lieu</u> :</li> <li>⇒ <u>Description de la salle</u> :</li> <li>⇒ <u>Heure</u> :</li> <li>⇒ <u>Instructeurs</u> :</li> <li>⇒ <u>Nombre de sujets</u> :</li> <li>⇒ <u>Type de sujets</u> :</li> <li>⇒ <u>Activités précédentes des sujets</u> :</li> <li>⇒ <u>Familiarité des sujets avec l'environnement</u> :</li> <li>⇒ <u>Familiarité des sujets avec les tâches</u> :</li> <li>⇒ <u>Position des sujets dans la salle</u> :</li> <li>⇒ <u>Recrutement des sujets</u> :</li> <li>⇒ <u>Matériel audio</u> :</li> <li>⇒ Magnétophone et Enceintes :</li> <li>⇒ Volume :</li> </ul>	<ul style="list-style-type: none"> <li>⇒ 09.08.03.</li> <li>⇒ Ecole de langue privée à Tokyo.</li> <li>⇒ Grande salle, un peu sombre.</li> <li>⇒ 18H45</li> <li>⇒ Trois : nous-mêmes (qui nous sommes occupés de la cassette, de la distribution des documents et de l'écriture au tableau), K. K. (qui a lu les consignes à voix hautes) et T. D. (assis à l'arrière de la classe).</li> <li>⇒ 22</li> <li>⇒ Profils très variés. Etudes intensives de français.</li> <li>⇒ 8h de cours de français. Conjugaison verbale.</li> <li>⇒ Très bonne.</li> <li>⇒ Discrimination : bonne. Les autres : moins bonne.</li> <li>⇒ Répartis sur 4 colonnes et 7 rangs.</li> <li>⇒ Sujets semi-volontaires (on leur a proposé de rester à la fin du cours, certains sont partis).</li> <li>⇒ 2 enceintes séparées, situées face aux étudiants en hauteur</li> <li>⇒ « Normal »</li> </ul>
--	---

<p>⇒ <u>Matériel visuel</u> :</p>	<p>⇒ Nous avons écrit les stimuli visuels à l'aide d'une craie blanche sur un grand tableau noir de la salle de classe. Nous les avons écrits en grandes lettres majuscules telles que nous le ferions dans une classe standard. A peine écrits, nous les avons ensuite effacés, comme prévu dans notre protocole.</p>
<p>⇒ <u>Documents utilisés</u> :</p>	<p>⇒ Nous avons distribué les 5 feuilles destinées aux étudiants au fur et à mesure.</p>
<p>⇒ <u>Durée des pré-tests</u> :</p>	<p>⇒ Environ 20 minutes.</p>

### ***Stimuli employés***

Pour chaque tâche, nous avons sélectionné des cibles linguistiques (par exemple les segments /b/ et /v/), pour lesquelles nous avons ensuite élaboré des non-mots. Pour ces prétests, les cibles ont été choisies au hasard (position initiale ou médiale (et dans ce cas parfois double : anté-pénultième et pénultième), fréquence indéterminée, segments au hasard). Nous leur avons ensuite adjoint des distracteurs placés dans un ordre aléatoire. Les barres obliques / / renvoient aux stimuli auditifs, les guillemets droits " " aux stimuli orthographiques et les accolades { } aux stimuli audiovisuels, c'est-à-dire auditifs et orthographiques simultanés.

### **Stimuli auditifs**

#### Discrimination

<b>CIBLES</b>	<b>DISTRACTEURS</b>
/baliko/ /valiko/	/bakedo/ /bakedo/
/siliko/ /shiliko/	/shiloka/ /shiloka/
/psaliko/ /pusaliko/	/pusogeda/ /pusogeda/

Choix de transcription

SEGMENTALE	
CIBLE 1	CIBLE 2
/bagade/ → "bagade" ou "vagade"	/vatosi/ → "vatosi" ou "batosi"
/øpaki/ → "upaki" ou "%paki"	/usego/ → "usego" ou "%sego"

SYLLABIQUE (orthographique)			
CIBLE 1		CIBLE 2	
/ptanabo/ → "ptanabo" ou "putanabo"		/putebala/ → "ptebala" ou "putebala"	
"p-tanabo" ou "pt-anabo" ou "pta-nabo" ou "ptan-abo"	"pu-tanabo" ou "put-anabo" ou "puta-nabo" ou "putan-abo"	"p-tebala" ou "pt-ebala" ou "pte-bala" ou "pteb-ala"	"p-utebala" ou "pu-tebala" ou "put-ebala" ou "pute-bala"

SYLLABIQUE (symbole)
/ptab/ → "O", "OO" ou "OOO"

Conditions d'enregistrement des stimuli auditifs

Date : Le 05.08.03-16h30

Lieu : Salle insonorisée (studio d'enregistrement) (Tokyo)

Matériel : Micro Sony ECM-MS907 (120°) ; Net MD walkman Sony MZ-N1 ; Mini-disc Sony puis copié sur Casette audio TDK

Contenu (enregistré par nous-même) :

<b>Contenu</b>
Dai-ichi-bu // Ichi-ban/ Shiloka/ Shiloka // Ni-ban/ Baliko/ Valiko // Sam-ban/ Psaliko/ Pusaliko // Yom-ban/ Bakedo/ Bakedo // Go-ban/ Pusogeda/ Pusogeda // Roku-ban/ Siliko/ Shiliko // STOP
Dai-ni-bu // Ichi-ban/ Putebala //// Ni-ban/ %paki // Sam-ban/ Bagade // Yom-ban/ Ptanabo //// Go-ban/ Usego // Roku-ban/ Ptab // Nana-ban/ Vatosi // STOP
Dai-go-bu // Ichi-ban /// Okotle //// Ni-ban /// Adotole //// Sam-ban /// Afremo //// Yom-ban /// Ikrush // Go-ban /// Afuregu //// STOP

Consignes : Les numéros des stimuli ont été donnés en japonais, comme cela est indiqué dans le contenu : « Dai-ichi-bu » signifie « partie 1 » et « ichi-ban » signifie « numéro 1 ».

Durées approximatives :

- 1) / = 2 sec ; // = 5 sec
- 2) / = 2 sec ; // = 8 sec ; //// = 20 sec
- 3) // = 6 sec ; /// = 8 sec ; //// = 20 sec

Durée de l'enregistrement : environ 5 min 10.

Aspects suprasegmentaux : nous avons essayé d'adopter une intonation (type assertion) et un rythme (débit « normal » de lecture à voix haute) identiques pour tous les stimuli.

Document distribué au sujet : la traduction du document original français en japonais a été délicate et nous n'aurions pu la réaliser sans la patience et le dévouement de notre informante principale, Y. D.-O. La traduction a par ailleurs permis de soulever plusieurs points problématiques et a conduit au réexamen de certaines questions. Nous avons par ailleurs dû établir une nouvelle version en anglais à partir de la version

japonaise de manière à bien cerner ce qui avait été indiqué aux sujets (consignes et informations).

En ce qui concerne les transcriptions et les réponses proposées, nous avons dû effectuer certains choix. En particulier, nous avons établi les correspondances suivantes, en léger décalage avec le système phonographémique français standard :

[e] → "e"

[u] → "u"

[ø] → "ø"

Ces choix ont été pris en compte dans la composition des consignes et dans la présentation des épreuves aux sujets.

### **Stimuli visuels**

Les stimuli 2, c'est-à-dire se trouvant sur la feuille de l'étudiant (par opposition aux stimuli 1, c'est-à-dire à ceux présentés fugacement en lettres majuscules sur le tableau par l'enseignant) ont été imprimés en lettres minuscules, en police « Normal, MS Sans Serif », taille 14, en italiques, de manière à bien distinguer les formes visuelles (et non les signes graphémiques) des stimuli 1 et 2. Afin de minimiser l'effet de persistance visuelle, et valoriser au contraire la stratégie de subvocalisation, pour avoir accès aux représentations phonologiques activées par les stimuli visuels, nous avons placé un distracteur visuel précédant verticalement le stimulus 2 pour chaque stimulus sur la feuille des sujets.

Dans le cas du schwa, le problème du signe graphique "e" ou "eu" pour représenter le schwa phonologique s'est posé à nous. En caractères japonais romaji, le premier correspond au phonème /e/, tandis que le deuxième correspond un digraphe qui aurait fort pu perturber les sujets et les entraîner vers un genre de diphtongaison (courant en japonais pour transcrire certains sons étrangers). Puisque le signe arobase « @ »,

parfois utilisé pour représenter le schwa, notamment dans le code SAMPA (Caelen-Haumont, 2002), évoque trop le graphème « a », nous avons choisi le signe des pourcentages « % » qui nous a semblé assez neutre quant à l'évocation phonique.

Pour la tâche de comptage syllabique symbolique à partir d'un stimulus visuel, nous avons conservé l'orthographe française du mot "plateau" à dessein. Pour la transcription du phonème /ʃ/ (premier phonème du mot français "chat"), il aurait peut-être fallu préférer la transcription "SH" à "CH", car en romaji (par exemple dans *le Tuttle new dictionary of loanwords in Japanese* de Kamiya, 1994) le digraphe "CH" est utilisé pour transcrire l'affriquée /tʃ/, comme dans le mot anglais « choice », alors que le digraphe "SH" est utilisé pour transcrire la fricative /ʃ/ telle qu'on la trouve dans le mot anglais « English ». C'est le système de transcription Hepburn.

### Discrimination

<b>CIBLES</b>	<b>DISTRACTEURS</b>
"abigu / avigu"	"obeka / obeka"
"ogabi / ogavi"	"ekobu / ekobu"

### Choix de transcription

<b>CIBLE</b>
"plateau" → "OO" ou "OOO"
"apta" → "ap-ta" ou "a-pta" ou "apt-a"

### **Stimuli audiovisuels : choix de transcription**

Pour les stimuli audiovisuels, les voyelles épenthétiques insérées dans les groupes /CC/ ont été choisies en fonction des schémas d'adaptation phonologique japonais et sont essentiellement /u/ et /o/. Pour ces épreuves, nous avons laissé une pause plus longue,

dans l'enregistrement des stimuli audio, entre le numéro du stimulus et le stimulus lui-même de manière à laisser le temps à l'instructeur d'écrire le stimulus visuel correspondant, de manière à l'effacer à peine le stimulus audio terminé et ainsi éviter la prépondérance du stimulus visuel. C'est aussi pour cela que nous demandions aux étudiants de répéter à voix haute ou basse, selon leur choix.

CIBLE 1		CIBLE 2	
{afRemo} → "afremo" ou "afuremo" ?		{afuRegu} → "afregu" ou "afuregu" ?	
a-fremo ou af-remo ou afr-emo ?	a-fu-remo ou af-u-remo ou a-fur-emo ?	a-fregu ou af-regu ou afr-egu ?	a-fu-regu ou af-u-regu ou a-fur-egu ?
{okotle} → "okotle" ou "okotole" ?		{adotole} → "adotle" ou "adotole" ?	
oko-tle ou okot-le ou okotl-e ?	oko-to-le ou okot-o-le ou oko-tol-e ?	ado-tle ou adot-le ou adotl-e ?	ado-to-le ou adot-o-le ou ado-tol-e ?
{ikRuch} → "OO", "OOO" ou "OOOO"		NUL	

### **III. Résultats et interprétation**

Si l'on examine à présent chaque tâche en détail, voici les résultats que l'on obtient, étant entendu que nous ne présentons ici essentiellement que les plus significatifs d'entre eux. Notons que si l'analyse statistique permet de dégager certaines tendances au niveau du groupe, il conviendrait également, dans une perspective didactique, d'établir le profil de chaque apprenant afin d'avoir une vue détaillée de la situation de leur(s) système(s) interphonologique(s), en perception ou en production, selon les tâches.

Tâches 1-2-3 : Discrimination auditive.

Il s'agit ici d'une simple tâche de discrimination auditive, visant à déterminer certaines propriétés du système interphonologique des apprenants, dans ce type de tâche. Nous nous intéressons particulièrement au niveau segmental (/b/ vs /v/), mais aussi syllabique (/si/ vs /shi/ et /ps/ vs /pus/).

Résultats :

Perception	Globalement (22)	Degré d'expertise	
		Débutants (6)	Autres (16)
<b>B/V identique</b>	68%	50%	75%
<b>B/B différent</b>	22%	0	31%
<b>SI/SHI identique</b>	4, 5%	16%	0%
<b>SHI/SHI différent</b>	13%	16%	12%
<b>PS/PuS identique</b>	4, 5%	16%	0%
<b>PuS/PuS différent</b>	54%	66%	50%

Commentaires :

Nous nous attendions à des confusions entre /b/ et /v/, /si/ et /shi/, /ps/ et /pus/. Globalement, seule la confusion b/v a été largement réalisée. Nous devons noter que, après avoir nous-mêmes écouté les stimuli, la distinction b/v n'était auditivement pas très nette dans l'enregistrement. Cela pourrait peut-être expliquer l'importance de la confusion. En revanche, nous avons été surpris par le résultat concernant le distracteur pus/pus. Il n'est pas impossible que le critère de différenciation ne nous soit pas apparu (par exemple la longueur de la voyelle, non-pertinent en français mais pertinent en japonais). Ces discriminations erronées pourraient peut-être relever du domaine allophonique, et non phonémique. En ce qui concerne l'absence de palatalisation du /si/ en /shi/ (la syllabe /si/ est absente de la distribution syllabique du japonais), on peut



l'attribuer soit à la distinction production (palatalisation)/perception (absence de palatalisation), soit à une question de strate lexicale : si les non-mots sont assignés à la classe des mots étrangers, il n'est pas impossible que la palatalisation n'y soit pas réalisée. Nous pouvons donc résumer les résultats comme suit :

<b>Résumé</b>
<ul style="list-style-type: none"> <li>- Confusion segmentale b/v importante, tant chez les novices que chez les experts ;</li> <li>- Hypercorrection chez les experts : discrimination erronée b/b ;</li> <li>- Confusion si/shi faible, nulle chez les experts ;</li> <li>- Hypercorrection chez les novices et les experts : discrimination erronée shi/shi ;</li> <li>- Confusion syllabique ps/pus faible, nulle chez les experts ;</li> <li>- Hypercorrection forte chez les novices et les experts : discrimination erronée pus/pus.</li> </ul>

Tâche 4 : Appariement segmental auditif/visuel orthographique.

Il s'agit ici d'une tâche d'appariement phonographémique destinée à nous renseigner sur le système phonographémique des apprenants et, à travers lui, sur leur système interphonologique.

Résultats :

<b>Appariement</b>	<b>Globalement (22)</b>	<b>Degré d'expertise</b>	
		<b>Débutants (6)</b>	<b>Autres (16)</b>
[bagade]/ « vagade »	32 %	50 %	25 %
[vatosi]/ « batosi »	27 %	17 %	31 %
[øpaki]/ « upaki »	59 %	67 %	56 %
[usego]/ « %sego »	36 %	50 %	31 %

Commentaires :

Globalement nos prévisions ont été réalisées, à savoir un taux de confusion phonographémique non négligeable pour les segments considérés. Nous remarquons,

d'une part que ce sont les segments vocaliques qui donnent lieu à la plus grande confusion, et d'autre part que les novices commettent plus d'erreurs que les experts, excepté pour le couple (/v/, "b"). Il n'est cependant pas impossible que le choix du symbole « % » pour représenter le /ø/ ait eu une influence sur ce résultat.

<b>Résumé</b>
<ul style="list-style-type: none"> <li>– Confusion phonographémique non négligeable pour les segments considérés ;</li> <li>– Confusion phonographémique vocalique plus élevée que confusion consonantique ;</li> <li>– Globalement plus de confusions chez les novices que chez les experts.</li> </ul>

Tâche 5-A : Appariement syllabique auditif/visuel orthographique.

Il s'agit ici d'une tâche d'appariement phonographémique orthographique, destinée à nous renseigner sur le système phonographémique interlangagier des apprenants dans une telle tâche, et à travers elle, sur leur système interphonologique. Nous nous intéressons aux épenthèses et effacements vocaliques dans un environnement biconsonantique (/pt\_/ et /p\_t/).

Résultats :

<b>Appariement</b>	<b>Globalement (22)</b>	<b>Degré d'expertise</b>	
		<b>Débutants (6)</b>	<b>Autres (16)</b>
<b>[Putebala]/ « ptebala » (effacement/dévoisement)</b>	20 %	17 %	19 %
<b>[Ptanabo]/ « putanabo » (insertion/épenthèse)</b>	32 %	50 %	25 %

Commentaires :

Nous nous attendions à l'effacement de la voyelle [u] dans le contexte /p\_t/ et à son insertion dans le groupe /pt/. Globalement nos prévisions ne se sont pas réalisées, excepté l'insertion vocalique pour les novices. L'absence d'effacement vocalique

graphémique peut s'expliquer par la nature de la tâche : même si le /u/ est phonétiquement dévoisé, il est phonologiquement présent et donc rétabli lors d'une transcription graphémique, le code graphémique étant, pour l'essentiel, et en japonais au moins, phonologique et non phonétique (de la même manière que [pti] s'écrira "petit"). Les facteurs prosodiques jouent par ailleurs un rôle important dans le dévoisement, et n'y étaient peut-être pas favorable ici, étant donné le rythme assez lent que nous avons choisi. Concernant l'épenthèse, l'apprentissage de la syllabation du français apparaît ici clairement dans l'écart entre les résultats des novices et des experts.

### Résumé

- Peu d'effacement vocalique graphémique (/CvépC/ → "CC") ;
- Insertion vocalique graphémique élevée chez les novices et faible chez les experts (/CC/ → "CvépC").

#### Tâche 5-B : Segmentation syllabique orthographique.

Il s'agit ici d'une tâche de segmentation syllabique orthographique, destinée à nous renseigner sur la manière dont les sujets placent les frontières syllabiques *via* la représentation orthographique. Cette tâche fait suite à la tâche précédente dont elle dépend : la segmentation dépend donc alors de la perception initiale des stimuli (/CC/ ou /CvépC/) par les sujets.

#### Résultats :

Pour la segmentation, nous considérons que nous avons eu X occurrences de perception de /CvépC/ et Y occurrences de perception de /CC/ (avec 22 sujets et 2 stimuli = 44 occurrences). A partir de cela, nous avons deux groupes : ceux qui ont perçu /CVC/ et les autres /CC/. Nous considérons alors le choix de segmentation orthographique.

Segmentation orthographique	Globalement		Degré d'expertise			
	Perçu CVC (25)	Perçu CC (19)	Débutants		Autres	
			Perçu CVC (9)	Perçu CC (4)	Perçu CVC (16)	Perçu CC (15)
C—VCVCVCV	4 %	16 %	0		6 %	20 %
CV—CVCVCV	48 %	0	44 %		50 %	0
CC—VCVCV		10 %		0		13 %
CCV—CVCV		74 %		100 %		67 %
CVC—VCVCV	4 %		0		6 %	
CVCV—CVCV	44 %		56 %		38 %	

Commentaires :

Cette épreuve était tout à fait exploratoire, motivée empiriquement par les problèmes récurrents de nos étudiants concernant le découpage typographique des mots en tâche d'écriture, lorsque l'espace est insuffisant pour écrire le mot complet sur une seule ligne. Les informations fournies par cette tâche sont globalement assez limitées. Si nous pouvons être surpris par la proportion de sujets experts ayant isolé la première consonne (découpage /C-/) le point particulièrement intéressant porte finalement moins sur le découpage orthographique que sur la longueur syllabique pour les formes /CvépC/ : alors que nous nous attendions à un découpage monosyllabique du type /CV-CVCVCV/, il apparaît que les formes dissyllabiques /CVCV-CVCV/ ont été préférées par les novices et choisies dans une proportion encore importante par les experts.

**Résumé**

- En segmentation orthographique, les formes dissyllabiques /CVCV-CVCV/ sont préférées par les novices aux formes monosyllabiques /CV-CVCVCV/.

**Tâche 6 : Appariement syllabique auditif/visuel symbolique.**

Il s'agit ici d'une tâche d'appariement destinée à évaluer, d'une part le taux d'épenthèse pour le groupe consonantique considéré, et d'autre part, en partie au moins, l'influence des représentations orthographiques sur une telle tâche, en comparant ces résultats à ceux de la tâche d'appariement syllabique précédente.

**Résultats :**

Appariement [ptab] (1 syllabe)	Globalement (22)	Degré d'expertise	
		Débutants (6)	Autres (16)
O (1 syllabe)	41 %	17 %	50 %
OO (2 syllabes)	45 %	66%	37,5 %
OOO (3 syllabes)	14 %	17%	12,5 %
<i>Total épenthèses</i>	59%	83%	50%

**Commentaires**

Concernant le taux d'épenthèse pour le groupe /pt/ on voit qu'il est élevé (59%), et supérieur à celui de la tâche orthographique (32%). Cela est vrai tant pour les novices (83%>50%) que pour les experts (50%>25%). Il faut cependant souligner que le stimulus est monosyllabique, ce qui lui confère un caractère exceptionnel par rapport aux stimuli trisyllabiques utilisés dans la tâche d'appariement précédente. Nous remarquons également que le traitement de la coda (consonne finale) donne lieu à un taux d'épenthèse non négligeable, en particulier chez les novices (17%).

<b>Résumé</b>
<ul style="list-style-type: none"> <li>- Taux d'épenthèses important pour le groupe /CC/ dans cette tâche ;</li> <li>- Taux d'épenthèses plus élevé dans la tâche de comptage avec support symbolique que dans celle avec support orthographique ;</li> <li>- Taux d'épenthèses plus élevé chez les novices que chez les experts ;</li> <li>- Insertion vocalique après consonne finale dans une proportion non négligeable.</li> </ul>

Tâche 7 : Discrimination visuelle (orthographe/orthographe).

La tâche a été présentée aux sujets en attirant leur attention sur l'existence d'homophones et d'homographes en français. La consigne était donc d'indiquer si le mot présenté fugacement au tableau et le mot écrit sur leur feuille se prononçaient de la même manière ou non.

Résultats :

Perception	Globalement (22)	Degré d'expertise	
		Débutants (6)	Autres (16)
<b>Bi/Vi (fin<sup>164</sup>) Identique</b>	9 %	17 %	6 %
<b>Bu/Bu (fin) Différent</b>	14 %	50 %	0
<b>Vi/Vi (méd) Différent</b>	9 %	33 %	0
<b>Be/Be (méd) Différent</b>	14 %	33 %	6 %

Commentaires

Globalement, les taux de confusions sont assez faibles, quasi-nuls chez les experts. Cependant, chez les novices, le taux de confusion de type hypercorrectif (distinction erronée de stimuli identiques) est loin d'être négligeable, soulignant ainsi l'instabilité du système phonographémiques chez de tels sujets. Nous ne pouvons en revanche déterminer si cette hypercorrection relève du domaine phonologique (hypercorrection erronée) ou phonétique (hypercorrection allophonique). Bien sûr, malgré les distracteurs visuels, on doit prendre en compte l'effet de persistance visuelle, ainsi que le rôle des savoirs déclaratifs des sujets à propos du système phonographémique français dans une telle tâche. Notons également qu'il y a eu une erreur méthodologique, sans gravité, dans notre protocole : au lieu de présenter les couples "ogaBi / ogaVi", "ekoBu / ekoBu", "aBigu / aVigu" et "oBeka / oBeka", nous avons présenté les couples "ogaBi / ogaVi", "ekoBu / ekoBu", "avigu / avigu" et "oBeka / oBeka", d'où les résultats, corrects, indiqués ci-dessus.

---

<sup>164</sup> fin = position finale ; méd = position médiale.

<b>Résumé</b>
<ul style="list-style-type: none"> <li>– Peu de confusion phonographémique segmentale chez les novices dans une telle tâche</li> <li>– Confusion de type hypercorrectif élevé chez les novices, indiquant une possible instabilité du système phonographémique interlangagier.</li> </ul>

Tâche 8 : Appariement syllabique visuel orthographe/symbole.

Il s'agit ici d'une tâche d'appariement syllabique dont les résultats visent à être comparés avec les tâches similaires précédentes, afin de déterminer si la modalité et le format de présentation des premiers stimuli influent ou non sur le taux d'épenthèse dans le traitement du groupe consonantique /pt/.

Résultats :

<b>Appariement "PLATEAU" (2 syllabes)</b>	<b>Globalement (22)</b>	<b>Degré d'expertise</b>	
		<b>Débutants (6)</b>	<b>Autres (16)</b>
OOO (3 syllabes)	37 %	67 %	25 %

Commentaires

Même si le respect de l'orthographe française du mot « plateau » et le trigraphe « eau » qu'il contient peuvent mettre en doute le taux d'épenthèse présenté ici (on pourrait penser à un découpage du type « pla-te-au » ou « plat-e-au » par exemple), nous supposons que le découpage adopté par les sujets qui ont compté trois syllabes résulte d'une épenthèse conduisant au découpage « pu-la-teau ». Dans ce cas, on voit que le taux est élevé chez les novices et non négligeable chez les experts. Si l'on compare ces taux (37%, 67%, et 25%) avec ceux de la tâche d'appariement auditif/orthographe (32%, 50%, et 25%) et auditif/symbole (45%, 66% et 37, 5%) on voit que certaines modalités sont plus propices à l'épenthèse que d'autres. Nous résumerons ces résultats plus bas, lorsqu'ils seront complétés.

<b>Résumé</b>
– Taux d'épenthèse élevé chez les novices et non négligeable chez les experts, dans une telle tâche.

Tâche 9 : Découpage.

Dans cette tâche, il s'agit de déterminer les frontières syllabiques *via* la représentation orthographique, lorsque le stimulus initial est lui aussi orthographique.

Résultats :

Segmentation orthographique "APTA"	Globalement (22)	Degré d'expertise	
		Débutants (6)	Autres (16)
V—CCV	27%	17%	31,5%
VC—CV	64%	83%	56%
VCC—V	9%	0	12,5%

Commentaires :

Globalement, et pour chaque sous-groupe, les formes monoconsonantiques sont largement préférées aux formes complexes. On pourrait donc penser que les contraintes portant sur les syllabes complexes (attaque et coda) sont classées plus haut que celles portant sur les codas simples.

<b>Résumé</b>
– Dans une telle tâche, les sujets préfèrent les formes simples (VC-CV) aux formes complexes (V-CCV, CCV-V).

Tâche 10-A : Appariement syllabique audio-visuel (audio-orthographique / orthographique).

Dans cette tâche, les stimuli initiaux sont audiovisuels et le comptage syllabique est effectué *via* leur représentation orthographique.



Résultats :

Rappelons que seuls les résultats les plus significatifs sont présentés ici, d'où les pourcentages non complets.

Appariement	Globalement (22)	Degré d'expertise	
		Débutants (6)	Autres (16)
{okotle}/ «okotole » (insertion/épenhèse)	45, 5%	50%	44%
{adotole}/ «adotle » (délétion/dévoisement)	0	0	0
{afremo}/ « afuremo » (insertion/épenhèse)	4, 5%	0	6%
{afuregu}/ « afregu » (délétion/dévoisement)	4, 5%	16%	6%

Commentaires :

Il apparaît dans ces résultats qu'une analyse grossière des groupes biconsonantiques (/tl/ et /fr/) n'est pas satisfaisante puisque, lorsque que la nature des groupes, la position des groupes et l'environnement vocalique varient, les résultats diffèrent considérablement. En effet, alors que le taux d'épenhèse est élevé pour le premier groupe (/tl/ en position médiale antépénultième avec un environnement /o\_e/) il est tout à fait négligeable pour le second.

<b>Résumé</b>
<ul style="list-style-type: none"> <li>– En présentation bimodale (audio/orthographe), le taux d'épenhèses est globalement élevé pour certains groupes ;</li> <li>– Ce taux varie considérablement selon le groupe (nature, position, environnement).</li> </ul>

**Tâche 10-B : Découpage.**

Cette tâche fait suite à la précédente.

**Résultats :**

Segmentation orthographique	Globalement		Degré d'expertise			
			Débutants		Autres	
	Perçu CVC (54)	Perçu CC (36)	Perçu CVC (14)	Perçu CC (12)	Perçu CVC (40)	Perçu CC (24)
VC—vép <sup>165</sup> CV	2%		0		2,5%	
VC—vép—CV	4%		0		5%	
VCvép—CV	85%		93%		82,5%	
VCvépC—V	9%		7%		10%	
V—CCV		75%		50%		87,5%
VC—CV		22%		50%		8,5%
VCC—V		3%		0		4%

**Commentaires :**

Les points à souligner ici sont les suivants : concernant les stimuli de type VCvépCV, le découpage VCvép—CV est largement majoritaire, ce que l'on peut interpréter comme un évitement des codas ; concernant les stimuli du type VCCV, le découpage V-CCV est privilégié pour les experts et équivalent au découpage VC-CV pour les novices. Ce dernier résultat ne correspond pas à celui de la tâche de découpage avec stimuli initiaux orthographique où les formes simples étaient préférées aux formes complexes, mais il confirme en revanche l'évitement des codas, puisque ce sont les attaques et non les codas complexes qui sont préférées.

---

<sup>165</sup> vép = voyelle épenthétique.

### Résumé

- Dans une telle tâche, les codas sont évitées ;
- Tandis qu'en présentation orthographique les formes simples sont préférées, en présentation bimodale, les formes complexes sont préférées pour les experts et équivalentes pour les novices.

Tâche 11 : Appariement syllabique audio-visuel (audio-orthographique/symbolique).

Dans cette tâche les stimuli sont audiovisuels, mais le comptage syllabique s'effectue *via* une représentation symbolique.

Résultats :

Appariement {IKRUSH} (2 syllabes)	Globalement (22)	Degré d'expertise	
		Débutants (6)	Autres (16)
OOO (3 syllabes)	18%	17%	19%
OOOO (4 syllabes)	18%	17%	19%
<i>Total épenthèses</i>	36%	34%	38%

### Commentaires

Dans cette tâche, le taux d'épenthèse n'est globalement pas très élevé. On peut noter, d'une part la répartition égale des réponses à une et deux épenthèses, et d'autre part le nombre légèrement supérieur de confusion chez les experts.

### Résumé

- Dans une telle tâche, en présentation bimodale, le taux d'épenthèse n'est pas très élevé et similaire chez les novices et les experts.

#### **IV. Résumé général et discussion**

Ces prétests avaient été conçus essentiellement à visée exploratoire, afin de mesurer les durées de passation et leur faisabilité d'une part, et de déceler des problèmes d'ordre méthodologique d'autre part. Ces deux objectifs ont été atteints. Nous avons en outre pu récolter certaines informations, tout à fait limitées bien sûr, étant donné le nombre réduit de sujets et les faiblesses du protocole, que nous résumons ci-dessous :

<b>Tâches</b>	<b>Résultats</b>
<p><u>Tâches I.1. :</u>  <u>Discrimination auditive</u></p>	<p><u>I.1.1. Segmentale : couple /b, v/</u></p> <ul style="list-style-type: none"> <li>- Confusion segmentale b/v importante, tant chez les novices que chez les experts.</li> <li>- Hypercorrection chez les experts : discrimination erronée b/b.</li> </ul> <p><u>I.1.2. Syllabique : couple /si, shi/</u></p> <ul style="list-style-type: none"> <li>- Confusion si/shi faible chez les novices, nulle chez les experts.</li> <li>- Hypercorrection chez les novices et les experts : discrimination erronée shi/shi.</li> </ul> <p><u>I.1.3. Syllabique-Q : couple /ps, pus/</u></p> <ul style="list-style-type: none"> <li>- Confusion syllabique ps/pus faible, nulle chez les experts.</li> <li>- Hypercorrection forte chez les novices et les experts : discrimination erronée pus/pus.</li> </ul>

<p><u>Tâches I.2. :</u>  <u>Appariement auditif → visuel</u></p>	<p><u>I.2.1. Segmental : couples /b, v/ et /u, ø/</u></p> <ul style="list-style-type: none"> <li>– Confusion phonographémique non négligeable.</li> <li>– Confusion phonographémique vocalique plus élevée que confusion consonantique.</li> <li>– Globalement plus de confusions chez les novices que chez les experts.</li> </ul> <p><u>I.2.2. Syllabique-Q/-D orthographique : couple /pt, put/</u></p> <p><u>Q :</u></p> <ul style="list-style-type: none"> <li>– Insertion vocalique graphémique élevée chez les novices et faible chez les experts (/CC/ → "CvépC").</li> <li>– Peu d'effacement vocalique graphémique (/CvépC/ → "CC").</li> </ul> <p><u>D :</u></p> <ul style="list-style-type: none"> <li>– En segmentation orthographique, les formes dissyllabiques /CVCV-CVCV/ sont préférées par les novices aux formes monosyllabiques /CV-CVCVCV/.</li> </ul> <p><u>I.2.3. Syllabique-Q symbolique : groupe /pt/</u></p> <ul style="list-style-type: none"> <li>– Taux d'épenthèses important pour le groupe /CC/ dans cette tâche.</li> <li>– Taux d'épenthèses plus élevé dans la tâche de comptage avec support symbolique que dans celle avec support orthographique.</li> <li>– Taux d'épenthèses plus élevé chez les novices que chez les experts.</li> <li>– Insertion vocalique après consonne finale dans une proportion non négligeable.</li> </ul>
--	--

<p><u>Tâches II. 1. :</u> <u>Discrimination visuelle</u></p>	<p><u>Segmentale : couple "b, v"</u></p> <ul style="list-style-type: none"> <li>- Peu de confusion phonographémique segmentale chez les novices dans une telle tâche.</li> <li>- Confusions de type hypercorrectif nombreuses chez les novices.</li> </ul>
<p><u>Tâches II.2. :</u> <u>Appariement visuel → visuel</u></p>	<p><u>II.2.1. Syllabique-Q symbolique : groupe "pl"</u></p> <ul style="list-style-type: none"> <li>- Taux d'épenthèse élevé chez les novices et non négligeable chez les experts, dans une telle tâche.</li> </ul> <p><u>II.2.2. Syllabique-D orthographique : groupe "pt"</u></p> <ul style="list-style-type: none"> <li>- Dans une telle tâche, les sujets préfèrent les formes simples (VC-CV) aux formes complexes (V-CCV, CCV-V).</li> </ul>
<p><u>Tâches III. :</u> <u>Appariement audiovisuel → visuel</u></p>	<p><u>III.1. Syllabique-Q/-D orthographique : couples {tl, tol} et {fR, fuR}</u></p> <p><u>Q :</u></p> <ul style="list-style-type: none"> <li>- En présentation bimodale (audio/orthographe), le taux d'épenthèses est globalement élevé pour certains groupes.</li> <li>- Ce taux varie considérablement selon le groupe (nature, position, environnement).</li> </ul> <p><u>D :</u></p> <ul style="list-style-type: none"> <li>- Dans une telle tâche, les codas sont évitées.</li> <li>- Tandis qu'en présentation orthographique les formes simples sont préférées, en présentation bimodale, les formes complexes sont préférées pour les experts et équivalentes pour les novices.</li> </ul>

	<p><u>III.2. Syllabique-Q symbolique : groupe {kR}</u></p> <p>– Dans une telle tâche, en présentation bimodale, le taux d'épenthèse n'est pas très élevé et similaire chez les novices et les experts.</p>
--	--

De ces résultats, émergent quatre conclusions provisoires :

- Hypercorrection et variation : le nombre élevé de confusions de type hypercorrectif (discrimination erronée de stimuli identiques) peut s'expliquer de plusieurs manières :
  - 1) Cause externe - la qualité de l'enregistrement : il est possible que la qualité acoustique des stimuli ait conduit à de telles erreurs.
  - 2) Cause interne - éléments discriminants : comme nous le détaillons plus bas à propos de la longueur vocalique, il est possible que des stimuli identiques pour un locuteur français ne le soient pas pour un locuteur japonais, de par la variation dans les stimuli de traits pertinents en japonais (longueur vocalique, accent) et non en français, d'où ces résultats.
  - 3) Cause interne - la variation en japonais : comme nous l'avons vu dans la partie phonologique descriptive du système japonais, l'inventaire phonémique du japonais se caractérise par sa grande variation phonétique (Labrune, 2001a). La richesse du répertoire allophonique en japonais pourrait donc conduire les sujets à effectuer une discrimination phonétique fine, qui n'était pas attendue dans le protocole des prétests.
- Effets de l'apprentissage : globalement, on peut affirmer que, d'après ces résultats, les experts ont commis moins d'erreurs que les novices, indiquant ainsi l'existence d'un apprentissage et de son application en contexte.

- Modalité et discrimination :

<b>Taux de confusion</b>			
<b>Couple b/v</b>	<b>Globalement</b>	<b>Novices</b>	<b>Experts</b>
<b>Modalité auditive</b>	68%	50%	75%
<b>Modalité visuelle (orthographique)</b>	9%	17%	6%

D'après ces résultats, il semble que la *modalité auditive* soit plus propice aux confusions segmentales que la modalité visuelle (orthographique).



- Modalité et appariement :

			Taux d'épenthèses		
Modalité Stimulus	Modalité rendu	Groupe	Globalement	Novices	Experts
Auditive	orthographique	/pt/	32%	50%	25%
	symbolique	/pt/	59%	83%	50%
Visuelle (Orthographique)	symbolique	/pl/	37%	67%	25%
Audiovisuelle (audio-orthographique)	orthographique	/tl/	45, 5%	50%	44%
		/fR/	4, 5%	0%	6%
	symbolique	/kR/	36%	34%	38%

D'après ces résultats, il semblerait que :

- 1) Pour un rendu symbolique (phonologiquement le plus fidèle, sans interférence orthographique) :
  - Globalement : la modalité auditive (59%) est plus propice aux épenthèses que la modalité visuelle (37%) et audiovisuelle (36%). Cela est vrai pour les novices.

- Pour les experts : la modalité auditive (50%) est suivie par la modalité audiovisuelle (38%) et la modalité visuelle (25%).

2) Pour un rendu orthographique (avec interférence orthographique) :

- Globalement : la modalité audiovisuelle (45, 5%) est plus propice aux épenthèses que la modalité auditive (32%) pour certains groupes, et le contraire est vrai pour d'autres. Cela est vrai pour les experts.
- Pour les novices : les deux modalités ne se distinguent pas (50%) pour certains groupes, tandis que la modalité auditive est plus propice aux épenthèses que la modalité audiovisuelle pour d'autres.

Cependant, ces résultats ne peuvent être utilisés que si l'on considère grossièrement les groupes consonantiques /CC/, ce qui est largement insuffisant. Seule une comparaison portant sur les mêmes groupes consonantiques peut être instructive, et ces résultats n'ont donc qu'une valeur illustrative.

## **V. Aspects méthodologiques et perspectives**

### Durée et faisabilité

Concernant la durée de passation des prétests, il nous a fallu environ 20 minutes pour l'ensemble de la procédure. Si nous avions voulu effectuer les mêmes tests en incluant cette fois-ci l'ensemble des variables indépendantes que nous souhaitions manipuler, alors les calculs, minimaux, auraient été les suivants :

- 1) (Audit-Discr-Segm) \*4 (position) \* 2 (fréquence) \*4 (couples cibles)
- 2) (Audit-Discr-Syll)\*3 (position) \*6 (couples cibles)
- 3) (Audit-Discr-Métri-Q)\*4 (position) \* 2 (fréquence) \*48 environ (CC)
- 4) (Audit-Trans-Seg) \*4 (positions) \* 2 (fréquence)\*2 (couples cibles)
- 5) (Audit-Trans-Métri-Q/-D) \* 4 (positions) \* 3 (fréquence) \* 48 (CC)

- 6) (Audit-Trans-Métr-Q-Symb) \*3 (positions) \* 48 (CC)
- 7) (Vis-Discri-Segm) \*2 (positions) \* 2 (fréquence) \*2 (couples cibles)
- 8) (Vis-Trans-Métr-Q-Symb) \* 3 (position) \* 2 (fréquence) \* 48 (CC)
- 9) (Vis-Trans-Métr-D) \* 3 (position) \* 2 (fréquence) \* 48 (CC)
- 10) (AVSimul-Trans-Métr-Q/-D) \* 2 (positions) \* 3 (fréquence) \* 48 (CC)
- 11) (AVSimul-Trans-Métr-Q-Symb) \* 3 (positions) \* 2 (fréquence) \* 48 (CC)

Cela aurait été beaucoup trop coûteux en temps. Par ailleurs, ces prétests nous ont permis d'évaluer, en partie au moins, la rentabilité informative des tâches proposées : les épreuves de segmentation graphique, en particulier, ainsi que les épreuves avec stimuli initiaux orthographiques semblent relativement peu productives mais coûteuses temporellement. Du point de vue de la faisabilité, comme prévu, les tâches impliquant des stimuli visuels sont plus délicates à mener. Cependant, les tâches audiovisuelles nous ont semblé satisfaisantes, et il nous a donc semblé important de les maintenir, puisqu'elles sont au cœur de notre étude.

#### Problèmes et remarques

Concernant les problèmes méthodologiques, cette passation nous a permis d'émettre diverses remarques, générales ou précises. Globalement, certains de ces prétests présentent trop de variables incontrôlées, conduisant à une absence de rendement informatif. Ils nous laissent cependant penser :

- Perception/ Production : qu'il pourrait y avoir chez certains sujets une dissociation entre perception (les sujets sont conscients des distinctions et les perçoivent) et production (en production, ils commettent cependant des « erreurs », aléatoires ou systématiques). Cette dissociation correspond à ce que certains appellent l'effet « fis »<sup>166</sup>.

---

<sup>166</sup> Cet « effet » renvoie au cas de cet enfant anglophone, incapable de produire correctement le mot anglais « fish » (= poisson) auquel il substitue la séquence « fis », mais parfaitement capable de percevoir la distinction entre les deux lorsque un autre locuteurs que lui prononce ces deux items.

- Non-mots/ Discours : qu'il y aurait une différence entre une tâche purement phonologique (non-mots) de discrimination ou d'appariement (l'attention est alors pleinement concentrée sur la tâche) et une situation de discours (partage attentionnel entre les différentes tâches à accomplir). D'où une meilleure performance de nos sujets lors des prétests qu'en situation d'observation informelle. Il nous semble ainsi qu'une tâche ne portant que sur des non-mots isolés est plus « facile » à réaliser qu'une tâche portant sur des non-mots dans un environnement phrastique (qui implique les aspects suprasegmentaux, ainsi que des questions de frontières morphophonémiques). Dans nos prétests, l'influence des facteurs suprasegmentaux (en particulier intonatifs) a été totalement négligée, puisque nous avons adopté la même intonation pour tous les stimuli auditifs.
- Ordre de présentation : que lorsque l'on propose un choix de segmentation graphique, par exemple, la direction de présentation des réponses (verticale ou horizontale) peut influencer sur le choix des sujets, probablement par un effet d'ordre.

Plusieurs problèmes spécifiques sont également apparus :

- Traduction des consignes : la consigne relative à la segmentation graphique s'est révélée délicate lors de la traduction en japonais, selon le choix entre les deux instructions suivantes : « le découpage le plus naturel » ? ou « le découpage le plus naturel quand vous lisez ? ». Cette différence pouvait considérablement modifier la réponse.
- Savoirs déclaratifs et consignes : pour la tâche d'appariement auditif/visuel segmental (I.2.1. : /upaki/"upaki") par exemple, une collègue<sup>167</sup> nous a fait remarquer que la plupart des étudiants de français, à partir d'un niveau faux-débutant, disposaient d'une certaine connaissance des règles orthoépiques françaises, et qu'ils allaient associer d'une part la lettre "u" au son /y/ et d'autre

---

<sup>167</sup> Nous remercions K. K. pour sa remarque pertinente.

part le digraphe "ou" au phonème /u/. Il ne fallait alors pas s'attendre à ce qu'ils associent la lettre "u" au phonème /u/ ou à son équivalent japonais, contrairement à nos prévisions. Dans cette tâche, il s'agissait d'appariement entre stimuli 1 auditifs et stimuli 2 visuels, avec pour consigne de choisir la transcription la mieux adaptée. Il est néanmoins raisonnable de supposer que ce problème figure en effet dans le protocole, impliquant une réponse identique pour deux stimuli différents, et cela même si les stimuli sont perçus différemment (par exemple, /ø/ ou /œ/ perçus comme /u/). Car, pour que les étudiants associent la lettre "u" au phonème /y/, encore faut-il qu'ils soient parvenus à le percevoir, alors que l'acquisition de ce phonème constitue justement une difficulté importante dans l'apprentissage du français, puisqu'il n'existe pas de phonème correspondant en japonais, et qu'il est souvent perçu/reproduit /u/ par les apprenants. Il pourrait donc sembler nécessaire de rajouter une consigne supplémentaire. Par exemple : « Seul le test II.2. utilise les règles d'orthographe française complexes. Pour tous les autres, chaque lettre correspond à un seul son. En particulier, le son [u], qu'on écrit souvent "ou" en français et qui ressemble au "u" en romaji, est ici transcrit avec une seule lettre : la lettre "u". Il en est de même pour la lettre "e" ou "E", qui se prononce comme en romaji [e], quelle que soit sa position dans le mot (début, milieu, fin) ». Mais finalement nous nous sommes abstenus, car cela aurait compliqué les consignes, qui présentaient déjà certaines difficultés pour les non-initiés. L'enseignant qui a passé les prétests nous a en effet fait part de son incompréhension partielle face à certaines consignes. Il y a donc eu nécessité de simplification ou d'éclaircissement des consignes. En ce qui concerne l'épreuve d'appariement, elle nous permet d'observer, au niveau collectif, les comportements existants et elle ne perd donc pas de sa validité. Il conviendrait de la compléter en y associant une tâche de discrimination afin de vérifier l'identité des phonèmes perçus associés aux signes graphiques.

- Variation inter-tâche et dissociation : dans l'une des réponses, pour la tâche I.2.2. (choix de transcription auditif/visuel), un étudiant débutant choisit "ptanabo" pour

la transcription, mais ensuite "puta-nabo" pour le découpage. Cela nous semble refléter une variation intertâche, la deuxième faisant particulièrement appel à la conscience (méta)-phonologique du sujet. Nous retrouvons un cas de dissociation similaire pour un étudiant avancé qui choisit "putanabo" en tâche d'appariement et "pta-nabo" en tâche de segmentation. Cette dissociation nous paraît particulièrement instructive.

- Non-mots, mots et phonotactique : en ce qui concerne la tâche de discrimination auditive (I.1.), pour les étudiants avancés, tous semblent faire la distinction entre /siliko/ et /shiliko/ et nous supposons ainsi qu'ils distinguent /si/ et /shi/. Mais il serait utile de noter s'ils distinguent /sinema/ de /shinema/ ou non, en perception et/ou en production, le mot « cinema » faisant partie des emprunts japonais (strate lexicale étrangère). En effet, il y aurait alors une opposition entre non-mot et mot, et l'on pourrait alors opposer un cas de phonologie lexicale (« addressed phonology ») et l'autre de phonologie prélexicale (« assembled phonology »). Nous avons en effet à plusieurs reprises constaté dans nos classes l'incapacité, relative mais très marquée chez certains, des étudiants à répéter le mot « cinéma » correctement, substituant /shi/ à /si/, alors qu'ils étaient tout à fait capable de produire le non-mot « sanéma », ou encore le mot « assis ». Considérons l'exemple suivant, authentique, en tâche de répétition corrective :

Monsieur → Monshieur

Sa → Sa

SsssssI → SsssssI

Sieu → Shieu

Il semblerait que des contraintes d'ordre phonotactique, et non phonématique, soient ici à l'œuvre, la phonotactique étant déterminée par la structure syllabique du système en question.

- Facteurs de discrimination et transcriptions ouvertes : en tâche de discrimination, il pourrait être utile de demander aux étudiants de transcrire ce qu'ils pensent

entendre, en katakana et/ou en romaji. En effet, dans l'une des tâches de discrimination, l'un des sujets a indiqué « différent », mais il a également indiqué dans la marge une transcription en hiragana, et l'on s'aperçoit alors qu'il a perçu une différence, non de mode ou de lieu d'articulation pour le segment consonantique cible (b/v), mais une différence de longueur pour la dernière voyelle du non-mot (la longueur étant un trait non-pertinent en français, mais pertinent en japonais). Le résultat semblerait donc partiellement faussé, même si cette différence de longueur pourrait précisément résulter de la différence consonantique. Dans ce cas, la différence perçue, même si elle n'est pas directement issue de la substitution consonantique, lui serait cependant tout à fait attribuable. La feuille de résultats de cet étudiant est très instructive : pour les stimuli n°6 (/siliko/, /shiliko/), l'étudiant a noté en hiragana : /shiriko/, /shiiriko/ (sensibilité à la longueur (syllabes bimoraïques, lourdes)). Superficiellement, on aurait donc localement ici : SI → SHI et SHI → SHII. Cette observation nous invite fortement à coupler, à la tâche de discrimination (par exemple), une tâche de transcription en hiragana ou en romaji, avec la possibilité de signaler l'incapacité à transcrire fidèlement les stimuli (ex : « SI »). Pour cet étudiant, voici les transcriptions données par lui en romaji, pour les stimuli perçus comme étant différents :

- [baliko/valiko] → « barigo »/ « barigou » : on observe une distinction de longueur et l'on note ici qu'il est impossible de vérifier la perception du [b] et du [v] puisque les hiragana ne permettent pas d'effectuer graphiquement la distinction.
- [psaliko/pusaliko] → « kusariko »/ « pusariko » : on note ici l'interprétation du [psa] en « kusa ». Il faudrait ici évaluer les fréquences d'occurrence des combinaisons [pusa] et [kusa], en s'interrogeant sur le dévoisement de la voyelle neutre en japonais. Si ce dévoisement est beaucoup plus fréquent après [k] qu'après [p], on peut envisager ici une influence de la fréquence d'occurrence syllabique sur la tâche de perception.
- [pusogeda/pusogeda] → « kusogeda »/ « pusogeda » : ici le [pus] a été perçu comme le [ps] précédent, à savoir comme [kus]. En revanche, le second [pus]

est toujours perçu correctement. Cela provient-il d'un manque d'homogénéité des stimuli, ou bien d'un phénomène d'ordre, qui permet de mieux percevoir le second stimulus grâce à la référence que constitue le premier, ou encore d'un simple effet de récence ? Ces tests ne nous permettent pas de répondre à la question.

- [siliko/ shiliko] → « shiriko »/ « shiiriko » : voir plus haut.

Pour les tests généraux subséquents, nous avons donc dû :

- Opérer une sélection des tâches ;
- Opérer une sélection des stimuli ;
- Rectifier les imperfections méthodologiques ;
- Vérifier que les tests permettaient d'éclairer nos hypothèses ;
- Redéterminer les durées de passation.





## **Partie Expérimentale II :**

### **Tests - Protocole**

Une fois nos prétests effectués et analysés, nous avons revu notre protocole initial, afin de pouvoir effectuer les nouveaux tests sur une plus large population de sujets<sup>168</sup>.

Parmi les caractéristiques de ces tests, nous pouvons mentionner :

- L'importance quantitative : un total initial de 182 sujets, séparés en deux groupes de tests, 120 sujets d'une part et 62 d'autre part.
  
- La validité écologique des conditions de passation : puisqu'il s'agissait de tests psycholinguistiques à pendant didactique, portant sur l'apprentissage, il nous semblait essentiel de tester la perception (et *a fortiori* l'apprentissage) des sujets dans les conditions où ceux-ci perçoivent/apprennent *effectivement*, c'est-à-dire en salle de classe, avec leurs pairs, dans un contexte d'apprentissage qui leur soit *familier*.
  
- Le ciblage de la population étudiée : les étudiants universitaires japonais au Japon, avec certaines variables (degré d'expertise linguistique en français par exemple) permettant de balayer un certain pan de cette population.

---

<sup>168</sup> Nous tenons à très vivement remercier ici les deux collègues qui nous ont permis de faire passer les tests aux étudiants en question.

## **Table des Matières Partielle**

I. METHODOLOGIE .....	- 695 -
II. DESCRIPTIF GENERAL DE LA POPULATION .....	- 696 -
III. DESCRIPTIF GENERAL DES PROTOCOLES .....	- 704 -
IV. PREMIERE PARTIE : AXE SEGMENTAL (GROUPE 1) .....	- 710 -
IV.1. <i>Test A (Identification)</i> .....	- 710 -
IV.2. <i>Test F (Identification et Discrimination)</i> .....	- 716 -
V. DEUXIEME PARTIE : AXE SYLLABIQUE (GROUPES 1 ET 2).....	- 721 -
V.1. <i>Test C1 (Groupe 1 – Unimodalité Auditive)</i> .....	- 721 -
V.2. <i>Test C2 (Groupe 2 – Unimodalité Auditive)</i> .....	- 728 -
V.3. <i>Test D (Groupe 2 – Unimodalité Visuelle)</i> .....	- 732 -
V.4. <i>Test E (Groupe 2 – Bimodalité Audio-Visuelle)</i> .....	- 735 -
V.5. <i>Méta-test C2-D-E (Groupe 2 – Effet de la Modalité)</i> .....	- 743 -
VI. VARIABLES ET CONSIDERATIONS SUPPLEMENTAIRES .....	- 744 -

## **I. Méthodologie**

Tous les tests ont été passés par la population qui constituait la cible majeure de notre étude, à savoir celle des étudiants universitaires japonais au Japon. Avant de préciser ci-dessous les caractéristiques de ces sujets, rappelons (Loonis, 2002) la définition de quelques notions importantes de psychologie expérimentale, nécessaires à la compréhension de notre protocole :

- Variables Indépendantes (VI) : ce sont les variables manipulées par le chercheur. Dans notre cas, ces variables apparaissent surtout dans la constitution des stimuli ainsi que dans les modalités de présentation desdits stimuli.
  
- Variables Dépendantes (VD) : ce sont les réponses des sujets aux tests. Dans notre cas, il s'agissait essentiellement de choix de réponses et d'erreurs.
  
- Variables Parasites : comme les VI, elles peuvent influencer les VD. Parmi celles-ci figurent les :
  - Variables d'ordre descriptif : démographiques (lieu de résidence, situation familiale, catégorie socioprofessionnelle d'appartenance, niveau d'études, etc.) ; organismiques (sexe, âge, etc.) ; historiques (événements significatifs de la vie, conditions d'éducation, etc.).
  
  - Variables de dispositif : celles qui, tout en étant parasites, relèvent du dispositif expérimental (expérimentateur, état des sujets, conditions matérielles, durée de passation, etc.). Dans notre cas, étant donné l'approche écologique adoptée, celles-ci sont assez nombreuses, mais sont intégrées à l'interprétation.

Les variables prises en compte dans notre étude ont été codées numériquement et renommées afin d'assurer le stockage informatique des données dans des classeurs Excel.

Enfin, (au moins) trois types d'hypothèses sont convoqués dans une telle étude :

- Hypothèses Générales : correspond au thème général de la recherche
- Hypothèses De travail : correspond à l'influence générale des VI sur les VD.
- Hypothèses Opérationnelles : correspond à l'influence précise des VI sur les VD.

Concernant notre plan de recherche, il s'agit plutôt d'un plan à mesures répétées, c'est-à-dire qu'un seul et même groupe a été soumis à toutes les valeurs des VI. Dans ce cas, le problème d'équivalence des groupes ne se pose pas, même si, dans une perspective plus « expérimentale », il aurait fallu exclure les étudiants les meilleurs ainsi que les moins bons de manière à homogénéiser le groupe.

Enfin, puisque nous nous intéressons à l'apprentissage du système phonéico-phonologique du français par des étudiants japonais, les stimuli utilisés étaient des non-mots, afin de permettre aux étudiants, autant que possible (car l'on ne peut éviter les effets de cohorte, d'homophonie et de ressemblance phonologique), de se concentrer sur la forme phonéico-phonologique (sublexicale) des stimuli (présentés aux étudiants comme des mots français) et d'échapper ainsi, en partie au moins, à un partage attentionnel entre forme et sens, d'une part, et à certains effets lexicaux (fréquence lexicale, etc.) d'autre part.

## **II. Descriptif général de la population**

La population étudiée se compose d'étudiants universitaires japonais, inscrits dans des universités de Tokyo ou de sa banlieue. Les tests ont eu lieu durant leurs cours de français habituels, afin de préserver la dimension écologique des tests et de ne pas avoir de recrutement à effectuer. Tous les étudiants des classes concernées, sans

exclusion, ont passé les tests, après avoir rempli un questionnaire d'ordre essentiellement sociolinguistique. Les motifs justifiant une telle ouverture sont d'ordre écologique et didactique : étant donné la perspective pédagogique que nous avons assignée à nos tests psycholinguistiques, nous avons délibérément choisi de saisir les classes d'étudiants « tel quel », afin de nous confronter, comme y sont obligés les enseignants, au réel des classes de langues, avec l'ensemble des variables incontrôlées qu'elles impliquent, au niveau humain en particulier. Les résultats étant d'ordre statistique, une telle démarche nous a semblé acceptable. En ce qui concerne la langue « maternelle » de nos sujets (ou sa variante sociolinguistique), nous avons supposé que tous les étudiants, même s'ils n'étaient pas originaires de la région du Kanto, avaient une certaine maîtrise du « japonais standard » tel que pratiqué dans la région de Tokyo, étant donné leur appartenance générationnelle et leur degré d'éducation. Ici encore nous avons privilégié l'inclusion dans la sélection des sujets, à l'instar des enseignants qui ne peuvent refuser les étudiants bilingues ou ceux de régions éloignées. Les données numériques relatives aux variables descriptives sont données plus bas. Il s'agissait d'étudiants de la première jusqu'à la quatrième année d'études (A) à l'université, étudiant le français soit comme première, soit comme seconde langue étrangère (L) à l'université. Ces étudiants se répartissent en deux grands groupes<sup>169</sup>, chaque groupe ayant passé trois tests :

Groupe 1 : 1L-3A : 13 sujets ; 2L-1A : 76 sujets ; 2L-2A : 10 sujets ; 2L-3A : 21 sujets.

Total : 120 étudiants

Groupe 2 : 1L-1A : 20 sujets ; 1L-3A : 9 sujets ; 1L-4A : 7 sujets ; 2L-1A : 21 sujets ; 2L-2A : 5 sujets.

Total : 62 étudiants

---

<sup>169</sup> Il s'agit ici des effectifs initiaux : par la suite, les réponses des sujets n'ayant pas respecté les consignes n'ont pas été prises en compte dans l'analyse (cf. infra. « Problèmes méthodologiques »). Il ne s'agit néanmoins que de quelques rares sujets (entre un et quatre selon les tâches).

Chacun de ces groupes correspond à plusieurs classes, dans lesquelles les sujets étaient répartis de la manière suivante :

<b>Sujets répartis par classe</b>	
<b>Groupe 1 (de 1 à 120)</b>	1-3, 4-31, 32-64, 65-85, 86-96, 97-120
<b>Groupe 2 (de 1 à 62)</b>	1-21, 22-26, 27-46, 47-55, 56-62

Les variables descriptives sont de trois types :

- Quantitative (Qt) (continue, numérique)
- Qualitative nominale (Qln) (catégorielle non-ordonnable)
- Qualitative ordinale (Qlo) (catégorielle ordonnable)

<b>Variable Descriptive</b>	<b>Type</b>	<b>Nom Excel</b>	<b>Code</b>
Sexe	Qln	sexe	1=homme, 2=femme
Age	Qt	age	1=17, 2=18, 3=19, 4=20, 5=21, 6=22, 7=23, 8=24, 9=25, 10=+ de 25 ans
Langue maternelle	Qln	langmate	1=japonais, 2=autre
Niveau d'études universitaires	Qlo	nivetuni	1=1 <sup>ère</sup> année, 2=2 <sup>ème</sup> année, 3=3 <sup>ème</sup> année, 4=4 <sup>ème</sup> année
Domaine de spécialité « français »	Qln	specifra	1=oui, 2=non
Durée d'étude du français	Qlo	duretfra	1=0-6 mois, 2=6 mois-1 an, 3=1-2 ans, 4=2-4 ans, 5=+ de 4 ans
Etude du français au lycée	Qln	etfralyc	1=oui, 2= non
Déjà eu des classes de prononciation	Qln	claspron	1= oui, 2=non
Langue étrangère majeure	Qln	langmaje	1= français, 2 = langue européenne/romane, 3=langue non-européenne

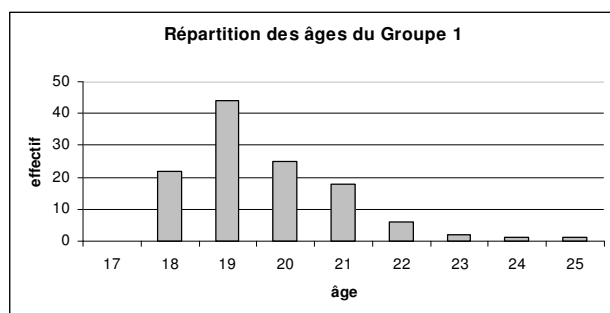
Bilingue	Qln	biling	1=oui langue européenne, 2= oui autre langue, 3=non
Parent bilingue	Qln	parbilin	1=oui langue européenne, 2= oui autre langue, 3=non
Déjà vécu à l'étranger 6 mois	Qln	vecetran	1=oui francophone, 2=oui pays de langue européenne non francophone, 3=oui autre pays, 4=non
Main dominante	Qln	maindomi	1=droitier, 2=gaucher
Actuellement cours de français supplémentaires	Qlo	courfsup	1=oui, 2=un peu, 3=non

### Répartition et Taux :

#### *Groupe 1*

	<b>Masculin</b>	<b>Féminin</b>
<b>Sexe</b>	49, 58%	50, 42%

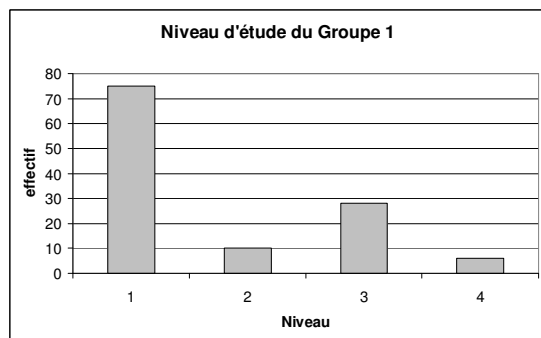
<b>Age</b>	<b>17</b>	<b>18</b>	<b>19</b>	<b>20</b>	<b>21</b>	<b>22</b>	<b>23</b>	<b>24</b>	<b>25</b>	<b>Plus</b>
<b>Répartition</b>	0	18,5%	37,%	21%	15,1%	5%	1,7%	0,8%	0	0,8%



	<b>Japonais</b>	<b>Autre</b>
<b>Langue maternelle</b>	98, 31%	1, 69%

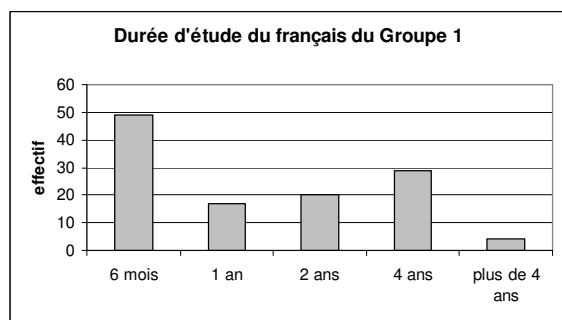


Niveau d'études	1 <sup>ère</sup> année	2 <sup>ème</sup> année	3 <sup>ème</sup> année	4 <sup>ème</sup> année
Répartition	63,02%	8,4%	23,52%	5,04%



	Français	Autre
Domaine de spécialité	73,1%	26,9%

Durée d'étude du français	Moins de 6 mois	Moins de 1 an	Moins de 2 ans	Moins de 4 ans	Plus de 4 ans
Répartition	41,17%	14,28%	16,8%	24,36%	3,36%



Etudiants ayant étudié le français au lycée	12,6%
---	-------

Etudiants ayant déjà eu au moins une classe de prononciation	73,94%
--	--------

<b>Langue étrangère majeure</b>	<b>Français</b>	<b>Européenne</b>	<b>Autre</b>
<b>Répartition</b>	10, 92%	87, 39%	1, 68%

<b>Bilingue</b>	<b>Langue européenne</b>	<b>Autre langue</b>	<b>Non</b>
<b>Répartition</b>	3, 36%	0, 84%	95, 79%

<b>Parent bilingue</b>	<b>Langue européenne</b>	<b>Autre langue</b>	<b>Non</b>
<b>Répartition</b>	6, 72%	1, 68%	91, 59%

<b>Etudiants ayant déjà vécu à l'étranger plus de 6 mois</b>	<b>En France</b>	<b>En Europe</b>	<b>Autre</b>	<b>Non</b>
<b>Répartition</b>	0, 84%	15, 12%	3, 36%	80, 67%

	<b>Gauche</b>	<b>Droite</b>
<b>Main dominante</b>	3, 37%	96, 63%

<b>Suivent actuellement des cours de français extra-universitaires</b>	<b>Oui</b>	<b>Un peu</b>	<b>Non</b>
<b>Répartition</b>	1, 68%	4, 2%	94, 11%

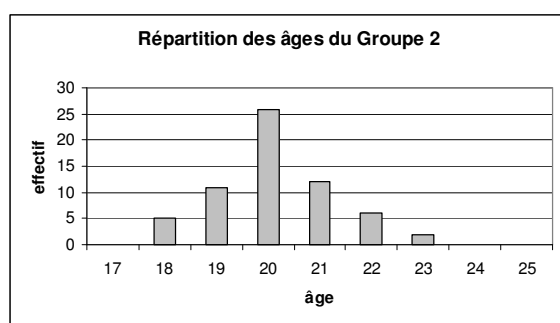
Pour le groupe 1, au vu de ces données, il serait envisageable de prendre ultérieurement en compte deux variables descriptives dans notre analyse :

- Le sexe : puisque nous avons une répartition à peu près équivalente de sujets féminins et masculins ;
- L'expertise en français, via la durée d'étude, en considérant deux sous-groupes : ceux ayant étudié le français moins de 6 mois d'une part, et les autres d'autre part.

**Groupe 2**

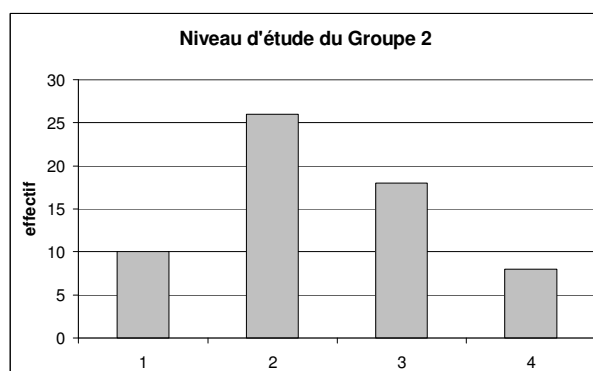
	<b>Masculin</b>	<b>Féminin</b>
<b>Sexe</b>	12, 91%	87, 09%

<b>Age</b>	<b>17</b>	<b>18</b>	<b>19</b>	<b>20</b>	<b>21</b>	<b>22</b>	<b>23</b>	<b>24</b>	<b>25</b>	<b>Plus</b>
<b>Répartition</b>	0	8,1%	17,7%	41,9%	19,3%	9,7%	3,2%	0	0	0



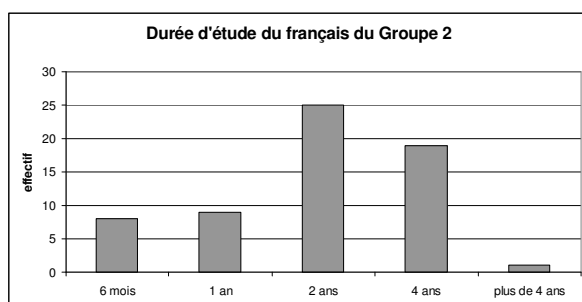
	<b>Japonais</b>	<b>Autre</b>
<b>Langue maternelle</b>	98, 38%	1, 62%

<b>Niveau d'études</b>	<b>1<sup>ère</sup> année</b>	<b>2<sup>ème</sup> année</b>	<b>3<sup>ème</sup> année</b>	<b>4<sup>ème</sup> année</b>
<b>Répartition</b>	16, 12%	41, 93%	29, 03%	12, 9%



	<b>Français</b>	<b>Autre</b>
<b>Domaine de spécialité</b>	48, 38%	51, 62%

<b>Durée d'étude du français</b>	<b>Moins de 6 mois</b>	<b>Moins de 1 an</b>	<b>Moins de 2 ans</b>	<b>Moins de 4 ans</b>	<b>Plus de 4 ans</b>
<b>Répartition</b>	12, 9%	14, 51%	40, 32%	30, 64%	1, 61%



<b>Etudiants ayant étudié le français au lycée</b>	6, 45%
--	--------

<b>Etudiants ayant déjà eu au moins une classe de prononciation</b>	69, 35%
---	---------

<b>Langue étrangère majeure</b>	<b>Français</b>	<b>Européenne</b>	<b>Autre</b>
<b>Répartition</b>	51, 61%	48, 38%	0

<b>Bilingue</b>	<b>Langue européenne</b>	<b>Autre langue</b>	<b>Non</b>
<b>Répartition</b>	3, 22%	0	96, 78%

<b>Parent bilingue</b>	<b>Langue européenne</b>	<b>Autre langue</b>	<b>Non</b>
<b>Répartition</b>	9, 67%	4, 83%	85, 48%

<b>Etudiants ayant déjà vécu à l'étranger plus de 6 mois</b>	<b>En France</b>	<b>En Europe</b>	<b>Autre</b>	<b>Non</b>
<b>Répartition</b>	3, 22%	3, 22%	3, 22%	90, 32%

	<b>Gauche</b>	<b>Droite</b>
<b>Main dominante</b>	4, 84%	95, 16%

<b>Suivent actuellement des cours de français extra-universitaires</b>	<b>Oui</b>	<b>Un peu</b>	<b>Non</b>
<b>Répartition</b>	4, 83%	1, 61%	93, 54%

Pour le groupe 2, au vu de ces données, il serait envisageable de prendre ultérieurement en compte une variable descriptive dans notre analyse :

- L'expertise en français, via le « domaine de spécialité », puisque les proportions de « spécialistes » et de « non-spécialistes » sont à peu près équivalente, une équivalence que l'on retrouve par ailleurs pour la variable « langue étrangère majeure ».

### **III. Descriptif général des protocoles**

Les tests ont été passés dans le cadre des cours de français habituels des étudiants, et ont été présentés comme des activités d'étude / d'apprentissage de la prononciation française. Les conditions de passation leur étaient donc tout à fait familières et les tests ont été passés l'un à la suite de l'autre afin de pallier, notamment, le problème d'absentéisme. Une telle démarche nécessite alors évidemment de prendre en compte dans l'analyse l'effet d'ordre (relatif à l'ordre de passation des tests), l'effet de la fatigue, lié à celui de la démotivation, et enfin l'effet de l'habitude (relatif à l'habitude aux tests). Tout en tenant compte de ces variables parasites, nous défendons de nouveau ce choix par l'optique didactique qui est la nôtre : les activités

d'apprentissage proposées aux étudiants dans leurs conditions normales d'enseignement s'enchaînent bel et bien, un effet dont nous tolérons donc la réplique dans nos expérimentations. Cette approche et ce milieu écologiques, s'ils confèrent une certaine faiblesse à la validité interne de nos tests, leur assurent ainsi en revanche une forte validité externe.

Les tests, qui sont détaillés plus bas, consistaient essentiellement en tâches d'identification, de discrimination, d'appariement et de segmentation de stimuli (pseudo-) linguistiques, en présentation auditive, visuelle ou audiovisuelle, envisagées comme tâches de perception et *a fortiori* comme tâches d'apprentissage. Voici ci-dessous un premier tableau de présentation simplifiée, permettant d'avoir une vision globale des tests effectués :

<b>Première Partie : Axe Segmental</b>	<b>Deuxième partie : Axe Syllabique</b>
<p><b><u>Sujets</u> :</b>  <u>Groupe 1</u> : 120</p> <p><b><u>Tests</u> :</b></p> <p><u>Test A) Appariement audiovisuel</u>: Les sujets entendent un non-mot et doivent choisir une transcription orthographique, parmi deux qui leur sont proposées.</p> <p><u>Test F) Appariement audiovisuel et discrimination auditive</u> : Les sujets entendent deux non-mots l'un à la suite</p>	<p><b><u>Sujets</u> :</b>  <u>Groupe 1</u> : 120  <u>Groupe 2</u> : 62</p> <p><b><u>Tests</u> :</b></p> <p><u>1) Groupe 1 :</u>  <u>Test C1) Comptage auditif</u> : Les sujets entendent un non-mot et doivent indiquer le nombre de syllabes entendues, à choisir entre trois propositions.</p> <p><u>2) Groupe 2 :</u>  <u>Test C2) Comptage auditif</u> : Les sujets entendent un non-mot et doivent indiquer le nombre de syllabes entendues, à choisir entre trois</p>

<p>de l'autre. Ils doivent indiquer s'ils sont différents ou identiques en choisissant une transcription pour chaque mot, entre deux qui leur sont proposées.</p>	<p>propositions.</p> <p><u>Test E) Comptage audiovisuel</u> : Les sujets entendent un non-mot et en même temps lisent ce non-mot écrit. Parfois le mot auditif et sa transcription sont identiques, parfois ils sont différents. Les sujets doivent indiquer le nombre de syllabes entendues, à choisir entre trois propositions.</p> <p><u>Test D) Comptage visuel</u> : Les sujets lisent un non-mot écrit et doivent indiquer le nombre de syllabes lues, à choisir entre trois propositions.</p>
---	--

### **Stimuli**

Les stimuli auditifs ont tous été enregistrés dans une salle insonorisée (studio d'enregistrement) par un locuteur natif français, spécialiste de sciences du langage, à l'aide d'un Microphone Sony ECM-MS907 (120°) et d'un Net MD walkman Sony MZ-N1 sur Mini-disc Sony, puis copiés ensuite sur une cassette audio TDK. Cette cassette a ensuite été utilisée en classe à l'aide des magnétophones normalement utilisés pour ces classes.

Ces stimuli consistaient en non-mots, tri-, tétra- ou pentasyllabiques, composés de séquences et de segments légaux en français, présentés cependant aux étudiants comme des mots français. D'un point de vue prosodique, notre locuteur s'est efforcé de garder le même schéma prosodique (intonatif et accentuel) pour tous les stimuli, à savoir celui d'une affirmation « neutre » du type "J'aime le thé". Les stimuli étaient séparés les uns

des autres à intervalles de temps réguliers par un indicateur sonore ("Bip") produit par le locuteur.

**Instructions :**

Toutes les consignes ont été traduites en japonais. Au début de chaque feuillet figuraient les consignes générales suivantes :

**Pour tous les exercices, il faut IMPERATIVEMENT :**

- NE PAS REFLECHIR (tous les mots présentés sont inconnus de vous) et répondre SPONTANEMENT (il n'y a pas de règle à trouver, ni de bonne ou mauvaise réponse ; il faut juste choisir la réponse de manière naturelle)
- Répondre DANS TOUS LES CAS, même si vous n'êtes pas sûr,
- Répondre IMMEDIATEMENT, car vous n'avez que quelques secondes entre chaque épreuve ;
- Répondre SIMPLEMENT d'après CE QUE VOUS ENTENDEZ ET VOYEZ ;
- ENTOURER les réponses que vous choisissez ;
- Lire les documents donnés au fur et à mesure de l'exercice : ne TOURNEZ PAS les pages en avance.
- Bien RESPECTER les consignes.

Ces consignes ont été lues et présentées aux étudiants par leur enseignant. Au début de chaque test figuraient ensuite les consignes spécifiques à l'épreuve, avec un ou deux exemples afin d'assurer leur bonne compréhension. Ici encore, ces consignes et ces exemples ont été lus et présentés par les enseignants.



### **Conditions de passation**

<b>Groupe</b>	<b>Sujets</b>	<b>Conditions de passation</b>
<b>1 (120)</b>	<b>Tous</b>	<u>Lieu</u> : Salle grande et lumineuse. <u>Matériel</u> : Deux enceintes face aux étudiants. Volume magnétophone normal.
	<b>1→3</b> (1A-2L)	<u>Date</u> : 27.10.2003. Début d'après-midi. Après la pause repas. <u>Durée</u> : 13H05-14H45. <u>Matériel</u> : magnétophone Sony CFD-6.
	<b>4→31</b> (1A-2L)	<u>Date</u> : 27.10.2003. Milieu d'après-midi. Après une classe. <u>Durée</u> : 14H50-15H40. <u>Matériel</u> : magnétophone Sony CFD-6.
	<b>32→64</b> (3A-1L et 4A- 1L)	<u>Date</u> : 30.10.2003. Début d'après-midi. Après la pause repas. <u>Durée</u> : 15H-15H45. <u>Matériel</u> : magnétophone Sony TCM.
	<b>65-85</b> (1A-2L)	<u>Date</u> : 31.10.2003. Début d'après-midi. Après la pause repas. <u>Durée</u> : 13H30-14H10. <u>Matériel</u> : magnétophone Sony TCM.
	<b>86→120</b> (1A-2L et 2A- 2L)	<u>Date</u> : 01.11.2003. Fin de matinée. Après une classe. <u>Durée</u> : 11H20-12H15. <u>Matériel</u> : magnétophone Sony TCM.

<b>2 (62)</b>	<b>Tous</b>	<p><u>Lieu</u> : Salle grande et lumineuse.</p> <p><u>Matériel</u> : Deux enceintes face aux étudiants. Volume magnétophone normal.</p>
	<p><b>1 → 21</b> (1A-2L)</p>	<p><u>Date</u> : 05.11.2003. Début d'après-midi. Après la pause repas.</p> <p><u>Durée</u> : 13H-14H30.</p> <p><u>Matériel</u> : enceintes Voss, magnétophone Victor, rétroprojecteur Elmo.</p>
	<p><b>22 → 26</b> (2A-2L)</p>	<p><u>Date</u> : 05.11.2003. Milieu d'après-midi. Après une classe.</p> <p><u>Durée</u> : 14H40-16H10.</p> <p><u>Matériel</u> : enceintes Voss, magnétophone Victor, rétroprojecteur Elmo.</p>
	<p><b>27 → 46</b> (1A-1L)</p>	<p><u>Date</u> : 11.11.03. Milieu d'après-midi. Après une classe.</p> <p><u>Durée</u> : 15H00-16H10.</p> <p><u>Matériel</u> : enceintes Ramsa, magnétophone Panasonic, rétroprojecteur Panasonic.</p>
	<p><b>47 → 55</b> (1A-1L)</p>	<p><u>Date</u> : 17.11.03. Début d'après-midi. Après la pause de midi.</p> <p><u>Durée</u> : 13H30-14H40.</p> <p><u>Matériel</u> : enceintes Ramsa, magnétophone Panasonic, rétroprojecteur Panasonic.</p>
	<p><b>56 → 62</b> (4A-1L)</p>	<p><u>Date</u> : 25.11.03. Fin d'après-midi. Après une classe.</p> <p><u>Durée</u> : 16H45-18H00.</p> <p><u>Matériel</u> : enceintes Ramsa, magnétophone Panasonic, rétroprojecteur Panasonic.</p>

### **Rendus :**

Les étudiants ont reçu des feuillets contenant les consignes (en japonais), les exemples et les espaces de réponses : il s'agissait dans la plupart des cas d'entourer la réponse choisie parmi plusieurs proposées.

Toutes les réponses ont ensuite été codées numériquement et stockées sous forme de classeurs Excel. A chaque test correspond un classeur, à chaque feuille de ce classeur correspond un sujet. Sur chaque feuille, les stimuli sont en ligne, tandis que les variables indépendantes (VI) et dépendantes (VD) sont en colonnes. Ce sont ces données qui ont été utilisées pour les calculs et les tests statistiques.

## **IV. Première partie : axe segmental (Groupe 1)**

Cette première partie est consacrée au traitement des représentations phonographémiques segmentales établies entre le plan phonologique (/b, v, R, l/) et graphémique ("b, v, R, l") pour nos sujets japonais apprenant la phonologie du français.

### **IV.1. Test A (Identification)**

#### **Tâche**

Appariement phonographémique segmental (B/V, R/L) (et dans une certaine mesure EU/OU). Les étudiants doivent choisir une transcription parmi deux (ou quatre lorsque deux segments sont associés) pour le non-mot qu'ils ont entendu.

Exemple 1 : [løkegi] → "leukegi"/ "reukegi"/ "loukegi"/ "roukegi"?

Exemple 2 : [ʁadebo] → "ladebo"/"radebo" ?

L'espace de rendu se présentait ainsi :

1. MEJASOUB	MEJASEUV	MEJASOUV	MEJASEUB
2. RADEBO	LADEBO		

Rappelons que nous nous intéressons au traitement des couples phonématiques (/R/, /l/), (/v/, /b/), et dans une moindre mesure (/ø/, /u/), en rapport avec leurs représentations orthographiques associées. L'objectif de ce premier test était de récolter des données statistiques sur les représentations phonographémiques des sujets japonais apprenant le français : quels sont les graphèmes alphabétiques mentalement associés de manière automatique par les sujets aux phonèmes en question ? Cette question générale a ensuite été affinée via les variables indépendantes formulées plus bas.

### Stimuli

Les non-mots ont été composés en faisant varier la position du segment cible (d'où le choix de stimuli trisyllabiques), ainsi que la nature de l'environnement vocalique. En ce qui concerne les réponses orthographiques proposées (qui constituent des stimuli visuels à part entière), nous avons effectué des choix :

- "e"=[e] (seule valeur en japonais et l'une des valeurs possibles en français)<sup>170</sup> ;

Graphème	Japonais	Français
"e"	[e]	[e], [ø], [ɛ], etc.

- "eu"=[ ø ] (digraphe existant en français, de valeur différente en japonais ; permet de distinguer "e" et "eu")

Digraphe	Japonais	Français
"eu"	[eu]	[ø] ou [œ]

<sup>170</sup> Pour les tâches de comptage syllabique avec présentation visuelle des non-mots (tests D et E), la possibilité que le graphème "e" en finale soit interprété comme un « e muet » (par exemple dans "plamote", ne comptant alors que deux syllabes si ce « e » est « muet ») et qu'il puisse ainsi brouiller nos résultats a été rejetée puisque, d'une part le nombre minimal de syllabes proposé aux sujets est trois, et d'autre part les non-mots étaient écrits en lettres majuscules ("PLAMOTE"), rendant ainsi plus probable l'association "E"/[e] de par l'absence générale d'accent sur les lettres majuscules ("é" ↔ "E").

- "ou"=[u] (digraphe existant en français, de valeur différente en japonais ; permet de distinguer "u" et "ou")

<b>Digraphe</b>	<b>Japonais</b>	<b>Français</b>
"ou"	[o:] ou [oɯ]	[u]

Il faut également garder à l'esprit le tableau suivant :

<b>Graphème</b>	<b>Japonais</b>	<b>Français</b>
"u"	[ɯ]	[y]

Les choix relatifs aux autres graphèmes (par exemple "g"=[g]) ont moins d'importance. Par ailleurs, sur la feuille de réponse, les transcriptions proposées, correctes et incorrectes, ont été réparties de manière aléatoire (en première, deuxième, ou troisième position lorsqu'il y a quatre transcriptions proposées) de manière à éviter la mise en œuvre de stratégie de repérage des transcriptions correctes par les sujets.

Voici les stimuli avant leur répartition aléatoire :

<u>Environnement</u> <u>Vocalique (V)C(V)</u>		<u>POSITION</u>		
		<u>INITIALE</u> (#_VCVCV)	<u>MEDIALE pénultième</u> (CVCV__V)	<u>FINALE</u> (CVCVCV_#)
<b>B</b>	<b>A</b>	Bagade	PageBa	PotigaB
	<b>EU</b>	Beusoki	BenaBeu	NeleteuB
	<b>I</b>	Bitane	MoteBi	ZashakiB
	<b>OU</b>	Bousine	DemaBou	MejasouB
<b>V</b>	<b>A</b>	Vagade	NageVa	PotigaV
	<b>EU</b>	Veusoki	SenaVeu	NeluteuV
	<b>I</b>	Vitane	ZoteVi	ZashakiV
	<b>OU</b>	Vousine	JemaVou	MejasouV
<b>R</b>	<b>A</b>	Radebo	KepiRa	KometaR
	<b>EU</b>	Reukegi	GineReu	SatodeuR
	<b>I</b>	Ripisa	PejaRi	JodeniR
	<b>OU</b>	Roudoke	MineRou	PekezouR
<b>L</b>	<b>A</b>	Ladebo	TepiLa	KometaL
	<b>EU</b>	Leukegi	DineLeu	SatodeuL
	<b>I</b>	Lipisa	NejaLi	JodeniL
	<b>OU</b>	Loudoke	ZineLou	PekezouL

Ces stimuli ont ensuite été placés de manière aléatoire. Les intervalles de temps laissés aux sujets pour répondre étaient organisés de la manière suivante : BIP-(1 seconde)-XXX-(5 secondes pour répondre)-BIP-(1 seconde)-XXX-(5 secondes pour répondre)-etc.

Le choix d'un environnement vocalique stable (A, I et EU/OU) pour tous les stimuli visait à réduire et à pouvoir éventuellement évaluer *l'effet de la fréquence* d'occurrence

syllabique/moraique (certaines unités, par exemple /ba/, pouvant être plus fréquentes que d'autres, par exemple /bi/, dans certaines positions, en japonais) dans nos tests.

Les stimuli visuels étaient présentés en caractères majuscules et les couples de non-mots visuels étaient toujours présentés dans le même ordre (par exemple, "R"- "L", pour le stimulus [R], et pour le stimulus [I]).

### **Variables**

<b>Variable Indépendante</b>	<b>Type</b>	<b>Nom Excel</b>	<b>Code</b>
Position du segment consonantique cible	Qln	position	1=initiale (#__V), 2= avant-dernier segment (syllabe finale __V#), 3=dernier segment (syllabe finale V__#)
Nature des segments consonantiques	Qln	natsegc	1=B, 2=V, 3=R, 4=L
Nature des segments vocaliques	Qln	natsegv	1=A, 2= I, 3=EU, 4=OU

<b>Variable Dépendante</b>	<b>Type</b>	<b>Nom Excel</b>	<b>Code</b>
Confusion phonographémique consonantique	Qln	confusion	1=oui, 2=non
Choix du graphème consonantique existant en japonais (B/R)	Qln	choijapc	1=oui, 2=non
Si V= EU/OU, choix du graphème vocalique existant en japonais (OU)	Qln	choijapv	1=oui, 2=non, 3=non pertinent

## Hypothèses

<b>Générale</b>	Le système segmental phonographémique (inventaire phonologique segmental et représentations phonographémiques) en L1 (japonais) influence la perception segmentale phonographémique (segments phonologiques et graphèmes) en L2 (français).
<b>De Travail</b>	Lorsque les systèmes graphémiques L1 et L2 convergent, au moins partiellement, tandis que les systèmes phonologiques divergent, au moins partiellement, les graphèmes communs L1-L2 vont être, préférentiellement, automatiquement interprétés comme des graphèmes L1. Les unités phonologiques qui leur sont respectivement associées en L1 et en L2 vont alors être préférentiellement perçues comme des unités phonologiques de L1, ce qui est opérationnalisé en termes de taux de confusion phonographémique. Notre hypothèse de travail peut donc être formulée ainsi : une confusion phonographémique significative va être entraînée par des unités de L2, partagées, soit au plan graphémique (« R/B »), soit au plan phonético-phonologique <sup>171</sup> , entre les systèmes L1 et L2.
<b>Opérationnelles</b>	<ol style="list-style-type: none"> <li>1. La position (lexicale et syllabique) du segment consonantique en L2 n'influence pas de manière significative le taux de confusion phonographémique ;</li> <li>2. La nature (en termes de famille et de sonorité phonologique) du segment consonantique en L2 n'influence pas de manière significative le taux de confusion phonographémique ;</li> <li>3. La nature du segment vocalique adjacent au segment consonantique en L2 (existant ou non en L1) n'influence pas de manière significative le</li> </ol>

<sup>171</sup> Le [l] et le [v] français pouvant, sous certaines conditions, être respectivement perçus comme un [r] et un [b] japonais au plan *phonétique*, donnant lieu ensuite à une identification *phonologique* erronée.



	<p>taux de confusion phonographémique consonantique ;</p> <p>4. Dans tous les cas, l'étudiant choisit préférentiellement les graphèmes existant en L1.</p>
--	--

## **IV.2. Test F (Identification et Discrimination)**

### **Tâche**

Discrimination phonologique et appariement phonographémique segmental (B/V, R/L) (et dans une certaine mesure EU/OU) : les étudiants entendent deux non-mots consécutivement, parfois identiques, parfois différents. Ils doivent choisir une transcription parmi deux pour chaque non-mot et ainsi indiquer leur identité ou leur différence.

Ex. : [potigav]/ [potigab] → "potigav" ou "potigab" / "potigav" ou "potigab" ?

Une possibilité de transcription personnelle a été intégrée aux rendus, de manière à vérifier, si nécessaire, l'identité de l'élément discriminatif (voir à ce propos les résultats de nos prétests). Une tâche de transcription ouverte aurait été irréalisable dans nos conditions et ne nous aurait peut-être pas apporté les informations recherchées. Cependant, étant donné les contraintes temporelles et attentionnelles, il n'y a eu que très peu de transcriptions personnelles proposées par les sujets.

L'espace de réponse se présentait ainsi :

N°	1 <sup>er</sup> MOT	2 <sup>ème</sup> MOT	Transcriptions éventuelles	
			1 <sup>er</sup> mot	2 <sup>ème</sup> mot
1.	POTIGAV	POTIGAV		
	POTIGAB	POTIGAB		

Le test F se pose en complément du test précédent : tandis que le test A ne visait à obtenir que des informations d'ordre descriptif sur les représentations phonographémiques chez nos sujets, le test F se propose d'évaluer l'influence de la tâche sur ces représentations. Dans cette tâche, plus complexe que la tâche A, deux stratégies peuvent théoriquement être mises en œuvre par les sujets :

- Une stratégie phonologique : discrimination phonologique, puis transcription du premier ou du non-mot, et ainsi du non-mot restant par différenciation ;
- Une stratégie graphémique : transcription des deux non-mots, sans se soucier de leur ressemblance ou différence phonologique.

Si le couple de tests A-F ne nous permet pas d'évaluer l'influence possible du support orthographique sur la tâche de discrimination phonologique (pour cela nous aurions besoin de deux tâches de discrimination : l'une sans support orthographique et l'autre avec), les tâches A et F nous permettent cependant d'évaluer la variabilité inter-tâche (*via* la variation du taux de confusion phonographémique) des représentations phonographémiques chez les sujets. Si nos résultats indiquaient que ces représentations ne sont pas stables, on pourrait alors penser que le support orthographique pourrait ne pas être un atout, mais plutôt un élément parasite pour nos sujets en tâche de discrimination. Mentionnons que le test F visait également à nous apporter des renseignements sur le système (inter-) phonologique des apprenants, *via* la tâche de discrimination phonologique.

### **Stimuli**

Les non-mots utilisés pour le test F sont les mêmes que dans le test A. En ce qui concerne les paires de stimuli consécutifs, nous avons placé un nombre égal de couples identiques (CV1) et différents (CV2) pour chaque phonème cible (/b, v, R, l/) de manière à compenser *l'effet d'ordre* dans la présentation des couples de stimuli.

Environnement vocalique (V)C(V)		POSITION		
		INITIALE (#_VCVCV)	MEDIALE pénultième (CVCV__V)	FINALE (CVCVCV__#)
B	A1	Bagade1	PageBa1	PotigaB1
	A2	Bagade2	PageBa2	PotigaB2
V	A1	Vagade1	NageVa1	PotigaV1
	A2	Vagade2	NageVa2	PotigaV2
B	EU1	Beusoki1	BenaBeu1	NeleteuB1
	EU2	Beusoki2	BenaBeu2	NeleteuB2
V	EU 1	Veusoki1	SenaVeu1	NeleteuV1
	EU 2	Veusoki2	SenaVeu2	NeleteuV2
B	I1	Bitane1	MoteBi1	ZashakiB1
	I2	Bitane2	MoteBi2	ZashakiB2
V	I1	Vitane1	ZoteVi1	ZashakiV1
	I2	Vitane2	ZoteVi2	ZashakiV2
B	OU1	Bousine1	DemaBou1	MejasouB1
	OU2	Bousine2	DemaBou2	MejasouB2
V	OU1	Vousine1	JemaVou1	MejasouV1
	OU2	Vousine2	JemaVou2	MejasouV2
R	A1	Radebo1	KepiRa1	KometaR1
	A2	Radebo2	KepiRa2	KometaR2
L	A1	Ladebo1	TepiLa1	KometaL1
	A2	Ladebo2	TepiLa2	KometaL2
R	EU1	Reukegi1	GineReu1	SatodeuR1
	EU 2	Reukegi2	GineReu2	SatodeuR2
L	EU1	Leukegi1	DineLeu1	SatodeuL1
	EU2	Leukegi2	DineLeu2	SatodeuL2
R	I1	Ripisa1	PejaRi1	JodeniR1
	I2	Ripisa2	PejaRi2	JodeniR2

L	I1	Lipisa1	NejaLi1	JodeniL1
	I2	Lipisa2	NejaLi2	JodeniL2
R	OU1	Roudoke1	MineRou1	PekezouR1
	OU2	Roudoke2	MineRou2	PekezouR2
L	OU1	Loudoke1	ZineLou1	PekezouL1
	OU2	Loudoke2	ZineLou2	PekezouL2

Ces stimuli ont ensuite été distribués de manière aléatoire. Les intervalles de temps laissés aux sujets pour répondre étaient organisés de la manière suivante : BIP-(1 seconde)-XXX/(1 seconde)/XXX-(6 secondes pour répondre)-BIP-(1 seconde)-XXX/(1 seconde)/XXX-(6 secondes pour répondre)-etc.

### Variables

<b>Variable Indépendante</b>	<b>Type</b>	<b>Nom Excel</b>	<b>Code</b>
Position du segment consonantique cible	Qln	position	1=initiale (#__V), 2= avant-dernier segment (syllabe finale __V#), 3=dernier segment (syllabe finale V__#)
Nature du segment consonantique	Qln	natsegc	1=B, 2=V, 3=R, 4=L
Nature du segment vocalique	Qln	natsegv	1=A, 2= I, 3=EU, 4=OU
Nature des couples de stimuli consonantiques	Qln	natcoupl	1=identique, 2=différent
Existence du premier graphème consonantique cible	Qln	exgraun	1=en français seulement (/v, l/, par exemple : Vageda/Bageda, Vageda/Bageda), 2=en japonais et en français (/b, R/, par exemple : Bageda/Bageda, Bageda/Vageda))

<b>Variable Dépendante</b>	<b>Type</b>	<b>Nom Excel</b>	<b>Code</b>
Confusion phonologique	Qln	confph	1=ooui, 2=non
Confusion phonographémique consonantique sur stimulus 1	Qln	cophgrun	1=ooui, 2=non
Choix du graphème consonantique existant en japonais (B/R) sur le stimulus 1	Qln	choijapc	1=ooui, 2=non
Si V= EU/OU, choix du graphème vocalique existant en japonais (OU) sur le stimulus 1	Qln	choijapv	1=ooui, 2=non, 3=non pertinent

### **Hypothèses :**

<b>Générale</b>	Le système segmental phonographémique (inventaire phonologique segmental et représentations phonographémiques) en L1 (japonais) influence la perception segmentale phonographémique (segments phonologiques et graphèmes) en L2 (français).
<b>De Travail</b>	<ol style="list-style-type: none"> <li>1) La performance des étudiants japonais en tâche de discrimination phonologique pour les couples de segments cibles (/b/, /v/) et (/R/, /l/) en français est faible.</li> <li>2) Une confusion phonographémique significative va être entraînée par des unités de L2, partagées, soit au plan graphémique (« R/B »), soit au plan phonético-phonologique, entre les systèmes L1 et L2.</li> <li>3) La complexité de la tâche (appariement et discrimination phonologique) entraîne un partage attentionnel conduisant à un taux de confusion phonographémique plus élevé que pour une tâche simple (appariement)</li> </ol>

<b>Opérationnelles</b>	<ol style="list-style-type: none"><li>1) La position du segment cible en L2 n'influence pas de manière significative les taux de confusion phonologique et phonographémique ;</li><li>2) La nature du segment cible en L2 n'influence pas de manière significative les taux de confusion phonologique et phonographémique ;</li><li>3) La nature des couples de stimuli en L2 n'influence pas de manière significative les taux de confusion phonologique et phonographémique ;</li><li>4) L'existence du graphème consonantique cible du premier stimulus en L1 ("b" et "R") n'influence pas de manière significative les taux de confusion phonologique et phonographémique ;</li><li>5) Le taux global de confusion phonographémique est plus élevé pour la tâche F que pour la tâche A.</li></ol>
------------------------	---

## **V. Deuxième partie : axe syllabique (Groupes 1 et 2)**

Cette deuxième partie est consacrée au traitement des représentations phonographémiques syllabiques, et plus particulièrement aux procédures de segmentation syllabique/moraïque appliquée par nos sujets japonais à des non-mots. En particulier nous nous intéressons au traitement des groupes consonantiques CCV et VCC français (attaques et codas syllabiques branchantes) avec et sans épenthèses vocaliques (CvCv et VCvCv). L'un des objectifs essentiels de cette deuxième partie est d'évaluer *l'effet de modalité* sur la procédure de segmentation syllabique, *via* la comparaison des résultats obtenus selon trois modalités différentes : unimodalité auditive (tests C1 et C2), unimodalité visuelle orthographique (test D) et bimodalité audiovisuelle orthographique (test E).

### **V.1. Test C1 (Groupe 1 – Unimodalité Auditive)**

#### **Tâche**

Comptage syllabique en unimodalité auditive : les étudiants entendent un non-mot et doivent entourer sur leur feuille le nombre de cercles correspondant au nombre de

syllabes perçues, à choisir entre trois groupes de cercles : trois, quatre et cinq. Le choix d'un format en cercle permet d'éviter, d'une part une interférence directe avec des représentations graphémiques externes, et d'autre part un effet parasite direct possible dû à un comptage numérique vraisemblablement en japonais, puisque le format choisi dans notre protocole encourage une procédure de segmentation par pointage et non par comptage numérique subvocalisé.

Exemple : [tedisats] → OOO/OOOO/OOOOO ?

L'espace de rendu se présentait ainsi :

1.	OOO	OOOO	OOOOO
----	-----	------	-------

L'objectif principal de ce test est de récolter des informations sur la segmentation syllabique de non-mots « français » effectuée par nos sujets japonais sans support orthographique. Les implications du choix du terme « syllabe » dans les consignes et sa traduction en japonais, ainsi que de la structure à l'étude, vis-à-vis de la « more » japonaise, seront discutées plus bas (Kubozono, 1989 ; Otake et *al.*, 1993 ; Labrune, 2001a, 2001b).

### **Stimuli**

Les non-mots ont été composés à partir de la distribution des groupes biconsonantiques en français telle que donnée par Dell (1995), groupes consonantiques que nous avons répartis entre nos groupes de sujets 1 et 2. L'inconvénient d'un tel choix est bien évidemment la non-homogénéité de la répartition des groupes, qui peut poser des difficultés par la suite, notamment dans l'analyse statistique. Cependant, dans une perspective écologique, il nous a semblé essentiel de respecter la réalité de la langue française, et de ne présenter aux sujets que des combinaisons consonantiques qu'ils pourraient *effectivement* rencontrer au cours de leur apprentissage. L'un des objectifs de tels tests étant à terme de proposer un classement de difficulté pour les apprenants

en question, il serait *a priori* non-pertinent, et en tous cas non-économe, d'y intégrer des combinaisons absentes de la langue française<sup>172</sup>. Pour le groupe 1, qui a passé le test C1, en voici la liste :

<b>Position</b>	<b>Groupes biconsonantiques -Test C1</b>
<b>Initiale / Médiale</b>	/pn, ps, pf, pt, kn, km, kv, ks, kt, tl, tm, ts, tsh/ (Plosive sourde + C)
<b>Finale</b>	/pt, kt, ps, km, ks, tl, tm, ts, tsh/ (Plosive sourde + C) + /ft/

Les groupes choisis ici sont donc essentiellement de la forme /Plosive Sourde + C/, auxquels vient s'ajouter le groupe /ft/ en finale.

Afin de parer à toute stratégie de repérage de la part des sujets, nous avons ensuite élaboré une série de distracteurs de la manière suivante : à chaque non-mot correspond un distracteur, qui correspond au non-mot tel qu'il pourrait vraisemblablement être après avoir subi une épenthèse au cours de son traitement par nos sujets en vue de respecter la syllabation du japonais : le groupe /CC/ devient alors /CvC/ ou /CvCv/ en position finale. Le choix des voyelles épenthétiques et leur insertion dans les groupes consonantiques ont été établis d'après l'abondante littérature sur le sujet, notamment les travaux de Dupoux et ses collègues (1999) d'une part, et de Shinohara (1997) d'autre part. Le /u/ (voyelle épenthétique principale) correspond ici à une voyelle neutre, de type schwa. Notons que ce protocole de comptage syllabique diffère sensiblement de celui de Nesterenko (2002) dans lequel des stimuli filtrés sont présentés aux sujets afin d'attirer leur attention sur les prééminences que représentent les syllabes dans l'organisation prosodique de l'énoncé.

Nous avons par ailleurs ici encore privilégié la stabilité de l'environnement phonologique en établissant de nombreuses similitudes entre les stimuli (paires minimales (axe vertical), contexte vocalique immédiat (antérieur ou postérieur)

---

<sup>172</sup> Notre démarche se situe donc à mi-chemin entre celle des grammaires artificielles, avec l'utilisation de non-mots, et une démarche plus écologiquement, avec le respect des contraintes phonotactiques et de la distribution réelle des éléments linguistiques à l'étude dans la langue française.



identique (axe horizontal)), de manière à pouvoir effectuer par la suite des comparaisons phonologiques suffisamment rigoureuses. C'est précisément ce contrôle des stimuli qui permet de compenser, du point de vue de la validité des tests, la variation inhérente à l'environnement écologique choisi pour la passation des tests.

Voici la liste des stimuli, avant leur répartition aléatoire :

Groupes Consonantiques sans (CC) et avec (CvC) épenhèse	POSITION		
	INITIALE (#_VCVCV)	MEDIALE pénultième (CVCV_V)	FINALE (CVCVCV_#)
PN	Pnotoda	todaPNo	
PuN	PuNotoda	todaPuNo	
PS	Psisheza	ShezaPSi	ShekozaPS
PuS	PuSisheza	ShezaPuSi	ShekozaPuSu
PF	Pfenika	NikaPFe	
PuF	PuFenika	NikaPuFe	
PT	Ptideso	DesoPTi	DekasoPT
PuT	PuTideso	desoPuTi	DekasoPuTo
KN	Knesagi	PagiKNe	
KuN	KuNesagi	PagiKuNE	
KM	Kmesagi	BagiKMe	KadagiKM
KuM	KuMesagi	BagiKuMe	KadagiKuMu
KV	Kvesagi	MagiKVe	
KuV	KuVesagi	MagiKuVE	
KS	Ksesagi	TagiKSe	MapegiKS
KuS	KuSesagi	TagiKuSE	MapegiKuSu
KT	Ktesagi	DagiKTe	NalogiKT

KuT	KuTesagi	DagiKuTE	NalogiKuTo
TL	Tletosa	TosaTLe	TopisaTL
ToL	ToLetosa	TosaToLe	TopisaToLu
TM	Tmetosa	TosaTMe	TopisaTM
ToM	ToMetosa	TosaToMe	TopisaToMu
TS	Tsetosa	TosaTSe	TedisaTS
ToS	ToSetosa	TosaToSe	TedisaToSu
TSH	TSHetosa	TosaTShe	TidesaTSH
ToSH	ToSHetosa	TosaToShe	TidesaToSHu
FT			MezakoFT
FuTo			MezakoFuTo

Pour les groupes consonantiques, nous avons donc plusieurs structures syllabiques *a priori* illicites en japonais :

- ⇒ Des attaques branchantes non-accentuées pour la colonne « Initiale »,
- ⇒ Des attaques branchantes accentuées pour la colonne « Médiale »,
- ⇒ Des codas branchantes accentuées pour la colonne « Finale ».

Les intervalles de temps laissés aux sujets pour répondre étaient organisés de la manière suivante : BIP-(1 seconde)-XXX-(7 secondes pour répondre)-BIP-(1 seconde)-XXX-(7 secondes pour répondre)-etc.

### Variables

Variable Indépendante	Type	Nom Excel	Code
Position des groupes	Qln	position	1=initiale (syllabe initiale #__V), 2= avant-dernière position (syllabe finale __V#), 3=dernière position (syllabe finale V__#)

Famille générale des groupes	QIn	famigcc	1=Obs+Liqu, 2=Obs+NonLiqu, 3=Sib+Stop, 4=Sib+NonStop
Famille des groupes	QIn	familcc	1=Obs+Lat, 2=Obs+Rhot, 3=Obs+Nas, 4=Obs+Cont, 5=Obs+Sib, 6=Obs+Stop, 7=Sib+Stop, 8=Sib+Nas, 9=Sib+Cont, 10=Sib+Lat
Nature des groupes	QIn	naturecc	1=pn, 2=ps, 3=pf, 4=pt, 5=kn, 6=km, 7=kv, 8=ks, 9=kt, 10=tl, 11=tm, 12=ts, 13=tsh ; 36=ft <sup>173</sup>
Type de sonorité pour le groupe CC	QIn	typsono	1= Montant (M), 2= Plateau (P), 3= Descendant (D)
Ecart de sonorité entre les deux consonnes du groupe	Qlo	ecarsono	0=0, 1=1, 2=2, 3=3, 4=4, 5=5, 6=6
Nombre de syllabes stimuli	QIn	nombsyls	1=3 syllabes (CC), 2=4 syllabes (CvC), 3=5 syllabes (CvCv)

Variable Dépendante	Type	Nom Excel	Code
Nombre de syllabes comptées	Qlo	x=nombsylc	1=3 syllabes (aucune épenthèse), 2=4 syllabes (une épenthèse), 3=5 syllabes (deux épenthèses)
Confusion syllabique	QIn	x=confsyll	1=oui, 2=non

<sup>173</sup> Le code un peu surprenant du groupe /ft/ n'est dû qu'à des raisons techniques d'organisation du codage et n'appelle aucune autre remarque.

## Hypothèses

<b>Générale</b>	Le système phonographémique syllabique en L1 (japonais) influence la perception phonographémique syllabique en L2 (français).
<b>De Travail</b>	Le classement des contraintes phonologiques rendant compte de la structure syllabique du lexique phonologique en L1 influe sur la conscience phonologique des apprenants de L2. Si la structure syllabique des stimuli en L2 est conforme à ce classement, alors la conscience phonologique des apprenants (exemple : CVCV) est conforme à leur perception phonétique/identification phonologique des stimuli (exemple : CVCV). Si la structure syllabique des stimuli n'est pas conforme à ce classement, alors la conscience phonologique des apprenants (CVCV) n'est pas conforme à leur perception phonétique/identification phonologique (exemple : CCV). Nous supposons donc que les stimuli du type /CC/ provoqueront des épenthèses du type /CvC/, étant entendu que, d'un point de vue principal, la structure /CVCV/ est moins marquée que la structure /CCV/.
<b>Opérationnelles</b>	<ol style="list-style-type: none"> <li>1) La position du groupe cible en L2 n'influence pas de manière significative le taux d'épenthèses ;</li> <li>2) La nature (famille et sonorité) du groupe cible en L2 n'influence pas de manière significative le taux d'épenthèses ;</li> <li>3) L'insertion de voyelle(s) épenthétique(s) propre(s) à la L1 dans les groupes cibles en L2 donne lieu au respect de la structure syllabique des stimuli (aucune confusion syllabique).</li> </ol>

## **V.2. Test C2 (Groupe 2 – Unimodalité Auditive)**

Le test C2 ne se différencie du test C1 que sur les points suivants :

- Les sujets : il s'agit du Groupe 2 (62 sujets, différents du Groupe 1)<sup>174</sup>.
- Les stimuli : les non-mots composant les stimuli sont différents ; les groupes de consonnes du test C2 relèvent des familles générales OBLI (obstruantes + liquides) et SC (s + consonne) :

<b>Position</b>	<b>Groupes biconsonantiques - Test C2</b>
<b>Initiale / Médiale</b>	<ul style="list-style-type: none"> <li>• /pl, pr, bl, br, fl, fr, vl, vr, tr, dr, kl, kr, gl, gr/ (OBLI) ;</li> <li>• /sp, st, sk/ (S + Plosive) ;</li> <li>• /sl, sm, sn, sv, sf/ (S + C non-plosive).</li> </ul> <p>(Tous ces groupes sont considérés par Dell comme des attaques syllabiques bien formées, tandis que les autres qu'il relève sont qualifiés de « rares ». Rappelons également que sont exclus de la famille OBLI les groupes composés de deux segments coronaux (par exemple /tl/) ainsi que ceux dont le premier membre est une coronale fricative (donc /sl/).</p>
<b>Finale</b>	<ul style="list-style-type: none"> <li>• /pl, pr, bl, br, fl, fr, vl, vr, tr, dr, kl, kr, gl, gr/ (OBLI) ;</li> <li>• /sp, st, sk/ (S + Plosive) ;</li> <li>• /sm/ (S + M).</li> </ul> <p>(En finale /vl/ est absent, mais il s'agit d'un « accidental gap » (Dell, 1995, p. 18))</p>

Les deux grandes familles OBLI et SC ont déjà reçu une attention particulière dans la littérature en interphonologie et elles présentent pour nous un intérêt particulier :

---

<sup>174</sup> Pour la tâche C2, les sujets 48 et 53 n'ayant pas correctement répondu, leurs résultats ont été retirés de ceux du groupe.

- **OBLI** : car ces groupes présentent un double problème (syllabique et segmental ; Paradis, 1996) pour les sujets japonais ;
- **SC** : car plusieurs études ont conduit à établir une distinction entre Obstruantes+ Sonorantes d'une part et Sibilantes + Plosives d'autre part, étant donné que certains groupes SC violent en effet parfois le principe de sonorité (ex : /st/), tandis que d'autres non (ex : /sn/).

Voici la liste des stimuli, avant leur répartition aléatoire :

Groupes Consonantiques sans (CC) et avec (CvC) épenhèse	<b>POSITION</b>		
	<b>INITIALE</b> (#_VCVCV)	<b>MEDIALE</b> <b>pénultième</b> (CVCV__V)	<b>FINALE</b> <b>pénultième</b> (CVCVCV_#)
PL	Plokama	TamaPLo	TamakoPL
PuL	PuLokama	TamaPuLo	TamakoPuLu
PR	Prokama	TamaPRo	TamakoPR
PuR	PuRokama	TamaPuRo	tamakoPuRu
BL	Blokama	TamaBLo	tamakoBL
BuL	BuLokama	TamaBuLo	TamakoBuLu
BR	Brokama	TamaBRo	TamakoBR
BuR	BuRokama	TamaBuRo	TamakoBuRu
FL	Flamote	MoteFLa	MotesaFL
FuL	FuLamote	MoteFuLa	MotesaFuLu
FR	Framote	MoteFRa	MotesaFR
FuR	FuRamote	MoteFuRa	MotesaFuRu
VL	Vlamote	MoteVLa	MotesaVL
VuL	VuLamote	MoteVuLa	MotesaVuLu
VR	Vramote	MoteVRa	MotesaVR
VuR	VuRamote	MoteVuRa	MotesaVuRu

TR	Trosema	SemaTRo	SemagoTR
ToR	ToRosema	SemaToRo	SemagoToRu
DR	Drosema	SemaDRo	SemagoDR
DoR	DoRosema	SemaDoRo	SemagoDoRu
KL	Klanape	NapeKLa	NapetaKL
KuL	KuLanape	NapeKuLa	NapetaKuLu
KR	Kranape	NapeKRa	NapetaKR
KuR	KuRanape	NapeKuRa	NapetaKuRu
GL	Glanape	NapeGLa	NapetaGL
GuL	GuLanape	NapeGuLa	NapetaGuLu
GR	Granape	napeGRa	NapetaGR
GuR	GuRanape	napeGuRa	NapetaGuRu
SP	Spijena	jenaSPi	JenakiSP
SuP	SuPijena	jenaSuPi	JenakiSuPu
ST	Stodime	dimeSTo	DimeloST
SuT	SuTodime	dimeSuTo	DimeloSuTo
SK	Skamose	moseSKa	ModenaSK
SuK	SuKamose	moseSuKa	ModenaSuKu
SL	Slotipa	TipaSLO	
SuL	SuLotipa	TipaSuLo	
SM	Smotipa	TipaSMO	TipakoSM
SuM	SuMotipa	TipaSuMo	TipakoSuMu
SN	Snotipa	tipaSNO	
SuN	SuNotipa	tipaSuNo	
SV	Svotipa	tipaSVO	
SuV	SuVotipa	tipaSuVo	
SF	Sfotipa	tipaSFO	
SuF	SuFotipa	tipaSuFo	

Excepté pour le groupe /vl/ en final, nous avons respecté la distribution des groupes donnée par Dell (1995).

- Les variables indépendantes : étant donné ces groupes consonantiques, il a semblé utile de rajouter à notre classification la variable indépendante « Sous-famille », comme cela apparaît dans le tableau ci-dessous :

<b>Variable Indépendante</b>	<b>Type</b>	<b>Nom Excel</b>	<b>Code</b>
Position des groupes	Qln	position	1=initiale (syllabe initiale #__V), 2= avant-dernière position (syllabe finale __V#), 3=dernière position (syllabe finale V__#)
Famille générale des groupes <sup>175</sup>	Qln	famigcc	1=Obs+Liqu, 2=Obs+NonLiqu, 3=Sib+Stop, 4=Sib+NonStop
Famille des groupes	Qln	familcc	1=Obs+Lat, 2=Obs+Rhot, 3=Obs+Nas, 4=Obs+Cont, 5=Obs+Sib, 6=Obs+Stop, 7=Sib+Stop, 8=Sib+Nas, 9=Sib+Cont, 10=Sib+Lat
Sous-famille des groupes	Qln	sfamicc	1=Stop+Lat, 2=Cont+Lat, 3=Stop+Rhot, 4=Cont+Rhot, 5=Obs+Nas, 6=Obs+Cont, 7=Obs+Sib, 8=Obs+Stop, 9=Sib+Stop, 10=Sib+Nas, 11=Sib+Cont, 12=Sib+Lat

---

<sup>175</sup> Notons dès à présent que, pour les tâches C2, D et E, nous n'avons utilisé qu'un seul représentant de la classe des Sibilantes, à savoir /s/. Il sera donc entendu que nos conclusions seront contraintes par cette limitation.



Nature des groupes	QIn	naturecc	14=pl, 15=pR, 16=bl, 17=bR, 18=fl, 19=fR, 20=vl, 21=vR, 22=tR, 23=dR, 24=kl, 25=kR, 26=gl, 27=gR, 28=sp, 29=st, 30=sk, 31=sl, 32=sm, 33=sn, 34=sv, 35=sf
Type de sonorité pour le groupe CC	QIn	typsono	1= Montant (M), 2= Plateau (P), 3= Descendant (D)
Ecart de sonorité entre les deux consonnes du groupe	Qlo	ecarsono	0=0, 1=1, 2=2, 3=3, 4=4, 5=5, 6=6
Nombre de syllabes stimuli	QIn	nombsyls	1=3 syllabes (CC), 2=4 syllabes (CvC), 3=5 syllabes (CvCv)

Pour le reste, le test C2 est identique au test C1.

### **V.3. Test D (Groupe 2 – Unimodalité Visuelle)**

Le test D ne se différencie du test C2 que sur les points suivants :

- La modalité de présentation des stimuli : Les stimuli sont présentés en unimodalité visuelle, sous leur forme orthographique alphabétique, selon les codes phonographémiques présentés plus haut (cf. test A). Les enseignants ont présenté ceux-ci à l'aide d'un rétroprojecteur habituellement utilisé dans leur salle de classe. Les non-mots (police Normale, caractères Times New Roman, taille 65, couleur Noir, fond blanc) étaient ainsi projetés l'un à la suite de l'autre sur écran, tandis qu'une cassette sur laquelle étaient enregistrés à intervalles réguliers des indicateurs sonores (« BIP »), permettait aux étudiants et à l'enseignant de progresser de manière régulière dans la présentation des stimuli. Les intervalles de temps se présentaient ainsi : BIP-(1 seconde)-XXX-(7 secondes pour lire le non-mot et pour répondre)-BIP-etc. La durée d'exposition aux stimuli est donc évidemment plus longue que dans la tâche C2. Cela ne pose aucun problème puisque c'est

précisément le rôle des représentations orthographiques en tant que représentations permanentes (et non fugaces, comme le sont les stimuli auditifs) qui nous intéresse : nous envisageons en effet ici système graphémique du japonais comme « système de fixation et de conscientisation » des unités phonologiques en L1, et ainsi du classement phonologique des contraintes en L1. Voici un exemple de stimulus visuel :

## PROKAMA

- Le nombre et la nature des stimuli : afin d'éviter de rendre le test trop long, et parce que la présence de distracteurs basés sur l'insertion de voyelles épenthétiques comme dans les tests C1 et C2 ne nous aurait sans doute pas apporté d'informations essentielles, nous n'avons sélectionné pour le test D que les non-mots cibles du test C2 (groupe CCV ou VCC) auxquels nous avons adjoint cinq distracteurs, également répartis dans la liste des stimuli, afin de maintenir éveillée l'attention des sujets et éviter toute stratégie de repérage. Ces distracteurs étaient du type : « SDOGABATE ».

Voici la liste des stimuli avant leur répartition aléatoire :

Groupe consonantique	POSITION		
	INITIALE (#_VCVCV)	MEDIALE pénultième (CVCV_V)	FINALE pénultième (CVCVCV_#)
PL	Plokama	TamaPLo	TamakoPL
PR	Prokama	TamaPRo	TamakoPR
BL	Blokama	TamaBLo	tamakoBL
BR	Brokama	TamaBRo	TamakoBR
FL	Flamote	MoteFLa	MotesaFL

FR	Framote	MoteFRa	MotesaFR
VL	Vlamote	MoteVLa	
VR	Vramote	MoteVRa	MotesaVR
TR	Trosema	SemaTRo	SemagoTR
DR	Drosema	SemaDRo	SemagoDR
KL	Klanape	NapeKLa	NapetaKL
KR	Kranape	NapeKRa	NapetaKR
GL	Glanape	NapeGLa	NapetaGL
GR	Granape	napeGRa	NapetaGR
SP	Spijena	jenaSPi	JenakiSP
ST	Stodime	dimeSTo	DimeloST
SK	Skamose	moseSKa	ModenaSK
SL	Slotipa	TipaSLO	
SM	Smotipa	TipaSMO	TipakoSM
SN	Snotipa	tipaSNO	
SV	Svotipa	tipaSVO	
SF	Sfotipa	tipaSFO	

- Les hypothèses :

<b>Générale</b>	Le système phonographémique en L1 (japonais) influence la perception graphémique en L2 (français).
<b>De Travail</b>	Le recodage phonologique syllabique/moraïque des graphèmes français/japonais va être effectué en fonction du stock et des procédures de segmentations phonologiques japonais, sans variation phonético-phonologique, et ainsi entraîner une confusion syllabique importante.
<b>Opérationnelles</b>	1) La position du groupe cible graphémique en L2 n'influence pas de manière significative le taux d'épenthèses ; 2) La nature du groupe cible graphémique en L2 n'influence pas de manière significative le taux d'épenthèses.

Pour le reste, le test D est identique au test C2.

L'objectif principal de ce test est de récolter des informations sur la segmentation syllabique de non-mots « français » effectuée par nos sujets japonais lorsque les stimuli sont orthographiques, nécessitant donc au préalable un recodage phonologique. Ces informations seront ensuite comparées à celles obtenues au test C1.

#### **V.4. Test E (Groupe 2 – Bimodalité Audio-Visuelle)**

Le test E correspond, en partie au moins, à la fusion des tests C2 et D, puisque les non-mots, identiques à ceux utilisés pour la tâche C2 (CC et CvC) sont présentés en bimodalité audiovisuelle de manière synchrone. L'objectif principal de ce test est à terme de pouvoir comparer les taux de confusion syllabique phonologique selon la modalité : unimodalité visuelle, unimodalité auditive et bimodalité audiovisuelle. Une autre composante de la tâche E est la suivante : tandis qu'une partie des stimuli est composée de couples audiovisuels de longueur syllabique égale (par exemple : (/plamote/ (3 syllabes) ; "plamote" (3 syllabes)) l'autre partie est composée de couples divergents (par exemple : (/spinoka/ (3 syllabes) ; "supinoka" (4 syllabes))<sup>176</sup>. L'objectif est de déterminer, d'une part si les étudiants perçoivent ou non la différence, et d'autre part, dans le cas où la différence ne serait pas perçue, la modalité favorisée en tâche de comptage syllabique.

#### **Tâche**

Comptage syllabique en bimodalité audiovisuelle : les étudiants entendent un non-mot qu'ils lisent simultanément sur écran. Ils doivent ensuite entourer le nombre de cercles

---

<sup>176</sup> Comme nous l'avons indiqué plus haut la voyelle épenthétique /u/ du japonais correspond à une voyelle neutre de type schwa, et est sujette en japonais à un dévoisement fréquent. Il n'est donc pas impossible qu'il y ait discordance entre syllabes phonétiques et syllabes phonologiques, comme cela est parfois le cas en japonais avec la voyelle /u/ et en français avec le « e muet ». Dans notre cas, nous nous intéressons particulièrement aux syllabes phonologiques, mais pouvons également apprécier les données issues de ces écarts puisqu'en situation d'apprentissage la superposition (« mapping ») de ces deux plans constitue une étape cruciale dans l'apprentissage, que l'on pourrait apparenter à celle de correspondance entre input (stimulus) et output (représentation mentale).

correspondant au nombre de syllabes perçues, en privilégiant, dans une certaine mesure<sup>177</sup>, ce qu'ils entendent<sup>178</sup>. Parfois les stimuli auditifs et visuels sont divergents (CC versus CvC). Les étudiants doivent alors indiquer s'ils perçoivent la différence.

L'espace de réponse se présente ainsi :

1.	<b>OOO</b>	<b>O000</b>	<b>00000</b>	DIFFERENT
----	------------	-------------	--------------	-----------

### Stimuli

Les non-mots sont identiques à ceux utilisés dans la tâche C2 (unimodalité auditive). Seule leur combinaison, ainsi que le nombre de stimuli est différent. En voici les types ainsi que leur liste avant leur répartition aléatoire :

### Types :

<u>Combinaison identique</u>	Un mot est présenté auditivement (CC) et orthographiquement (CC)
<u>Combinaison croisée 1</u> (CvépC-Phonologique / CC-Orthographique)	Un mot est présenté auditivement (CvépC) et orthographiquement (CC)
<u>Combinaison croisée 2</u> (CC-Phonologique / CvépC-Orthographique)	Un mot est présenté auditivement (CC) et orthographiquement (CvépC).

### Liste :

Les groupes sont ordonnés selon l'ordre des combinaisons ci-dessus. Par exemple pour le groupe /pl/ : PL1 (combinaison identique), PuL (combinaison croisée 1) et PL2 (combinaison croisée 2).

<sup>177</sup> Voir les consignes de la tâche.

<sup>178</sup> Puisque nous nous intéressons à l'influence des représentations orthographiques sur la perception des représentations phonologiques.

<b>Groupes</b> <b>Consonantiques</b> <b>phonologiques sans</b> <b>(CC) et avec (CvC)</b> <b>épenthèse</b>	<b>POSITION</b>		
	<b>INITIALE</b> <b>(#_VCVCV)</b>	<b>MEDIALE</b> <b>pénultième</b> <b>(CVCV_V)</b>	<b>FINALE</b> <b>pénultième</b> <b>(CVCVCV_#)</b>
PL1	Plokama1	TamaPLo1	TamakoPL1
PuL	PuLokama	TamaPuLo	TamakoPuLu
PL2	Plokama2	TamaPLo2	TamakoPL2
PR1	Prokama1	TamaPRo1	TamakoPR1
PuR	PuRokama	TamaPuRo	TamakoPuRu
PR2	Prokama2	TamaPRo2	TamakoPR2
BL1	Blokama1	TamaBLo1	TamakoBL1
BuL	BuLokama	TamaBuLo	TamakoBuLu
BL2	Blokama2	TamaBLo2	TamakoBL2
BR1	Brokama1	TamaBRo1	TamakoBR1
BuR	BuRokama	TamaBuRo	TamakoBuRu
BR2	Brokama2	TamaBRo2	TamakoBR2
FL1	Flamote1	MoteFLa1	MotesaFL1
FuL	FuLamote	MoteFuLa	MotesaFuLu
FL2	Flamote2	MoteFLa2	MotesaFL2
FR1	Framote1	MoteFRa1	MotesaFR1
FuR	FuRamote	MoteFuRa	MotesaFuRu
FR2	Framote2	MoteFRa2	MotesaFR2
VL1	Vlamote1	MoteVLa1	MotesaVL1
VuL	VuLamote	MoteVuLa	MotesaVuLu
VL2	Vlamote2	MoteVLa2	MotesaVL2
VR1	Vramote1	MoteVRa1	MotesaVR1
VuR	VuRamote	MoteVuRa	MotesaVuRu

VR2	Vramote2	MoteVRa2	MotesaVR2
TR1	Trosema1	SemaTRo1	SemagoTR1
ToR	ToRosema	SemaToRo	SemagoToRu
TR2	Trosema2	SemaTRo2	SemagoTR2
DR1	Drosema1	SemaDRo1	SemagoDR1
DoR	DoRosema	SemaDoRo	SemagoDoRu
DR2	Drosema2	SemaDRo2	SemagoDR2
KL1	Klanape1	NapeKLa1	NapetaKL1
KuL	KuLanape	NapeKuLa	NapetaKuLu
KL2	Klanape2	NapeKLa2	NapetaKL2
KR1	Kranape1	NapeKRa1	NapetaKR1
KuR	KuRanape	NapeKuRa	NapetaKuRu
KR2	Kranape2	NapeKRa2	NapetaKR2
GL1	Glanape1	NapeGLa1	NapetaGL1
GuL	GuLanape	NapeGuLa	NapetaGuLu
GL2	Glanape2	NapeGLa2	NapetaGL2
GR1	Granape1	NapeGRa1	NapetaGR1
GuR	GuRanape	NapeGuRa	NapetaGuRu
GR2	Granape2	NapeGRa2	NapetaGR2
SP1	Spijena1	JenaSPi1	JenakiSP1
SuP	SuPijena	JenaSuPi	JenakiSuPu
SP2	Spijena2	JenaSPi2	JenakiSP2
ST1	Stodime1	DimeSTo1	DimeloST1
SuT	SuTodime	DimeSuTo	DimeloSuTo
ST2	Stodime2	DimeSTo2	DimeloST2
SK1	Skamose1	MoseSKa1	ModenaSK1
SuK	SuKamose	MoseSuKa	ModenaSuKu
SK2	Skamose2	MoseSKa2	ModenaSK2
SL1	Slotipa1	TipaSLO1	

SuL	SuLotipa	TipaSuLo	
SL2	Slotipa2	TipaSLO2	
SM1	Smotipa1	TipaSMO1	TipakoSM1
SuM	SuMotipa	TipaSuMo	TipakoSuMu
SM2	Smotipa2	TipaSMO2	TipakoSM2
SN1	Snotipa1	TipaSNO1	
SuN	SuNotipa	TipaSuNo	
SN2	Snotipa2	TipaSNO2	
SV1	Svotipa1	TipaSVO1	
SuV	SuVotipa	TipaSuVo	
SV2	Svotipa2	TipaSVO2	
SF1	Sfotipa1	TipaSFO1	
SuF	SuFotipa	TipaSuFo	
SF2	Sfotipa2	TipaSFO2	

L'organisation temporelle des stimuli était la suivante : BIP- (1 seconde)- XXX-(9 secondes pour répondre)-BIP-(1 seconde)-XXX-etc. Chaque non-mot était présenté en double modalité en synchronie : à chaque « BIP » l'enseignant projetait le non-mot orthographique, et le retirait aussitôt après avoir entendu ce non-mot prononcé, de manière à ce que les deux formes du stimulus soient présentées simultanément et pendant 2 secondes environ.

### **Variables**

<b>Variable Indépendante</b>	<b>Type</b>	<b>Nom Excel</b>	<b>Code</b>
Position des groupes	Qln	position	1=initiale (syllabe initiale #__V), 2= avant-dernière position (syllabe finale __V#), 3=dernière position (syllabe finale V__#)



Famille générale des groupes	QIn	famigcc	1=Obs+Liqu, 2=Obs+NonLiqu, 3=Sib+Stop, 4=Sib+NonStop
Famille des groupes	QIn	familcc	1=Obs+Lat, 2=Obs+Rhot, 3=Obs+Nas, 4=Obs+Cont, 5=Obs+Sib, 6=Obs+Stop, 7=Sib+Stop, 8=Sib+Nas, 9=Sib+Cont, 10=Sib+Lat
Sous-famille des groupes	QIn	sfamicc	1=Stop+Lat, 2=Cont+Lat, 3=Stop+Rhot, 4=Cont+Rhot, 5=Obs+Nas, 6=Obs+Cont, 7=Obs+Sib, 8=Obs+Stop, 9=Sib+Stop, 10=Sib+Nas, 11=Sib+Cont, 12=Sib+Lat
Nature des groupes	QIn	naturecc	14=pl, 15=pR, 16=bl, 17=bR, 18=fl, 19=fR, 20=vl, 21=vR, 22=tR, 23=dR, 24=kl, 25=kR, 26=gl, 27=gR, 28=sp, 29=st, 30=sk, 31=sl, 32=sm, 33=sn, 34=sv, 35=sf
Type de sonorité pour le groupe CC	QIn	typsono	1= Montant (M), 2= Plateau (P), 3= Descendant (D)
Ecart de sonorité entre les deux consonnes du groupe	Qlo	ecarsono	0=0, 1=1, 2=2, 3=3, 4=4, 5=5, 6=6
Nature des couples de stimuli audio/visuel	QIn	natcoupl	1=stimuli identiques, 2=stimuli divergents
Nombre de syllabes stimuli visuelles	QIn	nomsylvi	1=3 syllabes (CC), 2=4 syllabes (CvC), 3=5 syllabes (CvCv)
Nombre de syllabes stimuli audio	QIn	nombsylau	1=3 syllabes (CC), 2=4 syllabes (CvC), 3=5 syllabes (CvCv)

<b>Variable Dépendante</b>	<b>Type</b>	<b>Nom Excel</b>	<b>Code</b>
Nombre de syllabes comptées	Qlo	nombsylc	1=3 syllabes (aucune épenthèse), 2=4 syllabes (une épenthèse), 3=5 syllabes (deux épenthèses)
Confusion syllabique phonologique	Qln	confsyll	1=ooui, 2=non
Perception de la différence stimuli audio/visuel	Qln	percdiff	1=ooui, 2=non

### Hypothèses

<b>Générale</b>	Le système phonographémique en L1 (japonais) influence la perception phonographémique en L2 (français).
<b>De Travail</b>	Notre hypothèse est que, en situation de comptage syllabique et de présentation bimodale, la modalité visuelle sera privilégiée, impliquant une interprétation graphémique selon le système phonographémique japonais et ainsi un recodage phonologique conforme au système japonais. Certaines combinaisons bimodales seraient donc plus propices à une confusion syllabique que d'autres.
<b>Opérationnelles</b>	A) <u>Lorsque les longueurs syllabiques des stimuli auditifs et visuels sont identiques :</u> 1) La position du groupe cible en L2 n'influence pas de manière significative le taux de confusion syllabique. 2) La nature du groupe cible en L2 n'influence pas de manière

	<p>significative le taux de confusion syllabique.</p> <p><u>B) Lorsque les longueurs syllabiques des stimuli auditifs et visuels sont différentes :</u></p> <ol style="list-style-type: none"><li>1) La position du groupe cible en L2 n'influence pas de manière significative le taux de discrimination phonographémique.</li><li>2) La nature du groupe cible en L2 n'influence pas de manière significative le taux de discrimination phonographémique.</li><li>3) La longueur syllabique du non-mot n'influence pas le taux de discrimination phonographémique : les taux de discrimination phonémique correcte sont semblables lorsque le stimulus visuel est plus long que le stimulus auditif et inversement.</li><li>4) Lorsqu'il y a discrimination : le taux de discrimination phonographémique correct est supérieur au taux de discrimination phonographémique erronée.</li><li>5) Lorsqu'il n'y a pas discrimination : c'est le mot graphémique (et donc la modalité visuelle) qui est privilégié pour le comptage syllabique phonologique.</li></ol> <p><u>C) Globalement :</u></p> <ol style="list-style-type: none"><li>1) Couple ([CC], "CC") : taux de confusion syllabique phonologique élevé ;</li><li>2) Couple ([CC], "CvC") : taux de confusion syllabique phonologique élevé ;</li><li>3) Couple ([CvC], "CC") : taux de confusion syllabique phonologique faible<sup>179</sup>.</li><li>4) Si le taux de discrimination correcte est élevé, alors le taux de confusion syllabique est élevé (conflit procédural d'ordre attentionnel) ;</li></ol>
--	---

---

<sup>179</sup> Nous n'avons pas testé le cas de figure ([CvC], "CvC").

	si le taux de discrimination correcte est faible, alors le taux de confusion syllabique est faible.
--	---

### **V.5. Méta-test C2-D-E (Groupe 2 – Effet de la Modalité)**

Les tests C2-D-E constituent en effet un « méta-test » en ce qu'ils permettent de comparer trois types de présentation pour les mêmes non-mots et, dans une certaine mesure, pour le même type de tâche.

**Variables Indépendante** : Modalité de présentation.

**Stimuli** : Les non-mots du type /CC/ composant les stimuli (c'est-à-dire ceux du test D, moins les distracteurs).

**Hypothèses** : En tâche de segmentation, syllabique, si l'on suppose que la modalité visuelle (qui impliquerait l'activation de représentations phonologiques via le système phonographémique, système fixe) est privilégiée par rapport à la modalité auditive (qui impliquerait l'activation de représentations phonologiques via le système phonético-phonologique, système variable), on peut s'attendre à ce que, pour les stimuli [CC] communs aux trois tests, le taux de confusion syllabique (correspondant pour nous au taux d'épenthèses) soit :

- Plus élevé dans le test D que dans le test C2 ;
- Moins élevée (ou identique) dans le test E que dans le test D, lorsque les stimuli audiovisuels y sont de longueur syllabique égale ([CC] et "CC") (puisque l'input phonétique ([CC] → /C(v)C/) et l'input phonographémique ("CC" → /CvC/) sont potentiellement divergents), mais plus élevé que dans le test C2 ;
- Plus élevé (ou identique) dans le test E que dans les tests D et C2 lorsque les stimuli audiovisuels y sont de longueurs différentes (puisque la modalité visuelle ("CvC") domine la modalité auditive ([CC])).

## **VI. Variables et considérations supplémentaires**

Comme dans tout test psycholinguistique, les résultats obtenus sont toujours le fruit des interactions de nombreuses variables, certaines contrôlées, d'autres non. Il convient donc de les prendre en compte, sinon dans l'analyse, au moins dans l'interprétation.

- **Effet des variables descriptives**

Pour chaque expérience, il conviendrait d'évaluer l'influence des variables descriptives sur la (les) variable(s) dépendante(s). Se poserait alors la question d'équivalence des groupes. Dans notre cas, deux variables descriptives pourraient particulièrement nous intéresser : d'une part le sexe, et d'autre part le degré d'expertise en français (via les variables « Spécialité en français » et surtout « Durée d'étude du français ») qui nous permettrait de distinguer deux groupes (« Experts » et « Novices ») de manière à évaluer l'hypothèse de travail suivante : les « experts » obtiennent de meilleurs résultats (en termes d'évaluation de l'apprentissage du français) que les « novices » pour les tâches proposées. Enfin, il conviendrait de prendre en compte l'appartenance académique des étudiants, ceux-ci provenant de différentes universités, avec des profils langagiers différents.

- **Effet des variables parasites**

Il conviendrait également pour chaque expérience de recenser la plupart des variables parasites et d'en évaluer l'influence sur nos résultats. Parmi celles-ci on peut mentionner *l'ordre de présentation* des réponses possibles sur les feuillets des étudiants (en particulier pour les tests A et F), *la qualité du matériel* audiovisuel utilisé et *la position* spatiale des étudiants dans la salle de passation des tests. Mais nous devons surtout nous arrêter sur certains aspects des stimuli utilisés, notamment :

- La similarité phonologique des non-mots avec des mots japonais ou d'autres langues connus des étudiants, pouvant donner lieu à une reconnaissance lexicale ou morphémique ;

- La similarité orthographique des non-mots avec des mots japonais ou d'autres langues connus des étudiants, pouvant donner lieu à une reconnaissance lexicale ou morphémique ;
- La fréquence de co-occurrence syllabique et segmentale en japonais des syllabes/mores et segments utilisés dans la composition des non-mots ;
- Le degré de marque (universel ou en japonais) des segments utilisés dans la composition des non-mots ;
- La fréquence d'occurrence des graphèmes alphabétiques en japonais ;
- Le caractère phonologiquement homorganique ou non des consonnes associées pour former les groupes consonantiques.

Il convient également de signaler, outre ces effets parasites possibles, l'existence de problèmes méthodologiques et techniques rencontrés au cours de la passation ou de l'analyse des résultats, problèmes inattendus et pourtant difficilement évitables dans ce type d'expériences.

- **Problèmes méthodologiques et techniques**

Malgré nos efforts et notre attention, plusieurs problèmes ont surgi :

- Non-respect des consignes : malgré les explications et les exemples donnés avant le début des tests, il semble possible que certains n'aient pas immédiatement parfaitement compris les consignes. En particulier, dans le Groupe 2, pour le test C2, les sujets n° 43 et n° 58 n'ont pas correctement rempli leur feuillet de réponse, tandis que dans le Groupe 1, pour le test C1, le sujet 119 n'a produit aucune réponse. De la même manière, pour les tests A et F, dans le groupe C1, le sujet 64 n'a pas correctement rempli son feuillet de réponse. Il est possible que l'anonymat sous lequel ont été passés les tests a encouragé certains à ne pas respecter les consignes ou tout simplement à ne pas répondre. Ces cas sont cependant tout à fait exceptionnels dans nos groupes de sujets. Il convient de mentionner un manquement particulièrement important : dans le test E, lorsque les stimuli audiovisuels étaient

divergents, les sujets étaient censés indiquer s'ils percevaient une différence ou non, mais également indiquer dans tous les cas le nombre de syllabes « entendues ». Cette consigne n'a globalement pas été respectée, puisque la plupart des sujets se sont contentés d'indiquer une différence perçue, les cas échéant. Ce comportement, *a priori* gênant pour notre étude s'est finalement révélé intéressant, puisqu'il pourrait probablement permettre de mettre au jour un partage attentionnel dû à la complexité de la tâche.

- Stimuli manquants : en raison de problèmes techniques passés inaperçus jusqu'au dernier moment, certains stimuli n'ont pu être présentés aux sujets du Groupe 2. Pour le test C2, il s'agit des stimuli audio [91, 99] et [102, 110]. Pour le test D, il s'agit des stimuli visuels [63, 66].
  
- Manque d'entraînement et problème de fatigue : au vu de certains résultats, il semble possible que certains sujets aient manqué d'entraînement aux tâches proposées et se soient fatigués ou démotivés vers la fin de tests. Nous aurions donc dû mettre en place une période d'entraînement, ainsi qu'une série de distracteurs au début (manque d'habitude) et à la fin (fatigue et démotivation) de chaque test, afin d'obtenir des réponses convenables pour les stimuli qui nous intéressent. Leur absence n'est due qu'à des contraintes matérielles (longueur des tests) et temporelles (disponibilité des sujets).
  
- Remarques des enseignants : Pour le Groupe 1 en particulier, l'enseignant ayant fait passer les tests nous a fait remarquer certains problèmes :
  - L'absence de pause supplémentaire régulière dans l'enregistrement destinée à laisser le temps aux étudiants de changer de page de réponse, d'autant que certains pourraient ne pas en changer spontanément, attendant les directives de l'enseignant ;

- La présence d'éléments linguistiques inutiles dans l'enregistrement pouvant troubler les étudiants, à savoir le nom de chaque test (par exemple, « tâche C1 ») ;
  - La complexité de certaines consignes, en particulier celle de la tâche F qui, par souci de clarté, peut s'avérer trop complexe ou trop longue pour les étudiants et ainsi rendre imprécise leur représentation de la tâche à accomplir ;
  - La répartition ergonomique des espaces de réponse sur plusieurs feuillets pouvant conduire les étudiants à hésiter (une partie de la réponse sur une page et l'autre partie sur la page suivante).
- Groupe de référence : nous notons qu'il conviendrait bien évidemment de faire passer des tests identiques à un échantillon approprié de locuteurs français, afin de pouvoir comparer les résultats de nos sujets à ceux d'un groupe de référence de locuteurs natifs.
  - Répartition des échantillons : Dans les tâches C2, D et E en particulier, nous avons tenu à respecter la distribution « naturelle » des groupes de consonnes en français telle que donnée par Dell (1995). De ce fait, la répartition des échantillons a certainement manqué d'homogénéité vis-à-vis de certains critères « secondaires » (le critère primaire étant l'identité spécifique de chaque groupe consonantique), rendant nécessaire l'adoption d'une certaine distance critique quant aux résultats de certaines analyses statistiques inférentielles. A titre d'exemple, dans la tâche E, le critère du type de sonorité est en effet inégalement représenté, puisque 152 items possédaient le type « Montant », 5 le type « Plateau » et 26 le type « Descendant ». Cela ne remet cependant nullement en cause les résultats descriptifs.





## **Partie Expérimentale III :**

### **Tests - Résultats**

## Table des Matières Partielle

I. RESULTATS DES TESTS .....	- 753 -
<i>I.1. Considérations statistiques</i> .....	- 753 -
<i>I.2. Axe segmental</i> .....	- 758 -
TACHE A .....	- 758 -
A) Résultats descriptifs .....	- 758 -
A) 1) Axe syntagmatique .....	- 759 -
A) 2) Axe paradigmatique.....	- 762 -
B) Résultats inférentiels : analyses de variance (Test F) .....	- 767 -
C) Résultats complémentaires : choix de graphème existant/absent en L1 (41 sujets) .....	- 768 -
D) Résumé .....	- 771 -
TACHE F .....	- 773 -
A) Résultats descriptifs .....	- 773 -
A) 1) Confusion phonographémique (premier stimulus) .....	- 773 -
A) 1) 1. Axe syntagmatique .....	- 774 -
A) 1) 2. Axe paradigmatique.....	- 776 -
A) 1) 3. Variable liée à la tâche : nature du couple (identique/différent).....	- 780 -
A) 2) Confusion phonologique (entre les deux stimuli auditifs) .....	- 781 -
A) 2) 1. Axe syntagmatique .....	- 782 -
A) 2) 2. Axe paradigmatique.....	- 784 -
A) 2) 3. Variables liées à la tâche : nature du couple (identique/différent).....	- 788 -
B) Résultats inférentiels : analyses de variance (Test F) .....	- 789 -
D) Résumé .....	- 790 -
COMPARAISON TACHES A ET F.....	- 792 -
A) Résultats globaux .....	- 792 -
B) Position .....	- 792 -
C) Nature.....	- 794 -
D) Résultats inférentiels.....	- 797 -
E) Résumé.....	- 798 -

1.3. Axe syllabique.....	- 799 -
TACHE C1.....	- 799 -
A) Résultats descriptifs.....	- 799 -
A) 1) Axe syntagmatique .....	- 800 -
A) 2) Axe paradigmatique.....	- 805 -
B) Résultats inférentiels : analyse de variance (Test F) : .....	- 816 -
C) Résumé.....	- 819 -
C) 1) Résumé détaillé .....	- 819 -
C) 2) Résumé essentiel .....	- 826 -
TACHE C2.....	- 827 -
A) Résultats descriptifs.....	- 827 -
A) 1) Axe syntagmatique .....	- 828 -
A) 2) Axe paradigmatique.....	- 831 -
B) Résultats inférentiels : analyse de variance (Test F).....	- 851 -
C) Résumé.....	- 854 -
C) 1) Résumé Détaillé .....	- 854 -
C) 2) Résumé Essentiel .....	- 862 -
COMPARAISON TACHES C1 ET C2 .....	- 863 -
TACHE D .....	- 871 -
A) Résultats descriptifs.....	- 871 -
A) 1) Axe syntagmatique .....	- 871 -
A) 2) Axe paradigmatique.....	- 873 -
A) 3) Variation entre sujets .....	- 879 -
B) Résumé Essentiel .....	- 880 -
COMPARAISON TACHES C2 ET D .....	- 881 -
TACHE E.....	- 890 -
CAS 1 - STIMULI AUDITIFS ET VISUELS CONVERGENTS .....	- 890 -
A) Résultats descriptifs.....	- 890 -
A) 1) Axe syntagmatique .....	- 891 -
A) 2) Axe paradigmatique.....	- 892 -
B) Résultats inférentiels.....	- 903 -
C) Résumé Essentiel .....	- 903 -

CAS 2 - STIMULI AUDITIFS ET VISUELS DIVERGENTS .....	- 905 -
A) Modalité dominante en tâche de comptage syllabique.....	- 905 -
B) Perception de la divergence intermodale.....	- 909 -
1) Résultats Globaux .....	- 909 -
2) Influence de la position.....	- 909 -
3) Influence de la nature.....	- 910 -
4) Influence de la structure/longueur.....	- 918 -
Comparaison des performances : Comptage syllabique vs. Perception de la divergence inter-stimuli. -	921
-	
COMPARAISON DES TACHES C2 (AUDITIVE), D (VISUELLE) ET E (AUDIOVISUELLE) .....	- 924 -
1. Résultats globaux .....	- 924 -
2. Position .....	- 924 -
3. Famille Générale .....	- 926 -
4. Famille .....	- 927 -
5. Sous-Famille.....	- 928 -
6. Nature.....	- 929 -
7. Type de Sonorité .....	- 930 -
8. Ecart de Sonorité.....	- 931 -

## **I. Résultats des Tests**

Nous présentons ci-dessous les résultats de nos cinq tests<sup>180</sup>.

