



HAL
open science

Caractéristiques des langues et apprentissage de la lecture en langue première et en français langue seconde : perspective évolutive et comparative entre l'arabe et le portugais

Anne-Sophie Besse

► **To cite this version:**

Anne-Sophie Besse. Caractéristiques des langues et apprentissage de la lecture en langue première et en français langue seconde : perspective évolutive et comparative entre l'arabe et le portugais. Psychologie. Université Rennes 2, 2007. Français. NNT : . tel-00267662

HAL Id: tel-00267662

<https://theses.hal.science/tel-00267662>

Submitted on 28 Mar 2008

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 Rennes 2
UFR Sciences Humaines
CRPCC – EA 1285

**Caractéristiques des langues et apprentissage de la
lecture en langue première
et en français langue seconde :**

perspective évolutive et comparative
entre l'arabe et le portugais

THÈSE DE DOCTORAT

MENTION PSYCHOLOGIE

Présentée et soutenue le 11 décembre 2007

par Anne-Sophie BESSE

sous la direction du Professeur Jean-Émile GOMBERT

JURY :

Pr. Mohamed BERNOUSSI (Université de Nantes)

Pr. Philippe BLANCHET (Université Rennes 2)

Pr. Élisabeth DEMONT (Université Louis Pasteur – Strasbourg)

Pr. Jean-Émile GOMBERT (Université Rennes 2)

Pr. Annie MAGNAN (Université Lyon 2)

Remerciements

Je remercie tout d'abord les membres du jury :

Jean-Emile Gombert, qui a dirigé cette thèse, pour son soutien, sa disponibilité et ses nombreuses qualités humaines. J'ai trouvé lors de nos échanges, la formation, l'orientation et la liberté de réflexion qui m'ont encouragé à travailler avec détermination.

Annie Magnan et Mohamed Bernoussi qui me font l'honneur d'être les rapporteurs de ce travail.

Elisabeth Demont, pour avoir suivi de près et avec rigueur mes recherches depuis le DEA. Ses conseils, ses encouragements, sa confiance, son enthousiasme ont été de précieux stimulants.

Philippe Blanchet dont la présence apporte une approche disciplinaire différente.

Ce travail n'aurait pas été possible sans la collaboration des directeurs et enseignants qui ont accepté de me recevoir et sans la volonté des élèves à participer aux expériences. Je leur adresse à tous mes sincères remerciements.

Mon thème de recherche m'a permis de découvrir des pays et des personnes qui m'ont sur tous les plans enrichies, comment ne pas leur témoigner ma reconnaissance ?

Je remercie tout d'abord Maria Régina Maluf pour son accueil et son aide lors de mon séjour Brésilien et l'équipe du LEDA qui m'a accueilli à la USP.

Je souhaite témoigner toute mon affection à Fraulein et Sara. Nous étions déjà amies lors de leur séjour en France, elles sont maintenant plus que « minhas amigazinhas » !!

Je remercie Nouri Romdhane pour avoir pris le temps de m'accueillir, de répondre à mes sollicitations, de me guider dans mes démarches. Mes remerciements s'adressent également à l'ensemble des membres de l'équipe du laboratoire EDIPSY. Merci à Aida ainsi qu'à toute sa famille (à Fatma pour l'affection qu'elle m'a témoigné) : bonne chance pour les nouveaux projets. Bon courage à Nahed pour la fin de la thèse ! Merci à Fatma, à son mari, à sa famille et un petit clin d'œil : « la psychologie française salut la psychologie tunisienne ».

Merci à toutes les filles qui m'ont aidé à l'élaboration des épreuves, aux passations dans les établissements et au dépouillage des protocoles : Abir, Basma, Emêny, Neziha, Rafika, Shaima, Sirine.

J'en profite pour saluer la famille Ben Youssef dont la gaîté et la générosité m'ont rendu la tâche plus légère.

Je suis également reconnaissante aux enseignants et doctorants de l'Université de Nantes qui m'ont accueilli en tant qu'ATER. J'apprécie beaucoup l'ambiance de travail qui y règne.

Mon amitié se tourne bien évidemment vers l'équipe du laboratoire de psychologie du développement de Rennes : L'ensemble des enseignants chercheurs pour l'intérêt qu'ils m'ont témoigné. Les doctorants qui m'ont accompagnés et aidés : Abdallah, Anaïg, Cécile, Cyrille. Ils ont supporté avec humour mon angoisse, mes coups de gueules et autres défauts, c'est une grande preuve d'amitié ! Merci aux anciens qui ont tracé la route : Manu et Nathalie pour leur apport à tous les niveaux, Stéphane et Zdenka pour leurs conseils judicieux. Les doctorants de social et d'expé.

J'en viens maintenant à remercier les amis, nombreux, qui ont été présents, patients, drôles, réconfortants du début à la fin : La bande de Rennes, Camille, la « pulpeuse » de Caracas, Cédric « Schrukdefuur », Guillaume ou la langue française dans tous ses états, Nadia, qui n'est pas que la copine de galère, Vrinvrine, qui ne vient pas souvent à Rennes ! Marie et Franck dont je suis impatiente de découvrir la Yourte, Mademoiselle Marteau, Morgane et Laurent, Flash, Spin, et tous ceux qui sont partis mais qui finalement ne sont pas bien loin : Arno, Antoine (vive la Belgique bilingue !), Bruno, Erouan, Laurie, Steeve. Un énorme merci à Véro et Jeff qui me connaissent trop pour que « je leur raconte ». A Luc, Lucie et c'est quoi le prénom au fait ?

A ma grande famille qui est à mes côtés depuis toujours et sans qui je n'aurais jamais eu la force de continuer. A mes parents, qui m'ont accordé tout leur temps pour lire, relire, corriger et améliorer la qualité du manuscrit. Que petits (Lancelot, Océane, Mathilda), moyens (Kimi, Fleur, Quentin, Isatis) et grands (Anne-Marie, Luc, Lolo, Marie, Seb, Isa et bien sur Fifi) trouvent dans mon travail de thèse une preuve de mon amour.

A Jeannot. Cette thèse nous a parfois éloigné, souvent réunie. Merci d'avoir été aussi patient, aussi conciliant, aussi présent. Un gros bisou à ses enfants, Lucie et Cécile.

Sommaire

INTRODUCTION THEORIQUE

Avant propos.....	8
Chapitre 1 : L'apprentissage de la lecture : modélisation et traitements mis en œuvre	11
1. Le développement de la reconnaissance des mots écrits.....	11
1.1 Les modèles d'accès au lexique mental par étapes	11
1.1.1 Les stades de Frith (1985, 1986).....	12
1.1.2 Le modèle de la double fondation de Seymour (1997, 1999).....	12
1.2 Les modèles connexionnistes	15
1.2.3 Les modèles à « traitement parallèle distribué ».....	15
1.2.4 Les modèles tentant d'intégrer les traitements morphologiques.....	17
1.3 Le modèle de l'apprentissage implicite et explicite de la lecture	20
2. La compréhension de phrases	24
2.1 Les traitements en jeu dans la compréhension de phrases : quelles relations entre syntaxe et sémantique ?.....	24
2.1.1 Le modèle du garden-path de Frazier (1987).....	25
2.1.2 Le modèle de la guidance lexicale de Tanenhaus <i>et al.</i> (1989ab).....	26
2.1.3 Le modèle à satisfaction de contraintes multiples de Taraban et McClelland (1990).....	28
2.2 Analyse des indices de surface et stratégie d'interprétation : le modèle de compétition de Bates et MacWhinney (1989).....	29
2.3 Y-a-t-il un développement spécifique de la compréhension écrite ? La question du contrôle des traitements syntaxiques.....	32
Chapitre 2 : Lecture et spécificités linguistiques.....	34
1. Variations interlangues et décodage.....	34
1.1 Le degré de transparence des systèmes orthographiques.....	34
1.2 Les variations des caractéristiques morphologiques des systèmes d'écriture.....	38
1.3 Les spécificités de la morphologie non concaténative des langues sémitiques	39
2. Variations interlangues et analyse syntaxique	42
Chapitre 3 : Les caractéristiques des systèmes écrits des langues arabe et portugaise	48
1. La langue arabe	48
1.1 La situation de diglossie des pays arabes	48
1.2 Le système écrit de l'arabe standard	50
1.2.1 Caractéristiques orthographiques	50
1.2.2 Caractéristiques phonologiques.....	52
1.2.3 Caractéristiques morphologiques	53
1.2.4 Quelques caractéristiques syntaxiques.....	56
1.3 Ecriture vocalisée <i>versus</i> non vocalisée	57
2. La langue portugaise	58
2.1 Le système écrit du portugais.....	59
2.1.1 Caractéristiques orthographico-phonologiques.....	59
2.1.2 Caractéristiques morphologiques	60
2.1.3 Quelques caractéristiques syntaxiques.....	62
Chapitre 4 : Dimensions phonologique, morphologique, syntaxique et lecture : perspective interlangue	63
1. Phonologie et lecture	63
1.1 Préliminaire	63
1.2 Les recherches interlangues	65

1.3 Les données en langue arabe.....	67
1.4 Les données en langue portugaise.....	70
2. Morphologie et lecture.....	73
2.1 Préliminaire.....	73
2.2 Les recherches interlangues.....	79
2.3 Les données en langue arabe.....	82
2.4 Les données en langue portugaise.....	86
3. Syntaxe et lecture.....	89
3.1 Préliminaire.....	89
3.2 Les données en langue arabe.....	92
3.3 Les données en langue portugaise.....	94
4. Analyse critique.....	96
5. Synthèse et perspectives.....	99
Chapitre 5 Apprentissage de la lecture en langue seconde.....	100
1. Précisions terminologiques et définitions.....	100
1.1 Du bilinguisme aux bilinguismes.....	100
1.2 De la langue étrangère à la langue seconde.....	103
2. Apprentissage des langues secondes et concepts associés.....	104
2.1.1 L'évolution des points de vue : de l'analyse contrastive à l'interlangue.....	104
2.1.2 Existe-t-il une période critique pour les apprentissages ?.....	106
2.1.3 Acquisition naturelle ou apprentissage formel des langues secondes ?.....	109
3. Lecture, capacités métalinguistiques et bilinguisme.....	113
3.1 Structure et fonctionnement du lexique bilingue.....	113
3.2 Stratégies d'interprétation chez les bilingues.....	118
3.3 Les capacités métalinguistiques des bilingues.....	121
3.3.1 Le bilinguisme facilite le développement métalinguistique.....	123
3.3.2 Le bilinguisme ne facilite pas le développement métalinguistique.....	125
3.3.3 Synthèse et perspectives.....	127
4. Le rôle des acquisitions en langue première lors de l'apprentissage de la lecture en langue seconde.....	130
4.1 Modélisations de la relation entre langue première et seconde.....	130
4.2 Validations empiriques du transfert interlangue.....	132
4.2.1 Les recherches sur des langues à écriture latine.....	132
4.2.2 Les recherches sur des langues utilisant différents alphabets.....	133
4.2.3 Les recherches sur les langues premières non alphabétiques.....	136
4.2.4 Synthèse et perspectives.....	138
4.3 La perspective <i>cross-linguistique</i> anglo-saxonne : les caractéristiques de la langue première et sa proximité avec la langue seconde.....	140
Chapitre 6 : Problématique de la présente recherche.....	145
1. Problématique des études menées sur les langues premières.....	145
1.1 Objectifs généraux.....	145
1.1.1 Objectifs généraux de l'étude 1 sur la langue arabe.....	146
1.1.2 Objectifs généraux de l'étude 2 sur la langue portugaise.....	147
1.1.3 Mise en correspondance des données issues des études 1 et 2.....	148
1.2 Hypothèses.....	149
1.2.1 Hypothèses de l'étude 1 menée sur la langue arabe.....	149
1.2.2 Hypothèses de l'étude 2 menée sur la langue portugaise.....	152
2. Problématique de l'étude menée sur le français langue seconde.....	154
2.1 Synthèse.....	154
2.2 Objectifs et hypothèses de l'étude 3 menée sur la langue française (L2).....	159

PARTIE EMPIRIQUE

Chapitre 7 : Étude 1 - Connaissances graphophonologiques, grapho-morpho-dérivationnelles, grapho-morpho-flexionnelles, grapho-syntaxiques et lecture en arabe standard langue première.....	163
1. Méthode.....	163
1.1 Population.....	163
1.2 Matériel	164
1.2.1 Épreuves de lecture en arabe standard vocalisé	164
1.2.1.1 Épreuve de lecture de mots isolés en une minute.....	164
1.2.1.2 Épreuve de compréhension de phrases écrites	165
1.2.2 Épreuves implicites de détection d'intrus	165
1.2.2.1 Épreuve implicite de détection d'intrus graphophonologique	165
1.2.2.2 Épreuve implicite de détection d'intrus grapho-morpho-dérivationnel	166
1.2.2.3 Épreuve implicite de détection d'intrus grapho-morpho-flexionnel	167
1.2.2.4 Épreuve implicite de détection d'intrus grapho-syntaxique.....	168
1.2.3 Épreuves explicites de détection d'intrus.....	169
1.2.3.1 Épreuve explicite de détection d'intrus graphophonologique.....	169
1.2.3.2 Épreuve explicite de détection d'intrus grapho-morpho-dérivationnel.....	169
1.2.3.3 Épreuve explicite de détection d'intrus grapho-morpho-flexionnel	170
1.2.3.4 Épreuve explicite de détection d'intrus grapho-syntaxique	170
1.2.4 Épreuve d'identification de la construction morphologique de pseudo mots	171
1.3 Procédure.....	172
2. Résultats	175
2.1 Démarche statistique	175
2.2 Analyse des performances obtenues	177
2.2.1 Épreuve de lecture de mots : effet du niveau scolaire.....	177
2.2.2 Épreuve de compréhension de phrases écrites : effet du niveau scolaire.....	177
2.2.3 Épreuves implicites et explicites de détection d'intrus	178
2.2.3.1 Sélection des épreuves en fonction des résultats obtenus	178
2.2.3.2 Épreuves explicites de détection d'intrus grapho-sémantique : effets du niveau scolaire et du domaine linguistique.....	181
2.2.3.3 Épreuve implicite de détection d'intrus graphophonologique : effets du niveau scolaire et de l'unité phonologique	183
2.2.4 Épreuve d'identification de la construction morphologique de pseudo-mots : effet du niveau scolaire, de la nature de la construction morphologique et de la nature des règles morphologiques	186
2.2.5 Synthèse des résultats et récapitulatif des épreuves ou groupes d'items sélectionnés pour les analyses suivantes.....	188
2.3 Analyse du lien entre les scores aux différentes épreuves ou groupes d'items sélectionnés	189
2.3.1 Analyses de corrélation	190
2.3.2 Analyses en composantes principales	192
2.3.3 Synthèse	194
2.4 Recherche des performances déterminant le niveau de lecture.....	195
2.4.1 Analyses sur le niveau de lecture de mots.....	195
2.4.1.1 Analyses de corrélations entre le niveau de lecture de mots et les scores aux différentes épreuves sélectionnées	195
2.4.1.2 Analyses de régressions sur la variable « lecture de mots »	196
2.4.2 Analyses sur le niveau de compréhension de phrases écrites	198

2.4.2.1	Analyse de corrélations entre le niveau de compréhension de phrases écrites et les scores aux différentes épreuves sélectionnées et en lecture de mots.....	198
2.4.2.2	Analyse de régressions sur la variable « compréhension de phrases écrites »	200
2.4.3	Synthèse sur la contribution des différentes variables à l'explication du niveau de lecture de mots et de compréhension de phrases écrites	202
3.	Discussion	204
Chapitre 8 :	Étude 2 - Connaissances graphophonologiques, grapho-morpho-dérivationnelles, grapho-morpho-flexionnelles, grapho-syntaxiques et lecture en portugais langue première.....	209
1.	Méthode.....	209
1.1	Population.....	209
1.2	Matériel	210
1.2.1	Epreuves de lecture en portugais.....	210
1.2.1.1	Epreuve de lecture de mots isolés en une minute.....	210
1.2.1.2	Epreuve de compréhension de phrases écrites	210
1.2.2	Epreuves implicites de détection d'intrus	211
1.2.2.1	Epreuve implicite de détection d'intrus graphophonologique	211
1.2.2.2	Epreuve implicite de détection d'intrus grapho-morpho-dérivationnel	211
1.2.2.3	Epreuve implicite de détection d'intrus grapho-morpho-flexionnel	212
1.2.2.4	Epreuve implicite de détection d'intrus grapho-syntaxique.....	213
1.2.3	Epreuves explicites de détection d'intrus.....	214
1.2.3.1	Epreuve explicite de détection d'intrus graphophonologique.....	214
1.2.3.2	Epreuve explicite de détection d'intrus grapho-morpho-dérivationnel.....	214
1.2.3.3	Epreuve explicite de détection d'intrus grapho-morpho-flexionnel	215
1.2.3.4	Epreuve explicite de détection d'intrus grapho-syntaxique	215
1.2.4	Epreuve d'identification de la construction morphologique de pseudo mots	216
1.3	Procédure.....	216
2.	Résultats	219
2.1	Démarche statistique	219
2.2	Analyse des performances obtenues	221
2.2.1	Epreuve de lecture de mots : effet du niveau scolaire.....	221
2.2.2	Epreuve de compréhension de phrases : effet du niveau scolaire.....	221
2.2.3	Epreuves implicites et explicites de détection d'intrus	222
2.2.3.1	Sélection des épreuves en fonction des résultats obtenus	222
2.2.3.2	Epreuves implicites de détection d'intrus : effets du niveau scolaire et du domaine linguistique	224
2.2.3.3	Epreuves explicites de détection d'intrus grapho-sémantique : effets du niveau scolaire et du domaine grapho-sémantique	226
2.2.3.4	Epreuve implicite et explicite de détection d'intrus grapho-morpho-flexionnel et grapho-syntaxique : effets du niveau scolaire du niveau cognitif et du domaine syntaxique.....	227
2.2.3.5	Epreuve implicite de détection d'intrus graphophonologique : effets du niveau scolaire et de l'unité graphophonologique	229
2.2.4	Epreuve d'identification de la construction morphologique de pseudo-mots : effet du niveau scolaire, de la nature de la construction morphologique et de la nature des règles morphologiques	232
2.2.5	Synthèse des résultats et récapitulatif des épreuves sélectionnées pour les analyses suivantes.....	235
2.3	Analyse du lien entre les scores aux différentes épreuves sélectionnées.....	236
2.3.1	Analyses de corrélation	236
2.3.2	Analyses en composantes principales	239

2.3.3 Synthèse	240
2.4 Recherche des performances déterminant le niveau de lecture.....	242
2.4.1 Analyses sur le niveau de lecture de mots.....	242
2.4.1.1 Analyses de corrélations entre les scores aux différentes épreuves sélectionnées et le niveau de lecture de mots	242
2.4.1.2 Analyses de régressions sur la variable dépendante « lecture de mots »	243
2.4.2 Analyses sur le niveau de compréhension de phrases écrites	245
2.4.2.1 Analyses de corrélations entre les scores aux différentes épreuves sélectionnées, au test de lecture de mots et le niveau de compréhension de phrases écrites	245
2.4.2.2 Analyses de régressions sur la variable « compréhension de phrases écrites »	246
2.4.3 Synthèse sur la contribution des différentes variables à l'explication du niveau de lecture de mots et de compréhension de phrases.....	247
3. Discussion	249
Chapitre 9 : Étude 3 - Perspective évolutive et comparative <i>cross-linguistique</i> de l'apprentissage de la lecture en français langue seconde.....	254
1. Méthode.....	254
1.1 Population.....	254
1.2 Matériel	256
1.2.1 Epreuve de vocabulaire en français.....	256
1.2.2 Epreuves de lecture en français.....	257
1.2.2.1 Epreuve de lecture de mots isolés en une minute.....	257
1.2.2.2 Epreuve de compréhension de phrases écrites	257
1.2.3 Epreuves implicites de détection d'intrus	257
1.2.3.1 Epreuve implicite de détection d'intrus graphophonologique	257
1.2.3.2 Epreuve implicite de détection d'intrus grapho-morpho-dérivationnel	258
1.2.3.3 Epreuve implicite de détection d'intrus grapho-morpho-flexionnel.....	259
1.2.3.4 Epreuve implicite de détection d'intrus grapho-syntaxique.....	259
1.2.4 Epreuves explicites de détection d'intrus.....	260
1.2.4.1 Epreuve explicite de détection d'intrus graphophonologique.....	260
1.2.4.2 Epreuve explicite de détection d'intrus grapho-morpho-dérivationnel.....	261
1.2.4.3 Epreuve explicite de détection d'intrus grapho-morpho-flexionnel	261
1.2.4.4 Epreuve explicite de détection d'intrus grapho-syntaxique	261
1.2.5 Epreuve grapho-morphologique implicite de plausibilité lexicale écrite en français langue seconde	262
1.3 Procédure.....	263
2. Résultats	266
2.1 Démarche statistique	266
2.2 Analyses des performances obtenues	270
2.2.1 Epreuve de lecture de mots : effet de la L1 et du niveau	270
2.2.2 Epreuve de compréhension de phrase écrites : effet de la L1 et du niveau.....	271
2.2.3 Epreuves implicites et explicites de détection d'intrus	272
2.2.3.1 Sélection des épreuves en fonction des résultats obtenus	272
2.2.3.2 Epreuves implicites de détection d'intrus graphophonologique et grapho-morpho-flexionnel : effet de la L1, du niveau et du domaine linguistique	275
2.2.3.3 Epreuves grapho-morpho-flexionnelles implicites & explicites : effet de la langue 1, du niveau et du degré de contrôle.....	277
2.2.3.4 Epreuves explicites de détection d'intrus grapho-sémantique : effet de la L1, du niveau et du domaine linguistique.....	279

2.2.3.5 Epreuve implicite de détection d'intrus graphophonologique : effet de la langue première, du niveau et de l'unité phonologique	282
2.2.4 Epreuve de plausibilité lexicale : comparaison entre arabophones et lusophones selon le niveau	284
2.2.5 Synthèse des résultats et récapitulatif des épreuves sélectionnées pour les analyses suivantes	287
2.3 Analyse du lien entre les scores aux différentes épreuves en FL2 sélectionnées	288
2.3.1 Chez les arabophones	288
2.3.1.1 Analyses de corrélation	289
2.3.1.2 Analyses en composantes principales	290
2.3.2 Chez les lusophones	292
2.3.2.1 Analyses de corrélation	293
2.3.2.2 Analyses en composantes principales	294
2.4 Analyse du lien entre L1 et en FL2 pour chaque type de connaissances et chaque niveau en fonction de la langue première.	296
2.4.1 Chez les arabophones	296
2.4.2 Chez les lusophones	304
2.5 Recherche des performances déterminant le niveau de lecture en FL2	308
2.5.1 Analyses sur le niveau de lecture de mots en FL2	308
2.5.1.1 Chez les arabophones	308
2.5.1.2 Chez les lusophones	313
2.5.2 Analyses sur le niveau de compréhension de phrases écrites en FL2	316
2.5.2.1 Chez les arabophones	316
2.5.2.2 Chez les lusophones	320
3. Discussion	322
Chapitre 10 : Discussion générale – Conclusion.....	329
1. Discussion générale des études sur les langues premières : mise en correspondance des données issues des études 1 et 2	329
2. Discussion générale de l'étude sur le français langue seconde	334
3. Conclusion.....	338
Liste des figures	341
Liste des tableaux	342
Références bibliographiques	347

Avant propos

Dans tous les pays, la maîtrise de la lecture et la connaissance de plusieurs langues écrites constituent aujourd'hui des enjeux de plus en plus importants pour l'avenir scolaire des jeunes élèves.

Les recherches de psychologie cognitive sur l'apprentissage de la lecture, en langue maternelle mais aussi en langue seconde, se sont essentiellement focalisées sur l'anglais et quelques autres langues indo-européennes. Elles se sont également surtout intéressées aux contraintes internes (la variabilité des compétences cognitives, linguistiques et métalinguistiques des apprentis lecteurs) pour comprendre comment s'élaborent les habiletés de lecture. Les caractéristiques linguistiques de chaque langue à l'origine de la diversité linguistique des systèmes d'écriture posent pourtant des contraintes externes importantes lors de l'apprentissage de l'écrit, ces contraintes ont rarement été prises en compte par les chercheurs.

En langue maternelle, la multiplication des études tenant compte de cette diversité et des données dans différentes langues est susceptible de renouveler de façon importante les modèles d'acquisition proposés jusqu'à présent. Par exemple, les recherches comparant les stratégies d'apprentissage de la lecture dans des langues très différentes apportent des informations décisives sur la façon dont s'articulent les aspects universaux et spécifiques aux langues lors de l'acquisition de l'écrit.

Dans le domaine de l'identification des mots écrits, la perspective interlangue est encore trop limitée à des comparaisons entre des langues alphabétiques latines pour lesquelles le développement de la conscience phonologique est primordial. Les langues écrites comme l'arabe, faisant intervenir de façon privilégiée d'autres dimensions, font apparaître de nouvelles questions. Comme le font remarquer Romdhane, Gombert et Belajouza (2003), le système d'écriture employé par les langues sémitiques, renforçant la structure morphologique des mots, permet d'observer un fonctionnement particulier de l'apprenti lecteur.

Le premier objectif de ce travail est d'approfondir cette perspective en effectuant une comparaison interlangue entre l'arabe et le portugais. Les variations que nous souhaitons établir concernent non seulement les procédures de reconnaissance de mots écrits mais aussi l'élaboration de la signification des phrases écrites, la compréhension étant le but ultime de l'acte de lire. Les études menées dans ces deux langues s'attachent ainsi à déterminer dans quelle mesure et avec quelle importance, les connaissances phonologiques, morphologiques et syntaxiques interviennent lors de la lecture.

En langue seconde, la perspective comparative interlangue constitue également un moyen particulièrement adapté pour déterminer précisément l'influence des apprentissages préalables de l'écrit en langue première et par conséquent les variations dans l'élaboration des procédures de lecture en langue seconde, en fonction des langues sources de référence. Bien que la relation entre les compétences en lecture développées en langue première¹ et en langue seconde² est affirmée par plusieurs auteurs (*e.g.* Cummins, 1989, 1991 ; Geva & Siegel, 2000 ; Sparks & Ganschow, 1993ab), les travaux empiriques confirmant cette interdépendance concernent surtout la conscience phonologique. Ainsi, jusqu'à présent, les phénomènes de contact des langues ont rarement été envisagés de façon exhaustive, c'est-à-dire en tenant compte des spécificités et de la complexité des systèmes linguistiques considérés, sur de multiples dimensions.

Aujourd'hui, les études en langue seconde menées dans une perspective *cross-linguistique* (Koda, 2005) ont tendance à se multiplier pour appréhender spécifiquement comment les processus de lecture sont affectés par les différences entre les langues connues ou apprises par l'apprenant. Bien que ces recherches soient essentiellement axées sur la nature des traitements orthographiques et sur la comparaison entre langues idéographiques et alphabétiques, nous sommes convaincue que la distance orthographique, mais aussi les caractéristiques structurales (par exemple les caractéristiques morphologiques et leurs poids) entre certaines langues alphabétiques méritent d'être étudiées.

Le deuxième objectif de cette recherche doctorale est donc d'approfondir les différences dans la façon d'aborder la lecture en français langue seconde entre apprenants de langue première arabe et portugaise. Plus précisément, il s'agit d'analyser la mise en place différentielle des connaissances phonologiques, morphologiques et syntaxiques lors de la reconnaissance des mots et de la compréhension de phrases écrites, au cours des apprentissages.

Ce travail de thèse présente deux grandes parties : une introduction théorique et une partie empirique.

La partie théorique s'articule en six chapitres distincts : le premier chapitre s'attache à présenter les principaux modèles susceptibles de rendre compte de la façon dont s'élabore l'organisation et le fonctionnement du système lecteur et compreneur..

Le deuxième chapitre aborde plus spécifiquement les recherches envisageant l'existence de variations dans la mise en place des traitements au cours de la lecture et de l'élaboration de la signification des phrases, en fonction des spécificités des langues. L'étude de ces variations

¹ L'abréviation langue 1 ou L1 sera souvent employée pour désigner le terme de langue première

² L'abréviation langue 2 ou L2 sera souvent employée pour désigner le terme de langue seconde

interlangues et de leurs conséquences sur l'apprentissage de la lecture ne peut être menée sans décrire précisément les caractéristiques linguistiques des langues. La description des systèmes écrits de l'arabe et du portugais, langues étudiées et comparées dans cette recherche, fait ainsi l'objet du troisième chapitre.

Les recherches consacrées au rôle des connaissances phonologiques, morphologiques et syntaxiques lors de l'apprentissage de la lecture sont successivement développées dans le chapitre 4. Pour chacune des trois dimensions linguistiques analysées, après avoir présenté de façon synthétique l'état de la recherche sur le lien unissant ces connaissances aux habiletés en lecture chez les enfants d'âge scolaire, sont rapportés les résultats issus des recherches interlangues puis des recherches en langue arabe et portugaise. Au terme de ce chapitre, une analyse critique et une reprise synthétique des principaux éléments permettront de déterminer l'orientation de notre étude comparative des procédures de lecture entre l'arabe et le portugais langue première, intégrant ces trois dimensions linguistiques.

Le chapitre 5 aborde l'apprentissage des langues secondes. Outre les différentes conceptions, développées depuis les années 60, sur le contact des langues chez les bilingues et/ou apprenants en langue seconde/étrangère, ce chapitre rend compte d'une part, de certaines thèses sur l'accès au lexique bilingue et les spécificités des traitements interprétatifs dans une deuxième langue et d'autre part des théories et recherches axées spécifiquement sur le rôle des acquisitions en langue première lors de l'apprentissage de la lecture en langue seconde. Dans ce domaine, nous distinguons les travaux portant sur le transfert des compétences de la langue première vers la langue seconde et les recherches se situant davantage dans une perspective *cross-linguistique*, en comparant les compétences ou stratégies d'apprenants maîtrisant l'écrit dans des L1 différentes.

Le dernier chapitre de l'introduction théorique expose les objectifs et hypothèses spécifiques des trois études empiriques réalisées. Les études 1 et 2 concernent les procédures de lecture en langue première (*i.e.* en arabe standard pour l'étude 1 et en portugais pour l'étude 2). La méthodologie puis les résultats obtenus sont respectivement décrits et discutés dans les chapitres 7 et 8 de la partie empirique. L'étude 3, présentée dans le chapitre 9, explore pour sa part les particularités de l'apprentissage de la lecture en français langue seconde par les apprenants arabophones et lusophones.

En conclusion de notre travail doctoral, le chapitre 10 propose une synthèse permettant de mieux comprendre les différences et points communs entre apprenants, lors la lecture en langue première, puis en français langue seconde et d'émettre quelques réflexions pédagogiques.

Chapitre 1 : L'apprentissage de la lecture : modélisation et traitements mis en œuvre

Le domaine de la lecture, de son apprentissage et des activités cognitives qui y sont associées, a été largement étudié depuis une trentaine d'années. La maîtrise de l'écrit et son acquisition est bien évidemment trop complexe pour pouvoir être abordée de manière univoque et monolithique. Le lecteur qu'il soit débutant ou expert, doit mettre en place de multiples opérations de nature générale (*e.g.* l'attention, la mémoire) et linguistique pour traiter l'ensemble des informations écrites et aboutir à l'objectif final de l'acte de lire : la compréhension. Les modèles théoriques ont ainsi distingué deux composantes dans le traitement de l'écrit (pour une revue récente voir Magnan & Colé, 1992) : une composante spécifique qui a trait aux procédures de reconnaissance des mots écrits, impliquant des processus qualifiés de bas niveaux, et une composante de compréhension qui n'est pas spécifique à la langue écrite, impliquant des processus de plus haut niveau.

La première nécessité, dans l'acquisition de l'écrit, consiste donc à développer des compétences à identifier les mots écrits et à automatiser ces procédures de reconnaissance (Colé & Fayol, 2000). Le premier axe de ce chapitre présentera donc, succinctement, quelques modèles ayant décrit l'acquisition de ce décodage. Le deuxième axe donnera un aperçu non exhaustif des interrogations qui ont pu être soulevées pour appréhender les mécanismes de base de la compréhension.

1. Le développement de la reconnaissance des mots écrits

1.1 Les modèles d'accès au lexique mental par étapes

Les conceptions sur la lecture apparues dans les années 70 – 80 ont envisagé les procédures de reconnaissances des mots écrits selon une représentation modulaire d'accès au lexique mental. La description d'une double voie d'accès, directe et indirecte, proposée par Coltheart (1978) chez le lecteur expert a été très largement adoptée. Ainsi, les modèles sur l'apprentissage de la reconnaissance de mots écrits, développés à cette époque et encore aujourd'hui, se sont attachés à rendre compte des étapes par lesquelles passe l'apprenti lecteur pour installer ces deux voies d'accès.

1.1.1 Les stades de Frith (1985, 1986)

Le modèle le plus connu est sans doute celui de Frith (1985, 1986) qui décrit trois étapes successives du développement de la maîtrise de l'écrit³.

Le premier stade est dit logographique : avant ou au début de l'apprentissage formel de la lecture, l'enfant reconnaît ou « devine » les mots comme des objets picturaux et non linguistiques appris par cœur par stratégie visuelle ou globale. Cette procédure consiste à mémoriser le patron ou la forme visuelle des mots écrits à partir d'indices saillants, souvent extralinguistiques (*e.g.* la couleur, le contour, la présence de formes particulières de lettres).

Le stade alphabétique, spécifique aux systèmes alphabétiques, se caractérise par la médiation phonologique permettant les relations entre l'écrit et l'oral grâce à des traitements de nature linguistique. Cette procédure repose sur la mise en correspondance d'unités visuelles discrètes (les graphèmes) et d'unités phonémiques abstraites (les phonèmes). Les règles de correspondance étant conventionnelles et arbitraires, ce « décodage séquentiel » nécessite un enseignement explicite du code et correspond à l'acquisition de la voie indirecte ou d'assemblage du modèle de Coltheart (1978). La répétition de ce déchiffrage permettra la construction d'un lexique mental orthographique et la mise en place d'une procédure plus efficace (notamment pour la lecture de mots irréguliers) et plus économique à la phase suivante.

Le stade orthographique correspond à la voie directe d'accès lexical, par adressage. Le traitement du mot écrit s'effectue automatiquement à partir d'unités orthographiques sans recours à la conversion ou médiation phonologique. Les mots écrits déjà rencontrés sont mis directement en relation avec leur représentation lexicale mais contrairement à la procédure logographique ce traitement est de nature linguistique et analytique.

La stricte séquentialité de ces stades de même que l'utilité, voire l'existence, d'une étape initiale logographique ont été remises en cause (Content & Leybaert, 1992; Sprenger-Charolles, 1992; Sprenger-Charolles & Casalis, 1996). D'autres modèles, plus interactifs, et permettant d'envisager la coexistence de différentes procédures ont ainsi été élaborés.

1.1.2 Le modèle de la double fondation de Seymour (1997, 1999)

Le modèle de développement orthographique à double fondation de Seymour (1997, 1999) suggère que processus logographiques et processus alphabétiques participent conjointement à l'élaboration des représentations orthographiques dont dépend l'acquisition de la lecture et de l'orthographe (*cf.* figure 1). Il faut toutefois noter que le processeur logographique envisagé par

³ Ce modèle a également été utilisé pour décrire l'acquisition de l'orthographe.

Seymour fait référence aux processus visuo-orthographiques permettant un stockage et une représentation directe des mots entiers sous une forme plus ou moins complète. Le processeur alphabétique repose sur la « traduction graphèmes-phonèmes » consécutive à un apprentissage formel. Le processeur alphabétique agit en interaction avec la conscience linguistique responsable de l'émergence d'une conscience des phonèmes et intervenant également dans le développement d'autres formes de segmentation. Ces deux processeurs sont fondateurs, car ils sont indispensables à la formation d'une structure orthographique centrale à l'origine d'une représentation abstraite de l'orthographe. Toutefois, le système orthographique ne code que des mots simples, mono- ou bisyllabiques. Les mots plus complexes, comme les mots morphologiquement composés, sont traités par une composante supplémentaire : la structure morphographique, dépendante de la conscience linguistique et tributaire de l'achèvement des représentations orthographiques.

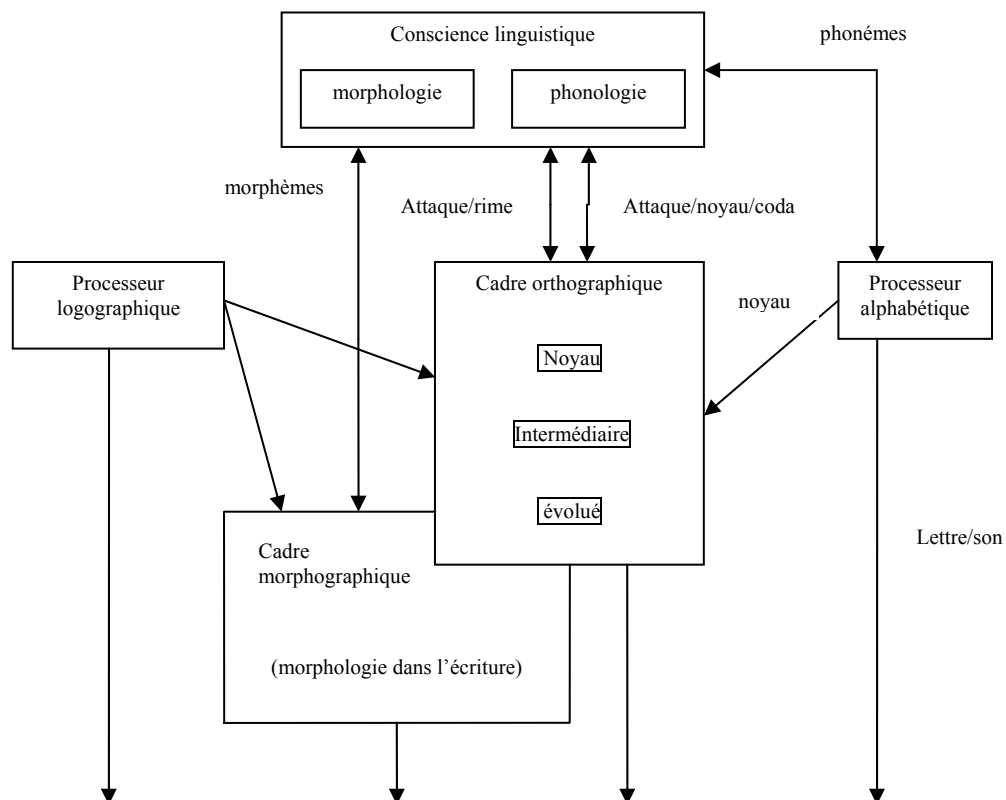


Figure 1: Le modèle à double fondation de Seymour (1997)

Bien que l'ordre de développement des fondations ne soit pas rigide, l'acquisition de la structure orthographique et donc de la lecture procède sinon par stades, au moins par étapes (cf. figure 2). L'étape initiale correspond à la mise en place des processus logographiques et alphabétiques. Le développement de la structure orthographique pourra alors s'effectuer sur la base de cette double fondation. Cette acquisition passe par une étape centrale, une étape intermédiaire et une

étape avancée, introduisant au fur et à mesure des structures de segmentation phonémique de plus en plus complexes. Le système lecteur opérant sera alors capable de lire les mots réels et les non-mots selon une forme de traitement direct.

La construction de la structure morphographique ne s'effectue qu'en dernier lieu et consécutivement au développement de la structure orthographique.

Dans ce modèle, l'acquisition de la lecture est parallèle au développement de la conscience linguistique qui suppose pour sa part une progression allant des petites unités telles que le phonème aux unités plus larges, telles que la rime, la syllabe et le morphème.

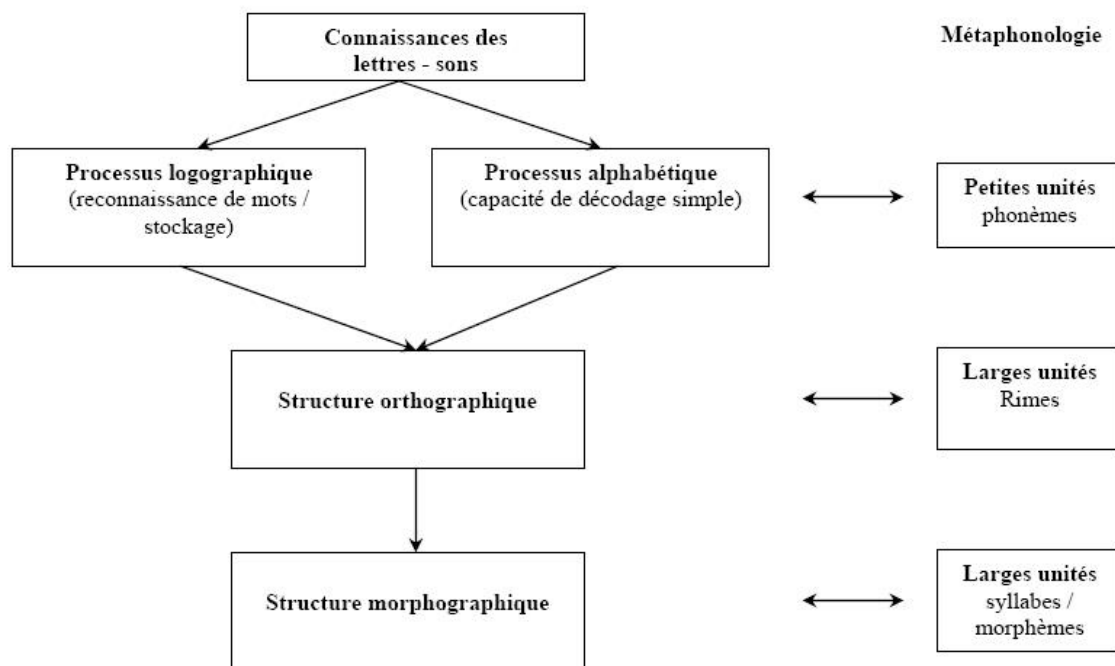


Figure 2: Représentation du modèle de Seymour intégrant le développement de la conscience linguistique (Seymour, Aro, & Erskine, 2003).

Que ce soit dans le modèle de Frith ou dans celui de Seymour, les traitements d'analyse segmentale les plus déterminants sont de nature phonologique. En effet, ces auteurs n'envisagent, implicitement ou explicitement, qu'une intervention tardive et secondaire des traitements morphologiques. De plus, ils privilégient l'installation de connaissances contrôlées, élaborées par le biais d'un enseignement explicite. Enfin, ces modélisations font référence à une architecture modulaire, plus ou moins encapsulée, du système cognitif.

Avec l'essor des neurosciences et de l'intelligence artificielle, et de façon radicalement opposée aux modèles symboliques, les conceptions connexionnistes se sont développées depuis une quinzaine d'année.

1.2 Les modèles connexionnistes

Dans les modèles connexionnistes, il n'y a plus à proprement parler de lexique mental. Les connaissances ou représentations se « matérialisent » par le poids des connexions, des interrelations entre les unités du système cognitif (Seidenberg & Gonnerman, 2000). L'activation des connexions entre les unités (et l'inhibition des autres connexions) forme ainsi des patterns ou des configurations en réseaux distribués variables, le poids des liaisons entre unités venant moduler le flux d'activation du réseau (Seidenberg, 2005). Pour éprouver ces modélisations, les auteurs proposent des simulations par ordinateur des mécanismes computationnels du système lecteur.

1.2.3 Les modèles à « traitement parallèle distribué »

Plaut, McClelland, Seidenberg, et Patterson (1996) et Seidenberg et McClelland (1989), ont décrit un modèle de traitements parallèles distribués de la reconnaissance visuelle des mots. Ce modèle propose une architecture minimale dans laquelle l'apprentissage⁴ joue un rôle majeur (*cf.* figure 3). Cette structure est constituée d'un ensemble d'unités représentationnelles primaires de nature orthographique, phonologique et sémantique, distribuées au travers du système et qui s'influencent les unes les autres. La construction de ces représentations serait également influencée par des facteurs contextuels de façon directe (sur le niveau sémantique) ou indirecte (sur les autres niveaux). Les unités orthographiques, activables dès la perception des lettres, envoient des signaux activateurs aux unités sémantiques et phonologiques par l'intermédiaire d'unités cachées qui calculent l'activation à propager. Réciproquement, les unités sémantiques et phonologiques renvoient un *feed-back* renseignant sur le niveau de compatibilité en modulant le poids des connexions.

⁴ Dans ce modèle, un algorithme de rétropropagation de l'erreur est utilisé comme méthode d'apprentissage. Il consiste à calculer pour chaque unité, l'écart entre l'activation réelle et l'activation correspondant à la forme correcte. Le processus de rétropropagation vient modifier le poids des connexions afin de minimiser cet écart.

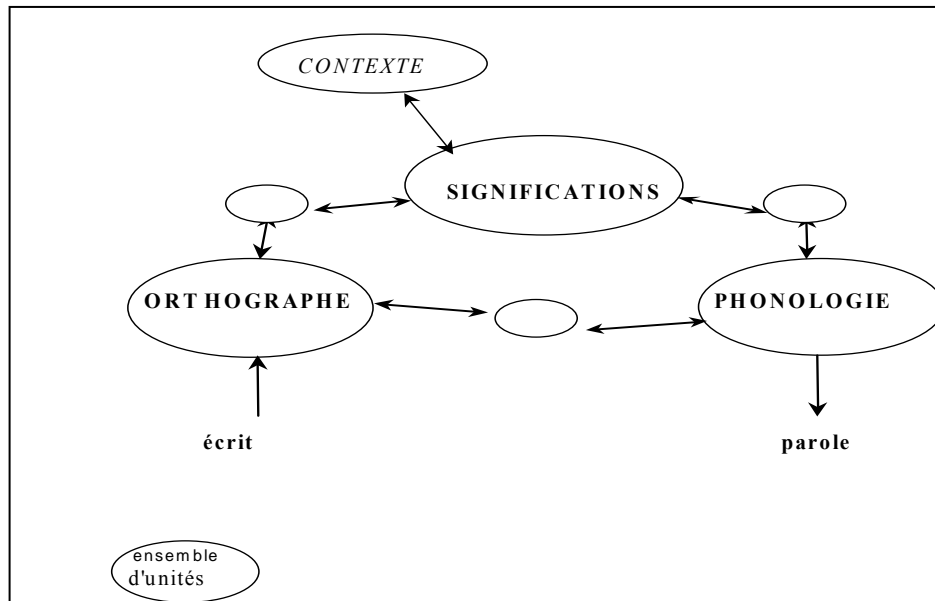


Figure 3 : Le modèle connexionniste de Seidenberg et McClelland (1989) proposé par Gombert (2003b)

Cette configuration composite d'unités activées correspond à la reconnaissance d'un mot. Ainsi, les auteurs ne postulent qu'une seule procédure de lecture puisque la reconnaissance repose sur les interactions entre les différentes informations issues du mot écrit. Seul le poids des connexions entre les unités est variable expliquant ainsi les différences entre la lecture des mots fréquents et rares ou réguliers et irréguliers. La reconnaissance des mots familiers est plus rapide du fait de configurations de connexions plus stables car activées plusieurs fois lors des lectures antérieures. Pour les mots réguliers, la similarité des connexions entre certaines unités orthographiques et phonologiques activées explique la diminution des temps de reconnaissance par rapport aux mots irréguliers.

Le modèle implémenté par Seidenberg et McClelland (1989) ne comprend que des unités phonologiques et orthographiques ainsi que des unités cachées les reliant et ne décrit que le traitement de mots monosyllabiques. Les résultats des simulations font apparaître de bonnes performances en lecture de mots, comparables à celles du lecteur humain : la rapidité et l'exactitude des prononciations du système sont sensibles aux effets de fréquence et de régularité et les erreurs produites sont de même type que celles du lecteur humain. Cependant, l'implémentation échoue en lecture de non-mots, ce qui restreint sa possibilité de supplanter les modèles à double voie en ne parvenant finalement qu'à simuler la voie d'accès direct aux mots écrits.

Plaut *et al.* (1996) parviennent à améliorer les performances du système en implémentant non pas des triplets, mais des unités de type infrasyllabique (attaque, voyelle et coda), pour éviter que les régularités entre les configurations de connexions ne soient trop dispersées. Les résultats sont

ainsi améliorés puisque le système parvient à une prononciation performante des non-mots. Les auteurs en concluent que de tels modèles sont en mesure d'appréhender les mécanismes de lecture sans faire appel à l'existence d'une structure à deux voies d'accès ou nécessitant plusieurs mécanismes différenciables.

La limite de ces travaux est que les simulations implémentent des unités orthographiques et phonologiques mais ne parviennent pas à encoder le niveau sémantique. Ainsi, comme dans les modèles symboliques d'accès au lexique, les traitements grapho-morphologiques sont négligés, bien que beaucoup d'auteurs connexionnistes reconnaissent à la morphologie des caractéristiques que leurs modèles sont en mesure de prendre en compte⁵ (Gonnerman, Seidenberg, & Andersen, 2007; Rueckl, Mikolinski, Raveh, Miner, & Mars, 1997).

1.2.4 Les modèles tentant d'intégrer les traitements morphologiques

Le premier modèle connexionniste qui a cherché à rendre compte des traitements portant sur les mots morphologiquement complexes est sans doute celui de Rumelhart et McClelland (1986). Ces chercheurs se sont attachés à rendre compte de l'acquisition des formes fléchies du passé des verbes en anglais et en particulier des modifications de performance lors des stades de cette acquisition (par exemple, la disparition des phénomènes de régularisations : *e.g. come* → *comed*). Leur modèle n'est en fait pas un modèle d'apprentissage de la lecture puisque seul le niveau phonologique des représentations est inclus dans le système. L'objectif est de faire apprendre au système la prononciation de formes verbales fléchies régulières et irrégulières à partir de leur racine. Dans ce modèle, les aspects flexionnels requièrent une correspondance entre les unités encodant la représentation phonologique de la racine et les unités décodant la représentation phonologique du passé des verbes. Cependant, compte tenu des irrégularités des verbes anglais, ces correspondances peuvent varier. En effet si la formation du passé d'une majorité de verbes anglais est parfaitement régulière (*e.g. need* → *needed*), celle de certains verbes entraîne une modification de la racine (*e.g. go* → *went*, *come* → *came*) et d'autres une modification de la flexion (*e.g. burn* → *burnt*, *look* → *looked*). Concrètement, les chercheurs ont simulé la prononciation de formes passées régulières ou irrégulières, nécessitant une modification phonologique de la racine ou une irrégularité dans la prononciation de la flexion et obtiennent de très bonnes performances.

⁵ Comme le code orthographico-phonologique, il s'agit d'un système quasi régulier faisant apparaître des règles strictes et des exceptions.

D'autres recherches se sont pour leur part centrées sur la morphologie dérivationnelle. Seidenberg et Gonnerman (2000) et Plaut et Gonnerman (2000) considèrent la morphologie comme le produit de l'effet conjugué du partage d'éléments sémantiques et formels. Comme dans le modèle de Rumelhart et McClelland (1986), les morphèmes ne sont pas présents en tant qu'unités dans le système. Ainsi, la structure morphologique est gouvernée par les mêmes principes de distribution que les autres codes du lexique (*cf.* figure 4).

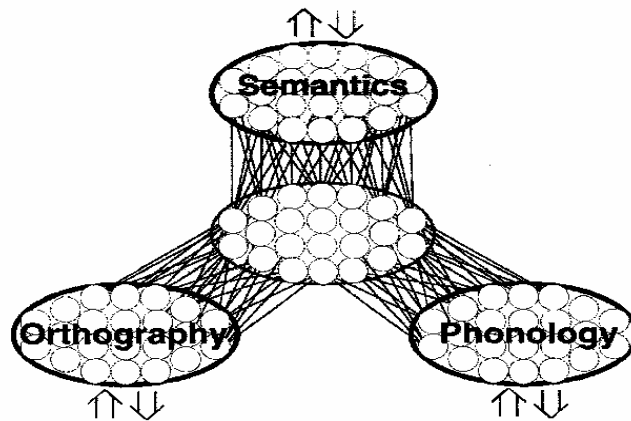


Figure 4: Représentation du modèle connexionniste de Plaut et Gonnerman (2000) intégrant la morphologie

L'idée de base est que la morphologie relève de la convergence entre les codes orthographique, phonologique et sémantique. Les unités cachées⁶ ou intermédiaires médiatisent le poids d'activation ou d'inhibition entre les trois autres groupes d'unités. Les représentations morphologiques distribuées varient donc en fonction du degré de recouvrement ou de similarité entre les patterns formels (phonologiques et orthographiques) et sémantiques des mots de la même famille. Par exemple en anglais, les mots *drink* et *drinker* entretiennent une relation sémantique plus importante que *liver* et *deliver* et une relation phonologique plus importante qu'entre *divine* et *divinity*.

Plaut et Gonnerman (2000) effectuent des simulations en intégrant des unités sémantiques et orthographiques⁷ et des variations en termes de richesse morphologique des langues⁸ qui corroborent leur modèle.

Ces recherches proposent ainsi une alternative par rapport aux autres modèles du lecteur expert défendant l'existence de représentations morphologiques discrètes, de processus de

⁶ Dans ce modèle, il n'y a pas trois mais un groupe d'unités cachées.

⁷ Les unités phonologiques n'ont pas été prises en compte car les auteurs considèrent qu'elles feraient apparaître une redondance avec la dimension orthographique.

⁸ La richesse morphologique des langues intervient pour rendre compte des phénomènes morphologiques de l'hébreu. Cette dimension est opérationnalisée par la fréquence des structures de mots implémentées de façon à simuler les résultats obtenus en anglais et en hébreu.

décomposition morphologique et donc d'une spécificité des traitements morphologiques (*cf.* encart 1).

Encart 1 : Les modèles et recherches sur la décomposition morphologique lors de la lecture experte

Depuis quelques années, plusieurs chercheurs se sont intéressés au rôle de la morphologie lors de l'accès au lexique chez le lecteur expert. S'ils s'accordent généralement sur l'existence d'une analyse morphologique, ils se distinguent quant à la façon d'entrevoir la décomposition morphologique lors de la reconnaissance de mots. Les hypothèses dans ce domaine avancent :

- soit une décomposition prélexicale obligatoire de tous les mots (Taft & Forster, 1975),
- soit une décomposition prélexicale parallèle à un traitement global : les procédures de décomposition et/ou le stockage des informations morphologiques ne peuvent concerner que certains mots : les mots sémantiquement opaques (Schreuder & Baayen, 1995), les mots peu fréquents ou inconnus (Caramazza, Laudanna, & Romani, 1988), les mots préfixés mais pas suffixés (Colé, Beauvillain, & Segui, 1989) etc.,
- soit encore une décomposition postlexicale (Giraud & Grainger, 2001).

Les recherches réalisées dans diverses langues sur les aspects morphologiques ont utilisé, le plus souvent, la technique d'amorçage ou d'enregistrement des mouvements oculaires (Frost & Grainger, 2000). Si toutes montrent l'importance des traitements morphologiques dans l'organisation du lexique, elles apportent plus spécifiquement des données en faveur de telle ou telle position, en soulignant l'influence de plusieurs facteurs sur les procédures de reconnaissance des mots morphologiquement complexes. Par exemple, l'effet de transparence orthographique⁹ et sémantique¹⁰ (pour une revue, voir Stolz & Feldman, 1995), l'effet de fréquence lexicale de surface et de fréquence cumulée de la racine (voir par exemple Giraud & Grainger, 2000), l'effet de la taille de la famille morphologique (Schreuder & Baayen, 1997) ou encore celui de l'ambiguïté et de la productivité¹¹ des affixes (Niswander, Pollatsek, & Rayner, 2000).

Malgré les débats et critiques qui ont encore lieu, les auteurs considèrent que leur modèle connexionniste est en mesure d'expliquer les effets de transparence phonologique, orthographique et sémantique sur la reconnaissance de mots plurimorphémiques, obtenus dans la littérature à l'aide de la technique d'amorçage rapide (pour une revue, voir Feldman & Soltano, 1999; Marslen-Wilson, Tyler, Waksler, & Older, 1994; Stolz & Feldman, 1995). Selon eux, le caractère robuste de l'amorçage morphologique par rapport à l'amorçage orthographique¹² ou sémantique¹³ n'est pas un argument suffisamment fiable pour réfuter leur conception.

⁹ Un mot dérivé est dit orthographiquement peu transparent lorsque le processus de dérivation entraîne une modification de la base (*e.g.* redig(er) → rédaction).

¹⁰ Un mot dérivé est dit sémantiquement peu transparent lorsque la relation sémantique avec la base est opaque (*e.g.* toile → toilette).

¹¹ Plus un affixe permet la formation de nombreux mots en se combinant avec de nombreuses bases, plus il est productif.

¹² Plusieurs études ont d'ailleurs observé une absence d'effet facilitateur voire un effet inhibiteur des amorces orthographiques (Drews & Zwitserlood, 1995 en allemand et en hollandais; Feldman & Fowler, 1987 en serbe; Grainger, Colé, & Segui, 1991 en français; Murrell & Morton, 1974; Rastle, Davis, Marslen-Wilson, & Tyler, 2000; Stolz & Feldman, 1995 en anglais).

¹³ L'effet d'amorçage sémantique n'apparaît en fait plus dans certaines conditions, par exemple lorsque les SOA - Stimulus Onset Asynchronies- sont courts (42 ou 72 ms) (Rastle *et al.*, 2000) ou lorsque les sujets doivent détecter une lettre dans l'amorce (Stolz & Besner, 1998). Les recherches menées en hébreu observent encore plus clairement une absence de cet effet sémantique facilitateur (*e.g.* Frost, Deutsch, Gilboa, Tannenbaum, & Marslen-Wilson, 2000b; Frost, Forster, & Deutsch, 1997).

Cependant, et bien que les auteurs ne le mentionnent pas, il semble que les résultats obtenus par Rastle *et al.* (2000) remettent en question le point de vue adopté par les connexionnistes. Ces auteurs montrent en effet que le partage d'éléments morphologiques produit davantage d'effets de facilitation que le partage conjoint de propriétés formelles et sémantiques, ce qui signifierait que le traitement du code morphologique n'est pas le simple produit de relations formelles et sémantiques entre les mots.

D'autres pierres d'achoppement peuvent être énoncées à l'encontre des modèles connexionnistes dans leur ensemble. Les modèles de simulation, forcément simplifiés, ne font jamais intervenir conjointement les dimensions orthographiques, phonologiques et sémantiques, si bien qu'ils ne permettent pas d'identifier à quels types de similarités (orthographico-phonologiques ou orthographico-sémantiques) le système est le plus sensible.

Par ailleurs, bien que les modèles connexionnistes soient aptes à rendre compte de l'apprentissage de la lecture, leur modélisation correspond finalement à un système « qui sait déjà lire » puisqu'il existe dès le départ des unités spécialisées dans le traitement orthographique (Gombert, 2003b).

Enfin, ils ne traduisent qu'un fonctionnement automatique et ne tiennent pas compte des possibilités du lecteur à porter son attention sur telle ou telle dimension de l'écrit. Autrement dit, ces modèles se limitent à l'apprentissage implicite de la lecture, au contraire des modèles symboliques qui eux s'intéressent principalement à l'installation de connaissances explicites dépendantes des enseignements.

En adoptant une perspective proche de la conception connexionniste, Gombert et ses collaborateurs ont décrit un modèle d'apprentissage de la lecture susceptible de dépasser ces écueils.

1.3 Le modèle de l'apprentissage implicite et explicite de la lecture

L'objectif de Gombert (Gombert, 2002a, 2003ab; Gombert, Bryant, & Warrick, 1997), est de développer un modèle prenant en compte le lien entre l'apprentissage de la lecture et les acquisitions linguistiques orales préalables, articulant les apprentissages implicites et explicites de la lecture et intégrant différentes dimensions des systèmes alphabétiques notamment les codes graphophonologique et grapho-morphologique.

Tout d'abord, Gombert et ses collaborateurs (Demont & Gombert, 2004; Gombert, 2002a, 2003ab; Gombert *et al.*, 1997) cherchent à déterminer la base de connaissances dont dispose l'apprenti lecteur lors de ses premiers contacts avec l'écrit. Le système cognitif de l'enfant est

considéré comme un système d'activation de concepts (*cf.* figure 5). Il comporte initialement quatre types de représentations : les représentations picturales, phonologiques, sémantiques et contextuelles. Il est ainsi capable de traiter d'une part le langage oral, ce qui lui servira de base pour élaborer des traitements spécifiques à l'écrit et d'autre part des stimuli visuels pouvant intervenir dans la perception d'un message linguistique. Par exemple, l'activation des unités sémantiques pertinentes pour le concept « chat » peut s'effectuer, soit par l'audition du mot [Σα], l'entrée se faisant via les unités phonologiques, soit par la vision ou la représentation figurative de l'animal via un processeur imagé ou pictural. Ainsi, le système est au départ, dépourvu de représentations spécifiques aux informations linguistiques écrites. Cependant, lors de son contact avec l'écrit, il parviendrait, grâce à un renforcement des relations entre les différentes dimensions du code, à effectuer une spécialisation orthographique des unités picturales. Pour expliquer ce phénomène, les auteurs avancent que le système cognitif perçoit les nombreuses régularités internes et contextuelles présentes dans le système écrit, comme dans le système oral (régularités orthographiques des configurations visuelles, régularités alphabético-phonologiques et phono ou grapho-morphologiques). La fréquentation répétée des mêmes configurations visuelles permet donc un premier type de traitement analogique, inconscient et spontané, mobilisant des connaissances implicites. Cette sensibilité ou ces connaissances implicites interviendraient lors de l'apprentissage de la lecture. Elles peuvent être orthographiques (Pacton & Fayol, 2000; Pacton, Fayol, & Perruchet, 2005) phonologiques (Demont & Gombert, 2004; Gombert, *et al.*, 1997) mais aussi morphologiques (Colé, Marec-Breton, Royer, & Gombert, 2003; Marec-Breton, Gombert, & Colé, 2005). Elles sont dites implicites car elles sont intégrées sans intention, par le seul contact prolongé avec le langage et sans recours à un apprentissage formel qui passerait par une explicitation des règles ou phénomènes linguistiques.

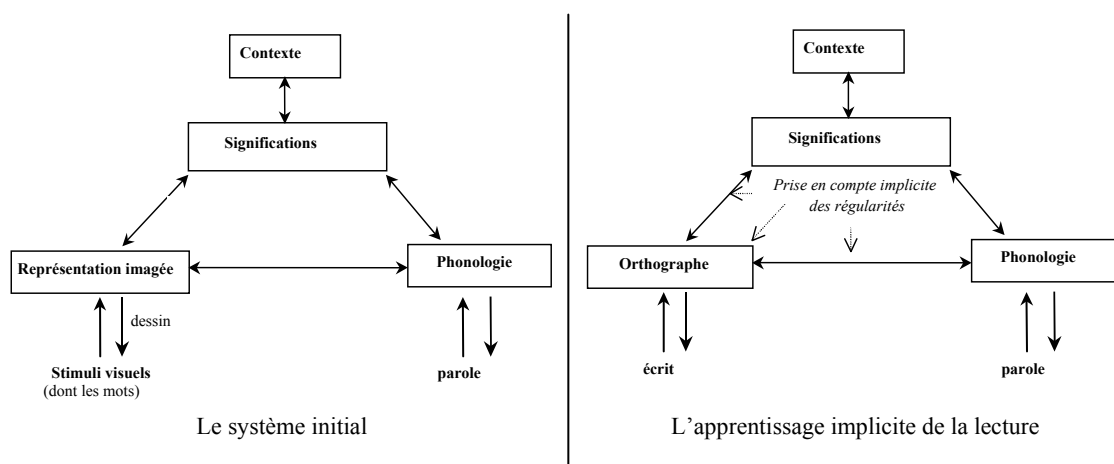


Figure 5: Le système initial et l'apprentissage implicite de la lecture (Gombert, 2003b)

Pour autant, l'enfant n'apprend pas à lire seul, le traitement du langage écrit exigeant un niveau d'abstraction, d'élaboration et de contrôle que le niveau implicite ne permet pas d'installer. La complexité du système écrit ainsi que les nombreuses irrégularités qui y apparaissent rendent en effet nécessaire un apprentissage explicite du code orthographique, de ses correspondances phonologiques et de la morphologie. Néanmoins, c'est sur la base de ces apprentissages implicites que viendront se greffer les apprentissages explicites (Gombert, 1990)¹⁴. Grâce à ces derniers, l'apprenti lecteur pourra décoder tout item écrit, même inconnu ou inexistant (*i.e.* des mots nouveaux et des pseudo-mots) et maîtriser l'orthographe des mots comprenant de nombreuses irrégularités (par exemple en identifiant la présence de lettres muettes). Les connaissances explicites, relevant de processus analytiques sont sous l'influence de contingences externes (enseignement scolaire, complexification des offres linguistiques de l'environnement) stimulant un effort cognitif pour apprendre, effort qui requiert donc un niveau élevé d'attention. Elles permettent également au lecteur de contrôler intentionnellement son activité, et ainsi de moduler les traitements automatiques, relevant des connaissances implicites (*cf.* figure 6). En effet, une des conséquences de l'apprentissage explicite est de répéter l'activité de lecture et de rendre de plus en plus fréquente l'attention portée aux caractéristiques des mots écrits, ce qui en retour va décupler les apprentissages implicites.

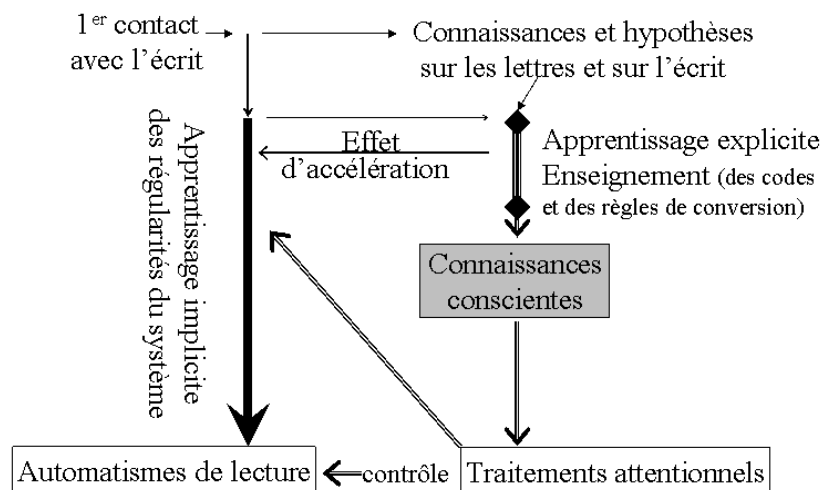


Figure 6: Apprentissage implicite et explicite de la lecture (Gombert, 2003b)

¹⁴ Le modèle précise les phases successives permettant le développement des connaissances implicites puis explicites. Dans la phase 1 se développeraient des habiletés linguistiques primaires sur la base de préprogrammations innées. Dans la phase 2, les connaissances accumulées dans la phase 1 se réorganiseraient pour former des connaissances implicites, encore inaccessibles à la conscience. Dans la phase 3, l'accès à la conscience s'effectuerait grâce à l'intervention de pressions environnementales déclenchant le développement de la conscience métalinguistique.

Ce modèle cherche également à expliquer la mise en place du caractère automatique de la lecture, indispensable pour atteindre une expertise optimale. Dans ce modèle, l'automatisme des traitements est issu des apprentissages implicites et ne relèverait pas d'une transformation des traitements contrôlés qui à force de répétition finirait par s'effectuer sans contrôle¹⁵. Cependant, les auteurs reconnaissent le rôle des traitements explicites dans l'expertise (puisqu'ils permettent la répétition de l'activité chez le lecteur débutant et renforcent les traitements implicites). Ces deux types de traitement implicites et explicites auraient ainsi la possibilité d'intervenir en parallèle.

Si ces deux types de connaissances implicites et explicites sont fortement différenciés au niveau théorique, leur mise en évidence empirique est plus délicate (*cf.* encart 2).

Encart 2 : L'évaluation des connaissances implicites et explicites

Les paradigmes destinés à évaluer les connaissances explicites doivent être contraignants en réflexion et en attention, contrairement à ceux utilisés pour évaluer les connaissances implicites qui doivent engager un coût cognitif minimum ou nul. Par exemple, les tâches de réplication d'anomalies surtout utilisées dans le domaine syntaxique et morphosyntaxique (*e.g.* appliquer la même erreur que « le voleuse emporte les bijoux » dans « le fugitif attaque un passant ») ainsi que les épreuves de suppression de phonème (ou de syllabe) ou de production de pseudo-mots par la fusion de certaines syllabes de deux mots cibles semblent attentionnellement coûteuses (Gaux & Gombert, 1999ab; Lecocq, 1991). Les épreuves de comptage (*e.g.* combien de phonèmes ou de syllabes sont entendus dans un mot ?) ou de segmentation (*e.g.* quels sont les phonèmes ou « autres mots » (morphèmes) entendus dans un mot ?) sont également considérées comme des tâches métalinguistiques, même si la charge cognitive qu'elle implique est plus faible.

Les tâches d'identification ou de jugement de similitude (*e.g.* parmi plusieurs mots proposés dont l'un ne partage pas le même phonème, la même rime, la même syllabe, ou n'est pas de la même famille, lequel ne va pas avec les autres ?) seraient pour leur part implicites. Cette affirmation est toutefois discutable.

D'autres exemples rendent compte de la difficulté à classer les épreuves selon le degré de contrôle qu'elles requièrent. L'étude de Gaux et Gombert (1999a) a montré, dans le domaine syntaxique, que le caractère délibéré des tâches de correction d'agrammaticalité dépend beaucoup de la saillance de l'anomalie : plus l'anomalie est saillante (*e.g.* inversion de l'ordre des mots dans la phrase), plus la correction peut être émise sans réflexion particulière, sur la simple base d'une impression de dissonance. Il semble donc que le caractère implicite ou explicite de l'épreuve est souvent déterminé par le domaine ou le type d'unité étudié.

Ces problèmes seront plus longuement abordés dans la partie « analyse critique » du chapitre 4.

Pour conclure, malgré ces difficultés méthodologiques, le modèle de Gombert a l'avantage d'aborder les mécanismes de lecture de mots dans leur ensemble et dans leur complexité. Il souligne l'enjeu, pour le lecteur, de parvenir à effectuer ces traitements de bas niveau de façon rapide et économique, pour pouvoir porter son attention sur les aspects de plus haut niveau, conduisant à la compréhension (Gombert, 1990, 2002b).

¹⁵ Ces connaissances explicites pourraient toutefois s'activer de plus en plus facilement mais jamais à coût attentionnel nul.

2. La compréhension de phrases

Comprendre ce qu'on lit est bien évidemment la finalité de l'activité de lecture, bien que la compréhension ne soit pas une caractéristique propre à l'écrit.

Comprendre un énoncé, oral ou écrit, passe par l'élaboration de la signification d'un ensemble structuré d'informations (micro-structurelles et macro-structurelles) permettant la construction progressive d'un modèle cohérent de la situation décrite (Fayol, 1992). Pour reprendre les propos de Golder et Gaonac'h (1998) concernant la lecture, la compréhension est une interprétation, car comprendre ce n'est pas extraire la signification d'un texte lu mais la reconstruire activement. Ainsi, la dimension de compréhension ne se limite bien sûr pas à la phrase écrite et fait interagir des mécanismes complexes de traitements des données linguistiques et extralinguistiques mais aussi de recouvrement de connaissances conceptuelles, de réalisation d'inférences, de sélection des informations pertinentes, de stockage en mémoire, etc. Nous ne nous intéresserons pour notre part qu'aux traitements micro-structurels mis en jeu dans la compréhension de phrases simples (et en particulier au rôle des informations syntaxiques et à leur relation avec les aspects sémantiques). Ces traitements constituent en effet une étape importante pour atteindre un niveau plus élevé d'élaboration de la signification des informations textuelles.

2.1 Les traitements en jeu dans la compréhension de phrases : quelles relations entre syntaxe et sémantique ?

Parce que la phrase constitue un énoncé syntaxiquement indépendant et complet du point de vue du sens, elle apparaît comme une unité importante pour appréhender les mécanismes de base de la compréhension. Parmi les mécanismes engageant le lecteur dans l'analyse du sens de la phrase dans son ensemble, deux grands ensembles de traitements peuvent être distingués selon la nature des informations analysées :

- les traitements des informations lexico-sémantiques englobant la signification des mots, leurs propriétés catégorielles (nom, verbes, mots fonction...) et des informations spécifiques tel que le caractère animé ou inanimé des noms, le nombre d'arguments des verbes (informations sous-catégorielles) ou encore les rôles thématiques de ces arguments (agent, patient, instrument...).
- les traitements des informations syntaxiques permettant le repérage des fonctions des différents mots et des relations qu'ils entretiennent entre eux. Ce sont ces traitements qui, de façon plus ou moins indépendante selon les auteurs, sont à l'origine du « calcul » de la structure syntaxique de la phrase.

- Enfin, et bien qu'ils soient moins accentués qu'à l'oral (Gombert & Colé, 2000), les traitements des informations contextuelles et pragmatiques sont également une composante intervenant dans la compréhension écrite (Golder & Gaonac'h, 1998).

A partir des années 80, et non sans rapport avec la vigueur et l'évolution de la théorie chomskienne sur la grammaire générative transformationnelle (Chomsky, 1957, 1965, pour les premières versions), beaucoup de psycholinguistes se sont attachés à comprendre comment l'analyse syntaxique ou *parsing* participe à l'élaboration de la signification de la phrase. Cependant, ni le caractère obligatoire de l'analyse syntaxique précédant la compréhension, ni la taille de l'unité de traitement, ni même le mode de fonctionnement de l'analyseur syntaxique ou *parser* et le moment de son intervention ne font consensus (Zagar, 1992). Ainsi, plusieurs conceptions, abordant le statut du traitement syntaxique en compréhension de phrases écrites, s'opposent. Une description schématique des points de vue, placerait aux deux extrémités de l'axe du débat, les théories structuralistes (prônant l'autonomie et la primauté de l'analyse syntaxique permettant de recouvrir la structure profonde d'une phrase en appliquant un certain nombre de règles) et les théories fonctionnalistes à orientation connexionniste défendant l'idée d'une intégration de l'ensemble des processus dans un système capable de traiter les relations formes-fonctions sur la base de leurs co-occurrences. Nous illustrerons ces prises de position en prenant quelques exemples de modélisations représentatives des courants autonomistes, interactifs et intégratifs qui ont émergé ces vingt dernières années.

2.1.1 Le modèle du garden-path de Frazier (1987)

Un des modèles les plus représentatifs du courant autonomiste et structuraliste de l'analyse syntaxique est sans doute celui de Frazier (1987). Pour lui, le système de compréhension du langage est doté d'un processeur spécialisé dans le traitement syntaxique. Les informations purement syntaxiques, telles que l'ordre des mots, la catégorie lexicale¹⁶, les règles de structuration syntaxique en constituants, sont traitées en premier lieu et indépendamment des autres caractéristiques linguistiques et extralinguistiques des items lexicaux. L'analyse syntaxique s'effectue donc en temps réel sans tenir compte de la signification. En cas d'ambiguïtés, ce sont les principes d'attachement minimal et de clôture tardive qui vont diriger le

¹⁶ Frazier (1989) précise que la seule connaissance lexicale que le parser utilise durant la phase initiale de structuration syntaxique est la catégorie lexicale.

choix, parfois erroné, vers une structure syntaxique simple et cognitivement économique¹⁷ (pour plus de précision sur les principes d'économie du processeur syntaxique, voir l'encart 3).

Encart 3 : Les deux principaux principes d'économie du processeur syntaxique

- ✓ La règle d'attachement minimal prédit que le système préfère une structure comportant le moins de nœuds syntagmatiques possibles : par exemple dans la phrase « *le garçon regarde le feu d'artifice près de la rivière* », le syntagme prépositionnel « *près de la rivière* » sera préférentiellement rattaché au verbe « *regarde* » car le nœud distribuant les constituants du syntagme verbal est déjà créé alors que le rattachement au syntagme nominal « *le feu d'artifice* » ajouterait un nouveau niveau syntagmatique.
- ✓ La règle de clôture tardive renvoie au fait que le système, au cours de son analyse, va attacher les nouveaux mots à la proposition ou au syntagme en cours de traitement. Par exemple dans la phrase « *le journaliste interviewait la sœur du colonel qui avait eu un accident* », le lecteur aura tendance à rattacher la proposition « *qui avait eu un accident* » au syntagme nominal « *du colonel* » en cours de traitement plutôt qu'au précédent.

Le processeur syntaxique ne propose donc qu'une seule structuration ou représentation syntaxique possible au processeur thématique ayant accès aux informations sémantiques et contextuelles. Si une incompatibilité est détectée, une nouvelle analyse syntaxique sera déclenchée¹⁸. Cet effet de *garden path* (traduisible comme l'induction en erreur par la structuration syntaxique initiale) a notamment été observé par des recherches mesurant les temps de lecture ou les mouvements oculaires (Ferreira & Clifton, 1986; Frazier & Rayner, 1982). La structuration syntaxique, guidée par ces règles d'attachement, s'effectuerait en dépit des préférences structurales sur les compléments pouvant émaner des caractéristiques sous-catégorielles du verbe (Ferreira & Henderson, 1990). Cependant, les résultats de ces recherches ont été à plusieurs reprises critiqués (pour une revue voir Schelstraete, 1993). Ainsi, l'universalité du principe de clôture tardive a été remise en cause du fait des difficultés à le mettre en évidence en espagnol et en français (e.g. Mitchell, Cuetos, & Zagar, 1990). D'autres études infirment l'hypothèse de l'autonomie du fonctionnement syntaxique en montrant par exemple le rôle désambiguïsant du contexte (e.g. Altmann, Garnham, & Dennis, 1992).

2.1.2 Le modèle de la guidance lexicale de Tanenhaus *et al.* (1989ab)

Le courant interactif postule, pour sa part, une interaction et une interdépendance plus ou moins importante entre les traitements syntaxiques et lexico-sémantiques. Le modèle de la guidance lexicale (Tanenhaus, Boland, Garnsey, & Carlson, 1989a; Tanenhaus, Carlson, & Trueswell, 1989b) peut être considéré comme une des conceptions les plus interactives de ce courant

¹⁷ Frazier décrit un autre principe d'économie : « la stratégie de l'antécédent potentiel le plus proche », que nous n'aborderons pas car il fait référence au traitement des anaphores et s'éloigne de notre thématique principale.

¹⁸ Dans une version plus récente, Frazier (1989) propose que le processeur thématique envisage, en cours de traitement et parallèlement aux processus syntaxiques, les différentes structures sémantiques possibles des phrases ambiguës.

(Schelstraete, 1993). Ce modèle s'est développé dans le cadre de la grammaire lexicale fonctionnelle (Kaplan & Bresnan, 1982) qui envisage les correspondances complexes entre les informations syntaxiques et les autres informations fournies par les items lexicaux. L'idée est que la représentation lexicale contient les informations linguistiques concernant un mot (sa transcription phonologique et graphique, sa signification, sa catégorie lexicale, etc.) et toutes informations de nature syntaxique, thématique, sémantique ou pragmatique renseignant sur les relations structurales qu'il peut entretenir avec les autres mots. L'ensemble de ces informations associées au lexique influencerait donc très largement le traitement syntaxique qui pourrait ainsi générer, temporairement et en cours de traitement, de multiples représentations syntaxiques. Plus précisément, Tanenhaus et ses collaborateurs (Tanenhaus *et al.*, 1989a; Tanenhaus *et al.*, 1989b) accordent une place primordiale aux rôles thématiques des arguments du verbe qui seraient pris en compte immédiatement, dès que le lecteur identifie le verbe¹⁹. En formulant certaines attentes à l'égard des compléments susceptibles d'apparaître, le système de compréhension de phrases pourrait alors assigner une signification provisoire à la phrase entière avant le terme de son traitement.

De nombreuses recherches ont obtenu des résultats allant dans le sens de ces hypothèses. Mitchell et Holmes (1985) montrent par exemple que le traitement syntaxique de certaines propositions subordonnées dépend de la nature des verbes de la proposition principale et que l'effet de *garden path* relève d'une incompatibilité sémantique entre les constituants (voir encart 4 pour un exemple).

Encart 4 : Exemples d'une incompatibilité sémantique entre les constituants de la phrase

Par exemple, la phrase « le gardien remarque la fille agitant une baguette dans sa main (à lui) » est traitée moins rapidement que la phrase « le gardien poursuit la fille agitant une baguette dans sa main (à lui) » car le syntagme « dans sa main » crée une incompatibilité avec la structure argumentale du verbe « remarquer » préférant l'attachement du syntagme « agitant une baguette » à la tête « fille » ce qui n'est pas le cas avec le verbe « poursuivre » qui admet davantage l'attachement d'une proposition adverbiale. Les auteurs interprètent cette différence en terme de dominance du verbe véhiculant des informations cruciales et spécifiques (par exemple son aspect dynamique) permettant de guider le *parsing* syntaxique.

Boland (1997) met, pour sa part, en évidence que les rôles thématiques des arguments du verbe sont identifiés en cours de traitement et permettent une interprétation provisoire de la phrase entière. Son étude montre par exemple que les phrases telles que « à quel enfant la baby-sitter a-t-elle eu une drôle de voix ? » sont généralement identifiées comme asémantiques dès l'apparition du mot suivant le verbe.

¹⁹ La reconnaissance du verbe rend en effet disponible la représentation sémantique de ce verbe, ainsi que les rôles thématiques des arguments qui lui sont associés, les types de constituants pouvant lui servir de compléments et la façon dont les rôles et les constituants sont associés entre eux.

Une telle modélisation envisage donc, au contraire de la théorie de Frazier, une interaction forte entre les aspects lexicaux-sémantiques et syntaxiques tout en les considérant comme des éléments de représentations distincts. Les modèles intégratifs franchissent en cela un pas de plus dans la façon d'envisager les relations entre les multiples informations véhiculées par les mots et les représentations cognitives impliquées dans les traitements interprétatifs des phrases.

2.1.3 Le modèle à satisfaction de contraintes multiples de Taraban et McClelland (1990)

Les modèles intégratifs, tel que celui de Taraban et McClelland (1988, 1990) dit de la « satisfaction des contraintes multiples », envisagent un système en interaction totale, dans lequel les facteurs syntaxiques et les facteurs lexico-sémantiques d'ordre supérieur influencent conjointement la construction de la représentation initiale de la phrase. Lors de cette construction, l'ensemble des informations disponibles, qu'elles soient linguistiques ou non linguistiques, est intégré simultanément. Dans ce modèle, l'analyseur syntaxique n'existe plus en tant que tel, puisque toutes les informations sont reliées les unes aux autres et que les processus de compréhension et de structuration syntaxique sont indissociables.

Plus précisément, chaque information est dotée d'un certain degré d'activation et impose des contraintes entrant en compétition les unes avec les autres. Ainsi, les contraintes syntaxiques pourraient dominer dans certaines phrases ou parties de phrase alors que les contraintes sémantiques pourraient dominer dans d'autres selon la puissance de chacune de ces sources d'informations.

Les auteurs avancent également que les informations sémantiques influençant l'agencement des constituants de la phrase ne sont pas exclusivement portées par les verbes, mais sont en fait plus largement distribuées. Ils insistent, par exemple, sur le fait que les têtes des syntagmes nominaux véhiculent des contraintes pouvant intervenir, aussi bien sur les attentes de nature syntaxique, portant sur l'attachement du constituant prépositionnel, que sur les attentes portant sur l'assignement du rôle sémantique de ce constituant (*cf.* encart 5).

Encart 5 : Résumé de l'expérimentation de Taraban et McClelland (1988)

Taraban et McClelland (1988) ont mené une expérimentation faisant apparaître deux résultats allant dans le sens de leur hypothèse : premièrement, ils observent que la lecture mot à mot d'une phrase comme « le dictateur regardait les troupes de la ville » est plus lente par rapport à la phrase « le dictateur regardait les troupes des marches » car dans le premier cas, le syntagme prépositionnel (« *from the city* » en anglais) ne respecte pas l'attachement contraint par la tête du syntagme nominal « les troupes » (qui ne nécessite pas particulièrement d'information sur sa source). De façon identique, la lecture de la phrase « le dictateur regardait les pétitions du podium » est plus lente que celle de la phrase « le dictateur regardait les pétitions des prisonniers » car dans ce cas, le syntagme prépositionnel « *from the podium* » ne respecte pas l'attente d'une précision sur la source requis par le syntagme nominal « les pétitions ». Pour les

auteurs, ces résultats donnent un exemple de l'intégration complexe entre de multiples informations de différentes natures.

Enfin, ils envisagent que les attentes des lecteurs soient surtout déterminées par la fréquence d'occurrence des relations entre les items lexicaux et leur fonction plutôt que par l'application de règles syntaxiques ou sémantiques indépendantes les unes des autres. Cependant, un tel modèle mériterait d'être étayé par davantage de données empiriques et d'études explorant les aspects computationnels déterminant la fréquence des relations envisagées.

Globalement, bien que l'intervention des traitements de nature syntaxique soit généralement acceptée, les divergences de points de vue et les controverses méthodologiques qui ont animé les auteurs semblent limiter les possibilités d'appréhender clairement les mécanismes en jeu lors de la compréhension de phrases.

Si ces modèles offrent tous l'avantage de s'intéresser aux traitements « *on-line* », ils se sont exclusivement attachés à comprendre les traitements syntaxiques réalisés chez le lecteur expert, en laissant de côté l'acquisition de ces traitements chez l'enfant et leur impact lors de l'apprentissage de la lecture.

Il peut également leur être reproché le fait qu'ils reposent sur des études portant, soit sur une structure syntaxique particulière, soit sur un matériel linguistique spécifique ou encore qu'ils se fondent sur des exemples de phrases relativement prototypiques. Certaines oppositions pourraient ainsi s'expliquer par le caractère « falsifiable » des postulats énoncés. Quoi qu'il en soit, ces théories se heurtent souvent à la difficulté de cerner et de catégoriser l'ensemble des informations disponibles dans une phrase, limitant de fait leurs capacités de généralisation. Par conséquent, les différences interindividuelles, de même que les différences interlangues, sont rarement clarifiées (Schelstraete, 1993). Les travaux initialisés par Bates et MacWhinney (1989) se situant dans une perspective intégrative et fonctionnaliste, ont tenté d'ouvrir une perspective pouvant lever certaines de ces difficultés, même si leur modélisation concerne plus spécifiquement le langage oral.

2.2 Analyse des indices de surface et stratégie d'interprétation : le modèle de compétition de Bates et MacWhinney (1989)

L'objectif de Bates et MacWhinney (1989) est d'élaborer un modèle pouvant rendre compte des variations quantitatives et qualitatives interlangues, prolongeant ainsi la perspective comparative initiée par Slobin (1980, 1985, cité par Kail, 1983, 1999).

Proches des conceptions de Taraban et McClelland, ces auteurs envisagent les connaissances linguistiques sous-jacentes à l'interprétation des phrases comme un réseau complexe de

correspondances, d'appariement direct entre les formes lexicales et les fonctions sémantiques qu'elles remplissent²⁰ (action, agent, patient, thème, etc.). Ils distinguent quatre types d'indices de surface pouvant représenter linguistiquement les relations entre les mots et donc permettre de leur attribuer une fonction dans la phrase.

1. La localisation des mots dans la phrase, l'ordre séquentiel des mots jouant un rôle dans l'établissement des relations locales.

2. Les marques morpho-flexionnelles portées par les mots assurant les relations coréférentielles. Il peut être distingué :

- les marques d'accord (déclinaison en nombre et/ou en genre pour les noms, pronoms et adjectifs et en nombre et en personne pour les verbes) établissant notamment les relations entre le sujet et le verbe ou entre le sujet et l'attribut ;
- les marques casuelles utilisées dans de nombreuses langues (nominatif, accusatif, datif, génitif etc.) permettant d'identifier les relations entre le verbe et ses arguments.

3. Les caractéristiques lexico-sémantiques et pragmatiques des mots : la principale propriété prise en compte dans ce domaine est le caractère animé ou inanimé des noms.

4. Les marques prosodiques à l'oral, qui sont souvent remplacées à l'écrit par la ponctuation.

Cependant, l'appariement entre les différents indices et les différentes fonctions n'est pas un appariement terme à terme car il existe incontestablement une multitude de relations formes / fonctions (ou encore un recouvrement partiel dans cet appariement) que cela soit au sein d'une même langue ou entre différentes langues²¹. En effet, dans les langues naturelles, les marques de surfaces portées par les formes lexicales peuvent avoir plusieurs fonctions (un nom animé et défini peut être agent ou patient) de même qu'une même fonction peut s'exprimer par de multiples indices de surfaces (la fonction d'agent peut être marquée par la position préverbiale du nom mais aussi par l'accord en personne et en nombre avec le verbe et dans certaines langues par une flexion casuelle).

Ainsi, les auteurs postulent que les divers indices, pouvant servir de sources d'information pour attribuer une même fonction, entrent en compétition les uns avec les autres. Dans une langue donnée, la préférence stratégique pour tel ou tel indice, dont dépendent les processus de compréhension, s'effectue selon la validité des indices et de leur poids.

²⁰ Les auteurs parlent plus exactement de fonction communicationnelle en référence aux significations et aux intentions exprimées dans un énoncé.

²¹ Les auteurs intègrent également la notion de coalition entre différents indices et entre différentes fonctions pour rendre compte de la combinatoire des appariements formes-fonctions.

La notion de validité renvoie à la valeur informative (objective et quantifiable) d'un indice (par exemple la position préverbale) pour attribuer une fonction ou une signification communicative particulière (par exemple le rôle d'agent). Elle se définit à l'aide de sa disponibilité et de sa fiabilité. Un indice est considéré comme disponible lorsqu'il est présent à chaque fois que le locuteur ou le lecteur en a besoin. Un indice est fiable si sa prise en compte donne toujours lieu à une interprétation correcte. En anglais par exemple, la position préverbale est un indice particulièrement valide car il est très souvent et sans ambiguïté lié à la fonction d'agent.

La notion de poids des indices est le reflet de leur validité, mais constitue une propriété subjective que l'on ne peut appréhender qu'à partir des comportements de l'utilisateur du langage. Elle relève ainsi d'une construction psychologique octroyant un certain degré d'applicabilité aux indices estimé à partir de leur validité.

En termes de validations empiriques, les nombreuses recherches menées dans le cadre de ce modèle ont particulièrement investi la perspective interlangue, au centre des préoccupations théoriques. Les conclusions des recherches abordant spécifiquement les variations interlangues, présentées dans le chapitre suivant, confirment que les stratégies d'interprétation et leur développement, dépendent globalement de l'ordre d'importance des indices (en termes de validité et de poids) dans une langue donnée pour assigner les fonctions aux éléments de la phrase.

D'un point de vue théorique général, ce courant a largement participé aux débats sur l'universalité et la spécificité de l'acquisition du langage, ainsi que sur les relations entre processus cognitifs et développement linguistique (*e.g.* Bronckart, 1977; Kail, 1983, pour une revue).

Il convient toutefois de rappeler que les stratégies d'interprétation et leur acquisition, envisagées dans cette perspective, concernent exclusivement le langage oral. Ce courant a certes permis de réfuter « le mythe de l'ordre » envisagé comme une prédisposition universelle dans le processus d'acquisition des langues (Kail, 2000) ; mais la possibilité de généraliser ces conclusions à la compréhension écrite mérite, à notre avis, de réexaminer le postulat d'une compréhension amodale du langage, d'approfondir l'articulation entre la compréhension orale et écrite et de replacer au centre de la réflexion les procédures de contrôle des aspects syntaxiques.

2.3 Y-a-t-il un développement spécifique de la compréhension écrite ? La question du contrôle des traitements syntaxiques

Egalement en jeu à l'oral, les mécanismes de compréhension ne sont généralement pas considérés comme spécifiques à l'apprentissage de l'écrit (Morais, 1994). Autrement dit, apprendre à lire consiste essentiellement à automatiser les procédures de reconnaissance des mots. Une fois le décodage maîtrisé, le lecteur est en mesure de comprendre ce qu'il lit, de la même façon que ce qu'il entend. De nombreuses recherches ont ainsi observé, d'une part le lien corrélationnel élevé entre la vitesse de reconnaissance de mots écrits isolés (mesurant le degré d'automatisme du décodage) et les capacités de compréhension en lecture, particulièrement chez les lecteurs débutants voir (pour une revue détaillée voir Stanovich, 1982ab) et d'autre part le lien corrélationnel entre les capacités de compréhension à l'oral et en lecture, d'autant plus marqué que le niveau d'expertise augmente (Stanovich, Cunningham, & Feeman, 1984).

Ces considérations expliquent que les modèles de la lecture se soient principalement centrés sur la mise en place de l'accès au lexique sans envisager l'évolution des procédures de calcul syntaxique et d'élaboration de la signification, considérant implicitement que ces mécanismes relèvent de l'acquisition préalable du langage oral. Il n'est pas ici question de remettre en cause la conclusion que la maîtrise de l'identification des mots écrits constitue une condition nécessaire à la compréhension (*cf.* paragraphe 1.3 supra), ni de nier que les mécanismes d'analyse syntaxique et sémantique sont en jeux à l'oral comme à l'écrit. Mais la maîtrise du décodage ne peut être considérée comme suffisante. De même, il existe un certain nombre de différences irréductibles entre le système oral et écrit pouvant affecter la compréhension particulièrement en début d'apprentissage.

Gombert et Colé (2000) soulignent quatre différences fondamentales entre l'oral et l'écrit résultant du mode physique des inputs.

- La première est que les caractéristiques extralinguistiques et prosodiques sont importantes à l'oral mais absentes à l'écrit.
- La deuxième concerne les contraintes mnésiques, plus réduites à l'écrit qu'à l'oral du fait du caractère permanent de la trace écrite.
- La troisième a trait aux différences de marquage des frontières des mots, physiquement identifiables à l'écrit mais pas ou moins clairement à l'oral.
- La quatrième oppose les aspects sociaux de la communication orale et écrite, liés à leur contexte de production : les messages oraux se situent généralement dans une dynamique contextuelle moins figée (facilitant le partage des cadres et références conceptuels entre les interlocuteurs). Le mode de communication à l'oral revêt ainsi un caractère plus naturel et plus

interactif, alors que l'écrit se caractérise habituellement par un mode plus conventionnel et par une activité individuelle qui le rend en ce sens relativement « décontextualisé ».

Gombert et Colé (2000) avancent ainsi que ces différences amènent à nuancer l'isomorphisme des traitements entre l'oral et l'écrit. Lors de la lecture, l'absence de certaines informations permettant à l'oral de structurer le discours (l'accentuation, l'intonation) et le rôle limité du contexte au sens large et des aspects pragmatiques, doivent être compensés. Cette compensation pourrait notamment s'effectuer par une prise en compte plus accentuée de la structure syntaxique. Par ailleurs, chez l'apprenti lecteur et contrairement à ce qui caractérise le lecteur expert, non seulement la reconnaissance des mots écrits est lente et coûteuse en attention, mais le calcul syntaxique automatique est impossible car il nécessite une reconnaissance rapide et quasi simultanée d'un ensemble de mots pour pouvoir se mettre en place sans effort attentionnel. Alors qu'à l'oral, les connaissances des aspects syntaxiques sont utilisées sans contrôle, la compréhension en lecture chez le lecteur débutant reposerait particulièrement sur une application volontaire des règles syntaxiques, qui relèvent de connaissances conscientes ou encore explicites. Paradoxalement, cette conscience syntaxique, qui n'est pas indispensable à l'oral, est conséquente à la manipulation de l'écrit et réciproquement, la compréhension écrite, particulièrement en début d'apprentissage, dépend de la conscience syntaxique (Gombert, Martinot, & Nocus, 1996; les recherches ayant étudié cette relation seront développées dans le chapitre 4).

Ainsi, il semble qu'il y ait bien une spécificité de la compréhension en lecture débutante résidant dans cette relation biunivoque avec la conscience syntaxique. Il est ainsi envisageable d'appliquer la conception de Gombert sur les apprentissages implicites et explicites à la syntaxe et à son influence sur la maîtrise de la compréhension en lecture (*cf.* paragraphe 1.3).

Chapitre 2 : Lecture et spécificités linguistiques

En lien avec l'internationalisation de la recherche en psycholinguistique, la diversité des langues constitue aujourd'hui un domaine d'étude particulièrement fécond pour formaliser des théories générales sur les mécanismes de lecture. Ces théories cherchent de plus en plus souvent à distinguer les universaux dans l'acquisition du langage de l'influence des aspects spécifiques à chacune des langues. Une telle démarche, qui se veut plus écologique, constitue un changement de perspective par rapport aux recherches menées sur l'anglais, langue qui révèle finalement de nombreuses exceptions.

1. Variations interlangues et décodage

1.1 Le degré de transparence des systèmes orthographiques

Assez récemment, une attention particulière a été accordée au fait que la maîtrise de l'écrit et l'acquisition de la lecture pouvaient varier d'une langue à l'autre en fonction du degré de « profondeur » ou de transparence de leur système orthographique (Seymour *et al.*, 2003; Treiman, 2005ab; Ziegler, 2004). Le degré de profondeur d'une langue dépend de la transparence ou de la régularité de ses correspondances graphèmes-phonèmes. Une orthographe est profonde lorsque les relations entre orthographe et phonologie sont équivoques : à un même graphème peuvent être associés plusieurs phonèmes (*e.g.* en français le graphème <s> peut être prononcé [z] ou [s] selon sa position et le graphème qui le suit, mais peut aussi ne pas se prononcer lorsqu'il marque le pluriel) et plusieurs graphèmes peuvent renvoyer au même phonème (*e.g.* toujours en français, les graphèmes <o>, <eau> et <au> se prononcent [o]). Une orthographe est superficielle lorsque les correspondances graphèmes-phonèmes sont biunivoques, c'est-à-dire qu'un graphème ne renvoie qu'à un seul phonème et réciproquement. Les différentes langues alphabétiques se situent donc sur un continuum, allant des langues à orthographe profonde, opaque ou encore irrégulière, aux langues à orthographe superficielle c'est-à-dire transparente ou régulière²².

Le degré d'irrégularité que peut revêtir ces correspondances a amené certains chercheurs à postuler que les orthographe transparentes se prêtent plus facilement à une reconnaissance des mots par la voie phonologique que les orthographe opaques, qui encouragent plutôt le lecteur à

²² Pour simplifier la lecture, la terminologie de langue opaque ou irrégulière *versus* langue transparente ou régulière sera utilisée bien qu'en toute rigueur, ce soit leur système d'écriture qui revêt ces caractéristiques.

procéder à un adressage direct sans passer par un décodage phonologique (Frost, Katz, & Bentin, 1987; Katz & Frost, 1992). Cette théorie, dite de la *transparence orthographique*, a cependant été remise en cause par plusieurs études montrant des effets phonologiques importants dans les langues opaques, comme l'anglais, même chez les lecteurs experts (*e.g.* Rayner, Sereno, Lesch, & Pollatsek, 1995; Van Orden, 1987; Ziegler, Van Orden, & Jacobs, 1997)²³.

Les études comparant souvent l'anglais, dont l'orthographe est particulièrement opaque, à d'autres langues indo-européennes plus transparentes, ont ainsi mis en évidence des différences notoires dans la rapidité d'acquisition de la lecture, sans remettre en cause l'universalité de la segmentation phonémique (Caravolas, Volin, & Hulme, 2005).

Thorstad (1991) a ainsi comparé l'acquisition des habiletés de compréhension en lecture d'enfants italiens et anglais âgés de 6 à 11 ans. Les résultats font apparaître une proportion plus importante d'erreurs chez les enfants anglophones, ayant suivi un enseignement orthographique traditionnel. Oney et Goldman (1984) indiquent des résultats allant dans le même sens en comparant des enfants de langue maternelle turque et américaine en 1^e et 3^e année de scolarisation. Cependant, en première année seulement, les enfants américains obtiennent de moins bons résultats en compréhension de texte. Par ailleurs, la lecture des mêmes pseudo-mots est effectuée plus lentement par les enfants américains. En 1^e année, leur précision en lecture est également plus faible mais cela n'est toutefois plus visible en 3^e année.

Goswami, Porpodas et Weelwright (1997) et Goswami, Gombert et Fraca de Barrera (1998) confirment le désavantage des enfants de langue anglaise âgés de 7 à 9 ans pour la lecture de pseudo-mots par rapport aux enfants grecs, français et espagnols de même âge.

Enfin, la comparaison anglais – hollandais menée par Patel, Snowling et De Jong (2004), auprès d'enfants de 6 à 11 ans, confirme les difficultés des enfants anglophones, en particulier pour la précision de la lecture de pseudo-mots chez les plus âgés.

Il semble toutefois délicat d'effectuer des comparaisons sur du matériel linguistique spécifique aux langues et de contrôler rigoureusement l'influence des facteurs visuels, sémantiques et articulatoires des items (Wimmer & Goswami, 1994). La comparaison entre l'anglais et l'allemand permet de réduire ces risques. Toutes deux germaniques, ces langues utilisent des codes orthographiques et phonologiques similaires mais se distinguent par leur régularité graphèmes-phonèmes (Ziegler, 2004). Les recherches comparatives entre ces deux langues ont ainsi l'avantage de pouvoir utiliser des mots quasi-identiques dans les deux langues. Les travaux

²³ Ces recherches ont notamment montré les effets d'homophonie en amorçage rapide sur différentes tâches de lecture (lecture avec enregistrement des fixations oculaires, catégorisation sémantique, tâche de recherche de lettre).

de Frith, Wimmer et Landerl (1998) ont ainsi montré des différences importantes et persistantes dans l'exactitude et la rapidité en lecture de mots et de pseudo-mots entre les enfants anglophones et germanophones ayant entre 7 et 12 ans. Le désavantage des anglophones n'est plus significatif pour la lecture de mots à partir de 9 ans mais persiste jusqu'à 12 ans, en ce qui concerne l'exactitude de la lecture de pseudo-mots trisyllabiques. Ces résultats se rapprochent de ceux obtenus par Wimmer et Goswami (1994)²⁴, Landerl (2000)²⁵ et Mann et Wimmer (2002)²⁶ auprès d'enfants de même langue.

Relativement peu de recherches se sont attachées à comparer d'autres langues. Celle de Defior, Martos et Cary (2002) s'est focalisée sur la lecture d'enfants espagnols et portugais de la 1^e à la 4^e année de scolarité. Les auteurs observent qu'en début d'apprentissage les enfants espagnols sont plus rapides et font moins d'erreurs sur les pseudo-mots que leurs pairs lusophones. Ils atteignent, dès la deuxième année, un niveau de lecture équivalent à celui des plus âgés. Cette facilité à lire en espagnol a également été relevée par Goswami *et al.* (1998) dans leur étude anglais/français/espagnol. Le pourcentage de pseudo-mots lus correctement, atteint des scores plafonds chez les hispanophones, sont plus faibles chez les francophones et encore plus chez les anglophones. L'espagnol apparaît donc comme une langue très transparente, ce qui facilite son acquisition.

Quelques recherches ont comparé plusieurs langues de manière à élaborer une classification en fonction de leurs caractéristiques orthographico-phonologiques.

L'étude d'Aro et Wimmer (2003) a ainsi pour objectif de confirmer ces effets de transparence orthographique sur un nombre de langues plus étendu incluant l'anglais, l'allemand, l'espagnol, le français, le hollandais, le suédois et le finnois. Cette recherche, réalisée auprès d'enfants de 1^e, 2^e, 3^e et 4^e année, utilise la technique des ratios²⁷. Les résultats obtenus confirment tout d'abord que, à l'exception des enfants anglais, la lecture des pseudo-mots avoisine, dès la fin de la 1^e année, 90 % de prononciation correcte. Par ailleurs, dès la 1^e année, les enfants finlandais, suédois et espagnols lisent plus rapidement les pseudo-mots alors que les enfants anglais, allemands, hollandais et français mettent toujours en 4^e année deux fois plus de temps pour lire les pseudo-mots que pour nommer les nombres. Enfin, seuls les enfants finlandais, et dans une moindre mesure les enfants suédois, parviennent dès la deuxième année à lire les mots plus vite qu'ils ne nomment les nombres. Les auteurs expliquent ces résultats en terme de transparence

²⁴ Recherche réalisée auprès d'enfants âgés de 7 à 9 ans.

²⁵ Recherche réalisée auprès d'enfants âgés de 6 à 9 ans.

²⁶ Recherche réalisée auprès d'enfants âgés de 7 et 8 ans.

²⁷ Ces ratios sont calculés à partir du temps de lecture de mots, de nombre et de pseudo-mots dérivés de ces derniers, de façon à obtenir deux valeurs numériques indépendantes de la longueur des mots, variable entre les langues.

orthographique mais en prenant également en compte d'autres caractéristiques des langues comme la présence de lettres muettes (en français notamment) ou de graphèmes complexes (nombreux en allemand).

L'article de Seymour *et al.* (2003), décrit pour sa part les résultats d'une collaboration européenne associant quatorze pays différents²⁸. Les données recueillies concernent 13 langues différentes : l'allemand, l'anglais, le danois, l'espagnol, le finnois, le français, le grec, le néerlandais, l'islandais, l'italien, le norvégien, le portugais et le suédois. Le niveau de lecture de mots familiers et de non-mots a notamment été évalué auprès d'élèves en fin de 1^e année. Les résultats confirment que, pour une majorité de langues, les enfants atteignent un niveau de lecture efficace dès la fin de la 1^e année, exception faite :

- de l'anglais : les performances y sont beaucoup plus faibles que dans les autres langues ;
- du français et du portugais : on y observe une réduction de l'exactitude et de la vitesse en lecture de mots familiers et de non-mots ;
- du danois : on y constate une réduction de l'exactitude en lecture de mots et de la rapidité et de l'exactitude de la lecture de non-mots ;
- dans une moindre mesure du néerlandais et du suédois, faisant apparaître une réduction de la rapidité en lecture de non-mots bisyllabiques.

Pour les auteurs, ces différences ne peuvent être attribuables à des différences d'âge ou de connaissance des lettres mais seraient dues à des différences linguistiques telles que la complexité syllabique et la transparence orthographique des systèmes d'écriture. Les variations interlangues, inégales entre la lecture de mots et de non-mots, permettent aux auteurs de conclure que la complexité syllabique affecte sélectivement le décodage alors que la transparence orthographique affecte conjointement la lecture de mots et de non-mots. Ils proposent également une classification des langues en fonction de leur transparence orthographique et de leur complexité syllabique (*cf.* tableau 1).

Tableau 1: Classification des langues proposée par Seymour *et al.* (2003)

		Transparence orthographique			
		transparente		opaque	
Structure syllabique	Simple	finnois	grec italien espagnol	portugais	français
	Complexe		allemand norvégien islandais	néerlandais suédois	danois anglais

²⁸ COST Action A8.

La conclusion la plus évidente qui émane de ces recherches concerne, ainsi, la spécificité de l'anglais, dont les caractéristiques orthographiques s'avèrent responsables de la lenteur dans l'acquisition du décodage généralement observée.

1.2 Les variations des caractéristiques morphologiques des systèmes d'écriture

Contrairement aux recherches sur la transparence des systèmes orthographiques, les études interlangues s'intéressant aux variations des aspects morpho-dérivationnels et leur impact sur la lecture sont rares. Cette lacune est sans doute liée aux difficultés méthodologiques incontournables pour comparer des procédures complexes et forcément associées à la dimension sémantique des mots.

Pourtant, les chercheurs admettent clairement que la complexité morphologique se réalise dans les différentes langues selon des principes différents (Frost, Kugler, Deutsch, & Forster, 2005b; Plaut & Gonnerman, 2000; Schreuder & Baayen, 1995). Par exemple, en italien, à la différence d'autres langues comme le français ou l'anglais, les bases morphologiques n'existent pas de façon indépendante (Colé, Segui, & Taft, 1997). De plus, les variations phonologiques des bases lors de la dérivation sont fréquentes en anglais mais inexistantes en serbo-croate (Frost, Grainger, & Rastle, 2005a). Dans les langues agglutinantes comme le turc ou le finnois, les morphèmes sont successivement ajoutés à une base, ce qui rend le processus de dérivation très transparent. Les langues non concaténatives comme les langues sémitiques révèlent aussi des caractéristiques morphologiques particulières. Contrairement aux langues concaténatives qui utilisent une construction morphologique linéaire²⁹, les langues sémitiques se caractérisent par une construction morphologique non linéaire. Les constructions morphologiques procèdent en effet par l'intrication d'un schème vocalique ou pattern à une racine, généralement composée de trois consonnes et ne pouvant exister dans sa forme nue. Schématiquement, l'intrication s'effectue par le placement de certains éléments du schème (une ou plusieurs lettres ou signes diacritiques) entre les consonnes de la racine. En terme perceptif, cette structuration en racine / schème implique, non pas une relation contiguë entre les unités morphémiques, mais au contraire une rupture de la forme orthographique et phonologique de ces éléments (par exemple, en hébreu, la racine ZMR se transforme en ZAMAR après l'application du schème -A-A-³⁰). Cette structuration rend le découpage morphologique peu transparent (Frost *et al.*, 2005a). Par ailleurs,

²⁹ Dans les langues concaténatives, le principe de dérivation s'effectue par une concaténation séquentielle d'unités morphémiques (par exemple l'ajout d'un préfixe ou d'un suffixe à une base). Ce processus peut parfois entraîner des modifications de la base mais l'intégrité orthographique de cette dernière reste intacte (Frost, Kugler, Deutsch & Forster, 2005b).

³⁰ Les tirets représentent la position des consonnes de la racine.

les morphèmes de ces langues sont plus abstraits puisque le sens exact des mots ne peut être identifié que par la combinaison de ces deux dimensions complémentaires (par exemple en hébreu la racine ZMR se rapporte au concept de « chant » mais le mot ZAMAR -chanteur- ne peut se comprendre sans l'information portée par le schème précisant qu'il s'agit d'un nom de profession (Frost *et al.*, 2005a). Ils sont en revanche très productifs car cette structure bidimensionnelle concerne la quasi majorité du lexique (Baccouche, 2003; Bentin, Frost, & Feldman, 1995; Cohen, 2007), même si le recouvrement sémantique entre les mots de la même famille peut être opaque (Frost *et al.*, 2005a).

Cette particularité a amené les chercheurs étudiant les langues sémitiques à s'interroger sur la façon dont l'espace lexical est organisé chez les lecteurs et sur le rôle respectif des racines et des schèmes dans cette organisation. Bien que les recherches menées concernent exclusivement la lecture experte, et ne soient pas toujours orientées vers une perspective interlangue, nous leur accorderons une place de choix, compte tenu de leur valeur heuristique dans le domaine des spécificités morphologiques des langues et de l'intérêt de cette étude pour les langues sémitiques.

1.3 Les spécificités de la morphologie non concaténative des langues sémitiques

Frost, Bentin et leurs collaborateurs ont été parmi les premiers à étudier les traitements morphologiques en hébreu chez le lecteur expert. Leurs travaux visaient à éprouver les hypothèses avancées à partir des données obtenues dans d'autres langues, voire même à proposer de nouvelles modélisations. Allant à l'encontre des conceptions connexionnistes (*cf.* paragraphe 1.2.2 du chapitre 1) et des études indiquant les effets de transparence sémantique (*e.g.* Marslen-Wilson *et al.*, 1994), ils ont notamment montré que les traitements morphologiques en hébreu sont indépendants des relations sémantiques existant entre les mots de la même famille.

Par exemple, lors de tâches de décision lexicale ou de dénomination précédées d'un amorçage rapide non détectable, Frost *et al.* (1997) ont montré un effet facilitateur, par l'amorçage d'une racine. Ils ont également fait apparaître un effet facilitateur avec, comme amorce, un mot dérivé de la même racine que le mot cible mais qui ne lui est pourtant pas sémantiquement lié. Frost *et al.* (2000b) retrouvent une facilitation avec une amorce sémantiquement opaque par un paradigme d'amorçage intermodal. Cette dernière recherche obtient, ainsi, des résultats très différents de celle de Marslen-Wilson *et al.* (1994) en anglais, avec une tâche et un matériel comparable. Selon les auteurs, cette différence n'est pas seulement liée à l'étendue et à la richesse du système morphologique des langues. Il pourrait s'expliquer par le fait que les

processus de construction morphologique en anglais ne sont pas suffisamment productifs pour exiger l'utilisation de la structure morphologique par le lecteur, ni pour stocker et retrouver l'information lexicale. En revanche, en hébreu, cette procédure est une condition nécessaire. Cependant, des données relatives à d'autres langues doivent être recueillies pour affirmer clairement les raisons de ces différences, pouvant être attribuées aux propriétés concaténatives ou non concaténatives des langues, ou plus généralement au rôle que joue la morphologie dans la structuration des mots.

Par ailleurs, si Frost *et al.* (1997) n'obtiennent pas d'effet d'amorçage, ni à l'aide de mots partageant le même schème nominal que la cible, ni à l'aide de noms pseudo-dérivés construits par la combinaison d'une racine existante et d'un schème légal, Deutsch, Frost et Forster (1998) parviennent à les faire apparaître sur le système de dérivation des verbes. Leurs résultats indiquent l'importance des schèmes verbaux dans l'accès au lexique hébreu, mais ne remettent pas en cause le rôle fondateur de la racine dans le système dérivationnel hébreu. Les résultats de Frost, Deutsch et Forster (2000a) viennent compléter ces observations en montrant que l'effet facilitateur d'un verbe amorce, partageant le même schème que le verbe cible, apparaît, à condition qu'il soit construit avec une racine complète, composée de trois consonnes. Cet effet facilitateur apparaît, d'ailleurs, en décision lexicale, même lorsque la racine triconsonantique de l'amorce n'existe pas (pseudo-racine). Le traitement de la structure consonantique serait ainsi une condition nécessaire à la décomposition morphologique des verbes hébreux mais pourrait s'effectuer sans que le sens de cette racine soit détecté.

A partir de ces données, les auteurs proposent un modèle des traitements morphologiques en jeu dans l'organisation lexicale en hébreu. Dans ce modèle, deux niveaux de représentation coexistent : le niveau des unités lexicales stockant les mots entiers, et un niveau prélexical morphologique représentant les racines, unités fondamentales du système verbal et nominal, mais aussi les schèmes verbaux. Ces deux niveaux sont interconnectés et les processus d'accès au lexique et de décomposition morphologique peuvent s'effectuer simultanément et s'influencer mutuellement à travers des connexions bidirectionnelles entre les deux niveaux. En d'autres termes, l'accès à la racine ou au schème verbal peut s'effectuer soit *via* le niveau lexical, par les mots contenant ces morphèmes, soit en suivant directement un processus de décomposition morphologique de la structure orthographique.

Une dernière recherche plus récente, menée par Frost *et al.* (2005b) auprès d'adultes de langue maternelle hébreu, arabe et bilingues hébreu-anglais accentue encore les différences entre ces deux langues. En plus d'un effet robuste d'amorçage morphologique, les résultats obtenus

montrent que l'effet d'amorçage orthographique apparaît en anglais et non en hébreu³¹, que la cible soit un mot contenant une racine productive ou non productive (non utilisée dans d'autres mots) ou qu'elle soit un mot ayant beaucoup ou peu de voisins orthographiques. Cette absence d'amorçage orthographique est également observée en arabe, langue qui partage les mêmes caractéristiques orthographiques que l'hébreu. Cette étude confirme ainsi le rôle des traitements purement morphologiques lors de l'accès lexical et l'importance des caractéristiques structurales dans son organisation en langues sémitiques.

Les autres recherches menées en langue arabe corroborent globalement ces résultats : Grainger, Dichy, El-Halfaoui et Bamhamed (2003) observent un effet facilitateur robuste de la présentation de la racine dans l'identification des mots. Pour les mots peu fréquents, cette facilitation est plus importante lorsque la racine est fréquente que lorsqu'elle est rare.

Boudelaa et Marslen-Wilson ont mené une série de travaux³² sur la décision lexicale de noms et verbes dérivés, avec amorçages de différentes natures (écrit, oral et intermodal) et plus ou moins rapides (SOA de 32 ms, 48 ms : amorce non détectable ; 64 et 80 ms : amorce détectable). D'une part ils vérifient les affirmations de Frost *et al.* (1997, 2000) en montrant un effet robuste d'amorçage morphologique par la racine quel que soit le degré de transparence sémantique entre l'amorce et la cible (Boudelaa & Marslen-Wilson, 2005). D'autre part, ils s'intéressent à l'effet d'amorçage par un schème : les résultats font apparaître un effet de facilitation du partage du schème verbal pour les trois types d'amorçages utilisés (Boudelaa & Marslen-Wilson, 2004) et du schème nominal en amorçage écrit (Boudelaa & Marslen-Wilson, 2005) mais pas dans toutes les conditions. En effet, cette facilitation par le schème, moins robuste que celle par la racine, n'apparaît pas pour des SOA les plus courts (32 ms) ou les plus longs (80 ms). Leurs recherches étayaient ainsi l'hypothèse que les informations de la racine et du schème ne jouent pas le même rôle dans la construction des représentations lexicales. En effet, l'orthographe arabe représente pleinement les informations sur la racine consonantique mais seulement partiellement les voyelles du schème. Par conséquent, le sens de la racine peut être extrait directement par le traitement orthographique des consonnes, alors que l'accès aux informations morphosyntaxiques du schème nécessite, selon les auteurs, une médiatisation phonologique retardant leur traitement. Par ailleurs, les informations transmises par la racine sont plus précises que celles du schème, ce

³¹ Les auteurs ont pris soin de contrôler la dominance d'une langue sur l'autre chez les participants bilingues. Ainsi, ces résultats ne peuvent être attribués à une différence de niveau d'efficacité entre les langues.

³² Dans les premiers articles, Boudelaa et Marslen-Wilson (2001) s'intéressent à une unité morphologique particulière, l'étymon composé d'une consonne et d'une voyelle, qui fait controverse (*cf.* la réponse de Bentin & Frost, 2001). Cette controverse et le statut de l'étymon dans les langues sémitiques ne seront toutefois pas abordés ici.

qui rend plus cohérent le traitement des mots par famille morphologique que par le partage d'un même schème.

En conclusion, bien que l'ensemble des recherches présentées ici en langues sémitiques souligne les caractéristiques morphologiques spécifiques aux langues concaténatives et leur impact sur les traitements lexicaux, leur restriction à l'étude de l'adulte invite à s'intéresser au lecteur débutant pour savoir comment ces caractéristiques idiosyncrasiques influencent l'apprentissage de la lecture.

2. Variations interlangues et analyse syntaxique

Depuis plusieurs décennies, et peut-être de façon encore plus importante que dans le champ des caractéristiques orthographiques des langues, de nombreuses recherches se sont intéressées à la diversité structurale des langues et de leurs conséquences sur les stratégies de compréhension (et de production) tant chez l'adulte que chez l'enfant. A l'instar du modèle probabiliste et fonctionnaliste de Bates et MacWhinney (1989), elles se sont attachées à tester la valeur heuristique des notions de validité et de poids des indices envisagés dans le modèle de compétition. Les paragraphes suivants présenteront quelques unes de ces recherches, dans l'objectif de rendre compte des différences entre certaines langues sur les traitements interprétatifs des phrases et leur mise en place. Là encore, il sera nécessaire de rappeler les différences entre l'oral et l'écrit avant d'appliquer au domaine de la compréhension écrite, et avec toutes les précautions nécessaires, des conclusions de recherches sur la compréhension orale.

La plupart des recherches développées dans le cadre du modèle de compétition, ont utilisé la même méthode d'investigation qui consiste à faire écouter des phrases transitives aux participants, qui doivent en déterminer l'agent. La tâche est généralement, pour les adultes, une simple tâche de décision, le temps de réponse étant également pris en compte, et pour les enfants un mime de l'action. Les indices que comportent ces phrases (ordre des mots, caractère animé ou inanimé des noms, marques flexionnelles, thématization ou topicalisation³³, accentuation etc.) sont mis, soit en convergence, soit en compétition, de façon à obtenir des phrases grammaticales ou semi-grammaticales³⁴.

³³ Par thématization, on entend l'usage d'un article indéfini ou défini accompagnant un nom, la présence de l'article défini faisant alors référence à un nom ayant fait l'objet d'un discours préalable (Kail, 1986).

³⁴ Par exemple en français, dans la phrase « le **chien** lèche un bâton », l'ordre, le caractère animé du premier nom et sa thématization par l'article défini ainsi que l'accentuation concourent pour le choix de « chien » comme agent. Par

Un premier groupe de recherches (Bates, Devescovi, & D'Amico, 1999; Bates, McNew, MacWhinney, Devescovi, & Smith, 1982) a permis de valider, chez l'adulte, l'impact de la validité et du poids des indices lors de l'assignation du rôle d'agent en anglais (américain) et en italien. Les résultats obtenus par Bates *et al.* (1982) montrent que l'ordre des mots et le caractère animé des noms sont deux indices valides dans ces deux langues. Toutefois, les stratégies des Américains et des Italiens concernant l'utilisation de ces deux types d'indices sont très différentes : chez les Américains, l'ordre est le facteur essentiel alors que chez les italiens, c'est l'information lexico-sémantique sur le caractère animé ou inanimé des noms qui l'emporte. Par ailleurs, les Italiens font un plus grand usage du thème et de l'accentuation ce qui signifie qu'ils doivent davantage s'appuyer sur d'autres facteurs que l'ordre des mots car l'italien informel permet des combinaisons d'ordre beaucoup plus variées que l'anglais, dominé par un ordre fixe. Bates *et al.* (1999) confirment cette dominance de l'ordre en anglais par rapport à l'italien sur l'interprétation de phrases complexes incluant une relative et en englobant des indices comme l'accord sujet/verbe. Lorsque l'ordre et l'accord sujet/verbe sont en compétition, l'ordre l'emporte en anglais, et les informations flexionnelles l'emportent en italien.

Dans une autre recherche concernant l'anglais, l'italien et l'allemand, MacWhinney, Bates et Kliegl (1984) retrouvent les mêmes dominances en anglais et en italien et approfondissent le rôle de la disponibilité, de la fiabilité et de l'applicabilité des indices pour comprendre leur ordre d'importance. En effet, les Italiens sont en particulier sensibles à l'accord marqué par le verbe car l'italien tolère largement l'omission du sujet, de sorte que la flexion verbale est plus souvent disponible. Cependant, cet indice n'est pas totalement fiable et peut donner lieu à ambiguïté. Dans de tels cas, les locuteurs auraient alors recours à certaines combinaisons du caractère animé, de l'ordre et de l'accent. En allemand en revanche, l'accord casuel de l'objet avec le verbe est l'indice le plus valide car dans cette langue, les variations de l'ordre des mots sont importantes³⁵, alors que les marques flexionnelles casuelles apparaissent avec une haute occurrence. Mais ces dernières peuvent parfois être insuffisantes car identiques pour des cas différents (*e.g.* l'article « *die* » pour le nominatif et l'accusatif des noms féminin). D'autres indices peuvent alors être utilisés, notamment le caractère animé ou inanimé des noms³⁶.

opposition, dans la phrase « un bâton lèche le **chien** », le caractère animé, la thématization et l'accentuation désignent « chien » alors que l'ordre des mots privilégie « bâton » comme agent).

³⁵ Cependant, l'ordre des mots en allemand est déterminé par des critères grammaticaux, par exemple la présence d'une proposition subordonnée, les relations entretenues entre les éléments de la principale et de la relative ou l'utilisation d'un pronom clitique ou d'un verbe auxiliaire). Par ailleurs, dans de nombreuses phrases le premier nom est généralement le sujet, ce qui peut se révéler un indice utile pour lever certaines ambiguïtés.

³⁶ Les auteurs avancent que l'accord verbal est moins fiable en allemand qu'en italien car les marques de conjugaisons pour distinguer les personnes sont plus ambiguës.

Les résultats obtenus chez l'adulte en anglais et en italien ont également été vérifiés auprès d'enfants âgés de 2;6 ans à 5;6 ans (Bates, MacWhinney, Caselli, Devescovi, Natale, & Venza, 1984). Dès le plus jeune âge et dans les deux langues, les stratégies des enfants sont dirigées selon la validité des indices dans la langue, c'est-à-dire par l'ordre des mots en anglais et par les indices lexico-sémantiques en italien (comme le caractère animé des noms).

Les recherches de Michèle Kail ont toutefois souligné l'intérêt de prendre en considération la notion de coût (Kail, 2000, 1999) pour rendre compte des procédures d'accès et des temps de traitement nécessaires pour assigner les fonctions aux éléments de la phrase. Le coût engendré par les indices serait en fait dépendant du caractère local ou topologique de ces derniers. Dans une étude portant sur le français, l'anglais et l'italien, chez des adultes et des enfants de même âge que dans celle de Bates *et al.* (1984), Kail (1986, 1989) observe que les adultes français fondent leurs stratégies interprétatives sur le contraste animé des noms plus que sur l'ordre. Ils se comportent ainsi comme les italiens, et s'opposent aux anglo-saxons, dont les stratégies sont toujours gouvernées par l'ordre. Ces résultats ne reflètent donc pas la validité de cet indice, alors que le français est, comme l'anglais une langue à ordre relativement fixe. Les profils des enfants anglophones et italophones sont identiques à ceux des adultes. La compréhension des enfants français est en revanche dominée par l'ordre à partir de 3 ans et demi. Les enfants francophones s'apparentent ainsi davantage aux enfants anglophones, tout en faisant plus d'usage de l'information lexico-sémantique que ces derniers et plus d'usage de l'ordre que les enfants italiens. L'auteur en déduit le statut intermédiaire du français, langue dans laquelle plusieurs indices valides n'ont pas forcément le même coût de traitement. Les indices locaux tels que les informations lexico-sémantiques seraient moins coûteux que les indices topologiques comme l'ordre, ce qui explique leur privilège chez les adultes. En ne faisant référence qu'à un seul élément bien localisé, les indices locaux peuvent conduire directement à l'assignement des rôles sans entraîner une charge mnémorique importante. Au contraire, les indices topologiques (ou relationnels), exigeant un traitement de l'ensemble de l'environnement phrastique, sont plus contraignants et donc moins économiques car ils ne peuvent être utilisés qu'une fois que l'ensemble des informations dispersées dans la phrase est disponible. En terme développemental, les changements orientés vers la prédominance des traitements les plus économiques (*i.e.* ceux des informations lexico-sémantiques locales), aux dépens des traitements topologiques, s'effectueraient par le biais d'une réorganisation du système de compréhension liée à l'évolution des capacités de traitement (Kail, 1989). En effet, l'adulte est en mesure de relever les conflits pouvant exister entre différents indices et se focaliser sur ceux se révélant les plus fiables et les plus économiques (en français, l'utilisation des pronoms clitiques viole l'ordre canonique Sujet-

Verbe-Objet et en atténue la validité en autorisant une plus grande variabilité de l'agencement des mots). En revanche, les jeunes enfants ne parviendraient pas à percevoir certaines violations syntaxiques car moins confrontés aux phrases complexes (par exemple utilisant des pronoms clitiques).

Le statut différent des indices locaux vs. topologiques ainsi que les hypothèses développementales sur le français ont pu être confirmés par une étude interlangue français-espagnol (Kail, 1989; Kail & Charvillat, 1988).

L'importance des indices locaux a notamment été mise en évidence dans les langues à morphologie riche comme le hongrois Pléh (1989), le turc (Slobin & Bever, 1982) ou les langues sémitiques. Concernant ces dernières, les travaux en hébreu ont souligné l'importance de la disponibilité de certains indices morphologiques dès le début de l'acquisition et la prédominance progressive de leur fiabilité (Sokolov, 1989). Frankel et ses collaborateurs ont étudié le rôle de l'ordre des mots, relativement libre en hébreu, ainsi que celui de deux types d'indices morphologiques plus locaux : la particule grammaticale de l'accusatif placée devant un nom défini et l'accord sujet/verbe. Dans une première expérience menée auprès d'adultes et d'enfants âgés de 3 à 11 ans, Frankel, Amir, Frenkel et Arbel (1980) montrent qu'en cas de compétition avec l'ordre des mots, les locuteurs de tout âge se basent sur les indices morphologiques pour interpréter les phrases. Les données d'une seconde expérience (Frankel & Arbel, 1981) recueillies auprès d'enfants de 4, 6, 8 et 10 ans ont permis d'estimer plus précisément la hiérarchie et le poids respectif de ces trois mêmes indices, au cours du développement. Premièrement, à tous les âges, le marquage de l'objet l'emporte sur les autres indices. Deuxièmement, avec l'âge, l'effet de l'ordre diminue alors que l'effet des indices morphologiques, en particulier la particule de l'objet direct augmente. Toutefois, et contrairement aux adultes, l'importance prise par l'accord sujet/verbe n'est jamais supérieure à celle de l'ordre.

Les auteurs interprètent ces résultats en termes de disponibilité et de moindre coût du marquage de l'objet. En revanche, l'accord sujet/verbe serait un indice moins local mais surtout moins valide en début d'apprentissage du fait de ses nombreuses irrégularités, et nécessiterait plus de temps pour être intégré. Enfin, l'ordre des mots n'est certes pas l'indice le plus pertinent, ni le plus économique en hébreu, mais est toutefois particulièrement détectable, ce qui expliquerait son poids chez les plus jeunes (Sokolov, 1989).

Les quelques études menées en langue arabe corroborent l'importance des aspects morphologiques pour l'interprétation des phrases. Taman (1993) étudie, auprès d'adultes, le rôle de plusieurs indices dans les stratégies d'interprétation de phrases simples en arabe standard,

présentées oralement et à l'écrit. Il constate que l'indice le plus valide est l'accord sujet/verbe, suivi de près par le marquage casuel, ces deux indices étant de loin plus valides que le caractère animé des noms, pour assigner le rôle d'agent. Taman (1993) interprète ce résultat par la régularité et la « simplicité » des désinences verbales en arabe. Cette étude montre par ailleurs qu'en l'absence de ces indices, les participants se basent sur l'ordre des mots pour orienter leur choix, ce qui confirme les résultats de Frankel *et al.* (1980) et Frankel et Arbel (1981) en hébreu. Selon les auteurs, bien que l'ordre des mots soit en arabe, encore plus variable qu'en hébreu, il peut toutefois apporter des informations pertinentes dans bon nombre de cas. L'utilisation de cet indice en dernier recours est nommée par cet auteur « stratégie dormante ». A l'aide d'un paradigme *on-line* appliqué au traitement de phrases complexes, l'étude de Bamhamed (1996) complète les données sur la prévalence des flexions verbales en arabe en précisant que les marques préverbaux (*i.e.* les flexions de la conjugaison arabe préfixées) sont traitées plus rapidement que les marques flexionnelles discontinues (conjointement préfixées et suffixées comme au féminin et au pluriel), leur position les rendant plus perceptibles au cours même du traitement de la phrase.