### **I.1. Considérations statistiques**

Lorsque l'on cherche à analyser les données quantitatives obtenues, on peut d'une part effectuer une analyse descriptive (par exemple calcul des moyennes et des écarts-types), et d'autre part effectuer une analyse inférentielle (mise à jour de corrélation ou d'indépendance des variables). Keller (1976) distingue ainsi deux types d'analyse des résultats :

- Les analyse statistiques de type *descriptif* (« Une bonne analyse statistique décrivant un ensemble de données doit donc faire état et de la moyenne (ou des modes) et de l'écart-type (indicateur de la variabilité) » (p. 48). Les statistiques descriptives servent essentiellement à présenter la distribution des données avec des indices de tendance centrale, comme la moyenne, et des indices de dispersion, comme l'écart-type.
- Les analyses statistiques de type *inférentiel* ou déductif (« dans la plupart des cas, les chercheurs veulent savoir : 1) si deux variables indépendantes se *distinguent* l'une de l'autre en termes de leur influence sur une variable dépendante ; et 2) si deux variables dépendantes ou indépendantes sont en *corrélation*, c'est-à-dire si elles ont une relation systématique ou non. » (p. 51)). Les statistiques inférentielles reposent sur des tests statistiques (par exemple de comparaison de moyennes d'échantillon, comme l'analyse de variance) qui évaluent la probabilité qu'une différence de moyenne ou qu'une corrélation entre variables soit due au hasard.

---

<sup>180</sup> Nous tenons à très vivement remercier Arnaud Gélas pour sa précieuse et indispensable contribution à l'analyse statistique de nos données.

A ces deux grandes catégories, il est possible d'ajouter celle des statistiques dites *corrélationnelles*, qui permettent de quantifier les relations possibles entre variables, mais qui est parfois incluse dans la catégorie inférentielle.

Trois types d'analyse sont ensuite envisagés :

- L'analyse *univarié* (descriptive) : elle porte sur une seule variable et sert à en dresser un portrait statistique. Il s'agit généralement de variables dépendantes, mais pas uniquement.
- L'analyse *bivariée* (comparée) : elle porte sur deux variables (généralement une indépendante et une dépendante) et permet de déterminer s'il existe une relation entre celles-ci qui ne soit pas due au hasard (en particulier une influence de la variables indépendante sur la variables dépendante).
- L'analyse *multivariée* (comparée) : le principe est le même que celui de l'analyse bivariée, mais porte sur un plus grand nombre de variables.

Dans notre cas, trois types de variables indépendantes peuvent être pris en compte :

- Linguistique : lié au matériel verbal composant les stimuli ainsi qu'aux systèmes langagiers en présence :
  - Position des groupes - cibles (axe syntagmatique)
  - Nature des groupes - cibles (axe paradigmatique)
- Psycholinguistique situationnel : lié aux tâches expérimentales et aux processus cognitifs associés :
  - Modalité de présentation
  - Tâche

- Psycholinguistique individuel : lié aux variables descriptives des sujets (sexe, âge, degré d'expertise) :
  - Degré d'expertise
  - Sexe

Ces variables ont déjà été présentées et décrites plus haut, dans ce qui constitue la première étape de l'analyse statistique des données quantitatives :

1. Analyse descriptive des variables indépendantes (univariée).
2. Analyse descriptives des variables dépendantes (univariée).
3. Analyse comparée des variables dépendantes en fonction des conditions des variables indépendantes (bi-/ multivariée).

En ce qui concerne la troisième étape, le choix du test statistique approprié dépend essentiellement de la nature (quantitative ou qualitative) des variables en jeu. Dans notre cas, les variables indépendantes sont essentiellement qualitatives (par exemple nature, position, modalité) tandis que les variables dépendantes sont essentiellement quantitatives (par exemple taux d'épenthèse, taux de confusion phonographémique). Dans ce cas, il convient d'effectuer une analyse de variance (Anova) à l'aide de tests d'inférence comme le test F, par exemple. Les résultats de ces tests peuvent être plus ou moins significatifs, c'est-à-dire qu'ils permettent de savoir si la relation établie entre les deux variables, dans l'échantillon de la population qui a passé les épreuves, est attribuable au hasard ou si elle constitue une relation stable, vraisemblablement pour toute la population représentée par l'échantillon. Ceci s'évalue via une comparaison entre les moyennes des résultats obtenus pour des valeurs différentes des variables en question.

Pour le test F par exemple, l'analyse de variance fournit trois valeurs particulières à partir des données qui lui ont été fournies : il s'agit des valeurs F, F critique et p, celle-ci correspondant à la probabilité que la différence des moyennes évoquée plus haut soit



due au hasard. En général, le seuil à partir duquel le hasard est exclu est 5%. Plus la valeur de  $p$  est faible, plus l'hypothèse nulle (selon laquelle les différences observées sont dues au hasard) est rejetée et plus les résultats semblent significatifs. Ainsi, dans le test  $F$ , si  $F > F$  critique avec  $p < 0,05$ , on peut considérer que les résultats obtenus sont significatifs, c'est-à-dire qu'il existe une relation non-aléatoire entre les deux variables en question. La procédure à suivre est similaire dans le cas d'analyse bi-variée, lorsque l'on cherche à évaluer l'existence d'une influence possible de deux variables indépendantes ou de leur interaction sur une variable dépendante. Il est ensuite possible d'effectuer une vérification de cette analyse à l'aide d'un autre test comme le test  $t$  de Student, utile pour les comparaisons de moyennes.

Les résultats statistiques visent ainsi généralement à *invalidier*, parfois valider, les hypothèses opérationnelles propres à chaque test et présentées plus haut. Ces hypothèses opérationnelles sont alors formulées en termes statistiques (taux, probabilité, coefficient) conduisant aux hypothèses statistiques du protocole.

Etant donné la quantité de tests, il nous a semblé utile d'intégrer quelques remarques au fur et à mesure de la présentation de ceux-ci. Tous les résultats numériques présentés seront arrondis à une ou deux décimales. Pour chaque test, la présentation se fera en trois temps : résultats descriptifs (analyse univariée), résultats inférentiels lorsque ceux-ci sont pertinents (analyse bi- ou multivariée) et résumé.

Il nous faut souligner qu'une analyse descriptive stimulus par stimulus permet de repérer ceux d'entre eux qui ont particulièrement « posé problème » aux sujets. Elle permet en particulier d'évoquer la possibilité de l'influence de facteurs parasites situationnels pendant la passation (par exemple : le bruit d'un téléphone portable qui aurait détourné l'attention de tous les sujets à un moment donné. Cet exemple est peu probable dans notre cas puisque les sujets étaient répartis dans des classes différentes. Les facteurs situationnels semblent donc exclus, mais cela n'est pas le cas de facteurs parasites propres au matériel de passation, comme par exemple un défaut d'enregistrement passé inaperçu). Cependant, une telle analyse n'apporte que peu

d'information dans un premier temps. Elle peut en revanche s'avérer utile *a posteriori* si les premières analyses ne sont pas satisfaisantes. Ainsi, dans le cas de la tâche C1, nous avons relevé, à titre illustratif, les stimuli qui ont suscité plus de 70% d'erreurs chez tous les sujets. Ces erreurs correspondent à chaque fois à des épenthèses.

Stimuli			Taux d'erreur moyen
Code	Non-mot auditif	Groupe Consonantique	
14	/ptideso/	/pt/	84, 87%
15	/ktesagi/	/kt/	86, 55%
22	/ksesagi/	/ks/	90, 76%
24	/kadagikm/	/km/	81, 36%
27	/mezakoft/	/ft/	72, 27%
28	/dagikte/	/kt/	83, 19%
34	/dekasopt/	/pt/	71, 43%
36	/tagikse/	/ks/	85, 71%
39	/nalogikt/	/kt/	75, 63%
42	/magikve/	/kv/	87, 39%
43	/tmetosa/	/tm/	79, 83%
56	/topisatm/	/tm/	79, 83%
58	/kvesagi/	/kv/	78, 99%
66	/kmesagi/	/km/	92, 44%
71	/knesagi/	/kn/	89, 92%
72	/pagikne/	/kn/	73, 11%

Les taux d'épenthèse pour ces stimuli sont effectivement remarquablement élevés.

## **I.2. Axe segmental**

Nous présentons ici les résultats relatifs au traitement des relations phonographémiques segmentales attachées aux segments /R, l, b, v/, soit les résultats des tâches A et F.

### **Tâche A**

#### **A) Résultats descriptifs**

Chaque consonne à l'étude a été présentée 12 fois, également réparties sur les trois positions.

#### **Résultats Globaux**

Le taux moyen de confusion phonographémique, pour l'ensemble des segments consonantiques concernés, est le suivant :

	<b>Taux de confusion phonographémique consonantique</b>
<b>Moyenne</b>	29, 37%
<b>Ecart-type</b>	10, 9%

#### **Remarques :**

- Le taux est légèrement inférieur au tiers des réponses, ce qui est non-négligeable.

#### **Résumé Essentiel Résultats Globaux**

**On observe 29% de confusion phonème-graphème au sein des couples de consonnes (B, V) et (R, L).**

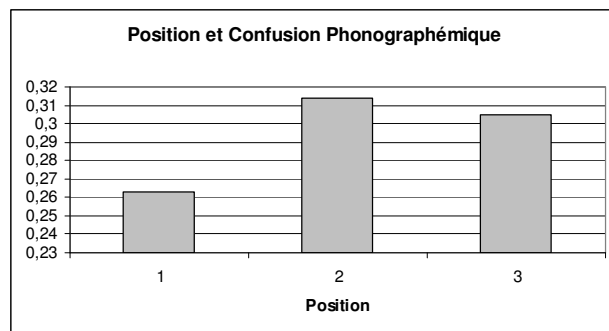
A) 1) Axe syntagmatique

Les nombres de stimuli, selon la position, étaient les suivants :

<b>Position</b>	<b>1</b>	<b>2</b>	<b>3</b>
<b>Nombre de stimuli</b>	16	16	16

Si l'on considère l'ensemble des quatre consonnes à l'étude, voici ce que l'on obtient :

	<b>Taux de confusion phonographique consonantique</b>		
<b>Position</b>	<b>1 (Attaque Initiale)</b>	<b>2 (Attaque Finale)</b>	<b>3 (Coda Finale)</b>
<b>Moyenne</b>	26, 3%	31, 35%	30, 47%
<b>Ecart-type</b>	14, 33%	13, 63%	14, 16%



Remarques :

- Peu de différences entre les positions 2 et 3.
- Le taux de confusion est le plus faible en position 1.

Cependant, si l'on considère chaque consonne en détail, on s'aperçoit que l'influence de la position varie davantage :

		<b>Taux de confusion phonographémique</b>		
<b>Consonne</b>	<b>Position</b>	<b>1</b>	<b>2</b>	<b>3</b>
		<b>(Attaque Initiale)</b>	<b>(Attaque Finale)</b>	<b>(Coda Finale)</b>
<b>B</b>	<b>Moyenne</b>	45, 21%	36, 25%	52, 92%
	<b>Ecart-type</b>	23, 44%	25, 9%	30, 26%
<b>V</b>	<b>Moyenne</b>	15%	34, 17%	27, 08%
	<b>Ecart-type</b>	21, 59%	27, 3%	24, 16%
<b>R</b>	<b>Moyenne</b>	26, 67%	27, 5%	16, 46%
	<b>Ecart-type</b>	28, 94%	27, 77%	22, 57%
<b>L</b>	<b>Moyenne</b>	18, 33%	27, 5%	25, 41%
	<b>Ecart-type</b>	21, 18%	22, 3%	28, 25%

Remarques :

- L'influence de la position est plus sensible pour certaines consonnes.
- On distingue 4 groupes :
  - 1) B Coda Finale et Attaque Initiale (53%-45%)
  - 2) B/V Attaque Finale (36%-34%)
  - 3) R Attaques Initiale et Finale, L Attaque et Coda Finales et V Coda Finale (27%-25%)
  - 4) L Attaque Initiale, R Coda Finale et V Attaque Initiale (18%-15%)
- Les extrêmes sont donc : B Coda finale (53%) > R Coda finale (16%) et V Attaque initiale (15%)

### **Résumé Essentiel Axe Syntagmatique**

**Globalement** : Peu de différences entre les trois positions, mais moins de confusions en attaque initiale (- 5%).

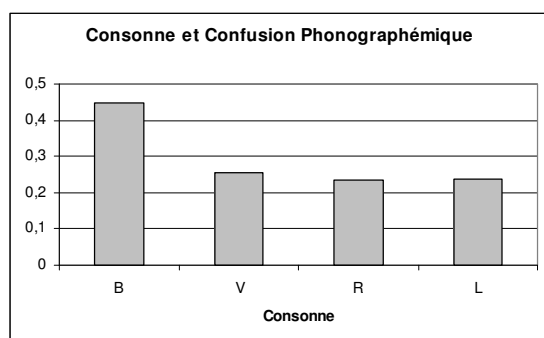
**Spécifiquement** : L'influence de la position varie selon la nature de la consonne.

- **B** : Coda finale (53%) > Attaque initiale (45%) > Attaque finale (36%)
- **R** : Attaque initiale et Attaque finale (27%) > Coda finale (16%)
- **V** : Attaque finale (34%) > Coda finale (27%) > Attaque initiale (15%)
- **L** : Attaque finale (27%) > Coda finale (25%) > Attaque initiale (18%)

A) 2) Axe paradigmatique

**α) Nature de la consonne**

Consonne	Taux de Confusion Phonographémique	
	Moyenne	Ecart Type
<b>B</b>	44, 8%	17, 8%
<b>V</b>	25, 4%	15, 2%
<b>R</b>	23, 5%	19, 7%
<b>L</b>	23, 7%	16, 8%



Remarques :

- Le classement est le suivant : B (45%) > V (25%) > L et R (24%)
- Les taux de V, R et L sont quasi-identiques (24, 5%)
- Dans 45% des cas : /b/ est transcrit <V>
- Dans 24% des cas : /v/ est transcrit <B>, /l/ est transcrit <R> et /R/ est transcrit <L>

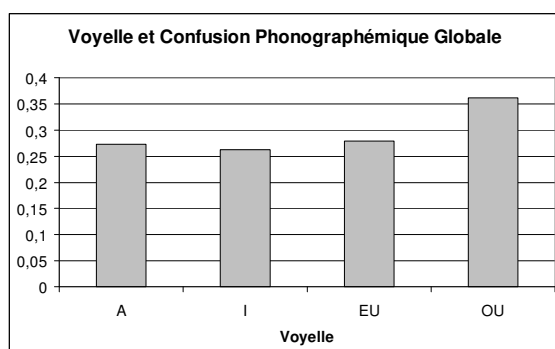
**La nature de la consonne influence le taux de confusion : B (45%) entraîne presque deux fois plus de confusion phonographémique que V, L et R, qui ont des taux similaires (24%).**

### **β) Nature de la voyelle**

#### Influence Globale :

Lorsque l'on considère l'ensemble des quatre consonnes à l'étude, l'influence globale de la voyelle tautosyllabique sur le taux de confusion s'exprime ainsi :

<b>Taux de Confusion Consonantique Phonographémique</b>		
<b>Voyelle</b>	<b>Moyenne</b>	<b>Ecart Type</b>
<b>A</b>	27, 29%	14, 57
<b>I</b>	26, 25%	14, 52%
<b>EU</b>	27, 85%	15, 26%
<b>OU</b>	36, 11%	13, 97%



#### Remarques :

- Le classement est le suivant : OU (36%) > EU (28%) > A (27%) > I (26%)
- Peu de différences, sauf pour la voyelle OU (+ 10% par rapport à I et + 8% par rapport à EU)

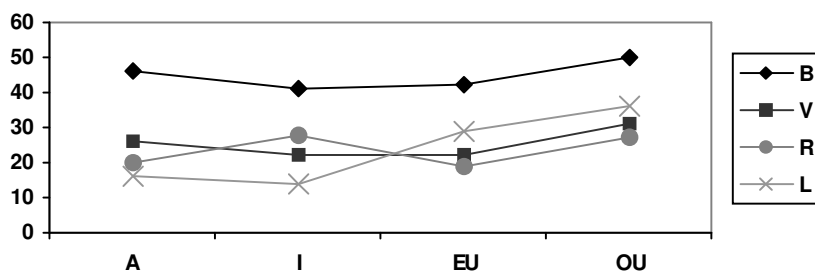
**Globalement, la nature de la voyelle adjacente semble influencer le taux de confusion consonantique phonème-graphème : OU (36%) déclenche plus de confusion que EU, A et I qui ont des taux similaires (27%).**



Influence Spécifique :

Cependant, lorsque l'on regarde l'influence de la voyelle adjacente pour chaque consonne, voici ce que l'on obtient :

		<b>Taux de Confusion Consonantique Phonographémique</b>			
<b>Consonne</b>	<b>Mesure</b>	<b>A</b>	<b>I</b>	<b>EU</b>	<b>OU</b>
<b>B</b>	Moyenne	46, 1%	40, 8%	42, 2%	50%
	Ecart-type	30, 3%	31%	32%	27, 3%
<b>V</b>	Moyenne	26, 4%	22, 5%	21, 7%	31, 1%
	Ecart-type	22, 8%	25, 2%	25%	27, 6%
<b>R</b>	Moyenne	20, 5%	28%	18, 6%	26, 9%
	Ecart-type	26, 3%	26, 6%	26, 6%	29, 4%
<b>L</b>	Moyenne	16, 1%	13, 6%	28, 9%	36, 4%
	Ecart-type	25, 2%	21, 4%	28%	29%



Remarques :

- L'influence de la voyelle varie selon la consonne.
- Le classement vocalique obtenu est le suivant :
  - BOU (50%) > BA (46%) > BEU (42%) > BI (41%)
  - VOU (31%) > VA (26%) > VI, VEU (22%)
  - RI (28%) > ROU (27%) > RA (20%) > REU (19%)

- LOU (36%) > LEU (29%) > LA (16%) > LI (14%)
  
- Le classement consonantique obtenu est le suivant :
  - BOU (50%) > LOU (36%) > VOU (31%) > ROU (27%)
  - BEU (42%) > LEU (29%) > VEU (22%) > REU (19%)
  - BA (46%) > VA (26%) > RA (20%) > LA (16%)
  - BI (41%) > RI (28%) > VI (22%) > LI (14%)
  
- On peut donc remarquer que :
  - BOU > BEU, VOU > VEU, ROU > REU et LOU > LEU
  - BA > BI, VA > VI et LA > LI, mais RI > RA

Les extrêmes sont :

- B et V : Max= OU ; Min= EU-I
- R : Max= I-OU ; Min= A-EU
- L : Max= OU ; Min= I

**De manière spécifique, la nature de la voyelle adjacente influence de manière différente le taux de confusion consonantique phonème-graphème selon la consonne.**

**En particulier :**

- **Pour A : B > V > R > L**
- **Pour I : B > R > V > L**
- **Pour OU et EU : B > L > V > R**

### **Résumé Essentiel Axe Paradigmatique**

**Consonne** : la nature de la consonne influence le taux de confusion : B (45%) entraîne presque deux fois plus de confusion que V, L et R (24%)

**Environnement vocalique** :

- **Globalement** : influence non-négligeable. Plus de confusions avec OU qu'avec EU, A et I.
- **Spécifiquement** : varie selon la nature de la consonne.

**Il semble donc y avoir une influence de l'interaction entre consonne et voyelle sur le taux de confusion**

## **B) Résultats inférentiels : analyses de variance (Test F)**

### **Univariée :**

Les trois variables indépendantes (position, nature consonantique, nature vocalique) semblent influencer significativement ( $p < 0,001$ ) la variable dépendante (confusion phonographémique).

### **Bivariée :**

D'après les analyses statistiques, il ne semblerait y avoir d'interaction significative ni entre la position et la voyelle ( $p = 0,69$ ), ni entre la position et la consonne ( $p = 0,68$ ). En revanche il semblerait y avoir une interaction significative entre la consonne et la voyelle ( $p < 0,001$ ).

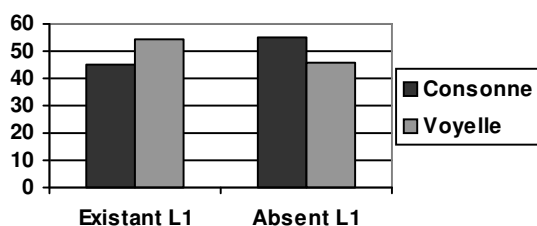
### **Résultats Inférentiels**

- **Influence des trois variables indépendantes (position, consonne, voyelle) sur la variable dépendante (confusion).**
- **Interaction entre la consonne et la voyelle, mais pas d'interaction avec la position.**

**C) Résultats complémentaires : choix de graphème existant/absent en L1 (41 sujets)**

Pour les 41 premiers sujets, nous nous sommes également interrogés sur le choix des graphèmes, tant consonantiques que vocaliques.

	Taux de choix du graphème			
	Consonantique		Vocalique	
	Existant en L1 (<B> ou <R>)	Absent en L1 (<V> ou <L>)	Existant en L1 (<OU>)	Absent en L1 (<EU>)
<b>Moyenne</b>	44, 97%	55, 03%	54, 37%	45, 63%
<b>Ecart Type</b>	9, 16%	9, 16%	16, 16%	16, 16%

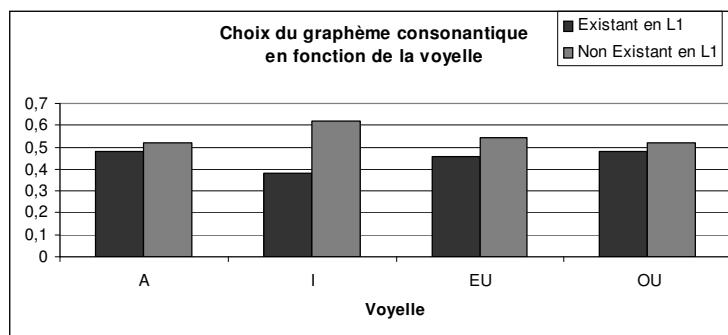


**Pour les consonnes : les sujets ont préférentiellement choisi les graphèmes absents en L1 (<V, L> 55% vs <B, R> 45%).**

**Pour les voyelles (digraphes) : les sujets ont préférentiellement choisi les graphèmes existant en L1 (<OU> 54% vs <EU> 46%).**

Lorsque l'on regarde ensuite l'influence des consonnes et des voyelles entre elles sur ce choix, voici ce que l'on obtient :

<b>Taux de Choix de Graphème Consonantique</b>				
<b>Voyelle</b>	<b>Existant en L1 (&lt;B&gt; ou &lt;R&gt;)</b>		<b>Absent en L1 (&lt;V&gt; ou &lt;L&gt;)</b>	
	<b>Moyenne</b>	<b>Ecart Type</b>	<b>Moyenne</b>	<b>Ecart Type</b>
<b>A</b>	48, 17%	14, 37%	51, 83%	14, 37%
<b>I</b>	38, 21%	15, 13%	61, 79%	15, 13%
<b>EU</b>	45, 53%	15, 6%	54, 47%	15, 6%
<b>OU</b>	47, 97%	16, 33%	52, 03	16, 33%



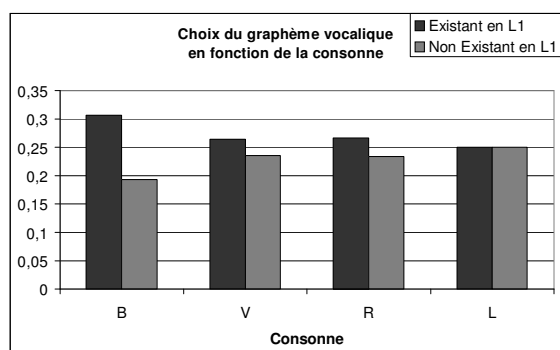
Remarques :

- La voyelle semble influencer le choix du graphème.
- Alors que A et OU semblent avoir une influence faible (48% vs 52%), I et EU ont une influence plus forte (38% vs 62%) et (46% vs 54%).

**Les voyelles A et OU semblent peu influencer le choix du graphème.**

**En revanche, EU (8% d'écart) et surtout I (24% d'écart) semblent influencer davantage le choix du graphème (absent en L1).**

<b>Taux de Choix de Graphème Vocalique</b>				
<b>Consonne</b>	<b>Existant en L1 (&lt;OU&gt;)</b>		<b>Absent en L1 (&lt;EU&gt;)</b>	
	<b>Moyenne</b>	<b>Ecart Type</b>	<b>Moyenne</b>	<b>Ecart Type</b>
<b>B</b>	30, 69%	9, 21%	19, 31%	9, 21%
<b>V</b>	26, 42%	11, 92%	23, 58%	11, 92%
<b>R</b>	26, 62%	12, 53%	23, 37%	12, 53%
<b>L</b>	25%	12, 64%	25%	12, 64%



Remarques :

- Le trigraphe <BOU> (31%) a été préféré à <VOU> (26%), alors que <VEU> (24%) a été préféré à <BEU> (19%).
- Le trigraphe <ROU> (27%) a légèrement été préféré à <LOU> (25%), alors que <LEU> (25%) a légèrement été préféré à <REU> (23%).
- Les écarts sont plus importants pour le couple B/V (5%) que pour R/L (2%).
- On obtient ainsi les classements suivants :
  - B/V : BOU > VOU > VEU > BEU
  - R/L : ROU > LOU > LEU > REU

**Les graphèmes consonantiques existant en L1 (B/R) semblent préférentiellement associés au graphème vocalique existant en L1 (OU).**

## D) Résumé

### Résumé Essentiel

**120 sujets**

On observe 29% de confusion phonème-graphème au sein des couples de consonnes (B, V) et (R, L).

#### Position :

Globalement : Peu de différences entre les trois positions, mais moins de confusions en attaque initiale (- 5%).

Spécifiquement : L'influence de la position varie selon la nature de la consonne :

-pour B : 3 > 1 > 2

-pour R : 1, 2 > 3

-pour V : 2 > 3 > 1

-pour L : 2 > 3 > 1

-Extrêmes : B Coda (53%) > R Coda / V attaque initiale (15, 5%)

#### Nature :

Consonne : elle influence le taux de confusion, plus importante pour B (45%) que pour V, L et R (24%).

#### Environnement vocalique :

Globalement : influence non-négligeable. Plus de confusions avec OU qu'avec EU, A et I.

Spécifiquement : varie selon la nature de la consonne :

-pour B, V, R, L : OU > EU

-pour B, V, L : A > I, mais RI > RA



**Résultats Inférentiels : On trouve une influence significative de chaque variable indépendante. Seule interaction : nature consonne / nature voyelle.**

**41 sujets**

**Choix des graphèmes : les sujets ont préférentiellement choisi :**

- les graphèmes consonantiques absents en L1
- les graphèmes vocaliques existants en L1

**Influence : dans le choix des graphèmes consonantiques :**

- A et OU ont peu d'influence
- EU et I ont plus d'influence (choix du graphème absent)

**Association : les graphèmes existants en L1 semblent préférentiellement être associés entre eux (B/R et OU) :**

- B/V : BOU > VOU > VEU > BEU
- R/L : ROU > LOU > LEU > REU

## **Tâche F**

### **A) Résultats descriptifs**

Il y avait en tout 96 couples de stimuli (48 couples de stimuli identiques et 48 couples de stimuli différents) également répartis sur les trois positions (32 couples de stimuli par position). Chaque consonne à l'étude a donc été présentée 24 fois (4 voyelles\*3 positions\*2 couples).

#### **A) 1) Confusion phonographémique (premier stimulus)**

Pour chaque couple de stimuli, nous nous sommes concentrés sur le premier stimulus : il s'agissait alors d'une tâche similaire à la tâche A.

### **Résultats Globaux**

Le taux moyen de confusion phonographémique, pour l'ensemble des segments consonantiques concernés est le suivant :

	<b>Taux de confusion phonographémique consonantique</b>
<b>Moyenne</b>	20, 86%
<b>Ecart-type</b>	8, 67%

#### **Remarques :**

- Le taux est légèrement inférieur au quart des réponses, ce qui est non-négligeable.

### **Résumé Essentiel Résultats Globaux**

**On observe 21% de confusion phonème-graphème au sein des couples de consonnes (B, V) et (R, L)**

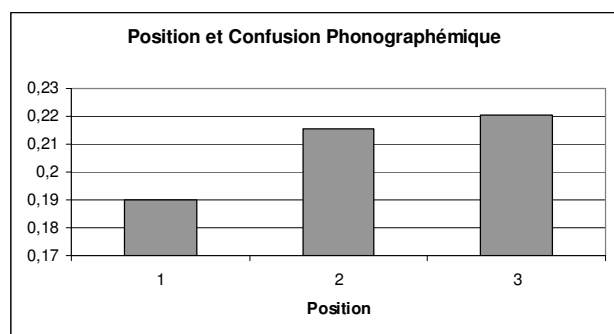
### A) 1) 1. Axe syntagmatique

Les nombres de stimuli, selon la position, étaient les suivants :

<b>Position</b>	<b>1</b>	<b>2</b>	<b>3</b>
<b>Nombre de stimuli</b>	32	32	32

Si l'on considère l'ensemble des quatre consonnes à l'étude, voici ce que l'on obtient :

<b>Taux de confusion phonographémique consonantique</b>			
<b>Position</b>	<b>1 (Attaque Initiale)</b>	<b>2 (Attaque Finale)</b>	<b>3 (Coda Finale)</b>
<b>Moyenne</b>	18,99%	21,56%	22,03%
<b>Ecart-type</b>	10,95%	10,07%	9,81%



#### Remarques :

- Peu de différences entre les positions 2 et 3.
- Le taux de confusion est le plus faible en position 1.

Si l'on considère chaque consonne en détail, voici ce que l'on obtient :

		<b>Taux de confusion phonographémique</b>		
<b>Consonne</b>	<b>Position</b>	<b>1 (Attaque Initiale)</b>	<b>2 (Attaque Finale)</b>	<b>3 (Coda Finale)</b>
<b>B</b>	<b>Moyenne</b>	25%	23, 3%	36, 6%
	<b>Ecart-type</b>	19, 8%	19, 9%	24, 8%
<b>V</b>	<b>Moyenne</b>	21, 8%	34, 3%	31, 2%
	<b>Ecart-type</b>	15, 8%	22, 2%	21, 4%
<b>R</b>	<b>Moyenne</b>	19, 4%	17, 7%	10, 9%
	<b>Ecart-type</b>	18, 5%	21, 2%	14, 8%
<b>L</b>	<b>Moyenne</b>	10, 1%	10, 2%	9, 3%
	<b>Ecart-type</b>	13%	14%	11%

Remarques :

- L'influence de la position varie selon la consonne :
  - B : Coda finale (37%) > Attaque initiale (25%) > Attaque finale (23%) → Différence importante (13%) (3 vs. 1/2)
  - V : Attaque finale (34%) > Coda finale (31%) > Attaque initiale (22%) → Différence importante (10%) (2/3 vs. 1)
  - R : Attaque initiale (19%) > Attaque finale (18%) > Coda finale (11%) → Différence non négligeable (7%) (1/2 vs. 3)
  - L : Attaque finale et Attaque initiale (10%) > Coda finale (9%) → Peu de différence
- Dans chaque cas : deux des trois taux sont toujours très proches.
- Les extrêmes sont donc : B Coda finale (37%) > L Coda finale (9%).

### Résumé Essentiel Axe Syntagmatique

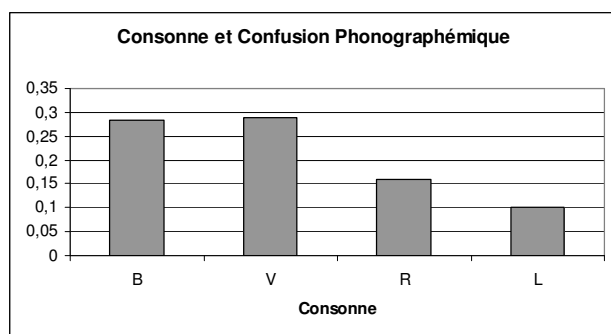
**Globalement** : Peu de différences entre les trois positions, mais moins de confusions en attaque initiale (- 3%).

**Spécifiquement** : L'influence de la position varie selon la nature de la consonne.

#### A) 1) 2. Axe paradigmaticque

##### α) Nature de la consonne

Consonne	Taux de Confusion Phonographémique	
	Moyenne	Ecart Type
<b>B</b>	28,46%	15,8%
<b>V</b>	28,84%	14,3%
<b>R</b>	16,03%	15,27%
<b>L</b>	10,08%	9,77%



#### Remarques :

- Le classement est le suivant : V (29%) > B (28%) > R (16%) > L (10%).
- Les taux de V et de B sont quasi-identiques (29%), alors que ceux de R et de L diffèrent de 6%.

- On obtient donc :
  - Dans 29% des cas : /v/ est transcrit <B> et /b/ est transcrit <V>
  - Dans 16% des cas : /R/ est transcrit <L>
  - Dans 10% des cas : /l/ est transcrit <R>

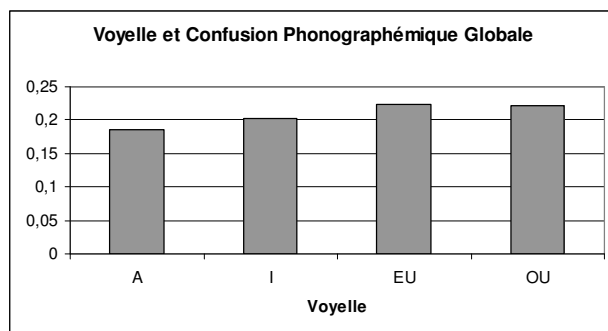
**La nature de la consonne influence le taux de confusion : V et B (29%) entraînent presque deux fois plus de confusion phonographémique que R (16%) et trois fois plus que L (10%).**

### β) Nature de la voyelle

#### Influence Globale :

Lorsque l'on considère l'ensemble des quatre consonnes à l'étude, l'influence globale de la voyelle tautosyllabique sur le taux de confusion s'exprime ainsi :

Voyelle	Taux de Confusion Consonantique Phonographémique	
	Moyenne	Ecart Type
A	18, 5%	9,65%
I	20, 3%	11, 4%
EU	22, 3%	11, 1%
OU	22, 2%	11, 8%



Remarques :

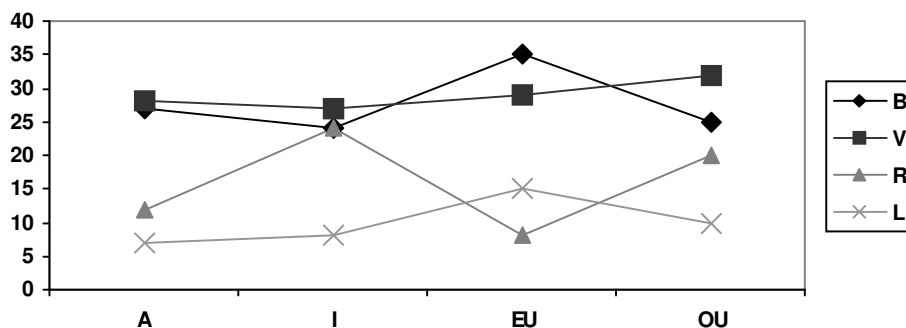
- Le classement est le suivant : EU, OU (22%) > I (20%) > A (18%).
- Peu de différences.
- OU et EU déclenchent davantage de confusion que A et I.
- A (18%) déclenche le moins de confusions.

**Globalement, la nature de la voyelle adjacente semble peu influencer le taux de confusion consonantique phonème-graphème.**

Influence Spécifique :

Cependant, lorsque l'on regarde l'influence de la voyelle adjacente pour chaque consonne, voici ce que l'on obtient :

		Taux de Confusion Consonantique Phonographémique			
Consonne	Mesure	A	I	EU	OU
B	Moyenne	27,4%	23,7%	35,3%	24,9%
	Ecart-type	22,3%	20,8%	23%	24%
V	Moyenne	27,7%	26,6%	28,6%	31,9%
	Ecart-type	20,5%	21,1%	23%	19,6%
R	Moyenne	12%	23,8%	7,8%	20,4%
	Ecart-type	15,6%	24,6%	14,5%	25,3%
L	Moyenne	6,8%	7,7%	15,4%	10,4%
	Ecart-type	12,9%	12,4%	16,4%	15,9%



Remarques :

- L'influence de la voyelle varie selon la consonne.
- Le classement vocalique obtenu est le suivant :
  - BEU (35%) > BA (27%) > BOU (25%) > BI (24%)
  - VOU (32%) > VEU (29%) > VA (28%) > VI (27%)
  - RI (24%) > ROU (20%) > RA (12%) > REU (8%)
  - LEU (15%) > LOU (10%) > LI (8%) > LA (7%)
- Le classement consonantique obtenu est le suivant :
  - VOU (32%) > BOU (25%) > ROU (20%) > LOU (10%)
  - BEU (35%) > VEU (29%) > LEU (15%) > REU (8%)
  - VA (28%) > BA (27%) > RA (12%) > LA (7%)
  - VI (27%) > BI et RI (24%) > LI (8%)
- On remarque que :
  - BEU > BOU et LEU > LOU, mais VOU > VEU et ROU > REU
  - BA > BI, VA > VI, mais RI > RA et LI > LA
- Les extrêmes sont donc :
  - B : Max= EU ; Min=I
  - V : Max= OU ; Min= I
  - R : Max= I ; Min= EU
  - L : Max= EU ; Min= A-I



**De manière spécifique, la nature de la voyelle adjacente influence de manière différente le taux de confusion consonantique phonème-graphème selon la consonne.**

**En particulier :**

- **Pour A : B, V > R > L**
- **Pour I : V > B, R > L**
- **Pour EU : B > L > V > R**
- **Pour OU : V > B > R > L**

### Résumé Essentiel Axe Paradigmatique

**Consonne :** la nature de la consonne influence le taux de confusion : V et B (30%) conduisent à deux fois plus de confusion que R (15%) et trois plus que L (10%).

**Environnement vocalique :**

- **Globalement :** peu d'influence, mais OU et EU déclenchent plus de confusions que I et A (minimum)
- **Spécifiquement :** l'influence varie selon la consonne.

#### A) 1) 3. Variable liée à la tâche : nature du couple (identique/différent)

Si l'on examine le taux global de confusion phonographémique consonantique en fonction de la nature (identique/différente) du couple de stimuli, voici ce que l'on obtient :

<b>Taux de Confusion Consonantique Phonographémique</b>		
	<b>Stimuli identiques</b>	<b>Stimuli différents</b>
<b>Moyenne</b>	23, 1%	18, 6%
<b>Ecart-type</b>	9, 6%	9, 4%

**Globalement, lorsque les stimuli sont identiques, il y a légèrement plus de confusions consonantiques phonographémiques (23%) que lorsqu'ils sont différents (19%).**

A) 2) Confusion phonologique (entre les deux stimuli auditifs)

Il s'agissait d'une tâche de discrimination phonologique via une tâche de choix de transcription orthographique (*cf.* tâche A).

Résultats Globaux

Le taux moyen de confusion phonologique, pour l'ensemble des segments consonantiques concernés, est le suivant :

	<b>Taux de confusion phonologique consonantique</b>
<b>Moyenne</b>	26, 8%
<b>Ecart-type</b>	7, 9%

Remarques :

- Le taux est légèrement supérieur au quart des réponses, ce qui est non-négligeable.

**Résumé Essentiel Résultats Globaux**

**On observe 27% de confusion phonologique au sein des couples de consonnes (B, V) et (R, L).**

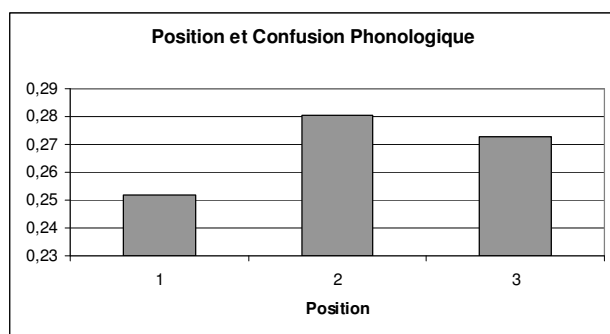
### A) 2) 1. Axe syntagmatique

Les nombres de stimuli, selon la position, étaient les suivants :

<b>Position</b>	<b>1</b>	<b>2</b>	<b>3</b>
<b>Nombre de stimuli</b>	32	32	32

Si l'on considère l'ensemble des quatre consonnes à l'étude, voici ce que l'on obtient :

<b>Taux de confusion phonologique consonantique</b>			
<b>Position</b>	<b>1 (Attaque Initiale)</b>	<b>2 (Attaque Finale)</b>	<b>3 (Coda Finale)</b>
<b>Moyenne</b>	25, 2%	28%	27, 2%
<b>Ecart-type</b>	9, 5%	10, 1%	9, 3%



#### Remarques :

- Le classement est le suivant : Attaque Finale (28%) > Coda Finale (27%) > Attaque Initiale (25%).
- Peu de différences.

Si l'on considère chaque consonne en détail, voici les résultats que l'on obtient :

		<b>Taux de confusion phonologique</b>		
<b>Consonne</b>	<b>Position</b>	<b>1 (Attaque Initiale)</b>	<b>2 (Attaque Finale)</b>	<b>3 (Coda Finale)</b>
<b>B</b>	<b>Moyenne</b>	31, 3%	38, 6%	38, 3%
	<b>Ecart-type</b>	13, 7%	14, 9%	16, 3%
<b>V</b>	<b>Moyenne</b>	36, 5%	36, 1%	42, 8%
	<b>Ecart-type</b>	16, 5%	17, 4%	16, 1%
<b>R</b>	<b>Moyenne</b>	18, 3%	17, 1%	12, 8%
	<b>Ecart-type</b>	15, 3%	15, 4%	15, 1%
<b>L</b>	<b>Moyenne</b>	14%	20, 3%	15%
	<b>Ecart-type</b>	12, 9%	16, 1%	16, 1%

Remarques :

- L'influence de la position varie selon la consonne :
  - B : Attaque finale (39%) > Coda finale (38%) > Attaque initiale (31%) → Différence non négligeable (7%) (2/3 vs. 1)
  - V : Coda finale (43%) > Attaque Initiale et Attaque finale (36%) → Différence non-négligeable (7%) (3 vs. 1/2)
  - R : Attaque initiale (18%) > Attaque finale (17%) > Coda finale (13%) → Différence négligeable (4%) (1/2 vs. 3)
  - L : Attaque finale (20%) > Coda finale (15%) > Attaque initiale (14%) → Différence négligeable (5%) (2 vs 1/3)
- Dans chaque cas : deux des trois taux sont toujours très proches.
- Les extrêmes sont donc : V Coda Finale (43%) > L Attaque Initiale (14%) et R Coda Finale (13%)

### Résumé Essentiel Axe Syntagmatique

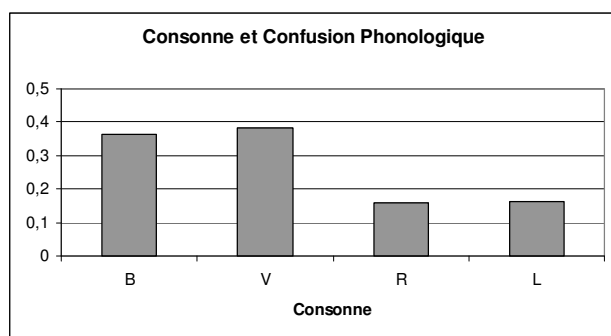
**Globalement** : Peu de différence entre les trois positions, mais moins de confusion en Attaque Initiale (- 2%).

**Spécifiquement** : l'influence de la position semble varier selon la consonne.

#### A) 2) 2. Axe paradigmatique

##### α) Nature de la consonne

Consonne	Taux de Confusion Phonologique	
	Moyenne	Ecart Type
<b>B</b>	36, 3%	9, 9%
<b>V</b>	38, 4%	11%
<b>R</b>	16, 1%	11, 8%
<b>L</b>	16, 4%	11%



#### Remarques :

- Le classement obtenu est le suivant : /v/ (38%) > /b/ (36%) > /l/, /r/ (16%).
- Les confusions /v, b/ (37%) sont plus nombreuses que les confusions /l, r/ (16%).

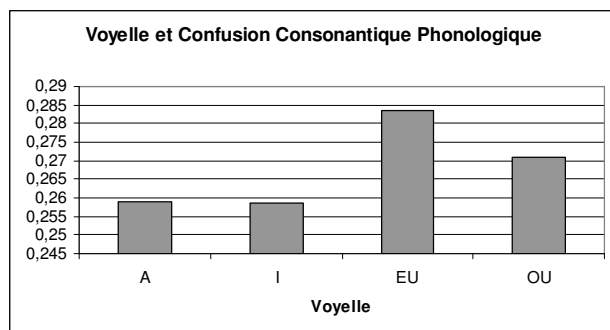
**La nature de la consonne influence le taux de confusion phonologique : les confusions /v, b/ sont plus nombreuses (37%) que les confusions /l, R/ (16%).**

### **β) Nature de la voyelle**

#### Influence Globale :

Lorsque l'on considère l'ensemble des quatre consonnes à l'étude, l'influence globale de la voyelle tautosyllabique sur le taux de confusion s'exprime ainsi :

Voyelle	Taux de Confusion Consonantique Phonologique	
	Moyenne	Ecart Type
<b>A</b>	25,9%	9,2%
<b>I</b>	25,9%	10,7%
<b>EU</b>	28,3%	10%
<b>OU</b>	27,1%	11,4%



#### Remarque :

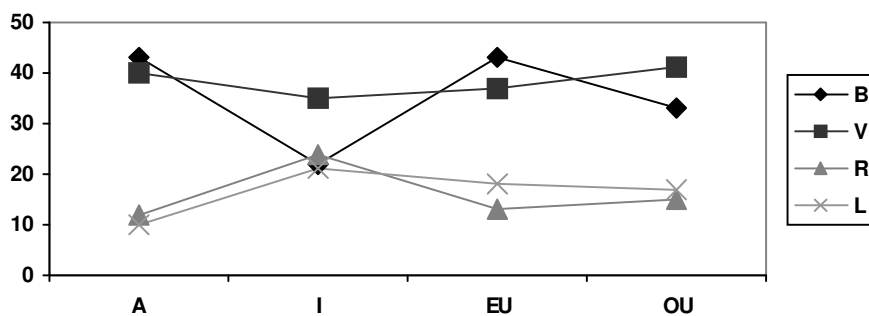
- Le classement obtenu est le suivant : EU (28%) > OU (27%) > A, I (26%).
- Peu de différence. Cependant, les voyelles /EU/ et /OU/ déclenchent légèrement plus de confusions que les voyelles /A/ et /I/.

**La nature de la voyelle semble peu influencer le taux de confusion consonantique phonologique.**

Influence Spécifique :

Cependant, lorsque l'on regarde l'influence de la voyelle adjacente pour chaque consonne, voici ce que l'on obtient :

		<b>Taux de Confusion Consonantique Phonologique</b>			
<b>Consonne</b>	<b>Mesure</b>	<b>A</b>	<b>I</b>	<b>EU</b>	<b>OU</b>
<b>B</b>	Moyenne	42, 6%	22, 3%	43%	33, 4%
	Ecart-type	16, 6%	17, 9%	16, 5%	18, 4%
<b>V</b>	Moyenne	39, 8%	35, 4%	37, 1%	40, 9%
	Ecart-type	19, 8%	19, 1%	19%	17, 2%
<b>R</b>	Moyenne	11, 6%	24, 4%	13%	15, 3%
	Ecart-type	15, 4%	18%	13%	15, 5%
<b>L</b>	Moyenne	9, 7%	20, 7%	17, 8%	17, 5%
	Ecart-type	13, 5%	16, 1%	16%	19, 4%



Remarques :

- L'influence de la voyelle varie selon la consonne.
- Le classement vocalique obtenu est le suivant :
  - BEU, BA (43%) > BOU (33%) > BI (22%)
  - VOU, VA (40, 5%) > VEU (37%) > VI (35%)
  - RI (24%) > ROU, REU (14%) > RA (12%)
  - LI (21%) > LOU, LEU (18, 5%) > LA (8%)

- On remarque :
  - Pour B et V : A au premier rang, I en dernier.
  - Pour R et L : I au premier rang, A en dernier.
  
- Le classement consonantique obtenu est le suivant :
  - VOU (41%) > BOU (33%) > LOU (17%) > ROU (15%)
  - BEU (43%) > VEU (37%) > LEU (18%) > REU (13%)
  - BA (43%) > VA (40%) > RA (12%) > LA (8%)
  - VI (35%) > RI (24%) > BI (22%) > LI (21%)
  
- On remarque que :
  - BEU > BOU, mais VOU > VEU
  - LOU = LEU et ROU = REU
  - BA > BI, VA > VI, mais RI > RA et LI > LA
  
- Les extrêmes sont donc :
  - B : Max= EU-A ; Min=I
  - V : Max= OU-A ; Min= I
  - R : Max= I ; Min= A
  - L : Max= I ; Min= A

**De manière spécifique, la nature de la voyelle adjacente influence de manière différente le taux de confusion consonantique phonème-graphème selon la consonne.**

**En particulier :**

- **Pour A : B, V > R > L**
- **Pour I : V > R > B > L**
- **Pour EU : B > V > L > R**
- **Pour OU : V > B > L > R**



### **Résumé Essentiel Axe Paradigmatique**

**Consonne** : la nature de la consonne influence le taux de confusion phonologique : les consonnes /b, v/ (37%) entraînent deux fois plus de confusion que les consonnes /l, R/ (16%).

#### **Environnement vocalique** :

- **Globalement** : peu d'influence, mais légèrement plus de confusion avec /EU, OU/ qu'avec /A, I/.
- **Spécifiquement** : l'influence varie selon la consonne.

#### **A) 2) 3. Variables liées à la tâche : nature du couple (identique/différent)**

Si l'on examine le taux global de confusion phonologique consonantique en fonction de la nature (identique/différente) du couple de stimuli, voici ce que l'on obtient :

	<b>Taux de Confusion Consonantique Phonologique</b>	
	<b>Stimuli identiques</b>	<b>Stimuli différents</b>
<b>Moyenne</b>	20, 3%	33, 3%
<b>Ecart-type</b>	14, 1%	13, 6%

On constate alors que :

- Lorsque les stimuli sont identiques, le taux de confusion phonologique est de 20%.
- Lorsque les stimuli sont différents, le taux de confusion phonologique est de 33%.

**Globalement, il y a une différence importante entre les deux cas (+13%) avec moins de confusions pour les stimuli identiques.**

## **B) Résultats inférentiels : analyses de variance (Test F)**

### **Univariée :**

#### 1) Taux de confusion phonographémique :

Les quatre variables indépendantes (position, nature consonantique, nature vocalique et nature du couple) semblent influencer significativement ( $p < 0,001$ ) la variable dépendante (confusion phonographémique).

#### 2) Taux de confusion phonologique :

Les quatre variables indépendantes (position, nature consonantique, nature vocalique et nature du couple) semblent influencer significativement ( $p < 0,001$ ) la variable dépendante (confusion phonologique).

### **Bivariée :**

#### 1) Taux de confusion phonographémique :

D'après les analyses statistiques, il ne semblerait y avoir d'interaction significative ni entre la position et la voyelle ( $p = 0,59$ ), ni entre la position et la consonne ( $p = 0,58$ ). En revanche il semblerait y avoir une interaction significative entre la consonne et la voyelle ( $p < 0,001$ ).

#### 2) Taux de confusion phonologique :

D'après les analyses statistiques, il ne semblerait y avoir d'interaction significative ni entre la position et la voyelle ( $p = 0,82$ ), ni entre la position et la consonne ( $p = 0,81$ ). En revanche il semblerait y avoir une interaction significative entre la consonne et la voyelle ( $p < 0,001$ ).

### **Résultats Inférentiels**

- **Influence des quatre variables indépendantes (position, consonne, voyelle, couple) sur les deux variables dépendantes (confusion phonographémique et phonologique).**
- **Interaction entre la consonne et la voyelle, mais pas d'interaction avec la position.**

## **D) Résumé**

### **Résumé Essentiel**

#### **A) Confusions Phonographémiques**

On observe 21% de confusion phonème-graphème au sein des couples de consonnes (B, V) et (R, L).

#### **Position :**

**Globalement** : peu de différence, mais moindre confusion en attaque initiale (-3%).

**Spécifiquement** : varie selon la consonne : influence plus importante pour B et V que pour R et L.

#### **Nature :**

**Consonne** : influence le taux de confusion : influence plus importante pour B et V (30%) que pour R (15%) et L (10%).

#### **Environnement vocalique :**

**Globalement** : peu d'influence, mais plus de confusion avec OU et EU qu'avec I et A (minimum).

**Spécifiquement** : varie selon la consonne.

**Couple** : plus de confusion pour les couples identiques (23%) que pour les couples différents (19%).

**Résultats Inférentiels** : influence de chaque variable indépendante. Seule interaction : nature consonne/voyelle.

### **B) Confusions Phonologiques**

On observe 27% de confusion phonologique au sein des couples de consonnes /B, V/ et /R, L/.

#### **Position** :

**Globalement** : peu de différence, mais moindre confusion en attaque initiale (-2%).

**Spécifiquement** : varie selon la consonne : influence plus importante pour B et V que pour R et L.

#### **Nature** :

**Consonne** : influence le taux de confusion : influence plus importante pour /b, v/ (37%) que pour /l, R/ (16%)

#### **Environnement vocalique** :

**Globalement** : peu d'influence, mais plus de confusion avec /OU, EU/ qu'avec /l, A/.

**Spécifiquement** : varie selon la consonne.

**Couple** : plus de confusion pour les couples différents (33%) que pour les couples identiques (20%).

**Résultats Inférentiels** : influence de chaque variable indépendante. Seule interaction : nature consonne/voyelle.

## Comparaison Tâches A et F

Lorsque l'on compare les résultats des tâches A et F (pour le premier stimulus) à propos des taux de confusion phonographémique, voici ce que l'on obtient.

### A) Résultats globaux

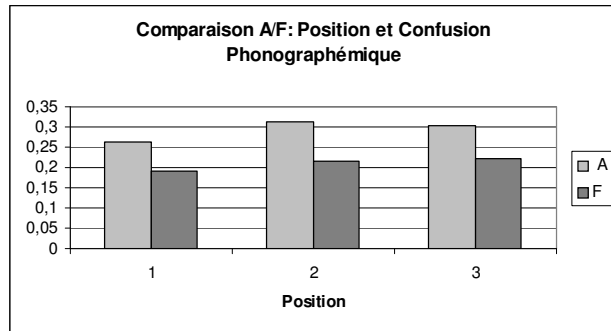
Taux de confusion phonographémique consonantique	Tâche A (identification)	Tâche F (discrimination)
Moyenne	29, 37%	20, 86%
Ecart-type	10, 9%	8, 67%

**Moins de confusion phonographémique en tâche de discrimination (21%) qu'en tâche d'identification (29%).**

### B) Position

#### Globalement :

Position	Taux de confusion phonographémique consonantique					
	1 (Attaque Initiale)		2 (Attaque Finale)		3 (Coda Finale)	
Tâche	A	F	A	F	A	F
Moyenne	26, 3%	18, 99%	31, 35%	21, 56%	30, 47%	22, 03
Ecart-type	14, 33%	10, 95%	13, 63%	10, 07%	14, 16%	9, 81%



**Classements relativement similaires en tâches A et F : moins de confusion en attaque initiale (position 1). Positions 2 et 3 similaires.**

**Spécifiquement :**

Tâche A	Tâche F
<ul style="list-style-type: none"> <li>• Influence forte pour les quatre consonnes.</li> <li>• Pour R, deux des trois taux sont similaires (Attaques initiale et finale), mais pas pour les autres.</li> <li>• <u>Extrêmes</u> :  <u>Max</u> : B Coda finale (53%)  <u>Min</u> : R Coda finale / V Attaque initiale (15, 5%)</li> </ul>	<ul style="list-style-type: none"> <li>• Influence forte pour B et V, mais moindre pour R et L.</li> <li>• Pour chaque consonne, deux des trois taux sont similaires (pour R : Attaques initiale et finale).</li> <li>• <u>Extrêmes</u> :  <u>Max</u> : B Coda finale (37%)  <u>Min</u> : L Coda finale (9%)</li> </ul>

**L'influence de la position varie selon la consonne et, dans une certaine mesure, la tâche :**

- **Cette influence est identique (en termes de classement) dans les deux tâches pour certaines consonnes :**

**B : Coda Finale > Attaque initiale > Attaque finale**

**V : Attaque finale > Coda Finale > Attaque initiale**

**R : Attaque initiale/Attaque finale > Coda finale**

- **Mais elle diffère selon la tâche pour L :**

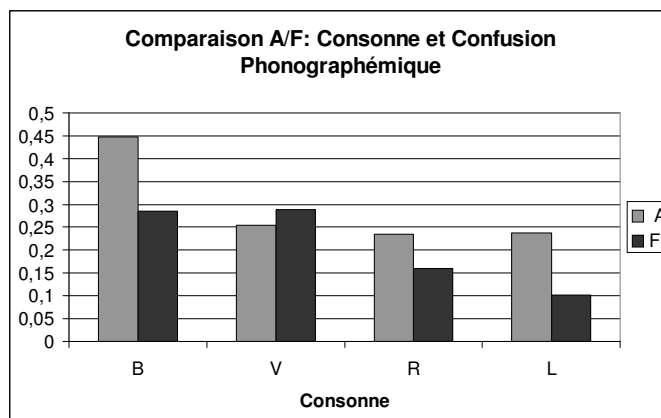
**Identification (A) : Attaque finale > Coda Finale > Attaque initiale**

**Discrimination (F) : pas de différence**

### **C) Nature**

#### **a) Consonne**

<b>Classement</b>	
<b>Tâche A</b>	<b>Tâche F</b>
B (45%) > V/R/L (24, 5%)	V/B (28, 5%) > R (16%) > L (10%)



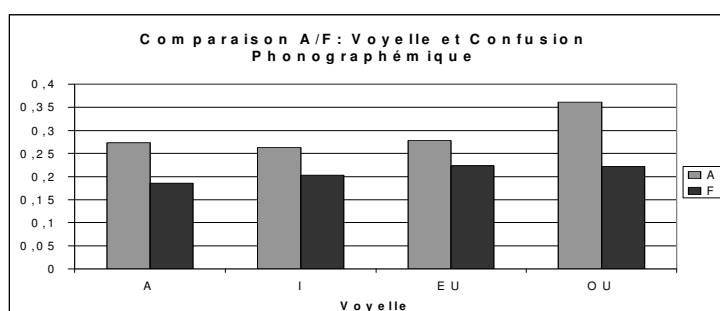
On constate alors que :

- b/ entraîne le plus de confusion dans les deux tâches, mais a un taux plus faible en tâche de discrimination.
- /v/ a un taux relativement similaire dans les deux tâches, mais il est plus élevé en tâche de discrimination.
- /R/ a un taux plus faible en tâche de discrimination.
- /l/ a un taux plus faible en tâche de discrimination.

**L'influence de la consonne varie selon la tâche, mais la tâche de discrimination semble, dans la plupart des cas (/b, R, l/), réduire les taux de confusion.**

b) Voyelle :

Classement Global	
Tâche A	Tâche F
OU (36%) > EU/A/I (27%)	OU/EU (22%) > I/A (19%)



**Globalement, l'influence de la voyelle semble moindre en tâche de discrimination qu'en tâche d'identification.**

**Dans les deux tâches, OU déclenche le plus de confusion, et OU et EU en déclenchent plus que A et I.**



<b>Classements Spécifiques</b>	
<b>Tâche A</b>	<b>Tâche F</b>
<p><u>Pour A et I</u> : B &gt; V et R &gt; L</p> <p><u>Pour OU et EU</u> : B &gt; V et L &gt; R</p> <p><u>Extrêmes</u> :</p> <p style="padding-left: 40px;"><u>B et V</u> : Max=OU ; Min= EU-I</p> <p style="padding-left: 40px;"><u>R</u> : Max= I-OU ; Min= A-EU</p> <p style="padding-left: 40px;"><u>L</u> : Max= OU ; Min= I</p>	<p><u>Pour A et I</u> : V &gt; B et R &gt; L</p> <p><u>Pour OU</u> : V &gt; B et R &gt; L</p> <p><u>Pour EU</u> : B &gt; V et L &gt; R</p> <p><u>Extrêmes</u> :</p> <p style="padding-left: 40px;"><u>B</u> : Max= EU ; Min=I</p> <p style="padding-left: 40px;"><u>V</u> : Max= OU ; Min= I</p> <p style="padding-left: 40px;"><u>R</u> : Max= I ; Min= EU</p> <p style="padding-left: 40px;"><u>L</u> : Max= EU ; Min= A-I</p>

On constate donc que l'influence de la voyelle peut être :

- Identique dans les deux tâches :

Pour A et I : R > L

Pour EU : B > V et L > R

Pour B et V : Min= I

Pour V : Max= OU

Pour R : Max= I et Min= EU

Pour L : Min= I

- Différente dans les deux tâches :

B Max : tâche A= OU / tâche F= EU

L Max : tâche A= OU / tâche F= EU

**L'influence de la voyelle varie selon la consonne, et dans une certaine mesure selon la tâche.**

#### **D) Résultats inférentiels**

Les résultats inférentiels sont identiques dans les deux tâches :

- Influence des variables indépendantes (position, consonne, voyelle) sur le taux de confusion consonantique phonographémique.
- Influence de l'interaction entre consonne et voyelle, mais pas d'influence de l'interaction avec la position, tant pour les consonnes que pour les voyelles.

**E) Résumé**

**Résumé Essentiel Comparaison tâches Identification (A) / Discrimination (F)**

<b><u>Identification (A)</u></b>	<b><u>Discrimination (F)</u></b>
<b><u>Confusion</u> : 29%</b>	<b><u>Confusion</u> : 21%</b>
<b><u>Position</u> :</b>	<b><u>Position</u> :</b>
<b><u>Globalement</u> :</b> Peu d'influence, mais moindre taux en Attaque initiale.	<b><u>Globalement</u> :</b> Peu d'influence, mais moindre taux en Attaque initiale
<b><u>Spécifiquement</u> :</b> Varie selon la consonne. Forte pour les quatre consonnes	<b><u>Spécifiquement</u> :</b> Varie selon la consonne. Forte pour B et V, moindre pour R et L.
<b>Max= B Coda</b> <b>Pour R : Min= Coda</b>	<b>Max= B Coda</b> <b>Pour R : Min= Coda</b>
<b><u>Consonne</u> : varie selon la consonne</b> <b>Max= /b/</b>	<b><u>Consonne</u> : varie selon la consonne</b> <b>Max= /b/</b>
<b><u>Voyelle</u> : varie selon la consonne</b> Plus de confusion avec OU et EU qu'avec A et I <b>R Max= L Min= I</b>	<b><u>Voyelle</u> : varie selon la consonne</b> Plus de confusion avec OU et EU qu'avec A et I <b>R Max= L Min = I</b>
<b>La tâche semble influencer les résultats : moins de confusion en tâche de discrimination que d'identification.</b>	

### **I.3. Axe syllabique**

Nous présentons ici les résultats relatifs au traitement des syllabes complexes biconsonantiques en mono- et bimodalité, soit les résultats des tâches C1, C2, D et E.

## **Tâche C1**

### **A) Résultats descriptifs**

#### **Résultats globaux**

Rappelons que nous nous intéressons spécifiquement aux épenthèses. La catégorie générale des erreurs inclut celles des épenthèses et des effacements (vocaliques). Les effacements vocaliques proviennent des stimuli initialement conçus comme distracteurs. Il s'agit cependant de résultats secondaires qui ne sont pas dénués d'intérêt, d'où leur mention.

<b>Type</b>	<b>TAUX</b>		
	<b>Erreurs</b>	<b>Epenthèses</b>	<b>Effacements</b>
<b>Moyenne</b>	44, 48%	81, 43%	18, 57%
<b>Ecart-type</b>	16, 34%	14, 34%	14, 34%

#### Remarques :

- Taux d'erreur globalement élevé (environ 44%).
- Une majorité d'épenthèses (environ 81%).
- Un taux d'effacement non négligeable (environ 19%).
- Des écarts-types élevés et similaires pour les trois catégories (environ 15%).

<b>Taux d'épenthèses pour les stimuli de type /CC/</b>	
Moyenne	63, 94%
Ecart-type	20, 32%

Remarques :

- Traitement des groupes /CC/ : environ 64% d'épenthèses.

**Résumé Essentiel Résultats Globaux**

**On observe 64% d'épenthèses pour les groupes /CC/ de la tâche C1  
(unimodalité auditive).**

A) 1) Axe syntagmatique

Les nombres de stimuli, selon la position, étaient les suivants :

<b>Position</b>	<b>1</b>	<b>2</b>	<b>3</b>
<b>Nombre de stimuli</b>	26	26	20

Cette apparente absence d'homogénéité résulte du respect de la distribution réelle des groupes biconsonantiques en français (Dell, 1995). Il était en effet pour nous exclu de présenter aux sujets des groupes consonantiques finaux qui n'apparaissent en réalité jamais en position finale en français.

Il importe de préciser ici la signification d'une notion employée dans la présentation des résultats, à savoir celle de *taux*.

Taux : on peut relier cette notion à celle de probabilité d'occurrence. Par exemple, lorsque l'on analyse les résultats des taux d'épenthèses en fonction de la position d'une part, et l'écart de sonorité d'autre part, voici ce que l'on peut dire : en position Z (par exemple 1), pour un écart de sonorité X (par exemple 2) la probabilité d'occurrence d'une épenthèse est de Y %. On peut alors représenter ces résultats à l'aide de trois histogrammes (combinés ou séparés) l'un pour la position 1, l'autre pour la position 2 et le dernier pour la position 3. Dans chacun de ces histogrammes, le taux d'épenthèses figure en ordonnées (axe vertical), tandis que l'écart de sonorité figure en abscisse (axe

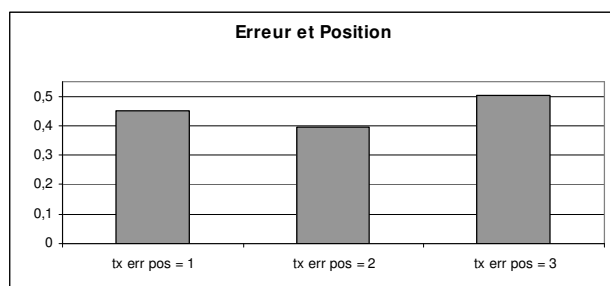
horizontal). Le calcul du taux d'épenthèse en position 1 se fait donc de la manière suivante :

Taux = nombre d'épenthèses en position 1 / nombre de stimuli de type « position 1 ».

Lorsque l'on considère globalement les résultats, sans se soucier de la nature des stimuli (classe générale des groupes /CC/), voici ce que l'on obtient en fonction de la position :

**Erreurs :**

Position	Taux d'erreurs		
	1	2	3
Moyenne	44,96%	39,56%	50,25%
Ecart-type	15,02%	16,51%	24%



**Remarques :**

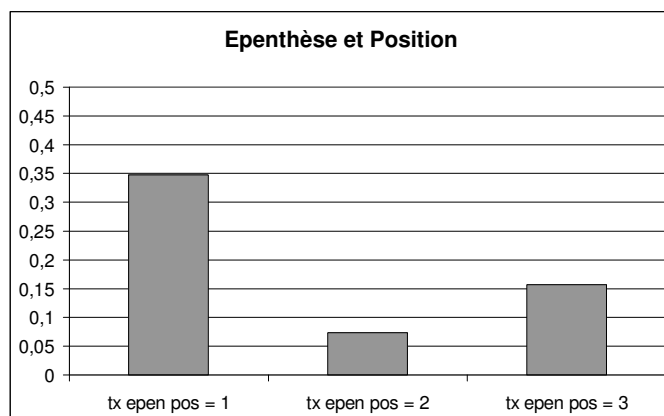
- Taux d'erreurs plus élevés en positions initiale (environ 45%) et finale (environ 50%) qu'en position médiale (environ 40%).
- Cependant, d'une part les écarts-types sont élevés, en particulier en position finale (24%), d'autre part la composition des stimuli en position finale (deux épenthèses possibles) rend délicate l'interprétation de la troisième position lorsque les écarts ne sont pas massivement inégaux.

- Le résultat à retenir est donc le suivant : légèrement davantage d'épenthèses en position 1 (taux : environ 45% ) qu'en position 2 (taux : environ 40% ). Mais cette différence reste très limitée.

**Légèrement moins d'erreurs en attaque de la syllabe finale (position 2) (entre 5 et 10% d'écart).**

### Epenthèses :

Position	Taux d'épenthèses		
	1	2	3
Moyenne	34,78%	7,4%	15,76%
Ecart-type	15,44%	6,13%	17,32%



### Remarques :

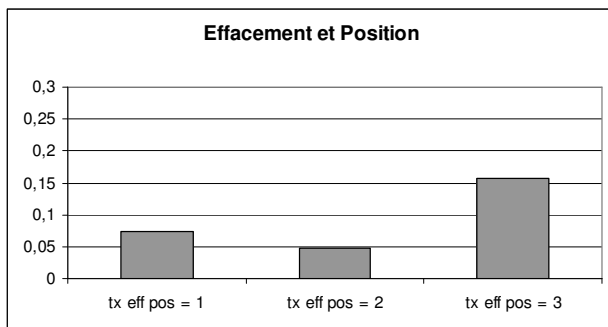
- Tout en gardant en tête les réserves émises à propos de l'interprétation des résultats de la troisième position, les résultats sur les épenthèses confirment la tendance observée pour les erreurs : taux d'épenthèses plus élevés en positions initiale (environ 35%) et finale (environ 16%) qu'en position médiale (environ 7%). Ces différences ne sont pas négligeables en ce que le taux d'épenthèses, entre chaque position, se voit doublé.
- Les écarts-types sont toujours élevés.

- L'influence de la position sur le déclenchement d'épenthèses se présente donc de la manière suivante, de la plus propice aux épenthèses à la moins propice (ordre décroissant) : position 1 (taux : environ 35% ), position 3 (taux : environ 16% ), position 2 (taux : environ 7% ).

- **Plus d'épenthèses en attaque de la syllabe initiale (position 1 : 35%).**
- **Moins d'épenthèses en attaque de la syllabe finale (position 2 : 7%).**

**Effacements :**

Position	Taux d'effacements		
	1	2	3
Moyenne	7,4%	4,78%	15,76%
Ecart-type	6,13%	6,01%	17,33%



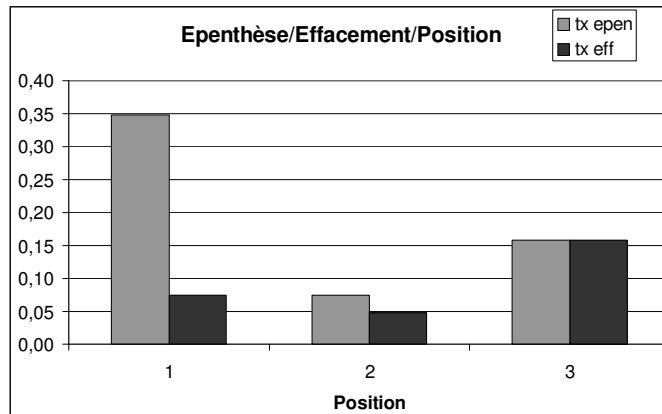
**Remarques :**

- D'après ces résultats, les effacements seraient favorisés d'abord en position 3 (taux : environ 16% ), puis en position 1 (taux : environ 7% ) puis en position 2 (taux : environ 5% ).

- **Plus d'effacements en coda de la syllabe finale (position 3 : 16%).**
- **Moins d'effacements en attaque de la syllabe finale (position 2 : 5%).**



### Comparaison Epenthèses/Effacements en fonction de la position



#### Remarques :

- Comme pour les erreurs et les épenthèses, la position 2 semble davantage préserver l'input que la position 1.
- Tandis que, pour les épenthèses le rapport entre position 1 et 3 est de 2/1 (plus d'épenthèses en position 1), ce rapport est inversé pour les effacements (plus d'effacements en position 3).

- **Attaque de syllabe finale (position 2) : moins d'épenthèses et moins d'effacements.**
- **Attaque de syllabe initiale (position 1) : plus d'épenthèses et moins d'effacements.**
- **Coda de syllabe finale (position 3) : moins d'épenthèses et plus d'effacements.**

#### Résumé Essentiel Axe Syntagmatique

L'influence de la position sur le taux d'épenthèse n'est pas évidente.

Cependant :

- Il y a plus d'épenthèses en attaque de la première syllabe (position initiale : 35%).
- Il y a moins d'épenthèses en attaque de la dernière syllabe (position médiale : 7%).

A) 2) Axe paradigmatic

La caractérisation paradigmatic des stimuli (globalement désignée jusqu'ici par le terme « nature ») a été effectuée selon deux grands modes : la nature phonologique et la sonorité. La nature phonologique peut elle-même être affinée, selon que l'on considère le groupe consonantique en tant qu'entité unique (la « nature » à proprement parler) ou la sous-famille, la famille et la famille générale à laquelle il appartient. Le critère de sonorité peut être affiné en considérant d'une part le type de sonorité du groupe, et d'autre part l'écart de sonorité entre les deux membres du groupe consonantique. Etant donné la répartition non-homogène de ces différents critères, issue de la répartition réelle des groupes consonantiques en français, nous ne présentons ci-dessous que les résultats les plus pertinents.

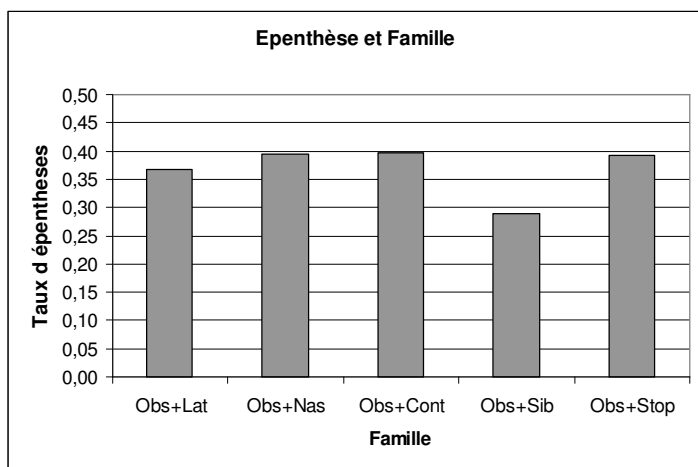
α) Nature phonologique

α). 1. Famille

Cinq familles étaient représentées dans cette tâche, dont une ne contenant qu'un seul représentant.

Epenthèses :

	<b>Taux d'épenthèses</b>				
<b>Famille</b>	<b>Obs + Lat</b>	<b>Obs+Nas</b>	<b>Obs + Cont</b>	<b>Obs + Sib</b>	<b>Obs + Stop</b>
	(exemple : /tl/) (code : 1)	(exemple : /tm/) (code : 5)	(exemple:/pf/) (code : 6)	(exemple : /ps/) (code : 7)	(exemple: /pt/) (code : 8)
<b>Moyenne</b>	36, 69%	39, 41%	39, 6%	29, 03%	39, 2%
<b>Ecart-type</b>	5%	0, 87%	3, 14%	0, 72%	1, 34%



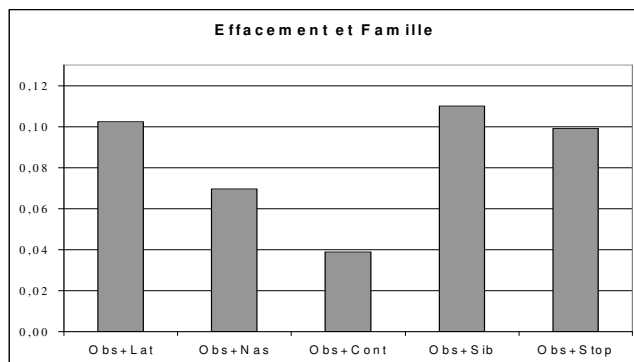
Remarques :

- Les taux d'épenthèses sont non négligeables.
- Les écarts-types sont faibles.
- Tandis que les taux d'épenthèses sont à peu près identiques (environ 39%) pour les familles Obs+Nas, Obs+Cont et Obs+Stop, on note un taux légèrement plus bas pour la famille Obs+Lat (environ 37%), mais surtout un taux plus faible de 10% pour la famille Obs+Sib (environ 29%), ce qui conduit à l'ordre décroissant suivant, selon la difficulté pour les sujets : Obs+Cont, Obs+Nas, Obs+Stop > Obs+Lat > Obs+Sib.
- Statut particulier de la famille Obs + Sib.

**Moins d'épenthèses pour la famille Obs+Sib (ex : /ps/) (29%) que pour les autres (37%-40%).**

Effacements :

Famille	Taux d'effacements				
	Obs + Lat	Obs+Nas	Obs + Cont	Obs + Sib	Obs + Stop
<b>Moyenne</b>	10, 22%	6, 97%	3, 89%	10, 99%	9, 9%
<b>Ecart-type</b>	2, 77%	0, 53%	1, 11%	0, 51%	1%

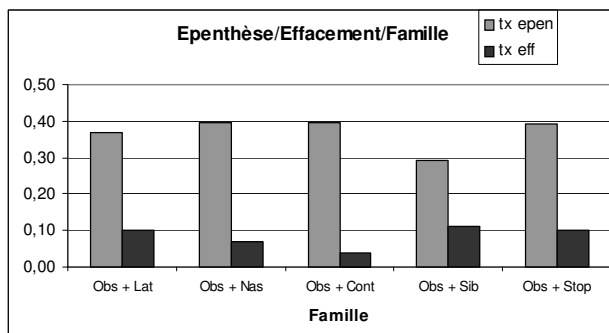


Remarques :

- Les taux d'effacements sont faibles.
- Les écarts-types sont faibles.

**Plus d'effacements pour la famille Obs+Sib (ex : /ps/) (11%) que pour les autres (4%-10%).**

Comparaison Epenthèses/Effacements en fonction de la Famille



Remarques :

- On remarque que, d'une part l'ordre est ici inversé par rapport aux épenthèses, et d'autre part que les écarts entre les trois dernières familles, même s'ils sont minimes, sont plus marqués que pour les épenthèses : Obs+Sib (environ 11%) >

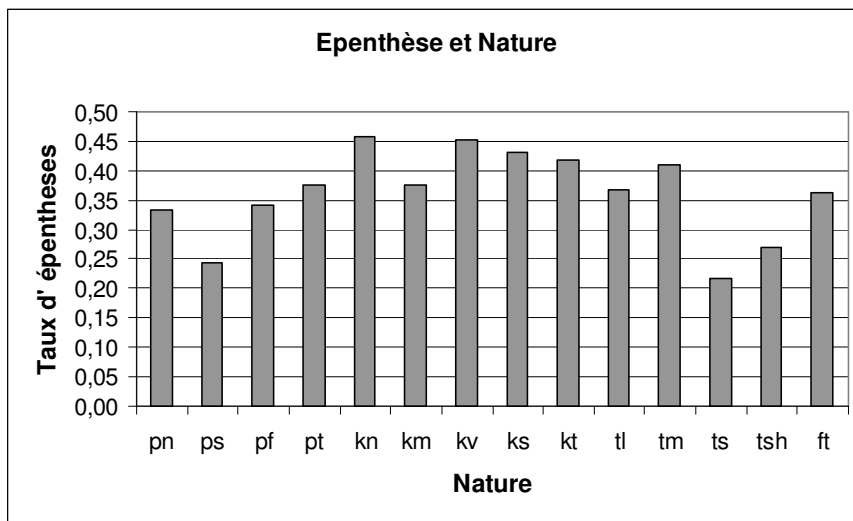
Obs+Lat (environ 10%), Obs+Stop (environ 10%) > Obs+Nas (environ 7%) > Obs+Cont (environ 4%).

**Famille Obs+Sib : moins d'épenthèses et plus d'effacements.**

*α). 2. Nature spécifique*

Quatorze groupes consonantiques étaient représentés dans cette tâche.

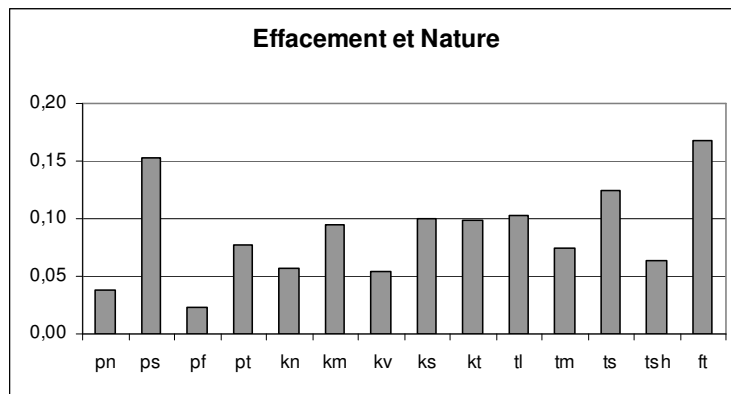
Groupe Consonantique	Taux			
	Epenthèses		Effacements	
	Moyenne	Ecart-type	Moyenne	Ecart-type
/tl/ (code : 10)	36, 69%	5, 06%	10, 22%	2, 77%
/pn/ (code : 1)	33, 4%	7, 9%	3, 78%	2, 95%
/kn/ (code : 5)	45, 8%	6, 34%	5, 67%	3, 98%
/km/ (code : 6)	37, 54%	3, 53%	9, 52%	3, 11%
/tm/ (code : 11)	41, 04%	4, 3%	7, 42%	2, 4%
/pf/ (code : 3)	34, 03%	9, 47%	2, 31%	2, 61%
/kv/ (code : 7)	45, 17%	5, 88%	5, 46%	3, 82%
/ps/ (code : 2)	24, 37%	4, 21%	15, 27%	3, 51%
/ks/ (code : 8)	43%	3, 78%	9, 94%	3, 12%
/ts/ (code : 12)	21, 71%	4, 32%	12, 46%	3, 22%
/tsh/ (code : 13)	27, 03%	4, 89%	6, 3%	2, 74%
/pt/ (code : 4)	37, 54%	4, 13%	7, 7%	2, 69%
/kt/ (code : 9)	41, 88%	3, 41%	9, 8%	2, 54%
/ft/ (code : 36)	36, 13%	11, 24%	16, 81%	11, 86%



Remarques :

- En ce qui concerne les épenthèses, voici le classement décroissant, selon le taux d'épenthèses : /kn/ (environ 46%) > /kv/ (environ 45%) > /ks/ (43%) > /kt/ (environ 42%) > /km/ (environ 38%), /pt/ (environ 38%) > /tl/ (environ 37%) > /ft/ (environ 36%) > /pf/ (environ 34%) > /pn/ (environ 33%) > /tsh/ (environ 27%) > /ps/ (environ 24%) > /ts/ (environ 22%).
- Résultats remarquables : statut particulier des groupes K+C (plus de 40%), statut particulier des groupes C+Sib (moins de 30%).
- Les écarts-types sont faibles, sauf pour le groupe /ft/.

- **Plus d'épenthèses pour les groupes /k/ + C (38% -46%).**
- **Niveau intermédiaire pour les groupes /p/ + C (33% -38%).**
- **Moins d'épenthèses pour les groupes C + Sib (22% -27%) (essentiellement /t/ + Sib).**

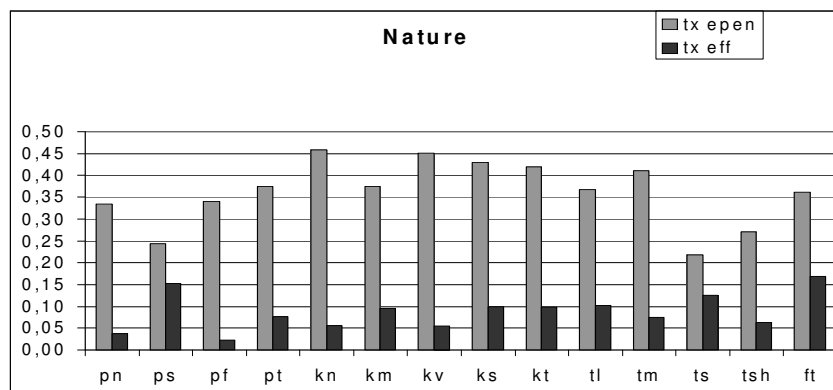


Remarques :

- En ce qui concerne les effacements, voici le classement décroissant, selon le taux d'effacements : /ft/ (environ 17%) > /ps/ (environ 15%) > /ts/ (environ 12%) > /tl/ (environ 10%), /ks/ (environ 10%), /kt/ (environ 10%), /km/ (environ 10%) > /pt/ (environ 8%) > /tm/ (environ 7%) > /tsh/ (environ 6%), /kn/ (environ 6%) > /kv/ (environ 5%) > /pn/ (environ 4%) > /pf/ (environ 2%).
- Pas de résultat remarquable.
- Les écarts-types sont faibles, sauf pour le groupe /ft/.

**Pas de régularité remarquable entre effacements et natures spécifiques.**

Comparaison Epenthèses/Effacements en fonction de la nature spécifique



Remarques :

- Pour les épenthèses et les effacements : écarts-types élevés pour le groupe /ft/.  
 Différence d'ordre pour les épenthèses et effacements, en particulier les groupes /kv/ et /kn/ situés en tête de classement pour les épenthèses mais en fin pour les effacements.

**Certains groupes déclenchent des réactions opposées :**

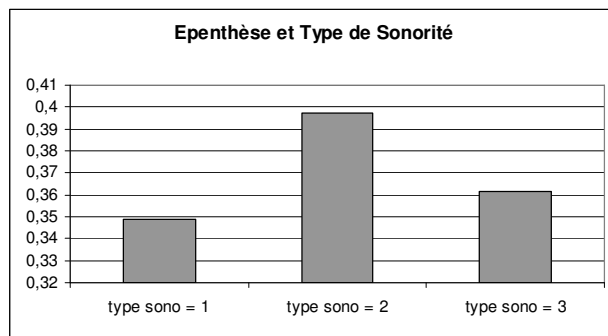
**Plus d'épenthèses et moins d'effacements (ex : /kn/ et /kv/) et vice-versa (ex : /ps/ et /ts/).**

β) Sonorité

β).1. Type de Sonorité

Type de sonorité	Taux			
	Epenthèses		Effacements	
	Moyenne	Écart-type	Moyenne	Écart-type
<b>Montant</b>	34, 86%	13, 31%	8, 44%	7, 03%
<b>Plateau</b>	39, 7%	14, 17%	9, 03%	10, 6%
<b>Descendant</b>	36, 13%	22, 48%	16, 8%	23, 72%

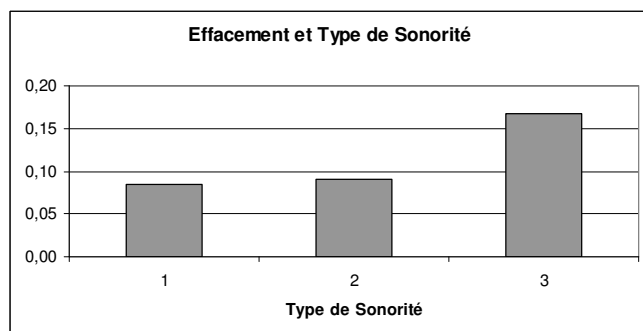




Remarques :

- En ce qui concerne les épenthèses, les écarts entre les trois types sont faibles, avec le classement suivant (par ordre décroissant) : Plateau (environ 40%) > descendant (environ 36%) > montant (environ 35%).

**Légèrement plus d'épenthèses (+ 5%) pour le type de sonorité « Plateau ».**

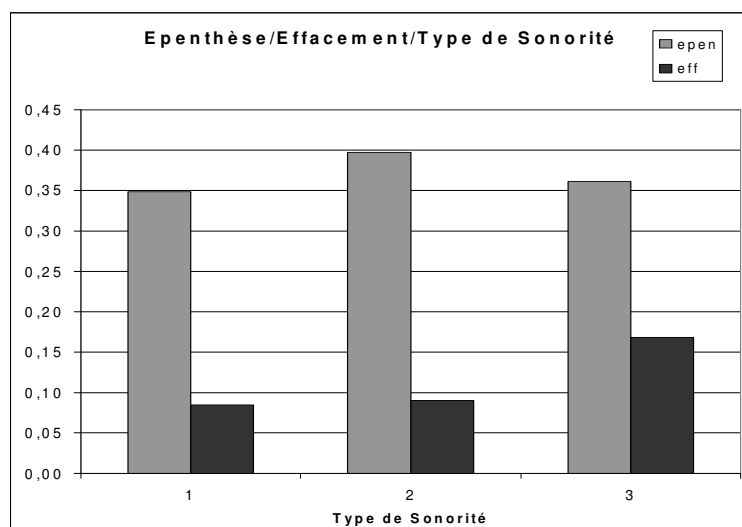


Remarques :

- En ce qui concerne les effacements, le classement est le suivant : descendant (environ 17%) > plateau (environ 9%) > montant (environ 8%).

**Légèrement plus d'effacements (+8%) pour le type de sonorité « descendant ».**

### Comparaison Epenthèses/Effacements en fonction du Type de Sonorité



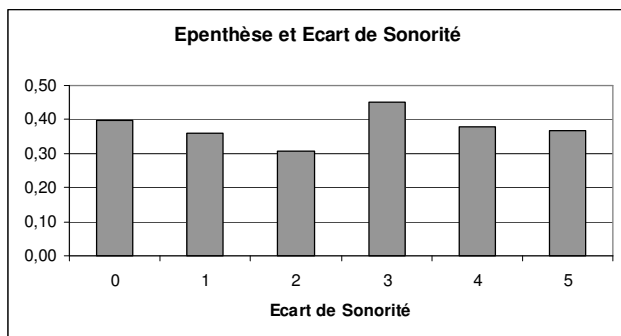
#### Remarques :

- Pas de résultat remarquable.

**Plus d'épenthèses pour le type « Plateau », mais plus d'effacements pour le type « Descendant ».**

#### β).2. Ecart de Sonorité

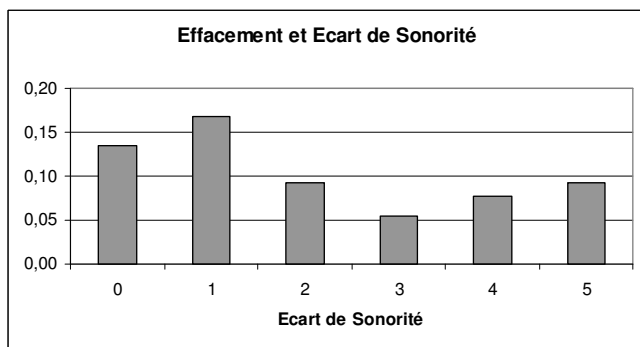
Ecart de sonorité	Taux			
	Epenthèses		Effacements	
	Moyenne	Ecart-type	Moyenne	Ecart-type
<b>0</b>	39, 71%	19, 25%	13, 55%	14, 38%
<b>1</b>	36, 13%	22, 48%	16, 81%	23, 72%
<b>2</b>	30, 53%	15, 27%	9, 27%	10, 71%
<b>3</b>	45, 17%	23, 51%	5, 46%	15, 28%
<b>4</b>	37, 94%	17, 2%	7, 72%	11, 16%
<b>5</b>	36, 69%	30, 35%	9, 24%	16, 64%



Remarques :

- En ce qui concerne les épenthèses, voici le classement décroissant, selon le taux d'épenthèses : écart 3 (environ 45%) > écart 0 (environ 40%) > écart 4 (environ 38%) > écart 5 (environ 37%) > écart 1 (environ 36%) > écart 2 (environ 31%).
- Résultats remarquables : statut particulier des écarts 3 (environ 45%) et 2 (environ 31%).
- Les écarts-types sont élevés, en particulier pour l'écart 5.

**Plus d'épenthèses pour Ecart 3 (45%) et moins pour Ecart 2 (31%).**

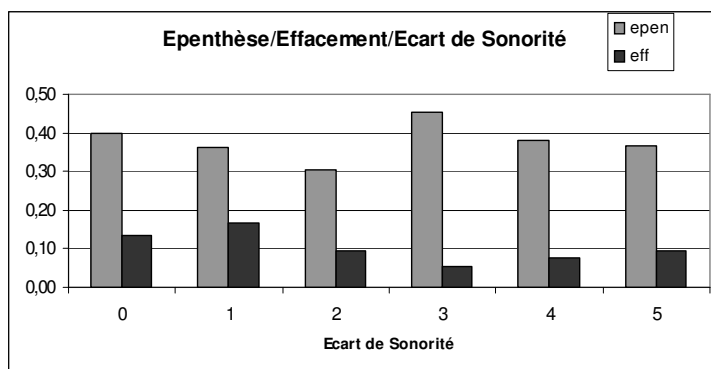


Remarques :

- En ce qui concerne les effacements, voici le classement décroissant, selon le taux d'effacements : écart 1 (environ 17%) > écart 0 (environ 14%) > écart 2 (environ 9%), écart 5 (environ 9%) > écart 4 (environ 8%) > écart 3 (environ 5%).
- Résultats remarquables : statut particulier des écarts 1 (environ 17%) et 3 (environ 5%).
- Les écarts-types sont élevés, en particulier pour l'écart 1.

**Plus d'effacements pour Ecart 1 (17%) et moins pour Ecart 3 (5%).**

### Comparaison Epenthèses/Effacements en fonction de l'Écart de sonorité



#### Remarques :

- Pour les épenthèses et les effacements : statut particulier de l'écart 3, situé en tête pour les épenthèses, mais en fin pour les effacements.

**Écart 3 déclenche plus d'épenthèses et moins d'effacements.**

### Résumé Essentiel Axe Paradigmatique

**La nature phonologique des groupes influence le taux d'épenthèses :**

- On observe davantage d'épenthèses pour : Obs+Cont (40%) (/kn, kv, ks, kt/).
- On observe moins d'épenthèses pour : Obs+Sib (29%) (/tsh, ps, ts/).

### **B) Résultats inférentiels : analyse de variance (Test F) :**

Rappelons que l'influence des variables indépendantes sur la variable dépendante est considérée significative lorsque  $F > F$  critique et  $p < 0,05$  (très significative lorsque  $p < 0,01$ ).

#### Un facteur :

Il s'agit ici d'une analyse de variance univariée destinée à déterminer si une variable indépendante influe significativement sur les taux d'erreurs, d'épenthèses et d'effacements.

<b>Variable Indépendante</b>	<b>Variable Dépendante</b>	<b>F</b>	<b>F critique</b>	<b>p</b>	<b>Relation</b>
Position	Taux d'erreurs	206,5	3,9	1,78E-29	Significative
	Taux d'épenthèses	80,8	3,9	1,38E-15	Significative
	Taux d'effacements	22	3,9	6,09E-06	Significative
Famille	Taux d'erreurs	188,1	3,9	8,45E-28	Significative
	Taux d'épenthèses	71,9	3,9	2,58E-14	Significative
	Taux d'effacements	11,7	3,9	0,00078	Significative
Nature	Taux d'erreurs	153,2	3,9	2,50E-24	Significative
	Taux d'épenthèses	56,5	3,9	5,59E-12	Significative
	Taux d'effacements	1,4	3,9	0,23	Non-significative
Type de Sonorité	Taux d'erreurs	212,2	3,9	5,55E-30	Significative
	Taux d'épenthèses	83,6	3,9	5,61E-16	Significative
	Taux d'effacements	25,8	3,9	1,13E-6	Significative
Ecart de Sonorité	Taux d'erreurs	201,1	3,9	5,34E-29	Significative
	Taux d'épenthèses	78,3	3,9	3,12E-15	Significative
	Taux d'effacements	18,8	3,9	0,000026	Significative

**Seule la relation Nature/Effacement semble non-significative.**

Deux facteurs :

Ici sont examinées les variations des épenthèses et des effacements en fonction de la position et la nature. Il s'agit donc d'une analyse de variance bivariée, puisqu'il s'agit de déterminer s'il existe une interaction entre ces deux variables indépendantes, susceptible d'influer significativement sur les taux d'épenthèses et d'effacements.

<b>Variable Indépendante 1</b>	<b>Variable Indépendante 2</b>	<b>Variable Dépendante</b>	<b>F</b>	<b>F critique</b>	<b>p</b>	<b>Relation</b>
Position	Famille	Taux d'erreurs	1, 44	3, 05	0, 24	Non-significative
		Taux d'épenthèses	0, 04	3,05	0, 95	Non-significative
		Taux d'effacements	0, 45	3,05	0, 63	Non-significative
	Nature	Taux d'erreurs	0, 66	3, 05	0, 51	Non-significative
		Taux d'épenthèses	0, 14	3,05	0, 86	Non-significative
		Taux d'effacements	1, 22	3,05	0, 3	Non-significative
	Type de Sonorité	Taux d'erreurs	1, 47	3, 05	0, 23	Non-significative
		Taux d'épenthèses	0, 049	3,05	0, 9	Non-significative
		Taux d'effacements	4, 13	3,05	0, 018	Significative

	Ecart de Sonorité	Taux d'erreurs	1, 61	3, 05	0, 20	Non-significative
		Taux d'épenthèses	0, 05	3,05	0, 94	Non-significative
		Taux d'effacements	4, 61	3,05	0, 011	Significative

**Seules semblent significatives les relations :**

- **Position/Type de Sonorité/Effacement.**
- **Position/Ecart de Sonorité/Effacement.**

**Résumé Essentiel Résultats Inférentiels**

- **Univariée : seule la relation Nature/Effacement semble non-significative.**
- **Bivariée : seule la relation Position/Sonorité (type et écart)/Effacement semble significative.**

## **C) Résumé**

### **C) 1) Résumé détaillé**

La tâche C1 était une tâche de comptage syllabique auditif. Il s'agissait en particulier d'observer le traitement des groupes consonantiques /pn, ps, pf, pt, kn, km, kv, ks, kt, tl, tm, ts, tsh/ (Plosive sourde + C) ainsi que /ft/ en position finale, par les sujets japonais. Nous supposons que ces groupes déclencheraient des épenthèses vocaliques, tandis que les stimuli pour lesquels ces groupes avaient été artificiellement séparés par l'insertion de voyelles épenthétiques propres au japonais pourraient donner lieu à des effacements/dévoisements vocaliques, étant donné le phénomène de dévoisement vocalique que l'on trouve en japonais. Nous voulions tester, d'une part l'influence de la position du groupe dans le mot, et d'autre part l'influence de la nature de ces groupes sur les réponses des sujets. Les taux indiqués ci-dessous sont des approximations, afin d'en faciliter la lecture. Le terme de « difficulté » employé ci-après, est à relier au taux d'erreurs (épenthèses ou effacements) et ainsi à l'incapacité des sujets à préserver l'identité de l'input dans la tâche considérée.

## **A) Résultats Descriptifs**

### **1) Erreurs globales :**

Le taux d'erreur est globalement élevé (44%), avec une majorité d'épenthèses (81%) et un taux d'effacement non négligeable (19%).

Position : Le taux d'erreur est légèrement plus élevé en positions initiale (taux : 45%) et finale (50%) qu'en position médiale (taux : 40%).

<b>Plus difficile pour les étudiants</b>	<b>Moins difficile pour les étudiants</b>
Position Initiale	Position Médiale
Position Finale	



## **2) Epenthèses vocaliques :**

Le taux global d'épenthèses pour les groupes /CC/ est élevé : 64%.

Position : En ce qui concerne la position, le taux d'épenthèses est plus élevé en positions initiale (taux : 35%) et finale (taux : 16%) qu'en position médiale (taux : 7%).

<b>Plus difficile pour les étudiants</b>	<b>Moins difficile pour les étudiants</b>
Position Initiale	Position Médiale
Position Finale	

Nature :

A) Famille : En ce qui concerne la nature, si l'on considère d'abord les familles, il apparaît que les taux d'épenthèses sont similaires (39%-37%) pour les familles Obs+Nas, Obs+Cont, Obs+Stop et Obs+Lat, alors que la famille Obs+Sib correspond à un taux nettement plus faible (29%).

<b>Plus difficile pour les étudiants</b>	<b>Moins difficile pour les étudiants</b>
Obs+Cont	Obs+Sib
Obs+Nas	
Obs+Stop	
Obs+Lat	

B) Nature spécifique : Si l'on considère maintenant individuellement les groupes consonantiques, voici ce que l'on obtient :

<b>Plus difficile pour les étudiants</b>	<b>Difficulté intermédiaire</b>	<b>Moins difficile pour les étudiants</b>
/kn/ (46%)	/pt/ (38%)	/tsh/ (27%)
/kv/ (45%)	/tl/ (37%)	/ps/ (24%)
/ks/ (43%)	/ft/ (36%) <sup>181</sup>	/ts/ (22%)
/kt/ (42%)	/pf/ (34%)	
/km/ (38%)	/pn/ (33%)	
/k/ + C (46% - 38%)	/p/ + C (38% - 33%)	C +Sib (27% - 22%) (en particulier /t/ + Sib)

C) Type/Profil de sonorité : Si l'on considère le profil de sonorité des groupes consonantiques, voici ce que l'on obtient :

<b>Plus difficile pour les étudiants</b>	<b>Moins difficile pour les étudiants</b>
Plateau (40%)	Descendant (36%)
	Montant (35%)

D) Ecart de sonorité : Si l'on considère l'écart de sonorité entre les deux segments composant le groupe, voici ce que l'on obtient :

<b>Plus difficile pour les étudiants</b>	<b>Difficulté intermédiaire</b>	<b>Moins difficile pour les étudiants</b>
Ecart 3 (45%)	Ecart 0 (40%)	Ecart 2 (31%)
	Ecart 4 (38%)	
	Ecart 5 (37%)	
	Ecart 1 (36%)	

<sup>181</sup> En position finale, avec un écart-type particulièrement élevé.

### **3) Effacements/Réductions vocaliques :**

Position : En ce qui concerne la position, le taux d'effacements est plus élevé en position finale (taux : 16%) qu'en positions initiale (taux : 7%) et médiale (taux : 5%).

<b>Plus difficile pour les étudiants</b>	<b>Moins difficile pour les étudiants</b>
Position Finale	Position Initiale Position Médiale

Nature :

A) Famille : En ce qui concerne la nature, si l'on considère d'abord les familles, il apparaît que les taux d'effacements sont similaires (11%-10%) pour les familles Obs+Sib, Obs+Lat et Obs+Stop, alors que la famille Obs+Cont correspond à un taux plus faible (4%). La famille Obs+Nas est intermédiaire (7%).

<b>Plus difficile pour les étudiants</b>	<b>Difficulté intermédiaire</b>	<b>Moins difficile pour les étudiants</b>
Obs+Sib Obs+Lat Obs+Stop	Obs+Nas	Obs+Cont

B) Nature spécifique : Si l'on considère maintenant individuellement les environnements consonantiques, voici ce que l'on obtient :

<b>Plus difficile pour les étudiants</b>	<b>Difficulté intermédiaire</b>	<b>Moins difficile pour les étudiants</b>
/ft/ (17%) <sup>182</sup>	/tl/ (10%)	/tm/ (7%)
/ps/ (15%)	/ks/ (10%)	/tsh/ (6%)
/ts/ (12%)	/kt/ (10%)	/kn/ (6%)
	/km/ (10%)	/kv/ (5%)
	/pt/ (8%)	/pn/ (4%)
		/pf/ (2%)

C) Type/Profil de sonorité : Si l'on considère le profil de sonorité des environnements consonantiques, voici ce que l'on obtient :

<b>Plus difficile pour les étudiants</b>	<b>Moins difficile pour les étudiants</b>
Descendant (17%)	Plateau (9%)
	Montant (8%)

---

<sup>182</sup> En position finale, avec un écart-type particulièrement élevé.

D) Ecart de sonorité : Si l'on considère maintenant l'écart de sonorité entre les deux segments consonantiques, voici ce que l'on obtient :

Plus difficile pour les étudiants	Difficulté intermédiaire	Moins difficile pour les étudiants
Ecart 1 (17%) Ecart 0 (14%)	Ecart 2 (9%) Ecart 5 (9%) Ecart 4 (8%)	Ecart 3 (5%).

#### **4) Comparaison entre épenthèses et effacements :**

Si l'on compare les résultats des épenthèses et des effacements, on remarque des convergences et des divergences (mais il ne s'agit pas d'incompatibilité ou d'incohérence, au contraire).

##### Position :

- ◆ Convergences : La position 2 semble davantage préserver l'input que la position 1.
- ◆ Divergences : Plus d'épenthèses en position 1 qu'en 3, mais plus d'effacements en position 3 qu'en 1.

##### Nature :

- ◆ Convergences :
  - Ecart-type : ils sont relativement faibles pour les résultats des sous-familles et des natures spécifiques<sup>183</sup>. Ils sont relativement élevés pour les résultats d'écarts de sonorité.
- ◆ Divergences :
  - Famille : le classement de difficulté est, dans une certaine mesure, inversé (Obs+Cont en tête d'épenthèse et fin d'effacement, et vice-versa pour Obs+Sib).

---

<sup>183</sup> Excepté pour le groupe /ft/.

-Nature spécifique : on constate également une différence de classement, en particulier une inversion pour /kv/ et /kn/ (tête d'épenthèse et fin d'effacement), tout comme, dans une moindre mesure, pour /ts/ et /ps/ (fin d'épenthèse, tête d'effacement).

-Type de sonorité : Le type « Plateau » est en tête pour les épenthèses, alors que c'est le type « Descendant » pour les effacements. Par ailleurs, le type « Descendant » se distingue peu (1%) du type « Montant » pour les épenthèses ; il s'en distingue davantage (plus du double, 9%) pour les effacements.

-Ecart de sonorité : L'écart 3 est en tête pour les épenthèses (45%), alors qu'il est en fin pour les effacements (5%).

## **B) Résultats Inférentiels**

### 1) Relation entre une variable indépendante et une variable dépendante :

Les analyses de variance (tests F), suggèrent des relations significatives entre les variables suivantes :

	<b>Position</b>	<b>Famille</b>	<b>Nature</b>	<b>Type de sonorité</b>	<b>Ecart de sonorité</b>
<b>Erreur</b>	Oui	Oui	Oui	Oui	Oui
<b>Epenthèse</b>	Oui	Oui	Oui	Oui	Oui
<b>Effacement</b>	Oui	Oui	Non	Oui	Oui

### 2) Relation entre deux variables indépendante et une variable dépendante :

Les analyses de variance (tests F), suggèrent des relations significatives entre les variables suivantes :

<b>POSITION</b>	<b>Famille</b>	<b>Nature</b>	<b>Type de sonorité</b>	<b>Ecart de sonorité</b>
<b>Erreur</b>	Non	Non	Non	Non
<b>Epenthèse</b>	Non	Non	Non	Non
<b>Effacement</b>	Non	Non	Oui	Oui

C) 2) Résumé essentiel

**Résumé**

**On observe 64% d'épenthèses pour les groupes /CC/ de la tâche C1 (unimodalité auditive).**

**L'influence de la position sur le taux d'épenthèse n'est pas évidente. Cependant :**

- **Il y a davantage d'épenthèses en attaque de la première syllabe (position initiale : 35%).**
- **Il y a moins d'épenthèses en attaque de la dernière syllabe (position médiale : 7%).**

**La nature phonologique des groupes influence le taux d'épenthèses.**

- **On observe davantage d'épenthèses pour : Obs+Cont (40%) (/kn, kv, ks, kt/).**
- **On observe moins d'épenthèses pour : Obs+Sib (29%) (/tsh, ps, ts/).**

**Les analyses inférentielles suggèrent :**

- **Une relation significative entre les variables VI et VD, excepté Nature et Effacement.**
- **Pas de relation significative entre la Position et les autres variables VI, excepté la Sonorité (type et écart), pour la variable Effacement seulement.**

## Tâche C2

### A) Résultats descriptifs

#### Résultats Globaux

Type	TAUX		
	Erreurs	Épenthèses	Effacements
Moyenne	37, 45%	77, 77%	22, 23%
Ecart-type	11, 17%	24, 21%	24, 21%

#### Remarques :

- Taux d'erreur globalement élevé (environ 37%).
- Une majorité d'épenthèses (environ 78%).
- Un taux d'effacement non négligeable (environ 22%).
- Des écarts-types élevés, similaires pour les épenthèses et les effacements (environ 24%) mais plus faible pour les erreurs (environ 11%).

	Taux d'épenthèses pour les stimuli de type /CC/
Moyenne	58, 5%
Ecart-type	25, 63%

#### Remarques :

- Traitement des groupes /CC/ : environ 58% d'épenthèses.

### Résumé Essentiel Résultats Globaux

**On observe 58% d'épenthèses pour les groupes /CC/ de la tâche C2  
(unimodalité auditive).**



A) 1) Axe syntagmatique

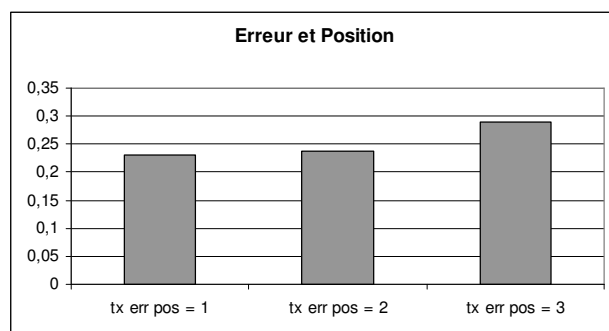
Les nombres de stimuli, selon la position, étaient les suivants :

Position	1	2	3
Nombre de stimuli	38	39	29

Lorsque l'on considère globalement les résultats, sans se soucier de la nature des stimuli (classe générale des groupes /CC/), voici ce que l'on obtient en fonction de la position :

Erreurs :

Position	Taux d'erreurs		
	1	2	3
Moyenne	22, 98%	23, 67%	29, 9%
Ecart-type	8, 13%	11, 53%	10, 89%



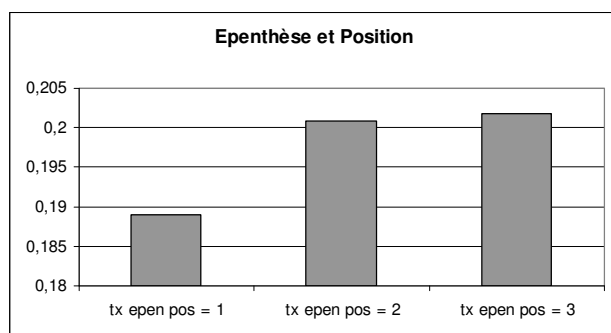
Remarques :

- Taux d'erreurs légèrement plus élevé en position finale (environ 30%) qu'en positions initiale et médiale (environ 23, 5%).
- Pas de résultats remarquables.

**Légèrement plus d'erreurs en coda de la syllabe finale (position 3) (+6, 5% d'écart).**

**Épenthèses :**

Position	Taux d'épenthèses		
	1	2	3
Moyenne	18, 9%	20, 08%	20, 17%
Ecart-type	9, 68%	12, 72%	10, 47%



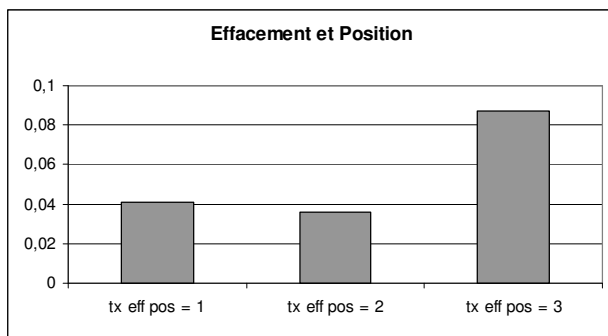
**Remarques :**

- Pas de résultats remarquables (position 1 : environ 18%, position 2 : environ 20%, position 3 : environ 20%, avec des écarts-types similaires).

**Pas de différence majeure entre les trois positions pour les épenthèses.**

**Effacements :**

Position	Taux d'effacements		
	1	2	3
Moyenne	4, 08%	3, 59%	8, 73%
Ecart-type	7, 14%	6, 74%	10, 7%

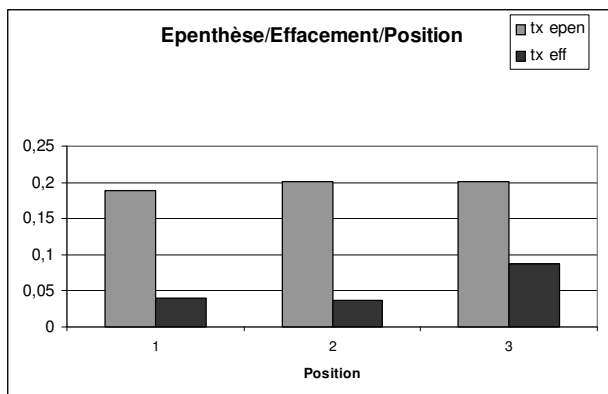


Remarques :

- Ecarts-types élevés.
- Légère dominance, en termes d'effacements, de la position 3 (taux : environ 9% contre 4% pour les positions 1 et 2).
- Pas de résultats remarquables.

**Légèrement plus d'effacements en coda de la syllabe finale (position 3) (+5% d'écart).**

Comparaison Epenthèses/Effacements en fonction de la position :



Remarques : pas de résultats remarquables.

**Résumé Essentiel Axe Syntagmatique**

**La position ne semble pas influencer le taux d'épenthèse.**

A) 2) Axe paradigmatique

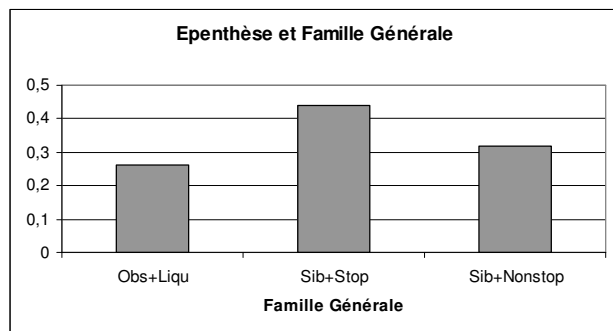
α) Nature phonologique

α) 1) Famille Générale

Trois familles générales étaient représentées dans cette tâche : Obs + Liqu (14 groupes /CC/), Sib + Stop (3 groupes /CC/), Sib + Nonstop (5 groupes /CC/).

**Épenthèses :**

Famille Générale	Taux d'épenthèses		
	Obs + Liqu (exemple : /kl/ (code : 1)	Sib + Stop (exemple : /sk/ (code : 3)	Sib + Nonstop (exemple : /sl/ (code : 4)
<b>Moyenne</b>	26, 13%	44, 02%	31, 9%
<b>Ecart-type</b>	19, 5%	20, 35%	20, 9%



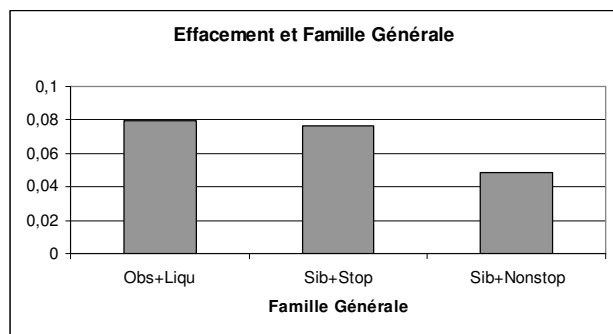
**Remarques :**

- Voici le classement décroissant, selon le taux d'épenthèses : Sib+Stop (environ 44% > Sib+Nonstop (environ 32%) > Obs+Liqu (environ 26%).
- Résultats remarquables : statut particulier des familles générales Sib+Stop (environ 44%) et Obs+Liqu (environ 26%).
- Les écarts-types sont élevés, mais similaires.

- **Plus d'épenthèses pour Sib+Stop (ex : /sk/) (44%).**
- **Moins d'épenthèses pour Obs+Liqu (ex : /kl/) (26%).**

**Effacements :**

Famille Générale	Taux d'effacements		
	Obs + Liqu	Sib + Stop	Sib + Nonstop
Moyenne	7,99%	7,65%	4,84%
Ecart-type	12,71%	12,96%	12,87%



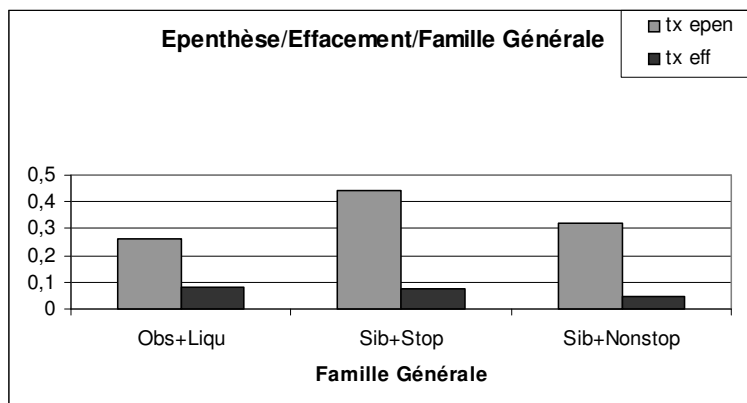
**Remarques :**

- Voici le classement décroissant, selon le taux d'effacements : Obs+Liqu (environ 8%), Sib+Stop (environ 8%) > Sib+Nonstop (environ 5%).
- Résultats remarquables : pas de différence pour les effacements entre les familles générales Sib+Stop et Obs+Liqu (environ 8%).
- Les écarts-types sont élevés, mais similaires.

**Légèrement moins d'effacements pour Sib+Nonstop (ex : /sl/) (-3%).**

**Pas de différence entre Sib+Stop et Obs+Liqu pour les effacements (8%).**

**Comparaison Epenthèses/Effacements en fonction de la famille générale :**



**Remarques :**

- On remarque que, pour les familles générales Sib+Stop et Obs+Liqu, alors que les taux d'épenthèses sont très différents, les taux d'effacements sont identiques.

**Pour Sib+Stop et Obs+Liqu :**

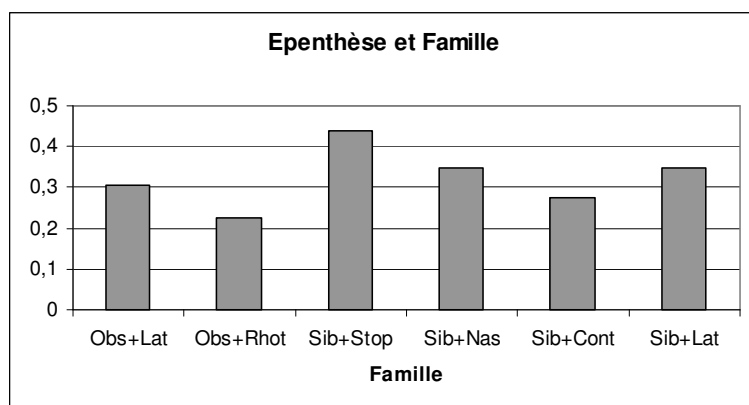
- Grande différence pour les épenthèses (44% et 26%).
- Pas de différence pour les effacements (8%).

**α 2) Famille**

Six familles étaient représentées dans cette tâche, dont une ne possédant qu'un seul représentant.

Famille	Taux			
	Epenthèses		Effacements	
	Moyenne	Ecart-type	Moyenne	Ecart-type
Obs + Lat (code : 1) (exemple : /bl/)	30, 54%	20, 24%	6, 51%	12, 49%
Obs + Rhot (code : 2) (exemple : /bR/)	22, 43%	19, 05%	9, 23%	12, 89%

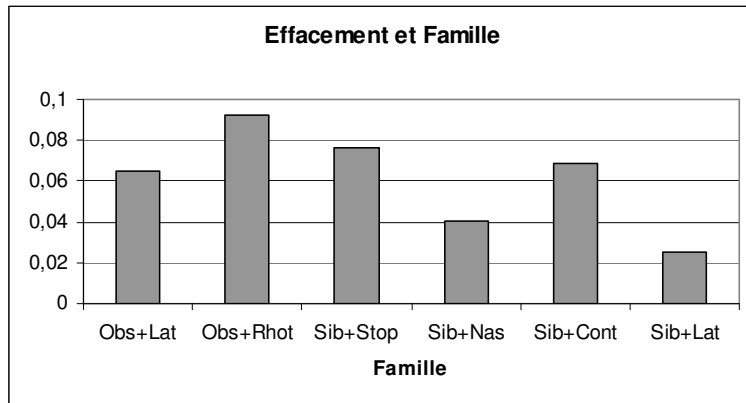
Sib +Stop (code : 7) (exemple : /sp/)	44, 02%	20, 35%	7, 65%	12, 96%
Sib + Nas (code : 8) (exemple : /sm/)	34, 63%	20, 27%	4, 07%	12, 3%
Sib + Cont (code : 9) (exemple : /sv/)	27, 5%	19, 06%	6, 88%	14, 55%
Sib + Lat (code : 10) (exemple : /sl/)	34, 58%	25, 98%	2, 5%	10, 81%



Remarques :

- En ce qui concerne les épenthèses, voici le classement décroissant, selon le taux d'épenthèses : Sib+Stop (environ 44%) > Sib+Nas (environ 35%), Sib+Lat (environ 35%) > Obs+Lat (environ 31%) > Sib+Cont (environ 28%) > Obs+Rhot (environ 22%).
- Résultats remarquables : Obs+Lat (environ 31%) > Obs+Rhot (environ 22%), statut particulier de Sib+Stop (environ 44%) et de Obs+Rhot (environ 22%).
- Les écarts-types sont élevés, mais similaires, sauf pour Sib+Lat (environ 26% contre environ 20% pour les autres).

**Plus d'épenthèses pour Sib+Stop (44%) et moins d'épenthèses pour Obs+Rhot (22%).  
 Le taux d'épenthèses de Obs+Lat (31%) est supérieur à celui de Obs+Rhot (22%).**



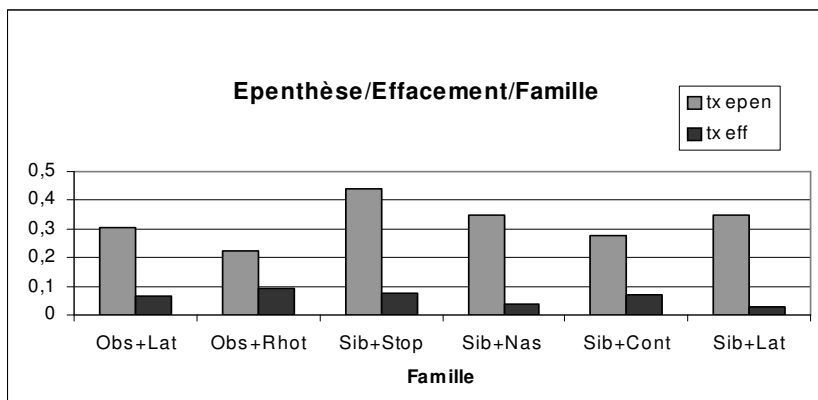
Remarques :

- En ce qui concerne les effacements, voici le classement décroissant, selon le taux d'effacements : Obs+Rhot (environ 9%) > Sib+Stop (environ 8%) > Sib+Cont (environ 7%), Obs+Lat (environ 7%) > Sib+Nas (environ 4%) > Sib+Lat (environ 2%).
- Résultats remarquables : statut particulier de Obs+Rhot (environ 9%) et de Sib+Stop (environ 8%) en tête, ainsi que de Sib+Lat (environ 2%) en fin.
- Les écarts-types sont élevés, mais similaires.

- **Plus d'effacements pour Obs+Rhot (9%) et Sib+Stop (8%).**
- **Moins d'effacements pour Sib+Lat (2%).**



### Comparaison Epenthèses/Effacements en fonction de la famille :



#### Remarques :

- Obs+Rhot est en tête de classement pour les effacements mais en fin pour les épenthèses ; Sib+Lat est en fin de classement pour les effacements mais en deuxième position pour les épenthèses. En ce qui concerne la famille Sib+Stop, elle est en deuxième position pour les effacements et en tête pour les épenthèses.
- Donc pour les épenthèses et les effacements : statuts particuliers de Sib+Stop et de Obs+Rhot.

#### Certains groupes déclenchent des réactions similaires (input non préservé) :

**Plus d'épenthèses et plus d'effacements pour Sib+Stop.**

#### Certains groupes déclenchent des réactions opposées :

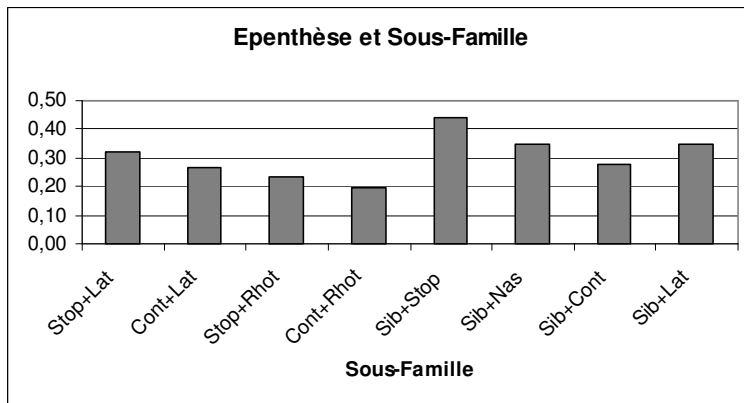
- **Plus d'épenthèses et moins d'effacements pour Sib+Lat.**
- **Moins d'épenthèses et plus d'effacements pour Obs+Rhot.**

#### α) 3) Sous-famille

Huit sous-familles étaient représentées dans cette tâche, dont une ne possédant qu'un seul représentant. Rappelons que les sous-familles 9, 10, 11 et 12 correspondent

exactement aux familles 7, 8, 9, 10 puisqu'il n'y a qu'un seul représentant à chaque fois. Toutes ces familles et sous-familles sont du type S + C.

<b>Sous-famille</b>	<b>Taux</b>			
	<b>Epenthèses</b>		<b>Effacements</b>	
	<b>Moyenne</b>	<b>Ecart-type</b>	<b>Moyenne</b>	<b>Ecart-type</b>
Stop + Lat (code : 1) (exemple : /bl/)	32, 12%	20, 47%	7, 2%	12, 57%
Cont + Lat (code : 2) (exemple : /fl/)	26, 67%	19, 68%	4, 81%	12, 28%
Stop + Rhot (code : 3) (exemple : /bR/)	23, 27%	18, 16%	8, 33%	11, 88%
Cont + Rhot (code : 4) (exemple : /fR/)	19, 81%	21, 14%	12, 04%	16, 03%
Sib + Stop (code : 9) (exemple : /sp/)	44, 02%	20, 35%	7, 65%	12, 96%
Sib + Nas (code : 10) (exemple : /sm/)	34, 63%	20, 27%	4, 07%	12, 3%
Sib + Cont (code : 11) (exemple : /sv/)	27, 5%	19, 06%	6, 88%	14, 55%
Sib + Lat (code : 12) (exemple : /sl/)	34, 58%	25, 98%	2, 5%	10, 81%



Remarques :

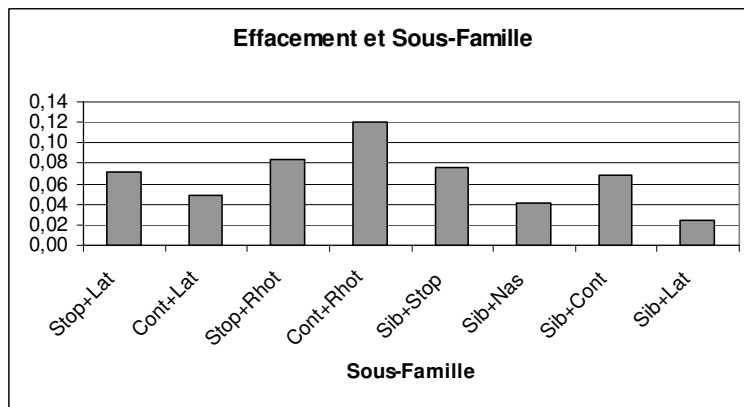
- En ce qui concerne les épenthèses, voici le classement décroissant, selon le taux d'épenthèses : Sib+Stop (environ 44%) > Sib+Nas (environ 35%), Sib+Lat (environ 35%) > Stop+Lat (environ 32%) > Sib+Cont (environ 28%) > Cont+Lat (environ 27%) > Stop + Rhot (environ 23%) > Cont + Rhot (environ 20%).
- Résultats remarquables : Statut particulier de Sib+Stop (environ 44%) et Cont+Rhot (environ 20%). Stop+Lat (environ 32%) > Cont+Lat (environ 27%) et Stop+Rhot (environ 23%) > Cont+Rhot (environ 20%).
- Les écarts-types sont élevés, mais similaires.

**Plus d'épenthèses pour Sib+Stop (44%) et moins d'épenthèses pour Cont+Rhot (20%).**

**Le taux d'épenthèses de Stop+Lat (32%) est supérieur à celui de Cont+Lat (27%).**

**Le taux d'épenthèses de Stop+Rhot (23%) est supérieur à celui de Cont+Rhot (20%).**

**Donc le taux de Stop+Liqu est supérieur à celui de Cont+Liqu pour les épenthèses.**



Remarques :

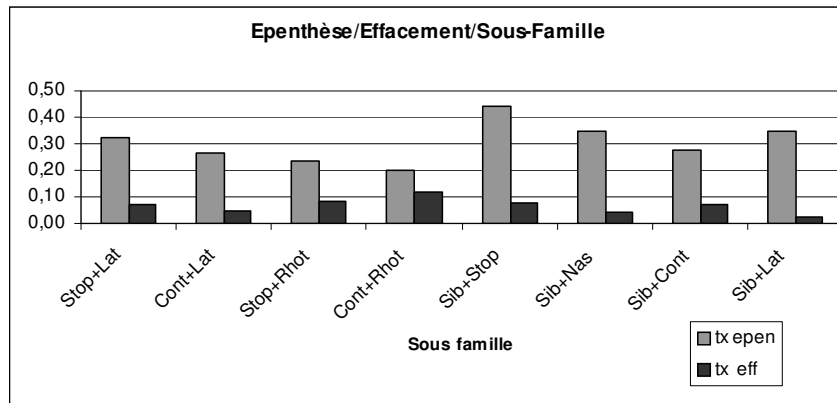
- En ce qui concerne les effacements, voici le classement décroissant, selon le taux d'effacements : Cont+Rhot (environ 12%) > Stop+Rhot (environ 8%), Sib+Stop (environ 8%) > Stop+Lat (environ 7%), Sib+Cont (environ 7%) > Cont+Lat (environ 5%) > Sib+Nas (environ 4%) > Sib+Lat (environ 2%).
- Résultats remarquables : Statuts particuliers de Cont+Rhot (environ 12%) et de Sib+Lat (environ 2%). Stop+Lat (7%) > Cont+Lat (5%) mais Stop+Rhot (8%) < Cont+Rhot (12%).
- Les écarts-types sont élevés, mais similaires, avec une légère dominance pour Cont+Rhot.

**Plus d'effacements pour Cont+Rhot (12%) et moins d'effacements pour Sib+Lat (2%).**

**Le taux d'effacement de Stop+Lat (7%) est supérieur à celui de Cont+Lat (5%).**

**Mais Le taux de Stop+Rhot (8%) est inférieur à celui de Cont+Rhot (12%).**

### Comparaison Epenthèses/Effacements en fonction de la sous-famille



#### Remarques :

- On remarque une inversion du classement pour les sous-familles Cont+Rhot et Stop+Rhot (en première et deuxième positions pour les effacements, en dernière et avant dernière position pour les épenthèses). En revanche, la sous-famille Sib+Stop, première pour les épenthèses, reste en deuxième position pour les effacements.
- Donc pour les épenthèses et les effacements : Statuts particuliers de Sib+Stop et de Cont+Rhot.

#### Certains groupes déclenchent des réactions similaires (input non préservé) :

**Plus d'épenthèses et plus d'effacements pour Sib+Stop.**

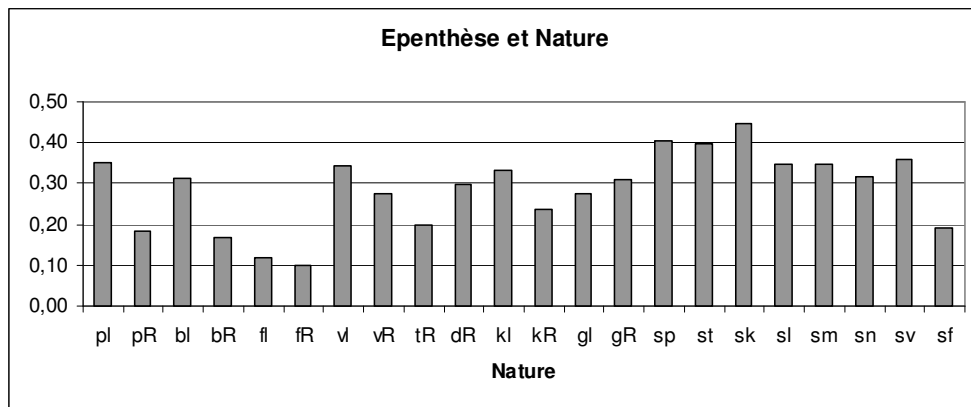
#### Certains groupes déclenchent des réactions opposées :

**Moins d'épenthèses et plus d'effacements pour Cont+Rhot et Stop+Rhot.**

α) 4) Nature spécifique

Vingt-deux groupes consonantiques étaient représentés dans cette tâche.

Groupe Consonantique	Taux			
	Epenthèses		Effacements	
	Moyenne	Ecart-type	Moyenne	Ecart-type
/p/ (code : 14)	35%	25, 84%	6, 67%	15, 43%
/b/ (code : 16)	31, 11%	26, 38%	8, 06%	18, 35%
/k/ (code : 24)	33, 33%	24, 16%	6, 67%	16, 63%
/g/ (code : 26)	27, 5%	23, 9%	7, 5%	16, 95%
/f/ (code 18)	11, 67%	24, 24%	2, 78%	13, 36%
/v/ (code : 20)	34, 17%	24, 56%	5, 83%	15, 74%
/pR/ (code : 15)	18, 33%	28, 09%	6, 67%	15, 03%
/bR/ (code : 17)	16, 67%	26, 82%	13, 89%	24, 42%
/tR/ (code : 22)	20%	26, 17%	9, 72%	18, 87%
/dR/ (code : 23)	29, 67%	21, 72%	11%	22, 26%
/kR/ (code : 25)	23, 67%	29, 21%	2%	8, 65%
/gR/ (code : 27)	30, 83%	34, 37%	8, 75%	18, 96%
/fR/ (code : 19)	10%	18, 17%	16, 25%	25, 66%
/vR/ (code : 21)	27, 67%	29, 54%	8, 67%	16, 1%
/sp/ (code : 28)	40, 28%	23, 37%	8, 61%	18, 36%
/st/ (code : 29)	39, 72%	22, 47%	10, 56%	17, 95%
/sk/ (code : 30)	44, 72%	24, 59%	2, 5%	8, 84%
/sm/ (code : 32)	36, 11%	24, 27%	4, 72%	13, 46%
/sn/ (code : 33)	31, 67%	19, 13%	2, 78%	13, 36%
/sv/ (code : 34)	35, 83%	25, 5%	3, 33%	12, 46%
/sf/ (code : 35)	19, 17%	22, 6%	10, 42%	21, 28%
/sl/ (code : 31)	34, 58%	25, 98%	2, 5%	10, 81%



Remarques :

- En ce qui concerne les épenthèses, voici le classement décroissant, selon le taux d'épenthèses : /sk/ (environ 45%) > /sp/ (environ 40%), /st/ (environ 40%) > /sm/ (environ 36%), /sv/ (environ 36%) > /sl/ (environ 35%), /pl/ (35%) > /vl/ (environ 34%) > /kl/ (environ 33%) > /sn/ (environ 32%) > /bl/ (environ 31%), /gR/ (environ 31%) > /dR/ (environ 30%) > /vR/ (environ 28%) > /gl/ (environ 27%) > /kR/ (environ 24%) > /tR/ (20%) > /sf/ (environ 19%) > /pR/ (environ 18%) > /bR/ (environ 17%) > /fl/ (environ 12%) > /fR/ (environ 10%).
- Résultats remarquables : d'abord, les statuts particuliers de /fl/ (environ 12%) et /fR/ (environ 10%), en fin de classement, et de /sk/ (environ 45%), /sp/ (environ 40%) et /st/ (environ 40%) en tête de classement. Ensuite, on peut distinguer trois grands groupes dans la première partie du classement : A) de /sk/ (environ 45%) à /sl/ (environ 35%) qui sont du type /s+C/, B) de /pl/ (35%) à /bl/ (environ 31%) qui sont du type /C+l/ et C) de /gR/ (environ 31%) à /bR/ (environ 17%) qui sont du type /C+R/. Cependant, si la série A) est homogène, les séries B) et C) ne le sont pas, puisque s'y insèrent le groupe /sn/ (environ 32%) pour la série B) et les groupes /gl/ (environ 27%) et /sf/ (environ 19%) pour la série C). Ce qui attire également l'attention dans ces résultats, est le rapport de dominance, en termes de taux d'épenthèses, entre consonnes voisées et non-voisées, lorsque cette distinction est pertinente :

Groupe Non-voisé	Rapport	Groupe Voisé
/p/ (code : 14) 35%	>	/b/ (code : 16) 31%
/k/ (code : 24) 33%	>	/g/ (code : 26) 27%
/f/ (code 18) 12%	<	/v/ (code : 20) 34%
/pR/ (code : 15) 18%	>	/bR/ (code : 17) 17%
/tR/ (code : 22) 20%	<	/dR/ (code : 23) 30%
/kR/ (code : 25) 24%	<	/gR/ (code : 27) 31%
/fR/ (code : 19) 10%	<	/vR/ (code : 21) 28%
/sf/ (code : 35) 19%	<	/sv/ (code : 34) 36%

- Les écarts-types sont élevés, particulièrement pour /vR/. Les écarts-types les plus réduits sont ceux pour /sn/ et /st/.

**Plus d'épenthèses pour /sk/ (45%) et moins d'épenthèses pour /fl/ (12%) et /fR/ (10%).**

- Plus d'épenthèses pour les groupes /s+C<sup>184</sup> (45%-35%).
- Un peu moins d'épenthèses pour les groupes C+/l/<sup>185</sup> (35%-31%).
- Moins d'épenthèses pour les groupes C+/R/ (31%-10%).

**Influence possible du voisement sur le taux d'épenthèse :**

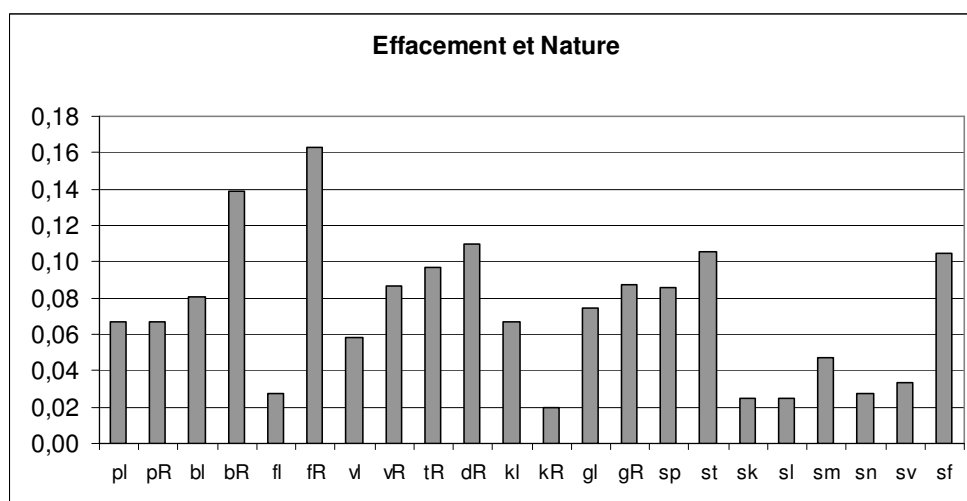
- C non-voisée + /l/ > C voisée + /l/
- C voisée + /R/ > (ou égal)<sup>186</sup> C non-voisée + /R/
- Quand la consonne est /f/ ou /v/, le taux est toujours supérieur pour la voisée /v/

<sup>184</sup> Excepté /sn/ (32%) et /sf/ (19%).

<sup>185</sup> Excepté /gl/ (27%) et /fl/ (12%).

<sup>186</sup> A 1% près pour /pR/ et /bR/.





Remarques :

- En ce qui concerne les effacements, voici le classement décroissant, selon le taux d'effacements : /fR/ (environ 16%) > /bR/ (environ 14%) > /dR/ (11%), /st/ (environ 11%) > /sf/ (environ 10%), /tR/ (environ 10%) > /sp/ (environ 9%), /gR (environ 9%), /vR/ (environ 9%) > /bl/ (environ 8%), /gl/ (environ 8%) > /pR/ (environ 7%), /pl/ (environ 7%), /kl/ (environ 7%) > /vl/ (environ 6%) > /sm/ (environ 5%) > /fl/ (environ 3%), /sv/ (environ 3%), /sn/ (environ 3%) > /sl/ (environ 2%), /sk/ (environ 2%), /kR/ (2%)
- Résultats remarquables : d'abord, les statuts particuliers de /fR/ (16%) et /bR/ (14%) en tête de classement, et ceux de /sl/ (2%), /sk/ (2%) et /kR/ (2%) en dernière position. En ce qui concerne le trait de voisement, voici ce que l'on obtient :

Groupe Non-voisé	Rapport	Groupe Voisé
/pl/ (code : 14) 7%	<	/bl/ (code : 16) 8%
/kl/ (code : 24) 7%	<	/gl/ (code : 26) 8%
/fl/ (code 18) 3%	<	/vl/ (code : 20) 6%
/pR/ (code : 15) 7%	<	/bR/ (code : 17) 14%
/tR/ (code : 22) 10%	<	/dR/ (code : 23) 11%
/kR/ (code : 25) 2%	<	/gR/ (code : 27) 9%
/fR/ (code : 19) 16%	>	/vR/ (code : 21) 9%
/sf/ (code : 35) 10%	>	/sv/ (code : 34) 3%

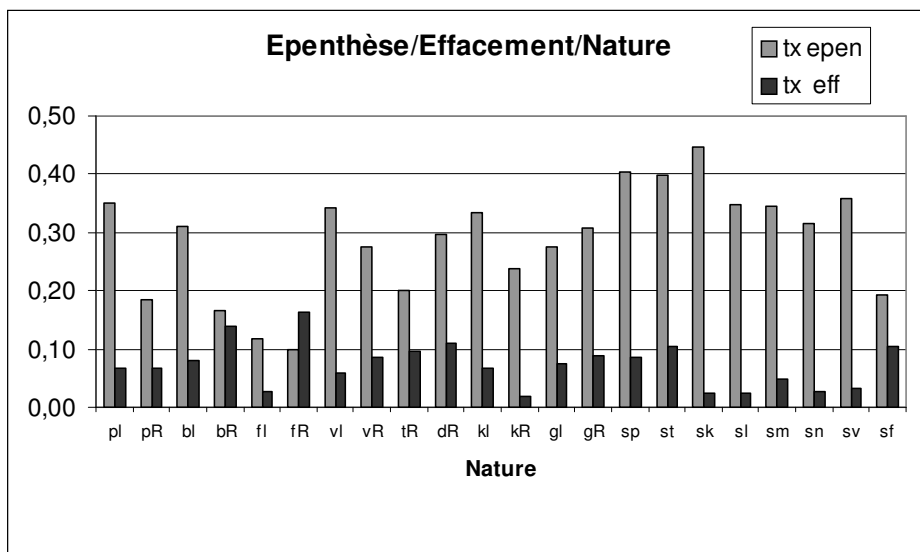
- Les écarts-types sont élevés, en particulier pour /bR/, /dR/, fR/, /sf/ (plus de 20%).

**Plus d'effacements pour /fR/ (16%) et /bR/ (14%) et moins d'effacements pour /sl/, /sk/ et /kR/ (2%). Mais pas de régularités remarquable.**

**Influence possible du voisement sur le taux d'effacements :**

- C non-voisée + /l/ < (ou égal) C voisée + /l/.
- C non-voisée + /R/ < C voisée + /R/.

**Comparaison Epenthèses/Effacements en fonction de la nature spécifique :**



**Remarques :**

- On remarque une inversion de classement pour /fR/ et /bR/, en tête de classement pour les effacements, mais en fin, respectivement en dernière et antépénultième position pour les épenthèses et pour /sk/, en dernière position pour les effacements, mais en tête pour les épenthèses. En revanche, on observe une tendance similaire pour /kR/, en dernière position pour les effacements et dans la deuxième moitié du

classement pour les épenthèses et pour /fl/, dans la deuxième moitié du classement pour les effacements et en avant-dernière position pour les épenthèses

- Donc pour les épenthèses et les effacements : statuts particuliers de /fR/, /sk/ et /fl/ et possible régularité relative des rapports entre groupes voisés et non-voisés.

**Certains groupes déclenchent des réactions opposées :**

- Plus d'épenthèses et moins d'effacements pour /sk/.
- Moins d'épenthèses et plus d'effacements pour /fR/ et /bR/.

**D'autres présentent la même tendance :**

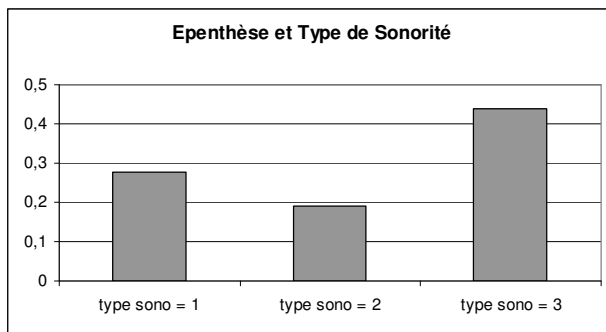
**Moins d'épenthèses et moins d'effacements pour /kR/ et /fl/.**

**Influence possible du voisement sur les taux, tant d'épenthèses que d'effacements.**

*β) Sonorité*

*β) 1) Type de Sonorité*

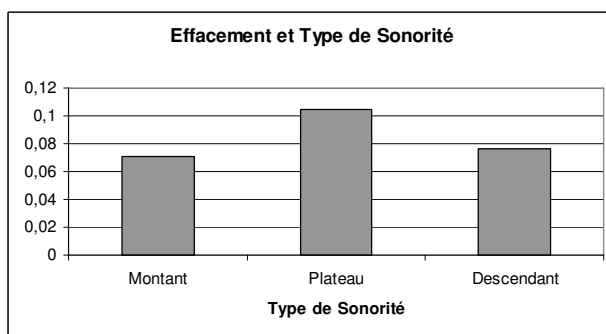
Type de sonorité	Taux			
	Epenthèses		Effacements	
	Moyenne	Ecart-type	Moyenne	Ecart-type
<b>Montant</b>	27, 88%	14, 52%	7, 1%	8, 08%
<b>Plateau</b>	19, 16%	16, 82%	10, 42%	17, 4%
<b>Descendant</b>	44, 02%	14, 75	7, 65%%	9, 82%



Remarques :

- En ce qui concerne les épenthèses, les écarts sont relativement assez marqués et le classement décroissant selon le taux d'épenthèses est le suivant : « Descendant » (environ 44%) > « Montant » (environ 28%) > « Plateau » (environ 19%).

- **Plus d'épenthèses pour type « Descendant » (44%).**
- **Moins d'épenthèses pour type « Plateau » (19%).**

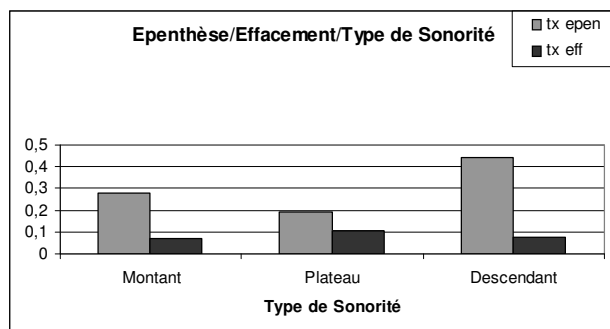


Remarques :

- En ce qui concerne les effacements, les écarts sont assez faibles et le classement décroissant est le suivant : « Plateau » (environ 10%) > « Descendant » (environ 8%) > « Montant » (environ 7%).

**Peu de différences pour les effacements entre les trois types (7%-10%).**

**Comparaison Epenthèses/Effacement en fonction du type de sonorité :**



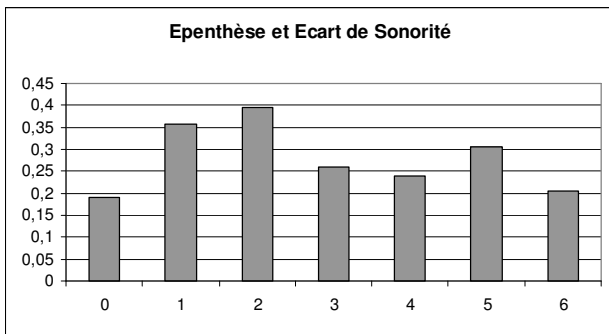
**Remarques :**

- Alors que le type « Plateau » est en tête pour les effacements, il en fin de classement pour les épenthèses.

**Le type « Plateau » déclenche moins d'épenthèses et plus d'effacements.**

**β)2) Ecart de Sonorité**

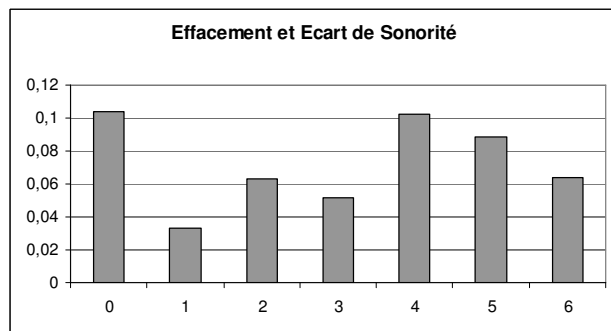
Ecart de sonorité	Taux			
	Epenthèses		Effacements	
	Moyenne	Ecart-type	Moyenne	Ecart-type
<b>0</b>	19, 17%	30, 07%	10, 42%	21, 28%
<b>1</b>	35, 83%	24, 70%	3, 33%	12, 46%
<b>2</b>	39, 53%	14, 52%	6, 3%	10, 36%
<b>3</b>	25, 97%	19, 88%	5, 14%	11, 4%
<b>4</b>	24, 05%	15, 18%	10, 24%	13, 59%
<b>5</b>	30, 49%	15, 81%	8, 82%	12, 65%
<b>6</b>	20, 63%	19, 74%	6, 35%	12, 01%



Remarques :

- En ce qui concerne les épenhèses, voici le classement décroissant, selon le taux d'épenhèses : écart 2 (environ 40%) > écart 1 (environ 36%) > écart 5 (environ 30%) > écart 3 (environ 26%) > écart 4 (environ 24%) > écart 6 (environ 21%) > écart 0 (environ 19%).
- Résultats remarquables : statut particulier de l'écart 2 (environ 40%) et de l'écart 0 (environ 19%).
- Les écarts-types sont élevés, en particulier pour les écarts 0 et 1. En revanche, l'écart-type de l'écart 2 est le plus réduit.

- **Plus d'épenhèses pour les écarts 2 (40%) et 1 (36%).**
- **Moins d'épenhèses pour les écarts 6 (21%) et 0 (19%).**



Remarques :

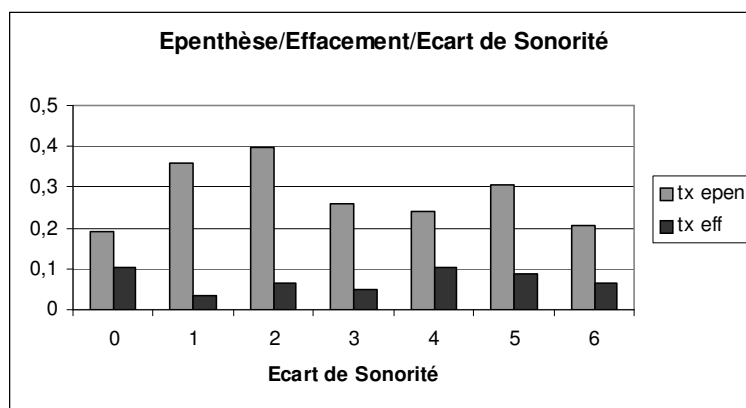
- En ce qui concerne les effacements, voici le classement décroissant, selon le taux d'effacements : écart 0 (environ 10%), écart 4 (environ 10%) > écart 5 (environ 9%) > écart 6 (environ 6%) > écart 2 (environ 6%) > écart 3 (environ 5%) > écart 1 (environ 3%).

9%) > écart 6 (environ 6%), écart 2 (environ 6%) > écart 3 (environ 5%) > écart 1 (environ 3%).

- Résultats remarquables : statut particulier des écarts 0 et 4 (environ 10%) et de l'écart 1 (environ 3%).
- Les écarts-types sont élevés, en particulier pour l'écarts 0.

- **Plus d'effacements pour les écarts 0 (10%), 4 (10%) et 5 (9%).**
- **Moins d'effacements pour l'écart 1 (3%).**

### Comparaison Epenthèses/Effacements en fonction de l'écart de sonorité



#### Remarques :

- Statut particulier de l'écart 0 (en fin de classement pour les épenthèses et en tête pour les effacements).

#### L'écart 0 déclenche des réactions opposées :

- **Moins d'épenthèses.**
- **Plus d'effacements.**

### Résumé Essentiel Axe Paradigmatique

**La nature phonologique des groupes influence le taux d'épenthèses :**

- **On observe davantage d'épenthèses pour : Sib+Stop (44%) (/sk, sp, st/).**
- **On observe moins d'épenthèses pour : Obs+Liqu (26%) (/pl, pR, kl, kR/).**

#### **B) Résultats inférentiels : analyse de variance (Test F)**

Rappelons que l'influence des variables indépendantes sur la variable dépendante est considérée significative lorsque  $F > F$  critique et  $p < 0, 05$  (très significative lorsque  $p < 0, 01$ ).

Un facteur :

Il s'agit ici d'une analyse de variance univariée destinée à déterminer si une variable indépendante influe significativement sur les taux d'erreurs, d'épenthèses et d'effacements.

<b>Variable Indépendante</b>	<b>Variable Dépendante</b>	<b>F</b>	<b>F critique</b>	<b>p</b>	<b>Relation</b>
Position	Taux d'erreurs	189,66	3,88	3,5E-31	Significative
	Taux d'épenthèses	81,9	3,88	9,66E-17	Significative
	Taux d'effacements	16,6	3,88	6,4E-5	Significative
Famille Générale	Taux d'erreurs	188,9	3,88	4,3E-31	Significative
	Taux d'épenthèses	81,74	3,88	1,04E-16	Significative
	Taux d'effacements	16,2	3,88	7,78E-5	Significative
Sous-Famille	Taux d'erreurs	127,31	3,88	2,18E-23	Significative
	Taux d'épenthèses	50	3,88	2,24E-11	Significative
	Taux d'effacements	0,96	3,88	0,33	Non-significative
Nature	Taux d'erreurs	0,6	3,88	0,44	Non-significative



	Taux d'épenthèses	8, 67	3, 88	0, 003	Significative
	Taux d'effacements	531, 5	3, 88	2E-59	Significative
Type de Sonorité	Taux d'erreurs	198, 83	3, 88	3, 28E-32	Significative
	Taux d'épenthèses	86, 15	3, 88	2, 13E-17	Significative
	Taux d'effacements	24, 38	3, 88	1, 6E-6	Significative
Ecart de Sonorité	Taux d'erreurs	158, 17	3, 88	2, 08E-27	Significative
	Taux d'épenthèses	64, 63	3, 88	6, 48E-14	Significative
	Taux d'effacements	1, 79	3, 88	0, 18	Non-significative

**Seules les relations Sous-famille/ Effacement, Nature/ Erreur et Ecart de sonorité/ Effacement semblent non significatives.**

Deux facteurs<sup>187</sup> :

Ici sont examinées les variations des épenthèses et des effacements en fonction de la position et la nature. Il s'agit donc d'une analyse de variance bivariée, puisqu'il s'agit de déterminer s'il existe une interaction entre ces deux variables indépendantes, susceptible d'influer significativement sur les taux d'épenthèses et d'effacements.

Variable Indépendante 1	Variable Indépendante 2	Variable Dépendante	F	F critique	p	Relation
Position	Famille Générale	Taux d'erreurs	1, 72	3, 03	0, 18	Non-significative
		Taux d'épenthèses	2, 12	3, 03	0, 12	Non-significative
		Taux d'effacements	0, 001	3, 03	0, 99	Non-significative

<sup>187</sup> Pour ces analyses, certains résultats ont dû être dupliqués afin d'harmoniser le nombre d'échantillons.

	Sous-Famille	Taux d'erreurs	4,68	3,03	0,01	Significative
		Taux d'épenthèses	0,9	3,03	0,32	Non-significative
		Taux d'effacements	7,36	3,03	0,0008	Significative
	Nature	Taux d'erreurs	4,22	3,03	0,015	Significative
		Taux d'épenthèses	0,83	3,03	0,43	Non-significative
		Taux d'effacements	5,2	3,03	0,006	Significative
	Type de Sonorité	Taux d'erreurs	1,5	3,03	0,22	Non-significative
		Taux d'épenthèses	0,25	3,03	0,78	Non-significative
		Taux d'effacements	5,43	3,03	0,005	Significative
	Ecart de Sonorité	Taux d'erreurs	3,78	3,03	0,025	Significative
		Taux d'épenthèses	0,59	3,03	0,55	Non-significative
		Taux d'effacements	5,71	3,03	0,003	Significative

**Les interactions Position/Sous-famille, Position/Nature, Position/Type de sonorité et Position/Ecart de sonorité semblent significatives pour les effacements.**

**Aucun résultat remarquable pour les épenthèses.**

#### **Résumé Essentiel Résultats Inférentiels**

- **Univariée** : l'effacement ne serait influencé ni par la sous-famille, ni par l'écart de sonorité.
- **Bivariée** : l'effacement semblerait subir l'influence de l'interaction entre position et nature.

## **C) Résumé**

### **C) 1) Résumé Détaillé**

La tâche C2 était une tâche de comptage syllabique auditif similaire à la tâche C1. Seuls différaient les groupes consonantiques cibles, et ainsi les non-mots stimuli (cibles et distracteurs). Ces groupes, pour C2, étaient les suivants : /pl,, pr, bl, br, fl, fr, vl, vr, tr, dr, kl, kr, gl, gr/ (OBLI), /sp, st, sk/ (S + Plosive non-voisée) et /sl, sm, sn, sv, sf/ (S + C non-plosive).

## **A) Résultats Descriptifs**

### **Erreurs globales :**

Le taux d'erreur est globalement élevé (37%), avec une majorité d'épenthèses (78%) et un taux d'effacement non négligeable (22%).

Position : Le taux d'erreur est légèrement plus élevé en position finale (taux : 30%) qu'en positions initiale et médiale (taux : 23% et 24% respectivement).

<b>Plus difficile pour les étudiants</b>	<b>Moins difficile pour les étudiants</b>
Position Finale	Position Initiale Position Médiale

### **Epenthèses vocaliques :**

Le taux global d'épenthèses pour les groupes /CC/ est élevé : 58%.

Position : En ce qui concerne la position, les taux d'épenthèses sont similaires en position initiale (taux : 19%), médiale (taux : 20%) et finale (taux : 20%). Il n'y a donc pas ici de résultat particulièrement remarquable.

Nature :

A) Famille Générale : En ce qui concerne la nature, si l'on considère d'abord les familles générales, les taux d'épenthèses se répartissent ainsi :

Plus difficile pour les étudiants	Difficulté intermédiaire	Moins difficile pour les étudiants
Sib+Stop (44%)	Sib+Nonstop (32%)	Obs+Liqu (26%)

B) Famille : Si l'on considère les familles, voici ce que l'on obtient :

Plus difficile pour les étudiants	Difficulté intermédiaire	Moins difficile pour les étudiants
Sib+Stop (44%)	Sib+Lat (35%) Sib+Nas (35%) Obs+Lat (31%)	Sib+Cont (27%) Obs+Rhot (22%)

C) Sous-famille : Si l'on considère les sous-familles, voici ce que l'on obtient :

Plus difficile pour les étudiants	Difficulté intermédiaire	Moins difficile pour les étudiants
Sib+Stop (44%)	Sib+Lat (35%) Sib+Nas (35%) Stop+Lat (32%)	Sib+Cont (27%) Cont+Lat (27%) Stop + Rhot (23%) Cont + Rhot (20%)

D) Nature spécifique : Si l'on considère enfin les groupes consonantiques individuellement, voici ce que l'on obtient :

Plus difficile pour les étudiants	Difficulté intermédiaire	Moins difficile pour les étudiants
/sk/ (45%) /sp/ (40%) /st/ (40%)  /sm/ (36%) /sv/ (36%) /sl/ (35%)	/p/ (35%) /v/ (34%) /k/ (33%) /sn/ (32%) /b/ (31%)	/f/ (12%) /fR/ (10%)
<b>/s/ + C (45% - 35%)</b>	<b>C + /l/<sup>188</sup> (35% - 31%)</b>	<b>/f/ + C (12% - 10%)</b>
	/gR/ (31%) /dR/ (30%) /vR/ (28%) /g/ (27%) /kR/ (24%) /tR/ (20%) /sf/ (19%) /pR/ (18%) /bR/ (17%)	
	<b>C + /R/<sup>189</sup> (31% - 17%)</b>	

E) Voisement : Les résultats ont conduit à l'observation des tendances suivantes (à 5% près), dans nos données :

<sup>188</sup> Excepté /sn/.

<sup>189</sup> Excepté /gl/ et /sf/.

Groupe	Plus difficile pour les étudiants	Moins difficile pour les étudiants
C + /l/	/C/ Non voisée	/C/ Voisée
C + /R/	/C/ Voisée	/C/ Non voisée
Avec /f/ ou /v/	Voisée (/v/)	Non voisée (/f/)

F) Type/Profil de sonorité : Si l'on considère le profil de sonorité des groupes consonantiques, voici ce que l'on obtient :

Plus difficile pour les étudiants	Difficulté intermédiaire	Moins difficile pour les étudiants
Descendant (44%)	Montant (28%)	Plateau (19%).

G) Ecart de sonorité : Si l'on considère maintenant l'écart de sonorité entre les deux segments consonantiques, voici ce que l'on obtient :

Plus difficile pour les étudiants	Difficulté intermédiaire	Moins difficile pour les étudiants
Ecart 2 (40%)	Ecart 5 (30%)	Ecart 6 (21%)
Ecart 1 (36%)	Ecart 3 (26%)	Ecart 0 (19%)
	Ecart 4 (24%)	

**Effacements/Réductions vocales :**

Position : En ce qui concerne la position, le taux d'effacements est plus élevé en position finale (taux : 9%) qu'en positions initiale et médiale (taux : 4%).

Plus difficile pour les étudiants	Moins difficile pour les étudiants
Position Finale	Position Initiale Position Médiale

Nature :

A) Famille Générale : En ce qui concerne la nature, si l'on considère d'abord les familles générales, les taux d'effacements se répartissent ainsi :

<b>Plus difficile pour les étudiants</b>	<b>Moins difficile pour les étudiants</b>
Obs+Liqu (8%) Sib+Stop (8%)	Sib+Nonstop (5%)

B) Famille: Si l'on considère ensuite les familles, voici ce que l'obtient :

<b>Plus difficile pour les étudiants</b>	<b>Difficulté intermédiaire</b>	<b>Moins difficile pour les étudiants</b>
Obs+Rhot (9%) Sib+Stop (8%)	Sib+Cont (7%) Obs+Lat (7%)	Sib+Nas (4%) Sib+Lat (2%)

C) Sous-Famille: En considérant les sous-familles, voici ce que l'obtient :

<b>Plus difficile pour les étudiants</b>	<b>Difficulté intermédiaire</b>	<b>Moins difficile pour les étudiants</b>
Cont+Rhot (12%)	Stop+Rhot (8%) Sib+Stop (8%) Stop+Lat (7%) Sib+Cont (7%)	Cont+Lat (5%) Sib+Nas (4%) Sib+Lat (2%).

D) Nature Spécifique: En considérant enfin la nature spécifique des environnements consonantiques, voici ce que l'obtient :

Plus difficile pour les étudiants	Difficulté intermédiaire	Moins difficile pour les étudiants
/fR/ (16%)	/sp/ (9%)	/sv/ (3%)
/bR/ (14%)	/gR/ (9%)	/fl/ (3%)
/dR/ (11%)	/vR/ (9%)	/sn/ (3%)
/st/ (11%)	/bl/ (8%)	/sl/ (2%)
/sf/ (10%)	/pR/ (7%)	/sk/ (2%)
/tR/ (10%)	/pl/ (7%)	/kR/ (2%)
	/kl/ (7%)	
	/vl/ (6%)	
	/sm/ (5%)	

E) Voisement : Les résultats ont conduit à l'observation des tendances suivantes (à 5% près), dans nos données :

Groupe	Plus difficile pour les étudiants	Moins difficile pour les étudiants
C + /l/	/C/ Voisée	/C/ Non-voisée
C + /R/	/C/ Voisée	/C/ Non voisée

F) Type/Profil de sonorité : Si l'on considère le profil de sonorité des groupes consonantiques, voici ce que l'on obtient :

Plus difficile pour les étudiants	Difficulté intermédiaire	Moins difficile pour les étudiants
Plateau (10%)	Descendant (8%)	Montant (7%)

G) Ecart de sonorité : Si l'on considère maintenant l'écart de sonorité entre les deux segments consonantiques, voici ce que l'on obtient :

Plus difficile pour les étudiants	Difficulté intermédiaire	Moins difficile pour les étudiants
Ecart 0 (10%)	Ecart 6 (6%)	Ecart 1 (3%)
Ecart 4 (10%)	Ecart 2 (6%)	
Ecart 5 (9%)	Ecart 3 (5%)	



### **Comparaison entre épenthèses et effacements :**

Si l'on compare les résultats des épenthèses et des effacements, on peut remarquer certaines convergences et divergences (mais il ne s'agit pas d'incompatibilité ou d'incohérence, au contraire).

#### Position :

- ◆ Convergences : les positions 1 et 2 semblent davantage préserver l'input que la position 3.

#### Nature :

- ◆ Convergences :
  - Famille Générale : Sib+Stop est plus difficile (input moins préservé) que la famille générale Sib+NonStop.
  - Famille : Sib+Stop se situe au premier rang de difficulté pour les épenthèses et au deuxième rang pour les effacements.
  - Nature Spécifique : /kR/ se trouve dans la deuxième moitié du classement pour les épenthèses et figure au dernier rang pour les effacements. De la même manière, la nature spécifique /fl/ figure à l'avant-dernier rang pour les épenthèses et dans la deuxième moitié du classement pour les effacements.
  
- ◆ Divergences :
  - Famille Générale : inversion du rang de Obs+Liqu (fin de classement pour les épenthèses, mais en tête pour les effacements).
  - Famille : inversion du rang de Obs+Rhot (fin de classement pour les épenthèses, mais en tête pour les effacements), de Sib+Lat (deuxième rang pour les épenthèses, mais fin de classement pour les effacements) et de Sib+Stop (tête pour les épenthèses, mais deuxième rang pour les effacements).

-Sous-Famille : inversion du rang pour Cont+Rhot et Stop+Rhot (dernier et avant-dernier rang pour les épenthèses, mais premier et deuxième rangs pour les effacements,).

-Nature Spécifique : inversion du rang pour /fR/ et /bR/, (fin de classement, respectivement dernier et antépénultième rang, pour les épenthèses, mais tête pour les effacements) et celui de /sk/ (tête pour les épenthèses, mais dernier rang pour les effacements).

-Type de Sonorité : le type « Plateau » est en tête pour les effacements, mais en fin pour les épenthèses.

-Écarts de Sonorité : l'écart 0 figure en fin de classement pour les épenthèses, mais et en tête pour les effacements.

## **B) Résultats Inférentiels**

### 1) Relation entre une variable indépendante et une variable dépendante :

Les analyses de variance (tests F), suggèrent des relations significatives entre les variables suivantes :

	<b>Position</b>	<b>Famille Générale</b>	<b>Sous-Famille</b>	<b>Nature</b>	<b>Type de sonorité</b>	<b>Ecart de sonorité</b>
<b>Erreur</b>	OUI	OUI	OUI	NON	OUI	OUI
<b>Epenthèse</b>	OUI	OUI	OUI	OUI	OUI	OUI
<b>Effacement</b>	OUI	OUI	NON	OUI	OUI	NON

### 2) Relation entre deux variables indépendantes et une variable dépendante :

Les analyses de variance (tests F), suggèrent des relations significatives entre les variables suivantes :

<b>POSITION</b>	<b>Famille Générale</b>	<b>Sous-Famille</b>	<b>Nature</b>	<b>Type de sonorité</b>	<b>Ecart de sonorité</b>
<b>Erreur</b>	NON	OUI	OUI	NON	OUI
<b>Epenthèse</b>	NON	NON	NON	NON	NON
<b>Effacement</b>	NON	OUI	OUI	OUI	OUI

C) 2) Résumé Essentiel

**Résumé**

**On observe 58% d'épenthèses pour les groupes /CC/ de la tâche C2 (unimodalité auditive).**

**La position ne semble pas influencer le taux d'épenthèse, même si celui-ci est maximal en coda de la syllabe finale.**

**La nature phonologique des groupes influence le taux d'épenthèses.**

- **On observe davantage d'épenthèses pour : Sib+Stop (44%) (/sk, sp, st/).**
- **On observe moins d'épenthèses pour : Obs+Liqu (26%) (/pl, pR, kl, kR/).**

**Les analyses inférentielles suggèrent que :**

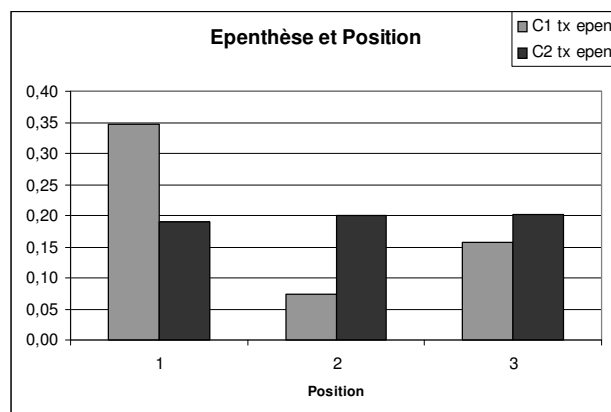
- **Il y a une influence des VI sur les VD. Mais l'effacement ne serait influencé ni par la sous-famille, ni par l'écart de sonorité.**
- **Il n'y a pas interaction entre la position et les autres VI. Mais l'effacement serait influencé par l'interaction position/nature.**

## Comparaison Tâches C1 et C2

Lorsque l'on intègre et compare les résultats des tâches C1 et C2 (qui ne différaient que par la composition et le nombre des stimuli, ainsi que par la population des sujets passant les épreuves), voici ce que l'on obtient :

	Stimuli	Population
<b>C1</b>	<ul style="list-style-type: none"> <li>• /pn, ps, pf, pt, kn, km, kv, ks, kt, tl, tm, ts, tsh/ (Plosive non-voisée + C)</li> <li>• /ft/ en position finale</li> </ul>	n=120
<b>C2</b>	<ul style="list-style-type: none"> <li>• /pl,, pr, bl, br, fl, fr, vl, vr, tr, dr, kl, kr, gl, gr/ (OBLI)</li> <li>• /sp, st, sk/ (S + Plosive non-voisée)</li> <li>• /sl, sm, sn, sv, sf/ (S + C non-plosive)</li> </ul>	n=60

### 1. Position



### Commentaire :

- L'influence de la position sur le taux d'épenthèse est beaucoup plus nette dans C1 que dans C2.

- Il n'y a pas de convergence des résultats entre C1 et C2, excepté les ordres de grandeur, relativement similaires, des taux en position 3 pour C1 et C2.

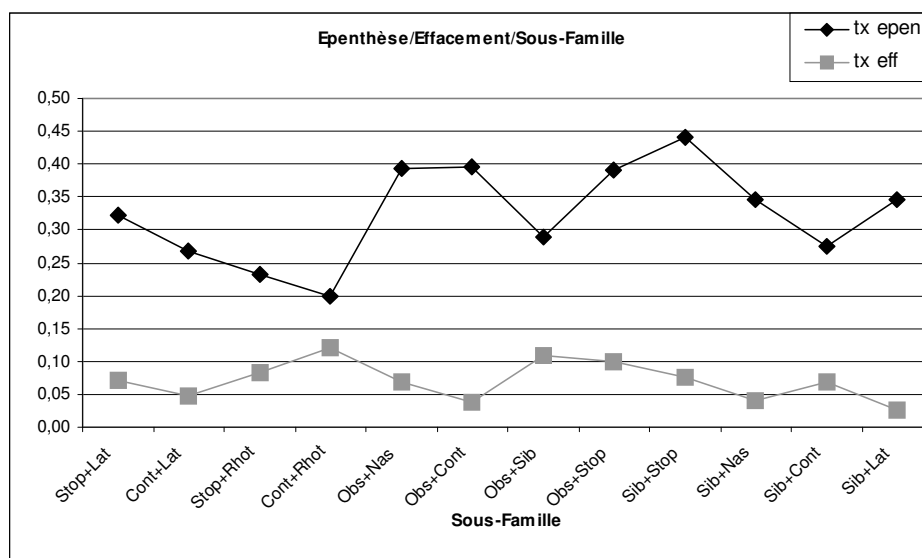


Commentaire :

- L'influence de la position sur le taux d'effacement est plus nette dans C1 que dans C2.
- Il y a convergence des résultats entre C1 et C2 : a) le classement de difficulté est identique ( $3 > 1 > 2$ ) (classement décroissant des taux d'effacements), b) les taux pour la position 2 sont relativement similaires entre C1 et C2.

- **Position et Epenthèse : pas de résultat convergent.**
- **Position et Effacement : résultats convergents : Coda finale > Attaque Initiale > Attaque Finale.**

## 2. Sous-Famille<sup>190</sup> :



### Commentaire :

Le classement de difficulté obtenu est le suivant :

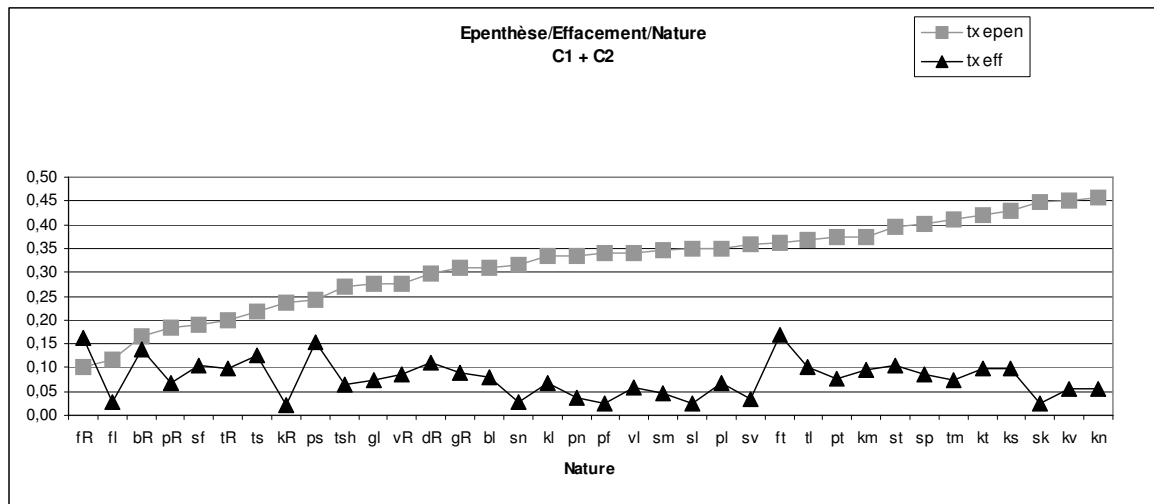
	Plus Difficile	Difficulté Intermédiaire	Moins Difficile
<b>Epenthèse</b>	Sib+Stop (44%)	Sib+Nas (35%)	Stop+Rhot (23%)
	Obs+Cont (40%)	Sib+Lat (35%)	Cont+Rhot (20%)
	Obs+Nas (39%)	Stop+Lat (34%) <sup>191</sup>	
	Obs+Stop (39%)	Obs+Sib (29%)	
		Sib+Cont (27%)	
	Cont+Lat (27%)		
<b>Effacement</b>	Cont+Rhot (12%)	Stop + Lat (8%) <sup>192</sup>	Cont+Lat (5%)
	Obs+Sib (11%)	Stop+Rhot (8%)	Obs+Cont (4%)
	Obs+Stop (10%)	Sib+Stop (8%)	Sib+Nas (4%)
		Obs+Nas (7%)	Sib+Lat (2%)
		Sib+Cont (7%)	

<sup>190</sup> Pour C1 et C2 les familles et les sous-familles étaient équivalentes (une sous-famille dans chaque famille), excepté pour les familles Obs+Lat et Obs+Rhot. Dans C1, seul le groupe /t/ (sous-famille : Stop+Lat ; taux d'épenthèses : 37% ; taux d'effacement : 10%) représentait la famille Obs+Lat.

<sup>191</sup> Correspondant à la moyenne des taux de Stop+Lat (C1) (37%) et de Stop+Lat (C2) (32%).

### 3. Nature

#### a) Par taux d'épenthèses croissant :

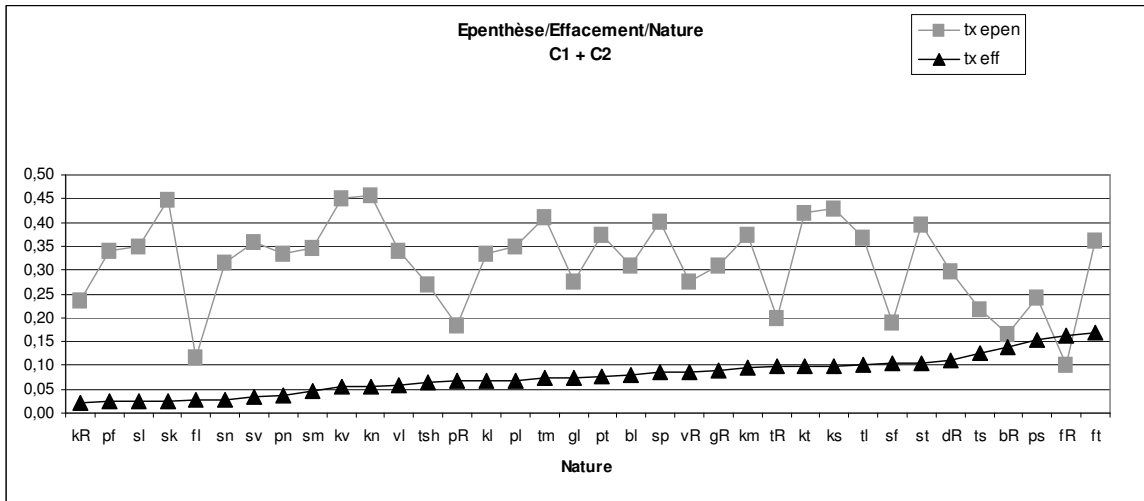


#### Commentaires :

Ce graphique permet de visualiser le classement des groupes consonantiques en fonction du taux d'épenthèse croissant, ainsi que les variations du taux d'effacement selon ce classement. On remarque que l'augmentation du taux d'épenthèse est relativement linéaire à travers l'ensemble de ces groupes consonantiques.

<sup>192</sup> Correspondant à la moyenne des taux de Stop+Lat (C1) (10%) et de Stop+Lat (C2) (7%).

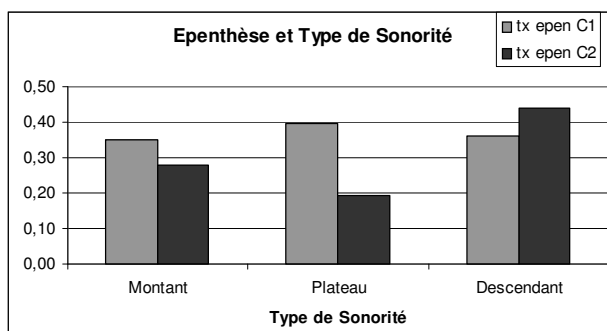
**b) Par taux d'effacements croissant :**



**Commentaires :**

Ce graphique permet de visualiser le classement des groupes consonantiques en fonction du taux d'effacement croissant, ainsi que les variations du taux d'épenthèse selon ce classement. On remarque que l'augmentation du taux d'effacement est relativement linéaire à travers l'ensemble de ces groupes consonantiques.

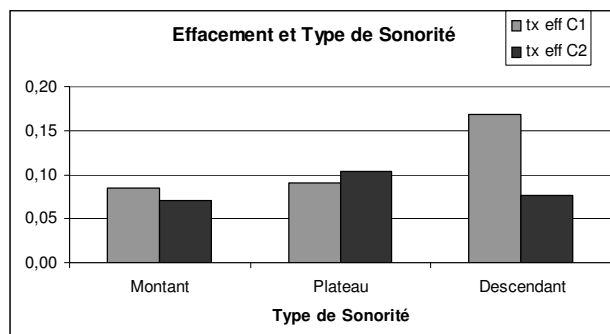
**4. Type de Sonorité**



**Commentaires :**



- L'influence du type de sonorité sur le taux d'épenthèse est beaucoup plus nette dans C2 que dans C1.
- Convergence partielle des résultats entre C1 et C2 : a) le type « Montant » se trouve en deuxième position du classement, b) les taux pour « Montant » d'une part, et « Descendant » d'autre part sont relativement similaires dans C1 et C2.

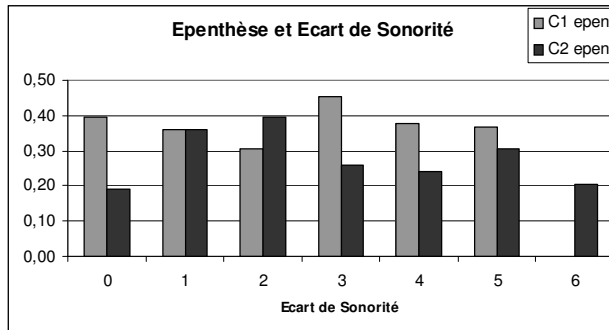


Commentaires :

- L'influence du type de sonorité sur le taux d'effacement est beaucoup plus nette dans C1 que dans C2.
- Convergence partielle des résultats entre C1 et C2 : a) le type « Montant » se trouve en dernière position du classement de difficulté, b) les taux pour « Montant » d'une part, et « Plateau » d'autre part sont relativement similaires dans C1 et C2.

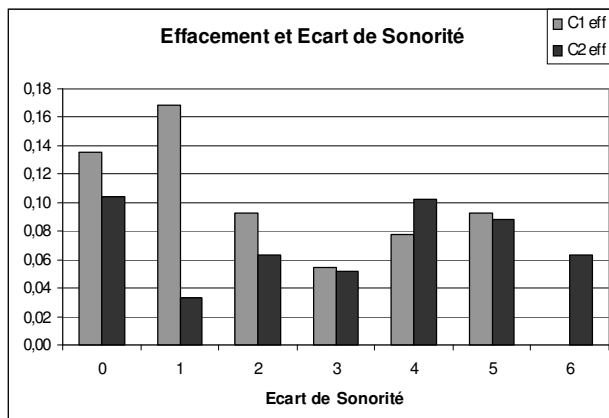
- **Type de Sonorité et Epenthèse : peu de résultats convergents (« Montant » au deuxième rang de difficulté).**
- **Type de Sonorité et Effacement : peu de résultats convergents (« Montant » au dernier rang de difficulté).**

## 5. Ecart de Sonorité<sup>193</sup> :



### Commentaires :

Il n'y a pas convergence des résultats : tandis que certains taux sont du même ordre (taux 1 et 5), d'autres sont tout à fait différents (taux 0 et 3 par exemple), conduisant à une différence de classement. L'écart de sonorité ne serait donc pas un indicateur stable du taux d'épenthèse.



### Commentaires :

Ici encore, il n'y a pas convergence des résultats : certains taux sont tout à fait différents (taux 1 et 4), conduisant à une différence de classement. Cependant, mis à part les écarts 1 et 4, les autres taux sont relativement similaires entre C1 et C2 (taux 0,

<sup>193</sup> Dans C1, l'écart 6 n'était pas représenté.

2, 3 et 5): les deux classements ne sont donc pas trop différents. L'écart de sonorité, sans être un indicateur stable, pourrait ainsi être un meilleur indicateur du taux d'effacement que du taux d'épenthèse.

- **Ecart de Sonorité et Epenthèse : pas de résultats convergents.**
- **Ecart de Sonorité et Effacement : résultats non identiques, mais davantage convergents que pour le taux d'épenthèse.**

### **Résumé Essentiel Comparaison C1/C2**

**La comparaison des résultats des deux tâches permet :**

- **D'établir un classement complet des groupes consonantiques en fonction des taux d'épenthèse et d'effacement.**
- **D'évaluer la valeur des indices « Position », « Type de sonorité » et « Ecart de sonorité » entre ces deux tâches.**

**Globalement, les résultats sont peu convergents, indiquant que l'épenthèse et l'effacement sont particulièrement liés à la nature phonologique (en terme de famille et de nature spécifique) des groupes en questions.**

## **Tâche D**

Rappelons que dans la tâche D, le nombre et la nature des distracteurs ont été modifiés par rapport à la tâche C2 : seuls les taux d'épenthèses ont été mesurés, la grande majorité des stimuli étant trisyllabique (par exemple /CCVCVCV/).

### **A) Résultats descriptifs**

#### **Résultats Globaux**

	<b>Taux d'épenthèses</b>
<b>Moyenne</b>	77,03%
<b>Ecart-type</b>	30,96%

#### **Remarques :**

- Le taux d'épenthèse moyen est très élevé (environ 77%).
- L'écart-type est élevé, mais relativement faible par rapport à la moyenne.

### **Résumé Essentiel Résultats Globaux**

**On observe 77% d'épenthèses pour les groupes /CC/ de la tâche D  
(unimodalité visuelle).**

#### **A) 1) Axe syntagmatique**

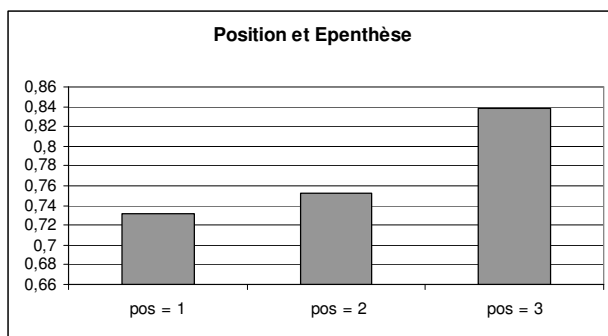
Les nombres de stimuli, selon la position, étaient les suivants :

<b>Position</b>	<b>1</b>	<b>2</b>	<b>3</b>
<b>Nombre de stimuli</b>	21	20	17

Lorsque l'on considère globalement les résultats, sans se soucier de la nature des stimuli (classe générale des groupes /CC/), voici ce que l'on obtient en fonction de la position :

**Epenhèses :**

Position	Taux d'épenhèses		
	1	2	3
Moyenne	73, 2%	75, 24%	83, 87%
Ecart-type	37, 06%	36, 04%	29, 92%



**Remarques :**

- Pas de résultats remarquables (position 1 : environ 73%, position 2 : environ 75%, position 3 : environ 84%, avec des écarts-types similaires), même si la position 3 est légèrement dominante, en particulier par rapport à la position 1, avec en outre l'écart-type le plus faible (environ 30% contre 37% et 36%). La dominance de la position 3 pourrait s'expliquer par la double épenhèse possible : CVCVCC → CVCVCvCv.

**Résumé Essentiel Axe Syntagmatique**

**Pas de différence majeure entre les trois positions pour les épenhèses, malgré une dominance de la position 3 (+10%).**

A) 2) Axe paradigmatic

α) Nature phonologique

Les groupes consonantiques représentés et leurs familles sont les mêmes que pour la tâche C2.

α). 1. Famille Générale

	Taux d'épenthèses		
Famille Générale	Obs + Liqu (exemple : /kl/) (code : 1)	Sib + Stop (exemple : /sk/) (code : 3)	Sib + Nonstop (exemple : /sl/) (code : 4)
<b>Moyenne</b>	76, 18%	79, 39%	78, 23%
<b>Ecart-type</b>	37, 8%	36, 02%	39, 33%

Remarques :

- Pas de résultats remarquables. Les taux sont très élevés et similaires (environ 76%, 79% et 78%).

**Pas de différence : taux élevés compris entre 76% et 79%.**

α). 2. Famille

Famille	Taux d'épenthèses	
	Moyenne	Ecart-type
Obs + Lat (code : 1) (exemple : /bl/)	77, 8%	37, 72%
Obs + Rhot (code : 2) (exemple : /bR/)	74, 93%	37, 86%
Sib +Stop (code : 7) (exemple : /sp/)	79, 39%	36, 02%

Sib + Nas (code : 8) (exemple : /sm/)	80, 65%	37, 97%
Sib + Cont (code : 9) (exemple : /sv/)	75%	40, 19%
Sib + Lat (code : 10) (exemple : /sl/)	79, 84%	40, 33%

Remarques :

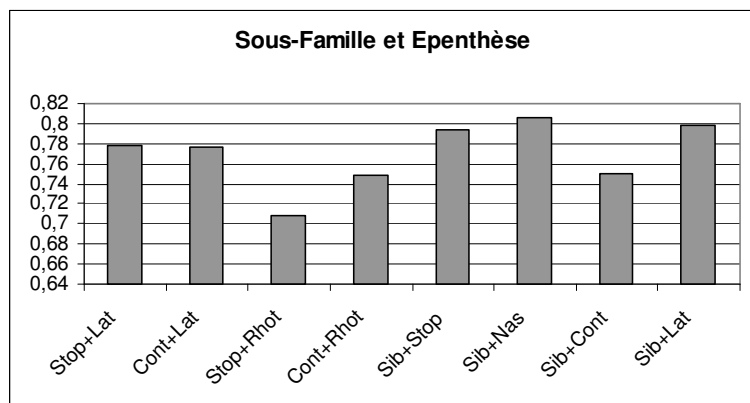
- Pas de résultats remarquables, puisque les moyennes sont comprises entre environ 75% et 80%.

**Pas de différence : taux élevés compris entre 75% et 80%.**

α). 3. Sous-famille

Sous-famille	Taux	
	Epenthèses	
	Moyenne	Ecart-type
Stop + Lat (code : 1) (exemple : /bl/)	77, 86%	36, 97%
Cont + Lat (code : 2) (exemple : /fl/)	77, 69%	39, 1%
Stop + Rhot (code : 3) (exemple : /bR/)	70, 79%	35, 11%
Cont + Rhot (code : 4) (exemple : /fR/)	74, 84%	40, 17%
Sib + Stop (code : 9) (exemple : /sp/)	79, 39%	36, 02%
Sib + Nas (code : 10) (exemple : /sm/)	80, 65%	37, 97%

Sib + Cont (code : 11) (exemple : /sv/)	75%	40, 19%
Sib + Lat (code : 12) (exemple : /sl/)	79, 84%	40, 33%



Remarque :

- Ici encore, les moyennes sont très élevées et assez proches, comprises entre environ 71% (Stop+Rhot) et 81% (Sib+Nas). On peut donc simplement noter le statut particulier de Stop+Rhot, situé en fin de classement.

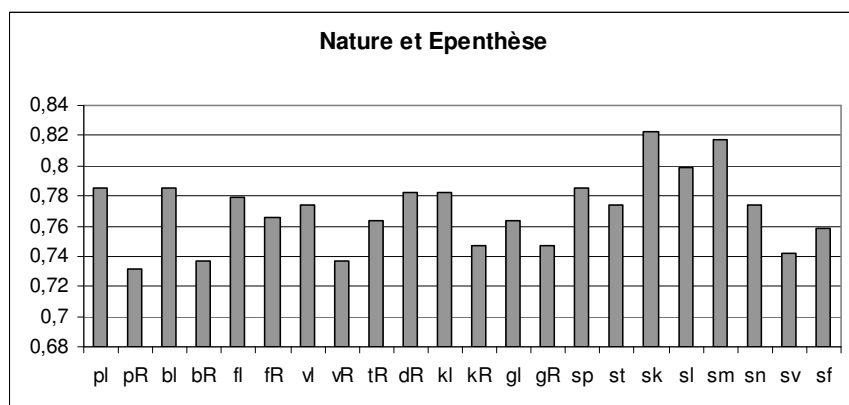
**Peu de différences : la plupart des taux sont compris entre 75% et 80%.**

**Seule se distingue la sous-famille Stop+Rhot avec le taux le plus bas (71%).**



*α). 4. Nature spécifique*

Groupe Consonantique	Taux d'épenthèses	
	Moyenne	Ecart-type
/pI/ (code : 14)	78, 49%	41, 12%
/bI/ (code : 16)	78, 49%	40, 88%
/kI/ (code : 24)	78, 23%	40, 64%
/gI/ (code : 26)	76, 34%	42, 66%
/fI/ (code 18)	77, 96%	41, 75%
/vI/ (code : 20)	77, 42%	42, 03%
/pR/ (code : 15)	73, 12%	44, 04%
/bR/ (code : 17)	73, 66%	42, 88%
/tR/ (code : 22)	76, 34%	42, 34%
/dR/ (code : 23)	78, 23%	40, 14%
/kR/ (code : 25)	74, 73%	42, 62%
/gR/ (code : 27)	74, 73%	42, 47%
/fR/ (code : 19)	76, 61%	41, 78%
/vR/ (code : 21)	73, 66%	43, 73%
/sp/ (code : 28)	78, 49%	36, 02%
/st/ (code : 29)	77, 42%	42, 12%
/sk/ (code : 30)	82, 26%	38, 15%
/sm/ (code : 32)	81, 72%	38, 59%
/sn/ (code : 33)	77, 42%	42, 15%
/sv/ (code : 34)	74, 19%	43, 76%
/sf/ (code : 35)	75, 81%	43, 13%
/sl/ (code : 31)	79, 84%	40, 33%



**Remarques :**

- Pas de résultats remarquables : les moyennes sont très élevées, variant entre environ 73% (/pR/) et 82% (/sk/), avec des écarts-types élevés.
- Pas de rapport de dominance particulier entre groupes voisé et non-voisé : les taux sont même identiques pour certains. Si l'on néglige un écart de 2%, voici ce que l'on obtient :

Groupe Non-voisé	Rapport (à 2% près)	Groupe Voisé
/pl/ (code : 14) 78%	=	/bl/ (code : 16) 78%
/kl/ (code : 24) 78%	=	/gl/ (code : 26) 76%
/fl/ (code 18) 78%	=	/vl/ (code : 20) 77%
/pR/ (code : 15) 73%	=	/bR/ (code : 17) 74%
/tR/ (code : 22) 76%	=	/dR/ (code : 23) 78%
/kR/ (code : 25) 75%	=	/gR/ (code : 27) 75%
/fR/ (code : 19) 77%	>	/vR/ (code : 21) 74%
/sf/ (code : 35) 76%	=	/sv/ (code : 34) 74%

**Peu de différences significatives : la plupart des taux sont compris entre 75% et 80%.**

- **Taux inférieurs** : /pR/ (73%), /bR/, /vR/ et /sv/ (74%).
- **Taux supérieurs** : /sk/ et /sm/ (82%).

**Pas d'influence significative du voisement. Dans le cas des groupes contenant /f/ ou /v/, /f/ déclenche davantage d'épenthèses.**

β) Sonorité

β).1. Type de Sonorité

Type de sonorité	Taux d'épenthèses	
	Moyenne	Ecart-type
Montant	77%	31%
Plateau	76%	38%
Descendant	79%	31%

Remarque :

- Peu de différence, avec une légère dominance pour le type « Descendant ».

**Pas de différence entre les trois types : taux compris entre 76% et 79%.**

β).2. Ecart de Sonorité

Ecart de sonorité	Taux	
	Epenthèses	
	Moyenne	Ecart-type
0	76%	43%
1	74%	44%
2	79%	35%
3	77%	38%
4	77%	37%
5	76%	36%
6	78%	37%

Remarques :

- Peu de différence, avec une légère dominance pour l'écart 2 (79%), dont l'écart-type est le plus faible.

**Peu de différences : taux compris entre 74% et 79%.**

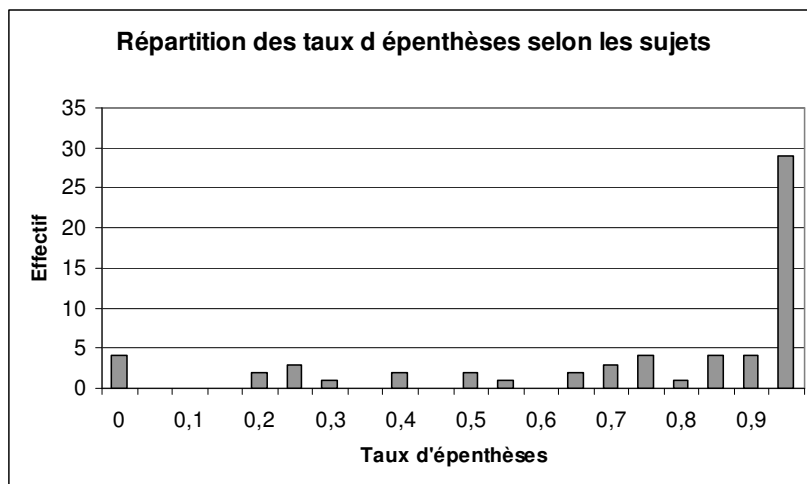
### **Résumé Essentiel Axe Paradigmatique**

**La nature phonologique des groupes ne semble que peu influencer le taux d'épenthèses en unimodalité visuelle.**

#### *A) 3) Variation entre sujets*

Au vu des résultats « massifs » obtenu pour cette tâche D, avec des taux d'épenthèses globalement très élevés sous toutes les conditions, il nous a paru utile de chercher à connaître la répartition des ces taux en fonction des sujets. Notre hypothèse était celle du « Tout-Ou-Rien » : soit les sujets avaient un taux d'épenthèses quasi-nul, soit un taux avoisinant les 100%.

Les résultats, sous forme graphique, se présentent ainsi :



Notre hypothèse est donc confirmée :

- Seuls 4 sujets (6%) ont un taux d'épenthèses nul.
- Seuls 12 sujets (19%) ont un taux inférieur à 50%.

- 29 sujets (47%) ont un taux compris entre 95% et 100%.

#### **Résumé Variation Inter-Sujets**

- **81% des sujets ont un taux d'épenthèses supérieur à 50%.**
- **6% des sujets ont un taux nul.**
- **47% des sujets ont un taux compris entre 95% et 100%.**

#### **B) Résumé Essentiel**

#### **Résumé Essentiel**

**On observe 77% d'épenthèses pour les groupes /CC/ de la tâche D (unimodalité visuelle).**

**La position ne semble pas influencer le taux d'épenthèse, même si celui-ci est maximal en coda de la syllabe finale.**

**La nature phonologique des groupes ne semble pas influencer le taux d'épenthèses, même si la sous-famille Stop+Rhot semble déclencher moins d'épenthèses.**

**L'analyse inter-sujets indique que 47% des sujets ont un taux d'épenthèses compris entre 95% et 100%.**

## **Comparaison Tâches C2 et D**

Lorsque l'on compare les résultats des tâches C2 et D (qui ne différaient que par la composition, le nombre et la modalité de présentation des stimuli), voici ce que l'on obtient, pour les cibles suivantes :

- /pl, pr, bl, br, fl, fr, vl, vr, tr, dr, kl, kr, gl, gr/ (OBLI)
- /sp, st, sk/ (S + Plosive non-voisée)
- /sl, sm, sn, sv, sf/ (S + C non-plosive)

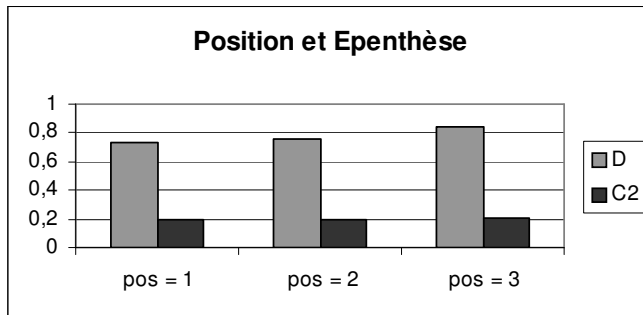
Etant donné les résultats obtenus pour D, nous ne présentons ci-dessous que les résultats les plus pertinents.

### **1. Résultats globaux**

<b>Taux d'épenthèses</b>	<b>Modalité Auditive</b>	<b>Modalité Visuelle</b>
<b>Moyenne</b>	58, 5%	77, 03%
<b>Ecart-type</b>	25, 63%	30, 96%

**La modalité visuelle (orthographique) déclenche plus d'épenthèses (77%) que la modalité auditive (58%).**

## 2. Position



### Commentaire :

- L'influence de la position sur le taux d'épenthèse est plus sensible dans D que dans C2.

**La position ne semble pas influencer grandement le taux d'épenthèses pour les groupes consonantiques en question. Cependant :**

- **En position d'attaque initiale (position 1) : le taux est le plus faible, dans les deux modalités.**
- **En position de coda finale (position 3), le taux est supérieur de 10% en modalité visuelle, alors qu'il est stable en modalité auditive.**

## 3. Famille Générale

Famille Générale	Obs+Liqu	Sib+Stop	Sib+Nonstop
Modalité Auditive	26%	44%	32%
Modalité Visuelle	76%	79%	78%

**Les écarts sont beaucoup moins marqués pour la modalité visuelle, mais le classement est identique dans les deux modalités : Sib+Stop > Sib+Nonstop > Obs+Liqu.**

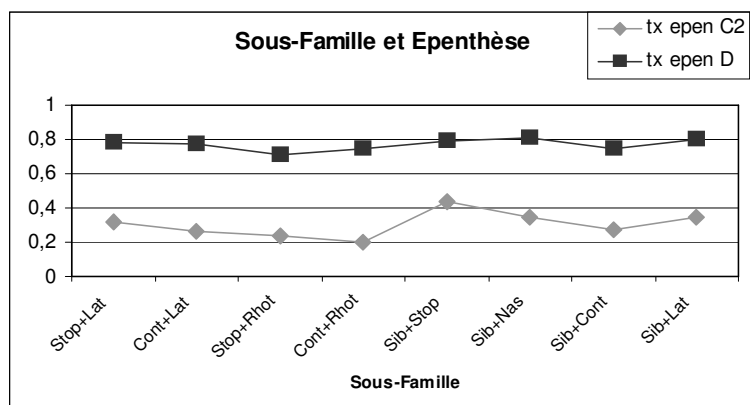
#### 4. Famille

Famille	Modalité Auditive	Modalité Visuelle
Obs+Lat	31%	78%
Obs+Rhot	22%	75%
Sib+Stop	44%	79%
Sib+Nas	35%	81%
Sib+Cont	27%	75%
Sib+Lat	35%	80%

Les écarts sont beaucoup moins marqués pour la modalité visuelle.

- Convergence de classement : Obs+Lat > Obs+Rhot.
- Divergences de classement : Sib+Stop > Sib+Nas et Sib+Lat en Auditif, mais Sib+Stop < Sib+Nas et Sib+Lat en Visuel.

#### 5. Sous-Famille :



Commentaire :

- Pour chaque sous-famille, les taux en modalité visuelle sont beaucoup plus élevés que ceux en modalité auditive.



- On remarque néanmoins un profil relativement similaire des deux courbes, avec des taux plus faibles pour Stop+Rhot et des taux plus élevés pour Sib+Stop.

**Taux plus élevés en modalité visuelle pour toutes les sous-familles.**

**Convergence relative inter-modale :**

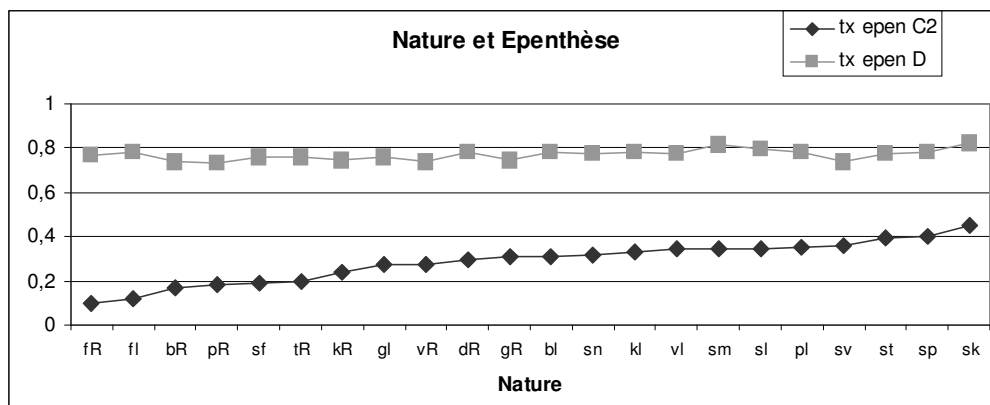
- Taux assez faibles pour Stop+Rhot.
- Taux assez élevés pour Sib+Stop.

**Divergence relative inter-modale:**

- Cont+Lat : plus élevé en visuel, moins élevé en auditif.
- Cont+Rhot : plus élevé en visuel, moins élevé en auditif.
- Sib+Nas : plus élevé en visuel, moins élevé en auditif.

**6. Nature :**

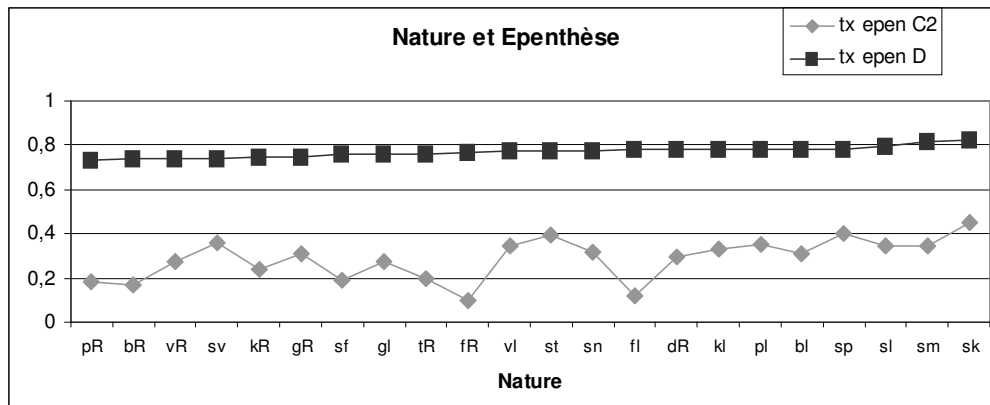
a) **Taux croissant selon C2**



**Commentaires :**

- Tandis que le taux d'épenthèses est croissant et relativement linéaire pour C2, il est relativement stable pour D.

**b) Taux croissant selon D**



Commentaires :

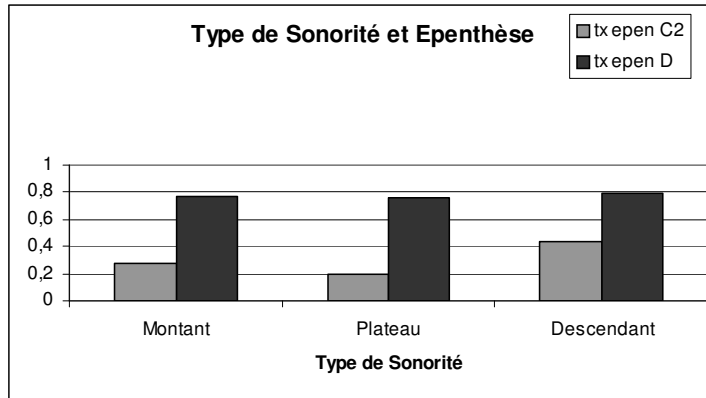
- Tandis que le taux d'épenthèses est légèrement croissant et quasi-linéaire pour D, il est se présente sous forme non-linéaire pour C2.
- Le taux d'épenthèse semble donc beaucoup plus sensible à la nature phonologique des groupes en modalité auditive qu'en modalité visuelle.
- Les classements de difficulté semblent donc être différents selon les modalités. Cependant, la comparaison des deux graphiques permet de distinguer les similarités déjà observées pour les sous-familles : globalement, les groupes C+/R/ semblent déclencher moins d'épenthèses que C+/l/ et que /s/+C.

**Le taux d'épenthèse est plus sensible aux propriétés phonologiques des groupes en modalité auditive qu'en modalité visuelle.**

**Les classements de difficulté dans les deux modalités semblent cependant convergents :**

**/s/+C > C+/l/ > C+/R/**

## 7. Type de Sonorité



### Commentaire :

- Les taux sont plus élevés en modalité visuelle qu'en modalité auditive.
- Les taux sont plus sensibles au type de sonorité en modalité auditive que visuelle (intra-modal).
- Convergence des classements : Descendant > Montant > Plateau (inter-modal).

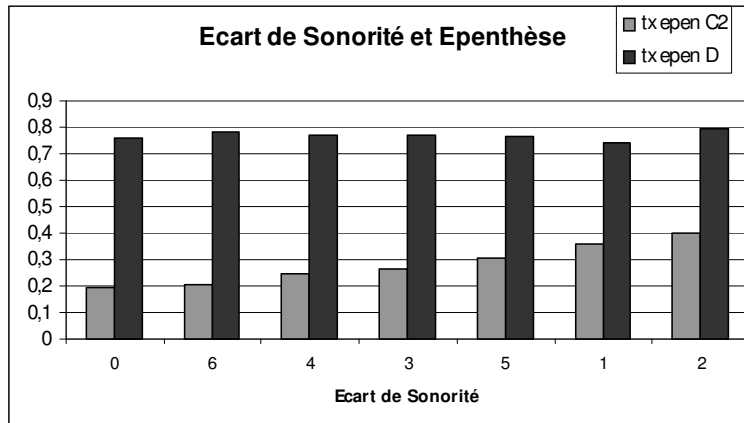
**Le taux d'épenthèse est plus sensible au type de sonorité des groupes en modalité auditive qu'en modalité visuelle.**

**Les classements de difficulté dans les deux modalités semblent cependant convergents :**

**Descendant > Montant > Plateau**

## **8. Ecart de Sonorité**

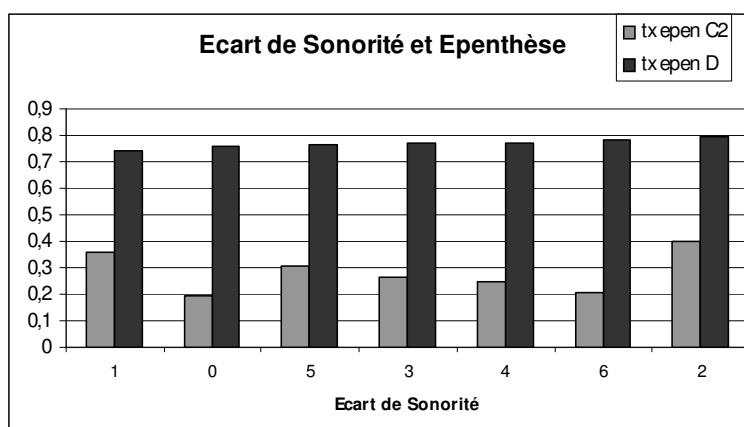
a) En fonction de C2 croissant :



Commentaire :

- Les taux sont plus élevés en modalité visuelle qu'en modalité auditive.
- Taux plus stable pour D que pour C2 (intra-modal).
- Les classements de difficultés sont différents (inter-modal).

b) En fonction de D croissant :



Commentaires :

- Les taux sont plus élevés en modalité visuelle qu'en modalité auditive.
- Taux plus stable pour D que pour C2 (intra-modal).
- Les classements de difficultés sont différents (inter-modal).

**Le taux d'épenthèse est plus sensible à l'écart de sonorité des groupes en modalité auditive qu'en modalité visuelle.**

**Les classements de difficulté dans les deux modalités ne sont pas convergents.**

### **Résumé Essentiel Comparaison C2/D**

**La comparaison des résultats des deux tâches permet d'évaluer :**

- L'impact de la modalité de présentation des stimuli sur les taux d'épenthèses.
- Les convergences et divergences des effets des variables indépendantes entre les deux modalités.

#### **Taux :**

**Dans tous les cas, la modalité visuelle (77%) déclenche beaucoup plus d'épenthèses que la modalité auditive (58%).**

#### **Position :**

**Peu d'influence, quelle que soit la modalité. Cependant :**

- Position 1 (Attaque initiale) : taux le plus faible dans les deux modalités.
- Position 3 (Coda finale) : taux plus élevé en modalité visuelle, mais stable en modalité auditive.

#### **Propriétés phonologiques :**

**Nettement plus d'influence en modalité auditive que visuelle. Il n'y pas toujours convergence inter-modale.**

**Convergences relatives :**

- **Famille Générale : Sib+Stop > Sib+Nonstop > Obs+Liqu**
- **Famille : Obs+Lat > Obs+Rhot**
- **Nature : /s/+C > C+/l/ > C+/R/**
- **Type de sonorité : Descendant > Montant > Plateau**

**Divergences particulières :**

- **Cont+Lat, Cont+Rhot et Sib+Nas : plus d'épenthèses en visuel, moins en auditif (comparaison intra-modale).**
- **Ecart de sonorité : classements de difficulté différents.**

## **Tâche E**

Deux cas sont à distinguer dans cette tâche, selon que les deux modalités sont convergentes ou divergentes.

### **Cas 1 - Stimuli Auditifs et Visuels Convergen**

Dans ce cas, les stimuli auditifs et visuels sont de même longueur syllabique, c'est-à-dire trisyllabiques (comme dans la tâche D). Par exemple : /vlamote/ et <VLAMOTE>.

#### **A) Résultats descriptifs**

##### **Résultats Globaux**

	<b>TAUX</b>	
	<b>Confusion syllabique (épenthèses)</b>	<b>Discrimination erronée</b>
<b>Moyenne</b>	Entre 64, 07% et 69%	7, 7%
<b>Ecart-type</b>	27, 8%	10, 8%

##### **Remarques :**

- Traitement des groupes /CC/ : environ 66, 5% d'épenthèses.
- Le taux de 64% correspond à *l'ensemble des réponses erronées*, c'est-à-dire :
  - 1) Perception d'une différence intermodale illusoire (discrimination erronée, qui présuppose une confusion syllabique intramodale, soit visuelle soit auditive),
  - 2) Confusion syllabique intermodale (similarité des stimuli correctement perçue).

Cependant, si l'on écarte les sujets ayant perçu une différence illusoire (discrimination erronée), alors le taux de confusion syllabique est d'environ 69%. Pour la suite des résultats, nous nous situerons dans ce cas de figure : nous laisserons de côté les cas où il y a eu perception d'une différence illusoire et ne

considèrerons que les cas où la perception du rapport entre les deux stimuli est correcte (pas de différence intermodale) mais où il y a confusion syllabique (erreur de comptage).

- Dans environ 8% des cas, les sujets ont jugé différents des stimuli qui, en fait, étaient similaires.

### **Résumé Essentiel Résultats Globaux Stimuli Convergentes**

**On observe entre 64% et 69% d'épenthèses pour les groupes /CC/ de la tâche E (bimodalité synchrone convergente auditive/orthographique).**

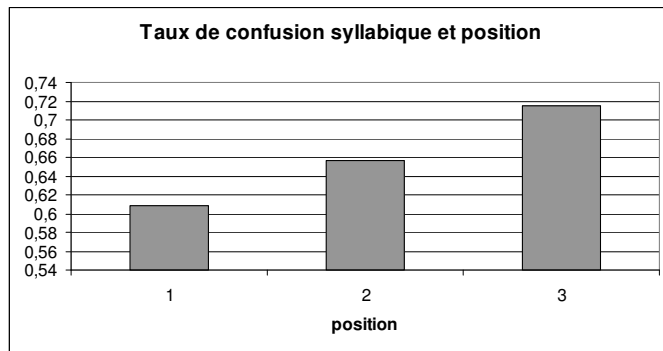
**On observe également 8% de perception illusoire d'une différence au sein des couples de stimuli, pourtant similaires.**

#### *A) 1) Axe syntagmatique*

Lorsque l'on considère globalement les résultats, sans se soucier de la nature des stimuli (classe générale des groupes /CC/), voici ce que l'on obtient en fonction de la position :

<b>Taux de Confusion Syllabique (épenthèses)</b>			
<b>Position</b>	<b>1</b>	<b>2</b>	<b>3</b>
<b>Moyenne</b>	60, 8%	65, 7%	71, 5%
<b>Ecart-type</b>	27, 2%	33, 8%	29, 7%





Remarques :

- Les taux sont compris autour de la moyenne, entre environ 61% et 71%.
- Il semblerait que la position influe sur le taux de confusion syllabique selon le classement décroissant suivant : Position 3 > Position 2 > Position 1.

**Résumé Essentiel Axe Syntagmatique Stimuli Convergen**

**La position semble influencer le taux d'épenthèse, avec un écart de 5% entre chaque position :**

- **Attaque initiale : 61%**
- **Attaque finale : 66%**
- **Coda finale : 71%**

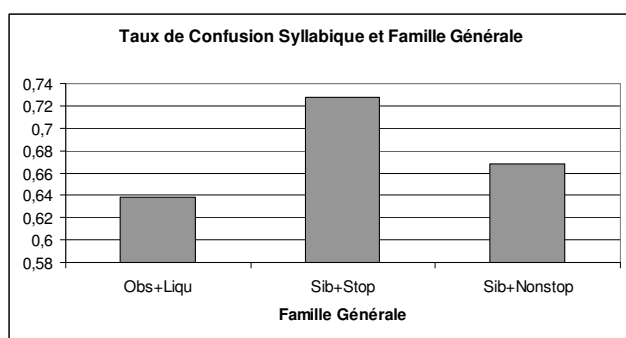
A) 2) Axe paradigm

α) Nature phonologique

Les groupes consonantiques représentés et leurs familles sont les mêmes que pour les tâches C2 et D.

*α). 1. Famille Générale*

	<b>Taux de Confusion Syllabique (épenthèses)</b>		
<b>Famille Générale</b>	<b>Obs + Liqui</b>  (exemple : /kl/) (code : 1)	<b>Sib + Stop</b>  (exemple : /sk/) (code : 3)	<b>Sib + Nonstop</b>  (exemple : /sl/) (code : 4)
<b>Moyenne</b>	63, 8%	72, 8%	66, 9%
<b>Ecart-type</b>	30%	33, 1%	21, 8%



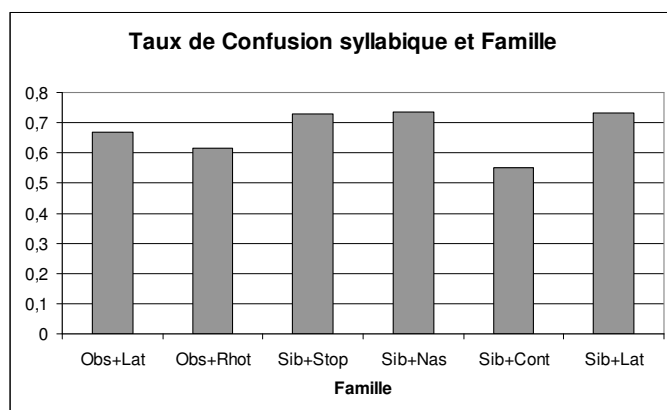
Remarques :

- Les taux sont élevés.
- Le classement obtenu est le suivant : Sib+Stop > Sib+Nonstop > Obs+Liqui.

**Variation des taux entre 64% et 73% : Sib+Stop (73%) > Sib+Nonstop (67%) > Obs+Liqui (64%).**

α). 2. Famille

Famille	Taux de Confusion Syllabique (épenthèses)	
	Moyenne	Ecart-type
Obs + Lat (code : 1) (exemple : /bl/)	66, 9%	28, 9%
Obs + Rhot (code : 2) (exemple : /bR/)	61, 5%	32, 2%
Sib +Stop (code : 7) (exemple : /sp/)	72, 8%	33, 1%
Sib + Nas (code : 8) (exemple : /sm/)	73, 5%	22, 3%
Sib + Cont (code : 9) (exemple : /sv/)	55, 2%	25, 3%
Sib + Lat (code : 10) (exemple : /sl/)	73, 4%	39, 2%



Remarques :

- Le classement obtenu est le suivant : Sib+Nas, Sib+Lat, Sib+Stop (73%) > Obs+Lat (67%) > Obs+Rhot (61%) > Sib+Cont (55%).
- Statut particulier de Sib+Cont (55%).

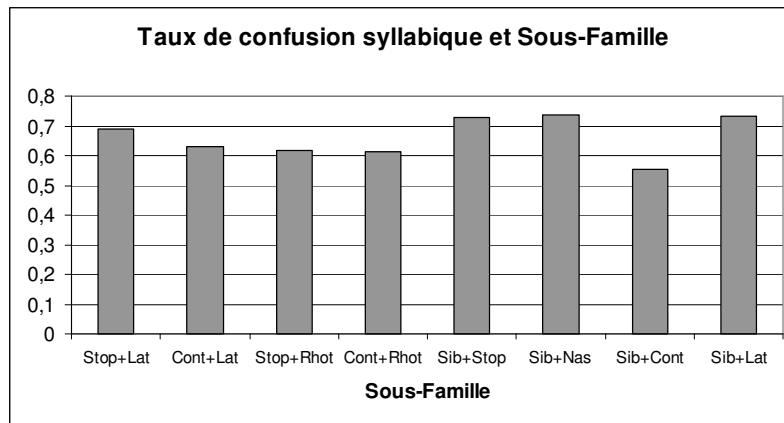
**On distingue quatre groupes (écart fixe de 6% entre eux) :**

- **Sib+Nas, Sib+Lat, Sib+Stop : 73%**
- **Obs+Lat : 67%**
- **Obs+Rhot : 61%**
- **Sib+Cont : 55%**

**On remarque que Obs+Lat > Obs+Rhot.**

*α). 3.Sous-famille*

Sous-famille	Taux de Confusion Syllabique (épentèses)	
	Moyenne	Ecart-type
Stop + Lat (code : 1) (exemple : /bl/)	68, 9%	32, 5%
Cont + Lat (code : 2) (exemple : /fl/)	62, 9%	25, 5%
Stop + Rhot (code : 3) (exemple : /bR/)	61, 6%	31, 7%
Cont + Rhot (code : 4) (exemple : /fR/)	61, 3%	36, 3%
Sib + Stop (code : 9) (exemple : /sp/)	72, 8%	33, 1%
Sib + Nas (code : 10) (exemple : /sm/)	73, 5%	22, 3%
Sib + Cont (code : 11) (exemple : /sv/)	55, 2%	25, 3%
Sib + Lat (code : 12) (exemple : /sl/)	73, 4%	39, 2%



Remarques :

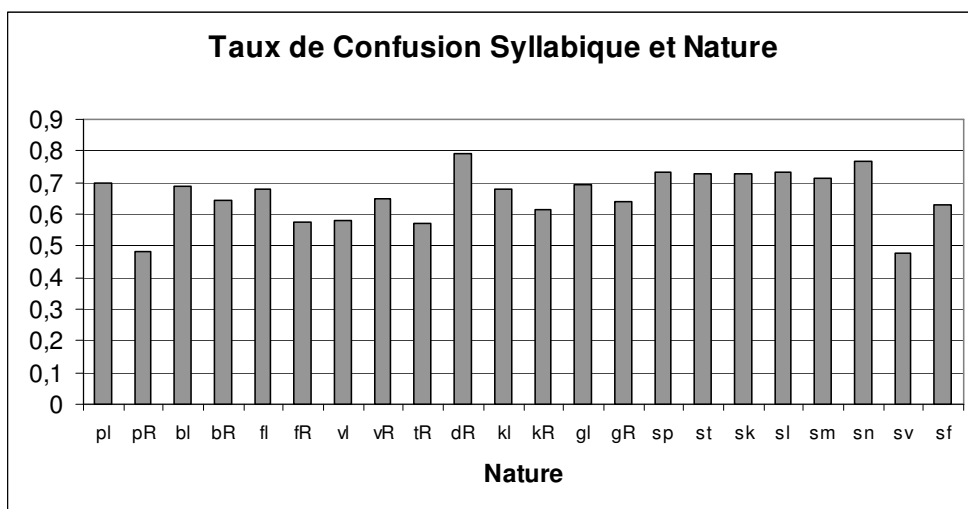
- Le classement obtenu est le suivant :  
Sib+Nas, Sib+Lat, Sib+Stop (73%) > Stop+Lat (69%) > Cont+Lat (63%) > Stop+Rhot (62%) > Cont+Rhot (61%) > Sib+Cont (55%).
- On remarque que : 1) Stop+Lat (69%) > Cont+Lat (63%), 2) Cont+Lat (63%)  $\cong$  Stop+Rhot (62%)  $\cong$  Cont+Rhot (61%).

**On remarque que :**

- **Stop+Lat (69%) > Cont+Lat (63%).**
- **Stop+Rhot, Cont+Rhot et Cont+Lat ont des taux équivalents (62% à 1% près).**

*α). 4. Nature spécifique*

Groupe Consonantique	Taux de Confusion Syllabique (épenthèses)	
	Moyenne	Ecart-type
/pl/ (code : 14)	69, 9%	34%
/bl/ (code : 16)	68, 8%	37, 6%
/kl/ (code : 24)	67, 7%	37, 2%
/gl/ (code : 26)	69, 3%	34, 3%
/fl/ (code 18)	67, 7%	36, 7%
/vl/ (code : 20)	58%	23, 3%
/pR/ (code : 15)	48, 4%	36%
/bR/ (code : 17)	64, 5%	39, 5%
/tR/ (code : 22)	57%	37%
/dR/ (code : 23)	79%	34, 7%
/kR/ (code : 25)	61, 3%	37, 3%
/gR/ (code : 27)	64%	37, 3%
/fR/ (code : 19)	57, 5%	39, 1%
/vR/ (code : 21)	65%	38, 8%
/sp/ (code : 28)	73, 1%	36, 1%
/st/ (code : 29)	72, 6%	34, 9%
/sk/ (code : 30)	72, 6%	35, 4%
/sm/ (code : 32)	71, 5%	20, 1%
/sn/ (code : 33)	76, 6%	37%
/sv/ (code : 34)	47, 6%	26, 3%
/sf/ (code : 35)	62, 9%	42, 4%
/sl/ (code : 31)	73, 4%	39, 2%



Remarques :

- Classement :

/dR/ (environ 79%) > /sn/ (environ 77%) > /sl/, /sp/, /sk/, /st/ (environ 73%) > /sm/ (environ 71%) > /pl/ (environ 70%) > /gl/, /bl/ (environ 69%) > /kl/, /fl/ (environ 68%) > /vR/ (environ 65%) > /bR/, /gR/ (environ 64%) > /sf/ (environ 63%) > /kR/ (environ 61%) > /vl/ (environ 58%) > /tR/, /fR/ (environ 57%) > /pR/, /sv/ (environ 48%).

- Groupes :

A)	/dR, sn/	78%
B)	/sl, sp, sk, st, sm/	72%
C)	/pl, gl, bl, kl, fl/	69%
D)	/vR, bR, gR, sf, kR/	63%
E)	/vl, tR, fR/	57, 5%
F)	/pR, sv/	48%

On peut globalement considérer que, mis à part les extrêmes et cas particuliers, on a :  
 S+C > C+L > C+R.

- Extrêmes :

<u>Tête</u>	/dR/(79%) et /sn/ (77%)
<u>Queue</u>	/pR/ et /sv/ (48%)

- Différences particulières :

- 1) /sn/ (77%) et /sm/ (71%),
- 2) /dR/ (79%) et /tR/ (57%)

- Ressemblances particulières : /st/ et /sk/ (72, 6%) et /sp/ (73, 1%) ont quasiment le même taux.

- Classement des fricatives /f, v/ :

/fl/ (68%) > /vR/ (65%) > /sf/ (63%) > /vl/ (58%) > /fR/ (57%) > /sv/ (48%).

Donc :

- 1) /fl/ > /vl/, mais /vR/ > /fR/.
- 2) /fl/ > /fR/, mais /vR/ > /vl/.
- 3) /sf/ > /sv/.

- Voisement :

Groupe Non-voisé	Rapport (à 1% près)	Groupe Voisé
/pl/ (code : 14) 70%	=	/bl/ (code : 16) 69%
/kl/ (code : 24) 68%	=	/gl/ (code : 26) 69%
/fl/ (code 18) 68%	>	/vl/ (code : 20) 58%
/pR/ (code : 15) 48%	<	/bR/ (code : 17) 64%
/tR/ (code : 22) 57%	<	/dR/ (code : 23) 79%
/kR/ (code : 25) 61%	<	/gR/ (code : 27) 64%
/fR/ (code : 19) 57%	<	/vR/ (code : 21) 65%
/sf/ (code : 35) 63%	>	/sv/ (code : 34) 48%



- Pour les groupes C+L : pas de différence (excepté quand C=f/v).  
Pour les groupes C+R : Taux Cnonvoisée+R < Taux Cvoisée+R.  
Pour les groupes avec f/v : pas de régularité remarquable.

**Les taux sont compris entre 78% (/dR, sn/) et 48% (/pR/, /sv/).**

**On peut distinguer trois grands groupes : S+C (73%) > C+L (69%) > C+R (62%).**

**Tandis que /sp, st, sk/ ont le même taux (73%), /sn/ (77%) et /sm/ (71%) ont des taux différents. De la même manière /dR/ est à 79%, tandis que /tR/ est à 57%.**

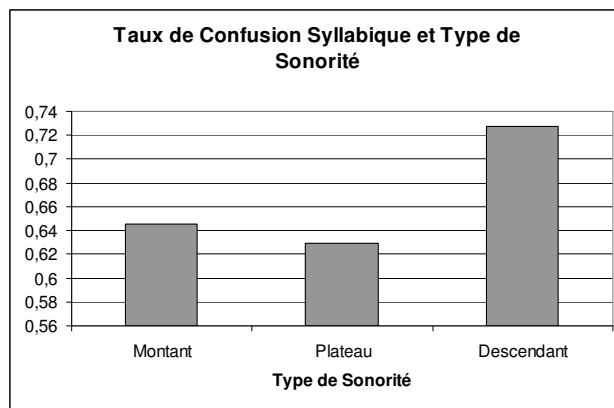
**Le voisement de la consonne semble influencer les taux des groupes C+R (Voisé > Nonvoisé), mais pas ceux des groupes C+L.**

**En ce qui concerne les fricatives /f, v/, il n'y a pas de régularité remarquable.**

β) Sonorité

β).1. Type de Sonorité

Type de sonorité	Taux de Confusion Syllabique (épenthèses)	
	Moyenne	Ecart-type
Montant	64,5%	27,5%
Plateau	62,9%	28,2%
Descendant	72,7%	33%



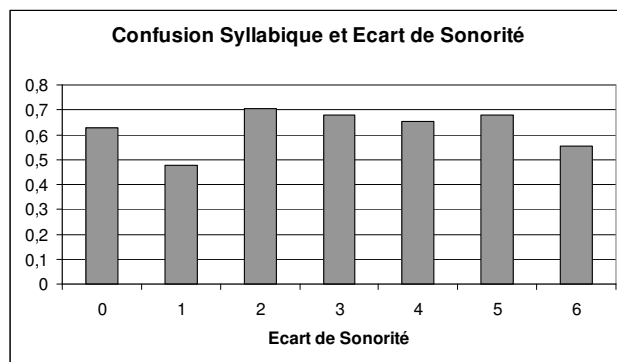
Remarques :

- Les écarts sont très marqués entre les trois types.
- Le classement obtenu est le suivant : Descendant (environ 73%) > Montant (environ 64%) > Plateau (environ 62%).

**Nette différence entre le type Descendant et les deux autres : Descendant (73%) > Montant (64%) > Plateau (62%).**

β).2. Ecart de Sonorité

Ecart de sonorité	Taux	
	Epenthèses	
	Moyenne	Ecart-type
<b>0</b>	62, 9%	42, 4%
<b>1</b>	47, 6%	26, 3%
<b>2</b>	70, 4%	24, 6%
<b>3</b>	68, 1%	34, 1%
<b>4</b>	65, 2%	32, 8%
<b>5</b>	68, 1%	32, 3%
<b>6</b>	55, 6%	32, 2%



Remarques :

- Certaines différences sont nettes.
- Le classement obtenu est le suivant : Ecart 2 (environ 70%) > Ecarts 3 et 5 (environ 68%) > Ecart 4 (environ 65%) > Ecart 0 (environ 63%) > Ecart 6 (environ 56%) > Ecart 1 (environ 48%).

**Taux compris entre 69% (Ecart 2, 3 et 5) et 48% (Ecart 1). Pas de régularité remarquable, même si les écarts 6 (56%) et 1 (48%) se trouvent en fin de classement.**

### **Résumé Essentiel Axe Paradigmatique Stimuli Convercents**

**La nature phonologique des groupes semble influencer le taux d'épenthèse en bimodalité auditive-orthographique.**

#### **B) Résultats inférentiels**

Les analyses de variances univariées (Test F) suggèrent que toutes les variables indépendantes à l'étude (position, famille générale, famille, sous-famille, nature, type de sonorité, écart de sonorité) influencent les taux d'épenthèses ( $F > F_{critique}$  et  $p < 0,01$ ).

#### **C) Résumé Essentiel**

### **Résumé Essentiel Stimuli Convercents**

**On observe :**

- **Entre 64% et 69% d'épenthèses pour les groupes /CC/ de la tâche E (bimodalité synchrone auditive - orthographique).**
- **8% de perception illusoire de différence intermodale.**

**La position semble influencer le taux d'épenthèse :**

- **Attaque initiale : 61%**
- **Attaque finale : 66%**
- **Coda finale : 71%**

**La nature phonologique semble influencer le taux d'épenthèse :**

- **Famille générale : Sib+Stop (73%) > Sib+Nonstop (67%) > Obs+Liqu (64%)**
- **Nature : /s/+C (73%) > C+/l/ (69%) > C+/R/ (62%)**
- **Type de sonorité : Descendant (73%) > Montant, Plateau (63%)**

**On remarque que :**

- **Famille** : Obs+Lat (67%) > Obs+Rhot (61%)
- **Sous-Famille** : Stop+Lat (69%) > Cont+Lat (63%) mais Stop+Rhot (62%) = Cont+Rhot (61%)
- **Nature** : Taux compris entre 78% (/dR/, /sn/) et 48% (/pR/, /sv/)
- **Voisement** : influence possible sur les groupes C+/R/, mais pas sur C+/l/

**Les analyses inférentielles suggèrent que chaque variable indépendante influence significativement le taux d'épenthèse.**

## **Cas 2 - Stimuli Auditifs et Visuels Divergents**

Dans ce cas, les stimuli auditifs et visuels ont des longueurs syllabiques différentes. Par exemple : /vlamote/ et <VULAMOTE>, ainsi que /vulamote/ et <VLAMOTE>. Le stimulus le plus long est donc légal en L1, contrairement au stimulus le plus court. Deux pistes d'investigation nous préoccupent alors ici :

- A) La modalité dominante en comptage syllabique,
- B) La capacité à percevoir la différence intermodale.

### **A) Modalité dominante en tâche de comptage syllabique**

Deux cas sont à distinguer ici :

- 1) A>V : Les stimuli auditifs sont plus longs, en termes syllabiques, que les stimuli orthographiques.  
(ex : /pulokama/ vs <PLOKAMA>)
  
- 2) A<V : Les stimuli orthographiques sont plus longs, en termes syllabiques, que les stimuli auditifs.  
(ex : /plokama/ vs <PULOKAMA>)

Il s'agit d'évaluer le nombre de fois où les sujets semblent compter les syllabes en modalité auditive, en modalité visuelle ou ni l'une ni l'autre. Il faudra néanmoins ensuite évaluer, dans la discussion, les différentes possibilités interprétatives : est-ce un effet de modalité, ou bien simplement un effet d'épenthèse dans la modalité présentant un groupe illégal ?

Deux autres sous-cas sont à distinguer ici :

- a) Lorsque, dans les calculs, tous les résultats sont pris en compte (c'est-à-dire lorsque l'on divise par le nombre total de stimuli divergents), incluant ceux qui ont été perçus

comme différents par les sujets. La somme totale des taux n'est alors pas égale à 100%, mais le nombre d'échantillons est identique pour tous les sujets.

b) Lorsque, dans les calculs, seuls les résultats effectifs sont pris en compte (c'est-à-dire lorsque l'on divise par le nombre de stimuli divergents perçus identiques par les sujets). La somme totale des taux est alors égale à 100%, mais le nombre d'échantillon varie selon les sujets.

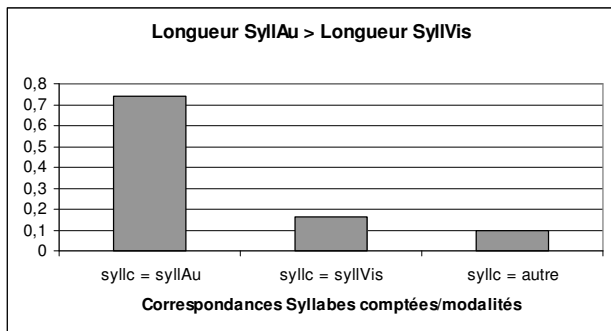
1) Longueur des stimuli auditifs > Longueur des stimuli visuels

a) Sur l'ensemble des échantillons

<b>Syllabes comptées</b>	<b>Moyenne</b>	<b>Ecart-type</b>
<b>En modalité auditive</b>	66, 3%	30, 2%
<b>En modalité visuelle</b>	13, 4%	23, 1%
<b>Ni l'une, ni l'autre</b>	8, 2%	6, 8%

b) Sur l'ensemble des échantillons perçus identiques

<b>Syllabes comptées</b>	<b>Moyenne</b>	<b>Ecart-type</b>
<b>En modalité auditive</b>	74%	30, 2%
<b>En modalité visuelle</b>	16, 6%	28%
<b>Ni l'une, ni l'autre</b>	9, 4%	7, 8%



Remarques :

- Dans les deux cas, il semblerait que la modalité auditive domine largement la modalité visuelle.
- On note environ 9% de comptage ne correspondant à aucune modalité.

**Quand le mot auditif est légal (plus long) en L1 et le mot visuel illégal (plus court) en L1 : la modalité auditive domine la modalité visuelle.**

2) Longueur des stimuli visuels > Longueur des stimuli auditifs

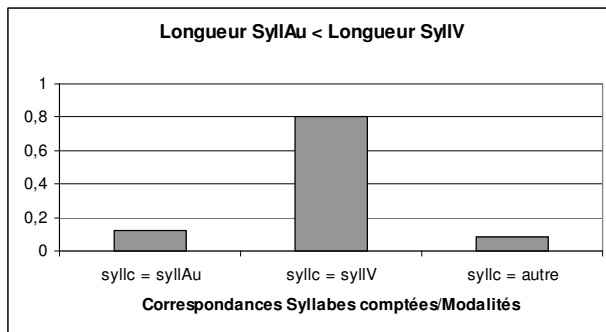
a) Sur l'ensemble des échantillons

Syllabes comptées	Moyenne	Ecart-type
En modalité auditive	10, 6%	15, 7%
En modalité visuelle	68, 4%	21, 3%
Ni l'une, ni l'autre	7, 4%	8, 2%

b) Sur l'ensemble des échantillons perçus identiques

Syllabes comptées	Moyenne	Ecart-type
En modalité auditive	11, 8%	17%
En modalité visuelle	80%	22, 5%
Ni l'une, ni l'autre	8, 2%	8, 8%





Remarques :

- Dans les deux cas, il semblerait que la modalité visuelle domine largement la modalité auditive.
- On note environ 8% de comptage ne correspondant à aucune modalité.

**Quand le mot visuel est légal (plus long) en L1 et le mot auditif illégal (plus court) en L1 : la modalité visuelle domine la modalité auditive.**

**Résumé Essentiel Stimuli Divergents Modalité Dominante**

- **Quand le mot auditif est légal en L1 (et donc plus long) et le mot visuel illégal en L1 (et donc plus court) : la modalité auditive domine largement la modalité visuelle.**
- **Quand le mot visuel est légal en L1 (et donc plus long) et le mot auditif illégal en L1 (et donc plus court) : la modalité visuelle domine largement la modalité auditive.**

**Cependant, la comparaison des résultats semble indiquer une légère dominance de la modalité visuelle :**

- **Illégal : Visuel (13%-17%) > Auditif (11%-12%)**
- **Légal : Visuel (68%-80%) > Auditif (66% - 74%)**

**Dans chaque cas on note entre 7% et 9% de décompte ne relevant d'aucune modalité de présentation**

## **B) Perception de la divergence intermodale**

Il s'agit d'évaluer le nombre de fois où les sujets ont effectivement perçu la différence inter-stimuli.

### *1) Résultats Globaux*

<b>Discrimination</b>	<b>Correcte</b>	<b>Erronée</b>
<b>Moyenne</b>	12, 9%	87, 1%
<b>Ecart-Type</b>	13, 9%	13, 9%

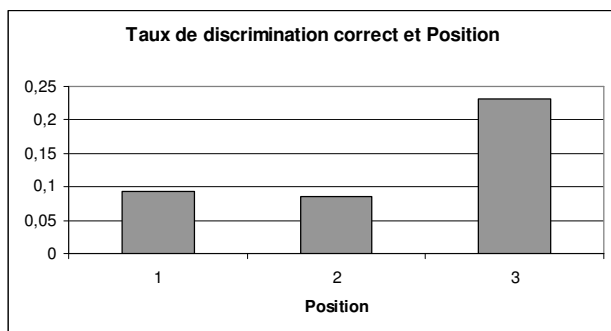
### Remarques :

- Le taux de perception correcte est très faible (environ 13%).

**Le taux de perception correcte de la divergence inter-stimuli est très faible (13%).**

### *2) Influence de la position*

<b>Position</b>	<b>Discrimination Correcte</b>		<b>Discrimination Erronée</b>	
	<b>Moyenne</b>	<b>Ecart-Type</b>	<b>Moyenne</b>	<b>Ecart-Type</b>
<b>1</b>	9, 2%	12, 5%	90, 8%	12, 5%
<b>2</b>	8, 5%	11, 3%	91, 5%	11, 3%
<b>3</b>	23, 1%	22, 6%	76, 9%	22, 6%



**Remarques :**

- En position 1 et 2 les taux de perception correcte sont très faibles (environ 9%).
- En position 3 le taux est plus élevé (environ 23%).

**La position semble influencer la perception de la divergence inter-stimuli.**

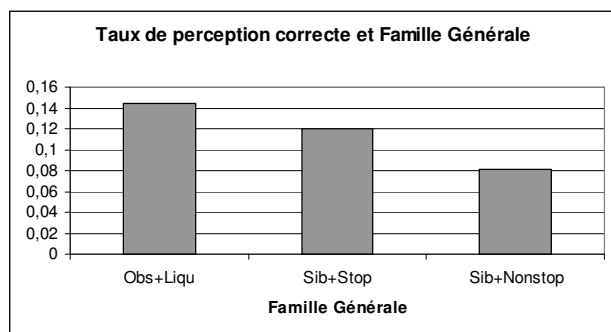
**Alors qu'en position d'attaque (syllabes initiales et finales) le taux de perception correcte est très faible (9%), il est plus élevé (23%) en position de coda de syllabe finale**

3) Influence de la nature

α) Nature phonologique

α). 1. Famille Générale

Famille Générale	Taux de Perception Correcte de la Divergence		
	Obs + Liqu (exemple : /kl/) (code : 1)	Sib + Stop (exemple : /sk/) (code : 3)	Sib + Nonstop (exemple : /sl/) (code : 4)
<b>Moyenne</b>	14, 44%	11, 95%	8, 06%
<b>Ecart-type</b>	15, 28%	18, 04%	9, 8%



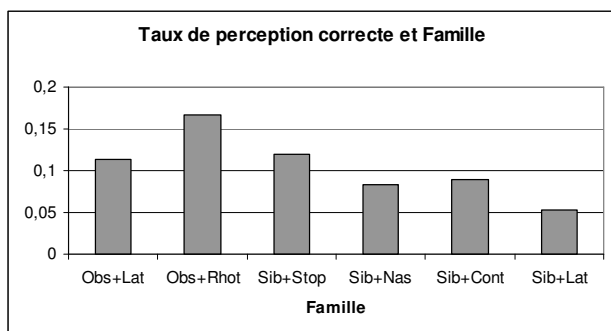
Remarques :

- Le classement obtenu est le suivant : Obs+Liqu (environ 14%) > Sib+Stop (environ 12%) > Sib+Nonstop (environ 8%).

**Divergence mieux perçue pour Obs+Liqu (14%), moins bien pour Sib+Nonstop (8%).**

α). 2. Famille

Famille	Taux de Perception Correcte de la Divergence	
	Moyenne	Ecart-type
Obs + Lat (code : 1) (exemple : /bl/)	11, 43%	13, 98%
Obs + Rhot (code : 2) (exemple : /bR/)	16, 73%	16, 93%
Sib +Stop (code : 7) (exemple : /sp/)	11, 95%	18, 04%
Sib + Nas (code : 8) (exemple : /sm/)	8, 39%	12, 44%
Sib + Cont (code : 9) (exemple : /sv/)	8, 96%	11, 35%
Sib + Lat (code : 10) (exemple : /sl/)	5, 24%	12, 09%



Remarques :

Le classement obtenu est le suivant : Obs+Rhot (environ 17%) > Sib+Stop (environ 12%) > Obs+Lat (environ 11%) > Sib+Cont (environ 9%) > Sib+Nas (environ 8%) > Sib+Lat (environ 5%).

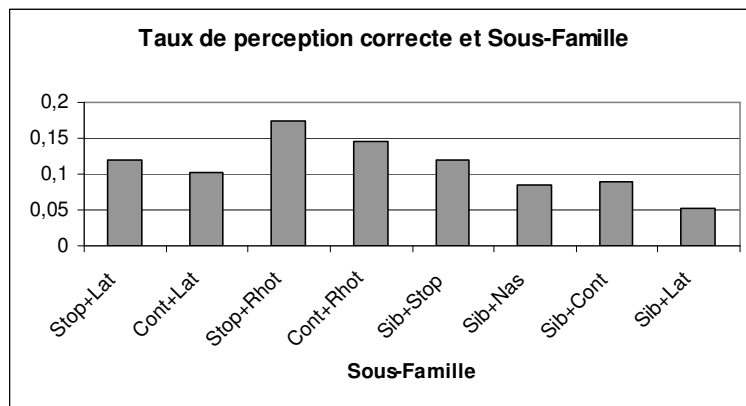
- Statut particulier de Obs+Rhot (17%) et Sib+Lat (5%).

**La divergence est mieux perçue pour les groupes Obs+Rhot (17%) et moins bien perçue pour les groupes Sib+Lat (5%).**

*α). 3.Sous-famille*

Sous-famille	Taux de Perception Correcte de la Divergence	
	Moyenne	Ecart-type
Stop + Lat (code : 1) (exemple : /bl/)	11, 96%	14, 4%
Cont + Lat (code : 2) (exemple : /fl/)	10, 26%	15, 83%
Stop + Rhot (code : 3) (exemple : /bR/)	17, 37%	17, 32%
Cont + Rhot (code : 4) (exemple : /fR/)	14, 66%	17, 21%

Sib +Stop (code : 9) (exemple : /sp/)	11, 95%	18, 04%
Sib + Nas (code : 10) (exemple : /sm/)	8, 39%	12, 44%
Sib + Cont (code : 11) (exemple : /sv/)	8, 96%	11, 35%
Sib + Lat (code : 12) (exemple : /sl/)	5, 24%	12, 09%



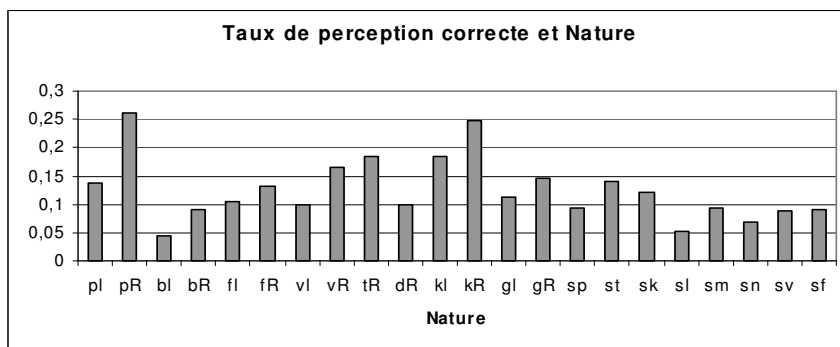
Remarque :

- Le classement obtenu est le suivant : Stop+Rhot (environ 17%) > Cont+Rhot (environ 15%) > Stop+Lat, Sib+Stop (environ 12%) > Cont+Lat (environ 10%) > Sib+Cont (environ 9%) > Sib+Nas (environ 8%) > Sib+Lat (environ 5%).
- Statut particulier de Stop+Rhot (17%) et de Sib+Lat (5%).

**La divergence est mieux perçue pour les groupes de type C+/R/ (16%) que pour ceux de type C+/L/ (9%).**

*α). 4. Nature spécifique*

<b>Groupe Consonantique</b>	<b>Taux de Perception Correcte de la Divergence</b>	
	<b>Moyenne</b>	<b>Ecart-type</b>
/pI/ (code : 14)	13, 71%	19, 21%
/bI/ (code : 16)	4, 3%	10, 43%
/kI/ (code : 24)	18, 55%	22, 4%
/gI/ (code : 26)	11, 29%	17, 28%
/fI/ (code 18)	10, 48%	17, 66%
/vI/ (code : 20)	10%	19, 08%
/pR/ (code : 15)	26, 08%	23, 69%
/bR/ (code : 17)	9, 03%	15, 23%
/tR/ (code : 22)	18, 55%	23%
/dR/ (code : 23)	9, 95%	16, 38%
/kR/ (code : 25)	24, 73%	26, 78%
/gR/ (code : 27)	14, 52%	19, 9%
/fR/ (code : 19)	13, 17%	17, 37%
/vR/ (code : 21)	16, 45%	22, 77%
/sp/ (code : 28)	9, 35%	19, 41%
/st/ (code : 29)	13, 98%	20, 06%
/sk/ (code : 30)	12, 1%	19, 83%
/sm/ (code : 32)	9, 41%	15, 86%
/sn/ (code : 33)	6, 85%	12, 94%
/sv/ (code : 34)	8, 87%	13, 74%
/sf/ (code : 35)	9, 14%	16, 16%
/sl/ (code : 31)	5, 24%	12, 09%



Remarque :

- Classement: /pR/ (environ 26%) > /kR/ (environ 25%) > /tR/, /kl/ (environ 19%) > /vR/ (environ 16%) > /gR/ (environ 15%) > /st/, /pl/ (environ 14%) > /fR/ (environ 13%) > /sk/ (environ 12%) > /gl/ (environ 11%) > /fl/, /dR/ (environ 10%) > /sm/, /sp/, /sf/, /bR/, /sv/ (environ 9%) > /sn/ (environ 7%) > /sl/ (environ 5%) > /bl/ (environ 4%).
- Aucun groupement clair n'apparaît, même si certains écarts sont très marqués.
- Extrêmes :  
Tête : /pR/ (26%), /kR/ (25%) et /tR/ (19%)  
Queue : /sl/ (5%) et /bl/ (4%)
- Classement de /f, v/ : /vR/ (16%) > /fR/ (13%) > /fl/ (10%) > /sf/ et /sv/ (9%).
- Voisement :

Groupe Non-voisé	Rapport (à 1% près)	Groupe Voisé
/pl/ (code : 14) 14%	>	/bl/ (code : 16) 4%
/kl/ (code : 24) 19%	>	/gl/ (code : 26) 11%
/fl/ (code 18) 10%	=	/vl/ (code : 20) 10%
/pR/ (code : 15) 26%	>	/bR/ (code : 17) 9%
/tR/ (code : 22) 19%	>	/dR/ (code : 23) 10%
/kR/ (code : 25) 25%	>	/gR/ (code : 27) 15%
/fR/ (code : 19) 13%	<	/vR/ (code : 21) 16%
/sf/ (code : 35) 9%	=	/sv/ (code : 34) 9%



On remarque alors que pour tous les groupes C+/R/ et C+/L/, la divergence est mieux perçue lorsque la consonne est non-voisée que lorsqu'elle est voisée, avec des écarts parfois très significatifs. En revanche lorsqu'une fricative /v/ ou /f/ est impliqué dans le groupe, la perception est soit identique, soit meilleure pour la consonne voisée /v/.

**Pas de catégories claires.**

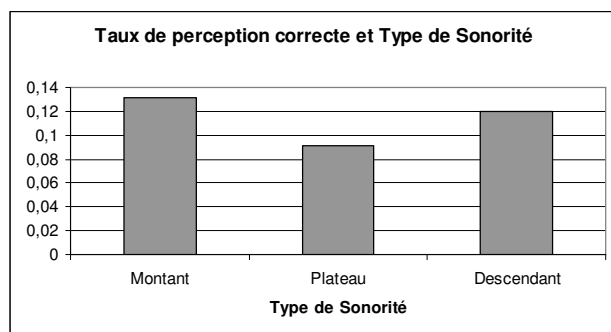
**La divergence est mieux perçue pour les groupes /pR, kR, tR/ (23%) et moins bien pour les groupes /sl, bl/ (5%).**

**Le voisement apparaît influencer la perception : pour tous les groupes C+/R/ et C+/L/ la divergence est mieux perçue lorsque la consonne est non voisée, excepté lorsqu'un fricative /v/ ou /f/ fait partie du groupe. Dans ce cas le rapport est égal, ou la divergence est mieux perçue pour la voisée /v/.**

*β) Sonorité*

*β).1. Type de Sonorité*

Type de sonorité	Taux de Perception Correcte de la Divergence	
	Moyenne	Ecart-type
Montant	13, 14%	13, 78%
Plateau	9, 14%	16, 16%
Descendant	11, 95%	18, 04%



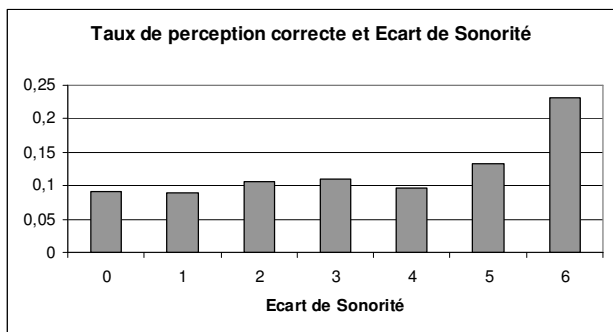
Remarque :

- Les écarts sont faibles.
- Le classement obtenu est le suivant : Montant (environ 13%) > Descendant (environ 12%) > Plateau (environ 9%).

**Peu de différence entre les trois types. La divergence est moins bien perçue pour le type Plateau (moins 3%).**

β).2. Ecart de Sonorité

Ecart de sonorité	Taux de Perception Correcte de la Divergence	
	Moyenne	Ecart-type
<b>0</b>	9, 14%	16, 16%
<b>1</b>	8, 87%	13, 74%
<b>2</b>	10, 5%	14, 65%
<b>3</b>	11, 08%	14, 58%
<b>4</b>	9, 59%	12, 77%
<b>5</b>	13, 29%	15, 87%
<b>6</b>	23, 12%	21, 03%



Remarques :

- Globalement, on remarque une tendance à une meilleure discrimination avec l'augmentation de l'écart.
- Le classement obtenu est le suivant : Ecart 6 (environ 23%) > Ecart 5 (environ 13%) > Ecart 3 (environ 11%) > Ecart 4, Ecart 2 (environ 10%) > Ecart 1, Ecart 0 (environ 9%).
- Statut particulier de l'écart 6 (23%).

**Globalement, plus l'écart de sonorité est grand plus la discrimination est meilleure, avec les taux compris entre 9% et 13% pour les écarts de 0 à 5 et un pic à 23% pour l'écart 6.**

4) Influence de la structure/longueur

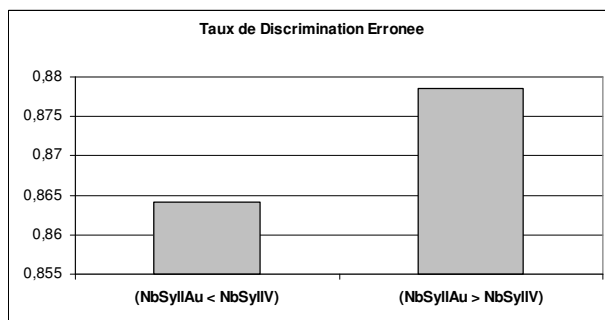
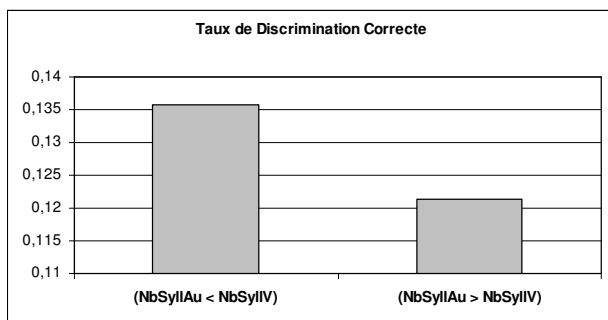
Deux cas dont à distinguer, selon le nombre de syllabe des stimuli (et ainsi selon leur caractère légal ou non en L1).

a) Longueur des stimuli auditifs > Longueur des stimuli visuels

Discrimination	Correcte	Erronée
Moyenne	12, 1%	87, 9%
Ecart-Type	16, 2%	16, 2%

b) Longueur des stimuli visuels > Longueur des stimuli auditifs

<b>Discrimination</b>	<b>Correcte</b>	<b>Erronée</b>
<b>Moyenne</b>	13, 6%	86, 4%
<b>Ecart-Type</b>	13, 4%	13, 4%



**Quelle que soit la modalité légal en L1 (et ainsi la longueur), les taux de perception correcte sont similaires (environ 13%).**

**Cependant, quand la modalité visuelle est légal en L1 (plus longue), le taux de perception correcte est plus élevé de 2%.**

**Résumé Essentiel Perception de la Divergence Inter-Stimuli**

**La position et la nature phonologique des groupes semblent influencer la perception de la divergence inter-stimuli.**

**Position :**

**Celle-ci est mieux perçue en position de coda de syllabe finale.**

**Nature :**

**Celle-ci est mieux perçue pour les groupes :**

- **OBLI (14%)**
- **Obs+Rhot (17%)**
- **/pR, kR, tR/ (23%)**

**Elle est moins bien perçue pour les groupes :**

- **Sib+Nonstop (8%)**
- **Sib+Lat (5%)**
- **/sl, bl/ (5%)**

**Globalement elle est mieux perçue pour les groupes C+/R/ que pour les groupes C+/L/.**

**Voisement :**

**Le voisement influence la perception de la divergence dans les groupes C+/R/ et C+/L/ :**

**Elle est mieux perçue lorsque C est non voisée.**

**Dans les groupes avec fricatives (/v,f/), elle est mieux ou également perçue avec la consonne voisée /v/.**

**Sonorité :**

- **Le type de sonorité a peu d'influence, même si la divergence est moins bien perçue pour le type Plateau.**
- **L'écart de sonorité semble influencer la perception de la divergence : globalement, plus l'écart est grand, plus la perception est meilleure, en particulier pour l'écart maximal (6) qui se distingue des autres.**

**En revanche, le rapport entre modalité et légalité (et ainsi longueur) des items ne semble pas influencer la perception de la divergence : quelle que soit la modalité légale en L1, les taux de perception sont similaires.**

*Comparaison des performances : Comptage syllabique vs. Perception de la divergence inter-stimuli*

**1) Taux de réussite**

Dans le test E, les sujets devaient effectuer deux sous-tâches simultanément :

- 1) Compter correctement le nombre de syllabes des stimuli identiques.
- 2) Percevoir correctement la divergence des stimuli différents.

Il est possible que les sujets aient mieux réussi une sous-tâche que l'autre, différence qui pourrait signaler un certain partage attentionnel dans l'accomplissement de la tâche complexe E. L'influence des consignes sur un tel partage sera discutée plus bas. Si l'on attribue un score de réussite à chacune de ces tâches, voici ce que l'on obtient en moyenne :

<b>Taux de réussite</b>	<b>Comptage syllabique</b>	<b>Perception de la divergence</b>
<b>Moyenne</b>	39, 3%	12, 88%
<b>Ecart-type</b>	15, 98%	13, 91%

Remarques :

- Le taux de réussite à la tâche de comptage (environ 40%) est nettement supérieur à celui de la tâche de perception (environ 13%).

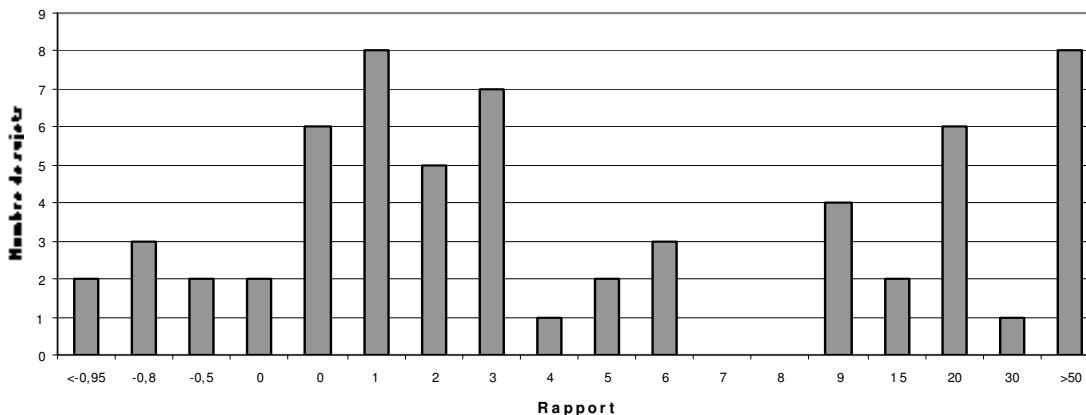
**La tâche de comptage est mieux réussie (40%) que la tâche de perception de la divergence (13%).**

## **2) Répartition des taux**

Afin de déterminer exactement le rapport entre les deux performances, il est nécessaire de déterminer si les sujets ayant bien réussi à une tâche ont également bien réussi l'autre, ou bien si le rapport entre les deux est différent. Pour cela, nous avons attribué une note de réussite à chaque tâche (comptage et perception), puis comparé pour chaque sujet ces deux notes. Nous obtenons ainsi la répartition des sujets en fonction du rapport des deux notes. Par exemple, nous pouvons voir dans le tableau ci-dessous que 6 sujets ont obtenu une note de comptage comprise dans l'intervalle [note de perception, note de perception + 100%] et que 8 sujets ont obtenu une note de comptage comprise dans l'intervalle [note de perception+100%, note de perception+200%], etc.

<b>Nombre de sujets dont la note de Comptage est supérieure à celle de Perception</b>	<b>Note de Comptage = Note de Perception + X pour cent</b>	<b>Nombre de sujets dont la note de Comptage est inférieure à celle de Perception</b>	<b>Note de Comptage = Note de Perception - X pour cent</b>
6	100%	2	50%
8	200%	2	80%
5	300%	3	95
7	400%	2	- de 95%
1	500%		
2	600%		
3	700%		
0	800%		
0	900%		
4	1500%		
2	2000%		
6	3000%		
1	5000%		
8	+ de 5000%		

### Répartition des résultats de Comptage syllabique par rapport à ceux de Perception de la divergence



#### Remarques :

- Alors que 8 sujets ont obtenu plus du double, 39 sujets ont obtenu, comme note de comptage, plus du triple de la note de perception.
- Seuls 9 sujets ont une note de comptage inférieure à la note de perception, mais le rapport est néanmoins inférieur au double.

**85,5% des sujets ont obtenu une note de comptage supérieure à leur note de perception. Parmi ceux-ci, 15% ont obtenu plus du double et 74% ont obtenu plus du triple.**

**Seuls 14,5% des sujets ont obtenu une note de comptage inférieure à leur note de perception. L'écart est cependant inférieur au double.**



## Comparaison des Tâches C2 (auditive), D (visuelle) et E

### (audiovisuelle)

Lorsque l'on compare les résultats des tâches C2, D et E, voici ce que l'on obtient<sup>194</sup> pour les cibles suivantes<sup>195</sup> :

- /pl, pr, bl, br, fl, fr, vl, vr, tr, dr, kl, kr, gl, gr/ (OBLI)
- /sp, st, sk/ (S + Plosive non-voisée)
- /sl, sm, sn, sv, sf/ (S + C non-plosive)

### 1. Résultats globaux

Taux d'épenthèses	Modalité Auditive	Modalité Visuelle	Modalité Audiovisuelle
<b>Moyenne</b>	58, 5%	77, 03%	66, 5%
<b>Ecart-type</b>	25, 63%	30, 96%	27, 8%

**Modalité Visuelle (77%) > Modalité Audiovisuelle (66%) > Modalité Auditive (58%)**

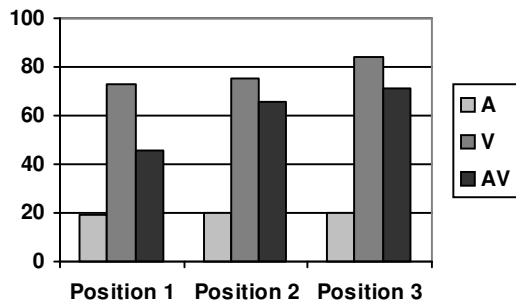
**Le taux d'épenthèse semble donc influencé par la modalité.**

### 2. Position

Famille Générale	Position 1	Position 2	Position 3
<b>Modalité Auditive</b>	19%	20%	20%
<b>Modalité Visuelle</b>	73%	75%	84%
<b>Modalité Audiovisuelle</b>	61%	66%	71%

<sup>194</sup> Par la suite, A=Auditif (→ C2), V=Visuel (→ D) et AV=Audiovisuel (→ E).

<sup>195</sup> Lorsque cela est approprié, nous ne présentons que les résultats les plus pertinents.



**Dans tous les cas :**

**Modalité V > Modalité AV > Modalité A.**

**Convergence intermodale :**

**Le classement est le même pour les trois modalités :**

**Position 3 (coda finale) > Position 2 (attaque finale) > Position 1 (attaque initiale).**

**Divergence intermodale :**

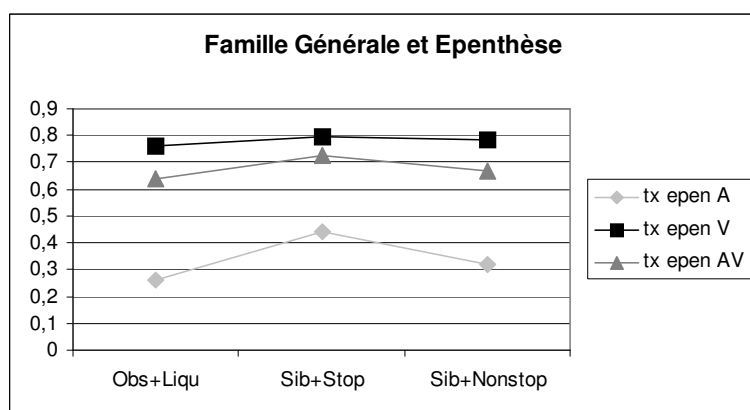
- **Ecart Position 1-Position 2 : peu marqué en A et en V (+2%), plus marqué en AV (+5%)**
- **Ecart Position 2-Position 3 : V (+9%) > AV (+5%) > A (+1%)**
- **Ecart Position 1-Position 3 : V (+11%) > AV (+10%) > A (+1%)**

**La position semble influencer les taux en modalités V et AV. Dans ce cas, la position 3 déclenche davantage d'épenthèse.**

**L'influence de la position sur les taux d'épenthèses semblerait donc dépendre de la modalité.**

### 3. Famille Générale

Famille Générale	Obs+Liqu	Sib+Stop	Sib+Nonstop
Modalité Auditive	26%	44%	32%
Modalité Visuelle	76%	79%	78%
Modalité Audiovisuelle	64%	73%	67%



**Dans tous les cas :**

**Modalité V > Modalité AV > Modalité A**

**Convergence intermodale :**

**Le classement est identique dans les trois modalités :**

**Sib+Stop > Sib+Nonstop > Obs+Liqu.**

**Divergence intermodale :**

**Ecart entre extrêmes :**

**Modalité Auditive (+18%) > Modalité Audiovisuelle (+9%) > Modalité Visuelle (+3%).**

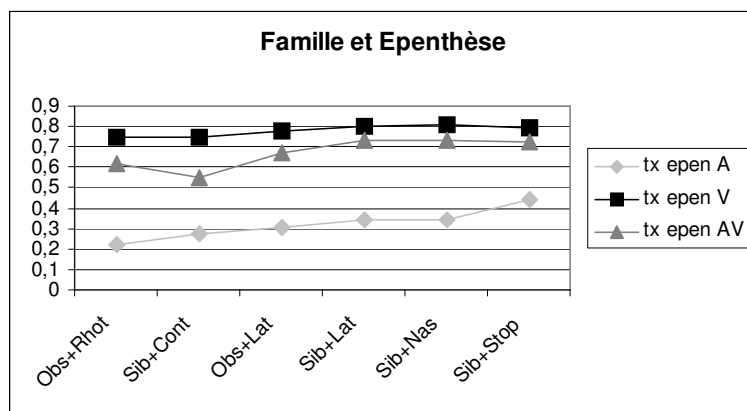
**L'influence de la famille générale sur les taux d'épenthèse semble donc dépendre en partie de la modalité : elle diminue selon le classement suivant :**

**Auditif > Audiovisuel > Visuel**

#### **4. Famille**

Désormais, les résultats sont présentés selon le taux croissant d'épenthèses en modalité auditive.

<b>Famille</b>	<b>Modalité Auditive</b>	<b>Modalité Visuelle</b>	<b>Modalité Audiovisuelle</b>
<b>Obs+Rhot</b>	22%	75%	61%
<b>Obs+Lat</b>	31%	78%	67%
<b>Sib+Stop</b>	44%	79%	73%



**Dans tous les cas :**

**Modalité V > Modalité AV > Modalité A.**

**Convergence intermodale :**

**Sib+Stop > Obs+Lat > Obs+Rhot.**

**Divergence intermodale :**

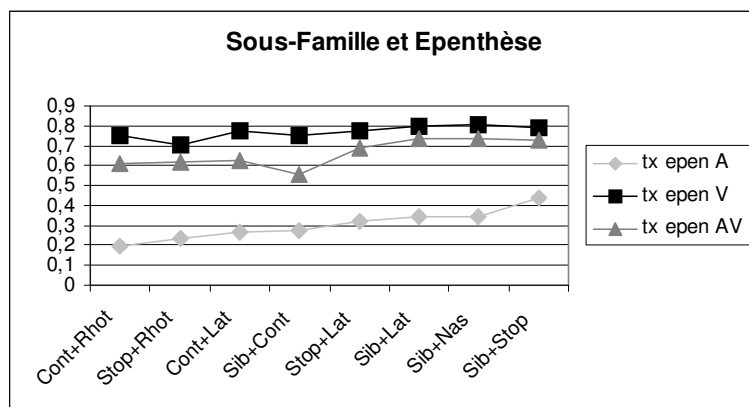
**On remarque certaines différences de classement pour :**

- **Sib+Cont (taux moindre au sein de la modalité AV).**
- **Sib+Stop (taux moindres au sein des modalités AV et V).**
- **Sib+Nas (taux supérieur (maximal) au sein des modalités AV et V).**

**L'influence de la famille sur les taux d'épenthèse semble dépendre en partie de la modalité. Elle diminue globalement ainsi : Auditif > Audiovisuel > Visuel.**

### 5. Sous-Famille

Sous-Famille	Modalité Auditive	Modalité Visuelle	Modalité Audiovisuelle
<b>Cont+Rhot</b>	20%	75%	61%
<b>Stop+Rhot</b>	23%	71%	62%
<b>Cont+Lat</b>	27%	78%	63%
<b>Stop+Lat</b>	32%	78%	69%



**Dans tous les cas :**

**Modalité V > Modalité AV > Modalité A.**

**Convergence intermodale :**

**Modalité V > Modalité AV > Modalité A.**

**Divergence intermodale :**

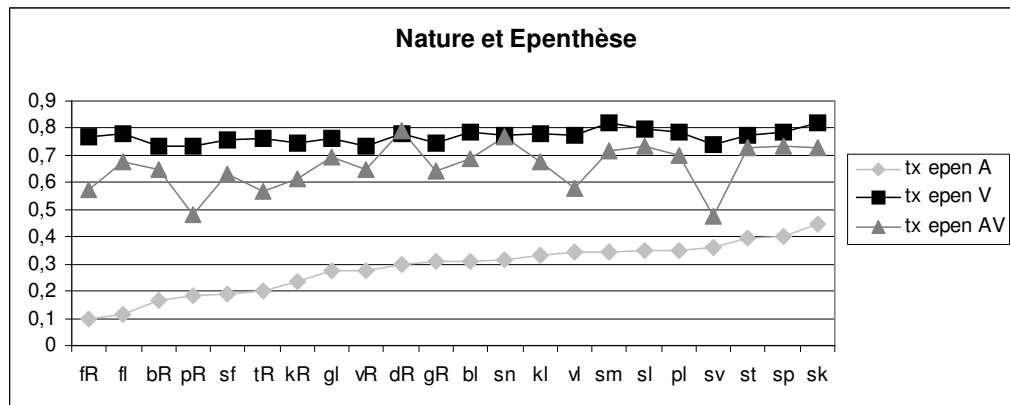
**On remarque certaines différences de classement : en particulier, pour les groupes C+/R/ et C+/L/, alors que les classements sont identiques pour A et AV, celui de V est différent :**

- **A et AV** : Stop+Lat > Cont+Lat > Stop+Rhot > Cont+Rhot.
- **V** : Stop+Lat et Cont+Lat > Cont+Rhot > Stop+Rhot.

L'influence de la famille sur les taux d'épenthèse semble dépendre en partie de la modalité : elle diminue globalement selon le classement suivant :

**Auditif > Audiovisuel > Visuel.**

## 6. Nature



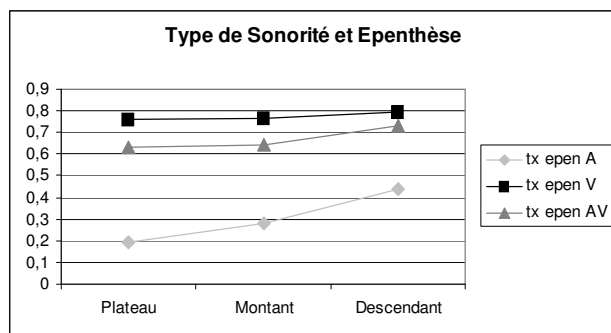
### Voisement :

<b>Modalité A</b>	<ul style="list-style-type: none"> <li>• <u>C+/l/</u> : C non-voisée &gt; C voisée</li> <li>• <u>C+/R/</u> : C voisée / &gt; (ou égal) C non-voisée</li> <li>• <u>Fricatives f/v</u> : taux supérieur pour /v/</li> </ul>
<b>Modalité V</b>	<ul style="list-style-type: none"> <li>• <u>C+/l/</u> : Pas d'influence du voisement</li> <li>• <u>C+/R/</u> : Pas d'influence du voisement</li> <li>• <u>Fricative f/v</u> : taux supérieur pour /f/</li> </ul>
<b>Modalité AV</b>	<ul style="list-style-type: none"> <li>• <u>C+/l/</u> : Pas d'influence du voisement</li> <li>• <u>C+/R/</u> : C voisée &gt; C non-voisée</li> <li>• <u>Fricative f/v</u> : pas d'influence</li> </ul>

**Dans tous les cas : Modalité V > Modalité AV > Modalité A, excepté pour /dR/ et /sn/ (taux identiques en modalités V et AV).**

- **Les classements sont différents** : alors que le taux augmente de manière quasi-linéaire en modalité A, et qu'il est quasi-constant en modalité V, il est beaucoup plus irrégulier en modalité AV.
- **Taux maximaux** :
  - A : /sk, sp, st/ (42%)
  - V : /sk, sm, sl/ (81%)
  - AV : /dR, sn/ (78%) et /sl, sp, sk, st/ (73%)
- **Taux minimaux** :
  - A : /pR, bR/ (17%) et /fl, fR/ (11%)
  - V : /pR, bR, vR, sv/ (74%)
  - AV : /pR, sv/ (48%) et /fR, tR, vl/ (57%)
- **Voisement** : l'influence du voisement semble varier avec la modalité.

## 7. Type de Sonorité

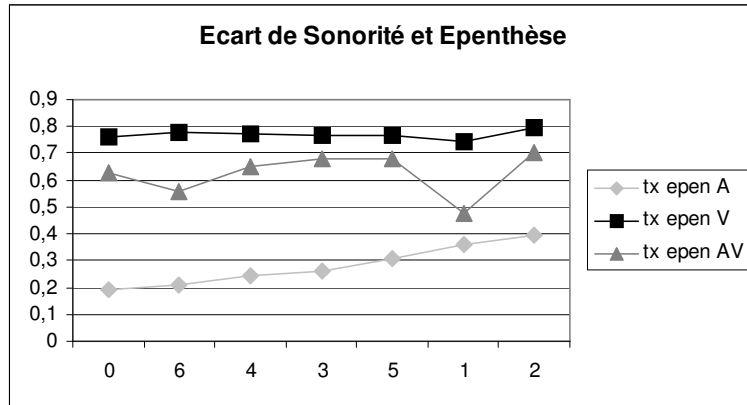


**Les classements sont identiques : Descendant > Montant > Plateau.**

**On constate une diminution des écarts entre les taux dans l'ordre suivant :**

**écart A > écart AV > écart V.**

## **8. Ecart de Sonorité**



**Les classements sont différents. Les modalités A et AV, malgré un classement assez similaire, divergent en particulier sur les écarts 6 et 1. Cependant les trois modalités convergent sur l'écart 2 (Taux maximal en A, V et AV).**



### **Résumé Essentiel Comparaison Tâches C2/D/E**

**La comparaison des résultats des trois tâches permet d'évaluer :**

- **L'impact de la modalité de présentation des stimuli sur les taux d'épenthèses.**
- **Les convergences et divergences des effets des variables indépendantes entre les trois modalités.**

#### **Taux d'épenthèses :**

**Modalité Visuelle (77%) > Modalité Audiovisuelle (66%) > Modalité Auditive (58%). Le taux d'épenthèse semble donc influencé par la modalité.**

#### **Position :**

**Pour les trois modalités : Position 3 (coda finale) > Position 2 (attaque finale) > Position 1 (attaque initiale).**

**Mais il y a variation dans les écarts : Ecart Position 1-Position 3 : Visuel (+11%) > Audiovisuel (+10%) > Auditif (+1%). L'influence de la position semble donc dépendre de la modalité : elle est beaucoup plus sensible en modalités V et AV.**

#### **Propriétés phonologiques :**

##### **Globalement :**

- **Modalité V > Modalité AV > Modalité A.**
- **La sensibilité aux propriétés phonologiques décroît ainsi :  
Modalité A > Modalité AV > Modalité V.**

#### **Convergences relatives :**

- **Famille Générale : pour les trois modalités :  
Sib+Stop > Sib+Nonstop > Obs+Liqu.**

- **Famille** : pour les trois modalités :  
Sib+Stop > Obs+Lat > Obs+Rhot.
- **Nature** : certaines tendances globales communes :  
/s/+C > C+/l/ > C+/R/,  
malgré un classement différent.
- **Type de sonorité** : classement identique :  
Descendant > Montant > Plateau.

**Divergences particulières** :

- **Famille générale** : Ecart entre extrêmes : Modalité Auditive (+18%) > Modalité Audiovisuelle (+9%) > Modalité Visuelle (+3%). L'influence (quantitative) de la famille générale semble donc dépendre de la modalité.
- **Famille** : certaines différences de classement pour Sib+Cont, Sib+Stop et Sib+Nas.
- **Sous-Famille** : certaines différences de classement. En particulier, pour C+/R/ et C+/l/ : classement en A et AV identiques, mais différent en V.
- **Nature** : courbe de taux pour A et V quasi linéaire, mais très irrégulière pour AV. Cependant les extrêmes sont identiques (Max : /sk/ ; Min : /pR/).
- **Voisement** : L'influence du voisement semble varier selon la modalité.
- **Ecart de Sonorité** : classement assez différent, même si les trois modalités convergent sur l'écart 2 (taux maximal dans les trois modalités).



## **Partie Expérimentale IV :**

### **Tests – Interprétation**

L'interprétation de résultats comme les nôtres n'est pas, comme pour toute interprétation de résultats de ce type, chose aisée : les données ne présentent pas toujours de régularités aussi remarquables qu'on l'aurait souhaité, et établir un scénario causal en vue de rendre compte de ces régularités souffre bien évidemment de la pluralité des scénarii explicatifs possibles. Si l'on songe en outre aux trois domaines (phonologie, psycholinguistique et didactique) auxquels nous nous référons, ainsi qu'aux différents niveaux d'analyse que nous sollicitons (notamment phonétique, phonologique et phonographémique), il apparaît que l'interprétation que nous proposons ci-dessous peut parfois manquer en partie d'une certaine finesse pourtant souhaitable. Rappelons enfin que certaines mesures, dans les tests segmentaux A et F en particulier, étaient tout à fait exploratoires, et visaient essentiellement à fournir des données descriptives quantitatives. Nous avons donc décidé de n'inclure dans l'interprétation que les résultats les plus essentiels et les plus informatifs, en privilégiant, dans une certaine mesure, ceux aux implications didactiques potentielles remarquables, puisque notre travail, rappelons-le, est fondamentalement ancré dans le domaine de la didactique du FLE.

Avant d'aborder l'interprétation des résultats proprement dite, il convient cependant d'apporter certaines précisions concernant le protocole, en particulier en ce qui concerne les tâches syllabiques.

#### **Consignes : le choix du terme « syllabe »**

En ce qui concerne les consignes, il convient de s'arrêter sur les implications du choix du terme « syllabe » (en japonais « onsetsu ») dans les tâches de comptage syllabique et de rappeler que, comme l'indique Labrune (2001b, p. 111) :

*« Pour les locuteurs japonais, l'objet « syllabe » tel que peut l'appréhender la linguistique occidentale ne possède aucune réalité intuitive [...] ce terme renvoyant tantôt à la more, tantôt à une unité supérieure à celle-ci de type syllabe lourde, tantôt aux deux à la fois ».*

Il semblerait donc que l'apprentissage de la phonologie du français pour les sujets japonais impliquerait, non seulement l'apprentissage de la « syllabation » du français, comme pour tout apprenant, « occidental » ou non, mais également celui de la « syllabation » du français en tant que principe interne d'organisation phonologique prosodique. Labrune (2001a, p. 90, note 6) indique que trois termes sont actuellement utilisés par les linguistes japonais : « haku » (terme d'origine musicale signifiant « battement, coup frappé »), « onsetsu » (« groupe de sons », terme pouvant signifier tout aussi bien syllabe que more) et « moura » (more), ce dernier étant généralement inconnu des non-spécialistes (Akamatsu, 1997, pp. 160-165). Nous ne pouvons donc exclure que l'interprétation du terme « onsetsu » par nos sujets ait pu influencer leurs performances, notamment dans une interprétation moraïque. S'ils sont en effet habitués à compter les unités de temps (« haku ») et si la consigne a été comprise dans ce sens, il se pourrait que certains éléments segmentaux aient pu être comptés comme unités moraïques, ne correspondant pas aux « syllabes occidentales ».

### **Méthodologie : tâche explicite**

Le choix d'une tâche explicite (les sujets sont conscients, de manière relative, de l'objectif réel, en totalité ou en partie, de la tâche) ou implicite (les sujets ne sont pas informés des objectifs réels de la tâche) est, comme pour les choix des consignes, du ressort de l'expérimentateur. Dans notre cas, pour les tests syllabiques, le choix d'une tâche de comptage syllabique explicite s'explique ainsi : étant donné l'attention portée, *a priori*, au contenu linguistique du matériel verbal utilisé dans le cadre didactique d'une situation d'enseignement/ apprentissage, nous nous intéressions précisément à l'influence (inconsciente pour les sujets) des relations entre l'orthographe et la phonologie du japonais (L1) sur la perception et le traitement *conscients* des unités phonologiques du français (L2). Notre hypothèse était que le traitement conscient des

unités perçues entraîne (inconsciemment) une modification (épenèses) desdites unités. De la même manière pour les tests segmentaux, nous cherchions à évaluer, d'une certaine manière, l'influence (inconsciente) des unités phonographémiques en L1 (japonais), en tant qu'unités, explicites et implicites, du système phonographémique des sujets, sur la perception des unités phonographémiques en L2 (français).

Ces deux points étant clarifiés, nous pouvons à présent aborder l'analyse des tâches et l'interprétation des résultats. Pour chaque tâche, après une introduction générale permettant de détailler les opérations cognitives potentiellement impliquées dans celles-ci, les résultats sont d'abord examinés de manière linéaire, avant d'être repris de manière plus synthétique pour tenter de répondre aux problématiques initiales.

Rappelons enfin deux hypothèses de base, telles que formulées dans la littérature, sur lesquelles nous nous appuyons dans notre interprétation :

- L'hypothèse du transfert plein : à l'état initial, l'interlangue des sujets est totalement issue du transfert de leur langue première, tant au niveau phonétique, phonologique que phonographémique.
- L'apprentissage phonétique : l'apprentissage des catégories phonologiques s'effectue sur la base d'exemplaires phonétiques.

## Table des Matières Partielle

I. TACHE A : APPARIEMENT SEGMENTAL PHONEME - GRAPHEME (L/R ; B/V).....	- 940 -
<i>I.1. Introduction</i> .....	- 940 -
I.1.1. Mode de traitement.....	- 940 -
I.1.2. Interphonologie : catégories phonétiques, phonologiques et graphémiques .....	- 943 -
<i>I.2. Interprétation linéaire des résultats</i> .....	- 945 -
I.2.1. La position.....	- 945 -
I.2.2. La nature de la consonne.....	- 947 -
I.2.3. La nature de la voyelle adjacente .....	- 949 -
I.2.4. Choix de graphèmes .....	- 953 -
II. TACHE F : DISCRIMINATION PHONOLOGIQUE ET APPARIEMENT SEGMENTAL PHONEME - GRAPHEME (L/R ; B/V).....	- 954 -
<i>II.1. Introduction</i> .....	- 954 -
<i>II.2. Interprétation linéaire des résultats phonographémiques (premier stimulus)</i> .....	- 955 -
II.2.1. La position .....	- 955 -
II.2.2. La nature de la consonne .....	- 956 -
II.2.3. La nature de la voyelle adjacente .....	- 958 -
II.2.4. La nature du couple de stimuli (identiques/différents) .....	- 960 -
<i>II.3. Interprétation linéaire des résultats phonologiques</i> .....	- 960 -
II.3.1. La position .....	- 960 -
II.3.2. La nature des consonnes .....	- 960 -
II.3.3. La nature de la voyelle adjacente .....	- 961 -
II.3.4. La nature du couple de stimuli (identiques/différents) .....	- 962 -
III. TACHES C1 ET C2 : COMPTAGE SYLLABIQUE EN UNIMODALITE AUDITIVE .....	- 963 -
<i>III.1. Tâche C1 - Epenthèses</i> .....	- 964 -
III.1.1. Position .....	- 964 -
III.1.2. Nature .....	- 966 -
III.1.3. Sonorité.....	- 969 -
<i>III.2. Tâche C1 – Effacements</i> .....	- 969 -
III.2.1. Position .....	- 971 -
III.2.2. Nature .....	- 971 -
III.2.3. Sonorité.....	- 972 -
<i>III.3. Tâche C1 – Comparaison Epenthèses/Effacements</i> .....	- 973 -
<i>III.4. Tâche C2 - Epenthèses</i> .....	- 974 -
III.4.1. Position .....	- 974 -
III.4.2. Nature .....	- 974 -
III.4.3. Sonorité.....	- 978 -
<i>III.5. Tâche C2 - Effacements</i> .....	- 979 -
III.5.1. Position .....	- 979 -
III.5.2. Nature .....	- 979 -

III.6. Tâche C2 – Comparaison Epenthèses/Effacements.....	- 980 -
IV. TACHE D : COMPTAGE SYLLABIQUE EN UNIMODALITE ORTHOGRAPHIQUE .....	- 980 -
IV.1. Position .....	- 982 -
IV.2. Nature .....	- 983 -
V. TACHE E : COMPTAGE SYLLABIQUE EN BIMODALITE AUDIO-ORTHOGRAPHIQUE .....	- 983 -
V.1. Stimuli auditifs et orthographiques convergents.....	- 987 -
V.1.1. Position .....	- 987 -
V.1.2. Nature .....	- 988 -
V.1.3. Sonorité.....	- 989 -
V.2. Stimuli auditifs et orthographiques divergents.....	- 989 -
V.2.1. Modalité dominante en tâche de comptage syllabique.....	- 989 -
V.2.2. Discrimination intermodale de stimuli divergents .....	- 991 -
V.2.2.1. Position .....	- 991 -
V.2.2.2. Nature.....	- 992 -
V.2.2.3. Sonorité .....	- 994 -
V.2.2.4. Longueur/Structure des stimuli divergents.....	- 994 -
V.3. Comptage syllabique vs. Discrimination intermodale.....	- 995 -
VI. COMPARAISON TACHES C2, D ET E : UNIMODALITE AUDITIVE VS. UNIMODALITE ORTHOGRAPHIQUE VS. BIMODALITE AUDIO-ORTHOGRAPHIQUE .....	- 995 -
VII. SYNTHÈSE ET PERSPECTIVES .....	- 1003 -
VII.1. Synthèse .....	- 1004 -
VII.2. Implications méthodologiques .....	- 1011 -
VII.3. Perspectives .....	- 1013 -
VII.3.1. Phonétique-Phonologie.....	- 1013 -
VII.3.2. Psycholinguistique .....	- 1016 -
VII.3.3. Didactique et Apprentissage.....	- 1019 -



## **I. Tâche A : appariement segmental phonème - graphème (L/R ; B/V)**

### **I.1. Introduction**

Globalement, le taux de confusion phonographémique pour les couples (B, V) et (R, L) est moindre que ce à quoi nous nous attendions, avec 29% de confusions. Cela semblerait signifier : 1) que les sujets font relativement peu de confusions phonologiques entre les phonèmes considérés, 2) qu'ils maîtrisent globalement le système de correspondances phonographémiques du français. Cependant, 29% reste un taux élevé de confusions, et la question se pose de savoir si ces confusions sont d'ordre phonético-phonologique (mauvaise « perception phonologique », mais bonnes correspondances phonographémiques), phonographémique (bonne « perception phonologique », mais mauvaises correspondances phonographémiques), ou encore si elles sont liées aux deux. Afin de pouvoir examiner plus en détail les résultats obtenus, il nous faut donc préciser le contenu « théorique » de la tâche accomplie par les sujets.

#### **I.1.1. Mode de traitement**

Cette tâche d'appariement phonème – graphème (ou plus exactement : stimulus phonétique – transcription orthographique, puisque les stimuli étaient des non-mots et non des phonèmes isolés) peut être abordée essentiellement de trois manières, dans un premier temps au moins :

- Mode A) Activation phone → phonème → graphème :

**Perception auditive**

(influence des caractéristiques acoustiques du signal et influence des connaissances phonotactiques)

**→ Activation d'une représentation phonologique**

(voie d'assemblage ; influence des contraintes phonologiques)

**→ Activation d'une représentation orthographique via le système de correspondances phonographémiques**

(influence de la fréquence d'association phonographémique)

**→ Choix de la transcription orthographique correspondant à cette représentation.**

Le mode A est donc un mode essentiellement linéaire, dans lequel *l'input* de chaque étape correspond à *l'output* de la précédente.

- Mode B) Activation phone → (phonème ; graphème) → graphème :

**Perception auditive**

**→ Activation d'une représentation phonologique et d'une représentation orthographique**

**→ Choix de la transcription orthographique correspondant au résultat de la « confrontation » entre ces deux dernières représentations.**

Le mode B est davantage interactif, dans le sens où la perception du stimulus entraîne l'activation parallèle de catégories phonologiques et graphémiques, qui peuvent rentrer en compétition ou s'influencer dans le choix de la représentation graphémique finale.

- Mode C) Comparaison phonème/phonème :

**Perception auditive**

→ **Activation d'une représentation phonologique**

(maintien en mémoire de travail)

→ **Perception visuelle des transcriptions orthographiques proposées et activation des représentations phonologiques correspondantes**

(conversion graphème – phonème)

→ **Comparaison des représentations phonologiques**

→ **Choix de la transcription orthographique issue de la comparaison phonologique.**

Le mode C implique plus fortement deux types de perception, auditive et visuelle, toutes deux activant certaines représentations phonologiques, similaires ou non.

Etant donné notre protocole, il semblerait possible de privilégier *a priori* le mode C), dans le sens où il semblerait mieux correspondre d'une part à la situation quasi-expérimentale (les sujets disposent des transcriptions avant d'entendre les stimuli) et d'autre part à l'élaboration du protocole (tâche d'*appariement* et non de *dictée*).