Globalement, ces recherches semblent révéler, encore une fois, la spécificité de l'anglais, qui accorde particulièrement d'importance à l'agencement des mots dans la phrase. Les aspects morphologiques constitueraient, dans cette langue, des indices moins décisifs pour assigner les fonctions sémantiques des éléments, notamment du fait de la pauvreté du système de conjugaison verbale (MacWhinney *et al.*, 1984). Par contraste, dans les langues à morphologie riche, les marques flexionnelles s'avèrent être un excellent indice de traitement. Cependant, plusieurs détails empiriques invitent à nuancer la dominance systématique de certains indices par rapport à d'autres en considérant, la complexité des phénomènes syntaxiques et sémantiques et les multiples interactions qui les sous-tendent. Parmi les caractéristiques pouvant affecter la prépondérance de tel ou tel indice, nous pouvons rappeler celles des contraintes procédurales (Kail, 1991, 1999). La facilité cognitive à prendre en considération certains indices peut ainsi dépendre de leur régularité, de leur forme continue ou dispersée, de leur redondance avec d'autres indices, de leur perceptibilité, etc. Cependant, il faut souligner que certaines de ces caractéristiques diffèrent beaucoup entre l'oral et l'écrit. L'une des différences les plus notables concerne sans doute les indices morphologiques, qui dans certains cas et dans certaines langues correspondent à des morphogrammes grammaticaux, marques graphiques ne correspondant à aucune réalité orale (sauf dans le cas particulier de liaisons).

Ainsi, la présence et l'absence de certaines informations perceptives, ainsi que le caractère plus complexe et moins « tolérant » de l'écrit, constituent des arguments permettant de supposer que

les indices les plus importants lors de la compréhension ne sont pas strictement identiques à l'oral et à l'écrit. En outre, la distance entre l'oral et l'écrit est particulièrement exacerbée dans les langues se caractérisant par la diglossie comme l'arabe. La diglossie arabe et les différences entre les langues orales et la langue écrite seront plus particulièrement développées dans le premier paragraphe du chapitre suivant mais il convient dès à présent de porter attention à cette particularité qui interroge les études précédentes sur le caractère écologique de présenter oralement une langue surtout écrite.

Enfin, l'investigation des stratégies d'interprétation s'est surtout focalisée sur les phrases simples et la fonction d'agent, mettant au premier plan le rôle central des informations véhiculées par le verbe (MacWhinney, 1989). Ces premiers constats devraient être enrichis par des recherches portant sur des structures plus complexes et axées sur d'autres fonctions sémantiques, qui, bien que présentes, restent toutefois moins abondantes.

Chapitre 3 : Les caractéristiques des systèmes écrits des langues arabe et portugaise

Le chapitre précédent a notamment permis de mettre en évidence l'existence de variations interlangues dans l'acquisition de la reconnaissance des mots écrits, l'appréhension de leur structure morphologiques et la mise en place des stratégies d'analyse syntaxique lors de la compréhension. Ces variations, dépendantes des caractéristiques idiosyncrasiques des langues, ne peuvent ainsi être envisagées sans considérer précisément et de façon multidimensionnelle les systèmes linguistiques étudiés. La langue arabe, comme l'hébreu, revêt des caractéristiques morphologiques spécifiques et très contrastées par rapport aux langues indo-européennes. Ces contrastes ne se limitent pas à la morphologie. Pourtant, il est rare que l'arabe soit pris en compte dans les recherches véritablement comparatives. Il en est de même pour le portugais. Néanmoins, cette langue a des caractéristiques intéressantes, notamment dans le cadre de cette étude qui envisage également l'impact des variations interlangues sur l'apprentissage du français langue seconde. Ainsi, nous nous attacherons, dans un premier temps, à décrire d'une part le système d'écriture de la langue arabe, et d'autre part celui de la langue portugaise.

1. La langue arabe

1.1 La situation de diglossie des pays arabes

Lorsque l'on s'intéresse à la pratique de l'écrit en langue arabe, la première caractéristique à signaler concerne la situation de diglossie. Elle est définie par Ferguson (1959) comme une situation où sont en usage deux langues apparentées, génétiquement et structurellement, et dont les distributions fonctionnelles sont complémentaires. Ainsi deux registres coexistent en arabe : l'arabe dialectal et l'arabe littéral ou standard. Ces deux idiomes sont tous deux sémitiques (comme l'hébreu, le maltais, l'amharique ou le tigrinya³⁷ pour ne citer que les langues officielles) mais ne sont pas utilisés dans les mêmes circonstances. L'arabe dialectal, pratiqué dans les conversations quotidiennes est essentiellement oral. Le terme d'arabe dialectal recouvre plusieurs dialectes, qui diffèrent d'un pays à l'autre voire d'une région à l'autre (*e.g.* l'arabe tunisien, marocain, algérois, oranais, libanais, syrien etc.). L'arabe standard, pratiqué dans les situations officielles et formelles est généralement réservé à l'écrit. C'est ainsi la langue de la

³⁷ L'amharique et le tigrinya sont principalement parlés en Éthiopie et en Erythrée.

littérature, des médias, de la culture en général³⁸ et généralement de la scolarisation³⁹ (Ammar, 1997; Baccouche, 2003). Il est en outre, identique, à quelques variations près, dans l'ensemble des pays arabes.

Du fait de son caractère usuel et utilitaire, l'arabe dialectal est évolutif et emprunte de nombreux vocables aux langues étrangères (en particulier au français, à l'italien, au castillan), alors que le lexique de l'arabe standard est beaucoup plus figé. Or, même si l'arabe dialectal s'est constitué un lexique propre, de nombreux mots sont toutefois assez similaires entre les deux registres, du fait de leurs bases linguistiques communes. Néanmoins, de nombreuses différences formelles subsistent. Le système phonologique de l'arabe dialectal est largement altéré par rapport à celui de l'arabe standard. Ces altérations portent en particulier sur les voyelles (changement de timbre, suppression ou ajout, transformation de leur longueur). De nombreuses marques flexionnelles (désinences verbales, existence de la forme duelle) disparaissent, ce qui limite les possibilités d'agencement des mots dans la phrase, plus diverses en arabe standard. La syntaxe de l'arabe dialectal est aussi simplifiée par rapport à celle de l'arabe standard (Baccouche, 2003).

La majorité des auteurs vont ainsi jusqu'à les considérer comme deux langues différentes, l'arabe standard correspondant finalement à une langue étrangère et la maîtrise des deux registres s'apparentant de fait au bilinguisme (Baccouche, 2003; Ibrahim, 2006; Saiegh-Haddad, 2004, 2005). Les relations qu'elles peuvent entretenir sont complexes. L'influence qu'elles exercent l'une sur l'autre favorise, en outre, l'émergence de niveaux intermédiaires pouvant s'échelonner selon un continuum, allant de l'arabe classique ancien, langue du Coran et des textes religieux, très archaïque, aux langues orales les moins standardisées. Celles, usant de nombreux mélanges et d'emprunts à d'autres langues, peuvent alors être définies comme des *sabirs* (Ben Rejeb, 2000 cité par Snoussi, 2003 ; Garmadi, 1968 cité par Ammar, 1997).

En ce qui concerne les apprentissages, il est généralement affirmé par les chercheurs que la langue orale interfère sur l'acquisition de la langue écrite (Ammar, 1997; Saiegh-Haddad, 2003ab, 2004, 2005) Ainsi, la distance, notamment phonologique, entre l'arabe dialectal et standard est souvent perçue comme étant à l'origine de nombreuses difficultés d'apprentissage de la lecture/écriture (Ayari 1996 ; Feitelson, Goldstein, Iraqi & Share, 1993 ; Snoussi, 2003)⁴⁰. Quoi qu'il en soit, le système linguistique de l'arabe standard, dont la maîtrise sous-tend l'acquisition de la lecture (Abu-Rabia, 2000), sera au centre de notre problématique.

³⁸ Il faut toutefois signaler que dans les productions audiovisuelles et les arts populaires, l'arabe standard n'est pas forcément dominant.

³⁹ Toutefois, pour des raisons historiques liées à la colonisation, de nombreux pays arabes utilisent d'autres langues (comme le français ou l'anglais) dans l'enseignement secondaire ou supérieur.

⁴⁰ Quelques études récentes dans le domaine des connaissances morphologiques cherchent pour leur part à rendre compte de l'articulation entre le registre tunisien et standard de façon plus positive (Boukadida & Besse, 2007 ; Najem & Besse, 2006).

1.2 Le système écrit de l'arabe standard

1.2.1 Caractéristiques orthographiques

L'arabe standard⁴¹ se caractérise par un système alphabétique orienté de droite à gauche et exclusivement écrit en caractères cursifs, sans majuscules. L'alphabet arabe se compose de 28 lettres (29 si l'on compte la hamza -ء-, qui est plus ambiguë⁴²) qui ne s'associent jamais en bigrammes ou trigrammes. La plupart d'entre elles, change de forme selon leur position dans le mot (initiale, médiane, finale) et suivant les règles d'attachement de la lettre qui la précède : 6 lettres (ا, د, ذ, ر, ز, و) se singularisent, en effet, par une absence de lien graphique à la lettre suivante, de sorte que même au milieu d'un mot, une lettre peut s'écrire sous sa forme initiale ou isolée. Certaines de ces lettres ne se distinguent que par des points infra ou supra linéaires (e.g. ب, ت, ث). On trouvera, dans le tableau 2, les lettres de l'alphabet, leur forme et leur transcription phonétique.

Tableau 2 : Liste des lettres de l'alphabet arabe

⁴¹ Dans la suite du texte, et pour simplifier, l'usage du terme arabe ou langue arabe renverra systématiquement et exclusivement à l'arabe standard.

⁴² Le statut de la hamza est particulier car graphiquement, elle nécessite souvent un support (apporté par les lettres ا, و ou ع) et se place alors au dessus ou en dessous de celui-ci (Blachère & Godefroy-Demombynes, 2000). Son statut est alors celui de signe diacritique. Cependant, contrairement aux autres signes diacritiques, il ne peut être omis.

Graphie				Nom (transcription phonétique adoptée par la Société Asiatique de Paris)	Transcription phonétique (Alphabet Phonétique International)
isolée	initiale	médiane	finale		
ء	أ, إ, و, ئ			hamza	[ʔ]
ا	—		ا	`alif	[a]
ب	ب	ب	ب	bā'	[b]
ت	ت	ت	ت	tā'	[t]
ث	ث	ث	ث	tā'	[θ]
ج	ج	ج	ج	ǧim	[ǧ]
ح	ح	ح	ح	ḥā'	[h]
خ	خ	خ	خ	ḫā'	[x]
د	—		د	dāl	[d]
ذ	—		ذ	ḏāl	[ð]
ر	—		ر	rā'	[r]
ز	—		ز	zāy	[z]
س	س	س	س	sīn	[s]
ش	ش	ش	ش	šīn	[ʃ]
ص	ص	ص	ص	ṣād	[sʰ]
ض	ض	ض	ض	ḏād	[dʰ], [ðʰ]
ط	ط	ط	ط	tā'	[tʰ]
ظ	ظ	ظ	ظ	zā'	[zʰ], [ðʰ]
ع	ع	ع	ع	`ayn	[ʔʰ]
غ	غ	غ	غ	ǧayn	[ɣ]
ف	ف	ف	ف	fā'	[f]
ق	ق	ق	ق	qāf	[q]
ك	ك	ك	ك	kāf	[k]
ل	ل	ل	ل	lām	[l]
م	م	م	م	mīm	[m]
ن	ن	ن	ن	nūn	[n]
ه	ه	ه	ه	hā'	[h]
و	—		و	wāw	[w]
ي	ي	ي	ي	yā	[j]

Outre ces lettres, la langue arabe utilise deux autres formes graphiques particulières, n'apparaissant qu'en fin de lemme : le 'alif maqṣûra (ﻻ ; [-α]), sorte de 'alif courbé qui indique, comme lui, un prolongement de la voyelle [-α] ; le tā' marbûṭa (ﺕ : [-α] ou [-ατ5]) dérivé de la lettre tā' (ﺕ) : porté exclusivement par des substantifs, il correspond souvent à la marque du féminin. En cas d'ajout d'un suffixe, sa forme est identique au tā' en position médiane.

Cet alphabet, essentiellement consonantique, peut être complété de signes diacritiques, placés au-dessus ou en dessous des lettres. Les principaux sont :

- la chadda « ّ » marquant le doublement de la consonne ou semi-consonne qui le porte,
- le soukoun « ْ » correspondant à l'absence de voyelle,
- la fatha « َ » codant la prononciation de la voyelle brève [A],
- la kasra « ِ » codant la prononciation de la voyelle brève [ɪ],
- la damma « ُ » codant la prononciation de la voyelle brève [ʊ].

Cependant, ces signes diacritiques sont facultatifs puisque la norme d'écriture employée communément ne note, à quelques rares exceptions près, que le squelette du mot, formé des lettres consonantiques. Toutefois, pour comprendre ce phénomène de suppression d'informations visuelles, et en particulier d'une majorité des éléments vocaliques, une description des systèmes phonologiques et morphologiques de l'arabe standard est nécessaire.

1.2.2 Caractéristiques phonologiques

Le système phonémique de l'arabe se caractérise par sa richesse consonantique et sa pauvreté vocalique (Cohen, 2007). Il existe en effet 26 phonèmes consonantiques en arabe, codés strictement par une et une seule lettre de l'alphabet dans une correspondance biunivoque. Sans reprendre l'ensemble des phonèmes utilisés, nous pouvons noter, par rapport aux langues indo-européennes, comme le français ou le portugais, l'absence de quelques phonèmes comme [p] et [v], mais surtout, une abondance de consonnes impliquant des articulateurs arrières (e.g. la pharyngale [ħ], la gutturale [ʔ^h], les emphatiques [s^h], [t^h], [d^h], [z^h] ; Baccouche (2003). Par ailleurs, deux lettres (wāw و [w] et yā ي [j]) sont nommées semi-consonnes. Chacune d'elles peut ainsi apparaître dans les deux seules diphtongues de l'arabe [aw] et [aj], comme consonne, ou encore comme le signe d'un allongement vocalique.

Comme il en a été question précédemment, l'arabe emploie seulement trois timbres vocaliques ([α], [ɪ] et [ʊ]) orthographiés, lorsque le système d'écriture est complet (i.e. vocalisé), par des signes diacritiques discrets. Or, ces trois voyelles peuvent être allongées, on les appelle alors des

voyelles longues. Celles-ci auront une existence graphique rendue respectivement par l'écriture des lettres **ا**, **ي** et **و** à la suite de celle correspondant à l'attaque consonantique. En effet, les syllabes commencent obligatoirement par une consonne et sont dominées par l'alternance consonne voyelle. Le système syllabique de l'arabe standard est ainsi réduit à quelques structures syllabiques du type CV⁴³, CVV, CVC et dans quelques cas CVVC (Cohen, 2007; Baccouche, 2003).

L'aspect discret et même déficitaire des voyelles par rapport aux consonnes, tant phonologiquement que graphiquement, est à mettre en relation avec les caractéristiques morphologiques spécifiques aux langues sémitiques et la richesse de cette dimension dans la production lexicale.

1.2.3 Caractéristiques morphologiques

Le système morphologique de l'arabe se caractérise par une structure bidimensionnelle composée d'une racine et d'un schème.

La racine⁴⁴ est exclusivement composée de consonnes (en général trois). Cette séquence trilitère code une signification générale et abstraite (par exemple, la suite **كتب** [kʌβ] signifie ce qui se rapporte à l'écrit). La notion de racine trilitère est loin d'être admise comme un principe absolu par les grammairiens arabes (Blachère & Gaudefroy-Demombynes, 2000; Dichy, 2006; Larcher, 2003). Elle a, malgré tout, une grande importance en arabe. C'est d'ailleurs en fonction de la racine que les mots sont traditionnellement classés dans les dictionnaires. Cependant, le squelette consonantique isolé n'a aucune existence lexicale. La formation des mots du lexique arabe nécessite donc obligatoirement l'ajout, ou plus exactement l'intrication d'un schème à la racine.

Le schème (ou thème pour Blachère & Gaudefroy-Demombynes, 2000; Cohen, 2007) contient des informations sémantiques et syntaxiques particulières. Son application fournit ainsi une forme lexicale et grammaticale aux mots en complétant le sens générique de la racine. Il s'agit d'un pattern dérivationnel interne principalement composé de voyelles. C'est donc le schème qui

⁴³ Abréviations de Consonne-Voyelle, les suivantes référant aussi à ces deux mêmes termes.

⁴⁴ Malgré les différences entre les définitions données par les dictionnaires (Arrivé, Gadet, & Galmiche, 1986 ; Dubois, Giacomo, Guespin, Marcellisi, Marcellisi, & Mevel, 1999 ; Gardes-Tamine, 1998), il convient de distinguer la racine du radical, correspondant généralement à la manifestation d'une racine ou encore à une de ses formes, prise une fois ôtées les affixes additionnels (e.g. cant- dans « cantatrice », chant- dans « chanteur ») et de la base, correspondant à l'élément lexical ou non, servant à former d'autres mots (e.g. « irréalisable » est construit sur la base de « réalisable » lui-même construit sur la base « réalis- »). Ces définitions sont issues des conceptions de Gardes-Tamine (1998).

véhicule les signes diacritiques absents en écriture non vocalisée. Il peut également contenir des consonnes et des voyelles longues, marquées graphiquement⁴⁵.

Pour désigner ces schèmes, les grammairiens arabes emploient, comme modèle, la racine **فعل** ([fɛsʕl]) existant en arabe⁴⁶ (qui signifie l'idée d'agir, de faire) ; ils reprennent aussi la distinction essentielle en arabe entre la catégorie verbale et nominale (Cohen, 2007).

Ainsi, dans l'exemple 5 du tableau 3, le nom **كاتب** ([kA]tɪβ) ; celui qui écrit) est formé par l'application du schème nominal du participe actif, désignant celui qui fait l'action, à la racine **كتب** ([kɪβ]).

Tableau 3 : Illustration de la structure morphologique des mots arabes à partir de la racine [kɪβ] كتب

Mot	Transcription phonétique (API)	Traduction	Schème appliqué ¹	Valeur ²	Type
1 كَتَبَ	[kɑtɑβɑ]	il a écrit	فَعَلَ [-α-α-α]	Forme simple (inaccompli)	
2 كَتَّبَ	[kɑttɑβɑ]	il a fait écrire	فَعَّلَ [-α- -α-α]	Forme II (valeur factitive)	Schémes verbaux
3 كَاتَبَ	[kɑ]tɑβɑ]	il a écrit à quelqu'un	فَاعَلَ [-α]-α-α]	Forme IV (valeur transitive)	
4 كِتَابَةٌ	[kɪtɑ]βɑt5]	écriture	فِعَالَةٌ [-i-α]-αt5]	Nom d'action	
5 كَاتِبٌ	[kɑ]tɪβ]	écrivain	فَاعِلٌ [-α]-i-]	Participe actif	Schémes nominaux
6 مَكْتَبَةٌ	[mɑkɑtɑβɑt5]	bibliothèque, librairie	مَفْعَلَةٌ [mɑ-°-α-αt5]	Lieu de l'action	
7 كِتَابٌ	[kɪtɑ]β]	livre	فِعَالٌ [-i-α]-]	Nom permettant l'action	
8 كُتُبٌ	[kɪtɪβ]	livres	فُعُلٌ [-v-v-]	Pluriel	Schémes flexionnels ³
9 كُتِبَ	[kɪtɪβɑ]	Il a été écrit	فُعِلَ [-v-i-α]	Passif (inaccompli)	

¹ Dans la transcription phonétique du schème, les tirets représentent l'emplacement des consonnes de la racine.

² La valeur sémantique présentée est, pour certains schèmes, celle la plus appropriée dans ce contexte d'application.

³ Bien que cette dénomination ne fasse pas consensus, elle évite l'amalgame avec les schèmes purement dérivationnels.

Comme l'illustre le tableau 3, il existe de nombreux schèmes. Les schèmes verbaux sont peu nombreux et relativement prévisibles. Les grammairiens comptent ainsi quatorze formes verbales dérivées ou « augmentées », mais le nombre de schèmes les plus fréquents est plus restreint (moins d'une dizaine). Chaque forme exprime des modalités verbales différentes telles que l'intensivité, la réciprocité, le but, le factitif, la réflexivité, qui peuvent être plus ou moins nuancées. Toutefois, l'usage de la langue fait qu'il est délicat de dégager une valeur univoque à ces schèmes (Neyreneuf & Al-Hakkak, 1996).

⁴⁵ Les consonnes sont placées, dans le schème, en position finale ou initiale, ce qui n'est pas le cas des voyelles longues internes.

⁴⁶ La transcription des schèmes, utilisée dans la traduction occidentale et représentant les consonnes de la racine par les codes C-C-C ou R¹-R²-R³ etc., ne sera pas employée ici.

Les schèmes nominaux sont, en revanche, très nombreux (plusieurs centaines) et moins prévisibles (Baccouche, 2003). En effet, les nombreuses variations pour exprimer une même idée complexifient la possibilité de les dénombrer. Toutefois, il est possible d'en isoler plusieurs dont la valeur est relativement fixe. On retrouve ainsi systématiquement les participes actifs et passifs, les noms d'action, régulièrement les noms d'instruments, de lieu, etc. (Blachère & Gaudefroy-Demombynes, 2000 ; Kouloughli, 1994).

Il existe également, en nombre très minoritaire, d'autres schèmes, qui, bien qu'ils n'appartiennent pas à proprement parler à la morphologie dérivationnelle, peuvent être considérés comme les vecteurs de la structure interne du mot (Mel'cuk, 1997). C'est le cas des pluriels internes et du passif.

Cette structuration en racine consonantique/schème vocalique concerne la quasi-totalité du lexique, hormis les particules. Les exceptions sont aussi, souvent des mots d'origine étrangère bien que la dichotomie racine/schème a parfois fini par s'y appliquer. Le démontre par exemple le terme philosophe « فَلَاسُوف » auquel s'applique un schème pluriel interne (فَلَاسِيف).

Outre le fait d'être une des rares langues à morphologie non concaténative, l'arabe comporte de nombreuses marques morpho-flexionnelles (Blachère & Gaudefroy-Demombynes, 2000 ; Kouloughli, 1994) généralement préfixées ou suffixées.

La conjugaison verbale est relativement pauvre comparée à celle des langues indo-européennes (Blachère & Gaudefroy-Demombynes, 2000). Le verbe arabe n'a en effet que deux paradigmes de base de conjugaison distinguant non pas le temps, mais l'aspect (accompli, inaccompli). Les formes verbales composées n'existent pas non plus en arabe. Ainsi, avec le mode impératif et les trois modes de l'inaccompli, six modèles de conjugaison se distinguent. Quelques variations sont à noter, selon l'irrégularité et/ou la forme augmentée des verbes. Par ailleurs, pour chaque conjugaison, les flexions préfixées ou suffixées apportent généralement des informations sur le genre (masculin ou féminin) et le nombre (singulier, duel ou pluriel) des trois personnes distinguées (1^e, 2^e, 3^e), de sorte qu'une conjugaison peut comporter jusqu'à 13 formes différentes. Les flexions portées par les noms et les adjectifs (et également par de nombreux pronoms) peuvent être des marques de genre (féminin ou masculin), de nombre (singulier, duel et éventuellement pluriel si celui-ci n'est pas un pluriel interne), de cas (sujet, direct, indirect) et de détermination. Ces marques, exclusivement des suffixes, s'intègrent les unes aux autres, nécessitant parfois quelques transformations. Pour donner un exemple, le féminin des noms se construit souvent par l'ajout du ta' marbûta (ة) à la fin du nom masculin (e.g. أمير [ʔɑmɪr]pov =

un prince ; أميرَة [ʔαμι]ρατυν]= une princesse). Mais sa forme graphique ou phonique peut être transformée (e.g. أميرتُه [ʔαμι]ρατυηυ]= sa princesse).

Dans cette abondance de marques flexionnelles, certaines révèlent l'originalité morpho-syntaxique de la langue arabe : la présence de flexions verbales préfixées, l'existence du duel, de flexions casuelles et de plusieurs formes fléchies indiquant la détermination. Là encore plusieurs d'entre elles prennent la forme de signes diacritiques absents à l'écrit standard.

1.2.4 Quelques caractéristiques syntaxiques

A cause de ces multiples formes fléchies et leur rôle syntaxique, l'ordre des éléments dans la phrase semble avoir relativement moins d'importance que dans d'autres langues (Baccouche, 2003 ; cf. paragraphe 1.1.1). Cependant, l'ordre canonique des mots dans les phrases verbales est l'ordre VSO⁴⁷ (يَسْكُنُ كَرِيمٌ فِي تُونِسَ [φασκυνυ κარი]μ φι] τυ]νισ] = habite Karim à Tunis) mais il n'est pas strict, et peut varier selon certaines règles (usage d'un pronom clitique), ou de façon à mettre en valeur certains constituants (Blachère & Gaudefroy-Demombynes, 2000 ; Kouloughli, 1994). Cet ordre a un impact assez complexe sur l'accord verbe/sujet dans les phrases conjuguées à la 3^e personne. En effet, le verbe ne s'accorde alors pas en nombre mais seulement en genre avec un cas particulier lorsque le sujet est un nom inanimé pluriel. Notons, par ailleurs, l'omission régulière des pronoms personnels sujets, limitant la possibilité d'identifier le genre et le nombre de l'agent aux seules marques verbales.

Compte tenu du fait que le temps est un élément sous-estimé dans les conjugaisons verbales, l'arabe requiert de nombreuses particules pour exprimer ses nuances, de même que pour exprimer la négation. Ces particules qui se fixent généralement devant le verbe, peuvent parfois en modifier la conjugaison. Ce phénomène d'agglomération n'est d'ailleurs pas spécifique aux particules temporelles. Par exemple, certaines prépositions s'accolent également au mot qui les suit. Il en est de même pour l'article défini ال ([ʔαλ]), qui a aussi la particularité d'être invariable, de sorte que seules les marques flexionnelles des noms ou adjectifs et les accords entre les constituants permettent, par exemple, d'identifier le genre ou le nombre.

Ainsi, la possibilité d'accoler des éléments grammaticaux aux mots, alliée à l'isolement de certaines lettres à l'intérieur de ceux-ci, rend compte d'un découpage lexical très différent des langues à écriture latine⁴⁸ (e.g. سَيَسْكُنُ بِالشَّمَالِ = il va habiter au nord⁴⁹).

⁴⁷ Abréviation Verbe - Sujet – Objet.

⁴⁸ Il faut toutefois noter que l'espacement « intra-mots » n'est pas aussi marqué que l'espacement « inter-mots ».

Notons enfin qu'il existe d'autres types de phrases simples en arabe, nommées phrases nominales, qui ne contiennent pas de verbe.

Globalement, il convient de souligner la richesse mais aussi la complexité des règles syntaxiques de la langue arabe sans omettre la très grande régularité de leur application.

1.3 Écriture vocalisée *versus* non vocalisée

Plusieurs éléments issus des descriptions précédentes permettent d'appréhender l'ambivalence du système d'écriture arabe, pouvant s'écrire avec ou sans signes diacritiques. Compte tenu du fait que ces signes codent essentiellement les vocalisations, l'on parle généralement d'écriture vocalisée ou non vocalisée. L'écriture vocalisée, utilisée en début d'apprentissage est orthographiquement très transparente⁵⁰. L'écriture non vocalisée, utilisée comme norme standard, se caractérise en revanche par son opacité. Le système dérivationnel se scindant en une racine consonantique et un schème vocalique, fait en effet apparaître de nombreux homographes sémantiquement et phonologiquement ambigus en écriture non vocalisée (par exemple, **كتب** peut correspondre à [καταβα] = il a écrit, [κατταβα] = il a fait écrire, [κτυβ] = livres ou [κτυβα] = il a été écrit). De même, les flexions casuelles et d'autres éléments morphosyntaxiques sont absents, limitant très largement les informations disponibles pour identifier les mots et leurs fonctions. Le système non vocalisé est introduit progressivement dans l'enseignement scolaire. En Tunisie, pays où ont eu lieu les expérimentations d'une partie de cette étude, il est introduit à partir de la 5^e année de l'enseignement de base, correspondant au niveau CM2 en France (Belajouza & Snoussi, 2006).

Plusieurs recherches en arabe et en hébreu ont été menées sur l'effet de cette absence de vocalisation lors de la lecture. Les conclusions des auteurs, n'allant pas toujours dans le même sens, invitent ainsi à nuancer l'appui que peut constituer les signes diacritiques en fonction du niveau d'expertise.

Plusieurs recherches ont mis en exergue l'effet facilitateur de la vocalisation et le recours au contexte pour palier l'« orthographe déficitaire » en version non vocalisée. Chez des enfants en 2^e et 6^e année de scolarisation, la compréhension écrite est améliorée lorsque les signes vocaliques sont présents (Abu-Rabia, 1999). Bouchafa & Magnan (2000) corroborent ces résultats sur la lecture de mots auprès d'enfants de 1^e année. Leur étude est, à notre

⁴⁹ L'analyse grammaticale des deux entités de la phrase est la suivante : 1) Particule du Futur/verbe « habiter » 3^e personne du singulier 2) préposition/article/nom « nord ».

⁵⁰ Quelques opacités subsistent toutefois, notamment l'assimilation du lām **ل** ([λ]) de l'article et du 'alif final **ا** non prononcé dans la conjugaison de l'accompli 3^e personne du masculin pluriel.

connaissance, la seule à aborder également le traitement des semi-consonnes et leur statut ambigu en arabe. L'analyse des erreurs réalisées montre la difficulté à les identifier comme voyelles longues. Les enfants ont tendance à les transformer en consonnes, permettant ainsi une correspondance graphophonologique plus transparente.

Abu-Rabia (1996, 1997ab, 1998, 2001) confirme cet effet facilitateur de la vocalisation, chez les jeunes adultes, avec plusieurs types de matériel (lecture de mots, de paragraphes, compréhension de textes). Par ailleurs, cet auteur observe à plusieurs reprises l'importance que joue le contexte lors de la lecture (Abu-Rabia, 1996, 1997ab) notamment en version non vocalisée (Abu-Rabia 2001). Ainsi, la vocalisation semble bien avoir un effet bénéfique, même chez le lecteur expert.

Cependant, d'autres résultats révèlent que cette assistance devient rapidement isolée, utilisée de manière complémentaire, dans un second temps, et essentiellement pour résoudre les ambiguïtés. En effet, de nombreuses recherches en hébreu n'ont pas mentionné une telle contribution des signes vocaliques en lecture de texte chez les adultes (Eshel, 1985; cité par Share & Levin, 1999; Navon & Shimron, 1985 cité par Shimron, 1999) Certaines ont même échoué à la faire apparaître chez des enfants de 2^e et 4^e année (Even, 1995; cité par Share & Levin, 1999).

Dans une recherche menée en langue arabe, Ammar (1997, 2002, 2003) s'oriente également vers l'hypothèse que les lecteurs ont rapidement tendance à délaissé les voyelles. Les résultats obtenus à l'aide d'une épreuve d'identification de mots écrits présentant des perturbations orthographiques, font apparaître d'une part que, dès la 1^e année d'apprentissage, les perturbations des voyelles diacritiques sont mal détectées et d'autre part que, dès la 2^e année, les perturbations sur les consonnes sont mieux détectées en arabe vocalisé que non vocalisé. Ces derniers résultats appuient ainsi l'idée que les signes diacritiques peuvent devenir source de difficultés. Certains chercheurs expliquent un tel phénomène par le fait que l'écriture non vocalisée nécessite une compensation par le contexte et privilégie les procédures top-down (Shimron, 1999). Quoi qu'il en soit, le rôle respectif de la vocalisation, du contexte et leur interaction, ainsi que le statut particulier des semi-consonnes mériteraient d'être étudiés de façon plus exhaustive au cours des apprentissages.

2. La langue portugaise

Le portugais est la langue officielle de plusieurs pays dits lusophones et subit de ce fait des variations importantes. Les grammaires portugaises distinguent ainsi deux normes linguistiques, celle du portugais du Brésil et celle du portugais européen, en vigueur au Portugal et dans les pays lusophones d'Afrique et d'Asie (Teyssier, 1992). Les différences entre le portugais

brésilien et européen concernent tous les aspects de la langue. Cette diversité ne remet toutefois pas en cause l'unité de la langue portugaise (Cortez Gomes, 1997; Teyssier, 1992). Ainsi, nous opterons pour la norme brésilienne, puisque c'est au Brésil qu'ont eu lieu les expérimentations. L'acception générique du terme portugais ou langue portugaise sera toutefois adoptée de manière à mettre en valeur l'unité de cette langue.

D'autre part, dans les pays d'Afrique et d'Asie où il est pratiqué, de même que dans certaines régions du Brésil et du Portugal, le portugais est en contact avec d'autres langues. Ce contact a favorisé la créolisation de la langue portugaise, à tel point que le portugais créole est une langue reconnue dans certains pays d'Afrique (Cap Vert, São Tomé et principe, Guinée-Bissau), qui connaissent des situations de diglossie. Cependant, ce phénomène reste relativement circonscrit et ne correspond pas à la situation linguistique majoritaire du Brésil et du Portugal.

Un bref aperçu descriptif du système écrit du portugais permettra de mettre en évidence les différences avec le système arabe et les relations de proximité avec les autres langues romanes notamment le français.

2.1 Le système écrit du portugais

2.1.1 Caractéristiques orthographico-phonologiques

Le portugais utilise, comme le français, l'alphabet latin. Le système de correspondance graphophonologique de ces deux langues est donc grossièrement assez similaire. Cependant, quelques précisions peuvent être apportées.

L'alphabet du portugais ne comporte, normalement, que 23 lettres, car les lettres K, W et Y n'en font pas partie. Elles sont toutefois employées pour transcrire des mots étrangers et leurs dérivés portugais (*wagneriano* = wagnérien). Les redoublements se limitent aux lettres S (de façon à obtenir, comme en français, le phonème [s] entre deux voyelles) R (qui devient grasseyé entre deux consonnes) et éventuellement C. Notons également l'utilisation de plusieurs digrammes, dont certains sont les mêmes qu'en français (*e.g.* <ch> se prononçant toujours [ʃ]) mais d'autres non. On trouve ainsi en portugais le digramme <lh> [x] et <nh> [ɲ], ce dernier étant de même valeur que le digramme <gn> français, cette séquence se prononçant pour sa part toujours [ɲv] en portugais. De façon beaucoup plus systématique qu'en français, la lettre H en situation isolée est muette (Cantel, 1999) et la seule que l'on peut clairement répertorier comme telle en brésilien. Mais, si le portugais présente très peu de lettres muettes, plusieurs graphèmes sont inconsistants. Parmi les exemples, témoignant en outre de différences par rapport au français, nous pouvons citer les graphèmes :

- T et D prononcés [τ] et [δ] mais pouvant aussi être palatalisés (*i.e.* prononcés [τσ] et [δz]),
- L correspondant au phonème [λ] lorsqu'il n'est pas dans un digramme mais qui subit une forte vocalisation en fin de syllabe ([ω] par exemple dans *Brasil*).
- X correspondant selon les cas aux consonnes ou groupes consonantiques [Σ] (*abacaxi* [αβαχσι] = ananas), [ζ] (*exercício* [ιζεχορσισφυ] = exercice) ou [κσ] (*táxi* [υτακσι] = taxi) (Teysnier, 1992).

Le système phonétique du portugais se caractérise, d'une part, par sa richesse et sa complexité vocalique. Les voyelles représentent, en effet, dans le discours, environ 47% des phonèmes (Morais-Borbosa, 2007). On y trouve une abondance de diphtongues et de voyelle nasalisées. Cette nasalisation est souvent notée graphiquement à l'aide du tilde « ~ » suscrit sur les lettres A et O (*mãe* [υμã] = mère ; *corações* [κωρãσ] = cœurs). Cependant, la structure syllabique des mots portugais reste simple, comparée à celle d'autres langues comme l'anglais (Seymour *et al.*, 2003).

Le portugais est une langue à forte accentuation tonique, à la fois en durée et en intensité (Cantel, 1999). Dans les mots polysyllabiques, l'accent tonique peut porter sur la dernière ou l'avant dernière syllabe selon la terminaison du mot, mais aussi sur la syllabe antépénultième. La syllabe accentuée peut être graphiquement marquée à l'aide d'un signe diacritique suscrit (un accent aigu ou circonflexe), notamment dans le cas où l'accentuation fait exception aux règles des terminaisons. L'accent tonique a une importance fondamentale en portugais. En effet, si l'on en fait abstraction, de nombreux mots deviennent homographes et inintelligibles (*e.g.* *cântara* [κãνταρα] = cruche, *cantara* [κνταρα] = il avait chanté, *cantará* [κνταρα] = il chantera).

Ainsi, l'écriture du portugais complète le système alphabétique de plusieurs marques diacritiques indispensables pour prononcer et lire correctement les mots, mais contrairement à l'arabe, ces éléments font partie intégrante du système d'écriture et ne peuvent être éludés.

2.1.2 Caractéristiques morphologiques

Comme cela est apparu dans le domaine graphophonologique, le portugais partage avec le français de nombreuses similitudes morphologiques. Ainsi, le système de construction interne des mots portugais est concaténatif. Les mots morphologiquement complexes sont composés

d'une base ou d'un radical auquel s'affixe un ou des morphèmes liés. Ces affixes peuvent être en portugais des préfixes (accolés avant la base) ou des suffixes (accolés après la base).

Bechara (2002) explique, d'une part qu'en portugais, le processus de dérivation lexicale s'effectue généralement à l'aide de radicaux de type latin et parfois de radicaux spécifiques au portugais, lorsque ces derniers reflètent une évolution propre à la langue, d'autre part que les principaux préfixes sont d'origine latine et grecque. Il en recense plus d'une soixantaine. L'application des suffixes apparaît en revanche plus complexe à définir et moins régulière. Plus de 150 suffixes sont répertoriés dans son ouvrage, dont de nombreux homophones et homonymes. Parmi ces affixes, et même parmi les bases, une ressemblance peut être trouvée avec la construction de mots en français (*in/corrigível* = incorrigible ; *en/terrar* = enterrer ; *re/ssaltar* = ressortir ; *dent/ista* = dentiste ; *sorvet/eria* = sorbetière ; *folh/agem* = feuillage). Ainsi, on retrouve dans ces deux langues de nombreux morphèmes identiques (e.g : *trans-*, *sub-*...) mais d'autres ont subi des modifications plus ou moins importantes ou sont étymologiquement différents (e.g. *sobre-* = *super-* en français ; *-inha* : diminutif⁵¹).

Concernant la morphologie flexionnelle, les principes sous-jacents aux règles de construction sont comparables à celles du français ou d'autres langues indo-européennes. Les marques flexionnelles du portugais sont toujours suffixées. Celles portées par les verbes peuvent apporter des informations sur le temps (passé, présent, futur), le mode (indicatif, impératif, subjonctif, infinitif), l'aspect (inconclu, conclu, habituel etc.) ainsi que la personne (1^e, 2^e et 3^e) et son nombre (singulier, pluriel) (Morais-Borbosa, 2007). Pour une même conjugaison, les diverses désinences sont toutefois moins nombreuses qu'en arabe (six au total), du fait de l'absence du duel et d'une même forme fléchie au masculin et au féminin.

En ce qui concerne les flexions nominales et adjectivales, il est possible de remarquer la relative productivité du marquage de genre « o/a » et la systématisme de la marque « -s » du pluriel, flexions que l'on retrouve également sur les articles et de nombreux pronoms (*um/uma/uns/unas* = un/une/des. ; *cujol/cuja/cujos/cujas* = dont) (Morais-Borbosa, 2007 ; Teyssier, 1992)

En portugais, contrairement au français les marques flexionnelles sont généralement perceptibles à l'oral (e. g : « Ele colhe uma bela flor » il cueille une jolie fleur ; « Eles colhem belas flores » ils cueillent des jolies fleurs). Le portugais se caractérise donc par une absence de morphogrammes.

⁵¹ Les grammaires portugaises soulignent la disposition du portugais à utiliser les suffixes augmentatifs ou diminutifs. Ces nuances affectives peuvent ainsi apparaître dans de nombreux morphèmes.

2.1.3 Quelques caractéristiques syntaxiques

La langue portugaise est une langue à ordre relativement fixe avec comme structure canonique la séquence SVO (Cortez Gomes, 1997). Elle n'échappe également pas aux règles d'accord sujet/verbe (*os gatos morreram* = les chats moururent) et article/nom/adjectif (*estas gatas pequenas* = ces petites chattes). Elle accepte favorablement l'omission des pronoms personnels sujets, les désinences verbales donnant généralement suffisamment d'informations pour identifier l'agent (Cantel, 1999). Il existe en portugais plusieurs temps composés ou périphrases verbales. Parmi les plus utilisés, se trouvent :

- le passé composé (*pretérito perfeito composto*) exprimant la continuité d'un évènement jusqu'à un moment présent. Il est formé à l'aide d'un seul auxiliaire (*ter*) et du participe passé, qui est, dans ce cas, toujours invariable (*e.g. Maria tem saído* = Marie est sortie).

- la forme *estar* + gérondif (ou participe présent, également invariable), exprimant le déroulement d'une action en cours (*e.g. estou escrevendo* = je suis en train d'écrire).

Enfin, notons la présence de formes contractées, telles que l'intégration d'un pronom à une préposition (*e.g. de + ele = dele*, de lui, *com + mim = comigo* avec moi).

Chapitre 4 : Dimensions phonologique, morphologique, syntaxique et lecture : perspective interlangue

1. Phonologie et lecture

1.1 Préliminaire

Pour apprendre à lire, il faut avoir compris le principe alphabétique, compréhension qui suppose la capacité à se représenter et à segmenter la structure phonologique de la parole. C'est en cela qu'intervient la conscience phonologique (*phonological awareness* en anglais) ou compétence métaphonologique, définie comme « *la capacité à identifier les composants phonologiques des unités linguistiques et à les manipuler de manière délibérée* » (Demont & Gombert, 2007 p. 51; Gombert, 1990 p. 112). Sans cette conscience, la connaissance totale du code graphophonologique ne peut s'élaborer. Il n'est alors pas étonnant de constater la richesse de la littérature consacrée à ces compétences segmentales, plus particulièrement à la segmentation en phonèmes. Toutefois, l'avancée de la recherche dans ce domaine ne s'est pas faite sans s.

Le premier point ayant donné lieu à discussion est la nature même de la relation entre lecture et conscience phonologique. Il vient d'être affirmé que la conscience phonologique est un prérequis de l'apprentissage de la lecture mais il faut souligner qu'elle peut également être considérée comme une de ses conséquences (pour une revue sur la question voir Demont & Gombert, 2007 ou Gombert, 1990,1992). L'argument supportant le mieux cette position est que la segmentation phonémique n'aurait aucune utilité avant que l'enfant ne soit confronté à l'écrit (Morais, Alegria, & Content, 1987)⁵². Toutefois, à l'instar des nombreuses études longitudinales (*e.g.* Demont & Gombert, 1996; Duncan, Seymour, & Hill, 1997) ou d'entraînements phonologiques (*e.g.* Byrne, Fielding-Barnsley & Ashley, 2000 ; Kjeldsen, Niemi & Olofsson, 2003)⁵³, le rôle facilitateur de ces compétences ne peut être nié. Le lien entre développement des compétences métaphonologiques et apprentissage de la lecture s'apparenterait donc davantage à une relation circulaire (Demont & Gombert, 2007).

⁵² Cet argument a notamment été confirmé par des recherches, menées chez les adultes analphabètes et les enfants prélecteurs, montrant leur difficultés, par exemple lors de tâche de suppression de phonème ou de segmentation libre (*e.g.* Alegria & Morais, 1979 ; Morais, Cary, Alegria & Bertelson, 1979, Lukatela, Carello, Shankweiler & Liberman, 1995 ; Martinot & Gombert, 1996).

⁵³ Cités par Demont et Gombert (2007).

Le second point, à l'origine de débats plus récents mais non sans relation avec la précédente question, concerne la taille de l'unité phonologique prise en compte au cours de l'apprentissage de la reconnaissance des mots écrits. En se fondant sur la structure hiérarchique de la syllabe⁵⁴, Goswami (1993, 2000, 2002) avance qu'une sensibilité aux unités de grandes tailles, telles que la syllabe, l'attaque et la rime, précède la conscience des unités plus petites comme le phonème et s'avère être un bon prédicteur des habiletés ultérieures en lecture. L'autre conception s'appuie sur les résultats de plusieurs recherches qui ont échoué à faire apparaître le rôle de la rime mais soutiennent que la conscience phonémique est le meilleur prédicteur en début d'apprentissage (e.g. Duncan *et al.*, 1997; Hulme, Hatcher, Nation, Brown, Adams, & Stuart, 2002; Hulme, Muter, & Snowling, 1998; Muter, Hulme, Snowling, & Taylor, 1997). Cette conception affirme ainsi que la progression part du phonème pour atteindre des unités de plus en plus larges (*cf.* le modèle de la double fondation de Seymour paragraphe 1.1.2 du chapitre 1). L'instruction des correspondances graphème-phonème permet l'apparition de la conscience phonémique. Au bout d'un ou deux ans, l'évolution dans la maîtrise du système écrit favorise l'émergence de la conscience à des unités plus larges, de façon à traiter des régularités plus composites mais aussi pertinentes pour prononcer les mots (Duncan, Seymour, & Hill, 2000; Seymour & Duncan, 1997).

Ces points de vue ne sont en fait pas contradictoires si l'on tient compte de la distinction entre apprentissage implicite et explicite (*cf.* paragraphe 1.3 du chapitre 1). La progression des larges unités vers les plus petites témoigne de l'évolution de la sensibilité implicite à traiter d'abord les unités les plus saillantes (la syllabe et la rime) puis les plus difficiles à isoler comme le phonème. La progression inverse, des unités minimales vers des unités plus composites, relève des apprentissages explicites, s'orientant vers une analyse de plus en plus complexe des séquences graphophonologiques (la rime, la syllabe ou les morphèmes) (Demont & Gombert, 2007 ;Gombert *et al.*, 1997).

Si beaucoup de recherches ont surtout étudié le rôle du phonème et de la rime, d'autres se sont également intéressées à la syllabe, justement car le lexique se compose également de mots plurisyllabiques et que cette unité, naturellement perceptible à l'oral, est un constituant articulatoire simple. Certaines études ont ainsi, d'une part confirmé qu'il est plus facile de manipuler les syllabes que les rimes et les phonèmes (Duncan, Colé, Seymour, & Magnan, 2006; Treiman & Zukowski, 1996) et d'autre part constaté que cette capacité peut également intervenir

⁵⁴ Les psycholinguistes envisagent en effet plusieurs niveaux de décomposition syllabique : une décomposition minimale en phonèmes, et des niveaux intermédiaires en attaque / noyau vocalique / coda (e.g. [kɔp/ɪ/πtɛ]) et en body / rime (e.g. [kɔp/ɪπtɛ]). A la conjonction entre ces deux derniers niveaux, la syllabe peut se segmenter en deux parties distinctes : l'attaque et la rime (e.g. [kɔp/ɪπtɛ]).

comme une unité de traitement lors de la reconnaissance (Colé, Magnan, & Grainger, 1999; Doignon & Zagar, 2006). Cependant, le rôle de la syllabe a souvent été controversé (e.g. Mann, 1984). Ainsi, certains auteurs considèrent que son rôle est limité (Castles & Coltheart, 2004) cette unité ne semblant pas supplanter l'importance du phonème, tout au moins en début d'apprentissage (voir Castles & Coltheart, 2004 pour une revue ; Aidinis & Nunes, 2001 en grec ; Demont & Gombert, 1996 en français; Gonzalez & Ortiz Gonzalez, 2000 en espagnol ; etc.).

Ainsi, sans forcément parler de compétition, il s'avère possible que plusieurs unités soient prises en compte lors des traitements phonologiques du langage écrit.

1.2 Les recherches interlangues

Les études tenant compte des différences interlangues ont apporté une contribution notable à la question de la taille des unités phonologiques. Plusieurs conclusions importantes semblent notamment s'en dégager :

- Tout d'abord, les résultats obtenus dans différentes langues tendent à montrer que le développement des connaissances phonémiques évolue plus rapidement lorsque les langues sont transparentes (pour une revue voir Durgunoğlu & Oney, 1999; Ziegler & Goswami, 2005). Durgunoğlu et Oney (1999) indiquent par exemple que la segmentation phonémique (de même qu'en syllabes) de lecteurs débutants turcs est plus développée que celle de leurs pairs anglais.
- Par ailleurs, les données recueillies dans une multitude de langues, transparentes ou non, appuient l'hypothèse de l'universalité du rôle de la conscience phonémique, (e.g. Caravolas *et al.*, 2005; Cossu, Shankweiler, Liberman, & Katz, 1988; Durgunoğlu & Oney, 1999). Néanmoins, l'étude comparative anglais-allemand de Mann et Wimmer (2002) invite à considérer que dans les langues transparentes, la contribution des connaissances phonémiques serait plus faible et diminuerait plus rapidement, du fait de la facilité à maîtriser un système graphophonologique régulier.
- Enfin, l'utilité de la segmentation en attaque rime semble également variable en fonction du degré de transparence du système orthographique. Goswami, Ziegler et Richardson (2005), ont comparé la saillance perceptive de la rime en anglais et en allemand auprès d'enfants prélecteurs et lecteurs débutants de 1^e année. A l'aide d'une épreuve implicite de détection d'intrus (e.g. détecter le mot qui est différent entre *house* [ηαΥσ] / *mousse* [μσΥσ] / *kiss* [κισ]), les auteurs examinent si les enfants sont sensibles au partage de la rime et du phonème vocalique. Les résultats indiquent que, de façon beaucoup plus marquée chez les enfants anglais, le jugement sur la rime est supérieur à celui sur le phonème. Par ailleurs, les enfants anglais obtiennent de

meilleures performances, en particulier lorsqu'il s'agit de détecter une rime consistante (*i.e.* dont l'orthographe est univoque).

Dans le même ordre d'idée, Goswami *et al.* (1998), dans leur comparaison anglais-français-espagnol, montrent notamment que, dans les langues dont l'orthographe est la moins transparente (*i.e.* en anglais et en français), la lecture est facilitée lorsque les pseudo-mots ont beaucoup de voisins de rime. En espagnol, la lecture de pseudo-mots, qu'ils aient ou non une rime familière, ne pose aucune difficulté. Des résultats similaires ont été obtenus en comparant des langues opaques versus transparentes, comme l'anglais et l'allemand (Goswami, Ziegler, Dalton, & Schneider, 2001, 2003) ou l'anglais et le grec (Goswami *et al.*, 1997). Cela suggère que dans les langues les moins transparentes, l'existence de voisin de rime peut inciter les lecteurs débutants à utiliser un décodage en unités de grande taille plutôt qu'un décodage graphémique. Par ailleurs, l'importance de la rime, attestée en anglais, tiendrait au fait que dans cette langue, le degré de consistance graphophonologique est plus élevé au niveau de la rime qu'au niveau du phonème, les irrégularités affectant en particulier les voyelles (Goswami, 2000). La saillance de cette unité serait également renforcée par le fait que 90 % du lexique anglais est composé de mots monosyllabiques (Goswami *et al.*, 2005).

Pour conclure, il semble possible d'avancer que le caractère transparent *versus* non transparent des langues n'affecte pas simplement la mise en place des procédures de décodage (*cf.* paragraphe 1.1, chapitre 2). Son impact concerne aussi le traitement des unités phonologiques de la langue et la possibilité d'utiliser des stratégies de décodage différentes selon le statut de ces unités.

Logiquement, toutes ces recherches se sont essentiellement centrées sur la composante du décodage lors de la lecture. La conscience phonologique, essentielle pour parvenir à identifier les mots dans toutes langues alphabétiques, aurait également un pouvoir prédictif non négligeable sur les habiletés de compréhension de phrase ou de texte (Caravolas *et al.*, 2005; Demont & Gombert, 1996; Gonzalez & Ortiz Gonzalez, 2000; Nocus & Gombert, 1997).

1.3 Les données en langue arabe

Les recherches menées en langue arabe sur les compétences phonologiques confirment le rôle primordial de la conscience phonémique lors de la reconnaissance des mots écrits. Par ailleurs certains chercheurs ont tenté de déterminer le statut d'autres unités telles que la syllabe et la rime. Mais à ce sujet, les preuves empiriques validant l'importance de la segmentation syllabique et l'absence de sensibilité à la rime sont encore limitées.

Concernant les compétences métaphonémiques, plusieurs recherches tendent à démontrer leur émergence et leur importance en début d'apprentissage de la lecture :

Ammar (1997) évalue les compétences phonologiques d'enfants tunisiens prélecteurs et bons ou faibles lecteurs de 1^e, 2^e et 3^e année, à l'aide d'épreuves de comptage, de segmentation et de suppression de phonème consonantique et de syllabe. Il constate les points suivants :

- Les enfants prélecteurs et en difficulté obtiennent de faibles performances en suppression phonémique
- Chez les apprentis lecteurs, les taux de réussite à cette tâche sont relativement équivalents pour les mots de l'arabe standard, de l'arabe dialectal ou pour des pseudo-mots.
- En 1^e année, la manipulation des phonèmes consonantiques s'avère meilleure que dans d'autres langues comme l'anglais.
- Par rapport aux non lecteurs, les lecteurs rencontrent des difficultés à isoler les syllabes de mots dans une tâche de comptage
- Avec l'avancée dans la scolarité, seules les compétences à manipuler les phonèmes augmentent
- Pour l'ensemble des élèves, les corrélations les plus importantes entre les habiletés phonologiques et d'identification de mots vocalisés⁵⁵ ou de lecture de mots et de pseudo-mots vocalisés concernent la suppression phonémique.

Ces observations permettent de confirmer la relation étroite et réciproque entre conscience phonémique et lecture vocalisée en arabe standard.

Par ailleurs, Abu-Rabia (1995), dans une étude auprès d'enfants palestiniens scolarisés de la 1^e à la 5^e année avance que les règles de conversion graphème-phonème s'avèrent particulièrement corrélées à la reconnaissance des mots vocalisés en arabe.

Abu-Rabia, Share et Mansour (2003) montrent pour leur part que, parmi diverses habiletés mesurées (orthographiques, syntaxiques, morphologiques, mémoire de travail, processus visuels), le déficit le plus sévère chez les lecteurs de 5^e année en grande difficulté concerne la

⁵⁵ Adaptation en arabe de l'épreuve collective d'identification de mots écrits de Khomsi (1994).

conscience phonémique. Les résultats obtenus pour ce groupe à une tâche de suppression phonémique sont en effet proches de zéro.

Ez-Zaher (2004) confirme l'importance de la conscience phonémique en arabe mais cherche également à la nuancer en tentant d'établir l'importance du rôle de la syllabe. A l'aide d'analyse de régression, il observe en 1^e année une contribution importante des compétences phonologiques aux habiletés initiales en lecture. La suppression de phonème apparaît être le meilleur prédicteur, mais les compétences en segmentation syllabique apportent également une contribution plus modique mais non négligeable. Entre le début et la fin de la 1^e année, les compétences phonologiques ont augmenté. Pour la suppression phonémique, cette augmentation serait, selon l'auteur, plus minime que dans d'autres langues alors que, pour la segmentation syllabique, elle serait en revanche supérieure. L'auteur rappelle que les unités graphiques en arabe correspondent le plus souvent à la syllabe, du fait du marquage diacritique des voyelles. Cette spécificité contraint selon lui, le développement des capacités métaphonémiques et privilégie, même à l'écrit, une analyse en syllabes.

Il semble que les propositions d'Ez-Zaher (2004) doivent être discutées. Son constat d'une faible augmentation des compétences phonémiques en arabe va à l'encontre du postulat de Ziegler et Goswami (2005) relatif à une évolution plus rapide des connaissances phonémiques lorsque les langues sont transparentes. Il se situe aussi en forte contradiction avec les résultats d'Ammar (1997). Ce dernier avait conclu, qu'avec l'expérience de l'écrit, les représentations liées au code grapho-phonémique dominant lors des opérations d'analyse phonologique. Cette procédure s'avère peu fructueuse lorsqu'il s'agit de manipuler les unités syllabiques. Cependant, bien que dans la recherche d'Ammar (1997) le lien entre lecture et conscience phonémique reste prépondérant, les corrélations entre la manipulation explicite de la syllabe et la lecture de mots sont élevées en particulier pour la lecture de mots non vocalisés et la lecture de pseudo-mots vocalisés. Ainsi, le rôle de la syllabe en arabe reste à démontrer plus précisément.

De plus, Ammar (1997, 2002, 2003) s'appuie sur différents résultats de ses recherches (*cf.* encart 6) pour avancer que les enfants arabophones ne seraient pas sensibles à la rime, car la structure bidimensionnelle en racine consonantique/schème vocalique serait incompatible avec le découpage attaque/rime. Par ailleurs qu'ils auraient précocement tendance à délaissé les voyelles au profit de la structure morphologique des mots.

Encart 6 : Les résultats de Ammar (1997) en faveur d'une décomposition en racine consonantique/schéma vocalique.

Les résultats obtenus à l'épreuve d'identification de mots ayant subi des perturbations orthographiques ou à des épreuves de conscience phonologique par des enfants de 1^e, 2^e, 3^e, 4^e, et 6^e année permettent à Ammar (1997) de constater les points suivants :

- Dès la 1^e année, lors de l'identification de mots, les perturbations portant sur les voyelles diacritiques sont mal détectées.
- Dès la 2^e année, les perturbations portant sur les consonnes sont plus faciles à détecter en version non vocalisée que vocalisée.
- Les scores de conscience phonémique sont faibles lorsqu'il s'agit de manipuler les phonèmes vocaliques (segmentation, comptage), et même plus faibles que les résultats trouvés en anglais.
- Les performances en suppression de phonèmes ne sont pas influencées par la position du phonème à supprimer, contrairement aux résultats trouvés par Rosner et Simon (1971)⁵⁶ en anglais.

L'assemblage phonologique en langue arabe s'effectuerait ainsi en deux cycles : les enfants décodent dans un premier temps les consonnes et si les ressources cognitives sont suffisamment disponibles ils décodent secondairement les voyelles (Ammar, 2002, 2003; Belajouza & Snoussi, 2006). La construction morphologique des mots arabes aurait ainsi un impact important sur les traitements phonologiques. Le rôle de la morphologie en arabe sera plus précisément envisagé dans le sous-chapitre suivant.

A l'issue de ces recherches, nous constatons que l'évolution du poids de la relation entre conscience phonémique et lecture n'a pas été précisée. A ce sujet, les recherches en hébreu soutiennent que très tôt dans l'apprentissage les corrélations entre la conscience phonémique et la lecture en version vocalisée sont plus faibles que dans une langue non transparente comme l'anglais. La maîtrise du décodage serait atteinte dès la fin de la 1^e année car l'hébreu vocalisé est une langue transparente et de structure syllabique relativement simple. La conscience phonémique apparaît donc moins cruciale pour lire (Share & Levin, 1999).

Par ailleurs, aucune recherche sur la lecture en arabe vocalisé n'a, à notre connaissance, étudié simultanément l'impact des connaissances segmentales à des unités subsyllabiques de petites et de grandes tailles. Les résultats récents d'Elbeheri et Everatt (2007) apportent toutefois quelques éléments de compréhension en s'intéressant à l'arabe non vocalisé. Les résultats d'élèves égyptiens de 4^e et 5^e année indiquent une relation relativement faible entre les habiletés de suppression de phonème et de détection d'intrus de rime d'une part, et la lecture de mots non vocalisés présentés en contexte d'autre part, mais une relation importante entre suppression de phonème et lecture de pseudo-mots. La détection de rime est pour sa part, une des épreuves les moins corrélées à cette dimension de décodage. La prise en compte de la rime chez les lecteurs déjà avancés ne semble ainsi pas s'établir, mais de telles conclusions doivent être avancées avec

⁵⁶ Cités par Ammar (1997).

précaution. En effet, Elbeheri et Everatt (2007) n'indiquent pas les taux exacts de signification des corrélations, ni ne précisent les caractéristiques des mots utilisés dans les tâches phonologiques.

L'hypothèse que la rime ne serait pas pertinente mais que le body pourrait être une unité infralexicale plus saillante dans les langues sémitiques, a été envisagée dans plusieurs recherches sur l'hébreu (pour une revue voir Share & Levin, 1999). La recherche de Share et Blum (2005) est particulièrement intéressante à citer : elle montre que les enfants d'âge préscolaire comme de 2^e année ont plus de facilité à segmenter les syllabes CVC en body/coda (CV/C) qu'en attaque/rime (C/VC). Selon les auteurs, cette préférence ne serait pas seulement le produit de l'exposition au système orthographique mais serait plutôt inhérent aux aspects phonologiques de l'hébreu. Cependant, leur analyse chez les plus jeunes indique que ni la décomposition en attaque/rime, ni celle en body/coda n'apparaît liée aux premières habiletés en lecture de mots vocalisés.

Pour conclure, bien qu'en arabe, l'importance de la conscience phonémique soit bien établie, l'évolution de son rôle et la référence éventuelle à des constituants plus composites sont moins claires. Aussi, il nous semble que l'investigation des connaissances phonologiques en relation avec la lecture vocalisée doit être approfondie.

1.4 Les données en langue portugaise

Les résultats recueillis en langue portugaise sur les compétences phonologiques et leur contribution à la lecture en début d'apprentissage sont moins sujets à discussions. Plusieurs travaux réalisés au Brésil se sont focalisés sur les phases initiales de l'acquisition de la lecture dans une perspective longitudinale ou de remédiation.

Les chercheurs ont clairement confirmé que la conscience phonémique se développe lors de la confrontation avec la langue écrite, alors que les unités telles que la syllabe peuvent plus facilement faire l'objet d'une analyse chez les prélecteurs (Manrique & Signori, 1988 ; Régo, 1984⁵⁷). Réciproquement que les connaissances phonologiques jouent incontestablement un rôle essentiel dès le début de l'apprentissage de la lecture.

Maluf et Barrera (1997) ont, ainsi, analysé les réponses d'enfants en école maternelle lorsqu'on leur demande de trouver le mot qui commence ou termine comme un mot cible et de justifier leur réponse. Dans cette épreuve, trois mots sont présentés, mais un seul partage la syllabe initiale ou

⁵⁷ Cités par Maluf et Barrera (1997).

finale avec un mot cible. Ces auteurs montrent d'une part que cette habileté phonologique augmente avec l'âge et d'autre part que, le cas échéant, les justifications données par les enfants de 4 ans et de 6 ans sont très différentes : en 1^e année de maternelle, la prise en compte de la syllabe domine, alors qu'en dernière année, niveau où commence l'initiation à la langue écrite, la prise en compte du phonème domine. Cela suggère que l'analyse phonémique se met en place lorsque les connaissances sur la langue écrite commencent à se développer.

De plus, Cardoso-Martins (1994) confirme qu'il est pratiquement impossible pour des enfants de niveau préscolaire de supprimer la consonne initiale de mots bisyllabiques. Les résultats à cette tâche augmentent en dernière année de maternelle et atteignent des scores très élevés en 1^e année de primaire. Morais *et al.* (1979), avaient, de la même façon, mis en évidence, de telles difficultés chez des adultes portugais analphabètes.

Guimarães (2002) observe que même au bout de 3 ou 4 ans d'apprentissage, les lecteurs en difficulté sont plus faibles que les enfants appariés en âge ou en niveau de lecture/écriture⁵⁸ à différentes tâches de conscience phonologique dont celle de suppression de phonème initial.

Capovilla et Capovilla (2000) se sont pour leur part attachés à évaluer l'impact d'un entraînement phonologique auprès d'enfants de 1^e année issus de milieu socio-économique défavorisé. L'entraînement consistait à effectuer des activités de jeux sur différentes unités phonologiques ainsi que sur les mots et les correspondances graphophonologiques. Les résultats obtenus à la suite de cet entraînement montrent son effet bénéfique sur le développement de la conscience phonologique, en particulier phonémique, et sur les habiletés en lecture de mots et de pseudo-mots. Carnio et Dos Santos (2005) de même que Paula, Mota et Keske-Soares (2005) mentionnent des résultats similaires.

Dans une étude longitudinale portant sur les normo-lecteurs, Barrera et Maluf (2003) ont étudié l'influence des connaissances d'unités plus grandes que le phonème, acquises en fin de maternelle, sur l'acquisition de la lecture en 1^e année. Leurs résultats indiquent que les habiletés à identifier la similarité phonologique entre les mots (*i.e.* détecter les deux mots entre trois partageant le même segment phonologique) sont fortement liées ($r = .53$) aux premières connaissances en lecture⁵⁹. Cependant, l'évaluation des connaissances phonologiques englobe,

⁵⁸ Les compétences en lecture de mots et en orthographe sont évaluées par Guimarães (2002) à l'aide du test TDE (Milnitsky Stein, 1994).

⁵⁹ Les enfants étant testés à la phase initiale de l'apprentissage, leur niveau de connaissance en lecture est évalué à l'aide d'une épreuve où sont présentés des mots ou phrases simples en contexte imagé. Les cotations des réponses s'effectuent selon la cohérence entre les réponses et les images, le respect de la segmentation lexicale du texte écrit, la reconnaissance de certaines lettres, les essais de décodage graphophonologique ou la reconnaissance de certains mots écrits.

dans cette recherche, plusieurs unités telles que la syllabe et la rime suprasyllabique, mais leur rôle respectif n'est pas analysé.

Dans une autre étude longitudinale, Cardoso-Martins (1995) s'y est attachée. Si la conscience phonémique (mesurée par une épreuve explicite de suppression du phonème consonantique initial), évaluée en fin de maternelle, reste le meilleur prédicteur du niveau de lecture de mots fréquents et non fréquents l'année suivante, la sensibilité à la syllabe (mesurée par une épreuve de détection d'intrus où un mot parmi trois se distingue par sa syllabe initiale) intervient également, de façon significative et indépendante. Plus spécifiquement, sur la lecture de mots fréquents, mesurée en milieu d'année, Cardoso-Martins (1995) note la contribution de la sensibilité à la rime (mesurée également par une épreuve de détection d'intrus utilisant des triplets de mots) après la prise en considération de la conscience phonémique mais elle disparaît lorsque cette dernière est remplacée par la sensibilité syllabique. Cependant ce résultat n'est guère étonnant compte tenu du matériel. La rime étudiée est en effet la rime suprasyllabique partagée par des mots bi ou trisyllabiques (*e.g.* chupeta / roleta) et non la rime intrasyllabique. Ainsi, une telle unité supérieure à la syllabe n'apporterait rien de plus que le partage syllabique.

A l'issue de ces différentes recherches, l'intervention des connaissances phonologiques et en particulier de la conscience phonémique en début d'apprentissage, ne fait pas de doute. Par ailleurs, l'expérimentation de Cardoso-Martins (1995) a permis de compléter ces résultats en montrant le rôle même faible, des traitements sur des unités plus larges et, selon l'auteur, de nature implicite. Toutefois, ces recherches se focalisent principalement sur la 1^e année d'apprentissage. Or, il s'agirait de déterminer, à plus long terme, comment ces connaissances peuvent intervenir, sachant qu'avec l'augmentation des représentations orthographiques, de plus en plus de mots (notamment les plus fréquents) peuvent être lus sans recours à la médiation phonologique (Pinheiro, 1995).

2. Morphologie et lecture

2.1 Préliminaire

La conscience morphologique ou les compétences métamorphologiques peuvent se définir comme des habiletés à réfléchir (sur) et à manipuler la structure morphémique des mots (Carlisle, 1995). De nombreuses raisons laissent penser que ces connaissances entretiennent des relations soutenues avec l'apprentissage de la lecture (Colé & Fayol, 2000; Kuo & Anderson, 2006).

- Tout d'abord, force est de constater que les systèmes écrits des langues ne se limitent pas à encoder des informations phonologiques mais représentent également la structure morphologique des mots. Cet encodage morphologique (dérivationnel et flexionnel) peut d'ailleurs être en partie responsable de l'absence de correspondances biunivoques entre les graphèmes et les phonèmes. En effet, dans certaines langues, les morphogrammes permettent de coder à l'écrit les relations morphologiques entre les mots (*e.g. gras* → *grasse* ; *lait* → *laitage* ; *acte* → *action*).

- Par ailleurs, les recherches menées chez le lecteur expert affirment l'utilisation des informations morphologiques lors de l'accès aux mots complexes. Ces derniers constituent, une proportion importante du lexique des langues alphabétiques⁶⁰ même celui acquis à l'âge scolaire⁶¹.

- Enfin, les morphèmes ayant des propriétés phonologiques, syntaxiques et sémantiques, la morphologie se trouve intégralement liée aux autres aspects linguistiques. Ainsi, Kuo et Anderson (2006) envisagent que la conscience morphologique peut rendre compte d'une « plus grande capacité métalinguistique générale » que la conscience phonologique ou syntaxique prises isolément.

Pourtant, comparées à la dimension graphophonologique, les recherches sur le développement des connaissances morphologiques et leur contribution à l'apprentissage de la lecture sont relativement moins nombreuses, et ont surtout été développées en anglais. De plus, l'étude des connaissances morfo-dérivationnelles chez l'enfant est plus récente que celle des connaissances morfo-flexionnelles. En fait, le développement et le rôle des connaissances morfo-dérivationnelles dans l'apprentissage de la lecture, ont longtemps été minimisés. De notre point de vue, cela peut encore une fois être rééquilibré, tout au moins en partie, si l'on tient compte du caractère explicite et implicite des connaissances. Enfin, ces dernières années, plusieurs thèses en

⁶⁰ Par exemple, 80 % du lexique français serait composé selon l'approche générative de Rey-Debove (1984).

⁶¹ Nagy et Anderson (1984) estiment que 60 % du vocabulaire anglais appris à l'âge scolaire est morphologiquement composé.

français consacrées à la morphologie dérivationnelle ont notamment permis d'améliorer l'investigation scientifique dans ce domaine (Marec-Breton, 2003; Rocher, 2005; Royer, 2004). A la suite des travaux initiateurs de Berko (1958) et de son célèbre test du « wug », il a été observé à plusieurs reprises que les connaissances des règles morpho-dérivationnelles se développent tardivement et plus lentement que les connaissances des règles morpho-flexionnelles (e.g. Carlisle & Nomanbhoy, 1993; Tyler & Nagy, 1989; voir encart 7 pour un exemple). Il faudrait même attendre le lycée pour que le sens de certains suffixes dérivationnels soit maîtrisé (Nagy, Diakidoy, & Anderson, 1993).

Encart 7 : « Ceci est un wug... »

La tâche de Berko (1958) consiste à demander aux enfants de créer des formes fléchies ou dérivées à partir de pseudo-mots (e.g. ceci est un wug. Maintenant, il y en a un autre, il y a deux __ ? Comment appellerait-on un petit wug, un __ ?). Par exemple, en anglais, les observations de Berko (1958) ont été répliquées auprès d'enfants de 4 à 14 ans par Selby (1972). Si les réponses correctes sont élevées à 4 ans et plafonnent à partir de 12 ans en ce qui concerne la formation d'une majorité de formes fléchies c'est loin d'être le cas pour les formes dérivées dont les productions sont quasi inexistantes avant 8 ans.

Cette différence d'acquisition s'explique généralement par le fait que beaucoup d'affixes dérivationnels sont moins productifs que les affixes flexionnels (Gordon, 1989). Toutefois, cette notion de productivité est souvent confondue avec les possibles altérations phonologiques et sémantiques de la base lors du processus de dérivation (Kuo & Anderson, 2006 ; cf. encart 8 pour plus de détails).

Encart 8 : les facteurs pouvant influencer les résultats aux tests de connaissances morphologiques

Gordon (1989) a élaboré une classification ordonnée des affixes à partir de la facilité des enfants de 5 à 9 ans à résoudre une tâche de décision lexicale :

- le 1^{er} niveau inclut les flexions régulières dont la productivité et l'applicabilité est maximale,
- le 2^e niveau inclut les affixes dérivationnels neutres, c'est-à-dire ne provoquant pas d'altération phonologique de la base⁶²
- le niveau 3 inclut les flexions irrégulières et les affixes dérivationnels non neutres, que l'auteur considère comme les moins productifs et les plus difficiles à appliquer.

D'autres recherches menées principalement sur la morphologie dérivationnelle aux mêmes âges ont observé un ordre d'apparition similaire dans l'utilisation des règles morphologiques (Carlisle, 1988, 1995; Tyler & Nagy, 1989; White, Power, & White, 1989).

Ces recherches, ainsi que d'autres ont également souligné deux facteurs pouvant influencer la résolution des tâches morphologiques (voir par exemple Carlisle, 1995; Carlisle & Nomanbhoy, 1993; Lecocq, Casalis, Leuwers, & Watteau, 1996) :

- un effet de la transparence sémantique : la production de mots dérivés où le jugement de relation morphologique serait significativement plus basse lorsque la dérivation apporte une nuance importante au niveau de la signification (e.g. toile → toilette)
- un effet de la transparence phonologique : de la même façon, la réussite aux épreuves morphologiques serait affectée par le fait que le processus de dérivation altère les caractéristiques phonologiques de la base ou de l'affixe (e.g. avion → aviation ; ouvrir → rouvrir)

⁶² Kuo et Anderson (2006) soulignent que pour ces affixes, la relation sémantique entre la base et le mot dérivé est généralement transparente.

Le fait que les performances aux épreuves morphologiques ne soient pas indépendantes de la structure formelle et sémantique des mots a engagé certain auteur à considérer que cette connaissance ne pouvait être que secondaire et tardive.

Malgré ce décalage dans l'utilisation de la morphologie dérivationnelle, les jeunes enfants (en général à partir de 2½, 3 ans) produisent spontanément des mots dérivés et en inventent à partir de leur connaissance de certaines bases (Clark & Cohen, 1984; Corbin, 1987). Cette capacité à repérer et appliquer spontanément des règles morphologiques ne serait pas strictement réservée à la morphologie flexionnelle⁶³. Bescond (1999) a ainsi observé, à l'aide d'une épreuve de plausibilité lexicale (choisir entre deux pseudo-mots, dont l'un respecte la structure morphologique affixe + base *-préfader-* et l'autre non *-pradéfer-*, lequel ressemble le plus à un mot) que, dès 5 ans, les enfants préfèrent les pseudo-mots préfixés⁶⁴.

Pour résumer, il existerait précocement une sensibilité aux aspects morphologiques flexionnels et dérivationnels, bien qu'elle soit moins développée que les connaissances phonologiques relevant de mécanismes linguistiques moins complexes et encore loin d'être achevée lorsque les élèves débutent l'apprentissage de la lecture (Carlisle, 1995; Rubin, 1988; Smith-Lock & Rubin, 1993). Néanmoins, ce serait sur la base de cette sensibilité et sous l'effet des apprentissages de l'écrit que les connaissances de nature explicite pourraient se mettre en place. Par exemple, dans le domaine de la morphologie dérivationnelle, Lecocq et al (1996) évaluent, entre autre, les connaissances morfo-dérivationnelles à l'aide d'une épreuve d'identification de mots dérivés. A partir d'un mot cible (*e.g. chat*) les enfants doivent choisir parmi 3 mots dont 2 distracteurs, lequel en constitue un dérivé (*e.g. chaton, chien, château*). Les résultats obtenus sont faibles en GSM (48 % de bonnes réponses) mais augmentent en milieu de CP (61 %), sans atteindre une valeur maximale en CE1 (67 %). Cependant cette étude ne concerne que les suffixes. Toujours en français, Casalis et Louis-Alexandre (2000) dans une étude longitudinale de la GSM à la 2^e année emploient différents types d'épreuves (implicites *vs.* explicites en production *vs.* en réception). Dans une tâche simple (implicite/réceptive) consistant à trouver l'image correspondant au mot dérivé parmi 4 images représentant des mots complexes (trouver enrouler parmi les dessins représentant enrouler, dérouler, rouler, rouleau) les enfants dès la GSM distinguent déjà relativement bien le sens des mots composés (64 % de réponses correctes). De plus, ces habiletés augmentent peu entre la GSM et le CP mais plafonnent dès le CE1. En revanche dans une tâche de synthèse (produire un mot dérivé à l'aide de deux morphèmes

⁶³ L'utilisation des règles morfo-flexionnelles dans les productions enfantines a par exemple été étudiée par Cazden (1968).

⁶⁴ Cependant, Rocher (2005) n'est pas parvenu à répliquer ces résultats dans une tâche similaire mais présentée à l'écrit. Dans son étude, cette préférence n'apparaît qu'en 2^e année de primaire. La modalité écrite peut éventuellement expliquer ces différences.

représentés par deux parties d'une poupée : haut = *réparer*, bas = *tion* → *réparation*), les réponses correctes sont de 50 % en GSM et augmentent largement par la suite, au point de plafonner.

Dans le domaine de la morphologie flexionnelle, la distinction implicite/explicite fait apparaître une course développementale similaire. D'une part, Demont et Gombert (1996) ont observé que si en GSM les enfants parvenaient assez bien à juger de l'agrammaticalité de phrases portant des erreurs morpho-flexionnelles (*e.g.* je mangeons des gâteaux), ils avaient plus de difficultés à les corriger, notamment lorsque les phrases sont sémantiquement étranges. D'autre part, Nocus et Gombert (1997) montrent qu'avant le CP, les enfants ont de grosses difficultés à reproduire une erreur morpho-flexionnelle (*e.g.* effectuer la même erreur que « je regardons la TV » dans la phrase « je trie des papiers »). Cette capacité à manipuler explicitement les règles morphosyntaxiques augmente toutefois en fin de CP passant à 82 % de réponses correctes puis à 87 % en CE1.

D'autres recherches ont également montré, que les connaissances morphologiques continuaient à se développer jusqu'à un âge avancé (Carlisle, 1988, 2000; Gaux & Gombert, 1999a; Singson, Mahony, & Mann, 2000)

Du fait de la maîtrise plus tardive des règles morphologiques, en particulier dérivationnelles, mais aussi et surtout de la nécessité absolue d'acquérir les règles de conversions graphophonologiques (que sous-tendent les connaissances phonologiques), la contribution de ces connaissances à la lecture a le plus souvent été envisagée tardivement (*i.e.* pas avant la 3^e année) et de préférence sur la compréhension. Or, plusieurs recherches suggèrent leur intervention dès le début de l'apprentissage sur les deux dimensions de la lecture : la reconnaissance de mots écrits et la compréhension. Il y aurait donc un traitement de la structure morphologique des mots lors de la lecture débutante (*cf.* encart 9).

Encart 9: Les traitements de la structure morphologique lors de la lecture débutante.

Dans une série d'expérience Colé *et al.*, (2003) Marec-Breton *et al.*, (2005), et Marec-Breton et Gombert (soumis) ont clairement montré que dès le CP, les enfants tiennent compte de la structure morpho-dérivationnelle des mots pour les lire⁶⁵. En effet, et de façon globale, les mots ou pseudo-mots construits (*e.g.* *débouder*) sont mieux lus (plus vite et avec moins d'erreur) que les mots non construits ou les pseudo-mots ne faisant apparaître qu'un seul élément morphémique (*e.g.* *débouver*, *cagarer*). Par ailleurs, les expérimentations menées à l'aide du paradigme d'amorçage montrent que la présentation de la base facilite la reconnaissance des mots complexes dès le CP (Colé, 2004; Colé *et al.*, 2003; Royer, 2004).

⁶⁵ Ces auteurs répliquent ainsi en français ce que Laxon, Rickard et Coltheart (1992) et Burani (2003) avaient observé en anglais et en italien.

En langue anglaise, Carlisle et Nomanbhoy (1993) rapportent que les connaissances morphologiques (évaluées par une tâche de production de mots dérivés et fléchis en contexte de phrases⁶⁶) expliquent une part de variance du niveau de lecture de mots dès la 1^e année. Cette contribution est, certes, beaucoup plus faible (4% de variance expliquée) que celle des mesures de conscience phonologique, mais toutefois significative⁶⁷. Les travaux de Marec-Breton (e.g. Marec-Breton, 2003 Marec-Breton & Gombert, soumis) confirment en français cette intervention précoce des connaissances morphologiques. Par exemple, elle trouve que les performances à une épreuve de création de néologisme (e.g. *arrêter de chauffer c'est ___ ?*) expliquent en CP, encore 15 % de la variance en lecture de mots après l'intervention de la conscience phonémique, dont le rôle est toujours plus important.

Dans leur étude longitudinale en français, Casalis et Louis-Alexandre (2000) s'intéressent simultanément aux connaissances morpho-dérivationnelles, morpho-flexionnelles ainsi qu'à la lecture de mots et la compréhension de phrases écrites en début d'apprentissage⁶⁸. En CP, le niveau de lecture de mots est surtout lié à la conscience phonémique mesurée au même âge. Cependant, les auteurs notent une contribution significative (5,7 % de variance expliquée) des performances en production de formes fléchies (e.g. produire le féminin de *boulangier*) mesurée en GSM. Au CE1 les compétences en lecture de mots sont principalement expliquées par la réussite à deux types de tâches morphologiques : celle, implicite, de complétement de phrases à l'aide de pseudo-mots dérivés (*celui qui sait plosser est un ___*) mesurée au même âge et celle, explicite, de synthèse morpho-dérivationnelle (*réparer + tion = réparation*) mesurée en GSM. A ce niveau, les connaissances phonologiques n'interviennent plus. Les compétences en compréhension de phrases écrites sont pour leur part expliquées par les habiletés de segmentation morphologique, de nature explicite (*réparation = réparer + tion*), au même âge, mais aussi par les connaissances morpho-flexionnelles du féminin mesurées en GSM.

A l'issue des résultats de Casalis et Louis-Alexandre (2000), il semble que les connaissances morphologiques d'abord flexionnelles puis dérivationnelles interviennent lors de la reconnaissance des mots écrits, dès le début de l'apprentissage. Sur la compréhension en lecture, ces deux types de connaissances interviennent conjointement, dès la 2^e année. Cette étude a par ailleurs souligné l'importance de la distinction implicite/explicite.

⁶⁶ Par exemple à partir de la base « ferme », l'enfant doit compléter la phrase « mon oncle est un ___ » ; réponse attendue : « fermier ».

⁶⁷ Dans cette étude, comme dans les recherches présentées par la suite et utilisant les analyses de régressions comme moyen d'investigation, l'influence de variables telles que les capacités intellectuelles non verbales, le vocabulaire ou les capacités mnésiques a été contrôlée.

⁶⁸ L'évaluation des habiletés de compréhension, à l'aide du test de l'ECOSSE, n'est administrée qu'en CE1.

Les recherches menées à des niveaux plus avancés renforcent l'importance des connaissances morpho-dérivationnelles. Ainsi, les études de Shankweiler *et al.* (1995) de Mahony, Singson et Mann (2000) et de Carlisle (2000) montrent que la contribution des connaissances morphologiques à la lecture de mots augmente de la 3^e à la 6^e année. Dans le même temps, la contribution des connaissances métaphonologiques décroît et d'après l'étude de Singson *et al.* (2000), elle cesserait d'être statistiquement significative en 4^e année.

Les travaux de Demont et Gombert (1996) et de Nocus et Gombert (1997) confirment la relation entre lecture et compétences morpho-flexionnelles dès le CP. Demont et Gombert (Demont & Gombert, 1996 voir également Demont, 1994) rapportent les résultats d'une étude longitudinale évaluant les compétences en lecture (décodage et compréhension de phrases écrites⁶⁹) d'enfants suivis de la GSM au début du CE2.

- Sur la dimension de décodage, les connaissances phonologiques et en particulier la conscience phonémique s'avèrent être le meilleur prédicteur tout au long de l'apprentissage. Mais les compétences à corriger les erreurs morphémiques ont un pouvoir explicatif qui est loin d'être négligeable et cela dès la fin du CP (entre 8 et 18 % de variance).

- Sur la dimension de compréhension de phrases écrites, les auteurs retrouvent, là encore, une contribution importante des capacités à corriger les erreurs morphémiques en fin de CE2, bien que ce soient les compétences syntaxiques portant sur l'ordre des mots qui expliquent le mieux des capacités de compréhension. En CE1, 1. La compréhension est surtout dépendante des connaissances phonologiques.

Dans une autre étude longitudinale cherchant en particulier à prédire les compétences en lecture en CE1, Nocus et Gombert (1997, voir également Nocus, 1997), complètent ces observations en multipliant les mesures du niveau de lecture. Dans leur étude, après avoir contrôlé le rôle des connaissances phonologiques, les compétences à juger de la grammaticalité de phrases comportant des erreurs morpho-flexionnelles interviennent sur la lecture de mots mais toutefois moins et plus tardivement que les compétences à corriger la violation de l'ordre des mots. Par ailleurs, à l'aide de la même méthode d'analyse, le niveau de compréhension est, pour une part importante, expliqué par les performances en correction d'anomalies morpho-flexionnelles mesurées en GSM.

De même, Gaux et Gombert (1999a) observent que la contribution des connaissances morpho-flexionnelles (par exemple les compétences à répliquer des erreurs morphémiques) à la lecture (en lecture de mots comme en compréhension) est toujours présente en 6^e collège.

⁶⁹ Partant du principe que la compréhension nécessite un minimum d'efficacité en décodage, le test de compréhension n'a été administré qu'à partir du CE1.

Il ressort donc de ces analyses :

- **d'une part, que la conscience morphologique est ultérieure à la sensibilité implicite aux règles morphologiques et procède des apprentissages de l'écrit,**
- **d'autre part que, sans supplanter le rôle essentiel des compétences phonologiques, ces connaissances morphologiques contribuent précocement et conjointement à la reconnaissance de mots écrits et à la compréhension de phrases écrites. Cette contribution apparaît par ailleurs perdurer tout au long des apprentissages.**

2.2 Les recherches interlangues

L'investigation des différences interlangues dans le développement ou l'utilisation des connaissances morphologiques lors de la lecture est particulièrement limitée. Quelques études comparatives existent toutefois, mais s'attachent plus à examiner des phénomènes spécifiques qu'à pointer des différences majeures dans le rôle que peut jouer la morphologie selon les caractéristiques des langues.

L'étude sur le serbe et le turc de Fowler, Feldman, Andjelkovic et Oney (2003), porte sur le développement des connaissances morpho-dérivationnelles et morpho-flexionnelles ainsi que sur les liens que ces connaissances peuvent entretenir avec la phonologie. Leurs résultats appuient la critique précédemment énoncée d'un amalgame entre productivité morphologique et transparence phonologique et/ou sémantique (Kuo & Anderson, 2006): les résultats d'enfants de 7 et 8 ans montrent, en effet, que, même dans les langues comme le Serbe et le Turc dont la morphologie dérivationnelle est tout aussi riche et productive que la morphologie flexionnelle, la supériorité de connaissances morpho-flexionnelles (mesurée à l'aide d'une tâche de complètement de phrases à partir d'une base) est toujours visible. Pour eux, cette différence relève du fait que les relations sémantiques entre formes fléchies sont plus fortes que celles entre formes dérivées. Cet avantage en faveur de la morphologie flexionnelle apparaît ainsi dans plusieurs langues.

Fowler *et al.* (2003) ont par ailleurs montré des différences entre les deux langues, concernant le lien entre phonologie et morphologie. Seule une corrélation partielle⁷⁰ entre les performances de suppression de phonème et de complètement de phrases par une forme dérivée apparaît pour le groupe des enfants serbes ; en revanche, pour le groupe des enfants turcs, la suppression de phonème rend compte d'un pourcentage de variance comparable pour les deux types de

⁷⁰ L'utilisation de corrélations partielles permet aux auteurs de contrôler l'effet de la variable vocabulaire.

morphologie (dérivationnelle et flexionnelle). Les auteurs expliquent ces résultats par le fait qu'en serbe, la structure morpho-dérivationnelle est phonologiquement moins prédictible alors qu'en turc, les altérations phonologiques apparaissent sur les formes fléchies comme dérivées. Ils infèrent de ces résultats que la prédictibilité phonologique joue un rôle central dans l'acquisition des règles morphologiques.

Ragnarsdóttir, Simonsen et Plunkett (1999), s'intéressent, pour leur part, à l'acquisition des règles morpho-flexionnelles du passé des verbes en islandais et norvégien. Leur étude consiste à analyser les performances d'enfants de 4, 6 et 8 ans à la tâche de Berko et les erreurs qu'ils commettent en terme de généralisation de formes verbales existantes. Ils retrouvent ainsi, dans les deux langues, plusieurs résultats obtenus dans la littérature notamment sur l'ordre d'acquisition des formes régulières et irrégulières et sur l'effet de fréquence du groupe des verbes et des formes fléchies. Cependant, ils obtiennent deux résultats inédits :

- Le premier est que, quelle que soit la langue, les enfants ne commettent pas seulement des erreurs en généralisant des formes régulières fréquentes mais aussi en appliquant des formes régulières non fréquentes. Selon les auteurs, ce résultat indique que la fréquence du groupe verbal, même si elle se manifeste clairement comme une caractéristique importante dans l'acquisition des règles, ne détermine pas à elle seule la productivité d'une forme flexionnelle donnée. La régularité des règles d'application des flexions peut également être un élément déterminant dans l'utilisation des formes fléchies.
- Le deuxième résultat, particulièrement intéressant dans le cadre de cette étude, concerne une différence interlangue dans le rythme d'acquisition des règles morpho-flexionnelles : à 4 ans, les enfants islandais sont moins performants, en particulier sur les verbes irréguliers, que leurs pairs norvégiens mais rattrapent nettement leur retard dès 6 ans. Les auteurs expliquent ce décalage par la plus grande complexité du système flexionnel islandais.

Enfin, la recherche de Ku et Anderson (2003) s'est attachée à comparer les connaissances des aspects morphologiques (par dérivation et composition) et les relations qu'elles entretiennent avec le vocabulaire et la compréhension en lecture sur des échantillons importants d'enfants chinois et américains en 2^e, 4^e et 6^e année. Dans une épreuve de discrimination de morphème, les enfants doivent détecter le mot dont la partie commune ne partage pas le même sens que dans les autres mots (*e.g. classroom – bedroom – mushroom*) Ku et Anderson (2003) observent de moindres performances chez les enfants chinois de 2^e année lorsque les items concernent la formation de mots dérivés. Pour l'ensemble des groupes sinophones, la discrimination morphémique des mots composés est d'ailleurs plus importante que celle des mots dérivés. Ces

profils de réponses sont encore plus nets dans une épreuve de jugement sémantique sur des pseudo-mots complexes⁷¹, pour laquelle les enfants chinois obtiennent, à partir de la 4^e année de meilleurs scores sur les items composés que leurs pairs anglophones. Cet avantage des Chinois à traiter les aspects compositionnels par rapport aux aspects dérivationnels s'explique par le fait que les processus compositionnels sont plus fréquents et probablement plus transparents en chinois qu'en anglais alors que les affixes sont en revanche moins productifs. Cette plus faible productivité des affixes en chinois expliquerait également que, pour l'ensemble des évaluations effectuées⁷², les enfants chinois de 2^e année soient légèrement plus faibles que les enfants anglais. Cependant cette différence reste très circonscrite et ne traduit pas une différence fondamentale dans le développement des connaissances morphologiques dans les deux langues. Également, les auteurs s'attendaient à ce que les connaissances morphologiques soient davantage liées aux habiletés en lecture en chinois qu'en anglais car, contrairement à l'anglais, la formation des mots complexes en chinois implique rarement des modifications phonologiques ou orthographiques et que la plupart des caractères chinois représentent seulement un morphème. Les résultats obtenus montrent une contribution importante des connaissances morphologiques à la compréhension en lecture dans les deux langues. Chez les enfants chinois, la contribution des connaissances morphologiques, après avoir contrôlé les connaissances lexicales, est particulièrement importante en 2^e et 4^e année, mais moins en 6^e année. Ainsi, malgré ces quelques nuances dans le développement et le poids des connaissances morphologiques au cours des apprentissages, les auteurs concluent que les connaissances morphologiques sont universellement importantes pour l'apprentissage de la lecture.

Pour résumer, bien que peu nombreuses, ces recherches interlangues mettent en évidence l'influence des caractéristiques des langues sur la nature et le développement des connaissances morphologiques. Cependant, les variations interlangues sur le rôle que peuvent jouer ces connaissances lors de l'apprentissage de la lecture sont très peu explorées. L'étude de Ku et Anderson (2003), tout en soulignant l'importance de la dimension morphologique dans des langues aussi différentes que le chinois et l'anglais, tend également à montrer que certains aspects de la morphologie sont importants dans certaines langues plus que dans d'autres ou encore à certains moments de l'apprentissage

⁷¹ Dans cette épreuve, les enfants doivent juger si oui ou non ils connaissent le sens des pseudo-mots construits (*e.g. hearthful, muchable, cowhouse*).

⁷² En plus des épreuves de discrimination de morphèmes et de jugement du sens de pseudo-mots construits, les auteurs utilisent une épreuve de sélection de l'interprétation de mots construits rares et une épreuve de jugement de lien morphologique (*e.g. déterminer si « enseignante » vient de « enseigner »?*).

plus qu'à d'autres. De telles perspectives nécessitent d'être envisagées dans d'autres langues dont les caractéristiques morphologiques sont encore différentes.

En outre, à l'issue des multiples interprétations proposées par les auteurs, il semble que la notion de productivité englobe des caractéristiques linguistiques très diverses telles que les relations phonologiques et sémantiques plus ou moins transparentes lors de la formation des mots, la régularité de ces processus morphologiques, leur fréquence ainsi que la complexité des règles morphologiques. Si cette notion apparaît comme fondamentale pour comprendre les différences interlangues, il conviendrait de la définir plus précisément.

2.3 Les données en langue arabe

Comme mentionné précédemment, les recherches auprès de lecteurs experts en langues sémitiques et celles de Ammar (1997, 2002, 2003) sur l'apprentissage de la lecture en arabe ont largement envisagé un traitement morphologique de la racine morphologique dans ces langues. A l'écrit, l'exploitation des indices morphologiques, permettant de repérer le squelette consonantique, serait déterminante lors de la reconnaissance des mots et par conséquent pour comprendre les énoncés écrits du fait de l'aspect déficitaire des voyelles essentiellement véhiculées par le schème. Les recherches qui se sont attachées à démontrer ce phénomène en analysant la contribution des connaissances morphologiques au cours de l'apprentissage de la lecture en arabe sont relativement peu nombreuses. Elles ont pour la plupart comparé les compétences de bons et faibles lecteurs, sans distinguer ce qui appartient à la morphologie dérivationnelle ou flexionnelle.

En collaboration avec le LabÉCD⁷³ de l'Université de Nantes, l'équipe du département de Psychologie de la Faculté des Sciences Humaines et Sociales de Tunis a travaillé sur la modélisation et l'évaluation de l'acquisition de la lecture en langue arabe⁷⁴. Lors de ces travaux, plusieurs outils d'évaluation ont notamment été élaborés. Ainsi, les tests en français d'identification des mots écrits, de lecture de mots en une minute et de compréhension de phrases écrites de Khomsi (1990, 1994) ont été adaptés en arabe. Les recherches de Belajouza (2003), de Boukadida (Boukadida, en cours ; Boukadida & Besse, 2007) et de Maaouia Belkhadhi (2003) menées en Tunisie et présentés dans ce paragraphe ont utilisé certaines de ces épreuves.

⁷³ Laboratoire Education, Cognition, Développement.

⁷⁴ Projet CMCU 97F0407 « Modèle d'acquisition de la lecture et évaluation de la compétence lexicale : l'apport de l'arabe aux études interlangues » (1999).

Par exemple, Belajouza (2003) observe tout d'abord une évolution entre la 2^e et la 3^e année des habiletés morphologiques, évaluées à l'aide d'une épreuve de jugement de lien morphologique⁷⁵. Les scores obtenus sont d'ailleurs en 2^e année non négligeables (67% de bonnes réponses) alors que le programme tunisien prévoit l'enseignement systématique et progressif des règles de formations des mots et de leurs formes fléchies à partir de la 3^e année. De plus, les résultats des faibles lecteurs arabophones de 5^e et 6^e année sont proches de ceux obtenus par des normo-lecteurs de 2^e et 3^e année appariés sur le niveau d'identification de mots écrits. Il en est de même pour les résultats à une épreuve de suppression de phonèmes. Ainsi, malgré leur avancée dans la scolarité, les connaissances phonologiques et morphologiques des lecteurs en difficulté n'ont pas progressé. Ces enfants présenteraient ainsi un déficit phonologique mais aussi morphologique.

Ces résultats sont congruents avec ceux obtenus en hébreu par Ben-Dror, Bentin et Frost (1995) auprès de faibles lecteurs de 5^e année, appariés à de bons lecteurs⁷⁶ sur l'âge chronologique ou le niveau de vocabulaire. Ces auteurs avaient par ailleurs révélé que, si les faibles lecteurs sont moins performants que les bons lecteurs à identifier les phonèmes des mots et les relations morphologiques que deux mots peuvent entretenir, les écarts les plus importants concernent la morphologie. Cette dernière observation confirme l'importance des traitements morphologiques lors du décodage en hébreu mais s'opposent aux résultats de Abu-Rabia *et al.* (2003) qui avaient observé, aux mêmes niveaux scolaires, un déficit plus sévère en conscience phonémique qu'en connaissance morphologique, cette dernière apparaissant toutefois comme le second prédicteur des différences entre normo- et faibles lecteurs.

Outre cette difficulté à s'accorder sur le rôle respectif des connaissances morphologiques et phonologiques, ces recherches ne tiennent pas compte de la distinction implicite/explicite, ni ne s'intéressent au tout début de l'apprentissage de la lecture.

Actuellement, les travaux de Boukadida (Boukadida, en cours ; Boukadida & Besse, 2007) ont pour objectifs de démontrer l'influence des connaissances morphologiques implicites et explicites dès la première année. Les premiers résultats font apparaître, dès la 1^e année, une contribution importante des habiletés morphologiques, en particulier celles de complètement de phrases par un mot dérivé à partir d'une base, à la lecture de mots vocalisés.

Les études tunisiennes se sont ainsi surtout focalisées sur la reconnaissance de mots écrits. A notre connaissance, seule la recherche très récente d'Abu-Rabia (2007) s'intéresse également à la compréhension en lecture.

⁷⁵ Dans cette épreuve, les enfants doivent se prononcer sur l'existence d'un lien morphologique entre trois mots, sachant que dans 2/3 des items, un des mots est un intrus soit phonologique soit sémantique.

⁷⁶ La distinction entre faibles et bons lecteurs s'est effectuée sur la base du nombre d'erreurs commises à un test de lecture de pseudo-mots.

Abu-Rabia (2007) étudie les connaissances morpho-dérivationnelles, les habiletés à ajouter les voyelles internes et flexionnelles à la fin des mots, ainsi que les compétences à orthographier les mots arabes chez des élèves normo-lecteurs ou en difficulté, scolarisés en 3^e, 6^e, 9^e et 12^e année. Ses objectifs consistent à déterminer les facteurs les plus impliqués lors de la lecture de mots vocalisés et la compréhension de textes vocalisés chez ces différents groupes d'apprenants. Il observe que, pour tous les niveaux et sur toutes les mesures effectuées, les lecteurs en difficultés sont moins performants que les normo-lecteurs. Les analyses de régression font apparaître que les connaissances morphologiques sont systématiquement le meilleur prédicteur des habiletés en lecture. Plus précisément, chez les normo-lecteurs, les performances en lecture de mots sont, de la 3^e à la 6^e année, très largement expliquées par les performances à une tâche d'identification morphologique qui consiste à détecter, parmi trois mots, ceux partageant la même racine. En 12^e année, la lecture de mots est dépendante des capacités à produire un nombre important de mots dérivés à partir d'une racine. Les performances en compréhension sont expliquées en 3^e année par les performances d'identification morphologique puis, aux niveaux suivants, par les performances en orthographe de mots, dans une tâche de dictée de phrases⁷⁷. Chez les lecteurs en difficulté, les niveaux de lecture de mots et de compréhension sont généralement prédits par les scores en orthographe et éventuellement par les habiletés de production morphologique. L'auteur interprète ces résultats comme une confirmation que la lecture en arabe est basée sur la structure morphographique du système écrit. Pour lui, les compétences à produire des mots dérivés et à orthographier correctement les mots, relèvent d'un haut niveau de connaissances sur la complexité morphologique de l'arabe. Plus le niveau de littéracie augmente, plus le lecteur doit faire face à la complexité morphologique de la langue. Les plus faibles lecteurs auraient, pour leur part, des difficultés à mettre en place une décomposition orthographique des unités morphologiques, traitement qui nécessite un effort cognitif important. En outre, l'absence de contribution des compétences portant sur la vocalisation des mots s'explique selon l'auteur par les contraintes posées par les instruments de mesures. Pour parvenir à effectuer les vocalisations des mots présentés en contexte, les lecteurs doivent avoir de très bonnes connaissances des règles grammaticales et syntaxiques des mots. Ces tâches exigent aussi que les sujets soient actifs, ce qui n'est pas le cas dans les études ayant révélé l'effet facilitateur de la vocalisation en comparant les niveaux de lecture en version vocalisée et non vocalisée (cf. paragraphe 1.1.3 du chapitre 3).

De façon générale, ces résultats confirment le rôle essentiel et durable des connaissances morpho-dérivationnelles impliquant des degrés de contrôle variables sur les deux dimensions de

⁷⁷ Dans cette épreuve, il n'est pas demandé aux élèves d'écrire les signes diacritiques des mots.

la lecture : la reconnaissance de mots écrits et la compréhension. Cependant malgré l'intérêt de cette étude, plusieurs remarques peuvent être émises à son égard :

- Tout d'abord, les épreuves de lecture étant administrées en version vocalisée, il n'est pas étonnant de remarquer que les épreuves de vocalisation n'expliquent aucune part de variance des résultats en lecture. En d'autres termes, en version vocalisée, les capacités d'attribuer correctement les voyelles ne servent à rien puisque l'information les concernant est disponible.
- De plus, la contribution de la conscience phonologique n'a pas été explorée. Cela tient sans doute au fait que l'auteur considère que l'épreuve d'ajout des voyelles internes relève de la phonologie. Nous pensons, pour notre part, que les capacités à vocaliser nécessitent également des connaissances morphologiques, au moins implicites, puisque les voyelles internes sont apportées par le schème.
- S'il est évident que les compétences à orthographier les mots requièrent un haut niveau d'attention, il nous semble qu'elles ne sont pas exclusivement dépendantes des habiletés à extraire la racine, mais peuvent également rendre compte d'une maîtrise du système alphabétique et/ou de l'écriture non vocalisée.
- Enfin, l'évaluation des compétences morpho-flexionnelles, par une épreuve d'ajout des signes diacritiques flexionnels terminaux, ne semble pas refléter de façon exclusive et exhaustive les connaissances des aspects morpho-flexionnels. D'une part, car la morphosyntaxe en arabe ne se limite pas seulement à des marques diacritiques apposées à la fin des mots. D'autre part car ces connaissances peuvent être évaluées par des épreuves moins coûteuses et moins biaisées par les mécanismes de production que celles utilisées par Abu-Rabia (2003). Il nous semble ainsi que les capacités à gérer les marques flexionnelles peuvent s'avérer un facteur important lors de l'acquisition de la lecture en arabe.

L'étude descriptive de Maaouia-Belkadi (2003) étaye cette dernière supposition. Les réponses obtenues à une épreuve de jugement et de correction d'agrammaticalité (proche de celle utilisée par Demont & Gombert, 1996 et Nocus & Gombert, 1997), indiquent que les lecteurs de 3^e année et les faibles lecteurs de 4^e année⁷⁸, rencontrent des difficultés à reconnaître et appliquer certaines règles flexionnelles comme les marques casuelles et l'accord sujet/verbe ou nom/adjectif. Les habiletés d'identification de mots ne seraient ainsi pas indépendantes des connaissances morpho-flexionnelles.

Cette dernière recherche, comme beaucoup d'autres, les considère comme des connaissances syntaxiques. Ainsi, dans la méthodologie employée, la dimension flexionnelle est intégrée, soit à

⁷⁸ Les enfants ont été identifiés comme faibles à partir de leurs résultats à l'épreuve d'identification de mots écrits adaptée de celle de Khomsi (1994).

l'évaluation des connaissances morpho-dérivationnelle soit à celle des connaissances syntaxiques portant sur l'ordre des mots. Les résultats la concernant, ne sont donc presque jamais présentés de façon isolée. En hébreu, la morphologie flexionnelle est généralement étudiée en lien avec l'orthographe (*e.g.* Ravid, 2001; Ravid & Zilberbuch, 2003). Il est donc difficile à l'heure actuelle d'avoir un point de vue clair sur le rôle spécifique de ces deux dimensions au cours de l'apprentissage de la lecture, de même que sur le développement des connaissances morpho-flexionnelles par rapport aux connaissances morpho-dérivationnelles. Sur ce dernier point, l'étude de Levin, Ravid et Rapaport (2001), en hébreu, apporte quelques éléments de réflexion. Ces auteurs confirment, dans cette langue, la plus grande facilité des enfants (testés en dernière année de maternelle puis en 1^e année) à produire des formes fléchies que des mots dérivés. Levin *et al.* (2001) expliquent cette différence par le fait qu'en hébreu, comme dans d'autres langues, les flexions sont sémantiquement plus prédictibles et plus transparentes.

Globalement, bien que la distinction entre morphologie dérivationnelle et flexionnelle soit rarement faite, l'ensemble de ces recherches sur la lecture en arabe, sans invalider la nécessité des traitements phonologiques lors du décodage, apporte des observations congruentes avec l'hypothèse d'un traitement privilégié de la morphologie dès le début et tout au long de l'apprentissage.

2.4 Les données en langue portugaise

Les études sur le portugais ont plus régulièrement étudié le rôle des connaissances morphologiques, en particulier morpho-flexionnelles, sur le développement de l'écriture (*e.g.* Da Mota, Moussatchè, Rabello de Castro, Seidl de Moura, & D'Angelis, 2000; De Queiroga, Lins, & De Andrade Lima Vasconcelos Pereira, 2006; Rego & Buarque, 1997 ; Rosa, 2005). La prise en compte des aspects morpho-dérivationnels lors de la lecture en portugais ne s'est effectuée que très récemment. A notre connaissance, seules les recherches de Vidigal de Paula (2007) et Mota (Mota, 2006 : Mota, Aníbal, & Lima, communication personnelle) en rendent compte.

Dans une thèse consacrée aux connaissances morphologiques implicites et explicites d'élèves brésiliens scolarisés en 1^e, 3^e, 5^e et 7^e année, Vidigal de Paula (2007) confirme l'existence d'une sensibilité morphologique et l'émergence plus tardive d'une conscience morphologique contrôlée. En 1^e année, les enfants brésiliens sont en mesure, sans que l'expérimentateur leur en ait donné la consigne, de détecter implicitement l'intrus entre deux mots, dont l'un partage la même structure morphologique qu'un mot cible et l'autre non (*e.g.* cible : *descolorir* = décolorer-, proposition : *destorcer* = détordre-, *deslizar* = glisser). Leurs scores ne se

différencient toutefois pas du hasard lorsque, dans une tâche identique, il est précisé, au préalable, sur quelle règle les enfants doivent porter leur attention. Cette version explicite de l'épreuve n'est réussie qu'en 5^e année. De même les résultats à une tâche de création de néologisme (*e.g. une petite lampe est une ___ ?*) sont faibles en 1^e année, mais progressent, particulièrement entre la 3^e et la 5^e année.

Les analyses de corrélation montrent que le lien entre conscience morpho-dérivationnelle et lecture de mots n'apparaît qu'à partir de la 3^e année, d'abord par le biais des performances obtenues à une épreuve d'identification de règles morphologiques dans une langue inventée ($r = .39$)⁷⁹, puis aux niveaux suivants, par l'intermédiaire des scores obtenus à l'épreuve de détection d'intrus explicite ($r = .30$ en 5^e année et $r = .21$ en 7^e année). Pour le dernier niveau, les résultats aux épreuves de création de néologisme et d'identification de règles morphologiques inventées sont, pour leur part, corrélés au niveau de compréhension de phrases écrites (respectivement $r = .47$ et $r = .38$).

Ces résultats confirment, en portugais, la relation existant entre la conscience morphologique et la lecture de mots isolés comme la compréhension en lecture. Cependant, Vidigal de Paula (2007) n'est pas parvenue à faire apparaître cette relation en tout début d'apprentissage.

Dans le même objectif, Mota *et al.* (communication personnelle) ont réalisé une étude auprès d'enfants de 1^e et 2^e année. Les résultats obtenus montrent, indépendamment de l'âge, une contribution des performances à une tâche d'analogie morphologique⁸⁰ ($\beta = .35$) à la lecture de mots, alors que les résultats à une épreuve de détection d'intrus phonologique (de rime suprasyllabique ou d'allitération), n'y sont pas corrélés.

Ces deux chercheurs se sont également intéressés à la morphologie flexionnelle :

Vidigal de Paula (2007), à l'aide d'épreuves de détection d'intrus morpho-flexionnel, identiques à celles axées sur la morphologie dérivationnelle (*cf. supra*), observe des réponses aléatoires, pour la version implicite comme explicite, en 1^e année. Sans atteindre un niveau plafond, les scores sont meilleurs à l'épreuve explicite à partir de la 3^e année, en particulier lorsqu'il s'agit de flexions de genre, mais progressent peu entre la 3^e et la 7^e année. Par ailleurs, ces performances, sur la dimension explicite des connaissances, apparaissent corrélées à la lecture de mots, de la 1^e

⁷⁹ Dans cette épreuve, les participants doivent choisir, parmi trois propositions, le pseudo-mot dont la construction correspond à une définition donnée. Au préalable, plusieurs exemples de construction sont fournis de façon que les règles de la langue inventée soient perceptibles. (*e.g.* : « en kazi, chat se dit XOT, petit chat XOTIKE et grand chat XOTUKO, vélo se dit GOF, petit vélo GOFIKE et grand vélo GOFUKO, pot se dit RUV, comment dit-on petit pot en kazi: RUVIKE, RUFIKE ou RUVUKO ? »). Cette épreuve utilise des items relatifs aux règles morpho-dérivationnelles mais aussi morpho-flexionnelles.

⁸⁰ La tâche d'analogie morphologique consiste à effectuer, à partir d'un mot donné, la même dérivation que dans un exemple (*e.g.* effectuer la même dérivation que *pedra-pedreiro* = pierre-maçon, à partir de *leite* = lait).

à la 5^e année ($r = .46$ en 1^e année, $r = .49$ en 3^e année et $r = .25$ en 5^e année). Leur lien avec le niveau de compréhension en lecture n'est cependant pas vérifié.

Mota (2006) compare, dans un premier temps, les résultats d'adultes illettrés, peu lettrés et lettrés à des tâches de jugement de phrases correctes, de jugement de phrases incorrectes, dont l'erreur porte sur l'accord flexionnel, et de correction de ces dernières. Après avoir contrôlé les possibles effets de l'âge et du niveau cognitif, Mota constate que les adultes lettrés parviennent mieux que les adultes illettrés à juger de l'agrammaticalité des phrases et sont plus en mesure de corriger les violations morpho-flexionnelles. Chez des enfants de 1^e et 2^e année, l'augmentation en fonction du niveau n'est pas significative mais la corrélation entre niveau de lecture et niveau de connaissances morpho-flexionnelles est importante ($r = .51$). Cela suggère que c'est l'expérience scolaire et non le niveau de littéracie qui contribue au développement des connaissances morphosyntaxiques.

Sousa (2005, voir également Maluf, Gombert & Sousa, 2004) cherche à mettre en évidence que les compétences épimorphosyntaxiques influencent l'acquisition de la lecture de mots qui en retour influence le développement des compétences métamorphosyntaxiques. Dans sa thèse, Sousa (2005) montre ainsi qu'en 1^e année, les habiletés à juger et corriger des phrases comportant des erreurs morpho-flexionnelles prédisent le niveau de lecture de mots et de compréhension (respectivement 25 % et 24 % de variance expliquée). De plus, les habiletés en lecture de mots sont un bon prédicteur des compétences à corriger des phrases avec violations morpho-flexionnelles après répétition (14 % de variance expliquée). Ces compétences à répéter les erreurs avant de les corriger sont jugées par l'auteur de nature explicite. Cependant les capacités de mémoire verbale s'avèrent avoir un effet plus important (39% de variance expliquée).

Ainsi, dans les études de Mota et Sousa, il apparaît que les épreuves de jugement et de correction ne sont pas considérées de la même façon par rapport au degré de contrôle qu'elles impliquent. Mota considère ces tâches comme des tâches de conscience morpho-syntaxique et Sousa comme des tâches épilinguistiques (*i.e.* implicites). De plus, l'hypothèse que la manipulation consciente des règles morpho-flexionnelles influence positivement l'acquisition de la lecture n'est pas envisagée.

Globalement, ces recherches ont permis de vérifier que la lecture de mots est liée aux connaissances morpho-flexionnelles implicite et explicite dès le début de l'apprentissage. Egalement, les résultats de Mota *et al.* (communication personnelle), semblent attester le rôle des connaissances morpho-dérivationnelles, avant la 3^e année, mais ce résultat reste isolé. D'autre part, la compréhension en lecture est peu étudiée en lien avec les

connaissances morphologiques en portugais. Vidigal de Paula n'a réussi à démontrer sa relation avec les connaissances morpho-dérivationnelles que très tardivement (à partir de la 7^e année), ce qui ne correspond pas à ce qui est généralement obtenu dans les autres langues de même origine. Les recherches dans ce domaine doivent ainsi être multipliées.

3. Syntaxe et lecture

3.1 Préliminaire

Depuis les années 80, plusieurs recherches se sont focalisées sur le développement des connaissances syntaxiques et leur lien avec la lecture. Cependant leur mode d'investigation est très différente de celui des recherches menées dans le cadre du modèle de compétition de Bates et MacWhinney (1989, *cf.* paragraphe 2. 2 du chapitre 1), de sorte que ces deux champs de recherches sont visiblement indépendants l'un de l'autre.

La conscience syntaxique ou les compétences métasyntaxiques renvoient à la possibilité pour le sujet de raisonner consciemment sur les aspects de la structure des phrases et de contrôler délibérément l'usage des règles de grammaire (Gombert, 1990). Ces compétences pourraient intervenir sur l'apprentissage du décodage en permettant au lecteur débutant d'utiliser le contexte pour identifier les mots les moins familiers et assistent ainsi l'application des correspondances graphophonologiques, partiellement maîtrisées. Par ailleurs leur rôle sur la compréhension, tiendrait au fait qu'en permettant la mise en articulation des mots, elles participent au calcul de la signification de la phrase.

Rappelons tout d'abord, comme cela a notamment été souligné par Gombert et Colé (2000) et par Gombert *et al.* (1996 *cf.* paragraphe 2.3 du chapitre 1), que les compétences syntaxiques entretiennent, là encore, une relation circulaire avec l'apprentissage de la lecture, relation qui de surcroît est fonction du niveau de contrôle impliqué lors des traitements. En effet, les jeunes enfants peuvent à l'oral utiliser correctement les règles de construction syntaxique de leur langue, sans pour autant être capables de mener une analyse explicite sur ces aspects (Gombert, 1990). Au niveau préscolaire, l'enfant n'arrive pas à différencier la structure des phrases de leur contenu et dirige donc généralement leur attention sur les aspects sémantiques (Pratt, Tunmer, & Bowey, 1984). La conscience syntaxique n'émergerait que lorsque l'enfant est confronté à l'écrit, c'est-à-dire généralement à son entrée dans la scolarité (Bialystok, 1986).

Considérant que l'ordre des mots est un aspect important de la syntaxe, les recherches dans ce domaine utilisent fréquemment des épreuves de jugement et/ou de correction de phrases dont les

mots sont mis en désordre. Certaines utilisent également, des épreuves de complètement de phrases (tâche de closure) dans lesquelles l'enfant doit analyser la structure de la phrase, pour la compléter du mot le plus approprié. Par ailleurs, de nombreuses d'entre elles étudient simultanément les connaissances qui ont trait aux règles morpho-flexionnelles. Pour les raisons évoquées dans l'avant propos, ces dernières peuvent en effet être considérées comme une branche de la syntaxe.

Par exemple, Demont (1994) et Nocus (1997), ont observé, en français, des résultats similaires à ceux obtenus par Pratt *et al.* (1984) en anglais. Les tâches de jugement de phrases désordonnées sont déjà assez bien réussies en GSM, et même mieux que celles portant des erreurs morpho-flexionnelles. Au contraire, lors de tâches de correction, la syntaxe est manipulée plus difficilement et plus tardivement que la morphosyntaxe (les bonnes réponses concernent, avant le CP, à peine la moitié des items). En effet, l'anomalie de l'ordre affecte plus largement le sens de la phrase. De telles phrases sont ainsi faciles à considérer comme acceptables mais difficiles à corriger car cette action nécessite de retrouver la signification de la phrase tout en focalisant son attention sur les relations structurelles que les mots entretiennent. Nocus (1997) retrouve un pattern de réponses identique dans une tâche de répliation de faute, très coûteuse cognitivement (*e.g.* répliquer l'erreur de « Paul suce pouce son » dans « Emilie mange son gâteau »). Encore plus largement que dans la version morpho-flexionnelle, les enfants échouent à cette tâche avant le CP. En CP, les scores sont encore faibles mais progressent en CE1, sans plafonner.

Ces résultats rendent ainsi compte de la relative précocité des connaissances tacites des règles grammaticales et la mise en place plus tardive, en relation avec les apprentissages scolaires de la conscience syntaxique. Ils permettent aussi de constater le décalage dans la manipulation de l'organisation des mots dans la phrase par rapport à celle des règles flexionnelles.

Pour revenir à la question de savoir si les connaissances syntaxiques sont impliquées en début d'apprentissage de la lecture, et surtout sur quelle dimension de la lecture, les points de vue des auteurs ne s'accordent pas toujours.

Les premières recherches qui se sont intéressées à ce lien ont suggéré que les connaissances syntaxiques influencent davantage et directement les habiletés de décodage mais pas ou indirectement les habiletés de compréhension (Tunmer, 1989 ; Tunmer, Nesdale & Wright, 1987 ; Willows & Ryan, 1986). Pour confirmer cette affirmation, Tunmer (1989) se base, par exemple, sur l'observation que les performances d'enfants de 1^e et 2^e année sont corrélées de façon importante au niveau de lecture de mots et de non-mots ($r = .50$). Ce lien est d'ailleurs aussi important que celui entre lecture et conscience phonémique. En revanche, seul le décodage de pseudo-mots et la compréhension orale contribuent directement à la compréhension en

lecture. Le rôle de la syntaxe sur la compréhension passe donc par l'intermédiaire de la compréhension orale et du décodage qu'elle influence directement.

De façon à nuancer ces premiers postulats Régo et Bryant (1993) ont souligné l'importance des connaissances syntaxiques dans l'usage que peut faire l'enfant du contexte, pour aider sa lecture. En utilisant une tâche de lecture de mots en situation de facilitation contextuelle (les mots que les enfants n'étaient pas parvenus à lire isolément étaient de nouveau présentés, mais après avoir fourni à l'oral des indices contextuels facilitant leur reconnaissance), les auteurs montrent que les multiples mesures des compétences syntaxiques (*e.g.* tâche de closure et de correction) contribuent à expliquer la variance en lecture mais seulement dans le cas où celle-ci est précédée d'une facilitation contextuelle alors que les compétences phonémiques, interviennent exclusivement sur la lecture de mots classiques.

Nocus et Gombert (1997) constatent pour leur part que les habiletés en lecture de CE1 sont prédites par les compétences à corriger les erreurs d'ordre des phrases mesurées en GSM puis par les compétences à répliquer des erreurs syntaxiques et les compétences phonologiques mesurées après le début de l'apprentissage de la lecture. Les meilleurs prédicteurs des habiletés de compréhension sont, en début de longitudinale, les compétences à corriger les erreurs morpho-flexionnelles puis, en fin de longitudinale, les connaissances phonologiques. Toutefois, à deux reprises, les compétences syntaxiques en CP (en particulier la réplification d'erreur) apportent une contribution significative sur la compréhension mesurée à l'aide du test de Khomsi (1994).

Demont et Gombert (1996) confirment, en fin de CE1, la contribution indépendante des habiletés à juger et corriger les erreurs d'ordre, à la reconnaissance de mots comme à la compréhension de phrases écrites, mais toutefois plus faiblement que les compétences phonologiques ou morpho-flexionnelles. En revanche, les compétences, en GSM, à corriger ces mêmes erreurs s'avèrent être un excellent prédicteur de la réussite en compréhension de phrases écrites en 3^e année. Bohannon, Warren-Leubecker et Hepler (1984), avaient obtenu des résultats allant dans le même sens en anglais. Enfin, Gaux et Gombert (1999ab) corroborent l'importance des connaissances syntaxiques, en particulier sur la compréhension chez des élèves de 6^e année collège.

Ces dernières études, permettent de conclure que les connaissances syntaxiques, d'abord peu contrôlées, puis de plus en plus explicites, peuvent jouer conjointement un rôle sur les deux dimensions de la lecture. Elles ne réfutent pas pour autant, le rôle des connaissances phonologiques essentiel au décodage, ni la prise en compte importante des indices morpho-flexionnels, facilement détectable à ces âges.

Cependant, le rôle des connaissances syntaxiques sur la lecture ne fait pas l'unanimité, contrairement à celui des connaissances phonologiques. Cela tient notamment au fait que certaines recherches, partisans de l'hypothèse de la « limitation phonologique » pour expliquer les dysfonctionnements en lecture (Shankweiler, Crain, Brady, *et al.*, 1992) ont échoué à en dégager rigoureusement le rôle (Shankweiler *et al.*, 1995; Bowey, 2005 ; Gottardo, Stanovich & Siegel, 1996).

Il convient enfin de souligner qu'il n'existe, à notre connaissance, aucune recherche interlangue qui ait pu aborder le développement des connaissances syntaxiques et leur rôle en comparant des langues différentes. Les recherches sur ces connaissances sont d'ailleurs surtout circonscrites aux langues anglaise et française. Elles se sont, de surcroît, heurtées à de nombreuses difficultés méthodologiques. Il est en effet souvent difficile de s'assurer sur quels aspects les enfants fondent leur jugement et d'évaluer précisément l'effort cognitif qu'ils fournissent lors de certaines épreuves. Ces remarques peuvent également être formulées à l'égard des quelques études sur les langues sémitiques et le portugais abordés maintenant. Nous reviendrons plus précisément sur ces problèmes méthodologiques ultérieurement.

3.2 Les données en langue arabe

Les recherches portant sur la dimension syntaxique lors de l'apprentissage de la lecture en arabe vocalisé sont rares et ne traitent pas toujours de façon séparée ce qui relève de l'ordre des mots et des règles morpho-flexionnelles. Soulignons à nouveau que les règles syntaxiques de l'arabe sont riches, complexes, et affectent beaucoup la structure interne des mots, mêmes si les marques flexionnelles ne correspondent parfois qu'à des signes diacritiques discrets. Ainsi, l'acquisition de telles règles, est souvent vécue par l'apprenant comme une difficulté, exacerbée par le phénomène de diglossie. Par ailleurs les linguistes et psycholinguistes ont suggéré que l'ordre des mots en arabe est sans doute moins déterminant que dans d'autres langues. Il est toutefois envisageable que les traitements sur les aspects organisationnels des éléments lexicaux interviennent lors de l'apprentissage de la lecture, notamment car l'utilisation du contexte est particulièrement pertinent dans cette langue (*cf.* Baccouche, 2003; Taman, 1993; voir paragraphes 2 du chapitre 2, et 1.2.3. du chapitre 3).

Maaouia-Belkadi (2003) montre ainsi une forte corrélation entre les performances en compréhension de phrases écrites et en jugement/correction de phrases agrammaticales chez des enfants tunisiens de 2^e, 3^e et 4^e année. De plus, l'analyse des réponses données à l'épreuve de

jugement / correction permet de constater que les anomalies d'ordre posent certaines difficultés aux faibles lecteurs de 3^e année, difficultés qui semblent se résorber en 4^e année. De façon générale, l'auteur retrouve aussi des résultats inverses par rapport aux études francophones ou anglophones (Demont & Gombert, 1996 ; Nocus & Gombert, 1997 ; Pratt *et al.*, 1984). En effet, les arabophones montrent une plus grande facilité à corriger les phrases dont l'ordre est altéré que celles portant des anomalies morphologiques. Maaouia-Belkadi (2003) interprète ces résultats par le fait que les règles morpho-flexionnelles de l'arabe standard sont complexes et différentes de celles du dialecte oral. L'interprétation que les modifications d'ordre affectent peu la signification des phrases arabes car les combinaisons sont plus libres que dans d'autres langues, n'est d'ailleurs pas contradictoire à celle de l'auteur.

Romdhane (2003) confirme également qu'en arabe, les compétences à juger et corriger l'agrammaticalité des phrases sont déjà élevées dès la 2^e année et apparaissent en 4^e année comme le deuxième prédicteur de la réussite en compréhension de phrases écrites, après les compétences d'identification de mots, mais avant la compréhension orale. Les résultats à une épreuve de suppression de phonème n'y apportent en revanche aucune contribution. Ces résultats attestent ainsi que les traitements syntaxiques sont indispensables à l'interprétation des énoncés écrits.

Toutefois, ces deux dernières recherches analysent le lien entre compréhension de phrases et jugement/correction d'agrammaticalité sans distinguer ce qui est relatif à l'ordre des mots et à leur forme flexionnelle et ne traitent pas de façon systématique l'impact de ces compétences sur la lecture de mots isolés.

Certains résultats de l'étude de Bentin, Deutsch et Liberman (1990), menée en hébreu et exclusivement sur les agrammaticalités d'inversion de l'ordre des mots, permettent de compléter ces données. Auprès d'enfants de 9 à 14 ans, les auteurs trouvent que les lecteurs en difficulté de décodage (lecture de pseudo-mots vocalisés) obtiennent des scores inférieurs aux bons lecteurs, aux tâches de jugement et de correction. De plus, les lecteurs maîtrisant le décodage mais ayant des difficultés de compréhension à l'écrit sont spécifiquement moins performants pour appliquer leurs connaissances syntaxiques pour corriger les erreurs.

L'ensemble de ces résultats laisse penser qu'en langues sémitiques également, la prise en compte des aspects organisationnels des phrases peut avoir un impact sur les habiletés de lecture, en reconnaissance comme en compréhension. Les données recueillies ne nous informent toutefois pas sur cette contribution lors des premières années d'apprentissage.

3.3 Les données en langue portugaise

A la suite des recherches de Régo (1993, 1997, 1999), plusieurs travaux sur la lecture au Brésil ont pris en considération le rôle des connaissances syntaxiques en début d'apprentissage.

Par exemple, Sousa (2005) constate, comme dans d'autres langues, la plus grande facilité des enfants brésiliens de 1^e année à corriger les phrases désordonnées (75% de bonnes réponses) par rapport aux phrases contenant des erreurs morpho-flexionnelles (61 %). Par ailleurs, les performances à corriger l'ordre des mots (habiletés épisyntaxiques) obtenues en début d'année sont corrélées ($r = .40$) aux scores de lecture de mots isolés mesurée 6 mois plus tard. Elles n'apparaissent toutefois pas prédire ce niveau de lecture, davantage dépendant des compétences épimorphosyntaxiques (*cf. supra*). Enfin, ces mêmes habiletés épisyntaxiques ne semblent pas liées aux habiletés de compréhension de phrases. L'auteur rapproche ces résultats de l'hypothèse de Shankweiler et al (1992) et des résultats de Tunmer (1989). Elle précise ainsi qu'à un stade aussi précoce, les enfants ne maîtrisent pas suffisamment les règles de décodage pour pouvoir utiliser des stratégies plus complexes lors de leur lecture de phrases.

En répliquant à plusieurs reprises l'étude menée en anglais par Régo & Bryant (1993) à l'aide du même paradigme Régo (1997) envisage pourtant différemment le rôle de la syntaxe en portugais. Au cours de ces recherches, elle teste les compétences phonologiques (tâche de détection d'intrus d'allitération) et syntaxiques (tâche de correction de phrases dont les mots sont mis en désordre) avant l'apprentissage puis évalue les compétences en lecture en début et fin de 1^e année⁸¹. Dans une première expérimentation Régo (1993) trouvait que les mesures des connaissances syntaxiques prédisent non seulement les compétences en compréhension de courts passages écrits en 1^e année, mais également les habiletés de lecture de mots dès l'entrée dans les apprentissages (que les mots soient présentés isolément ou en présence d'un contexte oral facilitateur). Les différences obtenues par rapport à l'étude de Sousa (2005), notamment sur la compréhension, peuvent être dues au type d'évaluation du niveau de compréhension ou au moment où cette évaluation a été effectuée⁸². Toutefois, ces résultats sont similaires avec ceux recueillis dans d'autres langues dont l'orthographe est plus opaque. Mais Régo souhaite également vérifier si cela ne provient pas de l'usage des méthodes d'apprentissages favorisant une approche holistique du langage aux stades initiaux. Elle s'intéresse alors aux enfants ayant exclusivement appris la lecture par la méthode phonique (Rego, 1997). Ses résultats mettent toujours en évidence que les habiletés à corriger les anomalies grammaticales apparaissent

⁸¹ Dans ces expériences, les habiletés de compréhension n'ont été évaluées qu'en fin de 1^e année.

⁸² Régo évalue la compréhension à l'aide d'une épreuve où les enfants doivent lire deux courts passages et répondre oralement à 5 questions, 10 mois après l'entrée en 1^e année. Sousa évalue ces mêmes compétences à l'aide d'une épreuve de désignation d'image correspondant à des phrases courtes et simples, en milieu d'année.

comme un très bon prédicteur de la compréhension en fin d'année, également un bon prédicteur de l'habileté en lecture de mots présentés avec un contexte oral facilitateur mais n'intervient pas sur la lecture de mots hors contexte, quel que soit le moment de son évaluation⁸³. Ainsi, l'auteur confirme l'impact des méthodes d'apprentissage, mais sur une seule épreuve de lecture, axée sur le décodage. Elle en conclut le lien essentiel entre les connaissances syntaxiques et les dimensions de la lecture qui requièrent l'usage d'information de haut niveau (*i.e.* la compréhension ou la lecture en contexte) puisqu'il apparaît dans différentes langues et quel que soit le type d'instruction reçu. En revanche, la connexion entre connaissances syntaxiques et habiletés de décodage est moins consistante car influencée par les instructions que les apprentis lecteurs reçoivent.