Cependant, il convient de bien souligner les contraintes exercées par le protocole sur la réalisation de la tâche : dès les premiers stimuli, le protocole permet aux sujets de comprendre qu'il s'agit essentiellement de distinguer deux catégories phonético-phonologiques placées ici en opposition (B/V et R/L). Pour des apprenants de niveau faux-débutant ou intermédiaire comme nos sujets (contrairement à des sujets monolingues non-apprenants), le problème n'est donc *a priori* pas un problème d'homophones hétérographes (chaque catégorie phonologique étant exclusivement associée à un graphème), mais plutôt un problème de perception et de correspondances entre catégorie phonétiques et phonologiques. Le format « fermé » binaire des réponses oriente donc fortement l'attention des sujets vers la perception auditive (puisque c'est

l'information dont ils ne disposent pas, contrairement aux transcriptions qui figurent sur leur document de réponse), et cette binarité leur permet en outre d'adopter une stratégie de résolution « par défaut », quel que soit le niveau d'analyse : « Si ce n'est pas X, alors c'est Y ».

L'évaluation par le sujet des informations disponibles vise ainsi à identifier ou à rejeter « X ». Deux procédures d'identification existent alors :

- Identification erronée : choix « délibéré » de « Y »
- Evaluation erronée (rejet) de « X » : choix « par défaut » de « Y »

Dans cette optique, l'attention des sujets semblerait donc effectivement concentrée sur la perception auditive des stimuli, tandis que les traitements phonologique et phonographémique ultérieurs seraient moins « contrôlés ». Plusieurs facteurs peuvent influencer les processus conduisant aux résultats erronés, et le mode A semblerait dans ce cas plus à même de décrire les processus impliqués dans la réalisation de la tâche.

### **I.1.2. Interphonologie : catégories phonétiques, phonologiques et graphémiques**

Tout au long des processus décrits plus haut, il convient de distinguer, lorsque cela est possible, ce qui relève de la L1 de ce qui relève de l'interlangue (IL). De plus, il faut clairement distinguer les catégories *phonétiques* et *phonologiques* de l'IL, de son système de correspondances *phonographémiques*, et l'on voit bien que les confusions peuvent être liées à différents niveaux d'analyse (perceptif, phonologique, phonographémique).

Si l'on considère que les apprenants disposent d'une certaine compétence interphonologique (IL), et si l'on distingue les plans perceptif (ex : [I]), phonologique (ex : /I/) et graphémique (ex : « l »), on doit alors se confronter à des triplets représentationnels : (ex : ([I], /I/, « l »)) entre la L1 et l'IL. La situation d'apprentissage des sujets peut alors conduire, dans le système interphonologique, à des associations

variées entre ces trois domaines, et plusieurs possibilités sont alors théoriquement envisageables.

Prenons l'exemple du couple /l, R/, en considérant que, pour tous les sujets, nous avons : L1 = ([l], [r], /r/, « r »)<sup>196</sup>. Plusieurs possibilités, *tout à fait hypothétiques*, sont alors envisageables et nous n'en mentionnons que quelques-unes ci-dessous, en laissant temporairement de côté le niveau phonétique :

**Exemple 1 : IL = L1 (système « nul »)**

Dans ce cas, le sujet est un débutant complet : le transfert est alors « plein » et l'apprentissage « nul ».

**Exemple 2 : IL  $\supset$  /R/ + (/R/, « r ») (système incomplet)**

L'IL inclut /R/ et (/R/, « r »), sans préjuger des autres modifications apportées au système initial (IL=L1).

- D'un point de vue phonologique, le sujet distingue /r/ de /R/.
- D'un point de vue phonographémique, le graphème « r » possède deux associés phonologiques (en IL et en L1) : il peut donc y avoir compétition entre les deux correspondances.

**Exemple 3 : IL  $\supset$  /R/ + (/R/, « l ») et /l/ + (/l/, « r ») (système erroné)**

L'IL inclut /R/ + (/R/, « l ») et /l/ + (/l/, « r ») sans préjuger des autres modifications apportées au système initial.

- D'un point de vue phonologique, le sujet distingue /r/, /l/ et /R/.
- D'un point de vue phonographémique, en raison de la similarité phonético-phonologique entre /r/ et /l/, le sujet associe /l/ à « r » et, par distinction, associe /R/ à « l ».

---

<sup>196</sup> En réalité, en L1, chacun de ces segments (phonétique, phonologique ou graphémique) constitue « l'attaque d'une more », et est donc foncièrement associé à un segment vocalique (A, I, U, E, O).

**Exemple 4 : IL = /R/ + (/R/, « r ») et /l/ + (/l/, « l ») (système complet)**

Dans ce cas, l'apprentissage tant phonologique que phonographémique est achevé, mais il peut y avoir compétition entre les systèmes (distincts) L1 et IL en tâche de reconnaissance notamment.

Comme cela apparaît donc ici, le croisement des trois variables impliquées (phonétique, phonologique, phonographémique) avec la distinction entre L1 et IL rend assez complexe l'approche des phénomènes en question, mais semble nécessaire afin de ne pas préjuger de la nature des « déficiences » constatées.

**I.2. Interprétation linéaire des résultats**

L'approche que nous adoptons ici s'accorde bien avec les principes sous-tendant le modèle de Massaro, dans le sens où le sujet est conduit à prendre une « décision » en vue de choisir l'une des transcriptions proposées. Trois types d'informations essentielles sont « évalués », puis « intégrés » pour prendre cette décision : information perceptive phonétique (acoustique/articulatoire), information phonologique et information phonographémique. Cette évaluation étant soumise à un ensemble de contraintes, dans une certaine mesure propres à chaque domaine, cette approche s'accorde également avec le cadre de la Théorie de l'Optimalité : cadre « classique » pour le domaine phonologique (Archangeli, 1997) et cadre probabiliste pour le domaine perceptif (Escudero et Boersma, 2004). Pour le domaine graphémique (évidemment lié aux deux premiers), on pourrait ainsi envisager d'adopter une approche similaire, quoique son autonomie vis-à-vis du domaine phonologique puisse être sérieusement mise en doute.

**I.2.1. La position**

D'après notre analyse de variance univariée, la position semble avoir une influence significative ( $p < 0,001$ ) sur le taux de confusion phonographémique. D'après nos résultats descriptifs, cette influence semble varier selon la nature du segment considéré, mais aucune interaction significative entre position et nature n'a été mise à jour

statistiquement. Rappelons le classement global obtenu, par ordre décroissant de confusion :

Coda Finale et Attaque Finale (31%) > Attaque Initiale (26%)

De manière générale, comme le rappelle Uffmann (2003) dans le cadre de la théorie de l'Optimalité, les positions proéminentes que sont les premières syllabes, les syllabes accentuées et les attaques syllabiques, de par leur saillance perceptive/psycholinguistique, facilitent la reconnaissance et sont ainsi, dans un cadre phonologique, particulièrement « protégées » par les contraintes de Fidélité. On parle alors de Fidélité de Position (« Positional Faithfulness »). Dans le domaine psycholinguistique, la première syllabe du stimulus est généralement « protégée » (Nespoulous, 2001), et ce « privilège » s'observe dans nos résultats globaux, puisque le taux global de confusion est moindre en attaque initiale (26% contre 31% pour les deux autres positions). La position initiale favorise donc globalement l'identification du segment, principalement en raison de la saillance perceptive qu'offre une telle position.

Afin de rendre compte de cette saillance, on pourrait évoquer trois niveaux d'analyse :

- A un niveau phonétique, le caractère post-pausal de l'attaque initiale (pas d'assimilation progressive due à la coarticulation d'un segment antérieur) qui la distingue de l'attaque finale et de la coda finale :

Attaque Initiale vs. Attaque Finale et Coda Finale

- A un niveau mnésique, le caractère initial de l'information en attaque initiale (pas d'information préalable pouvant entraver l'encodage en mémoire de travail) qui la distingue de l'attaque finale (mais aussi de la coda finale, quoique celle-ci puisse bénéficier de « l'effet de récence » en raison de sa position pré-pausale) :

Attaque Initiale vs. Attaque Finale

- A un niveau phonologique, le caractère moins marqué de l'attaque (qui la distingue de la coda) :

#### Attaque Initiale vs. Coda Finale

Ces trois arguments justifient ainsi *très globalement* la primauté de la position d'Attaque Initiale en termes de saillance perceptive. Cependant, si l'on souhaite analyser plus finement les raisons et les variations de cette saillance, il faut examiner l'influence de la position selon la nature des segments consonantiques. On s'aperçoit alors que cette tendance générale varie en fonction des segments considérés.

#### **I.2.2. La nature de la consonne**

D'après notre analyse de variance univariée, la nature du segment consonantique semble avoir une influence significative ( $p < 0,001$ ) sur le taux de confusion phonographémique, comme cela apparaît dans nos résultats descriptifs. Rappelons les résultats obtenus :

Dans environ 45% des cas : • /b/ est transcrit « v »

Dans environ 24% des cas :

- /v/ est transcrit « b » (25,4%)
- /l/ est transcrit « r » (23,7%)
- /R/ est transcrit « l » (23,5%)

De tels résultats pourraient alors être rattachés à l'hypothèse générale, selon laquelle les « nouveaux sons » sont plus faciles à acquérir que les « sons similaires » à ceux existant en L1 (Flege et Hillenbrand, 1987 ; Escudero et Boersma, 2004) : en effet alors que la catégorie phonologique /b/ existe en L1 (sous une forme moraique cependant), les catégories phonologiques /v/, /l/ et /R/ sont des catégories « nouvelles ». Les résultats s'expliquent alors de la manière suivante : si l'identification de /b/ est



« incertaine », cette incertitude conduit le sujet à privilégier, par défaut la catégorie opposée (en l'occurrence /v/) ; en revanche, si l'identification de /v/, /l/ ou /R/ est incertaine, il y a autant de raison de privilégier la catégorie « incertaine » que son opposée. En effet : alors que l'espace phonétique associé à la catégorie phonologique /b/ est clairement défini (de par son existence en L1), l'espace phonétique associé aux catégories phonologiques « nouvelles » l'est beaucoup moins (puisque celles-ci sont « nouvelles »). On pourrait ainsi considérer que, pour /b/, les sujets disposent d'un prototype, qui constitue le noyau de l'espace phonétique et par rapport auquel sont évalués les stimuli, tandis que pour les catégories « nouvelles », les apprenants ne disposent pas d'un tel prototype. Un même degré d'incertitude perceptive conduirait ainsi plus facilement au rejet de la catégorie phonologique /b/ qu'au rejet de catégories « inconnues ». Cela explique pourquoi le taux de confusion est plus élevé pour le phonème /b/ (rejet de /b/ et donc choix de « v » *par défaut*) que pour les autres phonèmes.

On pourrait cependant objecter que, à un niveau phonétique, contrairement à [v] et [R], la catégorie phonétique [l] n'est pas « nouvelle » (puisque le phone [l] est un allophone du /r/ japonais), et l'on pourrait ainsi distinguer deux groupes : [b] et [l] d'une part (catégories phonétiques « connues ») et [v] et [R] d'autre part (catégories phonétiques « inconnues »). D'un point de vue théorique, on pourrait ainsi s'attendre à davantage de confusion pour les stimuli auditifs de type [l] que ceux de type [R]. Or nos résultats ne correspondent pas à ces prédictions théoriques, puisque les taux de confusion pour [l] et [R] sont similaires. La différence de taux entre [b] et [l] s'explique alors de la manière suivante : à un niveau perceptif auditif, [l] et [R] seraient moins similaires que ne le sont [b] et [v]. L'incertitude perceptive serait donc fondamentalement moindre pour [l] (vis-à-vis de [R]) que pour [b] (vis-à-vis de [v]), ce qui explique qu'il y ait davantage de confusions pour [b] que pour [l].

Deux axes devraient donc être pris en compte : le caractère « connu » ou non des catégories phonologiques (/b/ vs. /v, l, R/) d'une part, et la similarité perceptive

auditive des catégories phonétiques ([b, v] vs. [l, R]). Ces deux arguments permettent ainsi de rendre compte des résultats obtenus.

### **I.2.3. La nature de la voyelle adjacente**

D'après notre analyse de variance univariée, la nature du segment vocalique adjacent semble avoir une influence significative ( $p < 0,001$ ) sur le taux de confusion phonographémique, comme cela apparaît dans nos résultats descriptifs. De plus, l'analyse de variance bivariée met à jour une interaction statistiquement significative entre la nature de la consonne et celle de la voyelle sur le taux de confusion. Cette interaction est particulièrement notable au vu de la structure moraïque du japonais : comme nous l'avons souligné précédemment, bien que nous analysions nos données en termes phonémiques, il est particulièrement probable que les unités effectivement traitées par les sujets soient moraïques (/ba/, /bi/, etc.) plutôt que phonémiques, pour les raisons présentées plus haut. L'influence de la structure moraïque et ainsi des propriétés phonotactiques de la L1 pourrait être particulièrement perceptible à travers cette interaction. Rappelons alors les résultats obtenus, par ordre décroissant de confusion :

OU (36%) > EU, A, I (27%)

La voyelle /OU/ semble donc entraîner davantage de confusion que les trois autres. Il faut alors rappeler que /OU/, qui est la seule voyelle postérieure (vélaire) des quatre, est également la seule voyelle arrondie, ce qui, d'après Lombardi (2002), en fait une voyelle particulièrement marquée<sup>197</sup>. Étant donné le phénomène d'anticipation vocalique en français, il se pourrait que l'arrondissement (labialisation) provoqué par OU augmente le risque de confusion au sein des couples consonantiques. De manière générale, la voyelle /OU/ serait donc plus « problématique » que les trois autres. Cependant, si l'on souhaite analyser plus finement ces résultats, au vu de l'interaction

---

<sup>197</sup> On doit également signaler que les voyelles arrondies, en particulier /u/ en anglais, sont celles qui masquent le plus l'information visuelle consonantique (Hardison, 1999, 2003).

entre consonne et voyelle, il faut examiner l'influence de la voyelle pour chaque segment consonantique considéré :

BOU (50%) > BA (46%) > BEU, BI (41%)

VOU (31%) > VA (26%) > VEU, VI (22%)

Dans le cas du couple /b, v/, le classement obtenu est identique. Trois facteurs peuvent alors être invoqués :

- 1) Le lieu d'articulation (antérieure = palatale/ postérieure = vélaire), que l'on peut, dans une certaine mesure, également lier à la tension<sup>198</sup> (tendu/relâché) ou encore à l'avancement de la racine de la langue (+ARL = avancée/ --ARL = rétractée) ;
- 2) L'aperture (fermé/ouvert) ;
- 3) La labialité (écarté/arrondi).

A un niveau articulatoire, il faut en outre noter que [b] et [v] sont toutes deux des consonnes labiales, ne faisant pas intervenir la langue dans leur articulation. Leur *lieu* d'articulation est donc très similaire, alors que leur *mode* d'articulation est différent (plosive vs. fricative) : dans le cadre phonétique décrit par Renard (1979) notamment, les plosives sont dites plus « tendues » que les fricatives. On pourrait donc penser qu'il n'y a pas d'interaction entre le *mode* d'articulation et la voyelle, à moins de regrouper ces deux phonèmes dans la famille des obstruantes.

Si l'on se penche sur le cas du couple /l, R/ cependant, la situation est quelque peu différente :

ROU, RI (27, 5%) > RA, REU (19, 5%)

LOU (36%) > LEU (29%) > LA (16%) > LI (13, 5%)

---

<sup>198</sup> Cependant, voir Durand (2005) pour une discussion à propos de ce trait.

Contrairement à [b] et [v], qui partagent, dans une certaine mesure, le lieu d'articulation (labial) mais s'opposent sur le mode (plosive vs. fricative), [R] et [l] partagent, dans une certaine mesure ([R] est une vibrante alors que [l] est une latérale), le même *mode* (fricative), mais diffèrent sur le *lieu* d'articulation ([l] = apico-alvéolaire (antérieure) vs. [R] = dorso-uvulaire (postérieure)), impliquant toutes deux la langue de manière différente (avancée vs. rétractée). Les classements obtenus, pour /b/ et /v/ d'une part et /l/ et /R/ d'autre part, semblent alors confirmer l'importance du *lieu* d'articulation dans l'interaction avec la voyelle.

Les résultats peuvent dès lors s'expliquer comme suit :

- Pour /b/, il semblerait possible de rendre compte des résultats observés via le trait « ARL » que nous lions au « timbre » vocalique : plus la langue est avant, plus la cavité buccale est réduite, plus le timbre de la voyelle est « clair » (Renard, 1979). Le classement des voyelles correspond alors à celui que nous avons obtenu :

<b>+ ARL (Timbre « Clair »)</b>	<b>--ARL (Timbre « Sombre »)</b>
I – EU – A – OU	

On pourrait alors penser que l'avancement de la langue (+ARL) favorise la tension articuloire de la consonne précédente et minimise ainsi la confusion de [b] en [v]. Si une telle interprétation permet de rendre compte des résultats obtenus pour [b], ceux obtenus pour [v] ne s'y accordent pas, puisque le classement devrait alors être inversé : VOU devrait entraîner moins de confusion que VI, alors que BOU devrait entraîner davantage de confusion que BI.

Si l'on considère cependant que, parmi les trois voyelles en question, seule OU est une voyelle postérieure (--ARL), alors il semble possible d'établir une première distinction sur cette base : OU (--ARL) vs. A, EU, I (+ARL). Puis, il est possible d'effectuer une nouvelle division sur la base de l'aperture : A (--Haute = ouverte) vs. EU, I (+Haute = fermée). On obtient alors le classement obtenu.

- En ce qui concerne le couple R/L, le trait « ARL » semble ici encore pertinent, puisque la voyelle OU est celle qui déclenche le plus de confusion. Cependant, pour [R], l'aperture permet de distinguer OU et I d'une part, et A et EU d'autre part : en effet, une aperture davantage ouverte (A et EU) conduit le phonème /R/ à se réaliser sous forme d'une approximante, qui se distingue davantage du [l]. Plus l'aperture serait ouverte, moins les confusions seraient nombreuses. En ce qui concerne [l], il faut se rappeler que les syllabes LA et LI existent, dans une certaine mesure, en japonais. Les deux combinaisons problématiques sont celles du OU et du EU, puisque ces deux voyelles chevauchent en partie l'espace acoustique de la voyelle japonaise [ɯ], à laquelle elles sont souvent ramenées, en production, par les apprenants japonais. Au sein de ces deux sous-groupes, le trait « ARL » (que l'on peut interpréter en termes de tension) permet d'obtenir le classement observé (OU vs. EU et A vs. I). Cela rejoint le fait que la position de la langue pour les voyelles OU et A est plus proche de celle du [R] que de celle du [l], d'où le taux plus élevé de confusions.

Enfin, le fait que les classements soient identiques pour /b/ et /v/, mais différents pour /R/ et /l/ soutient nos remarques précédentes concernant le traitement différent accordé à ces deux couples. B et V semblent en effet être traités de manière similaire, ce qui peut s'expliquer d'une part par leur similarité auditive (notamment *via* l'allophone [β] de /b/ en japonais, qui est une fricative bilabiale), et d'autre part par leur statut en japonais : /v/ appartient davantage au lexique japonais que le /R/, étant donné : a) son existence en anglais, bien plus répandu que le français, qui conduit à b) son émergence en tant que « nouvelle consonne » transcrite en katakana (contrairement au /R/ français). [l] et [R] sont en revanche beaucoup moins similaires auditivement et [R] est bien plus « étranger » (en termes de strates dans le lexique phonologique japonais) que ne l'est le [v] en japonais.

#### **I.2.4. Choix de graphèmes**

Parmi les 120 sujets ayant passé le test, 41 sujets ont été choisis au hasard, en vue d'examiner le type de graphème (existant ou absent en L1) préférentiellement choisi, sans s'occuper de leur valeur phonologique. Rappelons les résultats obtenus :

<b>Graphème consonantique</b>		<b>Graphème vocalique</b>	
<b>Existant en L1 (&lt;B&gt; ou &lt;R&gt;)</b>	<b>Absent en L1 (&lt;V&gt; ou &lt;L&gt;)</b>	<b>Existant en L1 (&lt;OU&gt;)</b>	<b>Absent en L1 (&lt;EU&gt;)</b>
45%	55%	54%	46%

Ainsi, alors que pour les consonnes, le graphème absent en L1 a été préférentiellement choisi, c'est le contraire qui apparaît pour les voyelles.

Pour les consonnes deux possibilités sont alors envisageables :

- A un niveau perceptif, globalement, les sujets auraient davantage perçu /v/ et /l/ que /b/ et /R/.
- A un niveau stratégique, globalement, les sujets auraient davantage choisi les graphèmes « v » et « l », en tant que graphèmes « par défaut », selon le raisonnement présenté plus haut, pour des sujets en situation d'incertitude perceptive. Il faut en effet souligner que ces deux graphèmes sont les graphèmes « inconnus » pour les sujets.

En ce qui concerne les voyelles, le résultat obtenu peut s'expliquer en examinant le protocole : il faut en effet se souvenir que la tâche était plus complexe pour les voyelles (quatre réponses possibles) que pour les consonnes (deux réponses possibles). De plus, l'attention des sujets était essentiellement orientée sur la perception et le choix des consonnes : on peut alors imaginer que le choix des graphèmes vocaliques (et la perception des voyelles associées) était beaucoup plus *automatique* que celui des consonnes (davantage sous contrôle attentionnel). Il n'est alors pas étonnant que les sujets aient davantage choisi le graphème plus familier « OU » et/ou aient davantage

perçu la voyelle OU (identification de EU en OU, classique en didactique du FLE pour des apprenants japonais qui écrivent régulièrement sous dictée le modèle « ils veulent » en « ils veulent »<sup>199</sup>).

## **II. Tâche F : discrimination phonologique et appariement segmental phonème - graphème (L/R : B/V)**

### **II.1. Introduction**

La tâche F constitue en quelque sorte une extension de la tâche A : alors que la tâche A consistait à identifier un phonème en choisissant une transcription orthographique, la tâche F consiste à discriminer deux phonèmes en choisissant une transcription orthographique pour chaque stimulus. Deux opérations sont alors évaluées :

- 1) Comme pour la tâche A : l'appariement entre le premier stimulus (phonème) et une transcription parmi deux (graphèmes). Nous évaluons alors le taux de confusion *apparemment* phonographémique.
- 2) La discrimination entre les deux stimuli (phonèmes), via les transcriptions choisies (graphèmes), identiques ou différentes. Nous évaluons alors le taux de confusion *apparemment* phonologique.

Sans reprendre en détail l'analyse de la tâche A, qui est applicable à la tâche F, il convient néanmoins de souligner certaines spécificités de cette dernière. Dans la tâche F, deux grandes stratégies de réponse sont envisageables :

- 1) Stratégie phonético-phonologique : le sujet évalue les deux représentations phonético-phonologiques (discrimination phonético-phonologique), puis choisit une transcription (appariement) pour la représentation phonémique dont il est le

---

<sup>199</sup> Le facteur morphologique y joue également un rôle non-négligeable (racine du verbe : « vou- »).

plus certain, puis choisit l'autre transcription *par défaut*<sup>200</sup>. La stratégie est donc en premier lieu phonético-phonologique.

- 2) Stratégie phonographémique : le sujet associe le premier stimulus à une transcription (appariement), puis évalue le deuxième stimulus auditif par rapport à cette transcription (discrimination phonème/graphème).

Par ailleurs, alors que la tâche A était une tâche d'identification sous-tendue par une tâche de discrimination (choix entre deux segments en opposition dans le protocole), la tâche F est, en revanche, une authentique tâche de discrimination (les deux stimuli doivent être qualifiés d'identiques ou de différents).

Ceci étant précisé, nous devons observer que les résultats globaux sont moindres que ce à quoi nous nous attendions : 21% de confusions phonographémiques et 27% de confusions phonologiques. Ces résultats pourraient soutenir l'idée que les sujets maîtrisent relativement mieux le système phonographémique que le système phonologique français, même si les taux de confusion pourraient être dus à des « erreurs » perceptives auditives et non phonologiques.

## **II.2. Interprétation linéaire des résultats phonographémiques (premier stimulus)**

### **II.2.1. La position**

D'après notre analyse de variance univariée, la position semble avoir une influence significative ( $p < 0,001$ ) sur le taux de confusion phonographémique. D'après nos résultats descriptifs, cette influence semble varier selon la nature du segment considéré, mais aucune interaction significative entre position et nature n'a été mise à jour statistiquement. Rappelons le classement global obtenu, par ordre décroissant de confusion :

---

<sup>200</sup> Le sujet peut également choisir une transcription pour chaque représentation, mais cette stratégie semble plus « coûteuse ».



### Coda Finale et Attaque Finale (22%) > Attaque Initiale (19%)

Le résultat obtenu est identique à celui de la tâche A) et l'on peut donc reprendre la même interprétation selon laquelle la position initiale favorise la préservation de l'input pour des raisons phonétique, mnésique et phonologique. Il faut cependant noter que la différence de taux entre les trois positions est très faible, et que l'influence de la position, malgré l'absence d'interaction statistique entre position et nature de la consonne, semble varier selon le segment considéré.

#### **II.2.2. La nature de la consonne**

D'après notre analyse de variance univariée, la nature du segment consonantique semble avoir une influence significative ( $p < 0,001$ ) sur le taux de confusion phonographémique, comme cela apparaît dans nos résultats descriptifs. Rappelons les résultats obtenus :

Dans environ 29% des cas :

- /b/ est transcrit « v »

- /v/ est transcrit « b »

Dans environ 16% des cas :

- /R/ est transcrit « l »

Dans environ 10% des cas :

- /l/ est transcrit « r »

On doit dans un premier temps souligner que les résultats sont bien plus distincts, pour chaque segment, que ceux de la tâche A, dans laquelle nous avons constaté une opposition apparemment simple entre catégories phonologiques « connues » (/b/) et « inconnues » (/v, l, R/). La tâche de discrimination F, en offrant un exemplaire phonétique des deux catégories en opposition, semble ainsi permettre de mieux mettre à jour les difficultés perceptives posées aux sujets par les segments considérés. Ces résultats confirment en outre le fait que la distinction [b]/[v] est moins évidente à accomplir que celle entre [R] et [l]. On pourrait néanmoins s'étonner de l'écart entre [b] et [l] qui sont toutes deux des catégories phonétiques « connues » : un tel écart

souligne l'importance du rapport de similarité perceptive et/ou phonologique entre les deux segments en opposition pour la réalisation de la tâche, mais doit également être rattaché à l'existence de l'allophone [β] en japonais, qui pourrait augmenter le risque de confusion entre [b] et [v].

- En ce qui concerne le couple (/b, v/), l'argument suivi dans la tâche A) peut être repris ici, en notant alors que la tâche de discrimination, en fournissant un exemplaire phonétique de chaque catégorie au sujet, permet de diminuer l'incertitude perceptive par rapport à la tâche d'identification, plaçant [b] et [v] dans une situation analogue à celle de [l] et [R] dans la tâche A) : cela expliquerait pourquoi /b/ et /v/ conduisent ici à des taux similaires, inférieurs à celui de /b/ dans la tâche A.
- En ce qui concerne le couple (/l/, /R/), le raisonnement est identique : tout en suivant l'argument présenté pour la tâche A), la tâche de discrimination permet de réduire l'incertitude perceptive, réduisant ainsi les taux de confusion pour les deux catégories. Cependant, le taux particulièrement faible obtenu pour /l/ est ici encore contraire à nos prévisions théoriques, lesquelles envisageaient soit des taux équivalents (*cf.* tâche A)), soit des taux de confusion supérieurs pour les catégories « connues » ([b] et [l]). Une possibilité d'explication serait alors la suivante : dans une tâche de discrimination comme la tâche F), à condition que la similarité perceptive ne soit pas trop élevée (comme dans le cas du couple [b]/[v]), le caractère « connu » de la catégorie phonétique faciliterait son identification par rapport à des catégories phonétiques « inconnues ».

Deux types de similarité perceptive devraient donc être pris en compte : similarité interlinguistique entre catégories de la L1 et catégories de la L2 (et qui peut être renforcée par l'existence d'allophones particuliers en L1) et similarité intralinguistique entre catégories de la L2. Leur influence varierait alors en fonction de la tâche (identification ou discrimination).

### **II.2.3. La nature de la voyelle adjacente**

D'après notre analyse de variance univariée, la nature du segment vocalique adjacent semble avoir une influence significative ( $p < 0,001$ ) sur le taux de confusion phonographémique, comme cela apparaît dans nos résultats descriptifs. De plus, l'analyse de variance bivariée met à jour une interaction statistiquement significative entre la nature de la consonne et celle de la voyelle sur le taux de confusion. Rappelons alors les résultats obtenus, par ordre décroissant de confusion :

OU, EU (22%) > A, I (19%)

Comme dans la tâche A) la voyelle /OU/ entraîne davantage de confusion que les voyelles /A/ et /I/. Cependant, dans la tâche F), aucune différence n'apparaît entre les voyelles /OU/ et /EU/. On peut alors distinguer ici, d'une part les voyelles phonétiquement quasi-identiques en L1 et L2 (/A/ et /I/) qui entraînent moins de confusions, et d'autre part les voyelles /OU/ et /EU/ qui présentent un cas de transfert négatif divergent : la voyelle japonaise [ɯ] se situe en effet sur un continuum aux extrémités duquel on pourrait placer [u] et [ø]. Ces deux voyelles françaises sont ainsi souvent confondues par les apprenants japonais, qui les identifient toutes deux à la voyelle japonaise [ɯ]. Dans le cadre TCRS, on pourrait alors considérer que la « malformation » des voyelles OU et EU augmentent le poids procédural de la « réparation » de l'input, augmentant alors le taux de confusion par rapport à A et à I.

Cependant, si l'on souhaite analyser plus finement ces résultats, au vu de l'interaction entre consonne et voyelle, il faut examiner l'influence de la voyelle pour chaque segment consonantique considéré :

BEU (35%) > BA, BOU, BI (25, 5%)

VOU (32%) > VA, VEU, VI (28%)

On remarque que le classement est presque identique à celui obtenu dans la tâche A. Cependant, l'influence de la voyelle semble moindre en tâche de discrimination (les taux sont davantage groupés), ce qui s'explique de nouveau par la présence des deux stimuli acoustiques dans la tâche F, qui conduit les sujets à concentrer leur attention sur la perception des consonnes (indices perceptifs consonantiques plus nombreux dans la tâche F que dans la tâche A). De plus, contrairement à la tâche A, la voyelle EU entraîne davantage de confusions que la voyelle OU. Cette différence souligne une fois de plus les difficultés posées aux sujets par cette situation de transfert négatif divergent entre la L1 (/u/) et la L2 (/ø/ vs. /y/ vs. /u/).

Si l'on se penche sur le cas du couple /l, R/, les résultats sont apparemment assez différents :

RI (24%) > ROU (20%) > RA (12%) > REU (8%)  
LEU (15%) > LOU (10%) > LA, LI (7, 5%)

Contrairement aux résultats du couple B/V, l'influence de la voyelle semble plus forte dans la tâche F que dans la tâche A (taux distincts). Cela souligne de nouveau la différence de traitement accordé aux deux couples de segments, et pourrait s'expliquer par le fait que [b] et [v] sont perceptivement plus similaires que [l] et [R], lesquels varient davantage, dans leur réalisation phonétique, selon l'environnement vocalique (en particulier pour le [R]). Les classements obtenus restent néanmoins similaires à ceux obtenus dans la tâche A, en dépit de deux différences, qui sont les suivantes :

- L'inversion de classement entre LOU et LEU, qui rejoint ce que nous avons affirmé dans le cas du couple B/V.
- L'écart entre RI et ROU, qui nécessiterait, pour pouvoir être expliqué, de données supplémentaires de type acoustique ou phonotactique, puisque la fréquence d'occurrence de ces syllabes en japonais pourrait jouer un rôle non négligeable dans ces résultats.

#### **II.2.4. La nature du couple de stimuli (identiques/différents)**

Globalement, le taux de confusion phonographémique est supérieur de 4% (23% - 19%) lorsque les stimuli sont identiques. Cela rejoindrait la différence de taux de confusion observée entre la tâche A) et la tâche F) : de la même manière que l'identification ne fournit aucun indice acoustique de distinction, conduisant à une plus grande incertitude perceptive, l'identité des stimuli dans la tâche F) conduirait également à une plus grande incertitude perceptive et ainsi à un taux de confusion plus élevé.

### **II.3. Interprétation linéaire des résultats phonologiques**

#### **II.3.1. La position**

D'après notre analyse de variance univariée, la position semble avoir une influence significative ( $p < 0, 001$ ) sur le taux de confusion phonologique. D'après nos résultats descriptifs, cette influence semble varier selon la nature du segment considéré, mais aucune interaction significative entre position et nature n'a été mise à jour statistiquement. Rappelons le classement global obtenu, par ordre décroissant de confusion :

Coda Finale et Attaque Finale (27, 5%) > Attaque Initiale (25%)

Ici encore, la position initiale présente globalement le taux de confusion le plus faible : on peut donc reprendre la même interprétation selon laquelle la position initiale favorise la préservation de l'input pour des raisons phonétique, mnésique et phonologique. Il faut cependant noter que la différence de taux entre les trois positions est très faible, et que l'influence de la position, malgré l'absence d'interaction statistique entre position et nature de la consonne, semble varier selon le segment considéré.

#### **II.3.2. La nature des consonnes**

D'après notre analyse de variance univariée, la nature du segment consonantique semble avoir une influence significative ( $p < 0, 001$ ) sur le taux de confusion

phonologique, comme cela apparaît dans nos résultats descriptifs. Rappelons les résultats obtenus :

Confusions /b, v/ (37%) > Confusions /l, R/ (16%)

Ces résultats confirment notre hypothèse précédente, selon laquelle [b] et [v] sont perceptivement plus similaires que ne le sont [l] et [R], notamment en raison de l'existence d'une bilabiale fricative allophone du /b/ en japonais, mais également en raison de la similarité du lieu d'articulation de [b] et [v] (labial).

### **II.3.3. La nature de la voyelle adjacente**

D'après notre analyse de variance univariée, la nature du segment vocalique adjacent semble avoir une influence significative ( $p < 0, 001$ ) sur le taux de confusion phonologique, comme cela apparaît dans nos résultats descriptifs. De plus, l'analyse de variance bivariée met à jour une interaction statistiquement significative entre la nature de la consonne et celle de la voyelle sur le taux de confusion. Rappelons alors les résultats obtenus, par ordre décroissant de confusion :

EU (28%) > OU (27%) > A, I (26%)

Les résultats descriptifs généraux ne permettent cependant pas d'apprécier l'influence de la voyelle en raison du faible écart qui sépare les taux. On note néanmoins ici encore la dominance des voyelles /EU/ et /OU/, qui, d'après le raisonnement présenté précédemment, augmente l'incertitude perceptive pour le sujet. Cependant, cette influence varie selon les segments considérés :

BEU, BA (43%) > BOU (33%) > BI (22%)

VOU, VA (40%) > VEU (37%) > VI (35%)

En ce qui concerne les segments /b/ et /v/, il faut d'abord noter que, comme pour le taux de confusion phonographique, la voyelle semble davantage influencer le taux

de confusion pour le segment /b/ (10% d'écart entre les trois groupes EU/A, OU et I), alors que les taux sont compris entre 40% et 35% pour le segment /v/. De plus, les classements sont relativement identiques à ceux du taux de confusion phonographémique, même si les écarts sont beaucoup plus marqués ici. Dans les trois cas (tâches A (1) et F (2)), on retrouve le même type d'opposition : A déclenche plus de confusions que I, tandis que les voyelles EU et OU semblent alterner en tête de classement pour /b/, avec OU toujours en tête de classement pour /v/. Nous pouvons donc reprendre en partie au moins l'explication proposée dans la tâche A, même si des données supplémentaires (en particulier la fréquence d'occurrence) seraient nécessaires pour bien rendre compte de ces résultats.

Les résultats relatifs à /R/ et /l/ sont particulièrement intéressants en ce qu'ils conduisent au même classement vocalique : I > OU, EU > A :

$$\begin{aligned} & \text{RI (24, 4\%)} > \text{ROU (15\%)} > \text{REU (13\%), RA (12\%)} \\ & \text{LI (20, 7\%)} > \text{LOU (17,5\%)} > \text{LEU (17,8\%)} > \text{LA (10\%)} \end{aligned}$$

L'équivalence des taux de OU et EU a déjà été commentée précédemment, et c'est donc la distinction entre I et A qui semble ici remarquable, d'autant que le classement pour /R/ et /l/ (I > A) est opposé à celui de /b/ et /v/ (A > I). Il semblerait donc que l'aperture de la voyelle influence différemment la perception des liquides (mieux perçues lorsque la voyelle est basse) et des labiales (mieux perçues lorsque la voyelle est haute). Cependant, ici encore, d'autres données (acoustique et phonotactiques) seraient nécessaires pour tenter d'expliquer les différences constatées entre les trois classements relatifs à l'influence de la voyelle adjacente.

#### **II.3.4. La nature du couple de stimuli (identiques/différents)**

Globalement, le taux de confusion phonologique est nettement supérieur (+10%) lorsque les stimuli sont différents. Ce résultat ne s'accorde pas avec l'idée formée jusqu'à présent selon laquelle la tâche de discrimination avec des stimuli différents serait plus facile à accomplir que celle d'identification ou celle de discrimination avec

stimuli identiques. On doit alors rappeler que, dans la tâche F, la tâche de discrimination phonologique était effectuée via une tâche d'appariement phonographémique. Il n'est donc pas impossible que le choix du graphème correct pour chacun des deux stimuli ait été facilité lorsque les stimuli étaient identiques : dans de telles conditions, le choix est en effet réduit à deux (choix d'un graphème correct parmi deux), alors qu'il est de quatre lorsque les stimuli sont différents (quatre combinaisons possibles). Cela expliquerait le taux plus important de confusions lorsque les stimuli sont différents.

### **III. Tâches C1 et C2 : comptage syllabique en unimodalité**

#### **auditive**

Rappelons que les tâches C1 et C2 ne différaient que par la nature des groupes CC cibles, ainsi que par la population des sujets testés. Dans chaque test, deux phénomènes étaient évalués : les épenthèses vocaliques et les effacements *a priori* vocaliques. Notre attention était cependant essentiellement portée sur les épenthèses, puisque, dans une perspective didactique, le problème rencontré par les apprenants japonais est celui du traitement des groupes CC, généralement résolu à l'aide de telles épenthèses. Nous traitons donc celles-ci en premier lieu.

La tâche de comptage syllabique, telle qu'implémentée dans notre protocole, visait à évaluer si la structure syllabique des items présentés était perçue et respectée par les sujets, et si certaines variables influençaient cette perception, en particulier la position et la nature du groupe CC. Cette « nature » pouvait être analysée de plusieurs manières : famille générale, famille, sous-famille ou nature spécifique phonologique ; profil de sonorité ou écart de sonorité.

Si l'on considère que, à l'état initial d'apprentissage chez nos sujets, IL=L1 (transfert plein), alors les résultats obtenus peuvent être abordés de la manière suivante :



- Absence d'épenthèse :
- Perception « erronée » du groupe CC : seule une consonne parmi les deux a été « perçue » ([CC] → [C]).
  - Représentation « erronée » du groupe CC : le groupe CC a été « perçu », mais la représentation phonologique est obtenue via un effacement consonantique ([CC] → /C/).
  - Apprentissage de la syllabation du français : la structure a été correctement perçue et traitée ([CC] → /CC/).

- Epenthèse :
- Représentation « erronée » du groupe CC : le groupe CC a été « perçu », mais la représentation phonologique est obtenue via l'épenthèse vocalique et/ou l'adaptation segmentale ([CC] → /CuC/).

### **III.1. Tâche C1 - Epenthèses**

Les groupes testés étaient les suivants : /pn, ps, pf, pt, kn, km, kv, ks, kt, tl, tm, ts, tsh/ (Plosive sourde + C) ainsi que /ft/ en position finale. Conformément à nos prévisions, le taux global d'épenthèses pour les groupes CC est élevé (64%).

#### **III.1.1. Position**

Les analyses inférentielles univariées suggèrent l'existence d'une influence significative de la position sur le taux d'épenthèses. Rappelons les résultats obtenus (par ordre décroissant d'erreurs) :

Attaque Initiale (35%) > Coda Finale (16%) > Attaque Finale (7%)

Pour le groupe /kt/, cela s'illustre par l'exemple suivant :

KTOSAGI > NALOGIKT > DAGIKTO.

Comme le rappelle Uffmann (2003) dans le cadre de la théorie de l'Optimalité, les épenthèses et les effacements varient généralement en fonction de la position. Les positions proéminentes que sont les premières syllabes, les syllabes accentuées et les attaques syllabiques, de par leur saillance perceptive/psycholinguistique, facilitent la reconnaissance et sont ainsi, dans un cadre phonologique, particulièrement « protégées » par les contraintes de Fidélité. On parle alors de Fidélité de Position (« Positional Faithfulness »). Cependant, dans les résultats obtenus ici, la position d'attaque initiale entraîne davantage d'épenthèses que les autres positions.

- D'un point de vue perceptif, ces données peuvent cependant être réconciliées avec les affirmations de Uffmann, en considérant que la position d'attaque initiale est *en effet* perceptivement plus saillante que les deux autres positions, ce qui oblige les sujets à traiter le groupe CC, alors que les deux autres positions sont plus propices à ce que nous avons nommé une perception « erronée » du groupe CC conduisant à un effacement consonantique « perceptif ». Seule une des deux consonnes est perçue, celle-ci étant notamment fonction de la sonorité des deux consonnes en présence. Cela expliquerait la différence de taux entre, d'un côté l'attaque initiale, et de l'autre les deux autres positions.
- D'un point de vue phonologique, le rapport entre Attaque initiale et Coda finale, pourrait conduire au classement suivant : \*AttaqueComplexe >> \*CodaComplexe. Mais il faut alors prendre en compte l'attaque finale. Pour cela, rappelons que le groupe CC que nous situons en position d'attaque finale pourrait, *a priori*, être syllabifié de quatre manières différentes : /CV.CV.CCV/ (canonique), /CV.CVC.CV/, /CV.CVCC.V/ ou même /CV.CVC-CCV/ (ambisyllabique) selon la nature, et ainsi la sonorité, des consonnes. Etant donné le classement précédent entre \*AttaqueComplexe et \*CodaComplexe, il serait possible de considérer que le segment est syllabifié de la manière suivante : /CV.CVC.CV/. Dans ce cas, seule la contrainte \*Coda est violée et nous parvenons au classement suivant :

\*AttaqueComplexe >> \*CodaComplexe >> \*Coda

Un tel classement correspond par ailleurs à nos observations empiriques selon lesquelles les apprenants parviennent plus facilement à traiter les codas simples (en perception et en production) que les structures complexes. Cela pourrait être dû au fait que, en japonais, les codas ne sont pas absolument inexistantes (Lombardi, 1998), même si ce qui semble être une coda à un niveau phonétique n'en est pas une à un niveau phonologique : en effet, si l'on suit l'analyse de Labrune (2001b) selon laquelle il n'y aurait que deux types de syllabe en japonais (type « plein » CV ou « déficient » C• ou •V) conduisant à l'équivalence entre more et syllabe et excluant toute coda en japonais, la coda phonétique [\_C#] correspondrait en fait à une syllabe phonologique déficiente /\_C•#. Ces « codas phonétiques » pourraient alors fournir une « base » phonétique à l'acquisition de codas phonologiques (correspondant à la disparition en diachronie (Brandão de Carvalho et Kawaguchi, 2002, p. 8) de la position « déficiente » inoccupée). Cependant, on pourrait également adopter le point de vue opposé : l'existence de telles codas « phonétiques » entraveraient l'apprentissage phonologique, en perpétuant l'effacement vocalique à un niveau *phonétique* (matériel segmental) et non *phonologique* (position). Il faudrait, pour éclaircir ce point disposer de données longitudinales. Dans tous les cas, le classement obtenu semble permettre de rendre compte de nos données.

### **III.1.2. Nature**

Les analyses inférentielles univariées suggèrent l'existence d'une influence significative de la famille et de la nature spécifique phonologique sur le taux d'épenthèses. Rappelons l'essentiel des résultats obtenus (par ordre décroissant d'erreurs) :

Famille phonologique : Obs + /Cont, Nas, Stop, Lat/ (38%) > Obs + Sib (29%)

Nature spécifique : /k/ + C (46%-38%) > /p/ + C<sup>201</sup> (38%-33%) > C + Sib (27%-22%)

---

<sup>201</sup> Non Sibilante.

- D'un point de vue perceptif, on pourrait dans un premier temps considérer que les groupes Obs + Sib sont plus difficilement perceptibles que les autres, conduisant à un effacement consonantique « perceptif », qui expliquerait le moindre taux d'épenthèse obtenu.
- D'un point de vue phonologique, il faut d'abord distinguer les groupes comportant le segment /t/ de ceux comportant /p/ et /k/ :
  - /t/ : comme le rappelle Labrune (2001a, p. 98) le segment /t/ en japonais présente en effet une grande variété allophonique et est souvent affriquée ([tʃ], [ts]) devant /i/ et /u/. Or, les trois groupes nommés « C + Sib » ci-dessus sont /tʃ/, /ts/ et /ps/. Il se pourrait donc que les groupes /tʃ/ et /ts/ soient traités comme des affriquées et ne conduisent donc pas à une épenthèse vocalique.
  - /p/ vs. /k/ : on doit en effet s'interroger sur la distinction entre /ps/ (24%) et /ks/ (43%) : /p/ et /k/ ne diffèrent en effet que par leur « point d'articulation » et sont tous deux licites en japonais. Deux arguments sont alors à considérer :
    - ◆ A un niveau phonétique articulatoire, les groupes [ps] et [ks] diffèrent notablement vis-à-vis des mouvements linguaux : alors que le segment bilabial [p] ne sollicite relativement pas la langue, le segment vélaire [k] la sollicite grandement. La transition de [p] à [s] et de [k] à [s] n'est donc pas la même d'un point de vue articulatoire, acoustique et ainsi *temporel*, puisque le groupe [ks] nécessite une modification de la position de la langue que le groupe [ps] ne nécessite pas. Cette différence pourrait ainsi conduire les sujets à percevoir

une « voyelle illusoire » (Dupoux et al., 1999) plus facilement dans le cas de [ks] que dans celui de [ps].

- ◆ D'un point de vue phonotactique, lorsque l'on examine le lexique japonais, ne serait-ce qu'à travers un dictionnaire ou le jugement de quelques locuteurs natifs, on s'aperçoit que la séquence /pus/ en initiale de mot est beaucoup moins fréquente (voire quasi-inexistante) en japonais que la séquence /kus/, celle-ci conduisant fréquemment à la séquence phonétique dévoisée [kʷs]. Cette différence de fréquence d'occurrence est par ailleurs renforcée par celle qui apparaît lorsque l'on consulte le nombre d'occurrence des caractères kana correspondant à <pu> et <ku>, tel qu'il émerge à travers un certain corpus<sup>202</sup> de la presse écrite nipponne : 2563 pour <ku> contre seulement 103 pour <pu>. Cette différence de fréquence pourrait ainsi jouer un rôle non-négligeable dans l'explication de nos résultats : alors que la séquence [ps] ne correspondrait à quasiment aucune séquence phonétique en japonais, la séquence [ks], au contraire, évoquerait bien plus fréquemment la séquence /kus/, via un dévoisement vocalique.

On peut donc noter que le classement approximatif obtenu pour la première consonne (/k/ > /p/ > /t/) apporte un complément d'information pertinent pour distinguer ces consonnes qui sont souvent traitées de manière homogène en termes phonémiques (plosives sourdes) : dans une approche en termes de traits, les oppositions entre /k/, /p/ et /t/ sont de type équipollent (c'est-à-dire que les segments ont chacun un trait distinct

---

<sup>202</sup> Ces données sont disponibles à l'adresse électronique suivante : [www7.plala.or.jp/dvorakjp/hinshutu.htm](http://www7.plala.or.jp/dvorakjp/hinshutu.htm). Le corpus est notamment issu de deux grands journaux quotidiens japonais (le *Asahi Shimbun* et le *Yomiuri Shimbun*), ainsi que d'autres types d'écrits japonais. Ces données nous ont été signalées par Laurence Labrune, que nous remercions ici très vivement.

qui les séparent). Ces trois segments se distinguent donc, non par leur geste catégoriel, mais par leur geste articulatoire (Durand, 2005) : alors que le segment /t/ présente le trait [dental] (et est un segment coronal, donc particulièrement non-marqué), le segment /p/ présente le trait [labial] et /k/ le trait [vélaire]. Le trait [vélaire], dans ce contexte, entraînerait davantage d'épenthèses que le trait [labial].

### **III.1.3. Sonorité**

Les analyses inférentielles univariées suggèrent l'existence d'une influence significative de la sonorité (profil et écart) sur le taux d'épenthèses. Rappelons les résultats obtenus (par ordre décroissant d'erreurs) :

Profil : Plateau=/kt, pt/ (40%) > Montant et Descendant (35, 5%)

Écart : Écart 3=/kv/ (45%) > Ecarts 0, 4, 5, 1 (40%-36%) > Écart 2 (31%)=/ps, pf, ks, ts, tʃ/

- En ce qui concerne le profil, d'un point de vue perceptif, on pourrait considérer que le type Plateau minimise la possibilité d'effacement consonantique chez les sujets (chaque segment étant aussi « sonore » que l'autre), expliquant ainsi le taux plus élevé d'épenthèses pour ce profil.
- En ce qui concerne les écarts, il faut noter que ce critère permet de confirmer le statut particulier des groupes Plosive + Sibilante, pour lesquels le taux d'épenthèse est moindre.

### **III.2. Tâche C1 – Effacements**

Le taux global d'effacement est très faible (8%) mais non nul, et varie selon la position et la nature de l'environnement consonantique. L'effacement vocalique peut être pertinent pour nous selon la manière dont il est envisagé :

- Effacement :
- Niveau « phonétique » : les deux syllabes /CuCV/ ont été perçues [CCV] (dévoisement de la voyelle épenthétique [u]) et le comptage syllabique est un comptage « phonétique ».
  - Niveau « phonologique » : les deux syllabes /CuCV/ ont été perçues [CCV] et ont activé une représentation /CCV/ chez les apprenants qui maîtrisent, en partie au moins, la syllabation du français.

Ce que nous avons nommé « effacement » vocalique jusqu'à présent doit en effet être divisé en deux catégories : l'une correspond à un dévoisement phonétique (ou encore à une désassociation du matériel segmental mais non structural (Labrune, 2001b)) mais non à un effacement phonologique, tandis que l'autre correspond à une perception phonétique « erronée » qui traduit néanmoins un certain apprentissage phonologique de la part des sujets. Dans notre tâche de comptage syllabique, l'effacement devrait alors correspondre à la deuxième catégorie, puisque, même si les voyelles peuvent être dévoisées en production chez nos sujets (voire à l'écrit, comme cette étudiante qui écrit « tempra » (en omettant le « u ») le mot japonais /te-N-pu-ra/ ou comme cette autre étudiante qui écrit « Atsmi » (en omettant le « u ») son propre nom de famille /a-tsumi/), elles gardent néanmoins une position squelettale : la syllabe phonétique [CCV] correspond à la syllabe phonologique /C•CV/. Si l'on aborde la question en termes de squelette et position vide, on peut en effet bien appréhender le phénomène de dévoisement en japonais et le décalage entre syllabe phonétique et phonologique chez les sujets (position vide). On peut dès lors considérer que les « effacements » observés sont bel et bien des effacements, témoignant de la capacité de sujets à activer une représentation phonologique de type /CCV/.

### **III.2.1. Position**

Les analyses inférentielles univariées suggèrent l'existence d'une influence significative de la position sur le taux d'effacement. Rappelons les résultats obtenus (par ordre décroissant d'erreurs) :

Coda Finale (16%) > Attaque Initiale et Finale (6%)

Pour le groupe /ps/, cela s'illustre par l'exemple suivant :

NALOGIPUSU > PUSSETAGI, DAGIPUSE.

Deux arguments permettent alors de rendre compte de ces résultats :

- D'un point de vue perceptif, les positions d'attaque favorisent une certaine saillance perceptive qui conduit les sujets à préserver l'input (pas d'effacement), contrairement à la position de coda<sup>203</sup>.
- D'un point de vue phonético-phonologique, le dévoisement des voyelles (épenthétiques) en japonais est particulièrement notable en position de coda (Vance, 1987).

Ces deux arguments permettent d'expliquer le taux supérieur d'effacement en position de coda.

### **III.2.2. Nature**

Bien que les analyses inférentielles univariées ne suggèrent l'existence que d'une influence significative de la famille, et non de la nature spécifique phonologique, sur le taux d'effacements, celui-ci semble varier selon l'environnement consonantique spécifique.

---

<sup>203</sup> Cet argument ne s'oppose pas à celui utilisé pour les épenthèses : l'attaque initiale conduirait les sujets à davantage préserver l'input (en termes d'effacement), mais pas à le respecter (en termes d'épenthèse).



Rappelons l'essentiel des résultats obtenus (par ordre décroissant d'erreurs) :

Famille                      Obs + /Sib, Stop, Lat/ (10, 5%) > Obs + Nas (7%) > Obs + Cont (4%)  
phonologique :

Bien qu'aucune régularité remarquable ne puisse apparaître pour les natures spécifiques, il faut néanmoins souligner que les environnements /p\_s/ et /t\_s/ conduisent aux taux les plus élevés d'effacement. D'un point de vue perceptif, cela rejoint notre affirmation précédente concernant les épenthèses, selon laquelle la distinction perceptive entre des groupes Obs + Sib et Obs + Nas, par exemple, [pusa] et [psa] d'une part, et [kuna] et [kna] d'autre part n'est pas équivalente : la perception d'une position (phonologique) et/ou d'un segment (phonétique) est moins évidente lorsque le second segment consonantique est une sibilante, en particulier /s/. D'un point de vue phonotactique, ces taux sont en outre à relier à l'existence de l'affriquée [ts] et de la quasi-inexistence de la séquence [pus] en japonais, comme nous l'avons signalé précédemment.

### **III.2.3. Sonorité**

Les analyses inférentielles univariées suggèrent l'existence d'une influence significative de la sonorité (profil et écart) sur le taux d'effacement. Rappelons les résultats obtenus (par ordre décroissant d'erreurs) :

Profil : Descendant = /ft/ (17%) > Montant et Plateau (8, 5%)

Ecart : Ecarts 1= /ft/ (17%) et Ecart 0= /kt, pt/ (14%) > Ecarts 2, 4, 5, (8, 75%) > Ecart 3 (5%) = /kv/

S'il est difficile de tirer des conclusions du groupe /ft/ (qui ne se trouvait qu'en position finale et qui était l'unique représentant du profil descendant), on peut souligner les résultats de l'écart 0 : la voyelle épenthétique étant par définition non accentuée, l'environnement de type « Plateau » (écart de sonorité = 0) dans lequel elle est insérée favorise son effacement, ce qui est en accord avec la nature non-voisée de

ces deux consonnes. Le pic de sonorité que devrait constituer la voyelle est ici diminué, en raison, d'une part de sa nature épenthétique, et d'autre part de son environnement consonantique non-voisée et surtout *identique* du point de vue de la sonorité.

### **III.3. Tâche C1 – Comparaison Epenthèses/Effacements**

Parmi les résultats obtenus, tant pour les épenthèses que pour les effacements, certains sont particulièrement éclairés par la mise en regard des deux évaluations.

#### Position :

En ce qui concerne la position, le rapport entre frontières de mots (attaque initiale et coda finale) pour les épenthèses et les effacements est cohérent : alors que la position initiale favorise la perception de l'input et ainsi oblige les sujets à traiter la séquence /C(u)C/ (« bonne » perception, donc peu d'effacements vocalique ou consonantiques, mais plus d'épenthèses), la position de coda finale favorise l'effacement vocalique et minimise le taux d'épenthèse (effacement consonantique dû à une « mauvaise » perception du groupe /CC/).

#### Nature :

En ce qui concerne la nature (famille, nature ou sonorité), le fait que les classements pour Obs+Cont et Obs+Sib (par exemple) soient inversés pour les épenthèses et les effacements est ici encore cohérent avec notre analyse : les groupes qui déclenchent le plus d'épenthèses (Obs+Cont) seraient ceux qui déclenchent le moins d'effacements vocaliques, lorsqu'ils sont séparés par une voyelle épenthétique. En effet, à un niveau phonologique, les groupes qui déclenchent le plus d'épenthèses sont ceux qui conduisent le plus au rétablissement du matériel segmental associé à une hypothétique « position vide » en leur milieu. Ce sont donc ceux qui tolèrent le moins les « positions vides » et donc les dévoisements vocaliques phonétiques conduisant à l'apparition de telles positions.

### **III.4. Tâche C2 - Epenthèses**

Les groupes testés étaient les suivants : /pl, pr, bl, br, fl, fr, vl, vr, tr, dr, kl, kr, gl, gr/ (OBLI), /sp, st, sk/ (S+Plosive non voisée) et /sl, sm, sn, sv, sf/ (S+C non-plosive). Conformément à nos prévisions, le taux global d'épenthèses pour les groupes CC est élevé (58%).

#### **III.4.1. Position**

Bien que les analyses inférentielles univariées suggèrent l'existence d'une influence significative de la position sur le taux d'épenthèses, les résultats descriptifs ne révèlent aucune différence importante entre les trois positions. Nous ne pouvons donc rien en conclure.

#### **III.4.2. Nature**

Les analyses inférentielles univariées suggèrent l'existence d'une influence significative de la famille générale, de la famille, de la sous-famille et de la nature spécifique phonologique sur le taux d'épenthèses. Rappelons l'essentiel des résultats obtenus (par ordre décroissant d'erreurs) :

---

Famille générale: Sib+Stop (44%) > Sib+Nonstop (32%) > Obs+Liqu (26%)

---

Famille: Sib+Stop (44%) > Sib+/Lat, Nas/ (35%) > Obs+Lat (31%) > Sib+Cont (27%) > Obs+Rhot (22%)

---

Sous-famille: Sib+Stop (44%) > Sib+/Lat, Nas/ (35%) > Stop+Lat (32%) > Sib+Cont, Cont+Lat (27%) > Stop+Rhot (23%) > Cont+Rhot (20%)

---

Nature spécifique: /s/ + C (45%-35%) > Obs + /l/ (35%-31%) > Obs + /r/ (31%-17%) > /f/+ Liqu (12%-10%)

---

Ces résultats permettent d'atteindre l'un des objectifs que nous nous étions fixés, à savoir la vérification des données relatives aux groupes Sib+Stop et Obs+Liqu que l'on

trouve dans la littérature phonologique<sup>204</sup> (Durand, 1990 ; Morelli, 1998, 1999 ; Fleischhacker, 2000) : les groupes Sib+Stop déclenchent bien plus d'épenthèses que les groupes Obs+Liqu<sup>205</sup>.

- Sib+Stop : dans le cas du japonais, ces résultats peuvent notamment s'expliquer en regards du phénomène de dévoisement des voyelles : comme le note Vance (1987, p. 48) la voyelle épenthétique /u/ est souvent dévoisée entre les consonnes sourdes /k, s, t, p/, en particulier lorsque la consonne qui la précède est une sibilante (/s/). Ainsi, les séquences phonétiques [sp, st, sk] ont naturellement tendance, pour des locuteurs japonais, à évoquer les représentations phonologiques /sup, sut, suk/, contrairement aux autres groupes. Cela expliquerait en partie le taux plus élevé d'épenthèses pour les groupes Sib+Stop.
- OBLI : les résultats permettent en outre de nous renseigner sur la difficulté relative des groupes Obs+/l/ et Obs+/R/, ceux-ci étant parfois traités sans distinction en tant que liquides. Il apparaît que le segment /l/ pose ici davantage de problèmes aux sujets japonais. Cela peut s'expliquer de plusieurs manières : si d'un point de vue perceptif, il se pourrait que le /R/, de par sa variation allophonique, soit moins aisément perçu que le /l/ dans un tel groupe, conduisant ainsi à un effacement consonantique « perceptif », d'un point de vue phonético-phonologique, le /l/ et le /R/ présentent également certaines différences, puisque la catégorie phonétique [l], en tant qu'allophone du /r/ japonais est connue des sujets, contrairement au [R]. En effet, En ce qui concerne le rapport entre le phonème /r/ en japonais et ses deux allophones principaux [r] et [l], Akamatsu (1997, p. 114) indique dans un premier temps que [l] n'est un allophone de /r/ qu'après la consonne [n] : [r] est un tap apico-alvéolaire voisé, tandis que [l] est une latérale apico-alvéolaire voisée. L'une des différences essentielles entre les deux tient donc au caractère maintenable ([l])

---

<sup>204</sup> On doit néanmoins souligner que, malgré une différence de profil de sonorité (« Descendant » vs « Montant »), les groupes /sk/ (Tâche C2) et /ks/ (Tâche C1) présentent des taux d'épenthèses équivalents (respectivement 45% et 43%).

<sup>205</sup> Notons cependant que seul un représentant de la famille « Sib » a été testé, à savoir le segment /s/.

ou non ([r]) de l'articulation. Cependant, Akamatsu (*op. cit.*, p. 105) note également que, alors que le [r] se trouve en contexte intervocalique, on trouve, en contexte post-pausal, une « variante » de [r], considérée par certains (*op. cit.*, p. 106) comme un tap apico-alvéolaire latéral (certains l'identifient même à [l]). Akamatsu souligne en outre l'importance de la variation sociolinguistique à ce sujet, les deux réalisations pouvant se trouver dans les deux contextes selon les locuteurs. Il signale également le fait que la plupart des locuteurs japonais produisent des sons relevant tantôt de la catégorie /l/, tantôt de la catégorie /r/ (*op. cit.*, p. 107) :

*« Some Japanese speakers, in pronouncing one and the same Japanese words, articulate on one occasion what a native speaker of English would perceive as some kind of r-sounds but on another occasion what he would perceive as some kind of l-sound ».*

Ce qui importe donc pour nous, c'est la similarité phonétique étroite entre le [l] et le [r], quel que soit le contexte (intervocalique ou post-pausal).

Ainsi, il se pourrait que, en raison de l'identification du [l] au /r/ japonais, la séquence [kla] puisse davantage être identifiée comme /kura/ (avec dévoisement vocalique) que la séquence [kRa], qui conduirait soit au respect de l'input, soit à un effacement consonantique (/ka/).

Par ailleurs, ces résultats confirment également *en partie* certaines prédictions effectuées précédemment dans le cadre TCRS : alors que les groupes Sib+Stop et Obs+/l/ ont un poids procédural égal à 2, les groupes Obs+/R/ et /f/+Liqu ont un poids procédural égal à 1. Or, ces deux dernières catégories présentent un taux d'épenthèse plus faible que les deux précédentes. Cependant, cette confirmation reste approximative, puisqu'il faut également prendre en compte la position (attaque ou coda).

Enfin, les résultats obtenus ont conduit à observer certaines tendances relatives au voisement des consonnes précédant les liquides. Rappelons les résultats obtenus :

<b>Groupe</b>	<b>Plus d'épenthèses</b>	<b>Moins d'épenthèses</b>
<b>C + /l/</b>	/C/ Non voisée	/C/ Voisée
<b>C + /R/</b>	/C/ Voisée	/C/ Non voisée
<b>/f/ ou /v/ + /l/ ou /R/</b>	Voisée (/v/)	Non voisée (/f/)

Rappelons alors que le dévoisement vocalique en japonais apparaît particulièrement lorsque la consonne qui le précède est non-voisée. Dans ce cas, la séquence [Cnonvoisée+l] semble plus à même d'activer une représentation /Cul/ (avec dévoisement vocalique) que la séquence [Cvoisée+l]. Cela expliquerait les résultats obtenus pour le /l/.

Pour le /R/, rappelons que la variation allophonique est plus importante que pour le /l/, l'allophone de /R/ après une consonne non-voisée étant généralement une fricative uvulaire non-voisée [χ], qui se distingue plus encore de la liquide [l] que la fricative uvulaire voisée [ʁ]. Ainsi, si d'un point de vue perceptif on pourrait considérer que la fricative non-voisée est plus propice à l'effacement consonantique « perceptif » que la fricative voisée (puisqu'elle est moins « sonore »), d'un point de vue phonético-phonologique, on pourrait considérer que le groupe non-voisé est moins propice à l'activation d'une représentation phonologique japonaise /Cur/ (et ainsi au rétablissement phonologique d'une voyelle phonétiquement dévoisée) que le groupe voisé (la similarité avec le [l] et ainsi le /r/ étant alors plus grande). On pourrait en outre également considérer que, le [R] étant un phone « inconnu », il est moins facilement assimilable au système L1 que ne l'est le [l] : dans ce cas, l'allophone non-voisé du /R/, en augmentant la distinction perceptive avec le [l] (par rapport à l'allophone voisé) serait plus facilement identifiable comme élément « étranger » et ne subirait ainsi pas les contraintes (syllabiques/phonotactiques) de la L1, qui

s'appliqueraient en revanche plus facilement au [l]. Ces trois arguments expliqueraient donc les résultats obtenus pour le /R/.

Enfin, pour les groupes dont le segment initial est /f/ ou /v/, on pourrait une fois de plus considérer que le segment non-voisé, en raison de sa moindre sonorité, conduirait à davantage d'effacement consonantique « perceptif » que le segment voisé. D'un point de vue phonético-phonologique, outre le fait que le voisement constitue généralement un degré de marque supplémentaire, il faut noter que la voyelle [v] possède une existence particulière en japonais, en raison d'une part de l'existence de l'allophone [β] de /b/ (si [v] est perçu comme [β] il pourrait activer le segment /b/), mais surtout en raison de son introduction récente en tant que « nouvelle consonne » (Labrune, 2001a) dans la société, et ainsi la langue japonaise (strate lexicale « étrangère »). Pour toutes ces raisons, le segment /v/ semble donc davantage « perceptible », et ainsi moins apte à l'effacement et davantage soumis aux contraintes syllabiques du japonais.

### **III.4.3. Sonorité**

Les analyses inférentielles univariées suggèrent l'existence d'une influence significative de la sonorité (profil et écart) sur le taux d'épenthèses. Rappelons les résultats obtenus (par ordre décroissant d'erreurs) :

Profil : Descendant = /sp, st, sk/ (44%) > Montant (28%) > Plateau (19%)

Ecart : Ecart 2 et 1=/vl, s+C<sup>206</sup>/ (40%-36%) > Ecarts 5, 3, 4 (30%-24%) > Ecart 6 et 0=/pR, tR, kR, sf/ (19%)

Ces résultats permettent de confirmer le caractère particulier des groupes /s/ + C (tant en termes de profil que d'écart), puisqu'ils déclenchent davantage d'épenthèses. Cependant, certains groupes ne respectent pas cette tendance, notamment le groupe /sf/

---

<sup>206</sup> Excepté /sl/ et /sf/.

qui constitue un « Plateau » de sonorité et qui explique ainsi le résultat obtenu à ce sujet (le groupe /sf/ est le seul de type « Plateau » ici).

On doit également noter que les groupes dont l'écart de sonorité est maximal sont /pR, tR, kR/, qui entraînent le moins d'épenthèses et qui confirment ainsi ce que nous avons affirmé précédemment à propos de la distinction perceptive entre ces groupes (non-voisés) et les groupes /pl, kl/.

### **III.5. Tâche C2 - Effacements**

Nous ne reviendrons pas ici sur l'approche générale des phénomènes d'effacements, qui a été présentée pour la tâche C1. Le taux global d'effacement est faible mais non nul (8%).

#### **III.5.1. Position**

Les analyses inférentielles univariées suggèrent l'existence d'une influence significative de la position sur le taux d'effacement. Rappelons les résultats obtenus (par ordre décroissant d'erreurs) :

Coda Finale (9%) > Attaque Initiale et Finale (4%)

Ce classement est identique à celui obtenu pour C1, et nous pouvons donc utiliser les mêmes arguments pour en rendre compte.

#### **III.5.2. Nature**

Bien que les analyses inférentielles univariées suggèrent l'existence d'une influence significative de la famille, de la nature spécifique phonologique et du profil de sonorité, il semble difficile de dégager certaines tendances générales de ces résultats, de par la faible variation entre les taux. Certains faits sont cependant à souligner : par exemple, alors que /s\_t/ entraîne 11% d'effacements et /t\_R/ 10%, les environnements /s\_k/ et /k\_R/ n'en entraîne que 2%. La différence entre /t/ et /k/, déjà évoquée précédemment



en termes de traits ([dental] vs. [vélaire]) pourrait donc ne pas être négligeable. Cependant, ces observations ne suffisent pas à établir certains principes généralisables.

### **III.6. Tâche C2 – Comparaison Epenthèses/Effacements**

Parmi les résultats obtenus, tant pour les épenthèses que pour les effacements, certains sont particulièrement éclairés par la mise en regard des deux évaluations. En particulier en ce qui concerne la nature, les inversions de classement observées entre les épenthèses et les effacements (notamment pour Obs+Liqu et pour le profil « Plateau ») rejoignent l'argument présenté dans la tâche C1 à ce sujet. On peut en revanche s'étonner du manque de cohérence pour les groupes /s/ + Plosive sourde : alors que cette classe déclenche le plus d'épenthèse de manière homogène, les résultats le sont beaucoup moins pour les effacements, puisque /s\_t/ et /s\_p/ présentent environ 10% d'effacements, alors que /s\_k/ n'en présente que 2%. Ainsi, alors que /sk/ respecte le principe d'inversion entre épenthèse et effacement, ce n'est pas le cas de /st/ et /sp/. Cela rejoint nos remarques précédentes concernant la distinction entre /p/, /t/ et /k/, qui nécessiterait plus ample investigation.

## **IV. Tâche D : comptage syllabique en unimodalité orthographique**

Les groupes testés dans la tâche D étaient les mêmes que ceux de la tâche C2, à savoir : /pl, pr, bl, br, fl, fr, vl, vr, tr, dr, kl, kr, gl, gr/ (OBLI), /sp, st, sk/ (S+Plosive non voisée) et /sl, sm, sn, sv, sf/ (S+C non-plosive). Dans cette tâche D, seule les épenthèses étaient évaluées.

La tâche D peut être envisagée de deux manières :

- Voie graphémique : la séquence orthographique est segmentée en graphèmes de gauche à droite, et le comptage syllabique s'effectue au fur et à mesure de ce séquençage.

- Voie phonologique : la séquence orthographique est globalement convertie en séquence phonologique, et le comptage syllabique s'effectue ultérieurement sur cette représentation phonologique.

Dans les deux cas, la tâche D impliquait la voie d'adressage et le système de conversion phonographémique. La question était alors de savoir si ces codes étaient ceux du français (pas d'épenthèse) ou ceux du japonais (épenthèses).

Les résultats globaux ont dépassé nos prévisions : avec 77% d'épenthèses, les taux d'épenthèses sont si élevés qu'il est difficile de repérer les éventuelles corrélations ou indépendance entre variables à l'intérieur de la tâche. De tels résultats peuvent nous faire douter de la validité même du test : n'y aurait-il pas eu d'erreurs de la part de l'expérimentateur, une incompréhension des consignes de la part des sujets, ou encore la mise œuvre d'une stratégie de réponse de leur part ? A cela, l'examen des réponses individuelles permet de répondre négativement, puisque certains, rares, sujets ont « très bien » réussi l'épreuve, avec un taux d'épenthèse quasi-nul. Cela exclut donc la possibilité d'une erreur méthodologique. En ce qui concerne les autres sujets, on constate que les réponses varient (mais généralement entre deux réponses erronées), ce qui exclut la possibilité de la mise en œuvre d'une stratégie de réponse systématique, d'autant que les distracteurs ont également été généralement correctement traités par les sujets, ce qui souligne le maintien de l'attention dans la réalisation de la tâche.

Par ailleurs, lorsque l'on examine certains résultats de Dupoux et al (1999b), nous nous trouvons confortés dans nos résultats. Ces derniers ont en effet élaboré deux ensembles de stimuli acoustiques, qu'ils ont ensuite demandé de transcrire phonétiquement à des sujets japonais à l'aide de l'alphabet romain sur un clavier d'ordinateur. Les stimuli ont été composés sous forme de triplets :

Groupe u :

- *sokdo* (non-mot avec groupe consonantique)
- *sokudo* (mot japonais avec la voyelle épenthétique /u/)
- *sokado* (non-mot sans la voyelle épenthétique /u/)

Groupe non-u :

- *mikdo* (non-mot avec groupe consonantique)
- *mikado* (mot japonais sans la voyelle épenthétique /u/)
- *mikudo* (non-mot avec la voyelle épenthétique /u/)

Or, les résultats de la tâche de transcription rejoignent en partie les nôtres : pour le « Groupe u », les non-mots avec groupe consonantique ont conduit à l'insertion d'un /u/ épenthétique dans 83% des cas, tandis que pour le « Groupe non-u » les insertions de /u/ ont été constatées dans 74% des cas, et l'insertion de la voyelle /i/ dans 10% des cas. Pour les deux groupes, seuls 16% des groupes consonantiques ont correctement été transcrits.

L'ampleur des taux obtenus semble donc être acceptable, et de tels résultats semblent par ailleurs aller dans le sens de nos hypothèses de travail. Il conviendrait néanmoins de répliquer le test D, afin de donner davantage de validité à ces résultats.

**IV.1. Position**

Bien que les trois taux soient élevés, on constate une augmentation du taux d'épenthèses pour la position de coda :

Coda Finale (84%) > Attaque Initiale et Médiale (74%)

On pourrait penser que cela est dû au fait que le groupe /CC/ en coda peut conduire à une double épenthèse (/CuCu/), mais notre protocole permet d'exclure cette hypothèse,

puisque la mesure concernait l'absence ou la présence d'une épenthèse, et non le nombre d'épenthèses.

Il se pourrait néanmoins que le caractère plus « marqué » (de par la double épenthèse possible) de la position de coda ait favorisé l'épenthèse rapport aux deux autres positions.

#### **IV.2. Nature**

Pour les familles générales, rappelons que le classement obtenu est le suivant :

Sib+Stop (79%) > Sib+Non-Stop (78%) > Obs+Liqu (76%)

L'absence d'écart important entre ces trois familles semble indiquer que les propriétés phonologiques sont peu influentes en modalité orthographique. On note néanmoins que la famille Sib+Stop déclenchent légèrement plus d'épenthèses que la famille Obs+Liqu.

A chaque niveau d'analyse (famille, sous-famille, nature spécifique, profil et écart de sonorité), les taux varient peu (généralement compris entre 75% et 80% d'épenthèse) et le voisement de la consonne /C/ dans les groupes C+/l/ et C+/R/ semble ne pas influencer le taux d'épenthèse. On note néanmoins, que le groupe à taux inférieur est /pR/, tandis que le groupe à taux supérieur est /sk/, et que le classement obtenu pour le profil de sonorité est : Descendant (79%) > Montant (77%) > Plateau (76%). Ces résultats rejoignent donc en partie ceux obtenus pour la tâche en modalité auditive C2.

#### **V. Tâche E : comptage syllabique en bimodalité audio-orthographique**

Les groupes testés dans la tâche E étaient les mêmes que ceux des tâches C2 et D, à savoir : /pl, pr, bl, br, fl, fr, vl, vr, tr, dr, kl, kr, gl, gr/ (OBLI), /sp, st, sk/ (S+Plosive non voisée) et /sl, sm, sn, sv, sf/ (S+C non-plosive).

La tâche E était cependant plus complexe que les deux autres étant donné :

- 1) Le double input (auditif et visuel) qui pouvait entraîner un certain partage attentionnel ;
- 2) La convergence/divergence aléatoire de ce double input qui pouvait entraîner certains conflits cognitifs, puisque la tâche s'apparentait à une tâche de perception dichotique.

Ces deux composants constituaient néanmoins précisément le cœur de la tâche puisque nous souhaitions évaluer la réaction des sujets face à un tel input.

La consigne, que nous rappelons ci-dessous, était par ailleurs suffisamment ambiguë pour ne pas exagérément orienter l'attention des sujets vers l'une ou l'autre des modalités :

*« Lisez les lettres montrées sur le rétroprojecteur et en même temps écoutez la prononciation de ce mot en provenance de la cassette. Divisez ce mot en syllabes (« onsetsu ») et entourez la colonne comportant le même nombre de cercles que le nombre de syllabes ».*

Par ailleurs, de manière à éviter qu'une modalité soit délibérément et systématiquement privilégiée par les sujets comme stratégie de réponse, il était offert aux sujets la possibilité de signaler d'éventuelles divergences entre les deux modalités. La tâche était donc double : segmenter/compter d'une part, et remarquer d'éventuelles différences d'autre part, ce qui pouvait entraîner d'éventuels « trade-off » attentionnels. Cela correspond à ce qu'évoquent Chen et Massaro (à paraître) dans le cas de stimuli bimodaux audio-faciaux (input visuel = visage du locuteur) :

*« Reporting audio-visual discrepancy as well as what was heard are two (possibly) opposing tasks. It is possible that between-group differences occurred in the amount of focus on one of these tasks, by either reporting incongruent*

*audiovisual (by looking more) or reporting what was heard (by visually-attending less) ».*

Cependant, la tâche n'était pas la même, puisque la nôtre impliquait un comptage syllabique et n'orientait pas explicitement l'attention des sujets vers la modalité auditive.

Par ailleurs, lorsque l'on présente un mot sous sa forme orale et écrite simultanément aux sujets, on se retrouve dans le même paradigme expérimental que celui évoqué par Katz et Frost (1992, p. 77), dans lequel les sujets devaient décider si les deux formes représentaient le même item lexical ou non. Deux possibilités sont alors évoquées pour la réalisation de la tâche : « The spoken word could have been imagined as a spelled (printed) word or subjects could have generated the phonology of the printed word ». Dans leur cas, les résultats montrent que les sujets ont choisi la deuxième option, ce qui pour eux n'est pas surprenant : « This was not surprising : subjects have had more practice reading than spelling. After converting the printed stimulus to a phonological representation, both phonological representations could then have been compared in order to determine if they matched ».

Deux opérations devaient donc être accomplies dans cette tâche E :

- 1) Segmenter/compter le nombre de syllabes du stimulus d'une ou des deux modalités.
- 2) Vérifier l'adéquation entre les deux stimuli.

En raison des contraintes temporelles imposées aux sujets, il est possible que ces derniers se soient concentrés sur l'une ou l'autre de ces deux opérations. Cependant, étant donné ces deux sous-tâches à accomplir, les sujets étaient conduits à générer deux représentations phonologiques (l'une issue du stimulus auditif, l'autre issue du stimulus orthographique), à les appairer (pour réaliser l'opération 2) puis à choisir l'une d'entre elle en cas de divergence (pour réaliser l'opération 1). Cependant, nous ne

pouvons préjuger de la chronologie des opérations, puisque trois possibilités sont envisageables en théorie :

A) Voie auditive dominante :

1. Représentation phonologique – auditive, segmentation, puis réponse - comptage
2. Représentation phonologique – orthographique, segmentation, comparaison avec la réponse – comptage précédente, puis réponse – adéquation

B) Voie visuelle dominante :

1. Représentation phonologique – orthographique, segmentation, puis réponse – comptage
2. Représentation phonologique – auditive, segmentation, comparaison avec la réponse – comptage précédente, puis réponse – adéquation

C) Pas de voie dominante *a priori* :

1. Représentations phonologiques auditive et orthographique, segmentation, comparaison, puis réponse – adéquation
2. Si adéquation : réponse-comptage ; Sinon : choix d'une modalité pour la réponse-comptage.

Notre protocole devait précisément nous permettre de soutenir l'une des ces procédures hypothétiques, tout en évaluant le degré de partage attentionnel.

Il faut alors noter que, malgré la complexité de la tâche E, les résultats globaux (entre 64% et 69% d'épenthèses pour les groupes /CC/ pour les stimuli convergents) restent inférieurs à ceux observés en unimodalité orthographique.

Nous ne reviendrons pas sur le détail des opérations de segmentation/comptage syllabique, tant en modalité auditive que visuelle, puisqu'ils ont été abordés dans les tâches C2 et D.

## **V.1. Stimuli auditifs et orthographiques convergents**

Dans ce cas, les stimuli cibles avaient la même longueur syllabique (trisyllabique), et la mesure essentielle portait sur le taux d'épenthèse. Les sujets avaient néanmoins la possibilité de signaler toute divergence qu'ils pouvaient percevoir entre les deux modalités. Le fait que 8% des stimuli identiques aient été rapportés comme différents, indique que l'une des modalités a été correctement perçue, alors que l'autre ne l'a pas été. Bien que ce taux soit faible, il constitue un premier indice de décalage possible entre les deux types de perception. Au vu des résultats obtenus en C2 et en D, on pourrait imaginer que la modalité auditive a été correctement perçue (trisyllabique, pas d'épenthèse), contrairement à la modalité visuelle (épenthèse). Cela expliquerait ce pourcentage de « discrimination intermodale » erronée.

### **V.1.1. Position**

Les analyses inférentielles univariées suggèrent l'existence d'une influence significative de la position sur le taux d'épenthèses. Rappelons le classement obtenu :

Coda Finale (71%) > Attaque Finale (66%) > Attaque Initiale (61%)

Comme pour la tâche D, la position de Coda entraîne davantage d'épenthèses que les deux autres. Si l'on considère, d'après les résultats précédents ainsi que nos hypothèses théoriques, que la modalité visuelle entraîne davantage d'épenthèses que la modalité auditive, on pourrait imaginer que, à mesure que le stimulus est décodé (sens de lecture droite - gauche), le stimulus auditif étant plus fugace que le stimulus visuel, la modalité visuelle est davantage privilégiée, conduisant ainsi les sujets à davantage reposer sur leur système de correspondance graphème - phonème (recodage phonologique du stimulus visuel selon le système moraïque japonais), alors que les positions initiales permettent un encodage phonétique auditif plus important. Une telle hypothèse (modalité auditive dominante en initiale, puis modalité visuelle dominante en finale en raison de la plus grande permanence du stimulus visuel) rendrait tout à fait compte du classement obtenu. Il se pourrait néanmoins, comme pour la tâche D, que le caractère



plus « marqué » (de par la double épenthèse possible) de la position de coda ait favorisé l'épenthèse rapport aux deux autres positions.

### **V.1.2. Nature**

Les analyses inférentielles univariées suggèrent l'existence d'une influence significative de la famille générale, de la famille, de la sous-famille et de la nature spécifique phonologique sur le taux d'épenthèses. Rappelons l'essentiel des résultats obtenus (par ordre décroissant d'erreurs) :

---

Famille générale: Sib+Stop (73%) > Sib+Nonstop (67%) > Obs+Liqu (64%)

---

Famille: Sib + /Stop, Lat, Nas/ (73%) > Obs+Lat (67%) > Obs+Rhot (61%) > Sib+Cont (55%)

---

Sous-famille: Sib + /Stop, Lat, Nas/ (73%) > Stop+Lat (69%) > Cont+Lat, Stop+Rhot, Cont+Rhot (62%) > Sib+Cont (55%)

---

Nature spécifique: /s/ + C (73%) > Obs + /l/ (69%) > Obs + /R/ (62%)

---

Ces résultats rejoignent globalement ceux obtenus pour la tâche auditive C2, et nous en reprenons donc en grande partie l'interprétation.

En ce qui concerne le voisement de la consonne /C/ dans les groupes C + Liqu, il semblerait, d'après nos résultats descriptifs que celui-ci influence les groupes C+/R/, mais non les groupes C+/l/, avec davantage d'épenthèses pour les consonnes voisées que non voisées. Si l'on considère que l'influence du voisement de la consonne est plus sensible pour le /R/ (allophone [χ]) que pour le /l/, il se pourrait alors que l'influence du facteur de voisement soit réduite en bimodalité : totalement réduite pour le /l/ mais encore sensible pour le /R/ puisque plus importante pour celui-ci en unimodalité auditive.

### **V.1.3. Sonorité**

Les analyses inférentielles univariées suggèrent l'existence d'une influence significative de la sonorité (profil et écart) sur le taux d'épenthèses. Rappelons les résultats obtenus (par ordre décroissant d'erreurs) :

Profil : Descendant (73%) > Montant (64%) > Plateau (62%)

Ecart : Ecart 2, 3 et 5 (69%) > Ecarts 4 et 0 (64%) > Ecart 6 (56%) > Ecart 1=sv/ (48%)

Le classement obtenu pour le profil de sonorité confirme la spécificité du profil Descendant (/s/ + Plosive sourde) qui déclenche davantage d'épenthèses que les deux autres. Les résultats relatifs à l'écart de sonorité confirment le taux supérieur pour les groupes /s/+ Plosive sourde (Ecart 2) et le taux inférieur pour les groupes Plosive sourde +/R/ (écart 6).

### **V.2. Stimuli auditifs et orthographiques divergents**

Dans ce cas, les deux stimuli ont des longueurs syllabiques différentes : le stimulus le plus long est légal en japonais (/CVCV/), tandis que le plus court est illégal en japonaise (/CCV/).

Deux aspects des réponses des sujets nous intéressent ici : 1) les taux de discrimination intermodale correcte et erronée ; 2) lorsque la divergence intermodale n'est pas perçue, la modalité apparemment dominante en tâche de comptage syllabique.

#### **V.2.1. Modalité dominante en tâche de comptage syllabique**

Nos résultats indiquent que :

- A) Quand le mot auditif est légal en L1 (et donc plus long) et le mot visuel illégal en L1 (et donc plus court) : la modalité auditive domine largement la modalité visuelle.
- B) Quand le mot visuel est légal en L1 (et donc plus long) et le mot auditif illégal en L1 (et donc plus court) : la modalité visuelle domine largement la modalité auditive.

On pourrait donc imaginer trois scénarii :

- Dominance de la modalité visuelle : le stimulus orthographique est converti selon le système de conversion moraique japonais conduisant à une représentation phonologique légale en L1. Cette procédure *donnerait l'impression* d'une dominance auditive dans le cas A), mais *masquerait* en fait une dominance visuelle.
- Dominance de la modalité auditive : le stimulus phonétique conduirait à une épenthèse vocalique dans la représentation phonologique (adaptation phonologique), qui serait alors légale en L1. Cette procédure *donnerait l'impression* d'une dominance visuelle dans le cas B), mais *masquerait* en fait une dominance auditive.
- Pas de dominance *a priori* : les deux stimuli conduiraient à deux représentations phonologiques qui seraient ensuite comparées, conduisant à la sélection de la représentation la plus conforme aux contraintes syllabiques du japonais. Si l'on considère que les représentations sont issues des deux processus que nous venons de décrire, alors elles seraient toutes deux légales en japonais. Cela expliquerait les résultats obtenus.

Si le troisième scénario semble *a priori* le plus « prudent », il faut néanmoins noter que la comparaison détaillée des résultats semble indiquer une légère dominance de la modalité visuelle. Si l'on considère en outre, d'après les résultats obtenus en C2 et en D, que la modalité auditive permet davantage d'éviter les épenthèses (et ainsi de violer les contraintes syllabiques du système japonais), nous sommes tentés de privilégier soit le scénario de la dominance visuelle, soit celui de l'absence de dominance *a priori*, en supposant que la représentation phonologique issue du stimulus visuel sera privilégiée.

Par ailleurs, il est intéressant de noter que, dans tous les cas, on observe entre 7% et 9% de décompte ne relevant d'aucune modalité de présentation, en particulier pour la

position de coda (celle-ci offrant la possibilité d'une double épenthèse). Il est bien sûr possible d'attribuer ces résultats au hasard (manque d'attention ou erreur des sujets dans les réponses), mais on pourrait également les attribuer à l'interaction de la perception des deux stimuli différents, conduisant à un effet de « fusion » comparable, de manière très relative, à l'effet McGurk, au niveau des structures syllabiques : un stimulus orthographique « kamapokl » couplé à un stimulus auditif /kamapokulu/ conduirait à une représentation intermédiaire du type /ka-ma-po-klu/, /ka-ma-pok-lu/ ou même /ka-ma-po-ku/.

### **V.2.2. Discrimination intermodale de stimuli divergents**

Globalement le taux de discrimination correcte est très faible (13%), ce qui peut s'expliquer, soit par une incapacité à percevoir la différence entre les deux stimuli (notamment en raison des contraintes temporelles), soit par un choix attentionnel des sujets, qui, en raison des consignes (la discrimination étant présentée comme une tâche secondaire), se sont essentiellement concentrés sur la tâche de comptage syllabique. Si la deuxième option est retenue, le faible taux de discrimination correcte indiquerait que les sujets ont effectivement privilégié une unique modalité perceptive dans la réalisation de la tâche de comptage, sans quoi on peut émettre l'hypothèse, que la prise en compte systématique des deux modalités aurait conduit à un meilleur taux de discrimination, les sujets étant plus à même de remarquer la différence entre les deux stimuli.

#### **V.2.2.1. Position**

Les résultats indiquent une bien meilleure discrimination en position de coda. Cela peut certainement être attribué au fait que la position de coda entraîne une double épenthèse : alors qu'en position d'attaque (initiale et finale) la différence entre les deux stimuli est de une syllabe (CCV vs. CVCV), elle est de deux syllabes en position de coda (VCC vs. VCVCV). La divergence entre les deux stimuli est donc plus importante, et ainsi plus sensible en position de coda, alors que les taux sont identiques pour les deux autres positions.

#### V.2.2.2. Nature

Rappelons l'essentiel des résultats obtenus (par ordre décroissant de discrimination correcte) :

---

Famille générale: Obs+Liqu (14%) > Sib+Stop (12%) > Sib+Nonstop (8%)

---

Famille: Obs+Rhot (17%) > Sib + /Stop, Lat/ (11, 5%) > Sib + /Cont, Nas/ (8, 5%) > Sib+Lat (5%)

---

Sous-famille: /Stop, Cont/ + /R/ (16%) > Sib + Stop (12%) > /Stop, Cont/ + /l/ (11%)

---

Nature spécifique: Plosive sourde + /R/ (23%) > Plosive voisée + /R/ (11, 3%) > C + /l/ (10, 5%) > /s/ + C (9, 5%)

---

Ces résultats rejoignent ceux obtenus dans les tâches de comptage syllabique : nous avons en effet vu que les groupes /s/ + C (en particulier /s/ + Plosive sourde) déclenchaient davantage d'épenthèses que les groupes C + /l/ qui eux-mêmes en déclenchaient davantage que les groupes C + /R/. Il est donc « logique » que le classement (positif) obtenu pour la tâche de discrimination intermodale soit : C + /R/ > C + /l/ > /s/ + C. Si l'on prend l'exemple d'un stimulus bimodal divergent (/CCV/, <CVCV>) (cas 1), voici le traitement que les sujets auraient pu lui accorder, en privilégiant dans un premier temps la modalité auditive (information la plus fugace) :

#### 1. Perception auditive : [CCV]

- /s/ + /C/ → épenthèse (/CVCV/);
- C + /l/ → moins d'épenthèses ;
- C + /R/ → pas d'épenthèse (/CCV/).

2. Perception visuelle : <CVCV>

- /s/ + /C/ → pas de divergence → discrimination erronée ;
- C + /l/ → faible divergence → discrimination faiblement erronée ;
- C + /R/ → divergence → discrimination correcte.

Ces résultats correspondent à ceux obtenus.

Si l'on prend l'exemple d'un stimulus bimodal divergent (/CVCV/, <CCV>) (cas 2), voici alors ce que l'on obtiendrait :

1. Perception auditive : [CVCV]

Input respecté.

2. Perception visuelle : <CCV>

- /s/ + /C/ → fort taux d'épenthèses → pas de divergence  
 → discrimination fortement erronée ;
- C + /l/ → épenthèse → discrimination erronée ;
- C + /R/ → épenthèse → discrimination erronée.

La différence de résultats entre C + /R/ et C + /l/ proviendrait donc essentiellement du premier cas. Par ailleurs, les résultats seraient identiques si l'on plaçait la perception visuelle en premier. Lorsque l'on prend en compte le facteur de voisement de la consonne des groupes OBLI, voici ce que l'on obtient (taux de discrimination correcte) :

<b>Non-Voisée</b>	<b>Voisée</b>
C non-voisée + /R/ (21%)	C voisée + /R/ (12, 5%)
C <sup>207</sup> non-voisée + /l/ (14%)	C voisée + /l/ (8%)

---

<sup>207</sup> Excepté /s/.

Les résultats obtenus ici pour le /R/ rejoignent ceux des tâches de comptage syllabique : les groupes C non-voisée + /R/ déclenchent moins d'épenthèses que leur équivalent voisé (*cf.* Tâche C2-épenthèses) et permettent ainsi une meilleure discrimination, comme nous venons de l'expliquer à travers l'exemple du cas 1.

En ce qui concerne le /l/, la situation est quelque peu différente puisque, en unimodalité auditive (tâche C2), les groupes non-voisés déclenchent davantage d'épenthèses, tandis que la modalité visuelle « neutralise » l'influence du voisement (tâche D et E), en particulier pour les groupes C + /l/. Il se pourrait alors que la tâche de discrimination n'implique pas exactement les mêmes procédures que la tâche de comptage syllabique : les sujets sembleraient alors pouvoir adopter des stratégies de discrimination qui ne nécessitent pas de comptage syllabique (et ainsi de représentation phonologique) précis. L'influence du voisement serait alors différente.

Cette différence souligne une fois de plus les différences de traitement selon les tâches entre les groupes C + /R/ et C + /l/.

#### V.2.2.3. Sonorité

On constate peu de différences entre les trois profils de sonorité (9%-13%). En ce qui concerne l'écart de sonorité, alors que la plupart des taux de discrimination correcte sont compris entre 9% et 13%, l'écart 6 présente un taux de 23%, ce qui confirme les résultats précédents sur les groupes Plosive sourde + /R/ : ces derniers facilitent la discrimination.

#### V.2.2.4. Longueur/Structure des stimuli divergents

D'après nos résultats, les taux de discrimination correcte sont identiques quelle que soit la modalité légale en L1. Si les sujets favorisaient, en termes attentionnels, la modalité auditive, alors la discrimination serait meilleure quand le stimulus auditif est illégal en L1 (moins d'épenthèse en modalité auditive). En revanche, si les sujets favorisaient la modalité visuelle, alors il y aurait peu de différences entre les deux cas. On pourrait

donc penser que les sujets ont, dans une certaine mesure favorisé la modalité visuelle. Il se pourrait en outre que, cette « préférence » ne soit pas due à un choix attentionnel des sujets (qui au contraire pourraient davantage se concentrer sur l'information la plus fugace, à savoir, l'information auditive), mais plutôt aux caractéristiques propres de la modalité visuelle qui, par nature, est plus « permanente » que la modalité auditive.

### **V.3. Comptage syllabique vs. Discrimination intermodale**

Comme nous l'avons indiqué, la tâche E combinait deux sous-tâches : l'une de comptage syllabique pour les stimuli convergents, l'autre de discrimination pour les stimuli divergents. Lorsque l'on compare les taux de réussite des deux tâches (40% pour la tâche de comptage contre 13% pour la tâche de discrimination), on s'aperçoit que la tâche de comptage a été nettement mieux réussie, comme en témoigne la répartition des résultats : seuls 14, 5% des sujets ont obtenu un résultat de comptage inférieur à l'autre, alors que 74% des sujets ont obtenu un résultat de comptage au moins trois fois supérieur à celui de la tâche de discrimination. Cela peut indiquer que les sujets se sont davantage concentrés sur la tâche de comptage, ce qui s'explique par la consigne qui leur a été donnée.

## **VI. Comparaison Tâches C2, D et E : unimodalité auditive vs. unimodalité orthographique vs. bimodalité audio-orthographique**

Cette comparaison constituait l'objectif central de notre travail, à savoir l'évaluation de l'influence de la modalité (auditive vs. visuelle orthographique) sur la conscience (méta-)phonologique des apprenants japonais. En ce qui concerne les tâches C2 et D, les résultats globaux ont dépassé nos attentes, puisque la modalité visuelle déclenche 77% d'épenthèses, alors que la modalité auditive n'en déclenche que 58%. Nous nous concentrons donc essentiellement sur la comparaison des tâches C2 (auditif) et D (visuel), avant de prendre en compte la tâche E (bimodalité) qui constitue, dans une certaine mesure, un cas intermédiaire entre les deux tâches, comme cela est confirmé



par les résultats de cette tâche (66%). En effet, il apparaît que, quelle que soit la variable indépendante prise en compte (position ou nature), les taux d'épenthèses s'ordonnent de la manière suivante :

Modalité Visuelle > Bimodalité > Modalité Auditive

On peut donc considérer que, dans la tâche E, les sujets ont tantôt focalisé leur attention sur la modalité visuelle, tantôt sur la modalité auditive (variation inter- et/ou intra-individuelle), ou bien que leurs réponses ont été influencées par l'interaction entre les deux modalités.

- Si l'on considère que la modalité auditive est dominante, alors on peut affirmer que l'ajout de la modalité visuelle a été « négative », puisque les taux d'épenthèses en bimodalité sont plus élevés qu'en unimodalité auditive.
- Si l'on considère que la modalité visuelle est dominante, alors on peut affirmer que l'ajout de la modalité auditive a été « positive », puisque les taux d'épenthèse en bimodalité moins élevés qu'en unimodalité visuelle.

Il apparaît alors que la consigne de réalisation de la tâche peut grandement influencer la performance en situation de bimodalité, en orientant l'attention des sujets. Dans notre protocole, les trois instructions relatives au comptage syllabique, que nous rappelons ci-dessous, étaient relativement similaires. Il va de soi que la version française donnée ci-dessous (traduction française de la version japonaise, elle-même traduite du français) ne peut prétendre à une identité parfaite avec les instructions originales qui, elles, étaient en japonais.

Tâche	Consigne de Comptage Syllabique
<b>C2</b>	Écoutez le mot sur la cassette, divisez le mot que vous avez entendu en syllabes (« onsetsu ») et entourez la colonne comportant le même nombre de cercles que le nombre de syllabes.
<b>D</b>	Lisez les lettres montrées sur le rétroprojecteur, divisez ce mot en syllabes (« onsetsu ») et entourez la colonne comportant le même nombre de cercles que le nombre de syllabes.
<b>E</b>	Lisez les lettres montrées sur le rétroprojecteur et en même temps écoutez la prononciation de ce mot en provenance de la cassette. Divisez ce mot en syllabes (« onsetsu ») et entourez la colonne comportant le même nombre de cercles que le nombre de syllabes.

Les trois grandes conclusions, relatives aux variables indépendantes, qui émergent de la comparaison des résultats sont les suivantes :

<b><u>Modalité</u></b> :	La modalité visuelle <sup>208</sup> augmente le taux d'épenthèses.
<b><u>Nature</u></b> :	La modalité visuelle diminue l'influence des propriétés phonologiques segmentales.
<b><u>Position</u></b> :	La modalité visuelle modifie l'influence de la position.

*A contrario*, la modalité auditive diminue le taux d'épenthèses et augmente l'influence des propriétés phonologiques segmentales.

---

<sup>208</sup> En uni- ou en bimodalité.

A un niveau général, six axes d'interprétation sont alors envisageables :

**Axe mnésique :**

Comme nous l'avons vu dans le modèle de mémoire de travail de Baddeley pour des lecteurs non-experts, l'information orthographique doit être recodée phonologiquement avant de pouvoir être stockée, contrairement à l'information auditive qui dispose d'un accès direct. L'information orthographique serait donc plus susceptible d'être influencée par la structure moraïque du système phonologique japonais que l'information auditive, d'où le taux plus élevé d'épenthèses pour la modalité visuelle.

**Axe métaphonologique :**

Comme nous l'avons vu, l'orthographe permet à la parole d'être clairement définie et segmentée en « unités » stables, contrairement à la parole qui est continue, variable dans ses réalisations phonétiques et fugace. C'est précisément ce qui fait de l'orthographe (lecture/écriture) un outil de développement particulièrement efficace de la conscience phonologique (phonémique). Or la tâche de comptage syllabique fait précisément appel à cette conscience phonologique, puisqu'il s'agit d'identifier et de compter ces « unités ». La conscience phonologique étant fondamentalement contrainte par l'apprentissage de la lecture en langue première, la structure moraïque du système d'écriture japonais exerce donc une influence particulière sur l'identification de ces unités. La modalité visuelle implique donc l'activation du système phonographémique des sujets, qui est contraint, non seulement par la structure phonologique de leur L1, mais également par la structure du système d'écriture, relativement transparent, de leur L1. Ainsi, alors que la modalité auditive permet, dans une certaine mesure, l'établissement de « nouvelles unités » (en particulier au niveau phonétique), la modalité visuelle semble impliquer davantage le système phonologique et phonographémique japonais qui sont fondamentalement moraïques. Ainsi, la modalité visuelle active des représentations phonographémiques qui sont moraïques, d'où le taux plus élevé d'épenthèses en modalité visuelle. Cet argument est particulièrement renforcé par l'axe suivant :

### **Axe sociophonologique :**

Comme nous l'avons indiqué tout au long de notre étude, la situation sociolinguistique, et plus particulièrement sociophonologique de la société japonaise a beaucoup évolué ces dernières décennies, en particulier en milieu urbain et parmi les jeunes générations, que représentent nos sujets : cela est notamment dû à la présence massive des emprunts étrangers, et plus particulièrement anglais, dans l'environnement japonais. Or, cette présence est particulièrement marquée sous sa forme écrite, dans le domaine des médias, de la publicité et de l'édition. Ce qui est alors frappant, c'est la « pluri-scripturalité » qui règne au Japon : sur une même affiche publicitaire, par exemple, figurent côte à côte un emprunt sous sa forme orthographique originale et la version « adaptée » de cet emprunt, le plus souvent en caractères katakana. Cela indique à quel point, la forme orthographique peut être intégrée dans la société, sans que le soit la forme phonologique. Une telle situation sociolinguistique semble alors conduire à l'automatisation des conversions entre formes orthographiques originales (alphabet) et formes « adaptées » en caractères kana. Un stimulus orthographique de type <CC> correspond alors tout naturellement à une forme phonologique /CvC/ pour un sujet japonais. On peut donc considérer qu'un nouveau système phonographémique se développe au Japon, sans qu'y soit, pour autant, intégré un nouveau système phonologique. Un tel argument éclaire alors fortement l'axe métaphonologique que nous venons de présenter, et il rend bien compte du fait que la modalité visuelle, en activant de manière *automatique* des codes de conversion graphème-phonème en L1, déclenche davantage d'épenthèses que la modalité auditive, qui demande en revanche un traitement davantage *contrôlé*. On comprend alors qu'une telle situation sociolinguistique, si elle peut sembler avoir l'avantage de familiariser les Japonais avec le lexique étranger, peut s'avérer tout à fait néfaste pour l'apprentissage phonologique. Les formes orthographiques deviennent « familières », mais non les formes phonologiques, ce qui peut entraîner des dysfonctionnements dans l'apprentissage lorsque le système phonographémique des sujets est activé.

### **Axe phonético-phonologique :**

Suivant toujours l'idée selon laquelle la modalité visuelle active les représentations phonologiques « profondes », les résultats obtenus pourraient possiblement être éclairés par la distinction entre syllabes *phonétiques* et syllabes *phonologiques*. Comme nous l'avons évoqué plus haut, en raison des phénomènes de dévoisement vocalique autorisés en japonais, la modalité auditive pourrait permettre aux sujets d'effectuer un comptage de syllabes phonétiques (moins d'épenthèses), alors que la modalité visuelle active des représentations phonologiques et ainsi conduit à un comptage de syllabes phonologiques (plus d'épenthèses). Cet argument permettrait donc de rendre compte des résultats obtenus, et il s'articule avec les deux axes suivants :

### **Axe perceptif :**

Alors que la modalité auditive, de par la fugacité et les caractéristiques (notamment prosodiques) du signal acoustique, peut conduire à une « mauvaise » perception du stimulus, conduisant ainsi par exemple à un effacement consonantique dans un groupe de consonnes, qui ne nécessite alors aucune épenthèse, la modalité visuelle, outre sa stabilité, ne subit pas l'influence de telles caractéristiques. Il serait donc beaucoup plus fréquent de « mal entendre » un groupe /CC/ que de « mal voir » un groupe /CC/ pour nos sujets. On pourrait donc considérer que la modalité auditive, de par le « bruit » qui lui est associé, favorise les effacements consonantiques et déclenche ainsi moins d'épenthèses que la modalité visuelle. Cet argument s'appuie particulièrement sur le dernier axe explicatif.

### **Axe attentionnel :**

Comme nous l'avons indiqué plus haut, on pourrait considérer que la modalité auditive, en raison de la fugacité, des caractéristiques acoustiques du signal, ainsi que de la situation sociolinguistique japonaise, conduit les sujets à lui accorder une attention plus soutenue (contrôle) que celle accordée à la modalité visuelle (automatique). Cependant, nous avons également indiqué que, malgré cette attention, la modalité auditive était plus à même de conduire les sujets à une « mauvaise perception » que la modalité

visuelle. Il semblerait alors que la modalité visuelle oblige, d'un point de vue attentionnel, à préserver davantage l'input de surface (pas d'effacement consonantique) que la modalité auditive. Cela s'illustre en particulier par l'influence de l'interaction entre modalité et position : en modalité auditive (tâche C2), la position de Coda déclenche autant d'épenthèses que la position d'attaque finale et légèrement plus d'épenthèses que l'attaque initiale. Or, cet écart de taux entre attaque initiale et coda finale augmente considérablement en modalité visuelle :

	<b>Attaque Initiale</b>	<b>Coda Finale</b>
<b>Modalité Auditive</b>	19%	20% → écart = 1%
<b>Modalité Audio-visuelle</b>	61%	71% → écart = 10%
<b>Modalité Visuelle</b>	73%	84% → écart = 11%

Alors que la modalité auditive faciliterait l'effacement consonantique en coda finale (« mauvaise perception », malgré le focus attentionnel), la modalité visuelle « obligerait » les sujets à traiter les groupes /CC/ en coda finale, étant donné leur saillance attentionnelle/perceptive, d'où le taux bien plus élevé d'épenthèses en modalité visuelle<sup>209</sup>. La modalité visuelle ne permettrait donc pas « d'ignorer » (négligence attentionnelle) les groupes /CC/, contrairement à la modalité auditive. L'épenthèse vocalique primerait donc en modalité visuelle, tandis que l'effacement consonantique pourrait être non négligeable en modalité auditive.

En ce qui concerne le rôle de l'attention, il est remarquable de noter que, malgré la demande attentionnelle accrue (voire le « partage attentionnel ») dans la tâche E en raison de la bimodalité des stimuli, les taux d'épenthèse restent inférieurs à ceux obtenus en unimodalité visuelle (tâche D). Deux arguments sont donc croisés ici :

- D'un point de vue cognitif, on peut imaginer que la demande attentionnelle de la modalité auditive est plus élevée, ce qui conduit ainsi à un traitement plus

---

<sup>209</sup> Nous remercions Jean-Luc Nespoulous pour ses suggestions très avisées sur la question.

« contrôlé » qui permet de mieux respecter l'input en fonction du système interphonologique, d'où le taux moindre d'épenthèses.

- D'un point de vue perceptivo-attentionnel, on peut considérer que la modalité visuelle diminue la possibilité de négligence attentionnelle et conduit ainsi les sujets à traiter les groupes /CC/, de manière toutefois plus « automatique » qu'en modalité auditive, d'où le taux plus élevé d'épenthèses.

Ces six axes explicatifs permettent donc clairement, nous semble-t-il, de rendre compte des résultats obtenus. Il nous semble en particulier important de souligner le rapport entre représentations phonologiques « profondes » et unités orthographiques : comme nous l'avons dit, l'apprentissage d'une orthographe conditionne fortement la conscience phonologique des sujets. Or, si l'on adopte le point de vue de la phonologie multilinéaire en distinguant, d'une part le squelette syllabique, et d'autre part le matériel segmental, les résultats que nous avons obtenus ici nous encouragent à penser que le lien entre représentations phonologiques « profondes » (squelette syllabique) et représentation orthographique est particulièrement fort dans le domaine de la segmentation de la parole : dans le cas du japonais en particulier, la représentation écrite (en particulier dans les syllabaires kana) « ancrerait », d'une certaine manière, la structure phonologique de leur langue première. L'apprentissage de la lecture/écriture en japonais par les syllabaires kana renforcerait donc, dans « l'esprit » des locuteurs, la structure moraique du système phonético-phonologique que ces syllabaires ont pour vocation de transcrire.

Malgré les imperfections et les faiblesses de l'interprétation que nous avons proposée des résultats obtenus, il est à présent possible de répondre, en partie au moins, à nos questions initiales, avant de nous plonger dans les implications didactiques, spécifiques et générales, de nos résultats et de la réflexion que nous avons conduite jusqu'à présent.

## **VII. Synthèse et Perspectives**

Avant de reprendre nos hypothèses de travail, en vue de synthétiser les résultats obtenus dans nos tests, il convient de reconnaître que notre étude, tant dans le protocole que dans notre interprétation des données, présente certaines faiblesses. Celles-ci sont dues à la fois au cadre « écologique » pour lequel nous avons opté, mais aussi à un manque de contrôle de certaines variables pourtant essentielles, à savoir en particulier les caractéristiques acoustiques des stimuli auditifs<sup>210</sup>, ainsi que la fréquence d'occurrence, en japonais, des syllabes et des segments utilisés, sous leur forme phonétique et/ou orthographique. D'autres facteurs, comme le voisinage orthographique et/ou phonologique, ont également été négligés. De plus, la multiplicité des niveaux d'analyse possibles (phonétique, phonologique, graphémique) impliquée dans notre protocole n'a pas permis, dans l'interprétation des résultats, de dégager clairement la nature et/ou le fonctionnement des interactions entre ces trois niveaux, en particulier dans le domaine segmental (tâches A) et F)). En ce qui concerne les variables descriptives, nous n'avons calculé que les taux plus importants en fonction des variables qui présentaient naturellement une valeur équilibrée dans les groupes :

<b>Taux de confusion phonographémique</b>	<b>Sexe Masculin</b>	<b>Sexe Féminin</b>
<b>Tâche A</b>	33%	25%
<b>Tâche F</b>	20%	22%

<b>Taux de confusion phonologique</b>	<b>Sexe Masculin</b>	<b>Sexe Féminin</b>
<b>Tâche F</b>	29%	25%

---

<sup>210</sup> Par exemple, pour la tâche F), celle-ci visait à évaluer la capacité de discrimination entre les segments /b/ et /v/, notamment en position de coda. Or, l'on sait bien (Wioland, 1991) que le segment /v/ en finale va provoquer l'allongement de la voyelle précédente (ou plutôt va se trouver en position faible et va présenter une « résistance » moins marquée, raison pour laquelle Wioland se refuse à parler de consonne allongée : il s'agit plutôt de « résistance » moindre), ce qui n'est pas le cas de /b/. On pourrait alors supposer que, étant donné la sensibilité des Japonais à la longueur vocalique, celle-ci pourrait constituer un indice perceptif permettant de distinguer [b] et [v]. Ce serait alors la perception de la longueur vocalique qui serait évaluée, et non celle du segment consonantique à proprement parler.



<b>Taux de confusion syllabique</b>	<b>Sexe Masculin</b>	<b>Sexe Féminin</b>
<b>Tâche C1</b>	63%	65%

<b>Taux de confusion syllabique</b>	<b>Sujets spécialisés en français</b>	<b>Sujets non-spécialisés en français</b>
<b>Tâche C2</b>	48%	69%
<b>Tâche D</b>	73%	80%
<b>Tâche E</b>	57%	69%

Ces derniers résultats soulignent une autre faiblesse de notre étude, à savoir le manque de profondeur et d'exhaustivité de nos analyses statistiques, pourtant indispensables à l'évaluation du degré de significativité des résultats.

Les imperfections de ce travail, si elles doivent en relativiser les résultats, permettent cependant de mettre en lumière les paramètres qui devront être mieux contrôlés lors de travaux ultérieurs, ainsi que les informations qui ont manqué pour une interprétation plus approfondie de nos données. Ces défauts méthodologiques ayant été mentionnés, nous pouvons à présent reprendre notre questionnement initial et évaluer nos hypothèses de travail face aux résultats obtenus.

### **VII.1. Synthèse**

Si l'on reprend les hypothèses opérationnelles que nous avons pu tester pour chaque tâche<sup>211</sup>, et qu'on les confronte aux résultats que nous avons pu obtenir, voici ce que l'on trouve :

---

<sup>211</sup> Car il est apparu que nous n'avons pu tester, au moins statistiquement, la totalité des hypothèses opérationnelles.

<b>Tâche A (pp. 771-772)</b>	
<ul style="list-style-type: none"> <li>• La position (lexicale et syllabique) du segment consonantique en L2 influence de manière significative le taux de confusion phonographémique ;</li> <li>• La nature (en termes de famille et de sonorité phonologique) du segment consonantique en L2 influence de manière significative le taux de confusion phonographémique ;</li> <li>• La nature du segment vocalique adjacent au segment consonantique en L2 (existant ou non en L1) influence de manière significative le taux de confusion phonographémique consonantique.</li> <li>• Dans tous les cas, l'étudiant choisit préférentiellement les graphèmes existant en L1 ;</li> </ul>	<ul style="list-style-type: none"> <li>• OUI</li> <li>• OUI</li> <li>• OUI</li> <li>• NON</li> </ul>

<b>Tâche F (pp. 790-791)</b>	
<ul style="list-style-type: none"> <li>• La position du segment cible en L2 influence de manière significative les taux de confusion phonologique et phonographémique ;</li> <li>• La nature du segment cible en L2 influence de manière significative les taux de confusion phonologique et phonographémique ;</li> <li>• Le taux de confusion phonographémique est plus élevé pour la tâche F que pour la tâche A.</li> </ul>	<ul style="list-style-type: none"> <li>• OUI</li> <li>• OUI</li> <li>• NON</li> </ul>

<b>Tâches C1 (p. 826) et C2 (p. 862)</b>	
<ul style="list-style-type: none"> <li>• La position du groupe cible en L2 influence de manière significative le taux d'épenthèses ;</li> <li>• La nature (famille et sonorité) du groupe cible en L2 influence</li> </ul>	<ul style="list-style-type: none"> <li>• OUI</li> <li>• OUI</li> </ul>

<p>de manière significative le taux d'épenthèses ;</p> <ul style="list-style-type: none"> <li>• L'insertion de voyelle(s) épenthétique(s) propre(s) à la L1 dans les groupes cibles en L2 donne lieu au respect de la structure syllabique des stimuli (aucune confusion syllabique).</li> </ul>	<ul style="list-style-type: none"> <li>• NON</li> </ul>
--	---

<b>Tâche D (p. 880)</b>	
<ul style="list-style-type: none"> <li>• La position du groupe cible graphémique en L2 influence de manière significative le taux d'épenthèses ;</li> <li>• La nature du groupe cible graphémique en L2 influence de manière significative le taux d'épenthèses.</li> </ul>	<ul style="list-style-type: none"> <li>• OUI</li> <li>• OUI</li> </ul>

<b>Tâche E</b>	
<p><u>A) Lorsque les longueurs syllabiques des stimuli auditifs et visuels sont identiques (p. 903) :</u></p> <ul style="list-style-type: none"> <li>• La position du groupe cible en L2 influence de manière significative le taux de confusion syllabique.</li> <li>• La nature du groupe cible en L2 influence de manière significative le taux de confusion syllabique.</li> </ul>	<ul style="list-style-type: none"> <li>• OUI</li> <li>• OUI</li> </ul>
<p><u>B) Lorsque les longueurs syllabiques des stimuli auditifs et visuels sont différentes (pp. 905-921) :</u></p> <ul style="list-style-type: none"> <li>• La position du groupe cible en L2 influence de manière significative le taux de discrimination phonographémique.</li> <li>• La nature du groupe cible en L2 influence de manière significative le taux de discrimination phonographémique.</li> <li>• Lorsqu'il y a discrimination : le taux de discrimination phonographémique correct est supérieur au taux de discrimination phonographémique erronée.</li> </ul>	<ul style="list-style-type: none"> <li>• OUI</li> <li>• OUI</li> <li>• NON</li> </ul>

<sup>212</sup> Nous n'avons pas testé le cas de figure ([CvC], "CvC").

<ul style="list-style-type: none"> <li>• Lorsqu'il n'y a pas discrimination : c'est le mot graphémique (et donc la modalité visuelle) qui est privilégié pour le comptage syllabique phonologique.</li> </ul> <p><u>C) Globalement :</u></p> <ul style="list-style-type: none"> <li>• Couple ([CC], "CC") : taux de confusion syllabique phonologique élevé ;</li> <li>• Couple ([CC], "CvC") : taux de confusion syllabique phonologique élevé ;</li> <li>• Couple ([CvC], "CC") : taux de confusion syllabique phonologique faible<sup>212</sup>.</li> </ul>	<ul style="list-style-type: none"> <li>• POSSIBLE</li> <li>• OUI</li> <li>• OUI</li> <li>• NON</li> </ul>
--	---

<b>« Méta-tâche » C2/D/E (pp. 924-931)</b>	
<p>Pour les stimuli [CC] communs aux trois tests, le taux de confusion syllabique (correspondant pour nous au taux d'épenthèses) est :</p> <ul style="list-style-type: none"> <li>• Plus élevé dans le test D que dans le test C2 ;</li> <li>• Moins élevée (ou identique) dans le test E que dans le test D.</li> </ul>	<ul style="list-style-type: none"> <li>• OUI</li> <li>• OUI</li> </ul>

Nous pouvons donc, à présent, répondre à nos questions initiales, que nous avons formulées en deux temps : d'abord dans une perspective strictement phonético-phonologique, puis dans une perspective plus largement psycholinguistique.

**Perspective phonético-phonologique :**

Domaine	Segmental (phonético-phonologique)		Syllabique (phonologique)		
	Syntagmatique	Paradigmatique	Syntagmatique	Paradigmatique	
Axes	Position	Nature	Position	Nature	Structure
Aspects	1	2	3	4	5
Questions	1	2	3	4	5
Expérience	(F)	(F)	(C1 et C2)	(C1 et C2)	(C1 et C2)
<b>1</b>	<i>La position du segment cible dans le mot phonologique influence-t-elle le taux de confusion segmentale ? → OUI</i>				
<b>2</b>	<i>La nature du segment cible dans le mot phonologique influence-t-elle le taux de confusion segmentale ? → OUI</i>				
<b>3</b>	<i>La position du groupe cible dans le mot phonologique influence-t-elle le taux de confusion syllabique ? → OUI</i>				
<b>4</b>	<i>La nature du groupe cible dans le mot phonologique influence-t-elle le taux de confusion syllabique ? → OUI</i>				
<b>5</b>	<i>La structure de surface du groupe cible dans le mot phonologique influence-t-elle le taux de confusion syllabique ? → OUI</i>				

**Perspective psycholinguistique :**

<b>Domaine</b>	<b>Segmental</b>	<b>Syllabique</b>
<b>Expériences</b>	<b>A et F</b>	<b>C2, D et E</b>
<b>A et F</b>	<p>La position du segment consonantique cible dans le mot phonologique influence-t-elle le taux de confusion phonographémique ?</p> <p><u>Tâche simple</u> : appariement → <b>OUI</b></p> <p><u>Tâche complexe</u> : appariement et discrimination → <b>OUI</b></p>	
<b>A et F</b>	<p>La nature du segment consonantique cible dans le mot phonologique influence-t-elle le taux de confusion phonographémique ?</p> <p><u>Tâche simple</u> : appariement → <b>OUI</b></p> <p><u>Tâche complexe</u> : appariement et discrimination → <b>OUI</b></p>	
<b>A et F</b>	<p>La nature du segment vocalique adjacent au segment cible dans le mot phonologique influence-t-elle le taux de confusion phonographémique ?</p> <p><u>Tâche simple</u> : appariement → <b>OUI</b></p> <p><u>Tâche complexe</u> : appariement et discrimination → <b>OUI</b></p>	
<b>F (et A)</b>	<p><u>Questions liées à la tâche complexe :</u></p> <p><u>Effet de différenciation :</u></p> <p>La nature des couples de mots phonologiques (identiques / différents) influence-t-elle le taux de confusion phonographémique ? → <b>Peut-être</b></p> <p><u>Effet de complexité :</u></p> <p>Le taux de confusion phonographémique est-il supérieur en tâche simple ou complexe ? → <b>Tâche Simple</b></p>	

<p><b>E</b></p>	<p><b>Tâche complexe</b> : comptage syllabique phonologique et discrimination syllabique phonographémique</p> <p><u>Cas 1</u> : longueur syllabique identique</p> <p>*La position du groupe cible dans le mot bimodal influence-t-elle le taux de confusion syllabique ? → <b>OUI</b></p> <p>*La nature du groupe cible dans le mot bimodal influence-t-elle le taux de confusion syllabique ? → <b>OUI</b></p> <p><u>Cas 2</u> : longueur syllabique différente</p> <p>*La position du groupe cible dans le mot phonologique influence-t-elle le taux de discrimination phonographémique ? → <b>OUI</b></p> <p>*La nature du groupe cible dans le mot phonologique influence-t-elle le taux de discrimination phonographémique? → <b>OUI</b></p> <p><u>2-A</u> : Lorsqu'il y a discrimination : le taux de discrimination phonographémique correct est-il supérieur au taux de discrimination phonographémique erronée ? → <b>NON</b></p> <p><u>2-B</u> : Lorsqu'il n'y a pas discrimination : quel mot (phonologique ou graphique) et donc quelle modalité est privilégiée pour le comptage syllabique ? → <b>Peut-être la modalité visuelle</b></p> <p><u>Effet de complexité</u> :</p> <p>Trouve-t-on un écart entre les taux de confusion syllabique (comptage) et phonographémique (discrimination) pouvant indiquer un conflit procédural d'ordre attentionnel ? → <b>OUI</b></p>
<p><b>D</b></p>	<p>La position du groupe cible dans le mot graphique influence-t-elle le taux de confusion syllabique ? → <b>OUI</b></p>
<p><b>D</b></p>	<p>La nature du groupe cible dans le mot graphique influence-t-elle le taux de confusion syllabique ? → <b>OUI</b></p>

<b>C2-D-E</b>	Le mode de présentation des mots (auditif / orthographique / audio-orthographique) influence-t-il le taux de confusion syllabique ? → <b>OUI</b> L'influence (potentielle) de la position et de la nature des groupes cibles varie-t-elle en fonction de la modalité ? → <b>OUI</b>
---------------	--

## **VII.2. Implications méthodologiques**

Tout en tenant compte des limites et des faiblesses de notre étude, les réponses auxquelles nous sommes parvenus, ainsi que notre réflexion préalable alimentée par la littérature spécialisée<sup>213</sup>, nous permettent dès lors d'insister sur la valeur que doivent accorder les chercheurs, tant en psycholinguistique qu'en phonétique ou en phonologie, au rôle des interactions entre représentations phonético-phonologiques et orthographiques, notamment dans les travaux sur la perception ou la représentation métalinguistique de la parole. Au vu de la variation des unités perceptives entre les langues, le fait qu'il puisse y avoir certaines interactions entre orthographe et phonologie, tant en perception qu'en production, oblige en effet les investigateurs à prendre des précautions supplémentaires soit dans leur protocole d'enquête ou expérimental, soit dans leur interprétation des phénomènes observés. C'est notamment ce que soulignent Damian et Bowers (2003) à propos des tâches d'amorce, en citant une étude qui visait à déterminer si les segments phonologiques ou les représentations de traits étaient les unités de base de la parole à un niveau phonologique. Dans cette étude, la forme orthographique des stimuli avait été négligée. Or, comme le montre leur propre étude, une incongruité orthographique peut, à elle seule, perturber l'effet d'amorce (2003, p. 129) :

*« The numerous studies that have employed this technique have not considered the relevance of spelling, and accordingly, effects attributed to phonology may have another source ».*

---

<sup>213</sup> Et parmi les travaux les plus récents, on peut citer ceux de Chéreau, Hallé et Ségui (1999) ou ceux de Ventura, Kolinsky, Brito-Mendes et Morais (2001).



Cette remarque peut s'avérer particulièrement importante en phonologie, puisque la question de l'orthographe en tant que variable dans la méthodologie phonologique nous semble généralement peu considérée. A titre d'illustration, on peut ainsi se pencher sur l'étude de Shinohara (1997), qui a présenté à ses informatrices une liste de mots écrits. L'auteur justifie sa démarche de la manière suivante (p. 10) :

*« La raison pour laquelle nous avons utilisé la liste écrite est d'éviter la variabilité phonétique qui peut apparaître dans chaque énoncé. Nous voulons obtenir les formes adaptées à partir des formes phonologiques de L2 que les informateurs stockent ».*

Au vu de notre étude, on peut émettre quelques réserves vis-à-vis de cette approche méthodologique, qui, en outre, reposait sur un nombre limité de sujets (p. 15) :

*« Trois informatrices, locutrices du dialecte de Tokyo, âgées de 30 à 40 ans, ayant vécu à Paris à l'âge adulte pour leurs études supérieures et ayant une bonne maîtrise du français. Nous leur avons présenté des listes de mots français et leur avons demandé d'adapter ces mots français de la façon dont elles les prononceraient en japonais. Dans la plupart des cas, elles devaient écrire en face de chaque mot français la forme adaptée en se servant du syllabaire japonais. Dans d'autre cas, nous leur avons demandé de faire l'adaptation oralement (soit pour contrôler le décalage entre leur prononciation réelle et leur transcription, soit pour les données de l'accentuation) et les formes adaptées ont alors été transcrites par l'auteur elle-même. [...] La raison pour laquelle nous avons utilisé des listes écrites est que les formes adaptées sont formées à partir d'une représentation phonologique de chaque mot français stockée dans leur vocabulaire français. Nous ne voulions pas l'adaptation de séquences de sons qu'elles entendent physiquement, parce que dans ce cas nous ne saurions pas si le son de l'input se réfère au système phonologique du français ou s'il se réfère aux propriétés acoustiques de l'énoncé prononcé ».*

Le rôle des représentations orthographiques dans certains travaux en phonétique et en phonologie devrait, nous semble-t-il, être ainsi davantage pris en considération par les chercheurs, dès que ceux-ci font appel à des sujets « lettrés ».

### **VII.3. Perspectives**

Ayant présenté tant les faiblesses que l'intérêt de nos protocoles semi-expérimentaux, l'étude générale de l'influence du système phonographémique de la langue première sur la perception et l'apprentissage du système phonologique d'une langue étrangère, nécessiterait de nombreuses études complémentaires et affinées. Les perspectives ouvertes par la présente étude sont donc très nombreuses, et nous présentons ci-dessous plusieurs pistes d'investigation permettant d'approfondir notre connaissance des phénomènes en jeu, que la perspective soit plutôt « théorique » (phonologie, psycholinguistique) ou « pratique » (didactique). Pour chacun des tests présentés ci-dessous, plusieurs variables doivent être considérées, en particulier celles que nous avons négligées dans nos tests (notamment fréquence et voisinage). L'ensemble des facteurs mentionnés l'est donc à titre d'exhaustivité : il va de soi que, *en pratique*, plusieurs expérimentations doivent être menées pour pouvoir tester la totalité des paramètres en jeu, puisque, par exemple, la tâche de production orale (libre ou guidée) ne permet pas de contrôler certains facteurs linguistiques (fréquence, etc.).

#### **VII.3.1. Phonétique-Phonologie**

Parmi l'ensemble des tests envisageables relatifs à notre problématique, certains concernent particulièrement les domaines phonétique et phonologique. Nous en donnons ci-dessous deux exemples.

**A) Domaine segmental : l'exemple du /R/ : variation phonétique en lecture vs. en répétition**

<b>Mesures Acoustiques</b>			
<b>Objectif</b>	Mesurer la variation phonétique des réalisations du /r/, en français et en japonais, selon la tâche (Lecture/ Répétition/Production).		
<b>Sujets et Input</b>	Français natifs et /R/ français	Variables possibles :	<ul style="list-style-type: none"> <li>• Tâche (Lecture vs. Répétition vs. Production)</li> <li>• Items (Mots vs. Non-mots)</li> <li>• Environnement Vocalique et Fréquence</li> <li>• Position Lexicale et Fréquence</li> <li>• Contexte (Lexical vs. Phrastique)</li> </ul>
	Japonais natifs et /r/ japonais	Variables possibles :	<ul style="list-style-type: none"> <li>• Idem. (mais en japonais)</li> <li>• Type de script pour la tâche de lecture (romaji vs. hiragana vs. katakana vs. kanji)</li> </ul>
	Japonais apprenant le français et /R/ français	Variables possibles :	<ul style="list-style-type: none"> <li>• Idem (mais en français, donc un script unique)</li> </ul>
<b>Output</b>	Enregistrements des productions des sujets, analyse acoustique (formantique, etc.)		

L'objectif principal de ce type de test serait de comparer la variation allophonique du segment /r/ (en français ou en japonais) selon la modalité de l'input (et donc la tâche) : lecture à voix haute vs. répétition. Notre hypothèse serait que la variation est moindre en tâche de lecture. Bien sûr, pour la tâche de répétition, deux conditions pourraient être proposées : dans la première, l'input serait « naturel », c'est-à-dire que l'input lui-même suivrait les règles allophoniques de la langue en question ; dans la seconde, l'input serait « contrôlé », c'est-à-dire qu'un allophone unique serait utilisé, via un

synthétiseur de parole, dans tous les contextes (voir à ce propos Peperkamp, Pettinato, et Dupoux, 2003).

Dans le cas où la variation serait effectivement moindre en tâche de lecture, on pourrait penser que l'écrit conduit à une sorte de « prototype » des réalisations phonétiques : la représentation orthographique pourrait alors, au niveau segmental, être considérée comme plus proche d'une représentation phonologique que phonétique. L'écrit serait donc « opposé » à la variation phonétique, ce qui pourrait avoir des implications didactiques soit « positives » soit « négatives » pour son utilisation en tant que support d'apprentissage de l'oral.

### **B) Domaine syllabique : Variété du français et schwa**

<b>Comptage Syllabique et Répétition</b>	
Objectif	Evaluer l'effet de la variété de français (« Nord » vs. « Sud ») vis-à-vis du « schwa » sur la perception de la structure syllabique d'une part, et sur la fidélité de la répétition d'autre part.
Sujets	Japonais apprenant le français
Variables possibles :	<ul style="list-style-type: none"> <li>• Tâche (Comptage syllabique vs. Répétition)</li> <li>• Variété de français (« Nord » vs « Sud »)</li> </ul>
Input	<ul style="list-style-type: none"> <li>• Liste de mots en modalité auditive</li> </ul>
Output	<ul style="list-style-type: none"> <li>• Document de comptage syllabique<sup>214</sup></li> <li>• Enregistrement des productions en répétition</li> </ul>

L'objectif principal de ce type de test serait d'évaluer l'influence du « type » de français présenté aux sujets japonais sur leur traitement de la structure syllabique des items français présentant un « schwa ». Notre hypothèse est que la variété méridionale

---

<sup>214</sup> Cf. nos tests C, D et E.

du français serait davantage respectée par les sujets, tant d'un point de vue métalinguistique (comptage syllabique), que linguistique (répétition).

Dans le cas où notre hypothèse serait confirmée, on pourrait penser que le français méridional constitue, sur la question syllabique qui nous occupe, une variété plus facile à appréhender pour des apprenants débutants, tant d'un point de vue phonétique que phonologique. En outre, il permettrait peut-être de faciliter l'acquisition des correspondances phonographémiques. Cependant, il se pourrait que, à long terme, l'apprentissage d'une telle variété gêne ensuite les apprenants dans leur traitement du français « non-méridional ».

### **VII.3.2. Psycholinguistique**

Dans une perspective plus largement psycholinguistique, l'ensemble des tests envisageables est considérable. Nous proposons ci-dessous quelques pistes d'investigations, à titre d'exemple.

<b>Répétition vs. Lecture dans le cas du /R/ : Français/Japonais/Anglais/Espagnol</b>	
Objectif	Evaluer l'influence de la modalité sur la production (contrainte) du segment /R/ pour des publics de L1 différentes.
Sujets	<ul style="list-style-type: none"> <li>Apprenants japonais/ anglais/ espagnols de français</li> </ul>
Variables possibles :	<ul style="list-style-type: none"> <li>L1 (orthographe et phonologie) et rapport à la L2 (famille linguistique, etc.)</li> <li>Niveau d'expertise des sujets (Novice vs. Expert)</li> <li>Tâche (Répétition vs. Lecture)</li> <li>Input (Mots vs. Non-mots ; Voisinage orthographique/phonologique entre la L1 et la L2 ; etc.)</li> </ul>
Output	<ul style="list-style-type: none"> <li>Enregistrements des productions des apprenants.</li> </ul>

L'objectif principal de ce test serait d'évaluer le taux « d'erreurs » (par exemple dans le cas du segment /R/ en français) commises par les sujets selon la tâche (et ainsi la modalité). Chacune de ses langues possède en effet, dans son système orthographique, le graphème <r>, mais son associé phonético-phonologique est différent. Un tel test permettrait donc de croiser deux variables essentielles : la modalité et la L1 (avec tout ce que cela implique : système d'écriture, script, profondeur orthographique, inventaire phonologique segmental et famille linguistique). Notre hypothèse est que le taux d'erreur sera moindre en répétition (modalité auditive). Si notre hypothèse est vérifiée, cela aurait des implications didactiques quant à l'attention à accorder à l'enseignement du système phonético-phonologique d'une part, et phonographémique d'autre part, du français.

Il faudrait par ailleurs, si l'on choisit un input lexical, approfondir la distinction entre « Mots similaires (orthographiquement et/ou phonologiquement) en L1 et L2 » et « Mots nouveaux en L2 » : il se pourrait en effet que la ressemblance des mots en L2 (mots de même forme<sup>215</sup> ou mots d'emprunts) conduisent à davantage « d'erreurs » que les mots « nouveaux », mais que cette augmentation varie en fonction de la modalité/tâche. Deux variables seraient alors à croiser : la modalité et la ressemblance lexicale L1/L2 (voir à ce propos Ziegler, Perry, Jacobs et Braun, 2001).

---

<sup>215</sup> Ce que l'on nomme en anglais « cognate words ».

Plusieurs autres types de tests pourraient ainsi être envisagés, dont nous ne donnons ci-dessous que les très grandes lignes :

Thème	Test
Le graphème <r> en français, en espagnol et en anglais pour des Japonais	<ul style="list-style-type: none"> <li>• Enregistrer les productions d'apprenants japonais en tâche de répétition, d'une part et de lecture d'autre part, pour des stimuli français, espagnols et anglais.</li> <li>• Comparer les taux d'erreurs pour le segment /r/.</li> <li>• Interpréter les résultats vis-à-vis des différences de statut (phonologique et orthographique) du segment /r/ et du graphème &lt;r&gt; entre la L1 et les L2.</li> </ul>
Variations des systèmes d'écriture en L1	<ul style="list-style-type: none"> <li>• Enregistrer les productions d'apprenants de L1 différentes en répétition et en lecture à voix haute d'items français. Quatre groupes sont envisageables : 1) alphabet identique (romain) et orthographe profonde (par exemple anglais), 2) alphabet identique (romain) et orthographe (relativement) transparente (par exemple espagnol), 3) alphabet différent (par exemple cyrillique) et 4) système non-alphabétique (par exemple mandarin).</li> </ul>
Rapport segment/syllabe et modalité pour des Japonais	<ul style="list-style-type: none"> <li>• Enregistrer les productions d'apprenants japonais en tâche de répétition, d'une part et de lecture d'autre part, pour des stimuli français contenant : 1) le segment /R/ en position d'attaque simple ; 2) le segment /R/ en position d'attaque complexe (par exemple de type OBLI).</li> <li>• Comparer les taux de substitution segmentale, d'épenthèse et/ou d'effacement consonantique selon la modalité.</li> </ul>
Bilingues franco-japonais et Novices en lecture de non-mots	<ul style="list-style-type: none"> <li>• Enregistrer les productions de sujets experts et novices franco-japonais en tâche de répétition, d'une part et de lecture d'autre part, pour des non-mots présentés, d'abord comme français, puis comme japonais.</li> <li>• Comparer le taux « d'erreurs » selon la modalité.</li> </ul>

Enfin, si l'on souhaitait réellement évaluer de manière directe l'influence que peut avoir le système orthographique japonais sur l'apprentissage de la phonologie du français, il faudrait disposer de sujets japonais illettrés et les comparer à un groupe de sujets lettrés, apprenant le français dans les mêmes conditions. Or de tels sujets sont bien rares, sinon impossibles, à trouver, dans ce pays où le taux d'individus sachant lire et écrire figure parmi les plus élevés de la planète.

### **VII.3.3. Didactique et Apprentissage**

Si l'on souhaite traiter d'apprentissage à proprement parler, il faut alors considérer la question de la formation fournie aux sujets, et mener des études de type longitudinal, avec pré- et post-tests. On pourrait donc imaginer une série de manipulations relatives à l'instruction (avec un groupe-contrôle qui recevrait une instruction standard, c'est-à-dire une instruction reposant, à un degré ou à un autre, sur un support écrit mais de manière globalement et généralement non-contrôlée), et en évaluer l'effet sur des points linguistiques particuliers, pour des tâches bien spécifiées. Ici encore, nous ne faisons qu'esquisser quelques pistes de recherche.

<b>Thème</b>	<b>Test</b>
Orthographe artificielle et chevauchement graphémique français /japonais	<ul style="list-style-type: none"> <li>• Deux groupes d'apprenants japonais totalement débutants sont formés à l'acquisition du phonème /R/ à l'aide de deux codes graphiques distincts : l'un est le graphème &lt;r&gt; (chevauchement avec le japonais), l'autre est un symbole abstrait inconnu en japonais.</li> <li>• Enregistrer les productions en répétition et comparer les taux « d'erreurs ».</li> </ul>
Enseignement oral vs. écrit pour la production des groupes /CC/ pour des apprenants japonais	<ul style="list-style-type: none"> <li>• Deux groupes : 1) instruction uniquement orale ; 2) instruction orale avec support écrit.</li> <li>• Enregistrer les productions en tâche de répétition, de lecture et de production guidée.</li> </ul>



Enseignement oral vs. écrit pour les distinctions B/V et R/L pour des apprenants japonais	<ul style="list-style-type: none"><li>• Deux groupes : 1) instruction uniquement orale ; 2) instruction orale avec support écrit.</li><li>• Enregistrer les productions en tâche de répétition, de lecture et de production guidée.</li></ul>
---	---

On pourrait en effet penser que les représentations orthographiques peuvent aider à entraîner les sujets à opérer certaines discriminations phonologiques. Pourtant, il se pourrait également que, pour des sujets familiers de l'alphabet romain, comme le sont les Japonais, les habitudes de conversion phonographémiques des emprunts étrangers (par exemple les graphèmes <R> et <L> qui activent tous deux le phonème japonais /r/) puissent, au contraire, entraver l'apprentissage de la discrimination phonético-phonologique. Ce point est, nous semble-t-il, crucial, pour la question qui nous occupe : s'agit-il d'acquérir une nouvelle opposition phonologique à l'aide d'une opposition graphique, ou bien s'agit-il d'assigner deux valeurs phonético-phonologiques différentes à des signes graphiques qui jusque là étaient homophones ?

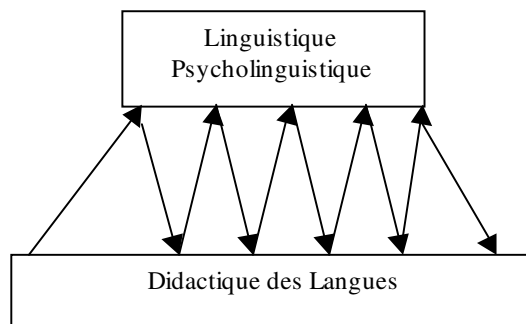
Comme cela apparaît donc à travers les quelques pistes d'études mentionnées jusqu'ici, les recherches sur lesquelles peut déboucher notre étude sont nombreuses et concernent à la fois la phonologie, la psycholinguistique et la didactique. Cependant, avant d'aller plus loin dans l'expérimentation psycholinguistique, il est d'ores et déjà possible de tirer certaines implications didactiques de nos résultats, implications qui nous ramènent à notre problématique centrale : celle de la multimodalité dans l'apprentissage du système phonético-phonologique du français langue étrangère.

## **Partie Didactique II :** **Implications pour l'enseignement de l'oral**

Après nous être plongés dans les domaines phonologique et psycholinguistique, afin de pouvoir mettre en place un protocole d'investigation susceptible d'apporter des réponses, au moins partielles, aux questions que nous avons formulées, il convient à présent de revenir à la perspective que nous n'avons cessé de maintenir dans notre approche, à savoir la perspective didactique. Il convient alors de prendre position, même brièvement, sur un point toujours discuté dans le domaine des sciences du langage, celui des relations, de contenus, de méthode, mais aussi de statut, formel ou académique, entre chacun des trois champs que nous invoquons dans notre étude : la linguistique, la psycholinguistique et la didactique des langues. Ce point, déjà évoqué plus haut, a son importance dans la formulation de l'intitulé de cette partie : doit-on en effet traiter ici d'*implications* ou d'*applications* didactiques ?

Le choix du terme *implications* reflète bien notre démarche. Si l'on ne peut effectivement conduire une recherche sur l'enseignement des langues sans avoir une connaissance approfondie de ce que l'on se propose d'enseigner (domaine linguistique, dans notre cas phonético-phonologique, pour les aspects strictement linguistiques) et de la manière dont ce contenu est appris (domaine psycholinguistique portant sur l'apprentissage), ces deux types de savoirs doivent, s'ils veulent effectivement pouvoir être réinvestis en didactique, être initialement et en permanence orientés dans une perspective didactique, tel que nous l'avons fait tout au long de notre étude, afin d'être perpétuellement confrontés à la réalité des situations d'enseignement / apprentissage et à leur variabilité. Il s'agit, comme toujours, de trouver la ligne d'équilibre entre données empiriques et savoirs théoriques, une ligne qui exclut toute idée d'applicationnisme.

Ce n'est pas tant l'applicationnisme en tant qu'idée qu'il convient de rejeter ici : on pourrait alors s'interroger sur les raisons d'un tel rejet. En revanche, ce sont les applications qui posent problème, puisqu'elles posent un problème de transposition des savoirs, problème que l'on retrouve tant à l'échelle individuelle qu'à l'échelle disciplinaire. Ces savoirs, s'ils sont bien évidemment généralement bâtis sur des données empiriques (*via* la constitution de corpus, de tests linguistiques, d'enquêtes sociolinguistiques, d'expérimentations (neuro-) psycholinguistiques de divers types, etc.), ne permettent cependant pas toujours à ceux qui souhaiteraient les « appliquer » de le faire avec succès, en raison de leur constitution même, dont l'orientation initiale n'était de toute évidence pas particulièrement didactique. Il nous semble donc nécessaire de figurer en termes d' « allers-retours » les échanges entre, d'un côté une didactique des langues qui, d'une part se voudrait scientifiquement fondée, et d'autre part désirerait davantage intégrer les données sur l'apprentissage à son analyse de l'enseignement, et de l'autre des disciplines connexes telles que la psycholinguistique (et donc la linguistique, puisque l'on ne peut faire de psycholinguistique sans une solide base linguistique), dont les avancées pourraient être profitables à la didactique, à condition d'être suffisamment orientées dès la mise en œuvre de leur constitution. Notre parcours pourrait donc s'illustrer de la manière suivante :



Il s'agirait donc finalement moins d'un problème disciplinaire que d'un problème humain : les linguistes et les psycholinguistes n'ont, *a priori* et généralement, ni les mêmes préoccupations, ni les mêmes connaissances de terrain que les didacticiens,

lesquels ont leurs propres limitations. Interroger la (psycho-)linguistique ne suffit donc pas : seule la mise en œuvre d'investigations par les didacticiens mêmes peut conduire à la récolte de réponses suffisamment adéquates, car constamment orientée dans une perspective didactique. C'est donc d'un problème de *formation* qu'il s'agit : les didacticiens, s'ils ne peuvent et ne doivent également être linguistes et psycholinguistes, doivent en revanche être foncièrement et *rigoureusement* interdisciplinaires, précisément en vue de fonder la didactique. Nous sommes donc loin de toute conception de « linguistique appliquée », et considérons en revanche avec grand intérêt les *implications* didactiques que nos investigations linguistiques et psycholinguistiques à *visée didactique* pourraient avoir pour l'enseignement de la phonétique et de la phonologie, et plus généralement de la langue, françaises.

Dans un premier temps (I), nous présentons donc, de manière non exhaustive, les principales *implications spécifiques* de nos tests pour l'enseignement de la prononciation du français aux apprenants universitaires japonais : d'abord (I.1.) dans une perspective segmentale (tests A et F), vis-à-vis des segments /b, v/ et /R, l/, en considérant les tâches d'identification (I.1.1.) et de discrimination (I.1.2.) phonémique, ainsi que la question des correspondances phonographémiques (I.1.3.) ; puis (I.2.) dans une perspective syllabique, vis-à-vis des groupes biconsonantiques, en considérant d'une part la tâche de comptage syllabique (tests C1 et C2), et d'autre part le rôle de la modalité (auditive *vs.* orthographique) (tests C2, D et E) dans l'apprentissage de la structure syllabique du français.

Dans un deuxième temps (II), nous reprenons l'ensemble de notre problématique initiale, afin de développer les *implications générales* de notre étude, basées à la fois sur les résultats de nos tests, mais également sur les données disponibles dans la littérature concernée et sur nos propres pratiques pédagogiques. Deux thèmes principaux émergent alors :

- Celui de la fonction de la modalité visuelle pour l'apprentissage de la prononciation, et plus généralement de l'oral, en L2 (II.1), dans lequel nous examinons, d'une part la fonction de l'écrit (II.1.1.), et d'autre part la fonction d'autres informations visuelles (II.1.2.), de nature notamment oro-faciale (II.1.2.1.), gestuelle (II.1.2.2.), picturale (II.1.2.3.) et graphique (II.1.2.4.), en soulignant que la fonction de l'écrit ne peut être définie que par rapport à certaines tâches didactiques précises (II.1.3.) ;
- Celui du type (explicite ou implicite) d'enseignement de l'oral (II.2.), qui nécessite en premier lieu d'examiner la nature (implicite ou explicite) de l'apprentissage (II.2.1.), avant de s'interroger sur la fonction de l'enseignement explicite (II.2.2.), vis-à-vis de l'activité métalinguistique des apprenants (II.2.3.), pour l'apprentissage de la prononciation et de l'oral en L2 (II.2.4.).