Barrera et Maluf (2003) ont aussi confirmé le rôle des connaissances syntaxiques (évaluées toujours avec le même paradigme de correction de phrases agrammaticales) en début d'apprentissage. Leur étude montre une corrélation significative mais toutefois faible de ces performances avec les scores des premières connaissances en lecture de début de 1^e année et une corrélation plus importante en fin de d'année mais toutefois inférieure à celle observée avec les connaissances phonologiques.

Enfin, Guimarães (2002), en s'intéressant aux difficultés en lecture de mots et d'écriture d'enfants de 3^e et 4^e année, rapporte que ces élèves sont moins performants que leurs pairs du même âge, à une tâche de corrections de phrases désordonnées.

Ainsi, si l'on fait abstraction de l'influence des méthodes d'enseignement⁸⁴, les données, même peu nombreuses, présentées précédemment semblent bien confirmer ce qui s'était dessiné dans d'autres langues alphabétiques : les connaissances syntaxiques jouent un rôle, conjointement sur l'acquisition de la lecture de mots et la compréhension. Cependant l'aspect universel de cette intervention ne peut être accepté en l'absence de recherches comparatives. Par ailleurs, les données recueillies auprès d'élèves de niveau déjà avancé restent toutefois absentes.

⁸³ Ces résultats ont été obtenus après avoir contrôlé l'effet du niveau cognitif, de connaissances orthographiques et de la mémoire verbale.

⁸⁴ Bien que ce facteur soit sans conteste capital pour comprendre les apprentissages, il semble difficile de l'évaluer rigoureusement. Nous pensons en effet que dans bon nombre de cas, l'application des méthodes par les enseignants est suffisamment diversifiée pour que l'on ne puisse considérer les instructions comme purement phoniques ou purement globales.

4. Analyse critique

Un des points critiques émergeant des études sur ces diverses connaissances, étudiées dans différentes langues est de nature méthodologique :

Premièrement, dans une perspective comparative, l'utilisation de multiples modes d'évaluation, administrés à des enfants de niveaux différents, complexifie la possibilité d'établir des comparaisons entre les résultats obtenus en arabe et en portugais. Il est ainsi difficile d'appréhender les différences fines dans l'apprentissage de la lecture entre ces deux langues, de même qu'entre ces langues et les autres langues étudiées dans la littérature.

Deuxièmement, les investigations sur les connaissances impliquées lors de la lecture sont systématiquement confrontées à la difficulté de distinguer sans ambiguïté, la nature implicite ou explicite des connaissances. Les commentaires, présentés dans l'encart 2, sur l'évaluation des connaissances implicites et explicites, proposent un bref aperçu du coût cognitif qu'entraînent certaines tâches et invitent à conclure que le caractère implicite ou explicite des épreuves est souvent déterminé par le domaine ou le type d'unité étudiés. Ainsi, dans la réalité expérimentale, la distinction est complexe et il est possible que le foisonnement des instruments de mesure soit, en partie, lié à cette difficulté. Or, bien que les choix des tâches proposées aux enfants soient d'une grande importance, il est rare d'obtenir un outil d'évaluation exempt de toute critique méthodologique. En outre, l'évaluation des connaissances morphologiques et syntaxiques se heurte également au risque de confusion avec la prise en compte des aspects strictement sémantiques. Une analyse sommaire de quelques paradigmes souvent employés permettra d'en illustrer les problèmes récurrents et de définir, à terme, les choix les plus pertinents pour évaluer ces connaissances auprès d'élèves de différents niveaux scolaires.

L'évaluation des connaissances à l'aide des épreuves de jugement ou d'identification de similitude ou de détection d'intrus nous semble assez équivoque. Ces épreuves, qu'elles évaluent les connaissances phonologiques ou morphologiques, sont souvent considérées comme implicites car la production de la réponse est en effet peu coûteuse. Il semble toutefois difficile de s'en assurer car la réponse peut être issue d'une analyse contrôlée surtout lorsque la consigne précise de quelle nature est la similitude ou la différence. En effet, en parlant des épreuves de jugement de relation morphologique Colé *et al.* (2003) précisent que ces dernières « *ne nécessitent pas une mobilisation consciente dans la mesure où la consigne ne focalise pas explicitement l'enfant sur la structure morphologique des mots* » (p 716). Pourtant, les consignes employées par la plupart des chercheurs utilisent la notion de famille de mots. De la même façon, pour évaluer les connaissances phonologiques, la demande porte sur les mots « qui

commencent ou terminent pareil ou par le même son ». De telles consignes peuvent, à notre avis orienter, les réponses des enfants. Le choix de Goswami *et al.* (2005) et Vidigal de Paula (2007) est plus clair puisque la consigne des épreuves de détection implicite d'intrus est beaucoup plus vague (détecter le mot qui est différent entre X, Y et Z ou qui va le moins avec X, Y ou Z).

L'aspect heuristique des tâches exigeant une manipulation des éléments linguistiques et donc un contrôle cognitif important n'est pas à remettre en cause. L'épreuve de suppression de phonème a largement fait ses preuves mais dans le domaine morphologique, la suppression ou même l'ajout d'un morphème s'apparente parfois à retrouver un mot à partir d'un autre, ce qui ne nécessite pas forcément une analyse en morphème mais s'apparente davantage à une application des connaissances lexicales. Bien souvent, la tâche du chercheur qui s'intéresse à la morphologie est justement d'éviter que les épreuves proposées ne soient exclusivement résolues sur la base des aspects sémantiques ou phonologiques Kuo & Anderson (2006) Par exemple, les réponses à la tâche de choix d'image représentant des mots de la même famille (*e.g.* rouler, dérouler, enrouler, rouleau), ou construit de façon identique (*e.g.* brûlure, chevelure, blessure, pelure) utilisée par Casalis et Louis-Alexandre (2000) dépendent autant de l'ampleur du lexique des enfants que de leur capacité à repérer la construction morphologique des mots. De même, les simples épreuves où l'enfant doit déterminer si deux mots appartiennent à la même famille (Ben-Dror *et al.*, 1995; Carlisle & Nomanbhoy, 1993) peuvent être résolues en se fondant sur la signification des mots connus, et sans prendre nécessairement en compte le lien morphologique. Enfin, dans les épreuves plus complexes de jugement de lien morphologique, présentant des mots de la même famille et un distracteur soit phonologique soit sémantique (*e.g.* cuisine-cousine-cuisinier ; fermer-refermer-ouvrir), souvent employées en arabe (Abu-Rabia, 2007; Abu-Rabia *et al.*, 2003; Belajouza, 2003), les réponses peuvent être fournies en éliminant les mots qui ne partagent pas les mêmes caractéristiques sémantiques ou phonologiques et non en discriminant simultanément ces deux aspects. Les performances à ce genre de tâches plafonnent d'ailleurs aux alentours de la 2^e année (Abu-Rabia *et al.*, 2003; Marec-Breton, 2003). Les épreuves de production de mots dérivés (sans ou en contexte de phrases) ou de néologismes font plus clairement intervenir une connaissance des règles de construction, même si là encore, les productions sont généralement spontanées.

L'évaluation des compétences métasyntaxiques est tout aussi délicate. Le risque d'amalgamer les réponses de nature syntaxique avec celles relevant de la sémantique est en effet important (Layton, Robinson, & Lawson, 1998) et s'ajoute à celui de confondre la nature implicite ou explicite des réponses.

Par exemple les épreuves de closure ou de complèvement de phrases par le mot syntaxiquement adéquat, relèvent tout autant de la gestion des contraintes syntaxiques que sémantiques.

Par ailleurs, il est difficile de savoir exactement de quelle façon sont résolues les tâches de jugement de grammaticalité de phrases. Les résultats obtenus dans la littérature semblent démontrer que les jugements des enfants s'effectuent la plupart du temps sur la base d'indices sémantiques. Cette interprétation est en effet cohérente avec les meilleures performances lorsqu'il s'agit de rejeter des phrases dont les erreurs affectent beaucoup le sens (*i.e.* les erreurs sur l'ordre, Demont & Gombert, 1996; Nocus & Gombert, 1997). Au contraire, la correction d'agrammaticalité apparaît, en particulier en début d'apprentissage, plus coûteuse lorsque le sens est affecté. Cela démontre, là encore, la forte intrication entre les aspects sémantiques et syntaxiques.

De plus, le problème du coût cognitif engagé dans les tâches de jugement et de correction n'est pas plus facile à résoudre que dans les autres tâches précédemment citées. Le démontre les différences d'interprétation entre Mota et Sousa à propos de la morphologie flexionnelles en portugais. Les jugements, de même que les corrections d'agrammaticalité peuvent en effet être effectués tout autant à partir d'une connaissance implicite activée à force de rencontrer les formes correctes récurrentes qu'à partir d'une réflexion consciente (Bialystok, 1986; Gaux & Gombert, 1999ab; Gombert, Gaux, & Demont, 1994). L'épreuve de répliation d'erreur apparaît en revanche suffisamment complexe pour exiger la mobilisation de connaissances contrôlées (Gaux et Gombert, 1999ab).

Il paraît donc extrêmement difficile d'identifier le mode d'évaluation le plus pertinent de ces différentes connaissances. Par ailleurs, les épreuves se situant sur le mode réceptif et productif impliquent des processus trop différents pour être comparables. Cependant, à l'instar de ces multiples remarques, l'usage du paradigme de détection d'intrus semble bien approprié pour trois raisons :

- 1) Il permet d'évaluer les choix spontanés des enfants lorsque ceux-ci ne sont pas dirigés par une consigne, mais également les réponses plus contrôlées lorsque l'expérimentateur explique les règles sous-jacentes à l'exercice.
- 2) Il peut s'appliquer à tous les domaines (phonologique, morphologiques et syntaxique).
- 3) Il a l'avantage d'être peu coûteux en termes de mise en place dans les classes.

5. Synthèse et perspectives

A l'issue de ces travaux, réalisés en arabe et portugais ou dans d'autres langues, plusieurs éléments généraux, communs aux langues alphabétiques, émergent. Les modèles théoriques généralement employés pour appréhender les mécanismes d'apprentissage de la lecture ne sont ainsi pas fondamentalement remis en cause (Pinheiro, 1995; Taouk & Coltheart, 2004). Cependant, plusieurs observations apportent des nuances non négligeables permettant d'envisager que les multiples connaissances pourraient, dans différentes langues, ne pas jouer le même rôle aux mêmes moments de l'apprentissage. De plus, le poids de chacune d'entre elles peut également être différent d'une langue à l'autre. La dimension phonologique, essentielle dans toutes les langues, doit être examinée dans ses multiples facettes pour comprendre l'impact de la transparence des systèmes écrits. La dimension morphologique vient de plus en plus pondérer la « suprématie » de la phonologie lors de l'apprentissage de la lecture mais les recherches interlangues dans ce domaine doivent être prolongées. Par ailleurs, les travaux sur la langue arabe, et sur l'hébreu, s'accordent sur le statut privilégié de la morphologie dans les langues sémitiques. Enfin, la dimension syntaxique se restreint généralement aux aspects organisationnels des mots dans les phrases, alors qu'elle constitue un domaine linguistique particulièrement vaste. Sa complexité et la diversité des formes syntaxiques entre les langues sont peut être à l'origine de l'absence d'études interlangues.

Finalement, ces différentes dimensions n'ont, à notre connaissance, jamais été considérées simultanément dans une perspective interlangue et évolutive. Notre recherche aura donc comme objectif de s'y attacher.

Chapitre 5 Apprentissage de la lecture en langue seconde

Un bref aperçu des situations linguistiques des pays du monde et des flux migratoires permet de remarquer que même chez les enfants, le monolinguisme ou l'utilisation d'une seule langue est loin d'être la norme. Le contact de deux langues ou plus, prend ainsi des formes multiples, allant presque jusqu'à rendre chaque situation unique. Par exemple, les langues en contact, le contexte de leur apprentissage, le statut et le niveau de maîtrise qu'elles acquièrent, sont souvent très différents. Rien d'étonnant, alors, que l'apprentissage des langues secondes et le bilinguisme soient devenus des problématiques de recherche à part entière et que la littérature scientifique, dans ces domaines, abonde. A première vue, les langues secondes et le bilinguisme réfèrent à un statut cognitif, une acquisition et une maîtrise des langues différents. Pourtant, les recherches qui s'intéressent aux sujets en contact avec plusieurs langues, souvent des populations migrantes allophones ou d'origine immigrée, font régulièrement appel à ces deux termes, sans toujours les distinguer clairement. En fait, l'acception de la notion de bilinguisme varie beaucoup selon les auteurs. De même, la frontière entre apprentissage d'une langue seconde ou d'une langue étrangère mérite d'être précisée. Une clarification terminologique s'avère donc nécessaire. Elle aura pour objectif de définir comment ces champs de recherches peuvent en effet être connectés et ce qu'ils peuvent respectivement apporter à notre propre problématique.

1. Précisions terminologiques et définitions

1.1 Du bilinguisme aux bilinguismes

Le foisonnement des définitions et l'émergence de multiples types de bilinguisme rendent compte de l'hétérogénéité de la population utilisant deux langues ou plus, mais reflètent également les options prises par les auteurs dans la façon de considérer cette notion.

Le sens populaire donné au bilinguisme, rejoint les définitions des dictionnaires courants et de plusieurs auteurs (Thiéry, 1976⁸⁵ ; Mazureau, 1961⁸⁶ ; Bloomfield, 1933 ; Halliday, McKintosh & Strevens, 1970⁸⁷) qui envisagent le bilinguisme comme la maîtrise parfaite de deux langues⁸⁸. Dans cette acception, les bilingues utilisent couramment et indifféremment deux langues, sans

⁸⁵ Cité par Bruyninckx & Harmegnies (1993).

⁸⁶ Cité par Titone (1972).

⁸⁷ Cités par Perregaux (1994).

⁸⁸ La maîtrise de plus de deux langues est nommée multilinguisme. Pour plusieurs auteurs le terme plurilinguisme désigne la coexistence de plusieurs langues dans un espace géographique ou politique donné (e.g. Hagège, 1996).

aptitude plus marquée pour l'une que pour l'autre. Certains théoriciens considèrent même que l'aisance des bilingues doit être comparable à celle des locuteurs de langue maternelle. Il n'y aurait, pour ainsi dire, pas de différences entre bilingues et monolingues. De même, il est affirmé, plus ou moins explicitement, un équilibre entre les langues connues des bilingues. Cependant, pour beaucoup d'autres auteurs, ces définitions et conceptions du bilinguisme « vrai »⁸⁹ ou « parfait »⁹⁰ sont trop restrictives pour appréhender la complexité du phénomène de contact des langues et se révèlent donc peu adaptées voire illusoirs (*e.g.* Mackey, 1976 ; Perregaux, 1994 ; Titone, 1972). Il est en effet aisé de rétorquer d'une part, que le degré de perfection dans la maîtrise des langues, que ce soit chez les locuteurs natifs monolingues ou les bilingues, est difficile à évaluer car extrêmement variable, et d'autre part, que la parité absolue dans le contrôle des deux langues est très rare.

Ainsi, la conception la plus souvent adoptée est celle qui considère le bilinguisme non pas comme un « tout ou rien » mais comme un processus individuel de développement de connaissances linguistiques et langagières, coexistant à des degrés variés (d'une compétence minimale à une maîtrise complète) et mettant en jeu deux codes distincts (*e.g.* Mackey, 1976; Perregaux, 1994). Ce processus s'élabore sur des éléments culturels, sociologiques et psychologiques et permet une adaptation permanente de l'activité langagière à la situation dans laquelle elle se déroule⁹¹.

Les descriptions de Mackey (1976) ont caractérisé assez précisément la complexité des phénomènes de contact de langues et les multiples facteurs en jeu pour expliquer la diversité des situations bilingues. Cet auteur dégage plusieurs variables affectant le degré d'aptitude, en compréhension, expression, lecture et écriture, atteint au niveau phonologique ou graphique, grammatical, lexical, sémantique et stylistique. Le niveau de maîtrise dans chacun de ces domaines, qui n'est d'ailleurs pas forcément identique, dépend entre autres :

- De la fonction interne et externe des langues : la représentation identitaire, le rôle communicationnel, la conception socioculturelle de la langue vont influencer le développement des compétences dans les deux langues et la motivation à les utiliser dans des contextes définis. Ces fonctions sont également à combiner avec la durée, la fréquence et la force⁹² du contact engagé entre le sujet et les langues qu'il connaît.

⁸⁹ Terme employé par Titone (1972).

⁹⁰ Terme emprunté à Jardel (1979 cité par Perregaux, 1994).

⁹¹ Parce que le bilinguisme appartient au domaine de l'activité linguistique, on parlera souvent du « parler bilingue » pour caractériser l'emploi individuel de deux langues, emploi qui dépend, au moins à l'origine, de la rencontre de deux communautés linguistiques différentes.

⁹² Sous le terme de force de contact, Mackey fait référence aux pressions économiques, culturelles, politiques, religieuses etc. pouvant expliquer l'enjeu de l'emploi des langues et de leur maîtrise.

- De l'influence d'une langue sur l'autre : elle est plus ou moins importante et s'applique dans différents domaines, lexical, structural, phonétique etc.
- De l'oscillation ou l'alternance qu'un sujet va effectuer en préférant l'une des deux langues suivant les interlocuteurs, les situations, les thèmes, les contextes etc.

D'autres facteurs comme l'âge, les aptitudes générales et intellectuelles des sujets, les langues impliquées sont également à l'origine de variations importantes dans les formes prises par le bilinguisme.

L'existence de ces variations est souvent évoquée dans la littérature. On retrouve ainsi régulièrement l'ajout d'un qualificatif précisant à quel type de bilinguisme il est fait référence (pour une revue voir par exemple Bruyninckx & Harmegnies, 1993; Cummins & Swain, 1986; Lucchini, 2002). Ces distinctions portent généralement :

- sur les modes d'acquisition du bilinguisme (*e.g.* bilinguisme scolaire *vs.* naturel) ou les conséquences de l'acquisition des langues (*e.g.* bilinguisme additif *vs.* soustractif⁹³),
- sur les aspects temporels de l'acquisition. On retrouve ainsi la distinction entre bilinguisme simultané et consécutif ou entre bilinguisme précoce et tardif (à partir du moment où la seconde langue n'intervient dans l'environnement enfantin qu'après l'âge de 3 ou 4 ans⁹⁴). La durée qui sépare l'apprentissage des deux langues a en effet un impact important sur l'acquisition du système linguistique (Berthoud, 1982) et peut ainsi légitimer l'association souvent faite entre l'apprentissage d'une langue seconde et le bilinguisme. Par exemple, Lucchini considère que « *la présence conjointe d'une langue d'origine et d'une langue seconde introduit l'enfant dans le monde du bilinguisme* » (Lucchini, 2002, p. 19).

Le bilinguisme peut donc être envisagé selon une acception assez large comme lié à un phénomène de contact entre deux langues et d'usage alterné de ces dernières. Dans cette perspective, il concerne notamment :

- les personnes issues de foyer bilingue, appartenant à une communauté bilingue⁹⁵, ou encore, immergées régulièrement dans différentes communautés linguistiques.
- les personnes immigrées ou d'origine immigrée, dont la langue d'origine, (employée plus ou moins strictement dans le cercle familial) n'est pas la langue du pays de résidence.
- Les enfants ou adolescents scolarisés dans des établissements bilingues ou internationaux.

⁹³ Dans le bilinguisme additif les deux langues se maintiennent et se cumulent, alors que dans le bilinguisme soustractif une des deux langues (en général la langue première) cède la place à l'autre.

⁹⁴ D'après McLaughlin (1984 cité par Perregaux, 1994).

⁹⁵ Nous entendons par communautés bilingues, les groupes sociaux vivant dans des zones géographiques où subsiste l'usage de plusieurs langues (*e.g.* le Liban, la Bretagne, l'Alsace etc.).

L'usage du terme « bilingue(s) » dans la littérature actuelle confirme d'ailleurs que ce concept scientifique est particulièrement vaste et finalement très relatif. Les exemples de recherches, cités ultérieurement, n'y font pas exception.

1.2 De la langue étrangère à la langue seconde

La dénomination d'apprenant en langue étrangère est généralement réservée à l'acquisition scolaire et relativement réduite d'une langue, sans objectif principal d'en faire une langue vernaculaire. En revanche, la langue seconde renvoie aux cas où la langue considérée est une langue officielle du pays de résidence, ce qui limite beaucoup son application (Lucchini, 2002). Une telle restriction exclurait par exemple les pays où les langues d'enseignement ne sont pas officielles. Cette différence, établie en référence au contexte social, n'a aussi que peu de réalité psycholinguistique et ne semble pas relever d'une distinction en termes d'apprentissages. Ainsi, Smith (1994) suggère que le terme de langue seconde peut être considéré comme un terme générique recouvrant toutes les langues autres que les langues premières. Cela inclut donc les langues étrangères et les langues qui ne sont pas langues maternelles mais néanmoins parlées par la communauté. Il est en effet courant de voir apparaître les sigles LS/E (ou ESL/EFL en anglais)⁹⁶, le terme « langue seconde » permettant finalement, de positionner les langues par rapport à la langue maternelle ou première, en particulier celles apprises consécutivement et ayant un autre statut.

Il faut cependant tenir compte, toujours selon Smith (1994), du type d'environnement dans lequel prennent place les apprentissages ainsi que des autres langues possédées par l'apprenant.

Ces précisions permettent de mieux comprendre pourquoi la littérature utilise régulièrement ces termes pour rendre compte des comportements linguistiques de diverses populations, confrontées à ce contact de langues. Nous considérerons ainsi que les données présentées maintenant appartiennent toutes à ce vaste champ de recherche qu'est l'acquisition/apprentissage des langues secondes. Toutefois, pour éviter les amalgames, nous prendrons la précaution de stipuler les caractéristiques des populations étudiées.

⁹⁶ LS/E = Langue Seconde / Etrangère. ESL/EFL = English Second language / English Foreign Language

2. Apprentissage des langues secondes et concepts associés

Depuis longtemps, l'apprentissage des langues secondes a subi un engouement particulier de la part des chercheurs de multiples disciplines (linguistes, psychologues, psycholinguistes, didacticiens, pédagogues etc.). Cet attrait réside, de notre point de vue, en au moins deux enjeux particulièrement déterminants : le premier est que l'apprentissage des langues secondes permet d'aborder la question fondamentale de la part d'inné et d'acquis dans l'acquisition du langage. Le second est davantage social et politique. Les difficultés des populations migrantes sont un sujet sensible et ne peuvent être niées. La maîtrise des langues autochtones constitue souvent un élément décisif dans les politiques d'intégration des pays occidentalisés où l'on observe, dans l'histoire contemporaine, des flux migratoires importants. Sans forcément participer à ces débats, la recherche sur l'apprentissage des langues secondes, qu'elle se soit focalisée sur le langage oral et/ou écrit a permis une évolution des points de vue et a soulevé plusieurs interrogations.

2.1.1 L'évolution des points de vue : de l'analyse contrastive à l'interlangue

Depuis les années 60, le champ de l'acquisition des langues secondes a été traversé par les conceptions théoriques classiques sur l'acquisition du langage (voir notamment Akiki, 1994; Cuq, 1996; Matthey, 1996; Smith, 1994; Vogel, 1995). Parmi les points de vue qui ont émergé dans le domaine de l'acquisition/apprentissage des langues secondes certains ont des moyens d'investigation différents, basés sur des bases théoriques parfois opposées.

On retrouve ainsi le courant de l'analyse contrastive qui s'est développé sur une base behavioriste (e.g. Lado, 1964). L'appropriation de la langue seconde se résume globalement par le transfert des habitudes acquises en langue première, le rôle de cette dernière est donc fondamental, en tant que source fondatrice des structures transférables. Ainsi, Lado (1964) distingue deux types de transfert :

- * le transfert positif : dans le cas où les structures des deux langues sont similaires, le transfert d'une expression de L1 sera correct en L2.
- * Le transfert négatif : lorsque les deux structures des deux langues n'ont pas de rapport, le transfert engendre une expression fautive.

Les travaux sur l'analyse contrastive se focalisent ainsi sur la comparaison des structures de la langue source (la langue première) et de la langue cible (la langue seconde) pour démontrer que l'on peut prédire et expliquer l'apparition des erreurs des apprenants. Quant au processus d'apprentissage, il se borne à faire disparaître les transferts négatifs.

En rejetant les positions skinneriennes et sur la base du « *Language Acquisition Device* » de la grammaire générative transformationnelle, plusieurs auteurs ont suggéré l'intérêt de se focaliser sur le système apprenant et d'analyser les erreurs commises. Ces dernières ne sont plus perçues comme négatives mais reconnues comme inévitables et normales ; elles ont même une utilité pour l'apprentissage. Partisan de ce courant, Corder (1967, 1980) indique que les erreurs systématiques (à distinguer des fautes qui sont des erreurs de performances) reflètent les compétences provisoires des apprenants, à un stade donné, et donc ses stratégies de construction du savoir interlingual. Corder fonde sa position sur le concept de « dialecte idiosyncrasique ». Ce concept renvoie au système transitoire de règles, reconstruites par l'apprenant lui-même. Plus exactement, ces connaissances instables résultent de l'élaboration et de la vérification d'hypothèses⁹⁷ qu'émet le sujet lorsqu'il tente de donner du sens et d'organiser l'information de l'input linguistique auquel il est exposé⁹⁸. Allant dans le même sens, Selinker (1972) est le premier à avoir utilisé le terme d'interlangue, largement repris par la suite. Son but est d'expliquer la particularité des apprenants en langue seconde de produire des énoncés qui ne sont ni identiques à ceux de la langue de départ, ni conformes à ceux de la langue cible. Il suggère que la langue propre à l'apprenant (*i.e.* l'interlangue) correspond à un système psycholinguistique dynamique régit par plusieurs types de processus. On retrouve des positions identiques dans les positions de Nemser (1971), qui parle pour sa part de « système approximatif », et dont le caractère déviant, par rapport aux règles de la langue cible, varie selon le niveau de compétence de l'apprenant.

L'analyse des erreurs a donné lieu à de nombreuses études descriptives qui ont pu classer les processus, sous-jacents aux erreurs commises, en différentes catégories selon les propositions théoriques des auteurs. Parmi les principaux on retrouve notamment :

- les processus de transfert entre L1 et L2, à l'origine d'erreurs d'interférences,
- les processus de surgénéralisation du matériel de la L2, faisant apparaître des erreurs de généralisation,
- les processus de transfert d'apprentissage pouvant être liés à une incompréhension lors des apprentissages de la L2.

Les modélisations du concept d'interlangue des années 70 ont depuis été complétées, clarifiées ou modifiées, notamment en ce qui concerne les processus d'apprentissage. Smith (1994) relève en effet plusieurs lacunes dans ces modélisations, par exemple sur la façon de conceptualiser les

⁹⁷ Ces processus d'élaboration et de vérification d'hypothèses sont directement issus de la théorie Chomskyenne (*cf.* Chomsky, 1965).

⁹⁸ Corder distingue l'*input* (*i.e.* les matériaux linguistiques auxquels l'apprenant est exposé) de l'*intake* qui correspond aux données effectivement saisies par l'apprenant, et donc susceptibles d'être prises en compte.

stratégies et processus sous-jacents aux apprentissages. La distinction entre ces deux derniers termes n'est en effet pas toujours claire, de même que la façon dont les connaissances linguistiques sont construites. Par ailleurs, l'interprétation des comportements linguistiques des apprenants peut parfois être ambiguë.

Ainsi, les orientations théoriques de la psychologie cognitive dans le domaine de l'apprentissage des langues secondes ont surtout concerné la mise en place des connaissances linguistiques et les facteurs influençant, au cours du développement, cette mise en place.

De nombreux chercheurs ont ainsi souligné l'importance du degré d'analyse ou de contrôle impliqué pour appréhender les connaissances et les processus d'acquisition (Bialystok, 1986, 1999; Ellis, 2002; Leow, 2000; MacWhinney, 1997; McLaughlin, 1987; Robinson, 1996, 1997).

Le rôle de la langue maternelle a également été réinvesti, mais davantage en termes de relations entre les langues des apprenants, bien que la notion de transfert translinguistique reste très employée (Cummins, 1987, 1989; Cummins & Swain, 1986; Odlin, 1989)

Ces deux perspectives seront présentées plus précisément ultérieurement, notamment en relation avec l'apprentissage de la lecture.

2.1.2 Existe-t-il une période critique pour les apprentissages ?

Une des idées dominantes dans le domaine de l'acquisition des langues secondes est l'existence d'une période critique, c'est à dire d'un âge particulièrement favorable à l'apprentissage des langues. L'hypothèse d'une période critique pour l'acquisition des langues (et pas seulement des langues secondes) trouve son origine dans les travaux en neurophysiologie de Penfield et Roberts (1959)⁹⁹. Elle a, par la suite été très largement relayée par Lenneberg (1967). Sur la base d'un postulat biologique et d'une conception innéiste stricte du fonctionnement cognitif, il est affirmé qu'après l'âge de 7/9 ans, et de façon encore plus catégorique après la puberté, l'acquisition naturelle d'une langue devient très difficile voire impossible. Les explications et arguments avancés sont d'abord neurofonctionnels. Après 9 ans, le cerveau humain a perdu sa plasticité initiale. Très rapidement, et de façon complète à la puberté, la spécification hémisphérique du langage est stabilisée. En cas d'aphasie, la large récupération du langage et la prise en charge par d'autres régions cérébrales, chez les enfants jusqu'à 8/9 ans mais beaucoup moins ensuite, confirment cette hypothèse d'une plasticité initiale qui s'estompe à l'approche de l'adolescence.

Cependant, les conclusions tirées de ces faits doivent être nuancées. Les observations n'indiquent pas de frontière nette à partir d'un âge où toute récupération serait impossible Les

⁹⁹ Cités par Hagège (1996).

contre-exemples existent de bonnes compétences linguistiques acquises par des apprenants tardifs. Par ailleurs, de nombreux facteurs non contrôlés peuvent intervenir. Enfin, cette affirmation d'un âge critique pour acquérir la langue première, fondée sur l'existence de processus de maturation et de capacités innées d'acquisition des langues, ne peut s'appliquer à la langue seconde sauf à postuler une équivalence entre l'acquisition de la langue maternelle et l'apprentissage d'une langue seconde.

Certes, de nombreuses recherches ont souligné le déclin précoce des compétences à discriminer et à reproduire les sons qui n'appartiennent pas à la langue maternelle (pour une revue, voir par exemple Gaonac'h, 2006 ou Hagège, 1996). De même, la relation inverse entre l'âge au moment de l'immersion et la maîtrise du système linguistique d'une langue seconde a été vérifiée à plusieurs reprises en ce qui concerne la prononciation mais aussi le jugement grammatical (*e.g.* DeKeyser, 2000; Johnson & Newport, 1989; Oyama, 1979). Outre la fossilisation progressive de la perception, certains auteurs avancent également qu'à l'adolescence, les transformations psychologiques caractéristiques de cet âge et la crainte de l'erreur qui s'installe, sont d'autres arguments en faveur d'un meilleur apprentissage des langues durant l'enfance (*cf.* Hagège, 1996).

Toutefois il convient de relativiser cette position. Tout d'abord, il est évident que les conditions d'apprentissage ainsi que les modalités d'exposition à la langue sont extrêmement variables selon les individus¹⁰⁰, et se révèlent être, bien souvent, des facteurs confondus à celui de l'âge d'immersion ou même, constituer des déterminants plus forts que ce dernier (Mägiste, 1992). De plus, le déclin n'est pas une preuve stricte de l'existence d'une période critique, cette dernière devant faire apparaître une cessation d'apprentissage ou un changement net de la courbe des apprentissages (Bialystok, 2002a; Hakuta, Bialystok, & Wiley, 2003). On retrouve aussi dans la littérature de nombreuses données soutenant que l'acquisition d'une langue seconde est plus efficace chez les adolescents ou les adultes que chez les enfants, les premiers bénéficiant de capacités cognitives permettant de mettre en place des stratégies spécifiques à l'acquisition d'une nouvelle langue. Par exemple les adolescents ou les enfants les plus âgés apprennent plus vite les règles morphologiques et syntaxiques que les plus jeunes (Ervin-Tripp, 1974; Fathman, 1975) tout au moins au début (Krashen, Long, Scarcella, 1979 ; Snow, Hoefnagel-Höle, 1978)¹⁰¹.

Une alternative au débat entre les partisans de l'existence d'une période critique et de leurs détracteurs semble toutefois possible si l'on tient compte des résultats de DeKeyser (2000) et d'Harley et Hart (1997). Ces deux études tendent à montrer que les apprentissages des langues

¹⁰⁰ Il peut être ajouté l'extrême variabilité des aspects motivationnels.

¹⁰¹ Cités par Gaonac'h (2006).

secondes par les enfants et les adultes ou adolescents ne relèvent pas exactement des mêmes mécanismes¹⁰².

DeKeyser (2000) a réalisé une recherche auprès d'adultes de langue maternelle hongroise, arrivés aux Etats-Unis à des âges variés allant de 1 à 40 ans. Il observe, outre la corrélation négative entre l'âge et les scores de jugement de grammaticalité, une corrélation entre ces derniers scores et le niveau d'aptitudes verbales, chez les plus âgés mais pas chez les plus jeunes (arrivés avant 16 ans). Selon les auteurs, le test d'aptitudes verbales donne un indice des habiletés analytiques explicites. Cette capacité à mener une réflexion consciente sur la langue serait un élément particulièrement déterminant lors de l'apprentissage des langues secondes chez les adultes les plus tardivement immergés. Pour ces apprenants, la seule façon d'acquérir un niveau de compétences grammaticales similaires à ceux des natifs est d'utiliser leurs habiletés analytiques et de résolution de problèmes parce que les mécanismes d'apprentissages implicites, facilement mis en place par les enfants, ne sont plus disponibles et accessibles. A partir des résultats obtenus à une tâche de découverte des structures linguistiques dans une nouvelle langue¹⁰³, par des lycéens de 11^e année, Harley et Hart (1997) avaient également montré que même avec un apprentissage scolaire, les capacités analytiques n'interviennent pas sur l'efficacité en langue seconde (vocabulaire, compréhension et production orale), chez les apprenants ayant suivi un programme d'instruction bilingue précoces (en 1^e année), mais constituent le principal prédicteur chez les enfants ayant débuté l'apprentissage du français plus tardivement, vers l'âge de 12 ans.

Certains de ces résultats ont toutefois été critiqués. Bialystok (2002a) reproche à l'étude de DeKeyser (2000) de ne pas tenir compte des distributions des scores (chez les apprenants précocement exposés à la langue, la distribution est tronquée, ce qui explique l'absence de corrélation) et considère les corrélations obtenues chez les apprenants tardivement exposés comme tautologiques. Ces recherches ont toutefois le mérite d'envisager l'intervention de processus d'apprentissage implicite et explicite en langue seconde.

¹⁰² Cette possibilité avait déjà été envisagée par Johnson et Newport (1991) qui pensaient que les apprentissages réalisés avant 7ans, dépendent de capacités naturelles et générales alors qu'au delà, ils sont davantage liés aux caractéristiques propres aux individus et aux situations d'apprentissage.

¹⁰³ Dans cette épreuve, un corpus linguistique dans une nouvelle langue est présenté avec une traduction en anglais. Pour évaluer comment les participants détectent le fonctionnement de cette langue, ces derniers doivent déterminer parmi plusieurs propositions comment exprimer correctement un énoncé donné en anglais.

2.1.3 Acquisition naturelle ou apprentissage formel des langues secondes ?

Depuis les années 70-80, cette question de l'apprentissage implicite versus explicite, bien que formulée différemment, est un élément essentiel des débats concernant les méthodes pédagogiques à promouvoir (MacWhinney, 1997). Plusieurs théories ont en effet tenté de définir ce qui est le plus déterminant pour un apprentissage efficace des langues secondes : l'acquisition naturelle ou l'enseignement formel des règles linguistiques (pour une revue sur la question des méthodes voir par exemple Gaonac'h, 2006).

L'auteur le plus favorable aux acquisitions naturelles et incidentes des langues secondes, et sans doute le plus influent, est Krashen (1985). Dans sa théorie du « modèle du moniteur », cet auteur, influencé par la linguistique nativiste, postule l'indépendance des processus subconscients de structuration des phénomènes langagiers (l'acquisition naturelle) et des connaissances explicites des règles de la langue relevant des apprentissages formels. Il revendique que les acquisitions naturelles (les mêmes que celles amenant l'enfant à maîtriser, sans apprentissages particuliers, sa langue maternelle) sont centrales, alors que les apprentissages formels et explicites sont périphériques ou secondaires et relèvent de l'utilisation d'un moniteur externe. L'utilité de ce moniteur n'est pas totalement nié mais n'apparaît toutefois que dans certaines conditions rarement satisfaites (le moniteur ne peut intervenir que s'il en a le temps, si l'apprenant est en mesure de se focaliser exclusivement sur la forme linguistique et si le besoin de connaître la règle est élevé). Cette instance de contrôle n'intervient, le cas échéant, qu'en fin de procédure, comme correcteur des productions de l'apprenant. De telles affirmations ont conduit à soutenir les thèses, parfois radicales, qu'une langue peut ou doit être acquise sans être enseignée, que l'enseignement explicite des règles de grammaire est inutile voire contre-productif et qu'il n'est pas nécessaire d'adapter les *inputs* linguistiques pour qu'il y ait acquisition (*cf.* MacWhinney, 1997 ; Gaonac'h, 2006). La théorie de Krashen a été très critiquée, entre autres pour le manque de clarté et de pertinence des concepts employés ainsi que pour l'impossibilité qui en découle de mettre en relation les différents types d'apprentissage (McLaughlin, 1987).

A l'opposée de cette conception, plusieurs auteurs ont avancé que la conscience des règles de la langue ont une influence positive sur les processus de traitements linguistiques et la maîtrise du système linguistique en langue seconde (pour une revue, voir Ellis, 2002 ou; Norris & Ortega, 2000). Dans cette perspective théorique, Schmidt (1992) a développé le modèle dit « *noticing hypothesis* » qui postule le rôle crucial du niveau de conscience sur les processus d'acquisition d'une langue seconde. Schmidt considère qu'aucun apprentissage ne peut se réaliser sans porter attention à la forme, et que cette attention rend possible que l'*input* linguistique soit converti en un *intake* intégré au système de l'apprenant. L'explicitation des règles permet donc à l'apprenant

de se focaliser sur les caractéristiques spécifiques de la langue seconde et cette prise de conscience est, dans cette perspective, une condition nécessaire et suffisante aux apprentissages. De nombreuses recherches ont d'ailleurs relevé que les instructions explicites apparaissent plus efficaces qu'un mode d'apprentissage implicite, en particulier pour la maîtrise des règles de grammaire, et que cette efficacité semble, en outre durable (De Graaff, 1997 ; DeKeyser, 1995 ; Leow, 2000 ; Robinson, 1996 ; Rosa & Leow, 2004 ; Rosa & O'Neil, 1999).

Dans une perspective proche de celle de Schmidt, d'autres recherches ont approfondi les relations entre apprentissages implicites et explicites dans l'objectif de mettre en évidence que l'explicitation des règles sert de socle aux apprentissages implicites, dépendants de l'exposition, en les amplifiant et en les accélérant (Ellis, 2002).

Ellis (1993) s'est intéressé à l'acquisition des règles de mutations phonologiques en gallois ¹⁰⁴ par des adultes anglophones recevant différents types d'instruction. Un des groupes d'apprenants reçoit un enseignement explicite sur les principes sous-jacents à l'utilisation de ces règles de mutation en dehors de tout contexte lexical alors qu'un autre groupe reçoit en plus de cette instruction décontextualisée, une séquence d'apprentissage sur l'application de ces règles à l'aide d'exemples. Les évaluations portent sur les connaissances explicites des règles de mutations (épreuves d'explicitation de la règle à appliquer selon la construction phrastique proposée) et les compétences implicites à juger du caractère correct de phrases (épreuve de jugement grammatical spontané). Ces évaluations intègrent en fin de procédure, l'utilisation de constructions nouvelles mais relevant des mêmes principes

Les résultats permettent de constater que les apprenants ayant reçu un enseignement explicite décontextualisé ont de solides connaissances explicites mais ont toutefois du mal à les appliquer pour identifier de façon spontanée les phrases mal formées. Les apprenants ayant également bénéficié d'une plus grande pratique, relative à l'application de ces règles, montrent simultanément de bonnes appréciations implicites et explicites de la structure phonologique du gallois. Leurs connaissances des règles se révèlent par ailleurs plus facilement généralisable à des constructions nouvelles. L'auteur en conclut qu'un enseignement explicite et abstrait des règles, allié à une pratique intensive, propice à la mise en place de compétences généralisables, constitue un apprentissage, long mais efficace, et favorise la transformation des connaissances explicites en connaissances implicites.

Michas et Berry (1994) apportent des conclusions similaires à partir de leur étude sur l'apprentissage des règles de prononciation en grec, par des adultes anglophones.

¹⁰⁴ Ces mutations phonologiques s'effectuent, dans l'expérience d'Ellis (1993) sur les consonnes initiales et sont dépendantes de la construction phrastique dans lequel le mot est inséré.

DeKeyser (1997) et Robinson (1997) envisagent, pour leur part, la pratique ou l'exposition à la langue en termes d'automatisation des traitements.

DeKeyser (1997) montre qu'après avoir enseigné les mêmes règles morpho-flexionnelles et le même vocabulaire d'une langue artificielle à plusieurs groupes d'apprenants, la pratique de ces règles et de ce vocabulaire permet de réduire le temps de réponse et le nombre d'erreurs en compréhension ou de production de phrases, même en paradigme de double tâche. Cependant, les étudiants ne progressent que dans les activités où la pratique a été effectuée (*i.e.* en compréhension lorsque la pratique concerne la compréhension, en production, lorsque c'est cette dernière activité qui a été pratiquée). Par ailleurs, le groupe qui a reçu moitié moins de pratique mais dans les deux activités cumulées, ne montrent pas de résultats plus faibles que les autres groupes ayant reçu une pratique plus ciblée. Ce résultat permet d'envisager les effets à long terme d'une pratique systématique et l'évolution asymptotique des processus d'automatisation. L'auteur en conclut que la pratique conduit à une automatisation graduelle et relève d'une compétence spécifique. Il reste prudent quant aux implications de cette étude car rien ne prouve que cette automatisation soit équivalente à celle mise en place naturellement par les locuteurs natifs.

Robinson (1997) envisage les choses encore différemment, en manipulant les instructions données à des adultes japonais apprenant l'anglais langue seconde. Les apprenants ayant reçu des instructions explicites les focalisant sur la structure syntaxique de l'anglais (construction de phrases à double objet) parviennent davantage que les apprenants exposés à un enseignement implicite (incident à une tâche de mémorisation de l'ordre d'occurrence des mots dans les phrases) à juger de la grammaticalité de phrases utilisant un matériel linguistique nouveau. Par ailleurs, leurs résultats ne sont pas plus faibles que ceux de l'autre groupe, lorsqu'il s'agit de juger de la grammaticalité de phrases déjà rencontrées. Ces résultats suggèrent que les apprentissages explicites augmentent la possibilité de généraliser les règles à d'autres types de phrases. D'autre part, une telle explicitation n'empêche pas que le jugement sur des formes connues puisse s'effectuer de manière implicite, par une récupération directe en mémoire des formes auxquelles l'apprenant a été exposé. L'auteur y voit une interaction possible entre les apprentissages explicites basés sur le repérage conscient des règles et les connaissances implicites basées sur la mémorisation.

Cependant, il faut souligner certaines différences dans la façon dont ces études considèrent la relation entre connaissances implicites/explicites et automatisation. Il semble, en effet, que Ellis (1993) assimile l'apprentissage implicite à l'automatisation d'Anderson (1983) pour qui l'activité contrôlée requiert de moins en moins d'attention et s'accélère en se répétant, permettant ainsi le

passage des connaissances explicites aux connaissances implicites. Cette position diffère de celle proposée par Logan (1988) et à laquelle se réfère l'étude de Robinson (1997), qui envisage les automatismes comme une récupération directe en mémoire d'une réponse, issue de la « procéduralisation » de connaissances déclaratives. Les positions de DeKeyser (1997) à ce sujet sont moins claires. Si l'auteur distingue explicitement les termes de procéduralisation et d'automatisation et considère que les automatismes relèvent de la pratique, leur relation avec les notions d'apprentissage explicite ou implicite n'est pas précisée.

Par ailleurs, à l'instar des remarques de MacWhinney (1997), il est risqué d'amalgamer instruction explicite et apprentissage explicite. Il est en effet illusoire de penser que les contenus des enseignements sont toujours et totalement intégrés par l'apprenant. De plus, même sans instruction explicite, un apprenant peut parfois extraire certaines règles sous-jacentes aux informations linguistiques qu'il rencontre.

Il convient aussi de préciser que les études sur cette question des apprentissages ont exclusivement été réalisées auprès d'apprenants adultes et dans des domaines linguistiques circonscrits à la grammaire et à la prononciation (Gaonac'h, 2006). Il s'agirait ainsi de développer davantage les recherches auprès d'enfants et dans différents domaines linguistiques de façon à appréhender la diversité et la complexité des mécanismes en fonction de l'âge et du domaine étudié.

Enfin, il est évident que les conclusions de ces recherches en situation de laboratoire doivent être confirmées dans des conditions plus écologiques pour être généralisables aux enseignements mis en place dans les classes (DeKeyser, 1997).

Globalement, et bien qu'il soit encore difficile de tirer des conclusions claires à partir des éléments présentés précédemment, il semble toutefois envisageable de souligner deux points :

- La notion de période critique ne peut être interprétée de la même façon qu'elle a pu l'être à l'origine, en biologie ou en éthologie. Par ailleurs elle ne semble pas représenter, en langue seconde, un déterminant aussi crucial que pour l'acquisition de la langue maternelle (Gaonac'h, 2006).

- Les apprentissages explicites sont nécessaires en langue seconde, mais doivent, toutefois, interagir avec une pratique intensive de la langue dont dépendra le caractère fluent des mécanismes de production et de compréhension des apprenants (Ellis, 2002). Ce qui apparaît donc important dans les modes d'acquisition des langues secondes se rapporte aux phénomènes de fréquence d'exposition à la langue, de généralisation de règles linguistiques et d'automatisation de leur application. Par conséquent, la distinction entre apprentissage

implicite et explicite, leurs mécanismes et leurs relations s'avèrent déterminants pour appréhender les processus d'acquisition des langues secondes à l'oral mais également à l'écrit.

3. Lecture, capacités métalinguistiques et bilinguisme

A l'issue des recherches précédentes, il est facile d'admettre que l'acquisition d'une langue seconde à l'oral diffère de façon importante de celle de la langue maternelle. Pourtant, cela est moins évident pour ce qui concerne l'apprentissage de la lecture. Dans le domaine de l'écrit, les auteurs ont en effet considéré, qu'à partir du moment où l'écrit est maîtrisé en langue première, ce qui est propre à la lecture des langues alphabétiques, c'est-à-dire l'utilisation du code orthographique pour représenter le langage, ne devrait pas poser de problème puisque le principe sous-jacent aux règles de conversions graphophonologiques, voire même, dans une certaine mesure, une partie du code orthographique lui-même, est connue (*cf.* Gaonac'h, 2000 ; Lucchini, 2002). En outre, les recherches sur la lecture en langue seconde se sont surtout intéressées aux processus de haut niveau ou ascendants, lors de la compréhension de texte (Goodman, 1971 ; Smith, 1971, cités par Gaonac'h, 2000). Ces dernières décennies ont en effet été très largement dominées par des conceptualisations *top-down* de la lecture en langue seconde (*cf.* Koda, 1996, 1999). Il en résulte une pauvreté des connaissances scientifiques sur les traitements linguistiques mis en œuvre lors de la lecture en langue seconde. Toutefois, les recherches s'intéressant aux processus descendants lors la lecture et au bilinguisme au sens large font apparaître certaines spécificités du mode d'acquisition bilingue, spécificités sur lesquelles les chercheurs ne semblent pas toujours s'accorder.

3.1 Structure et fonctionnement du lexique bilingue

Avant de s'intéresser spécifiquement à l'accès au lexique bilingue lors de l'identification des mots écrits, les chercheurs se sont tout d'abord interrogés, de façon générale, sur le mode de représentation en mémoire des deux langues chez les bilingues (pour une revue détaillée, voir par exemple De Groot, 1992). Là encore plusieurs hypothèses ont été émises :

- celle de représentations partagées ou interdépendantes : les équivalents de traduction des mots des deux langues partagent une même représentation, de nature conceptuelle ou supralinguistique,
- celle de représentations séparées ou indépendantes : des représentations distinctes et autonomes coexistent pour chaque langue connues du sujet.

Kolers (1963) est un des premiers auteurs à avoir formulé précisément ces deux hypothèses et plusieurs recherches ont tenté de mettre en évidence leur véracité à l'aide de diverses tâches de mémorisation (rappel libre, association libre, reconnaissance lexicale, compréhension et production etc. ; *e.g.* Kintsch, 1970; Kolers, 1963, 1966).

Dans les années 80, les recherches se sont dirigées vers des modélisations plus complexes en prenant en compte différents niveaux de représentations et en envisageant comment les connexions entre les représentations des mots dans les deux langues peuvent s'organiser (voir par exemple Beauvillain & Grainger, 1987; Grainger *et al.*, 1991). Potter, So, Eckhart et Feldman (1984) émettent l'hypothèse de la médiation conceptuelle, pour déterminer à quels niveaux de représentation il existe un système partagé ou séparé pour les deux langues. Cette hypothèse suggère que le niveau des représentations conceptuelles ou profondes est organisé de façon à intégrer les deux langues alors que les représentations formelles, ou superficielles, sont, par essence, spécifiques. En d'autres termes, il existe, pour Potter *et al.* (1984), un niveau de représentation conceptuelle commun aux deux langues mais un niveau de représentation formelle indépendant et propre à chacune d'elles. Ces affirmations découlent de leur propre expérience et de celles d'autres chercheurs (Caramazza & Bones, 1980 ; Shannon, 1982)¹⁰⁵, menée auprès d'adultes bilingues et montrant que les temps de réponse, pour juger l'appartenance des mots à une catégorie sémantique, présentée dans la même langue ou dans l'autre langue que le mot cible, sont équivalents. De plus, leur étude révèle que les sujets bilingues mettent approximativement le même temps pour traduire un mot donné dans l'autre langue que pour nommer le mot à partir de son image. Le traitement des formes lexicales et la traduction d'une langue à l'autre, passent donc obligatoirement par une représentation conceptuelle commune.

Kirsner, Smith, Lockhart, King et Jain (1984), en menant une expérimentation à l'aide du paradigme d'amorçage, envisagent, en revanche, l'existence de nœuds lexicaux entre les deux langues, intégrés dans un seul réseau de représentations lexicales (hypothèse de l'association lexicale). Les auteurs se basent sur les résultats, en français et en anglais, d'un effet similaire de facilitation sémantique interlangue et intralangue (*i.e.* lorsque le mot-amorce est relié sémantiquement au mot-cible que ce dernier appartienne à la même langue -arbre/feuille- ou à l'autre langue -*tree*/feuille-). Cependant, Grainger et Beauvillain (1987) ont émis certaines réserves sur les conclusions des auteurs. Ils argumentent que les temps de présentation de l'amorce sont trop longs pour démontrer strictement l'existence de relations structurelles entre les représentations lexicales dans les deux langues. Ils ont ainsi tentés de répliquer ces résultats

¹⁰⁵ Cités par Potter *et al.* (1984).

en situation d'amorçage rapide et masqué, de façon que les sujets ne perçoivent pas consciemment la relation entre les mots et ne puissent mettre en place des stratégies prédictives ou de traduction. Il ressort de leur étude une facilitation intralangue mais pas interlangue. Les auteurs soutiennent donc l'existence de deux systèmes lexicaux indépendants qui ne sont reliés que par l'intermédiaire d'un système conceptuel.

En admettant qu'il existe des réseaux représentationnels séparés entre les deux langues, il faut alors préciser à quel moment la langue est identifiée lors des traitements lexicaux. Cela revient donc à déterminer comment se mettent en place les procédures d'accès au lexique bilingue et comment s'effectue la sélection de l'une ou de l'autre langue. Grainger (1987, 1991) a pu repérer trois niveaux où la sélectivité de la langue peut s'opérer :

- Soit l'information sur la langue intervient à l'entrée et dirige les informations sensorielles vers les représentations lexicales de l'une ou de l'autre langue. Dans ce cas, les attentes du sujet à l'égard de la langue, en fonction du contexte, sont déterminantes. Les données empiriques allant dans le sens de cette présélection sont, par exemple, celles démontrant une plus grande difficulté à traiter des phrases mixtes, utilisant des mots des deux langues (Soares & Grosjean, 1984). En revanche, les effets de facilitation, pour les mots dans les deux langues, en paradigme d'amorçage écrit rapide par des homographes interlangues (*e.g.* « four » existe en anglais et en français), suggèrent que les deux acceptions des homographes sont disponibles indépendamment du contexte (Beauvillain & Grainger, 1987). Ainsi, ces derniers auteurs, considèrent que le ralentissement des traitements lors des mélanges de langue (*cf.* Soares & Grosjean, 1982), peut tout aussi bien être attribué à des procédures opérant pendant la phase de sélection (*i.e.* à un niveau lexical plutôt que prélexical).

- Soit la sélection s'effectuerait pendant la sélection des candidats lexicaux, une fois que l'information sensorielle a permis la détermination de candidats potentiels issus des deux langues. Dans cette hypothèse, la première phase de décodage est commune aux deux langues mais par la suite l'analyse lexicale tient compte de la langue. Pour Grainger et Beauvillain (1987), cette procédure semble la plus valide. Ils montrent, toujours avec le paradigme d'amorçage, que dans le cas où le mot cible se caractérise par des éléments orthographiques spécifiques à une langue (*e.g.* voix en français ou white en anglais), le ralentissement de la décision, lié à la présence d'une amorce dans l'autre langue, n'est plus observé.

- Soit la sélection ne s'effectue qu'à la sortie, c'est-à-dire après l'identification du mot. Le processeur lexical ne tient dans ce cas, pas compte de la langue à laquelle les représentations lexicales appartiennent, cette information s'activant en dernier lieu. Cette hypothèse, utile pour expliquer les interférences régulières entre les langues, ne permet toutefois pas de comprendre

l'influence du contexte ni pourquoi les changements de langue en paradigme d'amorçage classique (*e.g.* time/lire) ralentissent l'accès au lexique.

Ainsi, la deuxième hypothèse d'une sélection au cours de l'accès lexical, permet donc d'expliquer les interférences interlangues tout en rendant compte des sélections rapides de la langue, guidées par des indices linguistiques ou le mode langagier, que les mécanismes d'identification soient envisagés en termes de recherche séquentielle ou d'activation interactive (Grainger, 1987, 1991).

A partir de ses premiers résultats et dans des travaux plus récents, Grainger (2002) s'est attaché à adapter le modèle connexionniste d'identification des mots écrits de McClelland et Rumelhart (1981) au mode bilingue. Le modèle d'activation interactive bilingue (modèle de simulation BIA) qu'il propose avec ses collaborateurs (*e.g.* Van Heuven, Dijkstra, & Grainger, 1998) se centre en particulier sur l'organisation et l'activation des unités lexicales orthographiques¹⁰⁶. L'architecture du système lecteur proposé (*cf.* figure 7) intègre des représentations orthographiques de l'ensemble des mots, connectées à une unité de traitement codant la langue, à l'aide de « nœuds langagiers ». Ces nœuds langagiers permettent de contrôler le niveau d'activation des représentations lexicales dans chaque langue : si le mot correspond à la langue 1, les liens de nature excitatrice vers la langue 1 vont être activés. Les nœuds langagiers permettent également un contrôle ascendant de type inhibiteur vers toutes les représentations des mots de l'autre langue.

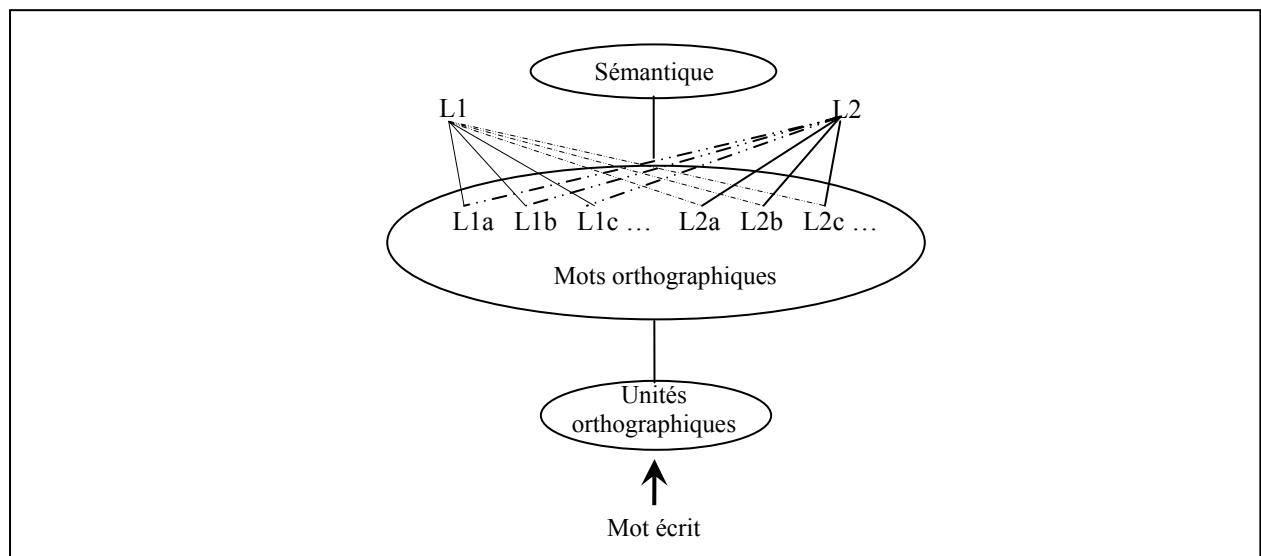


Figure 7: le modèle d'activation interactive bilingue de Grainger (2002)

¹⁰⁶ Les auteurs partent du principe qu'il existe très souvent un grand partage de représentations sémantiques entre les deux langues bien qu'il y ait aussi des structures sémantiques distinctes pour chaque langue lorsqu'il n'existe aucun équivalent de traduction. Par ailleurs, ils précisent que, pour deux langues partageant le même alphabet, les représentations des lettres ne seront guère modifiées en situation bilingue, hormis en ce qui concerne éventuellement le traitement des accents.

Les résultats empiriques favorables à cette architecture ont été obtenus à l'aide du paradigme de démasquage progressif (Van Heuven *et al.*, 1998) ou d'amorçage orthographique interlangue (Bijeljac-Babic, Biardeau, & Grainger, 1997) en situation expérimentale réelle ou en simulation par ordinateur. Dans une des expériences de Van Heuven *et al.* (1998), des adultes bilingues néerlandais-anglais devaient reconnaître un mot néerlandais, ayant un nombre plus ou moins élevé de voisins orthographiques soit en néerlandais (voisinage intralangue), soit en anglais (voisinage interlangue). La présentation des mots est faite en alternance avec un masque qui diminue en vitesse. Les temps de réponses, aussi élevés pour les mots ayant de nombreux voisins interlangues qu'intralangues, suggèrent qu'il existe un réseau intégré et interdépendant de représentations orthographiques pour chaque langue mais également que l'activation de la langue ne s'effectue pas avant le traitement du mot lui-même. Ces propositions sont en accord avec les résultats obtenus par Bijeljac-Babic *et al.* (1997) qui vérifient, en situation bilingue anglais-français, les effets inhibiteurs de l'amorçage orthographique sur la décision lexicale, que l'amorce soit ou non dans la même langue que le mot-cible.

Grainger (2002) s'interroge également sur la façon dont ce modèle peut rendre compte de la prise en charge bilingue des règles de décodage graphophonologique, variable entre les langues. Il envisage l'hypothèse d'une activation en parallèle des différentes correspondances plausibles dans les deux langues (*e.g.* la suite de lettre /coin/, correspondant à un mot dans les deux langues pourrait simultanément activé les représentations phonologiques correspondant aux sons [kɔɪv] ou [kɔE]). Cette hypothèse semble soutenue par les résultats de Brysbaert, Van Dyck et Van de Poel (1999) obtenus auprès d'adultes bilingues flamand-français. Dans leur expérience, les auteurs observent l'effet de deux types d'amorçage rapide sur l'identification de mots français : celui avec des amorces jouant le rôle de pseudo-homophones intralangue (*i.e.* dont la prononciation française correspond à celle du mot-cible, ex : fain/faim) et celui avec des amorces jouant le rôle de pseudo-homophones interlangue (*i.e.* dont la prononciation en flamand correspond à celle du mot cible, ex : soer/sourd). Les résultats révèlent un effet facilitateur même avec les pseudo-homophones flamands, ce qui suggère que les bilingues activent les correspondances phonologiques possibles dans les deux langues même s'ils n'ont à reconnaître que des mots d'une seule langue. Jared et Kroll (2001) en sont arrivés aux mêmes conclusions auprès de bilingues anglais-français et français-anglais à partir d'une tâche de lecture à voix haute.

Bien que le modèle d'activation interactive bilingue (Grainger 2002 ; Van Heuven *et al.*, 1998) soit un des seuls modèles connexionnistes développés sur l'identification des mots en mode bilingue, les confirmations se limitent aux procédures d'accès mises en place chez l'adulte, sans systématiquement préciser si, entre les deux langues connues des

participants, l'une domine sur l'autre (la seule recherche citée à avoir tenu compte de cette variable est celle de Jared & Kroll, 2001). Par ailleurs, les validations empiriques restent, de notre point de vue, trop parcellaires. Si les aspects orthographiques sont les plus investis, l'hypothèse sur les traitements de nature phonologique nécessite d'être consolidée et les aspects morphologiques sont totalement délaissés. Enfin, ce modèle n'envisage pas les cas de bilinguisme où les deux langues ne partagent pas le même système orthographique. Même en restant dans la sphère des langues alphabétiques, il est légitime de croire que lorsque les systèmes d'écriture diffèrent totalement, le choix de la langue s'effectue à l'étape initiale, dès le traitement perceptif des lettres.

3.2 Stratégies d'interprétation chez les bilingues

Bien que nous ayons déjà souligné les différences entre la compréhension orale et écrite, notamment en termes de contrôle syntaxique (paragraphe 2.3 du chapitre 1), les travaux axés sur le contact des langues 1 et 2, dans le cadre du modèle de compétition de Bates et MacWhinney (1989)¹⁰⁷, seront abordés car ils représentent un champ de recherche bien développé et toutefois connecté aux mécanismes de compréhension en lecture. Les aspects méthodologiques rejoignent donc ceux des recherches développées en langue maternelle (emploi de tâches d'assignement du rôle thématique de mot, dans des phrases où certains indices -l'ordre des mots, les aspects morpho-flexionnels etc.- sont mis soit en convergence, soit en compétition ; cf. paragraphe 2.2 du chapitre 1).

Le modèle de Bates et MacWhinney (1989) est en effet apte à modéliser les stratégies d'interprétation chez les bilingues. Ces derniers se particularisent par leur nécessité d'acquérir l'usage des indices spécifiques aux deux langues et de distinguer les différences interlangues quant à la validité de ces indices pour attribuer telles ou telles fonctions aux mots de la phrase. Ainsi, plusieurs types de phénomènes, intervenant lors des traitements interprétatifs des phrases en langue seconde, peuvent être étudiés dans le cadre du modèle de compétition (pour une revue récente, voir Koda, 2005). Ceux ayant été les plus examinés sont :

- Les processus de transferts de la langue première vers la langue seconde : dans ce cas, les stratégies correspondant à l'usage de certains indices sont également appliquées lors de l'interprétation des phrases en langue seconde.
- L'apparition de stratégies sur les correspondances forme-fonction exclusivement appliquées dans le contexte de la langue seconde. Kail (1991) ajoute que les stratégies développées lors de

¹⁰⁷ Ce modèle est, nous l'avons vu précédemment, surtout consacré à l'oral.

l'acquisition de la langue seconde peuvent éventuellement remplacer les stratégies en langue première ou encore être assimilées dans un ensemble applicable aux deux langues.

Elle signale également que ces différentes stratégies ne sont pas mutuellement exclusives et peuvent intervenir à certains moments de l'acquisition. Les données empiriques suivantes prouvent l'exactitude de cette remarque.

McDonald (1987) a par exemple comparé l'usage d'indices syntaxiques (*e.g.* l'ordre des mots, la présence de prépositions, les flexions casuelles) et lexico-sémantiques (*e.g.* le caractère animé des noms) lors de l'assignement des rôles thématiques en langue seconde chez des bilingues anglais-néerlandais et néerlandais-anglais de différents niveaux (novice, intermédiaire et avancé). Ses résultats montrent que les deux groupes se basent sur les indices prédominants dans leur langue première respective : l'ordre des mots pour les bilingues de langue première anglais et les flexions casuelles pour les bilingues de langue première néerlandais. Cependant, plus leur compétence en langue seconde augmente, plus les indices utilisés se déplacent vers ceux auxquels se réfèrent les natifs de la même langue (*i.e.* les indices flexionnels pour les bilingues anglais-néerlandais testés en néerlandais et l'ordre des mots pour les bilingues néerlandais-anglais testés en anglais). Les données de Kilborn et Ito (1989) confortent également ces résultats.

Dans une autre étude, impliquant des adultes de langue première anglaise, chinoise ou coréenne, apprenant le japonais, Koda (1993) confirme les observations de Harrington (1987) auprès de bilingues japonais-anglais qui révélaient à la fois un transfert des stratégies de la langue première mais aussi l'acquisition des indices de la langue seconde. Koda (1993) montre que l'assignement du rôle d'agent est plus difficile en l'absence d'indices fiables en japonais (*e.g.* les marques flexionnelles des particules) quelle que soit la langue première. Cependant, lorsque l'indice de l'ordre des mots (qui est également un indice fiable en japonais) n'est pas disponible, la chute des performances est particulièrement visible chez les apprenants dont la langue première se caractérise par la dominance de ce même indice (*i.e.* l'anglais et le chinois). Ces résultats rendent compte des interactions complexes unissant les traitements linguistiques à l'œuvre lors de l'interprétation des phrases dans les deux langues des sujets.

Gass (1987) approfondit également la problématique des transferts de stratégies d'une langue à l'autre en comparant l'usage d'indices comme l'ordre des mots et le contraste animé des noms, chez des bilingues italiens-anglais et anglais-italiens. En italien langue seconde, les bilingues de langue 1 anglais tendent à se fonder, comme les natifs italiens sur l'indice sémantique valide dans cette langue, c'est-à-dire le caractère animé des noms. En revanche, en anglais, les bilingues de langue 1 italien, à la différence des natifs, continuent à se référer à l'indice valide

dans leur langue maternelle (*i.e.* le caractère animé). Le transfert des stratégies de la langue première ne transparait donc pas de façon symétrique chez les deux groupes. Gass (1987) interprète finalement ces résultats comme une indication de la recherche de résolution de la compétition entre les deux langues. Lorsque les indices des deux langues sont en incongruité, les principes fondamentaux, dirigeant la mise en place de stratégies d'interprétation universelles, impliquent préférentiellement l'usage d'indices sémantiques.

Sasaki (1991) réplique les résultats de Gass (1987) auprès d'adultes japonais apprenant l'anglais et d'adultes américains apprenant le japonais. Là encore, les locuteurs anglophones se comportent lors de l'interprétation de phrases japonaises comme les natifs, en usant d'indices lexico-sémantiques comme le caractère animé, alors que les apprenants japonais transfèrent en anglais les stratégies de leur langue première. Sasaki (1991) argumente que le transfert ne prend place qu'à partir d'une langue peu centrée sur la syntaxe comme le japonais. Cependant, les résultats de Miao (1981) contredisent cette interprétation. Cet auteur avait en effet montré que les stratégies d'apprenants anglophones basées sur la syntaxe (*i.e.* l'ordre des mots, typiques de la langue anglaise), peuvent se transférer lors de la compréhension de phrases en mandarin. Sasaki (1991) interprète cette divergence en fonction de la distance entre les langues que peut percevoir l'apprenant. Le mandarin partage avec l'anglais l'ordre canonique SVO et l'absence de particules casuelles, ce qui n'est pas le cas du japonais. Le fait que la distance entre l'anglais et le japonais soit relativement plus perceptible qu'entre l'anglais et le chinois expliquerait que le transfert y soit également plus difficile. Sasaki reconnaît toutefois que cette notion de perception de la distance entre les langues n'est pas suffisamment opérationnalisée et est difficile à appréhender. Par ailleurs, les recherches n'indiquant par précisément le niveau des apprenants, le mode d'enseignement, le temps et la nature de l'exposition à la langue seconde, il est aussi possible que ces facteurs entraînent certaines différences dans les résultats. Signalons également que la référence aux aspects lexico-sémantiques dans les études de Gass et Sasaki, peut relever du fait que les phrases proposées sont dissonantes, à cause des manipulations effectuées sur les indices disponibles (notamment l'absence d'éléments syntaxiques typiques ; Koda, 1993). Les apprenants peuvent alors privilégier la recherche d'informations sémantiques pour donner du sens à des séquences de mots non naturelles. D'autre part, dans ces deux mêmes expériences, les indices les plus fiables dans les langues secondes (*i.e.* l'accord sujet/verbe en italien et le marquage casuel en japonais) n'ont pas été étudiés, de façon à obtenir une correspondance maximale entre les structures des langues comparées (L1 et L2). Mais en leur absence, le rôle déterminant du caractère animé des noms peut être amplifié.

En résumé, malgré qu'il puisse parfois y avoir plusieurs interprétations possibles pour une même observation, ces études montrent la diversité et la spécificité des stratégies syntaxiques et sémantiques pour interpréter les phrases en langue seconde. Pour reprendre les termes de Koda (2005) « l'interprétation de phrases en langue seconde implique une interrelation complexe entre plusieurs variables, incluant des principes universaux, la structure des correspondances forme-fonction de la langue première et les propriétés syntaxiques de la langue seconde » (p 114, notre traduction). Il apparaît toutefois nécessaire de développer les investigations chez l'enfant et sur un nombre de langues plus diversifié que l'anglais, le japonais et quelques autres langues.

3.3 Les capacités métalinguistiques des bilingues

Les recherches plus axées sur le bilinguisme infantin ont mis en évidence un autre aspect important, qui peut être mis en relation avec l'acquisition des langues écrites : celui de l'impact du contact des langues sur le développement cognitif en général et sur les compétences métalinguistiques en particulier.

Plusieurs auteurs ont avancé que la confrontation et l'emploi de plusieurs systèmes linguistiques, engendre une distanciation précoce entre les aspects formels et sémantiques du code, et donc de meilleures capacités d'abstraction et de manipulation conceptuelle ainsi qu'une facilité à porter attention aux aspects formels du langage (*e.g.* MacLaughlin, 1991 ; Peal & Lambert, 1962 ; Vygotsky, 1934 ; pour une revue détaillée voir entre autres Bialystok, 1991, 2005, 2007 ; Cummins & Swain ;1986, Perregaux, 1996). Il a ainsi été observé que les enfants bilingues perçoivent précocement la relation arbitraire entre les mots et leur signification (*e.g.* Cummins, 1978 ; Ianco-Worall, 1972). D'autres recherches ont également attesté la plus grande flexibilité des capacités cognitives des bilingues (*e.g.* BenZeev, 1977 ; Bialystok, 2005 ; Pearl & Lambert, 1962).

Plus spécifiquement, Vygotsky (1934) explique que la possibilité d'exprimer la même pensée dans deux langues différentes rend l'enfant capable de voir sa langue comme un système particulier parmi d'autres, ce qui l'amène à prendre conscience des opérations linguistiques propres aux langues. En d'autres termes, l'avantage des bilingues concernerait plus particulièrement le domaine des compétences métalinguistiques, dans le sens où les formes linguistiques se révèlent plus saillantes donc plus facilement accessibles à la conscience.

Bialystok (1991, 1993, 1999, 2001) considère pour sa part l'avantage bilingue en se basant sur la distinction de deux composantes :

- la composante d'analyse, renvoyant aux capacités à se représenter de façon de plus en plus explicite et abstraite la structure linguistique des langues.

- la composante de contrôle, renvoyant aux compétences à porter attention aux aspects spécifiques de ces représentations notamment en situation ambiguës ou trompeuses.

Pour Bialystok (1991, 1993, 1999, 2001), les changements des représentations mentales sont attribuables au développement de ces deux composantes, développement qui permet de comprendre la progression des simples compétences conversationnelles à un usage intentionnel des règles linguistiques lors de la lecture et la résolution de problèmes métalinguistiques. A l'issue de différentes expériences qu'elle a pu mener en anglais langue seconde, Bialystok avance que l'avantage bilingue concerne principalement les procédures de contrôle mais pas la dimension d'analyse. Par exemple, elle montre (Bialystok, 1986) qu'à 5, 7 et 9 ans, les bilingues équilibrés parviennent mieux que les monolingues à juger de la grammaticalité de phrases asémantiques, tâche évaluant les habiletés à porter attention à la forme plutôt qu'à la signification. Cependant, elle n'obtient pas cette supériorité des bilingues aux épreuves de corrections d'agrammaticalité, mesurant pour leur part les capacités d'analyse. Les bilingues seraient ainsi moins distraits par les dimensions non pertinentes lors de la résolution de certaines tâches linguistiques mais aussi non linguistiques (par exemple, une tâche de classification de cartes avec changement de catégories ; Bialystok, 1999), capacités dépendantes des processus de contrôle inhibiteur relevant des fonctions exécutives (Bialystok, 2007). La distinction entre analyse/contrôle, faite par Bialystok, a l'avantage de décrire les processus fondamentaux que les bilingues mettent en place aisément et d'établir les relations entre développement cognitif et métalinguistique. Cependant, cette distinction n'est pas toujours exploitée par les autres chercheurs.

Il faut enfin souligner que ces visions positives du bilinguisme n'ont pas toujours été adoptées (*e.g.* Saer, 1923 ; Darcy ; 1963 ; Palmer, 1972¹⁰⁸) et certaines recherches n'ont d'ailleurs pas permis d'aboutir à de telles conclusions. Nous présenterons donc tout d'abord quelques données empiriques vérifiant cet avantage bilingue sur certaines dimensions des habiletés linguistiques associées à la lecture. Les recherches qui ne sont pas parvenues à démontrer cette facilitation seront ensuite développées.

¹⁰⁸ Les deux premiers auteurs étant cités par Bialystok (2005), le dernier par Perregaux (1994).

3.3.1 Le bilinguisme facilite le développement métalinguistique

Dans une étude suisse réalisée auprès d'enfants de familles immigrées de tous horizons, Perregaux (1994, 1995) montre que, malgré leurs difficultés en compréhension et production lexicale en français, les bilingues prélecteurs de GSM réussissent aussi bien les tâches de segmentation phonémique que les monolingues, et même mieux celles portant sur les pseudo-mots. L'auteur retrouve également des performances significativement plus élevées en segmentation lexicale chez les bilingues. En fin de 1^e année, bilingues et monolingues présentent le même niveau en lecture de mots. Grâce à leurs habiletés métalinguistiques, les bilingues ont pu faire face à leur manque de maîtrise orale.

De même, Eviatar et Ibrahim (2000)¹⁰⁹ observent que les bilingues russe-hébreu (parlant russe avec leur famille et hébreu à l'école) en dernière année de maternelle obtiennent de meilleures performances que les monolingues hébreux aux tâches phonémiques (*i.e.* répéter le premier ou dernier phonème d'un mot). Cette différence n'est plus significative en 1^e année, mais pour ce niveau scolaire, les résultats à une tâche de suppression de phonème ou de syllabe réitèrent la supériorité des bilingues par rapport aux monolingues. Ces derniers résultats sont également retrouvés par Ibrahim, Eviatar & Aharon-Peretz (à paraître), auprès d'une même population de 1^e année. Comme dans l'étude de Perregaux (1994, 1995), ces auteurs remarquent que les connaissances lexicales en hébreu sont plus faibles chez les bilingues mais pas les performances en lecture de mots non vocalisés et de textes vocalisés. Ces habiletés sont expliquées, chez les bilingues comme les monolingues, par les compétences phonologiques ainsi que, mais chez les monolingues seulement, les compétences en vocabulaire. Ces résultats suggèrent, encore une fois, que la conscience phonologique permet aux bilingues de compenser leur manque de vocabulaire, lors de la lecture de textes.

Campbel et Sais (1995) corroborent cet avantage bilingue dans le domaine phonologique avec des enfants bilingues italien-anglais scolarisés dans une école maternelle bilingue. En anglais, les jeunes bilingues sont plus performants que les monolingues à détecter les intrus phonologiques (*i.e.* le mot, parmi trois, ne commençant pas par le même son) ainsi qu'à supprimer un morphème de mots composés (*e.g.* répéter *rain* dans *rainbow*) ou la première syllabe de pseudo-mots bisyllabiques. Bien que l'auteur considère la suppression morphémique comme une tâche

¹⁰⁹ Cette étude, de même que la suivante rapporte également les résultats d'enfants arabophones évalués en arabe de façon à déterminer l'impact de la diglossie (*cf.* paragraphe 1.1.1 du chapitre 3). Nous avons choisi de ne présenter dans ce paragraphe que les résultats comparant bilingues et monolingues en hébreu car il nous semble plus rigoureux d'analyser les performances dans une même langue et préférons traiter exclusivement dans cette partie du bilinguisme, le concept de diglossie ayant été préalablement abordé.

phonologique, les performances obtenues indiquent que l'avantage des bilingues concerne la conscience phonologique mais semble aussi s'appliquer aux connaissances morphologiques.

D'autres recherches se sont attachées à démontrer l'avantage des bilingues dans le domaine syntaxique et morpho-syntaxique. Par exemple, Cromdal (1999), réplique, en Suédois, une partie des résultats de Bialystok (1986) à l'aide d'épreuves similaires en comparant les résultats de monolingues, d'enfants partiellement bilingues (*i.e.* dont les compétences en vocabulaire en suédois sont plus faibles qu'en anglais et inférieures à celles des monolingues) et d'enfants hautement bilingues (*i.e.* dont les connaissances lexicales en suédois sont équivalentes à celles des monolingues)¹¹⁰. Comme dans l'étude de Bialystok (1986), les deux groupes bilingues, obtiennent des scores supérieurs aux monolingues lorsqu'il s'agit de juger la forme de phrases asémantiques, mais pour les phrases sémantiques, le jugement est seulement supérieur pour les enfants hautement bilingues. Le bilinguisme peut ainsi dans certains cas favoriser également le développement des compétences d'analyse métalinguistique. Les auteurs concluent globalement de leur étude que les représentations grammaticales des bilingues, même si elles révèlent parfois des hypothèses fausses (surtout dans le cas où la maîtrise de la langue seconde est encore imparfaite), atteignent un niveau explicite plus précocement que chez les monolingues.

Dans une étude menée en français et en Alsace sur le bilinguisme français-allemand, Demont (2001) observe également que les enfants de 1^e année, scolarisés sur un site bilingue, témoignent d'une conscience syntaxique plus élevée (évaluée par des épreuves de jugement et de correction de phrases dont les anomalies concernent l'ordre des mots et les affixes flexionnels) et d'une efficacité supérieure dans la reconnaissance des mots comparativement à leurs paires scolarisés dans une classe traditionnelle monolingue. En revanche, les capacités en compréhension de phrases n'apparaissent pas différencier les deux groupes. Il en est de même pour la conscience phonémique qui est liée à l'enseignement des correspondances graphophonologiques, enseignement réalisé dans toutes les classes. Il faut toutefois souligner qu'en l'absence d'évaluation de l'efficacité verbale dans les deux langues, il est difficile d'évaluer chez les bilingues, quelle est la langue première ou seconde. La description des situations des enfants, recevant un enseignement paritaire des deux langues à l'école mais semblant plutôt parler le français à la maison, laisse entendre que le français ne peut être considéré ici comme une langue seconde.

¹¹⁰ Les enfants rencontrés lors de cette étude ont entre 6 et 7 ans.

Ces différentes recherches ont ainsi observé une supériorité des bilingues sur certaines dimensions, notamment, sémantique, phonologique, morphologique et syntaxique, cet avantage ne semble toutefois pas attesté sur l'ensemble des tâches utilisées et à tous les âges.

3.3.2 Le bilinguisme ne facilite pas le développement métalinguistique

Contrairement aux recherches précédentes, Lucchini (2002) observe, chez des enfants d'origine sicilienne en début de 1^e année, des pourcentages de réussites aux épreuves d'identification de rime, de suppression de voyelle finale et de consonne initiale, largement plus faibles que ceux obtenus par d'autres recherches chez des enfants francophones plus jeunes (respectivement 49%, 53% et 16%). Cependant, cette étude n'a pas effectué de comparaisons bilingues/monolingues.

La récente recherche canadienne de Roy (2005) est à notre connaissance, la seule s'intéressant aux compétences morpho-dérivationnelles en comparant des enfants bilingues et monolingues de 1^e et 2^e année. Plusieurs épreuves morphologiques ont été administrées :

- une tâche de jugement de relation de mots (*e.g.* « est-ce que peur et peureux sont de la même famille ? »), évaluant les connaissances de la parenté entre les mots.
- une tâche de complèment d'énoncé avec choix de dérivés ou pseudodérivés (*e.g.* compléter la phrases « pourrais-tu me ___ ton livre », par *couter*, *coutable* ou *couteur*), évaluant les connaissances syntaxiques impliquées dans la compréhension du processus de dérivation.
- une tâche de plausibilité lexicale (*e.g.* qu'est ce qui ressemble le plus à un mot français *arrêteur* ou *arrétoume* ?), évaluant la sensibilité à la construction morphologique des mots français et aux contraintes distributionnelles qui s'y applique.

Les résultats obtenus révèlent une infériorité des enfants bilingues à l'épreuve de plausibilité lexicales mais une absence de différences entre les deux groupes pour ce qui concerne les épreuves de jugement et de complèment. Pour Roy (2005), tous les enfants ont ainsi une bonne compréhension de la famille morphologique et connaissent la catégorie grammaticale des suffixes courant. Cependant, il faut souligner que les scores aux deux épreuves évaluant ces compétences sont proches de l'effet plafond, en particulier chez les monolingues, ce qui peut biaiser les résultats obtenus. Les bilingues ont en revanche plus de difficultés à repérer la présence de suffixes dans les pseudo-mots et à déterminer si la combinaison base+affixe est légale. Roy (2005) interprète cette infériorité par un contact moins régulier avec le français, ce qui limite l'évolution de la connaissance des morphèmes et les contraintes distributionnelles qui s'appliquent aux règles de construction morphologique. Cette interprétation est cohérente avec le

fait que l'on retrouve généralement chez les bilingues ou apprenants en langue 2 un vocabulaire moins étendu.

Plusieurs études ont enfin investi le domaine des connaissances syntaxiques et morphosyntaxiques des enfants bilingues comparés aux monolingues.

Chiappe & Siegel (1999) ont réalisé une recherche en langue anglaise auprès d'enfants locuteurs punjabi¹ de 1^e année, exposés à l'anglais depuis deux ans. A une épreuve de complèment de phrases évaluant une sensibilité grammaticale, les élèves allophones obtiennent des scores plus faibles que leurs pairs monolingues anglophones. Les scores à plusieurs tâches métaphonologiques (identification, localisation, suppression et substitution phonémique) sont cependant équivalents pour les deux groupes. Toujours en anglais langue seconde, Da Fontoura et Siegel (1995) ont également observé cette faiblesse dans le domaine syntaxique avec des enfants d'origine portugaise plus âgés (en 4^e, 5^e et 6^e année) suivant un programme d'apprentissage de lecture en langue maternelle. Pourtant, la lecture de mots et de pseudo-mots en anglais, atteint un niveau d'efficience similaire à ceux des monolingues anglais.

Demont, Louvet et Nocus (2001), à l'aide des mêmes instruments de mesure et sur une population pourtant proche voir identique à celle de Demont (2001), n'obtiennent pas les mêmes résultats que cette dernière. En effet, les enfants scolarisés en 1^e année d'école internationale, recevant 5 ou 6 heures d'enseignement de l'allemand, commettent plus d'erreurs de jugement sur la grammaticalité de phrases en français. Par ailleurs, les enfants issus d'écoles bilingues français-allemand, n'ont pas de meilleurs résultats que les monolingues, ni en jugement et correction de phrases agrammaticales, ni en suppression phonémique, ni en lecture de mots, de pseudo-mots et en compréhension de phrases.

Au même niveau scolaire, Armand (2000) compare les compétences métalinguistiques d'enfants francophones monolingues et d'enfants immigrés au Canada ou d'origine immigrée, le niveau en lecture des deux groupes étant équivalent. Elle obtient une infériorité des bilingues dans la réussite aux épreuves de jugement et de correction d'agrammaticalité (portant sur l'ordre des mots et les morphèmes flexionnels) et non aux épreuves métasémantiques (jugement de la longueur de mots) et métaphonologiques (détection de rime et suppression de phonème). En outre, les bilingues ont des compétences langagières orales plus faibles que les monolingues (test de vocabulaire et de compréhension orale) que l'auteur explique par la moindre expérience linguistique. D'après les analyses de corrélation et de régression, le traitement des aspects syntaxiques de la langue seconde ne peut se manifester pleinement à cause de ces lacunes en

¹ Langue officielle de la région du Pendjab située au Nord de l'Inde.

compétences langagières. Enfin, elle en conclut un rôle fondamental des capacités métaphonologiques. Du fait de leur prédictibilité sur la reconnaissance des mots écrits, elles tiennent une place primordiale dans l'apprentissage de la lecture en L2 et permettent de compenser les faiblesses en termes d'habiletés langagières orales et de compétences métasyntaxiques.

3.3.3 Synthèse et perspectives

Les résultats de ces différentes recherches apparaissent ainsi contradictoires. Les dernières qui ont été cités ont surtout soulevé les insuffisances des bilingues concernant le traitement syntaxique et, sans surprise, le niveau d'efficacité verbale. La prise en compte des aspects phonologiques semble relativement préservée (hormis dans les propositions de Lucchini, 2002) sans pour autant faire émerger une habileté particulière. Plusieurs auteurs (Armand, 2000 ; Ibrahim *et al.*, à paraître ; Perregaux, 1994, 1995) ont par ailleurs suggéré l'importance de la conscience phonologique en langue seconde pour pallier le manque de vocabulaire. Ce déficit peut d'ailleurs en partie expliquer certaines difficultés à traiter les aspects morphologiques et syntaxiques en langue seconde (Armand, 2000; Roy, 2005).

La recherche de Lesaux et Siegel (2003) apporte des données intéressantes à ce propos. Les auteurs ont mené une étude longitudinale sur l'anglais langue seconde en suivant des enfants immigrés ou d'origine immigrée de différentes langues premières, de la dernière année de maternelle à la 2^e année. En dernière année de maternelle, les résultats des bilingues aux différentes tâches métalinguistiques sont soit proches de celles des monolingues (*e.g.* en identification de phonème, de syllabe et de suppression phonémique) soit inférieurs (*e.g.* en répétition de pseudo-mots, détection de rime et à l'épreuve syntaxique de complèment de phrase). En 2^e année, les bilingues ont rattrapés leur retard, sauf dans le domaine syntaxique où une légère différence subsiste. Ils obtiennent même de meilleurs scores en identification de mots et en rapidité de lecture de mots et de pseudo-mots. Par ailleurs, cette étude permet de rendre compte des similarités et différences en terme de facteurs impliqués en début d'apprentissage de la lecture en langue première et en langue seconde. Les résultats révèlent une contribution importante de la conscience phonologique (en particulier la détection de rime) aux compétences en lecture de mots et en compréhension de phrases écrites en anglais langue seconde. En revanche, les connaissances lexicales (évaluées par une épreuve de dénomination rapide) et la conscience syntaxique ne semblent pas intervenir, contrairement à ce qui apparaît chez les monolingues. Ce point résulte, selon les auteurs, de l'impact important des instructions axées sur le décodage, à l'étape initiale de l'apprentissage. Pourtant, bien que les résultats de Geva et

Yaghoub Zahed (2006) se rapprochent en de nombreux point de ceux de l'étude de Lesaux et Siegel (2003), ils font toutefois apparaître une contribution importante des compétences orales aux niveaux d'efficacité en lecture de mots et de texte, chez les bilingues de 2^e année. De même, pour ce qui concerne la composante de compréhension en lecture, les résultats de l'étude de Carlisle, Beeman, Davis et Spharim (1999) menée auprès de bilingues espagnol-anglais (américains d'origine mexicaine) de 1^e, 2^e et 3^e année suggèrent également la contribution de la conscience phonologique, mais toutefois moins que les connaissances lexicales. L'importance du niveau d'efficacité verbale a d'ailleurs été souligné par les résultats obtenus par Proctor, Carlo, August et Snow (2005) sur une même population bilingue espagnol-anglais de 4^e année, et par ceux relevés par Droop et Verhoven (2003) auprès d'enfants de 3^e et 4^e année, d'origine marocaine et turque et de langue seconde néerlandaise.

La synthèse de Bialystok (Bialystok, 2002b, 2007) permet aussi d'aboutir à une meilleure appréhension du problème. Après avoir souligné les divergences des données développées ces dernières années sur le bilinguisme, elle signale qu'il ne peut y avoir de réponse claire à un problème aussi complexe et multifactoriel. Elle tente toutefois d'identifier trois types d'habiletés pouvant intervenir dans l'acquisition de l'écrit et mettre en relation les langues connues des apprenants bilingues. Elle avance que les compétences orales ne sont généralement pas une habileté privilégiée chez les enfants bilingues et que ce déficit peut parfois compromettre l'apprentissage de la lecture. En revanche, de nombreuses études ont pu démontrer que le développement des concepts sur l'écrit était influencé par le bilinguisme. Les enfants bilingues semblent en effet particulièrement conscients de l'invariabilité de l'écrit pour représenter une forme symbolique et des modifications des correspondances forme-référent entre les systèmes des diverses langues écrites. La compréhension de ces concepts est sans conteste, une condition préalable à la lecture. Ainsi, les effets positifs les plus généralement visibles du bilinguisme dans le domaine de l'apprentissage de la lecture sont relatifs à l'établissement de ces concepts sur l'écriture, son caractère invariable mais arbitraire. Enfin, Bialystok avance que les données concernant les compétences métalinguistiques, en particulier la conscience phonologique ne permettent pas d'établir un effet favorable, systématique et généralisable du bilinguisme. Il semble toutefois possible de voir apparaître dans beaucoup d'études une précocité phonologique bilingue en maternelle, qui disparaît lorsque que l'apprentissage de la lecture commence. L'introduction d'un apprentissage formel de la lecture semble en effet aboutir à une égalisation des expériences et fait disparaître l'avantage initial des bilingues. D'autres facteurs tels que les instructions formelles auraient ainsi une plus grande importance que le bilinguisme sur le développement métaphonologique. Quoi qu'il en soit, ces compétences sont nécessaires à

l'acquisition de la lecture, en langues secondes comme en langue maternelle (Armand & Lefrançois, 2003 ; Geva et Yaghoub Zahed, 2006 ; Lesaux et Siegel, 2003).

Bialystok (2002b, 2007) prend aussi soin de souligner que les recherches dans ce domaine n'utilisent pas toujours les mêmes modes d'évaluation, ne portent pas toujours sur les mêmes unités ni sur des sujets dans les mêmes situations. Cette remarque est également valable pour les recherches sur la morphologie et la syntaxe. De même, les critiques relatives aux mesures des connaissances, développées lors du paragraphe sur les langues premières (*cf.* partie 6 du chapitre 4) peuvent être renouvelées ici (par exemple la difficulté d'attester la nature implicite ou explicite des tâches, d'isoler rigoureusement les connaissances morphologiques ou syntaxiques lors des tâches de jugement ou de complètement etc.). De plus, il nous paraît regrettable de ne pas mettre en avant le fait que les apprenants ont ou non appris à lire dans l'autre langue. Certaines études citées concernent en effet des lecteurs en langue première et d'autre des apprenants débutant directement la lecture en L2. Enfin, certaines de ces études n'apportent pas cette information.

Il est aussi important de constater que, de toute évidence, la majorité des recherches comparant bilingues et monolingues ne se soucient guère de la langue maternelle des sujets et intègrent, dans un même groupe bilingue, des apprenants de langue maternelle très disparate. Pourtant, l'impact sur le développement métalinguistique, des expériences en langue première, orientées en fonction des caractéristiques de ces dernières, ne peut être omis. Par exemple, Bialystok, Majumder et Martin (2003) ont observé que les bilingues espagnol-anglais de 1^e et 2^e année, sont plus performants que leurs pairs monolingues de langue anglaise à une tâche de segmentation phonémique alors que les bilingues chinois-anglais sont moins performants que ces derniers. La plus grande similarité des structures phonologiques entre l'espagnol et l'anglais, et éventuellement la facilité d'accéder à une conscience phonologique en espagnol car cette langue utilise un système phonétique simple permettent d'expliquer que l'analyse du système phonologique soit facilement menée par les bilingues de langue première espagnol mais pas par ceux dont la langue première n'a pas de telles caractéristiques. Les différences individuelles ne semblent là encore pas explicables par le seul facteur « bilinguisme ».

Le cadre conceptuel proposé par Bialystok (2002b, 2007) envisage ainsi le rôle des habiletés développées en langue premières en fonction de leurs caractéristiques et les possibilités de les transférer lors de l'apprentissage d'une langue seconde. Dans sa conception, les relations entre langues 1 et 2 ne sont pas unidirectionnelles, ni l'ordre séquentiel de l'apprentissage de ces deux langues fixe et stable. Au contraire, la progression dans une langue influence la progression dans l'autre. Plusieurs hypothèses théoriques ont d'ailleurs été émises sur cette influence dans le

domaine de la lecture et de son apprentissage, en particulier celle de la langue première sur la langue seconde.

4. Le rôle des acquisitions en langue première lors de l'apprentissage de la lecture en langue seconde

Lors de l'apprentissage de l'écrit en langue seconde, les apprenants ne peuvent être considérés comme indemnes de toutes connaissances sur l'écrit, dans la mesure, bien sur, où l'apprentissage de la littéracie a été amorcé en L1. Ainsi, de nombreux chercheurs ont mis en avant l'influence des expériences de la lecture en langue première sur l'apprentissage d'une langue seconde. Cet intérêt pour les apprentissages préalables n'a pas pour objectif de nier que les apprentissages propres à la langue seconde, interviennent pour parvenir à la maîtrise du système linguistique (e.g. Segalowitz & Segalowitz, 1993). Mais pour appréhender les spécificités des apprenants en langues secondes, et eu égard aux différences obtenues par Bialystok *et al.* (2003) entre bilingues de langues premières très différentes, il apparaît inévitable de tenir compte des connaissances développées par le biais de l'apprentissage de la lecture en langue première pour comprendre certaines difficultés ou atouts des apprenants en langue seconde.

4.1 Modélisations de la relation entre langue première et seconde

Dans les années 80, Cummins (1979) a développé un modèle théorique pour rendre compte des relations entre les compétences en lecture, en langue première et seconde. Dans les dernières versions de l'hypothèse de l'interdépendance développementale (Cummins, 1987, 1989, 1991 ; Cummins & Swain, 1986), deux types de compétences, associées à des situations de communication différentes, sont distingués :

- les compétences communicationnelles interpersonnelles de base (*basic interpersonal communicatives skills*), correspondant à la capacité communicative qui s'acquiert en situations discursives hautement contextualisées.
- les compétences cognitives et linguistiques académiques (*cognitive/academic language proficiency*) fortement requises lors de la lecture ou des activités linguistiques, cognitivement complexes et décontextualisées. Ainsi, elles se rapprochent clairement des compétences métalinguistiques que l'écrit nécessite. Ce sont ces compétences qui, une fois développées en langue première, servent de base à l'acquisition de l'écrit en langue seconde.

Il est en effet avancé que les langues 1 et 2 sont interdépendantes car elles sont toutes deux sous la dépendance de ces « compétences cognitivo-linguistiques » dénuées des aspects de surface et des fonctions de communication. En d'autres termes, ces compétences peuvent, du fait de leur

fondement cognitif et de leur dégagement des aspects purement formels, être sollicitées tout autant lors d'une activité de lecture en L1 qu'en L2. Elles représentent donc les compétences communes aux différentes langues ou encore transférables d'une langue à l'autre. Cummins soutient ainsi que la réussite en lecture en langue seconde dépend donc largement des compétences en littéracie acquises préalablement en langue première.

Cummins (1991, Cummins & Swain, 1986) évoque aussi la notion de seuil minimal d'efficacité linguistique, à atteindre dans les deux langues, pour que les compétences linguistiques et cognitives et donc la maîtrise de l'écrit se développent harmonieusement. L'existence de ce seuil, qui ne peut être déterminé de façon absolue, permet de comprendre les divergences précédemment évoquées entre les effets positifs ou négatifs du bilinguisme ou encore la distinction entre bilinguisme additif ou soustractif.

D'autres auteurs ont par la suite repris cette notion d'interdépendance entre langue première et langue seconde sous des formes un peu différentes.

Lucchini (2002) parle par exemple de « la langue de référence » définie comme « *la langue dans laquelle nous développons toutes les fonctions propres aux langues, y compris la fonction métalinguistique* » (Lucchini, 2002, p 79). Pour elle, ce qui distingue alors les bons apprenants en L2 des moins bons, concerne la construction de cette langue de référence. Par exemple, les enfants qui ont compris, dans leur langue maternelle, le principe des correspondances graphophonologiques pourront l'exploiter lors de l'apprentissage de la lecture en L2. Il ne sera alors question que d'apprendre un nouveau code de conversion et de nouveaux signifiants mais dont la logique sous-jacente aura été assimilée. Ainsi, chez les apprenants d'une L2, la base de cette construction est la langue maternelle puisqu'elle a fait l'objet d'un apprentissage préalable mais, réciproquement, l'apprentissage de la L2 va renforcer ces capacités de traitement linguistique. Le lien entre les deux langues est donc biunivoque.

Spark et ses collaborateurs (voir par exemple Kahn-Horwitz, Shimron, & Sparks, 2005, 2006; Sparks & Ganschow, 1993ab; Sparks, Patton, Ganschow, Humbach, & Javorsky, 2006) évoquent une « *Linguistic Coding Differences Hypothesis* » pour expliquer que les habiletés linguistiques de base (dans le domaine phonologique, orthographique, syntaxique et sémantique), aboutissant à la maîtrise de la lecture en langue première, ont un impact similaire sur l'acquisition de la lecture en langue seconde. Ainsi, les difficultés des faibles lecteurs en langue première, à porter attention aux aspects structuraux des mots écrits, se retrouveront en langue seconde.

Enfin, Geva et ses collaborateurs (*e.g.* Geva & Siegel, 2000; Geva, Wade-Woolley, Durgunoğlu, & Verhoeven, 1998; Geva, Wade-Woolley, & Shany, 1997; Gholamain & Geva, 1999) ; se basent sur l'hypothèse des « processus centraux » pour comprendre ce qui unit les langues des

apprenants¹¹¹. Selon leur point de vue, la maîtrise de la lecture dans toutes les langues, qu'elles soient première ou seconde, repose sur des mécanismes cognitifs et linguistiques communs, telles que les habiletés verbales, la mémoire verbale de travail, la rapidité et la précision des procédures d'accès lexical. Cependant, les auteurs ne nient pas que certaines composantes impliquées dans l'activité de lecture (par exemple les traitements visuels ou phonologiques) puissent être influencées par les variations de code entre les langues. Nous verrons de quelle manière dans la partie suivante.

Les validations empiriques de ces hypothèses renvoient généralement à la notion de transfert interlangue, bien que, parfois, ce terme ne soit, à notre avis, pas forcément le plus adéquat. Les recherches ont par ailleurs surtout investi le champ des habiletés en lecture, des connaissances orthographiques, phonologiques, et lexicales. Compte tenu du nombre important de recherches menées à ce sujet (pour une revue non exhaustive, voir Durgunoğlu, 2002 ou Lefrançois, 2001) la présentation qui va suivre ne pourra en relater qu'une infime partie, concernant les dimensions les plus pertinentes dans le cadre de cette étude.

4.2 Validations empiriques du transfert interlangue

Les études ayant montré, chez de jeunes apprenants, une corrélation importante entre L1 et L2 dans le domaine de la reconnaissance des mots écrits et de la conscience phonologique sont sans doute les plus nombreuses. Cependant quelques données, sur peu de langues, suggèrent également l'existence d'un transfert interlangue des compétences syntaxiques et de la compréhension en lecture

4.2.1 Les recherches sur des langues à écriture latine

Comeau, Cormier, Grandmaison & Lacroix (1999) ont réalisé une étude longitudinale auprès d'enfants anglophones de primaire (1^e, 3^e et 5^e année) en classe d'immersion précoce du français L2. Ces auteurs montrent non seulement que les compétences métaphonologiques (suppression de phonèmes) sont quasi symétriques en anglais et en français, qu'elles sont très fortement corrélées à la reconnaissance des mots dans les deux langues, mais aussi que le poids de prédiction de ces deux variables sur la reconnaissance des mots dans les deux langues est identique (83 % de variance expliquée en anglais et 80 % en français).

¹¹¹ L'hypothèse des « processus centraux » a en fait été développée par plusieurs chercheurs travaillant sur les traitements au cours de la lecture en langue maternelle.

Durgunoğlu et ses collaborateurs (Durgunoğlu, 1998, Durgunoğlu, Nagy & Hancin-Bhatt, 1993 ; Durgunoğlu, Mir & Ariño-Martí, 2002) ont mené une série de recherches auprès d'enfant latino-américains, parlant l'espagnol dans la famille et scolarisés dans des écoles ayant mis en place un programme d'éducation bilingue de transition¹¹². Leurs résultats font apparaître, en 1^e année, une corrélation entre la lecture de mots anglais et la lecture de mots espagnols ($r = .51$). De même, la conscience phonologique en espagnol (évaluée par une tâche de segmentation) explique à elle seule 47 % de la variance des résultats à une même tâche en anglais (Durgunoğlu, 1998). Toujours en 1^e année, les performances en lecture en anglais sont prédites par l'efficacité en lecture et la conscience phonologique en espagnol (75 % de variance expliquée Durgunoğlu *et al.*, 1993).

Concernant la conscience syntaxique et morpho-syntaxique, Durgunoğlu *et al.*, (2002) observent en 4^e année des corrélations entre les performances aux épreuves de jugement et de correction de phrases dont les erreurs portent sur les flexions ou l'ordre des mots dans les deux langues ($r = .44$). Les auteurs en concluent que les enfants qui sont capables d'analyser la structure syntaxique des phrases dans une langue sont aussi en mesure de mettre en place une telle analyse dans une autre langue.

Da Fontoura et Siegel (1995) ont confirmé certains de ces résultats auprès d'enfants de 4^e, 5^e et 6^e année, parlant le portugais dans le cadre familial et l'anglais dans le cadre scolaire mais suivant un programme intégrant l'apprentissage de la lecture dans la langue d'origine. Il apparaît là encore un lien entre le niveau de lecture de mots et de pseudo-mots¹¹³ en portugais et en anglais (respectivement $r = .52$ et $r = .64$) ainsi qu'entre les capacités syntaxiques à compléter des phrases à l'aide du mot adéquat dans les deux langues ($r = .63$).

Enfin, Van Gelderen, Schoonen, De Glopper, Hulstijn, Simis, Snellings et Stevenson (2004) ont observé, auprès d'élèves néerlandais de 8^e année, apprenant l'anglais L2 depuis 1 an et demi, que la compréhension de textes en néerlandais langue première est un bon prédicteur de la compréhension de textes en anglais langue seconde.

4.2.2 Les recherches sur des langues utilisant différents alphabets

D'autres recherches se sont intéressées au contact entre des langues première et seconde plus éloignées et dont les systèmes d'écritures sont différents.

¹¹² L'objectif de ce programme américain est de faire en sorte que les enfants issus de milieu ethnique minoritaire suivent progressivement et à partir de la 3^e année l'ensemble des apprentissages scolaires en anglais. Durant les deux premières années de scolarité, les enfants reçoivent la plupart des instructions, dont celle concernant l'écrit, en espagnol, l'anglais étant étudié comme une langue seconde.

¹¹³ Les pseudo-mots utilisés respectent la structure phonologique de l'une ou de l'autre langue.

Abu-Rabia et Siegel (2003), auprès de trilingues arabe-hébreu-anglais de 8^e année¹¹⁴, ont aussi vérifié l'existence de corrélations entre les compétences en lecture de mots et de pseudo-mots dans les trois langues¹¹⁵, malgré les différences orthographiques et structurelles qui les séparent. Les corrélations entre les compétences dans ces trois langues concernent également les habiletés syntaxiques évaluées à l'aide d'épreuve de complètement de phrases. Abu-Rabia et Siegel (2002) en étaient arrivés à la même conclusion d'une interdépendance entre les langues concernant les processus de décodage et les habiletés syntaxiques, auprès d'élèves bilingues arabe-anglais¹¹⁶ scolarisés entre la 4^e et la 8^e année.

En revanche, Abu-Rabia et Siegel (2003) ne parviennent pas à faire clairement apparaître une relation entre les habiletés phonologiques (*e.g.* lors de tâches de choix entre un pseudo-mot -*seaf*- et pseudo-homophone -*seet*-) en particulier entre l'anglais et les deux autres langues sémitiques, du fait des différences entre les systèmes orthographiques des langues. De notre point de vue, l'instrument de mesure employé évalue plus l'accès au lexique phonologique, spécifique aux langues, que les connaissances du principe alphabétique, inhérent à toute langue alphabétique.

Aussi, les résultats concernant les habiletés orthographiques semblent suggérer une spécificité des traitements : les performances aux tâches de choix entre deux pseudo-mots dont un seul respecte les règles orthographiques de la langue (*e.g.* *filk-filv*) diffèrent en effet beaucoup selon les langues et ne sont pas corrélées entre elles. Comme cela a été souligné précédemment par les auteurs, les langues sont certes, toutes trois alphabétiques mais emploient des alphabets différents pour lesquels la façon de transcrire les voyelles ajoute un contraste supplémentaire entre l'anglais et les langues sémitiques.

Les travaux de Geva et ses collaborateurs (Geva & Siegel, 2000; Geva *et al.*, 1998; Geva *et al.*, 1997) ont également concerné l'acquisition de la lecture en hébreu vocalisé par des locuteurs anglais, scolarisés dans des écoles juives canadiennes, en suivant ces enfants de la 1^e à la 5^e année. Les principaux résultats rapportés sont les suivants :

- Les habiletés de lecture de mots isolés ou en contexte de phrases et de lecture de pseudo-mots sont corrélées entre l'anglais L1 et l'hébreu L2 (Geva & Siegel, 2000 ; Geva & Wade-Woolley, 1998 ; Geva *et al.*, 1997).
- Le pourcentage de mots et de pseudo-mots lus correctement est plus important en hébreu L2 qu'en anglais L1 (Geva & Wade-Woolley, 1998), mais cette différence en faveur de l'hébreu

¹¹⁴ Dans cette étude menée à Haïfa, en Israël, la langue première des enfants est l'arabe. Dans le cadre scolaire, ils apprennent l'hébreu, leur deuxième langue à partir de la 3^e année et l'anglais à partir de la 4^e année.

¹¹⁵ Dans les études d'Abu-Rabia et Siegler, la lecture en langue sémitique concerne le système vocalisé.

¹¹⁶ Les élèves sont tous immigrés, parlant l'arabe dans leur famille et vivant au Canada depuis plus de deux ans.

tend à diminuer en 5^e année. Le profil développemental des performances en lecture diffère donc entre les deux langues (Geva & Siegel, 2000)

- En anglais L1, les enfants de 1^e et 2^e année sont plus performants pour lire des mots en contexte, alors que cet effet facilitateur du contexte n'apparaît pas lors de la lecture en hébreu L2 (Geva *et al.*, 1998, 1997).

Ces deux derniers résultats peuvent s'expliquer en termes de transparence orthographique et de complexité syntaxique. Premièrement, même avec un bagage minimal en hébreu, leur langue seconde, les enfants n'ont qu'à acquérir un nombre limité de règles de correspondances graphophonologiques simples et consistantes, ce qui facilite, dès le début, l'acquisition du décodage. En revanche l'acquisition des règles de correspondances, nombreuses et complexes en anglais, est loin d'être achevée en milieu de 1^e année. Ainsi, si la maîtrise de l'écrit en L1 reste un facteur contribuant à la lecture en L2, il semble que l'acquisition d'une L2 transparente est plus rapide que celle d'une L2 opaque. Cet effet de transparence a d'ailleurs été suggéré par plusieurs auteurs ayant comparé l'évolution de la lecture en anglais langue maternelle, par rapport à d'autres langues (*cf.* partie 1.1, chapitre 2).

Deuxièmement, si le contexte facilite la lecture en anglais L1 mais pas en hébreu L2, cela peut relever du fait que la structure morphosyntaxique de l'hébreu est complexe et exige des processus cognitif et linguistiques coûteux, difficiles à mettre en place en début d'apprentissage d'une L2. A ce stade, les enfants ont sans doute tendance à effectuer un décodage linéaire, même s'ils possèdent certaines connaissances syntaxiques et sémantiques dans la langue. Une autre interprétation plausible serait que le niveau de compétences linguistique en L2, notamment syntaxique, ne soit pas encore suffisant pour utiliser de façon productive ces informations, ce qui renvoie à l'existence d'un niveau seuil de compétences. D'autres recherches sont ainsi nécessaires pour examiner plus précisément ces deux propositions.

Dans le cas de figure inverse, chez des élèves natifs hébreux, en 4^e année, débutant l'apprentissage de l'anglais, Kahn-Horwitz *et al.* (2005), montrent également que les compétences développées en langue première interviennent sur la lecture en langue seconde. Même si la connaissance des lettres et des sons anglais explique de façon importante la reconnaissance des mots et la compréhension de texte en anglais, les auteurs notent que :

- 1) les compétences en lecture en hébreu vocalisé apportent leur contribution à la lecture de mots,
- 2) la conscience phonologique et morphologique (mesurée en hébreu à l'aide de tâches orales de suppression de phonème et de complètement de phrases par un mot dérivé à partir d'une

racine donnée) apporte leur contribution à la compréhension. Cependant, ces deux dernières compétences n'ont pas été évaluées en anglais.

Schiff et Calif (2007), en réalisant une recherche auprès de la même population, ont permis de résoudre en partie cette lacune et d'approfondir, entre autre, le rôle des connaissances morphologiques développées en L1 lors de la lecture L2. Leurs résultats font tout d'abord apparaître que les connaissances morphologiques, en hébreu L1, prédisent ces mêmes connaissances en anglais L2 (11% de la variance expliquée, ces connaissances étant évaluées à l'aide d'une tâche de jugement de relation morphologique entre deux mots ; *e.g. farm – farmer* en anglais¹¹⁷). Par ailleurs, les compétences en lecture de mots en anglais L2 dépendent des compétences en lecture de mots vocalisés en hébreu L1 mais cette relation n'est effective que lorsque les connaissances morphologiques (et orthographico-phonologiques) en hébreu sont faibles. Au contraire, lorsque ces mêmes connaissances sont bien développées, la relation entre le niveau de lecture dans les deux langues disparaît. Cependant, comme l'indiquent les auteurs, les compétences morphologiques mesurées en hébreu portent sur des structures morphologiques linéaires comme cela est le cas en anglais. Ils en déduisent que cette similarité morphologique est à l'origine de la relation entre les deux langues. Pourtant, l'hébreu utilise principalement une morphologie non linéaire (*cf.* paragraphe 1.3 du chapitre 2), ce qui rend le matériel linguistique employé dans l'épreuve de Schiff et Calif (2007) peu représentatif de l'ensemble des constructions morphologiques de cette langue. La relation observée peut ainsi être accentuée par l'utilisation de formes particulièrement proches en hébreu et en anglais et nécessite, à notre sens, d'être plus largement confirmée pour avoir une valeur écologique satisfaisante. Indépendamment de cette remarque, leurs résultats sur la lecture permettent de constater que les compétences liées à l'écrit en L1 interviennent, mais n'assurent pas, systématiquement la réussite en L2 et doivent être couplées avec les connaissances spécifiques à celle-ci.

4.2.3 Les recherches sur les langues premières non alphabétiques

Enfin, quelques recherches ont également mis en évidence la possibilité d'un transfert en particulier phonologique entre des langues idéographiques versus alphabétiques (exclusivement l'anglais L2).

Bialystok, MacBride-Chang & Luk (2005b) étudient, au niveau préscolaire et en 1^e année, la conscience phonologique (suppression de syllabe et comptage phonémique) et la lecture de mots d'enfants bilingues chinois-anglais vivant au Canada et d'enfants de Hong Kong apprenant l'anglais L2 à l'école. Leurs résultats suggèrent un transfert de la conscience phonologique entre

¹¹⁷ Les épreuves de jugement étant présentées à l'écrit, la version en hébreu utilisent des mots vocalisés.

les deux langues pour les deux groupes (r compris entre .52 et .68). En revanche, les performances en lecture de mots dans les deux langues sont corrélées chez les apprenants chinois ($r = .57$ ou $.39$) mais pas chez les bilingues canadiens. Leurs compétences en anglais étant bien développées, ces derniers ont déjà pu établir deux approches différentes pour identifier les mots écrits dans des systèmes aussi distants. La relation apparue chez les apprenants chinois, peut être attribuable à l'utilisation de stratégies de mémorisation identique aux deux systèmes écrits parce qu'ils n'ont pas suffisamment développé, en anglais, de compétences analytiques leur permettant de lire de nouveaux mots. Ces interprétations, suggèrent que, tout au moins chez des apprenants dont les principes d'écriture entre L1 et L2 sont éloignés, la référence aux stratégies de la L1 concernerait surtout les stades élémentaires de l'apprentissage de la lecture en L2.

Gottardo, Yan, Siegel et Wade-Woolley (2001) se sont intéressés, pour leur part, aux compétences d'enfants chinois immigrés au Canada anglophone depuis moins de 2 ans et ayant débuté leur scolarité en Chine. Ces enfants sont intégrés dans des classes allant de la 1^e à la 8^e année. Il apparaît, non seulement que les performances de détection de rime en chinois sont corrélées avec plusieurs autres scores de conscience phonologique en anglais (détection de rime : $r = .52$ et suppression de phonème : $r = .54$) mais aussi qu'elles contribuent à la variance en lecture de mots et de pseudo-mots en anglais, même après avoir contrôlé l'effet de la conscience phonémique en anglais (6 % de variance expliquée). Ce « transfert phonologique », même entre des langues qui ne partagent pas les mêmes principes d'écriture, suggère que ces connaissances ne relèvent pas de processus spécifiques mais sont liées à l'habileté générale des enfants à réfléchir sur la phonologie de n'importe quelle langue, à partir d'un minimum d'exposition.

Le rôle joué par la détection de rime en chinois sur la lecture en anglais rend compte de l'importance des compétences phonologiques développées en L1 pour décoder une orthographe alphabétique, en dépit des différences de système. L'intervention massive des compétences phonologiques confirme aussi les résultats d'une très grosse majorité de recherche sur la lecture en général. En revanche, il faut noter que dans cette étude, les connaissances syntaxiques évaluées à l'aide de tâche de complètement de phrases, ne sont pas corrélées entre les deux langues et n'apparaissent pas contribuer à la lecture en anglais. Les auteurs estiment que cette mesure n'est pas assez fiable pour faire apparaître des coefficients de corrélation significatifs.

Wang, Park et Lee (2006) obtiennent des résultats exactement similaires dans le domaine phonologique auprès de bilingues coréen-anglais en 1^e, 2^e et 3^e année, corroborant ainsi l'universalité des processus phonologiques lors de l'apprentissage de la lecture. Cependant, comme cela avait également été démontré par Gottardo *et al.* (2001), le transfert des connaissances orthographiques (mesurées à l'aide du même paradigme que celui de Abu-Rabia &

Siegel, 2003) semble inexistant entre les deux systèmes. L'absence de corrélation entre ces habiletés orthographiques dans les deux langues et de contribution de celles mesurées en chinois à la lecture de mots en anglais (alors que celles mesurées en anglais interviennent) peut être le reflet des différences visuelles du code et de sa transparence, entre les deux langues. Le système orthographique de l'anglais est linéaire et peu transparent alors que celui du coréen est non linéaire mais transparent. Les configurations visuo-spatiales sont donc très contrastées, ce qui peut entraver la possibilité de se référer aux acquisitions orthographiques préalables. L'absence de transfert orthographique peut également provenir du fait que les expériences réalisées dans une L1 transparente n'aident pas à mettre en place les traitements orthographiques dans une langue plus opaque. Quoi qu'il en soit, les auteurs résument leurs résultats en avançant que l'acquisition de la lecture relève de processus phonologiques généraux et universaux et de processus orthographiques spécifiques.

4.2.4 Synthèse et perspectives

Globalement, il semble se dégager de ces études des procédures de reconnaissance de mots écrits communs aux langues en contact chez les apprenants, dès le début de l'apprentissage. Les résultats et réflexions de Bialystok *et al.* (2005b) permettent toutefois de nuancer cette proposition lorsque les langues en questions ne relèvent pas des mêmes principes orthographiques. Dans ce cas, les stratégies de lecture peuvent être différentes, surtout lorsque les systèmes sont bien maîtrisés.

L'ensemble de ces données conforte aussi l'idée de compétences métaphonologiques transférables d'une langue à l'autre. Il serait toutefois intéressant de s'interroger sur la terminologie la plus pertinente pour rendre compte de ce phénomène. L'aspect universel des capacités à identifier le principe de correspondance entre forme orale et écrite et à l'utiliser pour lire dans de nombreuses langues, nous paraît relativement différent de ce que sous entend le transfert d'une langue à l'autre. La caractéristique principale de ces compétences nous semble plutôt être leur transversalité et leur caractère actualisable.

De plus, si les chercheurs semblent d'accord pour envisager l'existence de processus syntaxiques transversaux, tout au moins entre les langues alphabétiques, nous pensons que les modes d'évaluation de ces compétences pourraient être améliorés et diversifiés. En effet, hormis l'étude de Durgunoğlu *et al.* (2002) la tâche systématiquement utilisée est celle de complétement de phrases dont nous avons vu les limites dans la partie 4 du chapitre 4. D'ailleurs, les travaux de Geva laissent entendre que la complexité structurale des langues a un impact important sur l'élaboration des traitements des langues secondes écrites. Il reste alors à vérifier si cet impact,

de même que celui de la proximité syntaxique entre L1 et L2, ne concernerait pas également la construction de connaissances syntaxiques explicites générales, indépendantes des caractéristiques spécifiques aux langues, dont l'évaluation aura permis de les isoler des aspects sémantiques.

Il est également surprenant de remarquer le peu d'intérêt accordé à la dimension morphologique dans ce domaine de recherche. On connaît pourtant l'importance de ces traitements lors de la lecture dès le début de l'apprentissage (Marec-Breton *et al.*, 2005). Seules les recherches récentes de Kahn-Horwitz *et al.* (2005) et Schiff et Calif (2007) ont étudié la relation des compétences morpho-dérivatives entre L1 et L2 et leur rôle sur la lecture. Néanmoins, les limites méthodologiques qui les caractérisent invitent à multiplier les expérimentations. Dans le domaine morpho-syntaxique l'unique recherche de Durgunoğlu *et al.* (2002) ne peut être suffisante pour généraliser l'hypothèse d'une interdépendance des compétences morpho-flexionnelles entre des langues différentes, d'autant plus que ces compétences sont ici traitées avec les compétences syntaxiques portant sur l'ordre des mots. Il en est d'ailleurs de même pour les résultats sur le lien unissant les langues en contact sur la dimension de compréhension en lecture, qui restent encore très limités (Van Gelderen *et al.*, 2005).

Nous tenons également à souligner que ces recherches ont toujours envisagé les relations interlangues sur ce qui a trait à une analyse consciente des aspects linguistiques lors de la lecture (tout au moins en théorie) et mettent peu en évidence la façon dont cette relation évolue au cours des apprentissages. L'idée que les connaissances applicables aux langues 1 et 2 doivent être de nature analytiques pour pouvoir être dégagées des aspects formels spécifiques aux langues est, certes, cohérente. Mais il n'en est pas moins intéressant d'examiner d'une part leur évolution, et d'autre part si le lien, entre certaines langues, ne peut absolument pas s'élaborer, à un moment ou à un autre, sur un niveau plus implicite.

Enfin, plusieurs de ces recherches (Abu-Rabia & Siegel, 2003; Gottardo *et al.*, 2001; Wang *et al.*, 2006) avancent que, contrairement aux processus phonologiques, les traitements orthographiques sont spécifiques aux langues, notamment lorsque ces dernières n'ont pas le même système alphabétique ou lorsque la L1 est idéographique. Il faut toutefois souligner, là encore, que les instruments de mesure évaluent surtout une sensibilité aux règles de combinaisons orthographiques qui ne peut être que propre aux langues. Quoiqu'il en soit, ces données, de même que celles de Bialystok *et al.* (2005b) soulèvent le fait que les caractéristiques orthographiques des langues ont un effet important sur les relations que peuvent entretenir, à l'écrit, les L1 et L2. Pourtant, ces caractéristiques, et par conséquent la proximité entre L1 et L2

n'est pas toujours au centre des préoccupations. La perspective « *cross-linguistique* » insufflée en particulier par Koda (1988, 1989, 2005, 2007) s'est développée pour approfondir cette question. L'influence que peut avoir les L1 lors des acquisitions en L2, exclusivement l'anglais, en fonction des caractéristiques des langues, est, dans cette optique, étudiée par le biais de l'analyse comparative interlangue première.

4.3 La perspective *cross-linguistique* anglo-saxonne : les caractéristiques de la langue première et sa proximité avec la langue seconde

Dans cette perspective *cross-linguistic*, l'étude des variations linguistiques entre L1 et L2 s'est surtout orientée vers l'impact de la distance orthographique entre des langues partageant ou non le même système alphabétique que l'anglais.

Ainsi, Muljani, Koda et Moates (1998), dans une étude comparant des apprenants adultes de L1 indonésien et chinois, résidant aux Etats-Unis depuis environ 1;3 ans, affirment qu'une L1 alphabétique, comme l'indonésien, facilite la décision lexicale de mots en anglais L2 par rapport à une L1 non alphabétique, comme le chinois, spécialement pour les mots qui contiennent des patterns orthographiques présents en L1 (*e.g.* CVC). Pour les auteurs, les traitements orthographiques nécessaires aux langues alphabétiques, que partagent donc l'indonésien et l'anglais, sont à l'origine de cette supériorité.

Wang, Koda et Perfetti (2003) examinent, pour leur part, les habiletés d'identification de mots et la conscience phonologique d'étudiants chinois et coréens, immigrés aux Etats-Unis depuis moins d'un an et dont l'apprentissage de l'anglais est quantitativement équivalent. Ils utilisent une tâche de jugement de catégorisation sémantique de mots plus ou moins similaires phonologiquement ou orthographiquement à l'exemplaire correct de la catégorie (*e.g.* *beech* ou *bench* pour l'exemplaire *beach*) ainsi qu'une tâche écrite de suppression de phonème, suppression aboutissant à la formation d'un autre mot mais dont l'orthographe diffère (*e.g.* supprimer [t] de *might* → *my*). Les résultats à la première épreuve montrent que les Coréens font plus d'erreurs sur les homophones qu'ils n'en font sur les pseudo-homographes. Sur ces derniers items, ils obtiennent de meilleurs résultats que leur homologues chinois. Les chinois sont ainsi plus sensibles aux aspects orthographiques. Ils produisent d'ailleurs plus d'erreurs aux items orthographiquement similaires au mot de la catégorie qu'aux items pour lesquels la similarité est moins marquée. Les résultats à l'épreuve de suppression de phonème supportent également l'hypothèse que les chinois se reposent plus sur les informations orthographiques, et moins sur les informations phonologiques des mots que les Coréens. A cette épreuve, les chinois sont moins performants que les Coréens car ils font plus d'erreurs phonologiquement incorrectes mais

orthographiquement acceptables (e.g. *me* à partir de *might*). Wang *et al.* (2003) postulent que les différences de système d'écriture des L1, de même que le transfert des stratégies de lecture en L1 peuvent être responsables de certaines différences de performances en L2.

Cependant, Yamada (2004) a émis certaines critiques sur ces conclusions. Il reproche en effet à Wang *et al.* (2003) de se fonder essentiellement sur la distance orthographique des langues pour expliquer des effets différentiels au niveau des habiletés phonologiques et orthographiques lors de l'identification de mots anglais. Il propose que le système phonologique de la L1, plus que le système orthographique, affecte les processus cognitifs lors de la lecture en L2. Il considère que la pauvreté des combinaisons phonologiques du chinois peut expliquer le faible usage de la phonologie chez ces apprenants pour lire en anglais L2 et par conséquent leur préférence à utiliser les informations orthographiques. Wang *et al.* (2004) répondent qu'il existe en effet deux facteurs linguistiques importants dans l'apprentissage de la lecture en L2 : la phonologie des langues et l'orthographe des écritures ; mais qu'il s'avère très difficile d'isoler les effets respectifs de l'un ou l'autre de ces deux facteurs car il n'existe pas de données métriques permettant de les distinguer et que cette confusion est naturelle et inhérente aux langues. Ils reconnaissent donc que les différences observées proviennent non seulement de l'éloignement orthographique entre les systèmes écrits mais aussi des contrastes phonologiques entre les langues. Il conviendrait toutefois de dissocier ces effets, soit comme le propose Yamada (2004) en examinant l'identification de mots écrits en L2 chez des apprenants chinois ne sachant pas lire dans leur L1 soit en manipulant les variations phonologiques et orthographiques lors de comparaisons sur de multiples L1.

Quoi qu'il en soit, Koda (2005) considère que la distance typologique entre L1 et L2 n'explique pas seulement les différences de niveaux de lecture entre des apprenants dont la L1 se trouve orthographiquement liée ou non à la L2, mais permet aussi de rendre compte de la façon dont les expériences en L1 peuvent affecter les processus lexicaux lors de l'apprentissage de l'écrit en L2.

D'autre part, Bialystok, Luk et Kwan (2005a) ont pu confirmer une telle influence chez les enfants bilingues de 1^e année dont l'anglais est la langue scolaire et la langue familiale est soit le chinois, soit l'hébreu, soit l'espagnol, langue faisant également l'objet d'un enseignement de la lecture. Les auteurs veulent examiner si le transfert interlangue des habiletés de décodage (lecture de non-mots ayant plus ou moins de voisins orthographiques dans la langue considérée) s'effectue de la même façon selon les caractéristiques de la langue 1 des participants. Leurs résultats montrent, tout d'abord, de moins bonnes performances en décodage en anglais et en segmentation phonémique chez les enfants chinois. Par ailleurs, la corrélation, observée entre les

performances en lecture de non-mots dans les deux langues, est non significative pour le groupe chinois, mais élevée pour les groupes hébreu et espagnol (respectivement $r = .57$ et $.72$ pour ces deux derniers). Ainsi, chez ces enfants, le transfert des stratégies de lecture s'effectuerait mais seulement si les L1 et L2 sont toutes deux basées sur le même principe alphabétique.

D'autres recherches, plutôt que de pointer les déficits des apprentis lecteurs dont la L1 s'écrit dans un autre système, ont tenté de mettre en évidence des profils de lecture variés, selon les caractéristiques des L1. Par exemple, Wade-Woolley (1999) constate que les apprenants de langues 1 différentes mettent en place des stratégies différentes, propres à leur L1, pour parvenir aux mêmes niveaux en lecture L2. Elle observe notamment que les adultes russes et japonais, étudiant récemment l'anglais, parvenaient aussi bien à lire les mots anglais, les premiers grâce à leur capacité à manipuler les phonèmes, les seconds grâce à leur habileté à reconnaître les patrons orthographiques.

Dans le même ordre d'idée, Fender (2003) a examiné l'influence des traitements mis en place pour lire en L1 sur la lecture en anglais langue seconde en comparant les habiletés d'apprenants adultes, résidant aux Etats-Unis depuis moins d'un an, sachant lire en arabe standard ou en japonais kanji. Il ressort de leurs résultats que les arabophones rencontrent des difficultés dans les traitements prélexicaux lors d'une tâche de décision lexicale (lecture de mots et non-mots isolés), alors que les Japonais rencontrent des difficultés dans l'intégration on-line des mots dans des phrases, lors d'une tâche de compréhension de phrases par lecture mot à mot, le dernier lu disparaissant avec l'apparition du suivant. Les auteurs en déduisent que les apprenants japonais sont davantage influencés, lors de la lecture de mots isolés, par la nature visuelle et orthographique des mots sans être gênés par les aspects phonologiques car le « kanji » relève d'une écriture basée sur une représentation idéographique des morphèmes. Les arabophones sont plus dépendants des caractéristiques phonologiques, ce qui entraîne des difficultés à gérer les inconsistances orthographiques des mots anglais isolés lors de leur reconnaissance. En revanche, ils ont pu développer leur capacité à assembler les informations syntaxiques et sémantiques lors de la construction de phrases en anglais. L'arabe se caractérise en effet, comme l'anglais, par une configuration structurale stable et complète des éléments de la phrase, utilisable pour repérer les correspondances entre les fonctions syntaxiques et sémantiques des arguments lors de l'analyse des mots de la phrase. Ces éléments montrent l'implication des acquisitions préalables, les orientations et les besoins particuliers des apprenants, lors de l'apprentissage de l'écrit en langue seconde.