L'examen de ces deux thèmes permet alors d'amorcer certaines considérations didactiques méthodologiques générales (III), ayant trait notamment :

- Au rôle de l'attention sur la forme (III.1.), tant à un niveau cognitif (III.1.1.) que métacognitif (III.1.2.) ;
- Aux relations entre sciences du langage et didactique des langues (III.2.), en particulier au sujet de la transposition didactique des modèles et des données de la linguistique (III.2.1.), mais également au sujet de l'utilisation des corpus en classe de langue (III.2.2.) ;
- A la conception de la norme didactique (III.3.), de manière générale d'abord (III.3.1.), puis en tenant compte de la dimension sociolinguistique de la L2 et des caractéristiques psycholinguistiques des apprenants (III.3.2.), notamment dans le cas de la « prononciation » (III.3.3.).

Ces considérations nous permettent finalement d'ouvrir certaines perspectives méthodologiques (III.4.), en insistant sur la nécessité d'intégrer l'articulation oral/écrit à une méthodologie globale d'enseignement (III.4.1.), en vue de proposer un parcours d'apprentissage de l'oral (III.4.2.) dont les pratiques (III.4.3.) devraient être

cohérentes entre elles, une cohérence qui devrait notamment s'exprimer à travers l'indispensable complémentarité, nous semble-t-il, entre correction phonético-phonologique et correction phonographémique (III.4.4.). Ces perspectives indiquent à quel point les questions soulevées dans notre étude mériteraient d'être approfondies, et c'est donc sur une ouverture (III.4.5.) que nous achevons cette partie, ouverture sur la recherche (fondamentale et appliquée), sur la conception d'outils et de pratiques didactiques et surtout sur une réflexion didactique méthodologique à poursuivre, réflexion théoriquement informée mais véritablement ancrée dans le terrain des classes de langue.



## Table des Matières Partielle

I. IMPLICATIONS SPECIFIQUES .....	- 1029 -
<i>I.1. Axe segmental : tâches A (identification) et F (discrimination)</i> .....	- 1032 -
I.1.1. Tâche A (identification) .....	- 1034 -
I.1.2. Tâche F (discrimination) .....	- 1035 -
I.1.3. Tâches A et F .....	- 1038 -
<i>I.2 Axe syllabique : tâches C1, C2, D et E (comptage)</i> .....	- 1038 -
II. IMPLICATIONS GENERALES .....	- 1041 -
<i>II.1. Prononciation et modalité visuelle</i> .....	- 1042 -
II.1.1. Fonctions de l'écrit pour l'apprentissage de l'oral en classe de langue : fixation et segmentation de la parole .....	- 1044 -
II.1.2. Caractéristiques phonético-phonologiques et input visuel .....	- 1049 -
II.1.2.1. Préliminaires .....	- 1049 -
II.1.2.1.1. Systèmes d'écriture en L1 et homographes hétérophones L1/L2 .....	- 1051 -
II.1.2.1.2. Perception et mémorisation de l'oral .....	- 1057 -
II.1.2.1.3. Correction phonétique .....	- 1058 -
II.1.2.2. Information oro-faciale .....	- 1060 -
II.1.2.3. Information gestuelle non-faciale .....	- 1063 -
II.1.2.3.1. Axe segmental .....	- 1064 -
II.1.2.3.2. Axe rythmique/syllabique .....	- 1067 -
II.1.2.3.3. Axe accentuel/intonatif .....	- 1068 -
II.1.2.3.4. Compréhension et mémorisation .....	- 1068 -
II.1.2.4. Information picturale .....	- 1070 -
II.1.2.5. Information graphique .....	- 1080 -
II.1.2.5.1. Caractéristiques prosodiques .....	- 1080 -
II.1.2.5.2. Segmentation de la parole .....	- 1082 -
II.1.2.5.3. Code graphique et modifications orthographiques .....	- 1085 -
II.1.2.5.3.1. Elaboration d'un code graphique pour l'apprentissage de l'oral .....	- 1088 -
II.1.2.5.3.2. Le système Alfonic .....	- 1089 -
II.1.2.5.3.3. Conscience phonémique et correspondances phonographémiques .....	- 1091 -
II.1.2.5.3.4. L'Alphabet Phonétique International (API) .....	- 1093 -
II.1.3. Tâches et utilisation du support visuo-graphique .....	- 1096 -
II.1.3.1. Lecture vs. Transcription .....	- 1097 -
II.1.3.2. Utilisation passive vs. Utilisation active .....	- 1099 -
<i>II.2. Enseignement explicite ou implicite ?</i> .....	- 1102 -
II.2.1. Apprentissage implicite ou explicite ? .....	- 1105 -
II.2.2. L'enseignement explicite est-il bénéfique au développement des savoirs implicites ? .....	- 1109 -
II.2.3. Apprentissage explicite et réflexion métalinguistique .....	- 1116 -
II.2.4. Enseignement explicite et « prononciation » .....	- 1125 -
III. CONSIDERATIONS METHODOLOGIQUES : ATTENTION SUR LA FORME, OUTILS LINGUISTIQUES ET NORME DIDACTIQUE .....	- 1135 -
<i>III.1. Attention sur la forme et approches didactiques</i> .....	- 1135 -
III.1.1. Attention : niveau cognitif .....	- 1140 -
III.1.2. Attention : niveau métacognitif .....	- 1147 -
<i>III.2. Modèles linguistiques et corpus</i> .....	- 1150 -



III.2.1. Modèles linguistiques et modèles didactiques.....	- 1150 -
III.2.2. Corpus oraux .....	- 1166 -
<i>III.3. Norme didactique.....</i>	<i>- 1176 -</i>
III.3.1. Description vs. Prescription vs. Adaptation .....	- 1177 -
III.3.2. De la sociolinguistique à la psycholinguistique <i>pour la didactique</i> .....	- 1180 -
III.3.3. Le cas de la « prononciation » .....	- 1185 -
<i>III.4. Perspectives didactiques méthodologiques.....</i>	<i>- 1193 -</i>
III.4.1. Intégration méthodologique de l'articulation oral/écrit.....	- 1193 -
III.4.1.1. Module méthodologique.....	- 1200 -
III.4.1.2. Module explicite.....	- 1201 -
III.4.1.3. Module implicite .....	- 1202 -
III.4.2. Parcours d'apprentissage de l'oral.....	- 1203 -
III.4.2.1. Utilisation des « méthodes » de FLE.....	- 1206 -
III.4.2.2. Exercices, Lecture et API .....	- 1210 -
III.4.2.2.1. Les « exercices » .....	- 1211 -
III.4.2.2.2. La lecture à voix haute .....	- 1211 -
III.4.2.2.3. L'utilisation de l'Alphabet Phonétique International (API) .....	- 1212 -
III.4.2.3. Correction phonétique et correction phonographémique .....	- 1213 -
III.4.2.3.1. Utilisation du support visuo-graphique lors de la correction phonétique.....	- 1213 -
III.4.2.3.1.1. Identification/discrimination.....	- 1214 -
III.4.2.3.1.2. Segmentation .....	- 1217 -
III.4.2.3.2. Correction phonographémique en production (dictée) .....	- 1218 -
III.4.2.3.3. Correction phonographémique en perception (lecture à voix haute).....	- 1219 -
III.4.5. Ouvertures .....	- 1223 -

## **I. Implications spécifiques**

Avant de reprendre une perspective didactique plus générale concernant l'enseignement de l'oral, issue de l'ensemble de notre étude, il nous faut mentionner dans un premier temps les implications spécifiques des résultats de nos expérimentations. Ces implications spécifiques nous semblent mériter une certaine attention, puisqu'elles constituent une illustration claire de la manière dont certaines études très ciblées, comme celle que nous avons conduite, peuvent néanmoins conduire à des recommandations de conception et de pratiques didactiques « de terrain ». Plus encore, elles appuient l'argument présenté précédemment, selon lequel de telles études doivent effectivement être menées par des didacticiens, et non des spécialistes d'autres disciplines. Trois raisons, en particulier, justifient une telle affirmation :

- 1) La nécessité d'une perspective didactique initiale et finale : c'est celle que nous avons assignée à l'ensemble de notre étude, notamment lors de l'élaboration des hypothèses et du protocole de recherche, lesquels l'ont précisément été sur la base d'interrogations didactiques, aussi bien « théoriques », du point de vue de la recherche dans le domaine de l'acquisition des langues secondes, que « pratiques », issues de notre travail d'enseignement du FLE et de l'utilisation des matériels pédagogiques disponibles, tant en France qu'à l'étranger. Sans cette perspective et cette formation, initiales et continues, il est fort probable que la conception même de notre démarche de recherche n'aurait pu aboutir sur de telles implications spécifiques. Cela rejoint alors le point suivant.
- 2) La nécessité d'une connaissance de l'enseignement de terrain : sans celle-ci, les *implications* didactiques d'études quasi-expérimentales de ce type ne peuvent que rester dans l'ombre, puisque les non-didacticiens n'ont généralement que peu, voire souvent pas, de connaissance concernant d'une part les pratiques *concrètes*, contemporaines et passées, des enseignants de langue, et d'autre part les problèmes et interrogations rencontrés quotidiennement dans les salles de classes, classes sans

lesquelles la didactique des langues n'existerait sans doute pas. Un tel argument éclaire alors parfaitement la réflexion que nous avons menée jusqu'à présent sur le rapport entre applicationnisme et implicationnisme : étant donné leur relative ignorance tout a fait légitime du domaine, et plus précisément du terrain, les non-didacticiens ne peuvent, par définition, qu'envisager des *applications* de leurs propres études et préoccupations au champ de la didactique, et l'on ne peut le leur reprocher puisqu'ils ne sont précisément pas didacticiens, mais linguistes, psychologues ou sociologues par exemple. Seul le didacticien possédant à la fois une connaissance du champ théorique, en synchronie et en diachronie, ainsi qu'une expérience personnelle suffisante de l'enseignement de terrain, peut établir les connexions entre ces connaissances et les résultats d'études comme la nôtre, à condition que celles-ci y soient suffisamment adaptées (*cf.* point précédent). Si l'on considère que, pendant longtemps, la didactique des langues, en tant que discipline émergente, a « manqué de corps », notamment dans le domaine de la recherche de type expérimentaliste, on peut alors comprendre que celle-ci ait dû subir le « joug » de la *linguistique appliquée*, joug dont la responsabilité semble incomber tant aux linguistes (qui ne sont pas didacticiens et qui ne peuvent donc proposer que des applications et non des implications) qu'aux didacticiens qui, sans paradigme de recherche suffisamment établi, se tournaient vers eux, puisque la didactique était encore en construction, tout comme les sciences cognitives. Cette « linguistique appliquée » semblait donc inévitable, et sa remise en question ne remet en revanche absolument pas en doute, nous semble-t-il, l'importance des sciences du langage pour la didactique et plus précisément pour la *formation*, initiale et continue, des didacticiens. Cependant, depuis les vingt dernières années, les cursus universitaires en didactique des langues, d'une part, l'évolution des sciences cognitives d'autre part, et enfin la reconnaissance de la didactique des langues en tant que domaine autonome, quoique pluridisciplinaire dans une certaine mesure, permettent bel et bien d'envisager la constitution d'un domaine de recherche didactique capable de relier, de manière graduelle, connaissances théoriques en sciences cognitives et pratiques de classes, *via* plusieurs « palier » interprétatifs intermédiaires. Seuls les didacticiens, praticiens de l'enseignement des langues, nous semblent à même de

pouvoir retirer de véritables implications, spécifiques ou générales, d'études comme la nôtre. Il ne pourrait donc, dans cette perspective, y avoir de « linguistique impliquée » (sauf cas rares), mais plutôt une approche implicationniste de la didactique des langues, qui ne peut être menée que par des *didacticiens impliqués* dans la recherche, suffisamment formés, ou informés, en sciences du langage et sciences cognitives, pour pouvoir suivre leur propre paradigme de recherche, adapté, autant que possible, tant à la démarche scientifique contemporaine<sup>216</sup> (du côté des sciences cognitives) qu'aux réalités des classes de langue (du côté de la didactique). L'éclatement apparent de recherches généralement considérées trop « techniques », trop « limitées » ou encore trop « abstraites » par certains didacticiens, devrait ainsi être reconsidéré par les chercheurs du domaine, ces dernières pouvant être réassemblées de manière cohérente, à condition de l'être par des *didacticiens*. Cela nous conduit alors à notre troisième point.

- 3) L'interprétabilité des hypothèses et des résultats : en effet, afin de pouvoir « transposer » les interrogations didactiques en hypothèses de recherches précises (Juff, 2002), et pouvoir ensuite tirer des implications *effectivement* didactiques des résultats obtenus, il semblerait aujourd'hui qu'il faille correspondre au profil de didacticien impliqué que nous venons d'esquisser. Sans cela, les possibles implications risqueraient de rester sensiblement hors du champ didactique. Cela s'illustre, par exemple, dans le cas de certaines études en interphonologie, dont les auteurs, en vue de donner une valeur « appliquée » à leur travaux, à propos de la perception/production des segments en langue étrangère, et ainsi, plus généralement, de l'apprentissage de « la prononciation en LE », rédigent un court paragraphe, à la fin de leur article, pour conclure que les apprenants doivent

---

<sup>216</sup> Nous simplifions ici grossièrement la problématique: il faudrait en effet définir ce qu'est, ce qu'a pu être, ou ce que devrait être « la démarche scientifique » (et donc établir une réflexion épistémologique minimale sur la notion de « science » et sur la valeur heuristique des démarches actuelles, au moins en sciences cognitives), ce qui nous entraînerait évidemment bien trop loin. Il faudrait en outre émettre quelques réserves quant à l'utilisation parfois hâtive, ou naïve, de l'adjectif « scientifique » dans les travaux de recherche, en particulier vis-à-vis de l'articulation entre sciences dites « humaines » et sciences dites « pures » (Sokal et Bricmont, 1997).

s'entraîner à mieux percevoir le formant F2 pour tel segment donné (approche acoustique) ou encore à reclasser les contraintes de Fidélité plus haut que les contraintes de Marque dans leur système interphonologique (cadre de la théorie de l'Optimalité). De telles « conclusions » risquent fort de laisser passablement insensibles les enseignants de langue qui cherchent à résoudre, au quotidien, les problèmes de prononciation de leurs, étudiants...

La recherche en didactique des langues a donc une double vocation : utile aux sciences cognitives *via* les données qu'elle leur fournit, elle doit, tout en alimentant ses propres problématiques théoriques, être fondamentalement orientée, nous semble-t-il, vers les pratiques de classe et la conception de matériel et de démarches d'enseignement, afin de pouvoir y être effectivement réinvestie par les enseignants et leurs institutions.

Nous pouvons ainsi à présent mentionner les quelques implications spécifiques de nos propres tests, qui, bien que modestes, répondent à nos objectifs initiaux.

### **1.1. Axe segmental : tâches A (identification) et F (discrimination)**

Les tâches A et F renvoient, dans une certaine mesure, à deux tâches didactiques particulières, que l'on trouve dans de nombreux manuels d'apprentissage du FLE lorsque ceux-ci consacrent explicitement une certaine partie de leurs activités d'apprentissage à celui de la « prononciation » : il s'agit des tâches d'identification et de discrimination phonémique. Lorsque l'on effectue une comparaison des systèmes phonémiques français et japonais, on constate que les phonèmes français /v/, /R/ et /l/ sont absents du répertoire segmental japonais. A partir d'une telle analyse contrastive « classique », l'enseignant pourrait considérer que ces trois phonèmes doivent faire l'objet d'un travail d'apprentissage particulier de la part d'apprenants japonais de français.

Cependant, comme nous l'avons vu, une telle analyse est insuffisante, puisqu'elle ne prend en compte ni l'opposition /b/ vs. /v/, ni l'existence de l'allophone [β] en japonais,

ni la similarité phonétique entre le [l] et le [r] japonais, ni la variation allophonique en français, notamment pour le /R/, fonction de la position du segment et de son entourage vocalique. Bien sûr, il apparaît qu'une analyse phonético-phonologique plus approfondie, prenant en compte la problématique de la marque et la distinction entre catégories phonologiques et phonétiques, permet de mieux anticiper, dans une certaine mesure les difficultés d'apprentissage. Néanmoins, seule l'investigation psycholinguistique, qui permet notamment d'inclure la dimension phonographémique, offre la possibilité d'évaluer, avec une relative précision, les difficultés d'apprentissage probables que pourrait rencontrer une majorité d'apprenants. Car, comme le note Gaonac'h (1988, p. 84) :

*« On ne peut rendre compte de l'existence d'interférences par la seule analyse de la proximité formelle entre les deux langues : les relations entre systèmes linguistiques n'ont pas de valeur absolue, ne "préexistent" pas aux activités mises en œuvre par l'apprenant ».*

Il ne s'agit donc plus seulement d'effectuer des études comparatives linguistiques portant sur les structures, mais également d'études psycholinguistiques portant sur les processus.

De telles investigations visent à cibler les problèmes, en vue de minimiser le coût, notamment temporel, de l'apprentissage : on pourrait en effet se demander s'il est bien utile de consacrer 30 minutes à l'apprentissage du phonème /l/ pour des apprenants japonais. Si oui, et s'il s'agit de permettre l'émergence d'une nouvelle catégorie phonologique, doit-on conduire ce travail plutôt à l'aide d'une tâche d'identification ou de discrimination ? Doit-on proposer des items contenant ce segment de préférence en position initiale ou finale, et dans quel environnement vocalique ?

Comme on le voit à travers cette série de questions non hiérarchisées, les interrogations sont nombreuses et variées. Il va de soi que, dans un cadre didactique réaliste,

l'apprentissage de la « prononciation » du français se réalise de multiples manières, à travers l'ensemble des tâches d'apprentissage proposées par l'enseignant, que ce soit avec ou sans focalisation sur « la prononciation ». Cependant, si l'on souhaite élaborer un programme d'enseignement raisonné, comme cela peut être le cas lors de la conception de matériel pédagogique, en tenant compte des difficultés d'apprentissage pour un public ciblé, alors, dans le cas de la prononciation, de telles interrogations ne sont pas triviales. Il en va de même lorsque l'enseignant décide de mettre en oeuvre certaines activités complémentaires visant à pallier certaines difficultés de « prononciation » de ses étudiants. Les tâches A et F permettent, pour des apprenants japonais et pour les segments considérés, de répondre en partie à ces questions<sup>217</sup>.

Si l'on estime que les étudiants doivent consacrer davantage de temps et d'efforts à l'apprentissage des éléments qui leur posent le plus de difficultés, alors nous parvenons aux conclusions suivantes.

### **I.1.1. Tâche A (identification)**

Les résultats de la tâche A suggèrent que les étudiants japonais, en raison de la difficulté particulière que semblent présenter ces sous-tâches, doivent être prioritairement entraînés à :

- Position :**
- Identifier le segment /b/ en position finale (plutôt qu'en position initiale et médiale).
  - Identifier le segment /R/ en positions initiale et médiale (plutôt qu'en position finale).
  - Identifier les segments /v/ et /l/ en position médiale (plutôt qu'en positions initiale et finale).

---

<sup>217</sup> Nous n'indiquons ci-dessous que les grandes tendances observées dans nos résultats.

**Consonne :** • Identifier le segment /b/ (plutôt que /v, R, l/), en opposition à /v/.

**Environnement vocalique :** • Identifier les quatre segments considérés avec pour environnement vocalique la voyelle /u/ (plutôt que /ø/).  
• Identifier les segments /b/, /v/ et /l/ avec pour environnement la voyelle /a/ (plutôt que /i/).  
• Identifier le segment /R/ avec pour environnement la voyelle /i/ (plutôt que /a/).

De tels résultats soulignent l'importance de l'investigation quasi-expérimentale, puisque la consigne concernant l'identification de /b/ (en opposition à /v/) pourrait peut-être sembler contre-intuitive : le segment /b/ figurant dans le répertoire japonais, l'enseignant pourrait en effet négliger ce segment, négligeant ainsi l'opposition entre /b/ et /v/<sup>218</sup>.

### **I.1.2. Tâche F (discrimination)**

Les résultats de la tâche F suggèrent que les étudiants japonais, en raison de la difficulté particulière que semblent présenter ces sous-tâches, doivent être particulièrement entraînés à :

---

<sup>218</sup> Cela évoque la conclusion de Baqué (2005, p. 10) à propos du traitement du /ə/ et du /e/ par des apprenants hispano-catalophones de français:

*« Cette description [...] nous a permis de vérifier l'hypothèse communément admise, selon laquelle ces étudiants éprouvent des difficultés pour intégrer le « e muet » du français. Néanmoins, il est moins fréquent de s'intéresser à la prononciation du /e/ de ces apprenants. Or les résultats que nous avons obtenus indiquent que ce phonème pose également des problèmes, et qu'il est réalisé avec un F1 généralement trop haut et un F2 trop bas. Cela entraîne non seulement des productions peu conformes à celles des francophones, ou même la production de /ə/ comme des [e], mais surtout une centralisation des réalisations et, partant, une hypo-différenciation des deux phonèmes entre eux. Il en résulte que toute action pédagogique de correction phonétique devrait s'intéresser à la fois aux deux phonèmes, et, sans doute, impliquerait de corriger dans un premier temps le /e/ pour n'introduire qu'ensuite le /ə/ ».*



- Position :**
- Discriminer /b/ et /v/ selon différentes positions (plutôt que /l/ et /R/), et en particulier en position de coda finale.
- Consonne :**
- Discriminer /b/ et /v/ (plutôt que /l/ et /R/).
- Environnement vocalique :**
- Discriminer le segment /b/ de /v/ avec pour environnement les voyelles /a/ et /ø/ (plutôt que /i/ et /u/).
  - Discriminer le segment /v/ de /b/ avec pour environnement les voyelles /a/ et /u/ (plutôt que /i/ et /ø/).
  - Discriminer le segment /R/ de /l/ avec pour environnement la voyelle /i/ (plutôt que /a/, /ø/ et /u/).
  - Discriminer le segment /l/ de /R/ avec pour environnement les voyelles /i/, /ø/ et /u/ (plutôt que /a/).

Ici encore, ces résultats permettent d'indiquer certaines priorités d'apprentissage, *a priori* non évidentes : les couples /b, v/ et /l, R/ sont en effet des couples « classiquement difficiles » pour les apprenants japonais. Notre étude permet d'indiquer que, en perception au moins<sup>219</sup>, le couple /b, v/ doit être une cible d'apprentissage prioritaire par rapport au couple /l, R/ qui semble entraîner moins d'erreurs.

De telles directives permettent ainsi d'opérer, en termes de priorités d'enseignement, une *sélection*, voire une *progression, de contenus* pour l'enseignement de la prononciation du français, du moins lorsque cet entraînement phonético-phonologique est effectué avec des non-mots. Une telle démarche permet ainsi d'avancer un programme d'apprentissage véritablement *centré* sur les apprenants (Detey, 2003b),

---

<sup>219</sup> Il pourrait en effet en être autrement en production, bien que les deux aspects soient fortement interreliés.

avec comme critère directeur celui de la difficulté d'apprentissage, tel qu'établi auprès d'un échantillon réel d'étudiants, et non sur la base d'intuitions ou d'hypothèses théoriques. Croisé avec d'autres critères, comme par exemple celui de la fréquence d'occurrence des segments considérés en français, une telle analyse pourrait ainsi *a priori* permettre d'augmenter la rentabilité de l'enseignement et de l'apprentissage, en ciblant plus précisément les objectifs d'apprentissage établis par l'enseignant. Cela rejoint les critères mis en avant par Neri, Cucchiari et Strik (2002, p. 186) pour la sélection des erreurs sur lesquelles devraient être centrés les didacticiens d'entraînement à la prononciation, dans le cas d'apprenants de néerlandais comme L2 :

*« The selection of the errors to be addressed will be based on the following criteria : 1) frequency among learners of Dutch as a L2, 2) persistence, 3) perceptual importance for native speakers [...] ».*

Un tel programme doit en effet également prendre en compte la variation phonologique en L1, ainsi que le critère d'intelligibilité : d'un point de vue communicatif, les substitutions /b, v/ en français par des apprenants japonais sont-elles plus « dommageables » pour l'intercompréhension que les substitutions /l, R/ ? Comme le soulignent Neri, Cucchiari et Strik (2002, p. 180) :

*« It is fundamental for teachers and researchers alike to distinguish between two different dimensions of nonnative pronunciation : accentedness and intelligibility. These notions are related but fairly independent : a strong foreign accent does not always hinder intelligibility of speech and specific types of instruction do not necessarily lead to improvement of both these aspects [...] ».*

Plusieurs questions se posent donc ensuite, notamment axées sur la distinction, relative, entre perception et production. Néanmoins, le critère de difficulté d'apprentissage tel que représenté dans notre étude permet de fournir un premier axe d'orientation essentiel au curriculum didactique.

### **I.1.3. Tâches A et F**

Nous avons justifié précédemment le fait que les résultats de ces deux tâches puissent être interprétés en termes de perception erronée, que ce soit à un niveau phonétique ou phonologique. Cependant, il convient de rappeler que le protocole de ces deux tâches impliquait un appariement phonographémique. Quelle que soit l'interprétation que l'on donne des résultats, il apparaît que les étudiants ont commis des erreurs d'appariement de ce type : l'entraînement qui doit être donné aux apprenants, et que nous venons de décrire, ne doit donc pas être seulement phonétique mais également phonographémique. Les tâches de production sous dictée (dans la perspective qui est la nôtre) et/ou de lecture à voix haute doivent donc être proposées par les enseignants aux sujets concernés, en suivant les consignes précédemment indiquées. La comparaison des deux tâches indique également que les apprenants devraient davantage être entraînés à l'appariement phonographémique en tâche d'identification qu'en tâche de discrimination.

Dans tous les cas, ces résultats indiquent que les enseignants ne peuvent tenir pour acquise la connaissance procédurale, voire déclarative, des relations phonographémiques du français par les apprenants et que l'enseignement du système phonologique du français nécessite de prendre en compte celui du système phonographémique.

### **I.2 Axe syllabique : tâches C1, C2, D et E (comptage)**

Les tâches C1, C2, D et E, bien qu'elles puissent sembler assez éloignées du quotidien des classes de langue (quoique certains manuels proposent occasionnellement des tâches de comptage syllabique (Silveira, 2002, p. 109)), font appel à la conscience phonologique des apprenants, conscience impliquée dans l'apprentissage à la fois de la prononciation et de la lecture en français. Plus particulièrement, à travers ces tâches, c'est la maîtrise de la syllabation du français et la capacité de traitement des groupes consonantiques, relativement nombreux dans cette langue, qui sont évaluées. En outre, les trois modalités proposées pour les stimuli (acoustique, orthographique et audio-

orthographique) renvoient aux tâches de répétition, de lecture et d'oralisation mixte (répétition avec support orthographique et/ou lecture avec support acoustique). Les résultats de ces trois tâches soulignent alors, pour le public considéré :

- La nécessité de consacrer un certain travail à l'apprentissage de la structure syllabique du français (l'importance de la syllabe en français).
- La nécessité d'effectuer ce travail selon deux types de modalité, le transfert des compétences n'étant pas « automatique » (syllabation de l'oral et de l'écrit).
- La difficulté de la position de coda finale (qui doit donc recevoir une attention didactique particulière).
- La difficulté, hiérarchisée, des familles phonologiques suivantes (qui doivent donc recevoir une attention didactique particulière, fonction de cette hiérarchie) :  
Sib + Stop > Sib + Nonstop > Obs + /l/ > Obs + /R/.

Plus particulièrement, les tâches C1 et C2 permettent d'obtenir un classement de difficulté des groupes consonantiques, suggérant ainsi une progression d'enseignement (du plus difficile au moins difficile, ou vice-versa, selon les options méthodologiques de l'enseignant et selon les autres critères de sélection adoptés (par exemple la fréquence, ou encore l'intelligibilité<sup>220</sup>) :

---

<sup>220</sup> Le phénomène de dévoisement vocalique en japonais peut en effet compenser, selon le critère d'intelligibilité, la « difficulté » de certains groupes (par exemple /s/ + Plosive sourde), illustrant alors la distinction entre apprentissage phonétique et phonologique. Ici encore, c'est à l'enseignant d'assigner les objectifs, phonétiques ou phonologiques, de l'apprentissage, même s'il doit prendre garde à ce qu'une bonne performance phonétique en tâche de répétition par exemple, ne puisse dissimuler une déficience phonologique qui apparaîtrait en tâche de dictée par exemple, et qui pourrait se révéler dommageable pour l'apprentissage dans d'autres tâches.

<b>Plus difficile</b>	<b>1</b>	<ul style="list-style-type: none"> <li>• Sib+Stop</li> <li>• Obs+Cont</li> <li>• Obs+Nas</li> <li>• Obs+Stop</li> </ul>
	<b>2</b>	<ul style="list-style-type: none"> <li>• Sib+Nas</li> <li>• Sib+Lat</li> <li>• Stop+Lat</li> </ul>
	<b>3</b>	<ul style="list-style-type: none"> <li>• Obs+Sib</li> <li>• Sib+Cont</li> <li>• Cont+Lat</li> </ul>
	<b>4</b>	<ul style="list-style-type: none"> <li>• Stop+Rhot</li> <li>• Cont+Rhot</li> </ul>
<b>Moins difficile</b>		

La tâche D, quant à elle, souligne la différence de traitement syllabique, selon que les stimuli sont présentés de manière acoustique ou orthographique. Elle souligne en particulier la nécessité, pour ces apprenants, de s'entraîner à la lecture (à voix haute) tout en respectant la structure syllabique du français.

La tâche E, enfin, permet de donner un certain éclairage sur la bimodalité audio-orthographique en classe de FLE pour le public considéré : il apparaît en effet que, si l'objectif est de respecter (et *a fortiori* d'apprendre) la structure syllabique du français, alors la présentation bimodale devrait être abandonnée au profit d'une présentation unimodale acoustique. Le support orthographique ne faciliterait pas la répétition orale pour l'objectif donné. En revanche, si l'objectif est de respecter (et *a fortiori* d'apprendre) la structure syllabique du français telle qu'elle est représentée sous sa forme orthographique, alors la présentation bimodale devrait être préférée à une présentation unimodale orthographique. Le support acoustique faciliterait la lecture à voix haute pour l'objectif donné.

La conclusion est ici encore l'importance, pourtant souvent négligée, de l'apprentissage de la lecture (à voix haute) pour les étudiants considérés, en vue de maîtriser le système phonographémique français, utile à l'apprentissage du système phonético-phonologique, ainsi que le « danger » potentiel des représentations orthographiques en tant que support didactique lorsque l'objectif est l'acquisition de l'oral. Une telle conclusion ne conduit certainement pas au « rejet » de l'orthographe (qui serait tout à fait irréaliste d'un point de vue tant didactique que communicatif), mais au contraire à la nécessité d'accorder davantage d'attention à l'agencement spatio-temporel entre supports oraux et écrits, ainsi qu'à l'apprentissage des correspondances entre les deux types de formes.

Ces implications spécifiques pourront certes être considérées comme assez limitées. Néanmoins, « la science ne se bâtit pas en un jour » dit-on parfois, et ce n'est, nous semble-t-il, qu'à travers de telles investigations, laborieuses peut-être, mais relativement fines et rigoureuses, que les intuitions didactiques de certains pourront être confirmées, infirmées, rejetées ou au contraire remises « au goût du jour ».

Cependant, outre ces implications spécifiques, l'objectif de notre étude était également d'illustrer une problématique plus globale, celle de l'articulation entre oral et écrit pour l'apprentissage de l'oral, illustration qui nous permet à présent de nous pencher sur des implications didactiques plus générales.

## **II. Implications générales**

Parmi les implications générales, deux points en particulier doivent être soulignés : le rapport entre enseignement/apprentissage de la prononciation et modalité visuelle d'une part, et le rapport entre enseignement explicite et implicite du système oral du français en classe de langue d'autre part.

## **II.1. Prononciation et modalité visuelle**

Quelle que soit l'action entreprise volontairement<sup>221</sup> par l'enseignant en classe de langue, celle-ci devrait, dans l'idéal, systématiquement avoir une fonction didactique. La sélection du matériel pédagogique, linguistique ou non, et de son mode de présentation aux apprenants relèvent donc des choix didactiques de l'enseignant (du moins en théorie, car, en pratique, celui-ci ne peut bien souvent que se contenter de se conformer à ceux de l'institution dans laquelle il travaille). Ces choix sont bien évidemment tributaires des objectifs d'enseignement et d'apprentissage, objectifs qui, actuellement, tendent à suivre le principe d'une certaine modularisation, au moins partielle, des compétences<sup>222</sup>.

On pourrait considérer que cette modularisation, jugée illusoire par certains étant donnée la relative inséparabilité des micro-compétences, et des niveaux d'analyse, linguistiques, concerne davantage l'évaluation des compétences que leur apprentissage à proprement parler. Néanmoins, tout en souscrivant à une vision relativement unifiée de l'apprentissage linguistique, dans laquelle l'ensemble des paramètres sollicités oeuvrent en synergie, tant dans l'apprentissage que dans la communication, nous sommes également conscients des interrogations soulevées par la problématique du transfert inter-tâche des compétences, transfert souvent difficile, comme cela apparaît déjà dans nos propres résultats<sup>223</sup>.

---

<sup>221</sup> Car l'enseignant est lui aussi dépendant de processus contrôlés d'une part, et automatiques d'autre part, la procéduralisation des savoirs didactiques chez celui-ci étant passablement négligée par les chercheurs didacticiens, dont l'attention est prioritairement focalisée sur celle des savoirs linguistiques et/ou communicatifs chez l'apprenant de langue étrangère.

<sup>222</sup> Voir par exemple le « Cadre Européen Commun de Référence pour les Langues » (CECRL) (Béacco, 2004) ou la réforme du DELF et du DALF en 2005 qui est lui est liée (Riba, Lepage et Chevallier-Wixler, 2004).

<sup>223</sup> La prise en compte de la variation en performance et l'évolution des méthodes d'investigation dans le domaine de l'apprentissage, font en effet que l'on ne peut plus aujourd'hui traiter d'« apprentissage », de « savoir », de « compétence » ou de « performance » de manière *globale*. Il semble à présent nécessaire de mesurer l'impact de tel ou tel *facteur* sur la performance du sujet dans une *tâche* et dans des *conditions* particulières, étant donné que la généralisation des résultats et le transfert inter-tâche des compétences est loin d'être évident. Ce dernier point apparaît de manière particulièrement nette dans le domaine de la pathologie du langage, qui nous permet, dans une certaine mesure, de mieux cerner, à travers ses dysfonctionnements, le fonctionnement cognitif de l'être humain. Ainsi, comme le rapportent Nespoulous et Dordain (1994, p 266) :

Par conséquent, loin de soutenir une vue extrême de la modularisation des compétences dans le domaine de l'apprentissage des langues, nous reconnaissons néanmoins qu'il est possible, voire nécessaire, de bien distinguer les compétences mises en jeu lors de la réalisation de certaines tâches d'apprentissage<sup>224</sup>. L'accent n'est donc pas tant mis, dans notre perspective, sur la modularisation des compétences, que sur celle des tâches dans lesquelles ces dernières sont sollicitées. De cette manière, nous pouvons ainsi, à un niveau tant linguistique que psycholinguistique, distinguer, en partie au moins, ce qui relève de l'apprentissage de l'écrit de ce qui relève de l'apprentissage de l'oral, précisément en vue de pouvoir mieux souligner leurs interactions éventuelles.

L'objectif global d'apprentissage sur lequel nous nous concentrons donc ici est celui de l'oral. Nous pouvons alors nous interroger sur le rôle de la modalité visuelle, et plus

---

« C'est ainsi que l'on voit des aphasiques :

- Qui peuvent comprendre une phrase qu'ils ne peuvent produire ;
- Qui peuvent répéter une phrase qu'ils ne peuvent comprendre ;
- Qui peuvent produire une phrase verbalement alors qu'ils ne peuvent l'écrire ;
- Qui ne peuvent lire (et relire) un message qu'ils ont pourtant été capables d'écrire (= alexie sans agraphie !)
- Qui répètent mieux qu'ils ne lisent à voix haute ;
- Qui lisent mieux à voix haute qu'ils ne répètent (= l'inverse → « double dissociation » dont le neuropsychologue est friand dans la mesure où elle lui permet de postuler l'existence de modules de traitement distincts pour la répétition et la lecture...et ce, encore une fois, même si le matériau structural à traiter demeure identique !)
- Etc...etc... ».

<sup>224</sup> Notons à ce propos la remarque de Nespoulous (2000):

*« Il n'est point possible, particulièrement pour un apprenant, de gérer en même temps et au même niveau, les différents paramètres constitutifs de la langue qu'il est en train d'apprendre : paramètres phonétiques, phonologiques, morphologiques, lexiciaux, syntaxiques, pragmatiques... De la même manière, la démarche didactique, par nécessité, ne peut éviter de briser la synergie de l'acte verbal, en orientant tour à tour l'attention de l'apprenant tantôt vers un paramètre, tantôt vers un autre... espérant—au bout d'un certain temps—retrouver la synergie largement automatisée du locuteur natif ».*

Les questions qu'il pose sont donc les suivantes : « a) Comment briser cette pluralité [d'informations] pour travailler, en temps différé, sur tel ou tel aspect de la langue-cible ? b) Comment passer ensuite de n singularités—isolément travaillées—à la pluralité « naturelle » du locuteur expert ? ».



précisément orthographique, dans cet apprentissage, et plus particulièrement, dans celui de « la prononciation du français ».

### **II.1.1. Fonctions de l'écrit pour l'apprentissage de l'oral en classe de langue :**

#### **fixation et segmentation de la parole**

Il convient, avant de nous interroger sur le rôle spécifique de l'orthographe dans l'apprentissage de la « prononciation » en langue étrangère, d'établir quelques considérations générales sur les fonctions de l'écrit dans l'apprentissage de la parole en classe de langue. En effet, si l'on met de côté les habitudes et les traditions didactiques, on peut s'interroger sur la nécessité de recourir aux représentations orthographiques en classe de langue, lorsque l'objectif est l'acquisition de l'oral<sup>225</sup>.

Si l'on reprend les quatre grandes fonctions de l'écrit mentionnées par Coulmas (1989, pp. 11-13) et présentées précédemment, il apparaît que les deux fonctions cruciales, dans le cadre didactique qui est le nôtre, sont celles qu'il qualifie de « mnémonique » et de « réification » : la fonction de réification permet de fixer la parole dans l'espace (et ainsi, dans une certaine mesure, dans le temps), tandis que la fonction mnémonique nous intéresse en ce qu'elle renvoie à la fonction mnésique du support écrit<sup>226</sup>. Cependant, ces deux grandes fonctions doivent être ancrées dans le réel de la classe de langue, si l'on souhaite pouvoir en distinguer les spécificités en fonction des tâches à accomplir.

---

<sup>225</sup> Nous ne pourrions évidemment traiter ici du rôle de l'écrit vis-à-vis des aspects discursifs / textuels de l'énonciation orale, quoique ceux-ci puissent être également abordés par les deux grandes fonctions que nous évoquons ci-après.

<sup>226</sup> Fonction cependant quelque peu « paradoxale », puisque l'utilisation d'un support mnésique implique également une certaine « perte » de mémoire, comme le souligne Durand (2000b, p. 57) :

*« Si les dictionnaires permettent d'accéder à de nombreuses connaissances de façon rapide, on peut objecter que la mémoire individuelle a été considérablement affaiblie dans les communautés à écriture. Dans les situations où l'occidental moyen doit faire ses listes, consulter son agenda, mémoriser ses discours écrits, le "primitif" s'appuie sur sa mémoire (d'où la remarque que la mort d'un vieillard est l'équivalent d'une bibliothèque qui brûle dans de telles communautés). Elsa Gomez-Imbert nous faisait d'ailleurs remarquer que, dans les communautés tukanos où elle a travaillé en Amérique du Sud, les indigènes sont frappés par le manque de mémoire des occidentaux scolarisés qui leur rendent visite ».*

Trois grands types d'écrit peuvent généralement être distingués en classe :

- L'écriture de l'enseignant (généralement sur un « tableau ») ;
- L'écriture du support d'apprentissage (généralement celle de manuels, d'ouvrages, de textes ou de documents proposés par l'enseignant) ;
- L'écriture des apprenants (sous forme de notes personnelles prises par les sujets).

Dans un milieu hétéroglotte, comme celui des apprenants japonais de français au Japon, si l'on met de côté, temporairement au moins, la rencontre possible des sujets avec du texte français ou de langue romane hors de la classe de langue, on peut considérer que ces trois types d'écriture correspondent aux trois situations de lecture en français des sujets apprenants au cours de leur apprentissage de l'oral.

Si l'on cherche à définir à présent ce que nous avons nommé « apprentissage de l'oral », on peut considérer que celui-ci implique<sup>227</sup> :

- Une composante lexico-sémantique (apprentissage de formes lexico-sémantiques),
- Une composante morpho-syntaxique (apprentissage de fonctions morpho-syntaxiques),
- Une composante morpho-phonologique (apprentissage de formes morpho-phonologiques),
- Une composante phonético-phonologique (apprentissage de formes phonético-phonologiques).

Par « apprentissage », en tant que *résultat* d'un ensemble de processus cognitifs, il faut entendre : capacité stable à traiter les composantes sus-mentionnées dans l'ensemble des tâches didactiques/et ou communicatives auxquelles ont été confrontés les sujets apprenants.

---

<sup>227</sup> Plusieurs types de présentation et de découpage seraient bien évidemment possibles.

Parmi ces tâches, lorsque celles-ci impliquent un stimulus oral, nous pouvons distinguer certaines tâches fondamentales, comprises, à des degrés divers, dans les multiples activités didactiques que l'on peut trouver en classe de langue :

- Perception orale (composante phonético-phonologique) ;
- Compréhension orale (impliquant l'ensemble des composantes précédentes à des degrés divers) ;
- Répétition (impliquant l'ensemble des composantes précédentes à des degrés divers) ;
- Production orale, libre ou contrainte (impliquant l'ensemble des composantes précédentes à des degrés divers) ;
- Production écrite sous dictée (impliquant l'ensemble des composantes précédentes à des degrés divers).

Si l'on considère que, à la base de l'apprentissage de l'oral, se trouve celui de la composante phonético-phonologique (sans quoi on ne peut, dans une certaine mesure, comprendre ni produire en langue étrangère), on peut alors s'interroger sur le rôle de l'écrit dans un tel apprentissage.

La réponse à cette question, nous semble-t-il, se trouve fondamentalement dans la dimension de la temporalité : alors que l'oral est fugace, et ainsi généralement « non-contrôlable » par l'auditeur, l'écrit est permanent, et ainsi généralement « contrôlable » par le lecteur (dans le sens où celui-ci contrôle, au moins temporellement, le texte dans son activité de lecture). De par sa fonction de réification, l'écrit permet de fixer l'oral dans l'espace et ainsi dans le temps. Cette fixation permet alors la réalisation de la fonction mnémonique de l'écrit, le texte constituant alors une « mémoire externe », récupérable via le processus de lecture. Deux grands processus cognitifs sont donc concernés par l'écrit : la perception et la mémorisation. La réalisation des tâches de compréhension, de répétition et de production, orale ou écrite, mentionnées

précédemment, est fondamentalement dépendante de ces deux grands processus primordiaux, et c'est donc sur ceux-ci que nous nous concentrons ici<sup>228</sup>.

Du point de vue de la perception, l'écrit, considéré comme une simple transcription de l'oral, offre deux possibilités essentielles :

- Fixer les unités linguistiques, de manière à pouvoir les encoder plusieurs fois (les « retours en arrière » ou « saccades régressives », possibles en lecture, mais impossible en perception auditive) ;
- Segmenter les unités linguistiques, grâce à la fixation dans l'espace d'une part, et grâce au caractère discret des signes graphiques d'autre part.

Ces deux propriétés permettent alors de mettre en œuvre un certain processus de mémorisation de séquences (Ellis, N., 1996), et ainsi d'apprentissage :

- La fixation permet d'augmenter, *via* la lecture répétée, la fréquence d'exposition aux items, fréquence d'exposition cruciale pour la mémorisation ;
- La segmentation permet de satisfaire les contraintes de la mémoire de travail, dont l'empan est limité, et facilite l'identification des éléments à stocker et/ou déjà partiellement stockés.

En outre, si l'on considère le rôle des représentations orthographiques couplées aux stimuli oraux qu'elles transcrivent, et si l'on souscrit à la thèse de la pluralité des formats de stockage en mémoire, on peut également considérer que ces représentations constituent un deuxième format représentationnel (par rapport au format phonético-phonologique) et contribuent ainsi à la mémorisation des items *via* un double codage (visuo-orthographique d'une part et phonético-phonologique d'autre part).

---

<sup>228</sup> La perception pouvant étroitement être liée à la production, en particulier dans la perspective des théories motrices de la parole, pour lesquelles la perception implique une certaine récupération de gestes articulatoires.

D'un point de vue didactique, ce double format permet également de multiplier la variété des tâches didactiques et les possibilités d'utilisation et de manipulation du matériau linguistique, à travers lesquelles s'accomplit l'apprentissage de la langue.

Enfin, il nous faut également mentionner une dernière caractéristique, liée à la fonction de réification, de l'écrit : il s'agit de l'invariance, relative, des formes orthographiques, par rapport aux formes phonétiques, inévitablement soumises (excepté dans le cas d'enregistrements) à la variation du signal acoustique. Cette dernière propriété, qui renvoie à l'articulation entre norme et variation, sera traitée plus bas.

Il apparaîtra peut-être, à la lecture de ce qui précède, que les « items » ou « unités linguistiques » mentionnés semblent particulièrement renvoyer à la composante lexico-sémantique de l'oral. Néanmoins, si l'on considère que ces unités, que nous pourrions qualifier de morphèmes, constituent les unités significatives minimales que les apprenants doivent percevoir et mémoriser, il apparaît alors que la fixation et la segmentation de la parole, que permet l'écrit, jouent un rôle relativement similaire dans le cas des composantes morpho-syntaxiques et morpho-phonologiques, puisque l'écrit, de par les contraintes temporelles et attentionnelles différentes de celles de l'oral qu'il implique, permet une exposition répétée aux collocations morpholexicales. On pourrait bien sûr souligner le fait que, dans le cas du français, la linéarité de l'écriture peut parfois rentrer en conflit avec la non-linéarité de nombreux phénomènes morphosyntaxiques, à commencer par le cas classique de la négation en français : « ne...pas ». Mais une telle remarque s'appliquerait également, quoique de manière différente, à la parole orale.

Ainsi, il semblerait que deux des grandes fonctions didactiques de l'écrit en classe de langue, vis-à-vis de l'oral, soient liées à la fixation et à la segmentation de la parole, lesquelles sont impliquées dans la perception et la mémorisation des éléments linguistiques à apprendre. Les deux questions que l'on doit alors se poser sont les suivantes :

- Les représentations orthographiques sont-elles indispensables à la réalisation de ces deux fonctions ?
- Existe-t-il des « effets secondaires » de l'utilisation de l'orthographe, qui pourraient conduire à chercher d'autres procédures ou d'autres formats représentationnels pour la réalisation de ces deux fonctions ?

En vue de tenter de répondre à ces deux questions, nous devons, puisque le cœur de notre étude concerne la composante phonético-phonologique de la parole, recentrer celles-ci sur les caractéristiques phonético-phonologiques des items à percevoir et à mémoriser.

### **II.1.2. Caractéristiques phonético-phonologiques et input visuel**

Comme nous l'avons déjà indiqué, l'apprentissage du système phonético-phonologique du français implique la perception, la mémorisation et la capacité de solliciter ce système, en tâche de reconnaissance auditive, de compréhension orale, de répétition, de production orale ou encore de production écrite sous dictée. Il faut néanmoins rappeler que, dans une perspective réaliste, cet apprentissage ne s'effectue généralement pas isolé de l'apprentissage global du français : même si certaines activités ponctuelles peuvent y être spécifiquement consacrées, celui-ci s'effectue généralement *via* l'exposition à la langue, que ce soit à travers le discours de l'enseignant, à travers des documents sonores ou audiovisuels enregistrés ou encore, et c'est bien là ce qui nous préoccupe, à travers du texte, lequel est censé correspondre, chez le lecteur expert, à des séquences de formes phonologiques.

#### **II.1.2.1. Préliminaires**

Il est remarquable, pour nous, que le thème de la transcription de la parole soit un thème d'actualité en sciences du langage<sup>229</sup>, notamment en raison du développement

---

<sup>229</sup> Comme l'illustre la tenue du colloque « Transcription de la langue parlée. Aspects théoriques, pratiques et technologiques », organisé par la composante « linguistique sur corpus » de l'équipe GRELLANG à l'université de Perpignan en juin 2005, ou encore la publication, en 2002, d'un triple

des travaux de corpus, qui rencontrent précisément les difficultés inhérentes à une telle tentative de transcription. Loin de transcrire fidèlement la parole, notamment en français, l'écrit conduit, sur certains aspects, à un appauvrissement des énoncés originaux, notamment en termes prosodiques, un appauvrissement dont la fonction originelle est une fonction d'économie linguistique. D'un autre côté, l'orthographe française suit certains principes, en particulier les principes morphologiques (<d> : grand/grande/grandir) et distinctif (homonymes : poids, pois, poix), qui lui confèrent une valeur informative absente de l'oral.

Nous ne tenterons pas ici d'approfondir le thème complexe des interfaces entre écrit, oral et morphologie, même si celui-ci pourrait constituer l'une des suites logiques de notre travail sur la gestion de l'oral et de l'écrit pour les apprenants de FLE. Il nous semble qu'un véritable chantier est à entreprendre sur la question, et l'on ne manquera pas de remarquer, par exemple, que, dans leur chapitre consacré à la morphologie du français, Battye, Hintze et Rowlett (2000), traitent séparément la morphologie orale et la morphologie écrite. La question des alternances, évoquée par Goldsmith, est d'ailleurs l'une de celles qui se trouvent au cœur de l'interface entre morphologie et phonologie (« how do we account for alternations of morphemes in distinct morphological and phonological contexts ? » (Goldsmith, 1995, p. 8)), et leur gestion en classe de langue, vis-à-vis de l'oral et de l'écrit, constituerait un sujet d'étude ambitieux. Par ailleurs, dans le cadre de la théorie de l'Optimalité, la problématique des interfaces, entre syntaxe et phonologie par exemple, semble prendre un nouveau tour, celles-ci étant présentées non plus tant en termes d'interactions entre modules, qu'en termes de contraintes (Archangeli, 1997, p. 30) :

*« A logical extension of the OT model for language is that there is a single constraint hierarchy, which internally ranks all constraints, whether syntactic, morphological, phonological, phonetic or semantic. This possibility predicts*

*interactions between components (modules). For example, particular syntactic constraints might be violated in order to satisfy a phonological or morphological constraint, or vice-versa. This contrasts sharply with the view of grammar of having a separate and independent syntactic component, phonological component, etc. ».*

Pour notre part, nous nous contenterons ici de nous concentrer sur les aspects segmentaux et syllabiques de la phonologie du français. Cependant, comme l'illustre notre démarche, le rôle des représentations orthographiques dans l'apprentissage de l'oral ne peut être évalué globalement : il convient dans un premier temps de mettre en regard les systèmes phonético-phonologiques et phonographémiques en présence, puis de préciser les tâches didactiques que l'on examine, avant, enfin, de déterminer la fonction que l'on envisage de faire jouer, ou de ne pas faire jouer, au support orthographique dans l'apprentissage de l'oral.

#### II.1.2.1.1. Systèmes d'écriture en L1 et homographes hétérophones L1/L2

On doit alors distinguer quatre grandes catégories d'apprenants, en fonction de leur système d'écriture, vis-à-vis de celui du français :

Système alphabétique	1) Script romain
	2) Script non-romain
Système non-alphabétique	3) Avec système de transcription en alphabet romain
	4) Sans système de transcription en alphabet romain

Deux questions se posent alors : celle de la conscience phonémique (pour les groupes 3) et 4)) et celle des correspondances phonographémiques (pour l'ensemble des groupes). Cela s'illustre parfaitement dans le cas du japonais, qui relève de la catégorie 3) et dont les sujets doivent à la fois développer une conscience phonémique et apprendre de nouvelles correspondances phonographémiques. En revanche, un public



d'apprenants espagnols (groupe 1), par exemple, ne serait confronté qu'au second problème.

Si l'on tente à présent de caractériser le rôle de l'orthographe française dans l'apprentissage de la phonologie du français pour le groupe 1), voici ce que l'on pourrait obtenir *a priori*, de manière très globale et théorique<sup>230</sup> :

Conscience phonémique	Acquise	→ Orthographe non-nécessaire
Correspondances phonographémiques	Identiques	→ Orthographe utile (transfert positif)
	Différentes	→ Orthographe néfaste (transfert négatif)

Dans ce cas, l'orthographe n'est pas utile au développement de la conscience phonémique puisque le principe alphabétique est déjà acquis. Dans les cas, relativement rares, où les correspondances phonographémiques sont identiques en L1 et en L2, les apprenants peuvent bénéficier d'acquis en L1, qu'ils peuvent réinvestir en L2 (par exemple la correspondance (<f>, /f/) pour des apprenants anglais de français). Cependant, ces correspondances diffèrent généralement d'une langue à une autre : soit parce que les graphèmes (identiques) renvoient à des unités phonético-phonologiques différentes (homographes hétérophones L1/L2) (par exemple le graphème <r> pour des apprenants espagnols de français), soit parce que des graphèmes différents renvoient à des unités phonético-phonologiques identiques (hétérographes homophones L1/L2) (par exemple le graphème <eau> pour ces mêmes apprenants) ou différentes (hétérographes hétérophones L1/L2) (par exemple le graphème <th> pour des apprenants français d'anglais).

Alors que le cas des hétérographes L1/L2 concerne véritablement l'apprentissage de la lecture et de l'écriture (« l'apprentissage de l'orthographe »), celui des homographes nous concerne plus particulièrement puisque c'est précisément sur ce point que

---

<sup>230</sup> Puisqu'il faudrait davantage préciser la distinction entre niveaux phonétique et phonologique, ainsi que prendre en compte les variantes positionnelles à la fois allophoniques et graphémiques.

l'attention didactique semble faire défaut dans les pratiques pédagogiques contemporaines. En effet, en raison de la diffusion massive de l'alphabet romain à travers le monde, et en raison de l'absence fréquente de formation/d'information des enseignants de langue sur les processus à l'œuvre en tâche de lecture, par voie d'adressage ou d'assemblage, il est fréquent que ceux-ci tiennent pour acquise, chez un public d'adultes, la capacité de « lire » le français. Or, comme le souligne Catach (1995, p. 59) :

*« [L'alphabet traditionnel] ne saurait suffire à établir un bon découpage de la chaîne écrite. Le premier progrès à faire est de bien distinguer le phonème du son, le graphème de la lettre. [...] Il n'est pas étonnant qu'une approche lettre à lettre, coupée de l'oral et de la langue, apparaisse comme anarchique et démesurément difficile ».*

Le cas est sans doute différent pour les enseignants de chinois, de russe ou de japonais, par exemple, qui seraient plus enclins à consacrer un certain temps à l'enseignement des codes phonographémiques de leur langue, au vu de la moindre diffusion desdits codes. La distinction lettre/graphème, entre autres, semble donc (généralement) faire défaut aux pratiques pédagogiques contemporaines (connaître l'alphabet n'implique pas de connaître le répertoire graphémique de langues à orthographe profonde comme le français), et l'on peut ainsi entendre certains enseignants qualifier d'erreur de « prononciation » (ce qui n'est par ailleurs pas toujours faux) une erreur de correspondance phonographémique.

Lorsque l'on prend en compte le fait que ces homographies existent à un niveau tant lexical que sublexical, on réalise alors à quel point l'absence d'apprentissage systématique des correspondances phonographémiques peut avoir des conséquences sur l'apprentissage des formes lexicales et du système phonético-phonologique de la L2. Ces conséquences s'expliquent très schématiquement de la manière suivante : la force des associations phonographémiques en L1, aux niveaux lexical et sublexical, conduit à l'association des formes orthographiques de la L2 à des formes phonético-

phonologiques de la L1 (notamment *via* le processus de recodage phonologique dans la boucle articulatoire en mémoire de travail), qui sont alors considérées comme appartenant au lexique de la L2, conduisant à la construction du système phonético-phonologique de la L2 sur la base des unités de la L1.

Bien évidemment, la réalité est plus complexe, le processus précédemment décrit variant notamment en fonction des ressources attentionnelles du sujet et du degré de contrôle exercé par celui-ci sur l'accomplissement de la tâche. Cependant, cette schématisation a le mérite de souligner l'importance des homographes, lexicaux ou sublexicaux, dans l'apprentissage des formes linguistiques et ainsi dans la constitution du système phonético-phonologique en L2, impliquant, de manière quelque peu contre-intuitive, que l'apprentissage des correspondances phonographémiques pourrait se révéler, sous un certain angle, plus important pour des apprenants partageant le même type d'écriture et le même script (homographes nombreux) que pour d'autres. Cependant, et malgré la distinction entre les quatre groupes d'apprenants, la diffusion massive de l'alphabet romain, a, d'une certaine manière, complexifié la situation, comme cela apparaît nettement dans le cas du japonais, avec son système que nous avons qualifié de « pluriscriptural ».

En ce qui concerne les trois autres groupes, on pourrait considérer, d'après ce que nous avons affirmé, que, les graphèmes étant différents (hétérographes), l'utilisation de l'orthographe ne semble pas « néfaste » pour l'apprentissage des formes phonético-phonologiques. De plus, pour les groupes 3) et 4), l'utilisation de l'orthographe pourrait se révéler utile, voire nécessaire, pour le développement de la conscience phonémique, nécessaire à l'apprentissage de la phonologie du français<sup>231</sup>.

Néanmoins, comme nous l'avons déjà indiqué, une telle situation « théorique » dissimule le fait que, bien souvent, l'alphabet romain est déjà plus ou moins connu des apprenants, soit à travers un système officiel de transcription, comme en japonais ou en

---

<sup>231</sup> Par exemple dans le cas d'apprenants chinois (Bergère, 2005).

chinois, soit à travers l'apprentissage scolaire de l'anglais. Or, comme nous l'avons vu précédemment, la coexistence de ces systèmes n'assure en rien la connaissance des correspondances phonographémiques, voire pourrait même l'entraver, puisque cela ramène alors ces sujets à la catégorie 1). Dans cette perspective, on pourrait alors considérer que les Japonais ne sont pas « avantagés » par leur situation sociolinguistique contemporaine, puisque l'alphabet romain est déjà, partiellement au moins, utilisé pour transcrire leur propre lexique phonologique. Ce système graphique n'est donc plus « vierge » pour eux, puisqu'il est déjà inscrit dans un certain système phonographémique en L1 (ce qui nous ramène au groupe 1)), et, de plus, ce système étant moraïque, et non phonémique, l'alphabet perd alors son intérêt en tant qu'outil de développement de la conscience phonémique. Une telle situation, au niveau sublexical, est aggravée, au niveau lexical, par la présence massive d'emprunts étrangers, phonologiquement adaptés dans la langue japonaise, essentiellement, ou initialement diffusés sous leur forme orthographique (notamment dans le domaine médiatique et commercial). Ce dernier point évoque la question importante de la reconnaissance lexicale et/ou sublexicale pour les apprenants débutants, qui pourrait constituer un obstacle à l'apprentissage des nouvelles formes phonético-phonologiques.

Jusqu'à présent, nous avons affirmé que les apprenants de type 1), en raison des similarités graphémiques entre la L1 et la L2, risquaient d'effectuer des transferts négatifs si le support orthographique était utilisé pour appréhender l'oral. D'un autre côté, les apprenants des autres types semblaient moins susceptibles d'effectuer de tels transferts puisque les graphèmes étaient différents. Il faut cependant considérer l'autre versant de ce tableau : la familiarité des apprenants de type 1) avec l'alphabet romain, qui leur permet d'appréhender la lecture sans trop de difficultés de « bas-niveau », notamment d'un point de vue rythmique. Les autres apprenants, en revanche, doivent apprendre de nouveaux signes et combinaisons graphémiques, ce qui constitue une difficulté d'un autre genre.

Pourtant, en ce qui concerne le groupe 1), c'est précisément cette familiarité qui risque, en raison des automatismes de lecture, d'entraîner l'activation de correspondances

phonographémiques (lexicales ou sublexicales) de la L1, alors que ceux des autres groupes rencontreront des difficultés d'un autre ordre (notamment la délimitation des unités fonctionnelles, qui entraîne bien souvent chez les apprenants débutants une prépondérance de la voie d'assemblage dont l'unité de base serait la lettre et non le graphème, d'où des « erreurs de lecture », en particulier dans le cas des « lettres muettes » en français).

On pourrait donc considérer que les apprenants de type 1) doivent apprendre : a) de nouvelles catégories phonético-phonologiques et b) de nouvelles correspondances phonographémiques, tandis que les autres doivent apprendre, outre a) et b), de nouvelles catégories graphémiques. Dans le premier cas, il s'agirait donc de « modifier » des habitudes de lecture, tandis que dans le second il s'agirait d'un apprentissage « complet » de la lecture. Le problème se posant, dans les deux cas, est que, contrairement à l'apprentissage de la langue maternelle, les sujets ne maîtrisent pas encore le système phonético-phonologique de la L2, que le système orthographique est censé transcrire. « Lire » (à voix haute) en L2 reviendrait donc globalement, au niveau lexical ou sublexical, à associer une forme phonético-phonologique, qui n'est pas encore acquise, à une forme orthographique.

On comprend alors immédiatement que ces formes orthographiques ne puissent pas forcément aider l'apprentissage des formes phonético-phonologiques, bien au contraire. Cela n'enlève rien au fait que les formes orthographiques doivent également être apprises, mais il s'agit là d'une problématique différente.

Ayant récapitulé ainsi globalement la problématique, nous pouvons à présent, afin de mieux comprendre les implications générales de notre étude, partir d'exemple didactique simples et plausibles, afin d'examiner le rôle que l'on devrait ou ne devrait pas assigner aux représentations orthographiques pour l'apprentissage de l'oral.

### II.1.2.1.2. Perception et mémorisation de l'oral

Les enseignants sont habitués à ce que leurs apprenants leur demandent d'écrire au tableau un énoncé, un syntagme ou un morphème qui leur a été présenté à l'oral, mais qu'ils sont incapables de répéter, d'identifier ou, plus globalement, de percevoir. Une fois l'item écrit, les apprenants semblent alors capables d'identifier ou de reconnaître ce mot. Cependant, ils ne le lisent, ou ne le répètent, pas « correctement » pour autant.

On retrouve dans cette situation la fonction de réification de l'écriture, qui permet aux sujets de mieux « appréhender l'item ». Cependant, l'expression « appréhender l'item » est source de confusion : appréhender l'item orthographique ne revient pas à appréhender l'item phonétique, tant dans l'item en lui-même (caractéristiques intrinsèques) que dans les processus mis en œuvre (perception auditive *vs.* lecture). On doit alors rappeler que la parole, contrairement à l'écrit, est marquée d'une part par la coarticulation, d'autre part par la variation phonétique (aléatoire ou allophonique), et enfin par les caractéristiques prosodiques absentes de l'écrit.

Puisqu'il nous faut distinguer trois formes pour l'item considéré (phonétique, phonologique et orthographique), la question qui se pose à l'enseignant est alors la suivante : quel est l'objectif didactique de la tâche ?

S'agit-il d' :

1. Accéder au sens (comprendre) ?
2. Identifier l'item ? Si oui, sous quelle forme (phonétique, phonologique, orthographique) ?
3. Etre capable de répéter (forme phonétique/phonologique) ?
4. Mémoriser ? Si oui, quelle forme (phonétique, phonologique, orthographique) ?
5. Transcrire (forme orthographique) ?

Ce qui apparaît alors, c'est que l'apprentissage de l'oral en perception, en premier lieu phonétique, doit effectivement être phonétique et non orthographique : être capable de reconnaître une forme orthographique n'implique en aucun cas d'être capable de reconnaître une forme *phonétique*. Cela n'exclut pas que les formes orthographiques

puissent être utiles à cette formation, *via* les formes phonologiques auxquelles elles renvoient, mais de tels éléments démontrent clairement ce que les enseignants de langue savent depuis longtemps, à savoir que l'on n'apprend pas à entendre (et à parler) en L2 uniquement en lisant en L2<sup>232</sup>. Encoder une forme orthographique ne revient pas à encoder une forme phonétique, en particulier dans une perspective exemplariste (Hardison, 1999), ce qui nous renvoie au format de mémorisation.

### II.1.2.1.3. Correction phonétique

En ce qui concerne la « correction phonétique », il nous faut reprendre notre situation didactique de départ : des débutants complets, supposés ne pas avoir été exposés au texte, et ainsi ne pas avoir déjà stocké certaines formes lexicales orthographiques et correspondances phonographémiques erronées en L2. Les problèmes rencontrés sont alors d'ordre purement phonético-phonologique et renvoient à la problématique de l'interphonologie. Les problèmes de « prononciation » des apprenants, quelle que soit la tâche orale (perception, répétition, production), sont alors dus à des phénomènes de transfert, de marque, de différences d'inventaires phonémiques, allophoniques et phonotactiques, que l'on peut également exprimer en termes de classement de contraintes (en perception ou en production). Plusieurs méthodes et techniques existent dans ce cas pour tenter de remédier à ces problèmes, que ce soit dans une perspective plutôt auditive (la méthode verbo-tonale) ou plutôt articulatoire (avec présentation de schémas articulatoires par exemple).

---

<sup>232</sup> Ce qui correspond tout à fait à ce qu'écrit Coulmas à propos du linguiste et de l'étude « formelle » de la langue (1989, p. 268) : « *Every writing system is an abstraction, and every transcript is an artefact. Speech can be represented by graphical means very imperfectly. [...] every linguist should at least be aware of it, rather than assume, naïvely, that it is possible to study speech simply by writing it down* ». Comme il l'indique au début de son ouvrage (1989, p. 47) : « *The abstractness of writing is possible and functional because the typical reader –that is, the reader for whom the script is made – knows the language in which the written message is coded, and can thus rely on the redundancies of the language as an aid for deciphering (reading) written expressions which represent speech only incompletely or vaguely. Therefore it is possible, for example, to leave the tones of Tibetan unrepresented in the Tibetan script or the vowels in Semitic scripts* ». Or les apprenants de L2 ne connaissent justement *pas* la L2, et ne peuvent donc que très peu, voire ne doivent pas, par risque d'erreur, se servir de leurs connaissances phonético-phonologiques préalables pour « déchiffrer » le texte écrit.

Il faut alors noter que, dans toutes ces méthodes, la modalité visuelle est sollicitée : soit à travers la micro- ou macro-gestualité (pour la première), soit à travers des représentations picturales de type articulatoire (pour la seconde). De plus, même si l'approche adoptée est davantage de type « auditif », il peut parfois être utile, comme l'indique Wells (2000, p. 120), de fournir une explication de type articulatoire des processus phonétiques à l'œuvre, en particulier lorsque les organes articulatoire peuvent être facilement visualisés, comme c'est le cas pour la distinction entre [b] et [v].

Dans les deux cas, les représentations oro-labiales (vision des mouvements buccaux, externes ou internes, de l'enseignant) peuvent être sollicitées ou manipulées, et nous avons présenté précédemment certains résultats relatifs à ces manipulations. Deux cas de figure se présentent alors généralement : soit les apprenants se révèlent incapables de produire la séquence sonore attendue, soit cette production est irrégulière. Alors que dans le premier cas, le travail de l'enseignant est de conduire les apprenants à pouvoir percevoir et reproduire cette séquence, dans le second cas, il s'agit surtout de fournir des indices de rappel à l'apprenant de manière à ce que celui-ci puisse récupérer la séquence articulatoire qu'il a pu obtenir précédemment, de manière à ce que celle-ci puisse graduellement être fixée en mémoire.

Lorsque l'on considère des « erreurs de prononciation », celles-ci peuvent être généralement abordées en termes de « substitution » : substitution d'un trait à un autre, substitution de la présence d'un trait à son absence, substitution d'un phonème à un autre, substitution d'une structure syllabique à une autre (*via* l'épenthèse ou l'effacement), substitution d'une caractéristique prosodique à une autre, substitution d'un classement de contraintes à un autre. On peut donc imaginer que le travail de l'enseignant est, dans un premier temps, de faire prendre conscience aux apprenants de la distinction entre les deux items (le substitué et le substituant), d'où l'importance des tâches de discrimination. Puisque la modalité auditive ne semble pas toujours permettre aux apprenants de percevoir l'élément substitué (ce que certains pourraient qualifier de « surdité » phonologique ou phonétique (Billières, 1988 ; Borrell, 1991 ; Dupoux,



Pallier, Sebastian et Mehler, 1997)) l'enseignant peut tenter de faire appel à la modalité visuelle pour suppléer à cette déficience perceptive.

Cinq principaux types d'information visuelle sont alors envisageables :

- Information oro-faciale : la vision du visage (et donc de l'appareil buccal externe) de l'enseignant.
- Information gestuelle non-faciale : la vision des gestes (mains, bras et buste en particulier) de l'enseignant.
- Information picturale : la vision de schémas, d'icônes ou plus généralement d'images, soit normés comme des schémas phonétiques articulatoires, soit idiosyncratiques, comme des schémas géométriques, chromatiques ou non, propres à l'enseignant et pouvant alors recouper la catégorie de l'information graphique, mentionnée ci-dessous.
- Information graphique : la vision d'un code graphique non-orthographique (soit normé, comme l'Alphabet Phonétique International, soit idiosyncratique, comme un code abstrait propre à l'enseignant).
- Information orthographique : la vision des représentations orthographiques (en L2 ou en L1) des items.

#### II.1.2.2. Information oro-faciale

En ce qui concerne l'information oro-faciale, outre les résultats que nous avons déjà présentés à ce sujet (Hardison, 1999, 2000, 2003 ; Hazan, Anke et Faulkner, 2002 ; Massaro et Light, 2003 ; Ortega-Llebaria, Faulkner et Hazan, 2001 ; Sennema, Hazan et Faulkner, 2003), il est intéressant de noter que, d'après le modèle de Massaro, la perception de la parole implique une intégration des informations audiovisuelles, et que, lorsqu'une source d'information est ambiguë (en particulier la source auditive), alors la modalité visuelle est particulièrement sollicitée. Cela apparaît clairement aux enseignants dont les étudiants semblent explicitement focaliser leur attention visuelle sur la bouche de l'enseignant en tâche de répétition. Certains praticiens de la méthode verbo-tonale suggèrent ainsi parfois aux enseignants de masquer leurs lèvres, lors de la

correction phonétique, de manière à éviter certaines confusions et surtout à accroître, chez l'apprenant, l'attention portée sur la modalité auditive<sup>233</sup>.

Cependant, il nous semble qu'une telle pratique ne doit pas être généralisée, pour quatre raisons au moins :

- Entraîner à percevoir la parole en L2 implique également l'entraînement à la lecture oro-labiale.
- Entraîner à percevoir la parole, dans une perspective motrice, implique l'entraînement à la récupération des gestes articulatoires, notamment par la modalité visuelle (Schwartz, Abry, Boë et Cathiard, 2002).
- Pour certains sons (par exemple [b] vs. [v]) l'ambiguïté auditive peut être plus facilement résolue par le recours à l'information visuelle (labio-dentale dans notre exemple).
- L'entraînement à la résolution de l'ambiguïté visuelle peut entraîner l'augmentation de l'attention portée à la modalité auditive.

En bref, on peut dire qu'entraîner à la perception de la parole en L2 implique d'entraîner les deux modalités. On peut alors imaginer que, à l'instar de la technique de la « prononciation nuancée », dans laquelle l'input auditif est modifié, l'enseignant pourrait modifier la position de ses articulateurs (et plus globalement de son visage) de manière à mieux faire percevoir visuellement (modification des « visèmes ») ce qui ne semble pouvoir être perçu auditivement. Une telle modification visuelle, si elle ne conduit pas véritablement à l'identification auditive du segment en L2, servirait en revanche d'indice attentionnel pour signaler certaines distinctions entre deux segments de la L2 ou entre un segment de la L1 et un segment de la L2. C'est ce que nous pratiquons nous-mêmes parfois avec nos apprenants japonais, lorsque nous étirons exagérément les lèvres en vue de souligner la distinction entre le [ø] français et le [ɯ]

---

<sup>233</sup> Ce qui se justifie en particulier lorsque l'enseignant fait recours à la technique de la « prononciation nuancée », qui implique une modification de l'input acoustique, et donc possiblement des mouvements labio-buccaux visibles, en fonction des « erreurs » de répétition des apprenants.

japonais, une distinction qui passe autrement inaperçue, comme l'indique les commentaires des étudiants : « Je n'entends pas la différence ».

Il faut alors prendre en compte les différentes catégories de « visèmes » auxquelles renvoient les catégories phonémiques (Hardison, 2003) : tandis que certaines oppositions phonémiques peuvent bénéficier de la vision oro-faciale (distinction visuelle saillante), d'autres ne sont pas accessibles à la vision des apprenants (par exemple l'opposition [k] vs. [g]). L'information visuelle oro-faciale doit donc être sollicitée lorsque cela est approprié, et l'utilisation ponctuelle de schémas articulatoires (information picturale), que l'on trouve dans de nombreux manuels, peut alors se justifier, contrairement à leur utilisation systématique, qui ne prend pas en compte le caractère saillant ou non des traits distinctifs visuels.

C'est donc d'une question de « dosage » entre information auditive et visuelle qu'il s'agit, en fonction des caractéristiques, auditives et visuelles, des segments et des oppositions en question. Cependant, nous ne manquerons pas de noter, dans une perspective didactique réaliste, que la valeur de cette information est généralement relativement limitée par les contraintes spatio-temporelles de la situation didactique : dans un contexte universitaire en particulier, il est peu probable que les étudiants assis au fond d'un amphitéâtre de quarante places puissent avoir accès à une telle information. L'utilisation de celle-ci par l'enseignant implique donc une certaine organisation proxémique et kinésique, qui pourrait utilement faire appel à la technologie contemporaine, *via* un système de projection sur écran<sup>234</sup>.

---

<sup>234</sup> On pourrait également développer la question en mentionnant l'utilisation de laboratoires multimédia et de didacticiels appropriés.

### II.1.2.3. Information gestuelle non-faciale

En ce qui concerne l'information gestuelle non-faciale, que certains appelleront « macro-gestualité », celle-ci est également sollicitée dans la correction phonétique<sup>235</sup>, notamment dans la méthode verbo-tonale<sup>236</sup>. On peut distinguer physiquement trois grandes catégories de groupes gestuels :

1. Ceux liés aux mains et aux bras
2. Ceux liés au buste et à la posture
3. Ceux liés aux jambes et à la démarche

L'une des caractéristiques de cette macro-gestualité, par rapport à des supports visuels fixes (comme les schémas statiques), est certainement son caractère dynamique et adaptatif, qui en fait un outil particulièrement adapté à certains aspects prosodiques comme le rythme, l'accentuation, la syllabation ou l'intonation.

L'étude du rôle du geste dans la communication verbale, et donc dans l'enseignement/apprentissage des langues (Faraco, 2001 ; Tellier, 2005), constitue tout un domaine de recherche. Nous ne prétendons pas ici nous y rattacher, car cela nous éloignerait trop de notre problématique initiale et conduirait à un travail de recherche, notamment bibliographique, dépassant le cadre de cette étude. La catégorisation des gestes « communicatifs », en particulier, a donné lieu à de nombreux travaux, et plusieurs typologies ont ainsi vu le jour, distinguant notamment ceux qui peuvent fonctionner de manière indépendante de la parole (« quasi-linguistiques » chez des auteurs comme Cosnier, « emblèmes », « gestes autonomes » ou « geste sémiotiques » chez d'autres) de ceux qui accompagnent nécessairement la parole (« syllinguistiques » chez Cosnier, divisés en « phonogènes » et « coverbaux »). Parmi ces derniers, en fonction de leur lien au contenu linguistique de la parole (variable notamment en fonction du degré de concrétude sémantique des items linguistiques), il est possible de distinguer différentes sous-catégories de gestes (« paraverbaux », « expressifs » et

---

<sup>235</sup> Par l'enseignant ou dans certains manuels spécialisés (Kaneman-Pougatch et Pedoya-Guimbretière, 1989).

<sup>236</sup> Voir en particulier les travaux de Billières, Intraiva et Renard.

« illustratifs »). Ceux qui sont particulièrement liés à ce contenu sont nommés « illustratifs » ou « iconiques » chez Cosnier, qui distingue, au sein de cette catégorie, les « déictiques » (désignation référentielle), les « spatiographiques » (schématisation spatiale), les « kinémimiques » (mime de l'action) et les « pictomimiques » ou « pictographiques » (dessin référentiel dans l'espace).

Cependant, étant donné que nous nous intéressons exclusivement ici aux gestes « didactiques », comme par exemple ce que Tellier nomme « geste pédagogique » dans le cas de la compréhension verbale lexicale en langue étrangère, et que nous ne nous inscrivons pas dans le domaine de la kinésique, la terminologie employée ci-après ne vise donc pas à la précision scientifique qu'exigerait le rattachement à un tel domaine. Les termes employés ne renvoient donc à aucune typologie académique particulière et sont à interpréter dans leur acception commune, afin d'être à la portée de tout lecteur non-spécialiste. Ainsi, lorsque nous examinons notre propre pratique gestuelle en classe de correction phonétique<sup>237</sup>, nous parvenons à trois axes :

#### II.1.2.3.1. Axe segmental

Dans cas, les catégories 1) et 2) sont principalement sollicitées. Tandis que la catégorie 2) peut être utilisée dans le cadre phonétique de la méthode verbo-tonale (opposition Tension/Relâchement), elle l'est également, avec la catégorie 1) pour remplir une certaine fonction iconique : les gestes des bras, des mains et la posture permettent en effet de représenter, partiellement au moins, certains mouvements des articulateurs internes, en particulier de la langue (par exemple dans l'opposition entre /r/ et /l/ en français pour les Japonais, la main représentant la langue). Ils peuvent également permettre d'amplifier la micro-gestualité (en particulier la labialisation, représentée à l'aide des bras et des mains), de marquer dans l'espace l'opposition entre deux catégories phonémiques ou le passage progressif de l'une à l'autre (par exemple de droite à gauche et de haut en bas /u, ø, y/), ou encore simplement d'indiquer un certain

---

<sup>237</sup> Ce que l'on pourrait considérer comme une « théorisation des pratiques ».

lieu (par exemple la gorge pour la réalisation du [R] par opposition au [l]) ou trait (par exemple la gorge pour le voisement) d'articulation.

Ainsi, dans notre propre expérience de classe, nous avons constaté l'utilité, variable selon les sujets, de placer notre main sur notre gorge lorsque les apprenants devaient produire un [ʁ] et qu'ils avaient produit (« feedback correctif ») ou qu'ils risquaient de produire (« feedforward anticipateur ») un [l] ou un [r]. Cette technique pouvait avoir une fonction soit attentionnelle, soit mnésique en leur servant d'indice de rappel pour le geste articulatoire qu'ils devaient accomplir (et auquel nous les avons conduit progressivement auparavant).

Il apparaît alors que la gestualité, dans cette fonction iconique, s'apparente globalement aux représentations picturales de type schéma, avec pour fonction de mettre au jour, soit le fonctionnement des gestes articulatoires, soit l'organisation catégorielle de segments, en particulier grâce au dynamisme, contrôlé par l'enseignant, qu'elle permet, et surtout à l'adaptation instantanée à la situation qu'elle autorise, contrairement auxdits schémas.

La présentation, initiale et répétée, d'un énoncé ou d'un item en tâche de répétition, peut alors devenir tout à fait audiovisuelle, en ce que la présentation auditive s'accompagne de ce que l'on pourrait presque qualifier de pantomime articulatoire/gestuelle, destinée à mettre en relief tel ou tel trait phonétique de l'énoncé. Une telle pratique évoque, d'une part la *Réalité augmentée*, qui permet d'afficher certains des articulateurs non visibles sur un locuteur humain naturel, et d'autre part le *Langage Parlé Complété* (équivalent français du *Cued Speech*), qui permet de saisir à l'aide de la main et des doigts, des informations non-visibles (voisement, nasalité, mouvements de la langue), complétant ainsi la lecture labiale (Attina, Beautemps et Cathiard, 2002 ; Schwartz, Abry, Boë et Cathiard, 2002).

L'utilisation d'une symbolisation visuelle comme code-relai étant ainsi déjà exploitée au sein des techniques de rééducation en pathologie du langage<sup>238</sup>, on pourrait se risquer à imaginer que puisse être systématisé, dans un cadre didactique et pour un public donné, un ensemble de signes comparable à ceux des deux méthodes évoquées, même si une telle pratique est cependant exposée aux risques de surcharge cognitive pour les apprenants<sup>239</sup>, et doit donc être modulée en fonction de la situation, des items et du public.

Par ailleurs, outre son impact sur les caractéristiques infra-segmentales, la gestualité peut également servir, *via* les oppositions catégorielles dans l'espace, à opérer une segmentation phonémique de l'item présenté aux apprenants, et ainsi aider au développement de la conscience phonémique. On constate alors globalement que la gestualité remplit ici une fonction qu'assume, d'une certaine manière, l'écriture : la fonction de réification dans l'espace, en associant un percept visuel à un percept auditif pour l'apprenant. Cependant, contrairement à l'écriture, les symboles gestuels n'ont, a priori, pas de valeurs phonético-phonologiques déjà normées chez l'apprenant, échappant ainsi aux phénomènes de transferts phonographémiques. L'apprentissage des correspondances gestes-sons, destinés à faciliter celui du système sonore, évoque alors, de manière bien moins économe, celui des correspondances graphies-sons. On pourrait ainsi établir le parallèle suivant :

- Son → (micro) gestes articulatoires → (macro) gestes.
- Son → transcription phonétique fine → orthographe.

Cette dernière fonction de segmentation, si elle s'applique au niveau phonémique, peut également s'appliquer au niveau syllabique.

---

<sup>238</sup> Par exemple le système *AIRTAC*, décrit par Lambert (1997, p. 163), ou encore le *Visual Cue Programme* employé avec des patients agrammatiques (Van De Sandt-Koenderman et Bonta, 1998, p. 216).

<sup>239</sup> Et pour l'enseignant !

### II.1.2.3.2. Axe rythmique/syllabique

Dans ce cas, les trois catégories peuvent être sollicitées, et plus particulièrement la première et la dernière. En ce qui concerne le rythme, l'ensemble du corps peut être sollicité, avec ou sans déplacement dans l'espace, pour amplifier, et ainsi faire percevoir, le rythme de l'énoncé à répéter, que ce soit à travers un battement de bras, de main ou de pied par exemple. Dans le cas du français, ce « marquage rythmique » correspond également dans la plupart des cas à un certain découpage syllabique. La gestuelle rejoint alors ici encore, au niveau syllabique, l'une des fonctions essentielles de la représentation écrite : celle de la segmentation de la parole.

La segmentation syllabique peut parfois s'effectuer par « comptage » (les doigts de la main servent à compter le nombre de syllabes), mais aussi simplement par battement dans l'espace. L'apprentissage de la syllabation du français, qui dans le cas du japonais, est également lié à celui de la conscience phonémique, comme cela apparaît dans le cas des groupes consonantiques, peut en effet bénéficier de la gestualité, notamment à travers des techniques de « segmentation/concaténation », dans lesquelles la segmentation auditive de l'énoncé, puis la concaténation des éléments ainsi segmentés, s'accompagne de mouvements des bras et des mains. Dans ce cas, on constate que la gestuelle remplit à nouveau la fonction de réification de l'écriture, en permettant de délimiter et ainsi d'identifier les groupes syllabiques de l'item. Cependant, contrairement à une représentation orthographique qui possède un contenu segmental, la représentation gestuelle dans l'espace ne représente que ce que l'on pourrait associer à certaines « positions » sur l'axe syllabique, évitant ainsi toute interférence de contenu. Une telle représentation pourrait tout aussi bien être évoquée par une série de figures géométriques (comme les ellipses de nos tests C1, C2, D et E), mais perdrait alors, sur support traditionnel au moins, sa dynamique et son adaptativité.



### II.1.2.3.3. Axe accentuel/intonatif

Dans ce cas, ces sont les catégories 1) et 2) qui sont principalement sollicitées. Ici encore, la fonction du geste est essentiellement double : fonction de réification dans l'espace, destinée à faciliter la perception, et fonction d'amplification des caractéristiques phonétiques, destinée à orienter l'attention des apprenants. Lorsque l'on souligne le fait que les caractéristiques accentuelles et intonatives sont relativement<sup>240</sup> absentes de la représentation orthographique, on peut apprécier l'utilité du geste comme input visuel. Il faut alors noter que la visualisation de ces caractéristiques a reçu une attention particulière ces dernières années grâce au développement des outils multimédia, ce qui nous conduit au troisième type d'information visuelle, que nous avons qualifié de pictural. Cependant avant de conclure sur la gestualité, qui renvoie au domaine de la kinésique, et qui, en France, a notamment été traité par des auteurs comme Cosnier, il convient de mentionner rapidement son rôle dans le domaine de la compréhension et de la mémorisation.

### II.1.2.3.4. Compréhension et mémorisation

Le geste peut en effet servir à la compréhension des énoncés oraux, tant d'un point de vue sémantique, *via* sa fonction illustrative<sup>241</sup>, que « grammatical ». L'étude du rôle des gestes comme outils visuels de marquage grammatical conduit en effet Gullberg (1999, p. 69) à affirmer :

*« Gestures can handle grammatical problems. [...] By explicitly visualising grammatical relationships pertaining both to tense and co-reference, learners can reduce ambiguity and dramatically ameliorate listeners' comprehension ».*

---

<sup>240</sup> La ponctuation joue en effet un certain rôle dans la transcription de ces caractéristiques.

<sup>241</sup> Qui peut-être référentielle (représentation de l'objet, « pictomimique » dans la terminologie de Cosnier), déictique (repérage spatio-temporel, dans la situation d'énonciation ou dans le discours énoncé) ou autre encore.

On pourrait bien sûr ramener l'information visuelle gestuelle à une simple information picturale, mais rappelons qu'elle s'en distingue par son caractère dynamique (quoique certains outils informatisés permettent aujourd'hui de dynamiser les éléments et les supports picturaux), et surtout adaptatif : l'enseignant peut en effet rapidement ajuster son geste aux questions abordées et aux problèmes rencontrés en classe, contrairement à tout dispositif pré-programmé. Par ailleurs, le geste, étant par nature limité dans ses capacités représentationnelles, ne peut qu'évoquer certaines représentations, contrairement à certains supports picturaux beaucoup plus denses informativement. Si l'on ajoute à cela le fait que, tout en étant visible dans l'espace et pouvant être répété, le geste ne possède pas la permanence de l'écrit ou de l'image<sup>242</sup>, il s'apparente à un dispositif relativement fugace, optionnel, que l'on pourrait presque comparer, dans une certaine mesure aux escamots d'aide que l'on trouve dans certains didacticiels (Koulayan, Detey et Sagaz, 2003 ; Sagaz, Koulayan et Detey, 2004). Le geste entraînerait alors moins de risques de surcharge cognitive, comparé à l'image ou au schéma fixe.

En ce qui concerne la mémorisation, comme nous l'avons évoqué à propos de celle des gestes articulatoires dans le cas du [R], le geste peut également servir d'indice visuel utile à l'encodage mnésique puis au rappel. Tellier (2005) s'est ainsi particulièrement intéressée à l'effet des gestes de type iconique, visant à faciliter l'accès au sens, sur la mémorisation du lexique en classe de langue auprès de jeunes enfants, en se basant sur certains résultats avérés positifs auprès d'adultes. Sa conclusion est la suivante (p. 5) :

---

<sup>242</sup> Ce qui évoque la thèse de L.-J. Calvet (*Histoire de l'écriture*, Paris, Hachette, 1996), présentée par Durand (2000b, pp. 53-54). Pour Calvet, qui conteste la subordination de l'écrit à l'oral, l'ensemble des moyens d'expression humains peut être divisé en deux grandes catégories, décrites ainsi par Durand (p. 54) :

*« La gestualité, comme la parole, qui est par définition fugace, et la picturalité, comme la peinture pariétale préhistorique, qui permet la pérennité du message. La thèse que défend cet auteur [(Calvet)] est la suivante : "Langue et écriture procèdent de deux ensemble signifiants tout à fait différents à l'origine, la gestualité et la picturalité. Leurs rapports relèvent de la rencontre de ces deux ensembles qui continuent par ailleurs leurs vies autonomes : l'écriture est de la picturalité asservie à la gestualité" (1996 : 23) ».*

*« Il nous apparaît donc que le discours de l'enseignant gagne à être illustré visuellement par des images mais aussi, et surtout, par des gestes (support plus facile à produire spontanément et naturellement en interaction pédagogique) ».*

Cependant, si le geste se caractérise par son dynamisme, qui peut jouer un rôle attentionnel non-négligeable, ainsi que par son adaptabilité, essentielle dans un cadre didactique réaliste, la fonction de réification qu'il exerce peut, dans de nombreux cas, être ramenée à celle de l'information visuelle picturale.

#### II.1.2.4. Information picturale

Il ne s'agit pas pour nous de nous plonger dans l'analyse du rôle de l'image en didactique des langues, même en restreignant le domaine à celui de la « prononciation », car cela constituerait un sujet de thèse, voire de plusieurs, à part entière. En outre, cela nécessiterait de nous plonger davantage dans les domaines de la sémiotique et de l'interculturel, afin de mieux comprendre les phénomènes interprétatifs liés à la perception de l'image dans le cas d'apprenants de langue étrangère, une réflexion qui a été amorcée depuis longtemps par certains (Rivenc, 1973/2000). Cependant, il semble essentiel de relever certaines problématiques et certaines données relatives à ce thème, puisqu'il sous-tend une grande partie de notre problématique, en particulier au vu des développements des outils multimédiatisés en didactique des langues.

Ce que nous avons nommé « information picturale » englobe ici tout ce qui relève du domaine graphique, excepté l'orthographe et les codes de transcription phonétiques tels que l'A.P.I. Il existe donc de nombreux types d'informations de ce genre (Yazdani et Barker, 2000), de la photo au dessin en passant par le schéma. Deux fonctions essentielles nous occupent ici :

- Celle relative à l'accès au sens, pour des énoncés ou des items lexicaux, accès au sens qui pourrait faciliter l'encodage mnésique de la forme de l'item considéré.

Celle-ci renvoie donc à la compréhension et à la mémorisation de l'oral, avec un focus sur le sens. On pourrait mentionner les images référentielles par exemple, ou une série d'images évoquant l'item en question.

- Celle relative à l'accès à la forme, phonétique et/ou phonologique, cette forme pouvant s'exprimer de plusieurs manières, selon que l'on interprète celle-ci en termes d'unités phonologiques, de gestes articulatoires, ou de valeurs acoustiques. Celle-ci renvoie donc à la perception et à la mémorisation de l'oral, avec un focus sur la forme. On pourrait mentionner l'image (photo ou schéma) des articulateurs pour la production d'un segment donné, ou bien une série d'ellipses, symbolisant le nombre de syllabes de l'item, dont la dernière serait colorée (pour indiquer l'accentuation de la dernière syllabe).

De multiples possibilités sont bien évidemment envisageables, et l'on en trouve des échantillons dans l'ensemble des outils pédagogiques disponibles sur le marché. Il ne s'agit pas ici pour nous d'examiner à proprement parler le rôle de l'image dans la compréhension et la mémorisation de l'oral, un sujet qui a alimenté et continue d'alimenter de nombreuses recherches en didactique des langues (Merlet, 1998). En revanche, nous pouvons nous demander dans quelle mesure l'information picturale pourrait permettre de remédier aux possibles « problèmes » posés par les représentations orthographiques, au moins en début d'apprentissage, dans la perception et la mémorisation de l'oral, à travers les fonctions de réification et de « mémorisation » que nous avons présentées.

Nous pouvons alors reprendre en grande partie ce que nous avons présenté à propos du geste, tout en soulignant certaines des spécificités de l'information picturale, en particulier sa permanence, qui la rapproche de l'écrit et la distingue de l'oral. Il va de soi que la discussion ne peut être productive qu'à condition de spécifier les éléments (traits, phonèmes, gestes articulatoires, mot phonologique ou autres) à percevoir et à

mémoriser. Cependant, nous devons au préalable rappeler certaines généralités sur l'articulation entre modalité auditive et visuelle dans le traitement du langage.

Nous pouvons d'abord noter l'importance du domaine « visuel », prise, depuis quelques décennies, dans certains courants linguistiques. Ainsi, des travaux en grammaire cognitive, comme ceux de Langacker (Laks, 1996, p. 92), nous paraissent remarquables<sup>243</sup> pour l'importance qu'ils semblent accorder aux percepts visuels : Laks souligne que « tant du point de vue neurobiologique (Sereno, 1990), que du point de vue linguistique (Langacker, 1987) la grammaire cognitive pose donc une intégration horizontale entre vision et langage » (1996, p. 62). Cette caractéristique peut ensuite mener à une description grammaticale didactique beaucoup plus imagée sous forme de grammaire narrative (Lapaire, 2001), et donc possiblement plus accessible pour certains—et non pour l'ensemble—des apprenants. La distinction radicale entre la grammaire « externe » et la grammaire « interne » y étant par ailleurs fermement maintenue (Laks, 1996, p. 156). Dans une optique d'enseignement et de formation, cette distinction nous paraît fondamentale.

Comme l'indique Laks (1996, p. 54), les travaux de Sereno « montrent l'importance considérable de la participation des aires visuelles au développement du langage. De nombreux résultats neurophysiologiques récents suggèrent en effet que la compréhension du langage, tant du point de vue de la construction de sens que de l'interprétation phonologique, reposent sur des fonctionnalités de type visuel ». Il serait tout à fait incongru de ne pas prendre en compte ce genre de données lorsque l'on s'intéresse de près aux problèmes de format de présentation du matériau linguistique en classe de langue, en particulier dans le cadre de conception d'« outils techno-pédagogiques » (Koulayan, 1999) hypermédiatisés dans le domaine de l'enseignement/apprentissage des langues. L'intérêt pour nous des travaux sur la conceptualisation spatiale et sur la métaphore est lié au fait que cette dernière apparaît

---

<sup>243</sup> D'un point de vue descriptif et didactique au moins: nous ne nous préoccupons que très partiellement ici de leur valeur heuristique.

comme un outil utile de communication, car plus abstraite et donc moins spécifique (Collet, 2000, p. 39), et nous ne nous préoccupons pas ici de savoir si la « faculté de langage » est inextricablement liée à une modalité sensorielle ou non, certaines données suggérant le contraire, comme le rapporte Chomsky<sup>244</sup> :

*« Though highly specialised, the language faculty is not tied to specific sensory modalities, contrary to what was assumed long ago. Thus, the sign language of the deaf is structurally much like spoken language, and the course of acquisition is very similar. Large-scale sensory deficit seems to have limited effect on language acquisition. Blind children acquire language as the sighted do, even colour terms and words for visual experience like « see » and « look ». [...] The analytic mechanisms of the language faculty seem to be triggered in much the same ways whether the input is auditory, visual, even tactual, and seem to be localised in the same brain areas ».*

Dans le domaine de la didactique des langues, l'image a particulièrement été exploitée dans la problématique SGAV, qui en a justifié l'emploi et a su faire évoluer le type d'image exploité à travers les différentes « générations » de méthode issues de ladite méthodologie (Rivenc, 1973/2000, 2000, 2003, Rivenc-Chiclet, 1989, 2001)<sup>245</sup>. Nous-mêmes, dans notre propre pratique de classe, avons pu constater le rôle mnésique des images des principales méthodes SGAV, puisque celles-ci accompagnent la présentation des énoncés oraux afin d'en faciliter la compréhension et la mémorisation, comme en atteste leur rôle ultérieur d'indice de rappel dans des activités de classe très concrètes. Par ailleurs, motivés par le désir de mettre à profit l'ensemble des capacités perceptives des apprenants de langue étrangère, en fournissant un input informatif suffisamment riche et diversifié, les didacticiens sont fondamentalement concernés par la question du rôle de la multimodalité dans la perception, la compréhension et

---

<sup>244</sup> Chomsky, N., « Language and Nature », *Mind*, n° 104, p. 16, cité par Durand et Laks (2002, p. 19).

<sup>245</sup> L'un des modules de formation de la maîtrise FLE du Département des Sciences du Langage de l'Université de Toulouse-Le Mirail était ainsi entièrement consacré au rôle de l'image dans le domaine didactique.

l'apprentissage du matériel linguistique qu'ils présentent. L'une des interrogations fondamentales auxquelles ils doivent alors tenter de répondre est la suivante : sous quelles conditions la multimodalité est-elle plus efficace que l'unimodalité (Nespoulous, 2000) ?

On connaît en effet les risques de surcharge/saturation cognitive et de désorientation liée à une surabondance d'informations dont la conjonction rend caduque toute perspective de mémorisation ou d'apprentissage (informations *versus* connaissances), en raison, notamment, de conflits procéduraux entre les traitements inter-modaux (Bertin, 2001, pp. 130-131). Ainsi la présence d'images peut conduire à une surcharge cognitive, leur rôle (facilitateur ou complexifiant) étant, comme le souligne la théorie de la charge cognitive de Sweller (Tricot, 1998), « fonction du degré d'expertise » de l'apprenant en L2 (Merlet et Gaonac'h, 1995, p. 288). Cela est lié à « l'effet de redondance de l'information présentée », qui se distingue de « l'effet de dissociation de l'attention » de la manière suivante (Tricot, 1998, p. 53) :

*« L'effet de dissociation de l'attention s'exprime quand les sources d'informations ne sont pas compréhensibles de façon isolée, tandis que l'effet de la redondance s'exprime quand il est possible de comprendre ces informations de façon isolée. [...] Tout se passe comme si le même matériel considéré comme intégré pour des sujets novices et ayant un effet positif sur leurs performances, pouvait être considéré comme redondant pour des sujets experts, et aurait un effet négatif sur leurs performances ».*

On sait également que, si l'imagerie, lorsqu'elle est mise en œuvre par un individu, « occupe » une partie du dispositif de traitement que cet individu utilise pour traiter l'information perceptive, « la mise en œuvre simultanée d'une activité d'imagerie et d'une activité perceptive risque de se faire au détriment de l'une ou de l'autre » (Denis, 1994, p. 87). Il semblerait donc que la multimodalité doive être manipulée avec précaution.

La deuxième question essentielle pour le didacticien est la suivante : sous quelles conditions le rapport entre perception verbale et non verbale est-il le plus efficace ? (rapport : modalité/stimuli et combinaisons possibles) (Caro et Bétrancourt, 1998 ; Jamet, 2001 ; Merlet, 1998 ; Tricot, 1998). Comme l'indique Merlet (1998, pp. 153-154), dans une sous-partie intitulée « Discussion générale : les conditions de l'efficacité d'une présentation multimodale » :

*« L'efficacité de l'illustration semble en fait dépendre d'un certain nombre de facteurs ayant trait notamment aux caractéristiques du matériel présenté, mais aussi aux caractéristiques individuelles des apprenants ».*

Nous serons alors naturellement tentés d'évoquer les caractéristiques cognitives *spécifiques* à certains *groupes* d'apprenants (Detey, 2003), et de plaider pour une exploitation rigoureuse des analyses comme celle de Levin et de ses collaborateurs, qui ont dressé une typologie des illustrations en fonction de leurs rôles vis-à-vis du matériel verbal (Merlet, 1998). Cette réflexion sur l'image doit également bénéficier des travaux du domaine de l'imagerie mentale, puisque, comme l'indique Denis (1994, p. 79), tout en prenant garde de ne pas confondre l'objet et l'image mentale de cet objet, et en soulignant que l'image est à la fois « *produit* de processus et *lieu d'application* de processus »(p. 73) :

*« Il apparaît fondé, à l'examen des données, de reconnaître la similitude structurale des événements psychologiques engendrés par la perception des objets et de ceux qui sont produits par l'imagination de ces mêmes objets. L'image d'un objet reflète bien, dans sa structure, l'organisation interne de l'expérience perceptive de cet objet. Cette analogie au niveau de « l'apparence » est prolongée par l'analogie de certains au moins des processus mis en œuvre pendant l'activité perceptive et l'activité imaginative. Les images pourraient ainsi faire l'objet de traitements similaires à ceux qui s'appliquent aux stimulus physiques présents dans le champ perceptif ».*



La deuxième question que nous avons formulée, comme la première, renvoie donc à la problématique centrale du format de mémorisation, et donc aux différents modèles proposés (pour la mémoire permanente en particulier) (Mémoire Verbale *versus* Mémoire Imagée) (Merlet, 1998, p. 141 ; Merlet et Gaonac'h, 1995, p. 273).

Dans le domaine de la didactique et de l'ergonomie des hypermédias, et plus généralement de la psychologie ergonomique, plusieurs résultats commencent heureusement à prendre assise<sup>246</sup> (Tricot, 1998 ; Lemerrier et Tricot, en soumission). Comme le notent Tricot, Détienne et Bastien (2003, p. 7), l'ambition générale de certains travaux en ergonomie cognitive est de viser « la description de processus cognitifs généraux et l'élaboration de recommandations générales pour la conception de situations de travail ou d'outils », ce qu'ils illustrent à l'aide de la théorie de la charge cognitive de Sweller, qui conduit aux recommandations suivantes :

*« Par exemple, si un opérateur novice doit comprendre un schéma technique complexe, alors il faut intégrer les commentaires linguistiques dans le schéma, sur chaque point précis, plutôt que de disposer ces commentaires à l'extérieur du schéma. Quand cela est possible, ces commentaires linguistiques seront plutôt présentés oralement que visuellement ».*

Ces recommandations sont précisées par Tricot (1998), dans la présentation qu'il donne des travaux de Sweller, notamment les études conduites sur l'effet de « dissociation de l'attention » ou « dédoublement de l'attention », présenté comme tel par Tricot (1998, p. 44) :

*« L'effet de dissociation se produit quand les sujets doivent traiter des sources d'information multiples dont l'intégration doit être opérée mentalement afin que le sens puisse être inféré du matériel présenté ».*

---

<sup>246</sup> Voir à ce propos les travaux de A. Tricot, consultables à l'adresse suivante, en mars 2005:  
<http://perso.wanadoo.fr/andre.tricot/AndreTricot.html>.

Cet effet semblerait pouvoir être éliminé ou minimisé sous certaines conditions (pp. 51-52) : « en intégrant physiquement les informations textuelles et imagées mutuellement référencées [(principe de « proximité de présentation » ou encore principe de « contiguïté » de Mayer (Lemercier et Tricot, en soumission)], ou en utilisant le canal auditif pour présenter les informations verbales [(instruction en mode duel)] ». De tels travaux parviennent ainsi à fournir des recommandations didactiques très concrètes, notamment, en ce qui nous concerne, (Tricot, 1998, pp. 55-57) :

- 1) Le travail sur des exemples de problèmes résolus : « *Toute chose étant égale par ailleurs, travailler sur des exemples de problèmes résolus entraîne un meilleur apprentissage que travailler sur des exemples de problèmes non-résolus. [...] Il faut cependant noter que l'usage exclusif de problèmes résolus n'est pas une bonne solution. Cela peut en effet inhiber, chez le sujet, la génération de solutions nouvelles et créatives.* »
- 2) L'intégration des informations pour éliminer l'effet de dissociation de l'attention : « *L'élimination de l'effet de dissociation ne peut avoir lieu que quand il y a, au départ, un grand besoin d'attention, c'est-à-dire une situation difficile à traiter.* »
- 3) L'utilisation de deux modalités sensorielles plutôt qu'une : « *La charge en MDT [(mémoire de travail)] est moins importante quand le même matériel est présenté en utilisant les canaux auditifs et visuels, que quand on utilise le seul canal visuel. La phase d'étude de configurations "matériel verbal / matériel imagé" mutuellement référencé est plus efficace quand le matériel verbal est présenté oralement que quand il est présenté par écrit* »<sup>247</sup>.
- 4) La non-présentation d'information redondante : « *Le problème posé par l'étude de cet effet réside dans la définition de la redondance, qui, selon toute vraisemblance,*

---

<sup>247</sup> Notons que, à ce sujet, Tricot, à la suite de Sweller, fait référence (p. 44) au modèle mnésique de Baddeley à deux processeurs: le « calepin visuo-spatial » et la « boucle phonologique ».

*ne peut pas être considérée seulement de façon intrinsèque, mais dépend largement du niveau d'expertise du sujet dans le domaine de la situation traitée ».*

Certaines de ces études en psychologie ergonomique font notamment usage de dissociations entre modalités d'une part, et formats d'autre part, comme outils d'investigation, à l'instar de ce qu'évoque Prablanc à propos des dissociations perceptivo-motrices (1997, p. 95) : « La manipulation expérimentale des informations sensorielles est un outil extrêmement puissant : car il permet d'imposer soit des décalages relatifs d'une modalité sensorielle par rapport à l'autre, soit des décalages absolus lorsqu'une seule information sensorielle est disponible ». Il conviendrait ainsi de mener davantage expérimentations de ce type dans le cadre de la conception d'outils didactiques, comme le font, par exemple, les chercheurs engagés dans le projet « Modèle de conception de documents multimédia à destination des adultes », piloté par le Laboratoire Jacques Lordat de l'Université de Toulouse-Le Mirail, et dans lequel cinq conditions de présentation du matériel verbal associé à du matériel pictural (dessin) sont utilisées pour un public d'enfants :

- Oral seul,
- Ecrit seul,
- Oral+Ecrit (simultané),
- Oral, puis Ecrit (différé),
- Ecrit, puis Oral (différé).

Les questions abordées dans ce projet sont les suivantes :

- « Existe-t-il une modalité d'encodage plus prégnante qu'une autre pour la mémorisation ? »,
- « Deux modalités d'encodage facilitent ou sont un obstacle à la mémorisation lexicale ? »,
- « Si l'image sonore en L2 est proche de l'image sonore en L1, cela facilite-t-il la mémorisation ? »,

- « Si l'image écrite en L2 est proche de l'image écrite en L1, cela facilite-t-il la mémorisation ? ».

Certains résultats (Dat et Spanghero-Gaillard, 2005, p. 109) indiquent ainsi que :

*« La présentation conjointe de deux modalités d'encodage (orale et écrite) d'un support imagé ne favorise pas en soi la mémorisation ou la compréhension du lexique présenté. Cependant, [...] l'image induit la mémorisation en tâche de rappel, lorsque la phase de consultation est conduite à partir d'un support imagé ».*

Les auteurs soulignent également un point qui nous concerne, à savoir « la complémentarité de l'encodage à l'écrit par rapport à l'oral dans la mémorisation lexicale » et insiste sur le fait que, en situation de classe, l'on doit prendre garde à ce que la phase de mémorisation, « si elle se base sur des items écrits, n'influe pas négativement sur la correction phonétique par un mauvais rapport phonie-graphie ».

Une question didactique supplémentaire que l'on pourrait poser, et qui, à notre connaissance, n'a pas été traitée expérimentalement, est celle de la présentation, collective ou non, qui va de pair avec celle de la gestion, guidée ou non, de l'image en classe de langue : le bénéfice de l'image est-il modulé ou non par une présentation « libre » (par exemple sur un manuel individuel) ou « contrainte » (par exemple sur un écran mural, comme dans les méthode SGAV), voire « orientée » (à l'aide d'un « pointeur » manipulé par l'enseignant), des éléments picturaux. Il serait en effet souhaitable de pouvoir vérifier rigoureusement les intuitions didactiques que l'on peut se former lorsque l'on a la possibilité de manipuler les deux types de format, qui, du point de vue de la gestion interactionnelle et attentionnelle de la classe, ne présentent pas les mêmes caractéristiques.

### II.1.2.5. Information graphique

Dans cette sous-partie, nous nous concentrons davantage sur le rôle de l'information visuelle graphique dans la perception et l'apprentissage des caractéristiques phonéto-phonologiques de la parole.

#### II.1.2.5.1. Caractéristiques prosodiques

En ce qui concerne la perception de l'oral, plusieurs études ont été conduites ces dernières années sur l'apport des représentations visuelles du signal sonore linguistique à l'apprentissage de la langue, notamment dans le cadre de l'Entraînement à la Prononciation Assisté par Ordinateur (« Computer Assisted Pronunciation Training », CAPT) (Neri, Cucchiari et Strik, 2002). On constate que ces travaux se sont essentiellement concentrés sur les représentations de la dimension prosodique de la parole (amplitude, fréquence, etc.)<sup>248</sup>, en particulier l'intonation (Chun, 1998), et les travaux sur les segments dans ce domaine semblent moins nombreux<sup>249</sup>. A ces courbes et ces graphiques, en tant qu'input ou feedback visuel, est parfois ajoutée une dimension réflexive *via* des explications métalinguistiques en guise de retour correctif (Germain et Martin, 2000). D'autres auteurs (Cazade, 1999) se sont également interrogés plus généralement sur les différents types de présentation visuelle

---

<sup>248</sup> Cela apparaît nettement lorsque l'on parcourt la très riche et très utile bibliographie de Joaquim Llisterri de l'Université Autonome de Barcelone (Espagne), relative à l'enseignement de la prononciation assisté par ordinateur (<http://liceu.uab.es/~joaquim/home.html>). Il est également intéressant d'observer, dans l'organisation de cette bibliographie, très actualisée, que deux des cinq sous-parties sont consacrées, d'une part à l'input visuel (« Visual feedback as an aid for pronunciation teaching ») et d'autre part à la modification de l'input auditif (« Perceptual cues enhancement and modified auditory feedback for pronunciation teaching »), deux thèmes qui concernent donc, à la fois nos préoccupations dans cette étude, et d'autre part l'actualité du domaine.

<sup>249</sup> On citera néanmoins les travaux de Lambacher (2002) sur la prononciation des segments en L2, puisque ces derniers concernent l'utilisation de feedback visuel électronique (« Electronic Visual Feedback », EVF) pour l'amélioration de la prononciation d'apprenants japonais d'anglais, en particulier dans le cas du /t/ américain (Murakawa et Lambacher, 1996), pour lequel les auteurs semblent obtenir des résultats encourageants, comme ils l'indiquent dans leur conclusion : « *The results, as a whole, lend support to the claim that electronic visual feedback can effectively help learners improve their L2 pronunciation skills. EVF is effective as a teaching and learning tool because it allows learners to visualize their pronunciation through associating the acoustic patterns on the screen with the sounds they are producing* ». Cependant, comme le reconnaît Lambacher (2002), les représentations visuelles utilisées, malgré leur pluralité (courbes et graphiques représentant les analyses de fréquence, de durée, d'amplitude, etc., accompagnés d'une transcription phonétique fine) et le lien établi entre dimensions articulation et acoustique, restent parfois trop « techniques » pour les apprenants.

(essentiellement en termes de courbes et de schémas liés à la production articulatoire) pouvant être utiles à l'amélioration de la « prononciation » des apprenants de langue étrangère. La réflexion de l'auteur s'inscrit cependant dans un contexte didactique à options méthodologiques particulières, qui n'est pas directement transposable à toute situation d'enseignement/apprentissage du FLE, et dans lequel le rôle de l'écrit nous semble quelque peu mésestimé (pp. 21-24) :

*« La présence du texte, référent commun majeur à tous les types d'apprenants, peut fortement aider à étudier et à déchiffrer les courbes proposées. [...] Pour pouvoir tirer quelque parti que ce soit, aussi bien des courbes sonores de nature diverses qui ont été envisagées que des animations graphiques [...], il semble bien qu'une initiation aux bases élémentaires de la phonétique soit indispensable, en préalable ou/et en accès libre au moment où l'apprenant le jugera utile. S'il est donné à l'apprenant de comprendre avec les yeux autant qu'avec les oreilles les différences qui existent entre un bruit et un son, entre les sonorités simples et complexes, d'aborder les principes de la phonation, [...], etc., et si tous ces éléments lui sont proposés avec la possibilité de choisir [...], alors les explications phonétiques deviendront aussi indissociables des logiciels de langues que le sont [...] les oscillogrammes aujourd'hui ».*

Si nous partageons l'intérêt de l'auteur pour la dimension métalinguistique de l'enseignement/apprentissage, il nous semble néanmoins qu'une telle vision des choses est quelque peu simplificatrice, tant d'un point de vue didactique que d'un point de vue cognitif. Dans tous les cas, ces études reflètent l'attention portée aux aspects suprasegmentaux de la « prononciation » (Guimbretière, 2001), aspects essentiels tant dans l'apprentissage que dans la communication, qui avaient très longtemps été négligés en didactique au profit des phonèmes.

D'autres travaux se sont intéressés aux aspects accentuels des énoncés, en combinant une présentation auditive à une représentation orthographique, celle-ci étant chargée d'un codage visuel supplémentaire. Fraser (2000, p. 2), par exemple, a élaboré un

didacticiel dans lequel ce codage visuel supplémentaire a pour fonction d'augmenter la discriminabilité auditive :

*« This module lets them hear sentences with the stressed words visually highlighted in real time, to help them perceive the difference between the stressed words and the others. Module 3 carries the same methodology down to the level of syllables stress within words, giving learners practice in hearing stressed syllables with similar visual reinforcement ».*

Une question similaire a été abordée par Stenton, Péchou, Vaillant-Sirdey et Tricot (2005) lors de la conception d'un logiciel d'enseignement (basé sur le système auteur SWANS, c'est-à-dire *Synchronised Web Authoring Notation System*) destiné à faciliter la perception, et ainsi l'encodage mnésique puis la production en rappel, de l'accent tonique (et des syllabes non-accentuées) en anglais par des apprenants français. Les auteurs ont décidé de tirer parti de la possibilité d'un double encodage (auditif et visuel) de l'accent en présentant en synchronie l'énoncé oral et la représentation orthographique dans laquelle la syllabe accentuée est mise en exergue grâce à une couleur et un pas de police différents (la syllabe accentuée étant de taille plus grande et de couleur vive). Ils ont également évalué l'effet de ce double codage en distinguant les experts des novices, selon la théorie de la charge cognitive de Sweller qui conduit à l'effet de renversement de l'expertise (*expertise reversal effect*) : « ce qui facilite l'apprentissage chez les novices peut gêner le traitement chez les experts » (p. 308). Les résultats de leurs tests, qui incluaient trois modalités différentes de présentation des items lexicaux, semblent encourageants, puisqu'ils obtiennent un effet positif du double codage.

#### II.1.2.5.2. Segmentation de la parole

Alors que les aspects intonatifs et accentuels semblent donc être étudiés à l'heure actuelle dans le domaine de la didactique et des hypermédias, la question du rythme, de la syllabification et plus généralement de la segmentation de la parole, semble en

revanche moins attractive, même si celle de l'alignement entre séquences sonores, continues, et séquences orthographiques, discrètes, est un thème d'actualité en linguistique de corpus notamment<sup>250</sup>. La raison en est peut-être que l'orthographe est considérée par de nombreux didacticiens comme un outil « évident », voire « suffisant » de la segmentation de la parole. Or, comme nous l'avons démontré tout au long de notre étude, l'intérêt de l'utilisation de l'orthographe comme support de perception et d'apprentissage de la parole en langue étrangère, dans le cas du français au moins, est loin d'être « évident », négligeant en particulier la distinction entre syllabation de l'oral et de l'écrit. En ce qui concerne la segmentation rythmique d'un énoncé en tâche de répétition, il nous est ainsi apparu, de manière empirique et informelle, que la reproduction de l'énoncé par les étudiants était facilitée par une visualisation de type gestuel (battement des bras ou des mains par exemple), mais pas par la présentation de la transcription orthographique, qui entraînait l'activation de processus supplémentaires et alourdissait la charge cognitive de la tâche en question. Comme le note Cazade (1999, p. 25) :

*« La segmentation automatique, prenant en compte les baisses de volume ou d'intensité dans le flux sonore pour marquer l'endroit où il est possible d'insérer un repère ou un silence plus ou moins grand, propre à faciliter la compréhension ou la répétition partielle ultérieurement est également une piste à creuser et une fonctionnalité qui devrait être généralisée. Le didacticiel LAVAC [...] permet déjà ce travail aussi bien pour les productions sonores que pour l'étude de documents vidéos ».*

Le système LAVAC (« Laboratoire Audio-Visuel Actif-Comparatif »), que mentionne Cazade fait partie des réalisations du Laboratoire LAIRDIL (Laboratoire Interuniversitaire de Recherche en Didactique des Langues) de Toulouse. Comme l'indique Toma (2000, p. 119), l'un de ses concepteurs :

---

<sup>250</sup> Voir par exemple l'utilisation du logiciel PRAAT dans le projet PFC (Delais-Roussarie, Durand, Lyche, Meqgori et Tarrier, 2002; Tarrier et Auran, 2004).



*« Un bon système de création de cours devrait donc permettre le séquençage automatique en temps réel de la piste son et introduire après chaque séquence un temps de réponse proportionnel à la durée de chaque séquence créée. Ce taux doit de plus être modifiable même en présence des étudiants si la valeur du séquençage s'avère inadaptée au niveau des étudiants (séquence trop courte pour des niveaux forts ou trop longues pour des niveaux faibles). »*

Toma insiste donc sur l'adaptabilité du matériel et du séquençage au type d'apprenant. L'un des « pré-requis » à la conception de son système LAVAC est présenté comme tel (p. 121) : « Séquençage automatique du son avec visualisation des séquences et insertion après chaque séquence d'un temps de réponse d'une durée proportionnelle à la durée de la séquence créée ». Il nous faut alors noter l'utilisation du canal visuel pour faciliter la perception auditive (p. 122) : « La visualisation du son se trouve être un moyen efficace de permettre à l'apprenant de repérer des sons qu'il n'entendrait pas sans cela ».

Cette dernière remarque renvoie pleinement au sujet de notre thèse, puisque comme nous l'avons observé dans les résultats de nos tests A et F (identification et discrimination phonographémiques segmentales), d'une part, et C2, D et E (comptage syllabique uni- et bi-modal), d'autre part, la visualisation orthographique des items auditifs ne permet pas toujours de remédier à la « surdité » phonético-phonologique qu'évoque Toma. De manière tout à fait contre-intuitive, il est en effet apparu que la vision des graphèmes <r> et <l>, pour nos apprenants japonais de FLE, ne permettait pas toujours de distinguer ces deux catégories phonologiques, et que la représentation orthographique des groupes consonantiques <CC> (qui ne présente donc ni voyelle ni espace blanc graphique entre les deux graphèmes) ne permettait pas toujours d'empêcher l'apparition d'une épenthèse vocalique. Comme nous l'avons expliqué, les phénomènes en jeu sont en effet complexes, et nécessitent une analyse à la fois linguistique et psycholinguistique des deux systèmes langagiers en présence

(phonétique, phonologique et phonographémique), ainsi que des processus à l'œuvre en perception de la parole et en perception de l'écrit.

Néanmoins, malgré leurs faiblesses, de tels travaux soulignent l'importance potentielle du canal visuel dans la gestion temporelle de la perception auditive, *via* la réification dans l'espace, ainsi que dans la tâche fondamentale de segmentation et d'identification des unités de la chaîne parlée, particulièrement liée, pour des apprenants japonais ou chinois par exemple, au développement de la conscience phonémique. Outre cette dernière, essentielle à l'apprentissage du français, le canal visuel pourrait également aider à la perception de caractéristiques phonético-phonologiques, telles que le voisement ou la nasalité. On doit alors une fois encore s'interroger sur la nature du code visuel le plus à même de faciliter de telles opérations.

#### II.1.2.5.3. Code graphique et modifications orthographiques

Lorsque l'on examine certains manuels d'enseignement du FLE, on s'aperçoit que les représentations orthographiques sollicitées dans les activités dédiées à l'apprentissage de la « prononciation » subissent parfois certaines modifications destinées à permettre aux apprenants de mieux appréhender l'oral. Cela est particulièrement notable dans le cas du « e muet » (que l'auteur du manuel « barre » graphiquement, ou demande aux apprenants de « barrer ») ou de la liaison (que l'auteur signale par un « arc » graphique, ou demande aux apprenants de signaler par un « arc » graphique sous la représentation orthographique). Les effacements, vocaliques ou consonantiques, entre autres phénomènes notables en français parlé « familier », sembleraient ainsi pouvoir être « réifiés » à l'aide d'une modification, déjà accomplie ou à accomplir par les apprenants, de la représentation orthographique. De nombreux éléments graphiques sont ainsi sollicités pour rendre compte de phénomènes oraux initialement non transcrits, ou du moins théoriquement non perceptibles à l'écrit, voire à l'oral, par les apprenants, comme par exemple un trait horizontal, un soulignement ou un changement chromatique pour signaler l'accentuation de telle unité.

Ainsi, dans une tâche de répétition avec pour cible l'énoncé « Qu'est-ce que tu regardes à la fenêtre ? », si l'on souhaite proposer un support visuel aux apprenants, plusieurs possibilités sont envisageables, selon les caractéristiques phonéto-phonologiques que l'on souhaite mettre en exergue. En voici quelques exemples, qui pourraient peut-être permettre d'insister sur le respect des caractéristiques rythmiques, et dans une certaine mesure syllabiques, de l'énoncé :

Possibilité 1 :

Représentation orthographique dans laquelle sont barrés les « e muets » et dans laquelle les liaisons et enchaînements sont signalés par des « arcs » graphiques.

Possibilité 2 :

Modifier l'espacement des unités graphiques, avec ou sans signe graphique supplémentaire comme des tirets ou des barres obliques, de manière à se rapprocher de la syllabation de l'oral :

« Qu'est-ce / que / tu / re / garde / à / la / fe / nê / tre ? »

Mais, comme cela apparaît nettement ici, se pose le problème d'écart entre syllabation de l'oral (variable selon la variété de français, en particulier vis-à-vis du « e muet ») et de l'écrit (qui présente de nombreux signes « muets »), même si l'on combine les possibilités 1 et 2.

Possibilité 3 :

Proposer une transcription en orthographe « simplifiée » (que l'on pourrait également qualifier d'orthographiquement « erronée ») :

« ke skeu tur gar da laf netr ? »

Une telle transcription devrait alors être adaptée au public (valeur phonétophonologique du graphème <e> en L1 par exemple) et aux difficultés rencontrées (pas de simplification lorsque celle-ci n'est pas nécessaire).

De multiples possibilités sont donc envisageables, possibilités qui doivent notamment profiter des recherches en ergonomie (Fischer, 1979 ; Lefton, 1979 ; Tricot et Rouet, 1998, Yazdani et Barker, 2000), mais qui doivent impérativement être soumises à validation expérimentale dans des situations didactiques réalistes<sup>251</sup>.

Cependant, la dernière possibilité que nous avons évoquée, outre son effet vraisemblablement négatif sur l'apprentissage des correspondances phonographémiques et de l'orthographe française, nous rappelle ce que l'on peut trouver dans certains ouvrages commerciaux d'introduction aux langues étrangères, du type « Apprendre la langue X en Y jours », essentiellement destinés à fournir quelques rudiments linguistiques à des voyageurs, touristes ou professionnels, non-linguistes, qui vont brièvement séjourner dans le pays de la langue en question. Cela renvoie à ce que l'on pourrait qualifier d' « orthographe phonétique ». Or, de la même manière que notre argumentaire global ne vise en aucun cas à « bannir » l'orthographe du cours de langue, sous peine de déficiences sévères dans l'appréhension de l'écrit, tant en perception qu'en production (puisque notre argument vise au contraire à insister sur l'importance de l'enseignement/apprentissage des correspondances entre oral et écrit dès les débuts de l'apprentissage), l'utilisation d'un tel code ne semblerait pas résoudre les problèmes, voire même en rajouter.

---

<sup>251</sup> Au Japon, par exemple, certains enseignants, suffisamment formés en phonétique et en phonologie, pratiquent déjà, en partie au moins, cette intégration entre activités de lecture et enseignement explicite de la prononciation, en faisant usage de symboles, plus ou moins idiosyncrasiques, qui viennent modifier le texte orthographique, afin d'attirer l'attention des apprenants sur certaines caractéristiques comme les liaisons, les consonnes muettes, les enchaînements, les pauses, etc. (Azra, 2003). Généralement, c'est aux étudiants de porter ces symboles sur le texte, ce qui les place en situation d'analyse *active* des correspondances entre phénomènes oraux et réification graphique. D'autres (Abrial, 1995), en vue d'intégrer l'activité de lecture à l'apprentissage de la prononciation du français, ont repris l'idée de système phonologique minimal, en particulier pour les voyelles, et, en se limitant à 11 phonèmes vocaliques, ont élaboré un système graphémique reposant sur certains graphèmes de base du français et sur le format (minuscule/majuscule) des lettres, visant notamment à mettre en avant les phénomènes de liaison et d'enchaînement, et plus globalement la syllabation de l'oral *via* le canal écrit.

### II.1.2.5.3.1. Elaboration d'un code graphique pour l'apprentissage de l'oral

Trois démarches sont envisageables dans l'élaboration d'un tel code :

- 1) Prendre en compte le public auquel il s'adresse et s'y adapter. Cela pourrait notamment consister à utiliser certains codes phonographémiques de la L1 pour la transcription de la L2, ce qui, comme nous l'avons démontré, pourrait s'avérer tout à fait négatif, à court- comme à long-terme, pour l'apprentissage du système phonético-phonologique de la L2. En poussant cette logique à l'extrême, l'oral en L2 serait transcrit selon le système d'écriture de la L1<sup>252</sup>, ce qui nous renverrait alors au problème de l'adaptation phonographémique et phonologique des emprunts, et à la pratique, non-bénéfique selon nous, des étudiants japonais de français, qui transcrivent en katakana (et donc selon les contraintes phonologiques du système activé) les items français, oraux ou écrits, qui leur sont présentés.
- 2) Elaborer un code orthographique « simplifié » (selon quels critères ?), destiné à tous les publics, qui permettrait une « meilleure » (selon quels critères ?) appréhension de l'oral. Une telle démarche, si elle peut parfois s'avérer

---

<sup>252</sup> Ce que l'on observe dans de nombreux manuels de français conçus au Japon, qui proposent, sous chaque item orthographique français, une double « transcription », inévitablement approximative, en katakana. Une telle pratique, si elle peut être « acceptée » pour des ouvrages commerciaux destinés à des apprentissages auto-dirigés à très court terme, par exemple pour des séjours touristiques ou professionnels exceptionnels en France, ne peut en revanche être didactiquement justifiée pour des apprenants universitaires qui souhaitent se spécialiser en français, comme le démontre notre étude. Il est, à cet égard, intéressant de noter l'influence des traditions didactiques en L1 sur l'enseignement de la L2 : lorsqu'ils apprennent les « kanji », les Japonais utilisent en effet une « transcription » en hiragana, généralement placée au-dessus du kanji, pour connaître sa prononciation. Il s'agit des « furigana ». Ces furigana sont également utilisés au quotidien, notamment dans la presse, pour indiquer la prononciation de kanji dont la forme phonético-phonologique n'est pas évidente pour le lecteur, soit en raison de la rareté du kanji, soit en raison de son hétérophonie homographique. Il semblerait donc que les lecteurs japonais aient l'habitude d'un double « codage graphique » pour accéder à la « prononciation » des formes orthographiques. Il n'est dès lors pas surprenant que les étudiants soient tentés de transcrire l'orthographe française par des séquences en katakana, un fait qui s'explique notamment : 1) par l'utilisation des furigana en L1, et 2) par l'utilisation des mots d'emprunt en L1 (« gairaigo »), qui suit certains principes d'adaptation phonologique et de transcription en katakana. Ces deux faits relèvent à la fois du domaine sociolinguistique en L1, et du domaine didactique en L1, deux faits apparemment bien éloignés de la comparaison « structurale » des systèmes phonologiques en L1 et en L2, lorsque l'on s'intéresse aux facteurs pouvant conduire à des erreurs de « prononciation » chez les apprenants japonais.

*ponctuellement* utile pour tel ou tel problème spécifique, rentre cependant en contradiction avec, d'une part la cohérence interne, notamment phonéto-phonologique, de l'orthographe française (Catach, 1995)<sup>253</sup>, et d'autre part la nécessité pour les apprenants d'apprendre l'orthographe française : apprendre un deuxième code pourrait, sinon alourdir, du moins surtout perturber l'apprentissage de celle-ci.

3) Une combinaison des deux démarches précédentes.

#### II.1.2.5.3.2. Le système Alfonic

La deuxième démarche mentionnée appelle une remarque : celle-ci évoque en effet le système « Alfonic », élaboré dans les années soixante-dix par A. et J. Martinet, et présenté ainsi par Catach (1995, p. 106) :

*« L'Alphonic, alphabet "sous-phonologique" notant seulement les "archiphonèmes" des voyelles à double timbre avec possibilité pour l'enfant de noter certaines nuances s'il le désire. [...] Le véritable problème est pour l'Alfonic [...] le passage à l'orthographe traditionnelle. [...] En fait, on peut se demander s'il est bien nécessaire, pour le français, de faire un tel détour. L'API est absolument indispensable aujourd'hui, ne serait-ce que pour l'apprentissage des langues étrangères. Il est également indispensable de connaître l'alphabet traditionnel, tout aussi international. Enfin, l'orthographe du français est encore, et sans doute pour longtemps, le seul outil social reconnu ; faut-il y ajouter encore d'autres signes, surtout provisoires ? On peut se le demander, surtout quand on sait, comme nous l'avons dit, que les « phonogrammes » du français couvrent 80 à 85% des graphèmes, et que parmi*

---

<sup>253</sup> Même si Martinet (1986, p. 14) écrit à propos de l'orthographe française: « L'existence d'une graphie de ce type est un drame national pour la France et un drame à l'échelle du monde pour la francophonie. [...] Arguer, comme on le fait parfois, que l'orthographe comporte une certaine logique qui pourrait être formatrice pour l'esprit de l'enfant, c'est ne pas voir qu'il s'agit en l'occurrence d'une logique qui rappelle celle des déments, en ce qu'elle fait preuve d'une cohérence interne, mais n'est pas en prise sur le monde réel ».

*eux 33 signes fondamentaux constituent à leur tour 85% de ces « phonogrammes ». Ce sont ces signes qui doivent servir de bases au passage à l'écrit ».*

*Alfonic* est un système graphique biunivoque, dans lequel chaque lettre correspond à un seul phonème et vice-versa, ce qui conduit à un alphabet phonique comprenant 36 lettres (correspondant aux 36 phonèmes du français). Il s'agit d'un système « sous-phonologique », c'est-à-dire un « alphabet dans lequel certaines distinctions phonologiques ne sont pas marquées » (Bost, 1973), qui renvoie notamment aux conceptions de Martinet sur les rapports entre le parler et l'écrit (Martinet, 1974a, 1974b). Ce système a été élaboré en vue de répondre à certaines difficultés d'alphabétisation scolaires chez les enfants et son utilisation, pour le passage à l'écrit, se présentait ainsi :

Écriture (alfonic)	→ lecture (alfonic)	→ lecture (graphie traditionnelle)	→ écriture (graphie traditionnelle).
-----------------------	------------------------	---------------------------------------	---

Il s'agissait alors de dissocier l'apprentissage de la lecture-écriture de celui de l'orthographe.

Tandis que certains résultats se sont montrés décevants (Catach, 1995, p. 106), d'autres sembleraient avoir été plus positifs, notamment auprès d'enfants dyslexiques, conduisant certains formateurs à poursuivre son utilisation<sup>254</sup>. Il ne s'agira pas ici pour nous de rentrer dans le débat concernant les mérites et les torts d'un tel système (qui est présenté et analysé dans les revues *Alfonic* et *Liaison Alfonic*), puisque l'objectif et le public initialement visés par celui-ci diffèrent des nôtres au moins sur quatre points. En effet, nous nous préoccupons :

---

<sup>254</sup> Voir par exemple le projet « Je parle donc j'écris », initié par F.-X. Neve et porté par INFOREF en Belgique : <http://www.ulg.ac.be/cifen/inforef/projets/jeparledoncjecris/index.html>. Le site (consulté en mars 2005) offre une présentation très complète de l'*Alfonic*, sa conception, son utilisation, etc., ainsi que des textes de référence, en particulier de A. et J. Martinet, dont nous nous sommes inspirés ici.

- a) De non-francophonones, et non de francophonones,
- b) D'adultes, et non d'enfants,
- c) De sujets possédant déjà une compétence phonographémique (lecture et écriture), et non de sujets illétrés,
- d) De sujets apprenant le système oral du français, et non de sujets connaissant déjà ce système.

Il nous faut alors distinguer le cas des apprenants possédant déjà une conscience phonémique (en général à système d'écriture alphabétique) des autres (comme, dans une certaine mesure, nos apprenants japonais). Ce n'est que sur la question du développement de la conscience phonémique que les deux problématiques convergent. Comme le note Catach (1995, p. 106) :

*« Ce que l'enfant connaît à son arrivée à l'école, c'est l'oral, dont il faut lui rendre plus clairs les éléments constitutifs, les **phonèmes**. Non pas en les faisant pratiquer à l'écrit, mais en affichant dans la classe, à titre de référence, les tableaux des signes phonétiques du français et de leurs principaux correspondants écrits ».*

Or, si certains apprenants de FLE doivent également effectuer ce travail de phonémisation de la parole, dans lequel l'écrit peut constituer un outil de choix, ils se distinguent du public précoce évoqué par Catach, en ce qu'ils ne « connaissent » précisément pas « l'oral » en français.

#### II.1.2.5.3.3. Conscience phonémique et correspondances phonographémiques

Puisque notre problématique concerne le rôle de l'écrit dans l'apprentissage de l'oral, et non de l'écrit, il faut alors distinguer deux problèmes :

- 1) Celui du développement de la conscience phonémique, lié à la segmentation de l'oral : on peut dans ce cas s'interroger sur le choix du code graphique le plus approprié à la facilitation d'un tel développement. Outre la conscience phonémique,



on pourrait également considérer le rôle de ce code graphique dans le développement d'une conscience sub-phonémique, relative à la perception et la conscientisation des caractéristiques phonético-phonologiques des phonèmes en termes de traits, comme par exemple le voisement. Ainsi, dans le système kana japonais, il est remarquable que, pour certains couples comme ka/ga ou ta/da par exemple, le voisement soit transcrit graphiquement par l'ajout d'un signe, que l'on pourrait qualifier de diacritique (en l'occurrence deux traits obliques parallèles en haut à droite du caractère initial correspondant à la consonne non-voisée), à la consonne voisée, alors qu'il est transcrit par un graphème tout à fait différent dans l'alphabet romain. En revanche, il est également notable que le caractère kana correspondant à la nasale moraique /N/ puisse être considérée comme « sous-spécifié », étant parfois transcrit <n> ou <m> en caractère romaji afin de rendre compte de sa réalisation phonétique. Ce deuxième aspect concernerait alors tous les apprenants de FLE, quel que soit leur système d'écriture. La question serait alors de savoir si un système comme l'*Alfonic* serait plus à même de faciliter ces développements que l'orthographe française, voire un autre système comme l'*Alphabet Phonétique International*.

- 2) Celui de l'apprentissage des correspondances phonographémiques du français : comme nous l'avons montré, étant donné la réalité de l'utilisation du matériel écrit en classe de FLE, il semblerait nécessaire de mieux agencer l'apprentissage de l'oral et de ces correspondances, en vue d'éviter tout transfert phonographémique de la L1 vers la L2 potentiellement néfaste à l'apprentissage de l'oral en L2. Ce deuxième problème, loin de bannir l'orthographe, encourage au contraire à davantage insister sur l'apprentissage des correspondances phonographémiques telles qu'elles existent dans l'orthographe française, essentiellement dans le sens graphème-phonème, de manière à ce que les tâches de lecture, si nombreuses dans l'apprentissage contemporain d'une langue étrangère n' « interfèrent » pas avec l'apprentissage de l'oral. Dans ce cas, le système *Alfonic* ne résout en rien le problème du « chevauchement interscriptural » L1/L2, et pourrait même, dans certains cas, l'aggraver.

Ainsi deux constats se dessinent :

- 1) La nécessité de choisir un code visuo-graphique le mieux adapté à la résolution du premier problème (conscience phonémique et infra-phonémique).
- 2) La nécessité de davantage coordonner l'apprentissage de l'oral avec celui des correspondances graphèmes-phonèmes (en lecture au moins) afin de traiter le second problème.

En ce qui concerne le premier, il nous faut alors brièvement considérer le rôle potentiel de l'*Alphabet Phonétique International*.

#### II.1.2.5.3.4. L'Alphabet Phonétique International (API)

Celui-ci est clairement présenté par Durand (2005a), qui nous rapporte en particulier les six principes guidant l'emploi des symboles de l'alphabet en 1888, principes qui sont ensuite étiquetés et commentés par l'auteur :

<b>Principes de l'API de 1888</b>	<b>Etiquettes de Durand (2005a)</b>
(1) Il doit y avoir un signe séparé pour chaque son distinctif ; c'est-à-dire chaque son qui peut changer le sens d'un mot si on l'emploie à la place d'un autre son de la même langue.	<i>Le principe phonémique</i>
(2) Quand on trouve un son identique dans plusieurs langues, on utilisera le même signe pour toutes. Cela s'applique aussi aux nuances de sons proches les unes des autres.	<i>Le principe de similitude</i>
(3) L'alphabet sera autant que possible composé de lettres ordinaires de l'alphabet latin, en restreignant autant que possible l'utilisation de nouvelles lettres.	<i>Le principe latin</i>
(4) L'usage international décidera de la valeur à assigner aux caractères latins.	<i>Le principe international</i>

(5) Les nouvelles lettres devront suggérer les sons qu'elles représentent par ressemblance aux lettres déjà existantes.	<i>Le principe d'iconicité</i>
(6) On évitera les signes diacritiques, car ils sont incommodes pour la lecture et l'écriture.	<i>Le principe unitarien</i>

On remarque alors que le « principe phonémique » semble particulièrement bien se prêter au développement de la conscience phonémique et que l'usage des diacritiques, même s'il est limité par le « principe unitarien », pourrait permettre de rendre visuellement saillantes certaines caractéristiques phonétiques des items, lorsque l'on décide d'en effectuer une transcription phonétique (et non phonémique), que l'on peut encore appeler « allophonique » (Durand, 2001).

L'importance de l'utilisation des transcriptions phonétiques pour des apprenants de langue étrangère peut se justifier de plusieurs manières, à commencer par son rôle dans l'usage des dictionnaires, comme l'indique Wells (1996, p. 239) :

*« A passive acquaintance with phonetic transcription enables him or her to extract precise and explicit information on pronunciation from a dictionary, bilingual or monolingual. Without this information, a learner risks being misled either by an inadequately trained ear or by the dazzling effect of the ordinary spelling ».*

Wells recommande ainsi, pour des apprenants avancés au moins, la pratique de la transcription phonétique, ou du moins phonémique, de textes orthographiques. L'exercice de transcription phonétique serait donc propice à une certaine « prise de conscience » des caractéristiques phonético-phonologiques des items présentés, valable tant en langue étrangère qu'en langue maternelle : il est en effet intéressant de noter que la capacité de transcrire phonémiquement des items de sa propre langue n'est pas évidente. Comme le note Crookston (2001, p. 7) :

*« A naive reading of the evidence would suggest that a student who is literate in an alphabetic system and is beginning phonemic transcription in the language of literacy already possesses all the relevant knowledge bar the symbols [...]. Indeed, my own teaching experience suggests that there is a small minority of students for which this may really be the case ».*

L'auteur illustre notamment son propos à l'aide des fricatives anglaises : d'un côté le contraste entre /f/ et /v/, qui est représenté dans l'orthographe anglaise, et de l'autre celui entre /θ/ et /ð/ qui ne l'est pas (graphème <th>), et qui, dans les tests de l'auteur, conduit à davantage d'erreurs de transcription que le premier. La conclusion provisoire à laquelle il parvient est la suivante (p. 10) : « we seem to be building on a conscious representation of a full phoneme inventory, one which may be distinctive-feature-based, and one in which certain phonemic units including the non-labial fricatives are attenuated ».

En ce qui concerne les apprenants de langue étrangère, certains enseignants hésitent à utiliser l'API pour au moins deux raisons : 1) le coût temporel et cognitif de cet apprentissage, et 2) la difficulté à enseigner des signes graphiques correspondant à des « sons » qui ne sont, eux-mêmes, pas encore maîtrisés. Cette hésitation peut cependant être résolue lorsque l'on considère : 1) que l'apprentissage de ces « sons » peut précisément bénéficier de l'apprentissage et de l'utilisation de l'API et 2) que ce coût peut être rentable à long-terme vis-à-vis des formes lexicales stockées en mémoire par les apprenants. En effet, lorsque nous observons nos étudiants japonais « transcrire » spontanément en caractères kana les items français, oraux ou orthographiques, qui leur sont présentés<sup>255</sup>, et connaissant les effets dommageables d'une telle transcription sur

---

<sup>255</sup> Nous avons particulièrement observé ce comportement en tâche de lecture à voix haute, nous renvoyant ainsi à notre problématique centrale. Dans une telle tâche, l'enseignant « corrige » les apprenants en leur répétant le modèle « correct », et ces derniers, après avoir relu à voix haute l'item concerné, notent alors en katakana la transcription qu'ils jugent la plus correcte de ce qu'ils ont « répété ». C'est cette forme en katakana, inévitablement différente de la forme française, tant orale qu'orthographique, qui sera donc mémorisée par les étudiants, avec les conséquences que l'on peut imaginer sur leur apprentissage du français. Une transcription en API, produite par l'apprenant ou

l'apprentissage des formes lexicales, et ainsi sur la communication, orale ou écrite, en langue étrangère, il nous semble que l'apprentissage de l'API serait un « investissement » rentable, que l'on trouve parfois effectué par certains étudiants, avec un relatif succès.

Cependant, les principes « de similitude », « latin » et « d'iconicité » nous renvoient à notre second problème : celui des « chevauchements » phonographémiques : en raison de ces principes, la lecture d'un énoncé transcrit en API risque de susciter les mêmes difficultés que celles produites par un énoncé orthographique. On peut donc considérer que l'API présente un certain intérêt pour le développement de la conscience phonémique et infra-phonémique (pouvant également, dans une certaine mesure, permettre de mieux appréhender certains phénomènes de coarticulation), mais qu'il risque, en lecture au moins, de susciter les mêmes problèmes que ceux évoqués précédemment. L'entraînement à la transcription (phonémique et/ou phonétique) constitue donc une activité d'apprentissage permettant d'orienter l'attention des étudiants sur la dimension phonético-phonologique des items qui leur sont présentés, mais ne semble *a priori* ni constituer un outil d'apprentissage suffisant, ni traiter la question de l'apprentissage phonographémique.

### **II.1.3. Tâches et utilisation du support visuo-graphique**

Comme nous l'avons vu à propos de la distinction entre « conscience phonologique précoce » et « conscience phonémique » (Bentin, 1992), l'apprentissage d'une écriture alphabétique, permet le développement du deuxième type de conscience. Si l'apprentissage explicite d'un code abstrait (orthographique) permet ainsi d'acquérir, dans une certaine mesure, la capacité d'analyser la parole en constituants plus fins (phonèmes *vs.* rimes par exemple), on pourrait peut-être imaginer que la représentation visuelle, à l'aide de diacritiques orthographiques par exemple (Wells, 2001), de constituants infra-phonémiques (traits ou autres) en faciliterait la perception et ainsi

---

l'enseignant, serait sans doute, nous semble-t-il, plus profitable à l'apprentissage que cette transcription très approximative en caractère kana. Mais cela resterait à vérifier (quasi-)expérimentalement.

l'apprentissage. Cependant, comme nous l'avons vu dans nos propres tests, la simple présentation d'une représentation orthographique ne permet pas de faire prendre conscience aux apprenants des unités phonologiques (dans notre cas, l'absence de voyelle épenthétique entre les deux consonnes des groupes consonantiques). Ainsi, si l'écriture peut aider à la prise de conscience des unités phonologiques, cette aide à l'apprentissage dépend cependant des systèmes linguistiques en présence, du degré d'expertise des apprenants, ainsi que des conditions d'utilisation de ce support graphique (orthographique ou API), utilisation qui s'articule plus particulièrement autour de deux axes :

#### II.1.3.1. Lecture vs. Transcription

Comme nous l'avons indiqué, le rôle de l'écrit vis-à-vis de l'oral ne peut être évalué qu'en fonction de la tâche dans laquelle il est sollicité. Il est apparu que, en tâche de lecture, l'écrit pouvait entraver, dans une certaine mesure et sous certaines conditions, l'apprentissage des formes phonético-phonologiques de la L2. Cependant, en tâche de production, libre ou sous dictée, orthographique ou phonémique, l'écrit peut constituer une activité d'apprentissage de l'oral non-négligeable, selon deux points de vue. D'abord, de manière générale, parce que « apprendre l'oral » ne se limite pas à l'apprentissage de l'inventaire segmental et de la structure syllabique de la langue en question : cela concerne également l'apprentissage de la phonotactique<sup>256</sup>, de la

---

<sup>256</sup> Comme cela apparaît dans l'étude de Gathercole, Frankish, Pickering et Peaker (1999) : le rôle des connaissances statistiques (phonotactiques) sur l'apprentissage peut en effet également être abordé d'un point de vue mnésique, si l'on considère par exemple l'étude conduite par ces auteurs à propos de l'influence de ces connaissances sur la mémoire à court-terme. Travaillant sur une population enfantine anglaise avec des tâches de rappel sériel, les auteurs se sont en effet penchés, d'une part sur l'effet de lexicalité (le rappel de mots est supérieur à celui de non-mots), qui témoigne de l'influence de la mémoire à long-terme sur la mémoire à court-terme, mais surtout sur l'impact des connaissances phonotactiques (distribution phonémique ou syllabique et probabilité d'occurrence) sur le rappel, en utilisant des mots familiers et des non-mots non-familiers. Les résultats indiquent que, à valeurs phonotactiques égales, l'effet de lexicalité persiste, avec un meilleur rappel des mots sur les non-mots. En revanche, les non-mots à forte probabilité phonotactique sont mieux rappelés que ceux à faible probabilité. Cet effet de fréquence pour les non-mots semblerait lié aux constituants syllabiques des stimuli. La conclusion de cette étude indique que deux types de connaissances à long-terme peuvent influencer la tâche de rappel : l'une lexicale, l'autre phonotactique. Les auteurs en modélisent le processus de la manière suivante, en proposant que deux processus de reconstruction seraient à l'œuvre en tâche de rappel sériel (p. 93) : d'abord, la trace mnésique serait classifiée comme intacte, partielle ou

fréquence (lexicale et sublexicale), des voisinages (phonologique et orthographique) ainsi, sinon surtout, que l'apprentissage de la capacité à traiter des formes phonétiques pour en retirer une forme phonologique et/ou orthographique, ainsi que toutes leurs propriétés lexico-sémantiques et morpho-syntaxiques. Apprendre l'oral implique donc d'apprendre à traiter la variabilité (essentiellement phonétique) pour en retirer des invariants (essentiellement phonologiques) qui s'incarnent particulièrement bien dans leur forme orthographique.