Très peu de recherches ont en revanche analysé l'impact de la distance entre L1 et L2 dans le domaine des traitements morphologiques. A notre connaissance, seule la recherche de Koda (2000) a investi le champ des compétences morpho-dérivationnelles en L2 dans une perspective interlangue première. L'étude porte sur des adultes de langue 1 chinoise et coréenne résidant aux Etats-Unis depuis moins de 2 ans. Les résultats suggèrent que, pour un même niveau de lecture et de vocabulaire en anglais L2, les apprenants chinois sont moins performants à une tâche de jugement de séparation en plus petites unités significatives de mots composés que les apprenants coréens mais plus performants à une tâche d'intégration des informations morphologiques et contextuelles (détection de l'incongruence sémantique de phrases due à un mot morphologiquement composé). Ces différences sont explicables en terme de caractéristiques des L1 : le coréen écrit, le hangul, est une langue alphabétique non romane, sa morphologie fonctionne comme celle de l'anglais par processus de dérivation et ajout d'affixes. En chinois, la composition morphologique s'effectue selon un découpage totalement différent : les caractères idéographiques chinois se composent le plus souvent d'une partie constituant un radical sémantique et d'une autre partie codant « le radical phonologique ». Ce dernier vient compléter le radical sémantique qui n'apporte qu'une information morphologique partielle sur le caractère et est insuffisant pour identifier les caractères non familiers ou présentés hors contexte. Le lecteur doit ainsi intégrer les informations du radical sémantique aux indices phonologiques ou contextuels. Le processus dérivationnel est absent, de sorte que cette langue ne nécessite pas d'analyse morphologique interne des mots. L'influence des caractéristiques de la L1 lors des traitements morphologiques des mots de la L2 est ici parfaitement mise en évidence.

Ces différentes recherches démontrent ainsi que les caractéristiques linguistiques des systèmes d'écritures des L1 et leur degré de proximité avec la L2 ont un impact important sur les traitements élaborés lors de la lecture en langue seconde. Cependant, elles ne se sont pas intéressées aux deux dimensions de la lecture : la reconnaissance de mots écrits et la compréhension en lecture et n'ont pas analysé conjointement cet impact sur les différents aspects (phonologique, morphologique, syntaxique) qui sont associés à la lecture. Ainsi, l'unique recherche qui a pu étudier, de façon exclusive, les traitements morphologiques (Koda, 2000) est encore trop isolée.

De plus, il est important de remarquer que ces études portent le plus souvent sur des langues se différenciant quant à la nature du principe orthographique (logographique vs. alphabétique). Elles s'accordent ainsi sur le fait que les apprenants dont la L1 est idéographique ont tendance à se centrer sur les aspects orthographiques, ce qui n'est pas le cas des apprenants de L1 alphabétique. Il serait toutefois intéressant de développer les

comparaisons entre des langues alphabétiques ne partageant pas pour autant le même alphabet ni la même structure morphologique. Certes, les investigations de Bialystok *et al.* (2005a) ne font pas apparaître de différences notables entre la lecture et la conscience phonologique en anglais L2, chez des apprenants hébreux et espagnols, ni même en terme de relation plus ou moins fortes entre les habiletés en lecture en L1 et L2. Cependant, il nous semble qu'une analyse portant sur l'impact de ces différences sur de multiples dimensions (*i.e.* tant sur les connaissances phonologiques que morphologiques et syntaxiques) pourrait permettre de voir se dessiner des profils plus précis de lecteurs en fonction des caractéristiques de la langue première.

Par ailleurs, les données concernant l'apprentissage chez les enfants sont très rares (Bialystok *et al.*, 2005a) et méritent là encore d'être développées, en insistant sur le caractère évolutif de l'influence que peuvent avoir les caractéristiques typologiques des langues premières sur l'apprentissage d'une langue seconde.

Enfin, la totalité de ces recherches *cross-linguistiques* portant sur l'anglais langue seconde, il convient également d'élargir ce champ de recherche à d'autres langues secondes, en l'occurrence le français.

Ces différentes remarques invitent à multiplier les études interlangues s'intéressant à des L1 dont les caractéristiques sont spécifiques et originales par rapport aux différentes dimensions linguistiques pouvant être impliquées lors de la reconnaissance de mots mais aussi la compréhension de phrases (Koda, 2005, 2007). Une telle étude se devra également d'investir, vu le manque d'expérimentations dans ce contexte, l'évolution des compétences d'enfants apprenants le français langue seconde dans le cadre scolaire mais dans leur pays d'origine.

Chapitre 6 : Problématique de la présente recherche

1. Problématique des études menées sur les langues premières

1.1 Objectifs généraux

A l'issue des analyses menées dans différentes langues maternelles, il semble fondamental de tenir compte conjointement des multiples caractéristiques des systèmes écrits pour aborder les traitements impliqués lors de la lecture et leur évolution, dans différentes langues écrites. Toutefois, les études considérant simultanément plusieurs dimensions du langage, comme les connaissances phonologiques, morphologiques et syntaxiques, n'ont, à notre connaissance, jamais été envisagées dans une perspective comparative interlangues. Ces trois domaines constituent pourtant des éléments de connaissances déterminants pour la maîtrise de la lecture, et dont les caractéristiques, la complexité et l'importance varient d'une langue écrite à une autre¹¹⁸. En outre, le cadre théorique des apprentissages implicites et explicites (*e.g.* Gombert, 2002a, 2003ab), surtout établi sur la base de langues comme le français ou l'anglais, nécessite également d'être élargi à des langues dont les caractéristiques sont plus ou moins éloignées de celles déjà étudiées.

Notre intérêt sera donc d'approfondir les aspects spécifiques et universels des mécanismes de lecture, sur ces multiples dimensions et sur ces deux niveaux de contrôle cognitif (implicite *vs.* explicite) en étudiant deux langues très différentes, l'arabe et le portugais. Il est important de préciser dès à présent que l'arabe étudié ici est celui appris en Tunisie et le portugais celui du Brésil. L'arabe comme le portugais sont, en effet, les langues écrites de plusieurs pays ce qui implique certaines variations linguistiques mais aussi pédagogiques selon les cultures.

¹¹⁸ Plus exactement, il s'agira davantage, dans ces travaux, des connaissances grapho-linguistiques car les choix méthodologiques, pris pour mener à bien les passations à l'étranger, orientent les expérimentations dans ce sens. En effet, les évaluations des connaissances s'effectuant simultanément sur le mode oral et écrit (*cf.* infra), il sera alors préférable de dénommer ses connaissances : graphophonologiques, grapho-morpho-dérivationnelles, grapho-morpho-flexionnelles et grapho-syntaxiques. Cependant, bien que les traitements de nature graphique puissent intervenir, les mécanismes cognitifs sous-jacents à la résolution des épreuves employées, sont sur de nombreux aspects, communs à ceux abordés dans la littérature.

1.1.1 Objectifs généraux de l'étude 1 sur la langue arabe

Au niveau orthographique, la présence en arabe de deux formes d'écriture (avec ou sans voyelles) complexifie la façon de considérer le degré de transparence de cette langue par rapport au portugais. Cependant, en début d'apprentissage et jusqu'en 5^e année, les enfants tunisiens sont le plus souvent confrontés à l'écriture vocalisée. Ils apprennent donc à lire dans une langue plus transparente que leurs pairs brésiliens. Par ailleurs, les caractéristiques de l'arabe rendent particulièrement nécessaire l'analyse morphologique des mots et cela vraisemblablement plus précocement et plus fortement qu'en portugais, langue pour laquelle la dimension grapho-morphologique est moins privilégiée.

Ainsi, les données disponibles sur l'apprentissage de la lecture en langues sémitiques avancent des points communs avec celui dans d'autres langues alphabétiques mais également certaines spécificités des traitements graphophonologiques et grapho-morphologiques et en particulier l'importance de ces derniers. Cependant, beaucoup de points restent encore à approfondir, notamment en ce qui concerne la lecture en langue arabe et les connaissances qui lui sont associées, à certains moments de l'apprentissage.

- Par exemple, si le rôle de la conscience phonémique est bien établi, la comparaison entre les habiletés phonologiques segmentales de niveaux syllabique et subsyllabique, telles que la syllabe et la rime, n'a pas été clairement étudiée. D'autre part, le statut des voyelles longues mériterait d'être davantage examiné.

- Egalement, la distinction et l'évolution des connaissances morfo-dérivationnelles et morfo-flexionnelles, de même que la façon dont elles interviennent à l'écrit, tant sur la dimension de reconnaissance de mots écrits que sur celle de la compréhension de phrases, devra requérir une attention particulière. La morphologie tient indéniablement une place privilégiée lors de la lecture en arabe non vocalisé mais aussi vocalisé, du fait du statut différent accordé aux consonnes de la racine, aux voyelles du schème et aux flexions vocaliques finales qui s'y ajoutent. Pourtant, l'importance respective des aspects dérivationnels et flexionnels reste à définir plus précisément.

Concernant la structure morphologique des mots arabes écrits, la place accordée aux graphèmes affixés par rapport aux voyelles intriquées dans la racine pourrait également faire l'objet de plus d'attention. En effet, la composition morphologique du lexique arabe en racine/schème se caractérise souvent par l'intrication d'un schème non seulement vocalique mais auquel s'ajoute aussi des consonnes en position initiale et finale. Ces éléments graphiques, surajoutés à la racine

et également visibles graphiquement, s'apparentent ainsi, en quelque sorte, à une affixation. Leur appartenance au schème n'autorise toutefois pas à leur attribuer le statut d'affixe à proprement parler. En revanche, dans le domaine de la morphologie flexionnelle, on retrouve non seulement des suffixes mais aussi des préfixes dans la construction des mots fléchis.

- Enfin, il s'agira d'examiner l'évolution des connaissances syntaxiques par rapport aux autres dimensions et de leur contribution à la lecture (en reconnaissance et en compréhension), connaissances encore peu étudiées dans cette langue écrite. Nous tenterons dans ce domaine d'employer une méthodologie nouvelle, dans l'objectif de rendre comparable l'évaluation des différentes connaissances et d'aborder des connaissances syntaxiques d'une autre nature que celle de l'ordre des mots, c'est-à-dire davantage axées sur les fonctions syntaxiques de ces derniers. Cela se justifie aussi par le fait que cette étude s'intéresse non pas aux étapes initiales de l'apprentissage de la lecture en arabe mais à des niveaux plus avancés de l'acquisition.

1.1.2 Objectifs généraux de l'étude 2 sur la langue portugaise

La langue portugaise présente beaucoup de caractéristiques communes avec d'autres langues latines et se différencie donc des langues sémitiques tel que l'arabe. Ainsi, pour reprendre les propos de Seymour *et al.* (2003), nous retiendrons que cette langue est sous sa forme écrite, la plus proche du français ; en effet, bien que son système orthographique soit plus transparent, notamment du fait de la quasi-absence de lettres muettes, elle peut être classée parmi les langues à orthographe opaque.

L'objectif principal sera d'approfondir l'évolution connaissances graphophonologiques, grapho-morpho-dérivationnelles, grapho-morpho-flexionnelles et grapho-syntaxiques (évaluées en portugais par une méthode similaire à celle choisie en arabe), et leur rôle respectif dans l'installation de la maîtrise de l'écrit du portugais. Une telle analyse permettra de compléter les investigations de nombreuses recherches menées en portugais mais aussi dans d'autres langues à écriture latine, en tenant compte du degré de transparence du système orthographique qui les sépare.

- De la même façon que l'étude sur la langue arabe, il s'agira de rendre plus lisible le développement des connaissances graphophonologiques de nature implicite et explicite à différentes unités : la syllabe, la rime intrasyllabique et le phonème. La saillance de ces unités, en particulier celle de la rime intrasyllabique n'a en effet pas toujours été analysée conjointement en portugais et les recherches se situent en très grande majorité au niveau préscolaire ou en tout début d'apprentissage.

- Egalement, l'investigation des connaissances morphologiques en portugais est encore récente. Ainsi, dans la lignée des travaux de Mota (Mota, 2006 : Mota *et al.*, communication personnelle) et Vidigal de Paula (2007), et comme cela a largement été fait en français par exemple, nous tenterons de démontrer

1) que ces connaissances sont bien présentes en 3^e année de scolarité

2) qu'elles évoluent encore après quelques années d'enseignement de la lecture

3) que leur intervention pour lire les mots et comprendre les phrases écrites est loin d'être négligeable.

- Il en sera de même pour ce qui concerne les connaissances syntaxiques, pour lesquelles les données en portugais concernent là encore surtout l'ordre des mots et les premières phases de l'apprentissage.

1.1.3 Mise en correspondance des données issues des études 1 et 2

A partir des analyses menées dans ces deux langues, l'objectif final sera de dégager certaines spécificités de l'arabe par rapport au portugais, en tentant de discerner différents profils lecteurs en fonction de la langue mais aussi de l'avancée dans les apprentissages. La comparaison des mécanismes de lecture en arabe et en portugais n'a, d'ailleurs, jamais été effectuée. Nous nous contenterons, pour notre part, de mettre en parallèle les résultats car, compte tenu des différences très fortes, quelles que soient les dimensions linguistiques, entre l'arabe et le portugais, la comparaison expérimentale interlangue ne peut être envisagée. Néanmoins, dans ces deux études, les connaissances étudiées sont mesurées, le plus rigoureusement possible, par des paradigmes expérimentaux similaires, de façon à faciliter au maximum le rapprochement entre les résultats.

Notre objectif sera donc de décrire les différences apparentes, mais aussi certains points communs entre ces langues. Nous suggérons en effet que les dimensions étudiées n'interviennent pas de façon similaire dans les deux langues du fait de leur éloignement et de l'importance inégale de la morphologie entre l'arabe et le portugais. Plus précisément, il se pourrait que les connaissances grapho-morphologiques, bien qu'intervenant dans les deux langues, apportent une contribution plus importante à la lecture en arabe qu'en portugais.

Il convient de souligner dès à présent que cette description devra bien évidemment tenir compte systématiquement des variations, entre les deux études, en termes d'étendue des niveaux scolaires étudiés. Le facteur niveau d'apprentissage ne pouvant en effet pas avoir le même poids sur les performances en arabe et en portugais du fait d'une largeur d'échelle non comparable,

nous avons préféré traiter séparément, et dans les deux langues, chaque niveau, pour observer à chaque fois quelles peuvent être les compétences les plus déterminantes et en inférer certaines variations en fonction de la langue et du temps d'apprentissage.

1.2 Hypothèses

1.2.1 Hypothèses de l'étude 1 menée sur la langue arabe

Les objectifs spécifiques et les hypothèses de la recherche sur la lecture en langue arabe seront donc multiples et pourront se scinder en trois volets :

Le premier volet concerne l'évolution des connaissances entre la 3^e et la 5^e année d'apprentissage de l'arabe dans le domaine de la lecture et des traitements graphophonologiques, grapho-morphologiques et grapho-syntaxiques.

Premièrement, nous tenterons de confirmer, au cours des apprentissages, l'importance des connaissances morphologiques (dérivationnelles et flexionnelles) en les comparant simultanément aux connaissances graphophonologiques et grapho-syntaxiques tout en tenant compte du niveau implicite ou explicite qui les caractérise. Nous postulons en effet que les variations entre ces différents types de connaissances ne sont pas identiques au niveau implicite et explicite des connaissances.

Nous postulons que les connaissances implicites des aspects graphophonologiques sont moins développées que celles des aspects grapho-morphologiques, notamment car le découpage en racine/schème est particulièrement saillant en arabe. En revanche, sur le niveau explicite des connaissances, l'apprentissage des règles graphophonologiques aura sans doute permis que les connaissances graphophonologiques rattrapent voire même surpassent les autres connaissances (grapho-morphologiques et grapho-syntaxiques), plus complexes.

Concernant les connaissances morpo-dérivationnelles, morpo-flexionnelles et grapho-syntaxiques, nous souhaitons déterminer :

- D'une part si, sur le niveau implicite comme explicite, et de la même façon que cela a été démontré en hébreu par Levin *et al.* (2001), les connaissances des marques flexionnelles sont plus développées que les connaissances des aspects morpo-dérivationnels. Les résultats obtenus permettront en outre de réfléchir sur la notion de prédictibilité et de productivité des processus de formation des mots en arabe.

- D'autre part, comment les connaissances (implicites et explicites) grapho-morpho-flexionnelles évoluent par rapport aux connaissances grapho-syntaxiques. Les quelques données en arabe (Maaouia-Belkhadi, 2003) suggèrent que, contrairement à d'autres langues (par exemple, en français : Demont, 1994 ; Nocus, 1997), la manipulation délibérée de l'ordre des mots serait plus facile que celles des marques flexionnelles. En revanche, aucun résultat n'est actuellement disponible pour avancer que cette différence atteint également les connaissances de nature implicite. Pour ce niveau, les recherches de Demont (1994) et Nocus (1997) rendent, d'ailleurs compte d'une plus grande facilité à percevoir les dissonances liées à une perturbation de l'ordre des mots en français car cette perturbation affecte le sens de la phrase. Il s'agira ici d'investir ces phénomènes en s'intéressant à d'autres types de connaissances syntaxiques plus complexes, soit à des connaissances portant sur la construction du syntagme verbal.

Deuxièmement, nous nous focaliserons sur les connaissances graphophonologiques dans l'objectif de déterminer les unités phonologiques les plus pertinentes en arabe. Ce point est en effet sujet à discussions, les auteurs n'apportant pas les mêmes réponses à cette question (*cf.* par exemple les travaux d'Ammar, 1997 ou d'Ez-Zaher, 2004). Dans ce domaine, la distinction entre connaissances implicites et explicites est, là encore, décisive. Nous nous attendons, en effet, à ce que les unités les mieux identifiées ne soient pas les mêmes au niveau implicite et explicite. Si, lors d'une analyse consciente, le phonème consonantique est privilégié, sa taille est vraisemblablement trop petite pour que les apprentis lecteurs y soient sensibles implicitement. Au niveau implicite, une sensibilité aux unités plus larges telles que la syllabe pourrait être mieux développée. La rime ne semblant pas être pertinente en arabe (Ammar, 1997) comme dans d'autres langues transparentes, la sensibilité des arabophones à cette unité ne devrait pas apparaître.

Troisièmement, nous approfondirons l'analyse des compétences grapho-morphologiques lorsque celles-ci ne portent pas sur le lexique de la langue arabe. Il s'agira ici de déterminer si les spécificités et la diversité des constructions morphologiques en langue arabe ont un impact sur l'identification explicite de la construction morphologique en fonction des propriétés de cette construction. L'objectif de cette analyse sera d'évaluer les compétences à extraire les règles morphologiques sous-tendant la construction morphologique de pseudo-mots par préfixation, suffixation ou schématisation et selon que ces règles réfèrent à la morphologie dérivationnelle ou flexionnelle. Compte tenu des caractéristiques morphologiques de l'arabe (structure morpho-dérivationnelle en racine schème, présence dans de nombreux schèmes de graphèmes consonantiques placés avant ou après la racine, présence massive de préfixes et de suffixes dans la construction de mots par ajout de flexion en arabe), nous postulons que l'identification de la

construction de pseudo-mots par ajout de préfixe, suffixe ou schème sera globalement aussi bien réussie. En revanche, comme nous l'avons supposé précédemment, les règles morpho-flexionnelles devraient être mieux maîtrisées que les règles morpho-dérivationnelles, tout au moins en ce qui concerne l'identification des préfixes et suffixes. Cependant, il semble légitime de penser que plus les compétences morphologiques des enfants évoluent, moins elles seront dépendantes des conditions d'application propres à la langue arabe. Les différences de résolution, préalablement envisagées selon la nature des constructions, auront donc peut-être tendance à s'atténuer avec l'avancée dans les apprentissages.

Le deuxième volet se fixe comme objectif d'analyser les relations qu'entretiennent entre elles les différentes connaissances des aspects de la langue arabe ainsi que les compétences portant sur les principes grapho-morphologiques décontextualisés. Sachant que ces connaissances ou compétences sont susceptibles d'intervenir lors de l'apprentissage de la lecture, cette analyse permettra d'obtenir un aperçu préalable de la façon dont elles s'agencent, avant d'identifier plus spécifiquement comment elles contribuent à la lecture. Nous chercherons également à confirmer qu'il est cohérent d'englober, sous le terme de connaissances grapho-sémantiques, les connaissances concernant différentes notions telles que la morphologie dérivationnelle, flexionnelle et la syntaxe car elles engagent, d'une manière ou d'une autre, un traitement sémantique des éléments graphiques. Cependant, ces différentes connaissances n'entretiennent pas forcément les mêmes relations entre elles et n'ont pas forcément le même poids au cours de l'avancée dans les apprentissages. Il s'agira alors d'établir de quelle nature sont ces variations.

Le troisième volet de cette étude s'attachera, enfin, à déterminer si les connaissances graphophonologiques, grapho-morphologiques (dérivationnelles et flexionnelles) et grapho-syntaxiques interviennent lors de l'apprentissage de la lecture en arabe avec la même importance que dans d'autres langues. Il est avancé que les traitements grapho-morphologiques sont particulièrement nécessaires en arabe. Globalement, nous prévoyons d'observer une implication importante des connaissances grapho-morphologiques (en particulier morpho-dérivationnelle) lors de la reconnaissance de mots écrits et une implication importante des connaissances grapho-morphologiques (flexionnelle et dérivationnelle) et grapho-syntaxiques lors de la compréhension de phrases. Compte tenu des niveaux scolaires étudiés, nous postulons que les connaissances graphophonologiques nécessaires au décodage sont suffisamment bien installées pour ne plus intervenir.

1.2.2 Hypothèses de l'étude 2 menée sur la langue portugaise

L'étude portant sur la lecture en langue portugaise se fixe les mêmes objectifs que la précédente. Cependant, les caractéristiques du portugais étant très éloignées de celles de l'arabe, les hypothèses posées sont différentes.

Le premier volet s'intéressera à l'évolution des connaissances mais sur une étendue de niveaux scolaires plus large qu'en arabe, allant de la 3^e à la 7^e année d'apprentissage¹¹⁹.

1) Compte tenu des résultats obtenus en français et en portugais, voire en anglais, la comparaison des connaissances graphophonologiques, grapho-morpho-dérivationnelles, grapho-morpho-flexionnelles et grapho-syntaxiques, aussi bien implicite qu'explicite et l'étude de leur évolution tenteront de faire apparaître les points suivants :

- Au niveau implicite comme explicite, de meilleures connaissances graphophonologiques que grapho-morphologiques (*cf.* par exemple Carlisle, 1995).
- De meilleures connaissances implicites grapho-syntaxiques que grapho-morphologiques mais de meilleures connaissances explicites grapho-morpho-flexionnelles que grapho-syntaxiques (*cf.* Demont & Gombert, 1996 ; Nocus & Gombert, 1997).
- De meilleures connaissances grapho-morpho-flexionnelles que grapho-morpho-dérivationnelles (*cf.* par exemple Carlisle & Nomanbhoy, 1993). Si l'on tient compte des résultats de Vidigal de Paula (2007), cette différence pourrait ne concerner que le niveau implicite.

2) Concernant les unités phonologiques les mieux maîtrisées en portugais, nous posons l'hypothèse que seules la syllabe et la rime peuvent faire l'objet de connaissances implicites alors qu'au niveau explicite, la conscience phonémique peut être aussi développée que celle d'unités plus larges (*cf.* Cardoso-Martins, 1995 ; Demont & Gombert, 2007 pour un aperçu global sur la question).

3) Concernant l'analyse des compétences grapho-morphologiques dégagées des aspects de surface du portugais en fonction de la nature de la construction morphologique et de la règle étudiée, nous envisageons d'une part, que l'influence des caractéristiques morphologiques du portugais fasse apparaître un « déficit » à identifier les constructions basées sur la structure « schème vocalique » et d'autre part que les règles morpho-flexionnelles soient mieux maîtrisées que les règles morpho-dérivationnelles.

¹¹⁹ Cette différence entre les niveaux étudiés chez les arabophones et les lusophones est liée aux contraintes de l'apprentissage du français langue seconde en Tunisie et au Brésil, qui est l'objet de recherche de notre troisième étude.

Dans un deuxième volet, comme lors de l'étude en langue arabe, l'examen des connaissances associées à l'apprentissage de la lecture abordera les relations que ces connaissances entretiennent entre elles. Les analyses factorielles permettront de vérifier en portugais que la dimension grapho-sémantique se compose des aspects grapho-morpho-dérivationnels, grapho-morpho-flexionnels et grapho-syntaxiques malgré certaines variations dans les relations entre ces sous dimensions et leur poids respectif en fonction de l'avancée dans les apprentissages.

Le troisième et dernier volet de cette étude s'attachera à déterminer dans quelle mesure les connaissances graphophonologiques, grapho-morpho-dérivationnelles, grapho-morpho-flexionnelles et grapho-syntaxiques, interviennent lors de la lecture en portugais. Il s'agira de démontrer que les traitements grapho-morphologiques sont impliqués, notamment chez des apprenants ne se situant plus en tout début d'apprentissage de la lecture (par exemple en 3^e année). Pour les niveaux les moins avancés de l'étude, nous nous attendons éventuellement à ce que les connaissances graphophonologiques soient encore impliquées lors de la reconnaissance des mots écrits en portugais (*cf.* Singson *et al.*, 2000, qui observent en anglais une disparition de la contribution des connaissances phonologiques à partir de la 4^e année), ce qui ne signifie pas pour autant que la contribution des connaissances grapho-morphologiques, ne soit pas déjà requise (*cf.* en portugais : Mota, 2006 ; Mota *et al.*, communication personnelle ; Sousa, 2005 ; Vidigal de Paula, 2007). Egalement, les connaissances grapho-syntaxiques pourraient intervenir sur cette dimension (Régo, 1993). Nous tenterons également d'observer l'implication des dimensions grapho-syntaxiques puis grapho-morphologiques lors de la compréhension de phrases dès la 3^e année (*cf.* Régo, 1997, en ce qui concerne la syntaxe en portugais et Demont & Gombert, 1996 pour ce qui est de l'intervention précoce des connaissances syntaxiques puis morpho-flexionnelles en français).

2.Problématique de l'étude menée sur le français langue seconde

2.1 Synthèse

Les éléments présentés dans le chapitre 5 permettent clairement d'affirmer l'influence des apprentissages préalables de l'écrit en langue première lors de l'apprentissage de la lecture en langue seconde. Ils démontrent en même temps la complexité des phénomènes, intégrant la présence de traitements linguistiques universaux mais aussi une prise en considération des propriétés spécifiques aux langues. Cette influence complexe et multidéterminée pourrait se traduire de la façon suivante :

A l'issue des propositions de Bialystok (2002b, 2007) sur la lecture des enfants bilingues, il peut être avancé que les principes généraux sous-jacents au mode écrit des langues, appréhendés lors des premiers contacts avec l'écrit, interviennent en langue seconde en favorisant d'une part la compréhension du caractère relativement stable des relations entre une forme écrite et son référent symbolique et d'autre part, la perception des variations de ces correspondances entre les langues. L'acquisition de ces concepts liés à l'écrit permettrait notamment aux apprenants d'être particulièrement attentifs aux caractéristiques formelles des langues ou encore de mettre en place leur fonction exécutive de manière efficace.

Certains auteurs ont aussi mis en avant que lors de l'apprentissage d'une langue seconde, d'autres compétences métalinguistiques, déjà développées en langue première peuvent, dans certains cas, être actualisées et appliquées lors des traitements de l'écrit en langue seconde (*e.g.* Cummins, 1991 ; Durgunoğlu, 1998 etc.). Bien que les résultats des recherches comparant les compétences métalinguistiques de bilingues à celle des monolingues ne soient pas toujours identiques, les analyses concernant les relations entre les deux langues connues des apprenants semblent, pour leur part, attester le caractère transversal de certaines de ces connaissances.

Ceci est particulièrement visible dans le domaine de la conscience phonologique (Bialystok *et al.*, 2005b ; Comeau *et al.*, 1999 ; Durgunoğlu, 1998 ; Gottardo *et al.*, 2001). Par conséquent, et dans le cas des langues alphabétiques qui nous intéressent ici, le principe de correspondance graphème-phonème, normalement saisi lors de l'apprentissage de la lecture en langue première, pourrait servir de base à l'élaboration des capacités de segmentation et donc des habiletés de décodage dans une autre langue, même lorsque cette dernière utilise un alphabet fort différent.

Les rares études (*cf.* Durgunoğlu *et al.*, 2002 ; Schiff et Calif, 2007) sur les relations entre les connaissances morphologiques en langue première et en langue seconde apportent quelques

éléments en faveur de compétences morphologiques générales et transversales, même dans des langues aussi éloignées que l'hébreu L1 et l'anglais L2 (Schiff et Calif, 2007). Toutefois, les conclusions qui en émergent restent très limitées. Soit elles concernent des langues relativement peu éloignées en termes de structures morphologiques (morphologie de type linéaire), comme l'espagnol et l'anglais (Durgunoğlu *et al.*, 2002), soit leurs résultats nécessitent d'être vérifiés à l'aide d'un matériel expérimental plus adéquat par rapport au système des langues en contact (Schiff & Calif, 2007). Par ailleurs, concernant le rôle, peu étudié, des connaissances morphologiques en L1 sur la lecture en L2, une des interprétations données par Schiff et Calif (2007) nous semble intéressante à souligner. Celle-ci suggère que de bonnes connaissances morpho-dérivationnelles en L1 ne conduisent pas nécessairement à une meilleure réussite en lecture en L2 (ce qui avait pourtant été envisagé par Kahn-Horwitz *et al.*, 2005, tout au moins sur la dimension de compréhension de texte). Les connaissances requises en langue première ne constituent, en effet, pas à elles seules, une aide pour apprendre à lire en langue seconde, mais doivent être couplées avec des connaissances spécifiques à celle-ci. Cette proposition doit être complétée par les remarques de Koda (2000) relatives aux conséquences des processus morphologiques propres aux L1 sur la nature des analyses morphologiques en anglais L2. Il ne peut ainsi qu'être souligné la complexité des facteurs intervenant et l'intérêt d'approfondir les processus en jeu lors de la lecture, quand les langues en contact ont des caractéristiques morphologiques différentes (par exemple non-linéaire *vs.* linéaire comme l'arabe et le français). Dans le domaine de la syntaxe, les investigations n'aboutissent pas non plus à l'heure actuelle à un point de vue clair et facilement généralisable. Bien qu'il semble que l'on retrouve ici un lien d'interdépendance entre des langues comme l'espagnol et l'anglais (Durgunoğlu *et al.*, 2002), le portugais et l'anglais (Da Fontoura & Siegel, 1995) et l'arabe, l'hébreu et l'anglais (Abou-Rabia & Siegel, 2002, 2003), ces relations ont le plus souvent été établies à l'aide d'épreuves pour lesquelles la prise en compte des aspects syntaxiques est confondue avec celles des aspects sémantiques. Le rôle que peuvent jouer ou non les compétences syntaxiques en L1 lors de l'apprentissage de la lecture en L2 n'a, à notre connaissance, pas été envisagé, même dans des langues proches. Les travaux de Geva et Wade-Woolley (1998 ; Geva *et al.*, 1997) sur l'anglais L1 et l'hébreu L2 suggèrent d'ailleurs que la complexité et les spécificités des systèmes des langues (grapho-sémantiques mais aussi graphophonologiques) expliquent que certaines connaissances mettent du temps à s'installer en L2, ce qui n'est pas sans conséquence sur les mécanismes de lecture. Les lacunes au niveau des connaissances lexicales peuvent aussi être considérées pour comprendre la lenteur de l'installation des connaissances associées à la

dimension grapho-sémantique (dont les connaissances morphologiques et syntaxiques ; *cf.* Armand, 2000 ; Roy, 2005).

Ainsi, avant d'envisager la conclusion d'une relation systématique entre langue première et langue seconde quels que soient les domaines de connaissances, il est évident que d'autres recherches, plus exhaustives et mettant en relation d'autres langues, doivent être réalisées. La nature implicite et explicite des connaissances devra également être évaluée, cette distinction n'étant, la plupart du temps, pas considérée dans le champ des recherches interlangues. Enfin, la façon dont ces relations se mettent en place, perdurent ou diminuent pourra être déterminée au cours des apprentissages.

Par ailleurs, s'il n'est pas rare que les mêmes compétences (par exemple, les compétences phonémiques) peuvent intervenir lors de la lecture en langue première et en langue seconde, et qu'ainsi certaines de ces compétences, de même que les habiletés en lecture de mots voire en compréhension ne sont pas indépendantes dans les deux langues, plusieurs nuances doivent être apportées à ce postulat :

- il a été mis en évidence, à plusieurs reprises et parfois même simultanément aux effets des acquisitions préalables, que les apprentissages en L2, permettant par exemple l'acquisition du vocabulaire ou celle des règles de correspondances graphophonologiques, sont indispensables à l'actualisation des connaissances, actualisation consistant notamment en une prise en considération des propriétés spécifiques de la langue seconde.

Les connaissances lexicales nous semblent particulièrement importantes à considérer si l'on s'intéresse à l'acquisition du code grapho-sémantique des langues écrites. On connaît en effet, la relation que l'étendue du vocabulaire entretient avec les connaissances morphologiques mais aussi syntaxiques (*e.g.* Gleitman & Gleitman, 1992 ; McBride-Chang, Wagner, Muse, Chow & Shu, 2005). Son développement constitue également un élément sensible lors de l'apprentissage d'une langue seconde (Bialystok, 2002b). Aussi, dans le cadre d'une étude s'intéressant à des langues premières comme l'arabe et le portugais, il semble difficile de négliger le fait que le lexique du portugais et celui du français se ressemblent en de nombreux points.

Le rôle des apprentissages menés en L2 n'est en fait pas contradictoire avec les considérations précédentes, relatives à l'influence de la lecture en L1 sur l'apprentissage de la lecture en L2. Cependant, selon la complexité des domaines considérés dans les deux langues, et les similarités qu'ils font apparaître entre ces dernières, il semble envisageable que l'actualisation

des connaissances soit plus ou moins facile et que la nécessité des apprentissages réalisés en L2 soit plus ou moins amplifiée.

- Les recherches développées dans le cadre de la perspective *cross-linguistique* anglosaxonne, bien qu'elles se soient surtout orientées vers la comparaison de langues premières alphabétiques versus non alphabétiques, apportent plusieurs éléments de réflexion opportuns par rapport à notre recherche. Elles insistent justement sur l'importance à accorder aux caractéristiques des systèmes écrits des langues en contact et proposent une méthodologie adéquate pour l'étude des procédures d'apprentissage tenant compte des effets de proximité entre les langues premières et secondes. Plus précisément, il en ressort que les caractéristiques orthographiques des systèmes (logographique vs. alphabétique) affectent beaucoup les procédures de lecture. Cette affirmation vient là encore compléter les propositions de certains auteurs sur l'existence de processus orthographiques spécifiques (Abu-Rabia & Siegel, 2003 ; Gottardo *et al.*, 2001 ; Wang *et al.*, 2006), sans forcément les contredire. Bien qu'il semble qu'aucune étude n'ait montré que le simple changement d'alphabet puisse affecter fondamentalement la façon dont les enfants apprennent à lire en L2 (*cf.* Bialystok *et al.*, 2005a), il peut être intéressant de vérifier, avec l'arabe et le français, si cette confrontation à un nouvel alphabet n'a pas plus d'impact que celui d'un simple ralentissement dans l'élaboration du décodage. En effet, pour les apprenants sachant lire en langue arabe et apprenant une langue seconde à écriture latine l'ensemble du système de correspondance, même l'identité des lettres et le sens d'écriture est à réapprendre. Un tel réapprentissage peut, de notre point de vue, affecter la façon dont certains traitements interviennent au cours de l'acquisition de la lecture, par rapport aux profils d'apprenants dont la langue première écrite n'a pas de telles caractéristiques.

Pour finir, et sans oublier les précisions apportées ci-dessus, l'apparition d'une certaine symétrie entre les habiletés en lecture et de certaines similarités quant aux connaissances indispensables à ces acquisitions en L1 et L2 nous semble conciliable avec l'idée que les différentes composantes de la lecture n'interviennent pas de la même façon, au même moment et avec le même poids, selon les caractéristiques des langues premières écrites. Si l'on considère que ces dernières se distinguent quant à l'importance de ces composantes, le poids attribué à ces composantes lors de l'apprentissage d'une langue seconde pourrait alors varier. C'est ce que nous tenterons de démontrer en effectuant une comparaison interlangue première (arabe vs. portugais) des connaissances apparaissant contribuer le plus à la lecture de mots et la compréhension de phrases écrites en français langue seconde. Cette comparaison concernera donc des apprenants ayant pu

entamer la lecture par le biais d'une langue première qui varie en terme de poids accordé à la morphologie (contrairement au portugais, l'arabe est une langue qui privilégie la dimension morphologique des mots) mais aussi, pour les raisons évoquées précédemment, en termes de proximité avec le français (contrairement à l'arabe, le portugais écrit fait apparaître une certaine proximité de fonctionnement linguistique, à plusieurs niveaux, avec le français, *cf.* Seymour *et al.*, 2003). Nous prendrons soin en revanche à ce que le niveau de connaissances lexicales puisse être considéré comme similaire.

Un travail préparatoire s'est inscrit dans cette démarche (Besse, 2003 ; Besse, Demont & Gombert, 2007). Il a consisté à comparer les compétences morphologiques d'enfants ou adolescents arabophones, résidant en France et apprenant récemment à lire en français, à celles de lusophones. Les résultats révèlent que les connaissances explicites de la structure morpho-dérivationnelle des mots français (évaluées par une épreuve de création de néologisme par ajout de morphème à une base existante, *e.g.* : produire *bouteillette* pour signifier « une petite bouteille ») sont supérieures chez les arabophones en raison du poids de la morphologie dans leur langue maternelle. Cependant, cette supériorité n'est pas perceptible au niveau implicite (niveau évalué par une épreuve de préférence entre deux mots dont l'un est de la même famille qu'un mot cible et l'autre non, jouant ainsi le rôle d'intrus, soit sémantique, soit phonologique *e.g.* qu'est ce qui va le mieux avec peur : beurre ou peureux ?). Les arabophones ne seraient pas plus sensibles que les lusophones à la structure morphologique des mots français vraisemblablement car la langue arabe et la langue française sont trop distantes pour que les automatismes élaborés par le biais de la langue première puissent être réinvestis en L2 (*i.e.* le repérage non intentionnel de la construction interne des mots). De plus, chez les arabophones et contrairement à ce qui se passe chez les lusophones, les connaissances morphologiques explicites, pourtant bien développées, n'apparaissent pas contribuer à la lecture de mots en début d'apprentissage du français écrit. L'hypothèse émise à cet égard est la suivante : les connaissances sur la construction morpho-dérivationnelle des mots français, ne peuvent être immédiatement mobilisées (mais pourront l'être ultérieurement) car les différences de systèmes alphabétiques entravent le réinvestissement des acquisitions préalables, tant que les spécificités graphophonologiques du code ne sont pas suffisamment bien maîtrisées.

2.2 Objectifs et hypothèses de l'étude 3 menée sur la langue française (L2)

L'étude 3 a ainsi pour principal objectif d'approfondir cette thèse, en comparant les performances d'apprenants arabophones et lusophones en français ou leurs « profils de lecture » à différents niveaux de l'apprentissage en L2. En tenant compte des résultats obtenus en langue première arabe ou portugais lors des deux études précédentes, elle s'articulera en quatre parties distinctes.

La première partie concernera la comparaison des performances obtenues et suivra l'organisation suivante :

- Etudier les effets des caractéristiques linguistiques des langues premières sur les habiletés de lecture de mots et de compréhension de phrases écrites en français langue seconde selon l'avancée dans les apprentissages. Bien que les groupes d'élèves ne se différencient pas en terme d'efficacité verbale, nous postulons que les arabophones partent avec un désavantage initial, en apprenant l'écrit dans une langue seconde dont l'alphabet est différent de celui de leur langue première. Ce désavantage pourrait toutefois se résorber par la suite.

- Etudier les effets des caractéristiques linguistiques des langues premières sur les connaissances implicites et explicites, graphophonologiques, grapho-morphologiques et grapho-syntaxiques, et leur évolution, en français langue seconde. Outre la prédiction d'une progression des connaissances, nous pensons que la référence aux acquis en langue première ne s'élabore pas de façon similaire au niveau implicite et explicite des connaissances car la distance entre L1 et L2 n'y aura pas la même influence. Les connaissances implicites, liées à la fréquentation répétée des mêmes configurations ne pourraient être influencées par les contacts avec la langue première que dans le cas où celle-ci partage avec la langue seconde des configurations similaires. La continuité en terme de repérage non intentionnel ne pourrait alors s'établir que lorsque les systèmes linguistiques sont suffisamment proches. En revanche, les connaissances explicites seraient moins tributaires de la distance entre langue 1 et langue 2 car elles relèvent d'un niveau d'abstraction plus élevé permettant au système cognitif de se dégager des aspects formels des langues. Nous tenterons, ainsi, de confirmer ce qui s'était dessiné dans l'expérience préalable, soit de meilleures connaissances grapho-morphologiques explicites chez les arabophones par rapport aux lusophones mais pas de meilleures connaissances implicites. L'origine de cette supériorité au niveau des connaissances explicites grapho-morphologiques (dérivationale et flexionnelle) tient également au fait que la langue écrite arabe favorise davantage la dimension morphologique que la langue portugaise. Cet avantage ne devrait pas concerner les autres dimensions, car les connaissances phonologiques impliquées lors du décodage, et les

connaissances syntaxiques nécessaires pour agencer les mots et interpréter le sens des phrases, sont présentes en arabe comme en portugais.

- Approfondir les effets des caractéristiques linguistiques des langues premières dans le domaine de la phonologie et plus particulièrement sur les unités les mieux maîtrisées. Partant du postulat que la rime n'intervient pas de la même façon en arabe et portugais, nous pensons voir apparaître une différence très ciblée entre arabophones et lusophones en faveur des lusophones sur le repérage implicite de cette unité en français langue seconde.

- Approfondir les différences entre arabophones et lusophones quant au développement de leur appréhension spontanée de la construction interne des mots français. Plus exactement, nous souhaiterions déterminer à quelle unité les apprenants arabophones et lusophones de chaque niveau, sont respectivement sensibles, sensibilité évaluable par le biais de leur préférence entre des pseudo-mots non construits, construits à l'aide d'une base (un mot français), d'un affixe de la langue française ou de la conjonction des deux.

La deuxième partie aura pour objectif d'aborder les relations qu'entretiennent les différentes connaissances en FL2 elles-mêmes. De la même façon qu'en langues premières, nous souhaitons savoir comment les connaissances lexicales, graphophonologiques, grapho-morphologiques, grapho-syntaxiques, en français s'articulent les unes aux autres car il ne paraît pas possible de les considérer comme indépendantes. Nous examinerons donc, au cours des apprentissages et séparément pour nos deux populations d'apprenants, comment se construisent en un ensemble cohérent ces connaissances. Nous nous intéresserons plus particulièrement aux éléments participant à la maîtrise du code grapho-sémantique. En effet, l'acquisition de ce dernier peut être différemment associée aux connaissances grapho-morphologiques, grapho-syntaxiques et lexicales, selon l'exposition à la langue et bien sûr selon les caractéristiques de la langue première.

La troisième partie s'intéressera aux liens entre L1 et L2. Elle concernera l'analyse des relations entre les connaissances développées dans les deux langues (en arabe L1 et en français L2 ou en portugais L1 et en français L2) dans un même domaine (graphophonologique, grapho-morpho-dérivationnel, grapho-morpho-flexionnel et grapho-syntaxique) mais aussi selon un même niveau de contrôle (implicite vs. explicite). L'analyse consistera donc à établir successivement ce lien pour chaque type de connaissances, à chaque niveau d'apprentissage, chez les arabophones d'une part et les lusophones d'autre part.

Nous formulons l'hypothèse qu'il n'existe une relation entre les connaissances implicites en L1 et L2, que lorsque les langues entretiennent un rapport de proximité sur la dimension concernée, c'est-à-dire en particulier dans le domaine des connaissances graphophonologiques et graphomorphologiques entre le portugais et le français. En revanche, nous nous attendons à ce que les liens entre les connaissances de nature explicite en L1 et en L2 apparaissent chez les arabophones et les lusophones, en particulier dans les domaines de connaissances les plus nécessaires pour lire en L1 (*e.g.* la morphologie en arabe).

Il s'agira par ailleurs d'évaluer si la continuité des connaissances implicites, susceptible d'apparaître entre le portugais et le français, peut s'installer rapidement et perdure ; De même, nous tenterons de déterminer si l'actualisation des connaissances explicites est efficace approximativement au même moment au cours des apprentissages, chez les arabophones et les lusophones et reste stable pour ces deux groupes.

La quatrième et dernière partie s'attachera à analyser la contribution de ces différentes connaissances mesurées en L1 et en L2 à la reconnaissance des mots écrits et à la compréhension en lecture. Ainsi, en intégrant aux analyses de multiples dimensions, différents profils d'apprentis lecteurs pourront être dégagés en fonction de la langue première et de l'avancée dans les apprentissages.

Nous envisageons qu'il existe une différence entre arabophones et lusophones dans la façon dont les connaissances grapho-morphologiques, implicites et explicites, contribuent à la lecture (en reconnaissance de mots et en compréhension de phrases). Il s'agira tout d'abord d'approfondir l'hypothèse émise à l'issue de notre étude préalable (Besse, 2003 ; Besse *et al.*, 2007). Chez les arabophones, les connaissances grapho-morphologiques explicites (en attachant une attention particulière à la morphologie dérivationnelle développée en L1) pourraient contribuer moins immédiatement à la lecture de mots écrits en français que chez les lusophones. Les connaissances phonologiques appliquées aux mots français, indispensables pour acquérir le décodage en écriture latine seraient alors, dans un premier temps déterminantes, notamment car la reconnaissance des lettres et toutes les correspondances graphophonologiques sont à réapprendre. Une fois qu'ils auront dépassé cette période d'installation des traitements séquentiels, leurs capacités d'analyse morphologique bien développées en arabe, pourraient être particulièrement réinvesties pour lire en français langue seconde.

Au contraire, il s'avère possible que les lusophones abordent la lecture en français quasiment dans la continuité de leur apprentissage de la lecture en portugais. Dans ce cas, pour reconnaître les mots en portugais L1 comme en français L2, ils auront sans doute dépassé les traitements

séquentiels initiaux. Dès lors, dégagés des exigences relatives à la médiation phonologique, ils pourraient se consacrer aux aspects grapho-morphologiques, domaine de connaissances qu'il auront déjà appréhendé, implicitement et explicitement, par le biais de la lecture en portugais. Toutefois, parce que la dimension morphologique est moins accentuée en portugais qu'en arabe, nous supposons que l'intervention de ces connaissances est plus précoce mais moins importante que chez les arabophones.

De telles suppositions peuvent être réitérées, à plusieurs nuances près, dans le domaine de la compréhension en lecture. Il s'agira ici de vérifier le postulat d'une intervention plus tardive mais plus massive des connaissances grapho-morphologiques explicites chez les arabophones par rapport aux lusophones. Pour les arabophones, il semble logique de penser que la compréhension de phrases écrites puisse être surtout dépendante des habiletés de reconnaissance des mots et des compétences orales en français, en début d'apprentissage, les connaissances des règles syntaxiques du français étant, peut-être encore à ce niveau, en cours d'élaboration.

Cette recherche s'intéressera enfin à déterminer comment évolue le rôle de ces diverses acquisitions en langue première sur la lecture en français langue seconde, chez ces deux populations d'apprenants.

Chapitre 7 : Étude 1 - Connaissances graphophonologiques, grapho-morpho-dérivationnelles, grapho-morpho-flexionnelles, grapho-syntaxiques et lecture en arabe standard langue première

1. Méthode

1.1 Population

Cent soixante quatorze élèves tunisiens scolarisés en 3^e, 4^e et 5^e année de l'enseignement de base (correspondant aux classes françaises CE2, CM1 et CM2) dans deux établissements primaires publics, les écoles Habib Thameur et Farad Hached, du gouvernorat de l'Ariana (agglomération limitrophe de Tunis) ont participé à cette expérience d'avril à mai 2006. Ont été exclus des analyses, les élèves redoublants (concerne 1 élève de 3^e année, 6 élèves de 4^e année et 2 élèves de 5^e année), ceux ayant sauté une classe (concerne 2 élèves de 5^e année), ceux ayant participé à moins de la moitié des épreuves (concerne 4 élèves de 4^e année) ou n'ayant visiblement pas suivi les consignes de plusieurs épreuves (concerne 1 élève de 3^e année, 1 élève de 4^e année et 3 élèves de 5^e année). Les résultats d'un élève de 3^e année ont également été supprimés car ces performances étaient supérieures aux performances maximales obtenues par le groupe des 5^e année. L'effectif total une fois ces critères d'exclusion appliqués, s'élève ainsi à 153 participants. Les caractéristiques de cette population par niveau scolaire sont précisées dans le tableau 4.

Tableau 4 : Répartition des élèves en fonction de leur niveau scolaire

	Effectif	Nombre de garçons	Nombre de filles	Age chronologique moyen	Ecart-type (en mois)	Age du plus jeune	Age du plus âgé
3 ^e année ¹²⁰	59	22	37	8 ; 9 ans (105,76 mois)	3,18	8 ; 3 ans (99,77 mois)	9 ; 4 ans (112,43 mois)
4 ^e année ¹²¹	44	22	22	9 ; 8 ans (117,13 mois)	3,02	9 ; 2 ans (111,13 mois)	10 ; 1 ans (122,13 mois)
5 ^e année ¹²²	50	30	20	10 ; 9 ans (130,16 mois)	2,99	10 ; 3 ans (124,13 mois)	11 ; 4 ans (137,07 mois)

¹²⁰ Elèves provenant de deux classes, une de chaque école.

¹²¹ Elèves provenant de deux classes de l'école Habib Thameur.

¹²² Elèves provenant de deux classes, une de chaque école.

1.2 Matériel

Les élèves des trois niveaux scolaires ont été évalués à l'aide des mêmes instruments :

- Le premier type d'évaluation concerne les performances en lecture. Deux épreuves de lecture sont utilisées : une épreuve chronométrée de lecture de mots isolés à voix haute et une épreuve de compréhension de phrases écrites.
- Le deuxième type d'évaluation est composé de huit épreuves et concerne les connaissances graphophonologiques, grapho-morphologiques (dérivationnelle et flexionnelle) et grapho-syntaxiques mises en place implicitement et explicitement en arabe standard : le paradigme de détection d'intrus est utilisé pour l'ensemble de ces huit épreuves. Ces épreuves consistent à choisir parmi deux stimuli, celui se distinguant d'un stimulus-cible. Elles ont toutes été présentées oralement avec un support écrit vocalisé. Les niveaux implicite et explicite se distinguent par le type de consigne, consignes données en arabe dialectal.
- Le troisième type d'évaluation se focalise sur les compétences grapho-morphologiques (dérivationnelle et flexionnelle) générales à l'aide d'une épreuve d'identification de la construction morphologique de pseudo-mots nécessitant la prise en compte conjointe d'aspects orthographiques et sémantiques. Cette épreuve permet d'étudier le niveau de raisonnement sur l'application des règles de constructions morphologiques indépendamment des aspects de surface spécifiques à l'arabe ou aux autres langues connues des participants.

1.2.1 Epreuves de lecture en arabe standard vocalisé

1.2.1.1 *Epreuve de lecture de mots isolés en une minute*

Les compétences de reconnaissance de mots écrits en arabe standard vocalisé sont évaluées à l'aide d'une épreuve individuelle et chronométrée de lecture de mots isolés. Il s'agit de l'adaptation du test de « lecture en une minute » de Khomsi (1999) en langue arabe¹²³.

Une liste de 108 mots vocalisés est présentée au participant (*cf.* annexe 1). Il doit lire ces mots les uns après les autres à voix haute, le plus vite possible mais sans erreur jusqu'à ce que l'expérimentateur lui demande de s'arrêter. Un score de lecture de mots est calculé en comptabilisant le nombre total de mots lus correctement en une minute. Le niveau de performance est ainsi mesuré en fonction de l'exactitude et de la rapidité en lecture.

¹²³ Khomsi et l'équipe de l'Institut Supérieur de Sciences Humaines de Tunis (travaux non publiés).

1.2.1.2 *Epreuve de compréhension de phrases écrites*

Cette épreuve évalue les stratégies de compréhension de phrases de complexité variable. Il s'agit de l'adaptation (légèrement modifiée) en langue arabe du sub-test « compréhension de phrases » du test LMC-R (Khomsî, 1999)¹²⁴.

Elle est composée de 22 vignettes, chacune constituée de 4 images disposées en quadrant, sous lesquelles figurent un énoncé écrit en arabe standard vocalisé (*cf.* annexe 2).

Pour 12 items, le contenu des énoncés est suffisant pour construire une représentation imagée de ce qui est écrit, le choix s'effectue ainsi par une stratégie « imagée » (*e.g.* : le verre n'est pas cassé = الْكَاسُ غَيْرُ مُكَسَّرٍ).

Pour 10 items, le contenu des énoncés renvoie à un traitement temporel et causal de plusieurs actions pour aboutir à la construction d'un état final. La désignation correcte de l'image correspondant à cet état repose dans ce cas sur une stratégie « inférentielle » (*e.g.* : la fille qui a été frappée par le garçon a un chapeau = الْبَيْتُ الَّتِي ضَرَبَهَا الْوَلَدُ لَهَا طَرَبُوشٌ).

Les participants doivent lire l'énoncé puis entourer l'image leur semblant correspondre à la phrase écrite. Ils ont la possibilité de modifier leur réponse s'ils la jugent incorrecte. Le temps n'est pas limité mais les participants sont invités à travailler le plus rapidement possible.

1.2.2 *Epreuves implicites de détection d'intrus*

Le niveau implicite requiert une consigne qui ne spécifie pas l'unité linguistique ou le type d'analyse sur lequel ce choix doit se fonder. Par exemple, on demande aux participants de choisir entre deux mots, celui qui, à son avis, va le moins avec un mot-cible. Dans ces épreuves, les réponses sont engendrées par une association libre, les indices les plus prégnants déterminant le choix des participants à leur insu. Leur objectif est donc de déterminer la sensibilité à certains aspects formels des stimuli (les aspects sémantiques étant contrôlés), ce repérage ne s'effectuant pas par une analyse dirigée par l'expérimentateur qui orienterait l'attention sur les aspects manipulés.

1.2.2.1 *Epreuve implicite de détection d'intrus graphophonologique*

Cette épreuve est destinée à évaluer la sensibilité des participants aux unités syllabiques et intrasyllabiques des mots (substantifs ou verbes conjugués) écrits en arabe standard. Elle comporte 13 items, chacun constitué d'un mot cible et d'une paire de mots dont l'un partage une unité graphophonologique avec la cible et l'autre aucun, jouant alors le rôle d'intrus. Les participants doivent choisir entre les deux mots, celui « allant le moins » avec le mot cible (*e.g.* :

¹²⁴ Khomsî et l'équipe de l'Institut Supérieur de Sciences Humaines de Tunis (travaux non publiés).

qu'est ce qui va le moins avec « شَرُوق » ([ʃʊpʊθ] = lever de soleil) : « سِيَّاق » ([siβɑθ] = course) ou « دِفَاع » ([δʃiφɑθ] = défense) ?). 4 items portent sur la syllabe finale de mots trisyllabiques, 3 items sur la rime intrasyllabique de mots bisyllabiques, 3 sur le phonème consonantique final et 3 sur le phonème vocalique final de mots bisyllabiques. Les items ainsi que les paires de mots à l'intérieur des items ont été présentés selon un ordre aléatoire. La structure phonologique et graphique, la catégorie grammaticale, le rapport avec le lexique de l'arabe dialectal des mots utilisés ont été contrôlés. Les aspects sémantiques des items ont été contrôlés pour les items portant sur la rime, le phonème consonantique et le phonème vocalique (9/13 items). Le matériel, son mode de présentation, sa traduction, et ses caractéristiques sont présentés dans l'annexe 3. Un point est attribué à chaque item lorsque le choix porte sur l'intrus (réponse identifiée par la marque portée dans la case sous le mot écrit).

1.2.2.2 *Epreuve implicite de détection d'intrus grapho-morpho-dérivationnel*

Cette épreuve est destinée à évaluer la sensibilité des participants à la construction morphologique de mots écrits en arabe standard (substantifs ou verbes) qui se caractérise par la superposition d'un schème sur une racine consonantique. Elle comporte 12 items, chacun constitué d'un mot cible et d'une paire de mots. Dans cette paire, l'un des mots fait apparaître une racine identifiable, sémantiquement liée à son dérivé, et partage le même schème que la cible. L'autre mot, jouant le rôle d'intrus, a une racine non identifiable ou non liée sémantiquement avec le mot dérivé, si bien que le partage du schème avec la cible est inexistant ou opaque. Les participants doivent choisir entre les deux mots, celui « allant le moins » avec le mot cible (e.g. : qu'est ce qui va le moins avec « أُغْلِقَ » ([ʔɑφλɑθɑ] = faire fermer, vient de « غَلَقَ » [φɑλɑθɑ] = fermer) : « أَطَالَ » [ʔɑτ5zɑλɑ] = rendre plus long, vient de « طَالَ » [τ5zɑλɑ] = devenir plus long ou « أَفَادَ » [ʔɑφɑδ5ɑ] = apprendre quelque chose à quelqu'un, lui être utile, pas de racine identifiable) ?

La moitié des items concernent la formation de substantifs :

- 3 items portent sur les noms d'instruments (e.g. : cogner → marteau) formés soit sur le schème « مَفْعَلَةٌ » [μiφʃɑλɑτ5υv] constitué d'un graphème « préfixé », des voyelles diacritiques et d'un graphème « suffixé » (2 items), soit sur le schème « مَفْعَالٌ » [μiφʃɑλʊv] constitué d'un graphème « préfixé », des voyelles diacritiques, et d'un graphème, une voyelle longue, infixé (1 item).
- 3 items portent sur les substantifs (e.g. : nourrir → nourriture) formés sur le schème « فَعْلَةٌ » [φʊʃλɑτ5υv] constitué des voyelles diacritiques et d'un graphème « suffixé ».

L'autre moitié des items concernent la formation de verbes augmentés :

- 3 items portent sur la forme III « فاعل » [فَاَعَلَ] schème constitué des voyelles diacritiques, et d'un graphème, une voyelle longue, infixé. La valeur sémantique de cette forme est intensive avec insistance sur l'objet ou la réciprocité de l'action (e.g. : écrire une lettre → correspondre avec quelqu'un) (Larcher, 2003).

- 3 items portent sur la forme IV « أفعّل » [أَفْعَلَّ] schème constitué d'un graphème composé « préfixé » et des voyelles diacritiques. La valeur sémantique de cette forme est factitive (e.g. : étudier → faire étudier) (Larcher, 2003).

Les items ainsi que les paires de mots à l'intérieur des items ont été présentés selon un ordre aléatoire. Les aspects sémantiques des items, la structure phonologique et graphique, le rapport avec le lexique en arabe dialectal des mots utilisés ont été contrôlés, ainsi que la familiarité des paires de mots proposés. Le matériel, son mode de présentation, sa traduction et ses caractéristiques figurent dans l'annexe 4. Un point est attribué à chaque item lorsque le choix porte sur l'intrus (réponse identifiée par la marque portée dans la case sous le mot écrit).

1.2.2.3 *Epreuve implicite de détection d'intrus grapho-morpho-flexionnel*

Cette épreuve est destinée à évaluer la sensibilité des participants à la structure flexionnelle de mots écrits en arabe standard (substantifs ou verbes). Elle comporte 12 items, chacun constitué d'un mot cible et d'une paire de mots dont l'un fait apparaître une structure flexionnelle identique au mot cible et l'autre, jouant le rôle d'intrus, une forme non fléchie ou une structure flexionnelle graphophonologiquement semblable ou ressemblant à celle du mot cible mais syntaxiquement et sémantiquement différente. Les participants doivent choisir entre les deux mots, celui « allant le moins » avec le mot cible (e.g. : qu'est ce qui va le moins avec « عَجَلَةٌ » ([س١زلا٥٥٥٥٥٥] = brebis, féminin de « عَجَلٌ »), « فَيْلَةٌ » ([ف١٦لا٥٥٥٥٥٥] = éléphante, féminin de « فَيْلٌ ») ou « نَاقَةٌ » ([٥٥٦٥٥٥٥٥٥٥٥] = chamelle, pas de masculin) ?),

6 items concernent la formation du féminin de noms d'animaux. Les intrus sont des noms portant la marque du féminin (« ة » [-٥٥٥٥٥٥٥٥] ta marbouta) : mais qui ne sont pas la forme au féminin de leur homologue masculin, soit car ces derniers n'existent pas sous ce lexème, soit car cette forme fléchie correspond au pluriel du nom (e.g. : « لَبْوَةٌ » [لا٥٥٥٥٥٥٥٥٥٥] = lionne, « أُسْدٌ » [٥٥٥٥٥٥٥٥٥٥٥٥] = lion ; « سَمَكَةٌ » [سا٥٥٥٥٥٥٥٥٥٥٥٥], pluriel de « سَمَكٌ » [سا٥٥٥٥٥٥٥٥٥] = poisson).

6 items concernent la conjugaison des verbes à la première personne du singulier¹²⁵ de l'inaccompli (forme fléchie « أفعّل » [٥٥٥٥٥٥٥٥٥٥٥٥]). Les intrus sont des verbes hamzés conjugués à la 1^e P.S. de l'accompli (forme fléchie « فَعَلْتُ » [ف٥٥٥٥٥٥٥٥٥٥٥]), dont la première lettre du radical est

¹²⁵ Abréviation : 1^e P. S.

une hamza - « ء », graphème également utilisé comme la marque préfixée de la conjugaison de l'inaccompli à la 1^e P.S.- (e.g. : « أَمَّنَ » [ʔαμᾶνᾱ] = faire confiance « أَمَّنْتُ » [ʔαμᾶντ5ᵛ] = faire confiance, accompli 1^e P.S.). De façon que ces intrus ne soient pas reconnaissables par leur terminaison en « ت », spécifique à la conjugaison de l'accompli, 1^e P.S., les seconds verbes des paires conjugués, comme les cibles à l'inaccompli, sont des verbes simples dont la dernière lettre du radical est également le graphème « ت ».

Les items ainsi que les paires de mots à l'intérieur des items ont été présentés selon un ordre aléatoire. Les aspects sémantiques des items, la structure phonologique et graphique, le rapport avec le lexique en arabe dialectal des mots utilisés ont été contrôlés, ainsi que la familiarité des paires de mots proposés. Le matériel, son mode de présentation, sa traduction et ses caractéristiques sont présentés dans l'annexe 5. Un point est attribué à chaque item lorsque le choix porte sur l'intrus (réponse identifiée par la marque portée dans la case sous le mot écrit).

1.2.2.4 *Epreuve implicite de détection d'intrus grapho-syntaxique*

Cette épreuve est destinée à évaluer la sensibilité des participants à la structure syntaxique de phrases écrites en arabe standard. Elle comporte 12 items, chacun constitué d'une phrase cible et d'une paire de phrases dont l'une partage la même structure syntaxique de base : « verbe transitif direct – sujet – objet direct » que la phrase cible, et l'autre non car elle se compose d'un verbe transitif indirect qui nécessite de préfixer une préposition à l'objet, jouant ainsi le rôle d'intrus. Les participants doivent choisir entre deux phrases, celle « allant le moins » avec la phrase cible (e.g. : qu'est ce qui va le moins avec « زَارَ عَائِلَتَهُ » ([Zα]ρα 5αʔιλᾶτ5αηᵛ] = il a visité sa famille) : « حَلَقَ لِحَيْتَيْنِ » ([αλαθα λιφᾶτ5αφνι] = il a rasé deux barbes) ou « دَعَا لِبَحَّارَيْنِ » ([δ5α5α λιβαφᾶ]ραφνι] = il a prié pour deux marins) ?).

Pour 6 items, les intrus sont des phrases construites avec la préposition « لِ » [λι], dont la valeur sémantique peut se rapprocher de « pour » ou « à ». Pour les 6 autres items, les intrus sont des phrases construites avec la préposition « بِ » [βι], dont la valeur sémantique peut se rapprocher de « avec ». Pour éviter que la réponse puisse être orientée exclusivement par une reconnaissance graphique de ces prépositions, les autres phrases des paires utilisent comme objet direct des noms commençant par une syllabe identique à la préposition. De plus, de façon que les différences morphosyntaxiques, liées à la déclinaison casuelle des noms, n'apparaissent pas entre les deux paires de phrases, les syntagmes nominaux compléments d'objet sont toujours à la forme duelle pour laquelle on retrouve la même flexion au cas direct et indirect (« لَيْنِ » [αφνι]).

Les items ainsi que les paires de phrases à l'intérieur des items ont été présentés selon un ordre aléatoire. Le nombre de syntagmes, la structure phonologique et graphique des paires de phrases, la familiarité du lexique et la forme des verbes composant les paires de phrases ont été contrôlés. Le matériel, son mode de présentation, sa traduction et ses caractéristiques sont présentés dans l'annexe 6. Un point est attribué à chaque item lorsque le choix porte sur l'intrus (réponse identifiée par la marque portée dans la case sous la phrase écrite).

1.2.3 Epreuves explicites de détection d'intrus

Les épreuves explicites conservent le même principe de choix entre deux stimuli mais font appel à une consigne qui précise le critère de réponse. Les aspects phonologiques, morphologiques, ou syntaxiques sur lesquels porte l'épreuve sont préalablement expliqués. Les participants doivent ainsi déterminer consciemment lequel des deux stimuli de la paire respecte le critère précisé dans la consigne.

Chaque épreuve de détection explicite d'intrus est construite sur le même mode que leur homologue implicite. Ainsi elles comportent le même nombre d'items, les mêmes caractéristiques et le même mode de cotation que les épreuves implicites décrites précédemment. Le contrôle lors de leur construction est identique, seules les consignes diffèrent. Par commodité, il ne sera présenté dans les paragraphes suivants que le type de compétences évaluées, accompagné d'exemples précisant ce qui est demandé aux participants.

1.2.3.1 *Epreuve explicite de détection d'intrus graphophonologique*

Cette épreuve a pour objectif d'évaluer les connaissances graphophonologiques explicites permettant de discriminer consciemment les éléments graphophonologiques de mots écrits en arabe standard. Les participants doivent déterminer entre deux mots, lequel ne se termine pas par le même son qu'un mot cible. Par exemple, qu'est ce qui ne se termine pas par le même son que « أُجْرَاس » ([ʔαZpα]σ = sonnettes) : « بُرْنُوس » ([βυρνυ]σ = burnous) ou « طُرْبُوش » ([τ5≥υρβυ]Σ = chapeau) ? Le matériel, son mode de présentation, sa traduction et ses caractéristiques sont présentés dans l'annexe 7.

1.2.3.2 *Epreuve explicite de détection d'intrus grapho-morpho-dérivationnel*

Cette épreuve a pour objectif d'évaluer les connaissances morpho-dérivationnelles explicites permettant d'isoler consciemment la racine verbale d'un mot dérivé et d'identifier le lien sémantique entre la racine et ce dérivé. Dans cette épreuve, les participants doivent déterminer le mot, parmi deux proposés, qui malgré les apparences phonographiques ne fait pas clairement

apparaître la même structure morphologique racine/schème qu'un mot cible. Il leur est par exemple demandé quel est le mot qui ne vient visiblement pas d'un autre mot, qui n'a pas de racine identifiable comme « مرآة » ([μρ?ατ5υν] = miroir, racine voir : « رآى »), « مكنسة » ([μικνασατ5υν] = balai, racine balayer : « كَنَسَ ») ou « مئذنة » ([μινΔ5≥αδ5ατ5υν] = petite table, racine inusitée) ? Le matériel, son mode de présentation, sa traduction et ses caractéristiques sont présentés dans l'annexe 8.

1.2.3.3 *Epreuve explicite de détection d'intrus grapho-morpho-flexionnel*

Cette épreuve a pour objectif d'évaluer les connaissances morphosyntaxiques explicites permettant d'identifier consciemment les aspects syntaxiques des mots tels que le genre des substantifs ou l'aspect des verbes à partir d'éléments graphiques discriminants et en fonction des conventions lexicales de la langue. Dans cette épreuve, les participants doivent déterminer, parmi deux propositions et malgré les apparences phonographiques, le nom qui n'est pas assigné d'une marque de genre, ou le verbe qui n'est pas marqué par la conjugaison de l'inaccompli. Il leur est par exemple demandé quel est le mot qui n'est pas le féminin d'un mot masculin comme « أفعة » ([?αφ8ατ5υν] = vipère, féminin de « أفع »), « حمارة » ([□ιμ]αρατ5υν] = ânesse, féminin de : « حمار ») ou « ذبابة » ([δ5υβα]βατ5υν] = mouche, pluriel de « ذباب ») ? Ou encore quel est le verbe qui n'est pas à l'inaccompli comme : « أصل » ([?ασ5≥ιλν] = j'arrive du verbe « أصل »), « أسكت » ([?ασκντ5υ] = je me tais, du verbe « سكت ») ou « أكلت » ([?ακαλτ5υ] = j'ai mangé, du verbe : « أكل ») ? Le matériel, son mode de présentation, sa traduction et ses caractéristiques sont présentés dans l'annexe 9.

1.2.3.4 *Epreuve explicite de détection d'intrus grapho-syntaxique*

Cette épreuve a pour objectif d'évaluer les connaissances syntaxiques explicites permettant d'analyser consciemment la structure syntaxique de phrases en terme de transitivité des verbes de façon à identifier l'utilisation d'une préposition affixée au complément d'objet. Dans cette épreuve, les participants doivent déterminer, parmi deux propositions et malgré les apparences phonographiques, la phrase pour laquelle le verbe implique une préposition accolée au complément contrairement à la phrase cible. Il leur est par exemple demandé quelle est la phrase qui contient un « petit mot » devant le complément qu'il n'y a pas dans « دَخَنَ سِجَارَةً » ([δ5αΦΦανα σι]Zα]ρατ5αν] = il a fumé une cigarette), « باعَ لِأَخْتَيْنِ » ([βα]8α λιφα]φατ5αφνι] = il a vendu deux couvertures) ou « حَكَمَ لِأَخْتَيْهِ » ([□ακαμα λι?αΦτ5αφνι] = il a donné raison à ses deux soeurs) ? Le matériel, son mode de présentation, sa traduction et ses caractéristiques sont présentés dans l'annexe 10.

1.2.4 Epreuve d'identification de la construction morphologique de pseudo mots

Cette épreuve évalue les compétences explicites à extraire les règles morphologiques sous-jacentes à la construction de formes lexicales inconnues des participants. Même si dans la consigne, il n'est pas à proprement parlé précisé comment les participants doivent résoudre l'épreuve, les réponses correctes ne peuvent être trouvées sans une réflexion consciente et une manipulation attentive des aspects morphologiques. En effet, pour chaque item, il est demandé d'identifier à partir de deux exemples le pseudo-mot pouvant correspondre à une définition. Les deux exemples indiquent le sens d'une pseudo-racine et de deux pseudo-mots, composés de cette pseudo-racine et d'un pseudo-morphème préfixé, suffixé ou infixé de plusieurs voyelles (affixation de type « schème arabe »). Les deux mêmes pseudo-morphèmes sont utilisés dans les deux exemples. A la suite de ces deux exemples, une autre pseudo-racine est présentée ainsi qu'une question portant sur la forme lexicale en en formulant sa définition (définition dans laquelle on retrouve le sens de cette pseudo-racine). L'expérimentateur explique par exemple aux participants que dans une langue que personne ne connaît le « kazi », *chat* se dit XOT, *petit chat* XOTIKE et *grand chat* XOTUKO (exemple 1). Qu'en « kazi » *vélo* se dit GOF, *petit vélo* GOFIKE et *grand vélo* GOFUKO (exemple 2), enfin que *pot* se dit RUV (pseudo-racine). Alors, comment dirait-on *petit pot* en kazi : RUVIKE RUFIKE ou RUVUKO ?).

Les participants ont le choix entre trois réponses possibles, une utilisant la racine correcte et le morphème correct par rapport à la définition donnée (réponse correcte pour *petit pot* : RUV = *pot* + IKE = *petit*), une utilisant la racine correcte et un morphème présenté dans les exemples mais ne correspondant pas à la définition donnée (réponse incorrecte pour *petit pot* : RUV = *pot* + UKO = *grand*) et une dont la racine est modifiée et utilisant un des deux morphèmes présentés dans les exemples (réponse incorrecte pour *petit pot* : RUF \neq *pot* + IKE = *petit*)¹²⁶. Pour chaque item, les instructions et les définitions des pseudo-mots sont écrites en langue arabe. Cette épreuve se compose de 24 items (cf. annexe 11) : 12 font référence à des règles morpho-dérivationnelles de création lexicale (e.g. : faire/faire encore une fois, forêt/celui qui travaille dans la forêt) et 12 à des règles morphosyntaxiques comme l'accord en genre, en nombre, la conjugaison des verbes etc. (e.g. : propre pour un garçon/ propre pour une fille, c'est moi qui chante/ c'est elle qui chante). Pour un tiers des items, la construction des pseudo mots est de type racine/préfixe. Pour un autre tiers, la construction est de type racine/suffixe. Pour le dernier tiers, la construction est de type racine/schème vocalique. Ces trois formes de construction sont

¹²⁶ Un pré-test mené auprès d'enfants francophones de CP, CE1 et CE2 a permis d'observer que la construction des pseudo-mots proposés comme distracteurs n'influçait pas les réponses.

réparties équitablement entre les deux types de règles morphologiques. L'ordre de présentation des items en fonction de leurs caractéristiques ainsi que l'ordre de lecture des 3 réponses proposées, à l'intérieur de chaque item sont définis de façon aléatoire¹²⁷.

1.3 Procédure

Compte tenu du nombre important d'épreuves, plusieurs sessions se sont avérées nécessaires pour chaque classe¹²⁸. Lors de chaque session, la présentation du matériel s'effectuait de façon à respecter le rythme de la classe mais sans qu'aucune épreuve commencée ne soit inachevée.

L'ordre constant de présentation des épreuves ou groupes d'épreuves en arabe est le suivant :

- Epreuves implicites de détection d'intrus en passation collective. Pour chaque classe, l'ordre de présentation des épreuves graphophonologique, grapho-morpho-dérivationnelle, grapho-morpho-flexionnelle et grapho-syntaxique est différente de façon à contrebalancer l'effet possible d'une épreuve sur la suivante. Afin de s'assurer que les participants avaient compris ce que l'on attend d'eux, la première épreuve était précédée d'un exemple donné lors de la consigne et de deux items d'entraînements sans *feed-back* (cf. annexe 12) à remplir sur le cahier d'exercice. Ils sont constitués, selon l'épreuve, de substantifs, verbes ou phrases, présentés sous la même forme que les items de l'épreuve. Pour que les participants ne perçoivent pas les aspects étudiés dans cette épreuve, la distinction partage/non partage d'une unité linguistique avec le stimulus cible est absente. Du fait que la consigne ne présente aucune difficulté et se répète pour les quatre épreuves, il a été jugé inutile de proposer des exemples et un entraînement sur toutes les épreuves, présentées les une à la suite des autres.
- Epreuve d'identification de la construction morphologique de pseudo mots. La passation est collective. Deux items d'entraînement sans *feed-back* ont été utilisés de façon à familiariser les participants à cette tâche inhabituelle (cf. annexe 11). Dans ces items, la construction morphologique des pseudo-mots fait intervenir l'ajout d'infixes et non de préfixes, suffixes ou schèmes vocaliques comme cela est le cas dans les items cotés. L'expérimentateur, après avoir expliqué la consigne n'intervient pas au cours de sa résolution, qui s'effectue en silence.
- Epreuves explicites de détection d'intrus en passation collective. Pour chaque classe, l'ordre de présentation des épreuves graphophonologique, grapho-morpho-dérivationnelle, grapho-morpho-flexionnelle et grapho-syntaxique est randomisé et diffère de l'ordre fixé pour les épreuves implicites. Lors de la consigne, quatre exemples sont utilisés pour expliquer les aspects sur lesquels les participants doivent porter leur attention. De plus, chaque épreuve était précédée

¹²⁷ Les différentes constructions possibles des deux pseudo-mots incorrects proposés sont réparties de façon équitable sur les items de chaque type.

¹²⁸ Au cours de cette expérimentation, d'autres sessions proposant des épreuves en français ont été intégrées. Ces épreuves et les résultats obtenus sont présentés dans le chapitre 9

de deux items d'entraînement avec correction après chaque item (cf. annexes 7 à 10). Ils sont constitués, selon l'épreuve, de substantifs, verbes ou phrases. Lors de la correction, la distinction partage/non partage d'une unité linguistique avec le stimulus cible est de nouveau explicitée. Les unités choisies dans ces items ne sont pas les mêmes que celles utilisées dans l'épreuve mais appartiennent au même domaine linguistique.

- Epreuve de compréhension de phrases écrites. La passation est collective mais l'expérimentateur, après avoir expliqué la consigne n'intervient pas au cours de sa résolution, qui s'effectue en silence. Avant de débiter l'épreuve, un item d'entraînement sans *feed-back* est proposé sur la feuille d'exercice.

- Test de lecture de mots écrits en une minute en passation individuelle.

Les épreuves administrées en passations collectives se sont déroulées au cours de trois sessions d'environ 45 minutes sans dépasser le rythme d'une session par semaine. Les entretiens individuels au cours desquels le test de lecture de mots écrits était administré n'excédaient pas cinq minutes par enfant. Ils ont eu lieu sur une ou deux demi-journées pour chaque classe.

Le tableau 5 récapitule les types d'épreuves évaluant les connaissances sur les dimensions graphophonologique et grapho-sémantique¹²⁹

¹²⁹ Les instructions ayant été donnés en arabe, les consignes ont été traduites en français.

Tableau 5 : Epreuves évaluant les connaissances sur les dimensions graphophonologique et grapho-sémantique

	Dimension graphophonologique	Dimension grapho-sémantique		
		Morpho-dérivationnelle	Morpho-flexionnelle	Syntaxe
<p><u>Epreuves implicites de détection d'intrus</u>^{1,2}</p> <p>Présentation à l'oral + support écrit Entraînement au paradigme lors de la 1^o épreuve (2 items)</p>	Consigne implicite commune : Qu'est ce qui va le moins avec :			
	<p>شُرُوق [ʃʊʁuq]θ]</p> <p>دِفَاع ou سِيَاق [dɪfɑɑ]θ] [sɪjɑɑ]θ]</p> <p>cf. annexe 3</p>	<p>أَعْتَق (faire fermer)</p> <p>أَفَاد ou أَطَالَ (allonger) (être utile)</p> <p>cf. annexe 4</p>	<p>عَجَلَة (brebis)</p> <p>نَاقَة ou فِيلَة (éléphante) (chamelle)</p> <p>cf. annexe 5</p>	<p>زَارَ عَائِلَتَهُ (il a visité sa famille)</p> <p>دَعَا لِيَحَارِينَ ou حَقَّقَ لِحَيْتَيْنِ (il a rasé 2 barbes) (il a prié pour 2 marins)</p> <p>cf. annexe 6</p>
<p><u>Epreuve d'identification de la construction morphologique de pseudo-mots</u>²</p>		<p>En kazi : <i>chat</i> = XOT <i>petit chat</i> = XOTIKE <i>grand chat</i> = XOTUKO <i>vélo</i> = GOF, <i>petit vélo</i> = GOFIKE <i>grand vélo</i> = GOFUKO <i>pot</i> = RUV. Comment dirait-on <i>petit pot</i> en kazi :</p> <p style="text-align: center;"><u>RUVIKE</u>, RUFIKE ou RUVUKO ?</p> <p>cf. annexe 11</p>		
<p><u>Epreuves explicites de détection d'intrus</u>^{1,2}</p> <p>Présentation à l'oral + support écrit Entraînement systématique à la tâche (2 items)</p>	<p>Consigne explicite : Qu'est ce qui ne se termine pas par le même son que :</p> <p>أَجْرَاس [ʔɑzrɑɑ]θ]</p> <p>طَرَبُوش ou بَرَبُوش [tɑrɒuʃ]θ] [tɑrɒuʃ]θ]</p> <p>cf. annexe 7</p>	<p>Consigne explicite : Qu'est ce qui ne vient pas d'un autre mot comme :</p> <p>مِرَاة (miroir)</p> <p>مِنْدَدَة ou مَكْنَسَة (balai) (petite table)</p> <p>cf. annexe 8</p>	<p>Exemple de consigne explicite : Qu'est ce qui n'est pas à l'inaccompli comme :</p> <p>أَصَلَ (j'arrive)</p> <p>أَكَلْتُ ou أَسَكْتُ (je me tais) (j'ai mangé)</p> <p>cf. annexe 9</p>	<p>Consigne explicite : Quelle phrase a une préposition (petit mot) qu'il n'y a pas dans :</p> <p>دَخَنَ سِيَجَارَة (il a fumé une cigarette)</p> <p>حَكَّمَ لِأَخْتَيْهِ ou بَاعَ لِقَافَتَيْنِ (il a vendu 2 couvertures) (il a donné raison à ses 2 soeurs)</p> <p>cf. annexe 10</p>

¹ Randomisation de l'ordre de présentation des 4 épreuves

² Randomisation de l'ordre des items et des réponses dans chaque item

2. Résultats

2.1 Démarche statistique

En fonction des hypothèses posées, et des caractéristiques des outils d'évaluation, plusieurs types d'analyses statistiques ont été menés avec parfois certains critères d'exclusion selon les résultats obtenus :

Dans un premier temps, nous nous sommes attachée à analyser et comparer les performances de chaque groupe.

Afin de déterminer si les compétences en lecture et les connaissances de différents aspects du système d'écriture arabe augmentent entre la 3^e et la 5^e année, plusieurs analyses de variance ont été effectuées.

- Sur les deux épreuves de lecture (lecture de mots isolés en une minute et compréhension de phrases), des analyses de variance à un facteur ont été menées, afin de préciser les variations du niveau de lecture entre la 3^e et la 5^e année d'apprentissage.

- Concernant les épreuves implicites et explicites de détection d'intrus, notre objectif était de déterminer l'influence du niveau scolaire en fonction du degré de contrôle (implicite *vs.* explicite) et du domaine linguistique (graphophonologique, grapho-morpho-dérivationnelle, grapho-morpho-flexionnelle, grapho-syntaxique). Mais il s'est avéré que l'ensemble de ces épreuves n'a pas pu être retenu. Les résultats aux tests t pour échantillon unique, comparant les moyennes à la valeur du hasard, ont été utilisés pour déterminer les résultats pouvant faire l'objet d'analyses statistiques inférentielles. La résolution de ces épreuves utilisant le paradigme de choix forcé donne lieu à un choix entre deux propositions. Les scores avoisinant une bonne réponse sur deux¹³⁰ peuvent alors correspondre à des réponses au hasard, ce qui légitime l'élimination des épreuves concernées par ce problème, notamment si ce phénomène concerne l'ensemble des groupes. Ainsi, des analyses en sous-ordre à deux facteurs (niveau scolaire et domaine linguistique) ont été requises du fait de l'apparition de réponses aléatoires et de scores plafonds sur plusieurs épreuves et bien que cela dénature la cohérence méthodologique prévue. Les analyses *a posteriori* ont été systématiquement menées à l'aide du test de Newman-Keuls.

- L'étude des unités graphophonologiques les plus pertinentes en arabe n'a pu être menée que sur le niveau implicite des connaissances¹³¹. Nous avons donc cherché à déterminer si les

¹³⁰ Les termes de réussite ou de bonnes réponses aux items des épreuves de détection d'intrus seront employés pour qualifier le choix de l'intrus bien que pour les épreuves implicites, en l'absence de consigne, le choix corresponde à une préférence et non à une bonne ou une mauvaise réponse.

¹³¹ Cette analyse n'a pas pu être menée sur le niveau explicite des connaissances graphophonologiques compte tenu des scores plafonds observés.

performances des élèves de chaque niveau à l'épreuve de détection implicite d'intrus graphophonologique variaient selon le type d'unité impliqué. Les effets de la nature des items ont été analysés en fonction du niveau scolaire à l'aide de tests non paramétriques (test de Friedman, de Wilcoxon) car les critères de normalité et d'homocédasticité des résultats n'étaient pas toujours respectés.

- Afin de déterminer si la morphologie de la langue arabe a un impact sur les compétences à extraire explicitement les règles morphologiques sur un matériel lexical inconnu, il a été menée sur l'épreuve d'identification de la construction morphologique de pseudo-mots, une analyse à trois facteurs : le niveau scolaire {3^e, 4^e et 5^e année}, la nature de la construction morphologique {préfixe, suffixe, schème} et le type de morphologie auquel font référence les règles utilisées dans les items {dérivationnelle vs. flexionnelle}.

Ces analyses ont également permis de sélectionner les épreuves ou groupes d'items pouvant être retenus dans les études corrélatives, factorielles et de régressions ultérieures.

Dans un deuxième temps, afin d'identifier les relations entre les connaissances pouvant intervenir lors de l'apprentissage de l'écrit en arabe, des analyses de corrélations et des analyses factorielles à composantes principales ont été réalisées sur les résultats aux épreuves ou groupes d'items retenus lors des analyses présentées plus haut. Les analyses de corrélation donnent des indications sur le degré d'association entre certaines épreuves évaluant des domaines regroupables sous une même dimension. L'application d'une analyse factorielle en composantes principales permet à la fois de déterminer plus exactement l'existence d'un facteur commun sous-jacent à certaines de ces mesures et de rechercher celles qui expliquent la plus grande part de la variance entre les élèves. Ces deux types d'analyses ont été menés séparément sur chaque niveau scolaire. Il faut d'ailleurs souligner que les épreuves conservées sont parfois différentes d'un niveau à l'autre, ce qui constitue un argument supplémentaire pour traiter chaque niveau de façon isolée.

Dans un troisième temps, nous nous sommes attachée à rechercher les performances pouvant déterminer le niveau de lecture en arabe en 3^e, 4^e et 5^e année. Des analyses de corrélations ont d'abord été réalisées pour identifier le degré d'association entre les mesures de chaque domaine et les scores en lecture de mots et en compréhension de phrases. Les variables obtenant des coefficients de corrélation significatifs ont alors été entrées dans un modèle de régression pas à pas afin de déterminer la contribution des variables à l'explication du niveau de lecture de mots et de compréhension de phrases écrites. Pour les raisons évoquées plus haut et pour faciliter la

comparaison des résultats obtenus en arabe et en portugais, ces deux types d'analyses ont été menés séparément sur chaque niveau scolaire et sur les deux dimensions de la lecture¹³².

2.2 Analyse des performances obtenues

2.2.1 Epreuve de lecture de mots : effet du niveau scolaire

Le tableau 6 indique les moyennes, écart-types, minima et maxima obtenus à l'épreuve de lecture de mots en une minute, selon le niveau scolaire.

Tableau 6 : Performances moyennes, écart-types, minima et maxima au test de lecture de mots en une minute selon le niveau scolaire.

	Moyenne	Ecart-type	Minimum	Maximum
3 ^e année N = 59	55,53	18,08	9	87
4 ^e année N = 43	62,12	15,88	24	104
5 ^e année N = 50	73,84	18,94	30	108

L'analyse de variance (niveau scolaire {3^e année, 4^e année, 5^e année}) met en évidence un effet du niveau scolaire sur les performances en lecture de mots isolés [$F(2,149) = 14,51, p < .00001$] : les élèves de 5^e année lisent en moyenne plus de mots correctement que ceux de 4^e année ($p < .002$) et de 3^e année ($p < .0001$). La différence entre les 4^e année et 3^e n'est pas significative mais toutefois tendancielle ($p < .06$). **Les capacités de reconnaissance évoluent donc surtout entre la 4^e et la 5^e année.**

2.2.2 Epreuve de compréhension de phrases écrites : effet du niveau scolaire

Le tableau 7 indique les moyennes, écart-types, minima et maxima obtenus à l'épreuve de compréhension de phrases selon le niveau scolaire. Bien qu'ici les réponses sont à choisir entre quatre propositions, il n'a pas été nécessaire de conduire des tests t pour échantillon unique pour s'assurer que les réponses ne sont pas liées au hasard (1/4 de bonnes réponses). En effet, les scores s'élèvent systématiquement bien au dessus de 25 % (5,5/22) de bonnes réponses (la moyenne la plus basse est de 13,60/22).

¹³² Une approche par modélisation en équations structurales avait été envisagée, mais compte tenu du faible effectif des groupes comparativement au nombre important de variables, nous y avons renoncé.

Tableau 7 : Performances moyennes de réussite (sur 22), écart-types, minima, maxima obtenus au test de compréhension de phrases selon le niveau scolaire

	Moyenne	Ecart-type	Minimum	Maximum
3 ^e année N = 52	13,77	2,93	6	20
4 ^e année N = 43	13,60	2,61	8	20
5 ^e année N = 50	16,66	2,70	11	21

L'analyse de variance à un facteur (niveau scolaire {3^e, 4^e année, 5^e année}) met en évidence un effet du niveau scolaire sur les performances en compréhension [$F(2,149) = 18,93$; $p < .00001$] : les élèves de 5^e année comprennent en moyenne mieux que ceux de 4^e année ($p < .00001$) et de 3^e année ($p < .00001$). La différence entre les 4^e et 3^e année n'est pas significative.

Nous pouvons ainsi constater que les capacités de compréhension en arabe littéral évoluent entre la 4^e et la 5^e année sans pour autant atteindre à cet âge un score maximal.

2.2.3 Epreuves implicites et explicites de détection d'intrus

2.2.3.1 Sélection des épreuves en fonction des résultats obtenus

Le tableau 8 récapitule les performances moyennes obtenues par les trois niveaux scolaires aux épreuves implicites et explicites de détection d'intrus graphophonologique, grapho-morpho-dérivationnel, grapho-morpho-flexionnel et grapho-syntaxique.

Tableau 8 : Performances moyennes de réussite sur 12 (et écart-types) obtenues aux épreuves implicites et explicites de détection d'intrus en fonction du niveau scolaire, tests-t de comparaison à la valeur 6 (niveau du hasard).

Epreuves de détection d'intrus	3 ^e année	4 ^e année	5 ^e année	Tous niveaux scolaires Comparaison au hasard (6/12)
Implicite graphophonologique ¹	6,36 (1,84) N = 56	6,41 (2,17) N = 41	6,40 (2,18) N = 50	6,39 (2,04) t(146) = 2,29 ; p < .05
Implicite grapho-morpho-dérivationnel	5,82 (1,61) N = 56	6,02 (1,72) N = 41	5,80 (1,46) N = 50	5,87 (1,58) t(146) > -1
Implicite grapho-morpho-flexionnel	5,59 (1,97) N = 56	6,22 (1,29) N = 41	5,50 (1,79) N = 50	5,73 (1,76) t(146) = -1,83 ; ns
Implicite grapho-syntaxique	5,93 (1,41) N = 55	6,41 (1,77) N = 39	6,08 (1,60) N = 50	6,12 (1,55) t(141) = 1,24 ; ns
Explicite graphophonologique ¹	11,39 ^P (1,38) N = 56	11,27 ^P (1,52) N = 43	11,58 ^P (0,75) N = 50	11,42 ^P (1,25) t(148) = 52,97 ; p < .001
Explicite grapho-morpho-dérivationnel	7,11 (2,43) N = 56	7,67 (2,29) N = 43	9,10 (2,17) N = 50	7,94 (2,44) t(148) = 9,69 ; p < .001
Explicite grapho-morpho-flexionnel	6,80 (2,48) N = 56	6,70 (2,05) N = 43	8,98 (2,27) N = 50	7,50 (2,51) t(148) = 7,31 ; p < .001
Explicite grapho-syntaxique	8,06 (2,72) N = 54	8,35 (3,12) N = 43	10,16 ^P (2,73) N = 50	8,86 (2,98) t(146) = 11,63 ; p < .001

¹ Les deux épreuves graphophonologiques comportant 13 items, leurs scores ont été rapportés à un total sur 12 pour pouvoir être comparés à ceux des autres épreuves.

^P Scores plafonds.

En gras : les scores pouvant être retenus pour effectuer des analyses inférentielles.

La validité des mesures a été évaluée à l'aide des résultats aux tests t de comparaison à une valeur standard, sans distinction de groupe de manière à attester que, de façon générale, les réponses ne sont effectivement pas liées au hasard. Dans le cas contraire, il a toutefois été vérifié que pour les trois groupes, les réponses ne peuvent pas être attribuées à une sensibilité aux aspects manipulés avec certitude.

Les valeurs aux tests t, obtenues pour l'ensemble des niveaux scolaires font ainsi apparaître que les résultats aux épreuves implicites grapho-morpho-dérivationnelles, grapho-morpho-flexionnelles et grapho-syntaxiques ne se distinguent pas significativement de la valeur 6 (cf. tableau 8). Cette absence de différence par rapport au hasard concerne de surcroît les trois groupes distingués (cf. annexe 13). Ces observations justifient que, pour tous les groupes, ces résultats soient exclus des analyses statistiques suivantes.

Cependant, les résultats à l'épreuve implicite d'intrus graphophonologique sont faibles (moyenne générale de 53% de bonnes réponses) mais toutefois les seuls à se différencier positivement des

réponses données au hasard ($p < .05$, cf. tableau 8). Cela justifie qu'une analyse comparative soit menée sur ces résultats.

Ainsi, il semble que le paradigme de détection implicite d'intrus s'attachant ici à diverses caractéristiques linguistiques ne parvient à faire émerger pour l'ensemble des niveaux scolaires qu'une légère sensibilité aux aspects graphophonologiques de la langue arabe.

En revanche, lorsque la consigne est explicite, les épreuves utilisant ce même paradigme obtiennent des taux de réussite appréciables et supérieurs au niveau du hasard ($p < .001$, cf. tableau 8). Toutefois, les scores aux épreuves graphophonologiques explicites pour les trois niveaux scolaires et à l'épreuve grapho-syntaxique explicite pour le niveau 5^e année font apparaître des effets plafonds (respectivement 95% et 85 % de bonnes réponses). Pour l'épreuve graphophonologique, entre 86 et 88% des élèves répondent correctement à au moins 12 items sur 13. Ces résultats indiquent que dès la 3^e année, la discrimination consciente des éléments graphophonologiques arabes est parfaitement maîtrisée. Pour l'épreuve grapho-syntaxique, 68% des élèves de 5^e année répondent correctement à au moins 11 items sur 12. Là encore cela montre qu'à ce niveau scolaire les connaissances syntaxiques conscientes, permettant d'identifier la segmentation grammaticale des mots composant une phrase, sont acquises. De ce fait, ces deux épreuves ne pourront être prises en considération dans les analyses statistiques ultérieures aux niveaux scolaires indiqués.

Compte tenu des observations précédentes, parmi les épreuves implicites de détection d'intrus, seuls les résultats à l'épreuve graphophonologique pourront faire l'objet de tests inférentiels. Cette même dimension n'est toutefois pas analysable au niveau explicite du fait des scores plafonds observés quel que soit le niveau scolaire. Si la comparaison statistique ne peut être menée sur ces deux niveaux de contrôle, l'observation de moyennes proches du hasard dans un cas et proches du plafond dans l'autre suffit pour affirmer la différence. De même, la comparaison entre niveau implicite et explicite sur les connaissances grapho-sémantiques n'est pas non plus « statistiquement possible ».

Ainsi, concernant les épreuves explicites restantes, une analyse pour plan à deux facteurs (niveau scolaire {3^e, 4^e, 5^e} et type de connaissances morphologiques {morpho-dérivationnelle vs. morpho-flexionnelle}) sera menée. Elle sera complétée d'une seconde analyse se focalisant sur les niveaux 3^e et 4^e année de façon à inclure la modalité « connaissances syntaxiques » à l'analyse du facteur « connaissances grapho-sémantiques » (plan à deux facteurs : niveau

scolaire {3^e vs. 4^e année} et domaine grapho-sémantique {morpho-dérivationnel, morpho-flexionnel, syntaxique})¹³³.

Par ailleurs, une analyse sur les connaissances graphophonologiques implicites seulement, portant sur les facteurs « niveau scolaire » et « unité phonologique », sera réalisée.

2.2.3.2 *Epreuves explicites de détection d'intrus grapho-sémantique : effets du niveau scolaire et du domaine linguistique*

Une première analyse pour plan à deux facteurs (niveau scolaire {3^e, 4^e, 5^e} et type de morphologie {dérivationnelle vs. flexionnelle}) a été effectuée sur les résultats aux épreuves de détection d'intrus grapho-morphologiques. Ces derniers sont présentés dans le tableau 9.

Tableau 9 : Performances moyennes de réussite sur 12 (et écart-types) obtenues aux épreuves explicites de détection d'intrus grapho-morphologique en fonction du niveau scolaire et tests t de comparaison à la valeur 6 (niveau du hasard).

	3 ^e année N = 56	4 ^e année N = 43	5 ^e année N = 50	Moyenne tous niveaux scolaires N = 149
Explicite grapho-morpho-dérivationnel	7,11 (2,43) t(55) = 3,41, p < .001	7,67 (2,29) t(42) = 4,8, p < .001	9,10 (2,17) t(49) = 10,11, p < .001	7,94 (2,44)
Explicite grapho-morpho-flexionnel	6,80 (2,48) t(55) = 2,42, p < .05	6,70 (2,05) t(42) = 2,23, p < .05	8,98 (2,27) t(49) = 9,27, p < .001	7,50 (2,51)
Moyenne Explicite grapho-morphologique	6,96 (1,93)	7,19 (1,17)	9,04 (1,76)	7,72 (2,03)

Il peut être constaté un effet simple du niveau scolaire [$F(2,146) = 20,15, p < .0001$] : les élèves de 5^e année ont globalement des résultats supérieurs aux élèves de 3^e année ($p < .00001$) et aux élèves de 4^e année ($p < .0001$), la différence de performance entre les élèves de 3^e et de 4^e année n'est en revanche pas significative.

L'effet simple du type de morphologie est également significatif [$F(2,146) = 3,99, p < .05$] : les résultats sont supérieurs à l'épreuve grapho-morpho-dérivationnelle par rapport à l'épreuve grapho-morpho-flexionnelle. Les performances aux deux épreuves sont d'ailleurs supérieures au niveau du hasard pour les trois niveaux scolaires (cf. tableau 9).

L'interaction niveau scolaire * type de morphologie n'est pas significative [$F(2,146) = 1,15, ns$].

De façon à intégrer les résultats à l'épreuve grapho-syntaxique aux comparaisons, une seconde analyse pour plan à deux facteurs (niveau scolaire {3^e vs. 4^e}, et domaine grapho-sémantique {morpho-dérivationnel, morpho-flexionnel, syntaxique}) a été menée sur les résultats des deux

¹³³ Les analyses paramétriques en sous ordre ont été choisies car, hormis pour les scores supprimés, les critères de normalité et d'homocédasticité des résultats aux trois niveaux scolaires sont respectés.

premiers niveaux scolaires pour lesquels aucun effet plafond n'est observé. Les résultats comparés lors de cette analyse sont présentés dans le tableau 10¹³⁴.

Tableau 10 : Performances moyennes de réussite sur 12 (et écart-types) obtenues aux épreuves explicites de détection d'intrus grapho-sémantique en fonction du niveau scolaire et tests t de comparaison à la valeur 6 (niveau du hasard).

	3 ^e année N = 54	4 ^e année N = 43	Moyenne deux niveaux scolaires N = 97
Explicite grapho-morpho-dérivationnel	7,06 (2,46) t(53) = 3,15 p < .01	7,67 (2,29) t(42) = 4,80 p < .001	7,33 (2,37)
Explicite grapho-morpho-flexionnel	6,74 (2,50) t(53) = 2,17 p < .05	6,70 (2,05) t(42) = 2,23 p < .05	6,72 (2,30)
Explicite grapho-syntaxique	8,06 (2,72) t(53) = 5,55 p < .001	8,35 (3,12) t(42) = 4,94 p < .001	8,19 (2,89)
Moyenne explicite grapho-sémantique	7,29 (1,94)	7,57 (1,88)	7,41 (1,90)

Les résultats mettent en évidence un effet simple du domaine grapho-sémantique [$F(2,95) = 12,44$, $p < .00001$] : les résultats sont globalement supérieurs à l'épreuve grapho-syntaxique par rapport aux épreuves grapho-morpho-flexionnelle ($p < .0001$) et grapho-morpho-dérivationnelle ($p < .005$). De plus, comme dans l'analyse précédente, les résultats à l'épreuve grapho-morpho-dérivationnelle sont supérieurs à ceux obtenus à l'épreuve grapho-morpho-flexionnelle ($p < .05$). Les performances obtenues à l'épreuve explicite grapho-syntaxique sont d'ailleurs supérieures au niveau du hasard pour les deux niveaux scolaires ($p < .001$) comme cela avait été observé sur les épreuves explicites grapho-morpho-flexionnelle et grapho-morpho-dérivationnelle (cf. tableau 9).

La différence entre les 3^e et les 4^e années, pour l'ensemble des trois épreuves grapho-sémantiques, n'est toujours pas significative [$F(1,95) < 1$], ni l'interaction niveau scolaire * domaine grapho-sémantique [$F(2,190) < 1$].

La figure 8 reprend les résultats obtenus aux trois épreuves explicites de détection d'intrus grapho-sémantique en fonction du niveau scolaire, en incluant les scores des élèves de 5^e année.

¹³⁴ Les variations de l'effectif et donc des moyennes des 3^e année aux 2 épreuves grapho-morphologiques par rapport au tableau 4 et 5 s'expliquent par le fait que 2 élèves n'ont pas de scores à l'autre épreuve grapho-syntaxique.

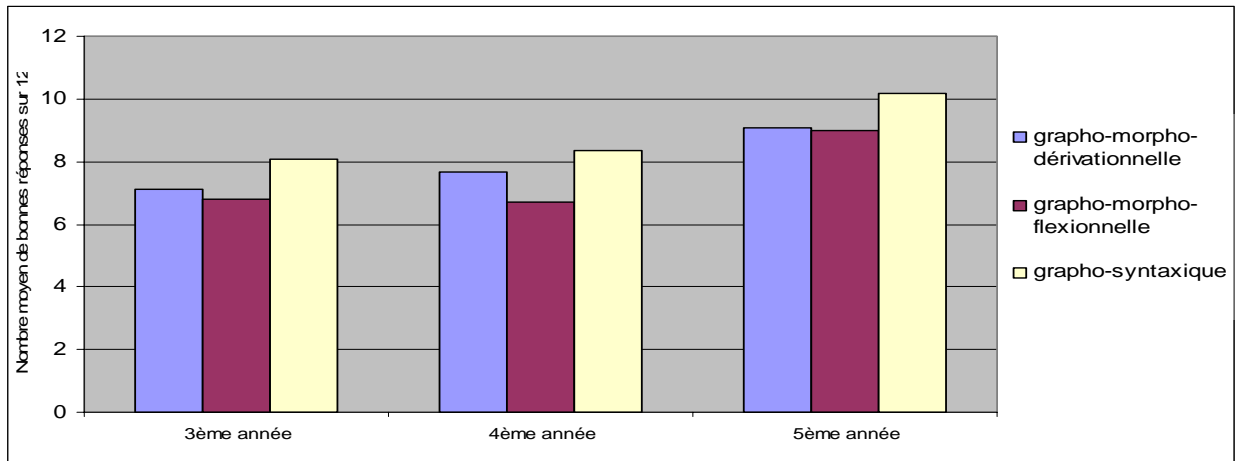


Figure 8: Performances moyennes obtenues aux épreuves explicites de détection d'intrus grapho-sémantique en 3^e, 4^e et 5^e année.

A l'issue de ces analyses, nous pouvons constater que les connaissances grapho-morphologiques (dérivationnelles et flexionnelles) évaluées par les épreuves explicites de détection d'intrus sont encore assez faibles en 3^e et 4^e année (résultats moyens inférieurs à 65% de réussite) mais augmentent en 5^e année (moyenne de 75% de réussite). De plus, que la détection d'intrus est plus facile lorsqu'elle porte sur la structure morpho-dérivationnelle que morpho-flexionnelle des mots arabes. L'épreuve de détection explicite d'intrus grapho-syntaxique apparaît encore plus facile. Elle serait même très bien résolue par une majorité d'élèves de 5^e année (moyenne de 85 % de réussite).

2.2.3.3 *Epreuve implicite de détection d'intrus graphophonologique : effets du niveau scolaire et de l'unité phonologique*

Les résultats obtenus à l'épreuve de détection implicite d'intrus graphophonologique en fonction du niveau scolaire et de l'unité phonologique ainsi que les comparaisons au hasard figurent dans le tableau 11.

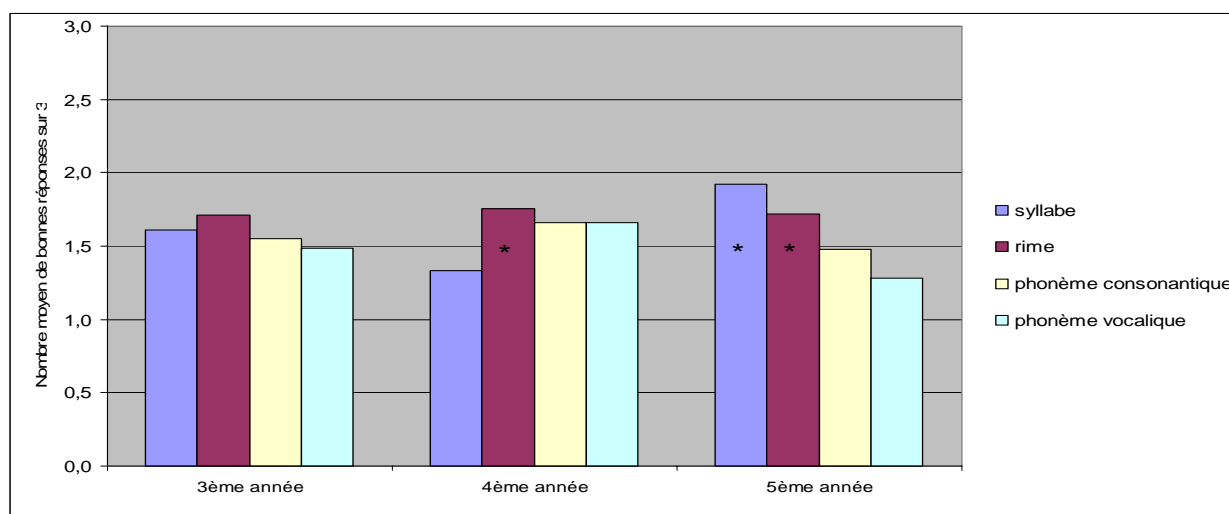
Tableau 11 : Performances moyennes de réussite (et écart-types) obtenues à l'épreuve implicite de détection d'intrus graphophonologique en fonction du type d'items et du niveau scolaire et tests t de comparaison à la valeur 1/2 (niveau du hasard).

	3 ^e année N = 56	4 ^e année N = 41	5 ^e année N = 50	Tous niveaux scolaires N = 147
Score Syllabe * (sur 3)	1,61 (0,83) t(55) < 1	1,34 (1,00) t(40) = -1,05 ns	1,92 (0,80) t(49) = 3,69 p < .001	1,64 (0,90) t(146) = 1,87 ns
Score Rime (sur 3)	1,71 (0,91) t(55) = 1,76 ns	1,76 (0,73) t(40) = 2,23 p < .05	1,72 (0,76) t(49) = 2,05 p < .05	1,73 (0,81) t(146) = 3,42 p < .001
Score Phonème consonantique (sur 3)	1,55 (0,63) t(55) < 1	1,66 (0,85) t(40) = 1,19 ns	1,48 (0,74) t(49) > -1	1,56 (0,73) t(146) < 1
Score Phonème vocalique (sur 3)	1,48 (0,89) t(55) > -1	1,66 (0,96) t(40) = 1,05 ns	1,28 (0,90) t(49) = -1,72 ns	1,46 (0,92) t(146) > -1
Score total (sur 12)	6,36 (1,84) t(55) = 1,46 ns	6,41 (2,17) t(40) < 1	6,40 (2,18) t(49) = 1,59 ns	6,39 (2,04)

* 4 items portent normalement sur la syllabe mais pour ce sous-groupe d'items, les scores ont été rapportés sur 3 pour pouvoir être comparés aux autres.

L'analyse du facteur « niveau scolaire » ne fait apparaître aucune différence significative entre les scores globaux des trois groupes à l'épreuve de détection implicite d'intrus graphophonologique [$F(1,144) < 1$]. Les comparaisons à la valeur 1,5 (50 %) montrent que les résultats des trois niveaux scolaires peuvent être liés au hasard (cf. tableau 11). Malgré ces faiblesses, il semble toutefois intéressant d'approfondir les analyses dans l'objectif de déterminer les unités phonologiques auxquelles les enfants arabophones sont les plus sensibles.

Les résultats obtenus aux items « syllabe », « rime », « phonème consonantique » et « phonème vocalique » peuvent être visualisés à l'aide de la figure 9.



* Scores supérieurs au hasard

Figure 9 : Performances moyennes (sur 3) obtenues à l'épreuve implicite de détection d'intrus graphophonologique en 3^e, 4^e et 5^e année en fonction de l'unité phonologique étudiée

Les résultats au test de Friedman révèlent un effet significatif de l'unité phonologique sur l'ensemble des groupes [$\chi^2(3, N = 147) = 8,55 ; p < .05$]. Les comparaisons deux à deux (test de Wilcoxon) permettent d'identifier que les résultats aux items portant sur la rime sont supérieurs à ceux portant sur le phonème consonantique ($T = 1370, z = 2,00, p < .05$) et vocalique ($T = 1563,5, z = 2,52, p < .05$). Les autres comparaisons ne font pas apparaître de différences significatives. Il est également à souligner que les seuls résultats significativement supérieurs au niveau du hasard sont ceux obtenus aux items « rime » ($p < .001$), les résultats aux items « syllabe » ne s'en distinguent pour leur part que tendanciellement ($p < .07$) (cf. tableau 11).

En 3^e année et en 4^e année, l'effet de l'unité phonologique n'est pas significatif [$\chi^2(3, N = 56) = 1,81, ns$ et $\chi^2(3, N = 41) = 3,29, ns$]. Les résultats obtenus en 3^e année, pour chaque groupe d'items pris séparément, ne sont d'ailleurs pas significativement différents du niveau du hasard. On note cependant une supériorité tendancielle des résultats aux items « rime ». En 4^e année, les scores obtenus aux items « rime » se distinguent significativement du hasard ($p < .05$) mais pas les autres (cf. tableau 11).

En 5^e année, l'effet de l'unité phonologique est significatif [$\chi^2(3, N = 50) = 23,99 p < .001$]. Les comparaisons deux à deux permettent d'identifier que les résultats aux items « syllabe » sont supérieurs à ceux aux items « phonème consonantique » ($T = 263,5, z = 3,18, p < .01$) et « phonème vocalique » ($T = 183, z = 3,91, p < .0001$). De même, que les items « rime » sont mieux résolus que les items « phonème vocalique » ($T = 137, z = 2,56, p < .05$). Les autres comparaisons ne font pas apparaître de différences significatives. Il est également à souligner que les seuls résultats supérieurs au niveau du hasard sont ceux obtenus aux items « syllabe » et « rime » ($p < .01$ et $p < .05 ; cf. tableau 11$).

Ainsi, ces analyses permettent de constater qu'en 3^e année, les enfants arabophones ne semblent clairement sensibles à aucune des unités phonologiques étudiées par le biais de l'épreuve de détection implicite d'intrus graphophonologique. Les résultats obtenus à l'ensemble de cette épreuve seront donc exclus des analyses ultérieures pour ce niveau. Il semble qu'en 4^e année, une sensibilité à la rime apparaisse avec un taux moyen de 59% de bonnes réponses à ces items. Cependant ces scores ne respectent pas le critère de distribution normale (Test de Kormogorov-Smirnov significatif : $d = 0,26, p < .01$) si bien qu'ils seront exclus des analyses corrélationnelles, factorielles et de régression ultérieures comme ceux sur les autres unités phonologiques ne se distinguant pas du hasard. En 5^e année, c'est à la syllabe et à la rime que les enfants sont les plus sensibles. De surcroît, les corrélations de rang de Spearman permettent

d'observer que c'est entre ces deux scores que les corrélations sont les plus élevées¹³⁵. Un score composite de 4,28/7 ($\sigma = 1,55$), englobant les items « syllabe et rime » peut ainsi être calculé pour ce niveau scolaire et être inclus dans les analyses ultérieures car il respecte une distribution normale (Test de Kormogorov-Smirnov non significatif : $d = 0,14$, $p > .20$).

2.2.4 Epreuve d'identification de la construction morphologique de pseudo-mots : effet du niveau scolaire, de la nature de la construction morphologique et de la nature des règles morphologiques

Le tableau 12 présente les résultats obtenus à l'épreuve d'identification de la construction morphologique de pseudo-mots (moyennes, écart-types, minima et maxima) pour chaque niveau scolaire. Pour cette épreuve, les réponses sont à choisir entre trois propositions ce qui rend là encore possible les choix aléatoires. Cependant, les scores s'élevant généralement bien au dessus de 1/3 de bonnes réponses (8/24), les tests t pour échantillon unique ne sont précisés que lorsque les résultats sont faibles pour vérifier si ces réponses sont effectivement liées au hasard.

Tableau 12 : Performances moyennes sur 24, écart-types, maxima, minima obtenus à l'épreuve d'identification de la construction morphologique de pseudo-mots en fonction du niveau scolaire.

	Moyenne	Ecart-type	Minimum	Maximum
3 ^e année N = 56	9,84	3,40	5	21
4 ^e année N = 41	13,59	5,27	4	22
5 ^e année N = 50	15,46	5,82	5	24

Les items de cette épreuve se distinguant quant à la nature de la construction des pseudo-mots et des règles morphologiques étudiées, il a été mené une analyse de variance à trois facteurs (niveau scolaire {3^e année, 4^e année, 5^e année}, construction morphologique {préfixe, suffixe, schème} et règle morphologique {morpho-dérivationnelle vs. morpho-flexionnelle}). Les résultats en fonction de la nature de la construction morphologique et du niveau scolaire figurent dans le tableau 13, ceux en fonction de la règle morphologique et du niveau scolaire dans le tableau 14¹³⁶.

¹³⁵ Coefficients de corrélation entre les scores rime/syllabe : $r_s = .40$, ; $p < .05$, syllabe/consonne et syllabe/voyelle : $r_s = .34$, ; $p < .05$, rime/consonne : $r_s = .22$; ns, rime/voyelle : $r_s = -.03$; ns et voyelle /consonne : $r_s = .26$, ; ns.

¹³⁶ Les interactions « construction morphologique » * « règle morphologique » et « construction morphologique » * « règle morphologique » * « niveau scolaire » n'étant pas significatives (*cf. infra*), nous avons préféré, pour ne pas alourdir la lecture du paragraphe, ne pas présenter les résultats distinguant simultanément ces 3 facteurs.

Tableau 13 : Performances moyennes sur 8 (écart-types) et en pourcentage, obtenues à l'épreuve d'identification de la construction morphologique de pseudo-mots en fonction du niveau scolaire et de la nature de la construction morphologique des pseudo-mots.

Construction morphologique	3 ^e année N = 56	4 ^e année N = 41	5 ^e année N = 50	Tous niveaux scolaires N = 147
"préfixe"	3,64 (1,68) 46%	4,76 (1,88) 59%	5,52 (2,05) 69%	4,59 (2,02) 57%
"suffixe"	3,39 (1,53) 42%	4,78 (1,78) 60%	5,36 (2,01) 67%	4,45 (1,96) 56%
"schème"	2,80 (1,48) 35% ^H	4,05 (2,24) 51%	4,58 (2,33) 57%	3,76 (2,15) 47%

Tableau 14 : Performances moyennes sur 12 (écart-types) et en pourcentage, obtenues à l'épreuve d'identification de la construction morphologique de pseudo-mots en fonction du niveau scolaire et de la règle morphologique impliquée.

Règle morphologique	3 ^e année N = 56	4 ^e année N = 41	5 ^e année N = 50	Tous niveaux scolaires N = 147
Morpho-dérivationnelle	4,89 (2,00) 41%	6,73 (2,85) 56%	8,24 (2,86) 69%	6,54 (2,92) 55%
Morpho-flexionnelle	4,95 (2,04) 41%	6,85 (2,87) 57%	7,22 (3,46) 60%	6,25 (2,99) 52%

Les résultats mettent en évidence :

- Un effet principal du niveau scolaire [$F(2, 144) = 18,39, p < .00001$] : Les 3^e année obtiennent en moyenne des résultats plus faibles que les 4^e année ($p < .001$) et les 5^e année ($p < .0001$). La différence entre les 4^e et 5^e année n'est pas significative (*cf.* tableau 12).
- Un effet principal « construction morphologique » [$F(2,288) = 20,29, p < .00001$] mais pas d'interaction « niveau scolaire » * « construction morphologique » [$F(4,288) < 1$] : Pour tous les niveaux scolaires, les items « schème » sont en moyenne moins bien résolus que les items « préfixe » ($p < .0001$) et « suffixe » ($p < .00001$), la différence entre ces deux derniers types n'étant pas significative (*cf.* tableau 13). Cependant, en 3^e année, les réponses aux items « schème » ne se distinguent pas du niveau du hasard [$t(55) = 0,73, ns$].
- L'absence de l'effet « règle morphologique » [$F(1,144) = 2,14, ns$] mais une interaction « niveau scolaire » * « règle morphologique » [$F(2,144) = 3,75, p < .05$] : en 5^e année, les items portant sur les règles morpho-dérivationnelles sont en moyenne mieux résolus que ceux portant sur les règles morpho-flexionnelles ($p < .003$) alors que cette différence n'est pas significative sur les deux autres niveaux scolaires (*cf.* tableau 14).
- L'absence d'interaction « construction morphologique » * « règle morphologique » [$F(2,288) < 1$] et d'interaction « construction morphologique » * « règle morphologique » * « niveau scolaire » [$F(4,288) = 1,98 ; ns$].

A l'issue de cette analyse, il est possible de relever que :

- 1) les performances à cette épreuve augmentent surtout entre la 3^e et la 4^e année,
- 2) la construction morphologique des pseudo-mots préfixés et suffixés est mieux identifiée que celle correspondant à une structure de type schème,
- 3) qu'en 5^e année seulement, la construction morphologique de pseudo-mots est mieux identifiée lorsqu'elle porte sur des aspects morpho-dérivationnels que morpho-flexionnels.

2.2.5 Synthèse des résultats et récapitulatif des épreuves ou groupes d'items sélectionnés pour les analyses suivantes

A l'issue des analyses précédentes, certaines épreuves implicites ou explicites, évaluant les connaissances graphophonologiques, grapho-morpho-dérivationnelles, grapho-morpho-flexionnelles et grapho-syntaxiques ont été éliminées et d'autres conservées. Dans un cas, seuls certains items ont été sélectionnés, et en 5^e année seulement.

Les résultats pris en compte dans les analyses corrélatives, factorielles et de régression ne seront donc pas exactement identiques d'un niveau scolaire à l'autre. Le tableau 15 permet d'avoir un aperçu synoptique des résultats obtenus et des épreuves ou groupes d'items inclus dans les analyses qui vont suivre¹³⁷.

¹³⁷ Ces épreuves ou groupes d'items sont mis en évidence dans le tableau par une police en gras.

Tableau 15 : Récapitulatif des épreuves ou groupes d'items sélectionnés pour les analyses suivantes à chaque niveau scolaire et synthèse des principaux résultats obtenus (en %)

Epreuves	3 ^e année	4 ^e année	5 ^e année
Det. Imp. G-phonologique	Supprimée		Score rime + syllabe 61,14 %
Det. Imp. G-morpho-dériv.	Supprimée (réponses aléatoires)		
Det. Imp. G-morpho-flex	Supprimée (réponses aléatoires)		
Det. Imp. G-syntaxique	Supprimée (réponses aléatoires)		
Det. Exp. G-phonologique	Supprimée (scores plafonds)		
Det. Exp. G-syntaxique	67,16 %	69,58%	Supprimée (scores plafonds)
Det. Exp. G-morpho-dériv.	59,25 %	63,92 %	75,83 %
Det. Exp. G-morpho-flex.	56,67 %	55,83 %	74,83 %
Ident. Const. morpho. PM	41 %	56,63 %	64,5 % « règle morpho-dériv » > « règle morpho-flex. »
	"préfixe" et "suffixe" > "schème"		

2.3 Analyse du lien entre les scores aux différentes épreuves ou groupes d'items sélectionnés

En 3^e année et en 4^e année, les épreuves retenues pour analyser le lien entre les différents types de connaissances sont les épreuves explicites de détection d'intrus grapho-morpho-dérivationnel, grapho-morpho-flexionnel, grapho-syntaxique et l'épreuve d'identification de la construction morphologique de pseudo-mots. En 5^e année, ce sont les items de l'épreuve implicite de détection d'intrus phonologique portant sur la rime et la syllabe, les épreuves explicites de détection d'intrus grapho-morpho-dérivationnel et grapho-morpho-flexionnel ainsi que l'épreuve d'identification de la construction morphologique de pseudo-mots.

2.3.1 Analyses de corrélation

Le tableau 16 précise les coefficients de corrélation de Bravais Pearson entre les scores à ces différentes épreuves pour chaque niveau scolaire.

Tableau 16 : Coefficients de corrélation de Bravais Pearson entre les épreuves ou groupe d'items sélectionnés en 3^e, 4^e et 5^e année.

3 ^e année N = 51	Dét. exp. G-morpho-dériv.	Dét. Exp. G-morpho-flex.	Dét. Exp. G-syntaxique	Ident. Const. morpho.
Dét. Exp. G-morpho-dériv.	----	.21	.37**	.44**
Dét. Exp. G-morpho-flex.		----	.48**	.30*
Dét. Exp. G- syntaxique.			----	.24

4 ^e année N = 40	Dét. exp. G-morpho-dériv.	Dét. Exp. G-morpho-flex.	Dét. Exp. G-syntaxique	Ident. Const. morpho.
Dét. exp. G-morpho-dériv.	----	.25	.40*	.29
Dét. Exp. G-morpho-flex.		----	.39*	.43**
Dét exp G- syntaxique			----	.39*

5 ^e année N = 50	Dét. Imp. G-phonologique	Dét. exp. G-morpho-dériv.	Dét. Exp. G-morpho-flex.	Ident. Const. morpho.
Dét. Imp. G-phonologique	----	-.11	-.15	.26
Dét. Exp. G-morpho-dériv.		----	.25	.13
Dét. Exp. G- morpho-flex.			----	.51**

** r significatif à $p < .01$ * à $p < .05$

En 3^e année, la matrice de corrélations (*cf.* tableau 16) permet de constater que les performances aux épreuves de détection explicite d'intrus grapho-morpho-flexionnel et grapho-syntaxique sont corrélées ($r = .48$, $p < .01$). On note également une corrélation entre les scores à l'épreuve d'identification de la construction morphologique de pseudo-mots et à l'épreuve de détection explicite d'intrus grapho-morpho-dérivationnelle ($r = .44$, $p < .01$), ces derniers étant également corrélés aux performances à l'épreuve grapho-syntaxique. ($r = .37$, $p < .05$). Enfin, la corrélation entre les scores à l'épreuve d'identification de la construction morphologique de pseudo-mots et à l'épreuve de détection grapho-morpho-flexionnelle est significative ($r = .30$, $p < .05$).

En 4^e année, les performances aux épreuves de détection explicite d'intrus grapho-morpho-flexionnel et d'identification de la construction morphologique de pseudo mots sont les plus corrélées ($r = .43$, $p < .01$). On note également une corrélation entre les scores aux épreuves de détection explicite d'intrus grapho-morpho-flexionnel et grapho-syntaxique ($r = .39$, $p < .05$), ces derniers étant également corrélés aux performances à l'épreuve grapho-morpho-dérivationnelle.

($r = .40$, $p < .05$) et aux scores à l'épreuve d'identification de la construction morphologique de pseudo-mots ($r = .39$, $p < .05$) (cf. tableau 16).

En 5^e année, seules les performances aux épreuves de détection explicite d'intrus grapho-morpho-flexionnel et d'identification de la construction morphologique de pseudo mots sont corrélées significativement ($r = .51$, $p < .01$).

Ces résultats mettent en évidence les points suivants :

En 3^e année, les connaissances morphologiques (dérivationnelles et flexionnelles) seraient liées aux connaissances syntaxiques (mesurées toutes deux à l'aide du même paradigme) mais aussi aux compétences d'identification de la construction morphologique de pseudo-mots.

Les connaissances morphologiques sont toujours en 4^e année liées aux connaissances syntaxiques (évaluées par le même paradigme). Cependant le lien établi entre l'identification de la construction morphologique sur des pseudo-mots et les autres mesures change par rapport au niveau scolaire précédent. Les résultats à cette épreuve seraient en 4^e année associés à la dimension morphosyntaxique des connaissances morphologiques (mais plus à la dimension morfo-dérivationnelle) et aux connaissances grapho-syntaxiques.

En 5^e année, l'épreuve évaluant les connaissances grapho-syntaxiques ayant été éliminée, le lien établi aux niveaux scolaires précédents entre ce type de connaissance et les connaissances morphologiques n'a pas pu être répliqué. Pour ce groupe, l'association entre les compétences à identifier la structure morphologique de pseudo-mots et les connaissances grapho-morpho-flexionnelles mesurées par le paradigme de détection d'intrus, déjà observé aux niveaux précédents, est particulièrement nette.

Ainsi, en 3^e année ce sont en priorité les élèves qui discriminent le mieux la structure morfo-dérivationnelle des mots arabes qui parviennent le mieux à s'abstraire des spécificités de l'arabe pour identifier la construction morphologique de pseudo mots alors qu'en 4^e et encore plus spécifiquement en 5^e année, ce sont les élèves dont les connaissances morphosyntaxiques en arabe sont les plus développées.

2.3.2 Analyses en composantes principales

Des analyses en composantes principales ont également été menées sur les épreuves ou groupe d'épreuves sélectionnées à chaque niveau scolaire. Le tableau 17 présente les résultats obtenus.

Tableau 17 : Contribution des performances aux différentes épreuves ou groupes d'items sélectionnés aux facteurs extraits des analyses en composantes principales pour chaque niveau scolaire

	3 ^e année	4 ^e année	5 ^e année	
	Facteur 1	Facteur 1	Facteur 1	Facteur 2
Dét. Imp. G-phonologique			-.05	.90*
Dét. Exp. G-morpho-dériv.	-.71*	-.65	-.50	-.44
Dét. Exp. G-morpho-flex.	-.70*	-.72*	-.85*	-.19
Dét Exp G-syntaxique	-.75*	-.77*		
Ident. Const. morpho.	-.69	-.74*	-.81*	.41
% de variance totale expliquée	50,60	52,07	40,64	30,14

* poids factoriel significatif (> .70)

En 3^e année, cette analyse permet de mettre en évidence un facteur, dont la contribution à la variance s'élève à 50,60 %, composé des épreuves de détection explicites d'intrus grapho-syntaxique (poids de -.75), grapho-morpho-dérivationnel (poids de -.71) et grapho-morpho-flexionnel (poids de -.70).

En 4^e année, l'ACP permet de mettre en évidence un facteur, dont la contribution à la variance s'élève à 52,07 %, composé des épreuves de détection explicite d'intrus grapho-syntaxique (poids de -.77), d'identification de la construction morphologique de pseudo-mots (poids de -.74) et de détection explicite d'intrus grapho-morpho-flexionnel (poids de -.72).

En 5^e année, deux facteurs se dégagent de l'ACP : le premier expliquant 40,64 % de la variance est composé des épreuves de détection explicite d'intrus grapho-morpho-flexionnel (poids de -.85), et d'identification de la construction morphologique de pseudo-mots (poids de -.81). Le deuxième, expliquant 30,14 % de la variance, est exclusivement composé de l'épreuve implicite de détection d'intrus graphophonologique (poids de .90).

En 3^e année, le facteur dégagé regroupe les trois épreuves de détection explicite grapho-sémantiques. De plus, l'épreuve apparaissant avoir le plus de poids dans la construction de ce facteur est l'épreuve de détection d'intrus grapho-syntaxique : les connaissances syntaxiques mesurées par le paradigme de détection d'intrus seraient en 3^e année un bon indicateur des compétences grapho-sémantiques. La participation de l'épreuve d'identification de la construction morphologique de pseudo-mots n'atteint pas à ce niveau le seuil de significativité. Toutefois le taux de réussite à cette épreuve est associé à celui aux deux épreuves de détection d'intrus grapho-morphologique, ce qui laisse supposer que cette mesure est un bon indicateur des

compétences morphologiques des élèves de 3^e année. **Le facteur dégagé à ce niveau scolaire pourrait ainsi être assimilé à une composante de discrimination appliquée aux aspects du code grapho-sémantique spécifiques à la langue arabe.**

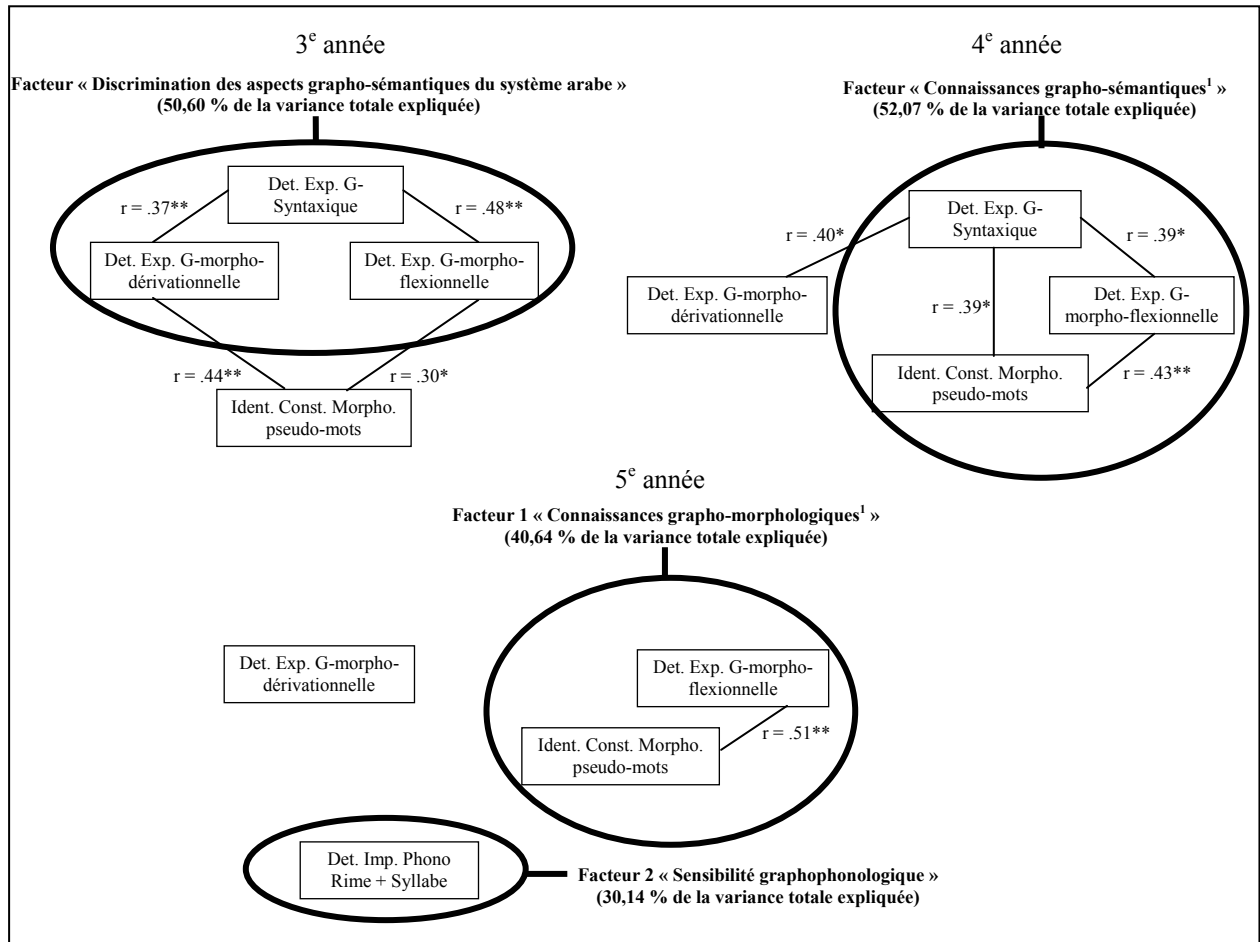
En 4^e année, le facteur dégagé regroupe les épreuves grapho-syntaxique et grapho-morpho-flexionnelle de détection explicite d'intrus et l'épreuve d'identification de la construction morphologique de pseudo-mots. L'épreuve apparaissant avoir le plus de poids dans la construction de ce facteur est toujours l'épreuve de détection d'intrus grapho-syntaxique. L'implication des connaissances grapho-morpho-dérivationnelle n'est plus significative. Cependant, ces connaissances ne sont pas totalement dissociées des autres connaissances grapho-sémantiques puisqu'elles sont notamment liées aux connaissances grapho-syntaxiques. De plus, ce qui apparaît nouveau par rapport au niveau précédent est la contribution significative de l'épreuve d'identification de la construction morphologique de pseudo-mots. **Ainsi, en 4^e année, le facteur extrait pourrait être assimilé à certaines connaissances grapho-sémantiques.**

On retrouve en 5^e année, un profil relativement proche dans la constitution du premier facteur bien que l'épreuve explicite de détection d'intrus grapho-syntaxique n'ait pas été intégrée dans l'analyse : ce sont là encore, les épreuves de détection explicite d'intrus grapho-morpho-flexionnel et d'identification de la construction morphologique de pseudo-mots qui interviennent avec un poids factoriel élevé dans la constitution de ce premier **facteur qui pourrait être assimilé à certaines connaissances grapho-morphologiques.** Comme en 4^e année, l'épreuve grapho-morpho-dérivationnelle n'intervient plus dans sa composition.

De plus, un second facteur se dégage pour ce niveau scolaire, constitué des items de l'épreuve graphophonologique implicite de détection d'intrus rimique et syllabique. La participation exclusive de cette sensibilité aux aspects phonologiques à la formation d'un deuxième axe établit le pouvoir discriminant secondaire et indépendant de certaines connaissances implicites graphophonologiques entre les élèves de 5^e année.

2.3.3 Synthèse

Le schéma 1 permet d'avoir un aperçu synoptique des principales données statistiques décrites précédemment.



* r significatif à $p < .05$; ** à $p < .01$

¹ A l'exclusion des connaissances grapho-morpho-dérivationnelles

Schéma 1 : Représentation des liens corrélationnels entre les épreuves ou groupe d'items sélectionnés et des facteurs dégagés de l'ACP en 3^e, 4^e et 5^e année.

2.4 Recherche des performances déterminant le niveau de lecture

2.4.1 Analyses sur le niveau de lecture de mots

2.4.1.1 Analyses de corrélations entre le niveau de lecture de mots et les scores aux différentes épreuves sélectionnées

Le tableau 18 présente les coefficients de corrélation entre le niveau de lecture de mots et les scores aux différentes épreuves ou groupe d'items sélectionnés, en 3^e, 4^e et 5^e année.

Tableau 18 : Coefficients de corrélation entre les scores aux épreuves ou groupes d'items sélectionnés et le niveau de lecture de mots en 3^e, 4^e et 5^e année.

	Dét.imp. G- phonologique (syllabe et rime)	Dét.exp. G- morpho-dériv.	Dét. exp. G- morpho-flex.	Dét. Exp. G- syntaxique	Ident. Const. morpho. PM
3 ^e année N = 51	X	.08	.25	.29*	.39**
4 ^e année N = 39	X	.40*	.54**	.49**	.53**
5 ^e année N = 50	.08	.28*	.26	X	.40**

** r significatif à $p < .01$ * à $p < .05$

En 3^e année, nous pouvons observer que les scores à l'épreuve d'identification de la structure morphologique, qui sont encore faibles à ce niveau scolaire, sont les plus corrélés aux performances en lecture de mots ($r = .39$, $p < .01$). Les scores à l'épreuve de détection grapho-syntaxique y sont également corrélés significativement ($r = .29$, $p < .05$). Les autres corrélations ne sont pas significatives.

En 4^e année, ce sont les scores aux épreuves de détection grapho-morphosyntaxique et d'identification de la structure morphologique qui sont les plus corrélées aux performances en lecture de mots ($r = .54$, $p < .001$ et $r = .53$, $p < .001$). Les scores aux épreuves de détection grapho-syntaxique et grapho-morpho-dérivationnelle présentent également un coefficient de corrélation significatif avec le niveau de lecture de mots ($r = .49$, $p < .01$ et $r = .40$, $p < .05$).

En 5^e année, les scores les plus corrélés aux performances en lecture de mots sont ceux obtenus à l'épreuve d'identification de la structure morphologique ($r = .40$, $p < .01$). Les scores à l'épreuve de détection grapho-morpho-dérivationnelle présentent un coefficient de corrélation significatif avec les performances en lecture de mots ($r = .28$, $p < .05$). Les autres corrélations ne sont pas significatives.

Ainsi, le niveau de lecture de mots est associé aux compétences à identifier la construction morphologique de pseudo-mots pour les trois niveaux scolaires. Le lien entre la lecture de mots et les connaissances grapho-syntaxiques est visible pour les niveaux où l'épreuve les

mesurant a été incluse dans les analyses (3^e et 4^e année). Celui entre la lecture de mots et les connaissances morpho-dérivationnelles n'apparaît qu'à partir de la 4^e année.

Plus précisément, la lecture de mots est surtout associée, en 3^e année, aux compétences à identifier la structure morphologique de pseudo-mots et aux connaissances explicites grapho-syntaxiques en arabe, qui contribuent le plus fortement à la constitution du facteur « discrimination des aspects grapho-sémantiques du code arabe » (cf. ACP supra).

En 4^e année, la lecture de mots serait surtout liée aux connaissances grapho-morpho-flexionnelles et aux compétences à identifier la structure morphologique de pseudo-mots mais également aux connaissances grapho-syntaxiques et grapho-morpho-dérivationnelles. Le facteur grapho-sémantique identifié à ce niveau scolaire (cf. ACP supra) jouerait donc un rôle important dans la reconnaissance des mots arabes.

Les connaissances les plus associées aux performances en lecture sont en 5^e année les connaissances grapho-morphologiques appliquées aux pseudo-mots (constituant avec les connaissances grapho-morpho-flexionnelles le facteur « connaissances grapho-sémantiques) mais également les connaissances grapho-morpho-dérivationnelles. En revanche, le deuxième facteur extrait de l'ACP ne fait apparaître aucun lien avec la lecture : les connaissances graphophonologiques implicites, telles qu'elles ont été évaluées dans cette étude, ne participeraient pas à la lecture de mots en arabe.

2.4.1.2 *Analyses de régressions sur la variable « lecture de mots »*

Afin de déterminer dans quelle mesure les connaissances grapho-sémantiques prédisent de façon indépendante le niveau de lecture de mots, à chaque niveau scolaire, des analyses de régression pas à pas ont été conduites.

En 3^e année, si l'on entre dans le modèle les deux variables présentant des corrélations significatives avec le niveau de lecture de mots (identification de la construction morphologique de pseudo-mots et détection explicite grapho-syntaxique), on observe que la seule variable dont la contribution est significative est celle mesurée par l'épreuve d'identification de la construction morphologique, expliquant 15,27% de la variance en lecture (cf. tableau 19). La contribution de la variable « détection explicite grapho-syntaxique » entrée dans un second temps est si faible (4,23%) qu'elle n'est pas significative ($\beta = .21$, ns).

Tableau 19 : Contribution des variables entrées dans le modèle de régression sur le niveau de lecture de mots en 3^e année

	R	R ²	R ² modifié	F	p
Pas 1 : Identification const. Morpho. Pseudo-mots	.39	.1527	.1527	8,83	<.01
Pas 2 : Det. Exp. G-Syntaxique	.44	.1950	.0423	2,52	ns

En 4^e année, si l'on entre dans le modèle les quatre variables corrélées au niveau de lecture de mots (détection explicite grapho-morpho-flexionnelle, identification de la construction morphologique de pseudo-mots, détection explicite grapho-syntaxique et grapho-morpho-dérivationnelle), on observe que les variables contribuant le plus à expliquer niveau lecture de mots sont celles mesurées par les épreuves grapho-morpho-flexionnelle et d'identification de la construction morphologique de pseudo-mots, qui expliquent ensemble 40,91% de la variance. La variable « détection explicite grapho-morpho-flexionnelle » apporte la contribution la plus importante (29,51%). La variable « identification de la construction morphologique », incluse dans un deuxième temps ajoute 11,40 % de variance expliquée (*cf.* tableau 20). Les contributions ajoutées des variables grapho-syntaxique (5,38 %) et grapho-morpho-dérivationnelle (2,05%) sont trop faibles pour être significatives ($\beta = .25$ et $\beta = .16$, ns).

Tableau 20 : Contribution des variables entrées dans le modèle de régression sur le niveau de lecture de mots en 4^e année

	R	R ²	R ² modifié	F	p
Pas 1 : Dét Exp G-morpho-flexionnelle	.54	.2951	.2951	15,49	<.001
Pas 2 : Identification construction morphologique	.64	.4091	.1140	6,94	<.05
Pas 3 : Det. Exp. G-syntaxique	.68	.4629	.0538	3,51	ns
Pas 4 : Det. Exp. G-morpho-dérivationnelle	.70	.4834	.0205	1,35	ns

En 5^e année, les variables « identification de la construction morphologique de pseudo-mots » et « détection explicite grapho-morpho-dérivationnelle », présentant des coefficients de corrélations significatifs avec les scores en lecture de mots sont entrées dans le modèle de régression pas à pas (*cf.* tableau 21).

Tableau 21 : Contribution des variables entrées dans le modèle de régression sur le niveau de lecture de mots en 5^e année

	R	R ²	R ² modifié	F	p
Pas 1 : Identification Const. Morpho. pseudo-mots	.40	.1579	.1579	7,00	<.01
Pas 2 : Det. Exp. G-morpho-dérivationnelle	.46	.2119	.0540	3,22	ns

Ce dernier met en évidence que seule la variable « identification de la construction morphologique » permet d'expliquer de façon significative 15,79 % la variance en lecture de mots à ce niveau scolaire. Les 5,40 % de part ajoutée par la variable grapho-morpho-dérivationnelle, entrée dans un second temps n'est pas significatif ($\beta = .23$; ns).

Nous retiendrons de ces analyses la contribution des compétences d'identification de la construction morphologique de pseudo-mots à la lecture de mots, pour les trois niveaux scolaires. Cette capacité à identifier les aspects morphologiques indépendamment des connaissances lexicales en arabe est, en 3^e et 5^e année, la seule contribuant à expliquer une part non négligeable (entre 15,27 et 15,79 %) de la variance en lecture de mots. En 4^e année, sa contribution explique 11,40 % de la variance et se dégage après celle de la variable « détection explicite grapho-morpho-flexionnelle » (mesurant les connaissances permettant d'analyser consciemment les marques flexionnelles portées par les verbes et les noms féminins) entrée en premier pas (29,51 % de variance expliquée). Ces deux variables, expliquant ensemble 40,91 % de la variance en lecture de mots constituent deux des trois dimensions du facteur « connaissances grapho-sémantiques » identifié à ce niveau scolaire.

2.4.2 Analyses sur le niveau de compréhension de phrases écrites

2.4.2.1 Analyse de corrélations entre le niveau de compréhension de phrases écrites et les scores aux différentes épreuves sélectionnées et en lecture de mots.

Le tableau 22 présente les coefficients de corrélation entre le niveau de compréhension de phrases écrites et les scores aux différentes épreuves sélectionnées ainsi qu'au test de lecture de lecture de mots, en 3^e, 4^e et 5^e année.

Tableau 22 : Coefficients de corrélation entre les différents scores et le niveau de compréhension de phrases en 3^e, 4^e et 5^e année.

:

	Dét.imp. G- phonologique (rime et syllabe)	Dét.exp. G- morpho-dériv.	Dét. exp. G- morpho-flex.	Dét. Exp. G- syntaxique	Ident. Const. morpho. PM	Lecture de mots
3 ^e année N = 44	X	.32*	.29	.54**	.10	.41**
4 ^e année N = 38	X	.39*	.01	.18	.04	.32*
5 ^e année N = 50	.12	.08	.36*	X	.49**	.40**

** r significatif à $p < .01$ * à $p < .05$

En 3^e année, nous pouvons observer que les scores au test de compréhension de phrases sont significativement corrélés à ceux obtenus à l'épreuve de détection explicite d'intrus grapho-syntaxique ($r = .55, p < .0001$), à l'épreuve de lecture de mots ($r = .41, p < .01$) et à l'épreuve de détection explicite d'intrus grapho-morpho-dérivationnelle ($r = .32, p < .05$), les autres corrélations n'étant pas significatives.

En 4^e année, ce sont les scores aux épreuves de détection grapho-morpho-dérivationnelle et en lecture de mots qui sont significativement corrélées aux performances en compréhension de phrases ($r = .39, p < .0001$ et $r = .32, p < .0001$). En revanche, aucune corrélation n'apparaît avec les performances aux épreuves d'identification de la construction morphologique de pseudo-mots et de détection explicite d'intrus grapho-morpho-flexionnel ($r = .04$ et $r = .01, ns$).

En 5^e année, les scores les plus corrélés aux performances en compréhension de phrases sont ceux obtenus à l'épreuve d'identification de la construction morphologique de pseudo-mots ($r = .49, p < .0001$). Le niveau de compréhension de phrases est également corrélé significativement au niveau de lecture de mots ($r = .40, p < .005$) et aux scores obtenus à l'épreuve explicite de détection grapho-morpho-flexionnelle ($r = .36, p < .02$), les autres corrélations n'étant pas significatives.

Pour les trois niveaux scolaires, la compréhension de phrases est associée à la lecture de mots. En 3^e année, il l'est également avec deux des trois dimensions du facteur « discrimination des aspects grapho-sémantiques en langue arabe » : les connaissances grapho-syntaxiques et grapho-morpho-dérivationnelles explicites. En 4^e année, en plus de la lecture de mots, seules les connaissances grapho-morpho-dérivationnelles explicites sont, comme au niveau précédent, liées à la compréhension de phrases. Cette mesure n'est pourtant pas en 4^e année incluse dans le facteur « connaissances grapho-sémantiques » identifié. L'association entre la compréhension de phrases et les connaissances grapho-syntaxiques, observée en 3^e année, n'est plus apparente.

En 5^e année, la compréhension de phrases serait surtout liée aux connaissances grapho-morpho-flexionnelles mais aussi aux compétences à identifier la structure morphologique de pseudo-mots (en plus du lien avec le niveau de lecture de mots). Le facteur grapho-morphologique identifié à ce niveau scolaire (cf. ACP supra) jouerait donc un rôle important dans la compréhension de phrases en arabe. Contrairement aux niveaux précédents, les mesures des connaissances grapho-morpho-dérivationnelles ne font apparaître aucun lien avec la compréhension de phrases, de même que le deuxième facteur

extrait de l'ACP : les connaissances graphophonologiques implicites, telles qu'elles ont été évaluées dans cette étude, ne participeraient pas à la compréhension de phrases en arabe.

2.4.2.2 Analyse de régressions sur la variable « compréhension de phrases écrites »

En 3^e année, sont entrées dans le modèle les trois variables présentant des corrélations significatives avec le niveau de compréhension de phrases (détection explicite grapho-syntaxique, lecture de mots et détection explicite grapho-morpho-dérivationnelle, cf. tableau 23).

Tableau 23 : Contribution des variables entrées dans le modèle de régression sur le niveau en compréhension de phrases en 3^e année

	R	R ²	R ² modifié	F	p
Pas 1 : Det. Exp. G-Syntaxique	.54	.2920	.2920	18,56	<.0001
Pas 2 : Lecture de mots	.60	.3566	.0646	4,42	<.05
Pas 3 : Det. Exp. G-morpho-dériv.	.61	.3712	.0147	1,00	ns

Cette analyse permet de constater que les variables contribuant le plus à expliquer la variance en compréhension de phrases sont celles mesurées par les épreuves de détection explicite grapho-syntaxiques et de lecture de mots, qui expliquent ensemble 35,66% de la variance. La variable « détection explicite grapho-syntaxique » apporte la contribution la plus importante (29,20%). La variable « lecture de mots », incluse dans un deuxième temps ajoute 06,46 % de variance expliquée (cf. tableau 23). La contribution ajoutée de la variable « détection explicite grapho-morpho-dérivationnelle » (1,47 %) ($\beta = .13$; ns) n'est pas significative.

En 4^e année, si l'on entre dans le modèle les deux variables corrélées significativement au niveau de compréhension de phrases (détection explicite grapho-morpho-dérivationnelle et lecture de mots ; cf. tableau 24), on observe que seule la variable « détection explicite grapho-dérivationnelle » contribue à expliquer 14,52 % de la variance en compréhension de phrases. L'ajout de 3,71 % de variance expliquée par la variable « lecture de mots » est trop faible pour être significatif ($\beta = .21$, ns).

Tableau 24 : Contribution des variables entrées dans le modèle de régression sur le niveau en compréhension de phrases en 4^e année

	R	R ²	R ² modifié	F	p
Pas 1 : Dét Exp G-morpho-dériv.	.38	.1452	.1452	6,62	<.05
Pas 2 : Lecture de mots	.43	.1823	.0371	1,72	ns

En 5^e année, les variables « identification de la construction morphologique » et « lecture de mots », présentant des coefficients de corrélations significatifs avec les scores en compréhension de phrases sont entrées dans le modèle de régression pas à pas (*cf.* tableau 25).

Tableau 25 : Contribution des variables entrées dans le modèle de régression sur le niveau en compréhension de phrases en 5^e année

	R	R ²	R ² modifié	F	p
Pas 1 : Identification construction morphologique de pseudo-mots	.49	.2397	.2397	15,13	<.0004
Pas 2 : Lecture de mots	.54	.2877	.0480	3,17	ns

Ce dernier met en évidence que seule la variable « identification de la construction morphologique » permet d'expliquer de façon significative 15,13 % de la variance en compréhension de phrases à ce niveau scolaire. Les 4,80 % de part ajoutée par la variable « lecture de mots », entrée dans un second temps sont trop faibles pour être significatifs ($\beta = .24$, ns).

Les connaissances apparaissant contribuer à la compréhension de phrases écrites sont différentes d'un niveau à l'autre.

En 3^e année, le niveau de compréhension de phrases est en premier lieu expliqué par les connaissances grapho-syntaxiques permettant d'analyser la composition des mots d'une phrase en arabe standard (29,20 % de variance expliquée). La reconnaissance de mots écrits, composante nécessaire mais pas suffisante à l'apprentissage de la lecture, contribue également et de façon indépendante à expliquer les capacités de compréhension mais toutefois faiblement (6,46 % de variance expliquée).

En 4^e année, les connaissances morpho-dérivationnelles permettant d'analyser consciemment la structure morphologique de mots arabes dérivés expliquent à elles seules le niveau de compréhension de phrases (14,52 % de variance expliquée).

En 5^e année, ce sont là encore les capacités à analyser la construction morphologique qui interviendraient, mais appliquées à des pseudo-mots c'est-à-dire sans rapport avec le lexique arabe (23,97 % de variance expliquée par la variable « identification de la construction morphologique de pseudo-mots »).

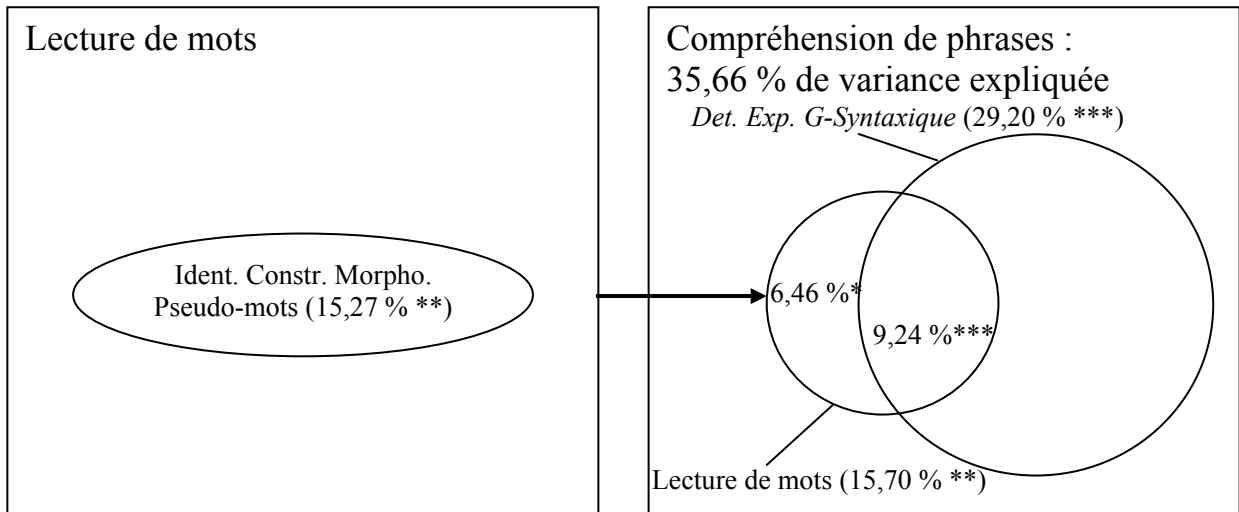
2.4.3 Synthèse sur la contribution des différentes variables à l'explication du niveau de lecture de mots et de compréhension de phrases écrites

A l'issue de ces multiples analyses, il est possible de dégager les points suivants :

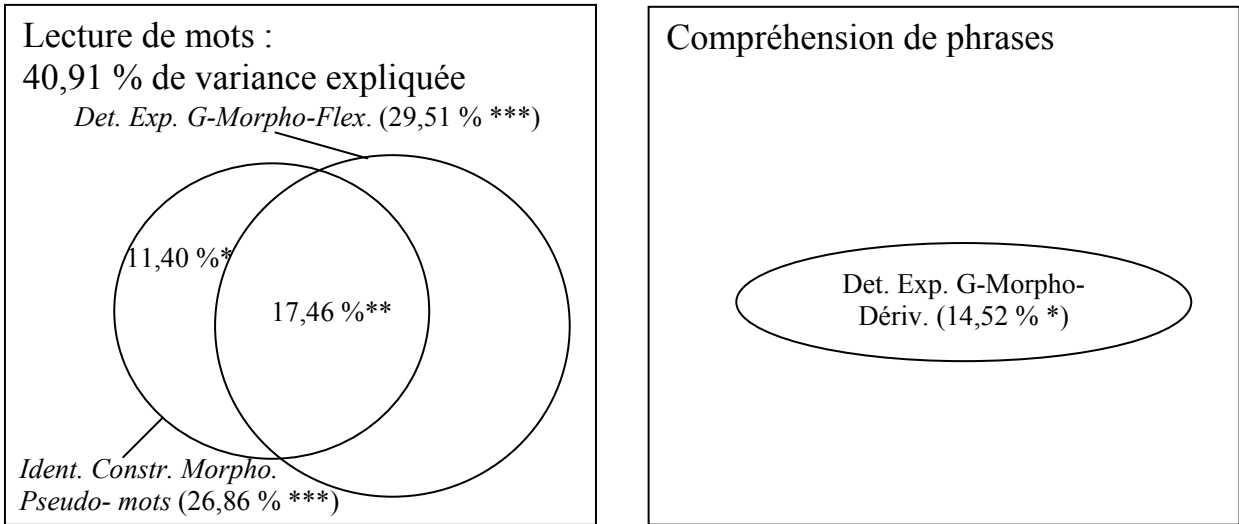
- En 3^e année, le facteur « connaissances grapho-sémantiques du système arabe » contribuerait à expliquer le niveau de compréhension en lecture par le biais des connaissances grapho-syntaxiques qui sont celles ayant le plus de poids dans la constitution de ce facteur (expliquent 29,20 % de la variance en compréhension). Telles qu'elles sont mesurées ici, ces connaissances sont notamment impliquées dans l'analyse de la composition des mots dans les phrases. La lecture de mots, elle-même expliquée par les capacités à identifier la construction morphologique de pseudo-mots (15,27 % de la variance en lecture de mots expliquée par cette variable), contribuerait faiblement à expliquer le niveau en compréhension de phrases (ajoute 6,46 % à l'explication de la variance, indépendamment des connaissances morphologiques), et cela exclusivement à ce niveau scolaire.
- En 4^e année, le facteur « connaissances grapho-sémantiques » identifié contribuerait à expliquer le niveau de lecture de mots par le biais des connaissances morpho-flexionnelles, permettant d'analyser la construction syntaxique interne des verbes et des noms féminins, et des compétences morphologiques appliquées aux pseudo-mots (ces deux variables expliqueraient ensemble 40,91 % de la variance en lecture de mots). La compréhension de phrases écrites serait pour sa part dépendante des connaissances morpho-dérivationnelles permettant d'identifier la structure interne de mots dérivés en arabe standard (14,52 % de variance expliquée).
- En 5^e année, seul le premier facteur « connaissances grapho-morphologiques » dégagé contribuerait à expliquer le niveau de compréhension de phrases mais aussi le niveau de lecture de mots par le biais des capacités à identifier la construction morphologique de pseudo-mots (expliquent respectivement 23,97 % et 15,79 % de la variance).

Le schéma 2 reprend les principales données statistiques établies par les modèles de régression menés sur les niveaux de lecture de mots et de compréhension de phrases pour chaque niveau scolaire. Les pourcentages de variance partagée par les variables significatives incluses dans les équations de régression sont également précisés.

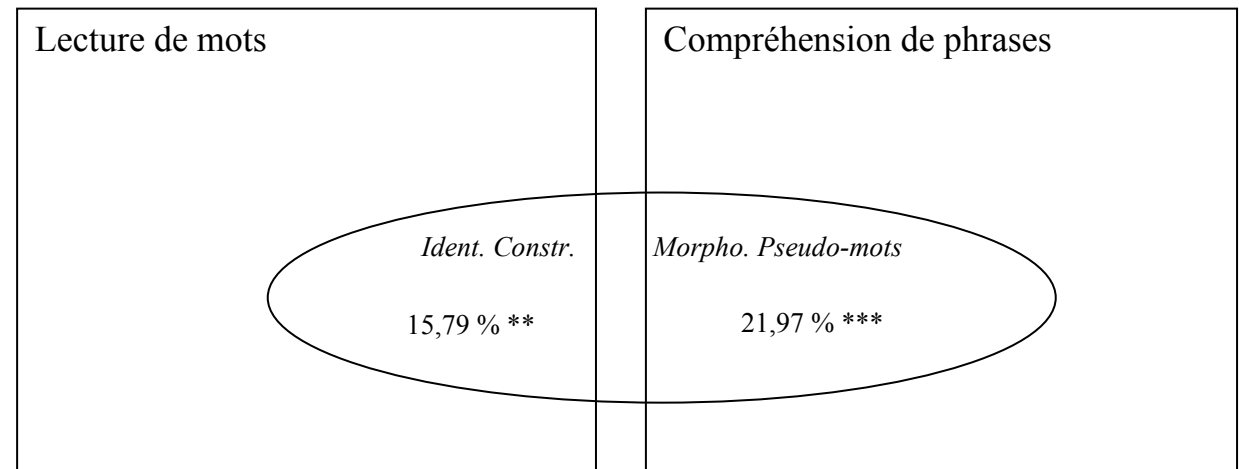
3^e année



4^e année



5^e année



*significatif à $p < .05$; **à $p < .01$; *** à $p < .001$. *En italique : Variables entrant dans la constitution du facteur dégagé au niveau scolaire indiqué.*

Schéma 2 : Représentation des variables expliquant le niveau de lecture de mots et de compréhension de phrases à chaque niveau scolaire.

3. Discussion

L'étude 1 avait comme objectif d'étudier l'évolution et le rôle de diverses connaissances sur l'apprentissage de la lecture en arabe vocalisée entre la 3^e et la 5^e année d'apprentissage.

Par rapport à l'évolution des connaissances, nous nous attendions d'une part à ce que les connaissances graphophonologiques explicites soient rapidement maîtrisées et, d'autre part, à ce que la sensibilité à la dimension morphologique des mots arabes soit particulièrement développée car les caractéristiques morphologiques y tiennent un rôle primordial à l'écrit, même vocalisé. Les scores plafonds ou au contraire liés au hasard ne nous ont pas permis de vérifier rigoureusement ces hypothèses. Néanmoins, les profils de résultats indiquent que le code graphophonologique de l'arabe est suffisamment bien maîtrisé pour que la détection consciente des unités phonologiques de l'arabe ne pose aucune difficulté dès la 3^e année. En revanche, la version implicite du paradigme de détection d'intrus n'a pas permis ici de faire émerger une sensibilité grapho-morphologique, ni d'ailleurs une sensibilité grapho-syntaxique en arabe. Cela ne signifie pas pour autant que de telles connaissances implicites n'existent pas. En effet, la consigne de ces épreuves laisse l'enfant déterminer spontanément les critères sur lesquels baser sa réponse. La prise en compte par l'enfant d'autres informations peut ainsi masquer la possibilité de faire référence aux aspects morphologiques pour orienter ses choix.

Les résultats obtenus aux épreuves évaluant les connaissances grapho-morpho-dérivationnelles, grapho-morpho-flexionnelles et grapho-syntaxiques explicites suggèrent quant à eux que l'ordre de complexité dans l'émission d'une analyse consciente sur ces différents aspects va du plus facile au plus difficile : la syntaxe, la morphologie dérivationnelle puis la morphologie flexionnelle. Le constat que les connaissances des aspects morpho-dérivationnels sont plus développées que celles des marques flexionnelles¹³⁸ va à l'encontre de l'affirmation de certains auteurs (*e.g.* Gordon, 1989). Cependant, dans les études antérieures une telle comparaison n'avait pas eu lieu en langue arabe. A notre connaissance, seuls les travaux de Levin *et al.* (2001) ont démontré un développement de cette nature en hébreu, langue qui présente vraisemblablement des différences importantes avec l'arabe, même si elle est également une langue sémitique. Il pourrait ainsi y avoir dans ce domaine une spécificité de la langue arabe écrite ou encore une singularité de l'enseignement de l'arabe, tout au moins en Tunisie. En effet, nous pouvons faire l'hypothèse que l'enseignement orienterait les apprentissages vers les aspects

¹³⁸ Cette affirmation semble étayée par les variations des compétences d'identification de la construction morphologique de pseudo-mots chez les élèves de 5^e année, pour qui les règles s'apparentant à des aspects morpho-dérivationnels sont plus facilement traitées que celles s'apparentant à des aspects morpho-flexionnels.

morpho-dérivationnels qui se révèlent déterminants dans cette langue. En outre, et de manière générale, la morphologie arabe (sous entendue dérivationnelle), se caractérisant par une structure en racine/schème systématique, pourrait se révéler particulièrement productive. Il s'agirait toutefois de s'interroger sur les caractéristiques des éléments morphologiques comparés, dont la productibilité intrinsèque peut biaiser les résultats obtenus en faveur de l'une ou l'autre des deux dimensions morphologiques.

L'observation d'une plus grande aisance à détecter les aspects syntaxiques que morpho-flexionnels confirme les suggestions de Maaouia-Belkhadi (2003) concernant l'ordre des mots mais pas celles des recherches classiques en langue indo-européenne (Demont, 1994, Nocus, 1997, etc.). Ici, cette observation ne peut pas être liée au fait souvent affirmé (Demont & Gombert, 1996 ; Gaux & Gombert, 1999a) que l'application des connaissances peut être entravée par des artefacts d'origine sémantique puisque les épreuves utilisées dans notre étude ne font pas directement intervenir cette dimension. Pour autant, la possibilité de faire référence aux informations sémantiques, contenues dans les noms, pour résoudre l'épreuve grapho-syntaxique ne peut pas être non plus totalement exclue. L'étude des connaissances syntaxiques a ici portée sur la transitivité des verbes mais peut finalement correspondre au repérage d'une forme lexicale à laquelle est accolée une préposition (e.g. : بنت = une fille ; لبنت pour une fille). Les difficultés méthodologiques pour contrôler l'usage des indices sémantiques dans la résolution de tâches syntaxiques sont d'ailleurs récurrentes (Gaux & Gombert, 1999b).

Un autre de nos objectifs était d'approfondir la pertinence de différentes unités phonologiques en arabe. Nos résultats permettent de vérifier que les phonèmes sont des unités segmentales trop petites pour pouvoir être l'objet d'analogies mises en place implicitement. Par ailleurs, les connaissances implicites sur la rime apparaissent légèrement à partir de la 4^e année. Cette apparition tardive de la sensibilité à des unités plus composites que le phonème est encore plus nette sur la syllabe. Le repérage implicite de la syllabe n'est visible qu'en 5^e année, niveau scolaire correspondant, en Tunisie, à l'introduction progressive de la lecture en version non vocalisée. Nous formulons l'hypothèse que l'évolution de la sensibilité à certaines unités phonologiques pourrait être liée non seulement à l'expérience de l'écrit mais plus précisément à la confrontation avec une forme d'écriture non vocalisée. En écriture non vocalisée, les consonnes ne seraient plus traitées comme des unités phonologiques minimales mais comme un ensemble indissociable des voyelles « invisibles » qui s'y superposent. En d'autres termes, apprendre à lire sans voyelles peut influencer les traitements phonologiques implicites vers une association des voyelles aux consonnes, constitutive d'unités plus larges telles que la rime ou la

syllabe. Il serait intéressant de continuer les investigations dans ce sens, de façon à déterminer à des niveaux plus avancés, si c'est effectivement l'exposition à l'écriture non vocalisée qui engendre cette sensibilité et comment cette dernière s'articule avec les aspects morphologiques responsables des configurations de voyelles à attribuer aux mots.

Concernant justement ces aspects morphologiques, certains éléments se dégagent des analyses sur l'identification de la construction morphologique de pseudo-mots en fonction du type de construction (structure en racine/schème ou structure en affixe/base). Il s'avère ainsi que les structures affixées sont mieux identifiées que les configurations par intrication de voyelles. Ces résultats ne confirment pas l'hypothèse que les arabophones appréhendent sans difficulté les constructions morphologiques similaires à celles de l'arabe écrit. De notre point de vue, cela va dans le sens de Frost *et al.* (2005) qui proposent qu'une structuration non linéaire rend le découpage morphologique peu transparent en provoquant une rupture de la forme orthographique (et phonologique) des éléments.

Nous nous sommes également intéressée à la façon dont les connaissances grapho-morpho-dérivationnelles, grapho-morpho-flexionnelles et grapho-syntaxiques explicites s'articulent les unes avec les autres au cours des apprentissages.

Au premier niveau scolaire étudié, la relation entre ces trois types de connaissances refléterait l'existence d'un « dénominateur commun » à la base du système grapho-sémantique de l'écriture arabe.

Par la suite, une sorte de réorganisation pourrait s'effectuer dans la mesure où, au niveau scolaire suivant, les variations interindividuelles s'expliquent essentiellement par les compétences à identifier la construction morphologique de pseudo-mots et non par les connaissances grapho-morpho-dérivationnelles¹³⁹.

La disparition de l'impact des connaissances grapho-morpho-dérivationnelles à ce niveau, suggère que la façon d'appréhender les aspects structuraux de l'arabe tendrait vers une distinction entre des traitements linéaires ou séquentiels (par exemple la concaténation des affixes flexionnels ou la segmentation des mots) *versus* non linéaires (*i.e.* le traitement dérivationnel en racine/schème). L'apparition, dans les analyses factorielles, des compétences à identifier la construction morphologique de pseudo-mots pose cependant problème, dans la mesure où ces connaissances amalgament l'identification de la structure de type affixation (donc

¹³⁹ Malheureusement, le fait que les mesures prises en compte lors des analyses factorielles ne soient pas identiques entre le troisième et les deux premiers niveaux d'apprentissage ne rend pas possible de suivre l'évolution de cette réorganisation rigoureusement et jusqu'au bout.

linéaire) et schématisation (donc non linéaire). Elle permet toutefois d'envisager que les connaissances sont moins dépendantes des aspects formels de l'arabe.

Pour déterminer plus précisément si ces deux types de connaissances structurales se distinguent et si oui, comment, il conviendrait de séparer ce qui relève des réponses aux items « schème » *versus* « affixe ». De cette manière, il serait possible de vérifier si c'est effectivement l'identification de la structure par affixation qui participe à la formation du facteur « dimension grapho-sémantique linéaire de la langue arabe ».

Nous nous attendions enfin à ce que les connaissances grapho-morphologiques contribuent de façon importante à la réussite en lecture de mots et en compréhension de phrases écrites en version vocalisée. Les résultats vont dans ce sens, même si l'absence d'épreuves suffisamment sensibles a sans aucun doute atténué le rôle des connaissances graphophonologiques dans la reconnaissance de mots écrits.

Plus précisément, les compétences morphologiques dégagées des aspects de surface des langues jouent un rôle, dès le premier niveau d'apprentissage et de façon durable, lors de la reconnaissance des mots écrits en arabe. La compréhension des principes d'application de règles morphologiques complexes, indépendamment des spécificités de formes, existe dès la 3^e année et interviendrait pour identifier les mots écrits même orthographiquement transparents, car vocalisés. Au niveau intermédiaire (4^e année), certaines précisions doivent être apportées. Pour ce groupe, les connaissances grapho-morpho-flexionnelles en arabe interviendraient de façon privilégiée, vraisemblablement car ces règles sont accentuées lors des apprentissages à ce niveau scolaire.

Concernant la dimension de compréhension en lecture, l'importance des connaissances grapho-syntaxiques explicites en 3^e année confirme le rôle déterminant dans l'élaboration d'une représentation sémantique, des compétences à identifier la frontière des mots et de l'organisation de ces derniers dans la phrase. A ce niveau, l'élaboration des stratégies d'interprétation dépendrait aussi à ce niveau de l'automatisation des processus de bas niveau, notamment des traitements graphophonologiques. Cet argument est en effet soutenu par le fait que les habiletés de reconnaissance des mots écrits contribuent à la compréhension de façon non négligeable.

Que l'analyse de la construction syntaxique des phrases soit un élément central lors de l'élaboration des procédures de compréhension en arabe (Romdhane, 2003) n'empêche pas que les compétences à analyser la structure interne des mots puissent également intervenir. Les résultats en 4^e et 5^e année confortent cette position. Cette intervention apparaît débuter en 4^e année par la mise en œuvre des connaissances grapho-morpho-dérivationnelles impliquées dans

la reconnaissance du schème et de son application à une racine arabe spécifique. Les enseignements qui soutiennent l'acquisition des schèmes de la langue arabe et se généralisent à ce niveau, ont visiblement une incidence sur l'élaboration de ces connaissances. De nouvelles recherches seront en revanche nécessaires pour tenter de comprendre pourquoi, contrairement à ce qui aurait pu être prédit, c'est plutôt les connaissances flexionnelles qui interviennent sur la lecture de mots et les connaissances dérivationnelles sur la compréhension.

En 5^e année, avec l'avancée dans les apprentissages, l'installation des connaissances se renforce et la maîtrise des principes de construction morphologique s'affine. Cela explique qu'interviennent lors de la compréhension évaluée à ce niveau scolaire les compétences morphologiques indépendantes des aspects formels de l'arabe et liées à une compréhension générale des processus de création lexicale.

En définitive, cette étude permet de confirmer le rôle décisif des traitements morphologiques lors de la lecture en arabe, établi par quelques autres auteurs chez l'enfant (*e.g.* Ammar, 1997 ; Belajouza & Snoussi, 2006 ; Boukhadida, en cours). Elle invite également à approfondir la façon d'envisager la spécificité de la lecture en langue arabe en termes de structure morphologique linéaire *versus* non linéaire. Cette distinction nous semble en effet importante pour mieux comprendre comment interagissent la gestion des voyelles brèves du schème et celle des ajouts graphiques à la racine (notamment les consonnes en position initiale et finale) dans le développement des connaissances morphologiques et pour maîtriser la lecture d'abord en version vocalisée puis en version non vocalisée.

Chapitre 8 : Étude 2 - Connaissances graphophonologiques, grapho-morpho-dérivationnelles, grapho-morpho-flexionnelles, grapho-syntaxiques et lecture en portugais langue première

1. Méthode

1.1 Population

Cent vingt-cinq élèves brésiliens scolarisés en 3^e, 5^e et 7^e année (correspondant aux classes françaises CE2, CM2 et 5^e) dans un établissement primaire et secondaire français privé, le lycée Pasteur de São Paulo (quartier villa Guarani, région sud ouest de São Paulo) ont participé à cette expérience d'août à septembre 2005. Ont été exclus des analyses les élèves redoublants (concerne 3 élèves de 3^e année, 3 élèves de 5^e année et 1 élève de 7^e année) et ceux ayant participé à moins de la moitié des épreuves (concerne 2 élèves de 5^e année et 2 élèves de 7^e année). Les résultats d'un élève de 3^e année ont également été supprimés du fait du comportement particulièrement agité de l'enfant. Egalement ceux d'un élève de 5^e année car ces performances en lecture étaient inférieures aux performances minimales obtenues par le groupe des 3^e années. L'effectif total une fois ces critères d'exclusion appliqués, s'élève ainsi à 112 participants. Les caractéristiques de cette population par niveau scolaire sont précisées dans le tableau 26.

Tableau 26 : Répartition des élèves en fonction de leur niveau scolaire

	Effectif	Nombre de garçons	Nombre de filles	Age chronologique moyen	Ecart-type (en mois)	Age du plus jeune	Age du plus âgé
3 ^e année ¹⁴⁰	48	30	18	9 ; 5 ans (113,98 mois)	3,68	8 ; 9 ans (106,17 mois)	9 ; 11 ans (120,13 mois)
5 ^e me année ¹⁴¹	30	17	13	11 ; 3ans (136,06 mois)	3,95	10 ; 6 ans (127,33 mois)	11 ; 11 ans (144,40 mois)
7 ^e me année ¹⁴²	34	18	16	13 ; 6 ans (163,33 mois)	4,53	12 ; 9 ans (124,13 mois)	14 ; 5 ans (137,07 mois)

¹⁴⁰ Elèves provenant de deux classes différentes.

¹⁴¹ Elèves provenant d'une même classe.

¹⁴² Elèves provenant d'une même classe.

1.2 Matériel

Les instruments utilisés pour évaluer les élèves des trois niveaux scolaires en portugais conservent les mêmes principes que ceux employés en langue arabe. La description des épreuves en portugais s'attachera donc plus particulièrement à rappeler leurs principaux fondements et à indiquer la spécificité de leur construction selon l'ordre suivant :

- L'évaluation des performances en lecture à l'aide d'une épreuve chronométrée de lecture de mots isolés à voix haute et d'une épreuve de compréhension de phrases écrites.
- L'évaluation des connaissances graphophonologiques, grapho-morphologiques (dérivationnelle et flexionnelle) et grapho-syntaxiques en portugais mises en place implicitement et explicitement à l'aide du paradigme de détection d'intrus qui consiste à choisir parmi deux stimuli, celui se distinguant d'un stimulus-cible. Comme en arabe, les épreuves ont été présentées oralement avec un support écrit.
- L'évaluation du niveau de raisonnement sur l'application des règles de constructions morphologiques indépendamment des aspects de surface spécifiques au portugais ou aux autres langues connues des participants. Ces compétences grapho-morphologiques (dérivationnelle et flexionnelle) générales ont été évaluées à l'aide de la même épreuve d'identification de la construction morphologique de pseudo-mots qu'en arabe.

1.2.1 Epreuves de lecture en portugais

1.2.1.1 *Epreuve de lecture de mots isolés en une minute*

Les compétences de reconnaissance de mots écrits en portugais sont évaluées à l'aide d'une épreuve individuelle et chronométrée de lecture de mots isolés qui reprend en partie le lexique du sub-test de lecture de la batterie TDE (Milnitsky Stein, 1994). Il s'agit de lire en une minute à voix haute et les uns à la suite des autres, le maximum de mots d'une liste proposée. Cette liste est composée de 108 mots de différentes catégories grammaticales et de différentes longueurs, variant en terme de fréquence, de régularité et de complexité (*cf.* annexe 14). Le niveau de performance est mesuré en fonction de l'exactitude et de la rapidité en lecture.

1.2.1.2 *Epreuve de compréhension de phrases écrites*

Cette épreuve (Zanella, 2005) est l'adaptation en portugais du sub-test « compréhension de phrases » du test LMC-R de Khomsi (1999). Elle est composée de 22 items, les participants devant, pour chaque item choisir l'image, parmi quatre, correspondant à une phrase écrite (*cf.* annexe 15). Pour 14 items, le choix de l'image peut s'effectuer par le biais d'une stratégie « imagée » et pour 10 items, il nécessite une stratégie « inférentielle ». Les modifications de

réponse en cours de résolution sont possibles et le temps imparti pour répondre n'est pas une contrainte stricte.

1.2.2 Epreuves implicites de détection d'intrus

La consigne utilisée pour évaluer le niveau implicite des connaissances est identique en arabe et en portugais puisque dans les deux cas, elle ne spécifie pas l'unité linguistique ou le type d'analyse sur lequel ce choix doit se fonder de façon que les réponses soient engendrées par une association libre et sans que l'attention ne soit dirigée par l'expérimentateur.

1.2.2.1 *Epreuve implicite de détection d'intrus graphophonologique*

Cette épreuve est destinée à évaluer la sensibilité des participants aux unités syllabiques et intrasyllabiques des mots (noms ou adjectifs) écrits portugais. Elle comporte 12 items. Pour chacun d'eux, les participants doivent choisir entre deux mots, celui « allant le moins » avec le mot cible (*e.g.* : qu'est ce qui va le moins avec « *chave* » ($[\cup\Sigma\alpha\tau\leftrightarrow]$ = clé) : « *monte* » ($[\cup\mu\sigma(\nu\tau\leftrightarrow)]$ = mont) ou « *tubo* » ($[\cup\phi\upsilon\equiv\upsilon]$ = tube) ?). 3 items portent sur la syllabe finale de mots trisyllabiques, 3 items sur la rime intrasyllabique de mots bisyllabiques, 3 sur le phonème consonantique final et 3 sur le phonème vocalique final de mots bisyllabiques. La structure phonologique et graphique, la catégorie grammaticale et la familiarité des paires de mots proposés ont été contrôlées sur l'ensemble du matériel. Ce dernier ainsi que son mode de présentation, sa traduction et ses caractéristiques sont présentés dans l'annexe 16. La randomisation et la cotation sont identiques à celles appliquées à l'épreuve arabe.

1.2.2.2 *Epreuve implicite de détection d'intrus grapho-morpho-dérivationnel*

Cette épreuve est destinée à évaluer la sensibilité des participants à la construction morphologique en base / affixe de mots écrits en portugais (noms ou verbes). Elle comporte 12 items, chacun constitué d'un mot cible dérivé et d'une paire de mots dont l'un est également dérivé et partage le même affixe que la cible et l'autre, jouant le rôle d'intrus, qui malgré les apparences phonographiques n'est pas construit (*e.g.* : *deslizar* = glisser, morphème libre) ou utilise une base non identifiable (*e.g.* : *doutor*), si bien que le lien de construction avec la cible est inexistant ou opaque. Les participants doivent choisir entre les deux mots, celui « allant le moins » avec le mot cible (*e.g.* : qu'est ce qui va le moins avec « *chaveiro* » ($[\Sigma\alpha\cup\tau\cup\iota\upsilon\upsilon]$ = porte-clé « *chave* » $[\cup\Sigma\alpha\tau\leftrightarrow]$ = clé) : « *cinzeiro* » ($[\sigma\iota\cup(\zeta\cup\iota\upsilon\upsilon)]$ = cendrier, vient de « *cinza* » $[\cup\sigma\iota(\zeta\cup)]$ = cendre) ou « *pendeiro* » ($[\pi\epsilon\cup(\delta\cup\iota\upsilon\upsilon)]$ = tambourin pas de racine identifiable) ?

La moitié des items concernent la formation de substantifs par ajout de suffixe :

- 3 items portent sur le suffixe « *-eiro* » ([ɨpυ]) permettant de former des noms d'agent, d'instrument ou de lieu.
- 3 items portent sur le suffixe « *-or* » ([op]) permettant de former des noms de d'agent, d'instrument ou de qualité.

L'autre moitié des items concernent la formation de verbes par ajout de préfixe :

- 3 items portent sur le préfixe « *des-* » ([δ↔σ] ou [δ↔ζ]) dont la valeur sémantique marque l'opposition ou la cessation par rapport au sens de la base verbale.
- 3 items portent sur le préfixe « *ré-* » ([↔]) dont la valeur sémantique marque la répétition, la réversibilité ou le retour à un forme ou un état antérieur.

Les aspects sémantiques des items, la structure phonologique et graphique ainsi que la familiarité des paires de mots proposés ont été contrôlés. Le matériel, son mode de présentation, sa traduction et ses caractéristiques figurent dans l'annexe 17. La randomisation et la cotation sont identiques à celles appliquées à l'épreuve arabe.

1.2.2.3 *Epreuve implicite de détection d'intrus grapho-morpho-flexionnel*

Cette épreuve est destinée à évaluer la sensibilité des participants à la structure flexionnelle de mots écrits en portugais (noms, adjectif ou verbes). Elle comporte 12 items, chacun constitué d'un mot cible et d'une paire de mots dont l'un fait apparaître une structure flexionnelle identique au mot cible et l'autre, jouant le rôle d'intrus, une forme non fléchie ou une structure flexionnelle graphophonologiquement semblable ou ressemblant à celle du mot cible mais syntaxiquement et sémantiquement différente. Les participants doivent choisir entre les deux mots, celui « allant le moins » avec le mot cible (*e.g.* : qu'est ce qui va le moins avec « *bailarina* » ([βαιλινριν]) = danseuse, féminin de « *bailarino* »), « *secretária* » ([σ↔κρ↔↔ταρφυ] = secrétaire, féminin de « *secretário* ») ou « *carioca* » ([κινρφκ] = habitant(e) de rio ?),

6 items concernent la formation du féminin de noms ou d'adjectifs. Les intrus terminent par le même graphème que celui de la marque du féminin (« *a* » [-]) mais sont neutres (*e.g.* *o/a poeta* [πωετ] = un(e) poète).

6 items concernent les désinences verbales :

- 3 items portent sur la conjugaison du prétérit parfait simple de l'indicatif à la première personne du singulier¹⁴³ de verbes du premier groupe (infinitif en « *-ar* » [-αρ] : désinence « *-ei* » [ɨ] *e.g.* : « *amar* » [αμαρ] = aimer, « *amei* » [αμη] = j'ai aimé). Les intrus sont des verbes du premier groupe conjugués à la 1^e P.S. du futur simple de l'indicatif dont la désinence est « -

¹⁴³ Abréviation : 1^e P. S.

arei » [-αρ□ι] (*i.e.* la même désinence qu’au prétérit mais accolé à la forme infinitive du verbe ; *e.g.* : « *amarei* » [□□μ□αρ□ι]= j’aimerai).

- 3 items portent sur la conjugaison du présent de l’indicatif à la troisième personne du pluriel¹⁴⁴ de verbes du premier groupe (désinence « -am » [□μ] *e.g.* : « *amam* » [□□μ□μ]= ils aiment). Les intrus sont des verbes du premier groupe conjugués à la 3^e P.P. du prétérit parfait simple dont la désinence est « -aram » [-αρ□μ] (*i.e.* la même désinence qu’au présent mais accolé à la forme infinitive du verbe ; *e.g.* : « *amaram* » [□□μ□αρ□μ]= ils ont aimé).

Pour que les intrus ne soient pas reconnaissables par leur terminaison spécifique, les seconds verbes des paires sont des verbes dont la racine termine par « r » ce qui permet que la dernière syllabe de leur forme conjuguée soit la même que celle des intrus (*e.i.* « -rei » pour les items sur le prétérit avec intrus au futur et « -ram » pour les items sur le présent avec intrus au prétérit).

Les aspects sémantiques des items, la structure phonologique et graphique ainsi que la familiarité des paires de mots proposés ont été contrôlés. Le matériel, son mode de présentation, sa traduction et ses caractéristiques figurent dans l’annexe 18. La randomisation et la cotation sont identiques à celles appliquées à l’épreuve arabe.

1.2.2.4 *Epreuve implicite de détection d’intrus grapho-syntaxique*

Cette épreuve est destinée à évaluer la sensibilité des participants à la structure syntaxique de phrases écrites en portugais. Elle comporte 12 items, chacun constitué d’une phrase cible et d’une paire de phrases dont l’une partage la même structure syntaxique que la phrase cible en utilisant une forme verbale composée d’un auxiliaire et d’un participe et l’autre non. Les participants doivent choisir entre deux phrases, celle « allant le moins » avec la phrase cible (*e.g.* : qu’est ce qui va le moins avec « *ele tem andado* » [(□ελ↔ □τεε)μ □]ν□δ□]ν[□] = il a marché « *ele tem rido* » [□ελ↔ □τεε)μ □]ν{□|□} = il a ri « *ele tem medo* » [□ελ↔ □τεε)μ □]ν□ε[□] = il a peur).

Pour 5 items, les phrases sont construites à l’aide de la forme composée « *estar* + gérondif » signifiant une action qui est en cours (*e.g.* *esta saindo* = il sort). Dans les phrases intruses le verbe « *estar* » est également utilisé mais n’est pas un auxiliaire puisqu’il est suivi d’un adjectif attribut terminant par les graphèmes « -ndo » comme le gérondif (*e.g.* *esta lindo* = il est joli).

Pour 7 items, les phrases sont construites au parfait composé de la forme « *ter* + participe passé » qui indique une action ayant débuté dans le passé mais qui n’est pas terminée (*e.g.* *ela tem mentido* = elle a menti). Dans les phrases intruses le verbe « *ter* » n’est pas un auxiliaire et est suivi d’un nom complément d’objet direct terminant par la même syllabe que le participe passé (*e.g.* *ela tem marido* = elle a un mari).

¹⁴⁴ Abréviation : 3^e P. P.

Les aspects sémantiques des items, le nombre de syntagmes, la structure phonologique et graphique des paires de phrases, la familiarité du lexique, le genre, le nombre des noms et la forme des verbes composant les paires de phrases ont été contrôlés. Le matériel, son mode de présentation, sa traduction et ses caractéristiques sont présentés dans l'annexe 19. La randomisation et la cotation sont identiques à celles appliquées à l'épreuve arabe.

1.2.3 Epreuves explicites de détection d'intrus

Les épreuves explicites conservent le même principe de choix entre deux stimuli mais font appel à une consigne qui précise le critère de réponse. Les aspects phonologiques, morphologiques, ou syntaxiques sur lesquels porte l'épreuve sont donc préalablement expliqués. Les participants doivent ainsi déterminer consciemment lequel des deux stimuli de la paire respecte le critère établi par la consigne.

Chaque épreuve de détection explicite d'intrus est construite sur le même mode que leur homologue implicite. Ainsi elles comportent le même nombre d'items, les mêmes caractéristiques et le même mode de contrôle que les épreuves implicites décrites précédemment, seules les consignes diffèrent. Comme pour les épreuves implicites, elles conservent globalement les mêmes principes que les épreuves explicites en arabe avec quelques modifications dans les consignes en fonction des règles manipulées en portugais. Nous ne présenterons donc, dans les paragraphes suivants, que le type de compétences évaluées, accompagné d'exemples précisant ce qui est demandé aux participants.

1.2.3.1 *Epreuve explicite de détection d'intrus graphophonologique*

Cette épreuve a pour objectif d'évaluer les connaissances graphophonologiques explicites permettant de discriminer consciemment les éléments graphophonologiques de mots écrits en portugais. Les participants doivent déterminer entre deux mots, lequel ne se termine pas par le même son qu'un mot cible. Par exemple, qu'est-ce qui ne se termine pas par le même son que « *baile* » ([βαιλ↔] = bal) : « *poste* » ([πσ↔] = poteau) ou « *curso* » ([κρσ↔] = cours) ? Le matériel, son mode de présentation, sa traduction et ses caractéristiques sont présentés dans l'annexe 20.

1.2.3.2 *Epreuve explicite de détection d'intrus grapho-morpho-dérivationnel*

Cette épreuve a pour objectif d'évaluer les connaissances morpho-dérivationnelles explicites permettant d'identifier la construction morpho-dérivationnelle de mots dérivés de façon à segmenter les mots en base / affixe. Dans cette épreuve, les participants doivent déterminer parmi deux mots proposés le mot, celui qui, malgré les apparences phonographiques, ne fait pas

apparaître la même construction morphologique qu'un mot cible. Il leur est par exemple demandé quel est le mot qui ne vient pas d'un autre mot comme « reabrir » ([{↔}∪=ριρ] = rouvrir vient de « abrir »), « reler » ([{↔}∪λερ] = relire, vient de « ler ») ou « regar » ([{↔}∪gαρ] = arroser, non construit) ? Le matériel, son mode de présentation, sa traduction et ses caractéristiques sont présentés dans l'annexe 21.

1.2.3.3 *Epreuve explicite de détection d'intrus grapho-morpho-flexionnel*

Cette épreuve a pour objectif d'évaluer les connaissances morphosyntaxiques explicites permettant d'identifier consciemment les aspects syntaxiques des mots tels que le genre des substantifs ou la conjugaison des verbes à partir d'éléments graphiques discriminants et en fonction des conventions de la langue. Dans cette épreuve, les participants doivent déterminer, parmi deux propositions et malgré les apparences phonographiques, le nom qui n'est pas assigné d'une marque de genre, ou le verbe qui n'est pas marqué par la même conjugaison que le mot cible. Il leur est par exemple demandé quel est le mot qui n'est pas seulement un féminin, qui ne sert pas seulement pour les filles, comme « *solteira* » ([σολ←∪τ□ιρ□] = célibataire, féminin de « *solteiro* »), « *dançarina* » ([δ□(σ□∪ριν□] = danseuse, féminin de : « *dançarino* ») ou « *camarada* » ([κ□μ□∪ρα□] = camarade, neutre) ? Ou encore quel est le verbe qui n'est pas conjugué comme ou qui n'est pas au même temps que : « *beliscam* » ([β←λισ∪κ□μ] = ils pincent), « *seguram* » ([σ←γυ∪ρ□μ] = ils tiennent) ou « *sonharam* » ([συ/□∪ρ□μ] = ils ont rêvé) ? Le matériel, son mode de présentation, sa traduction et ses caractéristiques sont présentés dans l'annexe 22.

1.2.3.4 *Epreuve explicite de détection d'intrus grapho-syntaxique*

Cette épreuve a pour objectif d'évaluer les connaissances syntaxiques explicites permettant d'analyser consciemment la structure syntaxique de phrases en terme de construction verbale de façon à distinguer les catégories constitutives du syntagme verbal et éventuellement leur fonction syntaxique. Dans cette épreuve, les participants doivent déterminer, parmi deux propositions et malgré les apparences phonographiques, la phrase pour laquelle le syntagme verbal n'est pas constitué d'une forme verbale composée. Il leur est par exemple demandé quelle est la phrase qui n'est pas construite avec deux verbes comme « *o furo esta crescendo* » ([υ ∪φυρυ ∪Στ□ κρ←Σ∪)σε)νδυ] = le trou s'agrandi), « *o frio esta vindo* » ([υ ∪φριυ ∪Στ□ ∪σι)νδυ] = le froid arrive) ou « *o fogo esta brando* » ([υ ∪φογυ ∪Στ□ ∪βρ□)νδυ] = le feu est faible) ? Le matériel, son mode de présentation, sa traduction et ses caractéristiques sont présentés dans l'annexe 23.

1.2.4 Epreuve d'identification de la construction morphologique de pseudo mots

Cette épreuve évaluant les compétences à extraire les règles morphologiques sous jacentes à la construction de formes lexicales inconnues des participants a déjà été présentée dans le chapitre 7, paragraphe 1.2.4, car c'est la même épreuve que celle présentée aux groupes arabophones mais en version portugaise. En effet, la consigne orale et les définitions écrites des pseudo-mots ont été données en portugais. Ainsi, le sens de lecture, y compris celui des trois propositions de réponses est inversé par rapport à son homologue arabe. De plus, les aspects sémantiques des définitions des pseudo-mots diffèrent parfois légèrement entre la version arabe et portugaise de façon à respecter les spécificités des langues dans leur utilisation naturelle et ainsi éviter l'apparition d'ambiguïtés (à cause de la présence d'homonyme par exemple) ou d'incongruences (par exemple pour définir un pseudo-mot dérivé signifiant un lieu il est plus adapté d'utiliser le mot « دَار » = maison, en arabe et « loja » = magasin, en portugais). Cependant, les différences entre les deux versions n'affectent pas le principe de l'épreuve et les caractéristiques contrôlées des items décrits dans le chapitre précédent (cf. annexe 24).

1.3 Procédure

La présentation du matériel expérimental aux quatre classes a nécessité plusieurs sessions¹⁴⁵. Lors de chaque session, chaque épreuve commencée était achevée tout en respectant le rythme de travail des élèves. La procédure choisie repose sur les mêmes arguments que ceux présentés paragraphe 1.3 du chapitre 7 concernant l'expérience menées en langue arabe. L'ordre constant de présentation des épreuves ou groupes d'épreuves en portugais est donc le même :

- Epreuves implicites de détection d'intrus en passation collective avec randomisation de l'ordre de présentation des épreuves graphophonologique, grapho-morpho-dérivationnelle, grapho-morpho-flexionnelle et grapho-syntaxique. La première épreuve implicite présentée était précédée d'un exemple et de deux items d'entraînements sans *feed-back* (cf. annexe 25) construits de façon à expliquer le principe de la tâche de détection d'intrus sans rendre possible la distinction partage/non partage d'une unité linguistique avec la cible.
- Epreuve d'identification de la construction morphologique de pseudo mots en passation collective. L'expérimentateur, après avoir expliqué la consigne et demandé de répondre aux deux items d'entraînement sans *feed-back*, n'intervient pas au cours de sa résolution, qui s'effectue en silence.

¹⁴⁵ Au cours de cette expérimentation, d'autres sessions proposant des épreuves en français ont été intégrées. Ces épreuves et les résultats obtenus sont présentés dans le chapitre 9.

- Épreuves explicites de détection d'intrus en passation collective. Pour chaque classe, l'ordre de présentation des épreuves graphophonologique, grapho-morpho-dérivationnelle, grapho-morpho-flexionnelle et grapho-syntaxique, défini au hasard, est différent de celui des épreuves implicites. En plus des quatre exemples donnés lors de la consigne, chaque épreuve était précédée de deux items d'entraînement avec correction portant sur d'autres unités du même domaine que celles utilisées dans l'épreuve (*cf.* annexes 20 à 23). Ils permettaient d'expliquer la distinction partage/non partage d'une unité linguistique avec la cible.
- Épreuve de compréhension de phrases écrites en passation collective. L'expérimentateur, après avoir expliqué la consigne, n'intervient pas au cours de sa résolution qui s'effectue en silence. La feuille d'exercice présente un exemple d'item avec la bonne réponse.
- Test de lecture de mots écrits en une minute en passation individuelle.

Les épreuves administrées en passations collectives se sont déroulées au cours de trois sessions d'environ 45 minutes sans dépasser le rythme d'une session par semaine. Les entretiens individuels au cours desquels le test de lecture de mots écrits était administré n'excédaient pas cinq minutes par enfant. Ils ont eu lieu sur une demi-journée pour chaque classe.

Le tableau 27 récapitule les types d'épreuves évaluant les connaissances sur les dimensions graphophonologique et grapho-sémantique¹⁴⁶

¹⁴⁶ Les instructions ayant été données en portugais, les consignes ont été traduites en français.

Tableau 27 : Epreuves évaluant les connaissances sur les dimensions graphophonologique et grapho-sémantique

	Dimension graphophonologique	Dimension grapho-sémantique		
		Morpho-dérivationnelle	Morpho-flexionnelle	Syntaxe
<p><u>Epreuves implicites de détection d'intrus</u>^{1 2}</p> <p>Présentation à l'oral + support écrit Entraînement au paradigme lors de la 1^e épreuve (2 items)</p>	Consigne implicite commune : Qu'est ce qui va le moins avec :			
	<p><i>chave</i> [∪Σαπ↔]</p> <p><i>monte</i> ou <i>tubo</i> [∪μο(ντ↔) [∪φυ≡ν]</p> <p>Cf. annexe 16</p>	<p><i>chaveiro</i> (porte-clé)</p> <p><i>cinzeiro</i> ou <i>pendeiro</i> (cendrier) (tambourin)</p> <p>Cf. annexe 17</p>	<p><i>bailarina</i> (danseuse)</p> <p><i>secretária</i> ou <i>carioca</i> (secrétaire) (habitant(e) de rio)</p> <p>Cf. annexe 18</p>	<p><i>ele tem andado</i> (il a marché)</p> <p><i>ele tem rido</i> ou <i>ele tem medo</i> (il a ri) (il a peur)</p> <p>Cf. annexe 19</p>
<p><u>Epreuve d'identification de la construction morphologique de pseudo-mots</u>²</p>		<p>En kazi : <i>chat</i> = XOT <i>petit chat</i> = XOTIKE <i>grand chat</i> = XOTUKO <i>vélo</i> = GOF, <i>petit vélo</i> = GOFIKE <i>grand vélo</i> = GOFUKO <i>pot</i> = RUV. Comment dirait-on <i>petit pot</i> en kazi :</p> <p style="text-align: center;"><u>RUVIKE, RUFIKE</u> ou <u>RUVUKO</u> ?</p> <p style="text-align: center;">Cf. annexe 24</p>		
<p><u>Epreuves explicites de détection d'intrus</u>^{1 2}</p> <p>Présentation à l'oral + support écrit Entraînement systématique à la tâche (2 items)</p>	<p>Consigne explicite : Qu'est ce qui ne se termine pas par le même son que :</p> <p><i>baile</i> [∪βαιλ↔]</p> <p><i>poste</i> ou <i>curso</i> [∪π∩Στ↔] [∪κρσν]</p> <p>Cf. annexe 20</p>	<p>Consigne explicite : Qu'est ce qui ne vient pas d'un autre mot comme :</p> <p><i>reabrir</i> (rouvrir)</p> <p><i>reler</i> ou <i>regar</i> (relire) (arroser)</p> <p>Cf. annexe 21</p>	<p>Exemple de consigne explicite : Qu'est ce qui n'est pas conjugué comme :</p> <p><i>beliscam</i> (ils pincent)</p> <p><i>seguram</i> ou <i>sonharam</i> (ils tiennent) (ils ont rêvé)</p> <p>Cf. annexe 22</p>	<p>Consigne explicite : Quelle phrase n'a pas deux verbes comme :</p> <p><i>o furo esta crescendo</i> (le trou s'agrandi)</p> <p><i>o frio esta vindo</i> ou <i>o fogo esta brando</i> (le froid arrive) (le feu est faible)</p> <p>Cf. annexe 23</p>

¹ Randomisation de l'ordre de présentation des 4 épreuves

² Randomisation de l'ordre des items et des réponses dans chaque item

2. Résultats

2.1 Démarche statistique

La démarche sous-tendant l'analyse des résultats est la même que celle de la première expérience. Ainsi, l'apparition de certains scores nous a obligé à modifier la structure méthodologique prévue pour vérifier les hypothèses posées :

La première partie concerne l'analyse comparative des performances. Afin de vérifier si ces dernières augmentent entre la 3^e et la 7^e année, plusieurs analyses de variance ont été effectuées.

- Sur les deux épreuves de lecture (lecture de mots isolés en une minute et compréhension de phrases), des analyses de variance à un facteur ont été menées afin de préciser les variations du niveau de lecture en 3^e, 5^e année et 7^e année d'apprentissage.

- Concernant les épreuves implicites et explicites de détection d'intrus, notre objectif était de déterminer l'influence du niveau scolaire en fonction du degré de contrôle (implicite vs. explicite) et du domaine linguistique (graphophonologique, grapho-morpho-dérivationnelle, grapho-morpho-flexionnelle, grapho-syntaxique) mais l'ensemble de ces épreuves n'a pas pu être retenu. Pour chaque épreuve, les résultats aux tests t pour échantillon unique comparant les moyennes à la valeur du hasard ont été utilisés pour déterminer si ces données peuvent faire l'objet d'analyses statistiques inférentielles. Ainsi, des analyses en sous-ordre à deux facteurs (niveau scolaire et domaine linguistique) ont été requises du fait de l'apparition de réponses aléatoires stables et de scores plafonds sur plusieurs épreuves¹⁴⁷. Les analyses *a posteriori* ont été systématiquement menées à l'aide du test de Newman-Keuls.

- L'étude des unités graphophonologiques les plus pertinentes en portugais n'a pu être menée que sur le niveau implicite des connaissances¹⁴⁸. Les variations des performances des élèves de chaque niveau à l'épreuve de détection implicite d'intrus graphophonologique en fonction du type d'unité ont été analysées à l'aide de tests non paramétriques (test de Friedman, de Wilcoxon, de Kruskal-Wallis et de Mann-Whitney¹⁴⁹) car les critères de normalité et d'homocédasticité des résultats n'étaient pas toujours respectés.

- Afin d'étudier les compétences à extraire explicitement les règles morphologiques sur un matériel lexical inconnu, il a été menée sur l'épreuve d'identification de la construction

¹⁴⁷ Les analyses paramétriques en sous ordre ont été choisies car, hormis pour les scores supprimés, les critères de normalité et d'homocédasticité des résultats étaient respectés.

¹⁴⁸ Cette analyse n'a pas pu être menées sur le niveau explicite des connaissances graphophonologiques compte tenu des scores plafonds observés.

¹⁴⁹ Ces deux derniers tests ont été utilisés pour déterminer si les différences significatives de scores entre deux items varient entre les niveaux scolaires.

morphologique de pseudo-mots, une analyse à trois facteurs : le niveau scolaire {3^e, 4^e et 5^e année}, la nature de la construction morphologique {préfixe, suffixe, schème} et le type de règle morphologique {dérivationnelle vs. flexionnelle}.

Ces analyses ont également permis de sélectionner les épreuves pouvant être retenues dans les études corrélatives, factorielles et de régressions ultérieures.

Dans un deuxième temps, afin d'identifier les relations entre les connaissances pouvant intervenir lors de l'apprentissage de l'écrit en portugais, des analyses de corrélations et des analyses factorielles à composantes principales ont été réalisées sur les résultats aux épreuves retenues. Ces deux types d'analyses ont été menés séparément sur chaque niveau scolaire. Les épreuves retenues sont d'ailleurs la plupart du temps différentes d'un niveau à l'autre, ce qui constitue un argument supplémentaire pour traiter chaque niveau de façon isolée.

Dans un troisième temps, nous nous sommes attachée à rechercher les performances pouvant déterminer le niveau de lecture en portugais en 3^e, 5^e et 7^e année. Des analyses de corrélations ont d'abord été réalisées pour identifier le degré d'association entre les mesures de chaque domaine et les scores en lecture. Les variables obtenant des coefficients de corrélation significatifs ou élevés ont alors été entrées dans un modèle de régression pas à pas afin de déterminer les variables contribuant à expliquer le niveau de lecture de mots et de compréhension de phrases écrites¹⁵⁰. Pour les raisons évoquées plus haut et pour faciliter la comparaisons des résultats obtenus en arabe et en portugais, ces deux types d'analyses ont été menés séparément sur chaque niveau scolaire et sur les deux dimensions de la lecture¹⁵¹.

¹⁵⁰ Dans le cas où une seule corrélation est identifiée, une régression simple est alors suffisante.

¹⁵¹ Une analyse structurale n'a pas pu être menée pour les mêmes raisons que dans l'expérience précédente.

2.2 Analyse des performances obtenues

2.2.1 Epreuve de lecture de mots : effet du niveau scolaire

Le tableau 28 indique les moyennes, écart-types, minima et maxima obtenus à l'épreuve de lecture de mots en une minute, selon le niveau scolaire.

Tableau 28 : Moyennes, écart-types, minima et maxima au test de lecture de mots en une minute selon le niveau scolaire.

	Moyenne	Ecart-type	Minimum	Maximum
3 ^e année N = 48	68,94	11,05	50	92
5 ^e année N = 24	82,67	12,63	51	101
7 ^e année N = 24	93,00	9,08	73	107

L'analyse de variance à un facteur (niveau scolaire {3^e année, 5^e année, 7^e année}) met en évidence un effet du niveau scolaire sur les performances en lecture de mots isolés [$F(2,93) = 40,53$; $p < .00001$] : les élèves de 7^e année lisent en moyenne plus de mots correctement que ceux de 5^e année ($p < .001$) et de 3^e année ($p < .001$). De même, le niveau de lecture de mots est plus élevé en 5^e année qu'en 3^e année ($p < .001$).

Les capacités de décodage évoluent donc avec l'avancée dans les apprentissages.

2.2.2 Epreuve de compréhension de phrases : effet du niveau scolaire

Le tableau 29 indique les moyennes, écart-types, minima et maxima obtenus à l'épreuve de compréhension de phrases selon le niveau scolaire¹⁵².

Tableau 29 : Performances moyennes de réussite (sur 22), écart-types, minima, maxima au test de compréhension de phrases selon le niveau scolaire.

	Moyenne	Ecart-type	Minimum	Maximum
3 ^e année N = 47	17,38	3,27	6	22
5 ^e année N = 29	18,34	1,88	15	22
7 ^e année N = 28	19,21	1,99	14	22

L'analyse de variance à un facteur (niveau scolaire {3^e année, 5^e année, 7^e année}) met en évidence un effet du niveau scolaire sur les performances en compréhension [$F(2,101) = 4,39$; $p < .05$] : les scores des élèves de 7^e année sont significativement supérieurs à ceux de 3^e année (p

¹⁵² Bien que pour cette épreuve les réponses sont à choisir entre quatre propositions, il n'a pas été nécessaire de conduire des tests t pour échantillon unique pour s'assurer que les réponses ne sont pas liées au hasard (1/4 de bonnes réponses) car les scores s'élèvent systématiquement bien au dessus de 25 % de bonnes réponses.

< .05) mais pas à ceux de 5^e année. De même, la différence de scores entre les 3^e et 5^e année n'est pas significative.

Nous pouvons ainsi constater que les capacités de compréhension en portugais évoluent entre la 3^e et la 7^e année mais de façon très progressive.

2.2.3 Épreuves implicites et explicites de détection d'intrus

2.2.3.1 Sélection des épreuves en fonction des résultats obtenus

Le tableau 30 récapitule les performances moyennes obtenues pour les trois niveaux scolaires aux épreuves implicites et explicites de détection d'intrus graphophonologique, grapho-morpho-dérivationnel, grapho-morpho-flexionnel et grapho-syntaxique.

Tableau 30 : Performances moyennes de réussite sur 12 (et écart-types) obtenues aux épreuves implicites et explicites de détection d'intrus en fonction du niveau scolaire, tests-t de comparaison à la valeur 6 (niveau du hasard)

Epreuves de détection d'intrus	3 ^e année	5 ^e année	7 ^e année	Tous niveaux scolaires Comparaison au hasard (6/12)
Implicite graphophonologique	8,44 (1,98) N = 45	6,89 (2,08) N = 27	8,22 (2,20) N = 32	7,97 (2,16) t(103) = 9,32 ; p < .0001
Implicite grapho-morpho-dérivationnel	6,09 (2,07) N = 45	4,83 (1,17) N = 28	5,76 (1,67) N = 34	5,65 (1,80) t(106) = -1,93 ; ns
Implicite grapho-morpho-flexionnel	7,11 (1,61) N = 45	6,48 (1,88) N = 27	7,19 (1,80) N = 32	6,97 (1,75) t(103) = 5,65 ; p < .0001
Implicite grapho-syntaxique	6,49 (1,49) N = 45	6,70 (1,86) N = 27	7,41 (2,05) N = 32	6,83 (1,80) t(103) = 4,69 ; p < .0001
Explicite graphophonologique	11,38 ^p (1,05) N = 47	11,28 ^p (1,25) N = 29	11,79 ^p (0,59) N = 34	11,48 ^p (1,01) t(109) = 56,85 ; p < .0001
Explicite grapho-morpho-dérivationnel	8,40 (1,87) N = 47	9,00 (1,58) N = 30	10,30 ^p (1,83) N = 33	9,14 (1,94) t(109) = 16,94 ; p < .0001
Explicite grapho-morpho-flexionnel	9,06 (2,44) N = 47	9,70 (2,34) N = 30	11,32 ^p (1,01) N = 34	9,93 (2,2) t(110) = 18,19 ; p < .0001
Explicite grapho-syntaxique	8,04 (2,26) N = 47	9,33 (1,84) N = 30	10,74 ^p (1,36) N = 34	9,22 (2,21) t(110) = 15,34 ; p < .0001

^p Scores plafonds.

En gras : les scores pouvant être retenus pour effectuer des analyses inférentielles.

Les valeurs des tests t de comparaison à la valeur 6 obtenues pour l'ensemble des niveaux scolaires rendent compte de la présence de réponses ne se distinguant pas significativement du hasard à l'épreuve implicite grapho-morpho-dérivationnelle. Cependant, les scores des élèves de 4^e année sont significativement inférieures au niveau du hasard ($p < .00001$), alors que celles des

deux autres niveaux ne s'en distinguent pas. Lors de la réalisation de cette tâche, les élèves de 4^e année seraient sensibles à autre chose qu'aux aspects morpho-dérivationnels des mots alors que les élèves des autres niveaux semblent répondre au hasard. De telles observations ne permettent de toute façon pas de conserver cette épreuve pour la suite des analyses.

En revanche, les résultats aux épreuves implicites d'intrus graphophonologique, grapho-morpho-flexionnel et grapho-syntaxique sont suffisamment élevés (moyenne générale entre 57 et 66 % de bonnes réponses) pour se différencier positivement des réponses données au hasard ($p < .0001$), ce qui justifie qu'une analyse comparative soit menée sur ces résultats.

Ainsi, il semble que le paradigme de détection implicite d'intrus, s'attachant ici à diverses caractéristiques linguistiques, parvienne à faire émerger, pour l'ensemble des niveaux scolaires, une sensibilité aux aspects graphophonologiques, grapho-morpho-flexionnels et grapho-syntaxiques de la langue portugaise mais pas aux aspects grapho-morpho-dérivationnels.

Lorsque la consigne est explicite, les épreuves utilisant ce même paradigme obtiennent des taux de réussite élevés (*cf.* tableau 30). Les réponses sont d'ailleurs systématiquement supérieures au niveau du hasard ($p < .001$).

Cependant, les scores de réussite aux épreuves graphophonologiques explicites pour les trois niveaux scolaires et à l'ensemble des épreuves pour le niveau 5^e année font apparaître des effets plafonds. Pour l'épreuve graphophonologique, entre 83 et 97 % des élèves répondent correctement à au moins 11 items sur 12. Ces résultats indiquent que dès la 3^e année, la discrimination consciente des éléments graphophonologiques du portugais est parfaitement maîtrisée. L'apparition de ces scores plafonds pour les trois niveaux scolaires nous obligera à supprimer cette épreuve des analyses statistiques suivantes.

En 7^e année, les scores moyens obtenus aux épreuves explicites grapho-morpho-dérivationnelle, grapho-morpho-flexionnelle et grapho-syntaxique dépassent systématiquement 85 % de réussite. A ces trois épreuves, une majorité d'élèves répondent correctement à au moins 11 items sur 12 [respectivement 53 % (+ 32 % avec un score de 10/12), 82 % et 62 % (+ 24% avec un score de 10/12)]. Cela signifie qu'à ce niveau scolaire, le paradigme de détection explicite d'intrus est trop facile pour continuer à être un mode d'évaluation des connaissances explicites discriminant. De ce fait, ces épreuves ne pourront être prises en considération dans les analyses statistiques ultérieures à ce niveau scolaire.

L'impossibilité de mener des analyses statistiques sur les résultats à l'épreuve implicite grapho-morpho-dérivationnelle, à l'épreuve explicite graphophonologique ainsi que sur les résultats des

7^e année aux épreuves explicites grapho-sémantiques nous obligera à réduire les comparaisons effectuées.

Ainsi, sur les épreuves implicites, à l'exclusion de l'épreuve grapho-morpho-dérivationnelle, une analyse pour plan à deux facteurs (niveau scolaire {3^e, 5^e et 7^e année} et domaine linguistique {graphophonologique, grapho-morpho-flexionnel, grapho-syntaxique}) sera effectuée.

Sur les épreuves explicites grapho-sémantiques, une analyse pour plan à deux facteurs (niveau scolaire {3^e vs. 5^e} et domaine grapho-sémantique {morpho-dérivationnel, morpho-flexionnel, syntaxique}) sera menée.

Sur les résultats aux épreuves grapho-morpho-flexionnelles et grapho-syntaxiques implicites et explicites les effets du niveau scolaire sur les deux premiers niveaux (3^e, 5^e année) du niveau cognitif (implicite vs. explicite) et du domaine linguistique (grapho-morpho-flexionnel vs. grapho-syntaxique) seront examinés par le biais d'une analyse à 3 facteurs.

Pour terminer, une analyse sur les connaissances graphophonologiques implicites seulement, portant sur les facteurs « niveau scolaire » et « unité phonologique » sera réalisée.

2.2.3.2 *Epreuves implicites de détection d'intrus : effets du niveau scolaire et du domaine linguistique*

Les résultats comparés lors de l'analyse pour plan à deux facteurs (niveau scolaire {3^e, 5^e, 7^e} et domaine linguistique {graphophonologique, grapho-morpho-flexionnel, grapho-syntaxique}) menée sur les épreuves implicites sélectionnées (*i.e.* pour lesquelles aucun effet du hasard n'est observé) sont présentés dans le tableau 31.

Tableau 31 : Performances moyennes de réussite sur 12 (et écart-types) obtenues aux 3 épreuves implicites de détection d'intrus sélectionnées, en 3^e, 5^e et 7^e année et tests t de comparaison à la valeur 6 (niveau du hasard).

	3 ^e année N = 45	5 ^e année N = 27	7 ^e année N = 32	Tous niveaux scolaires N =104
Implicite graphophonologique	8,44 (1,90) t(44) = 8,27, p<.0001	6,89 (2,08) t(26) = 2,22, p <.05	8,22 (2,20) t(31) = 5,72, p<.0001	7,97 (2,16)
Implicite grapho-morpho-flexionnel	7,11 (1,61) t(44) = 4,62, p<.0001	6,48 (1,88) t(26) = 1,32, ns	7,19 (1,80) t(31) = 3,72, p<.001	6,97 (1,75)
Implicite grapho-syntaxique	6,49 (1,49) t(44) = 2,21, p < .04	6,70 (1,86) t(26) = 1,97, ns	7,41 (2,05) t(31) = 3,89, p<.001	6,83 (1,80)
Moyenne	7,35 (0,96)	6,69 (1,14)	7,60 (1,30)	

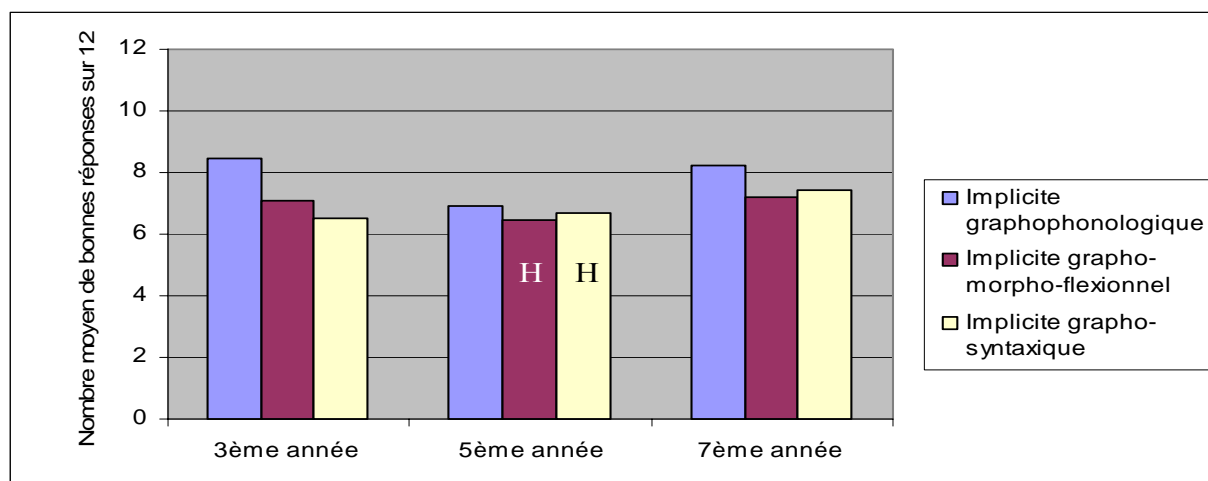
Cette analyse fait apparaître :

- un effet significatif du niveau scolaire [F (2,101) = 5,12, p < .008]. Les élève de 5^e année obtiennent des résultats plus faibles que les élèves de 3^e année (p < .02) et de 7^e année (p < .004) ;

- un effet significatif du domaine linguistique [$F(2,202) = 9,01, p < .0002$] : l'épreuve graphophonologique est mieux résolue que l'épreuve grapho-morpho-flexionnelle ($p < .0001$) et l'épreuve grapho-syntaxique ($p < .0001$) ;
- et une interaction tendancielle niveau scolaire * domaine linguistique [$F(4,202) = 2,246, p < .07$].

On remarque toutefois que les réponses des 5^e année aux épreuves grapho-morpho-flexionnelles et grapho-syntaxiques sont les seules à ne pas se différencier significativement du hasard (*cf.* tableau 31).

La figure 10 reprend les résultats obtenus aux trois épreuves implicites sélectionnées en fonction du niveau scolaire.



H : Scores non différents du hasard

Figure 10 : Performances moyennes (sur 12) obtenues aux épreuves implicites de détection d'intrus graphophonologique, grapho-morpho-flexionnelle et grapho-syntaxique en 3e, 5e et 7e année.

L'observation essentielle se dégageant de cette analyse a trait au fait que les connaissances les plus développées en portugais semblent être celles des aspects graphophonologiques. D'autre part, il est inattendu de voir apparaître une chute de la sensibilité en 5^e année, niveau pour lequel seules les réponses à l'épreuve implicite graphophonologique se différencient du hasard.

2.2.3.3 *Epreuves explicites de détection d'intrus grapho-sémantique : effets du niveau scolaire et du domaine grapho-sémantique*

Une analyse pour plan à deux facteurs (niveau scolaire {3^e vs. 5^e} et domaine grapho-sémantique {morpho-dérivationnel, morpho-flexionnel, syntaxique}) a été menée sur les résultats des deux premiers niveaux scolaires pour lesquels aucun effet plafond n'est observé. Les résultats comparés lors de cette analyse sont présentés dans le tableau 32.

Tableau 32 : Performances moyennes de réussite sur 12 (et écart-types) obtenues aux épreuves explicites de détection d'intrus grapho-sémantique en 3^e et 5^e année et tests t de comparaison à la valeur 6 (niveau du hasard)

	3 ^e année N = 47	5 ^e année N = 30	Moyenne deux niveaux scolaires N = 77
Explicite grapho-morpho-dérivationnel	8,40 (1,87) t (46) = 8,80 ; p<.0001	9,00 (1,58) t (29) = 10,43 ; p<.0001	8,64 (1,78)
Explicite grapho-morpho-flexionnel	9,06 (2,44) t (46) = 8,59 ; p<.0001	9,70 (2,34) t (29) = 8,67 ; p<.0001	9,31 (2,41)
Explicite grapho-syntaxique	8,04 (2,26) t (46) = 6,21 ; p<.0001	9,33 (1,84) t (29) = 9,90 ; p<.0001	8,55 (2,19)
Moyenne explicite grapho-sémantique	8,50 (1,47)	9,34 (1,28)	8,83 (1,45)

Les résultats mettent en évidence :

- Un effet simple du niveau scolaire [$F(1,75) = 6,62 ; p < .05$] : les résultats des 5^e année sont globalement supérieurs à ceux des 3^e année.
- Un effet simple du domaine grapho-sémantique [$F(2,150) = 3,10 ; p < .05$] : de façon générale, les résultats obtenus à l'épreuve grapho-morpho-flexionnelle sont supérieurs à ceux obtenus aux épreuves grapho-morpho-dérivationnelle et grapho-syntaxique ($p < .05$). Ceux obtenus à ces deux dernières épreuves ne sont pas significativement différents.
- Une absence d'interaction significative niveau scolaire * domaine grapho-sémantique [$F(2,150) < 1$].

Il est également à souligner que quelle que soit l'épreuve, les résultats se différencient positivement du hasard pour les deux niveaux scolaires ($p < .0001$, cf. tableau 32).

La figure 11 reprend les résultats obtenus aux trois épreuves explicites de détection d'intrus grapho-sémantique en fonction du niveau scolaire, en incluant les scores plafonds des élèves de 7^e année.

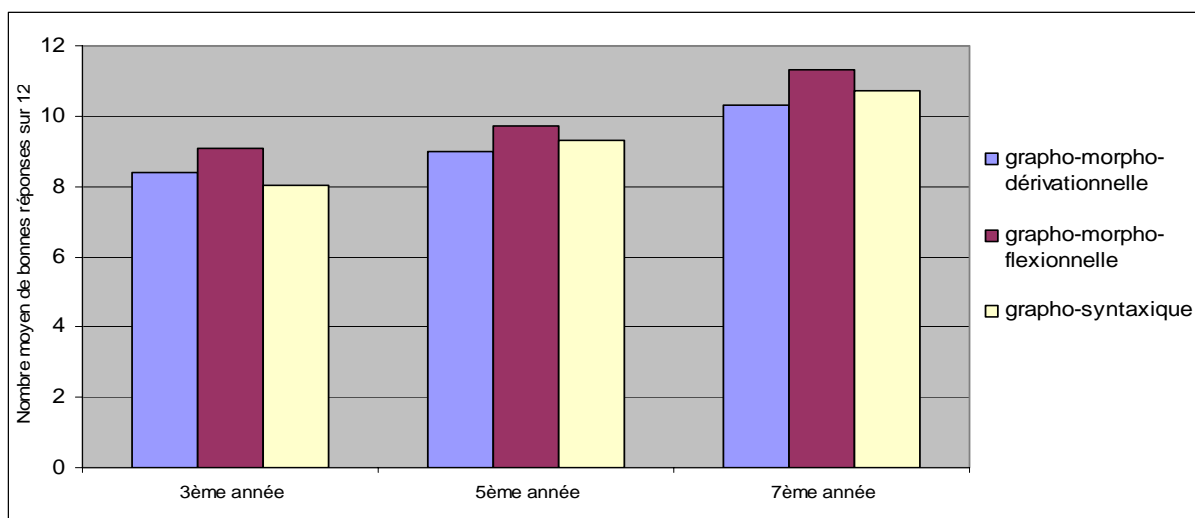


Figure 11 : Performances moyennes (sur 12) obtenues aux épreuves explicites de détection d'intrus grapho-sémantique en 3^e, 5^e et 7^e année

A l'issue de ces analyses, nous pouvons constater que les connaissances grapho-sémantiques (dérivationnelles, flexionnelles et syntaxiques) évaluées par les épreuves explicites de détection d'intrus sont non négligeables dès la 3^e (entre 67,02 et 75,53 % de moyenne de bonnes réponses) et augmentent avec l'avancée dans les apprentissages pour plafonner en 7^e année (entre 85,86 et 98,28 % de moyenne de bonnes réponses). De plus, la détection d'intrus qui apparaît la plus facile est celle qui porte sur la structure morpho-flexionnelle des mots portugais.

2.2.3.4 *Epreuve implicite et explicite de détection d'intrus grapho-morpho-flexionnel et grapho-syntaxique : effets du niveau scolaire du niveau cognitif et du domaine syntaxique*

Le tableau 33 rapporte les résultats obtenus aux épreuves implicites et explicites grapho-morpho-flexionnelle et grapho-syntaxique en fonction du niveau scolaire.

Tableau 33 : Performances moyennes de réussite sur 12 (et écart-types) obtenues aux épreuves implicites et explicites de détection d'intrus grapho-morpho-flexionnel et grapho-syntaxique en 3^e et 4^e année.

Dimensions / degré de contrôle		3 ^e année N = 45	5 ^e année N = 27	Deux niveaux scolaires N = 72
Grapho-morpho-flexionnelle	Implicite	7,11 (1,61)	6,48 ^H (1,89)	6,81 (1,73)
	Explicite	9,00 (2,40)	9,75 (2,45)	9,35 (2,42)
Grapho-syntaxique	Implicite	6,49 (1,49)	6,70 ^H (1,86)	6,59 (1,65)
	Explicite	8,05 (2,33)	9,32 (1,74)	8,54 (2,21)
Moyenne	Implicite	6,80 (1,18)	6,54 (1,22)	6,70 (1,21)
	Explicite	8,52 (1,76)	9,52 (1,59)	9,03 (1, 87)
Moyenne Totale		7,67 (1,01)	8,03 (1,19)	

L'analyse pour plan à 3 facteurs (niveau scolaire {3^e, 5^e}, domaine linguistique {grapho-morpho-flexionnel, grapho-syntaxique} et degré de contrôle {implicite vs. explicite}), montre que :

- l'effet du degré de contrôle véhiculé par la consigne est significatif [$F(1,70) = 89,82, p < .000001$], les épreuves explicites donnant lieu à des scores plus élevés que les épreuves implicites (cf. tableau 33) ;
- l'effet du domaine linguistique est significatif [$F(1,70) = 4,54, p < .05$] puisque, de façon générale, les scores aux épreuves grapho-morpho-flexionnelles sont supérieurs aux scores obtenus aux épreuves grapho-syntaxiques ;
- l'effet du niveau scolaire n'est pas significatif [$F(1,70) = 1,99, ns$] ;
- mais l'interaction degré de contrôle * niveau scolaire est significative [$F(1,70) = 6,81, p < .01$] : si les performances des élèves augmentent entre la 3^e et la 5^e année pour les épreuves explicites ($p < .0001$), les résultats ne sont pas significativement différents pour ce qui concerne les épreuves implicites. Rappelons toutefois les observations précédentes de scores non significativement différents du hasard aux épreuves grapho-morpho-flexionnelle et grapho-syntaxique implicites pour les 5^e année seulement¹⁵³.

Les interactions domaine linguistique * niveau scolaire, degré de contrôle * domaine linguistique et degré de contrôle * domaine linguistique * niveau scolaire sont non significatives [respectivement $F(1,70) = 2,07, ns$; $F(1,70) < 1$ et $F(1,70) < 1$].

On retrouve ici, ce qui est apparu dans l'analyse précédente, soit des connaissances grapho-morpho-flexionnelles mieux développées que les connaissances grapho-syntaxiques. Sans surprise, les épreuves dont la consigne est explicite sont mieux résolues que les épreuves sans consigne. L'implication de l'attention influence positivement les résultats et permet leur augmentation avec le niveau, si bien qu'en 7^e année, ces tâches deviennent trop faciles. Le fait de ne pas pouvoir s'assurer que les réponses aux deux épreuves implicites grapho-morpho-flexionnelle et grapho-syntaxique des élèves de 5^e année ne sont pas liées au hasard, nous obligera à éliminer ces données dans la suite des analyses, pour ce groupe seulement.

¹⁵³ Compte tenu du fait que les résultats des tests t ont été précédemment présentés pour ces groupes en ce qui concerne les résultats aux épreuves implicites, nous avons jugé inutile d'alourdir le tableau en les ajoutant. De même en ce qui concerne les épreuves explicites, bien que le tableau 7 rapporte les résultats d'une population légèrement plus importante de 3^e et 5^e année (ce phénomène s'explique par l'absence de 2 sujets de 3^e année et de 3 sujets de 5^e année les jours de passations des épreuves implicites), nous avons considéré que les scores étaient suffisamment proches pour ne pas modifier la différence par rapport au niveau du hasard.

2.2.3.5 *Epreuve implicite de détection d'intrus graphophonologique : effets du niveau scolaire et de l'unité graphophonologique*

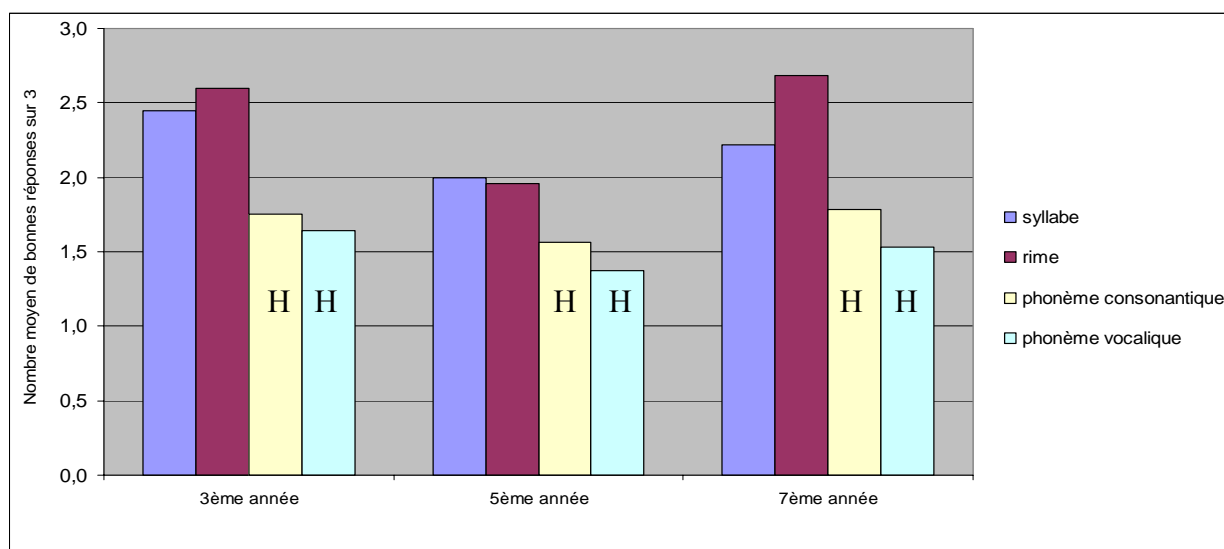
Les résultats obtenus à l'épreuve de détection implicite d'intrus graphophonologique en fonction du niveau scolaire et de l'unité phonologique ainsi que les comparaisons au hasard figurent dans le tableau 34.

Tableau 34 : Performances moyennes de réussite (et écart-types) obtenues à l'épreuve implicite de détection d'intrus graphophonologique en fonction du type d'items et du niveau scolaire et tests t de comparaison à la valeur 1/2 (niveau du hasard).

	3 ^e année N = 45	5 ^e année N = 27	7 ^e année N = 32	Tous niveaux scolaires N = 104
Score Rime (sur 3)	2,60 (0,75) t(44) = 9,83 ; p<.0001	1,96 (0,94) t(26) = 2,56 ; p<.05	2,69 (0,54) t(31) = 12,55 ; p<.0001	2,46 (0,93) t(103) = 12,26 ; p<.0001
Score Syllabe (sur 3)	2,44 (0,72) t(44) = 8,74 ; p<.0001	2,00 (1,04) t(26) = 2,50 ; p<.05	2,22 (0,87) t(31) = 4,67 ; p<.0001	2,26 (0,87) t(103) = 8,90 ; p<.0001
Score Phonème consonantique (sur 3)	1,76 (0,93) t(44) = 1,84 ; ns	1,56 (0,85) t(26) = 0,34 ; ns	1,78 (1,01) t(31) = 1,58 ; ns	1,71 (0,93) t(103) = 2,32 ; p<.05
Score Phonème vocalique (sur 3)	1,64 (0,91) t(44) = 1,07 ; ns	1,37 (0,74) t(26) = -0,91 ; ns	1,53 (0,98) t(31) = 0,18 ; ns	1,54 (0,89) t(103) = 0,44 ; ns
Score total (sur 12)	8,44 (1,98) t(44) = 8,27 ; p<.0001	6,88 (2,08) t(26) = 2,22 ; p<.05	8,22 (2,20) t(31) = 5,72 ; p<.0001	7,97(2,16)

Sur l'ensemble de cette épreuve, l'analyse du facteur « niveau scolaire » fait apparaître un effet significatif de ce facteur [$F(2,101) = 5,07 ; p < .001$] : les élèves de 5^e année obtiennent des résultats plus faibles que les élèves de 3^e et de 7^e année ($p < .01$ et $p < .05$). La différence de scores entre les 3^e et les 7^e année n'est en revanche pas significative. De plus, les comparaisons au niveau du hasard font apparaître des résultats supérieurs à la valeur 6 pour tous les niveaux scolaires (cf. tableau 34). Ainsi, plusieurs tests non paramétriques seront menés dans l'objectif de déterminer les unités phonologiques auxquelles les enfants lusophones sont les plus sensibles.

Les résultats obtenus aux items « rime », « syllabe », « phonème consonantique » et « phonème vocalique » peuvent être visualisés à l'aide de la figure 1.



H : Scores non différents du hasard

Figure 12 : Performances moyennes (sur 3) obtenues à l'épreuve implicite de détection d'intrus graphophonologique en 3^e, 4^e et 5^e année en fonction de l'unité phonologique étudiée.

Les résultats au test de Friedman révèlent un effet significatif de l'unité phonologique sur l'ensemble des groupes [$\chi^2(3, N = 104) = 71,67, p < .0001$]. Les comparaisons deux à deux (test de Wilcoxon) permettent d'identifier que les résultats aux items portant sur la rime et la syllabe sont supérieurs à ceux portant sur le phonème consonantique ($T = 235, z = 5,65 ; p < .00001$ et $T = 448, z = 4,20 ; p < .0001$) et vocalique ($T = 111, z = 6,56 ; p < .00001$ et $T = 545, z = 5,35 ; p < .00001$). Les autres comparaisons ne font pas apparaître de différences significatives. Il est également à souligner que les seuls résultats qui ne se distinguent pas significativement du niveau du hasard sont ceux obtenus aux items « phonème vocalique » (cf. tableau 34).

En 3^e année, on retrouve l'effet de l'unité phonologique [$\chi^2(3, N = 28) = 8,26 ; p < .05$] : les résultats aux items portant sur la rime et la syllabe sont supérieurs à ceux portant sur le phonème consonantique ($T = 51, z = 3,98 ; p < .00001$ et $T = 76, z = 3,22 ; p < .01$) et vocalique ($T = 10, z = 4,66 ; p < .00001$ et $T = 117, z = 3,81 ; p < .001$). Les autres comparaisons ne font pas apparaître de différences significatives.

En 5^e année, l'effet de l'unité phonologique est également significatif [$\chi^2(3, N = 27) = 8,96, p < .03$]. Les comparaisons deux à deux permettent d'identifier que les résultats aux items « rime » et « syllabe » sont significativement supérieurs à ceux aux items « phonème vocalique » ($T = 15, z = 2,56, p < .05$ et $T = 53,5, z = 2,37, p < .05$) mais pas aux items « phonème consonantique » ($T = 35, z = 1,71 ; ns$ et $T = 33, z = 1,81 ; ns$). Les autres comparaisons ne font pas apparaître de différences significatives.

En 7^e année, l'effet de l'unité phonologique [$\chi^2(3, N = 32) = 29,38 p=.0001$] est lié à un effet plafond des résultats aux items « rime » (89,67 % de bonnes réponses) qui sont de ce fait supérieurs à ceux obtenus aux autres items (items « syllabe », $T = 24, z = 2,28 ; p < .05$ et items « phonème consonantique », $T = 0,00, z = 3,82 ; p < .0002$ et items « phonème vocalique », $T = 9, z = 3,92 ; p < .00001$). Les items « syllabe » obtiennent de plus des scores supérieurs aux items « phonème vocalique » ($T = 30, z = 2,97 ; p < .01$). Les autres comparaisons ne font pas apparaître de différences significatives.

Il est également à souligner que, quel que soit le niveau, les résultats supérieurs au niveau du hasard concernent la rime et la syllabe mais pas les unités phonémiques (*cf.* tableau 34).

Ainsi, on retrouve pour les trois niveaux scolaires, une supériorité des scores aux items « rime » et « syllabe » par rapport aux items « phonème vocalique ». L'utilisation du test de Kruskal-Wallis indique que les différences de score entre les items « rime » et « phonème vocalique » et entre les items « syllabe » et « phonème vocalique » ne sont cependant pas plus marquées pour un niveau scolaire qu'un autre [$H(2, N = 105) = 5,44, p < .07$ et $H(2, N = 105) = 0,57, ns$].

La supériorité des scores aux items « rime » par rapport aux items « phonème consonantique » est significative en 3^e comme en 7^e année. Cependant, cette différence n'est pas significativement plus marquée en 3^e qu'en 7^e année (test de Mann-Whitney : $U = 705,5, z = -0,15, ns$).

Enfin, la supériorité des scores aux items « rime » par rapport aux items « syllabe » n'est visible qu'en 7^e année, ces performances atteignant de surcroît un niveau plafond.

Globalement, ces analyses permettent de constater que dès la 3^e année, les élèves lusophones repèrent davantage la rime et la syllabe terminale de mots écrits. En 7^e année, la perception de la rime intrasyllabique semble particulièrement développée ce qui va dans le sens de l'importance de la sensibilité à cette unité en portugais.

De plus compte tenu des faits suivants, aucun groupe d'items ne sera exclu des analyses à venir, bien que les résultats aux items « phonème consonantique » et « phonème vocalique » ne soient pas significativement différents du hasard :

- pour l'ensemble des niveaux scolaires, les coefficients de corrélation sont majoritairement significatifs entre les scores à ces quatre groupes d'items¹⁵⁴,

¹⁵⁴ Coefficients de corrélation de rangs de Spearman significatifs à $p < .05$: entre « voyelle » et « syllabe » $r_s = .21$, - « rime » $r_s = .26$, - « consonne » $r_s = .26$, entre « rime » et « consonne » $r_s = .22$.

- pour chaque niveau scolaire, les scores totaux à l'épreuve implicite de détection d'intrus graphophonologique sont normalement distribués¹⁵⁵ et significativement supérieurs au hasard (cf. tableau 34).

2.2.4 Epreuve d'identification de la construction morphologique de pseudo-mots : effet du niveau scolaire, de la nature de la construction morphologique et de la nature des règles morphologiques

Le tableau 35 présente les résultats à l'épreuve d'identification de la construction morphologique de pseudo-mots (moyennes, écart-types, minima, maxima) pour chaque niveau scolaire¹⁵⁶.

Tableau 35 : Performances moyennes sur 24, écart-types, maxima, minima obtenus à l'épreuve d'identification de la construction morphologique de pseudo-mots en fonction du niveau scolaire.

	Moyenne	Ecart-type	Minimum	Maximum
3e année N = 44	12,05	5,06	5	23
5ème année N = 29	16,90	5,42	6	23
7ème année N = 33	18,45	5,67	4	23

Les items de cette épreuve se distinguant quant à la nature de la construction des pseudo-mots et des règles morphologiques étudiées, une analyse de variance à trois facteurs (niveau scolaire {3^e année, 4^e année, 5^e année}, construction morphologique {préfixe, suffixe, schème} et règle morphologique {morpho-dérivationnelle vs. morpho-flexionnelle}) a été menée. Les résultats obtenus en fonction de la nature de la construction morphologique et du niveau scolaire figurent dans le tableau 36, ceux obtenus en fonction du type de règle morphologique et du niveau scolaire dans le tableau 37 et ceux obtenus pour l'ensemble des groupes en fonction de la construction et de la règle morphologique dans le tableau 38¹⁵⁷.

¹⁵⁵ Tests de Kormogorov-Smirnov non significatifs : $d = 0,12$, $p > .10$ en 3^{ème} ; $d=0,15$, $p > .20$ en 5^{ème} et $d=0,16$, $p > .20$ en 7^{ème} année.

¹⁵⁶ Pour cette épreuve, les réponses sont à choisir entre trois propositions ce qui rend là encore possible les choix aléatoires. Cependant, il n'a pas été nécessaire de conduire des tests t pour échantillon unique pour s'assurer que les réponses ne sont pas liées au hasard (1/3 de bonnes réponses) car les scores s'élèvent systématiquement bien au dessus de 33 % de bonnes réponses.

¹⁵⁷ L'interaction « construction morphologique » * « règle morphologique » * « niveau scolaire » n'étant pas significative (cf. infra), nous avons préféré ne pas présenter les résultats distinguant simultanément ces 3 facteurs.

Tableau 36 : Performances moyennes sur 24, écart-types, maxima, minima obtenus à l'épreuve d'identification de la construction morphologique de pseudo-mots en fonction du niveau scolaire.

Construction morphologique/8 et %	3 ^e année N = 44	5 ^e me année N = 29	7 ^e me année N = 33	Tous niveaux scolaires N = 106
Préfixe	4,16 (1,89) 51,99 %	6,00 (2,24) 75,00 %	6,39 (2,28) 79,92 %	5,36 (2,33) 66,98 %
Suffixe	3,98 (1,72) 49,72 %	5,59 (1,30) 69,83 %	5,85 (1,62) 73,11 %	5,00 (1,79) 62,50 %
Schème	3,91 (2,22) 48,86 %	5,31 (2,36) 66,38 %	6,21 (2,10) 77,65 %	5,01 (2,42) 62,62 %

Tableau 37 : Performances moyennes (écart-types) et en pourcentage, obtenues à l'épreuve d'identification de la construction morphologique de pseudo-mots en 3^e, 5^e et 7^e année, en fonction du type de règle morphologique.

Règle morphologique /12 et %	3 ^e année N = 44	5 ^e me année N = 29	7 ^e me année N = 33	Tous niveaux scolaires N = 106
Morpho-dérivationnelle	6,20 (2,72) 51,70 %	7,97 (2,73) 66,38 %	9,21 (2,51) 76,77 %	7,62 (2,94) 63,52 %
Morpho-flexionnelle	5,84 (2,92) 48,67 %	8,93 (3,05) 74,43 %	9,24 (3,35) 77,02 %	7,75 (3,46) 64,54 %

Tableau 38 : Performances moyennes sur 4 (écart-types) et en pourcentage, obtenues à l'épreuve d'identification de la construction morphologique de pseudo-mots pour l'ensemble des groupes (N = 106), en fonction de la nature de la construction et du type de règle morphologique

Règle morphologique	Construction morphologique		
	préfixe	suffixe	schème
Morpho-dérivationnelle	2,76 (1,34) 69,10 %	2,38 (0,81) 59,43 %	2,48 (1,35) 62,03 %
Morpho-flexionnelle	2,59 (1,29) 64,86 %	2,62 (1,28) 65,57 %	2,53 (1,36) 63,21 %

Les résultats de l'analyse à 3 facteurs mettent en évidence :

- Un effet principal du niveau scolaire [$F(2, 103) = 15,14, p < .00001$] : les 3^e année obtiennent en moyenne des résultats plus faibles que les 5^e année ($p < .01$) et les 7^e année ($p < .001$). La différence entre les 5^e et 7^e année n'est en revanche pas significative (cf. tableau 35).
- Un effet principal « construction morphologique » [$F(2,206) = 4,13, p < .05$] mais pas d'interaction « niveau scolaire » * « construction morphologique » [$F(4,206) < 1$] : Pour tous les niveaux scolaires, les items « préfixe » sont en moyenne moins mieux résolus que les items « suffixe » et « schème » ($p < .05$), la différence de scores entre ces deux derniers types n'étant pas significative (cf. tableau 36).
- Une absence de l'effet « règle morphologique » [$F(1,103) < 1$] mais une interaction « niveau scolaire » * « règle morphologique » [$F(2,103) = 3,37, p < .05$] : pour les items portant sur les règles morpho-flexionnelles, les scores des 5^e année sont significativement supérieurs à ceux des

3^e année ($p < .01$) mais pas pour les items portant sur les règles morpho-dérivationnelles, si bien qu'en 5^e année les items « morpho-flexionnels » sont mieux résolus que les items « morpho-dérivationnels » ($p < .01$) alors que cette différence n'est pas significative sur les deux autres niveaux scolaires (*cf.* tableau 37).

- Une interaction « construction morphologique » * « règle morphologique » [$F(2,206) = 3,62$, $p < .05$] : lorsque les règles réfèrent aux aspects morpho-dérivationnels, les items « préfixe » sont significativement mieux résolus que les items « suffixe » ($p < .01$) mais pas lorsque les règles réfèrent aux aspects morpho-flexionnels (*cf.* tableau 38).

- Une absence d'interaction « construction morphologique » * « règle morphologique » * « niveau scolaire » [$F(4,206) < 1$].

A l'issue de cette analyse, il est possible de relever que :

1) les performances à cette épreuve augmentent surtout entre la 3^e et la 5^e année, et principalement lorsque les règles s'apparentent aux aspects morpho-flexionnels. Ainsi, en 5^e année seulement, la construction morphologique de pseudo-mots est mieux identifiée lorsqu'elle porte sur des aspects morpho-dérivationnels que morpho-flexionnels.

2) la construction morphologique des pseudo-mots préfixés est mieux identifiée que celle des suffixés, en particulier dans le domaine morpho-dérivationnel, et que celle correspondant à une structure de type schème.

2.2.5 Synthèse des résultats et récapitulatif des épreuves sélectionnées pour les analyses suivantes

A l'issue des analyses précédentes, certaines épreuves implicites ou explicites, évaluant les connaissances graphophonologiques, grapho-morpho-dérivationnelles, grapho-morpho-flexionnelles et grapho-syntaxiques ont été éliminées et d'autres sauvegardées.

Le tableau 39 permet d'avoir un aperçu synoptique des épreuves incluses dans les différentes analyses qui vont suivre.

Tableau 39 : Récapitulatif des épreuves sélectionnées pour les analyses suivantes à chaque niveau scolaire et synthèse des principaux résultats obtenus (en %)¹

Epreuves	3 ^e année	5 ^e année	7 ^e année
Det. Imp. G-phonologique	70,33 % rime et syllabe > consonne - voyelle	57,42 % rime et syllabe > voyelle	61,14 % rime > syllabe > voyelle - consonne
Det. Imp. G-morpho-dériv.	Supprimée (réponses aléatoires)		
Det. Imp. G-morpho-flex	59,25 %	Supprimée (réponses aléatoires)	59,92 %
Det. Imp. G-syntaxique	54,08 %	Supprimée (réponses aléatoires)	61,75 %
Det. Exp. G-phonologique	Supprimée (scores plafonds)		
Det. Exp. G-syntaxique	67 %	77,75 %	Supprimée (scores plafonds)
Det. Exp. G-morpho-dériv.	70 %	75 %	Supprimée (scores plafonds)
Det. Exp. G-morpho-flex.	75,50 %	80,83 %	Supprimée (scores plafonds)
Ident. Const. Morpho. PM	50,21 %	70,42 % M-Flex. > M- dériv.	76,88 %
	sur règles morpho-dériv. : "préfixe" > "suffixe" et "schème"		

¹ L'ordre de présentation des épreuves suit celui choisi précédemment en arabe, de façon à mettre les deux tableaux en relation.

2.3 Analyse du lien entre les scores aux différentes épreuves sélectionnées

En 3^e année, les épreuves retenues pour analyser le lien entre les différents types de connaissances sont les épreuves implicites de détection d'intrus graphophonologique, grapho-morpho-flexionnel et grapho-syntaxique, les épreuves explicites de détection d'intrus grapho-morpho-dérivationnel, grapho-morpho-flexionnel, grapho-syntaxique et l'épreuve d'identification de la construction morphologique de pseudo-mots.

En 5^e année, seule l'épreuve de détection implicite d'intrus est conservée ainsi que les épreuves explicites grapho-morpho-dérivationnel, grapho-morpho-flexionnel, grapho-syntaxique. Enfin, l'épreuve d'identification de la construction morphologique de pseudo-mots sera également prise en compte.

En 7^e année, seules les épreuves de détection implicite d'intrus graphophonologique, grapho-morpho-flexionnel, grapho-syntaxique et d'identification de la construction morphologique de pseudo-mots obtiennent des résultats analysables.

2.3.1 Analyses de corrélation

Le tableau 40 précise les coefficients de corrélations de Bravais Pearson entre les scores à ces différentes épreuves pour chaque niveau scolaire.

Tableau 40 : Coefficients de corrélation de Bravais Pearson entre les épreuves sélectionnées en 3^e, 5^e et 7^e année.

3 ^e année N = 43	Dét. Imp. G-phono.	Dét. Imp. G- morpho- flex.	Dét. Imp. G-Synt.	Dét. exp. G- morpho- dériv.	Dét. Exp. G- morpho- flex.	Dét. Exp. G-synt.	Ident. Const. morpho.
Dét. Imp. G- phonologique	----	.00	-.21	.06	-.11	.06	.29 ^t
Dét. Imp. G-morpho-flex.		—	.22	-.04	.10	-.27	.01
Dét. Imp G-Syntaxique.			—	.16	.12	-.25	-.27
Dét. Exp. G-morpho- dériv.				----	.31*	.05	.23
Dét. Exp. G-morpho-flex.					----	.13	.34*
Dét. Exp. G- syntaxique.						----	.23

5 ^e année N = 26	Dét. Imp. G-phono.	Dét. exp. G- morpho- dériv.	Dét. Exp. G- morpho- flex.	Dét. Exp. G-synt.	Ident. Const. morpho.
Dét. Imp. G- phonologique	----	.01	.05	-.04	-.05
Dét. Exp. G-morpho- dériv.		----	.37 ^t	-.14	.39 ^t
Dét. Exp. G-morpho-flex.			----	.15	.19
Dét Exp G- syntaxique				----	.06

7 ^e année N = 32	Dét. Imp. G-phono.	Dét. Imp. G- morpho- flex.	Dét. Imp. G-Synt.	Ident. Const. morpho.
Dét. Imp. G- phonologique	—	.08	-.07	.19
Dét. Imp. G- morpho-flex		—	.40*	.24
Dét. Imp. G- syntaxique			—	.04

** r significatif à $p < .01$ * à $p < .05$ ^t tendanciel à $p < .07$

En 3^e année, la matrice de corrélation (cf. tableau 40) permet de constater que les scores à l'épreuve d'identification de la construction morphologique de pseudo-mots sont corrélés avec les scores à l'épreuve de détection explicite d'intrus grapho-morpho-flexionnel ($r = .34$, $p < .05$). Ces derniers le sont également avec ceux obtenus à l'épreuve de détection explicite d'intrus grapho-morpho-dérivationnelle ($r = .31$, $p < .05$). Enfin, la corrélation entre les scores à l'épreuve d'identification de la construction morphologique de pseudo-mots et à l'épreuve de détection implicite graphophonologique n'est que tendancielle ($r = .29$, $p < .06$). Les corrélations entre les autres performances ne sont pas significatives.

En 5^e année, aucune corrélation n'est significative, mais deux coefficients assez élevés et très tendancielles peuvent être relevés : celui concernant la corrélation entre les scores à l'épreuve

d'identification de la construction morphologique de pseudo mots et les scores à l'épreuve grapho-morpho-dérivationnelle. ($r = .39, p < .05$) et celui concernant la corrélation entre les scores à l'épreuve grapho-morpho-dérivationnelle et les scores à l'épreuve grapho-morpho-flexionnelle. ($r = .37, p < .06$) (cf. tableau 40).

En 7^e l'unique corrélation significative concerne les épreuves implicites grapho-morpho-flexionnelle et grapho-syntaxique ($r = .40, p < .05$).

Ces résultats mettent en évidence les points suivants :

En 3^e année, les connaissances morpho-dérivationnelles seraient liées à leurs homologues flexionnelles mais aussi aux compétences d'identification de la construction morphologique de pseudo-mots.

En 5^e année, les compétences d'identification de la construction morphologique de pseudo-mots ne seraient plus liées, comme au niveau précédent, aux connaissances morpho-flexionnelles mais plutôt aux connaissances morpho-dérivationnelles. Le lien établi précédemment entre les connaissances morphologiques dérivationnelles et flexionnelles n'est que tendanciel mais toutefois apparent.

En 7^e année, les résultats plafonds obtenus sur une majorité d'épreuves n'ont pas permis de rendre compte du lien entre les différentes connaissances que nous souhaitions analyser. Les associations entre les mesures des connaissances morphologiques, observées aux niveaux précédents, n'ont ainsi pas pu être vérifiées, les compétences à identifier la construction morphologique de pseudo-mots n'apparaissant en outre, liées à aucune des connaissances morphologiques implicites, à ce niveau scolaire. Il apparaît, en revanche, que les connaissances implicites grapho-morpho-flexionnelles et grapho-syntaxiques sont associées.

Ainsi, les résultats obtenus auprès des deux premiers niveaux scolaire indiquent que les deux types de connaissances morphologiques (dérivationnelles et flexionnelles) ne seraient pas indépendantes. De plus, en 3^e année, ce sont en priorité les élèves qui discriminent le mieux la structure morpho-flexionnelle des mots portugais qui parviennent le mieux à s'abstraire des spécificités de leur langue pour identifier la construction morphologique de pseudo mots alors qu'en 5^e ce serait davantage les élèves dont les connaissances morpho-dérivationnelles en portugais sont les plus développées. En 7^e année, la symétrie entre les réponses aux épreuves implicites grapho-morpho-flexionnelle et grapho-syntaxique suggère qu'ici, ces deux sensibilités relèvent de la prise en compte des aspects sémantiques des mots.

2.3.2 Analyses en composantes principales

Des analyses en composantes principales ont été menées sur les épreuves sélectionnées respectivement en 3^e, 5^e et 7^e année. Le tableau 41 présente donc les résultats obtenus pour chaque niveau.

Tableau 41 : Contribution des performances aux différentes épreuves sélectionnées aux facteurs extraits des analyses en composantes principales en 3e, 5e et 7e année.

	3 ^e année			5 ^e année			7 ^e année	
	Fact. 1	Fact. 2	Fact. 3	Fact. 1	Fact. 2	Fact. 3	Fact. 1	Fact. 2
Dét. Imp. G-phonologique	.44	-.21	.67	-.00	.34	-.89*	.24	-.76*
Dét. Imp. G-morpho-flex.	-.30	.45	.59				.83*	.14
Dét. Imp. G-syntaxique	-.52	.59	-.11				.67	.53
Dét. Exp. G-morpho-dériv.	.36	.60	-.11	-.81*	.28	.08		
Dét. Exp. G-morpho-flex.	.39	.69	-.20	-.69	-.23	-.32		
Dét. Exp. G-syntaxique	.61	-.20	-.44	-.06	-.92*	-.24		
Ident. Const. morpho.	.77*	.22	.27	-.70*	-.04	.25	.54	-.53
% de variance totale expliquée	25,73	21,79	16,26	32,90	22,00	20,44	37,52	29,12

* poids factoriel significatif (> .70)

En 3^e année, plusieurs facteurs se dégagent de l'ACP mais un seul rend compte dans sa composition d'un poids factoriel significatif pour l'épreuve la plus discriminante. Ce facteur expliquant 25,75 % de la variance totale est ainsi composé de l'épreuve d'identification de la construction morphologique de pseudo-mots (poids de -.77). Le poids des performances à l'épreuve grapho-morpho-flexionnelles explicite et à l'épreuve graphophonologique implicite dans la composition des deux autres facteurs dégagés (respectivement .69 et .67) ne permet pas d'aboutir à interprétation rigoureuse quant au caractère très discriminant de ces épreuves.

En 5^e année, trois facteurs se dégagent de l'ACP : le premier expliquant 32,90 % de la variance est composé des épreuves de détection explicite d'intrus grapho-morpho-dérivationnel (poids de -.81) et d'identification de la construction morphologique de pseudo-mots (poids de -.70). Le poids des résultats à l'épreuve grapho-morpho-flexionnelle n'est pas suffisant pour être significatif. Le deuxième, expliquant 22 % de la variance, est saturé par l'épreuve implicite de détection d'intrus graphophonologique (poids de -.92). Enfin, le troisième facteur, expliquant 20,44 % de la variance, est exclusivement composé de l'épreuve implicite de détection d'intrus graphophonologique (poids de .89).

En 7^e année, deux facteurs apparaissent : le premier est composé de l'épreuve implicite grapho-morpho-flexionnel (poids de .82) et explique à lui seul 37,52 % de la variance totale. Le second, indépendant du premier et composé de l'épreuve implicite graphophonologique (poids de .76) explique quant à lui 29,12 % de variance.

En 3^e année, l'épreuve qui apparaît la plus discriminante est l'épreuve morphologique portant sur les pseudo-mots constituant le facteur expliquant la part de variance totale la plus importante (25,73 %). Ce facteur pourrait ainsi être assimilé à une composante des connaissances morphologiques dégagées des aspects formels du portugais.

En 5^e année, trois facteurs apparaissent chacun contribuer à expliquer de façon indépendante une part de la variance totale allant de 20,44 à 32,90 %. Le premier facteur identifié et pouvant correspondre aux connaissances grapho-morphologiques n'est plus, comme en 3^e année, seulement composé de l'épreuve d'identification de la construction morphologique de pseudo-mots mais ferait intervenir en premier lieu les connaissances morpho-dérivationnelles en langue portugaise. Un deuxième facteur, nouveau par rapport au niveau précédent, se dégage. Il est constitué de l'épreuve grapho-syntaxique et pourrait ainsi être assimilés aux connaissances syntaxiques permettant d'analyser consciemment la structure verbale de phrases. Enfin, un troisième facteur assimilable aux « connaissances implicites graphophonologiques » contribuerait également à expliquer une part non négligeable de la variance des résultats en 5^e année.

En 7^e année, les deux facteurs expliquant la variance totale entre élèves font intervenir les connaissances implicites surtout grapho-morpho-flexionnelles mais aussi graphophonologiques, ces deux types de connaissances étant largement indépendantes l'une de l'autre.

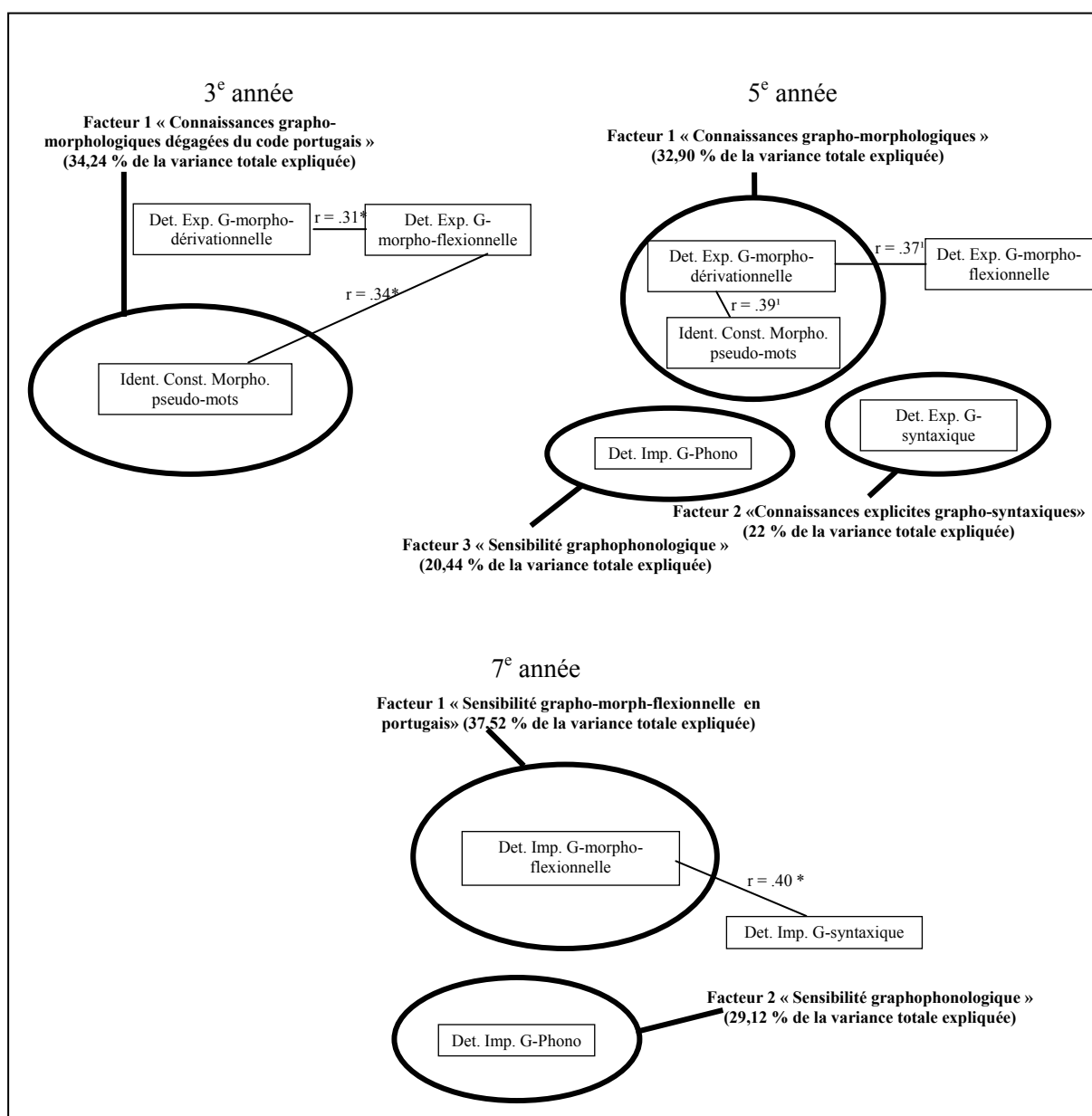
2.3.3 Synthèse

Le schéma 3 permet d'avoir un aperçu synoptique des principales données statistiques décrites dans les deux paragraphes précédents.