Ensuite, de manière plus spécifique, si l'on considère, aux niveaux segmental et syllabique, que l'écrit constitue, en quelque sorte, une forme de « prototype phonologique », celui-ci peut, certes, être négatif pour l'entraînement à la perception phonétique, mais il pourrait également être utilisé comme outil de « retour correctif » (« feedback ») pour délimiter un nouvel espace phonétique, et, à sa suite, de nouvelles catégories phonologiques. La « correction phonétique », telle qu'elle actuellement pratiquée, devrait alors être complétée d'une correction « phonographémique », qui viserait finalement non pas tant l'apprentissage de l'orthographe française (qui renverrait à l'exercice de dictée scolaire traditionnelle en langue maternelle), que celui des catégories phonologiques du français, *via* une délimitation phonético-graphémique. Une telle démarche pourrait donc impliquer à la fois des exercices de transcription phonémique, phonétique et orthographique. L'écrit constituant un certain « retour correctif » vis-à-vis de l'oral, il devrait donc être systématiquement contrôlé et corrigé, au risque de laisser les apprenants établir des correspondances écrit-oral et oral-écrit erronées, néfastes pour l'apprentissage de l'oral. De tels exercices devraient être déclinés en fonction du type et de la quantité d'information contenue par les items présentés (non-mots, mots ; information sémantique, morphologique, orthographique) et devraient également être élaborés en fonction du public auquel ils s'adressent. Ainsi dans le cas d'apprenants japonais, lorsqu'on prend connaissance du fait (Wioland,

---

manquante, puis un processus de reconstruction, d'ordre soit lexical, soit phonotactique, serait activé, pouvant faire usage de trace mnésique partielle. Les informations à long-terme seraient ainsi mises à profit, non lors de l'étape de stockage de l'information en mémoire à court-terme, mais lors de son rappel (p. 94).

1991, pp. 30-32) que le /R/ est la consonne la plus fréquente en français (7, 25% des phonèmes) et que le /l/ arrive en troisième position (5, 63%) après le /s/ (6%) on réalise qu'il est essentiel d'aborder les problèmes posés par le couple R/L immédiatement dans la progression d'apprentissage et ordonner le programme d'enseignement phonético-phonologique en fonction des critères de fréquence d'occurrence et d'intelligibilité.

### II.1.3.2. Utilisation passive vs. Utilisation active

Jusqu'à présent, nous avons mentionné l'utilisation du support viso-graphique en ne faisant que peu de cas de son *mode* d'utilisation<sup>257</sup>. Or, il semble évident que celui-ci constitue un facteur non-négligeable dans l'effet que ce support peut avoir sur l'apprentissage de l'oral. Ainsi, dans le cas des correspondances phonographémiques, l'utilisation « passive » du support orthographique, dans laquelle l'apprentissage des correspondances serait censé s'effectuer de manière « implicite », risque, *a priori*<sup>258</sup>, de ne pas conduire aux mêmes résultats, au moins en termes de durée d'apprentissage, qu'une utilisation « active », dans laquelle l'enseignant orienterait « explicitement » l'attention des apprenants sur lesdites correspondances<sup>259</sup>. Cette « utilisation passive »

---

<sup>257</sup> Il est d'ailleurs intéressant de noter que certains chercheurs en sciences cognitives, sur les questions qui nous concernent ici, semblent aujourd'hui ressentir la nécessité de prendre davantage en compte certains facteurs « externes » dans leur analyse de la cognition humaine, ce qui renvoie, d'une certaine manière, à la position « réaliste » que nous avons présentée plus haut. Ainsi Hutzler, Ziegler, Perry, Wimmer et Zorzi, dans le résumé qu'ils donnent de leur article intitulé « Do current connectionist learning models account for reading development in different languages » (*Cognition*, Vol. 91, n° 3, avril 2004, pp. 273-296), écrivent-ils :

*« Further investigations showed that the two-layer network could be brought to simulate the cross-language learning rate effect when cross-language differences in teaching methods (phonics versus whole-word approach) were taken into account. The present work thus shows that in order to adequately capture the pattern of reading acquisition displayed by children, current connectionist models must not only be sensitive to the statistical structure of spelling-to-sound relations but also the way reading is taught in different countries ».*

<sup>258</sup> Car un apprentissage tout à fait implicite peut également être réussi, selon, notamment, la motivation et la situation d'apprentissage.

<sup>259</sup> Ce qui renvoie, dans le domaine de l'apprentissage de la lecture en L1, à ce que Gombert (2003) nomme le "boosting effect" de l'enseignement/apprentissage explicite sur l'apprentissage implicite :



correspond à ce que l'on observe dans de nombreux types d'enseignement : une présentation initiale et succincte du système alphabétique et phonétique du français, à partir de laquelle les apprenants sont supposés « savoir lire » en français. L'enseignant utilise alors par la suite un ensemble de supports orthographiques (textes, exercices, etc.) sans davantage aborder la question, et les activités de correction phonétique, lorsqu'il y en a, ne sont que très ponctuellement reliées au domaine phonographémique. L'exercice de dictée, notamment en raison de son histoire didactique, est quant à lui jugé « rébarbatif » et très souvent négligé, une négligence que certains seraient tentés de justifier par son apparente absence de « perspective communicative ». Une telle position néglige ainsi tous les arguments que nous avons développés vis-à-vis du rôle de l'écrit dans l'apprentissage de l'oral. En ce qui concerne une hypothétique « utilisation active », celle-ci peut s'exprimer de diverses manières, à travers la méthodologie adoptée par l'enseignant, les tâches didactiques proposées, ainsi que le format ergonomique de ces dernières.

Cependant, ce dernier point soulève une question centrale en didactique des langues, celle de l'articulation entre le domaine de l'*implicite* et celui de l'*explicite*, tant du point de vue de l'apprentissage que de l'enseignement. On peut en effet se demander, comme nous l'avons évoqué, si *l'enseignement* des correspondances phonographémiques, et plus globalement du système oral (phonétique, phonologique et phonographémique) du français doit être effectué d'une manière que l'on pourrait qualifier d'implicite, d'explicite ou d'une combinaison des deux. Un tel thème pourrait faire l'objet de plusieurs thèses, étant donné que l'articulation entre explicite et implicite en didactique des langues continue d'être un sujet de recherche, de débats, et de très nombreuses publications<sup>260</sup> (Ellis, N., 1994), recoupant ainsi en bien des points les thèmes d'actualités du domaine d'étude de l'acquisition des langues secondes,

---

*« Overall, direct teaching is essential both in order to permit the elaboration of a control system which readers can use for conscious monitoring, and in order to boost implicit learning by multiplying the number of print-related activities they are engaged in ».*

<sup>260</sup> La revue *Studies in Second Language Acquisition* doit ainsi consacrer un numéro spécial au thème « Implicit and Explicit Learning » en juin 2005 (Vol. 27, n°2).

mentionnés par Valdman dans l'introduction d'un numéro spécial de la revue *Studies in Second Language Acquisition* consacré à la « fréquence » (juin 2002, p. 142) :

*« Current issues in SLA research : the relationship between associative learning and abstract rule systems ; the validity of postulated progression from formula to constructions via low-scope patterns ; the relationship between implicit and explicit learning ; the effect of form-focused instruction on implicit learning ; phonological reduction ; the relationship between frequency and saliency ; the importance of noticing and the learner attention ; the relationship between implicit learning and production ; and the isomorphy between comprehension and production grammars ».*

Pour notre part, nous nous contenterons donc de relever certains traits essentiels à notre problématique, en élargissant dans un premier temps notre thème initial à celui de l'apprentissage d'une L2. Un tel élargissement (de l'apprentissage du système phonético-phonologique à celui de la langue dans son ensemble) se justifie ici dans notre étude de trois manières : d'abord, 1) il nous semble essentiel, si l'on souhaite traiter du couple explicite/implicite dans l'apprentissage de la « prononciation du français », de mener en amont une réflexion suffisamment approfondie sur les notions d'implicite et d'explicite dans l'enseignement / apprentissage des langues. Or, 2) cette réflexion s'appuie sur une littérature qui concerne généralement davantage la langue dans son ensemble, et plus particulièrement le thème de la « grammaire », que celui de la « prononciation ». Enfin, 3) puisque notre étude vise à enrichir une réflexion *méthodologique* sur l'enseignement / apprentissage de l'oral en L2, et non simplement à proposer certaines techniques didactiques « universelles » ou, au contraire, spécifiques à certains problèmes particuliers, on ne peut, par souci de cohérence didactique méthodologique, traiter de « prononciation » sans se référer aux autres aspects de la compétence de communication à apprendre. Ces trois arguments justifient ainsi ce que nous pourrions considérer comme une « parenthèse » méthodologique dans notre propos, parenthèse « de fond » indispensable à notre approche. Ce sont les raisons pour lesquelles, le thème de l'explicite et de l'implicite dans notre étude nous conduit à

examiner une littérature qui s'est particulièrement concentrée sur l'enseignement/apprentissage de la « grammaire » en langue étrangère, les références concernant celui de la « prononciation » étant, dans une certaine mesure, moins nombreuses, et finalement relativement dépendantes de la problématique générale.

Lorsque l'on envisage un enseignement explicite de la « prononciation » du français, on doit en effet, au moins du côté de l'enseignant, s'interroger sur les représentations métalinguistiques que l'on doit avoir de cette « prononciation ». Les données des sciences du langage vont ainsi être sollicitées, complétées par celle de la psycholinguistique et de la psychologie ergonomique, afin de transposer ces savoirs explicites dans des activités didactiques qui soient le mieux adaptées au public et à la situation, de manière plus ou moins formelle, plus ou moins explicite et plus ou moins déclaratives, en fonction des choix méthodologiques de l'enseignant. Tant le choix des modèles que celui de leur « formatage » ergonomique et didactique, doit ainsi prendre en compte, à la fois les données linguistiques de références (issues des sciences du langage), que celles relatives à l'activité métalinguistique des apprenants (issues de la didactique et de la psycholinguistique).

Il nous semble donc justifié d'aborder dans un premier temps la problématique générale de l'explicite et de l'implicite dans l'enseignement/apprentissage des langues, en particulier dans le domaine de la « grammaire », auquel est lié celui de la phonologie, avant de revenir au thème de l'oral, de la prononciation et des correspondances phonographémiques.

## **II.2. Enseignement explicite ou implicite<sup>261</sup> ?**

La question est très vaste, et touche à la fois : 1) au domaine didactique, *via* les options méthodologie adoptées, parfois déclinables en fonction des sous-champs et objectifs d'enseignement/apprentissage, 2) à la psycholinguistique, *via* les modèles cognitifs (mnésique, d'apprentissage, etc.) sélectionnés, et 3) à la linguistique, *via* les théories

---

<sup>261</sup> Cette sous-partie s'inspire notamment de (Detey, 2003a, 2004).

associées (émergentiste, générativiste, etc.). D'un point de vue psycholinguistique, le débat porte naturellement plutôt sur la question de l'*apprentissage* que sur celui de l'enseignement. D'un point de vue didactique et linguistique, il est particulièrement attaché à la notion de « grammaire ».

Il est toujours délicat d'employer les termes « explicite » et « implicite » dans le domaine qui nous occupe, en ce qu'ils sont parfois abusivement associés, voire assimilés à d'autres, comme « déclaratif » et « procédural » lorsque l'on traite de « savoirs », ou encore « automatique » et « contrôlé » lorsque l'on traite de « processus ». En outre, la signification accordée à ces différents termes varie parfois en fonction des auteurs, des domaines et des modèles adoptés, et il convient donc de s'accorder sur ces significations, au risque de sombrer dans la confusion. La discussion menée par Laks (1996) à ce sujet nous semble éclairante, et nous emploierons donc ces termes en tenant compte des réserves que nous venons d'exprimer.

Etant donné notre perspective didactique, nous pouvons dans un premier temps noter les entrées « Explicite » et « Implicite » du *Dictionnaire de Didactique du FLES* (Cuq, 2003, p. 99 et p. 127) :

<b>Explicite</b>	<b>Implicite</b>
« La distinction entre explicite (on parle de connaissance, de grammaire, de savoir explicites) et implicite conduit à parler également de connaissances ou de grammaire explicitée, dans la mesure où les connaissances dites explicites ne sont le plus souvent qu'une formalisation de connaissances préalablement implicites. Ainsi, des connaissances grammaticales, qui peuvent être d'abord reçues sous une	« La distinction entre explicite et implicite, à propos de connaissance, de savoir ou de grammaire, provient en premier lieu de la psychologie cognitive et de la psycholinguistique (Gaonac'h, 1987). Elle correspond partiellement à la distinction, dans d'autres domaines que le langage, entre savoirs procéduraux et savoirs déclaratifs. On entend par grammaire implicite (ou connaissance implicite de la grammaire) la compétence grammaticale d'un locuteur (en langue

<p>forme explicite par des apprenants, sont issues de l'explicitation (par des linguistes ou par des enseignants) de connaissances implicites ».</p>	<p>première ou étrangère) indépendante ou distincte de connaissances métalinguistiques explicitant le système de cette grammaire. [...] On entend par grammaire explicite la représentation ou la formulation descriptives et explicatives de règles et de fonctionnements de la langue, au moyen de catégories métacognitives et métalangagières. [...] La nature et la part respective des connaissances implicites et explicites, et leurs interrelations même, varient, entre autres, selon les contextes, les étapes et les niveaux d'acquisition. En didactique des langues, la distinction implicite/explicite fonde [...] une diversité d'options méthodologiques, privilégiant, selon les cas, une approche « implicite » de la langue (SGAV, immersion, simulation, etc.), une approche « explicite » (méthodologie « grammaire-traduction », par exemple) ou des approches mixtes intégrant, sous des formes diverses, l'une et l'autre. Ainsi, la « grammaire explicitée » prend par exemple la forme d'activités de conceptualisation ».</p>
--	---

Trois grandes perspectives peuvent donc être développées à partir de ces définitions : l'une, à orientation psycholinguistique, concerne la question de l'apprentissage implicite et de l'effet de l'instruction explicite sur celui-ci, la seconde, davantage didactique, concerne l'enseignement de « la grammaire », la troisième, plutôt linguistique, concerne la nature des représentations utilisées pour l'enseignement précédent. Ces trois perspectives sont naturellement fortement interreliées, et il nous semble dès à présent essentiel d'insister sur la distinction entre *enseignement explicite/implicite* et *apprentissage explicite/implicite*.

En tant que didacticien, c'est bien la question de l'*enseignement* qui nous préoccupe au premier chef, en considérant ses conséquences sur l'apprentissage, que celui-ci soit « explicite » ou « implicite ». Dans le cas de l'apprentissage du système-phonéto-phonologique du français pour des étudiants japonais, l'interrogation essentielle est alors la suivante : l'enseignant doit-il aborder explicitement (avec ou sans réflexion métalinguistique) ou implicitement (avec ou sans activité épilinguistique) le thème de la « prononciation » (phonétique, phonologie et phonographème) en classe de langue ? Cette interrogation se décline ensuite selon l'approche et les supports adoptés.

Il convient alors, avant de s'interroger sur le rôle que peut avoir un « enseignement explicite », parfois nommé « instruction explicite », sur l'apprentissage implicite de la langue, de s'interroger sur la nature même de ce dernier (Winter et Reber, 1994).

### **II.2.1. Apprentissage implicite ou explicite ?**

Meulemans, qui a traité la question de l'apprentissage implicite dans une perspective à la fois cognitive, neuropsychologique et développementale, indique (1998, p. 15) :

*« La notion d'apprentissage « implicite » renvoie à notre capacité d'apprendre, sans que nous en soyons conscients, des informations de nature complexe, et au fait que la connaissance acquise est elle-même difficilement accessible à la conscience (ou, en tous cas, à l'évocation verbale) ».*

Pour Perruchet et Vinter (cités par Meulemans, 1998, p. 18) :

*« L'apprentissage implicite désigne un mode adaptatif par lequel le comportement des sujets se montre sensible aux caractéristiques structurales d'une situation à laquelle ils ont été préalablement exposés, sans que l'adaptation qui en résulte soit due à une exploitation intentionnelle de la connaissance explicite des sujets concernant ces caractéristiques ».*

En situation de classe, imaginons un énoncé, ou un ensemble d'énoncés, présenté « en situation », à l'aide de supports iconiques, et analysé explicitement d'un point de vue « formel » par l'enseignant, avant que ce dernier ne place les apprenants en situation de reconnaissance, de compréhension ou de manipulation de cet énoncé. Si le traitement de l'énoncé dans la deuxième phase est facilité sans « exploitation intentionnelle de la connaissance explicite des sujets », peut-on considérer qu'un *enseignement explicite* a débouché sur un *apprentissage implicite*, qui à son tour a favorisé le traitement d'informations présentées dans le cadre d'un *enseignement implicite* ?

En ce qui concerne la « grammaire », Meulemans (1998, pp. 26-39 et p. 186) présente trois grands points de vue relatifs à l'apprentissage implicite, particulièrement étudié à l'aide de « grammaires artificielles ». Il s'agit des approches abstractionniste, exemplariste et fragmentariste.

- Pour les abstractionnistes : l'apprentissage dans ce type de tâche serait sous-tendu par « un mécanisme inconscient ou implicite d'induction de règles »<sup>262</sup>. En 1993

---

<sup>262</sup> Notons dès à présent que la notion de « règle » doit être manipulée avec précaution, eu égard au concept de « régularité », dans le domaine linguistique au moins. Nombre d'enseignants de langue semblent en effet s'être forgés une représentation de la notion de « règle » bien loin de la réalité (socio- et psycho-) linguistique et scientifique du concept, qui doit être relié à ceux de norme, de variation, de régularité, de schéma et de modèle, entre autres. L'utilisation raisonnée du concept de « règle » nécessite ainsi une certaine réflexion préalable, en particulier sur le rapport entre « règle » et « régularité », réflexion utilement éclairée par l'analyse de Laks (1996) et qui, dans le domaine linguistique générativiste, se trouve également alimentée par l'opposition, toute relative, entre les concepts de « règles » et de « contraintes » (« violables »), tel qu'illustrés par exemple dans la théorie de l'Optimalité ou celle des Contraintes et des Stratégies de Réparation. Le développement des approches émergentistes et connexionnistes en sciences cognitives permet également d'enrichir cette réflexion, comme en témoigne la conclusion de McClelland et Rumelhart (1986, pp. 214-215) relative à l'apprentissage et à la mémoire distribuée :

*« The model [of distributed memory] thus stands as an alternative to the view that learning in cognitive systems involves the explicit formulation of rules and abstraction under the guidance of some explicit overseer. We do not wish to suggest that explicit rule formation has no place in a cognitive theory of learning and memory ; we only wish to suggest that such mechanisms need not be invoked whenever behavior is observed that appears to be described by some generalization or rule ».*

Cela renvoie en partie à « l'erreur d'isomorphie du premier ordre » (*First Order Isomorphism Fallacy*) de Kugler, Turvey et Shaw, cités par Laks (1996, p. 87) : « Cette erreur, récurrente en sciences

cependant, distinguant le point de vue de l'expérimentateur de celui du sujet, l'un des représentants de cette approche (Reber) écrit « there are no *rules* here, just a pattern of co-occurrence and covariation » (cité par Meulemans, 1998, p. 28), le problème étant notamment celui de l'appellation « règles » pour des connaissances non-accessibles à la conscience.

- Pour les exemplaristes : il n'y aurait pas de mécanisme d'abstraction, mais simplement un processus de mémorisation des exemplaires, donnant lieu, en tâche de classement de grammaticalité, à une comparaison entre les items stockés et ceux présentés. Il s'agit de « l'hypothèse de similarité », la similarité étant envisagée par rapport à un item spécifique et non à un quelconque prototype. On retrouve ici la problématique du « matching », c'est-à-dire, globalement, d'appariement entre percept et connaissances stockées en mémoire.
- Pour les fragmentaristes : il s'agirait de sous-groupes, ou de fragments, qui seraient mémorisés et non d'items entiers. Le modèle du « regroupement compétitif » de Servan-Schreiber et Anderson (cités par Meulemans, 1998, p. 188) par exemple, « repose sur le postulat général selon lequel, confrontés aux stimulations de notre environnement nous avons besoin d'y détecter des régularités ». Nous procédons ainsi à des découpages de fragments organisés hiérarchiquement, la facilité d'encodage des items étant fonction de la fréquence d'exposition, et l'encodage d'un nouvel item étant contraint par les représentations préexistantes.

Pour des auteurs comme Dulany, les sujets « apprennent un grand nombre de *micro-règles* explicites, formant des « grammaires corrélées » [...], c'est-à-dire, des relations structurales profondes qui ne sont pas strictement équivalentes à celles impliquées dans les items d'apprentissage, mais qui produisent des patterns de réponses similaires (et donc corrélés) » (Meulemans, 1998, p. 32). Ce dernier point de vue est intéressant,

---

cognitives, consiste à attribuer à un organisme la représentation rationnelle que nous nous construisons de son comportement et à conférer à cette représentation un statut causal du comportement observé ».



d'une part en raison des micro-systèmes évoqués, que nous rapprocherons des micro-systèmes qui structurent les interlangues des apprenants (Vogel, 1995), et d'autre part par la distinction établie entre les relations structurales « profondes » et celles, externes, qui régissent pour l'expérimentateur l'organisation du matériel présenté.

Meulemans et Van der Linden ont conduit une étude dont les résultats suggèrent « qu'en fonction des caractéristiques de la tâche, il est possible soit de favoriser un mécanisme d'apprentissage limité à l'acquisition d'informations liées aux bigrammes et aux trigrammes, soit de permettre à des processus d'abstraction plus complexes d'avoir lieu » (Meulemans, 1998, p. 38). Ces données peuvent être rapprochées de la suggestion de Reddington qui conçoit un « mécanisme « hybride » d'apprentissage » (fragments et règles), ou de celle de Perruchet et Pacteau, qui envisagent un élargissement progressif de la taille des fragments appris. On retrouve donc la figure d'un continuum, qui suggérerait d'enseigner *des* règles progressivement élargies et complexifiées, et non *les* règles les plus abstraites et les plus exhaustives.

Les auteurs ne s'accordent donc pas sur la nécessité d'établir une distinction entre deux systèmes, l'un d'apprentissage explicite et l'autre d'apprentissage implicite. En revanche, les uns (Broadbent et *al.*, Reber et *al.*, cités par Meulemans, 1998, p. 110) comme les autres (Whittlesea et Wright, cités par Meulemans, 1998, p. 114) reconnaissent comme déterminant sur la manière dont va être traité le matériel *le degré de saillance* de ce dernier et des « variables clés ». La question de *l'attention* est donc bien évidemment au centre de la problématique.

Des données disponibles sur l'apprentissage implicite, Meulemans conclue que, plutôt qu'une dichotomie stricte entre deux systèmes d'apprentissage distinct, l'un explicite et l'autre implicite, il serait plus réaliste de considérer (1998, p. 116) :

« *Un continuum où l'on trouve, à une extrémité, des processus de haut niveaux, du type testing d'hypothèse et planification consciente, gouvernés par les concepts (top-down) et, à l'autre extrémité, des processus probablement plus*

*élémentaires, reposant sur des mécanismes associatifs plus simples, dirigés par les données (bottom-up) ».*

Cela évoque, dans le domaine didactique, le modèle mixte de compréhension linguistique, à la fois onomasiologique (haut-bas) et sémasiologique (bas-haut) (Gremmo et Holec, 1990).

### **II.2.2. L'enseignement explicite est-il bénéfique au développement des savoirs implicites ?**

Cette absence de distinction claire entre apprentissages explicite et implicite ne remet pas en cause l'existence d'un *enseignement* explicite ou implicite, que l'on pourrait notamment définir en termes de facteurs attentionnels et métalinguistiques. La question de savoir lequel doit être pratiqué, sous quelles conditions et pourquoi demeure, même si « il n'existe pas de tâche "pure" et que, dès lors, il est vain de chercher à caractériser une situation donnée comme étant par exemple de nature soit implicite, soit explicite » (Meulemans, 1998, p. 201).

La vraie question, comme l'appelle Trévisse (1993, p. 48), demeure en effet « de savoir si une "bonne" description de la réalité du fonctionnement de la langue 2, ou des deux langues 1 et 2, "bien" simplifiée et adaptée aux représentations métalinguistiques préalables des apprenants et permettant des complexifications ultérieures, peut jouer un rôle dans les processus d'acquisition/apprentissage ». Trévisse semble répondre par l'affirmative, puisqu'elle conclut son article en écrivant (1993, p. 49) :

*« "A côté" du travail épilinguistique que représentent la perception active de langue cible, la fabrication et les tests d'hypothèses, l'élaboration et la définition du domaine des règles inconscientes et la mémorisation, il est certain que des explications métalinguistiques adéquates peuvent (parallèlement ou en composition ?) constituer une aide pour l'apprenant dans l'évolution de ses productions et de sa perception structurante de la langue cible ».*

Si ces propositions apparaissent aujourd'hui évidentes à certains, il reste encore, nous semble-t-il, à construire autour d'elles un véritable programme d'enseignement. Par ailleurs, cette préoccupation est courante parmi les enseignants de langue. Le titre d'un article de Narcy (1991) l'illustre bien : « La prise de conscience des problèmes linguistiques conduit-elle à une réduction des erreurs commises en production libre ? ». Dans cet article l'auteur rapporte les résultats d'une expérience qu'il a menée pour répondre à la question mentionnée : « Comme la plupart de nos collègues, nous estimions aussi qu'une réflexion formelle devait faciliter la structuration interne des données linguistiques, [...] il fallait nous en assurer » (p. 347). La réponse est positive, même si elle doit être nuancée selon les conditions.

Un chercheur qui a beaucoup travaillé sur les effets de l'instruction explicite et plus particulièrement sur le rôle que joue celle-ci dans le processus d'acquisition des savoirs implicites est Rod Ellis. Dans un article (Ellis, R., 2002), il examine 11 études portant sur l'influence de cette instruction sur la « production libre » des étudiants. Les conclusions sont clairement positives, et deux variables en particulier semblent déterminantes : le choix des structures étudiées et la durée de l'instruction. Il est remarquable que Ellis décide de commencer son article par la phrase suivante (p. 223) :

*« There is by now ample evidence to show that form-focused instruction (FFI) has a positive effect on second language (SL) acquisition. That is, by and large, learners seem to learn the grammatical structures they are taught ».*

Son appréciation positive est cependant très raisonnable et mesurée, puisqu'il passe en revue différents inconvénients et différentes faiblesses d'un tel enseignement (p. 224) :

*« FFI is invariably based on explicitly formulated rules derived from a reference grammar of some kind or another. Thus, what it seeks to teach, whether through implicit or explicit methods, may bear no resemblance whatsoever to how this feature is represented in the learner's store of implicit knowledge. Second, by artificially adjusting the frequency and distribution of a*

*property through instruction, FFI may artificially distort the input and actually mislead the learner ».*

Il prend également soin de noter que les résultats obtenus jusqu'à présent concernant l'utilité de « FFI » ne concernent que les apprenants de niveau intermédiaire à avancé, et non de niveau débutant (p. 231). Concernant les structures étudiées, il souligne (p. 232) :

*« FFI would seem to have a better chance of success if it is directed at simple morphological features (verb forms, articles, or formulaic items) than at more complex syntactic structures involving permutations of word order (e.g., word order involving Spanish clitic pronouns and passive sentences). Perhaps FFI [...] is less successful in the case of complex syntactic features because these require more complex processing operations that can only be mastered sequentially over a long period of time ».*

Ellis travaille depuis plusieurs années sur cette problématique, et il s'est ainsi efforcé de développer une théorie de l'instruction formelle. Il commence ainsi par distinguer « classroom language learning » de « naturalistic language learning » (1990, p. 1). D'un point de vue psycholinguistique, cette distinction est à rapprocher de celle entre apprentissage formel et informel, que l'auteur présente ainsi (p. 2) :

*« Formal learning involves some kind of studial activity on the part of the learner—for example an attempt to learn about the language by obtaining information about explicit rules of grammar. Informal learning takes place through observation and direct participation in communication—learning is a process of discovery which takes place spontaneously and automatically providing certain conditions have been met ».*

Même si Ellis se garde d'assimiler les deux couples, car par exemple une approche naturaliste peut aussi entraîner un apprentissage formel, par exemple lorsqu'une

requête métalinguistique est effectuée au milieu de la conversation, les uns et les autres présentent tout de même certaines affinités. Quatre ans plus tard, il écrit (Ellis, R., 1994, pp. 82-83) :

*« Code-oriented instruction may also provide the learner with a different kind of input—explicit information about the nature of linguistic features in the form of rule. Sharwood Smith (1981) suggests that deductive language teaching (which he calls « consciousness-raising ») varies according to the degree of explicitness and elaboration. Explicitness refers to the extent to which the teacher makes use of linguistic metalanguage. The teacher can simply hint with the help of an exemple or can provide a complete statement of the rule. Elaboration concerns the amount of time taken up in the presentation of a rule. Other variables relating to explicit grammar teaching include the source of the explanation (the teacher, the students, a textbook) and the manner of presentation (oral, written) ».*

Ici, l'élaboration d'un enseignement que l'on peut qualifier de « formel » implique donc de déterminer, d'une part le degré d'explicitation et d'autre part celui d'élaboration, qu'il est souhaitable d'obtenir, en regard des objectifs d'apprentissage.

On peut cependant noter deux points discutables dans cette citation : d'une part la notion de « règle » à laquelle Ellis semble systématiquement attacher toute information métalinguistique, et d'autre part la forme du « métalangage », qui est l'un des points cruciaux, d'après nous, de la conception de tout enseignement explicite. Néanmoins, ce qui nous intéresse ici, c'est bien de retenir l'existence de deux catégories de savoir, explicite et implicite, qu'il convient de ne pas confondre avec automatique et contrôlé (Ellis, R., 1994, tableau p. 86), ainsi que la position de l'auteur sur les rapports entre les deux catégories. Celui-ci adopte en effet une position d'*Interface Faible* (1994, figure 4, p. 97), c'est-à-dire qu'il considère que le savoir explicite, issu d'une instruction orientée vers le code (« code-oriented instruction »), joue un rôle essentiel dans l'acquisition de la L2.

Afin d'avoir une vue globale des différents savoirs linguistiques en L2 des apprenants, Ellis (1994, pp. 84-86) propose le tableau suivant, lequel permet d'éviter certaines confusions précédemment évoquées :

<b>Types of knowledge</b>	<b>Types of processing</b>	
	<b>Controlled</b>	<b>Automatic</b>
<b>Explicit</b>	(A) A new explicit rule is used consciously and with deliberate effort.	(B) An old explicit rule is used consciously but with relative speed.
<b>Implicit</b>	(C) A new implicit rule is used without awareness but is accessed slowly.	(D) A thoroughly learnt implicit rule is used without awareness and without effort.

En ce qui concerne les savoirs explicites (puisque c'est ce qui nous occupe principalement ici), l'auteur indique (1994, p. 84) :

*« Explicit knowledge refers to knowledge that is analysed (in the sense that it exists independently of the actual instances of its use), abstract (in the sense that it takes the form of some underlying generalisation of actual linguistic behaviour) and explanatory (in that the logical basis of the knowledge is understood independently of its application). Perhaps its defining characteristics is that it is available to the learner as a conscious representation [...]. However, as Bialystok points out, explicit knowledge is not necessarily the same as « articulated » knowledge, as it exists in the mind of the learner independently of its articulation ».*

Cette dernière remarque doit notamment faire réfléchir au passage d'un savoir apparemment non-articulé à un savoir articulé, car ce dernier implique l'émergence

d'un métalangage. Or le passage au métalangage est une étape délicate, et il peut prendre deux orientations :

- 1) Viser la monosémie : à travers l'élaboration d'un code métalinguistique unique et original. Cette orientation, qui a le mérite de favoriser la clarté dans la réflexion et la communication métalinguistique et l'homogénéité de celles-ci pour tous les apprenants, a le grave inconvénient de représenter un nouveau code linguistique en lui-même, dont l'apprentissage propre peut se révéler très difficile et nécessiter des capacités d'abstraction mentale et de gestion de l'abstraction que tous les apprenants n'ont pas nécessairement au même niveau.
  
- 2) Tolérer la polysémie : en utilisant une terminologie « non-vierge » : par exemple à travers des termes familiers aux apprenants, utilisés dans un domaine non-linguistique, voire non-académique, et dont le degré de concrétude ou d'imageabilité en favorise l'apprentissage et la manipulation. On notera alors le risque de confusion, ou de dérivation, sémantique entraîné par les termes employés, comme le note Collet dans le cas d'enseignements scientifiques, en particulier la physique : « Les connaissances véhiculées par un modèle physique ne se limitent donc pas aux définitions visibles, mais elles introduisent inévitablement, par le biais des mots employés, d'autres connaissances lexicales, extérieure à la théorie physique. Certaines de ces connaissances peuvent jouer un rôle "parasite" » (Collet, 2000, p. 21). L'autre possibilité est celle de la réutilisation d'un métalangage déjà connu des apprenants. Ici aussi, les risques sont similaires. L'avantage essentiel d'une telle approche est celui de l'économie cognitive.

Une autre étude qui peut nous aider à tenter de répondre à la question posée dans le titre de cette sous-partie est celle de Norris et Ortega (2000). Dans un article comptant 111 pages (pp. 417-528), les auteurs ont passé en revue 77 articles ou livres et analysé 49 études (p. 528, appendice B) publiées entre 1980 et 1998 dans les principales revues du domaine (par exemple *Studies in Second language acquisition* ou bien *Modern language journal*). Ces études, expérimentales ou quasi-expérimentales, portaient sur

l'efficacité de divers types d'enseignement d'une langue étrangère. Malgré certaines déficiences et un manque d'homogénéité méthodologique des enquêtes qui limitent la généralisabilité des résultats, les auteurs indiquent que les enseignements de type « explicite » (inductif ou déductif) sont plus efficaces que ceux de type « implicite », et que les effets d'un enseignement de type « Focus on Form » (étude des formes intégrée avec celle du sens dans le contexte d'activités communicatives) ou « Focus on FormS » (étude intensive des formes grammaticales artificiellement réparties dans le programme<sup>263</sup>), sont identiques. Bien sûr il convient de définir précisément chaque terme employé ici. Pour cela, nous renvoyons le lecteur à l'article en question<sup>264</sup>. Afin d'illustrer la distinction entre instruction explicite et implicite, Norris et Ortega donnent l'exemple d'un professeur qui lit à voix haute des extraits d'une histoire durant plusieurs classes avec pour seule consigne d'écouter puis de répondre à des questions de compréhension. Ce type d'instruction sera classé comme implicite. Ensuite est présentée une forme d'instruction explicite avec l'exemple d'un professeur qui oriente l'attention des étudiants en leur indiquant que l'histoire contient de nombreuses propositions relatives et en leur demandant donc de prêter une attention particulière à ces formes pendant l'écoute.

Il suffira, pour notre propos, de dire ici que, bien que la notion d' « efficacité » soit bien trop large lorsque l'on songe à la variété des tâches considérées et des méthodes d'évaluation des performances et de l'apprentissage des apprenants, il semble que la mise en focus attentionnel et la dimension métalinguistique qui caractérisent, d'une manière ou d'une autre, les enseignements explicites, soient profitables aux étudiants. Il faut cependant reconnaître que le terme « instruction formelle » peut recouvrir des réalités bien disparates.

Au vu de ces travaux, on comprend alors, que s'il est essentiel, pour diverses raisons (motivations, situations d'enseignement/ apprentissage, nécessité de pratiquer certaines

---

<sup>263</sup> Cette définition est celle de Long, cité par Norris et Ortega, (2000, p. 420). Pour plus de détails voir Norris et Ortega, 2000, p. 438) et R. Ellis (2002, p. 225 et p. 234, note 4 en bas de page).

<sup>264</sup> En particulier, pp. 436-437.



tâches, etc.), de favoriser le développement des savoirs implicites, comme cela peut être le cas des approches communicatives, de la problématique SGAV ou bien encore de la pédagogie par tâche, il ne faut pas dénigrer les savoirs explicites, au moins en tant qu'outils de développement des premiers. Une telle approche est particulière envisageable en situation d'apprentissage universitaire, contrairement à une situation d'immersion non-institutionnelle par exemple.

Il faut alors noter, comme nous l'avons évoqué précédemment, que la nature « explicite » de l'enseignement et/ou de l'apprentissage conduit à distinguer deux niveaux : l'un, que nous qualifierons de « linguistique » (et plus généralement de « cognitif »), l'autre de « métalinguistique » (et plus généralement de « métacognitif »).

### **II.2.3. Apprentissage explicite et réflexion métalinguistique**

Nous pouvons, dans un premier temps dans des termes très généraux, *voire grossiers*, tenter de cerner la distinction entre niveaux « cognitifs » et « métacognitifs » de la manière suivante :

- Au niveau cognitif : nous nous proposons d'englober tous les phénomènes de contact entre le monde extérieur et l'organisme (intégration des stimuli externes ou percepts)<sup>265</sup>. La cognition consisterait, pour nous, à intégrer des stimuli externes ou à produire des stimuli vers l'extérieur. Cette intégration/production peut parfois être appelée « pensée », en particulier lorsque l'intégration/production correspond à la compréhension/production de stimuli verbaux (dimension référentielle) (des phénomènes complexes de *projection* entre pensée et langage sont alors ici, comme toujours (Robert, 1997, p. 30), à l'œuvre).
- Au niveau métacognitif : nous nous proposons d'englober l'ensemble des phénomènes de contacts entre le domaine de la cognition et l'organisme. La

---

<sup>265</sup> Nous incluons dans ces percepts le cas de la proprioception (qui peut parfois s'avérer délicate à traiter dans le cadre que nous proposons ici), en particulier dans le domaine tactile, qui touche de près, au niveau des articulateurs internes, la question de la « prononciation ».

métacognition consisterait alors à prendre pour objet d'intégration / de production non plus des stimuli externes, mais des stimuli cognitifs (« pensée sur la pensée »). C'est ce que, dans l'expérience commune, on appelle « prendre du recul », « prendre conscience » ou bien encore simplement « penser ». Cette métacognition se traduit souvent par un « discours intérieur », ce dernier présentant de nombreuses singularités par rapport à un discours d'interaction communicative. Comme le note Trévis (1993, p. 41), « La langue 1 joue aussi un rôle important dans le discours intérieur du sujet, qui apparaît lors des résolutions de problèmes, des planifications, des schèmes opératoires de prise d'information, d'anticipations, d'inférence, de découverte de malentendu, de contrôle conscient de ces phénomènes ». La distinction entre langage et pensée devient ici bien difficile à appréhender. Dans cette conception, assez simpliste et séquentielle, nous l'admettons aisément<sup>266</sup>, la métacognition présenterait certaines propriétés de

---

<sup>266</sup> Le thème de la métacognition et de l'activité métalinguistique des sujets parlants a donné lieu à de très nombreux travaux et à une abondante littérature, qui dépassent largement le cadre de notre étude. C'est la raison pour laquelle nous adoptons une position relativement simpliste sur le sujet, quitte à faire parfois preuve d'imprécision, voire d'inexactitude. A titre d'exemple, nous pouvons mentionner les travaux de Gombert (2003), qui a proposé un modèle de développement métalinguistique en L1 chez l'enfant, relatif à l'apprentissage de la lecture et au développement de la parole, et qu'il résume dans son article de 2003. Gombert distingue également deux niveaux de contrôle cognitif, l'un qualifié d'épilinguistique, l'autre de métalinguistique, décrits en ces termes :

*« The first of these (epilinguistic level) is the control automatically exerted on linguistic processing by the linguistic organizations present in memory. The second (metalinguistic level) is the control consciously chosen, decided on and applied by the individual. While the first level of control is automatically present in all behaviour involving linguistic processing, the second level depends on an external demands or individual decisions ».*

Gombert considère qu'il existe une hiérarchie développementale entre ces deux niveaux (contrôle épilinguistique et conscience métalinguistique), niveaux qu'il inclut dans un modèle développemental plus global, présenté en ces termes :

*« 1) The phase of acquisition of first linguistic skills, in which, on the basis of pre-programmed formats, correspondences between linguistic forms and reinforced pragmatic contexts are stored in an implicit and instance-bound format ; 2) the phase of acquisition of epilinguistic control in which linguistic knowledge is reorganized in a multifunctional format which cannot be consciously accessed ; 3) the phase of acquisition of metalinguistic awareness, triggered by an external demand for intentional control of the organization established in phase 2 ».*

On retrouve, d'une certaine manière, dans ce modèle, le passage d'un apprentissage implicite à une conscience explicite, passage dans lequel l'enseignement et l'apprentissage explicites jouent un rôle d'accélérateur (« booster » dans les termes de Gombert) :

réurrence. La succession temporelle des étapes d'un raisonnement, par exemple, pourrait être envisagée comme une succession cyclique de niveaux, entre les niveaux cognitif et métacognitif, chaque pensée  $n$  (cognitif) devenant l'objet d'intégration de la pensée  $n+1$  (métacognitif). Il s'agit bien sûr là de niveaux essentiellement descriptifs mais aussi temporels, auxquels peuvent parfois être associés ce qu'on appelle grossièrement en psycholinguistique les bas et haut niveaux.

Parmi les différents savoirs déclaratifs dont dispose chaque individu, certains concernent plus particulièrement les systèmes de communication, le langage et les langues naturelles. Ces savoirs nourrissent donc une certaine activité d'observation, de réflexion et d'analyse des phénomènes langagiers, que cette activité soit consciente ou pas. Certains auteurs, comme Trévisse, se sont longuement penchés sur la question de l'activité métalinguistique et épilinguistique<sup>267</sup> des apprenants (1999, p. 92) : « Les apprenants, inconsciemment, feraient en quelque sorte la tâche d'un linguiste qui tente de représenter cette activité langagière ». Cette réflexion débouche sur des hypothèses (Trévisse et Demaizière, 1991, p. 356) :

*« On peut faire l'hypothèse qu'en milieu dit "guidé", les apprenants vont s'appuyer, plus ou moins consciemment, sur les représentations métalinguistiques qu'ils ont de leur langue maternelle dans des stratégies d'apprentissage contrastives [...]. Le système linguistique préalable et les représentations sur ce système jouent là, à notre avis, un rôle non négligeable ».*

Notons que nous tenons, comme Castelloti (2001, p. 78) à bien distinguer *activité métalinguistique* (processus conscient systématiquement opéré par les apprenants

---

*« One of the consequences of reading instruction is a huge increase in print-related activities. Attention to written words, previously occasional, becomes systematic and frequent. Therefore reading instruction boosts implicit learning ».*

<sup>267</sup> Terme forgé par A. Culioli en 1976 et qui correspond à l'activité métalinguistique inconsciente des individus.

d'une langue étrangère) de *réflexion métalinguistique* (ensemble organisé de description et d'analyse de faits de langue). Or, le didacticien, et à sa suite l'enseignant de langue, ne peut ignorer cette activité. Au contraire, il doit la prendre en compte, voire l'intégrer dans sa conception de l'apprentissage, et donc de l'enseignement (Trévisse et Demaizière, 1991, p. 362) :

*« Les apprenants, dans n'importe quel milieu (et dans tout type d'approche, y compris « communicative »), ont une activité métalinguistique. Le didacticien doit la cerner et la réguler ».*

Ceci est particulièrement vrai pour des apprenants adultes, qui, à la différence de l'enfant, sembleraient n'impliquer que leurs structures corticales et non plus sous-corticales dans l'apprentissage de la langue (Lebrun, 1984, p. 42). En situant l'activité métalinguistique de l'apprenant au niveau métacognitif, activité dont le noyau peut également s'apparenter aux phénomènes de « matching » évoqué précédemment, les questions qui se posent sont les suivantes :

- Peut-on considérer que cette activité puisse être caractérisée par des degrés d'explicitation, du point de vue de l'apprenant, sur une échelle aux extrémités de laquelle se trouvent les pôles « activité épilinguistique » et « réflexion métalinguistique », le degré variant en fonction du degré de verbalisation et de structuration des connaissances ?
- Existe-t-il des « interférences » / « conflits procéduraux » entre les activités situées au niveau cognitif et celles situées au niveau métacognitif ? Cette dernière question renvoie particulièrement au traitement en temps réel : en situation de classe, le flux des informations est imposé à l'apprenant qui doit mettre en œuvre les deux types d'activités. Bange, cité par Castelloti (2001, p. 62), évoque le « phénomène de bifocalisation », c'est-à-dire que :

*« L'attention des apprenants se focalise alternativement, de manière plus ou moins importante selon les contextes et les locuteurs, sur l'objet thématique de la communication (le contenu) et sur les problèmes rencontrés dans la réalisation des activités de communication (les moyens langagiers d'y parvenir) ».*

A l'oral en particulier, l'auditeur « ne peut exercer que peu de contrôle sur le débit et l'enchaînement des différents énoncés » (Merlet, 1998, p. 143), ce qui entraîne fréquemment des « courts-circuits », pour reprendre l'expression de Gaonac'h (1990b), entre processus de bas et de haut niveau, en raison de la trop grande attention accordée aux premiers. Sur cette question du contrôle en temps réel, Klein considère que lorsqu'un sujet est en train de parler, celui-ci est également toujours en train de réfléchir d'une manière ou d'une autre (« some reflection is always taking place while we speak »), et notamment sur sa propre production verbale. Il établit alors trois types de contrôle, définis en fonction de leur degré d'implication en temps réel (« on-line-ness ») (1986, p. 142) :

*« **Monitoring**, ou contrôle quasi-simultané sur notre discours, qui nous permet de modifier notre production langagière en cours ;*

***Feedback**, ou contrôle légèrement subséquent débouchant, dans bien des cas, sur une correction plus ou moins réussie ;*

***Reflection**, ou contrôle général sur notre performance langagière, qui ne coïncide pas avec notre production langagière »<sup>268</sup>.*

Ces types de contrôle peuvent constituer, d'une certaine manière, un exemple représentatif de ce que nous entendons par « interférence » entre niveaux cognitif

---

<sup>268</sup> « *Monitoring*, or virtually simultaneous control over our speech, which enables us to modify ongoing speech production ; *feedback*, or slightly delayed control resulting, in many cases, in more or less successful correction ; *reflection* or overall control over our language performance, not coincident with speech production » (notre traduction).

et métacognitif, même s'ils apparaissent ici comme des activités positives pour l'apprentissage.

- Quel degré d'explicitation des savoirs déclaratifs *de l'enseignant* est le plus propice à l'intégration cognitive des éléments de la L2 ? Cette question nous renvoie au rôle de l'instruction explicite (savoirs déclaratifs) dans l'apprentissage de savoir-faire langagiers (savoirs procéduraux). Nous pouvons alors noter la remarque de R. Ellis (2002, p. 234) :

*« There is general agreement and strong empirical evidence to show that FFI can affect explicit knowledge. It seems reasonable to conclude that it is easier to teach explicit knowledge than implicit knowledge. Elsewhere (Ellis, 1993, 1997) I have argued that explicit knowledge can facilitate the subsequent acquisition of implicit knowledge (e.g. by helping to make forms salient to learners). This being the case, another and possibly more tenable route to implicit knowledge might be to use FFI to develop an explicit understanding of how problematic structures work and then allow the human categorization ability to build implicit knowledge through the input made available in unfocused tasks and naturalistic exposure. Perhaps we do not have to bother with trying to teach implicit knowledge directly ».*

Une telle remarque conduit alors à la question suivante : est-il possible d'explicitier (du côté de l'enseignant) l'activité métalinguistique de l'apprenant, afin, d'une part de remédier aux éventuels conflits mentionnés plus haut, et d'autre part d'orienter cette activité pour qu'elle soit le plus économique possible pour le sujet apprenant ? Il s'agit donc de l'intérêt potentiel, pour l'apprentissage de savoir-faire linguistiques (et pas seulement de savoirs métalinguistiques), d'une analyse linguistique explicite des phénomènes langagiers en jeu, qui prennent en compte à la fois les données scientifiques de la linguistique, mais également le type d'activité méta- ou

épilinguistique des apprenants, et dont l'un des objectifs pourrait être de faire « prendre conscience » aux apprenants desdits phénomènes.

«Prendre conscience» correspondrait à ce que nous envisageons comme le passage d'un niveau cognitif (que nous pourrions très grossièrement décrire comme un niveau de traitement de l'information extéro- et proprio-ceptive résultant en une information « cognitive ») à un niveau métacognitif (niveau de traitement de l'information cognitive), et non uniquement à une simple mise en focus attentionnel, les deux aspects étant bien évidemment compatibles, voire nécessairement associés. On doit cependant probablement envisager ce passage non de manière binaire, mais plutôt comme une mise en conscience progressivement accrue, le long d'un continuum entre des prises de conscience quasi-procédurales et d'autres extrêmement explicites qu'on pourrait qualifier d'hyperdéclaratives<sup>269</sup>. L'état consécutif à une « prise de conscience » peut en effet donner lieu à une verbalisation, ou bien laisser le sujet dans l'incapacité de verbaliser ce dont il a pourtant bien « pris conscience ». Si l'intervention de l'enseignant préalable à, et cause de, celle-ci s'exprime sous forme déclarative, et si cette prise de conscience entraîne, comme c'est souvent le cas, une modification comportementale de l'apprenant, on peut considérer que le phénomène correspond à une influence, indirecte, des savoirs déclaratifs sur l'activité procédurale de l'apprenant. Ce constat personnel rejoint d'ailleurs certaines problématiques d'un courant relativement récent en didactique des langues et baptisé «Language Awareness», et qui est parfois lié aux publics précoces (Dabène, 1991, p. 57 ; Trévisse, 1993, p. 46 ; Pendanx, 1995, pp. 123-126).

La nécessité de réintégrer « l'activité pensante » des apprenants semble donc apparaître au sein de diverses écoles didactiques et le cœur du débat est présenté ainsi par Bérard (1995, p. 15) :

---

<sup>269</sup> Ceci nous est apparu lors d'une discussion avec Jean-Luc Nespoulous.

*« Le débat se situe au niveau de la nécessité ou non d'explicitier des règles, et du rôle de cette activité d'explicitation dans l'acquisition. La problématique développée autour de la notion de « awareness of language » peut recouvrir plusieurs zones : la nécessité de sensibiliser aux concepts en langue maternelle et en langue étrangère et à leurs réalisations, le développement d'activités de réflexion sur la langue étrangère ».*

L'intérêt, pour l'apprentissage, des « prises de conscience » des apprenants, relatives en particulier aux *écarts* entre leurs propres performances linguistiques en L2 et celles qu'ils visent à accomplir, est bien connu des didacticiens : « Une telle transition [d'un lecte d'apprenant au suivant] ne peut avoir lieu que tant que l'apprenant se trouve dans une position qui lui permette de remarquer un écart entre son lecte d'apprenant particulier et la cible elle-même »<sup>270</sup> (Klein, 1986, p. 57). Comme l'indique plus loin l'auteur (p. 139), « l'écart (dans un domaine particulier) n'est pas la même chose que la *perception* de cet écart [...]. Ce qui importe vraiment est l'évaluation subjective par l'apprenant de cet écart »<sup>271</sup>. Cela est particulièrement notable dans le domaine de la prononciation, comme le note Neri, Cucchiarini et Strik (2002, p. 180) :

*« The L1 influence can be so overwhelming that simple comparison of input with output may not lead to the perception of the discrepancies between the learner's interlanguage and L2 standards [...]. Feedback must then come into play. Through the provision of feedback teachers can bring the students to focus on specific problems, which hopefully stimulates them to attempt self-improvement ».*

---

<sup>270</sup> «Such a transition [from one learner variety to the next] can take place only as long as the learner is in a position to notice a discrepancy between his particular language variety and the target itself» (notre traduction).

<sup>271</sup> « Discrepancy (in a particular domain) is not the same as perception of that discrepancy [...]. What really matters is the learner's subjective evaluation of the discrepancy » (notre traduction).



Et si la « prise de conscience » de l'écart ne suffit pas à elle seule à résoudre le problème, la raison en est peut-être le manque *d'expertise* de l'apprenant, manque que ce dernier ne peut généralement combler sans l'aide de l'enseignant.

Au terme de ce rapide balayage de la problématique de l'enseignement/apprentissage explicite/implicite dans le domaine de l'apprentissage des langues étrangères, on peut globalement considérer que :

- Selon différents facteurs (tâche, situation, etc.), l'apprentissage peut être plus ou moins qualifié d'explicite ou d'implicite. Les deux types de mécanisme peuvent donc être mis en jeu en situation d'apprentissage, de la même manière que le sont les savoirs déclaratifs et procéduraux.
- Les apprenants de langue étrangère ont une activité métalinguistique relative à leur apprentissage.
- L'enseignement explicite peut être bénéfique à l'acquisition, tant de savoirs déclaratifs que de savoirs procéduraux, et influencer positivement sur les deux types d'apprentissage (explicite et implicite).
- L'activité métalinguistique des apprenants consiste notamment en une activité comparatiste entre la L1 et la L2, ainsi, sinon surtout, qu'entre leur interlangue (*via* leurs performances) et la L2, bien que la *perception* des contrastes pertinents constitue l'un des enjeux majeurs de cette activité.
- Cette activité métalinguistique concerne donc à la fois le « haut-niveau » (construction de catégories cognitives, etc.) et le « bas-niveau » (perception, réalisation articulo-motrice, etc.) cognitifs.

Partant de ces considérations générales, nous pouvons à présent revenir à notre domaine de prédilection, à savoir celui de la « prononciation » du français. Nous pouvons alors nous interroger sur la valeur d'un enseignement explicite pour l'apprentissage de la compétence visée, en insistant sur la nécessité de préciser la signification concrète d'un tel enseignement.

#### **II.2.4. Enseignement explicite et « prononciation »**

Même si le thème de l'instruction explicite a longtemps été attaché à celui de la « grammaire », certains chercheurs se sont également penchés sur le cas de la prononciation. Silveira (2002) a ainsi passé en revue plusieurs travaux concernant l'enseignement de la prononciation de l'anglais, à travers l'analyse de pratiques de classe, de manuels et de recherche empirique. Nous présentons ci-dessous les conclusions principales de son étude.

Comme elle le note, la place accordée à, et la conception de, l'enseignement de la prononciation a varié selon l'évolution des méthodologies didactiques dominantes (pp. 95-96) : imitation par répétition après une période d'écoute silencieuse ; enseignement de la phonétique et recours à l'alphabet phonétique international ; utilisation d'exercices de répétition basés sur les paires minimales ; enseignement à l'aide d'outils originaux (par exemple les baguettes colorées dans la « Méthode Silencieuse ») ; absence totale d'enseignement ; utilisation de laboratoires de langue pour que les apprenants puissent enregistrer et écouter leurs propres productions, insistance sur le domaine segmental au détriment du prosodique ou vice-versa.

En ce qui concerne les approches communicatives, sur lesquelles insiste l'auteur (pp. 97-101), l'objectif est non pas tant une prononciation « parfaite », qu'un degré acceptable d'intelligibilité et d'aisance (« fluency ») dans une perspective communicative. Plusieurs auteurs ont proposé certaines techniques et certains principes relatifs à l'enseignement de la prononciation dans le cadre de telles approches, en insistant notamment sur la motivation des apprenants, et ainsi sur la nécessité de donner une visée sémantique et communicative aux activités de prononciation. Silveira cite ainsi un exemple de Bowen (p. 98), qui mentionne en outre la prise en compte du facteur de fréquence d'occurrence dans la langue, lors de la conception de telles paires minimales (afin d'avoir une probabilité d'occurrence relativement similaire pour les deux énoncés) :

« *This pen leaks. (Then don't write with it.)*

*This pan leaks. (Then don't cook with it.)* »

D'autres auteurs, comme Celce-Murcia, citée par Silveira (p. 100), proposent de passer graduellement d'activités plutôt phonétiquement orientées (écoute et imitation ; pratiques de paires minimales ; pratiques de paires minimales contextualisées ; pratiques de « tongue twisters », etc.) à des activités plutôt communicativement orientées (jeux de rôles, résolutions de problèmes, etc.), selon le schéma suivant d'une « leçon de prononciation » :

- a) Description et analyse,
- b) Discrimination auditive,
- c) Pratique contrôlée avec retour correctif,
- d) Pratique guidée avec retour correctif,
- e) Pratiques communicatives avec retour correctif.

La conception de telles tâches communicatives, selon l'auteur, implique quatre étapes :

- a) Identifier les problèmes des étudiants,
- b) Trouver des contextes lexicaux/grammaticaux contenant de nombreuses occurrences des sons problématiques,
- c) Développer des tâches communicatives qui incorporent lesdits mots,
- d) Développer au moins trois ou quatre tâches en vue de recycler les problèmes et de pratiquer les sons problématiques dans de nouveaux contextes.

Analysant plusieurs manuels d'apprentissage de l'anglais, généraux ou bien spécialisés dans la prononciation, en distinguant d'une part le contenu et d'autre part le type d'activités proposés, Silveira résume dans le tableau reproduit ci-dessous les possibilités envisagées (pp. 102-103) :

<b>Pronunciation content</b>	<b>Teaching procedures</b>
a) Sound discrimination (vowels and consonants).	a) Connecting the pronunciation material with the language class or work environment.
b) Word and sentence stress.	b) Using visual aids such as mouth and lip illustration of sound articulation.
c) Rhythm.	c) Offering explicit instruction about pronunciation, including phonetic transcription.
d) Intonation.	d) Comparing the L1 and the L2.
e) Reduction.	e) Constant recycling of the pronunciation points.
f) Linking.	f) Providing learners with rule induction and deduction activities.
g) Deletion.	g) Listen and repeat activities.
h) Substitution.	h) Minimal pair drills.
i) Non-verbal behavior.	i) Contextualized minimal pairs.
j) Connection to vocabulary and grammar.	j) Tongue twisters, rhymes, poems, etc.
k) Inflectional endings (-s and -ed).	k) Silent practice.
l) Consonant clusters.	l) Reading aloud and recitation.
m) Practice of vowel shifts and stress shifts related to affixation.	m) Recording learners' speech samples.
	n) Practicing at the word, sentence and paragraph level.
	o) Developing self-monitoring strategies.
	p) Contrasting spelling and sounds.
	q) Working in pairs or groups so that learners receive peer feedback.
r) Working with naturalistic speech samples as used by native speakers.	

Si l'on peut regretter que les deux colonnes du tableau ne soient pas véritablement mises en correspondance ou davantage hiérarchisées, il s'agit néanmoins d'une grille d'analyse fort utile, et qui prend en compte à la fois l'utilisation des supports visuels (b), la comparaison de la L1 et de la L2 (d), ainsi que le rapport entre sons et orthographe (p).

Parmi certains traits qu'elle relève dans son analyse de manuels, et qui s'avèrent particulièrement pertinents vis-à-vis des lignes que nous avons suivies jusqu'ici, on peut noter que certaines de ces « méthodes » tentent de prendre en compte la L1 des apprenants dans les activités proposées, soit de manière théorique, ce qui ne conduit pas toujours aux résultats escomptés<sup>272</sup> (p. 104), soit lors de la conception des contenus du manuel, basée sur une analyse préalable d'échantillons de parole d'apprenants de L1 différentes (p. 107), soit en proposant, en annexe et dans certaines unités, certaines activités de contrastes phonémiques destinées à un public de L1 spécifique (p. 109).

Par ailleurs, certains auteurs accordent une importance non-négligeable à la syllabe dès le début de leur manuel, soit à travers des activités de comptage syllabique (p. 109), soit à travers une description détaillée (définition, comptage, relation à l'accentuation, à la longueur et à la réduction syllabiques, etc.) de la syllabe (p. 110). Il est également particulièrement intéressant de noter que, dans le dernier cas mentionné, les auteurs du même manuel<sup>273</sup> ont conçu des tâches de discrimination auditive de mots non-transcrits orthographiquement. Leur objectif, d'après Silveira, était le suivant (p. 111) : « Hagan and Grogan are trying to prevent learners from being dependent on spelling ». Si cela rejoint pleinement nos préoccupations, il faut néanmoins noter que l'ouvrage en question est un manuel spécialisé dans l'apprentissage de la prononciation, et non un

---

<sup>272</sup> « *The exercises are not designed to suit all of the learners with different L1 background, since the way sounds are contrasted is not always appropriate for some L1 learners. [...] For example, the sounds /j/ and /dʒ/ are supposed to be a problem for Portuguese learners, but actually, for Brazilian Portuguese learners, the problematic contrast is /ʒ/ and /dʒ/* » (Silveira, 2002, p. 104, note 4).

<sup>273</sup> Hagan, S. A. et Grogan, P. E., *Sound advantage: a pronunciation book*, Englewood Cliffs, NJ: Prentice Hall Regents, 1992, cités par Silveira (2002).

manuel d'apprentissage général. Au terme de son analyse, Silveira reste réservée quant aux pratiques proposées (p. 111), vis-à-vis du cadre communicatif dont elle se préoccupe :

*« There is a lot to be done to develop materials that approach pronunciation with a focus on communication and intelligibility. [...] Despite this effort to include a great deal of pronunciation items, the manual still stop short of offering tasks that range from the more controlled to the more communicative ».*

Après cette analyse de manuels, l'auteur se propose de passer en revue un certain nombre de travaux de recherche sur le thème qui nous occupe plus précisément ici, à savoir les effets de l'enseignement sur la prononciation des apprenants (pp. 112-122). Parmi certaines informations remarquables, nous pouvons noter les résultats positifs, en tâche de discrimination auditive de stimuli naturels et synthétiques, de Jamieson et Morosan (cités par Silveira, p. 113), qui incluent trois composants dans leur module d'entraînement : « a) acoustic training appropriate for normal speech ; b) identification training with immediate feedback ; and c) acoustic uncertainty provided by increasing variability in the signal ».

Ce dernier point correspond à l'utilisation de la « fading technique », qui nous intéresse particulièrement en raison de la manipulation de l'input, et ainsi de la norme, phonétique, qu'elle implique, et qui évoque, en partie au moins, certaines techniques de la méthode verbo-tonale. Cette technique est décrite ainsi par Jamieson et Morosan (cités par Silveira, 2002, p. 113) :

*« This technique attempts to train a perceptual contrast, without subject errors, by beginning with clearly discriminable stimuli which may exaggerate the normal perceptual differences or add other salient features. Progress in*

*training is made by slowly reducing the magnitude of the perceptual contrast, in small steps, so that the task never becomes too difficult and errors remain infrequent ».*

Il est également intéressant de noter que l'effet de l'enseignement peut varier dans le temps : comme le rapporte Silveira (2002, p. 114), certains chercheurs ont en effet montré que les performances des apprenants en tâche de discrimination phonémique suivaient une progression « en U » (suivant l'idée selon laquelle les performances diminuent pendant que les représentations internes se complexifient pour ensuite s'améliorer une fois cette complexification effectuée, en direction de l'expertise), suite à la réception d'une instruction explicite concernant la prononciation.

Une autre étude que l'on peut mentionner ici, malgré la tâche particulière qu'elle contient (tâche de lecture) qui fait intervenir la dimension phonographémique, est celle de Baqué (2005), qui s'est intéressée au traitement du /ə/ et du /e/ par des apprenants hispano-catalophones de français, en fonction de la tâche (lecture d'un texte littéraire vs. lecture d'un exercice grammatical à trous « où les mêmes mots se trouvaient tantôt en situation de focalisation sur la forme tantôt dans la phrase qui sert de contexte ») (p. 8) et de la focalisation sur la forme phonique (« a) lecture libre et continue d'un texte (focalisation minimale sur la forme (LL)), b) lecture des parties de phrases de l'exercice grammatical qui ne font pas l'objet d'une attention particulière (focalisation intermédiaire sur la forme (L)), c) lecture des mots complétés par l'étudiant durant l'exercice grammatical (focalisation maximale sur la forme (GRAM)) » (p. 8)).

La conclusion de cette étude est globalement positive pour la condition GRAM, puisque, au niveau du groupe global, celle-ci permettrait de mieux différencier les deux phonèmes en question. Cependant, Baqué observe également une importante variation intra-individuelle, qui la conduit à distinguer trois sous-groupes en fonction de l'effet

(« phonologique », « phonétique » ou « ultracorection ») de la focalisation sur la forme (p. 11) :

1. Effet phonologique : « *Le degré de focalisation sur la forme entraîne une amélioration progressive des productions à la fois de /ə/ et de /e/ et, partant, une meilleure (bien qu'insuffisante) différenciation entre les deux. [...] Ainsi, l'augmentation de la focalisation sur la forme a un effet de différenciation phonologique* ».

2. Effet phonétique : « *L'augmentation de la focalisation sur la forme entraîne des valeurs supérieures de F1 et de F2, ce qui se manifeste par une amélioration des réalisations de /e/, mais également par un éloignement de celles de /ə/ du [ə]-cible. Il semblerait qu'il s'agisse donc là d'un simple effet phonétique – facilitateur pour l'un des phonèmes et difficile pour l'autre – sans réelle valeur de distinction phonologique* ».

3. Effet d'ultracorection : « *L'augmentation de la focalisation sur la forme entraîne les effets contraires que ceux observés dans le deuxième sous-groupe [...]. Il en découle que les réalisations de /ə/ connaissent une nette amélioration, mais que celles de /e/ s'éloignent encore plus du [e]-cible, se confondant parfois avec des [ə]. Il semble donc que [...] ce groupe opère une ultracorection sur le micro-système /ə-e/* ».

Baqué conclut dès lors son étude par certaines recommandations pédagogiques, nécessitant d'adapter les activités de correction phonétiques proposées par l'enseignant, ainsi que le degré de focalisation sur la forme qu'elles impliquent, à chacun de ces trois



sous-groupes d'apprenants. Si globalement la focalisation sur la forme semble donc être bénéfique, elle ne semble pas l'être de la même manière pour tous.

Cette conclusion mitigée rejoint celle de Silveira (2002), issue de tous les travaux qu'elle a examinés, dont l'hétérogénéité ne permet pas, essentiellement en raison d'aspects méthodologiques, de retirer de conclusions claires (pp. 119-120) :

*« The controversial results yielded by the studies reviewed [...] come as no surprise if we observe their heterogeneous design. [...] Data can only be obtained if future research succeeds at specifying the target pronunciation aspect and the teaching procedures, as well as at measuring the effects of the treatment adequately ».*

L'auteur insiste alors sur la nécessité d'une meilleure formation en phonétique et phonologie appliquée pour les enseignants et les chercheurs du domaine, indispensable à une meilleure connexion entre la théorie, la recherche et la pratique :

*« It seems that the link between theory and research, added to careful research designs and abundant research results may have to effects. First, it could help educators to realize the importance of pronunciation instruction in the language curriculum. Second, writers of pronunciation and general language materials might be able to re-evaluate the extent to which their work is appropriately addressing the pronunciation component ».*

En conclusion, Silveira indique que, outre la nécessité de davantage intégrer la dimension communicative aux activités de prononciation, il semble important que les concepteurs et enseignants fournissent aux apprenants des activités motivantes, et, pour cela, que les points abordés puissent être reconnus par les apprenants comme

importants pour l'amélioration de leur prononciation, ce qui correspond tout à fait nos observations personnelles de classe. De cette analyse, sont issues certaines propositions méthodologiques, qui rejoignent pleinement nos préoccupations et notre démarche, et qui sont formulées ainsi par l'auteur (p. 123) :

*« That is why I suggest that ideally, at least in FL [Foreign Language] settings, the pronunciation manuals should have different versions, with exercises that address specific difficulties posed by L1 interference. The selection of pronunciation difficulties could be done with help of interphonology research and learners' pronunciation profiles ».*

Cela évoque donc une démarche de pédagogie différenciée que nous soutenons pleinement, et qui, de manière très générale, « s'organise à partir d'un ou plusieurs éléments caractéristiques de l'hétérogénéité des élèves comme leurs différences cognitives, socioculturelles, psychologiques. [...] La pédagogie différenciée se définit donc comme : a) une pédagogie individualisée qui reconnaît l'élève comme une personne ayant ses représentations propres de la situation de formation ; b) une pédagogie variée qui propose un éventail de démarches, s'opposant ainsi au mythe identitaire de l'uniformité, faussement démocratique, selon lequel tous doivent travailler au même rythme, dans la même durée, et par les mêmes itinéraires » (Przesmycki, 1991). Dans le cas de la prononciation et de l'établissement des « profils d'apprenant », on peut noter, d'après notre expérience pédagogique personnelle, que le laboratoire de langue constitue un très bon poste d'observation des erreurs individuelles, que ce soit en tâche de lecture à voix haute ou de répétition, permettant d'établir de tels profils.

La démarche que propose Silveira vis-à-vis de la sélection et de l'organisation des contenus correspond ainsi tout à fait à celle que nous avons suivie, faisant appel au domaine de l'interphonologie, en vue, notamment, d'établir un classement de

difficultés des groupes consonnantiques du français pour les apprenants japonais (tâches C1 et C2 dans notre protocole) ainsi que des couples B/V et R/L dans un certain type de tâche (tâches A et F dans notre protocole). De plus, les deux traits qu'elle mentionne, dans le cas d'apprenants brésiliens d'anglais, soulignent l'actualité des problématiques que nous traitons dans notre étude, puisqu'il s'agit de la syllabe, des groupes consonnantiques et du rapport entre sons et orthographe en L1 et en L2 :

*« In the case of Brazilian learners, interphonology research [...] indicates that pronunciation materials should deal with the syllable component, exploring, for example, the difficulties posed by consonants clusters and syllable-final consonants, the negative effects of using epenthesis to pronounce such clusters, as well as plenty of activities that call attention to the distinction between spelling and sound across languages ».*

Il faut alors souligner, dans le cas du français langue étrangère (comme dans toutes les langues à orthographe « profonde »), que cette distinction entre sons et orthographe prend une dimension toute particulière en français lorsque l'on considère l'écart existant entre les codes oral et écrit, dans le domaine de la morphologie verbale en particulier. L'attention portée aux problèmes de prononciation concerne donc également, dans une certaine mesure, l'acquisition du code écrit et la réalisation de multiples tâches : répétition, production orale (libre ou contrainte), production écrite, et bien sûr compréhension orale, qui renvoie au pendant perceptif de la boucle audio-phonatoire.

Cependant, il reste encore, comme le souligne Silveira à propos de « pronunciation instruction », à définir ce que l'on entend par « enseignement explicite » dans le cas de la prononciation. L'une des manières d'aborder la question est précisément de passer par la notion d'*attention* qu'elle mentionne dans sa conclusion, et dont le traitement renvoie plus généralement aux *options méthodologiques* adoptées dans le programme d'enseignement.

### **III. Considérations méthodologiques : attention sur la forme, outils linguistiques et norme didactique**

Ayant examiné les fonctions potentielles de l'information visuelle et de l'enseignement explicite pour l'apprentissage de l'oral, il nous faut à présent amorcer l'intégration de ces données dans une perspective méthodologique plus globale, perspective qui nécessite au préalable d'examiner certaines options, notamment, dans notre cas, celles relatives 1) à l'orientation attentionnelle des apprenants, 2) à l'utilisation des outils fournis par les sciences du langage (modèles et corpus) et 3) à l'élaboration d'une norme didactique.

#### **III.1. Attention sur la forme et approches didactiques**

Comme cela est apparu précédemment, le rôle joué par ce facteur transversal qu'est l'attention semble crucial dans la caractérisation de l'enseignement en termes d'explicite et d'implicite. Avant même de considérer le degré d'attention porté par les apprenants à la dimension phonético-phonologique de la L2, qui permettra de caractériser d'implicite ou d'explicite l'apprentissage, il nous faut considérer celui qu'y porte l'enseignant, puisque c'est, dans la plupart des cas, de celui-ci que dépend le premier, en fonction de l'approche et des tâches didactiques mises en œuvre en classe. En effet, lorsque l'on traite d'explicite en classe de langue, on doit se demander : 1) si ce qui est explicite l'est pour l'enseignant, pour les apprenants ou pour les deux, et 2) si ce qui est explicite est identique pour les deux.

Cinq cas de figure sont globalement envisageables, que nous schématisons ci-dessous (avec Ens = enseignant ; App = apprenants ; Dém = démarche ; Cont = contenu ; Imp = implicite ; Exp = explicite) :

Code	Ens		App		Nom	Description
	Dém	Cont	Dém	Cont		
I	Imp	Imp	Imp	Imp	Approche totalement implicite	Ni l'enseignant, ni les apprenants ne prêtent une attention particulière à la dimension phonético-phonologique du matériel qu'ils sont en train de traiter. <u>Exemple</u> : écoute et explication sémantico-grammaticale d'un dialogue oral transcrit à l'écrit.
DE-I	Exp	Imp	Imp	Imp	Démarche explicite - implicite	L'enseignant propose des activités dont l'objectif d'enseignement est globalement l'apprentissage de la prononciation du français (sans plus de précision), objectif non perçu comme tel par les apprenants. <u>Exemple</u> : tâche de répétition d'un dialogue, perçue comme une tâche de « mémorisation » par les apprenants.

DCE-I	Exp	Exp	Imp	Imp	Démarche et Contenu explicites - implicite	Comme la précédente, mais cette fois-ci l'enseignant vise un point phonético-phonologique précis. <u>Exemple</u> : l'enseignant fait répéter plusieurs fois une partie d'un énoncé contenant ce point.
DE-DE	Exp	Imp	Exp	Imp	Démarche explicite – explicite	L'enseignant propose des activités dont l'objectif d'enseignement est globalement l'apprentissage de la prononciation du français (sans plus de précision), objectif perçu comme tel par les apprenants. <u>Exemple</u> : tâche de répétition d'énoncés divers en cours de correction phonétique.
DCE-DE	Exp	Exp	Exp	Imp	Démarche et Contenu explicites – Démarche explicite	Comme la précédente, mais cette fois-ci l'enseignant vise un point phonético-phonologique précis, sans que ce point soit perçu comme tel par les apprenants. <u>Exemple</u> : l'enseignant fait répéter plusieurs fois une partie d'un énoncé contenant ce point.

DCE-DCE	Exp	Exp	Exp	Exp	Démarche et Contenu explicites – explicites	Comme la précédente, mais cette fois-ci les apprenants sont conscients du point visé par l'enseignant.  <u>Exemple</u> : l'enseignant fait répéter plusieurs fois une partie d'un énoncé contenant ce point, en l'expliquant ou en l'indiquant, d'une manière ou d'une autre, aux apprenants.
---------	-----	-----	-----	-----	---	---

Nous laissons de côté ici les cas où les apprenants se concentreraient sur la prononciation indépendamment de l'enseignant, car cela relèverait des stratégies d'apprentissage individuel, dont nous ne nous préoccupons pas ici, dans notre perspective méthodologique d'enseignement. Notons également que la situation d'apprentissage spécifique dans laquelle se trouvent les apprenants constitue un facteur déterminant dans ce descriptif : il va de soi qu'une séquence didactique consacrée à la prononciation durant un cours de langue général, une classe de correction phonétique ou une classe consacrée à l'oral n'ont pas le même effet sur la dimension explicite de l'enseignement/apprentissage qu'une classe intitulée « cours de grammaire » ou « cours de civilisation ». Pourtant, l'apprentissage de l'oral s'effectue bel et bien à travers l'ensemble de ces enseignements. C'est également la raison pour laquelle, même une méthode de correction phonétique comme la méthode verbo-tonale, qui prône « l'absence d'intellectualisation » de la part des apprenants (Renard, 1979), est bien une méthode que l'on peut globalement qualifier d'« explicite ». Cette dernière remarque permet donc de mettre en garde contre certaines assimilations abusives, dans la perspective qui est la nôtre, entre « explicite » et « déclaratif » ou « explicite » et « formel » par exemple.





### **III.1.1. Attention : niveau cognitif**

Etant donné les deux axes méthodologiques d'enseignement qui sont jusqu'ici ressortis de notre examen, à savoir « faire percevoir » et « faire prendre conscience », on comprend que le rôle de l'attention, aux multiples niveaux auxquels elle intervient, soit si crucial en didactique des langues, et ceci tout particulièrement pour les concepteurs d'outils multimédiatisés, en regard, notamment, des problèmes de limitations et de partage attentionnels. On sait en effet que l'apprentissage, et, dans une perspective mnésique, le rappel, d'informations est limité par la quantité de capacité attentionnelle disponible, si bien que l'accomplissement simultané d'une autre tâche nécessitant une certaine attention peut dégrader la performance dans la première. Cela renvoie au problème de conflits procéduraux d'ordre attentionnels (Baddeley, 1986, p. 38). Dans le cas de l'attention spatiale, par exemple, il faut noter la distinction entre attention globale et attention locale, distinction qui nous renvoie aux bases cérébrales de ces processus (Siéhoff, 1994, p. 144). Si le concepteur/l'enseignant élabore un matériel adapté à son public et à la cible d'enseignement, encore faut-il que les apprenants y « fassent attention ». On peut bien sûr percevoir sans « faire attention » (perception subliminale, entre autres), ce qui nous renvoie à la dichotomie entre attention passive et attention focalisée. Mais il est alors très difficile, pour l'enseignant, d'évaluer et d'exercer un contrôle sur les éléments perçus ou non. Il est, nous semble-t-il, plus facile de conduire un sujet à « faire attention », qu'à « ne pas faire attention », sans que le résultat soit garanti pour autant.

Berry et Dienes, cités par Meulemans (1998, p. 25), indiquent que la mémoire implicite et l'apprentissage implicite sont fortement dépendants des caractéristiques de surface de l'information. Ceci est important, car quels que soient les modèles et les théories, linguistiques ou cognitifs, sous-jacents aux cours dispensés par l'enseignant, l'accent est mis ici sur le format de surface de l'information traitée par les apprenants. Dans les tâches classiques d'apprentissage de grammaire artificielle, on remarquera ainsi que la présentation se fait sous forme graphique sans métalangage linguistique (voir Fig. 1, Meulemans, 1998, p. 25). Aussi, de quelle manière l'enseignant doit-il orienter

l'attention des apprenants et selon quel médium (verbal/gestuel ; visuel/auditif) ? On sait par exemple que (Hamon, 1998, p. 105) :

*« Ce sont l'intensité, la nouveauté et la charge affective des stimulus qui provoquent à la fois le réveil (passage du sommeil à l'éveil sur le continuum de la vigilance) et l'éveil de l'attention. Il est en effet connu qu'un stimulus intense élève la vigilance en provoquant une réaction d'orientation. De même, les modifications qui affectent les propriétés physiques des stimulus ou leur succession temporelle, ainsi que le déplacement des sources de stimulation par rapport aux organes des sens (déplacements spatiaux) ont des effets proxigènes (éveillants) ».*

Quelles règles l'enseignant doit-il respecter pour ne pas créer de « parasites attentionnels » ? En situation de production orale, par exemple, l'apprenant doit effectuer plusieurs tâches de rappel, et l'on peut se demander si la présentation de règles de production sous forme linguistique (métalangage) ne va pas diminuer les ressources cognitives affectées au rappel des items lexicaux par exemple, produisant des parasites entre le rappel des « règles » formulées linguistiquement et la production du message, linguistique lui aussi. Ces parasites ne seraient-ils pas supprimés si les « règles » étaient présentées sous un format autre que linguistique ? Si oui, quel format ? Le coût cognitif d'intégration des informations en coréférence, par exemple, en fonction du format de présentation (intégration texte/image ; texte/texte) a notamment été évalué par Sweller et ses collaborateurs (Tricot, 1998) qui parviennent à émettre des recommandations ergonomiques de conception des documents, en vue de réduire les effets de partage attentionnel et diminuer les temps d'apprentissage (Bétrancourt et Caro, 1998, p. 160 ; Virbel, 1997, pp. 256-257). En situation de classe « traditionnelle » se pose ici la question du comportement, verbal et non verbal, de l'enseignant. Comme l'indique Juan de Mendoza (1998, p. 84) :

*« La notion de "filtre attentionnel" est souvent invoquée pour rendre compte du fait que l'information sur laquelle est dirigée l'attention (l'information cible)*

*prend un relief particulier qui la fait apparaître perceptivement plus saillante que les autres informations simultanément présentes dans le champ perceptif (les distracteurs) ».*

Renversant la citation de Mendoza, il s'agirait, nous semble-t-il, de *faire apparaître perceptivement plus saillant l'input issu de la cible d'enseignement par rapport aux distracteurs*. C'est ici que l'outil multimédia prend toute sa valeur, et que les études en ergonomie liées aux hypermédias dans le domaine de l'enseignement/apprentissage se révèlent riches d'enseignements (Virbel, 1997, p. 260 ; Bétrancourt et Caro, 1998, pp. 158-159 ; Caro et Bétrancourt, 1998, pp. 123-136 ; Tricot, 1998, p. 57 ; Jamet, 2001). Il faut cependant ici noter que, si les processus attentionnels peuvent être qualifiés *d'universaux*, et donc de communs à tous les apprenants, les formes et les principes d'organisations auxquels ils s'appliquent varient en fonction des groupes visés. *Attirer l'attention* de tel apprenant peut nécessiter un ensemble de dispositions qui diffère de celui requis pour attirer l'attention de tel autre, en fonction des groupes d'origine en particulier (culturels, linguistiques, géographiques, sociologiques, etc.).

Ainsi, qu'il s'agisse de stimuli auditifs ou visuels dans l'enseignement des formes orales du français (quelle que soit la tâche d'apprentissage), il semblerait que les caractéristiques de surface de l'input (matériel ou comportement de l'enseignant) jouent un rôle non négligeable dans sa perception par les apprenants. La notion qui semble alors cruciale ici est celle de *saillance*. Cependant, une telle notion est bien souvent utilisée de manière assez générale, alors qu'elle doit être caractérisée plus précisément si l'on souhaite en faire un concept didactique opérationnel.

C'est ce que fait Landragin (2004), en confrontant des travaux issus de disciplines variées, qui distingue deux grands types de saillance : l'une physique (« P-saillance »), l'autre cognitive (« C-saillance ») (p. 2) :

*« Du point de vue physique, être saillant, c'est ressortir en premier de la perception d'un message visuel ou linguistique compte tenu de la forme que*

*prend ce message, c'est-à-dire des propriétés de ses composants (et compte tenu des caractéristiques physiologiques de l'organe de perception). Du point de vue cognitif, être saillant, c'est ressortir en premier de la perception d'une scène compte tenu des représentations mentales de l'utilisateur, de ses intentions de perception, de son attention visuelle et de sa mémoire à court-terme, de ses expériences personnelles et de ses émotions ».*

Après avoir positionné la notion de *saillance* par rapport à celles d'*attention*<sup>274</sup>, de *pertinence*<sup>275</sup> et de *prototype*<sup>276</sup>, l'auteur entreprend une classification des facteurs de saillance physique et cognitive, dans le domaine linguistique et visuel. Dans le domaine linguistique, qui nous intéresse plus particulièrement, les facteurs retenus sont les suivants<sup>277</sup> :

---

<sup>274</sup> « L'attention du sujet sur une certaine entité peut augmenter la saillance de cette entité, et inversement, la saillance d'une unité peut entraîner l'attention du sujet sur celle-ci, mais il n'y a en aucun cas équivalence entre saillance et attention » (p. 3).

<sup>275</sup> « La saillance nous semble intervenir [...] au niveau des processus cognitifs inférieurs (attention, perception et mémoire), alors que la pertinence relève des processus cognitifs supérieurs (représentation mentale, raisonnement, jugement, décision, résolution de problèmes) » (p. 4).

<sup>276</sup> « S'il est vrai que des caractéristiques saillantes peuvent être à l'origine d'un prototype, et que la connaissance d'un prototype renforce la saillance de certains objets, il n'en reste pas moins que la saillance est liée à un objet précis dans un environnement précis, alors que le prototype est lié à une représentation mentale durable et culturelle. [...] L'influence d'un prototype est l'un des facteurs cognitifs de saillance pouvant intervenir lors de la perception immédiate » (p. 4).

<sup>277</sup> Pour le détail des facteurs, voir Landragin (2004, pp. 4-15).

Saillance	Liée	Facteurs
<b>Physique</b>	<b>A la forme de l'énoncé</b>	Saillance intrinsèque au mot.
		Saillance due à une mise en avant explicite lors de l'énonciation.
		Saillance due à une construction syntaxique dédiée.
		Saillance syntaxique liée à l'ordre et à la fréquence d'apparition des mots.
		Saillance liée aux fonctions grammaticales.
		Saillance indirecte par transfert grammatical de saillance.
	<b>Au sens de l'énoncé</b>	Saillance liée à la sémantique des mots.
		Saillance liée au rôle thématique.
		Saillance liée au thème et au topique de l'énoncé.
		Saillance liée au propos de la conversation.
		Saillance liée à des inférences.
		Saillance indirecte par transfert sémantique de saillance.
<b>Cognitive</b>	Saillance liée à l'intention.	
	Saillance liée à l'attention.	
	Saillance liée à la mémoire à court terme.	
	Saillance liée à la mémoire à long terme.	
	Saillance liée à la personnalité.	
	Saillance liée aux affects et aux émotions.	

Une telle grille, malgré d'inévitables limitations, fournit un socle intéressant aux concepteurs et aux enseignants qui souhaiteraient, comme nous l'avons nous-mêmes exprimé, rendre *plus saillants* certains éléments de l'input en classe de langue. Il importe en effet de déterminer, autant que possible, la saillance la plus pertinente en fonction de l'objectif visé.

Procédant de la même manière dans le domaine visuel<sup>278</sup>, puis comparant les résultats, Landragin parvient à établir une caractérisation générique de la notion de saillance, qu'il présente ainsi (p. 17) :

<b>Facteur générique de saillance</b>	<b>Exemple visuel</b>	<b>Exemple linguistique</b>
Critère intrinsèque aux unités	Luminosité inhérente	Nom propre
Simplicité	Bonne forme	Interjection
Isolement	Isolement spatial	Apposition
Mise en évidence explicite	Eclairage spot	Accentuation
Construction dédiée	Valorisation du sujet	Présentatif
Placement à un endroit stratégique	Tiers du cadre	Début de l'énoncé
Répétition	Régularité	Répétition
Symétrie	Symétrie	Chiasme
Rupture dans une continuité	Disposition linéaire	Rythme d'élocution
Infraction d'une règle implicite	Disposition incongrue	Raté, erreur
Exploitation d'une norme	Composition classique	Construction usuelle
Structuration du message	Ligne de force	Construction à topique

<sup>278</sup>Facteurs physiques: a), saillance liée aux contrastes de luminosité et de couleur, b) saillance liée à une singularité dans un ensemble d'objets, c) saillance liée à la cohésion et à la structuration de la scène. Facteurs cognitifs: a) intention, b) attention, c) mémoire à court terme, d) mémoire à long terme, e) affect et émotions. L'auteur donne en particulier certaines informations relatives aux contrastes de couleurs qui pourraient être fort utiles aux concepteurs d'outils pédagogiques (p. 13) :

« La saillance liée aux variations de perception des couleurs recouvre deux aspects : un premier relatif aux temps de latence, et un second relatif à l'effet spatial des couleurs. Le temps de latence, c'est-à-dire le décalage entre le début de l'excitation et celui de la sensation varie selon les couleurs. Baticle (1985) donne les chiffres suivants : rouge (22.6 millièmes de seconde) ; vert (37.1 ms) ; gris (43.4 ms) ; bleu (59.8 ms) ; jaune (96.3 ms). Ces chiffres qui constituent une échelle de saillance doivent néanmoins être relativisés : pendant le jour, le maximum de sensibilité est dans le jaune et le rouge, alors que c'est le bleu le plus saillant pendant la nuit. D'autre part [...] une couleur appelle sa complémentaire. Un objet qui prend la couleur complémentaire de celle du fond en devient par exemple très saillant. [...] (Cocula & Peyrouet, 1986) propose ainsi une hiérarchie de couples de couleurs en fonction de l'importance de leur impact visuel : noir sur blanc ; puis noir sur jaune ; puis rouge sur blanc ; etc. Enfin, (Itten, 1985) note l'importance de l'effet spatial des couleurs : contrairement à du violet, du jaune sur fond noir a l'air d'avancer et voit ainsi sa saillance augmenter. Ces données nous donnent les bases pour un modèle de la saillance chromatique d'un objet en fonction de la luminosité ambiante, de la couleur dominante du fond et de la couleur dominante de l'objet ».

Tout en notant les imperfections de cette première approche, l'auteur conclue en insistant sur l'importance de la saillance (p. 19) :

*« La saillance ne fait pas explicitement partie du message communiqué, mais tout le message se fonde sur elle, se structure en fonction d'elle, et s'explique par elle ».*

Sans vouloir accorder une importance exagérée à la saillance, il nous semble que, dans un cadre didactique, ce facteur joue un rôle essentiel dans l'apprentissage, et que l'enseignant peut agir sur celui-ci, notamment lorsqu'il souhaite attirer l'attention des apprenants sur la forme du matériau linguistique, dans notre cas, essentiellement phonético-phonologique et phonographémique.

A titre d'exemple illustratif minimal, si l'on prend le facteur d' « isolement » ou de « mise en évidence explicite », il serait souhaitable que, dans les manuels de français destinés à des publics japonais, la présentation de l'alphabet que l'on trouve fréquemment en colonne dans les premières pages de l'ouvrage « isole » spatialement ou « mette explicitement en évidence » visuellement certains graphèmes classiquement problématiques pour ce public, en particulier le <R> et le <L>.

Ces données sur la saillance et l'attention au niveau que nous avons qualifié de « cognitif » doivent donc orienter la conception de matériel pédagogique, notamment en vue de permettre la « perception » et la « prise de conscience » des écarts entre la L1 et la L2 d'une part, et ceux entre L2 et l'interlangue des apprenants d'autre part. Cependant, de la même manière que la saillance implique une dimension physique et une dimension cognitive, nous devons à présent nous pencher, non plus tant sur le format de l'input que sur son contenu, ce qui nous conduit ainsi au niveau que nous avons qualifié de « métacognitif ».

### **III.1.2. Attention : niveau métacognitif**

Nous rejoignons ici la problématique de l'enseignement explicite, voire « formel », en considérant l'activité métalinguistique des apprenants et le rôle que peut y jouer l'attention. En particulier, nous nous demandons si une réflexion métalinguistique dirigée par l'enseignant peut orienter l'attention des apprenants sur certains aspects de la langue autrement plus difficilement accessibles.

Comme nous l'avons évoqué à propos de la perception, l'orientation de l'attention pour faciliter l'apprentissage peut se faire de deux manières : soit à travers l'organisation spatiale et temporelle du matériel présenté, soit à travers la mention verbale explicite de consignes ou d'instructions. La première manière se rapporte au niveau du traitement cognitif, la seconde au métacognitif. La question de la « familiarité » évoquée par Jacoby (Van der Linden, 1994, p. 297) est intéressante, puisqu'elle touche à la subjectivité du sujet, ce qui nous permet de retrouver le point de vue de l'apprenant. La facilitation du traitement, perceptif ou sémantique, grâce aux expériences passées et aux traces mnésiques, provoque un « sentiment de familiarité subjective », qui est lié à des mécanismes inférentiels ou d'attribution inconscients.

Si nous rattachons ceci à la distinction entre niveau cognitif et métacognitif, la question que nous posons est la suivante: en situation didactique classique, le traitement métacognitif d'un matériel verbal, peut-il en faciliter le traitement cognitif consécutif par effet de familiarité inconscient ? Un cours de linguistique contrastive appliquée à un phénomène langagier particulier peut-il faciliter la mémorisation, la compréhension et la manipulation des éléments verbaux en langue étrangère correspondant à ce phénomène langagier ? Soulignons la différence que nous maintenons entre « comparatif » (qui compare) et « contrastif » (qui contraste, c'est-à-dire qui fait *saillir* les zones d'écart et de recouvrement des deux items ou systèmes en présence). Un enseignement explicite déclaratif pourrait-il ainsi produire, sinon un apprentissage implicite, un effet de familiarité bénéfique à un apprentissage procédural ? Comme le note Hoyt (2001, p. 1) :



*« At the early level of second language acquisition teachers and students need to pay direct attention to phonetics and phonology. One big step that introductory language students must make is to simply realize that there is a pronunciation difference between the L1 and the L2 ».*

Si l'on se range du côté de ceux pour qui l'évaluation subjective par l'apprenant de l'écart entre sa production et le modèle à atteindre constitue une étape indispensable, la perspective contrastive nous semble tout à fait indiquée, favorisant ainsi ce qu'on appellera la « prise de conscience » de cet écart (sans laquelle l'apprenant ne peut pas même tenter de réduire cet écart). C'est davantage cette « prise de conscience », plutôt que la mise en place d'un système de contrôle sur ses propres productions ou bien de tentative de procéduralisation des savoirs qui nous préoccupe ici, les trois aspects étant bien évidemment intimement liés.

Une approche, en partie au moins, métalinguistique semble d'ailleurs plutôt d'actualité, puisque Fraser (2000), dans la présentation de son CD-ROM d'enseignement de l'anglais langue seconde, écrit :

*« It is not only possible but essential to develop a paradigm for pronunciation teaching which extends the general communicative approach by: 1) encouraging the learner to think of their own pronunciation as communication; and 2) providing more effective ways for teachers to communicate metalinguistically with learners about pronunciation. [...] It is essential that teachers are able to communicate effectively with learners about aspects of pronunciation, in order to point out errors and suggest improvements. This does not mean that teachers need to have an in-depth knowledge of English phonetics and phonology, though certainly the deeper their knowledge the better. Indeed, the danger of technical knowledge of this kind, is that it is tempting to impart it directly to learners. This is not useful because it is not easily comprehensible by learners ».*

Une telle préoccupation n'est d'ailleurs pas totalement absente des manuels d'enseignement du FLE. Ainsi, dans le guide pédagogique du manuel *Initial 1* (p. 5, leçon 0)<sup>279</sup>, on trouve :

*« L'enseignant peut, dans la langue maternelle des apprenants, engager une discussion sur la manière dont ils perçoivent les sons du français, l'intonation et le rythme des phrases, en comparaison avec leur langue maternelle et avec les langues étrangères qu'éventuellement ils connaissent ».*

Une telle recommandation implique donc : a) l'utilisation de la L1, b) une comparaison éventuelle avec d'autres langues, et c) une approche explicite, métalinguistique, du système phonético-phonologique du français et des langues en présence. Cependant, une recommandation si concise, ne permet pas de pallier les insuffisances, à ce sujet, des contenus des « méthodes », insuffisances parfois perçues par les enseignants<sup>280</sup>. Fraser (2000), quant à elle, insiste sur l'utilisation d'une terminologie appropriée (p. 4) :

*« Better information about which descriptive terms are most appropriate to learners require proper empirical research. [...] Testing the appropriacy of descriptors for learners is a very different thing from testing their accuracy in an abstract linguistic sense. I hope it is obvious that this is not to advocate inaccuracy in descriptive terms ! ».*

---

<sup>279</sup> Poisson-Quinton, S. et Sala, M, *Initial 1*, Paris, CLE international, 2000.

<sup>280</sup> Comme l'illustre, dans le cas du Japon, la critique de Richardot (1988, p. 26), qui concerne pourtant des « manuels » destinés à un public particulier :

*« Généralement les manuels de français utilisés au Japon ne tiennent aucun compte des ressemblances et des différences entre le français et le japonais, des difficultés qu'ont les Japonais à percevoir et à prononcer le français ; ils ne proposent aucun exercice sur la syllabation et sur les syllabes, pourtant véritables unités de base de la prononciation. Ils mentionnent simplement dans leurs premières pages la liste des phonèmes utilisés en français ».*

Avec ce dernier point, Fraser touche la question la plus délicate de notre problématique. Si l'on admet l'idée qu'un traitement conscient explicite peut faciliter l'apprentissage de la langue, de manière générale, la question cruciale qui se pose est la suivante : quel enseignement (formats et contenus) serait le plus à même d'atteindre les objectifs fixés ? Dans le domaine de la prononciation, les méthodes articulatoires ont souvent été montrées du doigt : on leur reprochait notamment de développer les « savoirs sur » mais pas les « savoir-faire ». Comme souvent, il y a là un problème d'extrémisme, et notre propos n'est pas de tenter de remettre le passé au goût du jour. Au contraire, il s'agit pour nous de viser un certain rééquilibrage. Vouloir prendre en compte l'activité métalinguistique attestée des apprenants de langue vis-à-vis du domaine de la prononciation, en offrant dans l'enseignement des activités descriptives, explicatives ou de conceptualisation relatives à ce domaine, implique, pour l'enseignant, de faire appel dans un premier temps aux données des sciences concernées, à savoir les sciences du langage, avant d'en envisager une transposition didactique la plus appropriée à son public.

### **III.2. Modèles linguistiques et corpus**

Deux questions se posent à nous ici : 1) celle de la transposition entre modèles linguistiques et modèles didactiques, 2) celle de l'utilisation des corpus, au vu des récents développements de la linguistique de corpus d'une part, et des perspectives émergentistes dans le domaine de l'acquisition des langues secondes d'autre part.

#### **III.2.1. Modèles linguistiques et modèles didactiques**

Si l'on considère le rapport entre connaissances « formelles », déclaratives, en phonétique et/ou phonologie d'un côté, et la prononciation des apprenants de l'autre, notre questionnement initial pourrait se formuler ainsi : 1) l'étude de la phonétique et la phonologie du français faciliterait-elle l'apprentissage des gestes articulatoires et du système phonologique de la langue en question ? 2) Si oui, cette phonétique et cette phonologie doivent-elles être plutôt françaises, générales ou contrastives entre la L1 et la L2 (bien que les trois soient évidemment fortement interreliées) ?

On pourrait par exemple noter ce que proposent, assez abruptement, Radwanska-Williams et Yam (2001) suite à leur étude sur l'acquisition des plosives, essayant, même de manière très succincte, d'établir un lien entre recherche fondamentale en interphonologie et application didactique (p. 38). C'est précisément sur l'aspect métalinguistique qu'elles insistent, par exemple en proposant de rendre accessible aux apprenants les connaissances métalinguistiques relatives aux phénomènes d'épenthèse et de délétion, sans bien sûr se perdre dans des aspects techniques, mais pour leur permettre d'en prendre (plus facilement) conscience. Elles citent ainsi Swain<sup>281</sup> et son *Hypothèse de l'Output*, dans laquelle trois niveaux de conscience sont distingués (*noticing, hypothesis testing, metalinguistic awareness*), lesquels sont tous bénéfiques au développement de la précision/exactitude linguistique dans la production en langue étrangère.

Cependant, si l'enseignant décide de faire appel aux domaines académiques concernés, il va être confronté à un problème de formation d'une part, de choix d'autre part, puisque les représentations métalinguistiques sont loin d'être uniformes, et enfin, sinon surtout, d'adéquation entre ces représentations et celles, intuitives ou apprises en L1, des apprenants.

Coleman (2002) en offre une illustration assez frappante, à propos de ce que nous pourrions appeler ici la « règle de prononciation »<sup>282</sup> des allomorphes de la marque du génitif en anglais (pp. 115-116). L'auteur propose en effet quatre formulations verbales différentes assez techniques pour décrire une même « règle » : une version phonémique, une version correspondant à une approche par traits, une version articulatoire et enfin

---

<sup>281</sup> Swain, M., (1995). « Three functions of output in second language learning ». In G. Cook and B. Seidlhofer (eds), *Principles and Practice in Applied Linguistics*. Oxford and New-York : Oxford university Press.

<sup>282</sup> Même si l'auteur, reprenant l'argument de Halle en faveur d'une approche en traits et non plus phonémique, indique justement que la version phonémique (1a)) ne constitue pas une bonne règle de prononciation pour les emprunts se terminant par des phonèmes non-voisés autres que ceux indiqués dans la liste (comme le mot allemand Bach, qui se termine par une fricative vélaire non-voisée [x]). Cette formulation « does not tell us how to pronounce the genitive « Bach's » (p. 115).

une version acoustique. La question est alors : doit-on présenter une de ces formulations aux apprenants, et si oui *laquelle* ?

<b>Phonologie</b>	<b>Version phonémique</b> (« phonemic version »)	<i>« If the last phoneme of the word is a member of the set {/p/, /t/, /k/, /tʃ/, /f/, /θ/}, the genitive allomorph is /s/ ».</i>
	<b>Version en traits</b> (« featural version »)	<i>« If the last segment of the word bears the features [+consonantal, –vocalic, –sonorant, –voice] (and if also [+continuant, +coronal] then [–strident]), the genitive allomorph is [+consonantal, –vocalic, –sonorant, +continuant, +coronal, +strident, –voice] ».</i>
<b>Phonétique</b>	<b>Version articuloire</b> (« articulatory version »)	<i>« If at the end of the word there is an obstruction within the vocal tract, but the vocal cords are not vibrating (and if furthermore there is a critical constriction between the tongue tip and the alveolar ridge, but the teeth are covered so that any turbulence is generated along the walls of the vocal tract rather than at the teeth), the genitive is formed by a critical constriction between the tongue tip and the alveolar ridge with the teeth uncovered so that turbulence is caused at the teeth, the vocal cords continuing not to vibrate ».</i>

	<p style="text-align: center;"><b>Version acoustique</b>          (« <b>acoustic version</b> »)</p>	<p><i>« If at the end of the word there is an interval of relatively low amplitude, either silence or with aperiodic excitation, and if in the latter case higher frequencies predominate in the spectrum and the lower limit of the frequency distribution is relatively high, but with comparatively low amplitude relative to other aperiodic intervals and a stronger resonant structure), the genitive is formed by aperiodic excitation with higher frequencies predominating in the spectrum, the lower limit of the frequency distribution being relatively high, comparatively high amplitude and a weak or indiscernible resonant structure ».</i></p>
--	---	--

Comme on le voit clairement ici, les manières et les angles d'analyse pour décrire une même réalité sont multiples, et si l'on décide de privilégier l'un ou l'autre, il faut le justifier. L'article de Coleman, entre autres, indique aussi que ces choix sont en grande partie déterminés par les options théoriques (dans le domaine de la phonologie et dans celui de la perception de la parole en particulier) auxquelles l'on se rattache, ce qui nous fait repasser du domaine didactique à celui de la linguistique et de la psycholinguistique.

Dans notre cas, nous pensons que les justifications doivent être didactiques, et viser ainsi la plus grande compréhensibilité en vue d'atteindre nos objectifs d'enseignement/apprentissage. Répondre à la question « Quel enseignement adéquat ? » constitue tout un programme de recherche, et nous n'avancerons donc pas de réponse ici, même si nous engageons la réflexion. Notons simplement qu'il n'existe peut-être pas de réponse générale—une quelconque panacée—mais plutôt des consignes méthodologiques en

vue d'obtenir des réponses spécifiques à chaque situation didactique (langue première des apprenants, contexte institutionnel, profils et savoirs des apprenants, passé métalinguistique des apprenants etc.).

Il semble également que l'évolution des sciences du langage ne doit pas être laissée de côté par le corps didactique, au vu des implications que celle-ci peut avoir sur la pratique pédagogique d'une part, mais également parce qu'elle permet de mieux appréhender les données issues de la psycholinguistique touchant de très près l'apprentissage des langues (par exemple l'importance de la syllabe en français en perception de la parole, qui relativise, en partie au moins, celle du phonème). Une fois encore, l'examen des représentations métalinguistiques des enseignants et de didacticiens constitue une étape préalable à toute tentative d'enseignement de la prononciation de leur part, car il se pourrait que ces représentations soient insuffisantes, ou insuffisamment informées, pour traiter de manière optimale le sujet.

A titre d'exemple, on peut ainsi s'interroger sur la persistance (nous serions tentés de dire la « perpétuation ») de présentations grammaticales traditionnelles, pourtant remises en cause par différentes études, qui ont en outre le mérite de relier certains plans d'analyse linguistique, effectivement reliés dans la compétence langagière des locuteurs. Notre exemple concerne l'interface entre morphologie et phonologie, dans le cas des « conjugaisons verbales » (inflexions verbales pour être précis) *orales* et de l'alternance Consonne/Morphème nul (C/Ø) que l'on trouve au présent de l'indicatif (par exemple : dort/dorment ; finit/finissent) mais aussi dans d'autres contextes morphologiques (gand/grande (genre); gros/grossir (dérivation)) et syntactiques (petit/petit amour (liaison)). Il s'agit d'une étude menée par Paradis et El Fenne (1995) dans le cadre de la théorie des contraintes et des stratégies de réparation (TCRS), qui conduit à une remise en question des catégories verbales (1<sup>er</sup>, 2<sup>nd</sup> et 3<sup>ème</sup> groupes) que l'on trouve par exemple dans des ouvrages comme le Bescherelle de 1985 (Paradis et El Fenne, p. 198). Comme l'écrivent les auteurs, ce type de classification ne permet pas

de rendre compte de l'alternance C/∅ qui doit être considérée d'un point de vue phonologique (p. 198) :

*« This type of classification, which aims essentially at establishing "morphological" links between verbal stems and their inflexional suffixes, cannot capture the phonologically induced behavior of the C/∅ alternation in verbs ».*

L'étude est certes trop technique pour le didacticien, mais les conclusions auxquelles ils arrivent sont assez claires : excepté certains cas (avec racines terminées par voyelles du type « voir, créer, distraire, conclure, etc. ») qui ne représentent que 4, 3% des verbes français, les racines verbales sont divisées en deux grandes catégories :

- Celles qui se terminent par une Consonne Flottante (« a consonant without a timing slot » (p. 201)) (par exemple : part/partent),
- Celles qui se terminent par une Consonne Permanente (par exemple : aime/aiment) et qui constituent 94% des verbes français.

La valeur prédictive de ces hypothèses a été vérifiée dans un test de production par El Fenne (p. 200) dans lequel on présentait, par exemple, aux participants le non-verbe « \*vanitre » ainsi que ses 3èmes formes : [vani/vanis]. On demandait ensuite aux sujets de restituer les formes du présent de l'indicatif. L'argument selon lequel un processus d'analogie est mis en œuvre est rejeté en note de bas de page (p. 201), parce qu'il impliquerait que le référent soit le système de consonnes flottantes lui-même, et donc un raisonnement circulaire. Bien sûr, les travaux et les résultats de ces auteurs sont discutables, mais ils illustrent bien l'écart séparant, d'une part le contenu des enseignements traditionnels (grammaticaux) et d'autre part celui des recherches contemporaines en linguistiques.

Cet exemple est particulièrement pertinent pour nous en ce qu'il renvoie, d'une part à l'interface entre phonologie et morphologie (soulignant ainsi l'importance de la



« prononciation » dans l'apprentissage général de la langue), et d'autre part à la distinction entre morphologie orale et écrite du français, distinction traitée par Battye, Hintze et Rowlett (2000) dans leur introduction linguistique au français<sup>283</sup>. Le cas de la morphologie verbale en français représente un bon exemple des implications futures de notre travail : celle-ci constitue en effet l'une des difficultés majeures de la langue pour de nombreux étudiants, notamment les apprenants japonais, et le rapport entre code oral et écrit dans l'apprentissage de *l'oral* devrait, nous semble-t-il, être un sujet de recherche didactique non-négligeable. Nous sommes en effet malheureusement habitués à entendre nos étudiants prononcer les <-r> graphiques des infinitifs des verbes dits « du premier groupe » (entre autres)<sup>284</sup>, ou encore poser des questions témoignant de leur incompréhension des rapports entre morphologie, phonologie et orthographe (« Pourquoi prononce-t-on le <-t> et pas le <-s-> dans "il est allé" ? » ; « Pourquoi prononce-t-on le <-t> dans "il est allé" et pas dans "et aussi" ? », etc.).

L'ouvrage de Battye, Hintze et Rowlett (2000), rédigé en anglais et destiné à un public d'anglophones, constitue une bonne tentative de présentation métalinguistique du français pour des apprenants de français, comme l'indiquent son titre *The French language today : a linguistic introduction*. La table des matières (simplifiée) se présente ainsi : 1) The external history of the French language : the standardisation of French and its distribution in the world today, 2) The sound system of French, 3) French word structure, 4) The sentence structure of French, 5) Varieties of French.

Si l'on s'arrête sur la partie phonético-phonologique de l'ouvrage, l'approche adoptée est plutôt traditionnelle (descriptions phonétiques, approche phonématique), mais les auteurs couvrent un large éventail de phénomènes spécifiques au français, et ce faisant

---

<sup>283</sup> Notons à ce propos la remarque de Coulmas (1989, p. 269) : « Descriptive linguistics is pervaded by notions originating in normative grammar. And normative grammar is based on written language ».

<sup>284</sup> Ce qui évoque clairement « l'effet Buben » observé en diachronie en français, comme le rapporte Catach (1995, p. 67) : « Notre langue, qui ne prononçait il y a deux siècles pratiquement aucun groupe de consonnes, si ce n'est avec les liquides (r et l), se surcharge actuellement, sous les actions combinées des emprunts, de l'école, de l'orthographe, etc., de nombreux groupes internes et finaux (cheptel, dompter, sangsue, etc.) ».

établissent des liens entre phonétique (« prononciation ») et morpho-syntaxe (« tu l'apprends vs. tu la prends » ; « je mourrai (futur) vs. je mourais (imparfait) »). De la même manière, après avoir été indiqué dans la partie sur le système phonique, l'écart entre codes écrit et oral en français est mentionné une seconde fois dans le chapitre sur la morphologie, afin de bien distinguer orthographe et prononciation (p. 120), mettant ainsi en relief les étroites relations existant entre les différents niveaux d'analyse linguistique, étroitesse parfois négligée dans les descriptions grammaticales traditionnelles. Suivant certains travaux d'André Martinet, les auteurs proposent ainsi une présentation simplifiée de la morphologie verbale orale, conduisant à une classification en neuf groupes des verbes français (pp. 139-146).

La partie sur le système phonique se présente ainsi :

- La norme de référence
- Les correspondances graphie-phonie
- La syllabification, l'allophonie, la transcription phonétique,
- La phonétique articulatoire
- Les classements consonantique et vocalique
- La gémination consonantique
- L'assimilation
- Les consonnes finales (tout/tous)
- Le « e muet »
- Des questions de distribution et d'alternances (eu/o fermés et ouverts)
- Les voyelles nasales (pardon/pardonne)
- Les semi-voyelles (pays/payé)
- Liaison, élisions, enchaînement
- Aspects suprasegmentaux : accentuation et intonation

Du point de vue de l'apprentissage, un tel ouvrage présente plusieurs problèmes. On peut mentionner, entre autres, l'absence d'activités d'apprentissage (il s'agit d'une simple présentation), la limitation aux domaines traditionnels (phonologie,

morphologie, syntaxe, même s'il y a une perspective diachronique et synchronique sur la variabilité) et l'absence de perspective linguistique générale qui permettrait peut-être de faciliter la comparaison avec d'autres langues, en particulier la langue première. A toutes ces critiques, il sera facile de répondre que les objectifs des auteurs et de l'ouvrage sont différents, et cette réponse sera parfaitement justifiée. En réalité, le vrai problème est sans doute celui qui figure à la fin de l'ouvrage, à savoir le glossaire, qui renvoie au thème du métalangage. La question est la suivante : *doit-on apprendre une métalangue pour apprendre une langue ou bien existe-t-il un moyen d'en faire l'économie ?* La question est délicate et surtout elle varie en fonction du niveau, du public et des situations didactiques et nous ne prétendons donc pas y apporter de réponse ici.

Il faut par ailleurs noter à ce sujet que l'on constate bien souvent un certain flottement terminologique dans les manuels de langue, lorsqu'il s'agit de l'intitulé de la section qui traite des aspects qui nous intéressent : parfois baptisée section « Phonétique », parfois « Prononciation ». Puisqu'il s'agit de manuels d'apprentissage de la langue, on peut s'étonner d'y trouver une rubrique « Phonétique », ce terme renvoyant a priori davantage à une discipline des sciences humaines qu'à un aspect de la compétence à acquérir en langue étrangère. On pourrait cependant considérer, comme nous l'avons suggéré jusqu'ici, que l'étude de la phonologie et de la phonétique du français est nécessaire à une bonne acquisition de la morphologie du français, *via* celle d'une bonne prononciation, comme cela nous est apparu à travers les travaux de Paradis et El Fenne, ou bien l'ouvrage de Battye, Hintze et Rowlett. Ces travaux nous montrent aussi à quel point l'approche grammaticale traditionnelle tend à séparer différents niveaux d'analyse qui sont pourtant étroitement liés dans le système linguistique.

Dans le département des études anglophones de l'université de Toulouse-Le Mirail par exemple, les étudiants de première et de deuxième année doivent suivre des enseignements obligatoires en phonétique-phonologie de l'anglais, enseignement, disons théorique, couplé à des travaux pratiques en classe et en laboratoire en deuxième année, avec notamment des exercices de transcription phonétique (Durand,

2000b). Ce module, pour les étudiants qui prennent la peine de fournir le travail demandé, semble tout à fait profitable aux étudiants, en ce qui concerne leur prononciation, mais aussi leur compétence orale plus générale (marques du pluriel et du passé à l'oral, en compréhension et en production, par exemple). Nous pourrions donc envisager un module similaire pour l'enseignement du français aux étudiants japonais. A nouveau, nous devons alors nous interroger sur l'approche qu'il convient d'adopter : faut-il enseigner une phonétique française, une phonétique générale ou une phonétique contrastive franco-japonaise à nos apprenants ? Le choix incombe ici encore aux responsables des curricula, lesquels, pour ce faire, doivent être suffisamment informés des enjeux didactiques et institutionnels.

Ainsi, si l'évolution des sciences du langage présente de nouvelles données et de nouvelles approches qui pourraient potentiellement être utiles au corps didactique, a *contrario*, l'évolution des connaissances en didactique permet aujourd'hui de mieux sélectionner les données pertinentes de la linguistique, et surtout d'en faire un usage approprié, qui n'apparaîtrait pas forcément au linguistique, comme cela est illustré, selon nous, par les propos de Grunig dans un ouvrage de 1974, qui touche de très près notre propos puisqu'il s'intitule : *De la théorie linguistique à l'enseignement de la langue*. Voici ce qu'écrit Grunig dans cet ouvrage, à propos de la langue étrangère (p. 212) :

« [Pour que l'élève] connaisse cette langue, il faut qu'on la lui ait présentée. Aussi dirons-nous d'une grammaire d'enseignement qu'elle doit être la présentation d'une langue [...] dans sa totalité et non la présentation d'une partie seulement de cette langue : aucun des éléments de la description linguistique ne devrait en être exclu ».

Cette affirmation de Grunig pourra aujourd'hui certainement susciter le doute parmi les didacticiens, étant donné notamment ce que l'on sait des interlangues, d'une part et du rôle de l'action dans l'apprentissage d'autre part. En ce qui concerne la prononciation, voici ce qu'elle indique, à propos d'élèves français apprenant l'allemand (p. 212) :

*« Progrès facile : la théorie de la neutralisation [phonologique] est suffisamment élaborée et rigoureuse pour que l'élève l'admette, l'effort de mémorisation est faible (la liste des [consonnes] sourdes est faible en particulier la prononciation de mots comme « Krug », prononcé [kru : k]) pour un grand rendement ».*

Ici encore, une telle « approche », sans être « erronée », reste cependant assez limitée d'un point de vue didactique, privilégiant apparemment la « mémorisation » d'une liste de consonnes à la réflexion métalinguistique ou à la pratique. Nous partageons donc l'opinion de Cuq (1996, p. 26) lorsqu'il écrit que, pour le didacticien :

*« [Il s'agit moins] d'importer du produit linguistique, au prix de faire d'un savoir-savant un savoir pédagogique (vulgarisé, mais surtout décontextualisé), que de revitaliser les concepts linguistiques jugés intéressants en les plaçant dans le contexte de l'enseignement/apprentissage ».*

Si le didacticien doit suivre les enseignements des linguistes pour comprendre le fonctionnement « authentique » du code linguistique, et donc alimenter sa compétence métalinguistique en L1 et en L2 en tant que spécialiste en la matière, il peut en revanche s'en détacher pour remplir sa fonction de « facilitateur d'apprentissage ». Comme le rappelle Cuq (1996, p. 36) :

*« L'objet linguistique n'est pas une réalité donnée mais construite. Ce n'est pas en effet sur l'expérience commune du langage que le didacticien peut s'appuyer, ni même le plus souvent sur les modèles fournis par la tradition scolaire, mais sur les modélisations orientées qu'il en propose. C'est dans l'orientation du modèle vers les problèmes posés par l'enseignement/apprentissage que se trouve la divergence méthodologique fondamentale avec le linguiste « pur » qui n'a pas ce souci ».*

Gardons aussi à l'esprit qu'il ne s'agit pas en premier lieu pour le didacticien d'*expliquer*, mais de *faire apprendre*, une distinction clairement illustrée, nous semble-t-il, dans le domaine grammatical, par la classification de Porquier (2000, pp. 70-71) :

*«La diversité des champs désormais couverts ou évoqués par le terme de grammaire, par ses dérivés (grammatical, grammaticalisation) et ses collocations (grammaire descriptive, grammaire pédagogique, grammaire scolaire, grammaire d'apprenant, [ainsi que grammaire générative, grammaire fonctionnelle, grammaire textuelle, grammaire du discours, grammaire du locuteur], etc.), peut être globalement triée selon quelques acceptions principales :*

- 1. Le fonctionnement interne d'une langue (système, règles internes) ;*
- 2. L'explicitation et la (une, des) description(s) de ce fonctionnement ;*
- 3. Un corps de savoir à enseigner/acquérir ;*
- 4. Des matériaux (produits et outils) relevant de ce corps de savoir et servant à l'enseigner/acquérir ;*
- 5. La compétence langagière d'un individu ou d'un ensemble d'individus ».*

La question des représentations « externes » employées par l'enseignant va donc s'avérer aussi, sinon plus, importante que celles des représentations « profondes », puisque le mode et le format de présentation joueront un rôle capital quant à *l'utilité* et *l'efficacité*, pour les apprenants, des modèles et représentations externes employés.

Le problème de la transposition didactique des modèles linguistiques (dans le cas de la grammaire, *via* les descriptions des grammairiens vers les grammaires d'enseignement et d'apprentissage) n'est pas nouveau. Mais la remarque de Cuq nous semble cependant ouvrir la voie vers un dépassement de ce problème : le didacticien doit orienter son modèle en fonction des contraintes qui pèsent sur lui dans le cadre de situations d'enseignement/apprentissage pour un public déterminé. Il possède donc à la fois plus et moins de degrés de liberté que son collègue, le « linguiste pur » :

- Plus de liberté, car ses modèles ne s'articulent pas autour des mêmes critères : ils ne servent pas les mêmes buts et ils ne devraient être ni rigides, ni fossilisés au cours de leur propre élaboration ; ils devraient pouvoir être adaptables aux situations rencontrées.
- Moins de liberté, car des contraintes supplémentaires, nombreuses et variées, propres à la situation didactique s'imposent à lui.

Il ne s'agit pas pour le didacticien de proposer un modèle causal à visée heuristique (décrire pour décrire, décrire pour expliquer), mais de proposer un modèle didactique utile (outil) (décrire pour faire apprendre et pour enseigner un savoir-faire). Les finalités sont différentes ; il nous semble donc que le didacticien devrait se suffire de ses contraintes, sans avoir en outre à subir celles du « linguiste pur » qui l'entravent plus qu'autre chose. Si ses objectifs requièrent l'orientation et l'adaptation de ses modèles, alors pourquoi le didacticien devrait-il se conformer aux cadres strictement normés des linguistes « pure laine » ? Comme l'écrivent Porquier et Vivès (1993, p. 68) :

*« Les règles de fonctionnement que dégagent et décrivent les linguistes, avec leurs propres outils, ne sont qu'une tentative d'explicitation et de rationaliser un mode de fonctionnement qu'ignorent pour l'essentiel, sous cette forme, tous les usagers de la langue, pour le connaître autrement ».*

Dans cette perspective, le didacticien ne devrait donc pas se laisser imposer ses modes de représentations langagières, mais plutôt élaborer ses propres modèles en fonction de ce à quoi ils sont destinés. Nous allons ainsi dans le même sens que Besse et Porquier (1991), avec leurs « exercices de conceptualisation », destinés à favoriser la conceptualisation grammaticales *par les apprenants* à partir de *leurs observations*, puisque comme l'indique Vivès (1988, p. 32) :

*« Ce qui est visé n'est pas un savoir linguistique mesurable en termes de connaissances quantifiables, mais l'acquisition d'une compétence d'utilisation de l'ensemble des connaissances linguistiques ou une prise de conscience personnelle de principes d'organisation sous-jacents ».*

Nous retrouvons ici cette notion de « prise de conscience », même si, comme l'écrit Trévisse (1993, p. 43) :

*« Ce n'est pas parce qu'il y a conscience qu'il y a travail "à bon escient". Cependant, on peut penser que les prises de conscience spontanées d'un décalage (Klein, 1986) entre l'interlangue du sujet et ce qu'il perçoit de langue 2, qu'elles soient provoquées par un sentiment d'incapacité à transmettre un contenu de pensée, ou par une attention au code à un moment donné, constituent un des facteurs qui peuvent déclencher le processus d'acquisition sur un point du système et le rééquilibrer à terme ».*

Ainsi, si l'on peut considérer qu'un enseignement favorisant le développement de l'activité métalinguistique des apprenants grâce à une réflexion métalinguistique organisée en fonction des objectifs d'apprentissage est recommandé, il est clair qu'il ne s'agit ni de présenter des extraits de théories linguistiques hermétiques, ni de retomber dans un enseignement magistral « traditionnel » peu profitable au développement des capacités communicatives des étudiants.

On pourrait rétorquer que le bénéfice d'un tel enseignement dépendra des styles d'apprentissage des apprenants (en particulier l'opposition grossière<sup>285</sup> *analytique* vs. *global*) et que certains étudiants manient mal l'abstraction (d'ailleurs, en France, de nombreux lycéens se tournent vers les langues vivantes par manque de goût ou d'aisance en mathématiques ou en physique par exemple). Sans nier ce point, il semble néanmoins que si le style d'apprentissage constitue une variable difficile à maîtriser,

---

<sup>285</sup> Pensons à la variation intra-individuelle et inter-tâches.



les oppositions entre savoirs déclaratifs et savoir-faire procéduraux d'une part, et entre instruction formelle et informelle d'autre part, restent des constantes bien établies. Par ailleurs, une telle critique pourrait être applicable à la plupart, sinon tous, les modes d'enseignement institutionnels, disons standards. De plus, il ne s'agit en aucun cas de vouloir verser dans l'abstraction et la complexité, loin s'en faut. Il s'agit plutôt de donner un espace, un temps et un encadrement didactique à l'activité métalinguistique des étudiants, cadre, qui peut, voire doit, être bâti dans le domaine des sciences du langage. Et si de nombreux lycéens apprécient l'interactionnisme vivant et communicatif de la majorité des cours de langue vivantes—qu'il convient de ne surtout pas abandonner !—d'autres lycéens attribuent leur échec en langue au manque d'explication et d'analyse...

On pourrait aussi reprocher au type d'enseignement que nous évoquons la surcharge potentielle de métalangage. Tout en souscrivant pleinement à la critique, et sans reprendre la discussion menée plus haut, s'il convient effectivement de minimiser autant que se peut tout excès de métalangage, il suffit cependant d'ouvrir la plupart des manuels de FLE actuellement produits en France<sup>286</sup> pour constater, non seulement l'abondance de métalangage, mais surtout l'absence de prise de position réelle face à ce problème, qui se traduit notamment par le recours à des présentations grammaticales bien trop familières. Ne faudrait-il alors justement pas un module spécifiquement destiné à aborder de front les aspects métalinguistiques de l'apprentissage, précisément pour en libérer les autres modules ? Dans le cas d'étudiants japonais apprenant le français par exemple, on pourrait également affirmer que ces apprenants ont longtemps été considérés comme peu brillants dans le domaine des langues vivantes, précisément parce qu'ils favorisaient les « savoirs sur la langue » aux « savoir-faire » avec la langue, et qu'il conviendrait donc de ne pas poursuivre ou retomber dans cette voie. Sans vouloir nous arrêter sur les causes de cette image quelque peu stéréotypée, ni sur la

---

<sup>286</sup> Par exemple *Studio 100*, p.16 (Lavenne, C., Bérard, E., Breton, G., Canier, Y. et Tagliante, C., Paris, Les Editions Didier, 2001) dont les auteurs sont pourtant fortement engagés dans la mouvance communicativiste, ce dont témoigne le manuel lui-même dans son ensemble. Si l'on regarde le premier encart de grammaire (« pronoms, articles, adjectifs »), on peut tout de suite s'interroger non seulement sur le métalangage utilisé, mais aussi sur les fonctions mêmes d'un tel encart.

réalité à laquelle elle peut renvoyer, nous répondrons : certes, non, il ne s'agit pas de renforcer cette tendance, mais au contraire de la transformer. Transformer la présentation et le contenu des « savoirs sur », mais aussi transformer leur mise en œuvre dans des activités nécessitant des « savoir-faire ».

Cette dernière critique soulève en effet le point le plus essentiel : celui des contenus d'un tel enseignement, mais aussi et surtout, celui de l'intégration, voire du couplage, d'un tel module avec des modules à orientation beaucoup plus pratique et/ou communicative. Car loin de défendre une approche formelle et disons « intellectuelle » de l'enseignement des langues, approche bien évidemment depuis longtemps obsolète, mais tout en étant conscient des faiblesses d'une approche « informelle » de l'enseignement/ apprentissage des langues, il serait souhaitable d'en développer un pendant nécessairement complémentaire, mais cependant indispensable, celui qui viserait à donner les moyens à l'apprenant de développer son activité métalinguistique de manière efficace, précisément en vue de se libérer de certains doutes et hésitations, pour s'investir ensuite pleinement dans d'autres activités. Cela passe par un important travail d'analyse, de programmation et de préparation en amont de l'élaboration des curricula. En particulier, l'élaboration d'un module contrastif entre la L1 et la L2 de sensibilisation métalinguistique pour apprenants débutants et de développement métalinguistique pour apprenants intermédiaires ou avancés nécessiterait un travail précis et important, impliquant une collaboration et une coordination étroite entre didacticiens, linguistes, enseignants et coordinateur du programme.

Si l'informel et le communicatif (comme d'autres avant, en particulier dans la problématique SGAV) ont supplanté un enseignement trop peu axé sur la communication réelle, il convient à présent de rééquilibrer la balance dans un monde qui change, et avec lui les étudiants. Adapter le mode d'enseignement aux apprenants, tel est ce qui nous préoccupe (Detey, 2003), mais une telle adaptation ne peut porter ses fruits qu'à condition d'être intégrée dans une perspective méthodologique cohérente et forte (Rivenc, 2003), qui prenne en compte l'ensemble de la formation linguistique.

Cependant, avant d'aborder les perspectives méthodologiques globales qui ressortent de notre étude, il convient au préalable de traiter deux derniers points concernant le rapport entre linguistique et didactique dans le thème qui nous occupe : il s'agit de la question des corpus d'une part, et de la norme d'autre part.

### **III.2.2. Corpus oraux**

La référence aux corpus oraux s'impose ici pour trois raisons au moins : d'abord, 1) d'un point de vue didactique, comme par exemple dans les « exercices de conceptualisation » (Besse et Porquier, 1991), l'utilisation des corpus en classe de langue n'est pas nouvelle, puisqu'elle est à la base des démarches inductives (essentiellement grammaticales, mais pas seulement)<sup>287</sup> ; ensuite 2) d'un point de vue globalement psycholinguistique, le développement des perspectives exemplariste d'une part, et émergentiste d'autre part, en souligne l'importance potentielle dans l'apprentissage d'une langue étrangère, notamment du point de vue de l'encodage mnésique des éléments linguistiques ; enfin 3) d'un point de vue linguistique, le développement des études de corpus oraux, en relation avec la sociolinguistique, font que les sciences du langage offrent de nouveaux outils pour la didactique, d'une certaine manière plus « concrets », estimeront certains, que les modèles et théories formelles.

A l'entrée « corpus » du *Dictionnaire de Didactique du Français Langue Etrangère et Seconde* (Cuq, 2003, p. 57), voici ce que l'on peut trouver :

*« On appelle corpus un ensemble de données collectées par enregistrement, par observation directe, par questionnaire ou entretien, et réunies pour décrire et analyser un phénomène. En didactique des langues on peut ainsi étudier le processus d'acquisition en se fondant sur les performances enregistrées*

---

<sup>287</sup> Dans le domaine de l'Apprentissage des Langues Assisté par Ordinateurs (ALAO, en anglais CALL) en particulier, l'utilisation de corpus fournit souvent la base linguistique des didacticiens de « concordancing », qui sont essentiellement liés aux problèmes de collocations lexicales et qui peuvent être utilisés dans des démarches de découverte grammaticale inductive.

*d'apprenants ou cerner les représentations et les attitudes des apprenants et des enseignants en analysant les réponses recueillies par des questionnaires ».*

Or, comme l'indique J. Véronis en 2004 dans un appel à communication pour un numéro spécial de la revue *TAL*, consacrée au « Traitement automatique des corpus oraux »<sup>288</sup> :

*« Alors qu'on dispose de plusieurs centaines de millions de mots de textes écrits [...], on dispose de très peu de données sur l'oral. Les corpus de langue oral spontanée sont pourtant d'une importance fondamentale pour l'étude linguistique, comme pour la mise au point de nouvelles technologies vocales ».*

Cette remarque, apparemment sans grande importance, nous paraît au contraire capitale pour notre problématique : la « domination de l'écrit » dans les classes de langue ne peut en effet s'expliquer sans la prise en compte des moyens et du matériel dont disposent les enseignants et les apprenants<sup>289</sup>.

De fait, comme nous l'avons vu, enseigner l'oral et enseigner l'écrit représentent finalement deux tâches assez distinctes, tant ces deux voies possèdent un code qui leur est propre, tant au niveau syntaxique, que discursif par exemple (Blanche-Benvéniste, 1997). Comme le note Durand (2000b, p. 55) :

---

<sup>288</sup> <http://www.up.univ-mrs.fr/~veronis/edito/index.html> (consulté en mars 2005).

<sup>289</sup> Ce qui renvoie au travail pionnier du *Français Fondamental*, dans la préface duquel est indiqué (1959) : « *Le français fondamental devant être aussi proche que possible des conditions réelles d'utilisation, on a pris pour base le français parlé. Des études de fréquences des mots avaient déjà été effectuées sur les langues écrites. C'est la première fois qu'on a fait une étude analogue sur la langue parlée. Cette innovation a été rendue possible par les perfectionnements du magnétophone* ».

Par ailleurs, il est également intéressant de relever, à ce propos, l'influence de ces moyens sur la représentation de la langue elle-même, comme le note Durand (2000b, p. 42) : « *Les développements scientifiques, technologiques et technocratiques favorisés par l'invention de l'imprimerie ont créé chez de nombreux locuteurs une identification entre la langue et sa matérialisation dans l'écriture* ». Cela renvoie pleinement, à grande échelle, à « l'effet Buben » qui n'a cessé de nous préoccuper jusqu'ici.

*« Dès que l'usage d'un système d'écriture se répand dans une société et constitue l'armature du système politique, juridique et social [...], l'écriture s'autonomise en grande partie à l'égard de la langue parlée. Elle n'est plus un système greffé en permanence sur l'oral mais un système parallèle avec ses règles propres de fonctionnement. Tout utilisateur d'un système graphique aussi ancien que celui du français sait qu'il y a un lexique et un phrasé qui sont typiques de l'écrit et qui, loin de refléter passivement l'usage oral, mettent très souvent ce dernier à distance ».*

L'écart entre les codes oral et écrit semble d'ailleurs, dans notre époque contemporaine, s'accroître de manière accélérée, chez les jeunes générations en particulier, dont « le langage » constitue une abondante source d'étude pour les sociolinguistes (Koulayan et Sayah, 2002).

La nécessité de développer la *disponibilité* et l'*accessibilité* du matériel oral pour l'enseignement des langues (sous la forme de corpus notamment) se fait donc sentir, et il serait temps que les *enregistrements de textes écrits oralisés*, que l'on trouve habituellement dans les méthodes, cèdent le pas à des *enregistrements de textes oraux transcrits*. Cela concerne tant les aspects phonético-phonologiques que les aspects pragmatico-énonciatifs, car bien que ces derniers aient été davantage pris en compte récemment, les dialogues « aseptisés » des méthodes (que l'on peut tout à fait justifier pour des apprenants débutants) constituent toujours, nous semble-t-il, la règle.

L'évolution technologique semble donc aujourd'hui ouvrir de nouvelles voies, en particulier en ce qui concerne le stockage de documents sonores. Si l'on ne peut enseigner l'écrit sans « écrit », il semble tout aussi évident que l'on ne peut enseigner l'oral sans « oral ». Or, alors qu'il est aujourd'hui possible d'accéder facilement à une très grande variété de documents écrits, il n'en est pas encore de même pour les documents oraux. Il est alors remarquable de constater que, à travers le monde, de nombreux projets de constitution de corpus oraux sont en cours, donnant lieu à des réalisations comme le « Corpus C-ORAL-ROM », présenté ainsi par J. Véronis sur le

site<sup>290</sup> de l'équipe DELIC (Description Linguistique Informatisée sur Corpus), qu'il dirige :

*« C-ORAL-ROM est un corpus multilingue de parole spontanée [...] composé de quatre collections d'enregistrements comparables de sessions de parole spontanée pour l'italien, le français, le portugais et l'espagnol (environ 300 000 mots par langue). [...] [Il] offre la source acoustique de chaque session ainsi que les annotations principales suivantes : a) la transcription orthographique [...], c) la synchronisation du texte par rapport à la parole [...] ».*

L'intérêt de tels corpus apparaît d'emblée si l'on prend en compte : 1) la valeur sociolinguistique de tels outils, qui renvoie à la question de la norme en classe de langue et que nous traitons plus bas, et 2) la perspective épisodique acoustique que nous avons présentée précédemment, notamment à travers les travaux de Hardison (1999), et qui insiste sur l'intérêt, pour l'apprentissage de la discrimination phonémique et plus généralement celui de la perception de la parole et de la « prononciation » en langue étrangère, d'un entraînement perceptif à haute variabilité, comme l'indiquent Iverson, Kuhl, Akahane-Yamada, Diesch, Tohkura, Kettermann et Siebert (2003, p. B54) :

*« Training procedures involving smaller stimulus sets are easier to learn, but do not readily generalize to new stimuli [...]. Training with larger stimulus sets may generalize better, because the variability provides information about which cues are more robust and trains individuals to ignore irrelevant variation ».*

Certes, disposer d'enregistrements ne résoud pas tous les problèmes. Cependant, lorsque ces corpus sont couplés à certaines technologies, on peut alors commencer à

---

<sup>290</sup> <http://delic-actu.blogspot.com/2005/01/corpus-c-oral-rom-4-langues-romanes.html> (consulté en mars 2005).

envisager de quelle manière, *en classe d'oral*, l'écrit cesse d'être *indispensable*, puisqu'il deviendrait possible pour l'apprenant de stocker (fonction mnésique) les données sonores<sup>291</sup> et de les visualiser, de les segmenter et de les réécouter plusieurs fois en en modifiant éventuellement certains paramètres acoustiques (fonction de réification) à conditions de disposer d'outils appropriés.

En classe de langue, l'une des fonctions clés de l'écrit a également été l'accès au dictionnaire. Or, à cet égard, on ne peut que saluer la diffusion des versions sonores des dictionnaires électroniques<sup>292</sup> : ceux-ci sont en effet munis d'écouteurs permettant aux étudiants d'entendre les mots énoncés oralement, ce qui est excellent pour l'apprentissage des formes sonores des lexies concernées (même si de nombreux dictionnaires traditionnels comportent une transcription phonémique en Alphabet Phonétique International). De nombreuses possibilités seraient à envisager pour le développement de ces outils dans un cadre didactique (contextualisation phrastique des items, possibilité de modifier le signal (« ralentissement »), modification du signal préprogrammé pour un certain public, etc.)<sup>293</sup>.

Cependant, avant d'envisager de telles extensions futures, les corpus oraux présentent dès à présent un intérêt certain pour la didactique, par delà les données linguistiques de référence qu'ils procurent aux linguistes et, à leur suite, aux didacticiens. La publication, en 2002 d'un ouvrage (papier et cédérom) intitulé « Choix de textes de français parlé, 36 extraits », édité par C. Blanche-Benveniste, C. Rouget et F. Sabio aux éditions Honoré Champion (Paris), et présenté par J.-C. Beacco dans *Le français dans le monde* du mois de juin 2003 (n° 327), rend d'ailleurs bien compte de la

---

<sup>291</sup> Nombreux sont nos étudiants qui enregistrent nos classes, puis retravaillent ensuite chez eux sur ces enregistrements sonores, nous demandant même parfois des précisions lors de la classe suivante.

<sup>292</sup> Très répandus au Japon, voire en passe de détrôner les dictionnaires « papier », notamment en raison de leur format « de poche », beaucoup plus pratique, au moins en ce qui concerne le transport, et dans une certaine mesure, la manipulation.

<sup>293</sup> Il serait également souhaitable de voir apparaître des dictionnaires à reconnaissance vocale, ce qui permettrait de ne plus systématiquement devoir transiter par l'écrit pour pouvoir ouvrir un dictionnaire, lourde contrainte que connaissent, par exemple, les apprenants français de chinois ou de japonais pour les caractères kanji.

demande. Dans son compte-rendu de la publication, Beacco écrit : « disposer de productions orales aussi facilement accessibles et transcrites avec cohérence est une opportunité [...] rare pour les enseignants de langue » et il conclut qu'il s'agit là d'un « échantillon qui fait attendre avec encore davantage d'impatience que ne se constitue, en France, un corpus d'oral facilement consultable, en règle avec les exigences juridiques ». Cela était déjà exprimé par M. Bilger dans la présentation qu'elle donnait du projet « Corpus de référence de français parlé » dans un numéro spécial du *français dans le monde* consacré au thème « Oral : variabilité et apprentissages » (janvier 2001, p. 103) :

*« En ce qui concerne le français de France, il n'existe toujours pas de "corpus de référence", ni pour l'écrit ni pour l'oral, c'est-à-dire que nous n'avons toujours pas à notre disposition un grand corpus "échantillonné" qui permettrait d'avoir une bonne représentation de tous les secteurs d'activité langagière du français ».*

Cette absence a donc suscité l'apparition de plusieurs projets : le projet « Corpus de référence de français parlé », initialement mené au sein de l'équipe CNRS ESA 6060, dirigée par C. Blanche-Benveniste, puis pris en charge par l'équipe DELIC de l'université de Provence, dirigée par J. Véronis et dont fait partie M. Bilger ; le projet « Phonologie du Français Contemporain » (PFC), coordonné par J. Durand (Université de Toulouse II), B. Laks (Université de Paris X) et C. Lyche (Université d'Oslo) que nous avons déjà présenté précédemment, ou encore le projet « Corpus clef » dirigé par B. Habert, cité par Bilger (p. 103), qui devrait lui aussi proposer un échantillonnage oral et écrit.

Nous reprenons ci-dessous en partie la présentation que fait Bilger du projet « Corpus de référence de français parlé » (pp. 103-105) : il s'agit de rassembler 160 échantillons, de 15 minutes chacun environ, de français parlé dans 40 villes différentes. Les critères sont : a) la zone géographique (Nord/ Sud ; Est/ Ouest), b) l'âge (18-30/ 30-65/ + de 65), c) le niveau d'étude (trois niveaux) et d) la situation de parole (« Privée » (récit de



vie et présentation d'un savoir-faire), « Publique » (réunions, débats, etc.) et « Professionnelle » (commerciale, etc.)). La transcription des enregistrements suit les conventions du GARS d'Aix-en-Provence. Parmi les objectifs du projet, on peut mentionner :

- La constitution d'un échantillonnage contemporain du français parlé dans l'Hexagone, basé sur des critères sociologiques mais aussi sur différentes activités langagières, ce qui donne accès à différents « genres » de discours.
- La caractérisation grammaticale de ces « genres » à travers l'exploitation du corpus.
- L'utilisation du corpus pour des études comparatives du type français général/régional ou bien dans l'Hexagone/ hors de l'Hexagone.

Les résultats devraient donc se présenter sous la forme de 40 heures d'enregistrement et environ 2000 pages de transcription, avec un outil informatique facilitant la navigation entre fichiers « textes » et fichiers « sons ».

L'utilisation des outils informatiques pour la mise en correspondance des données sonores et de leur transcription orthographique, de manière à ce que cette correspondance entre les deux inputs soit facilement perceptible et manipulable par l'utilisateur est particulièrement pertinente pour notre problématique, puisque de tels outils offrirait précisément la possibilité de travailler les correspondances entre oral et écrit en tenant compte de la variabilité de l'oral. Un tel dispositif se retrouve également dans le projet PFC, qui devrait mettre à la disposition du didacticien un corpus écrit et parlé, avec d'une part une liste de mots et un texte, communs à l'ensemble des enquêtes et lus à voix haute par tous les locuteurs enregistrés, et d'autre part l'enregistrement et la transcription de conversations « libres » et « guidées ». Ces données sont intégrées grâce à un outil informatique de présentation des données, le logiciel PRAAT.

Il nous semble ainsi que le projet PFC, notamment, peut constituer la première étape d'un parcours allant de l'actualisation des données sur le français oral pour

l'enseignement du français (contenu) à des exercices oraux relatifs à la variation francophone (mise en activités). L'intérêt didactique du projet pourrait ainsi, en première approximation, être présenté de la manière suivante :

Les apports du PFC :

- Des enregistrements oraux récents et variés, reflétant la réalité du français oral actuel.
- Des transcriptions orthographiques de ces enregistrements.
- Un outil de manipulation et de visualisation de ces données : PRAAT.
- Une remise à jour des contenus pédagogiques (inventaire phonémique, liaison et schwa).

Intérêt des enregistrements :

- Perspective sociolinguistique : l'approche de la variation est rigoureuse, et celle-ci peut être ainsi étudiée explicitement en classe, selon différents critères.
- Perspective phonétique : les apprenants doivent être graduellement entraînés à la variation phonétique (allophones), en particulier à un niveau intermédiaire.
- Perspective francophone : l'oral y est francophone et non français.
- Perspective générale : le contenu (en particulier schwa, liaisons, mais aussi caractéristiques morphosyntaxiques et autres) reflète la réalité du français tel qu'il est parlé aujourd'hui. Ce contenu est ainsi riche du point de vue des marqueurs pragmatiques et énonciatifs.

Intérêts des transcriptions :

- Elles permettent de visualiser les productions orales, ce qui, pour des apprenants est parfois essentiel, en particulier dans le cas d'enregistrements authentiques (segmentation, identification, etc.).
- Elles permettent de travailler les relations phonographémiques, ainsi que les différences entre oral et écrit.

- Elles permettent, d'un point de vue didactique, de mettre en place certaines activités d'apprentissage (approche de l'oral par l'écrit et vice-versa).

Intérêts de PRAAT (les plus simples):

- Il permet la combinaison temporelle et le découpage des formes sonores et graphiques des énoncés.
- Il permet (surtout) une manipulation efficace des bandes sonores (répétitions en boucle, retour en arrière, etc.) avec une précision que ne possèdent pas les outils classiques, alors que les enseignants en ont besoin, pour des raisons pratiques (perte de temps) et didactiques (parasites attentionnels).

Perspectives (exemple) :

On pourrait, à partir du corpus, concevoir un didacticiel du français pour des étudiants de niveau intermédiaire à avancé (étudiants entrant à l'université après avoir étudié le français au lycée par exemple), comportant des exercices sur la variabilité. Il s'agirait d'élaborer une palette d'exercices portant sur la variabilité, en s'inspirant des types existants, en tirant profit des apports de PFC. Deux types de plans principaux peuvent être adoptés : dominé par la variation (1) ou non (2).

Plan (1) (exemple) :

I Le français de Paris : activités.

II Le français de Toulouse : activités.

III Le français de Montréal : activités.

Plan (2) (exemple) :

I Discrimination et variation phonétique : lectures du texte, écoutes et observations.

II Questions de Compréhension : sélection de conversations (formelles et informelles).

III Oral-Ecrit : schwas, liaisons et dictée : observations et exercices.

IV Variation sociolinguistique : examen de quelques systèmes phonéto-phonologiques francophones.

Par ailleurs, il faut noter que l'utilisation de corpus dans l'enseignement, et ainsi l'exposition des apprenants à ces derniers, implique, non seulement de toucher à la variation inter-individuelle, mais aussi à la variation allophonique et aux phénomènes de variantes combinatoires dues, essentiellement mais non exclusivement, aux phénomènes d'assimilation. Les allophones du phonème /R/ à l'initiale ou en finale de mot, précédant ou suivant une voyelle, ne vont pas avoir la même substance phonétique, et ainsi être appréhendés de la même manière par les apprenants, car l'influence qu'exercent les sons entre eux jouent un rôle majeur dans la perception et donc la (re)production de ces derniers. Ces variantes peuvent par ailleurs fonctionner comme indices dialectaux ou sociaux, ou bien comme signaux phonostylistiques (Léon, 1992, pp. 72-76). Comme le note Hoyt (2001, p. 2) :

*« Using many different sources of language input will help the students' register flexibility in listening comprehension. Students need to know that target language has many different speakers and pronunciations beyond the way that the teacher speaks ».*

A cette remarque, nous aurions envie d'ajouter : « and beyond the way the teacher writes ». En effet, en ce qui concerne notre problématique, une telle remarque est tout à fait pertinente, puisque l'écrit (et les graphèmes) constitue une norme fixe par excellence (dans une certaine mesure), dont l'oralisation doit se confronter à la variation de l'oral dans l'espace francophone. Les outils proposés par les projets présentés sont donc tout à fait reliés à notre thème d'étude, en ce que le logiciel PRAAT, par exemple, pourrait permettre aux apprenants, de manière plus ou moins explicite selon les options de l'enseignant, de confronter en perception, sur un même support, la variabilité des réalisations orales face à la fixité de l'écrit. Ce travail serait bénéfique, à la fois pour l'apprentissage des correspondances phonographémiques du français, mais également pour celui des catégories et des phénomènes phonético-phonologiques propres au français oral.

Ce dernier point, tout comme celui des corpus en général, renvoie donc à la notion de « norme », et ainsi à l'une des questions posées au début de notre thèse : quel français enseigner ? Bien souvent, dans le cas d'apprenants en milieu hétéroglotte, « la norme » est inexorablement réduite à celle de l'enseignant, que celui-ci soit francophone natif ou non-natif, ce qui renvoie une fois encore à la question de la variabilité dans l'espace francophone, puisque les enseignants ne partagent pas tous la même « norme ». La question de l'utilisation des « accents » dans les méthodes de français langue étrangère a été notamment examinée par Billières et Borrell (1990), qui distinguent quatre types d'accent (p. 20) : a) Etranger, b) Francophone, c) Régional, d) De référence. Or, il faut également prendre en compte les souhaits et les attentes des apprenants qui veulent apprendre « le » français (Billières et Borrell, 1990, p. 37) :

*« L'écoute d'un accent étranger leur montre [(aux apprenants)] que toutes les prononciations sont acceptées ce qui peut les sécuriser. Le côté négatif existe aussi. Ces accents étrangers risquent d'être ressentis comme dévalorisants en fonction du critère phonético-esthétique évoqué plus haut ».*

Ce thème de la « norme » et des « corpus » présente donc des dimensions à la fois linguistiques, sociolinguistiques, mais aussi psycholinguistique *via* la notion de « norme didactique » : entre la « norme (socio-)linguistique » (abordée dans sa relative fixité *via* l'écrit ou dans sa variabilité *via* l'oral) et la « norme didactique » (adaptée à ce que l'on sait des processus d'apprentissage et au public auquel on s'adresse), le didacticien doit donc, comme dans le cas de la transposition entre modèles linguistiques et modèles didactiques, effectuer des choix, choix qui nécessitent une certaine réflexion préalable.

### **III.3. Norme didactique**

Lorsque l'on traite de « norme » didactique, de manière générale dans un premier temps, puis phonético-phonologique dans un second temps, il nous faut d'abord prendre position par rapport à la description et à la prescription linguistique, avant de

considérer l'adaptation de cette norme au cadre didactique *via* un passage de la sociolinguistique à la psycholinguistique.

### **III.3.1. Description vs. Prescription vs. Adaptation**

Sans reprendre la discussion amorcée précédemment sur la « norme linguistique » de référence (Laks, 2003b ; Girard et Lyche, 2003) à enseigner en classe, dans notre cas particulièrement vis-à-vis de la dimension phonético-phonologique, il convient en revanche d'insister sur les contraintes propres au cadre didactique, qui transforment l'opposition binaire « descriptivisme/prescriptivisme » (François, 1974), en une problématique à trois versants :

- 1) *Décrire la langue*, notamment pour assurer à la norme didactique issue de cette description une adéquation maximale à la réalité linguistique, dans sa variabilité, des locuteurs natifs ;
- 2) *Prescrire la langue*, notamment pour que la norme didactique issue de cette prescription puisse assurer une intercommunication maximale entre locuteurs natifs et non-natifs ;
- 3) *Adapter la norme didactique*, à la fois à la description et à la prescription, mais aussi, sinon surtout, au public auquel elle est destinée (niveau de compétence en L2, objectifs d'apprentissage, L1, processus d'apprentissage, etc.), afin d'être optimale pour l'apprentissage progressif de la langue, visant, à terme, à ce que l'interlangue « finale » des sujets, la norme didactique en fin d'apprentissage et la norme linguistique de référence puissent être relativement identiques.

Tel que nous l'avons abordé, le thème de la norme didactique concerne donc à la fois les niveaux linguistique (cognitif), en termes de contenus, et métalinguistique (métacognitif) en termes de représentations explicites. Si les deux premiers versants de la problématique ont été abondamment commentés, tant par les linguistes que par les didacticiens, il nous semble en revanche que le troisième versant reste encore relativement peu traité dans le domaine, alors qu'il renvoie précisément à la dimension psycholinguistique de l'apprentissage.

La question que l'on doit poser à ce sujet est formulée ainsi par Van Patten (2002, p. 105) :

*« How are pedagogical norms constructed in foreign language instruction [...] ? What is the relationship between these norms and what we understand about the psycholinguistic processing mechanisms that learners take to the task of acquisition ? These are fundamental questions that seldom receive discussion in the professional literature ».*

Une telle réflexion conduit inévitablement à conceptualiser la « norme didactique » en classe comme une norme *évolutive* (Detey, 2002), en particulier *via* la sélection des contenus linguistiques, évoluant au cours de la progression des apprenants. Il ne s'agit cependant en aucun cas de simplification ou de réductionnisme abusif, tel qu'on pourrait le croire ou le craindre : un appauvrissement de l'input langagier qui réduirait le potentiel d'apprentissage des sujets<sup>294</sup>. Une telle critique était déjà de mise, sous diverses formes (Rivenc, 1979/2000), lorsque les premières enquêtes sociolinguistiques de terrain à finalité pédagogique, en premier lieu le *Français Fondamental*, avaient finalement permis de progresser dans un certain paradigme de l'optimalité (Rivenc, 1961/2000, p. 32). Il est depuis clairement apparu, en particulier à une époque où le terme de « fréquence » ressurgit comme l'un des termes clés de nombreuses approches en recherche sur l'acquisition d'une langue seconde (Ellis, N., 2002), que les enquêtes de corpus permettaient au contraire de mettre concrètement en œuvre l'un des quatre principes d'élaboration de normes pédagogiques formulés par Valdman (1989) : « elles devraient refléter la parole réelle de locuteurs natifs dans des situations de communications authentiques »<sup>295</sup>. Car c'est bien de « norme pédagogique » qu'il s'agit dans notre propos, un sujet toujours d'actualité, comme l'indique la récente

---

<sup>294</sup> Comme cela est indiqué dès le début de la préface du *Français Fondamental* (1<sup>er</sup> degré) (1959) : « *Le français fondamental [...] n'est pas une langue différente du français "normal". On a seulement effectué un choix méthodique dans le vocabulaire et la grammaire de façon que l'essentiel puisse être enseigné d'abord* ».

<sup>295</sup> Notre traduction.

publication du volume *Pedagogical norms for second and foreign language learning and teaching* (Gass, Bardovi-Harlig, Magnan et Walz, 2002).

Rappelons ces quatre principes (Valdman, 1989, p. 21) :

1. Elles devraient refléter la parole réelle de locuteurs natifs dans des situations de communications authentiques.
2. Elles devraient se conformer à a vision idéalisée qu'ont les locuteurs natifs de leur propre utilisation de la langue.
3. Elles devraient se conformer aux attentes des locuteurs natifs et à celles des apprenants étrangers concernant le type de comportement linguistique approprié pour des apprenants étrangers.
4. Elles devraient prendre en compte les facteurs de traitement et d'apprentissage<sup>296</sup>.

Tandis que le premier renvoie à une dimension strictement linguistique, les deuxième et troisième relèvent du domaine sociopsychologique. Le dernier, enfin, est clairement d'ordre psycholinguistique (Magnan et Walz, 2002, p. 31). Si les enquêtes comme celle du *Français Fondamental* ont notamment permis de travailler sur le premier principe, il semble que, dans le domaine grammatical par exemple, des propositions comme celles de Santacroce (1999) de « grammaires transitionnelles » tendent à appliquer également le quatrième principe, les trois autres s'appliquant en diachronie (*via* le concept de « norme évolutive ») selon la hiérarchie suivante :

1. Quatrième principe
2. Troisième principe
3. Premier principe
4. Deuxième principe

---

<sup>296</sup> « 1) They should reflect the actual speech of target language speakers in authentic communicative situations ; 2) they should conform to native speakers' idealised view of their speech use ; 3) they should conform to expectations of both native speakers and foreign learners concerning the type of linguistic behaviour appropriate for foreign learners ; 4) they should take into account processing and learning factors » (notre traduction).



Ce changement de hiérarchie pourrait facilement correspondre au passage d'une approche sociolinguistiquement orientée à une approche psycholinguistiquement orientée, un passage déjà auguré par Valdman (1987, p. 145, cité par Magnan et Walz, 2002, p. 29) :

*« Là où existe une certaine diversité sociale ou géographique, les facteurs sociolinguistiques seront abandonnés en faveur de facteurs psycholinguistiques : ce seront les variantes les plus facilement traitées par les apprenants de langue seconde qui seront sélectionnées »<sup>297</sup>.*

### **III.3.2. De la sociolinguistique à la psycholinguistique pour la didactique**

Il ne s'agit pas ici de donner dans les luttes d'école : il s'agit au contraire d'intégration des résultats et des approches, une intégration qui nécessite inévitablement dans un premier temps un rééquilibrage des tendances. Ce passage du « sociolinguistique » au « psycholinguistique » n'est lui aussi qu'apparent et provisoire, car la notion de « norme évolutive » autorise précisément l'intégration méthodologique des dimensions socio-interactionnelles et linguistiques des contenus d'enseignement. Il ne s'agit pas de privilégier exagérément ce que l'apprenant *pourrait* apprendre au détriment de ce qu'il *devrait* apprendre (sans quoi toute communication réelle serait impraticable), mais bien de *faciliter* l'apprentissage *dans les premiers temps*, avant d'atteindre un stade langagier ou la pertinence sociolinguistique pourrait de nouveau reprendre le pas sur la pertinence psycholinguistique. Il s'agit aussi et surtout de reconnaître l'existence *de facto* dans de nombreuses classes, hors de l'hexagone en particulier, d'une telle norme évolutive, principalement implémentée dans le discours de l'enseignant (« teacherese »)<sup>298</sup>, afin de lui donner davantage de cohérence et de finalités.

---

<sup>297</sup> « Where social or geographical diversity exists, sociolinguistic factors will be abandoned in favor of psycholinguistic ones : those variants will be selected that are most easily processable by second language learners » (notre traduction).

<sup>298</sup> Les études sur la conversation exolingue, révélant une « asymétrie significative variable entre interlocuteurs du point de vue du partage du code » (Py, 1988, p. 10), dévoilent des stratégies conversationnelles qui pourraient tout à fait trouver écho en classe de langue, en tant que stratégies d'enseignement ou d'apprentissage, cette asymétrie renvoyant en partie à celle qui existe entre enseignant et apprenant (Py, 1988, p. 10):

Dans leur introduction, Bardovi-Harlig et Gass indiquent, à juste titre, que la conception de normes pédagogiques est un travail nécessitant au préalable la réalisation d'enquêtes concernant l'usage réel de la langue par les locuteurs natifs ; mais qu'en est-il de l'usage réel du discours interlangagier des apprenants (en tant que système langagier et communicatif à part entière, et non en tant que produit instable d'apprentissage) d'une part, et du discours, sinon simplifié du moins adapté, des enseignants dans leurs classes ? Nos suggestions ne font que rejoindre ce que posent les auteurs (2002, p. 3) :

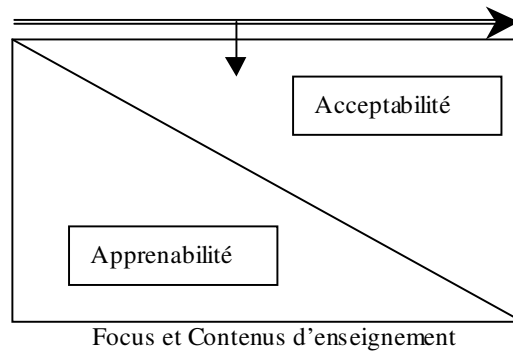
*« Les normes pédagogiques représentent une étape intermédiaire, ou une série d'étapes intermédiaires, pour les apprenants au cours de leur progression dans l'acquisition des normes natives de la langue en question [...]. [Le principe de base est le suivant] : sélectionner et enseigner une forme de langue qui soit acceptable aux locuteurs natifs mais plus facile à apprendre que le système langagier natif complet »<sup>299</sup>.*

Si l'équilibre entre ces deux pôles, *acceptabilité* et *apprenabilité*, est parfois difficile à trouver en synchronie, c'est bien en diachronie que se trouve la solution, telle qu'on peut l'illustrer ainsi :

---

*« Simplification de son propre système par le natif (« foreigner talk »), simplification du système linguistique du natif par l'apprenant (interlangue), ponctuation (fractionnement du discours propre ou du discours de l'autre en séquences accessibles), reformulations, mentions (verbalisations partielles du message) ».*

<sup>299</sup> « Pedagogical norms represent a mid-point, or series of mid-points, for learners as they progress towards acquiring native language norms [...]. [The basic principle is] : select and teach a form of language that is acceptable to native speakers but easier to learn than the full native language system » (notre traduction).



En déplaçant le « curseur » d'enseignement de gauche à droite, il s'agirait là du parcours que l'on pourrait proposer dans le cas d'apprenants débutants. Cela correspond à la définition de norme pédagogique que proposent Magnan et Walz (2002, p. 15) :

*« Les normes pédagogiques sont des abstractions qui servent de médium entre les réalités complexes de la variation linguistique et les difficultés d'apprentissage linguistique typiquement rencontrées par les apprenants de langue étrangère »<sup>300</sup>.*

Cette distinction peut renvoyer à une autre, non moins importante, celle entre norme productive et norme réceptive. La modularisation des compétences étant par ailleurs d'actualité, dans le domaine européen au moins, on doit en effet s'arrêter sur la séparation entre pôle perceptif et pôle productif. Ainsi dans le domaine phonético-phonologique, Valdman écrit<sup>301</sup> (cité par Magnan et Walz, 2002, p. 33) :

*« La présentation de documents authentiques, qui sert d'ancrage aux approches actuelles de l'enseignement du FLE ne requiert-elle pas que l'on*

---

<sup>300</sup> « Pedagogical norms are abstractions that mediate the complex realities of linguistic variation and typical language learning difficulties experienced by foreign language learners » (notre traduction).

<sup>301</sup> « Letter from the president », *American Association of Teachers of French, National Bulletin*, 21 (4), 1996, pp. 1-5.

*fasse écouter les voix réelles de la francophonie aux apprenants sans toutefois – et j'insiste là-dessus – leur demander de les imiter ».*

Cette question est tout à fait d'actualité lorsque l'on considère par exemple les travaux menés dans le cadre du projet Phonologie du Français Contemporain (PFC) (Durand, Laks et Lyche, 2003) par exemple. Si la norme réceptive doit impérativement être sociolinguistiquement orientée, comme cela est le cas depuis plusieurs décennies maintenant, la norme productive en revanche nous semble devoir être fondée psycholinguistiquement, des fondements auxquels nous nous attachons précisément dans notre réflexion. Cette distinction entre production et réception permet de résoudre le problème potentiel d'appauvrissement de l'input langagier évoqué plus haut. Les conclusions de Magnan et Walz s'associent tout à fait à nos perspectives (2002, p. 35) :

*« En termes d'enseignement, nous avons besoin de davantage d'informations sur les processus acquisitionnels, car [...] ces critères pour une norme pédagogique restent les plus difficiles à appliquer. [...] Les objectifs de l'instruction en langue étrangère sont, de nouveau, reconsidérés afin de d'insister sur la méta-cognition et sur la médiation intellectuelle entre les langues et les cultures »<sup>302</sup>.*

On ne peut alors ignorer les travaux de certains chercheurs qui se sont précisément concentrés sur le quatrième principe de Valdman, en particulier ceux de Van Patten (1996, 2002) et ses collègues (Lee, 2002) sur l'enseignement grammatical tels que présentés dans le cadre de ce qu'ils nomment « Processing Instruction » (P.I.). Ce cadre, basé sur de nombreux travaux psycholinguistiques dans le domaine de l'acquisition des langues secondes, et inséré dans une approche communicative de l'enseignement des langues, se présente non pas comme une « méthode » globale d'enseignement des

---

<sup>302</sup> « In terms of teaching, more information is needed on acquisitional processes, for [...] these criteria for a pedagogical norm remain the hardest to apply. [...] The goals of foreign language instruction are, again, being reconsidered to suggest greater focus on meta-cognition and on intellectual mediation between languages and cultures » (notre traduction).

langues, mais comme « une option pour toute approche d'obédience communicative qui souhaiterait intégrer un certain type d'accent-sur-la-forme »<sup>303</sup> (Van Patten, 2002, p. 115). Cette option est celle de la manipulation, ou plutôt la structuration, de l'input langagier dans les activités grammaticales proposées, précisément en fonction du traitement de cet input par les apprenants et des difficultés rencontrées telles que définies dans le cadre de P. I. Cette approche orientée vers l'input (« input-based approach to focus on form », Van Patten, 2002, p. 112) repose donc sur une certaine conception du traitement de l'input (« input processing ») par les apprenants, décrite comme telle par Lee (2002, p. 120) : « Nous pouvons définir le traitement de l'input comme le processus de connexion de la forme et du sens à partir des données linguistiques de l'input dans le but de construire un système linguistique »<sup>304</sup>. On le voit, une telle approche, dans le domaine grammatical, rejoint tout à fait nos propres préoccupations : fonder sur le quatrième principe de Valdman une méthodologie « cognitive » insistant sur la manipulation, ou la structuration de l'input, en vue de faciliter ou d'orienter son traitement, avec un souci de mise en focus attentionnel.

Ainsi, pour en revenir à la notion de norme, les scrupules que certains pourraient ainsi ressentir *a priori* à l'idée d'enseigner *une norme didactique*, qui pourrait parfois se trouver provisoirement en porte-à-faux avec « la norme » francophone, telle que représentée par celle de l'enseignant et des supports pédagogiques qu'il emploie, norme didactique que l'on pourrait éventuellement comparer à une certaine variété d'« interlangue », en dépit de la *réalité*, tant de l'apprentissage d'une langue étrangère<sup>305</sup>, que des conversations plurilingues en langues étrangères natif/étranger ou étranger/étranger, ne reposent que sur un sentiment de malaise axiologique ou idéologique. Bardovi-Harlig et Gass rappellent que (2002, p. 2):

---

<sup>303</sup> « An option for any communicatively-oriented approach to language that might want to incorporate some type of focus-on-form » (notre traduction).

<sup>304</sup> « We can define input processing as the process of making form-meaning connections from the linguistic data in the input for the purpose of constructing a linguistic system » (notre traduction).

<sup>305</sup> Puisque les enseignants de langue simplifient, *de fait*, leur propre discours dans la classe.

*« La notion selon laquelle tous les dialectes sont grammaticaux à égal niveau et selon laquelle la langue standard est la norme pour des raisons autres que linguistiques (par exemple sociales, économiques, politiques) est devenue un lieu commun et largement acceptée par la communauté scientifique »<sup>306</sup>*

Certains de ces réticents seraient d'ailleurs sans doute surpris s'ils examinaient attentivement leur *propre* discours didactique de terrain (à condition d'en avoir un), et l'écart entre ce discours et la représentation qu'ils s'en font. Cela s'applique parfaitement au cas de l'apprentissage de la « prononciation » en langue étrangère.

### **III.3.3. Le cas de la « prononciation »**

En ce qui concerne les contenus phonético-phonologiques, Blanchet (2002, p. 49) prône lui aussi un recours à la phonologie (fonctionnelle dans son cas) pour la mise en place d'une sélection et d'une prioritarisation desdits contenus. Pour lui (p. 49), « une langue n'est ni enseignable ni apprenable dans sa totalité, non seulement parce qu'elle est infinie, mais aussi dès lors qu'on la conçoit comme un ensemble de variétés toujours susceptibles de variations ». Il propose donc de « sélectionner l'essentiel phonologique avant les détails phonétiques, et, dans le phonologique, à la fois un système minimal à faire maîtriser et des variations auxquelles l'apprenant doit être sensibilisé, au moins en tant que récepteur ». Blanchet a donc élaboré un « système phonologique minimal », à adapter aux finalités de l'apprentissage, et qui élimine plusieurs oppositions : /a/ vs. /ɑ/ ; /ɛ/ vs. /œ/ ; /ŋ/ vs. /nj/ ; /R/ vs. toutes ses variétés ; consonnes simples vs. consonnes géminées, etc. Les conséquences sont immédiates : « cela invite à ne plus enseigner systématiquement une norme unique d'une langue conçue comme homogène » (p. 50), d'où des problèmes idéologiques et sociolinguistiques déclenchés par cette approche « plurinormaliste » phonologique, bien différente de celle, normative et phonétique, proposée par exemple par Léon (1992).

---

<sup>306</sup> « The notion that all dialects are equally grammatical and that the standard language is the norm for reasons other than linguistic ones (e.g. social, economic, political) became commonplace and widely-accepted » (notre traduction).

En ce qui nous concerne, par delà la question des inventaires segmentaux, minimaux ou non, il nous semble que la notion de norme dans le domaine de la prononciation doit également être envisagée à l'interface entre phonétique et phonologie, notamment dans le cadre d'activités perceptives ou de correction phonétique. A titre d'exemple, nous pourrions mentionner, dans le domaine qui nous occupe, le cas de l'apprentissage des catégories phonético-phonologiques /R/ et /l/ en français par des apprenants japonais. Nous avons en effet constaté à plusieurs reprises dans nos classes que la plupart de ces apprenants avaient moins de difficultés à percevoir et reproduire la fricative uvulaire non voisée [χ] que la fricative uvulaire voisée [ʁ] dont elle est un allophone<sup>307</sup>. L'opposition phonétique (et à sa suite phonologique) entre le [χ] et le [l] semblerait donc apparemment plus aisée, d'un point de vue perceptif et productif, que celle entre le [ʁ] et le [l]. On pourrait bien sûr s'interroger sur les raisons d'un tel phénomène, en particulier : est-ce parce que le [χ] n'appartient plus exactement à la classe des liquides (mais plutôt à celle des fricatives) ou bien parce que le [χ] constitue un allophone du [h] japonais ? Cependant, ce qui nous intéresse ici, d'un point de vue didactique, c'est d'imaginer que privilégier le travail en perception/reproduction sur un tel allophone secondaire (plus accessible aux apprenants) plutôt que sur la réalisation prototypique ([ʁ]) pourrait permettre de mieux établir l'opposition phonologique /R/ vs. /l/. Il resterait alors ensuite un travail phonétique de rétablissement des allophones, l'opposition phonologique principale étant bien mise en place. C'est ce que nous pratiquons nous-mêmes, de manière ponctuelle et non-systématique, dans nos activités de correction phonétique.

Doit-on/Peut-on ainsi privilégier un allophone au détriment d'un autre, sous couvert de telles justifications. Les sociolinguistes et les fonctionnalistes ne s'y opposeraient sûrement pas. Du point de vue de la théorisation des pratiques, on pourrait alors

---

<sup>307</sup> A ce propos, voir (Peperkamp, Pettinato et Dupoux, 2003).

considérer qu'on ne modifie pas tant les caractéristiques acoustiques de l'allophone (réalisation phonétique) (qui évoquerait la technique verbo-tonaliste de « prononciation nuancée », correspondant à un travail de phonétique corrective) que les caractéristiques distributionnelles du phonème (interface phonético-phonologique) (travail de « phonologie corrective »).

Cependant, on pourrait craindre qu'une telle démarche prive les apprenants des informations distributionnelles dont ils ont besoin pour l'apprentissage global du système phonético-phonologique, informations particulièrement cruciales si l'on se place dans une perspective émergentiste, considérant les unités phonologiques (phonèmes, syllabes) comme des objets émergents (issus de leurs schémas distributionnels et de leur fréquence d'occurrence) et non comme des primitives phonologiques. De la même manière, certaines règles phonético-phonologiques, par exemple de dévoisement, seraient ainsi inaccessibles aux apprenants et pourraient fausser le bon déroulement de leur apprentissage. Un tel problème devrait donc être soigneusement abordé en croisant les notions de norme en synchronie et en diachronie (progression d'apprentissage) et en prenant également en compte la thèse de « l'apprenabilité », selon laquelle certains éléments ne peuvent être appris ni avant d'autres, ni avant que l'apprenant ne soit prêt à les apprendre, en termes développementaux.

Néanmoins, comme le note Denis (1994, p.22), dans le domaine de l'imagerie mentale :

*« L'auteur d'une représentation peut se trouver conduit à introduire une certaine distorsion, s'il a des raisons d'en attendre une plus grande efficacité de la fonction que doit servir la représentation ».*

En ce qui concerne le domaine didactique, il ajoute (p. 32) :



*« En ce domaine [la didactique], un mouvement s'est très fortement dessiné en faveur de la notion de représentation, jugée préférable à celles de "savoir" ou de "connaissances", en ce qu'elle ne comporte pas nécessairement la caractéristique d'être "vraie" »<sup>308</sup>.*

Il faut alors constater que, à l'instar de la technique de la « prononciation nuancée » de la méthode verbo-tonale, la modification de l'input en fonction des cibles d'enseignement et du public auquel il s'adresse est une pratique relativement courante, à la fois dans l'enseignement des langues (« teacherese »)<sup>309</sup>, mais aussi dans certains domaines de recherche s'intéressant à l'apprentissage et à la perception de la parole, comme l'indique cette affirmation de Iverson, Kuhl, Akahane-Yamada, Diesch, Tohkura, Kettermann et Siebert (2003, p. B54) :

*« Likewise, enhancement of acoustic cues, in infant-directed [...] or signal-processed [...] speech, may improve early stages of learning by making the most robust acoustic cues more salient »*

Cela évoque également, d'une autre manière, la technique dite de « fading », qui nous intéresse particulièrement en raison de la manipulation de l'input, et ainsi de la norme, phonétique, qu'elle implique, et qui est décrite ainsi par Jamieson et Morosan (cités par Silveira, 2002, p. 113) :

---

<sup>308</sup> L'auteur précise: « Le terme de "représentation" est utilisé pour désigner à la fois un processus et le produit de ce processus » (p. 14).

<sup>309</sup> Nous pouvons également mentionner, à titre tout à fait anecdotique, le cas de cet enseignant de français au Japon, qui nous a indiqué qu'il traitait parfois le problème de la lecture à voix haute du /R/ français en indiquant, en japonais, aux étudiants dès les premiers cours : « Quand vous voyez un <R> français ne pensez pas à un <r> japonais en romaji, pensez que c'est un <h>. Par exemple, quand vous voyez <Paris>, pensez à <Pahi> ». Pour cet enseignant, la prononciation qui en résultait, en particulier en position intervocalique, était ainsi beaucoup plus proche de la « norme » attendue, même si, en position initiale de mot, les résultats étaient naturellement moins probants. Ce même collègue donnait par ailleurs la recommandation suivante aux étudiants qui butaient sur la prononciation de la nasale /ã/ mais pas sur celle de /õ/ : « Pour prononcer "an" (difficile), vous prononcez "on" (plus facile) en souriant ! »....

*« This technique attempts to train a perceptual contrast, without subject errors, by beginning with clearly discriminable stimuli which may exaggerate the normal perceptual differences or add other salient features. Progress in training is made by slowly reducing the magnitude of the perceptual contrast, in small steps, so that the task never becomes too difficult and errors remain infrequent ».*

Cette technique est également mentionnée par Hardison (2003, p. 498) :

*« An alternative training technique for L2 learners using computer-enhanced speech (exaggerated stimuli) and adaptative training [...]. This is reminiscent of the perceptual fading technique [...] used to train Canadian francophones to identify /θ/ and /ð/ by exaggerating the amount of frication. [...] Although adaptative training can facilitate learner's noticing of the critical features distinguishing two sounds and subsequent awareness of the contrast [...], the question remains as to the generalizability and retention of improvements in accuracy as a result of brief periods of training with exaggerated stimuli ».*

Tout en souscrivant pleinement aux réserves de Hardison, notamment vis-à-vis du transfert inter-tâche des compétences, il apparaît que nous retrouvons ici encore le facteur de saillance (Landragin, 2004). La question est alors la suivante : dans quelle mesure, dans le cadre didactique qui est le nôtre, ne serait-il pas possible d'adapter l'input (et donc la norme à laquelle renvoie ce dernier) en rendant certaines caractéristiques de celui-ci plus saillantes, cette saillance devant varier, pour ne pas dire être planifiée, en fonction du public, de la progression, des tâches, et surtout des objectifs d'enseignement ?

Il est intéressant de constater que de grands phonéticiens comme J. C. Wells (2000), au fait des variations allophoniques au sein d'une même communauté linguistique, n'hésite pas à « congédier » la norme pour faciliter l'apprentissage : à propos des

difficultés rencontrés en anglais par les apprenants japonais lorsqu'ils sont confrontés à des groupes consonantiques en position finale de mot (par exemple « help » ou « milk » souvent prononcés [he.rʊpɥ], [mi.rʊkɥ]), il écrit (p. 124) :

*« A radical solution which I think deserves consideration (though it may shock some teachers) is to follow the Estuary English habit of vocalizing the lateral, that is replacing it with a vowel of the [o] type. Learners could aim at [heop, beot, mɪok, heoθ, eos]. They would be in good company, since millions of English and American speakers do the same thing ».*

Wells insiste également sur l'importance du respect des structures syllabiques, et présente plusieurs stratégies *phonétiques* d'évitement des épenthèses et ainsi utiles à l'apprentissage *phonologique*, par exemple des stratégies de concaténations liées aux problèmes de coarticulation (p. 125) ou bien encore des stratégies de simplification, reflets des phénomènes authentiques en L2 d'élision et d'assimilation (p. 126) (par exemple « next time » prononcé [ˈneks ˈtaim]). Wells propose également à plusieurs reprises de se servir de certains phénomènes phonétiques existant, de manière limitée, en japonais, pour servir de repères aux apprenants dans leur apprentissage, faisant ainsi un usage raisonné des phénomènes linguistiques en L1 pour l'apprentissage de la L2.

De la même manière, en ce qui concerne la « norme correctrice » en correction phonétique, nous pouvons également mentionner l'une de nos propres pratiques de classe, qui, globalement, consiste, suite à une production erronée d'un étudiant, à, d'une part répéter (ou essayer de répéter) la production erronée de ce dernier en « exagérant » le trait erroné, et d'autre part à répéter le modèle initial en « exagérant » le trait qui ne semble pas avoir été correctement perçu, ou du moins reproduit. D'un point de vue phonologique, il s'agit donc de rendre saillante une opposition de trait ou de structure ; d'un point de vue psycholinguistique, il s'agit de faciliter la perception de l'écart entre les deux modèles ; d'un point de vue didactique, il s'agit de faire prendre

conscience à l'apprenant, en orientant explicitement son attention, notamment via la dimension humoristique, et donc émotionnelle<sup>310</sup>, de la technique, de cet écart.

Dans ce cas, le « modèle erroné », que l'apprenant doit contraster avec la cible d'apprentissage, est la production effective de l'apprenant (et non une production hypothétique basée sur l'analyse contrastive classique de la L1 et de la L2), ce qui permet de prendre en compte, à la fois la variation sociolinguistique potentielle de l'interlangue des sujets, mais également les cas d' « émergence du non-marqué », en particulier pour des apprenants de niveau intermédiaire ou avancé, plus à même que les débutants de percevoir de tels contrastes et d'analyser leurs propres productions, en vue de pallier les phénomènes de fossilisation. Comme cela apparaît clairement ici, ce n'est plus tant la « norme pédagogique » qui est modifiée que la « norme correctrice », adaptée aux apprenants.

Ainsi, une fois encore il apparaît que, dans un cadre didactique, le critère de *saillance*, à un niveau perceptif, mais également à un niveau métacognitif, doit primer sur la notion de norme, qui doit s'adapter aux objectifs d'enseignement et d'apprentissage. Si la modification de la norme telle que nous venons de l'envisager est psycholinguistiquement orientée, la modification de la norme telle qu'elle se présente dans l'utilisation de corpus est sociolinguistiquement orientée. « La » norme émerge alors, à la fois de la multiplicité et des régularités des exemplaires, et de la mise en valeur des indices les plus « robustes » pour la constitution de catégories plus abstraites.

Si cette approche n'est pas particulièrement nouvelles dans le domaine de la prononciation, envisagée sous un angle purement audio-articulatoire, le thème principal de notre étude permet en revanche de souligner l'important travail qu'il reste à accomplir dans le domaine de la perception visuelle, lorsque l'on considère que la perception de la parole est fondamentalement audiovisuelle, et que le domaine

---

<sup>310</sup> Puisque, comme l'indique Mesulam (1998, p. 1036) : « The modulatory effects of emotion and motivation influence the attentional resources allocated to a sensory stimulus ». Voir aussi (Blondel, 1998) pour le lien entre émotions et attention.

phonographémique est, dans nos sociétés « lettrées », tout à fait connecté au développement des compétences orales en langue étrangère<sup>311</sup>.

A titre illustratif, pourrait-on en effet envisager une « fading technique » bimodale, qui renverrait alors, d'une certaine manière à « l'effet McGurk » ? Ne serait-il pas envisageable de « jouer » sur la combinaison des saillances visuelles et auditives dans l'apprentissage de l'oral, une approche nous adoptons nous-mêmes, de manière relativement intuitive, dans nos classes de correction phonétique, en modifiant l'input auditif et visuel (oro-labial ou gestuel) que nous fournissons à nos apprenants en vue d'obtenir une production qui corresponde à l'objectif d'enseignement ? Une telle pratique est sans doute courante parmi les enseignants, mais il nous semble qu'un véritable travail de formalisation des procédures serait à considérer, en vue de les améliorer, voire de les systématiser.

Par ailleurs, dans le domaine phonographémique, cela nous renvoie à de nombreux travaux en psychologie ergonomique que nous avons déjà mentionnés, lesquels conduisent à la réalisation d'outils comme le didacticiel de Stenton et de ses collègues (2005), qui, pour « faire percevoir » l'accent lexical à des apprenants français d'anglais, modifient la taille et la couleur de la syllabe accentuée, la rendant ainsi plus saillante *via* une modification de l'input visuel orthographique<sup>312</sup>. On pourrait donc, en poussant le raisonnement à l'extrême, imaginer, à l'instar de la phonétique corrective, une « phonographémie corrective », destinée, non pas à l'apprentissage de l'orthographe, mais, pour le public qui nous intéresse, à l'utilisation du canal visuel pour la correction phonétique. Très grossièrement, l'idée serait la suivante : puisqu'on « déforme » les sons, pourquoi ne pas « déformer » les représentations visuo-graphiques ? On pourrait ainsi imaginer que, pour des apprenants allemands de français qui auraient tendance à dévoiser les consonnes finales, il conviendrait, dans l'utilisation de la lecture comme

---

<sup>311</sup> L'actualité du sujet est d'ailleurs illustrée par l'article de Billières (2005) dans le domaine de la phonétique corrective.

<sup>312</sup> Bien que se pose ici le problème des décalages entre syllabation de l'oral et syllabation de l'écrit, qui nécessiterait peut-être certaines modifications supplémentaires de l'input visuel.

tâche d'apprentissage du système phonético-phonologique du français, d'augmenter le pas de la consonne graphique concernée (« élève » → « élèVe »), la souligner ou tout autre dispositif de mise en avant explicite visuelle. De tels dispositifs sont particulièrement à l'ordre du jour dans le domaine des hypermédias, comme en témoigne l'exemple du didacticiel de Stenton et de ses collègues.

Ainsi, à travers cet examen de plusieurs facteurs essentiels dans l'apprentissage de l'oral en langue étrangère, il apparaît que la coordination de l'écrit et de l'oral dans l'enseignement/apprentissage de ce dernier ne peut se véritablement être efficace que si celle-ci est méthodologiquement intégrée de manière cohérente à l'ensemble de la formation linguistique, ce qui conduit à une réflexion méthodologique globale, agençant, de manière raisonnée, ces différents facteurs.

### **III.4. Perspectives didactiques méthodologiques**

Qu'il s'agisse d' « enseigner une langue étrangère » ou d' « aider à apprendre à communiquer en langue étrangère » (Rivenc, 2000), l'objectif relève dans les deux cas du domaine de la *formation*. Formation linguistique et/ou culturelle, il s'agit pour l'enseignant à la fois d'*informer* les étudiants (aspects théoriques et encyclopédiques) sur la société à découvrir et la langue à apprendre, et de les *former* (aspects pratiques et interactionnels) à certains comportements langagiers, perceptifs et communicatifs. Les formes et les contenus de cette formation sont variables, fonctions notamment du public d'apprenants, du contexte institutionnel (Rivenc-Chiclet, 1993) et des choix éducatifs de l'enseignant (Boogards, 1988). Il n'y a évidemment pas de solution unique au problème de l'apprentissage : chaque langue peut être apprise de multiples manières, avec une peine et un labeur tout autant variables que les résultats obtenus.

#### **III.4.1. Intégration méthodologique de l'articulation oral/écrit**

A une époque où la didactique des langues, après s'être difficilement extirpée de son carcan de linguistique appliquée, a, dans une certaine mesure, laissé place à la didactologie des langues et des cultures (DLC), le terme *méthodologie* se doit d'être manié avec précaution. Que l'on choisisse de définir la méthodologie comme ce que

Besse nomme « méthode », c'est-à-dire « un ensemble raisonné de propositions et de procédés (d'ordre linguistique, psychologique, sociopédagogique) destinés à organiser et à favoriser l'enseignement et l'apprentissage d'une langue naturelle » (Besse, 1985, p. 14), ou bien comme un « ensemble concerté, cohérent, souple et évolutif de recherches, de réflexions et d'actions fondé sur des bases théoriques (en linguistique, en psycholinguistique, en didactique, et dans d'autres disciplines concernées) qui inspirent, au travers de démarches interactives, des procédures pédagogiques dont la validité est confirmée par l'évaluation de pratique de terrain » (Rivenc, 2003, p. 3), on sent qu'une simple définition ne va pas de soi, prise dans un contexte disciplinaire actuel mouvant, dans lequel notions et étiquettes fluctuent au fil des écoles, parfois avec fracas (Puren, 1988). En dépit de cela, tout praticien de l'enseignement, sans être méthodologue de vocation, semble devoir avoir une idée de ce que peut être la dimension méthodologique de l'enseignement/apprentissage des langues et des cultures, dimension méthodologique qui renvoie au terme « méthodologie » dans son acception la plus générale et que l'on retrouve dans le *Dictionnaire de Didactique du FLES* (Cuq, 2003, p. 166) : « un domaine de réflexion et de construction intellectuelles ainsi que tous les discours qui s'en réclament ». Voilà une acception effectivement bien générale. Notre présent propos n'est donc pas de discourir à nouveau sur ce que l'on peut et/ou doit faire de ce terme sensible, mais plutôt de mettre en avant quelques implications issues de notre étude, que l'on peut aisément ainsi qualifier de *méthodologiques*.

Etant donné la réalité du terrain didactique, nous devons cependant distinguer trois types d'implications : les unes concernent la conception d'outils didactiques (manuels, didacticiels, etc.) ; les secondes concernent les techniques de classes pouvant être utilisées indépendamment du cadre méthodologique global (quand il y en a un) ; les dernières concernent véritablement un parcours méthodologique « idéal » d'enseignement/apprentissage, destiné à des apprenants « vierges » en français et disposant des ressources institutionnelles adéquates. Nous pouvons ici nous concentrer sur le troisième type, les deux premiers pouvant en être relativement aisément extraits.

Si l'on tente de résumer très brièvement l'essentiel des points clés évoqués jusqu'à présent, relatifs à l'apprentissage de l'oral, nous pouvons mentionner :

1. Le rôle de l'écrit, en tant qu'outil de réification, de segmentation et de mémorisation de la parole pour l'apprenant.
2. Le rôle de l'attention et de l'explicite, en tant que facteurs de « prises de conscience » des phénomènes linguistiques par l'apprenant
3. Le rôle de la comparaison entre la L1 et la L2 d'une part et entre l'interlangue des sujets et la L2 d'autre part, en tant qu'activité métalinguistique naturelle et incontournable de l'apprenant.
4. Le rôle de la variabilité de l'input oral, en tant que facteur d'apprentissage de la parole pour l'apprenant.

Tous ces points ont été développés dans les paragraphes précédents, mais puisque le cœur de notre étude concerne le premier d'entre eux, c'est par celui-ci que nous devons aborder nos perspectives méthodologiques. Rappelons qu'il ne s'agit pour nous ni de traiter de l'apprentissage de l'orthographe, ni de rejeter l'orthographe du cours de langue. Il s'agit en revanche d'examiner dans quelle mesure l'utilisation de l'orthographe dans les pratiques pédagogiques actuelles pour des apprenants débutants visant à acquérir une compétence orale, d'une part ne tient pas compte des « dangers » que comporte une utilisation naïve du code écrit comme simple transcription de l'oral, et d'autre part ne met pas à profit l'utilité potentielle du code écrit, ou plus généralement visuo-graphique, pour l'apprentissage de l'oral. Notre examen personnel de ces pratiques rejoint alors le constat de Billières (2005, pp. 1-2) :

*« La problématique du passage à l'écrit est complètement évacuée par la didactique du français langue étrangère (FLE) depuis l'avènement des approches communicatives au début des années 80. Dans la période antérieure, le courant structuro-global audio-visuel (SGAV) avait beaucoup réfléchi sur la question du passage à l'écrit dans les années 60-70. [...] Actuellement, toutes les méthodes prévoient des activités de lecture dès les tous débuts de l'apprentissage linguistique. [...] De fait, l'apprenant débutant en FLE est très*



*tôt invité à mettre en œuvre ses compétences textuelles [...]. Il semble aller de soi que ces compétences de haut niveau sont naturellement transférables telles quelles si l'élève est bon lecteur dans sa langue maternelle. Ce dernier argument permet également de considérer qu'il s'accoutumera rapidement au système graphique du français. [...] On soupçonne que les apprenants dont les alphabets sont fondés sur des représentations syllabiques ou idéographiques vont rencontrer de plus grandes difficultés, mais, à notre connaissance, il n'y a guère de propositions didactiques afin de venir en aide à ces personnes. Tout au plus certaines méthodes de FLE pour débutants proposent-elles quelques activités de découverte des lettres de l'alphabet latin »<sup>313</sup>.*

L'absence de réelle analyse actuelle de la question dans le domaine didactique, tirant parti des données récentes des sciences cognitives, était précisément l'un des éléments à l'origine de notre étude, articulé en deux interrogations :

1. Pouvait-on *didactiquement* justifier le fait que les « méthodes » actuelles de FLE semblaient globalement éluder la question de l'articulation oral/écrit, et ainsi ignorer les analyses didactiques antérieures, qui avaient débouché sur du matériel à la fois méthodologiquement « global », mais aussi plus « ponctuel »<sup>314</sup>, laissant

---

<sup>313</sup> On notera que le transfert des compétences de lecteur, tout à fait discutable, évoqué par Billières est précisément une problématique d'actualité, avec par exemple les travaux de N. Akamatsu qui s'intéresse aux effets de l'orthographe en L1 sur la lecture et la reconnaissance de mots en L2. Dans l'une de ses études (2003) l'auteur compare la performance en vitesse de lecture de sujets avancés dans l'apprentissage de l'anglais langue étrangère (L2), dont les systèmes orthographiques en L1 sont différents : deux groupes non-alphabétiques (chinois et japonais) et un groupe alphabétique (perse). Il apparaît que, en raison des différences en L1, les sujets à système alphabétique étaient plus performants que les autres. Cependant, dans une autre étude (2002) utilisant les mêmes sujets, il semble que l'orthographe en L1 n'influe pas sur la procédure de reconnaissance de mot, ce qui va à l'encontre des travaux spécifiant le rôle déterminant des unités représentationnelles et de la profondeur orthographique du système en L1 sur les procédures de reconnaissance de mots écrits en L2. Les performances varient en revanche en fonction de la fréquence et de la régularité des mots en L2. En ce qui concerne cette dernière étude, il nous semble important de relever le degré d'expertise des sujets qui sont qualifiés de « fluent ESL readers ». Que ce soit dans le modèle ontogénique de Major (2002) par exemple, ou dans une simple perspective de développement de l'interlangue, nous savons bien que l'influence de la L1 évolue considérablement au cours de l'apprentissage, et il conviendrait de pouvoir mener des études identiques pour des étudiants bien moins avancés dans l'apprentissage.

<sup>314</sup> Par exemple le « manuel » de C. Lavenne, *Passage à l'écrit : orthographe et expression*, publié chez CLE international (Paris) en 1975.

penser aux enseignants que les apprenants étrangers, quelle que soit leur L1, en termes phonético-phonologiques et phonographémiques, « savaient lire » le français, et que l'oral pouvait être appréhendé à travers l'écrit sans précaution ou démarche particulière, en particulier dans le cas du français ?

2. L'utilisation du support écrit, et plus généralement du support visuo-graphique dans le cadre des outils hypermédiatisés, en classe de langue ne contenait-elle pas d'autres atouts que celle de la simple transcription, imparfaite, de la parole, pour l'apprentissage de la prononciation, et plus généralement de l'oral en français ?

Il semble à présent que nous pouvons répondre négativement à la première, et positivement à la deuxième de ces interrogations. De telles réponses permettent alors de prendre certaines orientations méthodologiques, ancrées sur les travaux de nos prédécesseurs. En effet, comme le rappelle Billières, la problématique SGAV a présenté une argumentation didactique relativement précise, et novatrice pour son époque, pour différer l'introduction de l'écrit en classe de langue<sup>315</sup>, argumentation dont une version initiale figurait dans la préface du « Livre du Maître » de la méthode *Voix et Images de France*, reproduite dans la seconde édition (Rivenc et Boudot, 1962) :

*« Nous enseignons une langue "vivante" : une langue ne vit que si elle est parlée. Nous voulons donc que, pendant les premières heures de son initiation, l'élève consacre tous ses efforts à apprendre à écouter, à imiter, et à employer aussi spontanément que possible la langue parlée familière. [...] Au bout de quelques soixante ou soixante-dix heures de cours, il pourra commencer à lire, à faire des dictées, et bientôt après à rédiger. [...] Nous avons voulu mettre en garde le professeur contre la tentation de "donner le livre" : tentation héritée d'une pédagogie classique où le livre règne parfois presque seul, et où*

---

<sup>315</sup> Nous remercions ici M.-M. Rivenc-Chiclet pour ses informations à ce sujet.

*l'absence de texte imprimé plonge maître et élèves dans un cruel désarroi, ou tout au moins inspire une certaine mauvaise conscience ».*

La question du passage de l'oral à l'écrit y est explicitement abordée dans une sous-partie intitulée « Le passage à la langue écrite » et divisée en deux temps, « Place de la langue écrite dans la méthode audiovisuelle » et « Quand et comment procéder à ce passage » (pp. 42-43 de la seconde édition) :

*« Place de la langue écrite dans la méthode audiovisuelle. [...] Mettre au début le texte sous les yeux de l'étudiant, ce n'est pas l'aider [...]. [S'il] ne connaît pas les caractères d'écriture du français, il n'a devant les yeux qu'un indéchiffrable rébus. S'il les connaît, parce qu'il connaît déjà une autre langue (maternelle ou apprise) qui utilise les mêmes caractères, il pourra faire illusion, sembler articuler des consonnes, émettre certains sons qui paraissent avoir quelques analogies avec les voyelles du français, mais qui pourrait prétendre qu'il lit ? [...] La seule vue de caractères connus déclenche en nous des réactions automatiques d'articulations. Quel étudiant pourrait se soumettre à l'énorme, constant et souvent vain effort de résister à de tels mouvements puis de les modifier ? Si nous voulons que l'étudiant articuler et émette les sons du français, il ne faut pas lui placer devant les yeux une graphie qui pour lui ne les représente pas encore, mais qui au contraire, lui a servi jusqu'alors à transcrire d'autres sons, précisément ceux-là mêmes dont on veut le détacher».*

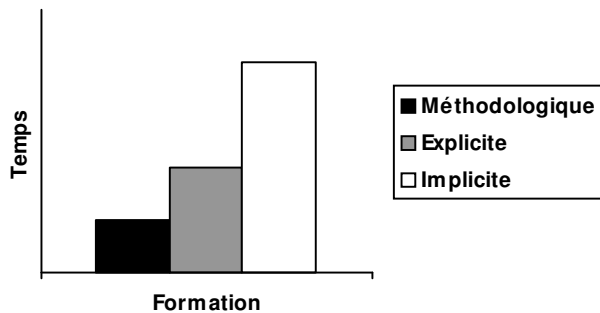
On ne peut ignorer la pertinence d'un tel propos pour notre propre étude, laquelle est bien évidemment inspirée des travaux antérieurs. Cependant, notre réflexion nous a conduit à analyser plus finement la question, à la fois en tirant parti des données récentes des sciences cognitives, et en ciblant plus précisément les enjeux de l'articulation entre oral et écrit dans l'apprentissage de l'oral. Nous pouvons alors envisager de possibles perspectives méthodologiques à nos travaux en nous inspirant, d'une part de nos propres pratiques pédagogiques, basées sur les méthodes existantes, et d'autre part sur l'ensemble des considérations rassemblées jusqu'ici. En particulier,

il apparaît qu'un parcours d'enseignement /apprentissage « idéal », destiné à un public d'adultes linguistiquement homogène en milieu hétéroglotte, comme les étudiants universitaires japonais au Japon, devrait comporter trois grands modules étroitement coordonnés, que l'on peut représenter ainsi :

<b>Module Implicite</b>	<b>Module Explicite</b>	<b>Module Méthodologique</b>
-------------------------	-------------------------	------------------------------

Si l'on considère, comme Gombert (2003) dans le cas de la lecture en L1, que les apprentissages explicite et implicite se renforcent (« boost ») mutuellement, notamment, dans notre cas, *via* les facteurs d'attention pour le premier et de fréquence pour le second, il est alors possible de se dégager, dans une certaine mesure, de l'opposition délibérément binaire de Krashen entre *apprentissage* et *acquisition*, puisque, comme le souligne Meulemans (1998, p. 201), « il n'existe pas de tâche "pure" et [...], dès lors, il est vain de chercher à caractériser une situation donnée comme étant par exemple de nature soit implicite, soit explicite » (Meulemans, 1998, p. 201).

L'importance respective de chacun de ces trois modules dans la formation varierait naturellement au cours du temps, les apprenants devant, à terme, pouvoir effectivement évoluer uniquement en L2, et accomplir les tâches didactiques qui sont aujourd'hui proposées par les approches communicatives ou encore par les tenants de l'apprentissage par l'action. En effet, ce qui semble parfois aujourd'hui manquer, c'est bien la *préparation* à une telle étape, d'où la schématisation suivante :



Nous pouvons alors, très brièvement, esquisser les grandes lignes de ces modules.

#### III.4.1.1. Module méthodologique

De la même manière que nous avons évoqué l'utilité potentielle d'une réflexion métalinguistique, correspondant à l'activité métalinguistique des apprenants, pour l'apprentissage explicite de la L2, il nous semble qu'une réflexion « métapsycholinguistique », conduite en L1 et en des termes appropriés, en début d'apprentissage au moins, serait indispensable à la mise en place réussie des deux autres modules. Dans le domaine didactique, cela évoque les notions de « apprendre à apprendre » et de « contrat d'apprentissage ». Il s'agit d'informer les étudiants : 1) sur les modes et les stratégies d'apprentissage (perception et compréhension orale et écrite, fréquence et mémorisation, interlangue et progression, etc.) ; 2) sur la méthodologie adoptée (articulation entre implicite et explicite, entre oral et écrit, entre la L1 et la L2, etc.). Le point 1) vise évidemment à soutenir le point 2), qui doit en effet s'accomoder des attentes, des habitudes et des représentations didactiques préalables des apprenants, lesquelles peuvent entraver les effets bénéfiques de la méthodologie en question. Cela renvoie, d'une manière très différente cependant, à ce que certains verbo-tonalistes nomment la « Leçon zéro » en correction phonétique, et qui consiste à expliquer, d'une manière très épurée, les principes méthodologiques adoptés pendant la classe, en vue de remporter l'adhésion, et ainsi la participation, des apprenants. Une telle « formation » à l'apprentissage des langues permettraient ainsi aux apprenants de mieux comprendre et gérer leur apprentissage, et, par exemple, de ne pas « plonger sur

la transcription » des enregistrements figurant souvent en fin de manuel lorsque l'enseignant propose une tâche de compréhension orale, après avoir pourtant demandé aux étudiants de « fermer leur livre », pour des raisons didactiques très claires au spécialiste, mais pas à l'apprenant novice en la matière. Permettre aux apprenants de tenter d'exercer un certain contrôle, dans une certaine mesure, sur leur apprentissage, et en particulier sur leurs « automatismes » didactiques et/ou cognitifs, tel serait l'un des objectifs de ce module, qui serait également tout à fait profitable aux dimensions socioculturelles et pragmatiques de l'apprentissage, puisque, avant d'aborder, de manière explicite et implicite, ces dimensions dans les domaines de la L1 et de la L2, il permettrait d'amorcer chez les étudiants une approche raisonnée des notions de culture, de civilisation, et surtout de la manière dont ces facteurs socioculturels peuvent influencer l'apprentissage et la communication en langue étrangère. Enfin, ce module, qui devrait entièrement être conduit en L1 ou dans une langue familière aux apprenants, vise très spécifiquement à remporter l'adhésion des apprenants<sup>316</sup> pour la mise en œuvre des deux autres modules, qui constituent évidemment le cœur de la formation.

#### III.4.1.2. Module explicite

Comme nous l'avons vu, un « enseignement explicite » n'est en aucun cas synonyme d'« enseignement formel ». Il s'agit plutôt d'un enseignement destiné à faciliter l'activité métalinguistique des apprenants, généralement axée sur la forme (« focus sur la forme »), mais également sur les connexions entre formes et significations (au sens large), en orientant *l'attention* des apprenants pour mener un apprentissage explicite. Cette activité métalinguistique étant, dans un premier temps essentiellement de nature contrastive entre la L1 et la L2, puis entre l'interlangue et la L2<sup>317</sup>, un tel enseignement

---

<sup>316</sup> Ce point est particulièrement important dans le cas de la prononciation, étant donné, notamment, la dimension sociale et identitaire de celle-ci : rien ne sert en effet de vouloir « corriger » la prononciation du /R/ d'un apprenant japonais, si celui-ci estime, pour diverses raisons, que le [r] qu'il utilise lui suffit pour atteindre la capacité communicative et le niveau de compétence qu'il vise. Les modules « méthodologique » et « explicite » devraient précisément permettre à l'enseignant de pouvoir aborder ces problèmes, et, peut-être, modifier les objectifs initiaux des apprenants, à condition qu'une telle modification soit didactiquement justifiable.

<sup>317</sup> A cet égard, il est intéressant de constater que certains outils multimédiatisés récents proposent précisément d'intégrer les productions des apprenants dans leur système, même si les justifications et les

viserait donc à donner aux apprenants les moyens d'optimiser leur activité métalinguistique. Selon les domaines et les points abordés, il pourrait être plus ou moins « formel », allant de la description métalinguistique « théorique », à de simples observations de corpus, en passant par des exercices de conceptualisation. Dans le cas de la prononciation et de la phonographémie, il viserait dans un premier temps à établir, en collaboration avec les apprenants, ce que nous-mêmes avons fait fait jusqu'ici, à savoir, une analyse contrastive des systèmes en présence et *surtout ses conséquences pour l'apprentissage*, avant, dans un deuxième temps, de se concentrer sur les spécificités du français. L'adaptation au public concerné (système d'écriture, système linguistique, voire représentations métalinguistiques préalables) semble ici essentielle. Un tel module, conduit en L1 ou dans une langue familière aux apprenants, devrait cependant être étroitement coordonné au troisième module, qui est évidemment, nous semble-t-il, le plus central dans la formation, sans quoi l'ensemble perdrait de sa cohérence et de son efficacité potentielle : les éléments ainsi « analysés » devraient en effet correspondre, préalablement ou ultérieurement, aux éléments introduits et « pratiqués » dans le module d'apprentissage implicite.

#### III.4.1.3. Module implicite

Comme nous l'avons vu, l'apprentissage implicite n'implique en aucun cas un « enseignement implicite », bien que les deux puissent souvent être liés. Sur ce point, nous ne prétendons pas apporter d'innovations particulières : depuis l'époque du SGAV, jusqu'à l'apprentissage par l'action en passant par les approches communicatives et toute autre approche privilégiant, d'une manière ou d'une autre, l'

---

objectifs d'une telle intégration ne correspondent pas forcément exactement à ceux que nous évoquons ici. On peut ainsi citer, pour le domaine prosodique, le système « WinPitch LTL » de Germain et Martin (2000, p.14) qui permet à l'apprenant de devenir son propre modèle grâce à une fonction de « synthèse » :

*« La fonction de synthèse disponible dans le programme permet à l'enseignant, à partir d'une production fautive, de redessiner le contour mélodique d'un énoncé ou de redéfinir la durée de tel ou tel segment pour permettre à l'apprenant de visualiser et d'entendre, grâce à la synthèse, le patron intonatif correct avec sa propre voix ».*

« action », que celle-ci soit plutôt orientée « haut-bas », *via* les processus interprétatifs et l'élaboration d'hypothèses, ou plutôt « bas-haut », *via* une automatisation des processus de bas niveau, toutes parviennent, à des degrés divers, à « faire apprendre » la L2 aux apprenants. Loin de nous l'idée d'une quelconque « panacée » des langues. Cependant, pour le public débutant qui nous intéresse, et dans l'optique de l'apprentissage de *l'oral*, nous pouvons néanmoins nous inspirer de nos propres pratiques pédagogiques, notamment en comparant notre utilisation des « manuels » et autre documents pédagogiques actuels disponibles sur le marché, avec celle de méthodes que nous pourrions qualifier de « multimédia », en l'occurrence certaines méthodes SGAV (*Voix et Images de France, De Vive Voix, Leçons de Transition*) ainsi qu'une méthode informatisée élaborée à l'Université Nationale des Langues et des Civilisations Etrangères de Tokyo (Japon). Il va en effet de soi que les multiples activités didactiques envisageables, recensées dans la littérature (Pendanx, 1998) et pratiquées sur le terrain par les enseignants peuvent être bénéfiques à l'apprentissage implicite de la langue. Nous sommes donc ici obligés de nous concentrer sur notre problématique, à savoir celle de l'articulation de l'oral et de l'écrit pour l'apprentissage de l'oral, en constatant que, entre autres, c'est cette articulation qui semble bien souvent faire défaut aux « méthodes » actuelles destinées aux apprenants débutants.

#### **III.4.2. Parcours d'apprentissage de l'oral**

Etant donné ce que nous avons présenté à propos de l'oral en tant que médium fondamentalement *variable* (dans le temps et dans ses caractéristiques intrinsèques) par opposition à l'écrit en tant que médium essentiellement *fixe*, il apparaît que la tâche du sujet apprenant l'oral (au moins en perception, étape préalable à toute production) est bien de « saisir » cette variabilité, l'écrit pouvant alors lui servir en tant qu'outil de *réification* et de *normalisation*. Il s'agit donc bien de « délimiter » puis de « fixer » la variation, et non le contraire. Dans cette perspective, l'écrit ne peut ni remplacer, ni précéder l'oral. En revanche, il peut servir d'outil d'appréhension de l'oral, comme nous l'avons exposé tout au long de cette partie. A partir de ce constat, il est possible, en perception/compréhension de l'oral, de suggérer, trois principes, que nous suivons nous-mêmes, lorsque cela est possible, dans nos classes :



- **L'oral avant l'écrit** : pour chaque séquence (et pas seulement du point de vue de la progression globale, quoique l'on puisse envisager une courte période initiale en début d'apprentissage consacrée à l'oral et à l'étude du système sonore du français), il conviendrait de travailler dans un premier temps l'oral (compréhension globale, répétition, compréhension détaillée, questions/réponses, etc.), avant de travailler, ensuite, sur la forme écrite des documents sonores, afin d'effectuer un véritable travail de perception orale. Concrètement, les documents oraux doivent systématiquement être abordés avant leur transcription écrite, ce qui remet en cause, sur ce point précis, la conception de nombreux manuels.
- **Le séquençage, la répétition et la variation de l'oral** : étant donné les limitations de la mémoire de travail d'une part, et la nécessité de préparer les apprenants à la variation de l'oral d'autre part, chaque séquence sonore devrait être : 1) segmentée de manière adéquate, variable en fonction des objectifs d'enseignement ; 2) répétée plusieurs fois successivement, de manière à éventuellement permettre le rafraîchissement de l'information dans la boucle phonologique ; 3) être énoncée successivement par différents locuteurs, c'est-à-dire, en pratique, au minimum par l'enregistrement du matériel pédagogique et par l'enseignant.

Le séquençage (1) de l'input oral concerne à la fois la compréhension et la mémorisation, mais il peut également concerner les aspects morphologiques et phonético-phonologiques, afin de développer, chez les apprenants une compétence de micro-segmentation morphémique, syllabique et phonémique, utile, comme nous l'avons vu, en lecture, mais également à l'identification morpholexicale.

La répétition (2) (et le séquençage) de l'input oral, qui concerne naturellement la mémorisation, vise à établir un parallèle entre les saccades régressives opérées en lecture, et le rafraîchissement de l'information auditive, également utile à la vérification des informations. La difficulté de l'oral est en effet étroitement liée à sa

gestion temporelle, un problème qui pourrait peut-être être résolu, notamment de deux manières : soit par un « ralentissement » de l'input oral (l'enseignant répète « plus lentement » la séquence sonore de l'enregistrement), soit par l'utilisation de saccades régressives auditives (répétition successives de brèves séquences).

La variation (3) (couplée au séquençage et à la répétition) de l'input oral vise à fournir plusieurs indices phonétiques pour une même représentation phonologique, et ainsi à faciliter l'accès à cette représentation, sans pour autant la réduire à une unique forme phonétique.

- **La fixation de l'oral dans l'écrit :** Lors de l'introduction de la forme écrite du document sonore, il serait nécessaire de consacrer un certain temps au repérage des correspondances entre les deux inputs (faire ré-écouter le document sonore en lisant le texte, faire lire à voix haute le document écrit à la suite de l'enseignant, faire ré-écouter ensuite le document sonore, sans le texte, dicter le document sonore, etc.), avant de travailler isolément sur le texte. Durant ce travail, il faudrait attirer explicitement l'attention des apprenants, en fonction de leur niveau, sur les écarts séparant la forme écrite de la forme orale des séquences (liaisons, élisions, prosodie, lettres « muettes », etc.), en particulier lors du travail de lecture à voix haute. Le deuxième principe doit être étroitement couplé à celui-ci, de manière à multiplier, dans un temps relativement bref, les tâches de mise en correspondance des inputs phonétiques, des représentations orthographiques et des formes phonologiques auxquelles ils renvoient tous deux. Cela implique donc d'exposer les apprenants, de manière conjointe, à une forme écrite fixe parallèlement à un input phonétique multiple, dans le temps (répétitions successives) et dans la forme (variation phonétique de l'input).

Ces trois principes balisent ainsi le parcours suivant :

1. Présenter les formes *orales* et essayer de les « fixer », en tant que formes phonétiques, à l'aide de plusieurs tâches, selon le schéma : perception (écoutes

- multiples) / répétition (mise en œuvre audio-articulatoire) / production (rappel de ces formes, à travers différents exercices).
2. « Fixer » ensuite les représentations phonologiques auxquelles elles renvoient dans leurs formes orthographiques, à l'aide de plusieurs tâches, selon le schéma : perception (lecture à voix haute, puis silencieuse) / production guidée (dictée, puis exercices) / rappel (mémorisation des séquences et « mise en scène » dans le cas d'un dialogue, afin de solliciter la récupération des gestes articulatoires en coarticulation).
  3. Réassembler finalement l'oral et l'écrit en bimodalité (relecture/réécoute), afin de tenter de réaliser un pluricodage en mémoire, selon un format phonétique et orthographique, tous deux renvoyant à une représentation phonologique unique.

Il va de soi qu'un tel parcours est particulièrement adapté aux documents oraux de type dialogique ou conversationnel, et que la perspective communicative qui lui est assignée implique d'effectuer ce travail en privilégiant les démarches d'accès au sens. Le pluricodage mnésique que nous évoquons, est naturellement centré autour de la valeur sémantique de l'item en question. Par ailleurs, tout au long de ce parcours, l'enseignant doit intervenir pour « rectifier » les performances erronées des apprenants, de manière à éviter l'encodage de formats « erronés ».

#### III.4.2.1. Utilisation des « méthodes » de FLE

Lorsque l'on examine, à la lueur de ce parcours et de ces principes, certaines de nos pratiques de classes, il semble que les méthodes que nous avons qualifiées de « multimédia » précédemment y correspondent assez bien, puisque la démarche que nous suivons, bien qu'elle ne corresponde pas exactement toujours à celle proposée par les auteurs (Rivenc-Chiclet, 1989, 2001, 2003), est globalement la suivante<sup>318</sup> :

---

<sup>318</sup> Nous ne rentrerons ici ni dans les discussions possibles concernant le détail des torts et des mérites de telles méthodes, car ce n'est pas ici notre objectif, ni dans le détail des justifications méthodologiques de tous les aspects du matériel (par exemple la fonction précise de l'image en fonction de sa nature).

A) Exemple d'utilisation personnelle des méthodes SGAV :

- 1) Présentation répétée des séquences sonores (de type « dialogue »), couplées à un support pictural, qui est présenté légèrement avant le son, de manière à « contextualiser » la séquence sonore. Une telle présentation vise donc un double encodage : image/son. L'enseignant répète également la séquence sonore plusieurs fois, de manière à fournir un input phonétique différent, voire adapté aux apprenants, *via* des procédures de modification du stimulus (ralentissement, segmentations de différentes natures, accentuation, etc.).
- 2) Répétition à voix haute individuelle des séquences par les apprenants, avec correction phonétique éventuelle par l'enseignant. Tout au long de ces deux étapes, l'essentiel du contenu linguistique, en particulier sémantique, mais aussi morphosyntaxique, est « expliqué » par l'enseignant à travers des interactions orales<sup>319</sup> et gestuelles avec les apprenants.
- 3) Distribution de la transcription du dialogue et des images correspondantes (image/texte) et ré-écoute des séquences sonores en « lisant » le texte (encodage son/texte).
- 4) Lecture à voix haute par l'enseignant et les apprenants (encodage texte/son).
- 5) Mémorisation des séquences à l'aide du texte et des images (encodage texte/image)<sup>320</sup> et « mise en scène » (encodage articulatoire à partir du texte, puis tâche de *rappel* « en situation » à l'aide des indices de rappel que constituent les images). Durant cette phase, les interactions entre apprenants permettent un encodage, en production, que nous pourrions qualifier d' « actionnel » (macro- et micro-gestualité) et d' « émotionnel » (vis-à-vis de l'importance des émotions dans l'encodage en mémoire).

---

<sup>319</sup> *Via* des questions ouvertes ("Que se passe-t-il?"), partielles ("Qui...?") ou fermée ("Est-ce que...?"), des commentaires et des reformulations.

<sup>320</sup> Il serait souhaitable, pour cette étape, de développer un format de présentation intégrant spatialement le texte et l'image, de manière à respecter le « principe de contiguïté » de Mayer (Lemerrier et Tricot, en soumission). On remarquera en revanche que la présentation auditive du matériel verbal lié à l'image correspond bien au mode « duel » d'instruction, qui serait davantage bénéfique à la compréhension lorsque le matériel verbal est présenté auditivement (plutôt que visuellement), un résultat établi par des chercheurs comme Sweller ou Mayer par exemple (Tricot, 1998 ; Lemerrier et Tricot, en soumission).

- 6) Ré-écoute des séquences sonores en « lisant » le texte (tâche de *reconnaissance* bimodale, en termes auditifs, mais aussi articulatoires, suite à la tâche de rappel, et dans la perspective des théories motrices de la parole).
- 7) Dictée des séquences sonores (tâche de *reconnaissance auditive* et de *rappel orthographique* : encodage en production son/texte), suivie d'une vérification immédiate par les apprenants (et l'enseignant) (auto-correction suivie par l'enseignant visant à favoriser la perception et l'évaluation *immédiate*<sup>321</sup> des écarts entre leur production et la cible visée). Cette tâche permet également ensuite de travailler sur les informations morphologiques et orthographiques notamment.
- 8) Ré-écoute des séquences sonores (tâche de *reconnaissance auditive*).

Ce parcours, qui est ensuite suivi d'activités de production orale plus « libres », ainsi que d'activités « grammaticales » supplémentaires, peut donc être schématisé ainsi, offrant un passage de l'écoute à la production, suivi d'un retour à l'écoute : son/image → son/image/texte → image/texte → texte → texte/son → son<sup>322</sup>. Il faut également

---

<sup>321</sup> Ce que ne permettrait pas une correction différée.

<sup>322</sup> Nous serions bien sûr tentés d'évoquer, d'un point de vue neuropsycholinguistique, la latéralisation progressive droite-gauche, au niveau cérébral et dans une perspective développementale, du traitement des stimuli linguistiques en fonction de l'expertise des sujets (appréhension globale de stimuli nouveaux à droite (parole continue et image), puis expertise analytique à gauche (parole segmentée et texte)) en référence notamment aux travaux de Goldberg et Costa (1981), mais cela ne serait pas certainement raisonnable d'un point de vue scientifique, en particulier à une époque où de nombreux spécialistes en sciences humaines semblent, en vue de donner une certaine assise à leurs propres réflexions, vouloir interpréter les données des neurosciences de manière fort hâtive, parfois naïve, et souvent abusivement simplificatrice. On peut toutefois, eu égard aux thèmes qui nous occupent ici, mentionner l'étude neuropsycholinguistique de Zhang, Kuhl, Imada, Iverson, Pruitt, Kotani et Stevens (2000) qui, à l'aide de la magnétoencéphalographie ont montré que, suite à l'entraînement d'un auditeur japonais dans la discrimination auditive des phonèmes anglais /l/ et /r/, on pouvait observer chez celui-ci un déplacement droite-gauche d'activités corticales dans l'établissement des deux catégories phonémiques (« a right-to-left hemispheric shift of cortical MMF [mismatch field] activities in the establishment of linguistic /l/ and /r/ categories »). Avant l'entraînement, l'hémisphère droit semblait davantage impliqué dans la détection des différences acoustiques, alors que c'est l'hémisphère gauche qui semblait dominer ensuite. Ce schéma diffère selon la nature des stimuli linguistiques (pleins ou tronqués), les stimuli tronqués étant perçus comme non-linguistiques par le sujet, ce qui rejoint la différence de traitement relevée par Celsis, Doyon, Boulanouar et Nespoulous (1997), selon que le stimulus est perçu comme linguistique (phonème) ou non (son non-linguistique). L'étude de Zhang, Kuhl, Imada, Iverson, Pruitt, Kotani et Stevens (2000) concerne donc le rôle de l'entraînement dans l'apprentissage de la prononciation, et ainsi l'augmentation de l'expertise linguistique du sujet, qui conduit à une plus grande activité analytique (que nous associons ici au texte vis-à-vis de l'image). Comme ils l'indiquent en conclusion:

noter que le recours à l'obscurité qui accompagne l'étape de présentation, outre l'intérêt (en particulier pour un public japonais), ou parfois les inconvénients, psychoaffectifs qu'il comprend, permet de focaliser l'attention des apprenants sur l'image puis sur l'input auditif. La dimension oro-faciale de la parole, quant à elle, est sollicitée dans les phases ultérieures.

#### B) Exemple d'utilisation personnelle d'une méthode multimédia franco-japonaise :

Nous ne rentrerons pas ici dans la présentation détaillée de cette méthode<sup>323</sup>. Il suffira de dire que le logiciel comprend : 1) de courts sketches vidéo, linguistiquement axés sur des actes de parole et 2) des fonctionnalités multiples (segmentation de la séquence audio et vidéo, transcription orthographique, traduction en japonais, etc.). Il est utilisé en classe à l'aide d'un projecteur et de matériel adapté, renvoyant ainsi, sur certains aspects, aux méthodes SGAV précédemment citées. L'une des démarches possibles est alors la suivante :

1. Présentation de la vidéo « muette » (en partie ou en totalité), accompagné d'un bref descriptif de la situation en japonais. Il s'agit pour les étudiants de se « mettre en situation » et d'émettre des hypothèses sur le contenu linguistique de la séquence, en japonais pour des débutants complets (1<sup>ère</sup> année d'étude), ou en français pour des non-débutants (2<sup>ème</sup> année d'étude).
2. Présentation de la vidéo « sonore » en séquences segmentées. Cette étape permet de travailler sur la répétition individuelle des étudiants (et donc la correction phonétique) et sur la compréhension orale, à travers les interactions avec l'enseignant.
3. Pour les étudiants débutants, présentation de la traduction en japonais (écrite) des séquences, afin de vérifier la compréhension globale.

---

*« Earlier we reported that an American control'MMFs showed a left hemispheric dominance for /l/ and /r/ whereas the Japanese subjects showed bilateral involvement. Training appeared to lead to more linguistic analysis of the speech stimuli in the left hemisphere ».*

<sup>323</sup> Accessible, en mars 2005, à l'adresse suivante:  
<http://www.coelang.tufs.ac.jp/modules/fr/dmod/index.html>.

4. Présentation en bimodalité de la vidéo « sonore » et des transcriptions orthographiques françaises.
5. Lecture à voix haute des transcriptions, après écoute de chaque séquence sonore.
6. Dictée du texte, suivie d'une vérification, puis d'une nouvelle lecture à voix haute.
7. En binômes, lecture à voix haute et mémorisation des dialogues, suivies par l'enseignant pour d'éventuelles corrections phonétiques.
8. « Mise en scène » des dialogues.

Ce parcours est ensuite suivi d'activités d'expression orale et écrite plus « libres ».

En anticipant quelque peu sur notre conclusion, à la vue de ces parcours, nous serions tentés d'affirmer que, *sur les points qui nous concernent ici*, ce qui semblait parfois manquer aux méthodes SGAV est à présent revisité (apprentissage explicite ; contrastivisme ; créativité et production linguistiques, etc.), tandis que ce qui semble manquer aux approches actuelles figurait, en partie au moins, dans la problématique SGAV (méthodologie cohérente ; articulation de l'oral et de l'écrit pour l'apprentissage de l'oral, intégration de la correction phonétique, importance de la fréquence d'exposition à la langue, etc.). Il semblerait donc que l'heure soit à la synthèse et au rééquilibrage.

Outre ces deux grands parcours, il nous faut également mentionner la portée des résultats rassemblés jusqu'ici pour d'autres aspects de l'approche méthodologique suggérée : 1) le format des « exercices » grammaticaux, 2) la pratique de la lecture à voix haute, 3) l'utilisation de l'API et enfin 4) l'indispensable complémentarité de la correction phonétique et de la correction phonographémique.

#### III.4.2.2. Exercices, Lecture et API

Avant de traiter d'une des implications essentielles de notre étude, à savoir l'importance de la correction phonographémique, nous pouvons formuler quelques

brèves recommandations concernant les « exercices », la lecture et l'utilisation de l'API, pour l'apprentissage de l'oral.

#### III.4.2.2.1. Les « exercices »

Si l'on ne peut, pour diverses raisons, en particulier matérielles et institutionnelles, mettre en œuvre un programme « idéal » d'enseignement/apprentissage de l'oral, on peut en revanche tirer d'autres implications des principes présentés précédemment. En ce qui concerne les « exercices de grammaire » par exemple, quelle que soit leur catégorie méthodologique (déductif, inductif, etc.), il semblerait essentiel de développer, parallèlement au matériel écrit qui existe à l'heure actuelle, davantage de matériel oral, puisque, comme nous l'avons vu avec l'ouvrage de Battye, Hintze et Rowlett (2000) par exemple, la grammaire de l'écrit et la grammaire de l'oral peuvent parfois sensiblement différer, par exemple dans le cas de la morphologie verbale, mais également de phénomènes comme les liaisons. En outre, le développement de ce type de matériel, répondrait à la suggestion de Toma (2000) de développer l'entraînement à la gestion du temps limité<sup>324</sup>. Si l'on prend en compte cette dimension temporelle, et la nature réflexive de telles activités, on pourrait alors proposer le parcours suivant :

1. Travail à l'écrit (temps libre)
2. Travail à l'oral avec texte (temps limité mais support orthographique)
3. Travail à l'oral (temps limité)
4. Travail à l'oral avec le texte (vérification).

#### III.4.2.2.2. La lecture à voix haute

Comme nous l'avons indiqué, l'oral impose un rythme d'écoute à l'auditeur, alors que l'écrit laisse davantage de liberté au lecteur. Or, si l'on souhaite que les apprenants puissent lire à voix haute un texte, en respectant les caractéristiques prosodiques, en particulier rythmiques, de l'énoncé oral qu'il transcrit, il conviendrait ici encore

---

<sup>324</sup> Ces deux aspects avaient été traités dans les méthodes SGAV *via* les exercices de laboratoire, parfois abusivement qualifiés de « structuraux » par leurs détracteurs, malgré leur contextualisation et leurs atouts. Ces exercices, de qualité variable, avaient bien évidemment aussi leurs défauts, particulièrement sensibles lorsqu'ils étaient isolés du cadre méthodologique dans lequel ils avaient initialement été conçus.



d'entraîner les apprenants à la gestion du temps limité en lecture. Il s'agirait alors d'imposer un rythme de lecture, de plus en plus rapide, aux apprenants, qui leur permettraient de respecter ces caractéristiques prosodiques (en particulier au niveau de la segmentation en groupes intonatifs et rythmiques). Il conviendrait pour cela d'entraîner à la lecture en bimodalité, soit synchrone (avec des indices auditifs, en premier lieu la forme orale de l'énoncé, selon différents rythmes), soit asynchrone (la lecture par l'apprenant suivant alors la forme sonore de la séquence, le temps de lecture disponible étant calqué sur celui du modèle, comme le proposait Toma (2000).

#### III.4.2.2.3. L'utilisation de l'Alphabet Phonétique International (API)

Si l'on considère que l'apprentissage des correspondances phonographémiques du français, dans la perspective qui est la nôtre, vise à « fixer » de nouvelles valeurs phonético-phonologiques sur des représentations graphiques plus ou moins nouvelles<sup>325</sup> selon les apprenants, on peut également envisager de recourir à l'API, en tant qu'*outil de conscientisation*, voire de conceptualisation, des relations phonographémiques, destiné non pas tant à l'apprentissage de la lecture qu'à celle du système oral de la L2. Un tel outil serait donc à inclure dans le module que nous avons qualifié d' « explicite », et il s'agirait alors d'exploiter le support visuo-graphique que celui-ci constitue pour mieux appréhender le fonctionnement de l'oral, à la fois en termes de structures (en particulier syllabes et phonèmes), mais également en termes de contenus (en particulier segmental et infrasegmental, ainsi que les phénomènes liés à la coarticulation). Couplé à une approche contrastiviste, un parcours possible serait alors le suivant :

1. Apprentissage de l'API pour la transcription de la L1.
2. Apprentissage en bimodalité (API/audio) de l'API pour la transcription de la L2.
3. Activités contrastives, réflexives et pratiques, entre la L1 et la L2 en bimodalité (API/audio).

---

<sup>325</sup> Ce qui évoque le système de baguettes colorées de la méthode *Silent Way*, conduisant, de manière très schématique, les apprenants à associer un son à une couleur.

4. Effectuer ensuite un « passage » non plus de « l'oral à l'écrit », mais de l'API au système orthographique français.

Il conviendrait alors, au vu de ce que l'on sait des processus de lecture, de travailler sur des unités phonographémiques de plusieurs tailles (graphèmes, corps, syllabes, mot), quitte à proposer, pour les unités de taille supérieure au segment, une symbolisation supplémentaire, comparable au système d'ellipses que nous avons utilisé dans nos tests C2, D et E (comptage syllabique), qui pourrait être couplée à l'utilisation de l'API. Si ce travail constituerait une certaine tentative d'utilisation du support visuo-graphique pour l'apprentissage du système phonético-phonologique, nous avons néanmoins noté précédemment les faiblesses de l'API, dans la perspective qui est la nôtre, en particulier en raison du principe « latin » (Durand, 2005), qui renvoie au problème des « interférences » phonographémiques entre la L1 et la L2. Ce dernier point renvoie alors à l'un des aspects que nous n'avons cessé de souligner au cours de notre réflexion, à savoir la nécessité de coupler au travail audio-oral d'intégration phonético-phonologique de la L2 (notamment *via* un travail de correction phonétique) un travail audio-graphique d'intégration phonographémique de la L2 (notamment *via* un travail de correction phonographémique).

#### III.4.2.3. Correction phonétique et correction phonographémique

Comme cela apparaît à présent nettement, il semblerait que, au travail de « correction phonétique », doit être impérativement *intégré* un travail de « correction phonographémique », sans quoi le premier ne pourrait que souffrir de l'utilisation massive du support écrit dès les premières heures d'apprentissage. Cette intégration peut prendre trois formes :

##### III.4.2.3.1. Utilisation du support visuo-graphique lors de la correction phonétique

Deux principales opérations correctives peuvent ici être envisagées : a) identification/discrimination ; b) segmentation.

#### III.4.2.3.1.1. Identification/discrimination

Si l'on se concentre dans un premier temps sur le support orthographique (et non visuo-graphique étant donné l'ampleur des possibilités), on peut en effet imaginer que, selon le principe de diagnostic/remédiation de la méthode verbo-tonale d'une part, qui propose de se baser sur la production de l'apprenant pour proposer un input correctif de remédiation, et selon l'idée que les apprenants doivent (explicitement) percevoir l'écart entre leur production et la cible d'autre part, on peut alors imaginer que cet input correctif, au lieu d'être auditif, puisse être orthographique, voire bimodal audio-orthographique. Il s'agit alors de faire prendre conscience à l'apprenant que sa production est « erronée » phonologiquement (car la « liberté » phonétique est évidemment plus tolérable que la « liberté » phonologique qui, elle, conduit à des commutations dommageables pour la communication), *via* une *réification graphémique des catégories phonologiques*<sup>326</sup>.

Il s'agit ainsi d'associer à chaque production d'apprenant une catégorie graphémique. Si la production est erronée, la catégorie graphémique le sera également. La perception de l'écart pourra ainsi être visuelle, si celui-ci n'est pas suffisamment perçu auditivement (notamment en raison de la fugacité de l'oral), mais, contrairement à un système symbolique arbitraire comme celui des couleurs de la méthode *Silent Way*, cette procédure permettrait d'intégrer le système graphémique du français au travail de correction phonético-phonologique, de manière à ne pas, ensuite, lors de la rencontre avec l'écrit, perdre les bénéfices d'un tel travail<sup>327</sup>. Il reviendrait donc à l'enseignant d'évaluer la production phonétique des apprenants et d'y assigner une catégorie graphémique, en indiquant simultanément la catégorie graphémique à atteindre.

Se pose alors ici le problème de l'inventaire graphémique, étant donné l'opacité relative de l'orthographe française et les possibilités d'homophonies. Pour cela, les

---

<sup>326</sup> Car comme le note Coulmas (1989, p. 45) : « *Writing systems are only rarely the result of conscious linguistic analysis, yet they are the expression and materialization of linguistic consciousness* ».

<sup>327</sup> Correspondant à l'expression populaire « reculer pour mieux sauter ».

travaux de Catach (1995), d'une part, et de Lange et Content (1998, 1999) d'autre part, pourraient être sollicités, en ne sélectionnant, dans un premier temps, que les graphèmes les plus fréquents (en cas d'homophonie), permettant de distinguer l'ensemble des phonèmes du français. L'apprentissage des autres correspondances phonographémiques, moins fréquentes, relèverait, d'une part de l'apprentissage de l'orthographe du français<sup>328</sup>, et d'autre part de la correction phonographémique à proprement parler, qui correspond aux deux autres formes d'intégration de la correction phonétique et phonographémique que nous évoquions, et que nous présentons plus bas. Il s'agirait donc ici essentiellement dans un premier temps de correction phonologique (catégories) et phonographémique (correspondances), comme nous pouvons l'illustrer à l'aide de l'exemple hypothétique suivant :

En tâche de répétition d'un item lexical ou d'un court énoncé, l'apprenant répète à la suite de l'enseignant. En cas de production erronée d'un segment, l'enseignant indique, répète le modèle initial en écrivant le graphème correspondant au tableau (par exemple <R>). L'apprenant répète à nouveau, et, selon le type d'erreur, l'enseignant trace une ligne aux extrémités de laquelle se trouvent le graphème visé (<R>) et le graphème *pouvant* correspondre en français à la production erronée (par exemple <l>). Cette ligne peut en outre être « graduée », et, à chaque production de l'apprenant, orientée par les modèles auditifs successifs de l'enseignant, ce dernier indique une position sur cette ligne, plus ou moins rapprochée du graphème auquel la production correspond. Cela permettrait à l'apprenant de visualiser, pour chaque production, la « distance » qui la sépare de la cible visée, et ainsi, par approximations audio-articulatoires successives de tenter de s'en « rapprocher ».

On pourrait bien sûr rétorquer qu'une telle visualisation « naïve » correspond à ce que l'on trouve dans les didacticiels proposant des « graphiques » de toutes sortes, issus

---

<sup>328</sup> A ce propos, Catach (1995, p. 94) indique : « Une écriture comme la nôtre peut être acquise de plusieurs façons, la plupart du temps complémentaires : la mémoire, la répétition (l'« usage ») ; la référence à l'étymologie ou à l'histoire [...] ; enfin, et c'est ce que nous préconisons, par référence constante à la langue actuelle (phonologie, mais aussi morphologie, syntaxe, lexic) ».

d'analyses acoustiques certainement plus « scientifiques » que le jugement intuitif de l'enseignant. A cela, il est facile de répondre : 1) que ces graphiques ignorent totalement la dimension phonographémique que nous abordons ici, 2) que ces graphiques ne peuvent concrètement être sollicités, de manière rapide et adaptée, en classe de langue « standard », 3) que ces graphiques sont beaucoup moins facilement *interprétables* que ne l'est la visualisation proposée, puisque, comme l'indique Cazade (1999), une interprétation efficace des ces « courbes » par les apprenants nécessiterait une formation en phonétique assez solide, qu'ils n'ont généralement bien évidemment pas. La difficulté posée par l'interprétabilité de ces visualisations (qu'il s'agisse d'analyses d'amplitude, de durée, de fréquence ou d'intonation) pour des apprenants non-spécialistes est soulignée par de nombreux chercheurs dans le domaine, comme par exemple Neri, Cucchiarini et Strik (2002, p. 181) :

*« While the level of detail with this type of feedback can be very high, the effectiveness of these systems is also questionable, as students will have a hard time deciphering these displays and even expert phoneticians may find it difficult to extract the information needed to improve pronunciation ».*

Cela nous renvoie pleinement à la distinction entre modèle linguistique et modèle didactique : l'enseignant doit être en effet suffisamment formé en phonétique et phonologie pour prendre les bonnes décisions, mais ce qu'il propose visuellement aux apprenants ne vise qu'à fournir des indices visuels (en accord avec le modèle de Massaro) pour baliser le processus d'apprentissage audio-phonatoire, *via* un « tâtonnement » articulo-auditif, éventuellement guidé, auditivement, gestuellement ou déclarativement, par l'enseignant, jusqu'à la réalisation de la cible visée, signalée par le retour correctif visuel positif de l'enseignant. Une telle « spatialisation » des graphèmes permettrait ainsi de délimiter explicitement pour l'apprenant des espaces phonético-phonologiques, ce qui pourrait également permettre, dans une certaine mesure, d'aborder la question de l'allophonie.

Une telle démarche viserait donc, initialement, moins une prononciation phonétiquement « parfaite » (prenant en compte la variation allophonique et la coarticulation) qu'une première reconstruction de l'espace phonético-phonologique segmental des apprenants, *via* l'établissement de catégories phonético-phonographémiques. Cependant, si l'on souhaite ensuite travailler, non plus sur le domaine phonologique segmental (inventaire segmental et frontières catégorielles), mais bien sur le domaine phonétique (allophones et coarticulation) et/ou phonologique à un niveau infra-segmental (traits et assimilation), on peut alors envisager le recours à d'autres outils visuo-graphiques comme l'API (par exemple dans le cas du dévoisement).

En ce qui concerne la prosodie, on peut également envisager de proposer une démarche similaire, c'est-à-dire celle d'une première catégorisation assez globale des « intonèmes », en faisant usage des signes de ponctuation : ., ?, !, :, ..., ; et , , en vue, ici encore, d'intégrer le travail phonético-phonologique au travail phonographémique <sup>329</sup>, pour que l'usage du texte ne soit pas un élément « perturbateur » dans le développement de l'oral.

#### III.4.2.3.1.2. Segmentation

Lorsque nous observons nos propres pratiques de classe, qu'il s'agisse de compréhension orale, de correction phonétique ou d'activités grammaticales, nous constatons que nous tendons spontanément à segmenter gestuellement tout énoncé complexe, à différents niveaux, quitte à utiliser des gestes « circulaires » lorsque les structures ne sont pas « linéaires », afin de mettre visuellement en relation ce qui doit

---

<sup>329</sup> Comme le note Coulmas (1989, p. 270) : « *The treatment in traditional and modern grammars of what [...] is called « suprasegmental features » such as stress, tone, intonation, melody and rhythm, illustrates the influence of writing on grammatical description most clearly. In traditional grammars, prosody is often ignored, and in modern ones it is treated in a separate chapter. In both cases it is abstracted from the discrete units that are distinguished in writing : sentences, words and letters. These units are treated as if they had an independent existence to which intonation is added in speech. But where does the intonationless sentence occur ? On the printed page* ». Il conviendrait donc de progressivement entraîner les apprenants à assigner un patron intonatif approprié aux énoncés écrits.

l'être, d'un point de vue tout aussi bien sémantique que grammatical ou autre. Si la didactique des langues, depuis l'époque SGAV en particulier, a « repoussé » l'écrit notamment par peur des interférences segmentales (graphèmes/phonèmes), nous avons vu qu'au niveau métrique, l'influence de l'écrit n'était pas moindre, que ce soit en termes de segmentation phonético-phonologique ou en termes d'unités fonctionnelles de lecture.

Lorsque l'on considère les différences de syllabation entre l'écrit et l'oral, particulièrement en français, un entraînement à la segmentation du français (en termes syllabiques, rythmiques voire intonatif), à l'écrit et à l'oral semble donc nécessaire, également utile, pour certains apprenants au développement de la conscience phonémique. Or, comme nous l'avons montré, l'écrit constitue précisément un support spatial de segmentation de l'oral. Il conviendrait donc de proposer des tâches de segmentation de l'oral *via* le support orthographique, ce qui permettrait en outre, d'une part d'aborder lien entre orthographe, phonologie et morphologie, mais également d'entraîner les apprenants à segmenter l'écrit en unités de tailles variables, en vue de l'entraînement à la lecture. Lorsque l'on prend également en compte les phénomènes de coarticulation, il serait envisageable de coupler à la segmentation orthographique une représentation visuo-graphique supplémentaire, soit de type phonétique (API), soit de type « diacritique », en « barrant » par exemple les « lettres » ou les « e » « muets » et en « encerclant » les unités de segmentation, syllabiques par exemple. Cela s'illustre très bien avec l'énoncé « Tout le monde est malade », qui, en API, pourrait être transcrit [tʉl-mɔ̃-dɛ-ma-lad].

#### III.4.2.3.2. Correction phonographémique en production (dictée)

Si l'on se penche à présent, non plus sur l'utilisation du support visuo-graphique dans la correction phonétique, mais dans l'apprentissage des correspondances phonographémiques à proprement parler, bien que cet apprentissage soit ultimement destiné à celui du système phonético-phonologique et non à celui de la « lecture », on peut envisager, en production, le rôle de la dictée. Ici encore, il ne s'agirait pas tant de

dictée « orthographique » que de dictée « phonographémique » faisant usage des graphèmes de base les plus fréquents du français mentionnés précédemment.

On pourrait alors imaginer la procédure suivante : l'enseignant dicte une liste d'items soigneusement conçus (mots ou non-mots), que l'apprenant doit transcrire. La correction, phonographémique et phonétique, doit intervenir juste après la production, puisque, à l'instar, de la correction phonétique, l'objectif est de faire percevoir l'écart phonographémique à l'apprenant. Concrètement, une telle activité pourrait se dérouler de la même manière qu'une séance de correction phonétique, l'enseignant intervenant successivement auprès de chaque apprenant. L'utilisation des graphèmes de base du français constituerait alors un moyen intermédiaire entre l'API (pas d'apprentissage des graphèmes du français) et l'Alfonic de Martinet (déformation, dans une certaine mesure, de l'orthographe française). De plus, une telle technique ne se veut pas systématique : il ne s'agit en aucun cas d'apprendre une orthographe simplifiée du français, mais plutôt d'effectuer un travail d'intégration phonético-phonologique, en utilisant, comme outil complémentaire et non exclusif, certains graphèmes du français, sélectionnés en fonction du public et des difficultés potentielles ou attestées, par exemple, dans le cas d'apprenants japonais, le <R> et le <l> ou encore les voyelles nasales. Envisagée de cette manière, la tâche de production sous dictée est tout à fait différente de la « dictée orthographique » traditionnelle<sup>330</sup>

#### III.4.2.3.3. Correction phonographémique en perception (lecture à voix haute)

Le pendant perceptif de la tâche précédente est évidemment celui de la lecture, dans un premier temps à voix haute, puis subvocalisée. La nécessité d'entraîner (ce qui implique un travail effectif de l'enseignant et des apprenants, et non une simple exposition au texte) les apprenants à lire (en tant que récupération des formes

---

<sup>330</sup> Quoique celle-ci puisse certainement, si elle traitée à bon escient, constituer un outil d'apprentissage des relations entre phonologie, morphologie et orthographe (voire sémantique) non-négligeable, ainsi utile au développement de la compétence grammaticale, et plus généralement langagière, des apprenants, *via* une certaine visualisation de ces relations, dans un code effectivement utilisé par la communauté francophone, à savoir l'orthographe française.



phonético-phonologique, et ainsi des gestes articulatoires) en français en début d'apprentissage semble aujourd'hui de nouveau considérée par les didacticiens (Billières, 2005). Cela est particulièrement important lorsque l'on songe que de nombreux enseignants, dont nous faisons partie, conseillent aux étudiants de « lire à voix haute » pour améliorer leur production orale. Dans notre cas, un tel conseil se justifie par le fait qu'il s'agit le plus souvent d'étudiants non-débutants, qui manquent cependant de « fluidité verbale » et de « rythme » : nous considérons alors que le texte écrit, en fournissant le contenu linguistique, permet aux apprenants de se concentrer précisément sur la forme (ou encore les processus de « bas-niveau »), notamment les caractéristiques prosodiques, des énoncés oraux correspondants.

Suivant la démarche présentée jusqu'ici, nous pouvons alors reprendre la procédure de correction phonétique évoquée plus haut, en remplaçant la tâche de répétition par une tâche de lecture à voix haute : l'enseignant propose une liste d'items soigneusement conçus (mots ou non-mots), éventuellement ciblés sur une difficulté phonographémique particulière, à lire à voix haute par l'apprenant. Suite à sa production, l'enseignant fournit alors un retour correctif audio-orthographique, en traçant une ligne graduée aux extrémités de laquelle se trouve le graphème modèle d'une part et le graphème correspondant à la production de l'apprenant d'autre part, de manière à ce que celui-ci puisse percevoir visuellement l'écart entre les deux. L'enseignant peut alors ensuite effectuer un travail de correction phonétique auditive. Il s'agit donc, ici encore, de coupler correction phonétique et correction phonographémique. On pourrait également envisager un recours à l'API si l'on souhaite davantage travailler sur certaines caractéristiques infra-segmentales ou allophoniques.

Une telle procédure permettrait en outre d'aborder l'apprentissage de l'orthographe française, en élargissant progressivement le répertoire graphémique des items soumis à lecture, des graphèmes les plus fréquents au moins fréquents, des non-mots aux mots. Cela serait également applicable au domaine syllabique (visualisation de la voyelle épenthétique lorsque l'apprenant effectue une épenthèse dans la lecture d'un groupe

consonantique), rythmique (visualisation d'espaces incorrectement placés entre les graphèmes, qui permettrait d'aborder la dimension morphosémantique), voir intonatif (*via* la ponctuation).

La mention rapide de l'utilisation des non-mots dans de telles procédures renvoie à une question peu abordée, nous semble-t-il, dans le domaine de la correction phonétique, et plus généralement de l'oral en didactique des langues, une absence qui s'explique par l'absence de démarche contrastive dans les méthodologies actuelles, quoique celle-ci semble redevenir un sujet d'actualité dans le domaine, notamment *via* le rôle de la L1 en classe de langue (Castellotti, 2001) : il s'agit du rôle des mots de même forme ou de même origine en L1 et en L2, que ceux-ci soient qualifiés de « mots transparents », de « vrais » ou « faux » « amis », ou encore de mots d'emprunts.

Afin d'illustrer la question, pensons à ces étudiants japonais qui, capables de répéter ou de lire correctement les syllabes /si-ne-ma/ (<ci-né-ma>) lorsqu'elles sont espacées dans le temps, sont en revanche incapables de répéter, ou plus encore de lire, correctement le mot /sinema/ (<cinéma>), qu'ils prononcent inmanquablement /ʃinema/. D'un point de vue phonético-phonologique, un tel phénomène renvoie à la question de la phonotactique et des contraintes phonologiques à l'œuvre dans le lexique phonologique en L1 ; d'un point de vue psycholinguistique, cela renvoie aux phénomènes de reconnaissance, auditive, par exemple dans la perspective des modèles à cohortes, ou visuelle, par exemple *via* la voie d'adressage dans les modèles à deux voies.

Il est alors intéressant de constater que, d'un point de vue didactique, ces unités lexicales constituent généralement un atout pour les apprenants en tâche de compréhension orale ou écrite, puisque, à l'instar des chiffres ou des noms propres par exemple, elles constituent des indices leur permettant d'appréhender globalement la signification du document et d'inférer celle des lexies et des énoncés inconnus. Or, il apparaît ici que, pour l'apprentissage du système phonético-phonologique (et

phonographémique) du français des apprenants débutants, de tels items semblent plutôt négatifs. On pourrait alors presque envisager *la reconnaissance morpholexicale comme un obstacle à l'intégration phonético-phonologique* (et ainsi à la correction phonétique), de la même manière que la reconnaissance phonémique et/ou allophonique constitue un obstacle à l'apprentissage de nouveaux phonèmes (selon le principe de crible phonologique entre autres).

Ainsi, apprendre l'oral revient également à apprendre la forme phonético-phonologique des *lexies*, et pas seulement l'inventaire segmental et la structure syllabique du système lexical. Vu sous cet angle, l'enseignement de la prononciation du français devrait donc, dans un premier temps, utiliser soit des non-mots, soit des mots non-reconnaissables par les apprenants, puisque l'apprentissage des catégories phonético-phonologiques s'effectue également par celui des formes lexicales. Or, si les formes phonético-phonologiques présentent une certaine variation entre les langues, qui permet ainsi d'éviter une reconnaissance auditive trop importante (*via* l'allophonie mais également l'adaptation phonologique), cela n'est pas le cas des formes orthographiques qui, pour les mots considérés, sont souvent graphiquement très ressemblants, étant donné la diffusion de l'alphabet latin à travers le monde. Cela s'illustre parfaitement avec le cas du mot « cinéma » ou bien du mot « football », qui, présentés auditivement, peuvent éventuellement être répétés correctement, mais qui, en présentation orthographique, entraîne de bien plus nombreuses productions erronées.

Cet aspect constitue donc le pendant morpholexical de ce que nous avons examiné au niveau segmental : la reconnaissance du morphème orthographique en L1 entraîne l'activation de la voie d'adressage et ainsi l'activation de la forme phonético-phonologique de ce mot en L1 et non en L2, activation renforcée par la valeur sémantique de l'item. Nous avons ainsi constaté à plusieurs reprises dans nos classes que, dès que le stimulus semble avoir été identifié (mots, morphèmes, graphèmes), la procédure d'adressage semble prendre le pas sur celle d'assemblage, du moins tel que cela apparaît dans la production orale des nos apprenants. Une telle observation permet alors de nuancer l'hypothèse forte que nous avons formulée dans l'interprétation de

nos résultats, à savoir que la forme orthographique était davantage liée à la forme phonologique lexicale de l'item qu'à sa forme phonétique en contexte, et que les contraintes phonotactiques qui s'appliquaient en tâche de répétition, résultant en des formes allophoniques, devaient moins s'appliquer en tâche de lecture à voix haute. Il semblerait ici qu'une telle hypothèse doive être modulée en fonction du caractère connu ou non de l'item : alors que les non-mots et les mots nouveaux activent la voie d'assemblage, plus contrôlée, qui correspond alors à notre hypothèse, la reconnaissance de mots familiers active la voie d'adressage, plus automatique, qui pourrait alors davantage directement connecter la forme orthographique à la forme phonétique (allophonique, coarticulée)<sup>331</sup>.

Un tel constat impliquerait donc, pour le didacticien, de prendre connaissance des ressemblances morphologiques existant entre la L1 et la L2, afin de ne travailler, dans un premier temps qu'avec des mots inconnus ou des non-mots, avant d'élargir ensuite la liste des items concernés.

### **III.4.5. Ouvertures**

Au terme de cette partie, il convient à présent de brièvement conclure, une conclusion qui ne peut être qu'une ouverture vers des travaux futurs. Comme nous l'avons vu, l'écrit est censé transcrire l'oral. Si l'on suit cette hiérarchie « naturelle », il semble donc évident que les étudiants doivent apprendre l'oral avant d'apprendre le moyen de le transcrire. Apprendre l'écrit, puis ensuite la manière de le prononcer, serait une démarche en contradiction avec ce que l'on sait des systèmes linguistiques et du traitement cognitif de l'écrit et de l'oral.

---

<sup>331</sup> Une telle analyse doit cependant être affinée, car il ne faudrait pas croire que les mots familiers peuvent être automatiquement reconnus « globalement ». Perea et Rosa (2002) ont effectué des études sur l'influence potentielle de la « forme globale des mots » en lecture, et leurs résultats semblent diminuer la portée de celle-ci, même s'ils reconnaissent également que cette « information globale » peut, à une certaine étape du traitement, jouer un rôle dans la procédure de reconnaissance sur l'existence de celle-ci : « *The results [...] weaken the view that high-frequency words can be identified via a fast word-level route as a function of visual familiarity. [...] The present findings suggest that recognition of familiar words does not seem to rely on word-shape information* ».

Cependant, cet agencement ne remet pas en cause l'importance de l'écrit dans l'apprentissage, et toute notre étude a donc consisté à souligner l'importance, non pas d'une séparation arbitraire, mais d'une intégration raisonnée des deux formats, sans quoi nous risquerions d'obtenir des apprenants qui « écrivent comme ils parlent » (ce que l'on a reproché à des approches pédagogiques trop « laxistes » sur l'apprentissage de l'écrit et de l'orthographe) ou bien qui « parlent comme ils écrivent » (ce que l'on a reproché à des approches par trop basées sur l'écrit, et que l'on trouve encore aujourd'hui, notamment en France, tant au niveau de la « prononciation » qu'au niveau syntaxique, lexical, et plus globalement discursif).

Ainsi, si l'on a longtemps considéré que les apprenants de langue étrangère étaient « sourds » à la parole en L2, il ne faut pas oublier que certains d'entre eux, pour ne pas dire la plupart, sont également, dans une certaine mesure, des « illettrés ». Il convient donc de leur apprendre à « entendre », mais également à « lire » en L2.

Cette intégration *méthodologiquement raisonnée* de l'oral et de l'écrit dans le curriculum d'enseignement nécessite cependant en premier lieu de clairement définir les objectifs d'enseignement/apprentissage, puisqu'il existe également des étudiants dont l'objectif premier, par exemple, est la lecture de textes français, scientifiques ou littéraires, et il se pourrait que l'approche que nous avons développée jusqu'ici, consacrée à l'apprentissage de l'oral, ne corresponde pas de manière optimale à leurs besoins ou à leurs attentes.

Néanmoins, pour le public qui nous intéresse, il semble que cette intégration soit aujourd'hui nécessaire, indiquant ainsi la nécessité de faire réémerger la problématique de l'articulation entre oral et écrit dans le domaine didactique, tant au niveau de la recherche, qu'à celui de la conception et des pratiques pédagogiques. Cette nécessité est particulièrement accentuée par le développement des outils multimédiatisés, qui offrent des possibilités jusqu'alors inaccessibles, notamment en termes de stockage et de manipulation d'input audiovisuel, dans l'espace et dans le temps.

Cependant, l'écrit restant encore un format informatif incontournable dans nos sociétés actuelles, les spécialistes du domaine devraient, nous semble-t-il, ne pas oublier que la multimodalité commence par la simple combinaison de l'oral et de l'écrit, de la même manière que les enseignants de langue ne devraient pas oublier que la multimodalité commence à la première page du manuel. Sans aucun doute, les travaux en psychologie ergonomique permettront, dans le futur, d'apporter davantage de réponses aux multiples questions que nous avons soulevées jusqu'ici, concernant l'optimalité, notamment spatiale et temporelle, des formats de présentation de l'input langagier pour l'apprentissage des langues étrangères.

C'est donc bien sur une ouverture multiple que nous achevons cette partie consacrée aux implications didactiques de notre étude : ouverture sur la recherche, ouverture sur la conception, ouverture sur l'innovation didactique, méthodologique et technique.



## **Conclusion Générale**

Conclure une thèse de doctorat n'est pas une tâche évidente : bien qu'un point final doive inévitablement être mis au fruit de ces trois années de recherche, certaines réflexions restent encore à approfondir, d'autres données semblent nécessaires pour développer ou affiner celles qui l'ont été et de nouveaux travaux concernant notre sujet ne cessent d'être conduits, et leurs résultats publiés, alors que nous rédigeons ces dernières lignes. Une telle conclusion se doit donc dans un premier temps, nous semble-t-il, de souligner les insuffisances du travail accompli : beaucoup reste à faire, et à refaire. Cependant, une telle conclusion est également prometteuse : de par ses insuffisances, mais surtout en raison de son sujet et de la pluridisciplinarité qu'elle invoque, notre thèse constitue également un certain tremplin vers des recherches ultérieures et variées, dont nous avons brièvement esquissé la teneur précédemment. Loin de constituer une fin, c'est donc bien d'un début qu'il s'agit, comme se devrait d'être tout travail de ce type.

Rappelons que le thème figurant au cœur de notre sujet de thèse, à savoir celui du rôle de l'écrit dans l'apprentissage de l'oral en langue étrangère, y a été traité en six temps :

- 1) Une amorce didactique, basée sur des observations de classe contextualisées de productions orales et écrites d'apprenants japonais de français ;
- 2) Une partie linguistique visant à contraster les systèmes phonético-phonologiques et orthographiques du français et du japonais ;
- 3) Une partie psycholinguistique rendant compte du fonctionnement cognitif, en particulier perceptif, des apprenants en question ;
- 4) Une partie de synthèse consacrée à la formulation de nos problématiques d'apprentissage liées aux interactions potentielles entre modalités auditive et visuelle dans la perception et *a fortiori* l'apprentissage de la parole en langue étrangère ;
- 5) Une partie expérimentale illustrant nos hypothèses de recherche ;



- 6) Une partie didactique relative aux implications de notre étude pour l'enseignement de l'oral en français langue étrangère.

Par delà les résultats spécifiques de notre partie expérimentale concernant le traitement des groupes consonantiques /CC/ et des segments /b, v, r, l/ par nos sujets japonais, notre étude nous a permis de montrer, à travers la revue de la littérature concernée, à travers notre argumentaire, et à travers nos résultats empiriques, que les représentations visuo-graphiques, et plus particulièrement orthographiques, pouvaient jouer un rôle dans le traitement, perceptif et mnésique, de la parole en langue étrangère. D'un point de vue phonético-phonologique et psycholinguistique, un tel constat n'est pas sans implications quant aux recherches menées dans ces deux domaines, puisque la variable « orthographe » devient alors non-négligeable dans les protocoles de recherche concernant la parole orale.

Cependant, puisque notre travail tient à se positionner en premier lieu dans le domaine de la didactique des langues et des cultures, ce sont les implications et conclusions didactiques qui retiendront particulièrement notre attention. En effet, notre étude à remis au goût du jour une problématique qui, notamment en raison des développements des outils didactiques multimédiatisés, ne peut plus être reléguée dans le passé, à savoir celle de l'articulation entre oral et écrit dans la progression d'enseignement / apprentissage pour un public adulte en milieu hétéroglotte. L'originalité de cette étude, outre le public (japonais) et la langue (française) en présence, tient à la fois au traitement pluridisciplinaire et actualisé de cette problématique, mais également, du point de vue des sciences cognitives, à la mise en relation de deux axes d'étude traités de manière relativement indépendante jusqu'à présent, nous semble-t-il, celui de l'interphonologie et celui de la lecture en langue étrangère. Ces deux axes nous ont conduits à parler d' « effet Buben en L2 », un effet aux implications didactiques non négligeables, et qui ouvre également la voie à de nombreuses recherches dans les domaines concernés, en déclinant ce thème selon différentes L1 et différentes L2. Notre étude souligne donc la nécessité de prendre en compte la dimension visuo-orthographique dans l'enseignement / apprentissage de l'oral, nécessité qui rend

caduque l'utilisation immédiate, naïve, et surtout *non-pensée*, du support écrit en classe d'oral, et plus généralement de langue.

Notre insistance sur la réflexion et l'analyse méthodologiques *en amont* des pratiques de conception et d'intervention didactique, bien qu'elles ne puissent être conduites, selon nous, sans interactions régulières et suffisantes avec le terrain, concret et contraignant, des classes de langue, au risque de se transformer en travail « de papier » sensiblement éloigné des réalités quotidiennes des enseignants et des apprenants, témoigne de notre conception non seulement de la formation des enseignants de langue, mais également de celle de la recherche en didactique des langues. Pour pouvoir enseigner une langue étrangère, il faut en effet, nous semble-t-il : 1) connaître ce que l'on doit enseigner (langue et culture), et 2) savoir qui l'apprend (public) et comment il l'apprend (apprentissage). Ce n'est qu'à partir de cette base que l'on peut envisager le « Comment enseigner », comme cela est grossièrement illustré ci-dessous :

<b>Objet / Enseigner</b>	<b>Sujet / Apprendre</b>
<u>Quoi</u> → Donné	<u>Qui</u> → Donné
<i>Comment : A élaborer</i>	<u>Comment</u> → Donné

Si ce tableau pourrait ainsi, de manière très globale, schématiser la formation qui devrait être dispensée aux enseignants, il pose également, d'une certaine manière, les prémisses de toute démarche de recherche en didactique des langues et des cultures.

Comme nous l'avons affirmé précédemment, il nous semble en effet que, afin d'assurer son intégrité légitime, la didactique des langues, en tant que domaine de recherche académique, a besoin de davantage se fonder scientifiquement. Cependant, cela n'implique en rien la constitution d'une didactique « théorique » illusoirement opposée

à une didactique « de terrain », car toute entreprise didactique se doit, nous semble-t-il, d'être fondamentalement « de terrain » et non « de papier » (*via* la glose, la paraphrase et le métadiscours, utiles pour la réflexion mais limitées à celle-ci) ou « de laboratoire » (*via* une quantification et une aseptisation des conditions d'investigation, utiles pour l'analyse des faits, mais trop coupées des conditions réelles d'enseignement / apprentissage). L'une des difficultés majeures posées à la didactique est la complexité et la multidimensionnalité de son objet d'étude, qui semblent impliquer un éclatement des problématiques et des approches, mettant en péril l'existence même de la didactique en raison du recours à d'autres disciplines de références qui risqueraient de se les accaparer. Or, il ne nous semble pas qu'un tel éclatement soit inéluctable : bien que pluridisciplinaire, la didactique n'est pas « tout », sans quoi elle ne serait « rien ». Si l'on ne peut effectivement, en l'état actuel des choses, conduire une recherche scientifique parfaitement « holistique », la recherche en didactique, plutôt que de souffrir d'éclatement ou de parcellisation, devrait, nous semble-t-il être « stratifiée », ces strates constitutives, et leurs interfaces, devant être « mises en réseaux », et non abusivement séparées les unes des autres. C'est ce que nous avons, dans une très modeste mesure, tenté d'accomplir, au moins partiellement, dans notre travail de thèse, en nous concentrant sur les aspects linguistiques et cognitifs de la question.

Lorsque l'on observe les mouvements les plus récents de la didactiques du FLE en France, en particulier la tendance « cognitive » liée au Colloque DidCog2005 tenu à Toulouse en janvier 2005, organisé par le laboratoire Jacques-Lordat, et la tendance « interdidactique et interculturelle » liée au Colloque de février 2005 à Saint - Etienne, organisé par le Centre d'Etudes en Didactique Comparée des Langues Et des Cultures (CEDICLEC), on ose espérer que ces deux grands courants, plutôt que de se placer en opposition, réalisent au contraire leur indispensable complémentarité dans la perspective didactique stratifiée que nous avons mentionnée. Puisque l'on ne peut « s'occuper de tout », il est en effet nécessaire, voire souhaitable, que se répartissent les thématiques de recherche, à condition que celles-ci puissent, à terme, être rassemblées en un tout cohérent pour les didacticiens et les enseignants de langue. Etant donné le dépassement actuel de la dichotomie Théorie/Pratique (la théorie ne pouvant subsister

sans pratique et la pratique ne pouvant être guidée sans théorie, ou du moins sans théorisation), il conviendrait donc, dans une perspective « floue » et non plus binaire, d'établir des « strates » thématiques de recherche, toutes fondées sur un continuum entre pratique didactique et recherche didactique. Si l'on ne peut être « spécialiste de tout », ce qui implique pour le chercheur - didacticien une certaine spécialisation, il n'en reste pas moins que la didactique des langues présente des caractéristiques qui lui sont propres : c'est donc bien, selon nous, aux didacticiens de prolonger leur formation initiale par une spécialisation adéquate, et non aux spécialistes d'autres disciplines de se former en didactique, si l'on souhaite assurer à la didactique un paradigme de recherche respectueux de son intégrité et de ses problématiques.

C'est dans cette perspective que nous plaçons nos propres travaux, travaux que nous espérons pouvoir poursuivre dans le futur, notamment autour d'un des thèmes qui nous préoccupent depuis plusieurs années et qui, nous semble-t-il, constitue également l'une des pistes porteuses de la didactique des décennies à venir, à savoir celui de l'*adaptation*, que celle-ci soit didactique (*via* les travaux sur l'interdidactique), culturelle (*via* les travaux sur l'interculturel) ou cognitive (*via* les travaux sur l'interlangue et les aspects cognitifs de l'apprentissage d'une langue étrangère). Lié au concept d'*adéquation* didactique, un tel thème remet alors également sur le devant de la scène, à l'instar de notre sujet de thèse, la question de la méthodologie didactique, délaissée ces dernières années face à l'éclectisme pragmatique des pratiques pédagogiques contemporaines.

Enfin, de manière plus spécifique, si nous espérons que ce travail de thèse sera bénéfique à la didactique du français au Japon, nous espérons que la réflexion que nous y avons conduite sera également utile à la didactique du FLE dans son ensemble, en soulevant quelques pistes de réponses et d'études relatives à l'enseignement / apprentissage de l'oral en langue étrangère. Ces quelques pistes ne prétendent en aucun cas apporter de quelconques « vérités » ou réponses absolues : tout au contraire, elles ne constituent qu'une bien modeste « étape » dans cette entreprise foncièrement collective qu'est la recherche scientifique, nous invitant dès lors, ainsi que les autres

chercheurs qui le souhaiteraient, à poursuivre nos travaux, nos investigations et notre réflexion sur la question.

## **Bibliographie**

Rappelons que nous avons choisi de mentionner dans le corps de la thèse (généralement en note de bas de page) la plupart des références citées par les auteurs que nous avons consultés, lorsque nous n'avons pas eu directement accès à ces textes. Celles-ci ne sont donc généralement pas répertoriées ci-dessous. Nous avons également restreint notre inventaire aux travaux explicitement mentionnés dans le corps de la thèse, et n'avons donc pas fait figurer ci-après les références des autres documents que nous avons lus au cours de notre recherche. Par ailleurs, lorsque certains documents n'ont été consultés qu'à titre de référence, c'est-à-dire lorsque nous n'avons pas lu l'ensemble de l'ouvrage ou de l'article en question, nous indiquons (REF) dans la bibliographie. Les autres documents ont été lus, soit en intégralité, soit en très grande partie. Enfin, nous indiquons, à la fin de cette bibliographie, un ensemble de références que nous n'avons pas pu consulter, mais qui concernent de près notre travail, et que nous avons notées à plusieurs reprises à travers nos lectures. Nous espérons, dans le futur, pouvoir y accéder. Cette bibliographie (504 références au total) est donc inévitablement limitée et inachevée : nous avons tenté, au cours de notre étude, de nous maintenir informés des publications les plus récentes concernant notre sujet, mais il va de soi que la quantité et l'évolution rapide des publications concernées rendent impossible une exhaustivité parfaite, impossibilité accentuée par le caractère relativement pluridisciplinaire de notre travail. De la même manière, certains ouvrages de référence plus anciens ne nous ont pas été accessibles, et nous assumons donc pleinement les insuffisances de notre recherche bibliographique.

## A

- ABRAHAMSEN, A. et BECHTEL, W., *Le connexionnisme et l'esprit, introduction au traitement parallèle par réseau*, Paris, Edition La Découverte, traduit de l'anglais par J. Proust, 1993.
- ABRIAL, J.-P., « Propositions en vue de l'enseignement de la lecture et de la phonétique dans les classes de débutants japonais », *Enseignement du français au Japon*, n° 24, 1995, pp. 52-55.
- ABRY, C. et SCHWARTZ, J.-L., « La parole cognitive », in Gordon, M. B. et Paugam-Moisy, H. (dir.), *Sciences cognitives : diversité des approches*, Paris, Hermès, 1997, pp. 103-115.
- AKAMATSU, N., « A similarity in word-recognition procedure among L2 readers with different L1 backgrounds », *Applied Psycholinguistics*, Vol. 23, 2002, pp. 117-133.
- AKAMATSU, N., « The effects of first language orthographic features on second language reading in text », *Language Learning*, Vol. 53, n° 2, 2003, pp. 207-231.
- AKAMATSU, T., *Japanese Phonetics : theory and practice*, LINCOS Studies in Asian Linguistics 03, Munich, LINCOS EUROPA, 1997.
- ALARIO, F.-X. et ZIEGLER, J., « The role of orthographic information in speech production revisited », poster, *AMLaP 2004*, 16-18 septembre 2004, Université d'Aix-en-Provence.
- ALLPORT, A., « Word Recognition in Reading », in Kollers, P. A., Wrolstad, M. E. et Bouma, H. (éd.), *Processing of Visible Language*, Vol. 1, New-York, Plenum Press, 1979, pp. 227-257.
- ANDERSON, J.-R., *Rules of the Mind*, Hillsdale, Lawrence Erlbaum, 1993 (REF).
- ARCHANGELI, D., « Optimality Theory : an introduction to linguistics in the 1990s », in Archangeli, D. et Langendoen, D. T. (éd.), *Optimality Theory : an overview*, Malden, Blackwell Publisher, 1997, pp. 1-33.
- ARCHIBALD, J., « Parsing procedures and the question of full access in L2 phonology », *New Sounds 2000*, Plenary address, University of Amsterdam, 2000.
- ATTINA, V., BEAUTEMPS, D. et CATHIARD, M.-A., « Organisation spatio-temporelle main-lèvres-son de séquences CV en Langage Parlé Complété », *XXIVèmes Journées d'Etude sur la Parole*, Nancy, 2002.
- AZRA, J.-L., « Difficultés posées par l'utilisation du "jugement de grammaticalité" », *Studies in Language and Culture*, 25, Osaka University, 1999.
- AZRA, J.-L., « Découpe syllabique et rythme en français : application pédagogique », communication, *Congrès d'Automne de la Société Japonaise de Langue et de Littérature Françaises*, Université des Langues Etrangères d'Osaka, 2003.

## B

- BADDELEY, A. D., *Working memory*, Oxford, OUP<sup>332</sup>, 1986.
- BADDELEY, A. D., GATHERCOLE, S. et PAPAGNO, C., « The phonological loop as a language learning device », *Psychological review*, Vol. 105, n°1, 1998, pp. 158-173.
- BAQUE, L., « Effet sur la prononciation d'apprenants de FLE de la tâche et de la focalisation sur la forme », in Billières, M., Gaillard, P. et Spanghero-Gaillard, N. (éd.), *Actes du Premier Colloque International de Didactique Cognitive, (Français Langue Etrangère / Langue Seconde / Langue Maternelle)* [CD-ROM], Université de Toulouse-le Mirail, 2005, pp. 7-13.
- BATTYE, A. HINTZE, M.-A. et ROWLETT, P., *The French Language Today : a linguistic introduction*, 2<sup>ème</sup> éd., Londres, Routledge, 2000.
- BAYLON, C. et MIGNOT, X., *La communication*, Paris, Nathan, 1994/1999 (REF).
- BEACCO, J.-C., « Influence du Cadre (CECR) sur les programmes et les dispositifs d'évaluation », *Le Français dans le Monde*, n° 336, 2004, pp. 25-28.
- BEACCO, J.-C. et CHEVALIER, J.-C. (Coord.), « Didactique des langues : quelles interfaces ? », *Etudes de Linguistique Appliquée* n°72, Paris, Didier Erudition, 1988.
- BEAUVILLAIN, C. et SEGUI, J., « Representation and processing of morphological information », in Frost, R. et Katz, L. (éd.), *Orthography, Phonology, Morphology and Meaning*, Amsterdam, Elsevier Science publishers, 1992, pp. 377-387.
- BEEBE, L. M., « Myths about interlanguage phonology », in Ioup, G. et Weinberger, S. H. (éd.), *Interlanguage Phonology : the acquisition of a second language sound system*, Cambridge, Newbury House, 1987, pp. 165-175.
- BENTIN, S., « Phonological awareness, reading, and reading acquisition : a survey and appraisal of current knowledge », in Frost, R. et Katz, L. (éd.), *Orthography, Phonology, Morphology and Meaning*, Amsterdam, Elsevier Science Publishers, 1992, pp. 193-210.
- BERARD, E., « La grammaire, encore...et l'approche communicative », in Atienza, J.-L., Bérard, E. et De Carlo, M. (Coord.), « Où en est le communicatif ? », *Etudes de Linguistique Appliquée*, n°100, Paris, Didier Erudition, 1995, pp. 9-19.
- BERGER-VACHON, C. et COLLET, L., « Le système auditif périphérique : physiologie et perception du signal », in Gordon, M. B. et Paugam-Moisy, H. (dir.), *Sciences cognitives : diversité des approches*, Paris, Hermès, 1997, pp. 29-41.
- BERGERE, A., « Adultes chinois apprenant le français : quelles compétences de départ et quelles compétences à acquérir pour la lecture et l'écriture alphabétiques ? », in Billières, M., Gaillard, P. et Spanghero-Gaillard, N. (éds.), *Actes du Premier Colloque International de Didactique Cognitive, (Français Langue Etrangère / Langue Seconde / Langue Maternelle)* [CD-ROM], Université de Toulouse-le Mirail, 2005, pp. 14-19.

---

<sup>332</sup> OUP = Oxford University Press ; CUP = Cambridge University Press.



- BERTIN, J.-C., *Des outils pour des langues, multimédia et apprentissage*, Paris, Ellipses, 2001.
- BESNER, D. et SMITH, M. C., « Basic processes in reading : is the orthographic depth hypothesis sinking ? », in Frost, R. et Katz, L. (éd.), *Orthography, Phonology, Morphology and Meaning*, Elsevier Science Publishers, 1992, pp. 45-66.
- BESSE, H., *Méthodes et pratiques des manuels de langue*, Paris, CREDIF, 1985.
- BESSE, H. et PORQUIER, R., *Grammaire et didactique des langues*, Paris, Didier, 1991.
- BETRANCOURT, M. et CARO, S., « Intégrer des informations en escamots dans les textes techniques : quels effets sur les processus cognitifs ? », in Tricot, A. et Rouet, J.-F. (éd.), « Les hypermédias : approches cognitives et ergonomiques », *Hypertextes et Hypermédias*, numéro hors série, Hermès, 1998, pp. 157-174.
- BIEDERMAN, « Higher-level vision », in Osherson D. N., Kosslyn, S. M. et Hollerbach, J. M. (éd.), *Visual cognition and action ; an invitation to cognitive science, vol. 2*, Cambridge, MIT Press, 1990, pp. 1-36.
- BILGER, M., « Le projet "Corpus de référence de français parlé" », *Le Français dans le Monde, Recherche et Applications*, numéro spécial « Oral : variabilité et apprentissages », 2001, pp. 103-105.
- BILLIERES, M., « Crible phonique, crible psychologique et interprétation phonétique en langue seconde », *Travaux de didactique du F.L.E*, Université de Montpellier III, n° 19, 1988, pp. 5-29.
- BILLIERES, M., « Psycholinguistique et méthodologie verbo-tonale », *Revue PArôle*, n° 11-12 : « Langage et cognition » (1<sup>ère</sup> partie), 1999, pp. 173-197.
- BILLIERES, M., Notes du Séminaire de DEA de Sciences du Langage, « La prosodie : nature et fonctions », Université de Toulouse-Le Mirail, 2002-2003.
- BILLIERES, M., « Codage phonologique et boucle articulatoire en mémoire de travail : un support pour la facilitation de l'accès à l'oral et à la lecture pour des publics débutants en français langue étrangère », *CORELA*, Numéro spécial : Colloque AFLS, 2005.
- BILLIERES, M. et BORRELL, A., « Quelques problèmes soulevés par les différentes variétés d'accents dans les méthodes de français langue étrangère », *Revue de Phonétique Appliquée*, n°94, 1990, pp. 17-40.
- BIRD, S. A. et WILLIAMS, J. N., « The effect of bimodal input on implicit and explicit memory : an investigation into the benefits of within-language subtitling », *Applied Psycholinguistics*, 23, 2002, pp. 509-533.
- BLANCHE-BENVENISTE, C., *Approches de la langue parlée en français*, Gap, Ophrys, 1997.
- BLANCHET, P., « Implications pratiques de la phonologie fonctionnelle : de la méthodologie de recherche à l'intervention didactique », *La linguistique*, Vol. 38, fasc. 2, 2002, pp., 37-52.
- BLEVINS, J., « The syllable in phonological theory », in Goldsmith, J. (éd.), *The Handbook of Phonological Theory*, Oxford, Blackwell, 1995, pp. 206-243.
- BLONDEL, J.-F., « Emotions et Attention », in Hamon, J.-F. (dir.), *Attention, vigilance, émotion*, Université de la Réunion, Paris, L'Harmattan, 1998, pp. 119-170.

- BOË, L.-J. et SCHWARTZ, J.-L., « L'émergence des structures phonologiques à la lumière des relations production-perception », in Lambert, J. et Nespoulous, J.-L. (dir.), *Perception auditive et compréhension du langage : état initial, état stable et pathologie*, Marseille, Solal Editeurs, 1997, pp. 49-65.
- BOERSMA, P., ESCUDERO, P. et HAYES, R., « Learning abstract phonological from auditory phonetic categories : an integrated model for the acquisition of language-specific sound categories », in Sole, M. J., Recansens, D. et Romero, J. (éds.), *Proceedings of the 15<sup>th</sup> International Congress of Phonetic Sciences*, Barcelone, 2003, pp. 1013-1016.
- BONNET, C., « La perception visuelle des formes », in Bonnet, C., Ghiglione, R. et Richard, J.-F. (dir.), *Traité de psychologie cognitive : perception, action, langage*, Paris, Bordas, 1989/ Dunod, 2003, pp. 1-82.
- BOOGARDS, P., *Aptitude et affectivité dans l'apprentissage des langues étrangères*, Paris, Hatier, 1988.
- BORRELL, A., « Correction phonétique de la prononciation et rééducation de l'audition par la méthode verbo-tonale », *Travaux de didactique du F.L.E.*, Université de Montpellier III, n° 8, 1982, pp. 21-28
- BORRELL, A., « Perception et crible phonologique », *Mélanges de phonétique et didactique des langues*, Hommage au Professeur R. Renard, Presses Universitaires de Mons et Didier Eruditions, Mons, Paris, 1991, pp. 31-42.
- BORRELL, A., « Relation entre les aspects articulatoires et les aspects acoustiques en phonétique. Quels outils utiliser en didactique des langues? », *Revue de Phonétique Appliquée*, n°107, 1993, pp. 97-111.
- BORRELL, A., « Parallèle entre perception et production. Complexité du phénomène de reconnaissance et de production des unités phonético-phonologiques », *La Linguistique*, Vol.v32, Fasc.2, 1996a, pp. 105-116.
- BORRELL, A., « Systématisation des erreurs de production et donc de perception chez les apprenants étrangers? », *Revue de Phonétique Appliquée*, 1996b, pp. 118-119.
- BORRELL, A., « Problèmes de distribution et correction phonétique », *Cahiers du Centre Interdisciplinaire des Sciences du Langage*, Université de Toulouse Le Mirail, n°12, 1997, pp. 53-65.
- BORRELL, A., « Peut-on dire qu'une langue est à prononciation tendue ou relâchée? Quelques réflexions à ce propos? » *La Linguistique*, Vol. 34, Fasc.1, 1998, pp. 51-64.
- BORRELL, A et BILLIERES, M., « L'évolution de la norme phonétique en français contemporain », *La Linguistique*, Vol. 25, Fasc. 2, 1989, pp. .
- BOST, S., « l'ortographe fonologique », *Le Français Aujourd'hui*, n°20, 1973.
- BOWERS, J. S., « Challenging the widespread assumption that connectionism and distributed representations go hand-in-hand », *Cognitive Psychology*, vol. 45, n°3, 2002, pp. 413-445 (REF).
- BRAND, M., REY, A. et PEEREMAN, R., « Where is the syllable priming effect in visual word recognition ? », *Journal of Memory and Language*, 48, 2003, pp. 435-443.

- BRANDAO DE CARVALHO J., « What are phonological syllables made of ? The voice/length symmetry », in Durand, J. et Laks, B. (éds.), *Phonetics, Phonology and Cognition*, Oxford, OUP, 2002, pp. 51-79.
- BRANDAO DE CARVALHO J., et KAWAGUCHI, Y., « Linéarité et variation : le cas du "n mouillé" en français », *Flambeau*, n° 28, Tokyo, Université Nationale des Langues Etrangères de Tokyo, 2002, pp. 1-20.
- BROEDER, B. et PLUNKETT, K., « Connectionism and second language acquisition », in Ellis, N. (éd.) *Implicit and Explicit Learning of Languages*, San Diego, Academic Press, 1994, pp. 421-453.
- BROSELOW, E., « Nonobvious transfer : on predicting epenthesis errors », in Gass, S. et Selinker, K. (éd.), *Language Transfer in Language Learning*, Rowley, Newbury house, 1983, pp. 269-280.
- BROSELOW, E., « Skeletal positions and moras », in Goldsmith, J. (éd.), *The Handbook of Phonological Theory*, Oxford, Blackwell, 1995, pp. 175-206.
- BRUILLARD, E. et DE LA PASSARDIERE, B., « Fonctionnalités hypertextuelles dans les environnements d'apprentissage », in Tricot, A. et Rouet, J.-F. (éd.), « Les hypermédias : approches cognitives et ergonomiques », *Hypertextes et Hypermédias*, numéro hors série, Hermès, 1998, pp. 95-122.
- BRUYER, R., « Les agnosies visuelles », in Séron, X. et Jeannerod, M. (dir.), *Neuropsychologie humaine*, Bruxelles, Mardaga, 1994, pp. 179-201.
- BULLIER, J., « Aspects temporels du traitement cortical de l'information visuelle », in Gordon, M. B. et Paugam-Moisy, H. (dir.), *Sciences cognitives : diversité des approches*, Paris, Hermès, 1997, pp. 77-89.
- BURGESS, N. et HITCH, G., « Toward a network model of the articulatory loop », *Journal of Memory and Language*, 31, 1992, pp. 429-460.
- BYUNG-JIN, L., « Effects of L1 phonology and orthography on L2 perception : the use of Korean Hangul in perception of English voicing contrasts », *9<sup>th</sup> Conference on Laboratory Phonology*, University of Illinois, juin 2004.

## C

- CALVERT, G. A., BULLMORE, E. T., BRAMMER, M. J., CAMPBELL, R. C., WILLIAMS, S. C. R., McGUIRE, P. K., WOODRUFF, P. W. R., IVERSEN, S. D. et DAVID, A. S., « Activation of auditory cortex during silent lipreading », *Science*, Vol. 76, 1997, pp. 593-596.
- CALVERT, G. A., CAMPBELL, R., BRAMMER, M. J., « Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex », *Current Biology*, 10, 2000, pp. 649-657.
- CARELLO, C., TURVEY, M. T. et LUKATELA, G., « Can theories of word recognition remain stubbornly Nonphonological ? », in Frost, R. et Katz, L. (éd.), *Orthography, Phonology, Morphology and Meaning*, Elsevier Science Publishers, 1992, pp. 211-226.
- CARO, S. et BETRANCOURT, M., « Ergonomie des documents techniques informatisés : expériences et recommandations sur l'utilisation des organisateurs

- para-linguistiques », in Tricot, A. et Rouet, J.-F. (éd.), « Les hypermédias : approches cognitives et ergonomiques », *Hypertextes et Hypermédias*, numéro hors série, Hermès, 1998, pp. 123-138.
- CARRELL, P.L., « Culture et contexte dans la lecture en langue étrangère : rôle des schémas de contenu et des schémas formels », (traduit de l'anglais par N. Lahmiti-Riou), in Gaonac'h, D. (Coord.), *Acquisition et utilisation d'une langue étrangère, l'approche cognitive*, Paris, Hachette, 1990, pp. 16-29 .
  - CARTON, F., *Introduction à la phonétique du français*, Paris, Bordas, 1974.
  - CASTELLOTTI, V., *La langue maternelle en classe de langue étrangère*, Paris, CLE International/HER, 2001.
  - CASTRO-CALDAS, A., PETERSSON, K. M., REIS, A., STONE-ELANDER, S. et INGVAR, M., « The illiterate brain : learning to read and to write during childhood influences the functional organization of the adult brain », *Brain*, 121, 1998, pp. 1053-63.
  - CATACH, N., *L'orthographe*, Paris, PUF, Col. Que-sais je ?, 1978, sixième édition 1995.
  - CATFORD, J. C., *A Practical Introduction to Phonetics*, Oxford, OUP, 1994.
  - CAZADE, A., « De l'usage des courbes sonores et autres supports graphiques pour aider l'apprenant en langues », *ALSIC*, Vol. 2, n°2, 1999, pp. 3-32.
  - CELSIS, P., DOYON, B., BOULANOUAR, K. et NESPOULOUS, J.-L., « Traitement de bas-niveau dans la perception auditive », in Lambert, J. et Nespoulous, J.-L. (dir), *Perception auditive et compréhension du langage : état initial, état stable et pathologie*, Marseille, Solal Editeurs, 1997, pp. 85-97.
  - CHAMBERS, K. E., ONISHI, K. H. et FISHER, C., « Learning different phonetic distributions in onset and coda position », poster, *9<sup>th</sup> Conference on Laboratory Phonology*, University of Illinois, juin 2004.
  - CHAMPAGNE, M., « Compréhension des actes de parole non-littéraires : approche psycholinguistique et perspectives neuropsycholinguistiques », *Revue Parole* n° 11-12, « Langage et cognition » (1<sup>ère</sup> partie), 1999, pp. 239-255.
  - CHAMPAGNE-MUZAR, C. et BOURDAGES, J. S., *La phonétique en didactique des langues*, Anjou, CEC, 1993.
  - CHAPELLE, C. A., *Computer Applications in Second Language Acquisition*, Cambridge, CUP, 2001 (REF).
  - CHEN, T.H. et MASSARO, D.W., « Mandarin speech perception by ear and eye follows a universal principle », *Perception and Psychophysics*, à paraître.
  - CHEREAU, C., HALLE, P. A. et SEGUI, J., « Interference between surface form and abstract representation in spoken word perception », *Proceedings of the Vth European Conference on Speech Communication and Technology*, Budapest, Hongrie, 1999.
  - CHEVROT, J.-P. et MALDEREZ, I., « L'effet Buben : de la linguistique diachronique à l'approche cognitive (et retour) », *Langue Française*, n°124, 1999, pp. 104-125.
  - CHIKAMATSU, N., « The effects of L1 orthography on L2 word recognition », *Studies in Second Language Acquisition*, 18, 1996, pp. 403-432.

- CHUN, D. M., « Signal analysis software for teaching discourse intonation », *Language Learning and Technology*, Vol. 2, n°1, 1998, pp. 61-77.
- COLANTONI, L. et STEELE, J., « Acquiring /R/ in context », *7èmes Rencontres Internationales du Réseau Français de Phonologie*, Université d'Aix-en-Provence, 2-4 juin 2005.
- COLEMAN, J., « Phonetic representations in the mental lexicon », in Durand, J. et Laks, B. (éds.), *Phonetics, Phonology and Cognition*, Oxford, OUP, 2002, pp. 96-130.
- COLLET, G., *Langage et modélisation scientifique*, Paris, CNRS éditions, 2000 (REF).
- CONTENT, A., MEUNIER, C., KEARNS, R. K. et FRAUENFELDER, U. H., « Sequence detection in pseudowords in French : where is the syllable effect ? », *Language and Cognitive Processes*, 16 (5-6), 2001, pp. 609-636.
- CORTES, J., « Les problèmes phonétiques de l'étudiant japonais », *Actes du premier colloque franco-japonais pour l'enseignement de la langue française au Japon*, 1970.
- COTE, M.-H., *Consonant cluster phonotactics : a perceptual approach*, PhD dissertation, MIT, septembre 2000, revised 2001, Rutgers Optimality Archives 548-1002, consulté le 21.03.03, <http://roa.rutgers.edu/view.php3?roa=548> (REF).
- COTE, M.-H. et MORRISON, G., « Experimental evidence and the nature of the schwa/zero alternation in French », poster, *9<sup>th</sup> Conference on Laboratory Phonology*, juin 2004.
- COULMAS, F., *The Writing Systems of the World*, Oxford, Blackwell 1989 (REF).
- CROOKSTON, I., « Alphabetic literacy and practical phonetics teaching : some preliminary connections », *Proceedings of the Phonetics Teaching and Learning Conference 2001*, Londres, UCL, 2001, pp. 7-10.
- CRYSTAL, D. *The Cambridge Encyclopedia of Language*, Cambridge, CUP, 1<sup>ère</sup> édition 1987, 2<sup>ème</sup> édition, 1997 (REF).
- CUQ, J.-P., *Une introduction à la didactique de la grammaire en français langue étrangère*, Paris, Didier/Hatier, 1996.
- CUQ, J.-P. (dir.), *Dictionnaire de didactique du français langue étrangère et seconde*, Paris, CLE international, 2003 (REF).

## D

- DABENE, L., « Enseignement précoce ou éveil au langage ? », *Le Français dans le Monde, Recherches et Applications*, numéro spécial « Enseignement/apprentissage précoce des langues », 1991, pp. 57-64.
- DAMIAN, M. F. et BOWERS, J. S., « Effect of orthography on speech production in a form-preparation paradigm », *Journal of Memory and Language*, 49, 2003, pp. 119-132.
- DAT, M.-A. et SPANGHERO-GAILLARD, N., « L'influence de la multimodalité sur la mémorisation lexicale en L2 : expériences menées auprès d'enfants en cycle

- 3 », in Billières, M., Gaillard, P. et Spanghero-Gaillard, N. (éds.), *Actes du Premier Colloque International de Didactique Cognitive, (Français Langue Etrangère / Langue Seconde / Langue Maternelle)* [CD-ROM], Université de Toulouse-le Mirail, 2005, pp. 105-110
- DE BLESER, R. et CHOLEWA, J., « Dissociations between inflection, derivation and compounding: neurolinguistic evidence for morphological fractionations within the lexical system », in Visch-Brink, E. et Bastiaanse, R., *Linguistic Levels in Aphasia*, San Diego, Singular Publishing Group, 1998, pp. 231-245.
  - DE BOT, K., « A bilingual production model: Levelt's « Speaking » model adapted ». *Applied Linguistics*, 13, 1992, pp. 1-24.
  - DEHAENE, S., NACCACHE, L., LE CLECH, G., KOECHLIN, E., MUELLER, M., DEHAENE-LAMBERTZ, G., VAN DE MOORTELE, P.-F. et LE BIHAN, D., « Imaging unconscious semantic priming », *Nature*, n°395, 1998, pp. 597-600.
  - DELAIS-ROUSSARIE, E., DURAND, J., LYCHE, C., MEQQORI, A. et TARRIER, J.-M., « Transcription des données : outils et conventions », in Durand, J., Laks, B. et Lyche, C. (éd.), *Bulletin PFC n°1, Protocole, conventions et directions d'analyse*, ERSS UMR 5610 & Université de Toulouse – Le Mirail, 2002, pp. 21-34.
  - DELL, F., *Les règles et les sons*, Paris, Hermann, 1973/1985.
  - DELL, F., « Consonant clusters and phonological syllables in French », in Durand, J. et Hintze M.-A. (éd.), *Lingua* 95, numéro spécial : « French phonology : morae, syllables, words », 1995, pp. 5-26.
  - DEMONET, J.-F., « Traitement phonologique et activation cérébrale en tomographie d'émission de positons » in Lambert, J. et Nespoulous, J.-L. (dir.), *Perception auditive et compréhension du langage : état initial, état stable et pathologie*, Marseille, Solal Editeurs, 1997, pp. 115-125.
  - DENIS, M., *Image et cognition*, Paris, P.U.F., 2<sup>e</sup> éd, 1994.
  - DERIVERY, N., *La phonétique du français*, Paris, Seuil, 1997.
  - DETEY, S., « Sciences Cognitives et Didactique des Langues: intégration des activités perceptives et métalinguistiques d'un apprenant adulte d'une langue étrangère. Fondements théoriques pour un Programme Adaptatif Intégré », Mémoire de D.E.A., Université de Toulouse-Le Mirail, 160 pages, 2002.
  - DETEY, S., « Enseigner la linguistique aux étudiants de français ? Approche cognitive », communication, *Congrès de Printemps de la Société Japonaise de Langue et Littérature Française*, résumé dans *Etudes de langue et littérature françaises*, n° 83, Tokyo, Hakusuisha, 2003a, p. 143.
  - DETEY, S., « Invariance, variations et "centration cognitive" : quelles nécessités pour la didactique des langues ? », *Revue PArole*, n° 25-26, 2003b, pp. 75-115.
  - DETEY, S., « Réalisme linguistique et/ou réalisme psycholinguistique : quelle nécessité ? Illustration phonologique franco-japonaise », Communication, *Colloque « Le langage, ce particulier, ce singulier »*, Université de Tsukuba, 13 novembre 2003c.
  - DETEY, S., « La pratique nécessite-t-elle l'analyse ? Questionnement méthodologique pour la didactique du français au Japon », *Enseignement du français au Japon*, n° 32, 2004, pp. 33-44.

- DETEY, S. et AKASHI, N., « La prononciation du français pour les apprenants japonais : peut-on faire quelque chose ? », Atelier Pédagogique, résumé dans *Compte-rendu de la 17<sup>ème</sup> Journée Pédagogique*, Tokyo, Université de Dokkyo, 2003, pp. 64-66.
- DETEY, S., DURAND, J. et NESPOULOUS, J.-L., « Apprentissage de la phonologie du français par des Japonais et multimodalité : une didactique nécessairement cognitive », in Billières, M., Gaillard, P. et Spanghero-Gaillard, N. (éd.), *Actes du Premier Colloque International de Didactique Cognitive, (Français Langue Etrangère / Langue Seconde / Langue Maternelle)* [CD-ROM], Université de Toulouse-le Mirail, janvier 2005, pp. 144-149.
- DEVELLOTTE, C., « Analyse et usages des supports multimédias : de nouveaux terrains de recherche en didactique des langues », in Koulayan, N. et Borrell, A. (éd.), *C.I.S.L. n°15*, 2000, Toulouse, Université de Toulouse-Le Mirail, pp. 19-35.
- DIJKSTRA, T., FRAUENFELDER, U. H. et SCHREUDER, R., « Bidirectional grapheme-phoneme activation in a bimodal detection task », *Journal of Experimental Psychology : human perception and performance*, Vol. 19, n°5, 1993, pp. 931-950.
- DRIVER, J. et SPENCE, C., « Multisensory perception : beyond modularity and convergence », *Current Biology*, 10, 2000, pp. 731-735.
- DUBOIS, D., « L'organisation du lexique mental et son traitement », in Nespoulous, J.-L. et Leclercq, M. (éd.), *Linguistique et Neuropsycholinguistique : tendances actuelles*, Paris, Edition de la Société de Neuropsychologie de Langue Française, 1990, pp. 35-47.
- DUBOIS, D. (dir.), *Sémantique et cognition : catégories, prototypes, typicalité*, Paris, Editions du CNRS, 1991a).
- DUBOIS, D., « Les catégories sémantiques "naturelles" : prototypes et typicalité », in Dubois, D. (dir.), *Sémantique et cognition : catégories, prototypes, typicalité*, Paris, Editions du CNRS, 1991b), pp. 14-27.
- DUBOIS, D. « Catégorisation et cognition : "10 ans après", une évaluation des concepts de Rosch », in Dubois, D. (dir.), *Sémantique et cognition : catégories, prototypes, typicalité*, Paris, Editions du CNRS, 1991c), pp. 29-54.
- DUBRAY, D. et KRAMER, V., « La méthode verbo-tonale appliquée à un groupe d'enfants sourds », *T.R.A.N.E.L.*, n°21, 1994, pp. 113-127.
- DUCHÊNE, A., « La gestion des inférences chez les cérébro-lésés droits », *Revue Parole* n° 11-12, « Langage et cognition » (1<sup>ère</sup> partie), 1999, pp. 257-287.
- DUCROT, O. et TODOROV, T., *Dictionnaire encyclopédique des sciences du langage*, Paris, Editions du Seuil, 1972.
- DUMAY, N., FRAUENFELDER, U. H., CONTENT, A., « The role of the syllable in lexical segmentation in French : word-spotting data », *Brain and Language*, Vol. 81, n° 1-3, 2002, pp. 144-161.
- DUPOUX, E., FUSHIMI, T., KAKEHI, K. et MEHLER, J., « Prelexical locus of an illusory vowel effect in Japanese », *Eurospeech '99 Proceedings, ESCA 7<sup>th</sup> European Conference on Speech Communication and Technology*, 1999.
- DUPOUX, E., KAKEHI, K., HIROSE, Y., PALLIER, C. et MEHLER, J., « Epenthetic vowels in Japanese: a perceptual illusion? », *Journal of Experimental*

- Psychology: human perception and performance*, Vol. 25, n° 6, 1999, pp. 1568-1578.
- DUPOUX, E., PALLIER, C., KAKEHI, K., MEHLER, J., « New evidence for prelexical phonological processing in word recognition », *Language and Cognitive Processes*, Vol. 16, n° 5-6, 2001, pp. 491-505.
  - DUPOUX, E., PALLIER, C., SEBASTIAN, N. et MEHLER, J., « A destressing « deafness » in French ? », *Journal of Memory and Language*, 36, 1997, pp. 406-21.
  - DURAND, J., *Generative and Non-Linear Phonology*, Londres, Longman, 1990.
  - DURAND, J., « Les traits phonologiques et le débat articulation/audition », *Points d'interrogation : Phonétique et phonologie de l'anglais*, Actes du colloque de décembre 1998, Université de Pau et des Pays de l'Adour, Publications de l'Université de Pau, 2000a.
  - DURAND, J., « Oral, écrit et faculté de langage », in Guillot, M.-N. et Kenning, M.-M. (éd.), *Changing Landscapes in Language and Language Pedagogy*, Londres, AFLS/CILT, 2000b, pp. 40-72.
  - DURAND, J., « La transcription phonétique de l'anglais », in Rotgé, W. (éd.), *Agrégation externe d'anglais : mode d'emploi*, Paris, Ellipses, 2001, pp. 181-219.
  - DURAND, J., « La phonétique classique : l'Association Phonétique Internationale et son alphabet », in Nguyen, N., Wauquier-Gravelines, S. et Durand, J. (éds.), *Phonologie et Phonétique : forme et substance*, Paris, Hermès, 2005a, ch. 1.
  - DURAND, J., « Les primitives phonologiques : des traits distinctifs ux éléments », in Nguyen, N., Wauquier-Gravelines, S. et Durand, J. (éds.), *Phonologie et Phonétique : forme et substance*, Paris, Hermès, 2005b, ch. 3.
  - DURAND, J., « Tense/Lax, the vowel system of English and phonological theory », in Carr, P., Durand, J. et Ewen, C. (éd.), *Headhood, Elements, Specification and Contrastivity*, John Benjamins, Amsterdam, 2005c, pp. 77-98.
  - DURAND, J. et LAKS, B. (éd.), *Phonetics, Phonology and Cognition*, Oxford, OUP, 2002a.
  - DURAND, J. et LAKS, B., « Phonology, phonetics and cognition », in Durand, J. et Laks, B. (éd.), *Phonetics, Phonology and Cognition*, Oxford, OUP, 2002b, pp. 1-30.
  - DURAND, J., LAKS, B. et LYCHE, C. (éd.), *Bulletin PFC n° 1*, « Protocole, conventions et directions d'analyse », Toulouse, Université de Toulouse-Le Mirail, ERSS, 2002.
  - DURAND, J., LAKS, B. et LYCHE, C. « Le projet Phonologie du Français Contemporain (PFC) », *La Tribune Internationale des Langues Vivantes*, n°33 : « La prononciation du français dans sa variation », AELPL, Editions Anagrammes, 2003, pp. 3-9.
  - DURAND, J. et LYCHE, C., « La phonologie : des phonèmes à la théorie de l'optimalité », in Escudier, P. et Schwartz, J.-L. (dir.), *La parole : des modèles cognitifs aux machines communicantes*, Paris, Hermès Science Europe, 2000, pp. 193-238.
  - DURAND, J. et LYCHE, C., « Le projet "Phonologie du Français Contemporain" (PFC) et sa méthodologie », in Delais-Roussarie, E. et Durand, J. (dir.), *Corpus et variation en phonologie du français : méthodes et analyses*, Toulouse, Presses Universitaires du Mirail, 2003, ch. 6.