La prise en compte simultanée des résultats obtenus aux analyses corrélationnelles et factorielles permet d'approfondir et nuancer les conclusions envisagées précédemment. En effet, l'observation de certains coefficients de corrélation, permet de compléter les résultats des ACP.

En 3^e année comme en 5^e année, l'épreuve de détection explicite d'intrus grapho-flexionnelle n'intervient pas dans la constitution du facteur « connaissances grapho-morphologiques ». Cependant le taux de réussite à cette épreuve n'est pas totalement dissocié de ceux obtenus à

d'autres épreuves entrant dans la composition de ce facteur (lien avec les scores d'identification de la construction morphologique de pseudo-mots en 3^e année et lien tendanciel avec les scores de détection d'intrus grapho-morpho-dérivationnel en 5^e année). Il est également intéressant de noter que le score à cette épreuve est, en 3^e année, également associé celui obtenu à la troisième épreuve évaluant les connaissances morphologiques (détection explicite d'intrus morpho-dérivationnel) et qu'en 5^e année son poids factoriel n'est pas négligeable, mais n'atteint pas le taux de significativité requis pour entrer dans la composition du facteur dégagée. Ces éléments laissent ainsi supposer que cette mesure serait toutefois un bon indicateur des compétences morphologiques des élèves de 3^e et de 5^e année.



* r significatif à $p < .05$; ¹ tendanciel à $p < .08$

Schéma 3 : Représentation des liens corrélationnels entre les épreuves sélectionnées et des facteurs dégagés de l'ACP en 3^e, 5^e et 7^e année.

2.4 Recherche des performances déterminant le niveau de lecture

2.4.1 Analyses sur le niveau de lecture de mots

2.4.1.1 Analyses de corrélations entre les scores aux différentes épreuves sélectionnées et le niveau de lecture de mots

Le tableau 42 présente les coefficients de corrélation entre les scores aux différentes épreuves sélectionnées et le niveau de lecture de mots en 3^e, 5^e et 7^e année.

Tableau 42 : Coefficients de corrélation entre les scores aux épreuves sélectionnées et le niveau de lecture de mots en 3^e, 5^e et 7^e année.

	Dét. Imp. G-phono	Dét. Imp. G-morpho- flex.	Dét. Imp. G-synt.	Dét. Exp. G-morpho- dériv.	Dét. Exp. G-morpho- flex.	Dét. Exp. G-synt.	Ident. Const. morpho. PM
3 ^e année N = 43	.06	-.05	.15	.29 ^t	.60**	.23	.51**
5 ^e année N = 20	-.33	X	X	.44 ^t	.31	.22	.34
7 ^e année N = 22	.44*	.36	-.03	X	X	X	.27

** r significatif à $p < .01$ * à $p < .05$ ^t tendanciel à $p < .07$

En 3^e année, nous pouvons observer que les scores à l'épreuve de détection explicite d'intrus grapho-morpho-flexionnel sont les plus corrélés aux performances en lecture de mots ($r = .60$, $p < .0001$). Les scores à l'épreuve d'identification de la construction morphologique de pseudo-mots y sont également corrélés significativement ($r = .51$, $p < .001$). Enfin, la corrélation entre les performances en lecture de mots et les performances à l'épreuve de détection grapho-morpho-dérivationnelles n'est que tendancielle ($r = .29$, $p < .06$). Les autres corrélations ne sont pas significatives.

En 5^e année, aucune corrélation n'est significative sans doute du fait du nombre limité d'observation pour ce groupe (N = 20). Cependant, la corrélation entre le niveau de lecture de mots et les performances à l'épreuve de détection d'intrus grapho-morpho-dérivationnelle est relativement élevée et tendanciellement significative ($r = .44$, $p < .06$).

En 7^e année, seule la corrélation entre les performances à l'épreuve de détection implicite d'intrus graphophonologique et le niveau de lecture de mots est significative ($r = .44$, $p < .04$)

Ainsi, l'association entre le niveau de lecture de mots et le niveau de connaissances dans différents domaines linguistiques évalués dans cette étude diffèrent d'un niveau scolaire à l'autre. Il convient de souligner que l'effectif de certains groupes ainsi que l'exclusion de plusieurs résultats obtenus en 5^e et 7^e année peuvent en être à l'origine. En 3^e année, le niveau de lecture de mots est surtout associé aux connaissances explicites morpho-

flexionnelles portant sur les mots portugais et aux compétences à identifier la construction morphologique de pseudo-mots. En 5^e année, le niveau de lecture serait plutôt associé aux connaissances explicites morpho-dérivationnelles appliquées à la langue portugaise et, en 7^e année, aux connaissances implicites des aspects graphophonologiques.

Ainsi, en 3^e et 5^e année, le facteur « connaissances grapho-morphologiques » identifié comme le plus discriminant jouerait donc un rôle dans la reconnaissance des mots portugais. Cette relation entre connaissances grapho-morphologiques et niveau de lecture n'a en grande partie pas pu être vérifiée en 7^e année.

2.4.1.2 Analyses de régressions sur la variable dépendante « lecture de mots »

Afin de déterminer dans quelle mesure les connaissances grapho-sémantiques prédisent de façon indépendante le niveau de lecture de mots, en 3^e et 5^e, et afin de préciser la contribution des connaissances graphophonologiques au niveau de lecture de mots en 7^e année, des analyses de régression pas à pas ont été conduites sur chaque niveau scolaire.

En 3^e année, si l'on entre dans le modèle les deux variables présentant des corrélations significatives avec le niveau de lecture de mots (détection explicite grapho-morpho-flexionnelle et identification de la construction morphologique de pseudo-mots) on observe que ces deux variables contribuent ensemble à expliquer 46,22 % de la variance en lecture de mots. La variable « détection explicite grapho-morpho-flexionnelle » apporte la contribution la plus importante (35,73 %). La variable « identification de la construction morphologique », incluse dans un deuxième temps ajoute 10,48 % de variance expliquée (*cf.* tableau 43).

Tableau 43 : Contribution des variables entrées dans le modèle de régression sur le niveau de lecture de mots en 3^e année

	R	R ²	R ² modifié	F	p
Pas 1 : Det. Exp. G-morpho-flex.	.60	.3573	.3573	22,80	.0001
Pas 2 : Identification construction morphologique	.68	.4622	.1948	7,80	.01

En 5^e année, si l'on entre dans le modèle l'unique variable tendanciellement corrélée au niveau de lecture de mots (détection explicite grapho-morpho-dérivationnelle), on observe qu'elle contribue à expliquer 23,18 % de la variance en lecture de mots (*cf.* tableau 44).

Tableau 44 : Contribution de la variable entrée dans le modèle de régression sur le niveau de lecture de mots en 5^e année

	R	R ²	R ² ajusté	F	p
Dét Exp G-morpho-dériv.	.48	.2318	.1952	6,34	.03

En 7^e année, si l'on entre dans le modèle l'unique variable corrélée au niveau de lecture de mots (détection implicite graphophonologique), on observe qu'elle contribue à expliquer 19,64 % de la variance en lecture de mots (*cf.* tableau 45).

Tableau 45 : Contribution de la variable entrée dans le modèle de régression sur le niveau de lecture de mots en 7^e année

	R	R ²	R ² ajusté	F	p
Dét Imp G-phonologique	.44	.1964	.1562	4,89	.04

A l'issue de ces analyses, nous pouvons constater, pour les deux premiers niveaux scolaires, une contribution des compétences morphologiques à la lecture de mots en portugais. Cependant, ces dernières ne sont pas identiques : pour le premier niveau étudié, la lecture de mots serait surtout expliquée par les connaissances explicites grapho-morpho-flexionnelles (35,73 % de variance expliquée) puis par les compétences d'identification de la construction morphologique de pseudo-mots (19,48 % de variance expliquée en deuxième pas), compétences qui étaient apparues comme les plus discriminantes pour ce groupe. En 5^e année en revanche, ce sont les connaissances explicites grapho-morpho-dérivationnelles, dégagées comme un bon indicateur des variations entre les élèves, qui apportent leur contribution (23,18 % de variance expliquée). Ces éléments permettent d'envisager une évolution dans la contribution des connaissances morphologiques (morpho-flexionnelles puis morpho-dérivationnelle) à la lecture de mots en portugais. Enfin, en 7^e année, et contrairement aux niveaux précédents, se sont les connaissances implicites graphophonologiques, pourtant apparues comme un facteur discriminant dès la 3^e année, qui contribueraient à expliquer une part plus modeste de la variance en lecture de mots (19,64 %).

2.4.2 Analyses sur le niveau de compréhension de phrases écrites

2.4.2.1 Analyses de corrélations entre les scores aux différentes épreuves sélectionnées, au test de lecture de mots et le niveau de compréhension de phrases écrites

Le tableau 46 présente les coefficients de corrélation entre les scores aux différentes épreuves sélectionnées et au test de lecture de mots et le niveau de compréhension de phrases écrites en 3^e, 5^e et 7^e année.

Tableau 46 : Coefficients de corrélation entre les différents scores et le niveau de compréhension de phrases en 3^e, 5^e et 7^e année.

	Dét. Imp. G-phono.	Dét. Imp. G-morpho- flex.	Dét. Imp. G-synt.	Dét. Exp. G-morpho- dériv.	Dét. Exp. G-morpho- flex.	Dét. Exp. G-synt.	Ident. Const. morpho. PM	Lecture mots
3 ^e année N = 42	-.14	-.21	-.28	-.25	.08	.22	.07	.04
5 ^e année N = 20	.49*	X	X	.03	.47*	-.22	.17	.20
7 ^e année N = 22	.36	.37	.42	X	X	X	.45*	.13

* à $p < .05$

En 3^e année, nous pouvons observer que les scores au test de compréhension de phrases ne sont significativement corrélés à aucun autre score mesuré par les différentes épreuves de détection, d'identification de la construction morphologique de pseudo-mots et de lecture de mots.

En 5^e année, ce sont les scores aux épreuves de détection implicite graphophonologique et de détection explicite grapho-morpho-flexionnelle qui sont significativement corrélés aux performances en compréhension de phrases ($r = .49$, $p < .05$ et $r = .47$, $p < .05$) les autres corrélations étant non significatives.

En 7^e année, seuls les scores à l'épreuve d'identification de la construction morphologique de pseudo-mots sont significativement corrélés aux performances en compréhension de phrases ($r = .45$, $p < .05$).

Ces résultats permettent de constater tout d'abord l'absence de relation entre lecture de mots et compréhension de phrases, quel que soit le niveau scolaire.

En 5^e année, les meilleurs compreneurs sont les élèves qui sont les plus sensibles à la structure phonologique des mots portugais mais aussi ceux qui parviennent le mieux à analyser les marques flexionnelles. Pourtant, ces connaissances, associées à la compréhension, ne sont pas celles qui sont les plus discriminantes (cf. ACP). En effet, la variable « sensibilité graphophonologique » avait été dégagée comme troisième et dernier facteur expliquant à peine plus de 20 % de la réussite aux différentes épreuves et l'épreuve

grapho-morpho-flexionnelle explicite n'été pas apparue avec suffisamment de poids pour intervenir dans la constitution du facteur « connaissances grapho-morphologiques ». En 7^e année, la compréhension de phrases serait pour sa part liée aux compétences à identifier la construction morphologique de pseudo-mots c'est-à-dire à des connaissances morphologiques plus décontextualisées.

2.4.2.2 Analyses de régressions sur la variable « compréhension de phrases écrites »

En 3^e année, l'analyse de régression n'a pu être menée puisque aucune variable n'est apparue corrélée au niveau de compréhension de phrases (*cf.* tableau 46).

En 5^e année, si l'on entre dans le modèle les deux variables significativement corrélées au niveau de compréhension de phrases (détection implicite graphophonologique et détection explicite grapho-morpho-flexionnelle), on observe qu'elles contribuent ensemble à expliquer 36,81 % de la variance en compréhension de phrases. La variable « détection implicite graphophonologique » explique à elle seule 20,18 % de la variance et la variable « détection explicite grapho-morpho-flexionnelle », incluse dans un deuxième temps ajoute 16,64 % de variance expliquée (*cf.* tableau 47).

Tableau 47 : Contribution des variables entrées dans le modèle de régression sur le niveau de compréhension de phrases en 5^e année

	R	R ²	R ² modifié	F	p
Pas 1 : Det. Imp. G-phonologique	.45	.2018	.2018	4,80	<.05
Pas 2 : Det. Exp. G-morpho-flex.	.61	.3681	.1664	4,74	<.05

En 7^e année, les performances d'identification de la construction morphologique de pseudo-mots (les seules significativement corrélées aux scores en compréhension de phrases) contribuent à expliquer 17,34 % de la variance en compréhension de phrases en portugais (*cf.* tableau 48).

Tableau 48 : Contribution de la variable entrée dans le modèle de régression sur le niveau de compréhension de phrases en 7^e année

	R	R ²	R ² ajusté	F	p
Identification construction morphologique de pseudo-mots	.42	.1734	.1404	5,25	.04

Les connaissances apparaissant contribuer à la compréhension de phrases écrites en portugais semblent évoluer entre la 5^e et la 7^e année.

En 5^e année, le niveau de compréhension de phrases est en premier lieu expliqué par les connaissances implicites à percevoir les relations ou les différences graphophonologiques entre les mots mais aussi par les connaissances explicites portant les marques morpho-flexionnelles des mots portugais. Les analyses de performances précédentes ont révélé que les connaissances graphophonologiques implicites étaient surtout développées sur les unités larges comme la rime et la syllabe. Cette sensibilité infra-lexicale pourrait ainsi participer à la compréhension à un niveau déjà avancé de l'apprentissage, de façon relativement importante.

En 7^e année, le rôle des aspects morphologiques se confirme puisque le niveau de compréhension de phrases est essentiellement expliqué par les compétences d'identification de la construction morphologique de pseudo-mots.

2.4.3 Synthèse sur la contribution des différentes variables à l'explication du niveau de lecture de mots et de compréhension de phrases

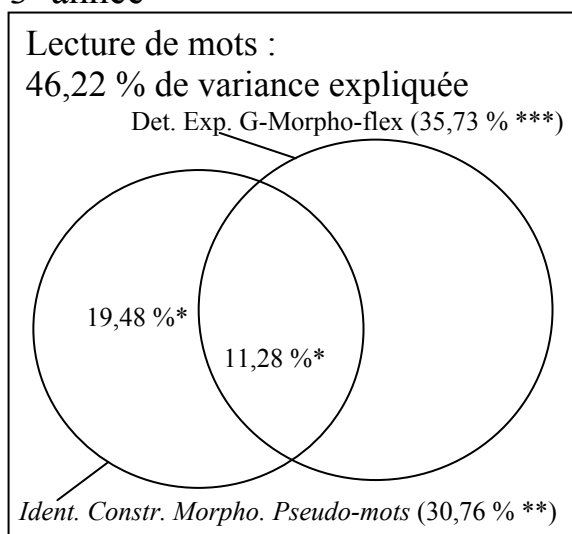
A l'issue de ces analyses, il est possible de dégager les points suivants (*cf.* schéma 4) :

- En 3^e année, le facteur « connaissances grapho-morphologiques dégagées du code portugais » contribuerait à expliquer le niveau de lecture de mots. Toutefois, un autre type de connaissances morphologiques interviendrait préalablement : les connaissances explicites des règles morpho-flexionnelles du portugais. Nous constatons également que l'ensemble des mesures obtenues échoue à expliquer le niveau de compréhension de phrases des élèves de ce niveau scolaire. Ce dernier serait vraisemblablement dépendant d'autres connaissances.

- En 5^e année, le premier facteur « connaissances grapho-morphologiques » identifié contribuerait à expliquer le niveau de lecture de mots par le biais des connaissances explicites grapho-morpho-dérivationnelles permettant d'analyser les mots en leurs composants morphémiques (23,18 % de variance expliquée par cette variable). La compréhension de phrases écrites serait pour sa part dépendante de la sensibilité graphophonologique mais aussi des connaissances grapho-morpho-flexionnelles permettant d'analyser la construction des verbes conjugués et des noms féminins en portugais (ces deux variables expliquant respectivement 20,18 % et 16,63 % de la variance).

- En 7^e année, seule la sensibilité graphophonologique contribuerait à expliquer le niveau de lecture de mots alors que la compréhension de phrases serait exclusivement expliquée par les connaissances grapho-morphologiques dégagées des aspects formels du portugais (les compétences d'identification de la construction morphologique de pseudo-mots expliquent en effet 17,38 % de la variance en compréhension).

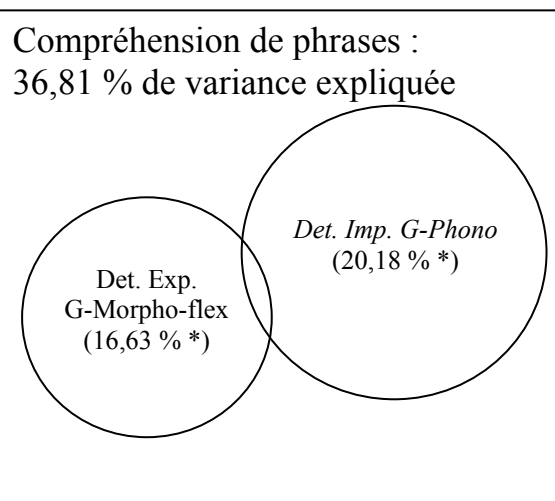
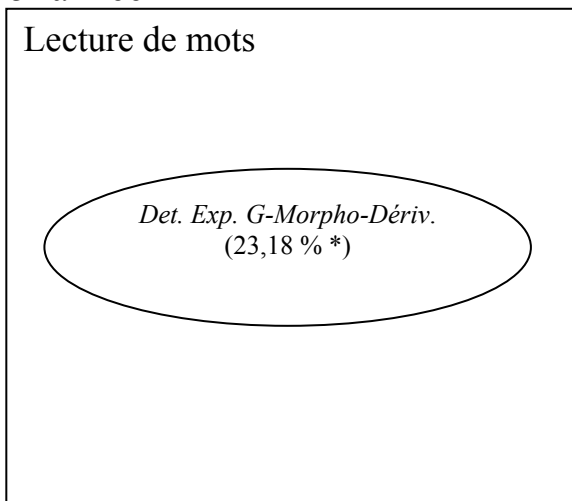
3^e année



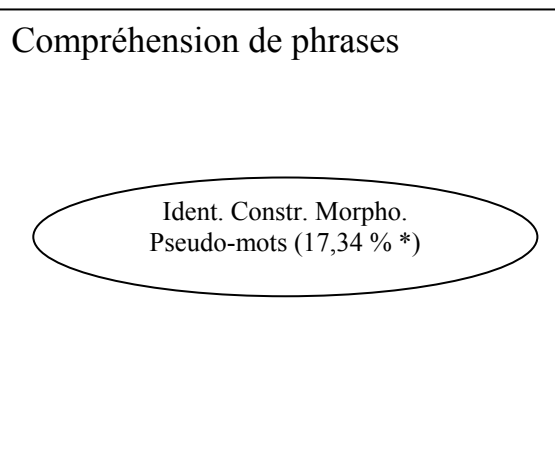
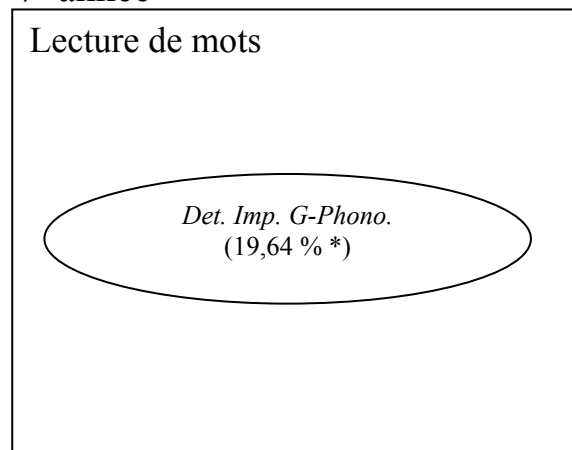
Compréhension de phrases

Aucune part de variance expliquée
par les épreuves

5^e année



7^e année



*significatif à $p < .05$; **à $p < .01$; *** à $p < .001$. *En italique : Variables entrant dans la constitution des facteurs dégagés au niveau scolaire indiqué.*

Schéma 4 : Représentation des variables expliquant le niveau de lecture de mots et de compréhension de phrases à chaque niveau scolaire

3. Discussion

Lors de l'étude 2, nous nous étions fixée comme objectif d'approfondir l'évolution respective des connaissances graphophonologiques, grapho-morphologiques et grapho-syntaxiques, implicites et explicites et leur rôle lors de la lecture en portugais en 3^e, 5^e et 7^e année d'apprentissage.

Concernant la dimension implicite des connaissances, nous avons pu vérifier que la sensibilité aux aspects phonologiques était particulièrement bien développée. Cette observation, associée à l'apparition d'une réussite totale pour une majorité d'élèves à détecter consciemment les intrus phonologiques, étaye le postulat d'une plus grande facilité à traiter cette dimension du code relevant de mécanismes linguistiques plus simples que la dimension grapho-sémantique (*e.g.* Carlisle, 1995).

Nous nous attendions en outre à ce que les connaissances grapho-morpho-flexionnelles soient plus développées en portugais que les connaissances grapho-morpho-dérivationnelles. Quel que soit le niveau scolaire, aucune sensibilité à la construction morpho-dérivationnelle du lexique portugais n'est apparue. En revanche, l'analogie entre les mots en fonction des marques flexionnelles semble généralement pouvoir s'effectuer. Cette supériorité a également pu être observée sur le niveau explicite des connaissances et permet d'aboutir aux mêmes conclusions : les règles morpho-flexionnelles s'avèrent être mieux maîtrisées que les règles morpho-dérivationnelles en portugais¹⁵⁸. Cette affirmation est cohérente avec celles des auteurs ayant travaillé sur l'anglais par exemple (*e.g.* Gordon, 1989 ; Carlisle & Nomanbhoy, 1993). Les arguments avancés pour expliquer ce phénomène font généralement référence à la productivité des processus de transformation (*e.g.* Gordon, 1989) ou encore au caractère prédictible de son application (*e.g.* Fowler *et al.*, 2003). En portugais, l'étude de Vidigal de Paula (2007) avait échoué à montrer cette différence sur la dimension explicite des connaissances. Notre recherche partage pourtant plusieurs caractéristiques avec cette dernière (mêmes épreuves, mêmes conditions de passation, mêmes niveaux scolaires¹⁵⁹). Ce contraste peut provenir des différences d'expériences linguistiques entre les populations. Même si ces expérimentations ont toutes deux été menées auprès d'élèves brésiliens scolarisés en établissement privé à São Paulo et dans une zone urbaine proche, notre étude a été réalisée dans un lycée français (le seul de São Paulo) où le

¹⁵⁸ Cet avantage à analyser les règles morpho-flexionnelles est également étayé par les variations des compétences d'identification de la construction morphologique de pseudo-mots chez les élèves de 5^e année pour qui les règles s'apparentant à des aspects morpho-flexionnels sont plus facilement traitées que celles s'apparentant aux aspects morpho-dérivationnels.

¹⁵⁹ La recherche de Vidigal de Paula (2007) porte en plus sur le niveau 1^e année.

français est enseigné dès la 2^e année. L'apprentissage des langues et les contacts avec l'écrit auxquels sont exposés les élèves de ces deux études sont donc difficilement comparables.

Concernant la comparaison entre connaissances grapho-morphologiques et grapho-syntaxiques, nos observations suggèrent que les élèves peuvent orienter leur choix selon la fonction d'un mot dans la phrase, alors que pour choisir entre deux mots, ils « préfèrent » se baser sur d'autres indices que celui des caractéristiques morphe-dérivationnelles. Cette sensibilité à la construction syntaxique des phrases ne semble en revanche, pas plus importante que celle concernant les marques flexionnelles portées par les mots. Ces résultats ne contredisent pas les conclusions des études ayant avancé qu'il est plus facile de juger incorrectes des phrases désordonnées que celles comportant des erreurs morphémiques (Demont & Gombert, 1996 ; Nocus & Gombert, 1997) et peuvent même donner lieu à une interprétation du même type. En étudiant ici d'autres aspects syntaxiques à l'aide d'un paradigme maintenant constant la nature sémantique des stimuli linguistiques, il se dégage qu'en l'absence de perturbations sémantiques le repérage inconscient de la fonction du mot dans une phrase n'est pas meilleur que celui du marquage flexionnel. Par ailleurs, l'analyse consciente s'avère pouvoir s'effectuer plus facilement sur les aspects morpho-flexionnels que syntaxiques. Pour reprendre la terminologie de Kail (1986), cette facilité peut vraisemblablement s'expliquer par le fait que les marques flexionnelles représentent des indices locaux, donc plus facilement détectables que les relations entre les mots de la phrase, de nature topologique.

Un autre objectif de cette recherche était de valider l'existence d'une sensibilité à la rime intrasyllabique et à la syllabe en portugais alors qu'une analyse consciente est nécessaire pour détecter les ressemblances entre les mots sur une unité phonémique. Les résultats valident clairement cette hypothèse. A tous les niveaux scolaires, les élèves brésiliens repèrent spontanément l'analogie de rime et de syllabe mais ne perçoivent pas intuitivement le partage du phonème final entre deux mots¹⁶⁰. La complète réussite à cette même tâche de détection mais avec une consigne explicite contraste donc très fortement avec l'absence de sensibilité phonémique ce qui renforce le postulat que les unités fines ne peuvent être traitées qu'explicitement.

D'autre part, la chute de la sensibilité phonologique observée chez les élèves de 5^e année peut dénoter un revirement dans la mise en place des connaissances graphophonologiques implicites.

¹⁶⁰ Le risque que certains enfants comprennent d'eux mêmes que les aspects graphophonologiques sont manipulés dans cette épreuve est possible. Leurs réponses, procédant alors d'un repérage conscient des caractéristiques des mots, pourraient alors surévaluer les moyennes de réussites.

Il semble ainsi que les enfants moins avancés que ces derniers répondent plus facilement intuitivement à une tâche sans consigne et se laissent porter par leur préférence. Cinq ans plus tard, les automatismes relatifs à l'application des correspondances graphophonologiques sont tels que le traitement des unités graphophonologiques les plus composites peut s'effectuer très facilement, presque à coût cognitif nul. Cette compétence relèverait donc davantage d'une procéduralisation des apprentissages explicites, telle que Logan (1988) la définit.

A un niveau intermédiaire, les enfants seraient suffisamment « entrés dans les apprentissages » pour être réticents à résoudre une tâche « sans réfléchir » mais ils n'auraient pas totalement, ou pas tous, automatisé les procédures de correspondances graphophonologiques. Prises conjointement, ces deux caractéristiques expliqueraient leur échec relatif à cette épreuve.

L'approfondissement des unités morphologiques (les préfixes, les suffixes, ou l'intrication non linéaire de voyelles) les mieux identifiées lorsque le matériel linguistique à manipuler n'est pas en portugais va dans le sens d'un meilleur traitement des préfixes que des suffixes. Ces résultats ne concordent pas avec ceux de certaines recherches françaises, par exemple celles de Colé (2004) ou Marec-Breton (2003). Ces auteurs trouvent en effet lors de la résolution de multiples tâches morphologiques (en particulier en contexte de phrases) une supériorité des suffixes par rapport aux préfixes. Cependant, Daigle et Demont (2006) n'ont pas noté une telle différence dans les réponses d'enfants du CP au CM2 à une épreuve de plausibilité¹⁶¹. Les résultats de Bescond (1999) tendent même de montrer un effet inverse lors de la résolution de tâche de plausibilité ou de détection implicite d'intrus. La nature des tâches et les caractéristiques des affixes employés peuvent être à l'origine de ces variations. Colé (2004) conclut d'ailleurs qu'il n'existe pas actuellement de données suffisantes pour expliquer les différences entre la conscience des préfixes et des suffixes.

Concernant l'articulation entre les connaissances grapho-morpho-dérivationnelles, grapho-morpho-flexionnelles et grapho-syntaxiques, nous ne sommes pas parvenue à faire émerger le facteur grapho-sémantique attendu. On a remarqué, par exemple en 5^e année, que les connaissances explicites grapho-morphologiques et grapho-syntaxiques étaient indépendantes. Les traitements de nature intra et inter-lexicale ne semblent ainsi pas dépendre des mêmes mécanismes, même s'ils peuvent tous deux faire intervenir la dimension sémantique des mots, ce qui ne nous semble pas contradictoire. Il faut aussi rappeler que l'ensemble des résultats aux

¹⁶¹ Les expérimentations de Colé (2004) n'ont d'ailleurs pas observé de telles différences lors d'épreuves de jugement ou d'extraction de la base

épreuves de détection explicite d'intrus plafonne en 7^e année, et que par conséquent l'analyse en composantes principales s'appuie à ce niveau scolaire sur peu de résultats. L'hypothèse de voir se dégager un facteur grapho-sémantique n'est ici pas totalement vérifiable.

Dans le domaine des mécanismes de lecture en portugais, il est possible de conclure tout d'abord que la conscience morphologique (dans un premier temps, surtout la conscience morpho-flexionnelle puis ultérieurement, la conscience morpho-dérivationnelle) intervient pour lire les mots portugais. Les conclusions de Vidigal de Paula (2007) se trouvent ainsi confortées. De plus en 7^e année, l'intervention des connaissances graphophonologiques implicites lors de la lecture de mots signifie que le décodage est parfaitement automatisé. Pour ce groupe, l'absence de contribution des connaissances grapho-morphologiques explicites en portugais, observée aux niveaux précédents, provient tout simplement du fait que les épreuves employées sont trop faciles pour évaluer correctement les compétences d'élèves aussi avancés dans les apprentissages. Le rôle de ces connaissances n'a ainsi pas pu être étudié mais serait probablement apparu en utilisant un moyen d'investigation plus adapté à l'âge.

L'intervention des connaissances grapho-morphologiques est également présente lors de la compréhension de phrases écrites. Elle semble toutefois plus tardive et moins importante que sur la reconnaissance des mots écrits. En effet, la conscience des règles grapho-morphologiques ne contribue à la compréhension de phrases qu'en 5^e année, mais toutefois dans une moindre mesure que la sensibilité graphophonologique. Là encore, il est mis en évidence la nécessité d'automatiser les procédures de décodage pour pouvoir mettre en place des mécanismes plus élaborés, utiles à l'accès à la signification des mots. Au niveau suivant, le profil des lecteurs semble encore avoir évolué. Les élèves de 7^e année ont visiblement dépassé les traitements séquentiels initiaux puisque la contribution des connaissances graphophonologiques implicites n'apparaît plus. Ils pourraient ainsi se consacrer, lors de leur lecture de phrases, à la mise en place des traitements grapho-morphologiques déjà bien élaborés, comme ceux relevant d'une analyse explicite des règles dégagée des spécificités formelles du portugais.

Ces résultats permettent, dans une certaine mesure, de dépasser la difficulté rencontrée par Vidigal de Paula (2007) qui n'avait pas réussi à confirmer en portugais le lien entre la conscience morphosyntaxique et la compréhension en lecture pourtant établie par plusieurs recherches en français (Casalis & Louis Alexandre, 2000 ; Demont & Gombert, 1996 ; Nocus & Gombert, 1997). Toutefois, nous ne parvenons pas à démontrer l'existence d'une contribution de la morphologie (dérivationnelle comme flexionnelle) à la compréhension en lecture avant la fin de primaire (5^e année) alors que les études précitées l'établissent bien plus précocement, au moins à

partir de la 3^e année. Le type de matériel employé peut expliquer ces différences. Néanmoins, force est de constater que les recherches établissant ce lien en début d'apprentissage ne concernent encore que très peu de langues. La prise en considération de la diversité linguistique dans ce domaine nous semble ainsi primordiale pour généraliser de tels résultats. A notre connaissance, seule l'étude de Ku et Anderson (2003) se situe dans cette orientation, en se focalisant sur l'anglais et le chinois. Dans nos travaux, la réalisation d'une étude similaire en arabe permettra, en guise de discussion générale, de mettre en évidence certains points communs et différences quant aux résultats les plus flagrants dans les deux langues.

Chapitre 9 : Étude 3 - Perspective évolutive et comparative *cross-linguistique* de l'apprentissage de la lecture en français langue seconde

1. Méthode

1.1 Population

Les élèves ayant participé à cette expérimentation sont les mêmes que ceux des études 1 et 2, soit des élèves tunisiens et brésiliens apprenant le français dans le cadre scolaire. Dans les établissements publics tunisiens, l'enseignement du français débute en 3^e année de scolarité à raison de 8 heures par semaine puis de 9 heures par semaine de la 4^e à la 6^e année. Au Brésil en revanche, cet enseignement ne fait pas partie du programme mais les écoles françaises le dispensent dès l'école primaire. Dans le cas de l'école privée sélectionnée (la seule école française de Sao Paulo), le français est enseigné formellement à partir de la 2^e année de scolarité à raison de 4 heures par semaine puis de 3 heures par semaine à partir de la 7^e année. Ainsi, dans les écoles publiques en Tunisie et dans l'école française brésilienne privée choisies, le nombre d'heures de cours consacrées au Français est loin d'être identique. De ce fait, les niveaux scolaires ont été choisis de façon que, d'une part le volume horaire d'apprentissage FL2, depuis le début de la scolarité, soit le moins éloigné possible entre les groupes lusophones et arabophones et d'autre part que l'efficacité verbale orale soit suffisamment proche pour permettre les comparaisons prévues¹⁶². Ainsi, 3 niveaux distincts ont été discriminés, en fonction du volume horaire reçu. A l'intérieur de ces trois niveaux, un appariement global entre arabophones et lusophones a été réalisé à partir des scores obtenus à l'épreuve réceptive de vocabulaire de Deltour et Hupkens (1980 ; cf. infra sa description dans la présentation du matériel).

Après la collecte des données mais avant d'effectuer les appariements, les données des élèves brésiliens n'ayant pas effectué l'ensemble de leur scolarité au lycée français ont d'abord été supprimées (concerne 6 élèves de 3^e année ou en 2^e année de français, 8 élèves de 5^e année ou en

¹⁶² Ce rapprochement a été établi à partir d'un pré-test réalisé auprès d'élèves tunisiens de 3^{ème}, 4^{ème}, 5^{ème} et 6^{ème} année, scolarisés dans la même localité que celle des écoles où ont eu lieu les expérimentations. Les élèves de ces classes ont été évalués avec le même test de vocabulaire en français que dans la présente étude (Deltour & Hupkens, 1980). Ce pré-test a permis de sélectionner *a priori* les niveaux scolaires pouvant s'apparenter au mieux aux groupes des brésiliens, en terme de connaissances lexicales.

3^e année de français et 7 élèves de 7^e année ou 5^e année de français), de même que celles des élèves ayant participé à moins de la moitié des épreuves en français (concerne 3 élèves brésiliens de 7^e année). Ensuite, l'appariement des groupes sur les connaissances lexicales nous a obligée à supprimer l'ensemble des résultats d'élèves n'ayant pas passé le test de vocabulaire (concerne 2 élèves brésiliens de 3^e année 1 de 5^e année, 1 de 7^e année et 5 élèves tunisiens de 1^e année). Il a également été nécessaire d'exclure les résultats de plusieurs élèves tunisiens dont le niveau de connaissances lexicales était, soit trop élevé, soit trop faible, par rapport à la distribution des résultats de leur pairs lusophones (concerne 13 élèves de 3^e année, 2 élèves de 4^e année et 8 élèves de 5^e année). Ainsi, parmi les 265 élèves déjà sélectionnés lors des expérimentations 1 et 2, 209 ont été retenus pour les analyses en français langue seconde (125 arabophones et 84 lusophones). Par conséquent, à l'issue de cet appariement, les analyses statistiques révèlent, d'une part, que les performances en vocabulaire augmentent avec le niveau [$F(2,203) = 52,74$, $p < .0000001$], les élèves du niveau A étant moins performants que les élèves du niveau B ($p < .00001$), eux-mêmes moins performants que ceux du niveau C ($p < .00001$) et, d'autre part que les différences entre arabophones et lusophones ne sont pas significativement différentes, quel que soit le niveau considéré (effet de la langue première non significative : $F(1,203) < 1$ et interaction langue première * niveau non significative : $F(2,203) = 2,27$, ns). Les caractéristiques de chaque groupe de participants, obtenues à la suite de cet appariement, figurent dans le tableau 49.

Tableau 49 : Caractéristiques de la population retenue pour l'expérimentation 3

	Arabo 3 ^e	Luso 3 ^e	Arabo 4 ^e	Luso 5 ^e	Arabo 5 ^e	Luso 7 ^e
Apprentissage FL2	160h 1 ^e année	224h 2 ^e année	436h 2 ^e année	512h 3 ^e année	724h 3 ^e année	780h 5 ^e année
Niv. Voca. FL2 /30	14,15 (3,24) Niveau A 14,33 (2,72)	14,53 (2,1)	17,05 (3,64) Niveau B 17,19 (3,43)	17,33 (3,02)	20,85 (3,89) Niveau C 20,22 (3,60)	19,04 (2,7)
Age chronologique	8;8 ans 106 mois (3,43)	9;5 ans 114 mois (3,71)	9;7 ans 117 mois (2,97)	11;5 ans 138 mois (4,38)	10;8 ans 130 mois (2,88)	13;7 ans 164 mois (4,66)
Effectif	41 13 garçons 28 filles	40 27 garçons 13 filles	42 21 garçons 21 filles	21 12 garçons 8 filles	42 25 garçons 17 filles	23 11 garçons 12 filles

1.2 Matériel

Tout d'abord, dans cette expérience, les données recueillies à l'aide des épreuves en arabe ou portugais langue première sont réexploitées. Mais au bloc d'épreuves présentées dans l'étude 1 et 2, s'ajoutent des instruments de mesures élaborés en langue française dont une grosse majorité conserve les mêmes principes que ceux construits dans les deux autres langues. Ainsi la description des épreuves propres au français présentera brièvement les tests ou épreuves qui sont similaires à celles déjà décrites et développera davantage les outils qui présentent une nouveauté par rapport aux précédentes expérimentations. Cette description suivra l'ordre suivant :

- L'évaluation des connaissances lexicales en français à l'aide du subtest de vocabulaire passif de Deltour et Hupkens (1980).
- Les évaluations des performances en lecture à l'aide d'une épreuve chronométrée de lecture de mots isolés à voix haute (LUM de Khomsi, 1999) et d'une épreuve de compréhension de phrases écrites (adaptation du subtest du LMC-R de Khomsi, 1999).
- L'évaluation des connaissances graphophonologiques, grapho-morphologiques (dérivationnelles et flexionnelles) et grapho-syntaxiques implicites et explicites, en français langue seconde, par le paradigme de détection d'intrus. Comme en arabe et en portugais, les épreuves ont été présentées oralement avec un support écrit.
- L'évaluation de la sensibilité aux différents éléments de la construction morphologique de pseudo-mots à l'aide d'une épreuve de plausibilité lexicale. Cette épreuve évalue en particulier les connaissances grapho-morphologiques implicites à différents morphèmes dérivationnels propres au français.

1.2.1 Epreuve de vocabulaire en français

Les connaissances lexicales des élèves sont évaluées à l'aide de la version du « test de vocabulaire passif » de Deltour et Hupkens (1980) destiné aux enfants natifs âgés de 3 à 5 ans. Ce test normalement prévu pour une passation individuelle a été adapté pour une passation collective (*cf.* annexe 26). Il s'agit d'un test de dénomination d'images. Les participants doivent choisir parmi 6 images en noir et blanc, celle correspondant au mot présenté oralement par l'expérimentateur. Ce test comporte 30 planches de 6 images. Les 30 mots oralisés sont soit des noms, soit des verbes à l'infinitif, de difficultés croissantes. Pour chaque planche, chaque élève entoure l'image choisie sur un livret qui lui a été distribué. 1 point est attribué à chaque bonne réponse sachant qu'il n'existe qu'une seule bonne réponse. Par exemple, les élèves doivent

entourer l'image correcte pour désigner le mot « parapluie » entre celles représentant une perruque, un parapluie, un imperméable, un parasol, une gazinière et un puits.

1.2.2 Epreuves de lecture en français

1.2.2.1 Epreuve de lecture de mots isolés en une minute

Les compétences de reconnaissance de mots écrits en français sont évaluées à l'aide du test de « lecture en une minute » (Khomsî, 1999). Il s'agit de lire en une minute à voix haute et les uns à la suite des autres, le maximum de mots d'une liste proposée. Cette liste est composée de 108 mots de différentes catégories grammaticales et de différentes longueurs, variant en termes de fréquence, de régularité et de complexité (cf. annexe 27). Ainsi, le niveau de performance est mesuré en fonction de l'exactitude et de la rapidité en lecture.

1.2.2.2 Epreuve de compréhension de phrases écrites

Cette épreuve est une adaptation du sub-test « compréhension de phrases » du test LMC-R de Khomsî (1999)¹⁶³. Elle est composée de 15 items, les participants devant, pour chaque item choisir l'image, parmi quatre, correspondant à une phrase écrite (cf. annexe 28). Pour 8 items, La réponse peut être déterminée par une stratégie « imagée » et pour 7 items, une stratégie « inférentielle » est nécessaire. Les participants peuvent en cas d'erreur modifier leur réponse et ne sont pas véritablement limités par le temps bien qu'ils soient invités à travailler le plus vite possible. 1 point est attribué à chaque bonne réponse.

1.2.3 Epreuves implicites de détection d'intrus

Comme en langue première les épreuves implicites de détection d'intrus se caractérisent par une consigne vague, du type « qu'est-ce qui va le moins avec ... ». L'unité linguistique ou le type d'analyse sur lequel ce choix doit se fonder ne sont ainsi pas spécifiés.

1.2.3.1 Epreuve implicite de détection d'intrus graphophonologique

Cette épreuve est destinée à évaluer la sensibilité des participants aux unités syllabiques et intrasyllabiques des mots (noms ou adjectifs) écrits en français langue seconde. Elle comporte 12 items. Pour chacun d'eux, les participants doivent choisir entre deux mots, celui « allant le moins » avec le mot cible (e.g. : qu'est ce qui va le moins avec « cheval » : « signal » ou « moteur » ?). 3 items portent sur la syllabe finale de mots trisyllabiques, 3 items sur la rime

¹⁶³ Une adaptation a été nécessaire de sorte que les items soient différents entre les épreuves en langue première et en français langue seconde. Ainsi, les items de l'épreuve originale qui avaient déjà été employés en langue première, selon une traduction plus ou moins stricte, ont été supprimés en français.

intrasyllabique de mots bisyllabiques, 3 sur le phonème consonantique final et 3 sur le phonème vocalique final de mots bisyllabiques. La structure phonologique et graphique, la catégorie grammaticale, les aspects sémantiques du lexique et la familiarité des paires de mots proposés ont été contrôlés sur l'ensemble du matériel. L'épreuve elle-même et les caractéristiques du matériel lexical sont présentées dans l'annexe 29. La randomisation et la cotation sont identiques à celles appliquées à cette même épreuve en arabe et en portugais.

1.2.3.2 *Epreuve implicite de détection d'intrus grapho-morpho-dérivationnel*

Cette épreuve est destinée à évaluer la sensibilité des participants à la construction morphologique en base / affixe de mots français (noms ou verbes). Elle comporte 12 items, chacun constitué d'un mot cible dérivé et d'une paire de mots dont l'un est construit à l'aide du même affixe et l'autre, jouant le rôle d'intrus, n'est pas construit ou utilise une base non identifiable (*e.g.* facteur). Les participants doivent choisir entre les deux mots, celui « allant le moins » avec le mot cible (*e.g.*: qu'est ce qui va le moins avec « revoir » : « replonger » ou « refuser » ?)

La moitié des items concerne la formation de substantifs par ajout de suffixe :

- 3 items portent sur le suffixe « -ier » permettant de former des noms d'agent ou d'instrument.
- 3 items portent sur le suffixe « -eur » permettant de former des noms de d'agent ou de qualité.

L'autre moitié des items concerne la formation de verbes par ajout de préfixe :

- 3 items portent sur le préfixe « dé- » dont la valeur sémantique marque l'opposition ou la cessation par rapport au sens de la base verbale.
- 3 items portent sur le préfixe « re- » dont la valeur sémantique marque la répétition, la réversibilité ou le retour à une forme ou un état antérieur¹⁶⁴.

Les aspects sémantiques des items, la structure phonologique et graphique ainsi que la familiarité des paires de mots proposés ont été contrôlés. L'épreuve elle-même, et les caractéristiques du matériel lexical figurent dans l'annexe 30. La randomisation et la cotation sont identiques à celles appliquées à cette même épreuve en arabe et en portugais.

¹⁶⁴ Ces 4 affixes ont été choisis pour leur productivité dans la langue mais aussi en fonction de la possibilité de pouvoir trouver parmi ces mots des pseudo-affixés. Leur ressemblance avec les affixes de la langue portugaise a aussi été considérée. Il se trouve que bien qu'il existe une proximité entre ces affixes dans les deux langues, ils ne se prononcent toutefois pas de façon identique et ne s'écrivent la plupart du temps pas de la même façon.

1.2.3.3 *Epreuve implicite de détection d'intrus grapho-morpho-flexionnel*

Cette épreuve est destinée à évaluer la sensibilité des participants à la structure flexionnelle de mots écrits en portugais (noms, adjectif ou verbes). Elle comporte 12 items, chacun constitué d'un mot cible et d'une paire de mots dont l'un fait apparaître une structure flexionnelle identique au mot cible et l'autre, jouant le rôle d'intrus, une forme qui malgré les apparences phonographiques est, en terme de structure flexionnelle, différente. Les participants doivent choisir entre les deux mots, celui « allant le moins » avec le mot cible (e.g.: qu'est ce qui va le moins avec « gratuite » : « absente » ou « adulte » ?).

- 6 items concernent la formation du féminin de noms ou d'adjectifs. Les intrus terminent par le même bigramme que celui des mots cibles et des mots fléchis (i.e. « te ») mais sont neutres (e.g. brute).

- 6 items concernent les désinences verbales de la conjugaison du présent de l'indicatif : 3 items portent sur la première personne du pluriel et trois sur la deuxième personne du pluriel. Les intrus sont systématiquement des verbes du troisième groupe conjugués au futur. Pour qu'ils ne soient pas reconnaissables par leur terminaison spécifique, les seconds verbes des paires sont des verbes dont la racine termine par les mêmes graphèmes précédant la désinence, ce qui permet que la dernière syllabe soit la même pour les deux verbes (e.g. « couvrons » –présent- et « suivrons » –futur-).

Les aspects sémantiques des items, la structure phonologique et graphique, la familiarité des paires de mots proposés, ainsi que le groupe des verbes pour ce qui concerne les items sur les désinences ont été contrôlés. L'épreuve elle-même et les caractéristiques du matériel lexical figurent dans l'annexe 31. La randomisation et la cotation sont identiques à celles appliquées à cette même épreuve en arabe et en portugais.

1.2.3.4 *Epreuve implicite de détection d'intrus grapho-syntaxique*

Cette épreuve est destinée à évaluer la sensibilité des participants à la structure syntaxique de phrases écrites en français. Elle comporte 12 items, chacun constitué d'une phrase cible et d'une paire de phrases dont l'une partage la même structure syntaxique que la phrase cible et l'autre non. Les participants doivent choisir entre deux phrases, celle « allant le moins » avec la phrase cible (e.g. : qu'est ce qui va le moins avec « il a peint » : « il a honte » ou « il a ri » ?).

- 6 items relèvent de la construction transitive de phrase du type SVO. Ici les cibles et les phrases syntaxiquement similaires aux cibles utilisent des verbes transitifs directs alors que les phrases intruses emploient un verbe transitif indirect suivi de la préposition « de ». De sorte que la simple présence de la préposition ne soit pas suffisante pour déceler l'intrus, une contraction

avec l'article pluriel suivant permet de la confondre avec l'article indéfini pluriel présent dans les autres phrases de l'item (e.g. « il chante des (article défini pluriel) chansons » –forme transitive directe et « il rentre des (de les) vacances » – forme transitive indirecte).

- Pour 3 items, les phrases concernent l'usage du verbe « avoir ». Dans ces items, la phrase cible et son homologue sont au passé composé alors que la phrases intruse contient une locution formée à l'aide du verbe avoir (e.g. il a peur).

- Pour 3 autres items, les phrases concernent l'usage du verbe « être ». Dans ces items, la phrase cible et son homologue sont composés du verbe être et d'un attribut du sujet (e.g. il est tranquille) alors que la phrases intruse utilise un verbe conjugué au passé composé à l'aide du verbe « être » (e.g. il est tombé).

Les aspects sémantiques des items, le nombre de syntagmes, la structure phonologique et graphique des paires de phrases, la familiarité du lexique, le genre, le nombre des noms et la forme des verbes composant les paires de phrases ont été contrôlés. L'épreuve et les caractéristiques du matériel lexical sont présentées dans l'annexe 32. La randomisation et la cotation sont identiques à celles appliquées à cette même épreuve en arabe et en portugais.

1.2.4 Epreuves explicites de détection d'intrus

Les épreuves explicites conservent le même principe de détection d'intrus entre deux stimuli mais font appel à une consigne qui explique sur quoi l'attention doit porter. Selon l'épreuve, les aspects phonologiques, morphologiques ou syntaxiques sont donc préalablement expliqués. Le choix des participants est déterminé cette fois par une prise en compte consciente des aspects manipulés dans l'épreuve.

Ces épreuves sont construites sur le même mode que leur homologue implicite. Le nombre d'items, les caractéristiques des items et le mode de contrôle sont donc exactement identiques à ceux des épreuves implicites décrites précédemment, seules les consignes diffèrent. Nous ne présenterons donc, dans les paragraphes suivants, que le type de compétences évaluées accompagné d'exemples précisant ce qui est demandé aux participants.

1.2.4.1 *Epreuve explicite de détection d'intrus graphophonologique*

Cette épreuve a pour objectif d'évaluer les connaissances graphophonologiques explicites permettant de discriminer consciemment les éléments graphophonologiques de mots écrits en portugais. Les participants doivent déterminer entre deux mots, lequel ne termine pas par le même son qu'un mot cible. Par exemple, qu'est ce qui ne se termine pas par le même son que

« vitrine » : « racine » ou « mesure » ? L'épreuve et les caractéristiques du matériel lexical sont présentées dans l'annexe 33.

1.2.4.2 *Epreuve explicite de détection d'intrus grapho-morpho-dérivationnel*

Cette épreuve a pour objectif d'évaluer les connaissances morpho-dérivationnelles explicites permettant d'identifier la construction morpho-dérivationnelle de mots dérivés de façon à segmenter les mots en base / affixe. Dans cette épreuve, les participants doivent déterminer parmi deux mots proposés, celui qui, malgré les apparences phonographiques, ne fait pas apparaître la même construction morphologique qu'un mot cible. Il leur est par exemple demandé quel est le mot qui ne vient pas d'un autre mot comme « cendrier » : « boîtier » ou « cahier » ? L'épreuve et les caractéristiques du matériel lexical figurent en annexe 34.

1.2.4.3 *Epreuve explicite de détection d'intrus grapho-morpho-flexionnel*

Cette épreuve a pour objectif d'évaluer les connaissances grapho-morpho-flexionnelles explicites permettant d'identifier consciemment les aspects syntaxiques des mots tels que le genre des substantifs ou la conjugaison des verbes en fonction des conventions de la langue. Dans cette épreuve, les participants doivent déterminer, parmi deux propositions et malgré les apparences phonographiques, le nom qui n'est pas assigné d'une marque de genre, ou le verbe qui n'est pas marqué par la même conjugaison que le mot cible. Il leur est par exemple demandé quel est le mot qui n'est pas seulement un féminin, qui ne sert pas seulement pour les filles, comme « forte » : « cuite » ou « triste » ? Ou encore quel est le verbe qui n'est pas conjugué comme, ou qui n'est pas au même temps que « vous cueillez » : « vous pleurez » ou « vous lirez » ? L'épreuve et les caractéristiques du matériel lexical sont présentées dans l'annexe 35.

1.2.4.4 *Epreuve explicite de détection d'intrus grapho-syntaxique*

Cette épreuve a pour objectif d'évaluer les connaissances grapho-syntaxiques explicites permettant d'analyser consciemment la structure syntaxique, en particulier celle du syntagme verbal, de phrases écrites. Dans cette épreuve, les participants doivent déterminer, parmi deux propositions et malgré les apparences phonographiques, la phrase pour laquelle le syntagme verbal n'est pas composé de la même façon que dans la phrase cible. Plus précisément, après explications, il leur est, par exemple, demandé quelle est la phrase qui n'est pas composée de deux verbes, n'est pas au même temps que « il a menti au voisin » : « il a joué au ballon » ou « il a droit au dessert » ? Ou encore, dans quelle phrase trouve-t-on le mot « de » caché, qu'il n'y a pas dans « il raconte des blagues » : « il reprend des pommes » ou « il revient des courses » ? L'épreuve et les caractéristiques du matériel lexical sont présentées dans l'annexe 36.

1.2.5 Épreuve grapho-morphologique implicite de plausibilité lexicale écrite en français langue seconde

Cette épreuve évalue la sensibilité à différents éléments utiles pour former des mots dérivés français, tels que les affixes (préfixes et suffixes), les bases et leur combinaison. Elle est proposée à l'écrit, sans limite de temps pour la réaliser et comporte 20 items. Chaque item est constitué d'une paire de pseudo-mots dont la construction morphologique est variable. Les participants doivent déterminer lequel des deux ressemble le plus à un mot français, en cochant la case sous le pseudo-mot écrit de leur choix. Il n'y a donc pas de bonnes réponses, le choix des participants permettant d'inférer les unités morphologiques, appartenant à la langue française, qu'ils repèrent. Les constructions morphologiques mises en compétition entre les deux pseudo-mots présentés, permettent de distinguer 5 types d'items (l'ordre droite, gauche, des pseudo-mots et la succession des items dans l'épreuve ont été randomisés) :

- 4 items présentent simultanément un pseudo-mot composé de deux morphèmes identifiables (affixe + base) et un pseudo-mot contenant seulement un affixe du français repérable (2M/A : *dépunir – déconvir*).
- 4 items présentent un pseudo-mot composé de deux morphèmes identifiables (affixe + base) et un pseudo-mot contenant seulement une base du français repérable (2M/B : *débouder – sobouder*).
- 4 items proposent un pseudo-mot composé seulement d'une base identifiable et un pseudo-mot contenant seulement un affixe du français (B/A : *sovoler – découner*).
- 4 items proposent un pseudo-mot composé seulement d'une base identifiable et un pseudo-mot ne relevant pas d'une composition morphologique (B/R : *socalmer – calsommer*).
- 4 items proposent un pseudo-mot composé seulement d'un affixe et un pseudo-mot ne relevant pas d'une composition morphologique (A/R : *dévoncir – vécondir*).

Hormis pour les items B/R, les affixes utilisés sont soit le préfixe « dé- » (la moitié des items), soit le suffixe « -ier » (l'autre moitié des items). Ce sont ainsi les deux mêmes affixes (« dé » ou « ier ») qui l'on retrouve dans les deux pseudo-mots des items 2M/A et les deux mêmes bases dans les deux pseudo-mots des items 2M/B. Quelles que soient les combinaisons étudiées, le nombre de syllabes, de phonèmes, de graphèmes, de même que la fréquence de bigrammes et trigrammes ont été contrôlés. L'épreuve est présentée en annexe 37. La cotation systématique d'un point à l'un des pseudo-mots et de 0 pour l'autre est choisie, pour chaque type d'item, de façon arbitraire puisqu'il n'y a pas de bonnes ou de mauvaises réponses. Ce choix s'est toutefois effectué de façon à rendre compte des orientations vers les constructions morphologiques les plus plausibles ou visibles (la combinaison base/affixe par rapport à la présence d'un seul élément, la

présence d'une base par rapport à un affixe ou à l'absence de morphème, la présence d'un affixe par rapport à l'absence de morphème).

1.3 Procédure

Les différentes sessions qui ont permis d'administrer l'ensemble des épreuves (en langue première et en français langue seconde) dans chaque classe se sont déroulées de façon à privilégier l'homogénéité des langues travaillées, tout en respectant le rythme de la classe et sans qu'une épreuve commencée ne soit inachevée.

L'ordre constant de présentation des épreuves ou groupes d'épreuves est le suivant¹⁶⁵ :

- Epreuves implicites de détection d'intrus en français langue seconde. La passation est collective et l'ordre de présentation des 4 épreuves graphophonologique, grapho-morpho-dérivationnelle, grapho-morpho-flexionnelle et grapho-syntaxique est randomisé. Avant de débiter la première épreuve, un exemple et deux items d'entraînement sans *feed-back* (*cf.* annexe 38), pour lesquels la distinction partage/non partage d'une unité linguistique avec le stimulus cible est absente, sont proposés.
- Epreuve de plausibilité lexicale écrite en français langue seconde. La passation est collective mais l'expérimentateur, après avoir expliqué la consigne n'intervient pas au cours de sa résolution, qui s'effectue en silence.
- Epreuves implicites de détection d'intrus en langue première, soit en arabe, soit en portugais, selon la population (*cf.* paragraphe 1.3 des chapitres 7 et 8).
- Epreuve d'identification de la construction morphologique de pseudo-mots (*cf.* paragraphe 1.3 des chapitres 7 et 8).
- Epreuves explicites de détection d'intrus en langue première, soit en arabe, soit en portugais, selon la population (*cf.* paragraphe 1.3 des chapitres 7 et 8).
- Epreuves explicites de détection d'intrus en français langue seconde. Là encore, la passation est collective. Pour chaque classe, l'ordre de présentation des épreuves graphophonologique, grapho-morpho-dérivationnelle, grapho-morpho-flexionnelle et grapho-syntaxique est randomisé et diffère de l'ordre fixé pour les épreuves implicites et les épreuves en langue première. Quatre exemples et deux items d'entraînements avec correction après chaque item (*cf.* annexes 33 à 36) sont utilisés pour expliquer la distinction partage/non partage d'une unité linguistique avec le

¹⁶⁵ La procédure des épreuves en langue première ayant déjà été présentée, le lecteur pourra se référer pour plus de détails aux paragraphes 1.3 des chapitres 7 et 8. Compte tenu de l'identité des modes de présentation des épreuves de détection d'intrus en langue première et en français langue seconde, la description du protocole de passations de ces épreuves en français est succincte car elle peut être éclaircie par la relecture des paragraphes 1.3 des chapitres précédents

stimulus cible sur laquelle les participants doivent porter leur attention. Les unités choisies ne sont pas les mêmes que celles utilisées dans l'épreuve mais appartiennent au même domaine linguistique.

- Epreuve de compréhension de phrases écrites en français langue seconde. La passation est collective mais l'expérimentateur n'intervient pas au cours de sa résolution, qui s'effectue en silence. Avant de débiter l'épreuve, un item modèle corrigé en groupe et un item d'essai sans *feed-back* sont présentés sur la feuille d'exercice.
- Epreuve de vocabulaire en français langue seconde, en passation collective. Avant de débiter l'épreuve, deux items d'entraînements sont utilisés afin de s'assurer que tous les enfants ont compris la consigne.
- Epreuve de compréhension de phrases écrites en langue première (*cf.* paragraphe 1.3 des chapitres 7 et 8).
- Tests de lecture de mots dans les deux langues, en passation individuelle. L'ordre d'administration des tests en langue première et en français est choisi de façon aléatoire.

Les épreuves administrées en passations collectives ont nécessité 6 sessions d'environ 45 minutes sans dépasser le rythme d'une session par semaine. Des entretiens individuels, d'une quinzaine de minutes au total par élève, ont été menés par deux expérimentatrices différentes, l'une native du pays, qui administrait le test de lecture de mots en langue première et récoltait des informations sur le statut du français dans la famille de l'enfant, et l'autre française faisant passer le test de lecture de mots en français. Ces expérimentatrices ont également participé au déroulement de l'ensemble des épreuves collectives. La présence de personnes formées à cette recherche était en effet importante pour qu'il soit toujours possible d'expliquer en langue première, les consignes des épreuves françaises.

Le tableau 50 récapitule les types d'épreuves évaluant les connaissances sur les dimensions graphophonologique et grapho-sémantique.

Tableau 50 : Epreuves évaluant les connaissances sur les dimensions graphophonologique et grapho-sémantique en langue première (arabe vs. portugais) et en français langue seconde.

	Dimension graphophonologique	Dimension grapho-sémantique		
		Morpho-dérivationnelle	Morpho-flexionnelle	Syntaxe
<p><u>Epreuves implicites de détection d'intrus en français langue seconde</u>^{1 2}</p> <p>Présentation à l'oral + support écrit Entraînement au paradigme lors de la 1^e épreuve (2 items)</p>	Consigne implicite commune : Qu'est ce qui va le moins avec :			
	<p>cheval signal ou <u>moteur</u></p> <p><i>cf. annexe 29</i></p>	<p>revoir replonger ou <u>refuser</u></p> <p><i>cf. annexe 30</i></p>	<p>gratuite absente ou <u>adulte</u></p> <p><i>cf. annexe 31</i></p>	<p>il a peint il a honte ou <u>il a ri</u></p> <p><i>cf. annexe 32</i></p>
<p><u>Epreuve implicite de plausibilité lexicale écrite en français langue seconde</u>²</p>		<p>Qu'est ce qui ressemble le plus à un mot français ?</p> <p><i>dépunir</i> ou <i>déconvir</i></p> <p><i>cf. annexe 37</i></p>		
<p><u>Epreuves implicites de détection d'intrus en langue première</u>^{1 2}</p>	<p><i>cf. tableau 5 et annexe 3 pour la version arabe</i> <i>cf. tableau 27 et annexe 16 pour la version portugaise</i></p>	<p><i>cf. tableau 5 et annexe 4 pour la version arabe</i> <i>cf. tableau 27 et annexe 17 pour la version portugaise</i></p>	<p><i>cf. tableau 5 et annexe 5 pour la version arabe</i> <i>cf. tableau 27 et annexe 18 pour la version portugaise</i></p>	<p><i>cf. tableau 5 et annexe 6 pour la version arabe</i> <i>cf. tableau 27 et annexe 19 pour la version portugaise</i></p>
<p><u>Epreuve d'identification de la construction morphologique de pseudo-mots</u>²</p>		<p><i>cf. tableau 5 et annexe 11 pour la version arabe</i> <i>cf. tableau 27 et annexe 24 pour la version portugaise</i></p>		
<p><u>Epreuves explicites de détection d'intrus en langue première</u>^{1 2}</p>	<p><i>cf. tableau 5 et annexe 7 pour la version arabe</i> <i>cf. tableau 27 et annexe 20 pour la version portugaise</i></p>	<p><i>cf. tableau 5 et annexe 8 pour la version arabe</i> <i>cf. tableau 27 et annexe 21 pour la version portugaise</i></p>	<p><i>cf. tableau 5 et annexe 9 pour la version arabe</i> <i>cf. tableau 27 et annexe 22 pour la version portugaise</i></p>	<p><i>cf. tableau 5 et annexe 10 pour la version arabe</i> <i>cf. tableau 27 et annexe 23 pour la version portugaise</i></p>
<p><u>Epreuves explicites de détection d'intrus</u>^{1 2}</p> <p>Présentation à l'oral + support écrit Entraînement systématique à la tâche (2 items)</p>	<p>Consigne explicite : Qu'est ce qui ne se termine pas par le même son que :</p>	<p>Consigne explicite : Qu'est ce qui ne vient pas d'un autre mot comme :</p>	<p>Exemple de consigne explicite : Qu'est ce qui n'est pas conjugué comme :</p>	<p>Exemple de consigne explicite : Quelle phrase n'a pas deux verbes comme :</p>
	<p>vitrine racine ou <u>mesure</u></p> <p><i>cf. annexe 33</i></p>	<p>cendrier boîtier ou <u>cahier</u></p> <p><i>cf. annexe34</i></p>	<p>vous cueillez vous pleurez ou <u>vous lirez</u></p> <p><i>cf. annexe 35</i></p>	<p>il a menti au voisin il a joué au ballon ou <u>il a droit au dessert</u></p> <p><i>cf. annexe 36</i></p>

¹ Randomisation de l'ordre de présentation des 4 épreuves

² Randomisation de l'ordre des items et des réponses dans chaque item

En grisé : les épreuves en français

2.Résultats

2.1 Démarche statistique

La démarche statistique sous-tendant l'analyse des résultats a été élaborée de manière à mener, lorsque cela est réalisable et a un sens, certaines comparaisons entre arabophones et lusophones.

Nous nous sommes, tout d'abord, attachée à examiner les différences entre les performances des apprenants en fonction de leur langue première, tout en tenant compte du niveau de français atteint au cours des apprentissages. Les mêmes tests ou épreuves construites en français ayant été administrés de façon identique aux groupes d'apprenants, l'utilisation d'analyses multivariées, prenant en compte le facteur langue première mais aussi le niveau de français¹⁶⁶ est envisageable.

- Ces analyses débutent par l'examen des performances en lecture (en lecture de mots¹⁶⁷ puis en compréhension de phrase). Malgré un niveau de compétences orales relativement similaire, la possibilité de voir apparaître, à l'issue de ces deux analyses, des différences importantes entre arabophones et lusophones est prévisible. Comme c'est bien le cas, ces deux dimensions de la lecture ont été incluses dans la suite des analyses comme des covariables, de façon à contrôler leurs effets. Il est en effet incontestable qu'elles occupent, dans le domaine des connaissances graphophonologiques, grapho-morphologiques et grapho-syntaxiques, une position prioritaire dans l'explication des résultats obtenus.

- Pour ce qui concerne les épreuves implicites et explicites de détection d'intrus, l'emploi du même paradigme permet d'englober ce sous-ensemble d'épreuves dans une même analyse. Notre objectif était donc de déterminer l'influence des caractéristiques de la langue première en fonction du niveau de français (niveau A, B ou C), du degré de contrôle impliqué par la tâche (implicite *vs.* explicite) et du domaine linguistique (graphophonologique, grapho-morpho-dérivationnelle, grapho-morpho-flexionnelle, grapho-syntaxique), en contrôlant les effets des performances en lecture (en lecture de mots et en compréhension). Malheureusement, les critères d'exclusion relatifs aux réponses aléatoires¹⁶⁸ et aux scores

¹⁶⁶ Lorsque nous parlons de niveau de français ou tout simplement de niveau, nous entendons les trois niveaux distingués après appariement et rendant compte d'un niveau de vocabulaire en constante augmentation. A ce niveau de vocabulaire, peut également être rapproché le temps d'apprentissage du français, bien que ce dernier ne soit pas égal entre arabophones et lusophones.

¹⁶⁷ Le critère d'homogénéité de la variance en lecture de mots n'étant pas respecté (test de Levene significatif : $F(5,200) = 6,00, p < .00001$), des tests non paramétriques (U de Mann-Whitney et ANOVA de Kruskal-Wallis) ont été choisis pour cette première analyse.

¹⁶⁸ Toujours, car la résolution de ces épreuves de choix forcé donne lieu à un choix entre deux propositions.

plafonds ont limité les analyses inférentielles prévues. Lorsque pour certaines épreuves, les résultats n'atteignent ni chez les arabophones ni chez les lusophones des scores supérieurs au niveau du hasard, et que de surcroît, pour les deux populations, ces scores ne se distinguent pas du niveau du hasard pour les trois niveaux, aucune analyse n'a été réalisée. Dans le cas de figure plus complexe où pour une des populations d'apprenants, sans distinction de niveau, les réponses à une épreuve, ne sont pas aléatoires, mais le sont pour l'autre population, les analyses ont malgré tout été menées, à condition toutefois qu'au moins un niveau parmi les trois ait des réponses supérieurs au hasard. En effet, en dépit d'un score global faible et proche du hasard, un ou plusieurs groupes peuvent avoir des résultats non négligeables, ce qui mérite d'être pris en compte dans les comparaisons.

Ainsi, plusieurs analyses en sous-ordre à trois facteurs (langue première, niveau et domaine linguistique ou degré de contrôle), avec toujours comme covariables les performances en lecture de mots et en compréhension, ont été requises. Les analyses *a posteriori* ont été systématiquement menées à l'aide du test de Newman-Keuls.

- Dans l'objectif d'examiner si les arabophones et les lusophones, maîtrisent de façon similaire les mêmes unités graphophonologiques en français et comment cette maîtrise évolue chez ces deux populations, une MANCOVA à trois facteurs (langue première, niveau, unité graphophonologique) et deux covariables (les performances en lecture de mots et en compréhension de phrases) a été menée mais seulement sur le niveau implicite des connaissances. En effet, les effets plafonds obtenus à l'épreuve graphophonologique explicite quels que soient les groupes, n'a pas permis d'inclure l'effet du degré de contrôle. Mais, contrairement aux données obtenues en langue première, les critères de normalité et d'homocédasticité des résultats à l'épreuve implicite sont respectés et permettent de mener des tests paramétriques.

- Afin d'approfondir les effets de la langue première sur la sensibilité à la morphologie des mots en français, une analyse spécifique a été mise au point de manière à détecter les éléments (base et/ou affixe de la langue française) de pseudo-mots construits, les plus repérables par les apprenants lors d'une tâche de plausibilité. Il a semblé délicat d'effectuer des analyses multivariées portant sur l'ensemble de cette épreuve car plusieurs items utilisent une structure différente et rendent compte de la sensibilité à certaines unités morphographiques par rapport à d'autres. Le choix des élèves s'effectue, ici encore, entre deux pseudo-mots dont la construction varie. L'intérêt de cette épreuve réside donc à comparer, pour chaque type d'item, le pourcentage de réponse attribuée à un des deux pseudo-mots par rapport à l'autre. La comparaison entre arabophones et lusophones en fonction du niveau et en contrôlant les effets des performances en lecture, a eu lieu lorsque,

pour les deux populations, les choix sont significativement plus marqués pour un de ces pseudo-mots (réponse moyenne différente de 50%). Dans le cas contraire, les réponses pouvant être liées au hasard, il a comme précédemment été vérifié si ce phénomène touche l'ensemble des niveaux. Compte tenu du faible nombre d'items spécifiques à une comparaison entre deux types de construction interne, se baser essentiellement sur l'effet du niveau nous est apparu insuffisant. Ainsi, tout en considérant les résultats avec précaution nous avons systématiquement examiné les valeurs aux tests t de comparaison au hasard pour chaque groupe pris séparément. La présentation des résultats commence par les items pour lesquels les MANCOVAs n'ont pas pu être effectuées.

L'ensemble de ces analyses a également permis de sélectionner les épreuves pouvant être retenues dans les études corrélatives, factorielles et de régressions ultérieures.

Dans le cas où les objectifs portent sur les associations entre connaissances (en FL2 ou entre L1 et L2) ou entre les connaissances et le niveau de lecture en français, les effets des acquisitions préalables en L1 ont été mis en évidence en décrivant les résultats aux analyses corrélationnelles, factorielles ou de régressions obtenus successivement par les 3 groupes d'arabophones d'une part, et les 3 groupes de lusophones d'autre part. Il s'agira de voir apparaître certaines tendances dans les ressemblances ou différences quant aux variables qui se dégagent. Une analyse statistique des effets de la langue première, et du niveau n'est, en effet, pas souhaitable dans le cadre de cette recherche pour plusieurs raisons. D'une part car il s'est avéré que les résultats sélectionnés ne sont jamais totalement identiques pour chaque groupe, les variables prises en considération dans les analyses étant par conséquent toujours en partie différentes. D'autre part car bien que les niveaux A, B et C, correspondent à un appariement en termes de connaissances lexicales entre arabophones et lusophones, ils englobent des élèves d'âge, de classe et donc d'exposition à l'écrit dans la langue première très différents entre ces deux populations. Ceci est particulièrement important pour comprendre les relations entre les connaissances développées en langue première et en français langue seconde mais aussi pour analyser les contributions de certaines connaissances en langue première à l'apprentissage de la lecture en français langue seconde. Il est ainsi essentiel que les discussions tentant de rapprocher les observations obtenues par les arabophones et les lusophones soulignent ces nombreuses variations.

Les analyses réalisées suivent donc l'organisation suivante :

Dans un premier temps, les liens existant entre les scores aux différentes épreuves, sélectionnées et mesurées en français, sont étudiés par le biais d'analyses corrélationnelles et factorielles (ACP). L'objectif principal est ici de déterminer sur quoi se fonde, au fur et à

mesure des apprentissages, l'acquisition de la dimension grapho-sémantique du système écrit français. Les résultats des 3 niveaux arabophones, puis des 3 niveaux lusophones sont présentés successivement.

Dans un deuxième temps, le lien existant entre langue première et langue seconde sur les performances à chaque type d'épreuve, de nature implicite ou explicite, est étudié successivement à condition bien sûr que les résultats obtenus dans les deux langues soient analysables (c'est-à-dire supérieurs au hasard et sans apparition d'effets plafonds). Ces analyses de corrélations ont également été complétées par d'autres analyses examinant toujours le lien entre L1 et français L2 mais en élargissant les domaines de connaissances pouvant être associées. Lorsque cela est nécessaire, des analyses de régression viennent également compléter les analyses corrélatives. L'objectif général est ici d'examiner dans quelle mesure et en fonction de quels facteurs, les connaissances développées en L1 ou en L2 entretiennent une relation d'interdépendance. Les résultats des 3 niveaux arabophones, puis des 3 niveaux lusophones sont présentés successivement.

Dans un troisième temps, nous nous sommes attachée à rechercher les performances (toujours parmi celles retenues) pouvant déterminer le niveau de lecture en français L2 des apprenants, en fonction de leur langue première et de leur niveau. Sur la variable « lecture de mots » tout d'abord, des analyses de corrélation ont été menées pour identifier le degré d'association entre les scores en lecture de mots et les mesures en L1 ainsi qu'en L2 dans les différents domaines retenus. Les variables obtenant des coefficients de corrélation significatifs ont alors été entrées dans un modèle de régression pas à pas¹⁶⁹. Ces deux analyses sont présentées successivement pour les 3 niveaux arabophones puis pour les 3 niveaux lusophones. Enfin la même démarche a été exactement employée pour ce qui concerne la variable « compréhension en lecture ».

¹⁶⁹ Dans les cas où une seule corrélation est identifiée, une régression simple est alors suffisante.

2.2 Analyses des performances obtenues

2.2.1 Epreuve de lecture de mots : effet de la L1 et du niveau

Les résultats obtenus à l'épreuve de lecture de mots français en une minute en fonction de la langue première et du niveau sont présentés dans le tableau 51.

Tableau 51 : Performances moyennes, en nombre de mots lus correctement en une minute (et écarts-types) au test LUM en FL2 en fonction de la L1 et du niveau

	Niveau A	Niveau B	Niveau C	Total
Arabophones	15,27 (8,41) N = 41	31,34 (12,94) N = 41	42,79 (16,78) N = 42	29,90 (17,33)
Lusophones	30,18 (8,98) N = 40	49,9 (8,58) N = 20	54,14 (9,62) N = 22	41,42 (14,29)
Tous groupes	22,63 (11,44)	37,43 (14,57)	46,69 (15,61)	

Le critère d'homogénéité de la variance en lecture de mots entre les groupes n'étant pas respecté (test de Levene significatif: $F(5,200) = 6,00$; $p < .00001$), des tests non paramétriques ont été utilisés. Les résultats au test de Mann-Whitney mené sur le facteur langue première {arabophones vs. lusophones} mettent en évidence une infériorité des arabophones par rapport aux lusophones ($U = 3023,00$, $z = 4,92$, $p < .0001$): les arabophones lisent significativement moins de mots français correctement en une minute que les lusophones.

Les résultats au test de Kruskal-Wallis mené sur le facteur niveau montrent également que ce dernier est significatif [$H(2, N = 206) = 71,4004$, $p < .001$]. Les comparaisons entre les trois niveaux indiquent que globalement, les élèves de niveau A sont moins performants que les élèves de niveau B ($p < .01$), eux-mêmes moins performants que les élèves de niveau C ($p < .01$), ces derniers étant de fait également plus performants que les élèves de niveau A ($p < .000001$). Cette augmentation en fonction du niveau est d'ailleurs observé chez les arabophones [$H(2, N = 124) = 53,3810$, $p < .0001$] comme chez les lusophones [$H(2, N = 82) = 50,7342$, $p < .0001$].

Plus précisément, chez arabophones, les élèves de niveau A ont en effet des scores plus faibles que les élèves de niveau B et C (respectivement $U = 253,50$, $z = -5,44$, $p < .000001$ et $U = 153,50$, $z = -6,44$, $p < .000001$) et les élèves de niveau B, des scores plus faibles que les élèves de niveau C ($U = 510,00$, $z = -3,20$, $p < .001$). Chez les lusophones, également, les scores sont plus faibles au niveau A qu'au niveau B et C (respectivement $U = 47,50$, $z = -5,53$, $p < .000001$ et $U = 30,00$, $z = -6,03$, $p < .000001$). Cependant, les scores ne sont pas significativement plus faibles au niveau B qu'au niveau C ($U = 168,00$, $z = -1,31$, ns).

De surcroît, les comparaisons deux à deux pour chaque niveau indiquent que les scores en lecture de mots des arabophones sont toujours plus faibles que ceux des lusophones, que les élèves soient de niveau A ($U = 184,00$ $z = 6,01$, $p < .0001$), de niveau B ($U = 805,00$ $z = 4,72$, $p < .0001$) ou de niveau C ($U = 270,00$ $z = 2,71$, $p < .01$).

Ces analyses permettent donc de constater qu’au cours des premières années d’apprentissage du français, les arabophones rencontrent plus de difficultés pour lire les mots en français que leurs pairs lusophones, même si leurs connaissances lexicales ne sont pas plus faibles. Leurs performances en lecture de mots ne cessent d’ailleurs d’augmenter alors que chez les lusophones cette augmentation est surtout visible entre le niveau A et B.

2.2.2 Epreuve de compréhension de phrase écrites : effet de la L1 et du niveau

Une analyse pour plan à deux facteurs (langue première {arabophones vs. lusophones} et niveau {A, B et C}) a été menée sur les scores à l’épreuve de compréhension de phrases écrites en français L2. Les résultats comparés lors de cette analyse sont présentés dans le tableau 52.

Tableau 52 : Performances moyennes, sur 15 (et écart-types) au test de compréhension de phrases écrites en FL2 en fonction de la L1 et du niveau

	Niveau A	Niveau B	Niveau C	Total
Arabophones	4,78 (2,29) N = 40	5,76 (2,24) N = 41	8,45 (2,89) N = 42	6,36 (2,93)
Lusophones	5,29 (1,92) N = 38	7,14 (2,06) N = 21	9,43 (1,90) N = 23	6,93 (2,06)
Tous groupes	5,03 (2,12)	6,23 (2,26)	8,80 (2,61)	

Les résultats permettent de constater :

- un effet de langue première [$F(1,199) = 8,28$; $p < .01$] : les résultats des arabophones sont globalement inférieurs à ceux des lusophones ;
- un effet du niveau [$F(2,199) = 49,23$; $p < .001$] : les performances des élèves de niveau A sont inférieures à celles des élèves de niveau B ($p < .01$), elles-mêmes inférieures à celles de niveau C ($p < .0001$), ces dernières étant de fait supérieures à celles de élèves de niveau A ($p < .00001$) ;
- mais pas d’interaction L1 * niveau [$F(2,199) < 1$, ns].

Comme cela est apparu en lecture de mots, les arabophones ont plus de difficultés que les lusophones pour comprendre les phrases écrites en français langue seconde, quel que soit le niveau et en dépit de leurs bonnes connaissances lexicales. Par ailleurs, pour les deux groupes, les capacités de compréhension augmentent avec l'avancée dans les apprentissages, sans atteindre au bout de quelques années un niveau maximal

Il apparaît donc que les niveaux de lecture de mots et de compréhension de phrases écrites en français sont différents entre les groupes arabophones et lusophones, pourtant appariés sur leurs connaissances lexicales. Ainsi, il s'est avéré nécessaire d'entrer ces deux facteurs comme co-variables dans les analyses suivantes.

2.2.3 Epreuves implicites et explicites de détection d'intrus

2.2.3.1 Sélection des épreuves en fonction des résultats obtenus

Le tableau 53 récapitule les performances moyennes obtenues par les arabophones et les lusophones, sans distinction de niveau, aux épreuves implicites et explicites de détection d'intrus graphophonologique, grapho-morpho-dérivationnel, grapho-morpho-flexionnel et grapho-syntaxique.

Tableau 53 : Performances moyennes de réussite sur 12 (et écart-types) obtenues aux épreuves implicites et explicites de détection d'intrus en fonction de la langue première et du niveau et tests-t de comparaison à la valeur 6 (niveau du hasard).

Epreuves de détection d'intrus	Arabophones	Comparaison au hasard (6/12)	Lusophones	Comparaison au hasard (6/12)
Implicite graphophonologique	6,49 (2,07) N = 119	t(118) = 2,57 ; p<.05	8,92 (2,32) N = 83	t(82) = 11,46 ; p<.0001
Implicite grapho-morpho-dérivationnel	6,07 (1,73) N = 119	t(118) = 0,42 ; ns	5,44 (1,81) N = 83	t(82) = -2,79 ; p<.01
Implicite grapho-morpho-flexionnel	6,44 (1,61) N = 119	t(118) = 2,95 ; p<.001	6,48 (1,56) N = 83	t(82) = 2,82 ; p<.01
Implicite grapho-syntaxique	5,69 (1,68) N = 119	t(118) = -2,02 ; p<.05	5,73 (1,82) N = 83	t(82) = -1,33 ; ns
Explicite graphophonologique	11,34 ^P (1,77) N = 123	t(122) = 33,44 ; p<.0001	11,63 ^P (0,74) N = 83	t(82) = 68,87 ; p<.0001
Explicite grapho-morpho-dérivationnel	7,91 (1,95) N = 125	t (124) = 10,96 ; p<.0001	7,78 (1,92) N = 81	t (80) = 8,35, p<.0001
Explicite grapho-morpho-flexionnel	7,19 (1,87) N = 125	t (124) = 7,11 ; p<.0001	6,15 (1,97) N = 82	t (81) = 0,67 ; ns
Explicite grapho-syntaxique	6,86 (2,08) N = 125	t (124) = 4,61 ; p<.0001	6,78 (2,12) N = 80	t (79) = 3,27 ; p<.01

^P Scores plafonds.

En gras : les scores pouvant être retenus pour effectuer des analyses inférentielles.

Les résultats aux tests t de comparaison à la valeur 6 (niveau du hasard), menés sur les deux populations, permettent de constater les points suivants :

- A l'épreuve implicite grapho-morpho-dérivationnelle, les réponses des arabophones ne se distinguent pas significativement du hasard, pour l'ensemble des groupes comme pour les groupes pris séparément et les réponses des lusophones sont significativement inférieures au hasard, en particulier pour les élèves de niveau C ($p < .01$), les réponses des deux autres groupes étant aléatoires (*cf.* tableau 53 et annexe 39¹⁷⁰). Ceci suggère d'une part, que l'ensemble des groupes arabophones n'est pas sensible aux caractéristiques morphologiques des mots écrits français, tout au moins telle que cette sensibilité est évaluée ici et d'autre part que les lusophones les plus âgés, lors de la résolution de cette tâche, sont sensibles à autre chose qu'aux aspects morpho-dérivationnels des mots.
- A l'épreuve implicite grapho-syntaxique on retrouve le même cas de figure mais dans le sens inverse (*cf.* tableau 53 et annexe 39). Ce sont cette fois les arabophones qui, quel que soit le niveau, donnent plus de mauvaises réponses que de bonnes réponses. Ce type de réponses est surtout produit par les élèves de niveau C ($p < .01$), leurs choix spontanés semblant ainsi se diriger vers autre chose que les aspects syntaxiques manipulés. Pour les deux autres groupes, les réponses ne se distinguent pas du hasard. Les lusophones obtiennent, quel que soit le groupe, des résultats qui ne se distinguent pas significativement des réponses aléatoires.

Ces différents résultats incitent donc à supprimer ces deux épreuves des analyses suivantes.

- Les deux autres épreuves implicites (graphophonologique et grapho-morpho-flexionnelle) ne sont en revanche pas concernées par ce problème, les scores, pour les deux populations, se situant significativement au dessus du niveau du hasard (*cf.* tableau 53).

Le paradigme de détection d'intrus, s'attachant ici à diverses caractéristiques linguistiques du français, parvient ainsi à ne faire émerger chez nos deux populations d'apprenants, qu'une sensibilité graphophonologique et grapho-morpho-flexionnelle mais pas une sensibilité aux aspects grapho-morpho-dérivationnels et grapho-syntaxiques de la langue française.

Si ces observations autorisent à mener d'autres analyses sur l'épreuve graphophonologique implicite, la présence de scores plafonds chez les arabophones comme les lusophones à son

¹⁷⁰ Dans l'objectif de simplifier les résultats, nous ne présentons, et ne présenterons, dans les analyses suivantes, que quelques valeurs par rapport au niveau du hasard, de façon à appuyer nos propos. Cependant, l'annexe 39 rapporte la totalité des résultats aux tests t de comparaison à la valeur 6, pour les différentes épreuves et selon la langue première et le niveau.

homologue explicite¹⁷¹ empêchera là encore qu'une analyse comparative entre ces deux types d'épreuves puisse s'effectuer. **La réussite massive à l'épreuve graphophonologique explicite, pour l'ensemble des participants permet d'avancer que, quelle que soit la langue première, la discrimination consciente des éléments graphophonologiques du français est parfaitement maîtrisée rapidement après avoir entamé l'apprentissage du français**

Concernant la dimension grapho-morpho-flexionnelle, l'étude des scores obtenus à l'épreuve explicite suggère qu'une analyse multivariée, prenant notamment en considération le degré de contrôle impliqué dans les tâches, est envisageable. Il est, certes, possible d'observer (*cf.* tableau 53) que les arabophones fournissent un taux de bonnes réponses supérieure à une sur deux à l'épreuve explicite grapho-morpho-flexionnelle mais pas les lusophones, dont les réponses ne se différencient pas du hasard (*cf.* tableau 53). Néanmoins, les élèves de niveau C donnent des réponses supérieures au hasard ($p < .0001$), contrairement aux deux autres niveaux (*cf.* annexe 39). Cette sensibilité même tardive chez les lusophones constitue ainsi un argument pour accepter d'intégrer les résultats à cette épreuve dans les analyses ultérieures.

A l'issue de cette sélection, les tests statistiques réalisés pour vérifier nos hypothèses sont les suivants :

- une analyse à trois facteurs menée sur les résultats aux deux épreuves implicites sauvegardées : l'épreuve graphophonologique et grapho-morpho-flexionnelle, de façon à examiner simultanément, l'effet de la langue première (arabe *vs.* portugais), du niveau (A, B, C) et du domaine linguistique (graphophonologique *vs.* grapho-morpho-flexionnel) ;
- une analyse à trois facteurs, réalisée sur l'ensemble des épreuves grapho-morpho-flexionnelles, de façon à déterminer conjointement l'effet de la langue première et du niveau selon le degré de contrôle impliqué dans les tâches ;
- une analyse à trois facteurs, sur les trois épreuves explicites (grapho-morpho-dérivationnelle, grapho-morpho-flexionnelle et grapho-syntaxique) qui ne présentent pas de scores plafonds de façon à examiner les effets de la langue première, du niveau et du domaine linguistique.

¹⁷¹ En effet, à l'épreuve explicite de détection d'intrus graphophonologique, une très grosse majorité d'apprenants (87 % des lusophones et 88 % des arabophones) répondent correctement à au moins 11 items sur 12. Plus précisément, 76 % des arabophones et 75 % des lusophones obtiennent 100 % de bonnes réponses.

- Enfin, une analyse ciblée sur les connaissances implicites graphophonologiques permettra de déterminer l'émergence de la sensibilité à différentes unités phonologiques en fonction de la langue première et du niveau.

2.2.3.2 *Epreuves implicites de détection d'intrus graphophonologique et grapho-morpho-flexionnel : effet de la L1, du niveau et du domaine linguistique*

Le tableau 54 rapporte les résultats obtenus aux épreuves de détection implicite d'intrus graphophonologique et grapho-morpho-flexionnel en français langue seconde, en fonction de la langue première et du niveau.

Tableau 54 : Performances moyennes sur 12 (et écart-types) obtenues aux épreuves de détection implicite d'intrus graphophonologique et grapho-morpho-flexionnelle FL2 en fonction de la L1 et du niveau.

Epreuve de détection implicite		Niveau A N = 79	Niveau B N = 58	Niveau C N = 65	Total
Arabo. N = 119	G-phonologique	6,68 (1,95)	6,08 ^H (2,25)	6,67 (2,01)	6,46 (1,39)
	G-morpho-flexionnel	6,80 (1,49)	6,03 ^H (1,69)	6,45 ^H (1,61)	
Luso. N = 83	G-phonologique	9,28 (2,31)	7,57 (2,11)	9,52 (2,11)	7,70 (1,27)
	G-morpho-flexionnel	6,69 (1,70)	6,43 ^H (1,47)	6,17 ^H (1,37)	
Tous groupes, toutes épreuves		7,35 (1,35)	6,40 (1,50)	7,02 (1,44)	6,97 (1,47)

^H non significativement différent du hasard

Les valeurs des tests statistiques obtenus à l'analyse à trois facteurs (langue première {arabophones vs. lusophones}, niveau {A, B, C} et domaine linguistique {graphophonologique vs. grapho-morpho-flexionnel} incluant comme covariables les performances en lecture de mots et en compréhension de phrases figurent dans le tableau 55.

Tableau 55 : Résultats de la MANCOVA sur les épreuves implicites graphophonologique et grapho-morpho-flexionnelle

	ddl	F	p
Intercept	1,187	325,77	.000001
Covariable lecture de mots	1,187	1,28	ns
Covariable compréhension	1,187	0,44	ns
Langue première	1,187	20,00	.00001
Niveau	2,187	9,16	.001
Domaine linguistique	1,187	0,10	ns
Langue première * Niveau	2,187	0,67	ns
Langue première * Domaine linguistique	1,187	22,49	.00001
Niveau * Domaine linguistique	2,187	1,80	ns
Langue première * Niveau * Domaine linguistique	2,187	2,46	ns

L'effet significatif de la langue première indique que globalement, les arabophones obtiennent des réponses plus faibles que les lusophones, à l'ensemble de ces deux épreuves. Cependant, l'interaction langue première * domaine linguistique révèle que la supériorité des lusophones concerne les performances de détection graphophonologique ($p < .0001$) mais pas celles de détection grapho-morpho-flexionnelle. Comme l'indique la figure 13 (voir également le tableau 53 des moyennes pour l'ensemble des deux populations), les différences de performances entre ces deux épreuves sont significatives chez les lusophones (l'épreuve graphophonologique est mieux résolue que l'épreuve grapho-morpho-flexionnelle, $p < .0001$), mais pas chez les arabophones.

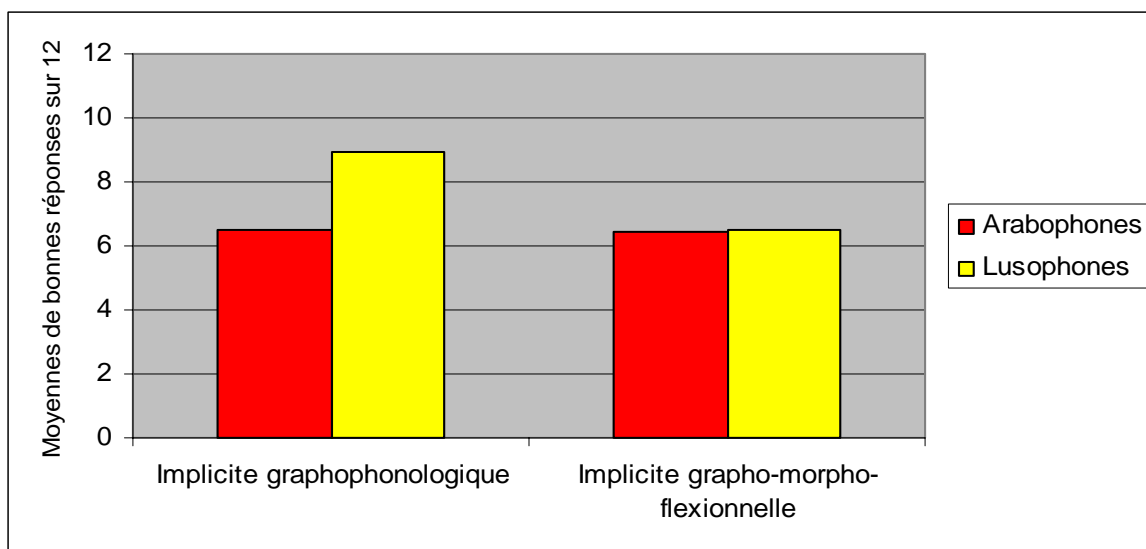


Figure 13 : Performances moyennes (sur 12) obtenues aux épreuves implicites de détection d'intrus graphophonologique, grapho-morpho-flexionnelle en fonction de la langue première

En outre, l'effet significatif du niveau est lié au fait que, sans distinction de langues, les élèves de niveau B ont des performances plus faibles que les élèves de niveau A et C (respectivement $p < .0001$ et $p < .001$), ces dernières n'étant pas significativement différentes.

Bien que l'interaction langue première * niveau * domaine linguistique ne soit pas significative, il faut constater que les élèves de niveau A, arabophones et lusophones, ont des résultats supérieurs au hasard, à toutes les épreuves (*cf.* annexe 39). Ce n'est pas le cas des élèves de niveau B, pour qui, seuls les scores des lusophones à l'épreuve graphophonologique se distinguent significativement du hasard ($p < .01$, *cf.* annexe 39). Ce n'est pas non plus le cas des élèves de niveau C, dont les scores ne sont différents du hasard qu'à l'épreuve graphophonologique, chez les arabophones comme chez les lusophones (respectivement $p < .04$ et $p < .0001$, *cf.* annexe 39 pour les autres résultats).

Ces résultats indiquent que les lusophones sont particulièrement sensibles aux aspects graphophonologiques du français, sensibilité qui n'apparaît d'ailleurs pas pour l'ensemble des groupes arabophones (plus précisément pas au niveau B). Il est également intéressant de constater, malgré l'absence d'interaction niveau * domaine linguistique, que seuls les élèves de niveau A sont sensibles aux aspects grapho-morpho-flexionnels du français, qu'ils soient arabophones ou lusophones.

2.2.3.3 *Epreuves grapho-morpho-flexionnelles implicites & explicites : effet de la langue 1, du niveau et du degré de contrôle.*

Le tableau 56 rend compte des résultats aux épreuves implicite et explicite de détection d'intrus grapho-morpho-flexionnel en fonction de la langue première et du niveau.

Tableau 56 : Performances moyennes sur 12 (et écart-types) obtenues aux épreuves de détection d'intrus grapho-morpho-flexionnel implicite et explicite en fonction de la langue première et du niveau

		Niveau A N = 77	Niveau B N = 58	Niveau C N = 65	Total
Arabophones N = 119	Implicite	6,80 (1,49)	6,03 ^H (1,69)	6,45 ^H (1,62)	6,82 (1,19)
	Explicite	6,63 (1,88)	7,42 (1,61)	7,67 (2,00)	
Lusophones N = 81	Implicite	6,68 (1,73)	6,43 ^H (1,47)	6,17 ^H (1,37)	6,31 (1,23)
	Explicite	5,54 ^H (1,80)	5,95 ^H (2,01)	7,35 (1,75)	
Total		6,42 (1,12)	6,50 (1,27)	6,95 (1,26)	6,61 (1,23)

^H non significativement différent du hasard

Le tableau 57 présente les valeurs statistiques obtenues à l'issue de l'analyse à 3 facteurs (langue première {arabophones vs. lusophones}, niveau {A, B, C} et degré de contrôle {implicite vs. explicite}) sur les épreuves grapho-morpho-flexionnelles, analyse incluant comme covariables les performances en lecture de mots et en compréhension de phrases.

Tableau 57 : Résultats de la MANCOVA sur les épreuves implicite et explicite grapho-morpho-flexionnelles

	ddl	F	p
Intercept	1,185	328,60	<.000001
Covariable lecture de mots	1,185	7,86	<.01
Covariable compréhension	1,185	<1	
Langue première	1,185	14,19	<.001
Niveau	2,185	1,84	ns
Degré de contrôle	1,185	<1	
Langue première * Niveau	2,185	< 1	
Langue première * Degré de contrôle	1,185	10,33	<.01
Niveau * Degré de contrôle	2,185	1,94	ns
Langue première * Niveau * Degré de contrôle	2,185	2,02	ns

Comme l'indiquent les tableaux 56 et 57, les arabophones obtiennent globalement de meilleures performances que les lusophones. Cependant, l'interaction langue première* degré de contrôle permet de constater que cette supériorité des arabophones par rapport aux lusophones concerne les résultats à l'épreuve explicite ($p < .0001$) mais pas à l'épreuve implicite. Comme l'indique la figure 14, les différences de performances entre les épreuves implicite et explicite sont significatives chez les arabophones (l'épreuve explicite est mieux résolue que l'épreuve implicite, $p < .01$), mais pas chez les lusophones.

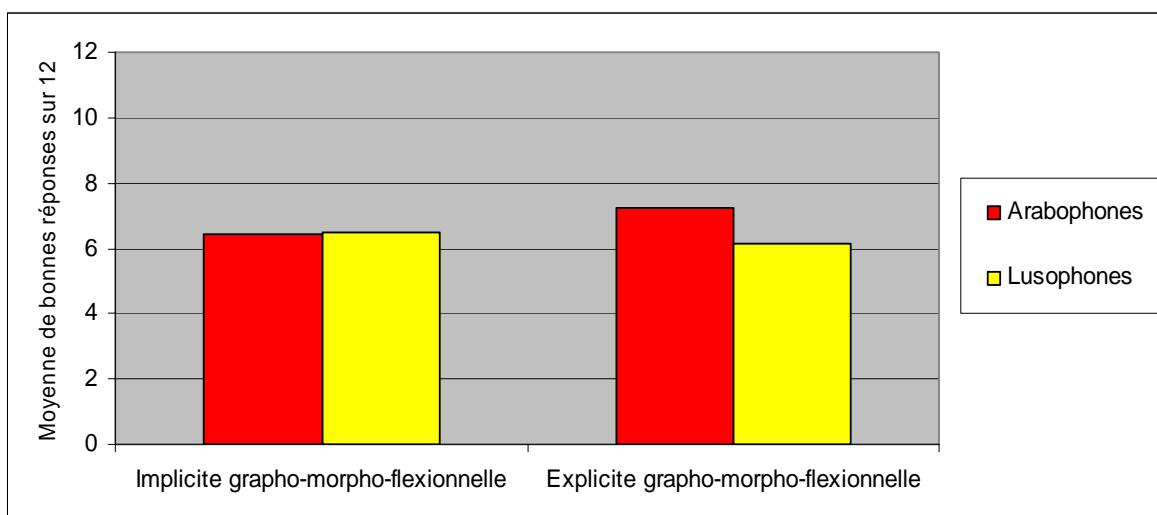


Figure 14 : Performances moyennes (sur 12) obtenues aux épreuves grapho-morpho-flexionnelles de détection d'intrus en fonction de la langue première

Bien que l'interaction langue première * niveau * degré de contrôle ne soit pas significative, et outre le fait déjà observé que seuls les élèves de niveaux A (arabophones et lusophones) obtiennent des performances différentes du hasard à l'épreuve implicite, les lusophones de niveaux A et B donnent, à l'épreuve explicite, des réponses qui ne se différencient pas significativement du hasard contrairement au niveau C et aux arabophones de tous les niveaux (*cf.* annexe 39)¹⁷².

De tels résultats suggèrent que les arabophones parviennent à analyser explicitement la structure morpho-flexionnelle des mots français écrits plus facilement que les lusophones, dont les résultats dénotent, jusqu'à un niveau avancé, un échec important dans la résolution de cette tâche.

Les résultats des groupes dont les scores ne s'écartent pas de façon significative des réponses au hasard seront supprimés des analyses corrélatives, factorielles et de régressions

¹⁷² Les moyennes des arabophones de niveau A et B à l'épreuve explicite et des lusophones de niveau A à l'épreuve implicite ne sont pas exactement identiques à celles de l'annexe 39 calculées à partir des scores de tous les élèves ayant participé à chaque épreuve (les analyses de variance supprimant les valeurs lorsque les sujets n'ont pas participé aux deux épreuves comparées). Toutefois, les tests t sur ces nouvelles valeurs n'ont pas été recalculés car celles-ci sont trop proches des premières pour envisager que ces différences infimes modifient la nature des réponses.

suivantes. Toutefois, pour ce qui concerne l'épreuve explicite grapho-morpho-flexionnelle, même si les lusophones de niveaux A et B obtiennent des résultats de ce type, nous intégrerons les données des trois niveaux dans l'analyse de variance portant sur l'ensemble des tâches explicites grapho-sémantiques, car ce phénomène ne concerne que 2 niveaux sur 3.

2.2.3.4 *Epreuves explicites de détection d'intrus grapho-sémantique : effet de la L1, du niveau et du domaine linguistique*

Les résultats aux trois épreuves explicites grapho-morpho-dérivationnelle, grapho-morpho-flexionnelle et grapho-syntaxique en fonction de la langue 1 et du niveau, sont présentés dans le tableau 58¹⁷³.

Tableau 58 : Performances moyennes sur 12 (et écart-types) obtenues aux épreuves de détection explicite d'intrus grapho-morphologiques et grapho-syntaxique en FL2 en fonction de la L1 et du niveau

		Niveau A N = 78	Niveau B N = 63	Niveau C N = 64	Total
Arabo. N = 125	G-morpho-dérivationnel	7,02 (1,54)	8,07 (1,90)	8,62 (2,06)	7,32 (1,38)
	G-morpho-flexionnel	6,61 (1,88)	7,29 (1,61)	7,67 (2,00)	
	G-syntaxique	6,65 (1,80)	6,38 ^H (2,05)	7,52 (2,23)	
Luso. N = 80	G-morpho-dérivationnel	7,00 (1,94)	8,86 (1,46)	8,09 (1,77)	6,91 (1,36)
	G-morpho-flexionnel	5,59 ^H (1,76)	5,95 ^H (2,01)	7,32 (1,78)	
	G-syntaxique	6,35 ^H (1,95)	6,38 ^H (2,22)	7,86 (1,98)	
Tous groupes		6,55 (1,13)	7,19 (1,13)	7,88 (1,55)	7,16 (1,38)

^H non significativement différent du hasard

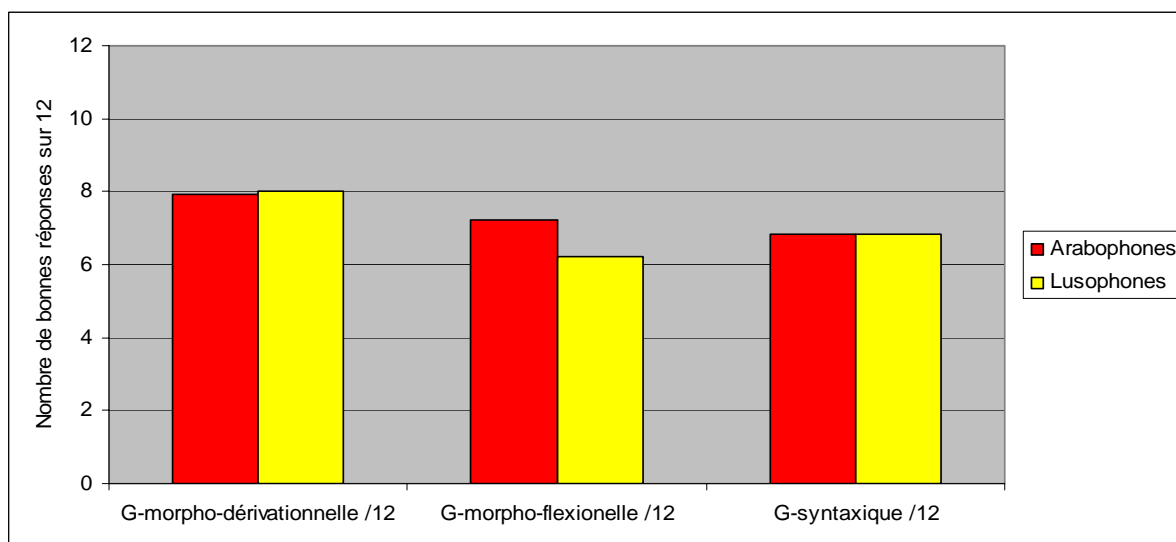
Les valeurs des tests statistiques obtenus à l'analyse à trois facteurs (langue première {arabophones vs. lusophones}, niveau {A, B, C} et domaine grapho-sémantique {grapho-morpho-dérivationnel, grapho-morpho-flexionnel, grapho-syntaxique}) incluant comme covariables les performances en lecture de mots et en compréhension de phrases figurent dans le tableau 59.

¹⁷³ Les petites différences de moyennes à l'épreuve grapho-morpho-flexionnelle pour certains groupes avec les valeurs du tableau 10 s'expliquent par le fait que les performances des sujets n'ayant pas participé à l'ensemble de ces trois épreuves n'ont pas été prises en compte. Ces différences n'étant pas très importantes, la nature des réponses apparues à l'aide des tests t (cf. annexe 38) a été conservée.

Tableau 59 : Résultats de la MANCOVA sur les épreuves explicites grapho-sémantiques

	ddl	F	p
Intercept	1,190	308,43	<.00001
Covariable lecture de mots	1,190	6,45	<.01
Covariable compréhension	1,190	3,09	ns
Langue première	1,190	10,58	<.001
Niveau	2,190	1,02	ns
Domaine grapho-sémantique	2,380	< 1	
Langue première * Niveau	2,190	< 1	
Langue première * Domaine G-sémantique	2,380	5,14	<.01
Niveau * Domaine G-sémantique	4,380	4,10	<.01
Langue première * Niveau * Domaine G-sémantique	4,380	1,86	ns

Si l'effet langue première est significatif, les arabophones obtenant globalement, de meilleurs scores que les lusophones, ce facteur interagit avec le type d'épreuve grapho-sémantique mais aussi avec le niveau. En effet, et comme cela est apparu dans les analyses précédentes, la supériorité des arabophones concerne la résolution de l'épreuve grapho-morpho-flexionnelle ($p < .01$), mais pas celle des autres épreuves (*cf.* figure 15). Ainsi, si l'épreuve grapho-morpho-dérivationnelle est la mieux résolue pour les deux groupes (scores supérieurs à ceux obtenus aux deux autres épreuves, $p < .01$), on observe également que l'épreuve grapho-syntaxique est mieux résolue que l'épreuve grapho-morpho-flexionnelle, mais chez les lusophones seulement ($p < .01$).

**Figure 15 : Performances moyennes (sur 12) obtenues aux épreuves explicites de détection d'intrus grapho-sémantiques en fonction de la langue première.**

L'interaction domaine linguistique * niveau (*cf.* figure 16) s'explique par la supériorité des performances à l'épreuve grapho-morpho-dérivationnelle par rapport aux deux autres épreuves, particulièrement visible au niveau B ($p < .0001$) mais également présente au niveau C ($p < .01$). En revanche, au niveau A, ces scores sont seulement supérieurs à ceux obtenus à l'épreuve morpho-flexionnelle ($p < .05$).

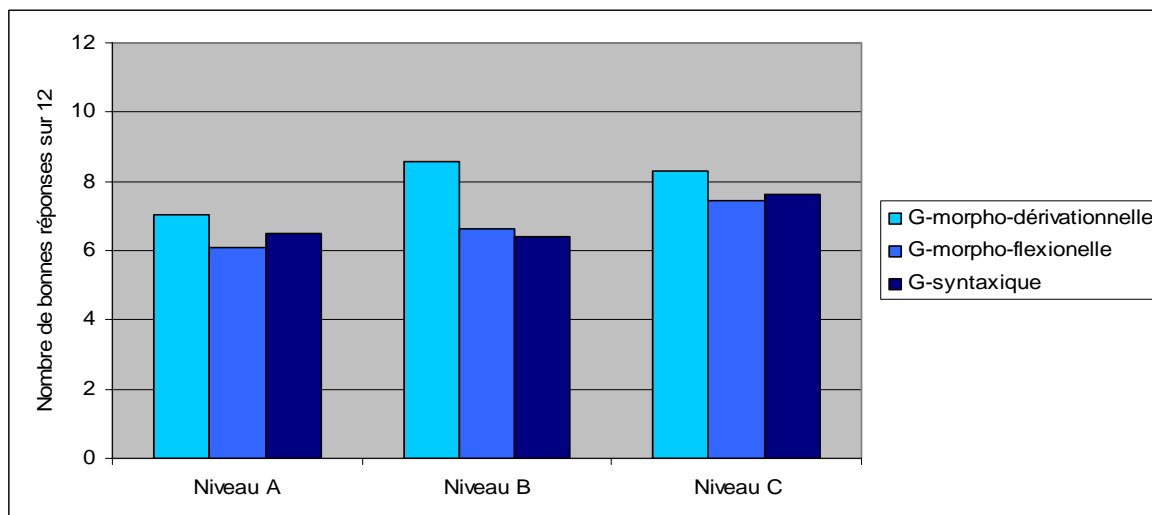


Figure 16 : Performances moyennes (sur 12) obtenues aux 3 épreuves explicites de détection d'intrus grapho-sémantiques en fonction du niveau

Bien que l'interaction Langue première * Niveau * Domaine grapho-sémantique ne soit pas significative, nous remarquons que les arabophones donnent généralement des réponses qui ne sont pas aléatoires, hormis les élèves de niveau B (*cf.* annexe 39). Au contraire les lusophones obtiennent des scores qui ne se différencient du hasard qu'au niveau C ($p < .0001$).

A l'issue de ces résultats, il est possible d'avancer que l'avantage des arabophones, en termes d'analyse de la structure grapho-sémantique du français, concerne essentiellement la morphologie flexionnelle. L'épreuve de détection d'intrus grapho-morpho-flexionnel est d'ailleurs, chez les lusophones, l'épreuve la moins bien résolue (ils produisent au niveau A et B des réponses pouvant être attribuées au hasard). Par conséquent, il est pour eux plus facile de travailler sur la construction syntaxique de phrases écrites en français que sur la structure flexionnelle des mots. Cependant, pour les deux premiers niveaux, les résultats, à ces deux épreuves, sont encore très faibles. Les arabophones ne montrent pas les mêmes profils de performances puisque l'on n'observe pas, chez eux, de différences marquées dans la résolution des épreuves grapho-morpho-flexionnelle et grapho-syntaxique, leurs réponses étant généralement

suffisantes, dès le niveau A pour ne pas être produites de façon aléatoire (hormis dans le cas des réponses des niveaux B à l'épreuve grapho-syntaxique).

En outre, chez les deux populations d'apprenants, la détection explicite d'intrus est plus facile lorsqu'elle porte sur les aspects morpho-dérivationnels des mots. Cette facilité à tenir compte de la construction morpho-dérivationnelle est d'autant plus marquée chez les niveaux les plus avancés.

Ces résultats complètent ainsi les données précédentes qui avaient déjà fait apparaître la difficulté des lusophones à mener une analyse explicite sur certaines formes fléchies en français L2, alors qu'ils semblent n'avoir aucune difficulté, contrairement aux arabophones à traiter de manière spontanée la structure phonologique des mots.

2.2.3.5 *Epreuve implicite de détection d'intrus graphophonologique : effet de la langue première, du niveau et de l'unité phonologique*

Le tableau 60 reprend les résultats obtenus à l'épreuve implicite graphophonologique en fonction de l'unité phonologique, de la langue première et du niveau.

Tableau 60 : Performances moyennes (et écart-types) obtenues à l'épreuve de détection d'intrus graphophonologique implicite en fonction de l'unité, de la langue 1 et du niveau

		Niveau A N = 79	Niveau B N = 58	Niveau C N = 65
Arabophones N = 119	Syllabe/3	1,58 ^H (.82)	1,78 ^H (.96)	1,90 (.88)
	Rime/3	1,68 ^H (.83)	1,22 ^I (.76)	1,57 ^H (.86)
	Consonne/3	1,80 (.88)	1,67 ^H (.86)	1,67 ^H (.75)
	Voyelle/3	1,60 ^H (.74)	1,42 ^H (.87)	1,52 ^H (.92)
Lusophones N = 83	Syllabe/3	2,54 (.72)	1,95 (.74)	2,70 (.56)
	Rime/3	2,00 (.92)	1,71 ^H (1,06)	2,13 (.92)
	Consonne/3	2,33 (.70)	2,05 (.67)	2,26 (.69)
	Voyelle/3	2,41 (.85)	1,86 ^H (.91)	2,43 (.90)
Tous groupes	Total /12	7,96 (2,49)	6,62 (2,30)	7,68 (2,45r)

^H non significativement différent du hasard ^I inférieur au hasard

Les résultats à l'analyse de variance à trois facteurs (langue première, niveau, unité phonologique) incluant deux covariables (lecture de mots et compréhension de phrases écrites) sont mentionnés dans le tableau 61.

Tableau 61 : Résultats de la MANCOVA sur l'épreuve implicite graphophonologique

	ddl	F	p
Intercept	1,186	44,80	<.000001
Covariable lecture de mots	1,186	2,88	ns
Covariable compréhension	1,186	1,88	ns
Langue première	1,186	28,44	<.000001
Niveau	2,186	7,88	<.0005
Unité phonologique	3,558	< 1	
Langue première * Niveau	2,186	2,56	ns
Langue première * Unité phonologique	3,558	1,04	ns
Niveau * Unité phonologique	6,558	< 1	
Langue première * Niveau * Domaine G-sémantique	6,558	1,43	ns

L'absence d'effet de l'unité phonologique et de son interaction avec d'autres facteurs, ne permet pas d'apporter de nouveaux éléments par rapport à ce qui était apparu précédemment, c'est-à-dire une sensibilité phonologique plus développée chez les lusophones que chez les arabophones et une chute des performances pour les élèves de niveau B par rapport aux deux autres niveaux. (A, $p < .0007$ et B, $p < .001$). De même, chez les arabophones la quasi-totalité des résultats sont trop faibles pour se différencier significativement du hasard. Seules les réponses des élèves de niveau A sur la consonne terminale et celles des élèves de niveau C sur la rime sont significativement supérieures au hasard. En revanche, chez les lusophones, tous les scores y sont supérieurs, hormis ceux sur la syllabe et la voyelle terminale des élèves de niveau B.

Les analyses de comparaison au niveau du hasard indiquent ainsi qu'aucun des groupes d'arabophones n'est sensible à la rime et à la voyelle terminale des mots français, alors que les lusophones des niveaux A et C le sont. Par ailleurs, les trois groupes de lusophones montrent une sensibilité à la syllabe et à la consonne terminale alors que ces deux types de sensibilité sont très limitées chez les arabophones, en particulier au niveau C pour la première et au niveau A pour la seconde.

Hormis les résultats très faibles des arabophones de niveau B, pour qui l'ensemble de l'épreuve ne sera pas pris en considération dans la suite des analyses, les résultats obtenus par les autres participants seront conservés et aucun groupe d'items ne sera exclu, bien que certains apparaissent statistiquement proches du niveau du hasard. Le maintien de cette épreuve chez tous les lusophones et chez les arabophones de niveau A et C se justifie car les

scores globaux sont normalement distribués¹⁷⁴ et significativement supérieurs au hasard (cf. annexe 39).

2.2.4 Epreuve de plausibilité lexicale : comparaison entre arabophones et lusophones selon le niveau

Dans cette épreuve chaque type d'items est considéré de façon isolée.

Les résultats présentés dans le tableau 62 révèlent que les arabophones comme les lusophones, sans distinction de niveau, ne donnent pas plus de réponses pour les pseudo-mots composés d'un affixe que pour ceux non construits (items A/R). Cependant, les lusophones de niveau C sont les seuls à donner significativement plus de réponses pour les pseudo-mots affixés ($p < .05$).

Tableau 62 : Réponses moyennes sur 4 (écart-types) et en % aux items A/R de l'épreuve de plausibilité lexicale en FL2 selon la L1 et tests t de comparaison au hasard (2/4 ou 50 %)

Groupe	Niveau A	Niveau B	Niveau C	Tous niveaux
Arabophone	2,12 (1,10) 53,05 % t (40) = 0,71, ns	2,00 (.94) 50 % t (36) = 0,00, ns	1,98 (1,04) 49,39 % t (40) = -0,15, ns	2,03 (1,02) 5,84 % t(118) = 0,36 ; ns
Lusophone	2,08 (1,22) 51,97 % t (37) = 0,40, ns	2,14 (1,06) 53,57 % t (20) = 0,62, ns	2,39 (.78) 59,78 % t (22) = 2,40, $p < .05$	2,18 (1,07) 54,57 % t(81) = 1,55 ; ns

Par ailleurs, les arabophones ne font pas de choix plus marqués pour les pseudo-mots composés d'une base et d'un affixe par rapport aux pseudo-mots seulement composés d'une base, et cela quel que soit le niveau concerné (cf. tableau 63 des résultats aux items 2M/B). En revanche, pour ces mêmes items, les lusophones donnent significativement plus de réponses pour les pseudo-mots constitués de deux morphèmes. Cette préférence ne concerne que les niveaux B et C, la proportion des réponses pour l'un ou l'autre des pseudo-mots étant, au niveau A comme pour l'ensemble des arabophones, non significativement différente du hasard.

¹⁷⁴ Tests de Kormogorov-Smirnov non significatifs, chez les lusophones : $d = 0,16$; $p > .20$ au niveau A ; $d = 0,18$; $p > .20$ au niveau B et $d = 0,24$; $p > .20$ au niveau C, chez les arabophones : $d = .13$ $p > .20$ au niv A et $d = .20$; $p > .05$ au niveau C.

Tableau 63 : Réponses moyennes sur 4 (écart-types) et en % aux items 2M/B de l'épreuve de plausibilité lexicale en FL2 selon la L1 et tests t de comparaison au hasard (2/4 ou 50 %)

Groupe	Niveau A	Niveau B	Niveau C	Tous niveaux
Arabophone	1,88 (1,08) 46,95 % t (40) = -0,73, ns	1,97 (.93) 49,32 % t (36) = -0,18, ns	2,12 (.87) 53,05 % t (40) = 0,90, ns	1,99 (0,96) 49,79 % t(118) = -0,10 ; ns
Lusophone	2,29 (1,04) 57,24 % t (37) = 1,72, ns	2,62 (1,20) 65,48 % t (20) = 2,36 p<.03	2,57 (1,20) 64,13 % t (22) = 2,26, p<.04	2,45 (1,12) 61,28 % t(81) = 3,64 ; p<.001

Pour chacun des trois autres types d'items (B/R, B/A et 2M/A), les scores étant supérieurs à 50 % pour les deux groupes, des analyses de covariance évaluant les effets de la langue première et du niveau ont pu être menées en contrôlant les effets des deux covariables : niveau de lecture de mots et de compréhension de phrase.

Pour les items B/R (*cf.* tableau 64), aucun des facteurs entrés dans l'analyse n'a d'effet significatif [F(1, 186) < 1 pour la compréhension de phrase et la langue première ; F(1, 186) = 1,64 ; ns pour la lecture de mots et F(2, 186) < 1 pour le niveau]. L'interaction langue première * niveau n'est pas non plus significative [F (2, 186) = 1,51 ; ns]. Une telle absence de différence peut être liée au faible nombre d'items testés.

Tableau 64 : Réponses moyennes sur 4 (écart-types) et en % aux items B/R de l'épreuve de plausibilité lexicale en FL2 selon la L1 et tests t de comparaison au hasard (2/4 ou 50 %)

Groupe	Niveau A	Niveau B	Niveau C	Tous niveaux
Arabo.	2,51 (.98) 62,80 % t (40) = 3,35, p<.01	2,68 (.82) 66,89 % t (36) = 5,02 p<.0001	2,90 (.97) 72,56 % t (40) = 5,96, p<.0001	2,70 (0,93) 67,44 % t(118) = 8,14 ; p<.0001
Luso.	2,84 (.95) 71,05 % t (37) = 5,49, p<.0001	2,71 (.78) 67,86 % t (20) = 4,18 p<.001	2,57 (1,08) 64,13 % t (22) = 2,51, p<.05	2,73 (0,94) 68,29 % t(81) = 7,02 ; p<.0001

Pour les items B/A (*cf.* tableau 65), on note un effet de la covariable compréhension en lecture FL2 [F(1, 186) = 9,86 ; p < .01] et un effet du niveau [F(2, 186) = 3,89 ; p < .05]. Les différences sont cependant non significatives dans les comparaisons par paire entre les trois niveaux. De plus, il faut là encore ajouter que seules les réponses données par les lusophones de niveau C ne s'orientent pas plus vers un type de pseudo-mots que vers un autre.

Tableau 65 : Réponses moyennes sur 4 (écart-types) et en % aux items B/A de l'épreuve de plausibilité lexicale en FL2 selon la L1 et tests t de comparaison au hasard (2/4 ou 50 %)

Groupe	Niveau A	Niveau B	Niveau C	Tous niveaux
Arabo.	2,56 (.95) 64,02 % t (40) = 3,78, p<.001	2,49 (.93) 62,16 % t (36) = 3,18 p<.01	2,63 (.86) 65,85 % t (40) = 4,73, p<.0001	2,56 (0,91) 64,08 % t(118) = 6,76; p<.0001
Luso.	2,47 (1,03) 61,84 % t (37) = 2,83, p<.01	2,57 (.93) 64,29 % t (20) = 2,83 p<.05	2,17 (.72) 54,35 % t (22) = 1,16, ns	2,41 (0,93) 6.36 % t(81) = 4,04; p<.0001

Pour les items 2M/A (*cf.* tableau 66), l'effet de la covariable lecture de mots en FL2 est significatif [$F(1, 186) = 4,61, p < .05$] mais pas celui du niveau de compréhension [$F(1, 186) < 1$], ni les effets principaux de la langue première [$F(1, 186) = 2,61, ns$] et du niveau. [$F(2, 186) = 2,56, ns$]. Cependant, l'interaction langue première*niveau est significative [$F(2, 186) = 4,11 ; p < .05$] : les arabophones de niveau C donnent plus de réponses en faveur des pseudo-mots composés de deux morphèmes que les arabophones des niveaux A ($p < .05$) et B ($p < .05$) et que les lusophones des niveaux B ($p < .05$) et C ($p < .05$), groupes pour lesquels les réponses ne sont d'ailleurs pas significativement différentes du hasard. En revanche, leurs performances ne se différencient pas de celles des lusophones de niveau A dont les réponses sont également supérieures à 50 %.

Tableau 66 : Réponses moyennes sur 4 (écart-types) et en % aux items 2M/A de l'épreuve de plausibilité lexicale en FL2 selon la L1 et tests t de comparaison au hasard (2/4 ou 50 %)

Groupe	Niveau A	Niveau B	Niveau C	Tous niveaux
Arabo.	2,12 (.87) 53,05 % t (40) = 0,90. ns	2,05 (.94) 51,35 % t (36) = 0,35, ns	2,78 (1,04) 69,51 % t (40) = 4,82, p<.0001	2,33 (1,00) 58,19 % t(118) = 3,57 ; p<.001
Luso.	2,42 (.83) 6.58 % t (37) = 3,14, p<.001	2,19 (.87) 54,76 % t (20) = 1,00, ns	2,13 (1,01) 53,26 % t (22) = 0,62, ns	2,28 (0,89) 57,01 % t(81) = 2,85 ; p<.01

Les résultats obtenus à cette épreuve de plausibilité lexicale permettent ainsi de constater de nombreuses différences entre arabophones et lusophones en fonction du niveau d'efficienc e atteint. Hormis le fait que l'ensemble des élèves apparaît sensible à la présence d'un mot de la langue française dans un pseudo-mot (items B/R), les réponses aux items B/A suggèrent que le repérage d'une base est plus facile que celui d'un affixe, moins saillant perceptivement. Ce constat n'est pas pourtant pas valable pour les lusophones de niveau C qui, de surcroît, sont les seuls à préférer les pseudo-mots affixés sans base aux pseudo-mots non construits (items A/R). Cette orientation vers l'affixe est décelable dès le niveau B. En effet, pour ces apprenants apparaît, à

partir de ce niveau, une préférence pour les pseudo-mots construits d'une base et d'un affixe par rapport à ceux contenant seulement une base (items 2M/B) alors que le repérage de la construction morphologique par la seule prise en compte de la base disparaît (items 2M/A). Chez les arabophones, la sensibilité à l'association base / affixe n'apparaît qu'au niveau C par une focalisation sur la base (réponses supérieures à 50 % aux items 2M/A et non différentes du hasard pour les items 2M/B) selon un profil assez identique aux lusophones de niveau A.

Globalement, il s'avère que les capacités à discriminer les unités morphologiques du français évoluent différemment entre les deux groupes d'apprenants puisque les lusophones semblent rapidement sensibles à la combinaison base/affixe alors que les arabophones restent encore au niveau C beaucoup focalisés sur la base.

2.2.5 Synthèse des résultats et récapitulatif des épreuves sélectionnées pour les analyses suivantes

A l'issue des analyses précédentes, certaines épreuves de détection d'intrus, implicites ou explicites, ont été éliminées et d'autres conservées. Les résultats pris en compte dans les analyses corrélatives, factorielles et de régression ne seront donc pas exactement identiques pour tous les groupes. Dans le cas de l'épreuve de plausibilité lexicale écrite, le choix de conserver les items donnant lieu à des réponses supérieures au niveau du hasard, selon le groupe, permet de pallier l'absence de données utilisables obtenues à l'aide de l'autre mesure des connaissances grapho-morpho-dérivationnelles implicites (épreuve de détection d'intrus). Cependant, en adoptant ce critère de sélection, les types d'items retenus et par conséquent, les scores globaux qui en découlent diffèrent quasi systématiquement entre les groupes. Il convient ainsi de souligner que le terme de sensibilité à la structure morphologique (plus précisément morpho-dérivationnelle) du français, employée pour décrire les connaissances sous-jacentes à ces évaluations, englobe des unités qui ne sont pas exactement identiques pour tous les groupes.

Le tableau 67 permet d'avoir un aperçu synoptique des résultats obtenus et des épreuves ou groupes d'items inclus dans les analyses qui vont suivre¹⁷⁵.

¹⁷⁵ Ces épreuves ou groupes d'items sélectionnés sont mis en évidence dans le tableau par une police en gras.

Tableau 67 : Récapitulatif des épreuves sélectionnées pour les analyses suivantes pour chaque groupe et synthèse des principaux résultats obtenus (en %)

Epreuves	Arabophones			Lusophones		
	Niveau A	Niveau B	Niveau C	Niveau A	Niveau B	Niveau C
Det. Imp. G-phonologique	55,63 %	Supprimée (hasard)	55,56 %	77,35 %	63,10 %	79,35 %
Det. Imp. G-morpho-dériv.	Supprimée (hasard)			Supprimée (< hasard)		
Det. Imp. G-morpho-flex	56,67 %	Supprimée (hasard)	Supprimée (hasard)	55,77 %	Supprimée (hasard)	Supprimée (hasard)
Det. Imp. G-syntaxique	Supprimée (< hasard)			Supprimée (hasard)		
Det. Exp. G-phonologique	Supprimée (plafond)			Supprimée (plafond)		
Det. Exp. G-morpho-dériv.	58,54 %	67,26 %	71,83 %	58,33 %	73,81 %	67,03 %
Det. Exp. G-morpho-flex.	55,08 %	60,71 %	63,89 %	Supprimée (hasard)		61,23 %
Det. Exp. G-syntaxique	55,49 %	Supprimée (hasard)	62,70 %	Supprimée (hasard)		65,53 %
Plausibilité	5,07 / 8 (B/R+B/A)	5,16 / 8 (B/R+B/A)	7,39 / 12 (B/R+B/A +2M/A)	7,74 / 12 (B/R+B/A +2M/A)	7,90 / 12 (B/R+B/A +2M/B)	7,52 / 12 (B/R+A/R +2M/B)

2.3 Analyse du lien entre les scores aux différentes épreuves en FL2 sélectionnées

2.3.1 Chez les arabophones

Chez les arabophones, les épreuves ou groupe d'items retenus pour analyser le lien entre les différents types de connaissances mesurées en français langue seconde sont les suivantes :

- Au niveau A, les épreuves implicites de détection d'intrus graphophonologique et grapho-morpho-flexionnel et explicites grapho-morpho-dérivationnel, grapho-morpho-flexionnel, grapho-syntaxique. Le score de sensibilité aux unités morphologiques du français, retenu après analyse des résultats à l'épreuve de plausibilité lexicale, correspond à la somme (sur 8) des réponses en faveur de la base, par rapport à l'absence de construction et par rapport à l'afixe.
- Au niveau B, uniquement les épreuves explicites de détection d'intrus grapho-morpho-dérivationnel et grapho-morpho-flexionnel. Le score de sensibilité aux unités

morphologiques du français retenu (épreuve de plausibilité lexicale) correspond, comme au niveau précédent, à la somme (sur 8) des réponses privilégiant la base par rapport à l'absence de construction ou à l'affixe.

- Au niveau C, l'épreuve de détection implicite graphophonologique et les épreuves explicites grapho-morpho-dérivationnelle, grapho-morpho-flexionnelle et grapho-syntaxique. Le score (sur 12) de sensibilité aux unités morphologiques du français a été calculé à partir des réponses aux mêmes items que précédemment (les deux types démontrant une sensibilité à la base) auxquelles sont ajoutés les scores obtenus aux items démontrant la préférence à la composition base/affixe par rapport à l'affixe seul.