## E

- ECKMAN, F. R., « Markedness and the contrastive analysis hypothesis », *Language Learning*, 31, 1977, pp. 195-216 / In Ioup, G. et Weinberger, S. H. (éd.), *Interlanguage Phonology : the acquisition of a second language sound system*, Cambridge, Newbury House, 1987, pp. 55-69.
- ECKMAN, F. R., « From phonemic differences to constraint rankings : research on second language phonology », *Studies in Second Language Acquisition*, Vol. 26, n°4, 2004, pp. 513-549.
- ELLIS, N (éd.), *Implicit and Explicit Learning of Language*, San Diego, Academic Press, 1994a.
- ELLIS, N, « Implicit and explicit language learning—an overview », in Ellis, N (éd.), *Implicit and Explicit Learning of Language*, San Diego, Academic Press, 1994b, pp. 1-31.
- ELLIS, N., « Sequencing in SLA : phonological memory, chunking and points of order », *Studies in Second Language Acquisition*, 18, 1996, pp. 91-126.
- ELLIS, N., « Emergentism, Connectionism, and Language Learning », *Language Learning*, Vol. 48, n°4, 1998, pp. 631-664.
- ELLIS, N., « Frequency effects in language processing : a review with implications for theories of implicit and explicit language acquisition », *Studies in Second Language Acquisition*, numéro spécial : « Frequency effects in language processing, a review with commentaries », Vol. 24, n°2, 2002a, pp. 143-179.
- ELLIS, N., « Reflections on frequency effects », in *Studies in Second Language Acquisition*, numéro spécial : « Frequency effects in language processing, a review with commentaries », Vol. 24, n°2, 2002b, pp. 297-339.
- ELLIS, R., « Informal and formal approaches to communicative language teaching », *ELT Journal*, Vol. 36, n°2, 1982, pp. 73-81.
- ELLIS, R., *Instructed Second Language Acquisition*, Blackwell, 1990.
- ELLIS, R., « A theory of instructed second language acquisition », in Ellis, N. (éd.), *Implicit and Explicit learning of language*, San Diego, Academic Press, 1994, pp. 79-113.
- ELLIS, R., « Does form-focused instruction affect the acquisition of implicit knowledge ? A review of the research », *Studies in Second Language Acquisition*, numéro spécial : « Frequency effects in language processing, a review with commentaries », Vol. 24, n°2, 2002, pp. 223-236.
- ERICSSON, K. A. et KINTSCH, W., « Long-term working memory », *Psychological Review*, 102, 1995, pp. 211-245.
- ESCUDERO, P. et BOERSMA, P., « Bridging the gap between L2 speech perception research and phonological theory », *Studies in Second Language Acquisition*, Vo. 26, n°4, 2004, pp. 551-585.

- ESCUDERO, P., « Revisiting the role of perception in segmental phonology », in Kerkhoff, K., Lange (de) J. et Sadeh Leicht, O. (éd.), *UiL OTS Yearbook 2004*, Utrecht, Utrecht Institute of Linguistics, 2004.
- EYCHENNE, J., *Prolégomènes à une étude comparative du schwa en français : aspects méthodologiques, empiriques et théoriques*, Mémoire de DEA, Université de Toulouse-Le Mirail, 2003a (REF).
- EYCHENNE, J., « Le schwa en français du Midi : quelques aspects », poster, Colloque *Des représentations aux contraintes*, Université de Toulouse-Le Mirail, 2003b.

## F

- FABBRO, F., *The Neurolinguistics of Bilingualism : an introduction*, Hove, Psychology Press Ltd, 1999.
- FARACO, M., « Geste et fonctionnalité dans une séquence narrative en langue seconde », *Travaux Interdisciplinaires du Laboratoire Parole et Langage d'Aix-en-Provence*, Vol. 19, Université d'Aix-en-Provence, 2001, pp. 63-72.
- FERRAND, L. et NEW, B., « Syllabic length effects in visual word recognition and naming », *Acta Psychologica*, 113, 2003, pp. 167-183.
- FISCHER, D., F., « Understanding the reading process through the use of transformed typography : PSG, CSG and automaticity », in Kolers, P. A., Wroldstad, M. E. et Bouma, H. (éd.), *Processing of Visible Language*, Vol. 1, New York, Plenum Press, 1979, pp. 77-88.
- FISHER, C., CHURCH, B. A. et CHAMBERS, K. E., « Learning to identify spoken words », in Hall, D. G. et Waxman, S. R. (éd.), *Weaving a Lexicon*, Cambridge, MIT Press, 2004.
- FLEGE, J. E. et HILLENBRAND, J., « Limits on phonetic accuracy in foreign language speech production », in Ioup, G. et Weinberger, S. H. (éd.), *Interlanguage Phonology : the acquisition of a second language sound system*, Cambridge, Newbury House, 1987, pp. 176-203.
- FLEISCHACKER, H., « Cluster-dependent epenthesis asymmetries », UCLA, Mémoire de Mastère, 2000 (REF).
- FOWLER, C. et DEKLE, D.J., « Listening with eye and hand : crossmodal contributions to speech perception », *Journal of Experimental Psychology : human perception and performance*, 17, 1991, pp. 816-28.
- FLEMMING, E., « Speech perception and phonological contrast », in Pisoni, D. et Remez, R. (éds.), *The Handbook of Speech Perception*, Blackwell, 2005.
- FRANCOIS, D., « La notion de norme en linguistique », in Martinet, J. (dir.), *De la théorie linguistique à l'enseignement de la langue*, Paris, PUF, 1974, pp. 145-159.
- FRASER, H., « Pronunciation as communication : introducing a new CD-ROM for teaching ESL pronunciation », *Australian Language Matters*, Vol. 8, n°3, 2000, pp. 10-12.

- FROST, R. et GRAINGER, J., « Cross-linguistic perspectives on morphological processing : an introduction », *Language and Cognitive processes*, Vol. 15, n°4-5, 2000, pp. 321-328.
- FROST, R. et KATZ, L. (éds), *Orthography, Phonology, Morphology and Meaning*, Amsterdam, Elsevier Science publishers, 1992.
- FUCHS, C. et ROBERT, S. (éd.), *Diversité des langues et représentations cognitives*, Paris, Editions Ophrys, 1997 (REF).
- FUKAZAWA, H., KITAHARA, M. et OTA, M., « Lexical stratification and Ranking invariance in constraint-based grammars », *Proceedings of the CLS*, 34, 1998, Rutgers Optimality Archives 267-0698, consulté le 27.03.03, <http://roa.rutgers.edu/view.php3?roa=267>.

## G

- GALLOWAY, L. M., « Etudes cliniques et expérimentales sur la répartition hémisphérique du traitement cérébral du langage chez les bilingues : modèles théoriques », in Paradis, M. et Lebrun, Y. (Coord.), *Langages*, n°72, numéro spécial : « La neurolinguistique du bilinguisme », 1983, pp. 79-113.
- GAONAC'H, D., « Psychologie et didactique des langues : perspectives de recherche en psychologie du langage », in Béacco, J.-C. et Chevalier, J.-C. (Coord.), *Etudes de Linguistique Appliquée* n°72, « Didactique des langues : quelles interfaces ? », Paris, Didier Erudition, 1988, pp. 83-93.
- GAONAC'H, D.(Coord.), *Acquisition et utilisation d'une langue étrangère, l'approche cognitive*, Paris, Hachette, 1990a.
- GAONAC'H, D., « Les stratégies attentionnelles », in Gaonac'h, D.(Coord.), *Acquisition et utilisation d'une langue étrangère, l'approche cognitive*, Paris, Hachette, 1990b.
- GATHERCOLE, S. E., « The structure and functioning of phonological short-term memory », *Surdité et accès à la langue écrite. De la recherche à la pratique*, Actes du colloque ACFOS 2, 27-28 novembre 1998, Paris, ACFOS.
- GATHERCOLE, S. E., FRANKISH, C. R., PICKERING, S. J. et PEAKER, S., « Phonotactic influences on short-term memory », *Journal of Experimental Psychology : learning, memory and cognition*, Vol. 25, n°1, 1999, pp. 84-95.
- GERMAIN, A., et MARTIN, P., « Présentation d'un logiciel de visualisation pour l'apprentissage de l'oral en langue seconde », *ALSIC*, Vol. 3, n°1, 2000, pp. 61-76.
- GIRARD, F. A. et LYCHE, C., « Norme, variation et enseignement du FLE », ms, 2003, <http://www.uio.no/studier/emner/hf/ilos/FRA1103/h04/undervisningsmateriale/FRA1103-norme.html> (consulté en mai 2005).
- GOLDSMITH, J. (éd.), *The Handbook of Phonological Theory*, Oxford, Blackwell, 1995a (REF).
- GOLDSMITH, J., « Phonological theory », in Goldsmith, J. (éd.), *The Handbook of Phonological Theory*, Oxford, Blackwell, 1995b, pp. 1-23.

- GOLDBERG, E. et COSTA, L. D., « Hemisphere differences in the acquisition and use of descriptive systems », *Brain and Language* n° 14, 1981, pp. 144-173.
- GOMBERT, J.-E., « Implicit and explicit learning to read : implication as for subtypes of dyslexia », *Current Psychology Letters/Behaviour, Brain and Cognition*, Numéro Spécial : « Language disorders and reading acquisition », 10, Vol. 1, 2003.
- GORDON, H. W. et WEIDE, R., « La contribution de certaines fonctions cognitives au traitement du langage, à son acquisition et à l'apprentissage d'une langue seconde », in Paradis, M. et Lebrun, Y. (Coord.), *Langages*, n°72, numéro spécial : « La neurolinguistique du bilinguisme », 1983, pp. 45-56.
- GORDON, M. B. et PAUGAM-MOISY, H. (dir.), *Sciences cognitives : diversité des approches*, Paris, Hermès, 1997.
- GOUGENHEIM, G., MICHEA, P., RIVENC, P. et SAUVAGEOT, A., *L'élaboration du français élémentaire*. Paris, Didier, 1956.
- GRAINGER, J., NGUYEN VAN KANG, M. et SEGUI, J., « Cross-modal repetition priming of heterographic homophones », *Memory and Cognition*, Vol. 29, n°1, 2001, pp. 53-61.
- GRANTZ, K. W. et SEITZ, P.-F., « The use of visible speech cues for improving auditory detection of spoken sentences », *Journal of the Acoustical Society of America*, Vol. 108, n°3, 2000, pp. 1197-1208.
- GREMMO, M. -J. et HOLEC, H., « La compréhension orale : un processus et un comportement », in Gaonac'h, D. (coord.), *Acquisition et utilisation d'une langue étrangère, l'approche cognitive*, Paris, Hachette, 1990, pp. 30-41.
- GRIJZENHOUT, J. et VAN ROOY, B., « Second language phonology : acquisition through gradual constraint demotion », working paper, *Theorie des Lexicons*, Université de Düsseldorf, 2001.
- GUIMBRETIERE, E., « Phonétique du français et enseignement : propositions pour une mise en forme didactique des savoirs phonétiques », *Enseignement du français au Japon*, n°23, 1994, pp. 8-17.
- GUIMBRETIERE, E., « La prosodie : un passage obligé pour la compréhension orale », *Le Français dans le Monde, Recherches et Applications*, numéro spécial : « Oral : variabilité et apprentissages », 2001, pp. 153-159.
- GULLBERG, M., « Communication strategies, gestures, and grammar », *Acquisition et Interaction en Langue Etrangère (AILE)*, numéro spécial, Actes du 8<sup>ème</sup> Colloque EUROSLA Paris, Volume 2 « Du mot à la structure », Université Paris VIII, 1999, pp. 61-70.
- GUPTA, P. et MACWHINNEY, B., « Is the phonological loop articulatory or auditory ? », *Proceedings of the 15<sup>th</sup> Conference of the Cognitive Science Society*, Lawrence Erlbaum, Hillsdale, New Jersey, 1993.

## H

- HALLE, P., SEGUI, J., FRAUENFELDER, U. et MEUNIER, C., « Processing of illegal consonant clusters : a case of perceptual assimilation ? », *Journal of Experimental Psychology : human perception and performance*, Vol. 24, n°2, 1998, pp. 592-608.
- HAMON, J.-F. (dir.), *Attention, vigilance, émotion*, Université de la Réunion, Paris, L'Harmattan, 1998a.
- HAMON, J.-F., « Concepts et domaines de l'attention », in Hamon, J.-F. (dir.), *Attention, vigilance, émotion*, Université de la Réunion, Paris, L'Harmattan, 1998b, pp. 17-30.
- HAMON, J.-F., « Vigilance et attention », in Hamon, J.-F. (dir.), *Attention, vigilance, émotion*, Université de la Réunion, Paris, L'Harmattan, 1998c.
- HAMON, J.-F. et JUAN DE MENDOZA, J.-L., « Les niveaux de contrôle attentionnel au cours du traitement de l'information auditive chez l'homme », in Hamon, J.-F. (dir.), *Attention, vigilance, émotion*, Université de la Réunion, Paris, L'Harmattan, 1998, pp. 59-82.
- HAMMOND, M., « Optimality theory and prosody », in Archangeli, D. et Langendoen, D. T. (éd.), *Optimality Theory : an overview*, Malden, Blackwell Publisher, 1997, pp. 33-58.
- HANDSCHIN, K., « L'influence de la langue de base dans la perception des alternances codiques : le cas de la consonne initiale du mot », *T.R.A.N.E.L.*, n° 21, 1994, pp. 51-59.
- HARDISON, D. M., « Bimodal speech perception by native and nonnative speakers of English : Factors influencing the McGurk effect », *Language Learning*, 49, 1999, pp. 213-283.
- HARDISON, D. M., « The neurocognitive foundation of second-language speech : a proposed scenario of bimodal development », in Swierzbien, B., Morris, F., Anderson, M. E., Klee, C. A. et Tarone, E. (éd.), *Social and Cognitive Factors in Second Language Acquisition*, Somerville, Cascadilla Press, 2000, pp. 312-325.
- HARDISON, D. M., « Acquisition of second-language speech : effects of visual cues, context, and talker variability », *Applied Psycholinguistics*, 24, 2003, pp. 495-522.
- HARNAD, S. (éd.), *Categorical Perception : the groundwork of cognition*, Cambridge, CUP, 1987 (REF).
- HARRIS, J., « Phonological universals and phonological disorders », in Visch-Brink, E. et Bastiaanse, R., *Linguistic Levels in Aphasia*, San Diego, Singular Publishing Group, 1998, pp. 91-115.
- HAUNZ, C., « The role of speech perception in the adaptation of loanwords », *Communication, Post-Graduate Conference (PGC) 2002*, University of Edinburgh, 2002.
- HAZAN, V., ANKE, S. et FAULKNER, A., « Audiovisual perception in L2 learners », *Proceedings of International Conference for Spoken Language Processing*, Denver, 2002, pp. 1685-1688.

- HIM CHEUNG, H.-C. et CHEN, H.-C., « Early orthographic experience modifies both phonological awareness and on-line speech processing », *Language and Cognitive Processes*, Vol. 19, n°1, 2004, pp. 1-28.
- HIROSE, Y. et DUPOUX, E., « Ignoring the signal : evidence from vowel epenthesis », poster, *AMLaP 2004*, 16-18 septembre 2004, Université d'Aix-en-Provence.
- HOLM, A. et DODD, B., « The effect of first written language on the acquisition of English literacy », *Cognition*, 59, 1996, pp. 119-147.
- HOYT, E., « Explicit instruction in the communicative method : pedagogical approaches for successfully teaching the sound system of a second language », 2001, ms, <http://www.stolaf.edu/depts/cis/wp/hoyt/phonpedigogy2.html> (consulté le 24.04.03).
- HUME, E. et JOHNSON, K., « A model of the interplay of speech perception and phonology », in Hume, E. et Johnson, K. (éd.), *The Role of Speech Perception in Phonology*, New York, Academic Press, 2001, pp. 3-26.

## I

- INTRAVAIA, P., *Formation des professeurs de langue en phonétique corrective : le système verbo-tonal*, Paris, Didier-Erudition, Mons, CIPA, 2000.
- ISHIKAWA, F., *L'interaction exolingue : analyse de phénomènes métalinguistiques. Continuité et discontinuité entre situation d'enseignement / apprentissage et situation « naturelle »*, Thèse de Doctorat, Université Paris III – Sorbonne Nouvelle, 2001 / Tokyo, Edition Shumpûsha, 2002. (REF).
- ITO, J. et MESTER, R. A., « Licensed segments and safe paths », in Paradis, C. et La Charité, D. (éd.), *The Canadian Journal of Linguistics*, Vol. 38, n°2, numéro spécial : « Constraint-based theories in multilinear phonology », 1993, pp. 197-213.
- ITO, J. et MESTER, R. A., « Japanese phonology », in Goldsmith, J. (éd.), *The Handbook of Phonological Theory*, Oxford, Blackwell, 1995, pp. 817-835.
- ITO, J. et MESTER, R. A., « The phonological lexicon », in Tsujimura, N. (éd.), *The Handbook of Japanese Linguistics*, Oxford, Blackwell, 1999, pp. 62-100.
- IVERSON, P., KUHL, P. K., AKAHANE-YAMADA, R., DIESCH, E., TOHKURA, Y., KETTERMANN, A. et SIEBERT, C., « A perceptual interference account of acquisition difficulties for non-native phonemes », *Cognition*, 87, 2003, pp. B47-B57.

## J

- JAMET, E., « Les effets de l'intégration spatiale sur le traitement cognitif des illustrations statiques et dynamiques », Communication, Journée Doctorale du Laboratoire J. Lordat, « Multimédia et Evaluation », Université de Toulouse-Le Mirail, 18 mai 2001.
- JOANETTE, Y. et ROCH-LECOURS, « Cerveau et Langage : nouvelles perspectives », *Méta*, Vol. 29, n°1, 1984, pp. 27-36.
- JUAN DE MENDOZA, J.-L., « Le paradigme de l'écoute dichotique : quelques réflexions à propos d'un modèle des mécanismes de l'attention », in Hamon, J.-F. (dir.), *Attention, vigilance, émotion*, Université de la Réunion, Paris, L'Harmattan, 1998, pp. 83-102.
- JACOBS, H. et GUSSENHOVEN, C., « Loan phonology : perception, Saliency, the lexicon and OT », in Dekkers, J., Leeuw, F. van der et Weijer, J. van de (dir.), *Optimality Theory*, Oxford, OUP, 2000, pp. 193-209.
- JAFFRE, J.-P., « La linguistique et la lecture-écriture : de la conscience phonologique à la variable "orthographe" », *Revue des Sciences de l'Education*, Vol. XXIX, n°1, 2003, pp. 37-49.
- JARVIS, H., « The changing role of computers in language teaching and the case for « study skills », *Modern English Teacher*, Vol. 9, n° 1, 2000, pp. 62-68.
- JETCHEV, G., « Le schwa français au miroir de la variation », *La Tribune Internationale des Langues Vivantes*, n°33 : « La prononciation du français dans sa variation », AELPL, Editions Anagrammes, 2003, pp. 102-107.
- JOHNSON, N. F., « On the role of cohorts or neighbors in visual word recognition », in Frost, R. et Katz, L. (éd.), *Orthography, Phonology, Morphology and Meaning*, Elsevier Science Publishers, 1992, pp. 147-164.
- JUFFS, A., « Formal linguistic perspectives on second language acquisition », in Kaplan, R. B. (éd.), *The Oxford Handbook of Applied Linguistics*, Oxford, OUP, 2002, pp. 87-103.

## K

- KAGER, R., *Optimality Theory*, Cambridge, CUP, 1999 (REF).
- KAIL, M., « Le traitement des données de langage : la prise des indices et leur utilisation », in Gaonac'h, D. (coord.), *Acquisition et utilisation d'une langue étrangère, l'approche cognitive*, Paris, Hachette, 1990, pp. 70-80.
- KAIL, M., « Variations linguistiques et contraintes cognitives dans le traitement et le développement du langage », in Fuchs, C. et Robert, S., *Diversité des langues et représentations cognitives*, Paris, Editions Ophrys, 1997, pp. 210-225.
- KAPLAN, R. B. (éd.), *The Oxford Handbook of Applied Linguistics*, Oxford, OUP, 2002 (REF).
- KAMIYA, T., *Tuttle New Dictionary of Loanwords in Japanese*, Rutland (Vermont), Charles E. Tuttle, 1994.

- KANEMAN-POUGATCH, M. et PEDOYA-GUIMBRETIERE, E., *Plaisir des sons*, Hatier, Paris, 1989.
- KATZ, L. et FROST, R., « The reading process is different for different orthographies : the orthographic depth hypothesis », in Frost, R. et Katz, L. (éd.), *Orthography, Phonology, Morphology and Meaning*, Elsevier Science Publishers, 1992, pp. 67-84.
- KEATING, P. and HUFFMAN, M. K., « Vowel variation in Japanese », *Phonetica*, 41, 1984, pp. 191-207.
- KEKENBOSCH, C., *La mémoire et le langage*, Paris, Nathan, 1994.
- KELLER, E., *Introduction aux systèmes psycholinguistiques*, Gaëtan Morin, Québec, 1985
- KENSTOWICZ, M., « Salience and similarity in loanword adaptation : a case study from Fijian », ms, MIT, 2003.
- KESS, J. F. et MIYAMOTO, T., *The Japanese Mental Lexicon : psycholinguistic studies of kana and kanji processing*, Amsterdam, John Benjamins, 1999.
- KEYS, K. J., « Interlanguage phonology. Theoretical questions and empirical data », *Linguagem & Ensino*, Vol. 5, N° 1, 2002, pp. 75-91.
- KIRCHNER, R., « Exemplar-based phonology and the time problem : a new representational technique », poster, *9<sup>th</sup> Conference on Laboratory Phonology*, University of Illinois, 24-26 juin 2004.
- KLEIBER, G., « Prototype et prototypes : encore une affaire de famille », in Dubois, D. (dir), *Sémantique et cognition : catégories, prototypes, typicalité*, Paris, Editions du CNRS, 1991, pp. 103-129.
- KLEIN, W., *Second Language Acquisition*, Cambridge, CUP, 1986.
- KLIR, G. J., *Fuzzy Sets*, Toulouse, Diderot Multimédia, 1996.
- KLOSE, K., « Perception catégorielle et sujets cérébro-lésés : données préliminaires », *T.R.A.N.E.L*, n°21, 1994, pp. 143-154.
- KOLERS, P. A., WROLSTAD, M. E. et BOUMA, H. (éd.), *Processing of Visible Language*, Vol. 1, New York, Plenum Press, 1979 (REF).
- KOLINSKY, R., MORAIS, J. et CLUYTENS, M., « Intermediate representations in spoken words recognition : evidence from word illusions », *Journal of Memory and Language*, 34, 1995, pp. 19-40.
- KONDO, T. et MAZUKA, R., « Prosodic planning while reading aloud : on-line examination of Japanese sentences », *Journal of Psycholinguistic Research*, Vol. 25, n°2, 1996, pp. 357-381.
- KONDO, T., AMANO, S. et MAZUKA, R., « Japanese lexicon databases for psycholinguistic research », (poster ) *9<sup>th</sup> Annual CUNY Conference on Human Sentence Processing*, New York, 22 mars 1996.
- KOSSLYN, S. M., « Mental imagery », in Osherson D. N., Kosslyn, S. M. et Hollerbach, J. M. (éd.), *Visual Cognition and Action ; an invitation to cognitive science*, vol. 2, MIT, 1990, pp. 73-96.
- KOUIDER, S. et DUPOUX, E., « A functional disconnection between spoken and visual word recognition: evidence from unconscious priming », *Cognition*, 82, 2001, B35-49.



- KOULAYAN, N., « Les NTIC : des outils pédagogiques ? Vers la notion « d'outils techno-pédagogiques » », *Actes du Séminaire SEFASE*, Paris-Singapour, 1999.
- KOULAYAN, N. et BORRELL, A. (éd.), Actes du Colloque International « Le multimédia et l'apprentissage du français langue étrangère et/ou seconde : de l'expérimentation à la réalité dans l'espace francophone » (mars 1999), *Cahiers du Centre Interdisciplinaire des Sciences du Langage* n°15, 2000, Toulouse, Université de Toulouse-Le Mirail.
- KOULAYAN, N., DETEY, S. et SAGAZ, M., « Les escamots d'aide sur cédéroms sont-ils bien perçus par les apprenants de langue étrangère ? Enquête préliminaire en France, en Angleterre et au Japon. », *ALSIC*, Vol. 6, n°2, 2003, pp. 101-115.
- KOULAYAN, N. et SAYAH, M., « Le parler urbain des *Moujs* toulousains », *Revue des Sciences Sociales*, n° 29, « Civilités, incivilités », 2002.
- KREMIN, H., « Perturbations lexicales : les troubles de la dénomination », in Séron, X. et Jeannerod, M. (dir.), *Neuropsychologie humaine*, Bruxelles, Mardaga, 1994, pp. 375-386.
- KROLL, J. F. et DIJKSTRA, T., « The bilingual lexicon », in Kaplan, R. B. (éd.), *The Oxford Handbook of Applied Linguistics*, Oxford, OUP, 2002, pp. 301-321.
- KUBOZONO, H., « The mora and the syllable structure in Japanese : evidence from speech errors », *Language and Speech*, 32, 1989, pp. 249-278.
- KUBOZONO, H., « Epenthetic vowels and accent in Japanese : facts and paradoxes », in Weijer, J. van der et Nishihara, T. (éd.), *Issues in Japanese Phonology and Morphology*, Berlin, Mouton de Gruyter, 2001, pp. 111-140.

## L

- LABRUNE, L., « Phonologie du japonais », in Blin, R. et Tamba, I. (éd.), *Faits de Langues « coréen-japonais »*, Paris, Ophrys, 2001a, pp. 89-110.
- LABRUNE, L., « Structure de la syllabe japonaise », in Blin, R. et Tamba, I. (éd.), *Faits de Langues « coréen-japonais »*, Paris, Ophrys, 2001b, pp. 112-122.
- LABRUNE, L., « Japanese /r/, the emergence of the unmarked, and cross-linguistic patterns of liquid initial prohibition », ms, à paraître.
- LACHARITE, D. et PARADIS, C., « The emergence of constraints in generative phonology and a comparison of three current constraint-based models », in Paradis, C. et La Charité, D. (éd.), *The Canadian Journal of Linguistics*, Vol. 38, n°2, numéro spécial : « Constraint-based theories in multilinear phonology », 1993, pp. 127-150.
- LACHARITE, D. et PARADIS, C., « Derivational residue : hidden rules in optimality theory », in Dekkers, J., Leeuw, F. van der et Weijer, J. van de (dir.), *Optimality Theory*, Oxford, OUP, 2000, pp. 211-233.
- LADO, R., *Linguistics Across Cultures*, Ann Arbor : University of Michigan Press, 1957/ In Gass, S. et Selinker, L., *Language Transfer In Language Learning*, Rowley, Newbury House, 1983.

- LAKS, B., « A connectionist account of French syllabification », in Durand, J. et Hintze M.-A. (éd.), *Lingua*, 95, numéro spécial : « French phonology : morae, syllables, words », 1995, pp. 169-204.
- LAKS, B., *Langage et cognition, l'approche connexionniste*, Paris, Hermès, 1996.
- LAKS, B., « Enquête sur la phonologie du français contemporain », *Le Français dans le Monde, Recherches et Applications*, numéro spécial : « Oral : variabilité et apprentissages », 2001, pp. 106-108.
- LAKS, B., « Les grandes enquêtes phonologiques en France », *La Tribune Internationale des Langues Vivantes*, n°33 : « La prononciation du français dans sa variation », AELPL, Editions Anagrammes, 2003a, pp. 10-17.
- LAKS, B., « L'oral et la variation », *La Tribune Internationale des Langues Vivantes*, n°33 : « La prononciation du français dans sa variation », AELPL, Editions Anagrammes, 2003b, pp. 96-101.
- LAMBACHER, S. G., « Improving Japanese Pronunciation of Segmentals Using Electronic Visual Feedback », *Communication, JALT 2002*, Shizuoka (Japon), novembre 2002.
- LAMBERT, J., « Troubles de la perception du langage parlé : approche cognitive et orientations thérapeutiques », in Lambert, J. et Nespoulous, J.-L. (dir.), *Perception auditive et compréhension du langage : état initial, état stable et pathologie*, Marseille, Solal Editeurs, 1997, pp. 151-164.
- LAMBERT, J. et NESPOULOUS, J.-L. (dir.), *Perception auditive et compréhension du langage : état initial, état stable et pathologie*, Marseille, Solal Editeurs, 1997.
- LANCIEN, T., *Le multimédia*, Paris, CLE International, 1998.
- LANDRAGIN, F., « Saillance physique et saillance cognitive », *CORELA*, Vol. 2, n° 2, 2004.
- LANGE, M., « Activation of multiple phoneme associates of graphemes in visual word recognition », *Brain and Language*, Vol. 81, n°1-3, 2002, pp. 610-620.
- LANGE, M. et CONTENT, A., « Length effects in pseudowords reading : letters or graphemes ? », *Communication, Annual Meeting of the Belgian Psychological Society*, Leuven, 8 mai 1998.
- LANGE, M. et CONTENT, A., « The grapho-phonological system of written French : statistical analysis and empirical validation », *Proceedings of the 37<sup>th</sup> Conference of the Association for Computational Linguistics*, Morristown, ACL, 1999, pp. 436-442.
- LAPAIRE, J.-R., *Imaginative Grammar*, Communication personnelle, présentée à Lodz (Pologne), 2001.
- LEATHER, J. et JAMES, A., « The acquisition of second language speech », *Studies in Second Language Acquisition*, 13, 1991, pp. 305-314.
- LEE, J. F., « The initial impact of reading as input for the acquisition of future tense morphology in Spanish », in Gass, S., Bardovi-Harlig, K., Sieloff Magnan, S. et Walz, J. (éd.), *Pedagogical Norms for Second and Foreign Language Learning and Teaching*, Amsterdam, John Benjamins, 2002, pp. 119-140.

- LEMERCIER, C. et TRICOT, A., « Multimédia, compréhension and the psychology of learning : a review of four cognitive models », manuscript, en soumission.
- LEON, P., *Prononciation du français standard : aide-mémoire d'orthoépée*, Paris, Didier, 1966, 4<sup>ème</sup> éd. 1978.
- LEON, P., *Phonétisme et prononciation du français*, Paris, Nathan, 1992.
- LEON, P. et M., *Introduction à la phonétique corrective*, 2<sup>nde</sup> édition, Hachette-Larousse, 1976.
- LEON, P. et M., *La Prononciation du français*, Coll.128, Nathan Université, 1997.
- LEFTON, L. A., « Reading and searching through geometrically transformed text », in Kolers, P. A., Wrolstad, M. E. et Bouma, H. (éd.), *Processing of Visible Language*, Vol. 1, New York, Plenum Press, 1979, pp. 91-96.
- LEVELT, W. J. M., *Speaking, from Intention to Articulation*, Cambridge, M.I.T. Press, 1989 (REF).
- LEVELT, W. J. M., « Spoken word production : a theory of lexical access », *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 98, n° 23, 2001, pp. 13464-13471.
- LEVELT W. J. M. et WHEELDON, L., « Do speakers have access to a mental syllabary ? », *Cognition*, 50, 1994, pp. 239-69.
- LEVY, M., *Computer-Assisted Language Learning : context and conceptualization*, Oxford, Clarendon Press, 1997.
- LEVY-SCHOEN, A. et O'REAGAN, K., « Control of Eye movements in reading », in Kolers, P. A., Wrolstad, M. E. et Bouma, H. (éd.), *Processing of Visible Language*, Vol. 1, New York, Plenum Press, 1979, pp. 7-30.
- LIBERMAN, A. M., « The relation of speech to reading and writing », in Frost, R. et Katz, L. (éd.), *Orthography, Phonology, Morphology and Meaning*, Elsevier Science Publishers, 1992, pp. 167-177.
- LIBERMAN, A.M. et MATTINGLY, I.G., « The motor theory of speech perception revised », *Cognition*, 21, 1985, pp. 1-36.
- LIEBERMAN, P. et BLUMSTEIN, S.E., *Speech Physiology, Speech Perception, and Acoustic Phonetics*, Bath, CUP, 1988 (REF).
- LIN, Y.-H., « Interphonology variability : sociolinguistic factors affecting L2 simplification strategies », *Applied Psycholinguistics*, Vol. 24, n° 4, 2003, pp. 439-464.
- LOMBARDI, L., « Evidence for MaxFeature constraints from Japanese », ms, janvier 1998, Rutgers Optimality Archives 247-0298, consulté le 21.03.03, <http://roa.rutgers.edu/files/247-0298/roa-247-lombardi-3.pdf>
- LOMBARDI, L., « Second language data and constraints on manner : explaining substitutions for the English interdental », ms, septembre 2000, Rutgers Optimality Archives 418-0900, consulté le 21.03.03, <http://roa.rutgers.edu/view.php3 ? roa=418>.
- LOMBARDI, L., « Markedness and the typology of epenthetic vowels », 2002, Rutgers Optimality Archives 578-0203, ms, consulté le 21.03.03, <http://roa.rutgers.edu/view.php3 ? roa=578>.

- LOONIS, E., *Les Stats pour les Nuls* [CD-ROM], Version 1.0., cédérom, Eric Loonis, 2002.
- LUELSDORFF, P.A. (éd.), *Orthography and Phonology*, Amsterdam, John Benjamins, 1987 (REF).

## **M**

- McCLELLAND, J. L., « The programmable blackboard model of reading », in McClelland, J. L., Rumelhart, D. E. et le PDP Research Group (dir.), *Parallel Distributed Processing :explorations in the microstructure of cognition, Vol. 2*, Cambridge, MIT press, 1986, pp. 122-128.
- McCLELLAND, J. L. et ELMAN, J. L., « Interactive process in speech perception : the TRACE model », in McClelland, J. L., Rumelhart, D. E. et le PDP Research Group (dir.), *Parallel Distributed Processing :explorations in the microstructure of cognition, Vol. 2*, Cambridge, MIT press, 1986, pp. 58-121.
- McCLELLAND, J. L. et RUMELHART, D. E., « A distributed model of human learning and memory », in McClelland, J. L., Rumelhart, D. E. et le PDP Research Group (dir.), *Parallel Distributed Processing :explorations in the microstructure of cognition, Vol. 2*, Cambridge, MIT press, 1986, pp. 170-215.
- McCLELLAND, J. L. et RUMELHART, D. E. and the PDP Research Group (dir.), *Parallel Distributed Processing :explorations in the microstructure of cognition, Vol. 2 : Psychological and biological models*, Cambridge, MIT press, 1986 (REF).
- McCONKIE, G., « On the role and control of eye movements in reading », in Kolars, P. A., Wrolstad, M. E. et Bouma, H. (éd.), *Processing of Visible Language*, Vol. 1, New York, Plenum Press, 1979, pp. 37-48.
- MAC NEILAGE, P. F., « The Frame/Content theory of evolution of speech production », *Behavioral and Brain Sciences*, 21, 1998, pp. 499-546.
- MAC WHINNEY, B., « Emergentist approaches to language », in Bybee, J. et Hopper, P. (éd.), *Frequency and the Emergence of Linguistic Structure*, TSL 45, Amsterdam, John Benjamins, 2001, pp. 449-470.
- MAJOR, R.C., «The phonology of the L2 user », in Cook, V. (éd.), *Portraits of the L2 user*, Clevedon, Multilingual matters, 2002, pp. 65-92.
- MARTIN, R., *Langage et Croyance*, Sprimont, Pierre Mardaga éditeur, 1987 (REF).
- MARTINET, A., « Le parler et l'écrit », in Martinet, J. (Dir.), *De la théorie linguistique à l'enseignement de la langue*, Paris, PUF, 1974a, pp. 53-71.
- MARTINET, A., « Langue parlée et code écrit », in Martinet, J. (Dir.), *De la théorie linguistique à l'enseignement de la langue*, Paris, PUF, 1974b, pp. 73-82.
- MARTINET, A., « Langue parlée et langue écrite », *Liaison alfonic*, Vol. 3, fasc. 3, Paris, 1986, pp. 3-15.
- MASSARO, D. W., « Reading and Listening », in Kolars, P. A., Wrolstad, M. E. et Bouma, H. (éd.), *Processing of Visible Language*, Vol. 1, New York, Plenum Press, 1979, pp. 331-351.

- MASSARO, D. W., « Categorical partition : a fuzzy-logical model of categorization behavior » in Harnad, S. (éd.), *Categorical Perception : the groundwork of cognition*, Cambridge, CUP, 1987, pp. 254-283.
- MASSARO, D. W., « Speech perception », in Smelser, N. M. et Baltes, P. B. (éd.) et Kintsch, W. (section éd.), *International Encyclopedia of Social and Behavioral Sciences*, Amsterdam, Elsevier, 2001, pp. 14870-14875.
- MASSARO, D. W. et COHEN, M. M., « Visual, orthographic and lexical influences in reading », *Journal of Experimental Psychology : human perception and performance*, Vol. 20, n° 6, 1994, pp. 1107-1128.
- MASSARO, D. W., COHEN, M. M. et SMEELE, P. M. T., « Perception of asynchronous and conflicting visual and auditory speech », *Journal of the Acoustical Society of America*, Vol. 100, n°3, 1996, pp. 1777-1786.
- MASSARO, D. W. et LIGHT, J., « Read my tongue movements : bimodal learning to perceive and produce non-native speech /r/ and /l/ », in *Proceedings of the 8<sup>th</sup> European Conference on Speech Communication and Technology, (Eurospeech'03/Interspeech'03)*, Genève, Suisse, 2003.
- MATSUNAGA, S., *The linguistic and psycholinguistic nature of kanji : do kanji represent and trigger only meanings ?*, PhD Thesis, Ann Arbor : Dissertation Information Service, 1994.
- MATTINGLY, I. G., « Linguistic awareness and orthographic form », in Frost, R. et Katz, L. (éds.), *Orthography, Phonology, Morphology and Meaning*, Elsevier Science Publishers, 1992, pp. 11-26.
- MAZET, C., « Fonctionnalité dans l'organisation catégorielle », in Dubois, D. (dir.), *Sémantique et cognition : catégories, prototypes, typicalité*, Paris, Editions du CNRS, 1991, pp. 89-100.
- MEDIN, L. D. et BARSALOU, L. W., « Categorization processes and categorical perception », in Harnad, S. (éd.), *Categorical Perception : the groundwork of cognition*, CUP, 1987, pp. 455-487.
- MERLET, S., « Niveaux de traitement et intégration des informations multimédias. L'exemple de la compréhension orale en langue étrangère », in Tricot, A. et Rouet, J.-F. (Eds.), « Les hypermédias : approches cognitives et ergonomiques », *Hypertextes et Hypermédias*, numéro hors série, Hermès, 1998, pp. 141-156.
- MERLET, S. et GAONAC'H, D., « Mise en évidence de stratégies compensatoires dans la compréhension orale en L.E. », *Revue de Phonétique Appliquée* n°115-116-117, « Psychologie et Didactique » (2<sup>ème</sup> partie), 1995, pp. 273-292.
- MEULEMANS, T., *L'apprentissage implicite : une approche cognitive, neuropsychologique et développementale*, Marseille, Solal Editeurs, 1998.
- Ministère de l'éducation nationale, Direction de la coopération, *Le français fondamental (1<sup>er</sup> degré)*, Paris, Institut Pédagogique National, 1959, 3<sup>ème</sup> édition, 1970.
- MOESCHLER, J. et AUCHLIN, A., *Introduction à la linguistique contemporaine*, Paris, Armand Colin/Masson, 1997 (REF).
- MOHANAN, K., P., « The organization of the grammar », in Goldsmith, J. (éd.), *The Handbook of Phonological Theory*, Oxford, Blackwell, 1995, pp. 24-69.

- MOIRAND, S., *Enseigner à communiquer en langue étrangère*, Paris, Hachette, 1982.
- MONAHAN, P. J., « Evidence of transference and emergence in the interlanguage », ms, mai 2001, Rutgers Optimality Archives 444-0701, consulté le 21. 03. 03, <http://roa.rutgers.edu/view.php3?roa=444>.
- MONDADA, L. et DUBOIS, D., « Construction des objets de discours et catégorisation : une approche des processus de référenciation », *T.R.A.N.E.L*, n°23, 1995, pp. 273-302.
- MORAIS, J., CONTENT, A., CARY, L., MEHLER, J. et SEGUI, J., « Syllabic segmentation and literacy », *Language and Cognitive Processes*, 4, 1989, pp. 57-67.
- MORELLI, F., « Markedness relations and Implicational Universals in the Typology of Onset Obstruent Clusters », *Proceedings of North Eastern Linguistics Society* 28, Vol. 2, 1998, pp. 107-120, consulté le 21. 03. 03, Rutgers Optimality Archives 251-0398: <http://roa.rutgers.edu/view.php3?roa=251>.
- MORELLI, F., *The Phonotactics and Phonology of Obstruent Clusters in Optimality Theory*, PhD dissertation, 1999, ROA 420-0900, consulté le 10.02.04 (REF).
- MULLER, A. et PRELAZ, A.-C., « Speech Viewer I: comment on s'en sert et comment s'en servir », *T.R.A.N.E.L*, n°21, 1994, pp. 165-177.
- MUNHALL, K. G. et TOHKURA, Y., « Audiovisual gating and the time course of speech perception », *Journal of the Acoustical Society of America*, 104 (1), 1998, pp. 530-539.
- MURAKAWA, H. et LAMBACHER, S. G., « Improving Japanese Pronunciation of American English [r] Using Electronic Visual Feedback », in Izzo, J. (éd.), *Center for Language Research 1995 Annual Review*, Aizu-Wakamatsu (Japon), Université de Aizu, 1996, pp. 79-97.

## N

- NAKAMURA, K. et KOUIDER, S., « Functional neuroanatomy of Japanese writing system », *Aphasiology*, 17 (6-7), 2003, pp. 667-683.
- NARCY, J.-P., « La prise de conscience des problèmes linguistiques conduit-elle à une réduction des erreurs commises en production libre ? Compte-rendu d'expérience », in *Acquisition et enseignement/apprentissage des langues*, Actes du VIIIème Colloque International « Acquisition d'une langue étrangère : perspectives et recherches », LIDILEM, Université Stendhal Grenoble III, mai 1991, pp. 345-353.
- NERI, A., CUCCHIARINI, C. et STRIK, H., « Feedback in Computer Assisted Pronunciation Training : when technology meets pedagogy », *Proceedings of the CALL Conference "CALL professionals and the future of CALL research"*, Anvers (Belgique), 2002, pp. 179-188.
- NESPOULOUS, J.-L., « Le Langage. Introduction : linguistique, neurolinguistique et neuropsycholinguistique. Un parcours en quatre étapes », in Séron, X. et Jeannerod, M. (éd.) *Neuropsychologie humaine*, Bruxelles, Mardaga, 1994, pp. 317-319.

- NESPOULOUS, J.-L., « Multimodalité et cognition : est-ce que le pluriel l'emporte toujours sur le singulier ? », in Koulayan, N. et Borrell, A.(éd.), *C.I.S.L.* n° 15, 2000, Toulouse, Université de Toulouse-Le Mirail, pp. 12-15.
- NESPOULOUS, J.-L., Notes de séminaire de neuropsycholinguistique de maîtrise et de DEA, Université de Toulouse-Le Mirail, 2001, 2002.
- NESPOULOUS, J.-L., « Linguistique, Pathologie du Langage et cognition : des dysfonctionnements langagiers à la caractérisation de l'architecture fonctionnelle du langage », in Fuchs, C. (éd.), *Les linguistiques cognitives*, Paris, Ophrys et Editions de la MSH, 2004.
- NESPOULOUS, J.-L. et DORDAIN, M., « Dissociations procédurales chez l'aphasique et variabilité inter-tâches chez l'apprenant de seconde langue », in *Revue de Phonétique Appliquée*, n° 112-113, 1994, pp. 263-278.
- NESPOULOUS, J.-L. et LECLERCQ, M. (éd.), *Linguistique et Neuropsycholinguistique : tendances actuelles*, Paris, Edition de la Société de Neuropsychologie de Langue Française, 1990.
- NESPOULOUS, J.-L. et MOREAU, N., « "Repair strategies" and the production of segmental errors in aphasia : epenthesis vs. syncope in consonantal clusters », in Visch-Brink, E. et Bastiaanse, R. (éd.), *Linguistic Levels in Aphasia*, San Diego, Singular Publishing Group, 1998, pp. 133-144.
- NESTERENKO, I., « French listeners counting syllables in read French and Russian : implications for the cognitive reality of syllable », *Communication, Speech Prosody Conference 2002*, Aix-en-Provence, 2002.
- NGUYEN, N., « La perception de la parole », in Nguyen, N., Wauquier-Gravelines, S. et Durand, J. (éd.), *Phonologie et Phonétique : forme et substance*, Paris, Hermès, 2005.
- NORRIS, D., « A quantitative multiple-levels model of reading aloud », *Journal of Experimental Psychology : human perception and performance*, 20, 1994, pp. 1212-1232.
- NORRIS, J. and ORTEGA, L., « Effectiveness of L2 instruction : a research synthesis and quantitative meta-analysis », *Language Learning*, 50, 2000, pp. 417-528.
- NOSKE, R., « Is French Optimal ? A question concerning phonological process order », in Durand, J. et Laks, B. (éd.), *Current Trends in Phonology*, CNRS, Paris-X et University of Salford : University of Salford publications, 1996, Rutgers Optimality Archives 109-0000, consulté le 21.03.03, <http://roa.rutgers.edu/view.php3 ? roa=109>.

## O

- ODLIN, T., *Language Transfer : cross-linguistic influence in language learning*, Cambridge, CUP, 1989.
- OLERON, P., « Compréhension verbale et mobilisation des savoirs », in *Actes du Colloque International et Multidisciplinaire sur la Compréhension du Langage : le point de vue des linguistes, le point de vue des traducteurs-interprètes, le point de vue des neuropsychologues*, Paris, Didier Erudition, 1981.
- ORTEGA-LLEBARIA, M., FAULKNER, A. et HAZAN, V., « Auditory-visual L2 speech perception : effects of visual cues and acoustic phonetic context for Spanish learners of English », *Speech, Hearing and Language : UCL Work in Progress*, Vol. 13, 2001, pp. 39-51.
- OSHERSON D. N., KOSSLYN, S. M. et HOLLERBACH, J. M. (éd.), *Visual Cognition and Action ; an invitation to cognitive science, vol. 2*, Cambridge, MIT, 1990 (REF).
- OTAKE, T., HATANO, G., CUTLER, A. et MEHLER, J., « Mora or Syllable ? Speech segmentation in Japanese », *Journal of Memory and Language*, 32, 1993, pp. 258-278.

## P

- PAAP, K. R., NOEL, R. W. et JOHANSEN, L. S., « Dual-route models of print to sound : red herrings and real horses », in Frost, R. et Katz, L. (éds.), *Orthography, Phonology, Morphology and Meaning*, Elsevier Science Publishers, 1992, pp. 293-318.
- PALLIER, C., « Word recognition : do we need phonological representations ? », in Cutler, A., McQueen, J. M. et Zondervan, R. (éd.), *Proceedings of the workshop on spoken word access processes (SWAP)*, Nijmegen, Max-Planck Institute for Psycholinguistics, 2000, pp. 159-162.
- PALLIER, C., *Rôle de la syllabe dans la perception de la parole*, Thèse de Doctorat Nouveau Régime, Paris, EHESS, 1994 (REF).
- PAPAGNO, C., VALENTINE, T. et BADDELEY, A., « Phonological short-term memory and foreign-language vocabulary learning », *Journal of Memory and Language*, 30, 1991, pp. 331-47.
- PARADIS, C., « The inadequacy of filters and faithfulness in loanword adaptation », in Durand, J. et Laks, B. (éds.), *Current Trends in Phonology*, University of Salford : University of Salford publications, 1996, Rutgers Optimality Archives 157-1196, consulté le 21.03.03, <http://roa.rutgers.edu/view.php3 ? roa=157>
- PARADIS, C. et BELAND, R., « Syllabic constraints and constraint conflicts in loanword adaptations, aphasic speech, and children's errors », in Durand, J. et Laks, B. (éd.), *Phonetics, Phonology and Cognition*, Oxford, OUP, 2002, pp. 191-225.



- PARADIS, C. et EL FENNE, F., « French verbal inflexion revisited : constraints, repairs and floating consonants », in Durand, J. et Hintze M.-A. (éd.), *Lingua*, 95, numéro spécial : « French phonology : morae, syllables, words », 1995, pp. 169-204.
- PENDANX, M., Enseignement communicatif et réflexion sur la langue : quelle complémentarité en contexte scolaire ? », in Atienza, J.-L., Bérard, E. et De Carlo, M. (Coord.), « Où en est le communicatif ? », *Etudes de Linguistique Appliquée*, n°100, Paris, Didier Erudition, 1995, pp. 117-126.
- PENDANX, M., *Les activités d'apprentissage en classe de langue*, Paris, Hachette, 1998.
- PEPPERKAMP, S., « L'acquisition de la phonologie et ses conséquences pour la théorie linguistique », *Dossier de candidature à l'Habilitation à Diriger des Recherches*, Vol. II : Synthèse. EHESS, Octobre 2002 (REF).
- PEPPERKAMP, S. et DUPOUX, E., « Reinterpreting loanword adaptations : the role of perception », *Proceedings of the 15<sup>th</sup> International Congress of Phonetic Sciences*, Université Autonome de Barcelone, 2003, pp. 367-370.
- PEPPERKAMP, S. et DUPOUX, E., « The acquisition of abstract phoneme categories », *9<sup>th</sup> Conference on Laboratory Phonology (LabPhon9)*, University of Illinois, 24-26 juin 2004.
- PEPPERKAMP, S., PETTINATO, M. et DUPOUX, E., « Allophonic variation and the acquisition of phoneme categories », in Beachley, B., Brown, A. et Conlin, F. (éd.), *Proceedings of the 27<sup>th</sup> Annual Boston University Conference on Language Development*, Vol. 2, Sommerville, Cascadilla Press, 2003, pp. 650-661.
- PEREA, M. et ROSA, E., « Does "whole word shape" play a role in visual word recognition ? », *Perception and Psychophysics*, Vol. 64, n°5, 2002, pp. 785-794.
- PERETZ, I., « Les agnosies auditives », in Séron, X. et Jeannerod, M. (dir.), *Neuropsychologie humaine*, Bruxelles, Mardaga, 1994, pp. 205-209.
- PERFETTI, C. A., ZHANG, S. et BERENT, I., « Reading in English and Chinese : evidence for a « universal » phonological principle », in Frost, R. et Katz, L. (éd.), *Orthography, Phonology, Morphology and Meaning*, Elsevier Science Publishers, 1992, pp. 227-248.
- PERRY, C. et ZIEGLER, J. C., « Beyond the two-strategy model of skilled spelling : effects of consistency, grain size, and orthographic redundancy », *The Quarterly Journal of Experimental Psychology*, 57A (2), 2004, pp. 325-356.
- PERRY, C., ZIEGLER, J. et COLTHEART, M., « A dissociation between orthographic awareness and spelling production », *Applied Psycholinguistics*, 23, 2002, pp. 43-73.
- PESCHANSKI, M., *Le cerveau en quatre dimensions*, Paris, Hachette, 1993.
- PORQUIER, R., « Grammaire(s) et grammaticalisation en langue non-maternelle », in *Le Français dans le Monde, Recherches et Applications*, « Une didactique des langues pour demain », 2000, pp. 70-77.
- PORQUIER, R. et VIVES, R., « Le statut des outils métalinguistiques dans l'apprentissage et l'enseignement au niveau avancé », in Béacco, J.-C., Perdue, C. et Vivès, R. (coord.), « Appropriation, descriptions et enseignements de langues », *Etudes de Linguistique Appliquée* n°92, 1993, Paris, Didier Erudition, pp. 65-77.

- PORTINE, H., « Le multimédia dans la méthodologie de la didactique des langues », in Chevalier, Y. (éd.), *Outils multimédia et stratégies d'apprentissage du Français Langue Etrangère*, Cahiers de la Maison de la Recherche, Lille, Université Charles-de Gaulle-Lille 3, 1996.
- POTHIER, M., *Multimédias, dispositifs d'apprentissage et acquisition des langues*, Paris, Ophrys, 2003.
- *Pour la Science*, Dossier hors-série, « Les langues du monde », octobre 1997.
- PRABLANC, C., « Les transformations visuomotrices », in Gordon, M. B. et Paugam-Moisy, H. (dir.), *Sciences cognitives : diversité des approches*, Paris, Hermès, 1997, pp. 89-96.
- PUREN, C., *Histoire des méthodologies*, Paris, Nathan, 1988 (REF).
- PY, B., « Didactique et modèles linguistiques », in Béacco, J.-C. et Chevalier, J.-C. (coord.), *Etudes de Linguistique Appliquée* n°72, « Didactique des langues : quelles interfaces ? », 1988, Paris, Didier Erudition, pp. 7-12.

## **R**

- RADWANSKA-WILLIAMS, J. et YAM, J. P. S., « The acquisition of English plosives by Chinese learners », *Proceedings of the Phonetics Teaching and Learning Conference 2001*, Londres, UCL, 2001, pp. 35-38.
- RAMSEY, W. STICH, S. P. et RUMELHART, D. E. (éd.), *Philosophy and Connectionist theory*, Hillsdale, Lawrence Erlbaum, 1991 (REF).
- RAMUS, F., DUPOUX, E. et MEHLER, J., « The psychological reality of rhythm classes : perceptual studies », *Proceedings of the 15<sup>th</sup> International Congress of Phonetic Sciences*, Barcelone, Université Autonome de Barcelone, 2003.
- RASTLE, K. et COLTHEART, M., « Serial and strategic effects in reading aloud », *Journal of Experimental Psychology : Human perception and performance*, 25, 1999, pp. 482-503.
- RASTLE, K., DAVIS, M. H. et NEW, B., « The broth in my brother's brothel : morpho-orthographic segmentation in visual word recognition », *Psychonomic Bulletin & Review*, à paraître.
- RAYNER, K., « Eye movements in reading », in Kolers, P. A., Wrolstad, M. E. et Bouma, H. (éd.), *Processing of Visible Language*, Vol. 1., New York, Plenum Press, 1979.
- REPP, B. H. et LIBERMAN, A. M., « Phonetic category boundaries are flexible », in Harnad, S. (éd.), *Categorical Perception : the groundwork of cognition*, Cambridge, CUP, 1987, pp. 89-111.
- RENARD, R. *La méthode verbo-tonale de correction phonétique*, Bruxelles, Mons, Didier, 1979.
- REY, A., JACOBS, A. M., SCHMIDT-WEIGAND, F., ZIEGLER, J. C., « A phoneme effect in visual word recognition », *Cognition*, Vol. 68, n°3, 1998, pp. B71-B80.

- REY, A., ZIEGLER, J. C. et JACOBS, A. M., « Graphemes are perceptual reading units », *Cognition*, Vol. 75, n° 1, 2000, pp. B1-B12.
- RIBA, P., LEPAGE, S. et CHEVALLIER-WIXLER, D., « La réforme du DELF et du DALF », *Le Français dans le Monde*, n° 336, 2004, pp. 29-31.
- RICHARDOT, D., « Syllabation comparée du français et du japonais », *Enseignement du français au Japon*, n°17, 1988, pp. 24-26.
- RIVENC, P., *Pour aider à apprendre à communiquer dans une langue étrangère*, Paris : Didier Eruditions – Mons : CIPA, 2000a.
- RIVENC, P. (avec G. Gougenheim), « Etat actuel du Français Fondamental », *Le Français dans le Monde*, n°1, 1961 / *Pour aider à apprendre à communiquer dans une langue étrangère*, Paris : Didier Eruditions – Mons : CIPA, 2000b, pp. 31-36.
- RIVENC, P., « Vers une approche sémiotique du « discours audiovisuel » dans les méthodes d'apprentissage linguistique », *Journal de Psychologie*, 1-2, 1973 / *Pour aider à apprendre à communiquer dans une langue étrangère*, Paris : Didier Eruditions – Mons : CIPA, 2000c, pp. 119-135.
- RIVENC, P., « Le Français Fondamental vingt-cinq ans après », *Le Français dans le Monde*, n°148, 1979 / *Pour aider à apprendre à communiquer dans une langue étrangère*, Paris : Didier Eruditions – Mons : CIPA, 2000d, pp. 71-89.
- RIVENC, P., (éd.), *Apprentissage d'une langue étrangère/seconde. Vol. 3 : La méthodologie*, De Boeck, 2003 (REF).
- RIVENC, P. et BOUDOT, J., Préface de « Voix et Images de France, 1<sup>er</sup> degré », CREDIF, Paris, Didier, 1962.
- RIVENC-CHICLET, M.-M., « Méthodologie, enseignement, apprentissage d'une langue étrangère », *Revue de Phonétique Appliquée*, 90, Mons, pp. 49-79, 1989.
- RIVENC-CHICLET, M.-M., « Didactique : le poids des institutions », *Le Français dans le Monde*, n° 261, 1993, pp. 40-44.
- RIVENC-CHICLET, M.-M., « SGAV, psycholinguistique et multimédia », Actes du Colloque *Les langues à travers le S.G.A.V.*, Collection Sources - Cibles, Université Saint-Joseph, Faculté des Sciences Humaines, Beyrouth, 2001, pp. 171-187.
- RIVENC-CHICLET, M.-M., « Méthodologie d'enseignement d'une langue vivante : parcours global au niveau 1 », in Rivenc, P. (éd.), *Apprentissage d'une langue étrangère/seconde. Vol. 3 : La méthodologie*, De Boeck, 2003, chap. 8.
- ROBERT, S., « Variations des représentations linguistiques : des unités à l'énoncé », in Fuchs, C. et Robert, S., *Diversité des langues et représentations cognitives*, Paris, Editions Ophrys, 1997, pp. 25-39.
- ROSEN, S. et HOWELL, P., « Auditory, articulatory, and learning explanations of categorical perception in speech », in Harnad, S. (éd.), *Categorical Perception : the groundwork of cognition*, Cambridge, CUP, 1987, pp. 113-157.
- ROUSSET, I. et VALLEE, N., « Vers une organisation syllabique des lexiques : tendances, dépendances et cooccurrences segmentales », *XXIVèmes Journées d'Etudes sur la Parole*, Nancy, 24-27 juin 2002.

## S

- SAGAZ, M., KOULAYAN, N. et DETEY, S., « Perception of help pop-up windows in language learning software by Japanese learners of French », *ICU Language Research Bulletin*, Vol. 19, 2004, pp. 129-138.
- SANTACROCE, M., *Grammaire, linguistique et didactique du français langue étrangère - Propositions pour une grammaire transitionnelle*, Thèse de Doctorat, Université de Paris III - Sorbonne nouvelle, 1999 (REF).
- SAUSSURE, F. de, *Cours de linguistique générale*, seconde édition 1955, Paris, Payot, 1916 (REF).
- SCHWARTZ, J.-L., ABRY, C., BOË, L.-J. et CATHIARD, M., « Phonology in a theory of perception-for-action-control », in Durand, J. et Laks, B. (éds.), *Phonetics, Phonology and Cognition*, Oxford, OUP, 2002, pp. 254-280.
- SCHWARTZ, J.-L., ABRY, C., BOË, L.-J., VALLEE, N. et CATHIARD, M.-A., « Economy and phonology in the perceptual Dispersion-Focalization Theory and the Theory of Perception for Action-Control », *First International Symposium on Linguistics (LICSSOLI)*, Institut des Sciences Cognitives, Lyon, octobre 1999.
- SEGUI, J., « La perception du langage parlé : données et théories », in Lambert, J. et Nespoulous, J.-L. (dir.), *Perception auditive et compréhension du langage : état initial, état stable et pathologie*, Marseille, Solal Editeurs, 1997, pp. 15-22.
- SEGUI, J., « La perception du langage parlé », in Bonnet, C., Ghiglione, R. et Richard, J.-F. (dir.), *Traité de psychologie cognitive : perception, action, langage*, Paris, Bordas, 1989/ Dunod, 2003, pp. 200-234.
- SEGUI, J. et FERRAND, L., « The role of syllable in speech perception and production », in Durand, J. et Laks, B. (éd.), *Phonetics, Phonology and Cognition*, Oxford, OUP, 2002, pp. 151-167.
- SELINKER, L., « Language transfer », in Gass, S. et Selinker, L. (éd.), *Language Transfer in Language Learning*, Rowley, Newbury House, 1983, pp. 33-53.
- SENNEMA, A., HAZAN, V. et FAULKNER, A., « The role of visual cues in L2 consonant perception », *Proceedings of the 15<sup>th</sup> International Congress of Phonetic Sciences*, Barcelone, 3-9 août 2003, pp. 135-138.
- SERON, X. et JEANNEROD, M. (dir.), *Neuropsychologie humaine*, Bruxelles, Mardaga, 1994 (REF).
- SERVICE, E., « Phonology, working memory, and foreign language learning », *Quarterly Journal of Experimental Psychology*, 45A, (1), 1992, pp. 21-50.
- SHAMS, L., KAMITANI, Y. et SHIMOJO, S., « What you see is what you hear », *Nature*, Vol. 408, 2000, p. 788.
- SHAMS, L., KAMITANI, Y., THOMPSON, S. et SHIMOJO, S., « Sound alters visual evoked potentials in humans », *Cognitive Neuroscience and Neurosychology, NeuroReport*, Vol. 12, n°17, 2001, pp. 3849-3852.
- SHEBILSKIE, W. L. et REID, L. S., « Reading eye movements », in Kolers, P. A., Wrolstad, M. E. et Bouma, H. (éd.), *Processing of Visible Language*, Vol. 1, New York, Plenum Press, 1979, pp. 97-110.

- SHIMOJO, S. et SHAMS, L., « Sensory modalities are not separate modalities : plasticity and interactions », *Current Opinion in Neurobiology*, 11, 2001, pp. 505-509.
- SHIMOJO, S., SCHEIER, C., NIJHAWAN, R., SHAMS, L., KAMITANI, Y. et WATANABE, K., « Beyond perceptual modularity : auditory effects on visual perception », *Acoustic Science and Technology*, Vol. 22, n°2, 2001, pp. 61-66.
- SHINOHARA, S., *Analyse phonologique de l'adaptation japonaise de mots étrangers*, Thèse de Doctorat, Université de la Sorbonne-Nouvelles Paris-III, 1997 (REF).
- SHINOHARA, S., « Emergence of Universal Grammar in foreign word adaptations », in R. Kager, J. Pater et W. Zonneveld (éd.), *Fixing Priorities : Constraints in Phonological Acquisition*, Cambridge, CUP, 2004, Rutgers Optimality Archives 622, consulté le 05-03-03 : <http://roa.rutgers.edu/ROA-622>.
- SIEROFF, E., « Les mécanismes attentionnels », in Séron, X. et Jeannerod, M. (dir.), *Neuropsychologie humaine*, Bruxelles, Mardaga, 1994, pp. 127-148.
- SIEROFF, E., « Les mécanismes attentionnels », in Séron, X. et Jeannerod, M. (dir.), *Neuropsychologie humaine*, Bruxelles, Mardaga, 1994, pp. 127-148.
- SILVEIRA, R., « Pronunciation instruction. Classroom practice and empirical research », *Linguagem & Ensino*, Vol. 5, N° 1, 2002, pp. 93-126.
- SMITH, J. L., « Loan phonology is not all perception : evidence from Japanese loan doublets », in Vance, T. J. (éd.), *Japanese/Korean Linguistics*, Vol. 14, Stanford, CSLI, 2005.
- SOKAL, A. et BRICMONT, J., *Impostures intellectuelles*, Paris, Odile Jacob, 1997.
- STEELE, J., « Assessing the role of orthographic versus uniquely auditory input in acquiring new L2 segments », *7èmes Rencontres Internationales du Réseau Français de Phonologie*, Université d'Aix-en-Provence, 2-4 juin 2005.
- STENTON, A., PECHOU, A., VAILLANT-SIRDEY, C. et TRICOT, A., « Effet du double codage synchrone de l'accentuation en L2 selon des modalités de restitution du sujet », in Billières, M., Gaillard, P. et Spanghero-Gaillard, N. (éd.), *Actes du Premier Colloque International de Didactique Cognitive, (Français Langue Etrangère / Langue Seconde / Langue Maternelle)* [CD-ROM], Université de Toulouse-le Mirail, janvier 2005, pp. 304-310.

## T

- TARONE, E., « The phonology of interlanguage », in Ioup, G. et Weinberger, S. H. (éd.), *Interlanguage Phonology : the acquisition of a second language sound system*, Cambridge, Newbury House, 1987, pp. 70-85.
- TARRIER, J.-M. et AURAN, C., « Fichiers mots : constitution, alignement et transcription », in Eychenne, J. et Mallet, G. (éd.), *Bulletin PFC n°3, Du segmental au prosodique : protocoles, outils, extensions et travaux en cours*, ERSS UMR 5610 & Université de Toulouse - Le Mirail, 2004, pp. 25-28.

- TAYLOR, I. et TAYLOR, M. M., *Writing and literacy in Chinese, Korean and Japanese*, Amsterdam John Benjamins, 1995 (REF).
- TELLIER, M., « L'utilisation des gestes en classe de langue : comment évaluer leur effet sur la mémorisation du lexique ? », in Billières, M., Gaillard, P. et Spanghero-Gaillard, N. (éd.), *Actes du Premier Colloque International de Didactique Cognitive, (Français Langue Etrangère / Langue Seconde / Langue Maternelle)* [CD-ROM], Université de Toulouse-le Mirail, janvier 2005.
- THIERRY, G., « La mémoire : approche par l'imagerie fonctionnelle », INSERM U455, Communication Personnelle, Séminaire de Neuropsychologie 2001, S. Ruff, Université Toulouse-Le Mirail.
- THORPE, S., « Modèles de traitement de l'information », in Séron, X. et Jeannerod, M. (dir.), *Neuropsychologie humaine*, Bruxelles, Mardaga, 1994, pp. 153-175.
- TIBERGHIE, G., « Psychologie cognitive de la mémoire humaine », in Séron, X. et Jeannerod, M. (dir.), *Neuropsychologie humaine*, Bruxelles, Mardaga, 1994, pp. 255-278.
- TIBERGHIE, G., « La mémoire humaine : connaître ou se souvenir ? », in Gordon, M. B. et Paugam-Moisy, H. (dir.), *Sciences cognitives : diversité des approches*, Paris, Hermès, 1997, pp. 139-153.
- TOMA, A., « La création de cours automatique à didactisation minimale », in Koulayan, N. et Borrell, A. (éd.), *Actes du Colloque International « Le multimédia et l'apprentissage du français langue étrangère et/ou seconde : de l'expérimentation à la réalité dans l'espace francophone »* (mars 1999), *CCISL n°15*, Université de Toulouse-Le Mirail, 2000, pp. 111-125.
- TOWELL, R. et HAWKINS, R., *Approaches to Second Language Acquisition*, Clevedon, Multilingual Matters Ltd, 1994 (REF).
- TRANEL, B., « Current issues in French phonology : liaison and position theories », in Goldsmith, J. (éd.), *The Handbook of Phonological Theory*, Oxford, Blackwell, 1995a, pp. 798-815.
- TRANEL, B., « French final consonants and non-linear phonology », in Durand, J. et Hintze M.-A. (éd.), *Lingua*, 95, numéro spécial : « French phonology : morae, syllables, words », 1995b, pp. 27-50.
- TREVISE, A., « Acquisition/apprentissage/enseignement d'une langue 2 : modes d'observation, modes d'intervention », in Béacco, J.-C., Perdue, C. et Vivès, R. (coord.), « Appropriation, descriptions et enseignements de langues », *Etudes de Linguistique Appliquée n°92*, Paris, Didier Erudition, 1993, pp. 38-50.
- TREVISE, A., « Schémas de base en langue 2 et invariants sémantico-cognitifs », *Acquisition et Interaction en Langue Etrangère (AILE)*, numéro spécial, Actes du 8<sup>ème</sup> Colloque EUROSLA Paris, Volume 2 « Du mot à la structure », Université Paris VIII, 1999, pp. 89-101.
- TREVISE, A. et DEMAIZIERE, F., « Imparfait, passé simple, prétérit : représentations métalinguistiques et didactique », in *Acquisition et enseignement/apprentissage des langues*, Actes du VIII<sup>ème</sup> Colloque International « Acquisition d'une langue étrangère : perspectives et recherches », LIDILEM, Université Stendhal Grenoble III, mai 1991, pp. 355-363.

- TRICOT, A., « Charge cognitive et apprentissage : une présentation des travaux de John Sweller », *Revue de Psychologie de l'Education*, 1, 1998, pp. 37-64.
- TRICOT, A., DETIENNE, F. et BASTIEN, J. M. C., « Recherches en psychologie ergonomique : introduction », *Psychologie Française*, 48 (3), 2003, pp. 1-8.
- TRICOT, A. et ROUET, J.-F. (dir.), « Les hypermédias : approches cognitives et ergonomiques », *Hypertextes et Hypermédias, numéro hors série*, Hermès, 1998 (REF).
- TROUBETZKOY, N. S., *Principes de phonologie*, Paris, Klincksieck, 1976 (REF).

## U

- UFFMANN, C., « Markedness vs. Transfer in SLA : a unified account », *XVIIth International Conference on Foreign/Second Language Acquisition*. Szczyrk, 22 mai 2003.

## V

- VALDMAN, A., « The problem of the target model in proficiency-oriented foreign language instruction », in Valdman, A. (éd.), *Proceedings of the symposium on the evaluation of foreign language proficiency*, Bloomington, Indiana University, 1987, pp. 133-149.
- VALDMAN, A., « The elaboration of pedagogical norms for second language learners in a conflictual diglossia situation », in Gass, S., Madden, C., Preston, D. et Selinker, L. (éd.), *Variation in Second Language Acquisition, Volume I : Discourse and Pragmatics*, Clevedon, Multilingual Matters, 1989, pp. 15-34.
- VALDOIS, S., « Tendances actuelles en phonologie et pathologie du langage », in Nespoulous, J.-L. et Leclercq, M. (éd.), *Linguistique et Neuropsycholinguistique : tendances actuelles*, Paris, Edition de la Société de Neuropsychologie de Langue Française, 1990.
- VALDOIS, S. et NESPOULOUS, J.-L., « Perturbations du traitement phonétique et phonologique du langage », in Séron, X. et Jeannerod, M. (dir.), *Neuropsychologie humaine*, Bruxelles, Mardaga, 1994, pp. 360-374.
- VAN DE SANDT-KOENDERMAN, W. M. E. et BONTA, E., « Agrammatism : Theory and Therapy », in Visch-Brink, E. et Bastiaanse, R. (éd.), *Linguistic Levels in Aphasia*, San Diego, Singular Publishing Group, 1998, pp. 211-231.
- VAN DER LINDEN, M., « Neuropsychologie de la mémoire », in Séron, X. et Jeannerod, M. (dir.), *Neuropsychologie humaine*, Bruxelles, Mardaga, 1994, pp. 282-309.
- VAN GELDER, « What is the "D" in "PDP" ? A survey of the concept of distribution », in Ramsey, W., Stich, S.P. et Rumelhart D. E. (éd.), *Philosophy and Connectionist Theory*, Hillsdale, Lawrence Erlbaum, 1991, pp. 33-59.

- VAN PATTEN, B., *Input Processing and Grammar Instruction : Theory and Research*, Norwood, Alex Publishing Corporation, 1996.
- VAN PATTEN, B., « Communicative classrooms, processing instruction, and pedagogical norms », in Gass, S., Bardovi-Harlig, K., Sieloff Magnan, S. et Walz, J. (éd.), *Pedagogical Norms for Second and Foreign Language Learning and Teaching*, Amsterdam, John Benjamins, 2002, pp. 105-118.
- VANCE, T. J., *An Introduction to Japanese Phonology*, New York, State University of New York Press, 1987.
- VENDELIN, I. et PEPPERKAMP, S., « Evidence for phonetic adaptation of loanwords : an experimental study », *Actes des Journées d'Etudes Linguistiques 2004 (JEL2004)*, Université de Nantes, 2004, pp. 129-131.
- VENDELIN, I. et PEPPERKAMP, S., « The influence of orthography on loanword adaptations », *7èmes Rencontres Internationales du Réseau Français de Phonologie*, Université d'Aix-en-Provence, 2-4 juin 2005.
- VENTURA, P., KOLINSKY, R., BRITO-MENDES, C. et MORAIS, J., « Mental representations of the syllable internal structure are influenced by orthography », *Language and Cognitive Processes*, 16 (4), 2001, pp. 393-418.
- VIGNAUX, G., « Catégorisation et schématisation : des arguments au discours », in Dubois, D. (dir.), *Sémantique et cognition : catégories, prototypes, typicalité*, Paris, Editions du CNRS, 1991, pp. 295-318.
- VILLAFANA, C., « Emergence of the unmarked in interlanguage coda production », in Villafaña C. (éd.), *George Mason University Working Papers in Linguistics*, 2000, pp. 33-48.
- VIRBEL, J., « Aspects du contrôle des structures textuelles », in Lambert, J. et Nespoulous, J.-L. (dir.), *Perception auditive et compréhension du langage : état initial, état stable et pathologie*, Marseille, Solal Editeurs, 1997, pp. 251-270.
- VISCH-BRINK, E. G. et BASTIAANSE, R. (éd.), *Linguistic Levels in Aphasia*, San Diego, Singular Publishing Group, 1998 (REF).
- VIVES, R., « Quand la didactique fait ses emplettes chez les linguistes », in Béacco, J.-C. et Chevalier, J.-C. (coord.), *Etudes de Linguistique Appliquée* n°72, « Didactique des langues : quelles interfaces ? », Paris, Didier Erudition, 1988, pp. 25-37.
- VOGEL, K., *L'interlangue : la langue de l'apprenant*, traduit de l'allemand par J.-M. Brohée et J.-P. Confais, Toulouse, Presses Universitaires du Mirail, Coll. Interlangues, 1995.

## W

- WALTER, H., « Diversité phonologique et communauté linguistique », in Martinet, J. (Dir.), *De la théorie linguistique à l'enseignement de la langue*, Paris, PUF, 1974, pp. 175-189.
- WALTER, H., « Le français en mouvement », in *Le Français dans le Monde, Recherches et Applications*, numéro spécial : « Oral : variabilité et apprentissages », 2001, pp. 8-17.



- WARSCHAUER, M., « Computer-assisted language learning : an introduction », in Fotos, S. (éd.), *Multimedia language teaching*, Tokyo, Logos International, 1996, pp. 3-20.
- WARSCHAUER, M. et HEALEY, D., « Computers and Language learning: an overview », *Language Teaching*, n° 31, 1998, pp. 57-71.
- WATANABE, K. et SHIMOJO, S., « When sound affects vision : effects of auditory grouping on visual motion perception », *Psychological Science*, Vol. 12, n°2, 2001, pp. 109-116.
- WELLS, J. C., « Orthographic diacritics and multilingual computing », *Language Problems and Language Planning*, Vol. 24, n° 3, 2001, pp. 249-272.
- WELLS, J. C., « Overcoming phonetic interference », *Journal of the English Phonetic Society of Japan*, 3, 2000, pp. 9-21.
- WELLS, J. C., « Why phonetic transcription is important », *Malsori (Phonetics)* (Journal of the Phonetic Society of Korea), n° 31-32, 1996, pp. 239-242.
- WILSON, M., « Brain mechanisms in categorical perception », in Harnad, S. (éd.), *Categorical Perception : the groundwork of cognition*, Cambridge, CUP, 1987, pp. 387-417.
- WINTER, B. et REBER, A. S., « Implicit learning and the acquisition of natural languages », in Ellis, N (éd), *Implicit and Explicit Learning of Language*, San Diego, Academic Press, 1994, pp. 115-145.
- WIOLAND, F., *Prononcer les mots du français*, Paris, Hachette, 1991.

## Y

- YAZDANI, M. et BARKER, P. (éd.), *Iconic communication*, Intellect Books, 2000 (REF).
- YUILLE, A. L. and ULLMAN, S., « Computational theories of low-level vision », in Osherson D. N., Kosslyn, S. M. et Hollerbach, J. M. (éd.), *Visual Cognition and Action ; an invitation to cognitive science, vol. 2*, Cambridge, MIT, 1990, pp. 5-39.

## Z

- ZESIGER, P. et PARTZ, M.-P. (De), « Perturbations du langage écrit : les dyslexies et les dysgraphies », in Séron, X. et Jeannerod, M. (dir.), *Neuropsychologie humaine*, Bruxelles, Mardaga, 1994, pp. 419-437.
- ZHANG, Y., KUHL, P., IMADA, T., IVERSON, P., PRUITT, J., KOTAMI, M. et STEVENS, E., « Neural plasticity revealed in perceptual training of a Japanese adult listener to learn american /l-r/ contrast : a whole-head magnetoencephalography study », *ICSLP-2000*, Vol. 3, 2000, pp. 953-956.
- ZIEGLER, J. C., MUNEAUX, M. et GRAINGER, J., « Neighborhood effects in auditory word recognition : phonological competition and orthographic facilitation », *Journal of Memory and Language*, 48, 2003, pp. 779-793.

- ZIEGLER, J. C., PERRY, C. et COLTHEART, M., « The DRC model of visual word recognition and reading aloud : an extension to German », *European Journal of Cognitive Psychology*, Vol. 12, n° 3, 2000, pp. 413-430.
- ZIEGLER, J. C., PERRY, C., JACOBS, A.M. et BRAUN, M., « Identical words are read differently in different languages », *Psychological Science*, Vol. 12, n° 5, 2001, pp. 379-384.
- ZORZI, M., HOUGHTON, G. et BUTTERWORTH, B., « Two routes or one in reading aloud ? A connectionist Dual-process model », *Journal of Experimental Psychology : human perception and performance*, Vol. 24, n°4, 1998, pp. 1131-1161.

**Sélection de références manquantes**  
**(à titre illustratif)**

**Articles et Contributions :**

- ARCHIBALD, J. « Second language phonology, phonetics and typology », *Studies in Second Language Acquisition*, 20 (2), 1998, pp. 189-211.
- BRADLOW, A. R., AKAHANE-YAMADA, R., PISONI, D. B. et TOHKURA, Y., « Training Japanese listeners to identify English /r/ and /l/ : long-term retention of learning in perception and production », *Perception & Psychophysics*, 61 (5), 1999, pp. 977-985.
- BRADLOW, A. R., PISONI, D. B., YAMADA, R. A. et TOHKURA, Y., « Training Japanese listeners to identify English /r/ and /l/ : IV. Some effects of perceptual learning on speech production », *Journal of the Acoustical Society of America*, 101, 1997, pp. 2299-2310.
- BROSELOW, E., SU-I C. et WANG, C., « The emergence of the unmarked in second language phonology », *Studies in Second Language Acquisition*, 20, 1998, pp. 261-280.
- COLEMAN, J. S., « Cognitive reality and the phonological lexicon : a review », *Neurolinguistics*, 11, 1998, pp. 295-320.
- DIJKSTRA, A. F. J., TIMMERMANS, M. et SCHRIEFERS, H., « On being blinded by your other language : effects of task demands on interlingual homograph recognition », *Journal of Memory and Language*, 42, 2000, pp. 445-464.
- ECKMAN, F. et IVERSON, G., « Sonority and markedness among onset clusters in the interlanguage of ESL learners », *Second Language Research* 9, 1993, pp. 234-252
- EDGE, B. A., « The production of word-final voiced obstruents in English by L1 speakers of Japanese and Cantonese », *Studies in Second Language Acquisition*, 13 (1), 1991, pp. 23-41.
- FROST, R., REPP, B. H. et KATZ, L., « Can speech perception be influenced by simultaneous presentation of print ? », *Journal of Memory and Language*, 27, 1988, pp. 741-755.
- HANCIN-BHATT, B. et BHATT, R. M., « Optimal L2 syllables : interaction of transfer and developmental effects », *Studies in Second Language Acquisition*, 19, 1997, pp. 331-378.
- HATTA, T., « Recognition of Japanese kanji and hiragana in the left and right visual fields », *Japanese Psychological Research*, 20, 1978, pp. 51-59.
- JONES, R.H., « Beyond "listen and repeat" : pronunciation teaching materials and theories of second language acquisition », *System*, Vol. 25, n° 1, 1997, pp. 102-112.
- KODA, K., « Effects of L1 orthographic representation on L2 phonological coding strategies », *Journal of Psycholinguistic Research*, 18, 1989, pp.201-222.
- KODA, K., « L2 word recognition research : a critical review », *Modern Language Journal*, 80, iv, 1996, pp. 450-460.

- KUHL, P. K. et IVERSON, P. « Linguistic experience and the « perceptual magnet effect », in Strange, W. (éd.), *Speech Perception and Linguistic Experience : issues in cross-language research*. Timonium MD, York Press, 1995.
- LAKS, B., « Description de l'oral et variation : la phonologie et la norme », *L'information grammaticale*, n°94, 2002, pp. 5-11.
- MACDONALD, D., YULE, G. et POWERS, M., « Attempts to improve English L2 pronunciation : the variables effects of different types of instruction. », *Language Learning*, Vol. 44, n°1, 1994, pp. 75-100.
- MASSARO, D., « Multiple book review of speech perception by ear and eye : a paradigm for psychological enquiry », *Behavioral and Brain Sciences*, 12, 1989, pp. 741-94.
- ROBERT-RIBES, J., LALLOUACHE, T. et ESCUDIER, P., « Complementarity and synergy in bimodal speech : auditory, visual and audiovisual identification of French oral vowels in noise », *Journal of the Acoustical Society of America*, 103, 1998, pp.3677-89.
- SASANUMA, S., « Kana and kanji processing in Japanese aphasics », *Brain and Language*, 2, 1974, pp. 369-83.
- TAFT, M., « An alternative to grapheme-phoneme conversion rules ? », *Memory and Cognition*, 10, 1982, pp. 465-74, 1982.
- ZIEGLER, J. C. et FERRAND, L., « Orthography shapes the perception of speech : the consistency effect in auditory word recognition », *Psychonomic Bulletin & Review*, 5, 1998, pp. 683-689.

#### **Ouvrages :**

- ARCHIBALD, J., *Second Language Phonology*, Amsterdam, John Benjamins, 1998.
- BERNHARDT, B.H. et STEMBERGER, J. P. (éd.), *Handbook of Phonological Development from the Perspective of Constraint-Based Non-Linear Phonology*. San Diego, Academic Press, 1998.
- BUBEN, V., *Influence de l'orthographe sur la prononciation du français*, Bratislava, Travaux de la faculté des lettres de l'université Komensky, 1935.
- CAMPBELL, R., DODD, B. et BURNHAM, D., *Hearing by Eyes, ii : perspectives and directions in research on audio visual aspects of language processing*. Hove, Psychology press, 1998.
- CELCE-MURCIA, M., BRINTON, D. et GOODWIN, J. M., *Teaching Pronunciation*. Cambridge, CUP, 1996.
- DALTON, C. et SEIDLHOFER, B., *Pronunciation*, Oxford, OUP, 1994.
- GATHERCOLE, S.E. et BADDELEY, A., *Working Memory and Language*, Hove, Erlbaum, 1993.
- HULST, H. VAN DER, et RITTER, N. (éd.), *The Syllable : views and facts*, Berlin, Mouton de Gruyter, 1999.
- JAMES, A. et LEATHER, J. (éd.), *Sound Patterns in Second Language Acquisition*, Dordrecht, Foris, 1987.
- KOLINSKY, R., MORAIS, J. et SEGUI, J. (éd.), *La reconnaissance des mots dans différentes modalités sensorielles*, Paris, PUF, 1991.

- MORLEY, J. A. (éd.), *Pronunciation Pedagogy and Theory : new ways, new directions*. Illinois, Pentagraph Print, 1994.
- PARADIS, M., HAGIWARA, H. et HILDEBRANDT, N., *Neurolinguistic Aspects of the Japanese Writing System*, Orlandon Academic Press, 1985.
- ROBINSON, P., *Cognition and Second Language Instruction*, Cambridge, CUP, 2000.
- STRANGE, W. (éd.), *Speech Perception and Linguistic Experience : issues in cross-language research*. Timonium MD, York Press, 1995.
- TULVING, E. et CRAIK, F. I. K. (éd.), *The Oxford Handbook of Memory*, Oxford, OUP, 2000.
- YAVAS, M. (éd.), *First and Second Language Phonology*, San Diego, Singular publishing group Ltd, 1994.

## **ANNEXES**

Dans ces annexes, nous présentons :

1. Le contenu d'ouvrages relatifs à la prononciation du français qui ont inspiré notre travail, *via* une comparaison des contenus et de leur présentation.
2. Une partie du matériel (papier) utilisé dans nos « pré-tests ».
3. Une partie du matériel (papier) utilisé dans nos « tests généraux ».

### **I. Manuels relatifs à la prononciation du français**

Voici ci-dessous la table des matières relativement détaillée de six ouvrages relatifs à la prononciation du français.

A) BATTYE, A., HINTZE, M.-A. et ROWLETT, P, *The French language today : a linguistic introduction (2<sup>nd</sup> edition), London, Routledge, 1992, 2000. Chapitre « The sound system of French » :*

- The point of reference
- The representation of sound
- The representation of sound through phonetic notation
- The physical production of speech sounds
- The classification of French consonants
- Consonants in contact : assimilation
- The pronunciation of double consonants and final consonants
- The classification of French vowels
- The special case of e-muet
- The mid vowels
- The nasal vowels
- Front [a] and back [ɑ]
- Minimal versus maximal vowel system

- The glides of French
- Syllables and sense groups
- Suprasegmentals : stress and intonation

B) LEON, P. et LEON, M., *La Prononciation du français*, Coll.128, Paris, Nathan Université, 1997.

L'oral et l'écrit

L'écrit en français moderne

Les deux codes

Le code phonétique. Les phones et les phonèmes

Phonèmes et variantes

L'alphabet phonétique international

Les autres alphabets phonétiques

Tableaux des symboles phonétiques

Signes diacritiques

La combinaison des phonèmes

Les usages et la norme

Classement articulatoire

Les organes de l'articulation

Articulation et timbres des voyelles

Classement des voyelles selon leurs traits articulatoires

Modes et lieux d'articulation des consonnes

Classement des consonnes d'après leurs traits articulatoires

Le coup de glotte

Articulation et timbre des semi-consonnes

Lois phonétiques

Observations et « lois »

Accentuation

## Syllabation

Division syllabique

Structure syllabique : syllabe ouverte/syllabe fermée

Distribution et timbre des voyelles

Orthoépique. Lois phonétiques et évolution

Exceptions à la loi de distribution complémentaire

Durée et allongement des voyelles

Problèmes d'orthoépique : E caduc et liaison

Les graphies et les sons

E caduc : problèmes de graphie

Quand le E caduc doit-il se prononcer ?

La liaison

La prononciation des consonnes de liaison

Liaison et groupe rythmique

Liaison et style

Aspects fonctionnels

La phonologie : une phonétique fonctionnelle

Commutations et oppositions phonologiques

Paires minimales

Oppositions phonologiques des voyelles à double timbre

Oppositions consonantiques

Distribution des phonèmes

Distribution, neutralisation et archiphonème

Changement et rendement phonologique

Aperçus phonologiques

Le système des phonèmes, valeurs oppositives et traits distinctifs

Séries vocaliques et corrélations consonantiques

Représentation des phonèmes par leurs traits distinctifs



## La simplification phonologique

Aspects psycho-physiologiques

Phénomènes de coarticulation

Mécanismes articulatoires et principe du moindre effort

Les consonnes et la loi du plus fort : les assimilations

Assimilations partielles

Facteurs psycholinguistiques

La loi du moindre effort face aux besoins de clarté et d'expressivité

Aspects acoustiques et auditifs

Degrés de périodicité des sons de la parole

Structuration harmonique des voyelles

Structuration acoustique des consonnes

Durée intrinsèque et allongement conditionné

Effets sonores et symbolisme des timbres vocaliques

Effets sonores des allongements vocaliques

Effets sonores et symbolisme des consonnes

Coloration sonore articulatoire du français

L'accentuation, les pauses et le rythme

L'accentuation sémantique

L'accentuation expressive : l'accent d'insistance

Groupe sémantique, groupe phonique et groupe de souffle

Pauses sémantiques, a-sémantiques et jonctures

L'accentuation, la durée syllabique et le rythme

L'utilisation poétique du rythme

Isochronie, diction classique et effets rythmiques

E caduc et diction classique des vers

Accentuation et liaison

## L'intonation

### La mélodie et l'intonation

#### Intonation, sémantique et phonologie

#### Rôle linguistique de l'intonation

#### Les niveaux et les patrons intonatifs du français

#### Enoncé déclaratif

#### L'incise et l'apposition

#### Questions de forme déclarative

#### Question avec inversion

#### Question avec mots interrogatifs

#### Ordre

#### Intonation de l'exclamation

#### Intonation et discours

### Les différents types de variation

#### Variation libre et conditionnée

#### Indices et signaux

#### Variation individuelle

#### Variation régionale

#### L'accent du Midi par rapport à celui du Nord

#### L'accent franco-québéco-canadien

#### Variation phonostylistique

#### Variation sociale valeur sociale de quelques variables

### C) CARTON, F.; *Introduction à la phonétique du français.*, Paris, Bordas 1974.

#### 1 Le phonétisme du français contemporain

##### 1.1 L'émission

###### 1.1.1 Généralités

###### 1.1.2 Caractéristiques physiologiques des consonnes

###### 1.1.3 Caractéristiques physiologiques des voyelles

##### 1.2 La transmission et la réception

- 1.2.1 Généralités
- 1.2.2 Caractéristiques acoustiques des voyelles
- 1.2.3 Caractéristiques acoustiques des consonnes
- 1.3 Le système phonologique du français contemporain
  - 1.3.1 Etude synchronique : terminologie et technique
  - 1.3.2 Système consonantique
  - 1.3.3 Système vocalique
  - 1.3.4 Divers types de description phonologique
  - 1.3.5 Fréquence d'apparition des phonèmes
- 1.4 Les combinaisons phonétiques du français
  - 1.4.1 Entre deux sons
  - 1.4.2 Entre deux syllabes
  - 1.4.3 Entre deux groupes accentuels et rythmiques
  - 1.4.4 Entre deux groupes de souffle
  - 1.4.5 Qu'est-ce qu'une syllabe ?
  - 1.4.6 Colorations secondaires
  - 1.4.7 Assimilation consonantique
  - 1.4.8 Harmonisation vocalique
  - 1.4.9 Liaison
  - 1.4.10 Elision
- 1.5 La prosodie du français
  - 1.5.1 Distinction entre accent et intonation
  - 1.5.2 L'intonation du français
  - 1.5.3 L'accent du français
  - 1.5.4 La quantité en français
- 1.6 Aspects phonostylistiques du français
  - 1.6.1 Peut-on définir une phonostylistique ?
  - 1.6.2 Les sons ont-ils une valeur stylistique par eux-mêmes ?
  - 1.6.3 L'intonation expressive
  - 1.6.4 L'accent d'insistance en français
- 1.7 Les caractères fondamentaux du français

- 1.7.1 Antériorité
- 1.7.2 Mode croissant
- 1.7.3 Tension
- 1.7.4 Egalité rythmique
- 2 Formation et évolution du phonétisme français
  - 2.1 [...]
  - 3 Comment prononcer le français contemporain ?
    - 3.1 Qu'est-ce que l'orthoépie ?
    - 3.2 Une norme du français ?
      - 3.2.1 [...]
    - 3.3 Les registres de langue orale
      - 3.3.1 [...]
    - 3.4 Les « lois » de position
    - 3.5 Quelques problèmes concernant les voyelles
      - 3.5.1 Les nasales
      - 3.5.2 Les deux A
      - 3.5.3 Le e muet
      - 3.5.4 [...]
    - 3.6 Divergences entre la graphie et le vocalisme
    - 3.7 Quelques problèmes concernant les consonnes
      - 3.7.1 [...]
      - 3.7.2 les semi-consonnes
      - 3.7.3 les consonnes finales
      - 3.7.4 divergences entre graphie et phonèmes consonantiques
    - 3.8 Quand faut-il allonger ?
      - 3.8.1 [...]
    - 3.9 Quand faut-il faire la liaison ?
    - 3.10 [...]

D) LEON, P., *Prononciation du français standard : aide-mémoire d'orthoépie, Paris, Didier, 1966, 4<sup>ème</sup> éd. 1978*

- 1 Voyelles
  - 1.1 Voyelles orales à un seul timbre
  - 1.2 Semi-consonnes
  - 1.3 Voyelles nasales
  - 1.4 Voyelles orales à double timbre
- 2 Consonnes
- 3 Liaisons
- 4 Mots spéciaux

E) WIOLAND, F., *Prononcer les mots du français, Coll F, Paris, Hachette, 1991*

- 1 L'identité des sons
  - 1.1 Le cadre rythmique
  - 1.2 L'identité articulatoire
  - 1.3 La fréquence des sons
  
- 2 L'unité rythmique
  - 2.1 Modèles rythmiques
  - 2.2 L'élément moteur
  - 2.3 Le modèle prosodique
  
- 3 Les lois de position
  - 3.1 Positions fortes/positions faibles
  - 3.2 Les voyelles en syllabe accentuée
  - 3.3 Les voyelles en syllabe inaccentuée
  
- 4 Les sons en contact
  - 4.1 Consonne + voyelle
  - 4.2 Voyelle + consonne
  - 4.3 Voyelle + voyelle
  - 4.4 Consonne + consonne

F) DERIVERY, N., *La phonétique du français*, Collection *Mémo*, Paris, Seuil, 1997

- 1 La phonétique et ses domaines
- 2 L'alphabet phonétique international
- 3 Eléments de phonétique articulatoire
- 4 Classement articulatoire des voyelles du français
- 5 Classement articulatoire des consonnes du français
- 6 Eléments de phonétique acoustique
- 7 Du phonème à la syllabe
- 8 Phonétique combinatoire
- 9 Eléments de phonologie
- 10 Faits prosodiques : l'accentuation et le rythme
- 11 Faits prosodiques : l'intonation
- 12 Les modes phonétiques du français

## **II. Pré-tests**

Voici ci-dessous d'une part le support de rendu distribué aux sujets ayant passé nos « pré-tests », et d'autre part les « consignes » de l'instructeur.

### **II.1. Document traduit en japonais et distribué aux étudiants**

Nous vous demandons :

- De NE PAS REFLECHIR ;
- De répondre DANS TOUS LES CAS, même si vous n'êtes pas sûr, et même si vous n'avez pas bien entendu ou bien vu ;
- De répondre SPONTANEMENT : ne cherchez pas à reconnaître des mots familiers car il n'y en a pas, ne cherchez pas à reconnaître des règles car il n'y en a pas, ne cherchez pas la bonne ou la mauvaise réponse car il n'y en a pas ;
- De répondre IMMEDIATEMENT, car vous n'avez que quelques secondes entre chaque épreuve ;
- De répondre simplement d'après CE QUE VOUS ENTENDEZ ET VOYEZ ;

- De bien RESPECTER les consignes ;
- D'ENTOURER les réponses que vous choisissez ;
- De lire ce document au fur et à mesure du test : ne TOURNEZ PAS les pages.  
ATTENDEZ que votre professeur vous l'indique.
- De remplir d'abord le tableau suivant (anonyme) :

<b><u>Sexe :</u></b>	M / F
<b><u>Age :</u></b>	17-24 / + de 25 ans
<b><u>Langue maternelle :</u></b>	Japonais / Autre (indiquez) : .....
<b><u>Année d'étude :</u></b>	1 / 2 / 3 / 4
<b><u>Discipline majeure :</u></b>	Français / Autre (indiquez) : .....
<b><u>Etude du français :</u></b>	0-6 mois / 6 mois-1 an / 1-2 ans / 2-4 ans / + de 4 ans

**I. Consigne :**

Ecoutez, puis indiquer si les deux mots que vous entendez ont une prononciation différente ou identique.

1)	Identique	Différent
2)	Identique	Différent
3)	Identique	Différent
4)	Identique	Différent
5)	Identique	Différent
6)	Identique	Différent

**II. Consigne :**

Ecoutez, puis indiquez la transcription qui vous semble la plus exacte, avec parfois un découpage des mots et parfois des simples cercles (cela est alors indiqué). Le symbole % est utilisé pour désigner un certain son.

1)	"ptebala" ou "putebala" ?	
	Si "ptebala", alors:  "p-tebala" ou "pt-ebala" ou "pte-bala" ou "pteb-ala" ?	Si "putebala", alors:  "p-utebala" ou "pu-tebala" ou "put-ebala" ou "pute-bala" ?
2)	"upaki" ou "%paki" ?	
3)	"bagade" ou "vagade" ?	
4)	"ptanabo" ou "putanabo"	
	Si "ptanabo", alors:  "p-tanabo" ou "pt-anabo" ou "pta-nabo" ou "ptan-abo" ?	Si "putanabo", alors:  "pu-tanabo" ou "put-anabo" ou "puta-nabo" ou "putan-abo" ?
5)	"usego" ou "%sego" ?	
6)	"O" ou "OO" ou "OOO" ?	
7)	"vatosi" ou "batosi" ?	

### **III. Consigne :**

En français, il y a de nombreux mots qui ont des homophones, c'est-à-dire qui s'écrivent avec des lettres différentes, mais qui se prononcent de la même manière. Vous devez dire, si les mots ci-dessous ont une prononciation identique ou différente. ATTENTION ! Vous devez bien respecter la consigne suivante :

- 1) Ne regardez pas ce qui est écrit sur votre feuille,
- 2) D'abord, regardez et lisez le mot n° 1 montré par votre professeur sur le tableau,
- 3) Regardez ensuite votre feuille : lisez d'abord la ligne avec les symboles, puis lisez la ligne suivante avec le mot n° 2.
- 4) Sans réfléchir et SANS Regarder le tableau une deuxième fois, indiquez immédiatement si la prononciation des mots n° 1 et n° 2 est identique ou différente.



1)	<ul style="list-style-type: none"> <li>- \$x\$x\$x\$x\$x\$x\$x\$x\$x\$x\$</li> <li>- <i>ogavi</i></li> <li>- Identique ou différent ?</li> </ul>
2)	<ul style="list-style-type: none"> <li>- *y*y*y*y*y*y*y*y*y*y*</li> <li>- <i>ekobu</i></li> <li>- Identique ou différent ?</li> </ul>
3)	<ul style="list-style-type: none"> <li>- μξμξμξμξμξμξμξμξμξ</li> <li>- <i>avigu</i></li> <li>- Identique ou différent ?</li> </ul>
4)	<ul style="list-style-type: none"> <li>- §Z§Z§Z§Z§Z§Z§Z§Z§Z§</li> <li>- <i>obeka</i></li> <li>- Identique ou différent ?</li> </ul>

**IV. Consigne :**

Respectez bien la consigne suivante :

- 1) Ici, on respecte l'orthographe française (par exemple "ou" prononcé [u]),
- 2) Regardez le tableau et lisez le mot montré par votre professeur sur le tableau,
- 3) Regardez votre feuille : sans trop réfléchir et SANS Regarder le tableau une deuxième fois, choisissez la transcription qui vous semble la plus appropriée.

1)	"OO" ou "OOO" ?
2)	"ap-ta" ou "a-pta" ou "apt-a" ?

**V. Consigne :**

Respectez bien la consigne suivante :

- Ecoutez et regardez en même temps le mot montré au tableau par votre professeur.
- Puis, répétez à voix haute.

- Puis, SANS Regarder une deuxième fois au tableau, et SANS trop réfléchir, choisissez sur votre feuille la transcription qui vous semble correspondre le plus exactement à ce que vous avez répété.

1)	<i>"okotle"</i> ou <i>"okotole"</i> ?	
	Si <i>"okotle"</i> , alors:  <i>"oko-tle"</i> ou <i>"okot-le"</i> ou <i>"okotl-e"</i> ?	Si <i>"okotole"</i> , alors:  <i>"oko-to-le"</i> ou <i>"okot-o-le"</i> ou <i>"oko-tol-e"</i> ?
2)	<i>"adotle"</i> ou <i>"adotole"</i>	
	Si <i>"adotle"</i> , alors:  <i>"ado-tle"</i> ou <i>"adot-le"</i> ou <i>"adotl-e"</i> ?	Si <i>"adotole"</i> , alors:  <i>"ado-to-le"</i> ou <i>"adot-o-le"</i> ou <i>"ado-tol-e"</i> ?
3)	<i>"afremo"</i> ou <i>"afuremo"</i>	
	Si <i>"afremo"</i> alors:  <i>"a-fremo"</i> ou <i>"af-remo"</i> ou <i>"afr-emo"</i> ?	Si <i>"afuremo"</i> alors:  <i>"a-fu-remo"</i> ou <i>"af-u-remo"</i> ou <i>"a-fur-emo"</i> ?
4)	<i>"OO"</i> ou <i>"OOO"</i> ou <i>"OOOO"</i> ?	
5)	<i>"afregu"</i> ou <i>"afuregu"</i> ?	
	Si <i>"afregu"</i> alors:  <i>"a-fregu"</i> ou <i>"af-regu"</i> ou <i>"afr-egu"</i> ?	Si <i>"afuregu"</i> alors:  <i>"a-fu-regu"</i> ou <i>"af-u-regu"</i> ou <i>"a-fur-egu"</i> ?

## **II.2. Notice de l'instructeur**

Cher collègue,

Vous avez bien voulu accepter de m'aider dans cette enquête et je vous en remercie très vivement. Votre aide est très appréciée. J'espère par ailleurs que ces tests vous intéresseront et sensibiliseront vos étudiants aux problèmes de perception, de prononciation et de lecture dans l'apprentissage du français langue étrangère.

Outre leur fonction de sensibilisation, ils pourront, je l'espère, présenter quelque aspect ludique pour vos étudiants, et leur passation pourra ainsi constituer un moment de détente. Enfin, j'espère qu'ils vous permettront d'apprécier les problèmes (ou l'absence de problème !) de perception des sons du français de vos étudiants, en rapport notamment avec la graphie, employée dans les manuels et autres documents pédagogiques.

Voici-ci dessous les indications et consignes nécessaires au bon déroulement de la passation des épreuves.

1. **Introduction** : présenter cette expérience comme une sensibilisation à la prononciation et à l'écriture du français.
  - Demander aux étudiants de BIEN lire les consignes, de les RESPECTER scrupuleusement, ainsi que de lire le document AU FUR ET A MESURE des épreuves (et non de parcourir le document une fois en entier avant de commencer).
  - Distribuer le document aux étudiants.
  - Si vous lisez le japonais, lire les premières consignes à voix haute devant les étudiants.
  - Dire aux étudiants qu'ils doivent répondre IMMEDIATEMENT car ils n'ont en moyenne que 5 secondes entre chaque épreuve.
  - Laissez les étudiants remplir le tableau d'information anonyme.

2. **Partie I (Discrimination auditive) :**

- ⇒ Passez la cassette. Le contenu de la cassette se présente ainsi : numéro de la partie – 2 sec - numéro du stimulus – 1 sec – stimulus 1 – 1 sec – stimulus 2 – 5 sec – numéro du stimulus suivant.
- ⇒ Quand vous entendez « STOP » (après 1min 10sec environ), arrêtez la cassette et faites une pause (20 secondes maximum) pour laisser les étudiants lire la consigne suivante, mais pas plus.
- ⇒ Dites aux étudiants de lire la consigne suivante.

3. **Partie II (Choix de transcription à partir de l'auditif) :**

- ⇒ Passez la cassette. Le contenu de la cassette se présente ainsi : - numéro de la partie – 2 sec – numéro du stimulus – 1 sec – stimulus 1 – 25 secondes - numéro du stimulus suivant.
- ⇒ Quand vous entendez « STOP » (après 2 min 50 sec. environ), arrêtez la cassette et faites une pause (20 secondes maximum) pour laisser les étudiants lire la consigne suivante, mais pas plus.
- ⇒ Dites aux étudiants de lire la consigne suivante.

4. **Partie III (Discrimination à partir du visuel) :**

Le principe est le suivant :

- ⇒ Vous écrivez un « non-mot » (voir liste plus bas) au tableau (ou sur transparent),
- ⇒ Puis après 1 seconde (pas plus !) vous l'effacez immédiatement.
- ⇒ Puis vous laissez 5 secondes aux étudiants pour regarder leur feuille et répondre.
- ⇒ Puis vous recommencez avec le non-mot suivant.

Voici les non-mots, à écrire au fur et à mesure (en grandes lettres majuscules) : 1) OGABI, 2) EKOBUI, 3) AVIGU, 4) OBEKA.

- ⇒ Ensuite, dites aux étudiants de lire la consigne suivante.

5. **Partie IV (Choix de transcription à partir du visuel):**

Le principe est le même que pour l'épreuve précédente.

- ⇒ Ecrivez au tableau, au fur et à mesure (en grandes lettres majuscules) : 1) PLATEAU, 2) APTA.
- ⇒ Ensuite, dites aux étudiants de lire la consigne suivante.

#### 6. **Partie V (Choix de transcription à partir de audiovisuel simultané) :**

Ici, vous allez combiner la cassette et l'écriture au tableau.

- ⇒ Passez la cassette. Lorsque vous entendez le numéro du stimulus, vous écrivez au tableau le « non-mot » correspondant au numéro, puis vous l'effacez aussitôt dès que vous entendez le stimulus lui-même. Vous laissez les étudiants répéter à voix haute 1 fois (c'est important, s'ils ne le font pas dites « répétez à voix haute 1 fois » en japonais si possible), puis vous laissez les étudiants répondre.
- ⇒ Quand vous entendez le numéro suivant, vous recommencez.
- ⇒ Quand vous entendez « STOP » (après 2 min 50 sec. environ), arrêtez la cassette et relevez les « copies ».

Le contenu de la cassette se présente ainsi : numéro de la partie – 2 sec – numéro du stimulus – 8 sec – stimulus 1 – 25 secondes - numéro du stimulus suivant.

Ecrivez au tableau, au fur et à mesure (en grandes lettres majuscules) : 1) OKOTLE, 2) ADOTOLE, 3) AFREMO, 4) IKRUSH, 5) AFUREGU.

### **III. Tests Généraux**

Le matériel utilisé par les instructeurs pour la passation des tests généraux était le suivant :

- \*La notice de l'instructeur (consignes et procédure).
- \*La fiche de description des conditions de passation (à remplir par l'instructeur).
- \*Le document de rendu en japonais (à remplir par les sujets) (10 pages).
- \*Le document de vérification du matériel pour l'instructeur (document de rendu des sujets en français et liste des stimuli).
- \*Les stimuli visuels (sur support papier, sous formes de bandes découpées à projeter par rétroprojecteur) (tests D et E).
- \*Les stimuli auditifs (une cassette audio).

### **III.1. La notice de l'instructeur**

#### **III.1.1. Tests A, F et C1**

##### **Procédure**

Heure : 0 min

Introduction :

1. Présentez les activités comme des exercices d'écoute et de prononciation (B/V, R/L, etc.).
2. Distribuez les documents aux étudiants et insistez pour qu'ils lisent bien les consignes et qu'ils remplissent bien les informations personnelles. Les retardataires ont le temps d'arriver.

Heure : 15 min

Exercice A :

Vous présentez l'exemple suivant pour que les étudiants comprennent bien (cet exemple est déjà indiqué sur leur feuille en japonais) :

*« Ecoutez les mots sur la cassette. Puis choisissez la transcription qui vous semble correspondre le mieux à ce que vous écoutez. Attention, vous n'avez que quelques secondes, donc répondez immédiatement. Par exemple : vous entendez le mot [soba] sur la cassette. Alors vous devez choisir l'écriture qui correspond à ce que vous entendez entre les deux choix proposés, par exemple : « zoba »/« soba » ( ?). Vous entourez la réponse choisie (par exemple « soba »). Parfois il y a quatre transcriptions proposées : vous n'en choisissez qu'une seule. Répondez immédiatement car il n'y a pas beaucoup de temps. Chaque mot est séparé par un « BIP ». Suivez les lignes sur votre feuille avec votre doigt pour ne pas vous tromper de ligne ».*

Passez la cassette (6 minutes environ) et ils font l'exercice. Pendant ce temps, s'il vous plaît, remplissez autant que possible la fiche de description des conditions de passation.

Heure : 21 min

Exercice F :

Vous présentez l'exemple suivant pour que les étudiants comprennent bien (cet exemple est déjà indiqué sur leur feuille en japonais) :

*« Sur la cassette, vous entendez deux mots prononcés l'un après l'autre (par exemple [soba]/[zoba] ou par exemple [kaki]/[kaki]). Vous devez dire si ces deux mots se prononcent de manière identique (l'exemple de « kaki ») ou différente (l'exemple de « soba ») : pour cela, vous choisissez la transcription qui vous semble la plus adaptée pour chacun des deux mots que vous entendez. Si exceptionnellement vous pensez que les transcriptions ne correspondent pas exactement (par exemple il y a « pake » et « take », mais vous entendez [taake]), choisissez d'abord la transcription en français qui correspond le mieux (par exemple « take »), puis écrivez votre propre transcription en katakana « taake ». Mais, pour chaque mot, choisissez toujours d'abord une transcription en français. Répondez immédiatement car il n'y a pas beaucoup de temps. Chaque mot est séparé par un « BIP ». Suivez les lignes sur votre feuille avec votre doigt pour ne pas vous tromper de ligne en répondant ».*

Numéro	1 <sup>er</sup> mot	2 <sup>ème</sup> mot	Transcription 1 <sup>er</sup> mot	Transcription 2 <sup>ème</sup> mot
<b>Exemple Différent</b>	Soba	Soba		
	Zoba	Zoba		
<b>Exemple Identique</b>	Kaki	Kaki		
	Kagi	Kagi		
<b>Exemple Précisé</b>	Take	Take	TAKE	TAAKE
	Pake	Pake		

Passer la cassette (13 minutes environ) et ils font l'exercice.

Heure : 33 min

Exercice C :

Vous présentez l'exemple suivant pour que les étudiants comprennent bien (cet exemple est déjà indiqué sur leur feuille en japonais) :

*« Ecoutez les mots sur la cassette. Puis, tout de suite, découpez ces mots en syllabes et choisissez le nombre de cercles qui vous semble correspondre au nombre de syllabes que vous entendez. Attention, vous n'avez que quelques secondes, donc répondez immédiatement. Par exemple : vous entendez le mot [soba]. Alors vous comptez les syllabes ([sob-a] ou [so-ba] (comme vous voulez) → 2 syllabes). Vous choisissez alors le nombre de cercles correspondant ([sob-a] ou [so-ba] → 2 syllabes → OO). Répondez immédiatement car il n'y a pas beaucoup de temps. Chaque mot est séparé par un « BIP ». Suivez les lignes sur votre feuille avec votre doigt pour ne pas vous tromper de ligne en répondant ».*

Vous passez la cassette (12 minutes environ) et ils font l'exercice.

Heure : 58 min.

Vous relevez les copies (l'important est que les feuilles de chaque étudiant soient bien ensemble (agrafées)) avec le descriptif des conditions de passation.



### III.1.2. Tests C2, D et E

#### Procédure

Heure : 0 min.

Introduction :

1. Présentez les activités comme des exercices d'écoute et de prononciation.
2. Distribuez les documents aux étudiants et insistez pour qu'ils lisent bien les consignes et qu'ils remplissent bien les informations personnelles. Pendant ce temps, les retardataires ont le temps d'arriver.

Heure : 15 min.

Exercice C :

Vous présentez l'exemple suivant pour que les étudiants comprennent bien (cet exemple est déjà inscrit sur leur feuille en japonais) :

*« Ecoutez les mots sur la cassette. Puis, tout de suite, découpez ces mots en syllabes et choisissez le nombre de cercles qui vous semble correspondre au nombre de syllabes que vous entendez. Attention, vous n'avez que quelques secondes, donc répondez immédiatement. Par exemple : vous entendez le mot [soba]. Alors vous comptez les syllabes ([sob-a] ou [so-ba] (comme vous voulez) → 2 syllabes). Vous choisissez alors le nombre de cercles correspondant ([sob-a] ou [so-ba] → 2 syllabes → OO). Répondez immédiatement car il n'y a pas beaucoup de temps. Chaque mot est séparé par un « BIP ». Suivez les lignes sur votre feuille avec votre doigt pour ne pas vous tromper de ligne en répondant ».*

Vous passez la cassette (21 minutes environ) et ils font l'exercice. Pendant ce temps, s'il vous plaît, remplissez autant que possible la fiche de description des conditions de passation.

Heure : 36 min

Exercice E :

Vous présentez l'exemple suivant pour que les étudiants comprennent bien (cet exemple est déjà inscrit sur leur feuille en japonais) :

*« Ecoutez et lisez les mots qui sont présentés. Puis, tout de suite, découpez ces mots en syllabes et choisissez le nombre de cercles qui vous semble correspondre au nombre de syllabes de ce que vous écoutez et lisez. Si exceptionnellement vous pensez remarquer des différences entre ce que vous lisez et écoutez, entourez la case « Différent » mais entourez quand même aussi le nombre de syllabes de ce que vous entendez. Attention, vous n'avez que quelques secondes, donc répondez immédiatement. Par exemple : vous lisez rapidement un mot sur l'écran / tableau (par exemple « soba ») et sur la cassette vous entendez ce mot prononcé (par exemple [soba]). Alors vous devez compter les syllabes ([sob-a] ou [so-ba] (comme vous voulez) → 2 syllabes), et vous choisissez alors le nombre de cercles correspondant à au nombre de syllabes ( [sob-a] ou [so-ba] (comme vous voulez) → 2 syllabes → OO). Répondez immédiatement car il n'y a pas beaucoup de temps. Chaque mot est séparé par un « BIP ». Suivez les lignes sur votre feuille avec votre doigt pour ne pas vous tromper de ligne en répondant ».*

Vous passez la cassette (31 min) : à chaque BIP, vous mettez sur le rétroprojecteur le mot qui est dans votre liste (155 mots) et dès que vous entendez le mot prononcé (environ 3 secondes ensuite) vous retirez la feuille du rétroprojecteur (9 secondes entre chaque BIP).

Heure : 1H07

Exercice D :

Vous présentez l'exemple suivant pour que les étudiants comprennent bien :

*« Lisez les mots au tableau. Puis, tout de suite, découpez ces mots en syllabes et choisissez le nombre de cercles qui vous semble correspondre au nombre de syllabes que vous lisez. Attention, vous n'avez que quelques secondes, donc répondez immédiatement. Par exemple : vous lisez le mot « soba ». Alors vous comptez les syllabes ([sob-a] ou [so-ba] (comme vous voulez) → 2 syllabes). Vous choisissez alors le nombre de cercles correspondant ([sob-a] ou [so-ba] → 2 syllabes → OO). Répondez immédiatement car il n'y a pas beaucoup de temps. Chaque mot est séparé par un « BIP ». Suivez les lignes sur votre feuille avec votre doigt pour ne pas vous tromper de ligne en répondant ».*

Passez la cassette : à chaque BIP, vous mettez sur le rétroprojecteur le mot qui est dans votre liste (62 mots). Ensuite vous comptez jusqu'à 3, vous retirez la feuille du rétroprojecteur et vous attendez le prochain BIP (7 secondes entre chaque BIP).

Heure : 1H16

Vous relevez les copies (l'important est que les feuilles de chaque étudiant soient bien ensemble (agrafées)), avec le descriptif des conditions.

### **III.2. Description des conditions de passation**

↻	<u>Type de classe</u> (ex : 1 <sup>ère</sup> année - 2 <sup>ème</sup> langue - Grammaire) :.....
↻	<u>Date</u> :.....
↻	<u>Lieu</u> (si possible) : Université.....
↻	<u>Horaire du cours</u> : .....
↻	<u>Salle</u> : grande / petite ---- lumineuse/sombre
↻	<u>Nombre d'étudiants</u> :
↻	<u>Activités précédentes des étudiants</u> : pause / divers / classe (de .....
↻	<u>Enceintes audio</u> : face aux étudiants / derrière / sur les côtés
↻	<u>Nombre d'enceintes</u> : .....
↻	<u>Type d'enceintes</u> (marque et code, si possible) :.....
↻	<u>Type de magnétophone</u> (marque et code, si possible) :.....
↻	<u>Volume</u> : .....
↻	<u>Type de rétroprojecteur</u> (marque et code, si possible) :.....
↻	<u>Heure Début des tests</u> : .....
↻	<u>Heure Fin des tests</u> : .....
↻	<u>Remarques</u> : .....

### **III.3. Documents de rendu des sujets**

Nous présentons ci-dessous les versions françaises des documents, dans lesquelles nous avons ici, afin de ne pas gonfler exagérément le volume de ces annexes, réduit en partie la taille de la police par rapport aux documents originaux.

#### **III.3.1. Tests A, F et C1 (en français)**

**Pour tous les exercices, il faut IMPERATIVEMENT :**

- NE PAS REFLECHIR (tous les mots présentés sont inconnus de vous) et répondre SPONTANEMENT (il n'y a pas de règle à trouver, ni de bonne ou mauvaise réponse ; il faut juste choisir la réponse de manière naturelle)
- Répondre DANS TOUS LES CAS, même si vous n'êtes pas sûr,
- Répondre IMMEDIATEMENT, car vous n'avez que quelques secondes entre chaque épreuve ;
- Répondre SIMPLEMENT d'après CE QUE VOUS ENTENDEZ ET VOYEZ ;
- ENTOURER les réponses que vous choisissez ;
- Lire les documents donnés au fur et à mesure de l'exercice : ne TOURNEZ PAS les pages en avance.
- Bien RESPECTER les consignes ;

**Veillez remplir COMPLETEMENT les deux tableaux d'information suivants :**

<u>Sexe :</u>	M / F
<u>Age :</u>	17 / 18 / 19 / 20 / 21 / 22 / 23 / 24 / 25 / + de 25 ans
<u>Langue maternelle :</u>	Japonais / Autre (indiquez) : .....
<u>Année d'étude universitaire :</u>	1 <sup>ère</sup> / 2 <sup>ème</sup> / 3 <sup>ème</sup> / 4 <sup>ème</sup>
<u>Discipline majeure :</u>	Français / Autre (indiquez) : .....

Etude du français :

Durée d'étude du français :	0-6 mois / 6 mois-1 an / 1-2 ans / 2-4 ans / + de 4 ans
Avez-vous étudié le français au lycée ?	OUI / NON
Si oui, combien de temps ?	1 an / 2 ans / 3 ans
Vous avez étudié	Surtout l'écrit/ Surtout l'oral / Les deux équilibrés
Avez-vous déjà eu des classes de prononciation ?	OUI / NON
Avez-vous déjà eu des classes de dictée ?	OUI / NON
Le français, c'est votre :	1 <sup>ère</sup> langue étrangère / 2 <sup>ème</sup> / 3 <sup>ème</sup>
Si le français n'est pas votre 1 <sup>ère</sup> langue étrangère, quelle est cette 1 <sup>ère</sup> langue ?	Anglais / Chinois / Coréen / Espagnol / Allemand / Autre :.....
Quel est votre niveau dans cette 1 <sup>ère</sup> langue ?	Faible / Moyen / Bon / Très Bon
Etes-vous bilingue ?	OUI / NON
Si oui, la deuxième langue est-elle une langue européenne ? (« American English » inclus)	OUI / NON Langue :.....
L'un de vos parents est-il bilingue ?	OUI / NON
Si oui, la deuxième langue est-elle une langue européenne (« American English » inclus) ?	OUI / NON Langue.....
Avez-vous vécu plus de 6 mois à l'étranger ?	OUI / NON
Si oui, dans un pays :	Francophone / Anglophone / Européen / Autre
Etes-vous gaucher ou droitier ?	Gaucher / Droitier
Étudiez-vous le français en plus des cours de l'université (école, cours privés, etc.) ?	NON / UN PEU / OUI
Entendez-vous/ Parlez-vous beaucoup le français ? (amis, TV, etc.)	NON / PARFOIS / OUI
Entendez-vous/ Parlez-vous beaucoup l'anglais? (amis, TV, classes, etc.)	NON / PARFOIS / OUI

**Exercice A - Consignes :**

Ecoutez les mots sur la cassette. Puis choisissez la transcription qui vous semble correspondre le mieux à ce que vous écoutez. Attention, vous n'avez que quelques secondes, donc répondez immédiatement. Par exemple : vous entendez le mot [soba] sur la cassette. Alors vous devez choisir l'écriture qui correspond à ce que vous entendez entre les deux choix proposés, par exemple : « zoba »/« soba » ( ?). Vous entourez la réponse choisie (par exemple « soba »). Parfois il y a quatre transcriptions proposées : vous n'en choisissez qu'une seule. Répondez immédiatement car il n'y a pas beaucoup de temps. Chaque mot est séparé par un « BIP ». **Suivez les lignes sur votre feuille avec votre doigt pour ne pas vous tromper de ligne en répondant.**

3. LEUKEGI	ROUKEGI	LOUKEGI	REUKEGI
4. MEJASOUB	MEJASEUV	MEJASOUV	MEJASEUB
5. RADEBO	LADEBO		
6. JODENIR	JODENIL		
7. NAGEVA	NAGEBA		
8. SATODOUL	SATODEUL	SATODEUR	SATODOUR
9. MOTEVI	MOTEBI		
10. PEKEZOUR	PEKEZOUL	PEKEZEUL	PEKEZEUR
11. LIPISA	RIPISA		
12. KOMETAR	KOMETAL		
13. MINELOU	MINEREU	MINEROU	MINELEU
14. JODENIR	JODENIL		
15. ZOTEBI	ZOTEVI		
16. ZASHAKIV	ZASHAKIB		
17. POTIGAV	POTIGAB		
18. NELETEUV	NELETOUV	NELETEUB	NELETOUB
19. GINEROU	GINELEU	GINEREU	GINELOU
20. VITANE	BITANE		
21. PEKEZOUR	PEKEZOUL	PEKEZEUL	PEKEZEUR
22. LIPISA	RIPISA		
23. SENAVOU	SENABOU	SENAVEU	SENABEU
24. LEUDOKE	LOUDOKE	REUDOKE	ROUDOKE
25. NELETEUV	NELETEUB	NELETOUV	NELETOUB
26. ROUDOKE	LOUDOKE	REUDOKE	LEUDOKE
27. BOUSINE	VOUSINE	BEUSINE	VEUSINE
28. DINEREU	DINELEU	DINELOU	DINEROU
29. RADEBO	LADEBO		
30. VOUSINE	BOUSINE	VEUSINE	BEUSINE
31. ZASHAKIB	ZASHAKIV		
32. KEPIRA	KEPILA		
33. BAGADE	VAGADE		
34. BOUSOKI	BEUSOKI	VEUSOKI	VOUSOKI

35. MEJASOUV	MEJASOUB	MEJASEUB	MEJASEUV
36. TEPIRA	TEPILA		
37. BENABEU	BENABOU	BENAVEU	BENAVOU
38. ZINELOU	ZINEROU	ZINEREU	ZINELEU
39. POTIGAB	POTIGAV		
40. BEUSOKI	VEUSOKI	VOUSOKI	BOUSOKI
41. PEJALI	PEJARI		
42. KOMETAR	KOMETAL		
43. NEJARI	NEJALI		
44. BAGADE	VAGADE		
45. LEUKEGI	ROUKEGI	REUKEGI	LOUKEGI
46. JEMABOU	JEMABEU	JEMAVEU	JEMAVOU
47. SATODEUL	SATODEUR	SATODOUR	SATODEUL
48. VITANE	BITANE		
49. DEMABOU	DEMAVOU	DEMAVEU	DEMABEU
50. PAGEBA	PAGEVA		

### **Exercice F - Consignes :**

Sur la cassette, vous entendez deux mots prononcés l'un après l'autre (par exemple [soba]/[zoba] ou par exemple [kaki]/[kaki]). Vous devez dire si ces deux mots se prononcent de manière identique (l'exemple de « kaki ») ou différente (l'exemple de « soba ») : pour cela, vous choisissez la transcription qui vous semble la plus adaptée pour chacun des deux mots que vous entendez. Si exceptionnellement vous pensez que les transcriptions ne correspondent pas exactement (par exemple il y a « pake » et « take », mais vous entendez [taake]), choisissez d'abord la transcription en français qui correspond le mieux (par exemple « take »), puis écrivez votre propre transcription en katakana « taake ». Mais, pour chaque mot, choisissez toujours d'abord une transcription en français. Répondez immédiatement car il n'y a pas beaucoup de temps. Chaque mot est séparé par un « BIP ». **Suivez les lignes sur votre feuille avec votre doigt pour ne pas vous tromper de ligne en répondant.**

Numéro	1 <sup>er</sup> mot	2 <sup>ème</sup> mot	Transcription 1 <sup>er</sup> mot	Transcription 2 <sup>ème</sup> mot
<b>Exemple Différent</b>	Soba	Soba		
	Zoba	Zoba		
<b>Exemple Identique</b>	Kaki	Kaki		
	Kagi	Kagi		
<b>Exemple Précisé</b>	Take	Take	TAKE	TAAKE
	Pake	Pake		



N°	1 <sup>er</sup> MOT	2 <sup>ème</sup> MOT	Transcriptions éventuelles	
			1 <sup>er</sup> mot	2 <sup>ème</sup> mot
2.	POTIGAV	POTIGAV		
	POTIGAB	POTIGAB		
3.	JODENIR	JODENIR		
	JODENIL	JODENIL		
4.	PEKEZOUL	PEKEZOUL		
	PEKEZOUR	PEKEZOUR		
5.	NAGEVA	NAGEVA		
	NAGEBA	NAGEBA		
6.	LEUKEGI	LEUKEGI		
	REUKEGI	REUKEGI		
7.	POTIGAV	POTIGAV		
	POTIGAB	POTIGAB		
8.	MOTEBI	MOTEBI		
	MOTEVI	MOTEVI		
9.	SATODEUR	SATODEUR		
	SATODEUL	SATODEUL		
10.	RADEBO	RADEBO		
	LADEBO	LADEBO		
11.	MEJASOUV	MEJASOUV		
	MEJASOUB	MEJASOUB		
12.	ZASHAKIB	ZASHAKIB		
	ZASHAKIV	ZASHAKIV		
13.	ZOTEVI	ZOTEVI		
	ZOTEBI	ZOTEBI		
14.	NELETEUB	NELETEUB		
	NELETEUV	NELETEUV		
15.	LIPISA	LIPISA		
	RIPISA	RIPISA		
16.	JODENIL	JODENIL		
	JODENIR	JODENIR		
17.	MINEROU	MINEROU		
	MINELOU	MINELOU		
18.	MEJASOUV	MEJASOUV		
	MEJASOUB	MEJASOUB		
19.	KOMETAL	KOMETAL		
	KOMETAR	KOMETAR		
20.	LEUKEGI	LEUKEGI		
	REUKEGI	REUKEGI		
21.	NELETEUB	NELETEUB		
	NELETEUV	NELETEUV		
22.	JODENIR	JODENIR		
	JODENIL	JODENIL		
23.	RADEBO	RADEBO		
	LADEBO	LADEBO		
24.	ZASHAKIB	ZASHAKIB		
	ZASHAKIV	ZASHAKIV		
25.	JODENIL	JODENIL		
	JODENIR	JODENIR		
26.	NAGEVA	NAGEVA		
	NAGEBA	NAGEBA		

27.	SATODEUL	SATODEUL		
	SATODEUR	SATODEUR		
28.	MOTEBI	MOTEBI		
	MOTEVI	MOTEVI		
29.	PEKEZOUL	PEKEZOUL		
	PEKEZOUR	PEKEZOUR		
30.	LIPISA	LIPISA		
	RIPISA	RIPISA		
31.	KOMETAL	KOMETAL		
	KOMETAR	KOMETAR		
32.	MINEROU	MINEROU		
	MINELOU	MINELOU		
33.	ZOTEVI	ZOTEVI		
	ZOTEBI	ZOTEBI		
34.	VAGADE	VAGADE		
	BAGADE	BAGADE		
35.	VOUSINE	VOUSINE		
	BOUSINE	BOUSINE		
36.	ROUDOKE	ROUDOKE		
	LOUDOKE	LOUDOKE		
37.	SENAVEU	SENAVEU		
	SENABEU	SENABEU		
38.	GINEREU	GINEREU		
	GINELEU	GINELEU		
39.	VAGADE	VAGADE		
	BAGADE	BAGADE		
40.	NELETEUV	NELETEUV		
	NELETEUB	NELETEUB		
41.	LOUDOKE	LOUDOKE		
	ROUDOKE	ROUDOKE		
42.	PEKEZOUR	PEKEZOUR		
	PEKEZOUL	PEKEZOUL		
43.	BITANE	BITANE		
	VITANE	VITANE		
44.	KEPIRA	KEPIRA		
	KEPILA	KEPILA		
45.	ZASHAKIV	ZASHAKIV		
	ZASHAKIB	ZASHAKIB		
46.	BEUSOKI	BEUSOKI		
	VEUSOKI	VEUSOKI		
47.	BOUSINE	BOUSINE		
	VOUSINE	VOUSINE		
48.	RIPISA	RIPISA		
	LIPISA	LIPISA		
49.	LADEBO	LADEBO		
	RADEBO	RADEBO		
51.	BITANE	BITANE		
	VITANE	VITANE		
52.	DINELEU	DINELEU		
	DINEREU	DINEREU		
53.	GINEREU	GINEREU		
	GINELEU	GINELEU		

54.	<b>BEUSOKI</b>	<b>BEUSOKI</b>		
	<b>VEUSOKI</b>	<b>VEUSOKI</b>		
55.	<b>VOUSINE</b>	<b>VOUSINE</b>		
	<b>BOUSINE</b>	<b>BOUSINE</b>		
56.	<b>PEKEZOUR</b>	<b>PEKEZOUR</b>		
	<b>PEKEZOUL</b>	<b>PEKEZOUL</b>		
57.	<b>KEPIRA</b>	<b>KEPIRA</b>		
	<b>KEPILA</b>	<b>KEPILA</b>		
58.	<b>RIPISA</b>	<b>RIPISA</b>		
	<b>LIPISA</b>	<b>LIPISA</b>		
59.	<b>SENAVEU</b>	<b>SENAVEU</b>		
	<b>SENABEU</b>	<b>SENABEU</b>		
60.	<b>LOUDOKE</b>	<b>LOUDOKE</b>		
	<b>ROUDOKE</b>	<b>ROUDOKE</b>		
61.	<b>NELETEUV</b>	<b>NELETEUV</b>		
	<b>NELETEUB</b>	<b>NELETEUB</b>		
62.	<b>ROUDOKE</b>	<b>ROUDOKE</b>		
	<b>LOUDOKE</b>	<b>LOUDOKE</b>		
63.	<b>BOUSINE</b>	<b>BOUSINE</b>		
	<b>VOUSINE</b>	<b>VOUSINE</b>		
64.	<b>DINELEU</b>	<b>DINELEU</b>		
	<b>DINEREU</b>	<b>DINEREU</b>		
65.	<b>LADEBO</b>	<b>LADEBO</b>		
	<b>RADEBO</b>	<b>RADEBO</b>		
66.	<b>ZASHAKIV</b>	<b>ZASHAKIV</b>		
	<b>ZASHAKIB</b>	<b>ZASHAKIB</b>		
67.	<b>DEMABOU</b>	<b>DEMABOU</b>		
	<b>DEMAVOU</b>	<b>DEMAVOU</b>		
68.	<b>JEMAVOU</b>	<b>JEMAVOU</b>		
	<b>JEMABOU</b>	<b>JEMABOU</b>		
69.	<b>KOMETAR</b>	<b>KOMETAR</b>		
	<b>KOMETAL</b>	<b>KOMETAL</b>		
70.	<b>POTIGAB</b>	<b>POTIGAB</b>		
	<b>POTIGAV</b>	<b>POTIGAV</b>		
71.	<b>MEJASOUB</b>	<b>MEJASOUB</b>		
	<b>MEJASOUV</b>	<b>MEJASOUV</b>		
72.	<b>DEMABOU</b>	<b>DEMABOU</b>		
	<b>DEMAVOU</b>	<b>DEMAVOU</b>		
73.	<b>PEJARI</b>	<b>PEJARI</b>		
	<b>PEJALI</b>	<b>PEJALI</b>		
74.	<b>VEUSOKI</b>	<b>VEUSOKI</b>		
	<b>BEUSOKI</b>	<b>BEUSOKI</b>		
75.	<b>BENABEU</b>	<b>BENABEU</b>		
	<b>BENAVEU</b>	<b>BENAVEU</b>		
76.	<b>TEPILA</b>	<b>TEPILA</b>		
	<b>TEPIRA</b>	<b>TEPIRA</b>		
77.	<b>VITANE</b>	<b>VITANE</b>		
	<b>BITANE</b>	<b>BITANE</b>		
78.	<b>SATODEUR</b>	<b>SATODEUR</b>		
	<b>SATODEUL</b>	<b>SATODEUL</b>		
79.	<b>PAGEBA</b>	<b>PAGEBA</b>		
	<b>PAGEVA</b>	<b>PAGEVA</b>		

80.	NEJALI	NEJALI		
	NEJARI	NEJARI		
81.	ZINELOU	ZINELOU		
	ZINEROU	ZINEROU		
82.	REUKEGI	REUKEGI		
	LEUKEGI	LEUKEGI		
83.	TEPILA	TEPILA		
	TEPIRA	TEPIRA		
84.	BAGADE	BAGADE		
	VAGADE	VAGADE		
85.	MEJASOUB	MEJASOUB		
	MEJASOUV	MEJASOUV		
86.	PAGEBA	PAGEBA		
	PAGEVA	PAGEVA		
87.	JEMAVOU	JEMAVOU		
	JEMABOU	JEMABOU		
88.	BENABEU	BENABEU		
	BENAVEU	BENAVEU		
89.	VITANE	VITANE		
	BITANE	BITANE		
90.	ZINELOU	ZINELOU		
	ZINEROU	ZINEROU		
91.	POTIGAB	POTIGAB		
	POTIGAV	POTIGAV		
92.	VEUSOKI	VEUSOKI		
	BEUSOKI	BEUSOKI		
93.	PEJARI	PEJARI		
	PEJALI	PEJALI		
94.	KOMETAR	KOMETAR		
	KOMETAL	KOMETAL		
95.	NEJALI	NEJALI		
	NEJARI	NEJARI		
96.	BAGADE	BAGADE		
	VAGADE	VAGADE		
97.	REUKEGI	REUKEGI		
	LEUKEGI	LEUKEGI		
98.	SATODEUR	SATODEUR		
	SATODEUL	SATODEUL		

### **Exercice C -Consignes :**

Ecoutez les mots sur la cassette. Puis, tout de suite, découpez ces mots en syllabes et choisissez le nombre de cercles qui vous semble correspondre au nombre de syllabes que vous entendez. Attention, vous n'avez que quelques secondes, donc répondez immédiatement. Par exemple : vous entendez le mot [soba]. Alors vous comptez les syllabes ([sob-a] ou [so-ba] (comme vous voulez) → 2 syllabes). Vous choisissez alors le nombre de cercles correspondant ([sob-a] ou [so-ba] → 2 syllabes → OO).

Répondez immédiatement car il n'y a pas beaucoup de temps. Chaque mot est séparé par un « BIP ». **Suivez les lignes sur votre feuille avec votre doigt pour ne pas vous tromper de ligne en répondant.**

	000	0000	00000
2.	000	0000	00000
3.	000	0000	00000
4.	000	0000	00000
5.	000	0000	00000
6.	000	0000	00000
7.	000	0000	00000
8.	000	0000	00000
9.	000	0000	00000
10.	000	0000	00000
11.	000	0000	00000
12.	000	0000	00000
13.	000	0000	00000
14.	000	0000	00000
15.	000	0000	00000
16.	000	0000	00000
17.	000	0000	00000
18.	000	0000	00000
19.	000	0000	00000

### **III.3.2. Tests C2, D et E (en français)**

#### **Pour tous les exercices, il faut IMPERATIVEMENT :**

- NE PAS REFLECHIR (tous les mots présentés sont inconnus de vous) et répondre SPONTANEMENT (il n'y a pas de règle à trouver, ni de bonne ou mauvaise réponse ; il faut juste choisir la réponse de manière naturelle)
- Répondre DANS TOUS LES CAS, même si vous n'êtes pas sûrs,
- Répondre IMMEDIATEMENT, car vous n'avez que quelques secondes entre chaque mot ;
- Répondre SIMPLEMENT d'après CE QUE VOUS ENTENDEZ ET VOYEZ ;
- ENTOURER les réponses que vous choisissez ;
- Lire les documents donnés au fur et à mesure de l'exercice : ne TOURNEZ PAS les pages en avance.
- Bien RESPECTER les consignes ;

#### **Veillez remplir COMPLETEMENT les deux tableaux d'information suivants :**

<u>Sexe :</u>	M / F
<u>Age :</u>	17 / 18 / 19 / 20 / 21 / 22 / 23 / 24 / 25 / + de 25 ans
<u>Langue maternelle :</u>	Japonais / Autre (indiquez) : .....
<u>Année d'étude universitaire :</u>	1 <sup>ère</sup> / 2 <sup>ème</sup> / 3 <sup>ème</sup> / 4 <sup>ème</sup>
<u>Discipline majeure :</u>	Français / Autre (indiquez) : .....

#### Etude du français :

Durée d'étude du français :	0-6 mois / 6 mois-1 an / 1-2 ans / 2-4 ans / + de 4 ans
Avez-vous étudié le français au lycée ?	OUI / NON
Si oui, combien de temps ?	1 an / 2 ans / 3 ans
Vous avez étudié	Surtout l'écrit/ Surtout l'oral / Les deux

	équilibrés
Avez –vous déjà eu des classes de prononciation ?	OUI / NON
Avez –vous déjà eu des classes de dictée ?	OUI / NON
Le français, c'est votre :	1 <sup>ère</sup> langue étrangère / 2 <sup>ème</sup> / 3 <sup>ème</sup>
Si le français n'est pas votre 1 <sup>ère</sup> langue étrangère, quelle est cette 1 <sup>ère</sup> langue ?	Anglais / Chinois / Coréen / Espagnol / Allemand / Autre :.....
Quel est votre niveau dans cette 1 <sup>ère</sup> langue ?	Faible / Moyen / Bon / Très Bon
Etes-vous bilingue ?	OUI / NON
Si oui, la deuxième langue est-elle une langue européenne ? (« American English » inclus)	OUI / NON Langue :.....
L'un de vos parents est-il bilingue ?	OUI / NON
Si oui, la deuxième langue est-elle une langue européenne (« American English » inclus) ?	OUI / NON Langue.....
Avez-vous vécu plus de 6 mois à l'étranger ?	OUI / NON
Si oui, dans un pays :	Francophone / Anglophone / Européen / Autre
Etes-vous gaucher ou droitier ?	Gaucher / Droitier
Étudiez-vous le français en plus des cours de l'université (école, cours privés, etc.) ?	NON / UN PEU / OUI
Entendez-vous/ Parlez-vous beaucoup le français ? (amis, TV, etc.)	NON / PARFOIS / OUI
Entendez-vous/ Parlez-vous beaucoup l'anglais? (amis, TV, classes, etc.)	NON / PARFOIS / OUI

**Exercice C - Consignes :**

Écoutez les mots sur la cassette. Puis, tout de suite, découpez ces mots en syllabes et choisissez le nombre de cercles qui vous semble correspondre au nombre de syllabes que vous entendez. Attention, vous n'avez que quelques secondes, donc répondez immédiatement. Par exemple : vous entendez le mot [soba]. Alors vous comptez les

syllabes ([sob-a] ou [so-ba] (comme vous voulez) → 2 syllabes). Vous choisissez alors le nombre de cercles correspondant ([sob-a] ou [so-ba] → 2 syllabes → OO). Répondez immédiatement car il n'y a pas beaucoup de temps. Chaque mot est séparé par un « BIP ». **Suivez les lignes sur votre feuille avec votre doigt pour ne pas vous tromper de ligne en répondant.**

1.	OOO	OOOO	OOOOO
2.	OOO	OOOO	OOOOO
3.	OOO	OOOO	OOOOO
4.	OOO	OOOO	OOOOO
5.	OOO	OOOO	OOOOO
6.	OOO	OOOO	OOOOO
7.	OOO	OOOO	OOOOO
8.	OOO	OOOO	OOOOO
9.	OOO	OOOO	OOOOO
10.	OOO	OOOO	OOOOO
11.	OOO	OOOO	OOOOO
12.	OOO	OOOO	OOOOO
13.	OOO	OOOO	OOOOO
14.	OOO	OOOO	OOOOO
15.	OOO	OOOO	OOOOO
16.	OOO	OOOO	OOOOO
17.	OOO	OOOO	OOOOO



**Exercice E - Consignes :**

Ecoutez et lisez les mots qui sont présentés. Puis, tout de suite, découpez ces mots en syllabes et choisissez le nombre de cercles qui vous semble correspondre au nombre de syllabes de ce que vous écoutez et lisez. Si exceptionnellement vous pensez remarquer des différences entre ce que vous lisez et écoutez, entourez la case « Différent » mais entourez quand même aussi le nombre de syllabes de ce que vous entendez. Attention, vous n'avez que quelques secondes, donc répondez immédiatement. Par exemple : vous lisez rapidement un mot sur l'écran/ tableau (par exemple « soba ») et sur la cassette vous entendez ce mot prononcé (par exemple [soba]). Alors vous devez compter les syllabes ([sob-a] ou [so-ba] (comme vous voulez) → 2 syllabes), et vous choisissez alors le nombre de cercles correspondant à au nombre de syllabes ( [sob-a] ou [so-ba] (comme vous voulez) → 2 syllabes → OO). Répondez immédiatement car il n'y a pas beaucoup de temps. Chaque mot est séparé par un « BIP »

**Suivez les lignes sur votre feuille avec votre doigt pour ne pas vous tromper de ligne en répondant.**

2.	OOO	OOOO	OOOOO	DIFFERENT
3.	OOO	OOOO	OOOOO	DIFFERENT
4.	OOO	OOOO	OOOOO	DIFFERENT
5.	OOO	OOOO	OOOOO	DIFFERENT
6.	OOO	OOOO	OOOOO	DIFFERENT
7.	OOO	OOOO	OOOOO	DIFFERENT
8.	OOO	OOOO	OOOOO	DIFFERENT
9.	OOO	OOOO	OOOOO	DIFFERENT
10.	OOO	OOOO	OOOOO	DIFFERENT
11.	OOO	OOOO	OOOOO	DIFFERENT
12.	OOO	OOOO	OOOOO	DIFFERENT
13.	OOO	OOOO	OOOOO	DIFFERENT
14.	OOO	OOOO	OOOOO	DIFFERENT
15.	OOO	OOOO	OOOOO	DIFFERENT
16.	OOO	OOOO	OOOOO	DIFFERENT

**Exercice D - Consignes :**

« Lisez les mots au tableau. Puis, tout de suite, découpez ces mots en syllabes et choisissez le nombre de cercles qui vous semble correspondre au nombre de syllabes que vous lisez. Attention, vous n'avez que quelques secondes, donc répondez immédiatement. Par exemple : vous lisez le mot « soba ». Alors vous comptez les syllabes ([sob-a] ou [so-ba] (comme vous voulez) → 2 syllabes). Vous choisissez alors le nombre de cercles correspondant ([sob-a] ou [so-ba] → 2 syllabes → OO). Répondez immédiatement car il n'y a pas beaucoup de temps. Chaque mot est séparé par un « BIP ». **Suivez les lignes sur votre feuille avec votre doigt pour ne pas vous tromper de ligne en répondant ».**

1.	OOO	OOOO	OOOOO
2.	OOO	OOOO	OOOOO
3.	OOO	OOOO	OOOOO
4.	OOO	OOOO	OOOOO
5.	OOO	OOOO	OOOOO
6.	OOO	OOOO	OOOOO
7.	OOO	OOOO	OOOOO
8.	OOO	OOOO	OOOOO
9.	OOO	OOOO	OOOOO
10.	OOO	OOOO	OOOOO
11.	OOO	OOOO	OOOOO
12.	OOO	OOOO	OOOOO
13.	OOO	OOOO	OOOOO
14.	OOO	OOOO	OOOOO