Toutes les analyses englobent de surcroît les performances au test de vocabulaire.

2.3.1.1 Analyses de corrélation

Le tableau 68 précise les coefficients de corrélations entre les scores à ces différentes épreuves sélectionnées, chez les arabophones des trois niveaux.

Tableau 68 : Coefficients de corrélation entre les épreuves en FL2 sélectionnées, aux niveaux A, B et C chez les arabophones

Niveau A (N = 40)	Dét. Imp. G-phono. FL2.	Dét. Imp. G- morpho- flex. FL2	Dét. Exp. G-morpho- dériv. FL2	Dét. Exp. G-morpho- flex. FL2	Dét. Exp. G-synt. FL2	Plausibilité FL2
Vocabulaire FL2	.29	.07	.21	.17	.00	-.17
Dét. Imp. G-phono.	----	.22	.00	.07	.20	-.02
Dét. Imp. G- morpho-flex.		----	.02	-.24	-.29	.01
Dét. Exp. G-morpho-dériv.			----	-.14	.00	-.21
Dét. Exp. G-morpho-flex.				----	.14	.40**
Dét. Exp. G- syntaxique.					----	.15

Niveau B (N = 37)	Dét. Exp. G-morpho- dériv FL2.	Dét. Exp. G-morpho- flex. FL2	Plausibilité FL2
Vocabulaire FL2	.23	.19	-.05
Dét. Exp. G-morpho-dériv.	----	.23	.12
Dét. Exp. G-morpho-flex.		----	-.11

Niveau C (N = 41)	Dét. Imp. G-phono. FL2	Dét. Exp. G-morpho- dériv. FL2	Dét. Exp. G-morpho- flex. FL2	Dét. Exp. G- synt.FL2	Plausibilité FL2
Vocabulaire FL2	.18	.06	.19	.16	.60***
Dét. Imp. G-phonologique	----	-.28	-.13	-.03	-.04
Dét. Exp. G-morpho-dériv.		----	.50***	.33*	.13
Dét. Exp. G- morpho-flex.			----	.43**	.56***
Dét. Exp. G- syntaxique.				----	.28

*** r significatif à $p < .001$ ** à $p < .01$ * à $p < .05$

Chez les arabophones de niveau A, la matrice de corrélation permet de constater une seule corrélation significative entre les performances à l'épreuve de détection explicite d'intrus grapho-morpho-flexionnel et à l'épreuve de plausibilité lexicale ($r = .40$; $p < .01$).

Si au niveau B, aucune corrélation entre épreuves n'apparaît, au niveau C en revanche, plusieurs coefficients importants sont à noter : les performances à l'épreuve de plausibilité lexicale (score composite B/R + B/A + 2M/A sur 12) sont fortement corrélées aux scores obtenus au test de vocabulaire ($r = .60$; $p < .0001$) et aux scores de détection explicite d'intrus grapho-morpho-flexionnel ($r = .56$; $p < .0001$). Ces derniers sont aussi corrélés aux scores de détection explicite grapho-morpho-dérivationnelle ($r = .50$; $p < .001$) et grapho-syntaxique ($r = .43$; $p < .01$). La corrélation entre les performances à des deux dernières épreuves est également significative ($r = .33$; $p < .05$).

Ces résultats mettent en évidence certaines relations entre les performances obtenues mais au niveau A et C seulement. Pour ces deux niveaux, on retrouve que la sensibilité aux unités morphologiques du français est liée aux connaissances grapho-morpho-flexionnelles explicites, bien que la composition de ce premier score ne soit pas la même pour les deux groupes. Cependant au niveau C, plus les élèves ont un bon niveau de vocabulaire, plus ils orientent leurs réponses de choix lexical vers des formes composées d'une base existante. Toujours pour ce dernier niveau, une relation « tripartite » entre connaissances grapho-morpho-dérivationnelle, grapho-morpho-flexionnelle et grapho-syntaxique apparaît également.

2.3.1.2 Analyses en composantes principales

Des analyses en composantes principales ont également été menées sur les épreuves ou groupe d'épreuves sélectionnées à chaque niveau. Le tableau 69 présente les résultats obtenus.

Tableau 69 : Contribution des performances aux différentes épreuves sélectionnées aux facteurs extraits des analyses en composantes principales pour chaque niveau chez les arabophones

	Niveau A			Niveau B		Niveau C
	Facteur 1	Facteur 2	Facteur 3	Facteur 1	Facteur 2	Facteur 1
Vocabulaire	-.19	.77*	-.42	-.68	.07	-.53
Dét. Imp. G-phonologique	-.04	.74*	-.31			.17
Dét. Imp. G-morpho-flex.	-.49	.10	-.74*			
Dét. Exp. G-morpho-dériv.	-.43	.28	.46	-.71*	-.40	-.59
Dét. Exp. G-morpho-flex.	.71*	.32	-.09	-.68	.27	-.83*
Dét Exp G-syntaxique	.49	.34	.40			-.64
Plausibilité	.69	-.06	-.42	.06	-.92*	-.77*
% de variance totale expliquée	24,06	20,70	17,19	35,82	27,22	39,27

* poids factoriel significatif ($> .70$)

Au niveau A, cette analyse permet de mettre en évidence trois facteurs : le premier, dont la contribution à la variance s'élève à 24,06 % est composé de l'épreuve de détection explicite grapho-morpho-flexionnelle (poids de .71). Le deuxième, expliquant de façon indépendante 20,70 % de la variance totale, est composée du test de vocabulaire (poids de .77) et de l'épreuve graphophonologique explicite (poids de .74). Le troisième facteur, incluant l'épreuve implicite grapho-morpho-flexionnelle (poids de -.74) ajoute 17,19 % de variance totale expliquée.

Au niveau B, l'ACP permet de dégager deux facteurs indépendants, dont la contribution à la variance s'élève respectivement à 35,82 et 27,22 %. Le premier est composé de l'épreuve de détection explicite d'intrus grapho-morpho-dérivationnel (poids de -.71), et le second est saturé par l'épreuve de plausibilité lexicale écrite (poids de -.92).

Au niveau C, l'unique facteur identifié explique 39,27 % de la variance et est composé des épreuves de détection explicite d'intrus grapho-morpho-flexionnel (poids de -.83), et de plausibilité lexicale (poids de -.77).

Pour résumé, il apparaît que l'épreuve la plus discriminante au niveau A est l'épreuve explicite grapho-morpho-flexionnelle correspondant au facteur « connaissances grapho-morpho-dérivationnelles explicites en français ». Un second facteur, intervenant pour une part non négligeable, est constitué des connaissances lexicales évaluées à l'oral et de la sensibilité graphophonologique, ce qui suggère que cette dimension pourrait correspondre à la reconnaissance des mots. Il est également important de souligner l'intervention d'un autre facteur, celui englobant les connaissances grapho-morpho-flexionnelles implicites, de façon indépendante de ces mêmes connaissances explicites.

Au niveau B, les ACP montrent que les épreuves les plus discriminantes sont ici l'épreuve grapho-morpho-dérivationnelle explicite et l'épreuve de plausibilité lexicale. Ces deux épreuves s'attachent toutes deux à évaluer les connaissances morpho-dérivationnelles en français, mais l'une de façon explicite et l'autre de façon implicite. Leur indépendance rejoint ainsi les résultats du niveau précédent montrant que les connaissances grapho-morpho-flexionnelles implicite et explicite ne font pas partie de la même composante.

Au niveau C, le seul facteur apparu pourrait être assimilé à des compétences morphologiques de différentes natures, car il est constitué de l'épreuve de détection d'intrus grapho-morpho-flexionnel et de l'épreuve de plausibilité axée sur la sensibilité implicite aux unités morpho-dérivationnelles comme la base. Ce

regroupement suggère une évolution par rapport aux niveaux précédents quant à l'indépendance entre les connaissances explicites et implicites, connaissances qui semblent, à ce niveau, commencer à entretenir des relations. Il suggère également la possibilité que s'établissent, avec l'avancée dans les apprentissages, des liens entre les connaissances morphologiques des deux types (dérivationnelle et flexionnelle).

Les connaissances implicites et explicites s'élaborent selon des mécanismes de récupération du savoir très différents. L'indépendance des niveaux implicites et explicites des connaissances, est ainsi apparue lors des ACP menées sur les niveaux A et B, tout au moins pour certaines dimensions linguistiques les plus discriminantes (au niveau A dans le domaine grapho-morpho-flexionnel, au niveau B dans le domaine grapho-morpho-dérivationnel). Toutefois, cette indépendance des connaissances implicites et explicites semblent évoluer dans le sens où, au bout d'un certains temps d'apprentissage (au niveau C étudié dans cette étude) la possibilité qu'elles puissent interagir l'une sur l'autre semble se dessiner.

2.3.2 Chez les lusophones

Chez les lusophones, les épreuves ou groupe d'items retenus pour effectuer l'analyse des liens entre les différents types de connaissances mesurées en français langue seconde sont les suivantes (*cf. supra*, tableau 67) :

- Au niveau A, les épreuves implicites de détection d'intrus graphophonologique et grapho-morpho-flexionnel et explicites grapho-morpho-dérivationnel. Le score de sensibilité aux unités morphologiques du français déterminé à l'issue de l'analyse des résultats à l'épreuve de plausibilité lexicale correspond à la somme (sur 12) des réponses en faveur de la base par rapport à l'absence de construction des réponses en faveur de la base par rapport à l'affixe et de celles en faveur de la combinaison base/affixe par rapport à l'affixe seul.
- Au niveau B, les épreuves implicites de détection d'intrus graphophonologique et grapho-morpho-flexionnel ainsi que l'épreuve explicite de détection d'intrus grapho-morpho-flexionnel. Le score de sensibilité aux unités morphologiques du français calculé (épreuve de plausibilité lexicale) correspond à la somme (sur 12) des réponses privilégiant la base par rapport à l'absence de construction, la base par rapport à l'affixe et la combinaison base/affixe par rapport à la base simple.
- Au niveau C, l'épreuve de détection implicite graphophonologique et les épreuves explicites grapho-morpho-dérivationnelle, grapho-morpho-flexionnelle et grapho-syntaxique. Le score (sur 12) de sensibilité aux unités morphologiques du français a été

calculé à partir des réponses aux items de l'épreuve de plausibilité démontrant la préférence à la base et à l'affixe par rapport à l'absence de construction et à la composition base/affixe par rapport à la base seule.

Toutes les analyses englobent, comme précédemment, les performances au test de vocabulaire.

2.3.2.1 Analyses de corrélation

Le tableau 70 précise les coefficients de corrélations entre les scores à ces différentes épreuves sélectionnées, chez les lusophones des trois niveaux.

Tableau 70 : Coefficients de corrélation entre les épreuves en FL2 sélectionnées, aux niveaux A, B et C chez les lusophones

Niveau A (N = 35)	Dét. Imp. G-phono. FL2.	Dét. Imp. G-morpho-flex. FL2	Dét. Exp. G-morpho-dériv. FL2		Plausibilité FL2
Vocabulaire FL2	-.13	.16	.21		.04
Dét. Imp. G-phono.	----	-.50**	.04		.07
Dét. Imp. G-morpho-flex.		----	-.17		.18
Dét. Exp. G-morpho-dériv.			----		.18

Niveau B (N = 21)	Dét. Imp. G-phono. FL2.	Dét. Exp. G-morpho-dériv. FL2.		Plausibilité FL2
Vocabulaire FL2	-.02	-.20		-.29
Dét. Imp. G-phono.	----	.04		.09
Dét. Exp. G-morpho-dériv.		----		-.29

Niveau C (N = 22)	Dét. Imp. G-phono. FL2	Dét. Exp. G-morpho-dériv. FL2	Dét. Exp. G-morpho-flex. FL2	Dét. Exp. G-synt. FL2	Plausibilité FL2
Vocabulaire FL2	.28	.37	.15	.37	.19
Dét. Imp. G-phonologique	----	-.07	-.30	-.05	.57**
Dét. Exp. G-morpho-dériv.		----	.41	.26	-.11
Dét. Exp. G-morpho-flex.			----	.32	-.15
Dét. Exp. G-syntaxique.				----	.00

*** r significatif à $p < .001$ ** à $p < .01$ * à $p < .05$

Au niveau A, la matrice de corrélations ne permet de constater qu'une seule corrélation négative significative entre les performances à l'épreuve de détection implicite d'intrus graphophonologique et l'épreuve implicite grapho-morpho-flexionnelle ($r = -.50$; $p < .002$). Si au niveau B, aucune corrélation entre épreuves n'apparaît, au niveau C, un seul coefficient apparaît significatif, celui entre les performances à l'épreuve de plausibilité

lexicale (score composite B/R + A/R + 2M/B sur 12) et à l'épreuve de détection implicite d'intrus graphophonologique ($r = .57$; $p < .01$).

Ainsi, il semble que chez les lusophones, bien peu de relations soient visibles entre les scores sélectionnés. Il est toutefois possible de retenir que les connaissances graphophonologiques sont au niveau A inversement liées aux connaissances grapho-morpho-flexionnelles implicites alors qu'elles sont, au niveau C, positivement liées à la sensibilité, également implicite, aux unités morpho-dérivationnelles telles que la base et la combinaison base/affixe.

2.3.2.2 Analyses en composantes principales

Les analyses en composantes principales viennent compléter ces premiers résultats (cf. tableau 71).

Tableau 71 : Contribution des performances aux différentes épreuves sélectionnées aux facteurs extraits des analyses en composantes principales pour chaque niveau chez les lusophones

	Niveau A		Niveau B		Niveau C	
	Facteur 1	Facteur 2	Facteur 1	Facteur 2	Facteur 1	Facteur 2
Vocabulaire	-.37	-.55	.53	.71*	.46	.67
Dét. Imp. G-phonologique	.80*	-.15	-.22	-.19	-.38	.80*
Dét. Imp. G-morpho-flex.	-.87*	.08				
Dét. Exp. G-morpho-dériv.	.12	-.81*	.44	-.82*	.73*	.19
Dét. Exp. G-morpho-flex.					.75*	-.11
Dét Exp G-syntaxique					.64	.29
Plausibilité	-.15	-.06	-.90*	.06	-.33	.75*
% de variance totale expliquée	31,74	25,89	33,22	30,29	32,96	29,62

* poids factoriel significatif ($> .70$)

Pour chaque niveau, deux facteurs se dégagent des ACP.

Au niveau A, le premier facteur, expliquant 31,74 % de la variance totale est constitué de deux épreuves se situant à l'opposé l'une de l'autre. A une extrémité de l'axe factoriel, se situe l'épreuve implicite grapho-morpho-flexionnelle (poids de $-.87$) et à l'autre extrémité, l'épreuve implicite graphophonologique (poids de $.80$). Le second, expliquant de façon indépendante 25,89 % de la variance totale, est exclusivement composé de l'épreuve de détection explicite grapho-morpho-dérivationnelle.

Au niveau B, Le premier facteur dégagé, contribuant à expliquer 33,20 % de la variance est saturé par l'épreuve de plausibilité lexicale (poids de $-.90$). Le second, expliquant à lui seul 30,29 % de variance est pour sa part constitué des épreuves de détection explicite grapho-morpho-dérivationnelle et de vocabulaire dont les scores « se situent » là encore, à l'opposé de l'axe (poids respectif de $-.82$ et $.71$).

Au niveau C, le premier facteur, contribuant à expliquer 32,96 % de la variance est constitué des épreuves explicites de détection d'intrus grapho-morpho-flexionnelle (poids de .75) et grapho-morpho-dérivationnelle (poids de .73). Le second facteur, permettant d'expliquer 29,62 % de variance est composé en premier lieu de l'épreuve graphophonologique (poids de .75) et en deuxième lieu de l'épreuve de plausibilité lexicale.

A l'issue de ces analyses, il semble que les variations de performances des lusophones de niveau A soient en partie liées aux résultats des épreuves implicites graphophonologique et grapho-morpho-flexionnelle, qui entretiennent une relation inverse. Autrement dit, si ces épreuves sont bien discriminantes, leurs résultats s'opposent. Cela signifie que, plus les enfants sont sensibles aux relations ou différences d'unités phonologiques entre les mots en français, moins ils sont sensibles à la présence et à la nature des flexions du lexique français. Ce facteur pourrait ainsi correspondre aux connaissances implicites à différentes dimensions linguistiques. En outre, et indépendamment de ce premier facteur, le pouvoir discriminant de l'épreuve explicite de détection d'intrus grapho-morpho-dérivationnel est attesté. Ce second facteur pourrait ainsi correspondre aux « connaissances explicites grapho-morpho-dérivationnelles en français ».

Au niveau B, l'épreuve discriminant le mieux les élèves est la plausibilité lexicale. Le premier facteur ainsi identifié pourrait être assimilé à la sensibilité à certaines unités morphologique du français. Par ailleurs, dans la formation du deuxième axe contribuant à la variance totale, à ce niveau, on retrouve la présence de deux épreuves à chaque extrémités : le vocabulaire d'un côté et les connaissances grapho-morphologiques de l'autre. Cet éloignement pourrait correspondre aux deux facettes orale et écrite d'une même composante, orientée vers les aspects sémantiques des mots de la langue française.

Enfin, au niveau C, il semble que le premier facteur (composé des épreuves grapho-morpho-flexionnelle et grapho-morpho-dérivationnelle explicites) reflète le domaine des connaissances morphologiques explicites. Le deuxième facteur, constitué des deux seules épreuves implicites sélectionnées (détection implicite d'intrus graphophonologique et plausibilité lexicale), correspondrait pour sa part à la nature non intentionnelle de la réalisation de ces tâches.

2.4 Analyse du lien entre L1 et en FL2 pour chaque type de connaissances et chaque niveau en fonction de la langue première.

L'utilisation dans les trois langues d'épreuves de détection d'intrus de construction identique permet particulièrement bien d'examiner le degré de correspondance entre les connaissances développées en L1 et FL2. Cependant, comme d'autres variables mesurées peuvent être associées et intervenir dans l'explication des performances mesurées, des matrices de corrélation, prenant simultanément en compte les multiples mesures disponibles et utilisables, ont été réalisées pour chaque groupe de chaque langue première. Une attention particulière sera toutefois accordée aux coefficients de corrélation entre les épreuves identiques en L1. Successivement les analyses de régression permettront de dégager la ou les performances expliquant le plus le niveau de connaissances en FL2 pour chaque dimension analysable.

2.4.1 Chez les arabophones

A cette étape de l'analyse, il convient non seulement de rappeler les épreuves de détection d'intrus en français langue seconde sélectionnées lors des analyses précédentes, mais surtout de préciser les mêmes épreuves en arabe langue première qu'il est possible de prendre en compte pour examiner le lien entre les connaissances mesurées dans les deux langues. Les résultats obtenus à ces épreuves ont été largement développés dans l'expérience 1. Toutefois, comme l'échantillon des enfants arabophones est plus faible dans cette expérience, les critères de validité des épreuves n'ont pas abouti à la même sélection. Le tableau 72 indique ainsi lorsque les épreuves (en L1 mais aussi en FL2) ont pu être sauvegardées, les moyennes des élèves sélectionnés dans cette expérimentation en fonction de leur niveau. Les suppressions qui ont été nécessaires sont également mentionnées.

Tableau 72 : Epreuves en arabe et en français entrées dans les analyses sur le lien entre L1 et FL2 en fonction des résultats obtenus

Epreuves	Niveau A		Niveau B		Niveau C	
	Arabe	Français	Arabe	Français	Arabe	Français
Det. Imp. G-phonologique	55,93 %	55,63 %	Supprimée (hasard)		Supprimée (hasard)	55,56 %
Det. Imp. G-morpho-dériv.	Supprimée (hasard)					
Det. Imp. G-morpho-flex	Supprimée (hasard)	56,67 %	Supprimée (hasard)		Supprimée (hasard)	
Det. Imp. G-syntaxique	Supprimée (hasard)		55,13 %	Supprimée (hasard)	Supprimée	
Det. Exp. G-phonologique	Supprimée (plafond)					
Det. Exp. G-morpho-dériv.	61,40 %	58,54 %	63,62 %	67,26 %	75,20 %	71,83 %
Det. Exp. G-morpho-flex.	58,33 %	55,08 %	56,71 %	60,71 %	75,00 %	63,89 %
Det. Exp. G-syntaxique	69,21 %	55,49 %	70,33 %	Supprimée (hasard)	83,13 %	62,70 %

Le tableau 73 présente, pour chaque niveau les coefficients de corrélation entre scores aux différentes épreuves de détections d'intrus sélectionnées en arabe ainsi qu'à l'épreuve d'identification de la construction morphologique de pseudo-mots (*cf.* étude 1) d'une part et les scores aux différentes épreuves de détection en français et à l'épreuve de plausibilité lexicale d'autre part.

Tableau 73 : Corrélations entre les épreuves en L1 et FL2 chez les arabophones de niveaux A, B et C

Niveau A (N = 33)	Dét. Imp. G-phono. FL2.	Dét. Imp. G- morpho- flex. FL2	Dét. Exp. G-morpho- dériv. FL2	Dét. Exp. G-morpho- flex. FL2	Dét. Exp. G-synt. FL2	Plausib. FL2
Dét. Imp. G-phono. L1	.12	-.03	.42*	.08	.18	-.22
Dét. Exp. G-morpho-dériv. L1	-.04	-.07	.08	.12	.02	.24
Dét. Exp. G-morpho-flex. L1	.12	.23	.06	.21	-.39*	.09
Dét. Exp. G-syntaxique. L1	.16	.15	.12	-.07	-.20	-.28
Ident. Constr. Morpho. Pseudo-mots	.22	-.01	.13	-.07	-.15	-.10

Niveau B (N = 34)	Dét. Exp. G-morpho- dériv FL2.	Dét. Exp. G-morpho- flex. FL2	Plausib. FL2
Dét. Imp. G-syntaxique. L1	-.15	.04	-.26
Dét. Exp. G-morpho-dériv. L1	.35*	-.12	.08
Dét. Exp. G-morpho-flex. L1	.20	.15	.01
Dét. Exp. G-syntaxique. L1	.32	.30	.48**
Ident. Constr. Morpho. Pseudo-mots	.35*	.04	.21

Niveau C (N = 41)	Dét. Imp. G-phono. FL2	Dét. Exp. G-morpho- dériv. FL2	Dét. Exp. G-morpho- flex. FL2	Dét. Exp. G-synt. FL2	Plausib. FL2
Dét. Exp. G-morpho-dériv. L1	-.02	.40**	.36*	.29	.09
Dét. Exp. G-morpho-flex. L1	-.00	.23	.35*	.44**	.32*
Dét. Exp. G-syntaxique. L1	.24	.40**	.33*	.16	.24
Ident. Constr. Morpho. Pseudo-mots	.11	.23	.32*	.54***	.39*

*** r significatif à $p < .001$, ** à $p < .01$ * à $p < .05$

Pour les arabophones de niveau A, aucune corrélation significative n'apparaît entre les performances aux épreuves identiques en L1 et FL2 (*cf.* corrélations encadrées dans le tableau 73 entre les scores aux deux épreuves implicites graphophonologiques, aux deux épreuves explicites grapho-morpho-dérivationnelles, grapho-morpho-flexionnelles et grapho-syntaxiques). En revanche on note une corrélation significative entre les scores à l'épreuve de détection explicite grapho-morpho-dérivationnelle en français et ceux à l'épreuve de détection implicite d'intrus graphophonologique en arabe L1 ($r = .42$; $p < .01$). Ces derniers contribueraient à expliquer 10,58 % de la variance des performances à l'épreuve grapho-morpho-dérivationnelle en français [$R = .3253$; $R^2 = .1058$; R^2 ajusté = $.0803$, $F(1,35) = 4,14$; $p < .05$].

La corrélation négative entre les scores de détection explicite grapho-syntaxique en français et de détection explicite grapho-morpho-flexionnelle en arabe ($r = -.39$; $p < .05$) est

également significative. La faiblesse des scores obtenus à cette dernière épreuve en arabe contribue ainsi à expliquer 14,60 % de la variance des résultats à l'épreuve grapho-syntaxique en français [$R = .3821$; $R^2 = .1460$; R^2 ajusté = .1223, $F(1,36) = 6,15$; $p < .05$]. Les scores des autres épreuves en FL2 ne sont corrélés à aucune épreuve en L1 (*cf.* tableau 73).

Pour les arabophones de niveau B, le lien corrélationnel entre les scores aux épreuves de détection explicite grapho-morpho-dérivationnelle dans les deux langues apparaît significatif ($r = .35$; $p < .041$), mais pas celui entre les scores aux épreuves grapho-morpho-flexionnelles ($r = .15$; ns). En français, ces derniers ne sont d'ailleurs corrélés avec aucune autre épreuve en arabe. En revanche, les performances de détection explicite grapho-morpho-dérivationnelle en français ne sont pas seulement corrélées à ces mêmes performances mesurées en arabe, mais également aux performances à identifier la construction morphologique de pseudo-mots. L'analyse de régression pas à pas, prenant en compte des deux variables, permet de constater que les performances expliquant le plus la variance des résultats à cette épreuve en français sont celles obtenues à la même épreuve grapho-morpho-dérivationnelle en arabe (13,38 % de variance expliquée). La part ajoutée par l'identification de la construction morphologique de pseudo-mots n'est pas significative ($\beta = .23$; ns ; *cf.* tableau 74).

Tableau 74 : Contribution des variables entrées dans le modèle de régression expliquant les performances en détection explicite d'intrus grapho-morpho-dérivationnel en FL2 chez les arabophones de niveau B

	R	R ²	R ² modifié	F	p
Pas 1 : Det. Exp. G-morpho-dériv. Arabe L1	.3658	.1338	.1338	5,56	<.05
Pas 2 : Ident. Constr. Morpho. pseudo-mots	.4248	.1805	.0463	1,99	ns

Par ailleurs, la corrélation significative entre les scores à l'épreuve de plausibilité lexicale et à l'épreuve de détection explicite grapho-syntaxique en arabe $r = (.48$; $p < .01$) permet de mener une analyse de régression qui montre que ces derniers scores contribuent à expliquer 23,25 % de la variance à l'épreuve de plausibilité [$R = .4821$; $R^2 = .2325$; R^2 ajusté = .2105, $F(1,35) = 1.60$; $p < .001$].

Pour les arabophones de niveau C, la corrélation entre les scores aux épreuves de détection explicite grapho-morpho-dérivationnelle dans les deux langues est significative ($r = .40$; $p < .01$), de même que celle entre les scores aux épreuves de détection grapho-morpho-flexionnel ($r = .35$; $p < .05$). En revanche, aucune corrélation n'apparaît entre les scores aux épreuves grapho-syntaxiques dans les deux langues ($r = .16$; ns).

En outre, ces trois épreuves de détection d'intrus en FL2 font apparaître d'autres corrélations significatives avec d'autres épreuves en arabe :

- Premièrement, les scores à l'épreuve de détection explicite grapho-morpho-dérivationnelle sont également corrélés à ceux obtenus à l'épreuve explicite grapho-syntaxique en arabe ($r = .40$; $p < .01$). Cependant, l'analyse de régression pas à pas montre que ce sont les scores à la même épreuve grapho-morpho-dérivationnelle en arabe qui contribuent à expliquer 16,32 % de la variance des résultats à cette épreuve en français, la contribution supplémentaire de la variable « détection explicite grapho-syntaxique en arabe » étant non significative ($\beta = .19$; cf. tableau 75).

Tableau 75 : Contribution des variables entrées dans le modèle de régression expliquant les performances en détection explicite d'intrus grapho-morpho-dérivationnel en FL2 chez les arabophones de niveau C

	R	R ²	R ² modifié	F	p
Pas 1 : Det. Exp. G-morpho-dériv. Arabe L1	.4039	.1632	.1632	7,80	<.01
Pas 2 : Det. Exp. G-syntaxique. Arabe L1	.4248	.2240	.0608	3,06	ns

- Deuxièmement, les scores à l'épreuve grapho-morpho-flexionnelle explicite en français sont pour leur part également corrélés à toutes les autres performances mesurées en arabe (corrélations significatives avec les performances de détection d'intrus grapho-morpho-dérivationnelle : $r = .36$; $p < .02$, grapho-syntaxique : $r = .33$; $p < .04$, et d'identification de la construction morphologique de pseudo-mots : $r = .32$; $p < .05$). L'analyse de régression, menée pour déterminer l'épreuve expliquant le mieux les performances à l'épreuve explicite grapho-morpho-flexionnelles en français, révèle que ce sont les variables « détection explicite grapho-morpho-dérivationnelle en arabe » et « identification de la construction morphologique de pseudo-mots » qui contribuent à expliquer significativement ces performances (expliquant respectivement 13,40 % et 8,72 % de la variance, cf. tableau 76). La part de variance expliquée (2,06 %) par la « détection explicite grapho-flexionnelle en arabe » est non significative ($\beta = .17,18$; ns).

Tableau 76 : Contribution des variables entrées dans le modèle de régression expliquant les performances en détection explicite d'intrus grapho-morpho-flexionnel en FL2 chez les arabophones de niveau C

	R	R ²	R ² modifié	F	p
Pas 1 : Det. Exp. G-morpho-dériv. Arabe L1	.3661	.1340	.1340	6,19	<.05
Pas 2 : Ident. Const. Morpho. Pseudo-mots	.4704	.2213	.0872	4,37	<.05
Pas 3 : Det. Exp. G-morpho.-flex. Arabe L1	.4918	.2419	.0206	1,03	ns

- Troisièmement, si les scores à l'épreuve grapho-syntaxique en français ne sont pas corrélés à leur homologue en arabe, ils le sont avec ceux obtenus à l'épreuve d'identification de la construction morphologique de pseudo-mots et avec ceux de l'épreuve de détection explicite grapho-morpho-flexionnelle en arabe. Là encore, l'analyse de régression pas à pas menée sur ces deux variables indépendantes (*cf.* tableau 77) montre que seule la variable « identification de la construction morphologique de pseudo-mots » contribue significativement à expliquer 27,97 % de la variance des résultats de détection d'intrus grapho-syntaxique en français, la contribution de la variable « détection explicite grapho-flexionnelle en arabe » étant non significative ($\beta = 23,45$; ns).

Tableau 77 : Contribution des variables entrées dans le modèle de régression expliquant les performances en détection explicite d'intrus grapho-syntaxique en FL2 chez les arabophones de niveau C

	R	R ²	R ² modifié	F	p
Pas 1 : Ident. Const. Morpho. Pseudo-mots	.5289	.2797	.2797	15,53	<.001
Pas 2 : Det. Exp. G-morpho.-flex. Arabe L1	.5660	.3203	.0407	2,33	ns

- Quatrièmement, les scores de plausibilité lexicale sont corrélés d'une part, aux scores d'identification de la construction morphologique de pseudo-mots ($r = .39$; $p < .01$) et d'autre part, aux scores de détection explicite grapho-morpho-flexionnel en arabe ($r = .32$; $p < .05$). Comme précédemment, l'analyse de régression pas à pas, menée sur ces deux variables montre que seule la variable « identification de la construction morphologique de pseudo-mots » contribue significativement à expliquer 15,46 % de la variance des résultats en plausibilité [$R = .3932$; $R^2 = .1546$; R^2 ajusté = .1329 ; $F(1,39) = 7,13$; $p < .01$], la variable « détection explicite grapho-flexionnelle en arabe » étant exclue du modèle.

Les analyses corrélationnelles et de régression portant sur le lien entre L1 et FL2 chez les arabophones mettent en évidence les points suivants :

- **Au premier niveau d'apprentissage du français, les relations entre les deux langues sur des connaissances identiques n'apparaissent pour aucune des dimensions étudiées (graphophonologique implicite, grapho-morpho-flexionnelle implicite¹⁷⁶ et explicite, grapho-morpho-dérivationnelle implicite et grapho-syntaxique explicite).**

- **A partir du niveau B, les connaissances grapho-morpho-dérivationnelles explicites seraient cette fois-ci dépendantes de ces mêmes connaissances développées par le biais de la L1. Cependant, la possibilité de faire référence aux mêmes connaissances**

¹⁷⁶ Le lien entre L1 et L2 sur les dimensions graphophonologique et grapho-morpho-flexionnelle implicite non observé au niveau A, n'a pas pu être vérifié au niveau B et C car ces épreuves n'ont pas été incluses dans les analyses.

appliquées en L1 n'apparaît pas sur la dimension grapho-morpho-flexionnelle au niveau B.

- Au niveau C, un lien corrélational est présent entre les connaissances grapho-morpho-flexionnelles explicites dans les deux langues mais toutefois pas assez fortement puisque ces connaissances en français sont davantage expliquées par d'autres types de connaissances, telles les connaissances grapho-morpho-dérivationnelles.

- Le lien entre L1 et FL2 n'a pas pu être analysé sur les connaissances grapho-syntaxiques explicites au niveau B. Mais au niveau C (comme au niveau A) ce lien n'est pas apparu, les connaissances grapho-syntaxiques étant, pour ce niveau déjà avancé, surtout expliquées par les compétences à identifier les règles de construction morphologique de pseudo-mots.

A plusieurs reprises, les acquisitions dans certains domaines développés en L1 semblent d'ailleurs intervenir sur d'autres types de connaissances en FL2. Nous avons en effet pu noter au niveau A, une association entre les connaissances graphophonologiques implicites en arabe et les connaissances explicites grapho-morpho-dérivationnelles en français. Ainsi, en début d'apprentissage, les traitements graphophonologiques en arabe auraient un impact sur l'installation des connaissances grapho-morpho-dérivationnelles en FL2. Ces traitements participeraient au repérage de la construction des mots français en assistant la prise en compte des relations graphophonologiques entre les mots composés des mêmes affixes. Il aurait en fait été plus probable de voir intervenir ces mêmes connaissances implicites mais spécifiques au code alphabétique français. Cependant, ces dernières ne sont sans doute pas encore suffisantes (du fait du contact encore récent avec l'écrit en FL2) pour relayer celles acquises par le biais de la L1. Cette relation entre phonologie et morphologie dérivationnelle confirme l'implication des connaissances phonologiques dans l'analyse des aspects morphologiques en français (Carlisle & Nomanbhoy, 1993, Casalis & Louis-Alexandre, 2000 etc.).

De plus, toujours chez les arabophones de niveau A, les connaissances grapho-syntaxiques en français apparaissent inversement associées aux connaissances grapho-morpho-flexionnelles en arabe. En d'autres termes, plus les élèves parviennent à détecter la forme fléchie des mots arabes, moins ils ont de chances de maîtriser la structure de formes verbales composées ou des verbes transitifs en français. Cette observation semble indiquer que les traitements de la structure morpho-flexionnelle

en arabe et de la structure des syntagmes verbaux en français sont de nature différente. Il serait peut être possible qu'en se focalisant sur la présence et le statut des éléments flexionnels, les élèves mettent en place en arabe une analyse segmentale non linéaire (par exemple, de façon à repérer la racine des verbes conjugués à l'accompli ou à l'inaccompli). Au contraire, le repérage de la structure syntaxique des phrases françaises utilisées ici, nécessite un traitement linéaire plus global, de façon à percevoir la relation que peut entretenir les différents éléments de la phrase.

Enfin, bien que l'épreuve de plausibilité lexicale utilisée pour évaluer la sensibilité à certaines unités morphologiques en français n'ait pas d'équivalent en arabe, les résultats qu'elle obtient montrent que la sensibilité aux unités morphologiques du français entretient une relation avec les connaissances grapho-syntaxiques au niveau B puis aux compétences à identifier la construction morphologique de pseudo-mots au niveau C. Ces associations peuvent s'expliquer par le fait que les réponses des élèves s'effectuent sans doute, dans ces trois cas, par le repérage de la frontière des mots ou la segmentation des configurations visuelles en unités pertinentes. En effet, pour répondre à certains items de l'épreuve de plausibilité lexicale, les participants semblent extraire une base saillante correspondant à un mot connu de la langue française. Dans l'épreuve de détection d'intrus grapho-syntaxique en arabe, également, il faut repérer la présence d'une préposition accolée au complément donc détecter les frontières lexicales. Enfin, identifier la construction morphologique de pseudo-mots nécessite de dissocier et synthétiser les différents éléments morphémiques sur un matériel linguistique dégagé des spécificités des langues.

Pour l'essentiel, bien que l'ensemble des relations n'ait pas été testé sur tous les niveaux, il semble que la relation entre les mêmes connaissances en arabe et en français puisse s'établir mais relativement tardivement et sur certains aspects seulement. En effet, l'implication directe des acquisitions en L1 n'a été observée que sur les connaissances morphe-dérivationnelles explicites à partir au niveau B et C. En revanche, elle n'a pas pu être testée sur la sensibilité graphophonologique pour ces deux derniers niveaux et n'est pas apparue sur la dimension grapho-syntaxique des connaissances, bien que cette relation n'ait pu être étudiée que sur deux niveaux sur trois (le niveau A et C).

2.4.2 Chez les lusophones

De la même façon que pour les arabophones, avant de présenter les résultats concernant le lien, en fonction du niveau, entre portugais langue première et français langue seconde, un bref récapitulatif des épreuves choisies dans les deux langues est utile. Le tableau 78 indique ainsi, lorsque les épreuves (en L1 mais aussi en FL2) ont pu être sauvegardées, les moyennes pour les échantillons qui nous concernent ici, en fonction du niveau. Les suppressions qui ont été nécessaires sont également mentionnées.

Tableau 78 : Epreuves en portugais et en français entrées dans les analyses sur le lien entre L1 et FL2 en fonction des résultats obtenus

Epreuves	Niveau A		Niveau B		Niveau C	
	Portugais	Français	Portugais	Français	Portugais	Français
Det. Imp. G-phonologique	72,07 %	77,35 %	Supprimée (hasard)	63,10 %	68,84 %	79,35 %
Det. Imp. G-morpho-dériv.	Supprimée (hasard ou <)					
Det. Imp. G-morpho-flex	59,68 %	55,77 %	Supprimée (hasard)		63,77 %	Supprimée (hasard)
Det. Imp. G-syntaxique	Supprimée (hasard)		Supprimée (hasard)		64,49 %	Supprimée (hasard)
Det. Exp. G-phonologique	Supprimée (plafond)					
Det. Exp. G-morpho- dériv.	67,74 %	58,33 %	78,97 %	73,91 %	81,35 %	67,03 %
Det. Exp. G-morpho-flex.	74,36 %	Supprimée (hasard)	82,14 %	Supprimée (hasard)	Supprimée plafond	61,23 %
Det. Exp. G-syntaxique	69,44 %	Supprimée (hasard)	78,97 %	Supprimée (hasard)	Supprimée plafond	65,53 %

Le tableau 79 présente, pour chaque niveau les coefficients de corrélation entre scores aux différentes épreuves de détections d'intrus sélectionnées en arabe ainsi qu'à l'épreuve d'identification de la construction morphologique de pseudo-mots (*cf.* étude 2) d'une part et les scores aux différentes épreuves de détection en français et à l'épreuve de plausibilité lexicale d'autre part.

Tableau 79 : Corrélations entre les épreuves en L1 et FL2 chez les lusophones de niveaux A, B et C

Niveau A (N = 31)	Dét. Imp. G-phono. FL2.	Dét. Imp. G- morpho- flex. FL2	Dét. Exp. G- morpho- dériv FL2	Plausib. FL2
Dét. Imp. G-phono. L1	.18	-.39*	.15	.08
Dét. Imp. G- morpho-flex. L1	.17	.17	.13	.25
Dét. Exp. G-morpho-dériv. L1	-.21	-.11	.17	-.06
Dét. Exp. G-morpho-flex. L1	-.13	.02	.27	.10
Dét. Exp. G-syntaxique. L1	-.14	.02	.05	.11
Ident. Constr. Morpho. Pseudo- mots	-.01	-.15	.39*	-.10

Niveau B (N = 21)	Dét. Imp. G-phono. FL2.	Dét. Exp. G- morpho- dériv FL2	Plausib. FL2
Dét. Exp. G-morpho-dériv. L1	-.26	-.07	.14
Dét. Exp. G-morpho-flex. L1	-.34	-.16	.16
Dét. Exp. G-syntaxique. L1	.03	-.12	-.42 [†]
Ident. Constr. Morpho. Pseudo- mots	.05	-.18	-.21

Niveau C (N = 20)	Dét. Imp. G-phono. FL2	Dét. Exp. G- morpho- dériv FL2	Dét. Exp. G- morpho- flex. FL2	Dét. Exp. G-synt. FL2	Plausib. FL2
Dét. Imp. G-phono. L1	.59**	.09	.08	.39	.60**
Dét. Imp. G-morpho-flex. L1	-.22	.29	.37	.47*	-.07
Dét. Imp. G-syntaxique. L1	-.08	.31	.33	.47*	.38
Dét. Exp. G-morpho-dériv. L1	-.10	.24	.08	.17	-.16
Ident. Constr. Morpho. Pseudo- mots	.07	.52*	.37	.47*	.09

** r significatif à $p < .01$ * à $p < .05$

Pour les lusophones de niveau A, aucune corrélation significative n'apparaît entre les performances aux épreuves identiques en L1 et FL2 (*cf.* corrélations encadrées dans le tableau 79). En revanche on note une corrélation négative significative entre les scores à l'épreuve de détection explicite grapho-morpho-flexionnelle en français et ceux à l'épreuve de détection implicite d'intrus graphophonologique en portugais L1 ($r = -.39$; $p < .05$). La faiblesse des résultats à cette dernière épreuve contribuerait à expliquer 17,20 % de la variance des performances à l'épreuve grapho-morpho-flexionnelle en français [$R = .4147$; $R^2 = .1720$; R^2 ajusté = .1477, $F(1,34) = 7,06$; $p < .01$].

Par ailleurs, la corrélation entre les scores de détection explicite grapho-morpho-dérivationnelle en français et d'identification de la construction morphologique de pseudo-mots est significative ($r = .39$; $p < .05$). Cette variable contribuerait à expliquer 12,85 % de la variance des résultats à l'épreuve grapho-morpho-dérivationnelle en français [$R = .3585$; $R^2 = .1285$; R^2 ajusté = $.1004$, $F(1,31) = 4,57$; $p < .05$]. Les scores des autres épreuves en FL2 ne sont corrélés à aucune épreuve en L1 (*cf.* tableau 79).

Pour les lusophones de niveau B, aucun coefficient de corrélation n'est significatif si bien qu'aucune analyse de régression n'a été réalisée pour ce groupe.

Pour les lusophones de niveau C, la corrélation entre les scores aux deux épreuves de détection implicites graphophonologiques en français et en portugais est significative ($r = .59$; $p < .01$). Ainsi, la variable « détection implicite graphophonologique en portugais » expliquerait 31,49 % de la variance de cette même variable mesurée en français [$R = .5611$; $R^2 = .3149$; R^2 ajusté = $.2822$, $F(1,21) = 9,65$; $p < .01$]. D'autre part, si la corrélation entre les scores de détection explicite d'intrus grapho-morpho-dérivationnels dans les deux langues n'est pas significative, on trouve toutefois une corrélation entre ces scores grapho-morpho-dérivationnels en français et ceux de l'épreuve d'identification de la construction morphologique de pseudo-mots ($r = .52$; $p < .02$). Cette dernière variable expliquerait ainsi 23,21 % de la variance des performances à l'épreuve de détection explicite grapho-morpho-dérivationnelle en français [$R = .4817$; $R^2 = .2321$; R^2 ajusté = $.1955$, $F(1,21) = 6,35$; $p < .05$].

De plus, aucune corrélation n'apparaît entre les performances à l'épreuve grapho-morpho-flexionnelle explicite en français et les autres mesures en portugais. En revanche, pour l'épreuve grapho-syntaxique explicite, plusieurs corrélations avec les mesures en langue portugaise sont significatives : celle avec les scores à l'épreuve de détection implicite grapho-syntaxique ($r = .47$; $p < .05$) et celle avec les scores à l'épreuve de détection implicite grapho-morpho-flexionnelle ($r = .47$; $p < .05$). On note également une corrélation avec les performances d'identification de la construction morphologique de pseudo-mots ($r = .47$; $p < .05$). Pourtant, l'analyse de régression pas à pas incluant ces trois variables ne révèle qu'une contribution tendancielle de la variable « détection implicite grapho-syntaxique en portugais » expliquant 16,09 % de la variance en détection explicite grapho-syntaxique en français (*cf.* tableau 80), la contribution de la seconde variable incluse « identification de la construction morphologique de pseudo-mots » est non significative.

Tableau 80 : Contribution des variables entrées dans le modèle de régression expliquant les performances en détection explicite d'intrus grapho-syntaxique en FL2 chez les lusophones de niveau C

	R	R ²	R ² modifié	F	p
Pas 1 : Det. Imp. G-syntaxique. Portugais L1	.4011	.1609	.1609	3,83	< .07
Pas 2 : Ident. Const. Morpho. Pseudo-mots	.4920	.2420	.0812	2,03	ns

Enfin, pour ce groupe, les scores à l'épreuve de plausibilité lexicale sont corrélés aux scores de détection implicite d'intrus graphophonologique ($r = .60$, $p < .001$) si bien que cette variable expliquerait 37,14 % de leur variance [$R = .6094$; $R^2 = .37,14$; R^2 ajusté = .3414, $F(1,21) = 12,41$; $p < .01$].

A l'issue de ces analyses, nous retiendrons les points suivants :

- **Le lien entre les connaissances implicites graphophonologiques en portugais et en français n'est pas apparu au niveau A, n'a pas pu être testé au niveau B mais intervient assez fortement au niveau C (plus de 30 % de la variance des résultats en français seraient en effet expliqués par les performances à cette même épreuve en portugais).**

- **Le lien entre les connaissances implicites grapho-morpho-flexionnelles dans les deux langues n'a pu être analysé qu'au niveau A mais n'y est pas apparu. Ces connaissances en français se révèlent toutefois inversement associées à la sensibilité graphophonologique en portugais. Autrement dit, plus les élèves repèrent implicitement les analogies graphophonologiques entre les mots, moins ils seraient sensibles au partage des unités flexionnelles en français, et réciproquement. Cette relation inverse avait déjà été observée avec cette même sensibilité en français. De tels résultats laissent ainsi sous-entendre qu'il pourrait exister une sorte de compétition entre les sensibilités aux aspects graphophonologiques et grapho-morpho-flexionnels.**

- **Le lien entre les connaissances grapho-morpho-dérivationnelles explicites entre le français et le portugais ne s'est vérifié à aucun des niveaux étudiés. Ces connaissances en FL2 apparaissent dépendre à plusieurs reprises (niveaux A et C) des compétences à identifier la construction morphologique des pseudo-mots, portant donc sur la même dimension linguistique mais dégagées des aspects de surface propres aux langues.**

- **Le lien entre les deux langues sur les connaissances grapho-morpho-flexionnelles et grapho-syntaxiques explicites n'a pas pu être testé. Au niveau C, seul niveau où ces connaissances en FL2 ont pu être intégrées dans les analyses, le rôle que peut avoir sur elles d'autres types de connaissances en portugais n'est guère probant, sans doute car ces analyses portent sur un échantillon de taille à peine suffisant ($N = 20$).**

Ajoutons enfin que la sensibilité aux unités morphologiques en français, évaluée à l'aide de l'épreuve de plausibilité lexicale, ne présente qu'une association avec les connaissances graphophonologiques implicites en portugais au niveau C. Cette relation avait déjà été observée avec ces mêmes connaissances en français et avait permis de dégager un facteur composé de ces deux types de connaissances implicites en français. Elle sous-entend également que les analogies graphophonologiques aident au repérage des unités morphémiques du français contenus dans des pseudo-mots. Cela confirme le rôle des connaissances phonologiques lors des traitements morphologiques mais à un niveau plus implicite des connaissances, encore peu pris en considération.

En définitive, bien que l'ensemble des relations n'ait pas été testé sur tous les niveaux, il semble que les correspondances entre les connaissances développées en L1 et en L2 n'émergent que tardivement et seulement sur la dimension graphophonologique implicite.

2.5 Recherche des performances déterminant le niveau de lecture en FL2

2.5.1 Analyses sur le niveau de lecture de mots en FL2

2.5.1.1 Chez les arabophones

2.5.1.1.1 Analyses de corrélations

De façon à déterminer dans un premier temps les relations entre les compétences en lecture et les autres compétences mesurées en L1 et en FL2, les analyses corrélationnelles ont inclu en tant que variable, non seulement les performances aux épreuves de détection implicite et explicite sélectionnées dans les deux langues (L1 et FL2) et aux épreuves d'identification de la construction morphologique de pseudo-mots et de plausibilité lexicale mais également les scores aux tests de lectures de mots en L1 et au test de vocabulaire en français.

Le tableau 81 présente les coefficients de corrélation entre le niveau de lecture de mots en FL2 les scores aux épreuves ou tests sélectionnées en L1 et en L2 et chez les arabophones de niveaux A, B et C.

Tableau 81 : Coefficients de corrélation entre les différents scores et le niveau de lecture de mots en FL2 chez les arabophones de niveaux A, B et C

	Détection d'intrus en arabe L1					Ident. Const. Morpho pseudo-mots	Lecture L1	Détection d'intrus en français L2					Plausi.	Voca.
	Imp G-phono.	Imp G-synt	Exp G-morpho-dériv	Exp G-morpho-flex	Exp G-synt.			Imp G-phono.	Imp G-morpho-flex.	Exp G-morpho-dériv.	Exp G-morpho-flex.	Exp G-syntax.		
Niveau A (N = 33)														
Niveau B (N = 33)														
Niveau C (N = 41)														

*** r significatif à $p < .001$, ** à $p < .01$ * à $p < .05$

Chez les arabophones de niveau A, nous pouvons observer que les scores en lecture de mots en arabe sont les plus corrélés aux performances en lecture de mots en FL2 ($r = .68$, $p < .0001$). Les scores à l'épreuve d'identification de la construction morphologique de pseudo-mots et à l'épreuve de détection explicite grapho-morpho-flexionnelle en arabe y sont également corrélés significativement (respectivement $r = .52$, $p < .002$ et $r = .46$, $p < .01$). On note également une corrélation négative significative entre le niveau de lecture de mots en FL2 et les performances à l'épreuve grapho-syntaxique en français ($r = -.41$, $p < .05$).

Chez les arabophones de niveau B, ce sont toujours les scores en lecture de mots en arabe qui sont les plus corrélés aux performances en lecture de mots en FL2 ($r = .71$, $p < .0001$), mais également les résultats à l'épreuve de vocabulaire en français ($r = .66$, $p < .0001$), les scores aux deux épreuves grapho-morpho-dérivationnelle et grapho-morpho-flexionnelle explicite en arabe (respectivement $r = .62$, $p < .0001$ et $r = .57$, $p < .001$), et enfin les score à l'épreuve d'identification de la construction morphologique de pseudo-mots ($r = .56$, $p < .001$).

Au niveau C, la corrélation entre les performances en lecture de mots en arabe et en FL2 est encore une fois significative ($r = .64$, $p < .0001$). On retrouve par ailleurs des corrélations significatives entre les score en lecture de mots en FL2 et :

- les score à l'épreuve de plausibilité lexicale en français ($r = .55$, $p < .0001$) ;
- le niveau de vocabulaire en français ($r = .54$, $p < .0001$) ;
- les scores à l'épreuve grapho-morpho-flexionnelle explicite en français ($r = .43$, $p < .01$) ;
- les scores à l'épreuve d'identification de la construction morphologique de pseudo-mots ($r = .39$, $p < .01$) ;
- les scores à l'épreuve grapho-morpho-dérivationnelle explicite en arabe ($r = .37$, $p < .05$) ;
- les scores à l'épreuve grapho-syntaxique explicite en arabe ($r = .33$, $p < .05$).

Ainsi, le niveau de lecture de mots en français langue seconde est systématiquement associé au niveau de lecture de mots en arabe. De même, les compétences à identifier la construction morphologique de pseudo-mots apparaissent liées à la lecture en français aux trois niveaux.

Pour les deux premiers niveaux, il peut être relevé une association entre la lecture en français et les connaissances grapho-morpho-flexionnelles en arabe.

La relation importante entre la lecture et les connaissances lexicales en français ainsi que l'association entre lecture en français et connaissance grapho-morpho-dérivationnelle en arabe est particulièrement visible au niveau B mais également présente au niveau C.

Pour ce dernier, un nombre important de connaissances est associé à la lecture, en particulier plusieurs connaissances mesurées en français, dont la sensibilité à certaines unités morphologiques du français évaluée à l'aide de la tâche de plausibilité lexicale.

2.5.1.1.2 Analyses de régressions

Afin de déterminer dans quelle mesure les connaissances considérées en arabe L1 et en français L2 ainsi que les compétences et/ou la sensibilité à tenir compte de la construction morphologique de pseudo-mots prédisent de façon indépendante le niveau de lecture de mots en français, des analyses de régression pas à pas ont été conduites à chaque niveau scolaire.

Au niveau A, si l'on entre dans le modèle les variables présentant des corrélations significatives avec le niveau lecture de mots en FL2 (lecture de mots en arabe, identification de la construction morphologique de pseudo-mots, détection explicite grapho-morpho-flexionnelle en arabe et grapho-syntaxique en français, *cf.* tableau 82), on observe que les deux variables contribuant le plus à expliquer le niveau lecture de mots en FL2 sont celles mesurées par le test de lecture de mots en arabe et l'épreuve d'identification de la construction morphologique de pseudo-mots, expliquant ensemble 52,11 % de la variance des scores en lecture. La variable « lecture de mots en arabe » apporte la contribution la plus importante (40,74 %). La variable « détection explicite grapho-morpho-flexionnelle en arabe », incluse dans un deuxième temps ajoute 11,37 % de variance expliquée. La contribution ajoutée de la variable « identification de la construction morphologique » est trop faible pour être significative ($\beta = .16$, ns), la variable « détection explicite grapho-syntaxique en français » étant, pour sa part, exclue du modèle.

Tableau 82 : Contribution des variables entrées dans le modèle de régression sur le niveau de lecture de mots en FL2 chez les arabophones de niveau A

Niveau A (N = 36)	R	R ²	R ² modifié	F	p
Pas 1 : Lecture de mots en arabe L1	.6383	.4074	.4074	23,37	<.0001
Pas 2: Det. Exp. G-morpho-flexionnelle L1	.7219	.5211	.1137	7,84	<.01
Pas 3 : Ident. Const. Morpho. pseudo-mots	.7362	.5383	.0172	1,19	ns

Au niveau B, si l'on entre dans le modèle les variables corrélées au niveau lecture de mots en FL2 (lecture de mots en arabe, vocabulaire, détection explicite grapho-morpho-dérivationnelle et grapho-morpho-flexionnelle en arabe et identification de la construction morphologique de pseudo-mots, *cf.* tableau 83), on observe que les variables contribuant le plus à expliquer le niveau lecture de mots en FL2 sont le vocabulaire en français, la détection explicite grapho-morpho-flexionnelle en arabe et la lecture de mots qui expliquent ensemble 66,56 % de la variance des scores en lecture FL2. La variable « vocabulaire FL2 » apporte la contribution la plus importante (40,98 %). La « détection explicite grapho-morpho-dérivationnelle en arabe », incluse dans un deuxième temps ajoute 17,94 % de variance expliquée. La « lecture de mots en arabe » apparaît au troisième pas et contribue à expliquer, de façon indépendante 7,64 % de variance. La contribution ajoutée de la variable « identification de la construction morphologique » est trop faible pour être significative ($\beta = .16$, ns), la variable « détection explicite grapho-morpho-flexionnelle explicite en arabe » n'ayant pas été retenue.

Tableau 83 : Contribution des variables entrées dans le modèle de régression sur le niveau de lecture de mots en FL2 chez les arabophones de niveau B

Niveau B (N = 37)	R	R ²	R ² modifié	F	p
Pas 1 : Vocabulaire FL2	.6401	.4098	.4098	24,30	<.0001
Pas 2: Det. Exp. G-morpho-dérivationnelle L1	.7676	.5892	.1794	14,85	<.001
Pas 3 : Lecture de mots L1	.8158	.6656	.0764	7,54	<.01
Pas 4 : Ident. Const. Morpho. pseudo-mots	.8261	.6824	.0168	1,70	ns

- Au niveau C, si l'on entre dans le modèle l'ensemble des variables corrélées au niveau lecture de mots en FL2 (lecture de mots en arabe, plausibilité, vocabulaire FL2, détection grapho-morpho-flexionnelle explicite en français, identification de la construction morphologique de pseudo-mots, détection explicite grapho-morpho-dérivationnelle et grapho-syntaxique en arabe, *cf.* tableau 84) on observe que les variables contribuant le plus à expliquer le niveau lecture de mots en FL2 sont le vocabulaire en français, la détection explicite grapho-morpho-flexionnelle en arabe et la lecture de mots qui expliquent

ensemble 66,56 % de la variance des scores en lecture FL2. La variable « vocabulaire FL2 » apporte la contribution la plus importante (40,98 %). La « détection explicite grapho-morpho-dérivationnelle en arabe », incluse dans un deuxième temps ajoute 17,94 % de variance expliquée. La « lecture de mots en arabe », apparaît au troisième pas et contribue à expliquer, de façon indépendante 7,64 % de variance (*cf.* tableau 84). La contribution ajoutée de la variable « identification de la construction morphologique » est trop faible pour être significative ($\beta = .16$, ns), la variable « détection explicite grapho-morpho-flexionnelle explicite en arabe » n’ayant pas été retenue.

Tableau 84 : Contribution des variables entrées dans le modèle de régression sur le niveau de lecture de mots en FL2 chez les arabophones de niveau C

Niveau C (N = 41)	R	R ²	R ² modifié	F	p
Pas 1 : Lecture de mots L1	.6417	.4118	.4118	27,31	<.00001
Pas 2: Plausibilité lexicale FL2	.7958	.6333	.2215	22,96	<.0001
Pas 3 : Det. Exp. G-morpho-dérivationel L1	.8266	.6833	.0499	5,83	<.05
Pas 4 : Vocabulaire FL2	.8428	.7103	.0270	3,36	ns

Ainsi, ces analyses révèlent que les compétences en lecture de mots ne sont pas indépendantes de ces mêmes compétences préalablement développées en arabe langue première. Pour deux groupes d’arabophones sur trois (niveau A et C), la lecture en L1 apparaît même comme le meilleur prédicteur de la lecture de mots en français L2, en expliquant environ 40 % de la variance. Pour le niveau intermédiaire (niveau B) les habiletés en lecture de mots en arabe interviennent également, mais toutefois après que les connaissances lexicales et les connaissances grapho-morpho-dérivationnelles en arabe y aient apporté leur contribution.

Plusieurs types de connaissances morphologiques contribuent d’ailleurs à expliquer les compétences en lecture de mots en français des autres groupes, avec certaines variations quant à la nature de ces connaissances en fonction du niveau. En effet, au niveau A, les connaissances grapho-morpho-flexionnelles mesurées en arabe participent à expliquer une part non négligeable des compétences en lecture de mots en français (11,37 % de variance expliquée après la lecture de mots en arabe).

Au niveau C, la sensibilité aux unités morphologiques du français, telle que la base et sa combinaison avec un affixe, interviennent de façon claire (variable expliquant indépendamment de la lecture de mots en arabe, 22,15 % de la variance). Interviennent ensuite, avec une contribution plus modérée mais toutefois significative,

les connaissances grapho-morpho-dérivationnelles en arabe, déjà apparues au niveau précédent.

Par ailleurs, le rôle de la sensibilité aux aspects graphophonologiques du code arabe, telle qu'elle est évaluée ici par le paradigme de détection d'intrus, n'apparaît pourtant pas intervenir, tout au moins au niveau A, puisque c'est seulement à ce niveau que cette connaissance a pu être prise en considération.

Il se dégage ainsi, outre un certain degré de symétrie dans les compétences en lecture de mots en arabe L1 et en français L2, une intervention appréciable et durable des connaissances morphologiques développées en langue première et pour le niveau le plus avancé, l'apparition du rôle de la sensibilité à certaines unités morphologiques propres aux français.

2.5.1.2 Chez les lusophones

Une démarche identique a été conduite sur les résultats des lusophones pour déterminer les connaissances en portugais L1 et en français L2 pouvant intervenir sur la lecture de mots en français langue seconde.

2.5.1.2.1 Analyses de corrélations

Le tableau 85 présente les coefficients de corrélation entre le niveau de lecture de mots en français langue seconde et les scores aux différentes épreuves sélectionnées en L1 et en L2, au test de lecture de mots en portugais ainsi qu'au test de vocabulaire, chez les lusophones du niveau A, B et C.

Tableau 85 : Coefficients de corrélation entre les différents scores et le niveau de lecture de mots en FL2 chez les lusophones de niveaux A, B et C

	Détection d'intrus en portugais L1						Ident. Const. Morpho pseudo-mots	Lecture L1	Détection d'intrus en français L2					Plausi. L2	Voca. L2
	Imp G-phono.	Imp G-morpho-flex.	Imp G-synt	Exp G-morpho-dériv	Exp G-morpho-flex	Exp G-synt.			Imp G-phono.	Imp G-morpho-flex.	Exp G-morpho-dériv.	Exp G-morpho-flex.	Exp G-syntax.		
Niveau A (N = 31)	.09	.06		.27	.42*	.12	.19	.79***	-.25	-.06	.28			.13	.15
Niveau B (N = 20)				.59**	-.03	.14	.05	.70***	-.04		.06			-.00	-.16
Niveau C (N = 19)	.69***	.21	.04	.08			.03	.75***	.31		.09	.17	.38	.28	.16

*** r significatif à $p < .001$, ** à $p < .01$ * à $p < .05$

Chez les lusophones de niveau A, il est possible de constater que les scores en lecture de mots en portugais sont les plus corrélés aux performances en lecture de mots en FL2 ($r =$

.79, $p < .0001$). La corrélation entre les scores en lecture de mots en FL2 et les scores à l'épreuve de détection explicite grapho-morpho-flexionnelle en portugais est également significative ($r = .42$, $p < .05$). Les autres corrélations sont non significatives.

Chez les lusophones de niveau B, les scores en lecture de mots en FL2 sont toujours corrélés aux performances en lecture de mots en portugais ($r = .70$, $p < .0001$), mais aussi aux performances à l'épreuve grapho-morpho-dérivationnelle en portugais ($r = .59$, $p < .01$). Les autres corrélations ne sont pas significatives.

Chez les lusophones de niveau C, on retrouve toujours une forte corrélation entre les scores en lecture de mots en portugais et en français langue 2 ($r = .75$, $p < .0001$) et cette fois-ci une corrélation significative entre les scores en lecture de mots en FL2 et les scores à l'épreuve de détection implicite d'intrus graphophonologique en portugais ($r = .69$, $p < .001$). Les autres corrélations ne sont pas significatives.

Quel que soit le niveau, il apparaît chez les lusophones une forte association entre les habiletés en lecture en portugais et en français. Conjointement, il convient de remarquer d'une part, que la lecture de mots en français n'est associée à aucune connaissance mesurée en français (par le biais des épreuves de détection d'intrus sélectionnées, cf. annexe 40) et d'autre part que les connaissances en portugais, liées au niveau de lecture en français, diffèrent systématiquement d'un niveau à l'autre¹⁷⁷. On note ainsi une relation entre la lecture de mots en français et les connaissances grapho-morphologiques en portugais pour les deux premiers niveaux (d'abord avec la morphologie flexionnelle au niveau A, puis avec la morphologie dérivationnelle au niveau B) et une relation entre la lecture de mots français et la sensibilité aux aspects graphophonologiques des mots portugais.

2.5.1.2.2 Analyses de régressions

Au niveau A, si l'on entre dans le modèle les variables présentant des corrélations significatives avec le niveau lecture de mots en FL2 (lecture de mots en arabe et détection explicite grapho-morpho-flexionnelle en portugais, cf. tableau 86) on observe que seule la variable « lecture de mots en portugais » permet d'expliquer de façon significative 59,65 %

¹⁷⁷ Bien que les épreuves évaluant ces trois types de connaissance n'aient pas toujours été sélectionnées simultanément pour les trois groupes, il faut toutefois souligner que chacune est présente dans les analyses dans un moins deux groupes sur trois et qu' hormis pour la dimension graphophonologique, il s'agit de groupes de niveaux successifs.

de la variance en lecture de mots en FL2, la variable « détection explicite grapho-morpho-flexionnelle en portugais » étant pour sa part exclue du modèle.

Tableau 86 : Contribution des variables entrées dans le modèle de régression sur le niveau de lecture de mots en FL2 chez les lusophones de niveau A

Niveau A (N = 36)	R	R ²	R ² modifié	F	p
Pas 1 : Lecture de mots L1	.7724	.5965	.5965	50,27	<.00001

Au niveau B, le modèle incluant les variables « lecture de mots en portugais » et « détection explicite grapho-morpho-dérivationnelle en portugais » (*cf.* tableau 87) permet là encore de constater que seule la variable lecture de mots en portugais permet d'expliquer de façon significative 48,64 % de la variance en lecture de mots en FL2. La contribution ajoutée de la seconde variable « détection explicite grapho-morpho-dérivationnelle en portugais » n'est pas significative ($\beta = .30$, ns).

Tableau 87 : Contribution des variables entrées dans le modèle de régression sur le niveau de lecture de mots en FL2 chez les lusophones de niveau B

Niveau B (N = 20)	R	R ²	R ² modifié	F	p
Pas 1 : Lecture de mots L1	.6974	.4864	.4864	17,05	<.001
Pas 2 : Det. Exp. G-morpho-dérivationnel L1	.7439	.5534	.0669	2,55	ns

Au niveau C, le modèle incluant les variables « lecture de mots en portugais » et « détection explicite grapho-morpho-dérivationnelle en portugais » (*cf.* tableau 88) permet de constater que les deux variables expliquent ensemble 68,93 % de la variance en lecture de mots en FL2. Dans un premier temps, le niveau de lecture de mots en portugais permet d'expliquer de façon significative 57,85 % de la variance en lecture de mots en FL2. La part ajoutée par la seconde variable « détection implicite graphophonologique en portugais » (11,08 %) est également significative.

Tableau 88 : Contribution des variables entrées dans le modèle de régression sur le niveau de lecture de mots en FL2 chez les lusophones de niveau C

Niveau C (N = 22)	R	R ²	R ² modifié	F	p
Pas 1 : Lecture de mots L1	.7606	.5785	.5785	27,45	<.0001
Pas 2 : Det. Imp. G-phonologique L1	.8302	.6893	.1108	6,77	<.05

A l'issue de ces analyses, il est clair que la réussite en lecture de mots en portugais intervient comme un prédicteur important de la réussite en lecture de mots en français langue seconde, quel que soit le niveau étudié. Par ailleurs, et bien que la contribution des connaissances graphophonologiques implicites n'ait pas pu être analysée au niveau

B, le rôle de ces dernières apparaît au niveau C, même après que la contribution des compétences en lecture en portugais ait été considérée (ajoute encore 11,08 % de la variance en lecture de mots français en deuxième pas).

2.5.2 Analyses sur le niveau de compréhension de phrases écrites en FL2

2.5.2.1 Chez les arabophones

2.5.2.1.1 Analyses de corrélations

Le tableau 89 présente, pour chaque niveau d'élèves arabophones, les coefficients de corrélation entre le niveau de compréhension de phrases écrites en français langue seconde et les scores aux épreuves en arabe et en français sélectionnés ainsi qu'aux épreuves de lecture en L1, aux tests de lecture de mots et de vocabulaire en FL2. Cependant, compte tenu du nombre important de variables, les résultats présentés ne concerneront que les épreuves évaluant les connaissances pour lesquelles apparaissent pour un niveau ou un autre des corrélations significatives. La matrice complète est présentée dans l'annexe 40.

Tableau 89 : Coefficients de corrélation entre les différents scores et le niveau de compréhension de phrases chez les arabophones de niveaux A, B et C

	Détection d'intrus en arabe L1				Ident. Const. Morpho pseudo-mots	Lecture mots L1	Comp. phrases L1	Détection d'intrus en français L2				Lecture mots FL2	
	Imp G-synt	Exp G-morpho-dériv	Exp G-morpho-flex	Exp G-synt.				Imp G-morpho-flex.	Exp G-morpho-dériv.	Exp G-morpho-flex.	Plausib.		Voca.
Niveau A N = 32		,11	-,10	-,14	,33 ^t	,33 ^t	,29	,25	,10	,34 ^t	,11	,30	,68***
Niveau B N = 32	-.45**	,21	,05	,16	,11	,06	,56***		,50**	-,31	,06	,48**	,30
Niveau C N = 41		,37*	,34*	,41**	,35*	,32*	,15		,26	,39*	,55***	,59***	,65***

*** r significatif à $p < .001$, ** à $p < .01$ * à $p < .05$ ^t tendanciel à $p < .07$

Les résultats à l'analyse corrélationnelle menée sur les scores des arabophones de niveau A permettent de constater que seules les performances en compréhension de phrases écrites en arabe sont significativement corrélées aux performances en compréhension de phrases écrites en français L2 ($r = .68$, $p < .0001$). Les corrélations de ces même performances en français avec les scores aux épreuves de détection explicite grapho-morpho-flexionnelle en français, d'identification de la construction morphologique de pseudo-mots et de lecture de mots en arabe ne sont que tendancielle ($p < .07$).

Au niveau B, les performances les plus corrélées au niveau de compréhension de phrases en FL2 sont celles obtenues à la même épreuve de compréhension de phrases écrites en arabe ($r = .56, p < .001$). Les scores à l'épreuve de détection explicite grapho-morpho-dérivationnelle en français et au test de vocabulaire y sont également corrélés significativement (respectivement $r = .50, p < .004$ et $r = .48, p < .01$). On note également une corrélation négative significative entre le niveau de compréhension de phrases en FL2 et les performances à l'épreuve implicite grapho-syntaxique en arabe ($r = -.45, p < .01$).

Au niveau C, plusieurs corrélations significatives apparaissent :

- celle avec les performances en lecture de mots en FL2 ($r = .65, p < .0001$) ;
- celle avec le niveau de vocabulaire ($r = .59, p < .0001$) ;
- celle avec les résultats à l'épreuve de plausibilité retenus ($r = .59, p < .0001$) ;
- celle avec les scores à l'épreuve explicite grapho-morpho-flexionnelle en français ($r = .39, p < .01$) ;
- celles avec les scores aux épreuves explicites grapho-syntaxique, grapho-morpho-dérivationnelle et grapho-morpho-flexionnelle en arabe (respectivement $r = .41, p < .008, r = .37, p < .05$ et $r = .34, p < .05$) ;
- celle avec les scores d'identification de la construction morphologique de pseudo-mots ($r = .35, p < .05$) ;
- celle avec le niveau de lecture de mots en arabe ($r = .32, p < .05$).

Les analyses de corrélation précédentes rendent compte du fait que chez les arabophones de niveau A, seul le niveau de lecture de mots en FL2 est associé aux compétences en compréhension de phrases écrites dans cette même langue. En revanche, au niveau B, les liens entre la compréhension en français et d'autres compétences mesurées émergent. Parmi les plus importants, nous pouvons retenir celui entre les habiletés de compréhension dans les deux langues et l'association du niveau de compréhension avec les connaissances grapho-morpho-dérivationnelles explicites ainsi qu'avec les connaissances lexicales en FL2.

Pour les arabophones de niveau C, si de nombreuses relations sont apparentes, les plus élevées se situent, comme au niveau A, entre la compréhension de phrases écrite dans les deux langues et, comme au niveau B, entre la compréhension de phrases écrite et le vocabulaire en français. De plus, la possibilité que les meilleurs compreneurs de ce groupe soient également les meilleurs à repérer les éléments morphologiques du français dans la tâche de plausibilité lexicale semble attestée.

Pour le reste, il est frappant de remarquer qu'à ce niveau, les correspondances entre la compréhension en lecture en français et d'autres connaissances (principalement celles des aspects grapho-sémantiques de la langue arabe) semblent radicalement différentes de celles observées aux niveaux précédents. Cette différence pourrait dénoter d'un net progrès dans l'utilisation éventuelle des connaissances en arabe lors de l'apprentissage d'une autre langue.

2.5.2.1.2 Analyses de régressions

Pour le groupe des arabophones de niveau A, l'analyse de régression simple avec comme unique variable le niveau de lecture de mots en FL2 indique que cette variable explique 38,37 % de la variance en compréhension de phrases écrites dans la même langue (cf. tableau 90).

Tableau 90 : Contribution de la variable lecture de mots en FL2 au niveau en compréhension de phrases en FL2 chez les arabophones de niveau A

Niveau A (N = 39)	R	R ²	R ² ajusté	F	p
Lecture de mots FL2	.6194	.3837	.3675	23,66	<.0001

Pour le niveau B, sont entrées, dans un modèle de régression pas à pas, les trois variables présentant des corrélations significatives avec le niveau de compréhension de phrases en FL2 (compréhension en arabe, détection grapho-morpho-dérivationnelle explicite en français, vocabulaire et détection implicite grapho-syntaxique en arabe, cf. tableau 91).

Tableau 91 : Contribution des variables entrées dans le modèle de régression sur le niveau de compréhension de phrases en FL2 chez les arabophones de niveau B

Niveau B (N = 38)	R	R ²	R ² modifié	F	p
Pas 1 : Compréhension phrases L1	.5676	.3221	.3221	17,11	<.001
Pas 2 : Det. Exp. G-morpho-dériv. FL2	.6591	.4344	.1122	6,94	<.05
Pas 3 : Det. Imp. G-syntaxique L1	.6855	.4699	.0355	2,28	ns

Cette analyse permet de constater que les deux variables contribuant à expliquer les résultats en compréhension de phrases en FL2 sont la compréhension de phrases en arabe et la détection explicite grapho-morpho-dérivationnelle en FL2, qui expliquent ensemble 43,44 % de la variance. La variable « compréhension de phrases en arabe » apporte la contribution la plus importante (32,21%). La variable « détection explicite grapho-morpho-dérivationnelle en FL2 », incluse dans un deuxième temps ajoute 11,22 % de variance expliquée (cf. tableau 91). La contribution ajoutée de la variable « détection implicite grapho-syntaxique en arabe » (3,55 %) n'est pas significative.

Au niveau C, si l'on entre dans le modèle l'ensemble des variables corrélées au niveau lecture de mots en FL2 (lecture de mots en FL2, vocabulaire, plausibilité retenus, détection explicite grapho-morpho-flexionnelle en français, détections explicites grapho-syntaxique, grapho-morpho-dérivationnelle et grapho-morpho-flexionnelle en arabe, identification de la construction morphologique de pseudo-mots et lecture de mots en arabe, cf. tableau 92), on observe que les deux variables contribuant significativement à expliquer la variance en compréhension de phrases en FL2 sont le niveau de lecture de mots en FL2 et le vocabulaire en français, la détection explicite grapho-morpho-flexionnelle en arabe et la lecture de mots (50,01 % de la variance expliquée au total). La variable « lecture de mots en FL2 » apporte la contribution la plus importante (42,04 %). La variable « vocabulaire FL2 », incluse dans un deuxième temps ajoute 7,97 % de variance expliquée. La détection explicite grapho-morpho-flexionnelle et la détection explicite grapho-syntaxique en arabe, qui apparaissent au troisième et quatrième pas apportent une contribution trop faible pour être significative (respectivement 3,80 et 2,33 %).

Tableau 92 : Contribution des variables entrées dans le modèle de régression sur le niveau de compréhension de phrases en FL2 chez les arabophones de niveau C

Niveau C (N = 41)	R	R ²	R ² modifié	F	p
Pas 1 : Lecture de mots FL2	.6484	.4204	.4204	28,28	<.00001
Pas 2: : Vocabulaire FL2	.7072	.5001	.0797	6,06	<.05
Pas 3 : Det. Exp. G-morpho-flexionnelle L1	.7336	.5381	.0380	3,01	ns
Pas 4 : Det. Exp. G-syntaxique L1	.7493	.5614	.0233	1,91	ns

Ces résultats montrent ainsi que les habiletés en lecture de mots en français langue seconde interviennent à plusieurs niveaux pour expliquer les compétences en lecture compréhension (au niveau A et C). En revanche au niveau intermédiaire, c'est surtout sur la base des habiletés de compréhension mises en place en langue première que semblent s'élaborer la compréhension en français. Dans cette élaboration, l'intervention des connaissances grapho-morpho-dérivationnelle n'est toutefois pas négligeable.

Pour le groupe le plus avancé (niveau C), outre la contribution de la lecture de mots, l'étendue des connaissances lexicales explique aussi et de façon indépendante les variations entre élèves en termes de niveaux de compréhension du français écrit.

2.5.2.2 Chez les lusophones

2.5.2.2.1 Analyses de corrélations

Le tableau 93 présente, pour les lusophones des trois niveaux, les coefficients de corrélation entre le niveau de compréhension de phrases écrites en FL2 et les scores aux épreuves en portugais et en français, à condition que ces dernières fassent parfois apparaître des coefficients significatifs, ainsi qu'aux épreuves de lecture en L1, aux tests de lecture de mots et de vocabulaire en FL2 (la matrice complète étant disponible dans l'annexe 40).

Tableau 93 : Coefficients de corrélation entre les différents scores et le niveau de compréhension de phrases chez les lusophones de niveaux A, B et C

	Ident. Const. Morpho pseudo-mots	Lecture mots L1	Comp. phrases L1	Exp G- morpho-dériv en FL2.	Voca.	Lecture mots FL2
Niveau A (N = 31)	.11	.09	.18	.20	.44*	.18
Niveau B (N = 20)	.50*	.31	.38	-.15	-.17	.20
Niveau C (N = 19)	.70***	.07	.40	.49*	.05	-.01

*** r significatif à $p < .001$, * à $p < .05$

Chez les lusophones de niveau A, nous constatons que les performances en compréhension de phrases écrites en FL2 sont exclusivement corrélées au niveau de vocabulaire en français ($r = .44$, $p < .05$).

Chez les lusophones de niveau B cette fois, l'unique corrélation significative avec le niveau de compréhension de phrases en français concerne les scores en identification de la construction morphologique de pseudo-mots ($r = .50$, $p < .05$).

Chez les lusophones de niveau C, les scores en compréhension de phrases en français sont significatifs corrélés aux scores d'identification de la construction morphologique de pseudo-mots ($r = .70$, $p < .001$) et aux scores de détection explicite d'intrus grapho-morpho-dérivationnelle en français ($r = .49$, $p < .05$). Les autres corrélations ne sont pas significatives (*cf.* annexe 40).

Ces analyses révèlent que, chez les lusophones, la compréhension en lecture en français est bien dépendante d'autres compétences relatives à la dimension lexicosémantique des langues mais que certaines différences apparaissent entre les niveaux étudiés.

Les meilleurs compreneurs sont au niveau A, ceux qui ont de bonnes connaissances lexicales en français, alors qu'au niveau B et C ce sont ceux qui parviennent le mieux à identifier la construction morphologique de pseudo-mots. Pour le dernier niveau, le

lien trouvé entre les habiletés de compréhension en lecture et les connaissances grapho-morpho-dérivationnelles en français, renforce la relation que peut entretenir la compréhension avec la conscience morphologique.

2.5.2.2.2 Analyses de régressions

Pour le groupe des lusophones de niveau A, l'analyse de régression simple, avec comme unique variable le niveau de lecture de mots en FL2, indique que cette variable explique, de façon significative 11,76 % de la variance en compréhension de phrases écrites dans la même langue (*cf.* tableau 94).

Tableau 94 : Contribution de la variable vocabulaire FL2 au niveau en compréhension de phrases en FL2 chez les lusophones de niveau A

Niveau A (N = 38)	R	R ²	R ² ajusté	F	p
Vocabulaire FL2	.3429	.1176	.0931	4,80	<.05

Pour le groupe des lusophones de niveau B, une analyse de régression du même type permet de constater que les performances à l'épreuve d'identification de la construction morphologique de pseudo-mots expliquent, de façon significative, 26,99 % de la variance en compréhension de phrases écrites en français langue seconde (*cf.* tableau 95).

Tableau 95 : Contribution de la variable identification de la construction morphologique de pseudo-mots au niveau en compréhension de phrases en FL2 chez les lusophones de niveau B

Niveau B (N = 21)	R	R ²	R ² ajusté	F	p
Ident. Constr. Morpho. pseudo-mots	.5195	.2699	.2315	7,02	<.05

Pour les lusophones de niveau C, l'analyse de régression pas à pas, incluant les variables « identification de la construction morphologique de pseudo-mots » et « détection explicite grapho-morpho-dérivationnelle en français » (*cf.* tableau 96) montre que seule la variable « identification de la construction morphologique de pseudo-mots » explique de façon significative 40,63 % de la variance en compréhension de phrases en FL2.

Tableau 96 : Contribution de la variable entrées dans le modèle de régression sur le niveau de compréhension de phrases en FL2 chez les lusophones de niveau C

Niveau C (N = 23)	R	R ²	R ² modifié	F	p
Ident. Constr. Morpho. pseudo-mots	.6374	.4063	.4063	13,00	<.001

A l'issue de ces analyses, il est possible de constater que si les compétences en compréhension de phrases écrites en français, sont au premier niveau surtout

dépendante de l'étendu du vocabulaire dans cette même langue, elles sont aux niveaux suivants, principalement liées aux compétences à analyser explicitement la structure morphologique, indépendamment des caractéristique du code grapho-sémantique spécifique aux langues (évaluées par l'épreuve d'identification de la construction morphologique de pseudo-mots).

3. Discussion

Le premier objectif de cette étude sur l'apprentissage de la lecture en français langue seconde était de déterminer dans quels domaines de compétences ou de connaissances linguistiques se situent les différences entre arabophones et lusophones. Les deux populations étaient appariées sur leurs connaissances lexicales et approximativement sur le volume horaire de français reçu depuis le début de leur scolarité.

La comparaison des performances a ainsi permis de vérifier que les habiletés de reconnaissance des mots écrits en français sont systématiquement moins développées chez les arabophones que chez les lusophones. Le changement de système alphabétique entre l'arabe et le français écrit est sans aucun doute à l'origine des difficultés de décodage des apprenants arabophones. Ces difficultés ont visiblement une incidence sur la compréhension en lecture puisqu'il apparaît aussi que les habiletés de compréhension de phrases écrites sont plus faibles chez les arabophones que chez les lusophones. Le décalage entre les compétences en lecture des deux populations est visible quels que soient les niveaux comparés, suggérant que les arabophones mettent du temps pour rattraper leur désavantage initial.

Les travaux ayant montré l'impact de la distance de systèmes entre L1 et L2 sur les processus lexicaux en langue seconde se sont surtout centrés sur l'adulte en comparant des langues premières alphabétique et non alphabétique (Bialystok *et al.*, 2005a ; Muljani *et al.*, 1998 ; Wang *et al.*, 2003, Wang *et al.*, 2004 ; Yamada, 2004). Nous le vérifions ici chez des élèves en situation d'apprentissage scolaire et dont les langues respectives sont différentes mais toutefois alphabétiques.

Egalement, nous avons supposé que la référence aux acquis en langue première ne s'élabore pas de façon similaire au niveau implicite et explicite des connaissances car la distance entre L1 et L2 n'y a pas la même influence. Les connaissances implicites, liées à la fréquentation répétée des mêmes configurations ne seraient influencées par les acquisitions préalables qu'avec une langue première comme le portugais qui partage de nombreuses

similarités avec le français (*e.g.* orthographe-phonologiques et orthographe-morphologiques). En revanche, les connaissances explicites seraient moins tributaires des relations formelles entre langue 1 et langue 2 car elles relèvent d'un contrôle cognitif plus élevé, permettant au système de se dégager des caractéristiques de surface des langues. Bien que l'ensemble des analyses prévues n'ait pas pu être réalisé, les résultats supportent globalement cette hypothèse.

Les lusophones apparaissent plus sensibles que les arabophones aux aspects graphophonologiques des mots français¹⁷⁸. Au contraire, ce sont les arabophones qui présentent un avantage dans l'analyse explicite de la structure morphe-flexionnelle, cet avantage ne concernant toutefois pas les autres domaines linguistiques de la langue française.

L'influence positive des acquisitions en portugais jouerait surtout sur la sensibilité graphophonologique car les systèmes graphophonologiques des deux langues sont particulièrement proches. Une continuité dans le repérage non intentionnel des configurations graphophonologiques pourrait alors s'effectuer entre les deux langues. Les similitudes formelles sont sans doute moins flagrantes pour ce qui concerne les règles morphe-flexionnelles, règles qui en français, ne font pas l'objet d'une sensibilité plus développée chez les lusophones, dans notre étude.

D'un autre côté, les caractéristiques de l'arabe standard et de son apprentissage auraient pour conséquence de focaliser l'attention sur la dimension morphologique des mots même dans une autre langue. En effet, grâce à la conscience morphologique, les apprenants sont capables de mener une analyse sur les principes d'application des règles quels qu'en soient les aspects de surface ; d'autant plus, si les premiers apprentissages de la lecture ont souligné son importance. Toutefois, cette hypothèse n'a pas pu être confirmée dans le domaine de la morphologie dérivationnelle, les lusophones parvenant, dans ce domaine, à un niveau de connaissance équivalent, probablement ici encore, grâce à la proximité entre le portugais et le français.

En analysant les réponses données lors d'une tâche de plausibilité lexicale écrite, il s'avère que les capacités de discrimination implicite des unités morphologiques du français évoluent différemment entre les deux groupes d'apprenants. Globalement, les lusophones semblent rapidement sensibles à la combinaison base/affixe et sont capables de tenir compte de la seule présence d'un affixe (en préférant les pseudo-mots composés ou les pseudo-mots

¹⁷⁸ Les comparaisons entre arabophones et lusophones ont également été menées sur chaque unité graphophonologique mais les résultats n'apportant rien de plus que cette analyse, ils ne seront pas discutés ici.

contenant un affixe de la langue française). En revanche, les arabophones restent longtemps attachés à la base (en montrant une préférence pour les pseudo-mots composés d'une base par rapport aux pseudo-mots contenant seulement un affixe par exemple). Chez les lusophones, la facilité à répéter les unités discrètes tels que les affixes tiendrait notamment à l'existence d'affixes proches en portugais (pour ceux utilisés ici : « dé- » et « -ier », proches de « *des-* » et « *-eiro* » en portugais). Le repérage des affixes serait plus difficile à mettre en place par les arabophones du fait de la distance formelle des morphèmes entre L1 et L2 mais aussi car leur langue première accepte souvent un autre type de découpage morphologique, non linéaire, qui favorise l'extraction de la racine. Les caractéristiques de la morphologie de l'arabe permettent donc d'expliquer également que la sensibilité à la présence d'une base, par rapport à celle d'un affixe, se maintienne chez les arabophones mais pas chez les lusophones.

Nous voulions également explorer les connaissances les plus déterminantes dans l'acquisition du code grapho-sémantique en français langue seconde, et voir s'il existe certaines différences dans cette acquisition en fonction de la langue première des apprenants. Du fait de la suppression de nombreuses épreuves, les analyses menées ne permettent pas d'obtenir une vue claire sur la façon dont diverses connaissances participent à la construction cognitive de cette dimension du code. Toutefois, chez les arabophones comme chez les lusophones, on remarque, tout au moins pour les niveaux les plus avancés, qu'une relation entre les deux types de connaissances grapho-morphologiques (flexionnelle et dérivationnelle) a tendance à s'établir, ce qui justifie de les étudier conjointement.

La possibilité de mettre en relation des connaissances mesurées par les mêmes instruments en langue première et seconde a permis d'approfondir la façon dont la référence aux connaissances développées en langue première influence le développement des connaissances en FL2. Il était supposé que les relations entre les deux langues peuvent s'établir au niveau implicite des connaissances à condition que ces langues soient, sur la dimension concernée, suffisamment proches. En revanche, nous nous attendions à observer des liens entre les connaissances explicites en L1 et en L2, en particulier dans les domaines linguistiques les plus nécessaires pour lire en L1.

Tout d'abord, plusieurs résultats soutiennent que les correspondances entre les connaissances développées en L1 et en L2 n'émergent que tardivement. Cette observation est cohérente avec l'hypothèse théorique de Cummins (1991) du seuil minimal d'efficiency linguistique à atteindre pour que les compétences se développent en symétrie dans les deux

langues. Pourtant cette affirmation ne permet pas de comprendre quels sont les facteurs responsables de cette efficacité linguistique.

En outre, le fait que l'ensemble des relations n'ait pas été testé sur tous les niveaux et chez les deux populations nous empêche de comparer les arabophones et les lusophones quant aux relations observées -ou non- entre les connaissances grapho-morpho-flexionnelles et grapho-syntaxiques dans les deux langues¹⁷⁹. Une comparaison rigoureuse pourrait être menée uniquement sur les connaissances grapho-morpho-dérivationnelles explicites car ces dernières sont les seules où la relation interlangue a été analysée pour tous les groupes. Toutefois, le fait que la sensibilité graphophonologique en français soit liée à cette même sensibilité en portugais suggère l'existence d'une continuité entre les connaissances implicites en langues première et seconde, tout au moins dans le cas où les langues sont proches et dans des domaines linguistiques importants pour la lecture dans ces langues. Il s'agira toutefois de confirmer plus rigoureusement qu'une telle relation n'existe pas lorsque les langues 1 et 2 sont graphophonologiquement plus éloignées et/ou lorsque la dimension étudiée est relativement moins essentielle pour la lecture. L'absence de sensibilité graphophonologique en arabe chez les apprenants les plus avancés de cette étude, va d'ailleurs dans ce sens. Si déjà en arabe, ils ne parviennent pas à repérer implicitement la relation graphophonologique entre les mots, tout porte à croire que leur sensibilité graphophonologique en français est liée à autre chose qu'à leurs acquisitions en langue première.

Les conclusions que l'on peut tirer des liens interlangues sur les connaissances grapho-morpho-dérivationnelles sont plus claires. Dès le deuxième niveau étudié, chez les arabophones, les connaissances en arabe sont associées à ces mêmes connaissances en français alors que la correspondance entre L1 et L2 n'apparaît pas chez les lusophones. Malgré la distance formelle et structurale entre l'arabe et le français, l'actualisation des connaissances grapho-morpho-dérivationnelles semble avoir lieu chez les apprenants tunisiens. Pour dépasser les contraintes posées par l'éloignement des règles morphologiques, il faut que ces connaissances aient été bien ancrées lors des apprentissages en arabe. C'est à cette condition que le lien entre les langues pourrait s'effectuer. Que ce phénomène ne soit pas présent chez les lusophones, alors que les règles de dérivation sont beaucoup plus proches, suggère que la morphologie dérivationnelle n'est pas suffisamment déterminante en portugais pour que les acquisitions en L1 favorisent le développement de la conscience morphologique en L2.

¹⁷⁹ Par ailleurs, les épreuves de détection grapho-morpho-dérivationnelle implicite et graphophonologiques explicites ont été supprimées dès le départ

Ces affirmations contredisent, à première vue, les interprétations données précédemment, suite à l'observation inattendue d'un niveau de connaissances grapho-morpho-dérivationnelles explicites relativement équivalent entre arabophones et lusophones. Pour comprendre l'absence de supériorité des arabophones, il était envisagé que les lusophones pouvaient se reposer sur leurs acquisitions en portugais du fait des similitudes entre les deux langues. Cet argument ne tient plus à la lumière de l'absence de relation entre les connaissances grapho-morpho-dérivationnelles en L1 et en L2 chez les lusophones. Cela signifierait alors que d'autres facteurs non considérés dans cette expérimentation interviendraient sur le développement des connaissances en L2, par exemple le contenu des apprentissages effectués en français.

Quoi qu'il en soit, ces interprétations supposent que les relations entre des langues mêmes éloignées peuvent s'effectuer, notamment sur les dimensions privilégiées lors des apprentissages spécifiques en langue première. Ce facteur n'a, à notre connaissance, jamais été pris en considération dans les recherches menées pour confirmer le transfert interlangue. De plus, les recherches ont surtout été développées sur la conscience phonologique et se sont peu intéressées à la morphologie. A partir des résultats de leur recherche sur une autre langue sémitique première, l'hébreu, et sur l'anglais langue seconde, Schiff et Calif (2007) avaient fait apparaître une relation entre ces deux langues à l'aide des deux mêmes tâches de jugement de relation morphologique entre deux mots. Cependant, ces épreuves ne portaient que sur des structures linéaires, les différences interlangues quant au processus de dérivation étaient ainsi minimisées. Les résultats des apprenants arabophones permettent ainsi de confirmer que la relation interlangue concerne aussi, dans le domaine morphologique, des structures qui n'emploient pas forcément les mêmes principes de construction.

L'ultime objectif de cette étude était de dégager les profils d'apprentissage de la lecture des arabophones et lusophones en analysant les connaissances contribuant le plus à la reconnaissance des mots écrits et à la compréhension de phrases. Toujours pour souligner l'importance des acquisitions de l'écrit en langue première, nous avons cherché à déterminer le poids de la morphologie lors de la lecture en français langue seconde. L'examen des connaissances associées à la lecture en fonction de l'avancée dans les apprentissages s'est effectué de façon à mettre en parallèle l'évolution des profils de lecture des arabophones et lusophones. Cette démarche est en effet préférable compte tenu du fait que les deux populations d'apprenants ne se situent pas au même niveau d'habiletés en reconnaissance de mots écrits en français.

Concernant la dimension de reconnaissance de mots écrits, un des phénomènes qui apparaît de façon récurrente est la relative symétrie des compétences en lecture de mots en langue première et en français langue seconde. De nombreuses recherches sur le transfert interlingual en lecture avaient ainsi fait les mêmes constations dans des langues premières proches et éloignées de l'anglais L2 (*e.g.* Abu-Rabia & Siegel, 2002, 2003 ; Kahn-Horwitz *et al.*, 2005).

De plus, les habiletés en lecture de mots en L1 expliquent une part très importante de ces mêmes habiletés en français chez les lusophones (entre 49 et 59 % de variance expliquée, contre 40 % environ chez les arabophones). Ainsi, cette interdépendance ne nous semble pas toujours exactement interprétable de la même façon chez les deux populations d'apprenants. En effet, la contribution d'autres types de connaissances invite à approfondir les variations entre arabophones et lusophones en termes de stratégies de lecture mises en place au cours des apprentissages.

Si l'intervention des connaissances morphologiques apparaît régulièrement chez les arabophones, l'évolution des habiletés de reconnaissance en français langue seconde se traduit chez eux par une implication des connaissances morphologiques de plus en plus orientée vers la morphologie dérivationnelle et de plus en plus axée sur les processus de construction propres aux français. En outre, les connaissances intervenant au dernier niveau étudié mettent en évidence la possibilité que les traitements morphologiques implicites soient suffisants pour supporter la reconnaissance des mots écrits en français. Ces éléments confirment donc le rôle non négligeable des traitements morphologiques implicites en français langue seconde que plusieurs recherches en français langue maternelle avaient démontré très tôt lors des apprentissages (*e.g.* Marec-Breton, 2003 ; Royer, 2003).

Le rôle de la morphologie est loin d'être aussi déterminant chez les lusophones. Chez eux la constance des mécanismes d'apprentissages de la lecture en portugais et en français semble attestée par la grande symétrie entre les habiletés de lecture dans les deux langues et par une évolution presque inexistante des profils. La seule nuance, dans ces profils successifs, concerne l'intervention, au dernier niveau, de la sensibilité aux aspects graphophonologiques du portugais suggérant que le décodage est parfaitement automatisé et que la proximité entre les langues joue un rôle important dans la mise en place de ces automatismes.

Concernant enfin la dimension de compréhension en lecture, on retrouve, de façon flagrante, des différences entre arabophones et lusophones. En début d'apprentissage, les arabophones ne font que déchiffrer, ce qui monopolise toutes leurs ressources cognitives.

Aucune connaissance ne peut ainsi venir supporter la compréhension, ces habiletés étant exclusivement dépendantes des habiletés de décodage présentes dans les procédures de reconnaissances des mots. Le deuxième niveau montre des différences très importantes quant aux compétences associées à la compréhension. Ce changement correspondrait à une recherche de nouvelles stratégies pour pallier leur difficulté de déchiffrage et accélérer leur compréhension. Ces nouvelles stratégies semblent s'appuyer sur les stratégies de compréhension de phrases élaborées en arabe, stratégies qui se basent notamment sur l'analyse explicite de la construction morpho-dérivationnelle des mots. On retrouve en effet comme prédicteur de la compréhension en français, surtout la compréhension en arabe, mais aussi les connaissances grapho-morpho-dérivationnelles explicites en français.

Les arabophones les plus avancés n'ont pas encore dépassé les traitements séquentiels initiaux permettant d'aborder une autre étape de la compréhension en lecture moins dépendante du décodage. Cependant, l'intervention secondaire, des connaissances lexicales en français permet de supposer que les apprenants commencent à progresser vers une lecture plus fluide. Leurs stratégies de compréhension ne sont toutefois pas aussi élaborées qu'en arabe. Elles pourraient encore rester limitées à un niveau d'analyse de la signification globale des mots isolés, sans pouvoir utiliser d'autres informations, comme leur construction interne ou la relation qu'ils entretiennent les uns avec les autres.

Au contraire, les lusophones en début d'apprentissage se situeraient à un niveau où la signification des phrases écrites est surtout extraite à partir de la signification globale des mots. Le décodage graphophonologique ne pose déjà plus de difficultés. Par la suite, les stratégies de compréhension progressent vers un profil de plus en plus similaire à celui rencontré en langue première (*cf.* étude 2). Il s'avère en effet que, dès le deuxième niveau étudié, la compréhension est basée sur des compétences morphologiques plus complexes et plus générales, comme celle d'identification de la construction morphologique de pseudo-mots.

Ainsi, les compétences en compréhension des arabophones et lusophones évoluent mais cette évolution ne semble pas débiter à partir du même « état initial ». En effet les lusophones ne sont plus contraints par les procédures de décodage alors que ce n'est pas le cas des arabophones, même les plus avancés. Là encore, les ressemblances entre le portugais et le français permettent aux lusophones d'aborder la compréhension en langue seconde en continuité avec les apprentissages débutés en langue première, sans avoir, contrairement aux arabophones, à réapprendre de nouvelles règles graphophonologiques.

Globalement, il apparaît assez évident que la distance typologique entre L1 et L2 affecte beaucoup et durablement les mécanismes de lecture (Koda, 2005).

Chapitre 10 : Discussion générale – Conclusion

1. Discussion générale des études sur les langues premières : mise en correspondance des données issues des études 1 et 2

Il nous semble difficile d'appréhender l'apprentissage de la lecture et les mécanismes cognitifs qui lui sont nécessaires sans tenir compte de la diversité des langues. L'influence des spécificités linguistiques sur les processus mis en jeu pour lire constitue d'ailleurs un champ de recherche fécond et extrêmement étendu. Dans ce domaine, la comparaison interlangue se révèle être un outil précieux, à même de déterminer ce qui relève des spécificités des langues ou des systèmes d'écriture et ce qui appartient à des processus généraux (Romdhane, Gombert & Belajouza, 2003).

En formalisant de la sorte les mécanismes communs ou transversaux entre les langues et les facteurs à l'origine des variations, les recherches interlangues sont susceptibles de renforcer l'étendue heuristique des modèles de l'apprentissage de la lecture. Celles réalisées jusqu'à présent se sont essentiellement situées dans le cadre des modèles classiques de l'apprentissage de la reconnaissance des mots écrits (*e.g.* Frith, 1985, 1986 ; Seymour, 1997) mettant en exergue l'importance de la maîtrise du code graphophonologique. Les investigations ont ainsi souvent concerné la rapidité d'acquisition des procédures de décodage. Si le rôle de la conscience phonologique est universel, les langues orthographiquement opaques font toutefois l'objet d'un apprentissage de la lecture plus lent et plus coûteux et peuvent inciter le lecteur à utiliser des unités orthographiques composites plus consistantes (par exemple la rime ; Ziegler, 2004 ; *cf.* pour une revue Demont & Gombert, 2007).

Or, les systèmes orthographiques représentent simultanément les niveaux phonologique et morphologique du langage oral. Le modèle théorique de Gombert (*e.g.* Gombert 2002a, 2003ab) suggère d'ailleurs non seulement l'intervention des traitements graphophonologiques mais aussi ceux du code grapho-sémantique lors de l'apprentissage de la lecture. Pourtant, les variations interlangues sur le rôle des connaissances morphologiques lors de l'apprentissage de la lecture sont très peu explorées. A notre connaissance, seule l'étude de Ku et Anderson (2003), tout en soulignant l'importance de la dimension morphologique dans des langues aussi différentes que le chinois et l'anglais, a pu montrer

que certains aspects de la morphologie sont importants dans certaines langues plus que dans d'autres¹⁸⁰ ou encore à certains moments de l'apprentissage plus qu'à d'autres¹⁸¹.

D'autre part, la comparaison des traitements syntaxiques participant à la lecture dans différentes langues n'a visiblement jamais été étudiée. En revanche, l'utilisation des informations syntaxiques lors de la compréhension orale et leurs variations en fonction des langues ont été analysées dans plusieurs recherches se référant au modèle de compétition de Bates et MacWhinney (1989). Il ressort de ces travaux que les caractéristiques syntaxiques des langues vont privilégier l'usage de certains indices mais que ces indices ne présentent pas tous les mêmes contraintes procédurales pour le système cognitif (voir par exemple Kail, 1991, 1999). De plus, la compréhension orale et la compréhension écrite ne posent pas les mêmes exigences à l'apprenant. En particulier, la gestion de l'attention y est très différente. Ainsi, au cours de l'apprentissage de l'écrit, les processus de compréhension pourront se déclencher facilement si, toutefois, la maîtrise du code est suffisante pour que la reconnaissance de l'ensemble des mots s'effectue rapidement, quasi simultanément et sans effort attentionnel (Gombert & Colé, 2000). En outre, la compréhension des énoncés écrits dépend plus spécifiquement qu'à l'oral d'une analyse syntaxique consciente car le recours à d'autres informations sur la structure du discours (*e.g.* aspects pragmatiques et contextuels) est réduit à l'écrit (Gombert *et al.*, 1996).

La question du degré de contrôle nécessaire pour traiter les informations orthographiques, phonologiques mais aussi morphologiques et syntaxiques de l'écrit s'avère donc cruciale pour comprendre comment s'élaborent les mécanismes de lecture. Le modèle de Gombert (*e.g.* Gombert, 2002a, 2003ab ; Gombert & Demont, 2004) en rend compte. Il est envisagé que dès les premiers contacts avec l'écrit, un travail de découvertes et d'hypothèses débute spontanément. Cet apprentissage de nature implicite constitue le socle sur lequel va s'appuyer l'enseignement explicite de la lecture et les premières connaissances conscientes sur l'écrit, essentielles à cette acquisition. Ces apprentissages explicites permettent, entre autres, de maîtriser les procédures permettant de répéter l'activité de lecture, donc l'exposition aux régularités de l'écrit. Cela a alors pour conséquence de renforcer les apprentissages implicites dont sont issus les automatismes.

Ce modèle s'applique à la reconnaissance des mots mais permet également de comprendre la façon dont s'installent les processus de compréhension, l'objectif même de la lecture.

¹⁸⁰ Par exemple les processus de composition en chinois.

¹⁸¹ Par exemple en début d'apprentissage toujours en chinois.

Considérant, l'état de la recherche sur l'apprentissage de la lecture dans une perspective comparative interlangue d'une part et nous situant dans le cadre du modèle des apprentissages implicites et explicites d'autre part, nous avons mené deux études similaires dans deux langues alphabétiques, se distinguant, entre autres, quant à l'importance relative des dimensions graphophonologique et grapho-sémantique : l'arabe et le portugais. En arabe standard, la dimension grapho-sémantique est déterminante, notamment du fait de la structure morphologique en racine-schème (*e.g.* Ammar, 2002 ; Belajouza & Snoussi, 2006 ; Taouk & Coltheart, 2004). En portugais, la dimension graphophonologique joue un rôle essentiel (*e.g.* Capovilla & Capovilla, 2000 ; Cardoso-Martins, 1995 ; Barrera & Maluf, 2003). Également, le système d'écriture du portugais est moins transparent que le système arabe, lorsque celui-ci est vocalisé.

Nous nous attacherons maintenant à décrire les similarités et les différences entre l'arabe et le portugais quant à la mise en place des connaissances de différents aspects de la langue écrite et leur contribution à la lecture. Il convient de rappeler au préalable que les groupes étudiés dans notre étude ne correspondent pas toujours au même niveau d'apprentissage scolaire, apprentissage qui, du fait des caractéristiques des dispositifs d'enseignement, est différent en Tunisie et au Brésil. En conséquence, la mise en correspondance des résultats ne pourra être que partielle et globale.

Concernant l'évaluation des connaissances, et parmi les éléments comparables, il semble que les épreuves de détection d'intrus sont plus à même de rendre compte d'une sensibilité à certains aspects linguistiques en portugais qu'en arabe¹⁸². Les Brésiliens ont, dès la 3^e année, une sensibilité aux unités graphophonologiques composites de leur langue (la rime et la syllabe), sensibilité qui n'apparaît clairement qu'en 5^e année chez les Tunisiens. Nous avons rapproché l'apparition tardive de la sensibilité aux unités larges en arabe avec le début de la lecture en version non vocalisée (lors de la 5^e année). Ces observations semblent cohérentes avec l'affirmation que les systèmes à orthographe peu transparente orientent les lecteurs vers des unités de traitements plus larges. Il semble également que la sensibilité grapho-morpho-flexionnelle aux marques du féminin et aux conjugaisons verbales, évaluée à l'aide d'une tâche de détection d'intrus, peut émerger chez les Brésiliens (par exemple en 3^e année) mais pas chez les Tunisiens de même niveau. La difficulté de ces derniers à détecter implicitement certaines régularités dans le système de l'arabe standard nous semble devoir être mise en relation avec la situation de diglossie. Parce que l'arabe standard diffère en

¹⁸² Cette remarque n'est pas valable pour la sensibilité aux aspects grapho-morpho-dérivationnels car les évaluations utilisant le paradigme de détection implicite d'intrus échouent systématiquement à la faire émerger.

partie de l'arabe dialectal (par exemple dans l'utilisation des marques flexionnelles), les Tunisiens peuvent plus difficilement se baser sur leur expérience de la langue orale pour effectuer des analogies sur les règles de l'écrit.

Concernant les connaissances « orthographe-syntaxiques » (implicites et explicites), il nous semble préférable de ne pas tenter de comparaisons dans la mesure où les règles étudiées dans les deux langues ne sont pas de même nature : les évaluations portent en arabe sur la structure des compléments, directs ou indirects et en portugais, sur des formes verbales composées.

Quant aux résultats aux épreuves grapho-morphologiques explicites, ils montrent une « inversion », entre l'arabe et le portugais, dans l'élaboration des connaissances des deux types de morphologie. En arabe, les constructions par dérivation sont généralement mieux détectées que les règles grapho-morpho-flexionnelles, alors qu'en portugais, c'est la morphologie flexionnelle qui apparaît plus facile¹⁸³. Les caractéristiques morphologiques de l'arabe expliquent sans doute que les apprentissages de la lecture accentuent les principes de formation des mots par un schème superposé à une racine. Conformément à nos attentes, le processus de transformation par dérivation serait moins privilégié lors des apprentissages de l'écrit en portugais car moins productif et moins prévisible que l'application de règles morpho-flexionnelles.

En ce qui concerne la contribution respective des connaissances précitées à la lecture, le rôle des connaissances morphologiques sur l'identification de mots écrits est attesté, de façon générale, en arabe comme en portugais¹⁸⁴. On perçoit également un profil similaire entre les Tunisiens de 4^e année et les Brésiliens de 3^e année. Pour ces deux groupes, les habiletés en lecture apparaissent surtout déterminées par les connaissances des règles morpho-flexionnelles de la langue, puis par les compétences à identifier la construction morphologique de pseudo-mots. Ce décalage d'un an peut être lié à l'organisation des interventions scolaires. Il est en effet vraisemblable que l'apprentissage des règles grammaticales ne soit pas amorcé au même moment dans la scolarité en Tunisie et au Brésil.

¹⁸³ Cet avantage des arabophones dans le domaine de la morphologie dérivationnelle et des lusophones dans le domaine de la morphologie flexionnelle est également étayé par les variations des compétences d'identification de la construction morphologique de pseudo-mots des deux groupes de même niveau scolaire (5^e année) pour qui les règles s'apparentant à des aspects morpho-dérivationnels sont mieux, ou au contraire moins bien traitées que celles s'apparentant aux aspects morpho-flexionnels.

¹⁸⁴ Nous ne prendrons pas en compte dans les comparaisons les résultats des lusophones de 7^e année, chez qui de surcroît, l'intervention des connaissances grapho-morphologiques n'a pas pu être testée. La comparaison ne se justifie pas car le niveau scolaire est trop éloigné de ceux des autres groupes arabophones.

Par ailleurs, un ensemble d'observation amène à certaines nuances quant à la nature des connaissances morphologiques intervenant.

Chez les Tunisiens de plusieurs niveaux scolaires (3^e et 5^e année), la seule compétence prédisant les habiletés en lecture de mots porte sur des constructions morphologiques diverses (dérivationnelles et flexionnelles, par préfixation, suffixation, et schématisation), indépendantes des aspects formels de l'arabe. En revanche, chez les Brésiliens de mêmes niveaux scolaires, ce sont des connaissances grapho-morphologiques spécifiques au matériel linguistique du portugais qui interviennent, seules ou de façon privilégiée. Ces différences pourraient signifier que la lecture en arabe est liée à des traitements morphologiques plus généraux qu'en portugais. La richesse et la diversité des constructions des mots arabes rendraient plus complexe la mise en œuvre des connaissances morphologiques spécifiques à l'arabe. En revanche, lors de la lecture, les compétences dans l'application des principes productifs de formation morphologique seraient plus économiques et permettraient d'identifier les mots, même si leur combinaison racine/schème n'a jamais été rencontrée ou s'ils correspondent à des formes fléchies peu fréquentes ou inconnues. La lecture en portugais ne semble en revanche pas ou peu requérir ce genre de compétences, les connaissances sur la construction interne du lexique portugais semblant généralement suffisantes.

En ce qui concerne la compréhension en lecture, les points communs sont inexistant, tout au moins à des niveaux comparables. Seuls les Tunisiens de 3^e année semblent mettre en place une analyse syntaxique permettant de repérer l'enveloppe des mots. Les connaissances grapho-syntaxiques évaluées dans cette langue concernent en effet la détection d'une préposition accolée au complément. Rappelons toutefois que celles évaluées en portugais sont d'une autre nature (elles concernent les formes verbales composées), ce qui peut expliquer la présence d'un côté et l'absence de l'autre¹⁸⁵ de connaissances grapho-syntaxiques référant à des aspects différents.

Une autre différence entre les profils de lecture en arabe et en portugais a trait à l'intervention des connaissances grapho-morphologiques spécifiques aux langues considérées. Les connaissances grapho-morpho-dérivationnelles semblent intervenir lors de la compréhension uniquement en arabe (elles apparaissent comme le meilleur prédicteur de la lecture en 4^e année). En portugais, ce serait davantage les connaissances des aspects

¹⁸⁵ En 3^e année, chez les lusophones, aucune connaissance évaluée n'apparaît d'ailleurs liée à la compréhension de phrases écrites. Face à cette constatation, il convient de souligner que l'évaluation de compétences verbales orales fait défaut. Il est probable que de telles mesures soient parvenues à expliquer une partie des habiletés de compréhension.

grapho-morpho-flexionnels qui interviendraient mais à un niveau scolaire supérieur (5^e année) et toutefois de façon moindre par rapport aux connaissances graphophonologiques implicites. A ce même niveau scolaire, la compréhension en arabe apparaît pour sa part dépendante de connaissances morphologiques libérées des aspects superficiels des langues, profil observé également chez les lusophones les plus âgés.

A l'issue de ces différents éléments, nous retiendrons essentiellement :

- *l'importance des processus d'automatisation encore en 5^e année d'apprentissage en portugais mais plus en arabe. L'usage d'un système graphophonologique plus opaque en portugais qu'en arabe vocalisé nous semble susceptible d'expliquer cette différence interlangue ;*
- *l'implication de la dimension grapho-morpho-dérivationnelle en arabe mais l'implication de la dimension grapho-morpho-flexionnelle en portugais. Même s'il est encore difficile de comprendre pourquoi la prise en compte des aspects dérivationnels plus que flexionnels joue un rôle dans la compréhension en arabe, son existence souligne le rôle primordial de la décomposition en racine/schème dans l'accès à la signification des mots et donc en partie à celle des phrases ;*
- *la possibilité d'utiliser rapidement en arabe les principes généraux de construction interne des mots.*

2. Discussion générale de l'étude sur le français langue seconde

Le principal objectif de l'étude 3 était de déterminer dans quelle mesure et en fonction de quels facteurs l'acquisition de l'écrit en langue première alphabétique influence le développement des connaissances linguistiques et l'élaboration des procédures de lecture en français langue seconde.

L'influence des apprentissages préalables en langue maternelle ou première est généralement affirmée dans les théories sur l'apprentissage de la lecture en langue seconde. Le modèle de l'interdépendance développementale (e.g. Cummins, 1989) est souvent cité pour comprendre comment les acquisitions en L1 permettent de développer des compétences cognitivo-linguistiques référentielles applicables ou transférables à la L2. Pour certains auteurs, l'interdépendance des langues peut être appréhendée au travers de compétences dépourvues d'objectifs communicationnels et portant sur les principes généraux du fonctionnement linguistique des langues, indépendamment de leurs

caractéristiques formelles (e.g. Cummins, 1991 ; Lucchini, 2002). Pour d'autres, la relation entre la maîtrise de l'écrit dans différentes langues repose sur des processus centraux, cognitifs et linguistiques¹⁸⁶ (e.g. Gholamain & Geva, 1999 ; Geva & Siegel, 2000).

Quel que soit le positionnement théorique des auteurs, la plupart de ces recherches souligne l'existence d'une interdépendance entre la reconnaissance des mots écrits dans les deux langues en contact (Durgunoğlu *et al.*, 2002 ; Da Fontoura & Siegel, 1995 etc.). Cette affirmation semble en particulier valable pour les langues alphabétiques, même si elles n'emploient pas le même alphabet (e.g. Abu-Rabia & Siegel, 2002, 2003 ; Bialystok *et al.*, 2005a ; Kahn-Horwitz *et al.*, 2005). L'observation d'une conscience phonologique « parallèle » dans les deux langues, permet en effet de soutenir que la reconnaissance des mots écrits nécessite des traitements phonologiques communs, dans leurs principes, aux langues en contact (e.g. Comeau *et al.*, 1999 ; Durgunoğlu, 1998, Durgunoğlu *et al.*, 1993).

En revanche, les relations interlangues, au niveau des traitements morphologiques mais aussi syntaxiques ne s'avèrent pas aussi convaincantes. La dimension morphologique n'a été étudiée que récemment par Durgunoğlu *et al.* (2002) et par Schiff et Calif (2007) toutefois à l'aide de mesures qui nous semblent devoir être améliorées (*cf.* paragraphes 4.2.2 et 4.2.4 du chapitre 5). La même remarque peut être formulée à propos des recherches sur la dimension syntaxique (e.g. Abu-Rabia & Siegel, 2002 ; Da Fontoura & Siegel, 1995 ; *cf.* paragraphe 4.2.4, chapitre 5).

Dans une autre perspective, l'influence de l'acquisition de l'écrit en langue première est étudiée par le biais des comparaisons interlangues. Les recherches menées dans ce cadre, souvent attachées aux différences entre les langues idéographiques et alphabétiques, ont démontré les difficultés ou les variations de performances à diverses tâches de lecture selon la typologie des langues 1 (e.g. Fender, 2003 ; Wade-Woolley, 1998 ; Wang *et al.*, 2003). L'étude de Koda (2000) comparant des adultes de langue première chinoise et coréenne a également pu mettre en évidence l'influence des caractéristiques morphologiques des langues premières sur les compétences morphologiques développées en anglais langue seconde. Ainsi, les caractéristiques orthographiques des langues (mais vraisemblablement aussi morphologiques, etc.) sont susceptibles de rendre compte comment les expériences en L1 affectent les traitements lors de l'apprentissage de l'écrit en langue seconde. Cependant, il apparaît encore difficile de savoir d'une part, si l'impact de la distance orthographique des systèmes se limite à des L1 qui impliquent, par rapport à la L2, des changements radicaux

¹⁸⁶ Par exemple, les habiletés verbales, la mémoire verbale de travail, la rapidité et la précision des procédures d'accès lexical.

dans l'organisation des configurations visuo-spatiales, et d'autre part quelle est la place du degré de transparence orthographique pour comprendre les différences entre les apprenants (Yamada, 2004).

Ainsi, la perspective d'allier ces deux orientations nous a semblé intéressante. L'analyse des relations entre L1 et L2 s'est donc effectuée en comparant les phénomènes entre des apprenants dont les expériences linguistiques en langue première sont très différentes. En outre, notre étude *cross-linguistique* s'est centrée sur l'apprentissage de la lecture en français langue seconde en situation scolaire à l'étranger (Tunisie et Brésil).

Les spécificités des L1 (l'arabe et le portugais) ont été envisagées sur la base des résultats des recherches réalisées dans les deux langues, résultats globalement confirmés par les deux précédentes études. Le degré de proximité que ces langues entretiennent avec le français a également été pris en compte. Ainsi, nous pouvons rappeler que les spécificités de l'arabe sont liées à l'importance de la dimension grapho-morphologique (du fait notamment d'une structure morphologique non linéaire quasi systématique) mais aussi que son système orthographique est très différent de celui du français. Les spécificités du portugais relèvent pour leur part de l'automatisation des traitements graphophonologiques, même si les traitements grapho-morphologiques peuvent également intervenir. Cette langue latine apparaît en outre dans sa forme écrite très proche du français (Seymour *et al.*, 2003).

Notre étude a permis d'approfondir quelques interrogations sur les relations entre les langues 1 et 2 en prenant en considération le domaine de connaissances accentué lors des apprentissages de la lecture en L1 mais également la proximité linguistique entre les langues 1 et 2.

Bien que l'ensemble des analyses souhaitées n'ait pas pu être mené, du fait d'effets plafonds ou de réponses aléatoires, nos résultats ont mis en évidence un lien d'interdépendance entre les deux langues des apprenants, soutenant ainsi l'hypothèse de Cummins (1989). Ce lien n'est cependant pas immédiat et ne semble se réaliser que dans certaines conditions. Plus particulièrement, il pourrait concerner des connaissances de nature analogique, élaborées sans intention, tout au moins lorsque les langues partagent, dans ce domaine, des relations de proximité évidente et/ou lorsque le domaine considéré s'avère privilégié lors des apprentissages (en l'occurrence le domaine graphophonologique en portugais).

Il pourrait également concerner des connaissances de nature analytique développées dans des langues aussi distantes que l'arabe et le français, tout au moins lorsque le domaine

linguistique considéré est particulièrement déterminant lors des apprentissage en langue première (en l'occurrence le domaine grapho-morphologique en arabe).

Ainsi, à l'issue de ces considérations, il nous semble nécessaire, pour mieux comprendre comment l'interdépendance linguistique s'élabore, de déterminer les aspects linguistiques privilégiés dans les langues premières et de les prendre en compte simultanément avec le degré de proximité interlangue et la nature des connaissances (implicite ou explicite).

Par ailleurs, nous avons préféré envisager le transfert des compétences, souvent avancé dans la littérature, en terme d'actualisation, en L2, de connaissances explicites élaborées en L1 (ou de transversalité des compétences analytiques). Cette option nous semble permettre de mieux rendre compte des mécanismes communs ou généraux, tout en rendant compte de « l'empreinte » des caractéristiques des langues premières sur les connaissances ultérieures (ne serait-ce qu'en termes de poids ou de focalisation de l'attention sur certaines dimensions). Les aspects implicites des connaissances nous semblent relever davantage d'une continuité. Nous postulons que cette continuité ne serait effective que dans le cas où les langues partagent suffisamment de similarités formelles pour rendre leurs analogies saillantes.

D'autre part, l'étude des connaissances contribuant à la lecture en français langue seconde en fonction de la langue première et du niveau nous amène à plusieurs réflexions :

Bien que, et comme nous pouvions nous y attendre, l'apprentissage d'un nouveau code orthographique complexifie durablement l'acquisition de la lecture en français langue seconde chez les arabophones, la relation entre les habiletés en lecture dans les deux langues apparaît. Cette relation reflèterait l'existence de mécanismes sous-jacents, communs aux langues alphabétiques. Cette relation s'avère cependant moins forte comparativement à ce qui est observé chez les lusophones, pour qui l'application de règles graphophonologiques proches de leur L1 ne peut être que plus facile.

En outre, les arabophones mobilisent aussi des connaissances explicites acquises lors de leurs premiers apprentissages de l'écrit en arabe. Il semble ainsi possible que les compétences nécessaires à la lecture en arabe puissent également intervenir lors de la lecture en français langue seconde. Plus tardivement, l'intervention d'une sensibilité à certaines unités morphologiques du français renforce le rôle de la morphologie et les progrès quant aux allocations de l'attention dans la mise en place de ces traitements.

Concernant les processus de compréhension, les interprétations que nous avons avancées soulignent un décalage dans la façon d'aborder et d'accéder à la signification des phrases. Les arabophones apparaissent au début complètement dépendants de leur capacité à décoder les mots puis progresseraient lentement vers un déchiffrement plus fluide. Les lusophones, au contraire dégagés des contraintes attentionnelles liées aux traitements séquentiels, parviennent plus vite à recourir aux mêmes stratégies que celles mises en place en portugais (*i.e.* mobilisant des compétences morphologiques décontextualisées). Ceci souligne là encore la correspondance des mécanismes de lecture dans des langues proches.

Au terme de nos analyses, les phénomènes observés amènent à une double conclusion :

- d'une part, les acquisitions réalisées en langue première jouent un rôle à plusieurs niveaux, même lorsque les langues ont des caractéristiques très différentes de la langue seconde ;

- d'autre part, la proximité entre les langues 1 et 2 facilite non seulement l'acquisition de la reconnaissance des mots écrits mais permet aussi de mettre en place facilement, lors de la compréhension, des processus de plus hauts niveaux.

3. Conclusion

A l'heure actuelle, les recherches interlangues ont tendance à se développer. Qu'elles se centrent sur l'apprentissage de la lecture en langue maternelle ou en langue seconde, elles constituent un domaine vaste et particulièrement fécond pour formaliser les théories générales sur les mécanismes de ces apprentissages et déterminer précisément les universaux dans l'acquisition du langage mais aussi les variations en fonction des spécificités des langues ou de l'ordre de leur acquisition.

L'intérêt heuristique des investigations interlangues va de pair avec la difficulté à les mettre en œuvre. Outre la nécessité de veiller à la comparabilité des outils d'évaluation et du matériel linguistique, les variations des contenus et de l'organisation des apprentissages scolaires constituent également un facteur à ne pas sous-estimer. Ces éléments posent ainsi certaines limites aux comparaisons que l'on peut effectuer.

Il n'en reste pas moins intéressant de multiplier les recherches sur des langues dont les caractéristiques sont spécifiques et originales par rapport aux différentes dimensions linguistiques impliquées lors de la lecture. Nos recherches se sont situées dans cette

perspective en comparant deux langues encore peu étudiées par rapport à d'autres langues indo-européennes.

Les études en langue première (étude 1 et 2) ont tenté de montrer que l'arabe et le portugais, de par leurs caractéristiques différentes, ne rendent pas les mêmes informations linguistiques aussi saillantes et nécessaires lors de la lecture.

Le système arabe se caractérise avant tout par l'importance de la morphologie. Les travaux sur cette langue sont ainsi susceptibles d'apporter des résultats intéressants quant au rôle des traitements morphologiques lors de l'apprentissage de la lecture.

L'élaboration d'une recherche ultérieure, visant à consolider et approfondir les résultats des études présentes, pourra s'intéresser à des niveaux scolaires moins avancés et davantage similaires entre les deux mêmes langues. Elle pourra aussi accorder une place plus importante à l'arabe dialectal, la véritable langue maternelle orale des tunisiens. Les différences entre l'arabe dialectal et standard interrogent en effet sur la façon dont les connaissances de nature implicite, liées à l'oral, interviennent comme support à l'apprentissage explicite de la lecture

Dans la continuité des expérimentations précédentes, l'étude 3 sur le français langue seconde s'est attachée à déterminer comment les expériences préalables en lecture, liées aux caractéristiques des langues premières, peuvent orienter le système cognitif de l'apprenti lecteur en langue seconde vers la prise en considération (non intentionnelle ou contrôlée) de certaines informations du système écrit français. Très brièvement, il est apparu une plus grande continuité entre les apprentissages (par exemple dans le domaine de la sensibilité graphophonologique et des habiletés en lecture) lorsque les systèmes écrits sont proches. Cependant, même lorsque les langues utilisent des systèmes très contrastés, la relation entre certaines connaissances explicites développées dans les deux langues et privilégiées en L1, est aussi visible. Ainsi, la mobilisation, lors de la lecture en français langue seconde, de connaissances essentielles pour la lecture en langue première s'avère possible.

Le rôle des expériences préalables de l'écrit sur le développement des connaissances et l'apprentissage de la lecture en français langue seconde apparaît donc important, même si nous reconnaissons que la maîtrise des langues secondes est également déterminée par d'autres facteurs.

Il faudra, pour l'avenir, approfondir et répliquer les résultats en tentant de contrôler au mieux les conditions d'apprentissage du français, notamment pour ce qui concerne le temps

écoulé entre le début de l'apprentissage de l'écrit en langue première puis en langue seconde (responsable d'une plus ou moins grande maîtrise de l'écrit en L1 au moment où sont initiés les apprentissages en L2) et la fréquence des séquences d'enseignement de la langue seconde.

Malgré tout, certaines réflexions émanant de ce travail nous semblent converger vers des préoccupations plus didactiques sur l'enseignement de la lecture en général, et l'enseignement des langues secondes/étrangères, du français à l'étranger ou par les populations allophones en particulier. Certains de nos résultats ont par exemple suggéré, outre l'importance de la morphologie des mots, l'intérêt de développer des compétences analytiques décontextualisées pour faciliter l'application des principes de construction productifs. En même temps, favoriser la fréquentation de l'écrit apparaît indispensable. Au sujet des apprentissages implicites, nous avons souligné l'influence de la proximité entre langue première et langue seconde. Notamment, les relations de proximité ou d'éloignement des systèmes nous semblent devoir être prises en considération, lors de la mise en œuvre des dispositifs les plus favorables aux apprentissages implicites en langue seconde dans le cadre d'une pédagogie différenciée.

Enfin, la prise en compte des caractéristiques et de l'influence des langues premières lors des apprentissages en langue seconde peut aider à comprendre à quelles difficultés se heurtent les apprenants mais aussi mettre en exergue les atouts pouvant servir de leviers pour l'apprentissage des langues. L'introduction de la langue première dans les dispositifs d'enseignement destinés à des enfants allophones ou de langues minoritaires par exemple, s'inscrit dans cette perspective. Il convient toutefois d'éviter de faire de la langue première une simple béquille aux apprentissages de la langue seconde qui peut être éventuellement plus valorisée. En ce sens, les dispositifs bilingues de transition, qui suppriment la langue première une fois qu'elle a rempli son rôle, n'ont pas obtenus, selon certains auteurs, les effets positifs escomptés (*cf.* Lucchini, 2002).

Liste des figures

Figure 1: Le modèle à double fondation de Seymour (1997).....	13
Figure 2: Représentation du modèle de Seymour intégrant le développement de la conscience linguistique (Seymour, Aro, & Erskine, 2003).	14
Figure 3 : Le modèle connexionniste de Seidenberg et McClelland (1989) proposé par Gombert (2003b).....	16
Figure 4: Représentation du modèle connexionniste de Plaut et Gonnerman (2000) intégrant la morphologie	18
Figure 5: Le système initial et l'apprentissage implicite de la lecture (Gombert, 2003b)....	21
Figure 6: Apprentissage implicite et explicite de la lecture (Gombert, 2003b).....	22
Figure 7: le modèle d'activation interactive bilingue de Grainger (2002)	116
Figure 8: Performances moyennes obtenues aux épreuves explicites de détection d'intrus grapho-sémantique en 3 ^e , 4 ^e et 5 ^e année.	183
Figure 9 : Performances moyennes (sur 3) obtenues à l'épreuve implicite de détection d'intrus graphophonologique en 3 ^e , 4 ^e et 5 ^e année en fonction de l'unité phonologique étudiée.....	184
Figure 10 : Performances moyennes (sur 12) obtenues aux épreuves implicites de détection d'intrus graphophonologique, grapho-morpho-flexionnelle et grapho-syntaxique en 3 ^e , 5 ^e et 7 ^e année.	225
Figure 11 : Performances moyennes (sur 12) obtenues aux épreuves explicites de détection d'intrus grapho-sémantique en 3 ^e , 5 ^e et 7 ^e année	227
Figure 12 : Performances moyennes (sur 3) obtenues à l'épreuve implicite de détection d'intrus graphophonologique en 3 ^e , 4 ^e et 5 ^e année en fonction de l'unité phonologique étudiée.....	230
Figure 13 : Performances moyennes (sur 12) obtenues aux épreuves implicites de détection d'intrus graphophonologique, grapho-morpho-flexionnelle en fonction de la langue première	276
Figure 14 : Performances moyennes (sur 12) obtenues aux épreuves grapho-morpho-flexionnelles de détection d'intrus en fonction de la langue première.....	278
Figure 15 : Performances moyennes (sur 12) obtenues aux épreuves explicites de détection d'intrus grapho-sémantiques en fonction de la langue première.....	280
Figure 16 : Performances moyennes (sur 12) obtenues aux 3 épreuves explicites de détection d'intrus grapho-sémantiques en fonction du niveau	281

Liste des tableaux

Tableau 1: Classification des langues proposée par Seymour <i>et al.</i> (2003)	37
Tableau 2 : Liste des lettres de l'alphabet arabe	50
Tableau 3 : Illustration de la structure morphologique des mots arabes à partir de la racine [κτβ] كتب	54
Tableau 4 : Répartition des élèves en fonction de leur niveau scolaire	163
Tableau 5 : Epreuves évaluant les connaissances sur les dimensions graphophonologique et grapho-sémantique	174
Tableau 6 : Performances moyennes, écart-types, minima et maxima au test de lecture de mots en une minute selon le niveau scolaire	177
Tableau 7 : Performances moyennes de réussite (sur 22), écart-types, minima, maxima obtenus au test de compréhension de phrases selon le niveau scolaire	178
Tableau 8 : Performances moyennes de réussite sur 12 (et écart-types) obtenues aux épreuves implicites et explicites de détection d'intrus en fonction du niveau scolaire, tests-t de comparaison à la valeur 6 (niveau du hasard)	179
Tableau 9 : Performances moyennes de réussite sur 12 (et écart-types) obtenues aux épreuves explicites de détection d'intrus grapho-morphologique en fonction du niveau scolaire et tests t de comparaison à la valeur 6 (niveau du hasard)	181
Tableau 10 : Performances moyennes de réussite sur 12 (et écart-types) obtenues aux épreuves explicites de détection d'intrus grapho-sémantique en fonction du niveau scolaire et tests t de comparaison à la valeur 6 (niveau du hasard)	182
Tableau 11 : Performances moyennes de réussite (et écart-types) obtenues à l'épreuve implicite de détection d'intrus graphophonologique en fonction du type d'items et du niveau scolaire et tests t de comparaison à la valeur 1/2 (niveau du hasard)	184
Tableau 12 : Performances moyennes sur 24, écart-types, maxima, minima obtenus à l'épreuve d'identification de la construction morphologique de pseudo-mots en fonction du niveau scolaire	186
Tableau 13 : Performances moyennes sur 8 (écart-types) et en pourcentage, obtenues à l'épreuve d'identification de la construction morphologique de pseudo-mots en fonction du niveau scolaire et de la nature de la construction morphologique des pseudo-mots.	187
Tableau 14 : Performances moyennes sur 12 (écart-types) et en pourcentage, obtenues à l'épreuve d'identification de la construction morphologique de pseudo-mots en fonction du niveau scolaire et de la règle morphologique impliquée.	187
Tableau 15 : Récapitulatif des épreuves ou groupes d'items sélectionnés pour les analyses suivantes à chaque niveau scolaire et synthèse des principaux résultats obtenus (en %) ...	189
Tableau 16 : Coefficients de corrélation de Bravais Pearson entre les épreuves ou groupe d'items sélectionnés en 3 ^e , 4 ^e et 5 ^e année.	190
Tableau 17 : Contribution des performances aux différentes épreuves ou groupes d'items sélectionnés aux facteurs extraits des analyses en composantes principales pour chaque niveau scolaire	192
Tableau 18 : Coefficients de corrélation entre les scores aux épreuves ou groupes d'items sélectionnés et le niveau de lecture de mots en 3 ^e , 4 ^e et 5 ^e année	195
Tableau 19 : Contribution des variables entrées dans le modèle de régression sur le niveau de lecture de mots en 3 ^e année	197
Tableau 20 : Contribution des variables entrées dans le modèle de régression sur le niveau de lecture de mots en 4 ^e année	197
Tableau 21 : Contribution des variables entrées dans le modèle de régression sur le niveau de lecture de mots en 5 ^e année	197
Tableau 22 : Coefficients de corrélation entre les différents scores et le niveau de compréhension de phrases en 3 ^e , 4 ^e et 5 ^e année	198

Tableau 23 : Contribution des variables entrées dans le modèle de régression sur le niveau en compréhension de phrases en 3 ^e année	200
Tableau 24 : Contribution des variables entrées dans le modèle de régression sur le niveau en compréhension de phrases en 4 ^e année	200
Tableau 25 : Contribution des variables entrées dans le modèle de régression sur le niveau en compréhension de phrases en 5 ^e année	201
Tableau 26 : Répartition des élèves en fonction de leur niveau scolaire	209
Tableau 27 : Epreuves évaluant les connaissances sur les dimensions graphophonologique et grapho-sémantique	218
Tableau 28 : Moyennes, écart-types, minima et maxima au test de lecture de mots en une minute selon le niveau scolaire	221
Tableau 29 : Performances moyennes de réussite (sur 22), écart-types, minima, maxima au test de compréhension de phrases selon le niveau scolaire	221
Tableau 30 : Performances moyennes de réussite sur 12 (et écart-types) obtenues aux épreuves implicites et explicites de détection d'intrus en fonction du niveau scolaire, tests-t de comparaison à la valeur 6 (niveau du hasard)	222
Tableau 31 : Performances moyennes de réussite sur 12 (et écart-types) obtenues aux 3 épreuves implicites de détection d'intrus sélectionnées, en 3 ^e , 5 ^e et 7 ^e année et tests t de comparaison à la valeur 6 (niveau du hasard)	224
Tableau 32 : Performances moyennes de réussite sur 12 (et écart-types) obtenues aux épreuves explicites de détection d'intrus grapho-sémantique en 3 ^e et 5 ^e année et tests t de comparaison à la valeur 6 (niveau du hasard)	226
Tableau 33 : Performances moyennes de réussite sur 12 (et écart-types) obtenues aux épreuves implicites et explicites de détection d'intrus grapho-morpho-flexionnel et grapho-syntaxique en 3 ^e et 4 ^e année	227
Tableau 34 : Performances moyennes de réussite (et écart-types) obtenues à l'épreuve implicite de détection d'intrus graphophonologique en fonction du type d'items et du niveau scolaire et tests t de comparaison à la valeur 1/2 (niveau du hasard)	229
Tableau 35 : Performances moyennes sur 24, écart-types, maxima, minima obtenus à l'épreuve d'identification de la construction morphologique de pseudo-mots en fonction du niveau scolaire	232
Tableau 36 : Performances moyennes sur 24, écart-types, maxima, minima obtenus à l'épreuve d'identification de la construction morphologique de pseudo-mots en fonction du niveau scolaire	233
Tableau 37 : Performances moyennes (écart-types) et en pourcentage, obtenues à l'épreuve d'identification de la construction morphologique de pseudo-mots en 3 ^e , 5 ^e et 7 ^e année, en fonction du type de règle morphologique	233
Tableau 38 : Performances moyennes sur 4 (écart-types) et en pourcentage, obtenues à l'épreuve d'identification de la construction morphologique de pseudo-mots pour l'ensemble des groupes (N = 106), en fonction de la nature de la construction et du type de règle morphologique	233
Tableau 39 : Récapitulatif des épreuves sélectionnées pour les analyses suivantes à chaque niveau scolaire et synthèse des principaux résultats obtenus (en %) ¹	235
Tableau 40 : Coefficients de corrélation de Bravais Pearson entre les épreuves sélectionnées en 3 ^e , 5 ^e et 7 ^e année	237
Tableau 41 : Contribution des performances aux différentes épreuves sélectionnées aux facteurs extraits des analyses en composantes principales en 3 ^e , 5 ^e et 7 ^e année	239
Tableau 42 : Coefficients de corrélation entre les scores aux épreuves sélectionnées et le niveau de lecture de mots en 3 ^e , 5 ^e et 7 ^e année	242
Tableau 43 : Contribution des variables entrées dans le modèle de régression sur le niveau de lecture de mots en 3 ^e année	243

Tableau 44 : Contribution de la variable entrée dans le modèle de régression sur le niveau de lecture de mots en 5 ^e année	243
Tableau 45 : Contribution de la variable entrée dans le modèle de régression sur le niveau de lecture de mots en 7 ^e année	244
Tableau 46 : Coefficients de corrélation entre les différents scores et le niveau de compréhension de phrases en 3 ^e , 5 ^e et 7 ^e année.....	245
Tableau 47 : Contribution des variables entrées dans le modèle de régression sur le niveau de compréhension de phrases en 5 ^e année.....	246
Tableau 48 : Contribution de la variable entrée dans le modèle de régression sur le niveau de compréhension de phrases en 7 ^e année.....	246
Tableau 49 : Caractéristiques de la population retenue pour l'expérimentation 3	255
Tableau 50 : Epreuves évaluant les connaissances sur les dimensions graphophonologique et grapho-sémantique en langue première (arabe vs. portugais) et en français langue seconde.....	265
Tableau 51 : Performances moyennes, en nombre de mots lus correctement en une minute (et écart-types) au test LUM en FL2 en fonction de la L1 et du niveau	270
Tableau 52 : Performances moyennes, sur 15 (et écart-types) au test de compréhension de phrases écrites en FL2 en fonction de la L1 et du niveau.....	271
Tableau 53 : Performances moyennes de réussite sur 12 (et écart-types) obtenues aux épreuves implicites et explicites de détection d'intrus en fonction de la langue première et du niveau et tests-t de comparaison à la valeur 6 (niveau du hasard).....	272
Tableau 54 : Performances moyennes sur 12 (et écart-types) obtenues aux épreuves de détection implicite d'intrus graphophonologique et grapho-morpho-flexionnelle FL2 en fonction de la L1 et du niveau.	275
Tableau 55 : Résultats de la MANCOVA sur les épreuves implicites graphophonologique et grapho-morpho-flexionnelle	275
Tableau 56 : Performances moyennes sur 12 (et écart-types) obtenues aux épreuves de détection d'intrus grapho-morpho-flexionnel implicite et explicite en fonction de la langue première et du niveau.....	277
Tableau 57 : Résultats de la MANCOVA sur les épreuves implicite et explicite grapho-morpho-flexionnelles	277
Tableau 58 : Performances moyennes sur 12 (et écart-types) obtenues aux épreuves de détection explicite d'intrus grapho-morphologiques et grapho-syntaxique en FL2 en fonction de la L1 et du niveau	279
Tableau 59 : Résultats de la MANCOVA sur les épreuves explicites grapho-sémantiques	280
Tableau 60 : Performances moyennes (et écart-types) obtenues à l'épreuve de détection d'intrus graphophonologique implicite en fonction de l'unité, de la langue 1 et du niveau	282
Tableau 61 : Résultats de la MANCOVA sur l'épreuve implicite graphophonologique ...	283
Tableau 62 : Réponses moyennes sur 4 (écart-types) et en % aux items A/R de l'épreuve de plausibilité lexicale en FL2 selon la L1 et tests t de comparaison au hasard (2/4 ou 50 %)	284
Tableau 63 : Réponses moyennes sur 4 (écart-types) et en % aux items 2M/B de l'épreuve de plausibilité lexicale en FL2 selon la L1 et tests t de comparaison au hasard (2/4 ou 50 %)	285
Tableau 64 : Réponses moyennes sur 4 (écart-types) et en % aux items B/R de l'épreuve de plausibilité lexicale en FL2 selon la L1 et tests t de comparaison au hasard (2/4 ou 50 %)	285
Tableau 65 : Réponses moyennes sur 4 (écart-types) et en % aux items B/A de l'épreuve de plausibilité lexicale en FL2 selon la L1 et tests t de comparaison au hasard (2/4 ou 50 %)	286

Tableau 66 : Réponses moyennes sur 4 (écart-types) et en % aux items 2M/A de l'épreuve de plausibilité lexicale en FL2 selon la L1 et tests t de comparaison au hasard (2/4 ou 50 %)	286
Tableau 67 : Récapitulatif des épreuves sélectionnées pour les analyses suivantes pour chaque groupe et synthèse des principaux résultats obtenus (en %)	288
Tableau 68 : Coefficients de corrélation entre les épreuves en FL2 sélectionnées, aux niveaux A, B et C chez les arabophones	289
Tableau 69 : Contribution des performances aux différentes épreuves sélectionnées aux facteurs extraits des analyses en composantes principales pour chaque niveau chez les arabophones	290
Tableau 70 : Coefficients de corrélation entre les épreuves en FL2 sélectionnées, aux niveaux A, B et C chez les lusophones	293
Tableau 71 : Contribution des performances aux différentes épreuves sélectionnées aux facteurs extraits des analyses en composantes principales pour chaque niveau chez les lusophones	294
Tableau 72 : Epreuves en arabe et en français entrées dans les analyses sur le lien entre L1 et FL2 en fonction des résultats obtenus	297
Tableau 73 : Corrélations entre les épreuves en L1 et FL2 chez les arabophones de niveaux A, B et C	298
Tableau 74 : Contribution des variables entrées dans le modèle de régression expliquant les performances en détection explicite d'intrus grapho-morpho-dérivationnel en FL2 chez les arabophones de niveau B	299
Tableau 75 : Contribution des variables entrées dans le modèle de régression expliquant les performances en détection explicite d'intrus grapho-morpho-dérivationnel en FL2 chez les arabophones de niveau C	300
Tableau 76 : Contribution des variables entrées dans le modèle de régression expliquant les performances en détection explicite d'intrus grapho-morpho-flexionnel en FL2 chez les arabophones de niveau C	300
Tableau 77 : Contribution des variables entrées dans le modèle de régression expliquant les performances en détection explicite d'intrus grapho-syntaxique en FL2 chez les arabophones de niveau C	301
Tableau 78 : Epreuves en portugais et en français entrées dans les analyses sur le lien entre L1 et FL2 en fonction des résultats obtenus	304
Tableau 79 : Corrélations entre les épreuves en L1 et FL2 chez les lusophones de niveaux A, B et C	305
Tableau 80 : Contribution des variables entrées dans le modèle de régression expliquant les performances en détection explicite d'intrus grapho-syntaxique en FL2 chez les lusophones de niveau C	307
Tableau 81 : Coefficients de corrélation entre les différents scores et le niveau de lecture de mots en FL2 chez les arabophones de niveaux A, B et C	309
Tableau 82 : Contribution des variables entrées dans le modèle de régression sur le niveau de lecture de mots en FL2 chez les arabophones de niveau A	311
Tableau 83 : Contribution des variables entrées dans le modèle de régression sur le niveau de lecture de mots en FL2 chez les arabophones de niveau B	311
Tableau 84 : Contribution des variables entrées dans le modèle de régression sur le niveau de lecture de mots en FL2 chez les arabophones de niveau C	312
Tableau 85 : Coefficients de corrélation entre les différents scores et le niveau de lecture de mots en FL2 chez les lusophones de niveaux A, B et C	313
Tableau 86 : Contribution des variables entrées dans le modèle de régression sur le niveau de lecture de mots en FL2 chez les lusophones de niveau A	315
Tableau 87 : Contribution des variables entrées dans le modèle de régression sur le niveau de lecture de mots en FL2 chez les lusophones de niveau B	315

Tableau 88 : Contribution des variables entrées dans le modèle de régression sur le niveau de lecture de mots en FL2 chez les lusophones de niveau C	315
Tableau 89 : Coefficients de corrélation entre les différents scores et le niveau de compréhension de phrases chez les arabophones de niveaux A, B et C.....	316
Tableau 90 : Contribution de la variable lecture de mots en FL2 au niveau en compréhension de phrases en FL2 chez les arabophones de niveau A	318
Tableau 91 : Contribution des variables entrées dans le modèle de régression sur le niveau de compréhension de phrases en FL2 chez les arabophones de niveau B	318
Tableau 92 : Contribution des variables entrées dans le modèle de régression sur le niveau de compréhension de phrases en FL2 chez les arabophones de niveau C	319
Tableau 93 : Coefficients de corrélation entre les différents scores et le niveau de compréhension de phrases chez les lusophones de niveaux A, B et C	320
Tableau 94 : Contribution de la variable vocabulaire FL2 au niveau en compréhension de phrases en FL2 chez les lusophones de niveau A	321
Tableau 95 : Contribution de la variable identification de la construction morphologique de pseudo-mots au niveau en compréhension de phrases en FL2 chez les lusophones de niveau B.....	321
Tableau 96 : Contribution de la variable entrées dans le modèle de régression sur le niveau de compréhension de phrases en FL2 chez les lusophones de niveau C	321

Références bibliographiques

- Abu-Rabia, S. (1995). Learning to read in Arabic: Reading, syntactic, orthographic and working memory skills in normally achieving and poor readers. *Reading Psychology: an International Quarterly*, 16, 351-394.
- Abu-Rabia, S. (1996). The role of vowels and context in the reading of highly skilled native Arabic readers. *Journal of psycholinguistic research*, 25(6), 629-641.
- Abu-Rabia, S. (1997a). Reading in Arabic orthography: The effect of vowels and context on reading accuracy of poor and skilled native Arabic readers. *Reading and Writing*, 9(1), 65-78.
- Abu-Rabia, S. (1997b). Reading in Arabic orthography: The effect of vowels and context on reading accuracy of poor and skilled native Arabic readers in reading paragraphs, sentences, and isolated words. *Journal of Psycholinguistic Research*, 26(4), 465-482.
- Abu-Rabia, S. (1998). Reading Arabic texts: Effects of text type, reader type and vowelization. *Reading and Writing*, 10(2), 105-119.
- Abu-Rabia, S. (1999). The effect of Arabic vowels on the reading comprehension of second- and sixth-grade native Arab children. *Journal of Psycholinguistic Research*, 28(1), 93-101.
- Abu-Rabia, S. (2000). Effects of exposure to literary Arabic on reading comprehension in a diglossic situation. *Reading and Writing*, 13(1), 147-157.
- Abu-Rabia, S. (2001). The role of vowels in reading Semitic scripts: Data from Arabic and Hebrew. *Reading and Writing*, 14(1), 39-59.
- Abu-Rabia, S. (2003). The influence of working memory on reading and creative writing processes in a second language. *Educational Psychology*, 23(2), 209-219.
- Abu-Rabia, S. (2007). The role of morphology and short vowelization in reading Arabic among normal and dyslexic readers in grades 3, 6, 9, and 12. *Journal of Psycholinguist Research*, 36, 89-106.
- Abu-Rabia, S., & Siegel, L. S. (2002). Reading, syntactic, orthographic, and working memory skills of bilingual Arabic-English speaking Canadian children. *Journal of Psycholinguistic Research*, 31(6), 661-678.
- Abu-Rabia, S., & Siegel, L. S. (2003). Reading skills in three orthographies: The case of trilingual Arabic-Hebrew-English-speaking Arab children. *Reading and Writing*, 16(7), 611-634.

- Abu-Rabia, S., Share, D., & Mansour, M. S. (2003). Word recognition and basic cognitive processes among reading-disabled and normal readers in Arabic. *Reading and Writing, 16*(5), 423-442.
- Aidinis, A., & Nunes, T. (2001). The role of different levels of phonological awareness in the development of reading and spelling in Greek. *Reading and Writing, 14*(1), 145-177.
- Akiki, P. C. (1994). *Les stratégies d'apprentissage du français écrit d'élèves libanais. Propositions méthodologiques* (Vol. XXXII). Kaslik Liban: Université Saint Esprit.
- Alegria, J., & Morais, J. (1979). Le développement de l'habileté d'analyse phonétique de la parole et l'apprentissage de la lecture. *Archive de Psychologie, 47*, 251-270.
- Altmann, G. T., Garnham, A., & Dennis, Y. (1992). Avoiding the garden path: Eye movements in context. *Journal of Memory and Language, 31*(5), 685-712.
- Ammar, M. (1997). *Les stratégies d'identification de mots écrits en arabe*. Unpublished Doctorat de Sciences de l'Education, Université de Nantes, Nantes.
- Ammar, M. (2002). L'assemblage phonologique: Sa nature et son fonctionnement chez des enfants lecteurs en arabe. *Enfance, 54*(2), 155-163.
- Ammar, M. (2003). Le fonctionnement de l'assemblage phonologique chez des enfants lecteurs en arabe. In M. N. Romdhane, J. E. Gombert & M. Belajouza (Eds.), *L'apprentissage de la lecture: Perspectives comparatives* (pp. 69-79). Rennes: Presses Universitaires de Rennes.
- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Armand, F. (2000). Le rôle des capacités métalinguistiques et de la compétence langagière orale dans l'apprentissage de la lecture en français langue première et seconde. *La revue canadienne des langues vivantes, 56*, 470-495.
- Aro, M., & Wimmer, H. (2003). Learning to read: English in comparison to six more regular orthographies. *Applied Psycholinguistics, 24*(4), 621-635.
- Arrivé, M., Gadet, F., & Galmiche, M. (1986). *La grammaire d'aujourd'hui*. Paris: Flammarion.
- Ayari, S. (1996). Diglossia and illiteracy in the Arab world. *Language, Culture and Curriculum, 9*, 243-252.
- Baccouche, T. (2003). La langue arabe: Spécificités et évolution. In M. N. Romdhane, J. E. Gombert & M. Belajouza (Eds.), *L'apprentissage de la lecture: Perspectives comparatives* (pp. 377-386). Rennes: PUR.

- Bamhamed, M. (1996). *Traitement en temps réel des énoncés complexes: Études comparatives interlangues (arabe/français)*. Unpublished Doctorat de Psychologie: Université de Paris V, Paris
- Barrera, S. D., & Maluf, M. R. (2003). Consciencia metalinguística e alfabetização: Um estudo com crianças da primeira série do ensino fundamental. *Psicologia: Reflexão e Crítica*, 16(3), 491-502.
- Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates. (Eds.), *The crosslinguistic study of sentence processing* (pp. 3-73). New York: Cambridge University Press.
- Bates, E., Devescovi, A., & D'Amico, S. (1999). Processing complex sentences: A cross-linguistic study. *Language and Cognitive Processes*, 14(1), 69-123.
- Bates, E., MacWhinney, B., Caselli, C., Devescovi, A., Natale, F., & Venza, V. (1984). A cross-linguistic study of the development of sentence interpretation strategies. *Child Development*, 55, 341-354.
- Bates, E., McNew, S., MacWhinney, B., Devescovi, A., & Smith, S. (1982). Functional constraints on sentence processing: A cross-linguistic study. *Cognition*, 245-299.
- Beauvillain, C., & Grainger, J. (1987). Accessing interlexical homographs: Some limitations of a language-selective access. *Journal of memory and language*, 26, 658-672.
- Bechara, E. (2002). *Moderna gramática portuguesa*. Rio de Janeiro: Lucerna.
- Belajouza, M. (2003). Stratégies d'identification des mots en arabe, compétences phonologiques et morphologiques des lecteurs en difficulté. In M. N. Romdhane, J. E. Gombert & M. Belajouza (Eds.), *L'apprentissage de la lecture : perspectives comparatives* (pp. 194-211). Rennes: Presses Universitaires de Rennes.
- Belajouza, M., & Snoussi, F. (2006). Apprentissage de la lecture en langue arabe et contraintes externes. *Muqabasat intersignes*, 1, 91-113.
- Ben-Dror, I., Bentin, S., & Frost, R. (1995). Semantic, phonologic, and morphologic skills in reading disabled and normal children: Evidence from perception and production of spoken Hebrew. *Reading Research Quarterly*, 30(4), 876-893.
- Bentin, S., & Frost, R. (1995). Morphological factors in visual word identification in Hebrew. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 271-292). Hillsdale, NJ: Erlbaum.
- Bentin, S., & Frost, R. (2001). Linguistic theory and psychological reality: A reply to Boudelaa & Marslen-Wilson. *Cognition*, 81(1), 113-118.
- Bentin, S., Deutsch, A., & Liberman, I. Y. (1990). Syntactic competence and reading ability in children. *Journal of Experimental Child Psychology*, 49(1), 147-172.

- BenZeev, S. (1977). The influence of bilingualism on cognitive strategy and cognitive development. *Child Development*, 48, 1009-1018.
- Berko, J. (1958). The child's learning of English morphology. *Word*, 14, 150-177.
- Berthoud, A. C. (1982). *Activité métalinguistique et acquisition d'une langue seconde; étude des verbes déictiques allemands*. Berne: Lang.
- Bescond, A. (1999). *Sensibilité et conscience morphologique chez les enfants de 5 à 8 ans*. Unpublished mémoire de DEA, Université de Haute-Bretagne Rennes 2.
- Besse, A. S. (2003). *Lecture et connaissances morphologiques des arabophones et lusophones en français langue seconde. Perspective comparative interlangues maternelles*. Unpublished Mémoire de DEA, Université Haute Bretagne Rennes 2, Rennes.
- Besse, A. S., Demont, E., & Gombert, J. E. (2007). Effet des connaissances linguistiques en langue maternelle (arabe vs. Portugais) sur les performances phonologiques et morphologiques en français langue seconde. *Psychologie Française*, 52(1), 89-105.
- Bialystok, E. (1986). Factors in the growth of linguistic awareness. *Child Development*, 57(2), 498.
- Bialystok, E. (1991). Metalinguistic dimensions of bilingual language proficiency. In E. Bialystok (Ed.), *Language processing in bilingual children*. (pp. 113-140). New York, NY, US: Cambridge University Press.
- Bialystok, E. (1993). Metalinguistic awareness: The development of children's representations of language. In C. Pratt & A. F. Garton, (Eds.), *Systems of representation in children: development and use*. (pp. 213-233). Chichester: John Wiley & sons.
- Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development*, 70(3), 636-644.
- Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. New York, NY, US: Cambridge University Press.
- Bialystok, E. (2002a). Acquisition of literacy in bilingual children: A framework for research. *Language Learning*, 52(1), 159-199.
- Bialystok, E. (2002b). On the reliability of robustness: A reply to dekeyser. *Studies in Second Language Acquisition*, 24(3), 481-488.
- Bialystok, E. (2005). Consequences of bilingualism for cognitive development. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches*. (pp. 417-432). New York, NY, US: Oxford University Press.

- Bialystok, E. (2007). Cognitive effects of bilingualism: How linguistic experience leads to cognitive change. *International Journal of Bilingual Education & Bilingualism*, 10(3), 210-223.
- Bialystok, E., Luk, G., & Kwan, E. (2005a). Bilingualism, biliteracy, and learning to read: Interactions among languages and writing systems. *Scientific Studies of Reading*, 9(1), 43-61.
- Bialystok, E., Majumder, S., & Martin, M. M. (2003). Developing phonological awareness: Is there a bilingual advantage? *Applied Psycholinguistics*, 24(1), 27-44.
- Bialystok, E., McBride-Chang, C., & Luk, G. (2005b). Bilingualism, language proficiency, and learning to read in two writing systems. *Journal of Educational Psychology*, 97(4), 580-590.
- Bijeljac-Babic, R., Biardeau, A., & Grainger, J. (1997). Masked orthographic priming in bilingual word recognition. *Memory & Cognition*, 25(4), 447-457.
- Blachère, R., & Gaudefroy-Demombynes, M. (2000). *Grammaire de l'arabe classique*. Paris: Maisonneuve & Larose.
- Bohannon, J. N., Warren-Leubecker, A., & Hepler, N. (1984). Word order awareness and early reading. *Child Development*, 55(4), 1541-1548.
- Boland, J. E. (1997). Resolving syntactic category ambiguities in discourse context: Probabilistic and discourse constraints. *Journal of Memory and Language*, 36(4), 588-615.
- Bouchafa, H., & Magnan, A. (2000). *Le double statut des voyelles longues en arabe: Une difficulté d'acquisition*. Paper presented at the 27ème Congrès International de Psychologie, 23-28 juillet 2000, Stockholm, Suède.
- Boudelaa, S., & Marslen-Wilson, W. D. (2001). Morphological units in the arabic mental lexicon. *Cognition*, 81(1), 65-92.
- Boudelaa, S., & Marslen-Wilson, W. D. (2004). Abstract morphemes and lexical representation: The CV-skeleton in Arabic. *Cognition*, 92(3), 271-303.
- Boudelaa, S., & Marslen-Wilson, W. D. (2005). Discontinuous morphology in time: Incremental masked priming in Arabic. *Language and Cognitive Processes*, 20(1), 207-260.
- Boukadida, N. (en cours). *Connaissances phonologiques et morphologiques chez des enfants en difficulté d'apprentissage de la lecture en arabe*. Unpublished Thèse de Doctorat de Psychologie, Université Rennes 2 Haute Bretagne, Rennes.
- Boukadida, N., & Besse, A. S. (2007). *L'influence des connaissances morphologiques sur l'apprentissage de la lecture en arabe standard et en français langue seconde chez*

- les élèves tunisiens*. Paper presented at the 2ème colloque international de didactique cognitive des langues, 19-21 septembre 2007, Université Toulouse Le Mirail, Toulouse.
- Bronckart, J. P. (1977). *Théories du langage, une introduction critique*. Bruxelles: Mardaga.
- Bruyninckx, M., & Harmegnies, B. (1993). Bilinguismes. *Revue de Phonétique Appliquée*, 108-109, 192-217.
- Brybaert, M., Van Dyck, G., & Van de Poel, M. (1999). Visual word recognition in bilinguals: Evidence from masked phonological priming. *Journal of Experimental Psychology*, 25(1), 137-148.
- Burani, C. (2003). *Derivational morphology, word comprehension and reading aloud in Italian children with and without developmental dyslexia*. Paper presented at the 3rd International Workshop on Morphological Processing. 11-14th June 2003, Aix-en-Provence.
- Campbell, R., & Sais, E. (1995). Accelerated metalinguistic (phonological) awareness in bilingual children. *British Journal of Developmental Psychology*, 13, 61-68.
- Cantel, R. (1999). *Précis de grammaire portugaise*: Paris: Vuibert.
- Capovilla, A. G. S., & Capovilla, F. C. (2000). Efeitos do treino de consciência fonológica em crianças com baixo nível socio-econômico. *Psicologia: Reflexão e Crítica*, 13(1), 7-24.
- Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition*, 28(3), 297-332.
- Caravolas, M., Volin, J., & Hulme, C. (2005). Phoneme awareness is a key component of alphabetic literacy skills in consistent and inconsistent orthographies: Evidence from Czech and English children. *Journal of Experimental Child Psychology*, 92(2), 107-139.
- Cardoso-Martins, C. (1994). Rhyme perception: Global or analytical? *Journal of Experimental Child Psychology*, 57(1), 26-41.
- Cardoso-Martins, C. (1995). Sensitivity to rhymes, syllables, and phonemes in literacy acquisition in portuguese. *Reading Research Quarterly*, 30(4), 808-828.
- Carlisle, J. F. (1988). Knowledge of derivational morphology and spelling ability in fourth, sixth, and eighth graders. *Applied Psycholinguistics*, 9(3), 247-266.
- Carlisle, J. F. (1995). Morphological awareness and early reading achievement. In L. B. Feldman (Ed.), *Morphological aspects of language processing*. (pp. 189-209). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

- Carlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing, 12*(3), 169-190.
- Carlisle, J. F., & Nomanbhoy, D. M. (1993). Phonological and morphological awareness in first graders. *Applied Psycholinguistics, 14*(2), 177-195.
- Carlisle, J. F., Beeman, M., Davis, L. H., & Spharim, G. (1999). Relationship of metalinguistic capabilities and reading achievement for children who are becoming bilingual. *Applied Psycholinguistics, 20*(4), 459-478.
- Cárnio, M. S., & Dos Santos, D. (2005). Evolução da consciência fonológica em alunos de ensino fundamental. *Pró-Fono Revista de Atualização Científica, 17*(2), 195-200.
- Casalis, S., & Louis-Alexandre, M.-F. (2000). Morphological analysis, phonological analysis and learning to read french: A longitudinal study. *Reading and Writing, 12*(3), 303-335.
- Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition, 91*(1), 77-111.
- Cazden, C. B. (1968). The acquisition of noun and verb inflections. *Child Development, 18*, 21-40.
- Chiappe, P., & Siegel, L. S. (1999). Phonological awareness and reading acquisition in English- and Punjabi-speaking Canadian children. *Journal of Educational Psychology, 91*(1), 20-28.
- Chomsky, N. (1957). *Structures syntaxiques*. Paris: Seuil.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge: MIT Press.
- Clark, E. V., & Cohen, S. R. (1984). Productivity and memory for newly formed words. *Journal of Child Language, 11*(3), 611-625.
- Cohen, D. (2007). Langue arabe. In *Encyclopædia universalis*. France.
- Colé, P. (2004). Le traitement des mots morphologiquement complexes au cours de l'acquisition de la lecture: Des données préliminaires. In L. Ferrand, J. Grainger & J. Segui (Eds.), *Psycholinguistique cognitive* (pp. 307-328). Bruxelles: De Boeck Université.
- Colé, P., & Fayol, M. (2000). Reconnaissance de mots écrits et apprentissage de la lecture: Rôle des connaissances morphologiques. In M. Kail & M. Fayol (Eds.), *L'acquisition du langage. Le langage en développement au-delà de trois ans* (Vol. 2, pp. 151-181). Paris: Presses Universitaires de France.
- Colé, P., Beauvillain, C., & Segui, J. (1989). On the representation and processing of prefixed and suffixed derived words: A differential frequency effect. *Journal of Memory and Language, 28*(1), 1-13.

- Colé, P., Magnan, A., & Grainger, J. (1999). Syllable-sized units in visual word recognition: Evidence from skilled and beginning readers of French. *Applied Psycholinguistics*, 20(4), 507-532.
- Colé, P., Marec-Breton, N., Royer, C., & Gombert, J. E. (2003). Morphologie des mots et apprentissage de la lecture. *Rééducation Orthophonique*, 213, 57-73.
- Colé, P., Segui, J., & Taft, M. (1997). Words and morphemes as units for lexical access. *Journal of Memory and Language*, 37(3), 312-330.
- Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing*. New York: Academic Press.
- Comeau, L., Cormier, P., Grandmaison, A. R., & Lacroix, D. (1999). A longitudinal study of phonological processing skills in children learning to read in a second language. *Journal of Educational Psychology*, 91(1), 29-43.
- Content A., & J, Leybaert. (1992.). L'acquisition de la lecture: Influence des méthodes d'apprentissage. In P. Lecoq. (Ed.), *La lecture: Processus, apprentissage, troubles*. Lille: Presses Universitaires de Lille.
- Corbin, D. (1987). *Morphologie dérivationnelle et structure du lexique*. Lille, Tübingen: Niemeyer.
- Corder, S. P. (1967). The significance of learner's errors. *International Review of Applied Linguistics*, 4, 161-170.
- Corder, S. P. (1980). Dialectes idiosyncrasiques et analyse d'erreurs. *Langages* (57), 17-27.
- Cortez Gomes, A. M. (1997). *Structure propositionnelle et ordre des mots en portugais brésilien et en portugais européen*. Unpublished Doctorat de Linguistique, Université Paris 8, Paris.
- Cossu, G., Shankweiler, D., Liberman, I. Y., & Katz, L. (1988). Awareness of phonological segments and reading ability in Italian children. *Applied Psycholinguistics*, 9(1), 1-16.
- Cromdal, J. (1999). Childhood bilingualism and metalinguistic skills: Analysis and control in young Swedish-English bilinguals. *Applied Psycholinguistics*, 20, 1-20.
- Cummins, J. (1978). Métalinguistic development of children in bilingual education programs: Data from Irish and Canadian ukrainian-English programs. In M. Paradis (Ed.), *Aspects of bilingualism* (pp. 127-138). [Columbia: Hornbeam Press](#).
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49, 222-251.
- Cummins, J. (1987). L'éducation bilingue: Théorie et mise en oeuvre. In Centre pour la recherche et innovation dans l'enseignement (Ed.), *L'éducation multiculturelle* (pp. 323-354). Paris: OCDE.

- Cummins, J. (1989). Language and literacy acquisition in bilingual contexts. *Journal of Multilingual and Multicultural Development*, 10(1), 17-31.
- Cummins, J. (1991). Interdependence of first-and second-language proficiency in bilingual children. In E. Bialystok (Ed.), *Language processing in bilingual children*. Cambridge: Cambridge University Press.
- Cummins, J., & Swain, M. (1986). *Bilingualism in education*. London: Longman.
- Cuq, J. P. (1996). *Une introduction à la didactique de la grammaire en français langue étrangère*. Paris: Didier Hatier.
- Da Fontoura, H. A., & Siegel, L. S. (1995). Reading, syntactic, and working memory skills of bilingual portuguese-English Canadian children. *Reading and Writing*, 7(1), 139-153.
- Da Mota, M., Moussatchè, A. H., Rabello de Castro, C., Seidl de Moura, M. L., & D'Angelis, T. (2000). Erros de escrita no contexto: Uma análise na abordagem do processamento da informação. *Psicologia: Reflexão e Crítica*, 13(1), 1-16.
- Daigle, D., & Demont, E. (2006). *Sensibilité aux traitements morphologique et visuo-orthographique en lecture chez des élèves du CP au CMI*. Paper presented at the Colloque International "Approche cognitive de l'apprentissage de la langue écrite", 19 et 20 octobre 2006, Université Rennes 2 - Haute Bretagne, Rennes.
- De Graaff, R. (1997). The esperanto experiment. Effects of explicit instruction on second language acquisition. *Studies in Second Language Acquisition*, 19, 249-276.
- De Groot, A. M. B. (1992). Bilingual lexical representation: A closer look at conceptual representations. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning*. Amsterdam: North-Holland.
- De Queiroga, B. A. M., Lins, M. B., & de Andrade Lima Vasconcelos Pereira, M. (2006). Conhecimento morfossintático e ortografia em crianças do ensino fundamental. *Psicologia: Teoria e Pesquisa*, 22(1), 95-100.
- Defior, S., Martos, F., & Cary, L. (2002). Differences in reading acquisition development in two shallow orthographies: Portuguese and Spanish. *Applied Psycholinguistics*, 23(1), 135-148.
- DeKeyser, R. M. (1995). Learning second language grammar rules. An experiment with a miniature linguistic system. *Studies in Second Language Acquisition*, 17, 379-410.
- DeKeyser, R. M. (1997). Beyond explicit rule learning. Automating second language morphosyntax. *Studies in Second Language Acquisition*, 19, 195-221.
- DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, 22(4), 499-533.

- Deltour, J.-J., & Hupkens, D. (1980). *Test de vocabulaire actif et passif pour enfants de 3 à 5 ans*. Liège: Université de Liège. Editions scientifiques et psychologiques.
- Demont, E. (1994). *Developpement metalinguistique et apprentissage de la lecture*. Unpublished Doctorat de Psychologie, Université de Bourgogne, Dijon.
- Demont, E. (2001). Contribution de l'apprentissage précoce d'une deuxième langue au développement de la conscience linguistique et à l'apprentissage de la lecture. *Journal International de Psychologie*, 36(4), 274-285.
- Demont, E., & Gombert, J. E. (1996). Phonological awareness as a predictor of recoding skills and syntactic awareness as a predictor of comprehension skills. *British Journal of Educational Psychology*, 66(3), 315-322.
- Demont, E., & Gombert, J. E. (2007). Relations entre conscience phonologique et apprentissage de la lecture: Peut-on sortir de la relation circulaire? In E. Demont & M. L. Metz-Lutz. (Eds.), *L'acquisition du langage et ses troubles* (pp. 47-79). Marseille: Solal.
- Demont, E., & Gombert, J.-E. (2004). L'apprentissage de la lecture: évolution des procédures et apprentissage implicite. *Enfance*, 56(3), 245-257.
- Demont, E., Louvet, E., & Nocus, I. (2001). L'enseignement bilingue en Alsace: Quelles influences sur l'apprentissage de la lecture & la conscience linguistique? In E. Regnault (Ed.), *Les enjeux du pluralisme linguistique pour les systèmes d'éducation et de formation*. Paris: L'Harmattan.
- Deutsch, A., Frost, R., & Forster, K. I. (1998). Verbs and nouns are organized and accessed differently in the mental lexicon: Evidence from Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(5), 1238-1255.
- Dichy, J. (2006). «Larcher Pierre, le système verbal de l'arabe classique, publications de l'Université de Provence (collection « didactilangue »), 2003, 191 p.», *Revue des mondes musulmans et de la Méditerranée [En ligne], N°115-116 - La Syrie au quotidien. Cultures et pratiques du changement, décembre 2006*.
- Doignon, N., & Zagar, D. (2006). Les enfants en cours d'apprentissage de la lecture perçoivent-ils la syllabe à l'écrit? *Canadian Journal of Experimental Psychology*, 60(4), 258-274.
- Drews, E., & Zwitserlood, P. (1995). Morphological and orthographic similarity in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 21(5), 1098-1116.
- Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly*, 38(1), 78-103.

- Dubois, J., Giacomo, M., Guespin, L., Marcellesi, C., Marcellesi, J. B., & Mevel, J. P. (1999). *Dictionnaire de linguistique et des sciences du langage*. Paris: Larousse.
- Duncan, L. G., Colé, P., Seymour, P. H. K., & Magnan, A. (2006). Differing sequences of metaphonological development in French and English. *Journal of Child Language*, 33(2), 369-399.
- Duncan, L. G., Seymour, P. H. K., & Hill, S. (1997). How important are rhyme and analogy in beginning reading? *Cognition*, 63(2), 171-208.
- Duncan, L. G., Seymour, P. H. K., & Hill, S. (2000). A small-to-large unit progression in metaphonological awareness and reading? *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, 53(4), 1081.
- Durgunoğlu, A. Y. (1998). Acquiring literacy in English and Spanish in the united states. In A. Y. Durgunoğlu & L. Verhoeven (Eds.), *Literacy development in a multilingual context: Cross-cultural perspective* (pp. 135-145). Hillsdale, NJ, England: Lawrence Erlbaum associates.
- Durgunoğlu, A. Y. (2002). Cross-linguistic transfer in literacy development and implications for language learners. *Annals of Dyslexia*, 52, 189-204.
- Durgunoğlu, A. Y., & Oney, B. (1999). A cross-linguistic comparison of phonological awareness and word recognition. *Reading and Writing*, 11(4), 281-299.
- Durgunoğlu, A. Y., & Verhoeven, L. (1998). *Literacy development in a multilingual context: Cross-cultural perspectives* (Vol. XVIII). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Durgunoğlu, A. Y., Mir, M., & Arino-Martí, S. (2002). The relationship between bilingual children's reading and writing in their two languages. In S; Ransdell & M. Barbier (Eds), *Psycholinguistic approaches to understanding second language writing* (pp. 81-100). Dordrecht: Kluwer.
- Durgunoğlu, A. Y., Nagy, W. E., & Hancin-Bhatt, B. J. (1993). Cross-language transfer of phonological awareness. *Journal of Educational Psychology*, 85(3), 453-465.
- Elbeheri, G., & Everatt, J. (2007). Literacy ability and phonological processing skills amongst dyslexic and non-dyslexic speakers of Arabic. *Reading and Writing*, 20, 273-294.
- Ellis, N. (1993). Rules and instances in foreign language learning: Interactions of explicit and implicit knowledge. *European Journal of Cognitive Psychology*, 5(3), 289-318.
- Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24(2), 143-188.

- Ervin-Tripp, S. M. (1974). Is second-language learning like the first? *TESOL Quarterly*, 8, 111-127.
- Eviatar, Z., & Ibrahim, R. (2000). Bilingual is as bilingual does: Metalinguistic abilities of arabic-speaking children. *Applied Psycholinguistics*, 21(4), 451-471.
- Ez-Zaher, A. (2004). *Representations metaphonologiques et apprentissage de la lecture en arabe*. Unpublished Doctorat de Psychologie, Université Toulouse-Le Mirail, Toulouse.
- Fathman, A. (1975). The relationship between age and second language productive ability. *Language Learning*, 25, 245-266.
- Fayol, M. (1992). Comprendre ce qu'on lit: De l'automatisme au contrôle. In M. Fayol, J-E. Gombert, P. Lecocq, L. Sprenger-Charolles & D. Zagar (Eds.), *Psychologie cognitive de la lecture* (pp. 73-105). Paris: Presses Universitaires de France.
- Feitelson, D., Goldstein, Z., Iraqi, J., & Share, D. L. (1993). Effects of listening to story reading on aspects of literacy acquisition in a diglossic situation. *Reading Research Quarterly*, 28(1), 71-79.
- Feldman, L. B., & Fowler, C. A. (1987). The inflected noun system in serbo-croatian: Lexical representation of morphological structure. *Memory and Cognition*, 15, 1-12.
- Feldman, L. B., & Soltano, E. G. (1999). Morphological priming: The role of prime duration, semantic transparency, and affix position. *Brain and Language*, 68(1), 33-39.
- Fender, M. (2003). English word recognition and word integration skills of native Arabic- and Japanese-speaking learners of English as a second language. *Applied Psycholinguistics*, 24(2), 289-315.
- Ferguson, C. (1959). Diglossia. *Word*, 15, 325-340.
- Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, 25(3), 348-368.
- Ferreira, F., & Henderson, J. M. (1990). Use of verb information in syntactic parsing: Evidence from eye movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(4), 555-568.
- Fowler, A. E., Feldman, L. B., Andjelkovic, D., & Oney, B. (2003). Morphological and phonological analysis by beginning readers: Evidence from Serbian and Turkish. In E. M. H. Assink & D. Sandra (Eds.), *Reading complex words. Cross language studies* (pp. 53-80). New-York: Kluwer academic.
- Frankel, D. G., & Arbel, T. (1981). Developmental changes in assigning agent relations in Hebrew: The interaction between world order and structural cues. *Journal of experimental child psychology*, 32, 102-114.

- Frankel, D. G., Amir, M., Frenkel, E., & Arbel, T. (1980). A developmental study of the role of word order in comprehending Hebrew. *Journal of Experimental Child Psychology*, 29(1), 23-35.
- Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance. The psychology of reading* (Vol. 12, pp. 559-585). Hillsdale N.J.: Lawrence Erlbaum.
- Frazier, L. (1989). Theories of sentence processing. In J. L. Garfield (Ed.), *Modularity in knowledge representation and natural-language understanding* (pp. 291-308). Cambridge: MIT Press.
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14(2), 178-210.
- Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. E. Patterson, J. C. Marshall & M. Coltheart (Eds.), *Surface dyslexia: Cognitive and neuropsychological studies of phonological reading* (pp. 301-330). London: Erlbaum.
- Frith, U. (1986). A developmental framework for developmental dyslexia. *Annals of Dyslexia*, 36, 69-81.
- Frith, U., Wimmer, H., & Landerl, K. (1998). Differences in phonological recoding in German- and English-speaking children. *Scientific Studies of Reading*, 2(1), 31-54.
- Frost, R., & Grainger, J. (2000). Cross-linguistic perspectives on morphological processing: An introduction. *Language and Cognitive Processes*, 15(4/5), 321-328.
- Frost, R., Deutsch, A., & Forster, K. I. (2000a). Decomposing morphologically complex words in a nonlinear morphology. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(3), 751-765.
- Frost, R., Deutsch, A., Gilboa, O., Tannenbaum, M., & Marslen-Wilson, W. (2000b). Morphological priming: Dissociation of phonological, semantic and morphological factors. *Memory & Cognition*, 28(8), 1277-1288.
- Frost, R., Forster, K. I., & Deutsch, A. (1997). What can we learn from the morphology of Hebrew? A masked-priming investigation of morphological representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(4), 829-856.
- Frost, R., Grainger, J., & Rastle, K. (2005a). Current issues in morphological processing: An introduction. *Language and Cognitive Processes*, 20(1), 1-5.

- Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance*, 13(1), 104-115.
- Frost, R., Kugler, T., Deutsch, A., & Forster, K. I. (2005b). Orthographic structure versus morphological structure: Principles of lexical organization in a given language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(6), 1293-1326.
- Gaonac'h, D. (2000). La lecture en langue étrangère: Un tour d'horizon d'une problématique de psychologie cognitive. *Acquisition et interaction en langue étrangère*, 13, 1-9.
- Gaonac'h, D. (2006). *L'apprentissage précoce d'une langue étrangère. Le point de vue de la psycholinguistique*. Paris: Hachette.
- Gardes-Tamine, J. (1998). *La grammaire. Phonologie, morphologie, lexicologie*. Paris: Armand Colin.
- Gass, S. M. (1987). The resolution of conflicts among competing systems: A bidirectional perspective. *Applied Psycholinguistics*, 8, 329-350.
- Gaux, C., & Gombert, J. E. (1999b). Implicit and explicit syntactic knowledge and reading in pre-adolescents. *British Journal of Developmental Psychology*, 17(2), 169-188.
- Gaux, C., & Gombert, J.-E. (1999a). La conscience syntaxique chez les preadolescents: Question de méthodes. *L'année Psychologique*, 99(1), 45-74.
- Geva, E., & Siegel, L. S. (2000). Orthographic and cognitive factors in the concurrent development of basic reading skills in two languages. *Reading and Writing*, 12(1), 1-30.
- Geva, E., & Wade-Woolley, L. (1998). Component process in becoming English-Hebrew biliterate. In Y. Durgunoğlu & L. Verhoeven (Eds.), *Literacy development in a multilingual context: Cross-cultural perspective* (pp. 85-110). Hillsdale, NJ, England: Lawrence Erlbaum associates.
- Geva, E., & Zadeh Yaghoub, Z. (2006). Reading efficiency in native english-speaking and english-as-a-second-language children: The role of oral proficiency and underlying cognitive-linguistic processes. *Scientific Studies of Reading*, 10(1), 31-57.
- Geva, E., Wade-Woolley, L., & Shany, M. (1997). Development of reading efficiency in first and second language. *Scientific Studies of Reading*, 1(2), 119-144.
- Geva, E., Wade-Woolley, L., Durgunoğlu, A. Y., & Verhoeven, L. (1998). Component processes in becoming English-Hebrew biliterate. In C (pp. 85-110): Lawrence Erlbaum Associates Publishers.

- Gholamain, M., & Geva, E. (1999). Orthographic and cognitive factors in the concurrent development of basic reading skills in English and Persian. *Language Learning*, 49(2), 183.
- Giraud, H., & Grainger, J. (2000). Effects of prime word frequency and cumulative root frequency in masked morphological priming. *Language and Cognitive Processes*, 15(4), 421-444.
- Giraud, H., & Grainger, J. (2001). Priming complex words: Evidence for supralexical representation of morphology. *Psychonomic Bulletin & Review*, 8(1), 127-131.
- Gleitman, L. R., & Gleitman, H. (1992). A picture is worth a thousand words, but that's the problem: The role of syntax in vocabulary acquisition. *Current Directions in Psychological Science*, 1(1), 31-35.
- Golder, C., & Gaonac'h, D. (1998). *Lire et comprendre. Psychologie de la lecture*. Paris: Hachette éducation.
- Gombert, J. E. (1990). *Le développement métalinguistique*. Paris: Presses Universitaires de France.
- Gombert, J. E. (1992). Activité de lecture et activités associées. In M. Fayol, J-E. Gombert, P. Lecocq, L. Sprenger-Charolles & D. Zagar (Eds.), *Psychologie cognitive de la lecture* (pp. 107-140). Paris: Presses Universitaires de France.
- Gombert, J. E. (2002a). L'apprentissage de la lecture: Processus et stratégie. In A. Florin & J. Morais (Eds.), *La maîtrise du langage* (pp. 79-94). Rennes: Presses Universitaires de Rennes.
- Gombert, J. E. (2002b). Préface: La modélisation cognitive de la lecture et ses implications pédagogiques. In J. Ecalle & A. Magnan (Eds.), *L'apprentissage de la lecture: Fonctionnement et développement cognitifs* (pp. 7-10). Paris: Armand Colin.
- Gombert, J. E. (2003a). Implicit and explicit learning to read: Implication. As for subtypes of dyslexia. *Current Psychology Letters*, 10(1).
- Gombert, J. E. (2003b). L'apprentissage des codes grapho-phonologique et grapho-sémantique en lecture. In M. N. Romdhane, J. E. Gombert & M. Belajouza (Eds.), *L'apprentissage de la lecture: Perspectives comparatives* (pp. 19-34). Rennes: Presses Universitaires de Rennes.
- Gombert, J. E., & Colé, P. (2000). Activités métalinguistiques, lecture et illettrisme. In M. Fayol & M. Kail (Ed.), *L'acquisition du langage. Le langage en développement au delà de 3 ans*. (Vol. 2, pp. 151-181). Paris: Presses Universitaires de France.

- Gombert, J. E., Bryant, P., & Warrick, N. (1997). Les analogies dans l'apprentissage de la lecture et de l'orthographe. In L. Rieben, M. Fayol & C.A. Perfetti (Eds.), *Des orthographes et leur acquisition* (pp. 319-333). Lausanne: Delachaux & Niestlé.
- Gombert, J. E., Gaux, C., & Demont, E. (1994). Capacités métalinguistiques et lecture. Quels liens? *Repères*, 9, 61-72.
- Gombert, J. E., Martinot, C., & Nocus, I. (1996). Conscience linguistique et traitement de l'information écrite. *TRANEL (Travaux neuchâtelois de linguistique)*, 25, 135-153.
- Gonnerman, L. M., Seidenberg, M. S., & Andersen, E. S. (2007). Graded semantic and phonological similarity effects in priming: Evidence for a distributed connectionist approach to morphology. *Journal of Experimental Psychology: General*, 136(2), 323-345.
- Gonzalez, J. E. J., & Ortiz Gonzalez, M. R. (2000). Metalinguistic awareness and reading acquisition in the Spanish language. *The Spanish Journal of Psychology*, 3(1), 37-46.
- Gordon, P. (1989). Levels of affixation in the acquisition of English morphology. *Journal of Memory and Language*, 28, 519-530.
- Goswami, U. (1993). Toward an interactive analogy model of reading development: Decoding vowel graphemes in beginning reading. *Journal of Experimental Child Psychology*, 56(3), 443-475.
- Goswami, U. (2000). Phonological representations, reading development and dyslexia: Toward a cross-linguistic theoretical framework. *Dyslexia*, 6, 133-151.
- Goswami, U. (2002). In the beginning was the rhyme? A reflection on Hulme, Hatcher, Nation, Brown, Adams, and Stuart (2002). *Journal of Experimental Child Psychology*, 82(1), 47-57.
- Goswami, U., Gombert, J. E., & de Barrera, L. F. (1998). Children's orthographic representations and linguistic transparency: Nonsense word reading in English, French, and Spanish. *Applied Psycholinguistics*, 19(1), 19-52.
- Goswami, U., Porpodas, C., & Wheelwright, S. (1997). Children's orthographic representations in English and Greek. *European Journal of Psychology of Education*, 12(3), 273-292.
- Goswami, U., Ziegler, J. C., & Richardson, U. (2005). The effects of spelling consistency on phonological awareness: A comparison of English and German. *Journal of Experimental Child Psychology*, 92(4), 345-365.
- Goswami, U., Ziegler, J. C., Dalton, L., & Schneider, W. (2001). Pseudohomophone effects and phonological recoding procedures in reading development in English and German. *Journal of Memory and Language*, 45(4), 648-664.

- Goswami, U., Ziegler, J. C., Dalton, L., & Schneider, W. (2003). Nonword reading across orthographies: How flexible is the choice of reading units? *Applied Psycholinguistics*, 24(2), 235-247.
- Gottardo, A., Stanovich, K. E., & Siegel, L. S. (1996). The relationships between phonological sensitivity, syntactic processing, and verbal working memory in the reading performance of third-grade children. *Journal of Experimental Child Psychology*, 63(3), 563-582.
- Gottardo, A., Yan, B., Siegel, L. S., & Wade-Woolley, L. (2001). Factors related to English reading performance in children with Chinese as a first language: More evidence of cross-language transfer of phonological processing. *Journal of Educational Psychology*, 93(3), 530-542.
- Grainger, J. (1987). L'accès au lexique bilingue: Vers une nouvelle orientation de recherche. *L'Année Psychologique*, 87(553-566).
- Grainger, J. (1991). L'identification des mots chez le sujet bilingue. In R. Kolinsky, J. Morais & J. Segui (Eds.), *La reconnaissance des mots dans les différentes modalités sensorielles. Etudes de psycholinguistique cognitive*. Paris: Presses Universitaires de France.
- Grainger, J. (2002). Structure et fonctionnement du lexique bilingue. In A. Florin & J. Morais (Eds.), *La maîtrise du langage* (pp. 41-59); Rennes: Presses Universitaires de Rennes.
- Grainger, J., & Beauvillain, C. (1987). Language blocking and lexical access in bilinguals. *Quarterly Journal of Experimental Psychology*, 39A, 295-319.
- Grainger, J., Colé, P., & Segui, J. (1991). Masked morphological priming in visual word recognition. *Journal of Memory and Language*, 30(3), 370-384.
- Grainger, J., Dichy, J., El-Halfaoui, M., & Bamhamed, M. (2003). Approche expérimentale de la reconnaissance du mot écrit en arabe. In J. P. Jaffré (Ed.), *Dynamique de l'écriture: Approches pluridisciplinaires, revue faits de langue n° 22* (pp. 77-86). Paris: Ophrys.
- Guimarães, S. R. K. (2002). Dificuldades no desenvolvimento da lectoescrita: O papel das habilidades metalinguísticas. *Psicologia: Teoria e Pesquisa*, 18(3), 247-259.
- Hagège, C. (1996). *L'enfant aux deux langues*. Paris: Editions Odile Jacob.
- Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence: A test of the critical-period hypothesis for second-language acquisition. *Psychological Science*, 14(1), 31-38.

- Harley, B., & Hart, D. (1997). Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition*, 19, 379-400.
- Harrington, M. (1987). Processing transfert: Language-specific processing strategies as a source of interlanguage variation. *Applied Psycholinguistics*, 8, 351-377.
- Hulme, C., Hatcher, P. J., Nation, K., Brown, A., Adams, J., & Stuart, G. (2002). Phoneme awareness is a better predictor of early reading skill than onset-rime awareness. *Journal of Experimental Child Psychology*, 82(1), 2-28.
- Hulme, C., Muter, V., & Snowling, M. (1998). Segmentation does predict early progress in learning to read better than rhyme: A reply to Bryant. *Journal of Experimental Child Psychology*, 71(1), 39-44.
- Ianco-Worall, A. (1972). Bilingualism and cognitive development. *Child Development*, 43, 1390-1400.
- Ibrahim, R. (2006). Morpho-phonemic similarity within and between languages: A factor to be considered in processing Arabic and Hebrew, *Reading and Writing*, 19, 563-586.
- Ibrahim, R., Eviatar, Z., & Aharon-Peretz, J. (à paraître). Metalinguistic awareness and reading performance: A cross language comparison. *Journal of Psycholinguistic Research*.
- Jared, D., & Kroll, J. F. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words? *Journal of Memory and Language*, 44, 2-31.
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21(1), 60.
- Johnson, J. S., & Newport, E. L. (1991). Critical period effects on universal properties of language: The status of subadjacency in the acquisition of a second language. *Cognition*, 39(3), 215-258.
- Kahn-Horwitz, J., Shimron, J., & Sparks, R. L. (2005). Predicting foreign language reading achievement in elementary school students. *Reading and Writing*, 18(6), 527-558.
- Kahn-Horwitz, J., Shimron, J., & Sparks, R. L. (2006). Weak and strong novice readers of english as a foreign language: Effects of firsts language and socioeconomic status. *Annals of Dyslexia*, 56(1), 161-185.
- Kail, M. (1983). L'acquisition du langage repensé: Les recherches interlangues. *L'année Psychologique*, 83, 225-258.

- Kail, M. (1986). Validité et coût des indices linguistiques dans la compréhension des phrases. Recherches interlangues sur l'acquisition. *Bulletin de psychologie*, XXXIX(375), 387-396.
- Kail, M. (1989). Cue validity, cue cost and processing types in sentence comprehension in French and Spanish. In B. MacWhinney & E. Bates. (Eds.), *The crosslinguistic study of sentence processing* (pp. 77-117). New York: Cambridge University press.
- Kail, M. (1991). Acquisition des langues premières et secondes dans une perspective interlangues. *Revue française de pédagogie*, 96, 57-78.
- Kail, M. (1999). Linguistic variations and cognitive constraints in the processing and the acquisition of language. In C. Fuchs & S. Robert (Eds.), *Language diversity and cognitive representations* (pp. 179-195). Amsterdam, Netherlands: Benjamins.
- Kail, M. (2000). Acquisition syntaxique et diversité linguistique. In M. Kail & M. Fayol (Eds.), *L'acquisition du langage. Le langage en développement au-delà de 3 ans* (pp. 9-44). Paris: Presses Universitaires de France.
- Kail, M., & Charvillat, A. (1988). Local and topological processing in sentence comprehension by French and Spanish children. *Journal of child language*, 15, 637-662.
- Kaplan, R. M., & Bresnan, J. (1982). Lexical-functional grammar: A formal system of grammatical representation. In J. Bresnan (Ed.), *The mental representation of grammatical relations*. Cambridge: MIT Press.
- Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning*. (pp. 67-84). Oxford, England: North-Holland.
- Khomsî, A. (1990). *Epreuve collective d'évaluation de la compétence en lecture: E20*. Paris: Edition du Centre de Psychologie Appliquée.
- Khomsî, A. (1994). *Epreuve Collective d'Identification de Mots en présentation écrite (ECIM-E)*: Travaux de Psycholinguistique, 5, Université de Nantes, Département de Psychologie.
- Khomsî, A. (1994). *Epreuve collective d'identification de mots en présentation écrite (ECIM-E)*: Travaux de Psycholinguistique, 5, Université de Nantes, Département de Psychologie.
- Khomsî, A. (1999). *Lecture de Mots et Compréhension – Révisée. Epreuve d'évaluation de la compétence en lecture*. Paris: Les Editions du Centre de Psychologie Appliquée.

- Kilborn, K., & Ito, T. (1989). Sentence processing strategies in adult bilinguals. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study in sentence processing* (pp. 257-291). New York: Cambridge University Press.
- Kintsch, W. (1970). Recognition memory in bilingual subjects. *Journal of verbal learning and verbal behavior*, 9, 405-409.
- Kirsner, K., Smith, M. C., Lockhart, R. S., King, M. L., & Jain, M. (1984). The bilingual lexicon: Language-specific units in an integrated network. *Journal of Verbal Learning and Verbal Behavior*, 23(4), 519-539.
- Koda, K. (1988). Cognitive process in second language reading: Transfert of L1 reading skills and strategies. *Second language research*, 4, 133-156.
- Koda, K. (1989). Effects of L1 orthographic representation on L2 phonological coding strategies. *Journal of Psycholinguistic Research*, 18 (2), 201-222.
- Koda, K. (1993). Transferred L1 strategies and L2 syntactic structure in L2 sentence comprehension. *Modern Language Journal*, 77 (4), 490-500.
- Koda, K. (1996). L2 word recognition research: A critical review. *Modern Language Journal*, 80 (4), 450-460.
- Koda, K. (1999). Development of L2 intraword orthographic sensitivity and decoding skills. *Modern Language Journal*, 83 (1), 51-64.
- Koda, K. (2000). Cross-linguistic variations in L2 morphological awareness. *Applied Psycholinguistics*, 21 (3), 297-320.
- Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. Cambridge; New York: Cambridge University Press.
- Koda, K. (2007). Reading and language learning: Crosslinguistic constraints on second language reading development. *Language Learning*, 57(1), 1-44.
- Kolers, P. A. (1963). Interlingual word associations. *Journal of verbal learning and verbal behavior*, 2, 291-300.
- Kolers, P. A. (1966). Interlingual facilitation of short-term memory. *Journal of verbal learning and verbal behavior*, 5, 314-319.
- Kouloughli, D. E. (1994). *Grammaire de l'arabe d'aujourd'hui*. Paris: Pocket. Collection Langue pour tous.
- Krashen, S. (1985). *The input hypothesis: Issues and implications*. London: Longman.
- Ku, Y.-M., & Anderson, R. C. (2003). Development of morphological awareness in Chinese and English. *Reading and Writing*, 16(5), 399-422.
- Kuo, L.-j., & Anderson, R. C. (2006). Morphological awareness and learning to read: A cross-language perspective. *Educational Psychologist*, 41(3), 161-180.

- Lado, R. (1964). *Language teaching, a scientific approach*. New-York: McGraw-Hill, Inc.
- Landerl, K. (2000). Influences of orthographic consistency and reading instruction on the development of nonword reading skills. *European Journal of psychology of Education, 15*(3), 239-257.
- Larcher, P. (2003). *Le système verbal de l'arabe classique*. Marseille: Publications de l'Université de Provence.
- Laxon, V., Rickard, M., & Coltheart, V. (1992). Children read affixed words and non-words. *British Journal of Psychology, 83*(3), 407-423.
- Layton, A., Robinson, J., & Lawson, M. (1998). The relationship between syntactic awareness and reading performance. *Journal of Research in Reading, 21*(1), 5-23.
- Lecocq, P. (1991). *Apprentissage de la lecture et dyslexie*. Liège: Mardaga.
- Lecocq, P., Casalis, S., Leuwens, C., & Watteau, N. (1996). *Apprentissage de la lecture et compréhension d'énoncés*. Villeneuve d'Ascq: Presses Universitaires du Septentrion.
- Lefrançois, F. (2001). De la langue maternelle à la langue seconde en lecture et en écriture. *La Revue de l'AQEFLS, 23* (1), 45-57.
- Lenneberg, E. H. (1967). *Biological foundations of language*. New York: John Willey & sons.
- Leow, R. P. (2000). A study of the role of awareness in foreign language behavior. Aware versus unaware learners. *Studies in Second Language Acquisition, 22*, 557-583.
- Lesaux, N. K., & Siegel, L. S. (2003). The development of reading in children who speak English as a second language. *Developmental Psychology, 39*(6), 1005-1019.
- Lété, B., Sprenger-Charolles, L., & Colé, P. (2004). MANULEX ; A Web-accessible lexical database from French primary school reading books. *Behavior Research, Methods, Instruments and Computers*.
- Levin, I., Ravid, D., & Rapaport, S. (2001). Morphology and spelling among Hebrew-speaking children: From kindergarten to first grade. *Journal of Child Language, 28*(3), 741-772.
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*(4), 492-527.
- Lucchini, S. (2002). *L'apprentissage de la lecture en langue seconde. La formation d'une langue de référence chez les enfants d'origine immigrée*. Cortil-Wodon: Editions Modulaires Européennes.
- Lukatela, K., Carello, C., Shankweiler, D., & Liberman, I. Y. (1995). Phonological awareness in illiterates: Observations from Serbo-Croatian. *Applied Psycholinguistics, 16*(4), 463-487.

- Maaouia Belkadhi, A. (2003). La compréhension de l'énoncé et la structure syntaxique de la phrase arabe. In M. N. Romdhane, J. E. Gombert & M. Belajouza (Eds.), *L'apprentissage de la lecture: Perspectives comparatives*. (pp. 143-160). Rennes: Presses Universitaires de Rennes.
- Mackey, W. f. (1976). *Bilinguisme et contact des langues*. Paris: Editions Klincksieck.
- MacWhinney, B. (1989). Competition and connectionism. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study in sentence processing* (pp. 422-457). New York: Cambridge University Press.
- MacWhinney, B. (1997). Implicit and explicit processes. *Studies in Second Language Acquisition, 19*, 277-281.
- MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning & Verbal Behavior, 23*(2), 127-150.
- Mägiste, E. (1992). Second language learning in elementary and high school students. *European Journal of Cognitive Psychology, 4*(4), 355-365.
- Magnan, A., & Colé, P. (1992). Les modèles d'apprentissage de la lecture: Évolution et perspectives. *Rééducation Orthophonique, 30*, 43-61.
- Mahony, D., Singson, M., & Mann, V. (2000). Reading ability and sensitivity to morphological relations. *Reading and Writing, 12*(3), 191-218.
- Maluf, M. R., & Barrera, S. D. (1997). Consciência fonológica e linguagem escrita em pré-escolares. *Psicologia: Reflexão e Crítica, 10*(1), 125-145.
- Maluf, M. R., Gombert, J. E., & Sousa, E. D. O. (2004). *Metasyntactic and reading abilities: A study with Brazilian Portuguese speaking children*. Paper presented at the XXVIII International Congress of Psychology, Beijing, China.
- Mann, V. A. (1984). Longitudinal prediction and prevention of early reading difficulty. *Annals of Dyslexia, 34*, 117-136.
- Mann, V., & Wimmer, H. (2002). Phoneme awareness and pathways into literacy: A comparison of German and American children. *Reading and Writing, 15*(7), 653-682.
- Marec-Breton, N. (2003). *Les traitements morphologiques dans l'apprentissage de la lecture*. Unpublished Doctorat de Psychologie, Université Haute-Bretagne Rennes 2, Rennes.
- Marec-Breton, N., & Gombert, J. E. (soumis). How young children read complex word?
- Marec-Breton, N., Gombert, J.-E., & Colé, P. (2005). Traitements morphologiques lors de la reconnaissance des mots écrits chez des apprentis lecteurs. *L'année Psychologique, 105*(1), 9-45.

- Marslen-Wilson, W., Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review*, 101(1), 3-33.
- Martinot, C., & Gombert, J. E. (1996). Le développement et le contrôle des connaissances phonologiques à l'âge préscolaire. *Revue de Neuropsychologie*, 6(2), 251-270.
- Matthey, M. (1996). *Apprentissage d'une langue et interaction verbale*. Bern; Berlin (etc.): Peter Lang.
- McBride-Chang, C., Wagner, R. K., Muse, A., Chow, B. W.-Y., & Shu, H. (2005). The role of morphological awareness in children's vocabulary acquisition in English. *Applied Psycholinguistics*, 26(3), 415-435.
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive-activation model of context effects in letter perception, part 1: An account of basic findings. *Psychological Review*, 88, 375-405.
- McDonald, J. L. (1987). Sentence interpretation in bilingual speakers of English and Dutch. *Applied Psycholinguistics*, 8, 379-413.
- McLaughlin, B. (1987). *Theories of second-language learning*. London New York Melbourne Auckland: Edward Arnold.
- McLaughlin, B. (1991). *Theories of second-language learning*. London: Edward Arnold.
- Mel'cuk, I. (1997). *Cours de morphologie générale (théorique et descriptive)* (CNRS Editions ed. Vol. IV). Montréal: Presses de l'Université de Montréal.
- Miao, X. C. (1981). Word order and semantic strategies in Chinese sentence comprehension. *International Journal of Psycholinguistics*, 23, 109-122.
- Michas, I. C., & Berry, D. C. (1994). Implicit and explicit processes in a second-language learning task. *European Journal of Cognitive Psychology*, 6(4), 357-381.
- Milnitsky Stein, L. (1994). *Teste de desempenho escolar*. São Paulo: Casa do Psicologo Livraria e Editora Ltda.
- Mitchell, D. C., & Holmes, V. M. (1985). The role of specific information about the verb in parsing sentences with local structural ambiguity. *Journal of Memory and Language*, 24(5), 542-559.
- Mitchell, D. C., Cuetos, F., & Zagar, D. (1990). Reading in different languages: Is there a universal mechanism for parsing sentence? In D. A. Balota, G. B. Flores d'Arcais & K. Rayner (Eds.), *Comprehension processes in reading* (Vol. 1, pp. 285-331). Hillsdale N.J.: Lawrence Erlbaum Associates.
- Morais, J. (1994). *L'art de lire*: Editions Odile Jacob.

- Morais, J., Alegria, J., & Content, A. (1987). The relationships between segmental analysis and alphabetic literacy: An interactive view. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, 7(5), 415-438.
- Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, 7(4), 323-331.
- Morais-Barbosa, J. (2007). Portugal (langue). In *Encyclopædia universalis*.
- Mota, M. (2006). *The effects of schooling on metalinguistic awareness*. Paper presented at the 26th international Congress of Applied Psychology, Athen, Greece.
- Mota, M., Aníbal, L., & Lima, S. (Communication Personnelle). A morfologia derivacional contribui para a leitura e escrita no português?
- Muljani, D., Koda, K., & Moates, D. R. (1998). The development of word recognition in a second language. *Applied Psycholinguistics*, 19(1), 99-111.
- Murrell, G. A., & Morton, J. (1974). Word recognition and morphemic structure. *Journal of Experimental Psychology*, 102(6), 963-968.
- Muter, V., Hulme, C., Snowling, M., & Taylor, S. (1997). Segmentation, not rhyming, predicts early progress in learning to read. *Journal of Experimental Child Psychology*, 65(3), 370-396.
- Nagy, W. E., Diakidoy, I.-A. N., & Anderson, R. C. (1993). The acquisition of morphology: Learning the contribution of suffixes to the meanings of derivatives. *Journal of Reading Behavior*, 25(2), 155-170.
- Nagy, W., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly*, 19(3), 304-330.
- Najem, A., & Besse, A. S. (2006). *L'impact des connaissances acquises à l'oral sur les traitements morphologiques implicites et explicites de l'arabe écrit chez les bons et faibles lecteurs dans un contexte de diglossie*. Paper presented at the Colloque International "Approche cognitive de l'apprentissage de la langue écrite", 19 et 20 octobre 2006, Université Rennes 2 - Haute Bretagne, Rennes.
- Nemser, W. (1971). Approximative systems of foreign language learners. *IRAL*, 9, 115-123.
- Neyreneuf, M., & Al-Hakkak, G. (1996). *Grammaire active de l'arabe littéral*. Paris: Livre de poche.
- Niswander, E., Pollatsek, A., & Rayner, K. (2000). The processing of derived and inflected suffixed words during reading. *Language and Cognitive Processes*, 15(4), 389-420.
- Nocus, I. (1997). *Conscience metasyntaxique et apprentissage de la lecture*. Unpublished Doctorat de Psychologie, Université de Bourgogne, Dijon.

- Nocus, I., & Gombert, J. E. (1997). Conscience morpho-syntaxique et apprentissage de la lecture. *Revue de psychologie de l'éducation*, 2, 71-102.
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50(3), 417-528.
- Odlin, T. (1989). *Language transfer. Cross-linguistic influence in language learning*. Cambridge: Cambridge University Press.
- Oney, B., & Goldman, S. R. (1984). Decoding and comprehension skills in Turkish and English: Effects of the regularity of grapheme-phoneme correspondences. *Journal of Educational Psychology*, 76(4), 557-568.
- Oyama, S. (1979). The concept of the sensitive period in developmental studies. *Merrill-Palmer Quarterly*, 25(2), 83-103.
- Pacton, S., & Fayol, M. (2000). The impact of phonological cues on children's judgements of nonwords: The case of double letters. *Current Psychology Letters: Behaviour, Brain & Cognition*, 1, 39-54.
- Pacton, S., Fayol, M., & Perruchet, P. (2005). Children's implicit learning of graphotactic and morphological regularities. *Child Development*, 76(2), 324-339.
- Patel, T. K., Snowling, M. J., & De Jong, P. F. (2004). A cross-linguistic comparison of children learning to read in English and Dutch. *Journal of Educational Psychology*, 96(4), 785-797.
- Paula, G. R., Mota, H. B., & Keske-Soares, M. (2005). A terapia em consciência fonológica no processo de alfabetização. *Pró-Fono Revista de Atualização Científica*, 17(2), 175-184.
- Pearl, E., & Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychological Monographs*, 76, 1-23.
- Perregaux, C. (1994). *Les enfants à deux voix. Des effets du bilinguisme sur l'apprentissage de la lecture*. Paris: Peter Lang.
- Perregaux, C. (1995). Reconnaître les atouts linguistiques des enfants bilingues pré-lecteurs. *TRANEL (Travaux neuchâtelois de linguistique)*, 19, 125-139.
- Pinhero, A. M. V. (1987). *Contagem de frequência de ocorrência de palavras. Expostas a crianças na faixa pré-escolar e séries iniciais de 1º grau*. São Paulo, Associação Brasileira de Dislexia.
- Pinheiro, A. M. V. (1995). Reading and spelling development in Brazilian Portuguese. *Reading and Writing*, 7(1), 111-138.

- Plaut, D. C., & Gonnerman, L. M. (2000). Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing? *Language and Cognitive Processes, 15*(4), 445-485.
- Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103*(1), 56-115.
- Pléh, C. (1989). The development of sentence interpretation in Hungarian. In B. MacWhinney & E. Bates. (Eds.), *The crosslinguistic study of sentence processing* (pp. 158-184). New York: Cambridge University Press.
- Potter, M. C., So, K. W., Von Eckart, B., & Feldman, M. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior, 23*, 23-38.
- Pratt, C., Tunmer, W. E., & Bowey, J. A. (1984). Children's capacity to correct grammatical violations in sentences. *Journal of Child Language, 11*(1), 129-141.
- Proctor, C. P., Carlo, M., August, D., & Snow, C. (2005). Native spanish-speaking children reading in English: Toward a model of comprehension. *Journal of Educational Psychology, 97*(2), 246-256.
- Ragnarsdottir, H., Simonsen, H. G., & Plunkett, K. (1999). The acquisition of past tense morphology in Icelandic and Norwegian children: An experimental study. *Journal of Child Language, 26*(3), 577-618.
- Rastle, K., Davis, M. H., Marslen-Wilson, W. D., & Tyler, L. K. (2000). Morphological and semantic effects in visual word recognition: A time-course study. *Language and Cognitive Processes, 15*(4), 507-536.
- Ravid, D. (2001). Learning to spell in Hebrew: Phonological and morphological factors. *Reading and Writing, 14*(5), 459-485.
- Ravid, D., & Zilberbuch, S. (2003). Morphosyntactic constructs in the development of spoken and written Hebrew text production. *Journal of Child Language, 30*(2), 395-418.
- Rayner, K., Sereno, S. C., Lesch, M. F., & Pollatsek, A. (1995). Phonological codes are automatically activated during reading: Evidence from an eye movement priming paradigm. *Psychological Science, 6*(1), 26-32.
- Rego, L. L. (1999). Phonological awareness, syntactic awareness and learning to read and spell in Brazilian portuguese. In M. Harris & G. Hatano (Eds), *Learning to read and write: A cross-linguistic perspective*. New York, US: Cambridge University Press.

- Rego, L. L. B. (1993). O papel da consciência sintática na aquisição da língua escrita. *Temas em psicologia, 1*, 79-87.
- Rego, L. L. B. (1997). The connection between syntactic awareness and reading: Evidence from portuguese-speaking children taught by a phonic method. *International Journal of Behavioral Development, 20*(2), 349-365.
- Rego, L. L. B., & Buarque, L. L. (1997). Consciência sintática, consciência fonológica e aquisição de regras ortográficas. *Psicologia: Reflexão e Crítica, 10*(2), 1999-1217.
- Rego, L. L., & Bryant, P. E. (1993). The connection between phonological, syntactic and semantic skills and children's reading and spelling. *European Journal of Psychology of Education, 8*(3), 235-246.
- Rey-Debove, J. (1984). Le domaine de la morphologie lexicale. *Cahier de Lexicologie, 45*, 3-19.
- Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studies in Second Language Acquisition, 18*, 27-67.
- Robinson, P. (1997). Generalizability and automaticity of second language learning under implicit, incidental, enhanced, and instructed conditions. *Studies in Second Language Acquisition, 19*, 223-247.
- Rocher, A.-S. (2005). *Régularités graphophonologiques, orthographiques, morphologiques: Apprentissage implicite et impact précoce sur la lecture*. Unpublished Doctorat de Psychologie, Université de Haute-Bretagne Rennes 2, Rennes.
- Romdhane, M. N. (2003). Déterminants de l'acquisition de la lecture compréhension en arabe. In M. N. Romdhane, J. E. Gombert & M. Belajouza (Eds.), *L'apprentissage de la lecture* (pp. 99-112). Rennes: Presses Universitaires de Rennes.
- Romdhane, M. N., Gombert, J. E., & Belajouza, M. (2003). *L'apprentissage de la lecture: Perspectives comparatives*. Rennes: Presses Universitaires de Rennes.
- Rosa, E. M., & Leow, R. P. (2004). Awareness, different learning conditions, and second language development. *Applied Psycholinguistics, 25*(2), 269-292.
- Rosa, E. M., & O'Neill, M. D. (1999). Explicitness, intake, and the issue of awareness. Another piece to the puzzle. *Studies in Second Language Acquisition, 21*, 511-553.
- Roy, C. (2005). *Comparaison des compétences en morphologie dérivationnelle chez les francophones et non-francophones apprentis lecteurs*. Unpublished mémoire en linguistique, Université du Québec, Montréal.
- Royer, C. (2004). *Connaissances morphologiques dérivationnelles et apprentissage de la lecture*. Unpublished Doctorat de Psychologie, Université de Savoie, Chambéry.

- Rubin, H. (1988). Morphological knowledge and early writing ability. *Language and speech*, 31(4), 337-355.
- Rueckl, J. G., Mikolinski, M., Raveh, M., Miner, C. S., & Mars, F. (1997). Morphological priming, fragment completion, and connectionist networks. *Journal of Memory and Language*, 36(3), 382-405.
- Rumelhart, D. E., & McClelland, J. L. (1986). On learning the past tenses of English verbs. In J. L. McClelland & D. E. Rumelhart (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Psychological and biological models* (Vol. 2, pp. 216-268). London: MIT Press.
- Saiegh-Haddad, E. (2003a). Bilingual oral reading fluency and reading comprehension: The case of Arabic/Hebrew (L1)-English (L2) readers. *Reading and Writing*, 16(8), 717-736.
- Saiegh-Haddad, E. (2003b). Linguistic distance and initial reading acquisition: The case of Arabic diglossia. *Applied Psycholinguistics*, 24(3), 431-451.
- Saiegh-Haddad, E. (2004). The impact of phonemic and lexical distance on the phonological analysis of words and pseudowords in a diglossic context. *Applied Psycholinguistics*, 25, 495-512.
- Saiegh-Haddad, E. (2005). Correlates of reading fluency in Arabic: Diglossic and orthographic factors. *Reading and Writing*, 18(6), 559-582.
- Sasaki, Y. (1991). English and Japanese interlanguage comprehension strategies: An analysis based on the competition model. *Applied Psycholinguistics*, 12, 47-73.
- Schelstraete, M.-A. (1993). La conception du traitement syntaxique en compréhension de phrases. *L'Année psychologique*, 93, 543-582.
- Schiff, R., & Calif, S. (2007). Role of phonological and morphological awareness in 12 oral word reading. *Language Learning*, 57(2), 271-298.
- Schmidt, R. (1992). Awareness and second language acquisition. *Annual review of applied linguistics*, 13, 206-226.
- Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 131-154). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Schreuder, R., & Baayen, R. H. (1997). How complex simple words can be. *Journal of Memory and Language*, 37(1), 118-139.
- Segalowitz, N. S., & Segalowitz, S. J. (1993). Skilled performance, practice, and the differentiation of speed-up from automatization effects: Evidence from second language word recognition. *Applied Psycholinguistics*, 14(3), 369-385.

- Seidenberg, M. S. (2005). Connectionist models of word reading. *Current Directions in Psychological Science*, 14(5), 238-242.
- Seidenberg, M. S., & Gonnerman, L. M. (2000). Explaining derivational morphology as the convergence of codes. *Trends in Cognitive Sciences*, 4(9), 353-361.
- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96(4), 523-568.
- Seidenberg, M. S., & McClelland, J.L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96(4), 523-568.
- Selby, S. (1972). The development of morphological rules in children. *British Journal of Educational Psychology*, 42(3), 293-299.
- Selinker, L. (1972). Interlanguage. *IRAL*(10), 219-231.
- Seymour, P. H. K. (1997). Les fondations du développement orthographique et morphographique. In L. Rieben, M. Fayol & C. A. Perfetti (Eds.), *Des orthographes et leur acquisition* (pp. 385-403). Lausanne, Paris: Delachaux et Niestlé.
- Seymour, P. H. K. (1999). Cognitive architecture of early reading. In I. Lundberg, E. E. Tonnenssen & I. Austrad (Eds.), *Dyslexia: Advances in theory and practice* (pp. 59-73). Dordrecht: Kluwer.
- Seymour, P. H. K., & Duncan, L. G. (1997). Small versus large unit theories of reading acquisition. *Dyslexia*, 3, 125-134.
- Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94(2), 143-174.
- Shankweiler, D., Crain, S., Brady, S., Macaruso, P., Gough, P. B., Ehri, L. C., & Treiman, R (1992). *Identifying the causes of reading disability*. Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.
- Shankweiler, D., Crain, S., Katz, L., Fowler, A. E., Liberman, A.M., Brady, .S.A., Thornton, R., Lundquist, E., Dreyer, L., Fletcher, J.M., Stuebing, K.K, Shaywitz, S.E. & Shaywitz, B.A. (1995). Cognitive profiles of reading-disabled children: Comparison of language skills in phonology, morphology, and syntax. *Psychological Science*, 6(3), 149-156.
- Share, D. L., & Blum, P. (2005). Syllable splitting in literate and preliterate Hebrew speakers: Onsets and rimes or bodies and codas? *Journal of Experimental Child Psychology*, 92(2), 182-202.
- Share, D., & Levin, I. (1999). Learning to read and write in Hebrew. In M. Harris & G. Hatano (Eds.), *Learning to read and write. A cross linguistic perspective*. (pp. 89-111). Cambridge: Cambridge University Press.

- Shimron, J. (1999). The role of vowel signs in Hebrew: Beyond word recognition. *Reading and Writing: An Interdisciplinary Journal*, 11, 301-319.
- Singson, M., Mahony, D., & Mann, V. (2000). The relation between reading ability and morphological skills: Evidence from derivation suffixes. *Reading and Writing*, 12(3), 219-252.
- Slobin, D. I., & Bever, T. G. (1982). Children use canonical sentence schemas: A crosslinguistic study of word order and inflections. *Cognition*, 12(3), 229-265.
- Smith, M. S. (1994). *Second language learning: Theoretical foundations*. London and New York: Longman.
- Smith-Lock, K. M., & Rubin, H. (1993). Phonological and morphological analysis skills in young children. *Journal of Child Language*, 20(2), 437-454.
- Snoussi, F. (2003). *Situation de diglossie et apprentissage de la lecture en arabe*. Unpublished mémoire de DEA de Psychologie, Université de Tunis, Tunis.
- Soares, C., & Grosjean, F. O. (1984). Bilinguals in a monolingual and a bilingual speech mode: The effect on lexical access. *Memory & Cognition*, 12(4), 380-386.
- Sokolov, J. L. (1989). The development of role assignment in Hebrew. In B. MacWhinney & E. Bates. (Eds.), *The crosslinguistic study of sentence processing* (pp. 118-157). New York: Cambridge University Press.
- Sokolov, J. L. (1989). The development of role assignment in Hebrew. In B. MacWhinney & E. Bates. (Eds.), *The crosslinguistic study in sentence processing* (pp. 118-157). New York: Cambridge University Press.
- Sousa, E. D. O. (2005). *Habilidades metasintáticas e aprendizagem da leitura: Estudo com crianças da 1ª série do ensino fundamental*. Unpublished Doctorat de Psychologie, Universidade de São Paulo, São Paulo.
- Sparks, R. L., & Ganschow, L. (1993b). The impact of native language learning problems on foreign language learning: Case study illustrations of the linguistic coding deficit hypothesis. *Modern Language Journal*, 77(1), 58-74.
- Sparks, R. L., Patton, J., Ganschow, L., Humbach, N., & Javorsky, J. (2006). Native language predictors of foreign language proficiency and foreign language aptitude. *Annals of Dyslexia*, 56(1), 129-160.
- Sparks, R., & Ganschow, L. (1993a). Searching for the cognitive locus of foreign language learning difficulties: Linking first and second language learning. *Modern Language Journal*, 77(3), 289-302.
- Sprenger-Charolles, L. (1992). L'évolution des mécanismes d'identification des mots. In M. Fayol, J-E. Gombert, P. Lecocq, L. Sprenger-Charolles & D. Zagar (Eds.),

- Psychologie cognitive de la lecture* (pp. 141-193). Paris: Presses Universitaires de France.
- Sprenger-Charolles, L., & Casalis, S. (1996). *Lire. Lecture et écriture: Acquisition et troubles du développement*. Paris: Presses Universitaires de France.
- Stanovich, K. E. (1982a). Individual differences in the cognitive processes of reading: I. Word decoding. *Journal of Learning Disabilities, 15*(8), 485-493.
- Stanovich, K. E. (1982b). Individual differences in the cognitive processes of reading: II. Text-level processes. *Journal of Learning Disabilities, 15*(9), 549-554.
- Stanovich, K. E., Cunningham, A. E., & Feeman, D. J. (1984). Relation between early reading acquisition and word decoding with and without context: A longitudinal study of first-grade children. *Journal of Educational Psychology, 76*(4), 668-677.
- Stolz, J. A., & Besner, D. (1998). Levels of representation in visual word recognition: A dissociation between morphological and semantic processing. *Journal of Experimental Psychology: Human Perception and Performance, 24*(6), 1642-1655.
- Stolz, J. A., & Feldman, L. B. (1995). The role of orthographic and semantic transparency of the base morpheme in morphological processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 109-129). Hillsdale, N. J.: Lawrence Erlbaum Associates.
- Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior, 14*(6), 638-647.
- Taman, H. A. (1993). The utilization of syntactic, semantic, and pragmatic cues in the assignment of subject role in Arabic. *Applied Psycholinguistics, 14*(3), 299-317.
- Tanenhaus, M. K., Boland, J., Garnsey, S. M., & Carlson, G. N. (1989a). Lexical structure in parsing long-distance dependencies. *Journal of Psycholinguistic Research, 18*(1), 37-50.
- Tanenhaus, M. K., Carlson, G., & Trueswell, J. C. (1989b). The role of thematic structures in interpretation and parsing. *Language and Cognitive Processes, 4*(3), SI211-SI234.
- Taouk, M., & Coltheart, M. (2004). The cognitive processes involved in learning to read in Arabic. *Reading and Writing, 17*(1), 27-57.
- Taraban, R., & McClelland, J. L. (1988). Constituent attachment and thematic role assignment in sentence processing: Influence of content-based expectations. *Journal of memory and language, 27*, 597-632.
- Taraban, R., & McClelland, J. L. (1990). Parsing and comprehension: A multiple-constraint view. In D. A. Balota, G. B. Flores d'Arcais & K Rayner (Eds.), *Comprehension processes in reading*. (pp. 231-263). Hillsdale, N. J.: Lawrence Erlbaum Associates.

- Teyssier, P. (1992). *Manuel de langue portugaise*. Paris: Editions Klincksieck.
- Thorstad, G. (1991). The effect of orthography on the acquisition of literacy skills. *British Journal of Psychology*, 82(4), 527-537.
- Titone, R. (1972). *Le bilinguisme précoce*. Bruxelles: Delessert.
- Treiman, R. (2005a). Introduction: Linguistic constraints on literacy development: Introduction to the special issue. *Journal of Experimental Child Psychology*, 92(2), 103-106.
- Treiman, R. (2005b). Linguistic constraints on literacy development: Introduction to the special issue. *Journal of Experimental Child Psychology*, 92(4), 303-306.
- Treiman, R., & Zukowski, A. (1996). Children's sensitivity to syllables, onsets, rimes, and phonemes. *Journal of Experimental Child Psychology*, 61(3), 193-215.
- Tunmer, W. E. (1989). The role of language-related factors in reading disability. In D. Shankweiler & I. Y. Liberman (Eds.), *Phonology disability: Solving the reading puzzle* (pp. 91-132). Ann Arbor: The University of Michigan Press.
- Tunmer, W. E., Nesdale, A. R., & Wright, A. D. (1987). Syntactic awareness and reading acquisition. *British Journal of Developmental Psychology*, 5(1), 25-34.
- Tyler, A., & Nagy, W. (1989). The acquisition of English derivational morphology. *Journal of Memory and Language*, 28(6), 649-667.
- Van Gelderen, A., Schoonen, R., De Glopper, K., Hulstijn, J., Simis, A., Snellings, P. & Stevenson, M. (2004). Linguistic knowledge, processing speed, and metacognitive knowledge in first- and second-language reading comprehension: A componential analysis. *Journal of Educational Psychology*, 96(1), 19-30.
- Van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, 39, 458-483.
- Van Orden, G. C. (1987). A row is a rose: Spelling, sound, and reading. *Memory & Cognition*, 15(3), 181-198.
- Vidigal de Paula, F. (2007). *Conhecimento morfológico implícito et explícito na linguagem escrita*. Unpublished Doctorat de Psychologie, Universidade de São Paulo - Université Haute Bretagne Rennes 2, São Paulo.
- Vogel, K. (1995). *L'interlangue: La langue de l'apprenant*. Toulouse: Presses Universitaires du Mirail.
- Vygotsky, L. S. (1934, trad. 1985). *Pensée et langage*. Paris: Editions sociales.
- Wade-Woolley, L. (1999). First language influences on second language word reading: All roads lead to Rome. *Language Learning*, 49(3), 447-471.

- Wang, M., Koda, K., & Perfetti, C. A. (2003). Alphabetic and nonalphabetic L1 effects in English word identification: A comparison of Korean and Chinese English L2 learners. *Cognition*, 87 (2), 129-149.
- Wang, M., Koda, K., & Perfetti, C. A. (2004). Language and writing systems are both important in learning to read: A reply to Yamada. *Cognition*, 93 (2), 133-137.
- Wang, M., Park, Y., & Lee, K. R. (2006). Korean-English biliteracy acquisition: Cross-language phonological and orthographic transfer. *Journal of Educational Psychology*, 98(1), 148-158.
- White, T. G., Power, M. A., & White, S. (1989). Morphological analysis: Implications for teaching and understanding vocabulary growth. *Reading Research Quarterly*, 24(3), 283-309.
- Willows, D. M., & Ryan, E. B. (1986). The development of grammatical sensitivity and its relationship to early reading achievement. *Reading Research Quarterly*, 21(3), 253-266.
- Willows, D. M., & Ryan, E. B. (1986). The development of grammatical sensitivity and its relationship to early reading achievement. *Reading Research Quarterly*, 21(3), 253-266.
- Wimmer, H., & Goswami, U. (1994). The influence of orthographic consistency on reading development: Word recognition in English and German children. *Cognition*, 51(1), 91-103.
- Yamada, J. (2004). An L1-script-transfer-effect fallacy: A rejoinder to Wang *et al.* (2003). *Cognition*, 93(2), 127-132.
- Zagar, D. (1992). L'approche cognitive de la lecture: De l'accès au lexique au calcul syntaxique. In M. Fayol, J-E. Gombert, P. Lecocq, L. Sprenger-Charolles & D. Zagar (Eds.), *Psychologie cognitive de la lecture* (pp. 15-72). Paris: Presses Universitaires de France.
- Zanella, M. S. (2007). *Leitura e aprendizagem da ortografia: Um estudo com alunos de 4ª a 6ª série do ensino fundamental*. Unpublished Tese de Doutorado em Psicologia da Educação, Pontifícia Universidade Católica de São Paulo, São Paulo.
- Ziegler, J. C. (2004). *Du lecteur débutant au lecteur expert: Perspectives cross-linguistiques*. Unpublished Habilitation à Diriger des Recherches, Université de Provence, Aix-Marseille 1.
- Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, 131(1), 3-29.

Ziegler, J. C., Van Orden, G. C., & Jacobs, A. M. (1997). Phonology can help or hurt the perception of print. *Journal of Experimental Psychology: Human Perception and Performance*, 23(3), 845-860.

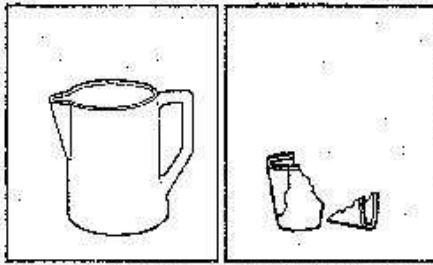
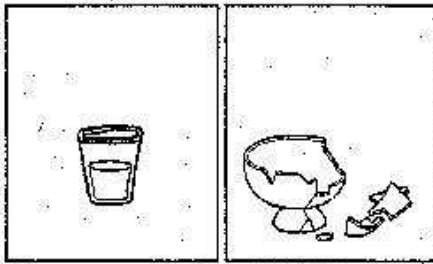
Annexe 1

Test de lecture de mots en une minute en arabe vocalisé

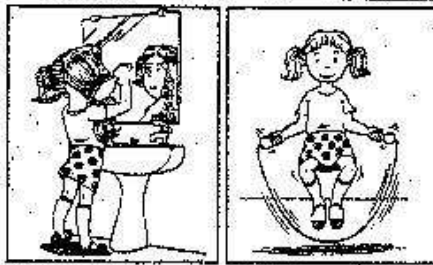
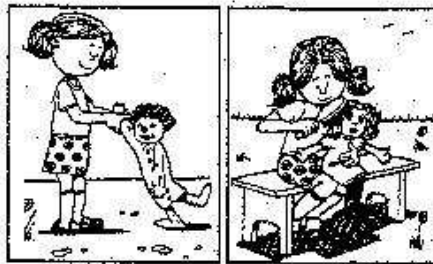
أَمَامَ	سَقَطَ	سِرَّ	كُرَّاسٌ	تَلْمِيزًا	لَا	أَطْفَالَ	بَحْرًا	كَمْ
لُحْمَةٌ	فَقَصٌ	نَمَّ	قِطَارٌ	يَقْفِرُ	فَمَّ	مَسْكِينٌ	طُيُورٌ	لَنْ
يَطْعِمُ	رَجُلٌ	خَطَّ	طَرَبُوشٌ	بَارِدٌ	جَا	حِسَابٌ	نَوَاءٌ	قَطٌّ
صَابُونَ	حَلِيبٌ	فَيْلٌ	صَفْرَاءٌ	أَرْسَمَ	يَدٌ	وَلَدٌ	قَوِيٌّ	مَاءٌ
مُحَفِّظَةٌ	قَلَمٌ	نَسَخَ	نَجَاجَةٌ	وَرَاءَ	رِيحٌ	قِرَاءَةٌ	أَرْزَبٌ	فِسْمٌ
رِسَالَةٌ	نُودَةٌ	قَرَدٌ	حَلْزُونٌ	صُورَةٌ	فَمٌ	حِكَايَةٌ	يَلْعَبُ	بِنْتُ
جَدِيدَةٌ	كِتَابٌ	تَمْرٌ	مَدْرَسَةٌ	جَمِيلٌ	كَلْبٌ	دَرَجَةٌ	فِرَاحٌ	نَهْرٌ
خِرَانَةٌ	مَنْزِلٌ	شَعْرٌ	حَدِيقَةٌ	نَكِيٌّ	تَحْتَ	سَبُّورَةٌ	يَجْرِي	مِشْطٌ
شَجَرَةٌ	نُؤْمِيَةٌ	بَيْكٌ	مُعَلِّمَةٌ	جِصَّانٌ	فَارٌ	زَرَافَةٌ	أَكَلَ	خُوتٌ
بَيْعَاءٌ	يَصِيحُ	فَوْقَ	سَيَّارَةٌ	رَاحَةٌ	شَمْسٌ	مُلُونَةٌ	صَغِيرٌ	حَبْلٌ
طَائِرَةٌ	مَاهِرٌ	فِصَّةٌ	وَرَقَةٌ	لُعْبَةٌ	حَلْوَى	سَمَكَةٌ	لَوْنٌ	ثَبٌّ
تَكَثَّرَتْ	كَثِيرٌ	فَتَحَ	طَاوِلَةٌ	سَاحَةٌ	كَتَبَ	عُصْفُورَةٌ	شَارِعٌ	أَحْمَرٌ

الإسم والتَّقْب :

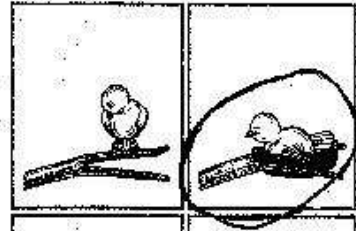
القِسْم :



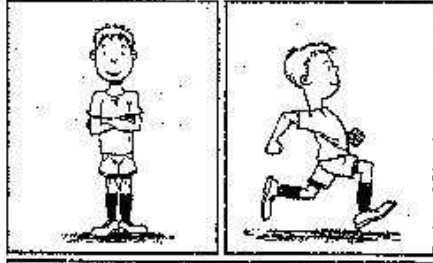
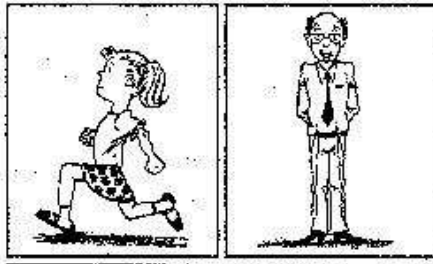
الكأسُ غيرُ مكسّرٍ.



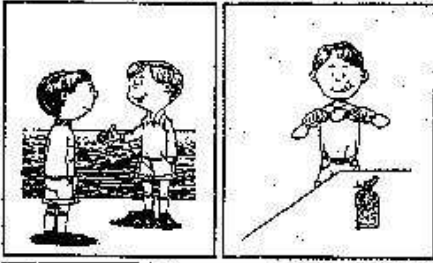
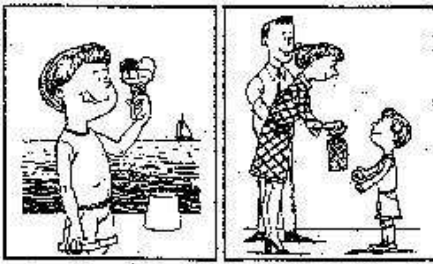
تلعبُ البنتُ بالحبْلِ.



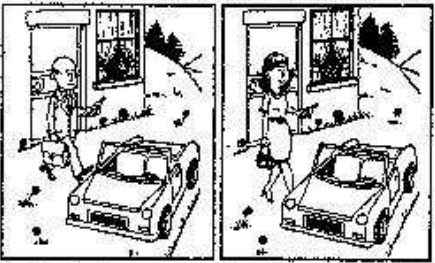
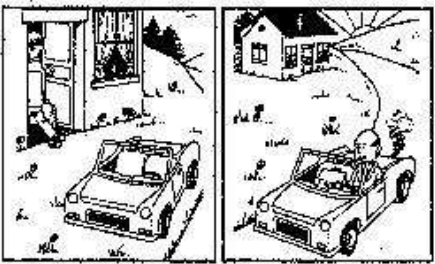
العصفورُ نائمٌ في عشِّهِ.



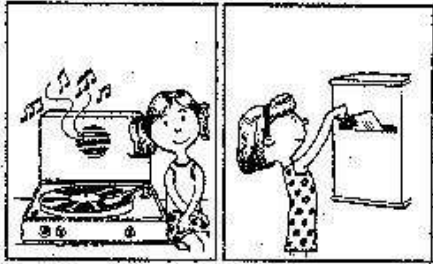
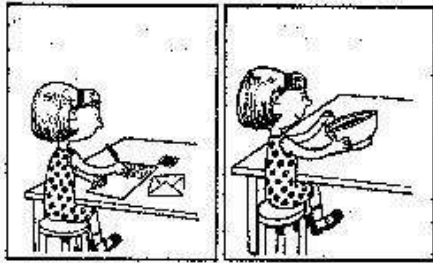
لا يجري الولدُ.



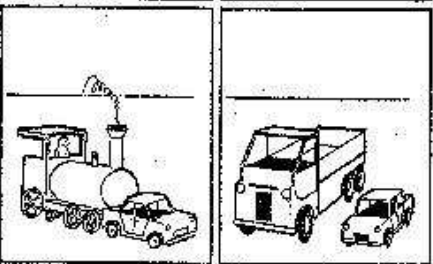
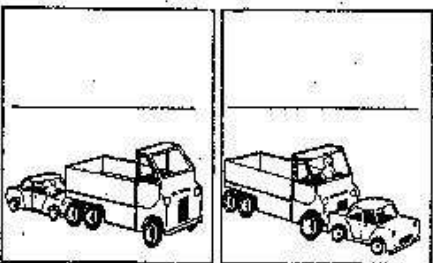
أرى أنك تأكلُ حلوى.



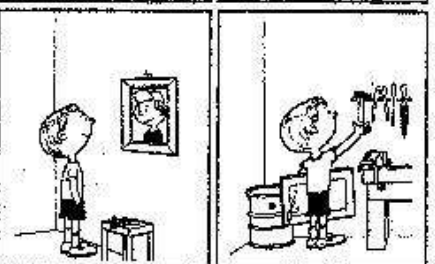
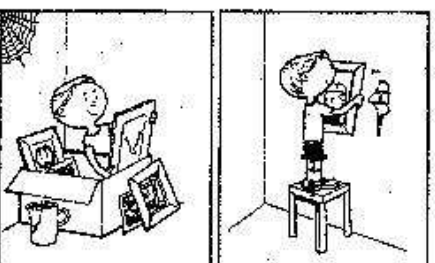
سيركبُ الرجلُ سيارتَهُ.



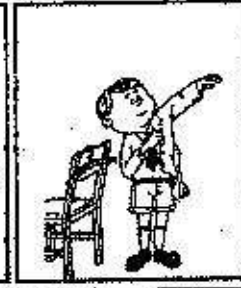
البنتُ كتبتُ رسالةً.



السيارةُ مصنوعةٌ من طرفِ الشاحنة.



علقَ الولدُ صورةً.



تتحدثين معها.

قالت لي أمي أن ألبس معطفي.

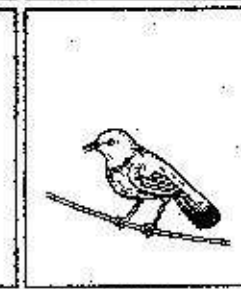
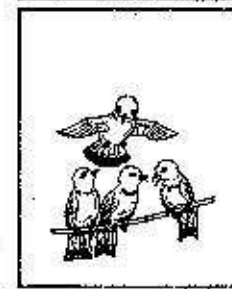
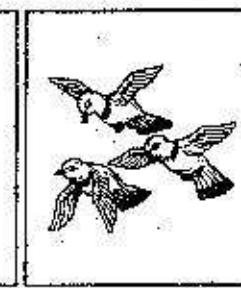
الوكذ الصغير سيقرا كتابا.



لماذا كسرت لعبة أختك ؟

الدب نائم.

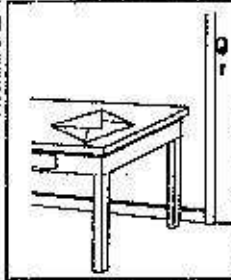
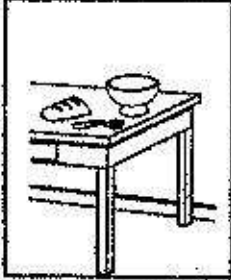
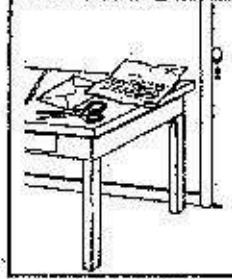
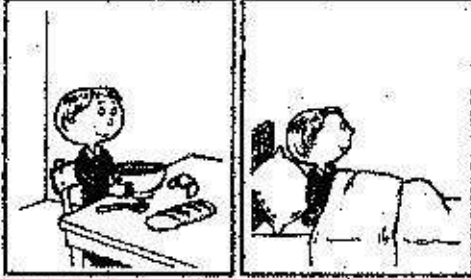
هل سقطت البنت ؟



البنت التي ضربتها الوكذ لها طربوش.

العصفوران واقفان فوق الحبل.

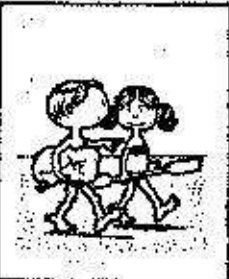
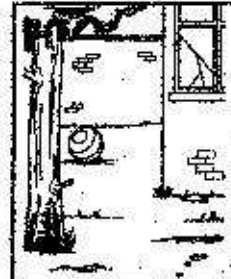
يعيشي التلميذان في الطريق.



أكل صالح فطور الصباح.

تقبل مَحَمَّدَ رسالة. بعد أن فتحها،
قرأها ثم وضعها على طرف الطاولة

أين السمكة التي وضعتها في هذا
الطبق؟



بعد أن لعب صالح و زينب في
البحر، رجعا إلى الدار.

قذف سامي الكرة بقوة فتكسر بلور
الشباك. غضبت أمه و عاقبته.

ثُرُوقٌ

سِبَاقٌ

دِفَاعٌ

عَبٌّ

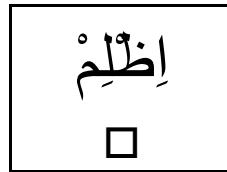
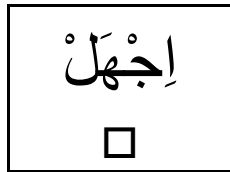
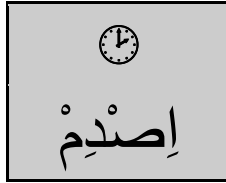
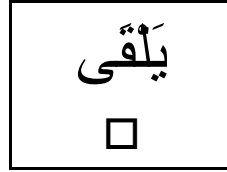
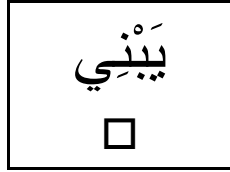
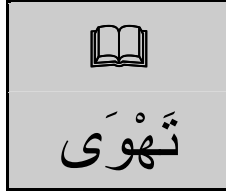
إِسْرَاقٌ

إِضْرَابٌ

تَطْوِي

يَبْكِي

يُنْسَى



✂
مُجَرَّبٌ

اِنْتَقَمَ
□

تَحَارَبَ
□

➔
خُنْزِيرٌ

أَسْتَاذٌ
□

طَبَشُورٌ
□

🕯
مُؤَشِّرٌ

اِنْتَشَرَ
□

تَقَابَلَ
□

✂
تَجُو

يَشْفِي

يَلْهُو

👂
بَشْر

مُدُن

حُقْر

👓
أَكْتُبْ

إِجْذِبْ

إِسْأَلْ

🚲
مُحَمَّلْ

إِنْتَخِبْ

تَعَامَلْ

Item	Intrus	traduction	C. G.	S. P.	N.S.	N.P.	N.G.	N.L.I. *	Lié	traduction	C. G.	S. P.	N.S.	N.P.	N.G.	N.L.I. *	Cible	traduction	C. G.	S. P.	N.S.	N.P.	N.G.	N.L.I. *	Unité phonologique	choix sémantique*
1	دِفَاع	défense	nom	CvCVC	2	5	4	2	سَبَاق	course	nom	CvCVC	2	5	4	1	شُرُوق	lever de soleil	nom	CvCVC	2	5	4	2	C terminale	40%
2	إِسْرَاقٌ	vole	verbe	CvCCvC	2	6	4	2	إِضْرَبْ	frapper	verbe	CvCCvC	2	6	4	2	إِلْعَابٌ	joue	verbe	CvCCvC	2	6	4	1	rime	80%
3	يُنْسَى	oublier	verbe	CvCCV	2	5	4	0	يَبْكِي	pleurer	verbe	CvCCV	2	5	4	0	تَطْوِي	plier	verbe	CvCCV	2	5	4	1	V terminale	38%
4	تَتَابَعٌ	continue	verbe	CvCVCvC	3	7	5	1	إِنْفَجِرْ	explose	verbe	CvCVCvC	3	7	5	1	مُؤَجَّر	propriétaire	nom	CvCVCvC	3	7	5	1	syllabe	
5	يَبْنِي	construire	verbe	CvCCV	2	5	4	0	يَلْقَى	rencontrer	verbe	CvCCV	2	5	4	0	تَهْوَى	désirer	verbe	CvCCV	2	5	4	1	V terminale	38%
6	إِجْهَلْ	ignore	verbe	CvCCvC	2	6	4	1	إِظْلِمْ	cause du tord	verbe	CvCCvC	2	6	4	1	إِصْدِمْ	choque	verbe	CvCCvC	2	6	4	2	rime	65%
7	إِنْتَقِمْ	venge toi	verbe	CvCCvCvC	3	8	5	1	تُحَارِبْ	fais la guerre	verbe	CvCCvCvC	3	8	5	2	مُجَرَّبٌ	expérimenté	adjectif	CvCCvCvC	3	8	5	1	syllabe	
8	أُسْتَاذٌ	professeur	nom	CvCCvC	2	6	5	2	طَبَسُّور	craie	nom	CvCCvC	2	6	5	1	خِزِيرٍ	cochon	nom	CvCCvC	2	6	5	1	C terminale	80%
9	تَقَابَلْ	rencontre	verbe	CvCVCvC	3	7	5	1	إِنْتَشِرْ	répond	verbe	CvCVCvC	3	7	5	0	مُؤَشِّرٌ	indicateur	nom	CvCVCvC	3	7	5	1	syllabe	
10	يَسْتَفِي	guérir	verbe	CvCCV	2	5	4	0	يَلْهُو	se distraire	verbe	CvCCV	2	5	4	0	تَنْجُو	se sauver	verbe	CvCCV	2	5	4	0	V terminale	67%
11	مُدُنٌ	villes	nom	CvCvC	2	5	2	1	حُفْرٌ	trous	nom	CvCvC	2	5	2	0	بَشَرٌ	gens	nom	CvCvC	2	5	2	0	rime	60%
12	إِسْأَلٌ	questionne	verbe	CvCCvC	2	6	4	2	إِجْدِبْ	tire	verbe	CvCCvC	2	6	4	2	أَكْتَبْ	écris	verbe	CvCCvC	2	6	4	1	C terminale	47%
13	إِنْتِخِبْ	vote	verbe	CvCCvCvC	3	8	5	1	تُعَامَلْ	fais des affaires	verbe	CvCCvCvC	3	8	5	1	مُحَمَّلٌ	chargé	adjectif	CvCCvCvC	3	8	5	0	syllabe	

C. G. = Catégorie Grammaticale ; S.P. = Structure Phonologique

N.S. = Nombre de syllabes; N.P. = Nombre de phonèmes; N.G. = Nombre de graphèmes; N.L.I.= Nombre de lettres isolées

* sans compter la lettre finale

** % moyen de réponses de 20 élèves de 2ème année en faveur de l'intrus

Moyenne choix sémantique = 57 %



تَابِعَ

نَاقَشَ



غَافَلَ



حَبَزَهُ

سُئِرَهُ



دُمِيَ



مِطْبَخَةٌ

مِنْقَلَةٌ



مِلْحَفَةٌ





أَشْرَبَ

أَسْعَفَ



أَنْهَضَ



مَادَحَ

قَاطَعَ



ضَايَقَ



مِطْرَقَةٌ

مِعْرَقَةٌ



مِحْرَمَةٌ



قُدْرَةٌ

سُفْرَةٌ

شُرْبَةٌ

أَغْلَقَ

أَطَالَ

أَقَادَ

قَاسَمَ

نَافَسَ

حَاصَرَ



أَلْبَسَ

أَوْقَفَ



أَضْرَبَ



عُقِدَ

حُجِرَ



أَكَلَتْ



مِنْظَرَ

مِنْقَارَ



مِسْمَارَ



Item	Intrus	Traduction	N.S.	N.P.	N.G.	N.L.I.*	Familiarité**	Liés	Traduction	N.S.	N.P.	N.G.	N.L.I.*	Familiarité**	Cible	Traduction	N.S.	N.P.	N.G.	N.L.I.*	Type de schème	choix sémantique***
1	نَاقَشَ	discuter	3	6	4	1	1,286	غَافَلَ	distraire qqun	3	6	4	1	0,483	تَابَعَ	suivre qqun	3	6	4	1	Verbe forme III	38%
2	دُمِيَّةٌ	poupée	3	8	4	1	1,760	سُتْرَةٌ	veste, manteau	3	8	4	1	0,217	خُبْزَةٌ	pain	3	8	4	1	Substantif	86%
3	مِيقَلَةٌ	équerre	4	10	5	0	0,652	مِخْفَقَةٌ	sorte de haik	4	10	5	0	0,522	مِطْبَخَةٌ	cuisinière	4	10	5	0	instrument	36%
4	أَسْعَفَ	aider; secourir	3	7	4	1	1,034	أَنهَضَ	faire lever	3	7	4	1	1,440	أَشْرَبَ	faire boire	3	7	4	2	Verbe forme IV	32%
5	ضَايِقٌ	embêter qqun	3	6	4	1	0,478	قَاطَعَ	couper les liens avec qqun	3	6	4	1	1,130	مَادَحٌ	complimenter qqun	3	6	4	2	Verbe forme III	55%
6	مِحْرَمَةٌ	mouchoir	4	10	5	1	0,864	مِعْرَفَةٌ	cuillère	4	10	5	1	0,818	مِطْرَقَةٌ	marteau	4	10	5	1	instrument	45%
7	سُفْرَةٌ	table dressée	3	8	4	1	0,414	شُرْبَةٌ	soupe	3	8	4	1	1,407	قُدْرَةٌ	capacité	3	8	4	2	Substantif	45%
8	أَقَادَ	être utile	3	6	4	2	0,769	أَطَالَ	prolonger, allonger	3	6	4	2	0,769	أَغْلَقَ	faire fermer	3	6	4	1	Verbe forme IV	59%
9	نَافَسَ	entrer en compétition	3	6	4	1	0,840	حَاصَرَ	encercler qqun	3	6	4	1	0,923	فَاسَمَ	partager	3	6	4	1	Verbe forme III	32%
10	أَضْرَبَ	faire grève	3	7	4	2	1,087	أَوْقَفَ	faire arrêter	3	7	4	2	1,609	أَلْبَسَ	faire habiller	3	7	4	1	Verbe forme IV	55%
11	حُجْرَةٌ	pièce	3	8	4	1	1,480	أَكْلَةٌ	nourriture	3	8	4	1	1,808	عُقْدَةٌ	nœud	3	8	4	1	Substantif	18%
12	مِيسْمَارٌ	clou	3	8	5	1	1,826	مِيقَارٌ	bec	3	8	5	1	1,696	مِيزَانٌ	lunettes, jumelles	3	8	5	1	instrument	59%

Moyenne familiarité 1,041

1,068

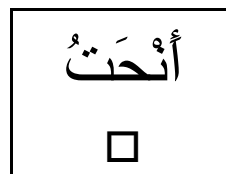
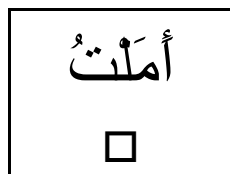
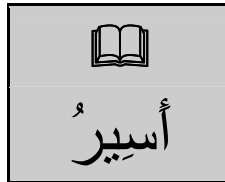
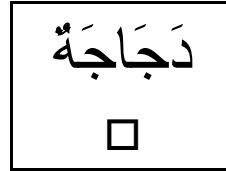
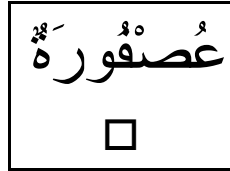
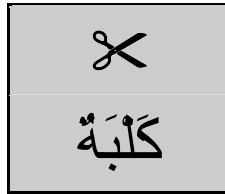
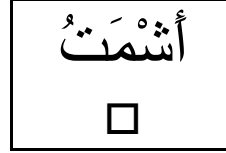
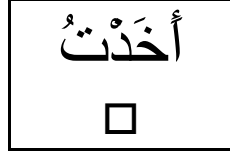
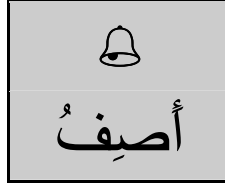
Moyenne choix sémantique = 47 %

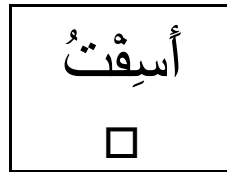
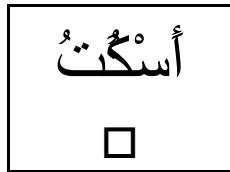
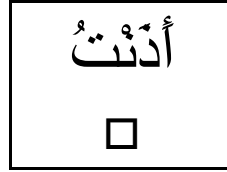
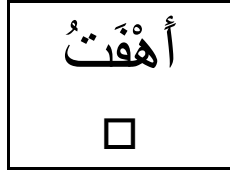
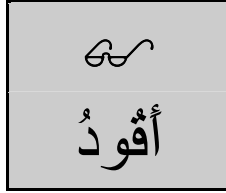
N.S. = Nombre de syllabes; N.P. = Nombre de phonèmes; N.G. = Nombre de graphèmes; N.L.I. = Nombre de lettres isolées

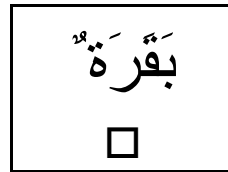
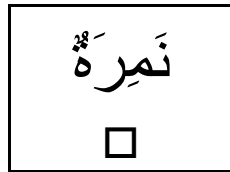
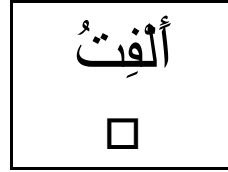
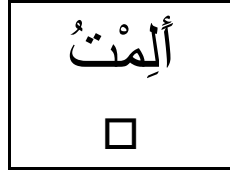
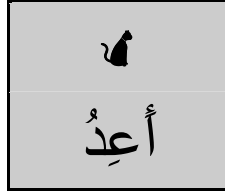
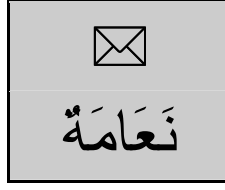
* sans compter la lettre finale

** réponses moyennes /3 d'une vingtaine d'élèves de 3ème année à un test de familiarité lexicale (échelle en trois points)

*** % moyen de réponses de 20 élèves de 2ème année en faveur de l'intrus









أَقُولُ

أَقَلْتُ



أَنْسَتُ



تَنْبِيْئًا

سَمَكَةً



ضَيْقَةً



فَرُوجَةً

بَعْلَةً



نَمْلَةً



Item	Intrus	Traduction	N.S.	N.P.	N.G.	N.L.I.*	Familiarité**	Liés	Traduction	N.S.	N.P.	N.G.	N.L.I.*	Familiarité**	Cible	Traduction	N.S.	N.P.	N.G.	N.L.I.*	Type de flexion	choix sémantique***
1	أَخَذْتُ	j'ai pris	3	6	4	2	2,000	أَشْمَتُ	je me réjouis du malheur des autres	3	6	1	1	1,040	أَصِفُ	je décris	3	6	3	1	Conjugaison verbale	52%
2	دَجَاجَةٌ	poule	4	9	5	2	1,816	عُصْفُورَةٌ	oiseau	4	10	6	2	2,842	كَلْبَةٌ	chienne	3	8	4	0	féminin	30%
3	أَمَلْتُ	j'ai souhaité	3	7	4	1	0,880	أَنْحَتُ	je sculpte	3	7	1	1	0,167	أَسِيرُ	je marche	3	6	4	1	Conjugaison verbale	48%
4	حَجَلَةٌ	perdix	4	9	4	0	0,079	عَرَبَةٌ	scorpion	4	10	5	0		تَعَلَبَةٌ	renarde	4	10	5	0	féminin	43%
5	أَدْنَتُ	j'ai permis	3	7	4	2	1,320	أَهْفَتُ	je voltige	3	7	1	1	0,250	أَقُودُ	je conduis	3	6	4	2	Conjugaison verbale	48%
6	أَسِفْتُ	j'étais désolé	3	7	4	1	1,391	أَسْكُتُ	je me tais	3	7	1	1	1,478	أَضَعُ	je mets	3	6	3	1	Conjugaison verbale	35%
7	لَبْوَةٌ	lionne	3	8	4	1	1,026	قِرْدَةٌ	guenon	3	8	4	2	1,974	نَعَامَةٌ	autruche	4	9	5	1	féminin	57%
8	أَلِمْتُ	j'ai souffert	3	7	4	1	0,870	أَلْفِتُ	j'attire	3	7	1	1	1,091	أَعِدُّ	je promets	3	6	3	1	Conjugaison verbale	52%
9	بَقْرَةٌ	vache	4	9	4	1	1,579	نَمِرَةٌ	tigresse	4	9	4	1	3,000	سِنْجَابَةٌ	écureuil	4	10	6	1	féminin	74%
10	أَيْسَتُ	j'étais accompagné	3	7	4	1	0,552	أَقْلِبْتُ	j'échappe	3	7	1	1	0,931	أَقُولُ	je parle	3	6	4	2	Conjugaison verbale	35%
11	سَمَكَةٌ	poisson	4	9	4	0	1,263	ضِفْدَعَةٌ	grenouille	4	10	5	1	0,632	تَيْيَنَةٌ	dragon	4	10	5	0	féminin	65%
12	نَمْلَةٌ	fourmi	3	8	4	0	0,079	بَعْلَةٌ	mule	3	8	4	0	0,079	فَرُوجَةٌ	poulette	5	10	5	2	féminin	48%

Moyenne familiarité 1,07

1,23

Moyenne choix sémantique = 49 %

N.S. = Nombre de syllabes; N.P. = Nombre de phonèmes; N.G. = Nombre de graphèmes; N.L.I. = Nombre de lettres isolées

* sans compter la lettre finale

** réponses moyennes /3 d'une vingtaine d'élèves de 3ème année à un test de familiarité lexicale (échelle en trois points).

Pour les items "féminin", le score a été calculé à partir de productions d'enfants de primaire à une tâche catégorielle (travaux non publiés de l'équipe EDIPSY)

*** % de réponses de 20 élèves de 2ème année en faveur de l'intrus



صَادَ سَمَكَةً

طَوَى لِحَافَيْنِ



شَكَا لِحَارَيْنِ



أَثَقْنَ الْعَمَلِيَّةَ

أَقْرَأَ بَغْلَطَتَيْنِ



أَرْسَلَ بِطَاقَتَيْنِ



غَافَ الْكُرَّاسَ مَعَ أُمَّهِ

شَرَبَ لِثَرَيْنِ مِنَ الْحَلِيبِ



أَذِنَ لِبَنَّتَيْنِ فِي الْخُرُوجِ





سَرَقَ تُقَّاحَةً

فَرَشَ بِسَاطِينَ



لَحِقَ بِقِطِينٍ



دَبِحَ بَقْرَةً

هَلَّلَ لِحَاجِّينَ



صَلَحَ لِحَامِيْنَ



لَازَمَ فِرَاشَهُ

شَاهَدَ بِلَاطِيْنَ



بَادَرَ بِلُفْظِيْنَ





تَأْمَلِ الْقَمَرَ

إِنَّهُمْ لَصَيِّن



هَتَفَ لِشَخَصَيْن



دَامَ لِحَطَّتَيْن

دَارَ لِشَيْخَيْن



كَوَى لِبَاسَيْن



هَدَمَتِ النَّيْرَانَ إِسْطَبْلَيْن

إِمْتَاذَ الْوَلَدِ يُكْتَتَيْن



إِخْتَارَ الْحَلْزُونَ بِيئَتَيْن





أَخَذَ جَائِزَةً

أَمْسَكَ بِسَطْلَيْنِ



أَفْرَعُ بِرْمِيلَيْنِ



زَارَ عَائِلَتَهُ

دَعَا لِبَحَّارَيْنِ



حَلَقَ لِحْيَتَيْنِ



جَعَّتِ الرِّيحُ الْقَمِيصَ

قَابَلَ الْمُعَلِّمُ بِنْتَيْنِ



طَالَبَ الْعَامِلُ بِأَجْرَيْنِ



Item	Intrus	Traduction	N.S. /mot*	N.G. /mot*	N.L.I. total**	Forme verbale	Familiarité ***	Lié	Traduction	N.S. /mot*	N.G. /mot*	N.L.I. total**	Forme verbale	Familiarité ***	Cible	Traduction	N.S. /mot*	N.G. /mot*	N.L.I. total**	Forme verbale	Préposition
1	شَكَا لِجَارَيْنِ	il s'est plaint à deux voisins	2-4	3-6	2	I	1,613	طَوَى لِحَافَيْنِ	il a plié deux couvertures lit	2-4	3-6	2	I	0,735	صَادَ سَمَكَةً	il a pêché un poisson	2-4	3-4	1	I	[li]
2	أَقْرَأَ بِعَطَائِنِ	il a reconnu deux fautes	3-5	3-7	1	IV	1,069	أَرْسَلَ بِطَاقَتَيْنِ	il a envoyé deux cartes	3-5	4-7	3	IV	1,517	أَتَقَّنَ الْعَمَلِيَّةَ	il a réussi l'opération	3-5	4-7	2	IV	[bi]
3	أَذِنَ لِبَنَاتَيْنِ فِي الْخُرُوجِ	il a autorisé à deux filles de sortir	3-4-1-3	3-6-2-6	5	I	1,765	شَرَبَ لِثَرَيْنِ مِنَ الْحَلِيبِ	il a bu deux litres de lait	3-3-1-4	3-5-2-6	3	I	1,884	عَلَّفَ الْكُرَّاسَ مَعَ أُمِّهِ	il a enveloppé le cahier avec sa mère	3-3-2-3	3-6-2-3	4	II	[li]
4	لَحَقَ بِعَظْمَيْنِ	il a rattrapé deux chats	3-4	3-5	0	I	1,715	فَرَشَ بِسَاطِنِ	il a étendu deux tapis	3-4	3-6	2	I	1,473	سَرَقَ تَفَاحَةً	il a volé une pomme	3-4	3-5	2	I	[bi]
5	هَلَّلَ لِحَاجِّينِ	il a poussé des cris de joies pour deux pèlerins	3-4	3-6	1	II	0,761	صَلَحَ لِحَامَيْنِ	il a réparé deux soudures	3-4	3-6	1	II	0,935	ذَبَحَ بَقْرَةً	il a égorgé une vache	3-4	3-4	2	I	[li]
6	بَادَرَ بِلَفْطَيْنِ	il a commencé de parler par deux mots	3-4	4-6	2	III	0,558	شَاهَدَ بِبَلَّاطَيْنِ	il a vu deux palais	3-4	4-6	2	III	1,173	لَا زَمَ فَرَّاشَتُهُ	il est resté dans son lit	3-4	4-5	4	III	[bi]
7	هَتَّفَ لِشَخْصَيْنِ	il a applaudi deux individus	3-4	3-6	0	I	1,715	إِثْمَمَ لِصَنَيْنِ	il a accusé deux voleurs	4-3	4-4	1	VIII	0,887	تَأَمَّلَ الْقَمَرَ	il a observé la lune	4-3	4-5	2	VIII	[li]
8	دَارَ لِشَيْخَيْنِ	il s'est retourné vers deux vieillards	2-4	3-6	1	I	1,260	كَوَى لِبَاسَتَيْنِ	il a repassé deux habits	2-4	3-6	2	I	1,160	دَامَ لِحَطَّتَيْنِ	il a duré deux instants	2-4	3-6	2	I	[li]
9	إِمْتَانَ الْوَلَدَ بِتُكْنَيْنِ	le garçon s'est distingué par deux blagues	3-3-5	5-5-7	5	VIII	1,594	إِخْتَارَ الْحِلْزُونَ بَيْنَتَيْنِ	l'escargot a choisi deux foyers	3-4-4	5-7-6	5	VIII	1,841	هَدَمَتِ النَّيْرَانَ إِسْطَبْلَيْنِ	les flammes ont détruit deux étables	4-3-4	4-7-7	5	II	[bi]
10	أَمْسَكَ بِسَاطِلَيْنِ	il a saisi deux sauts	3-4	4-6	1	IV	1,835	أَفْرَعُ بِرَمِيْلَيْنِ	il a vidé deux tonneaux	3-4	4-7	3	IV	1,704	أَخَذَ جَائِزَةً	il a obtenue une récompense	3-4	3-5	3	I	[bi]
11	دَعَا لِجَارَيْنِ	il a prié pour deux marins	2-5	3-7	3	I	1,357	حَلَقَ لِخَيْتَيْنِ	il a rasé deux barbes	3-4	3-6	0	II	1,132	زَارَ عَائِلَتَهُ	il a visité sa famille	2-5	3-6	3	I	[li]
12	طَالَبَ الْعَامِلَ بِأَجْرَيْنِ	l'ouvrier a réclamé deux loyers	3-3-4	4-6-6	5	III	1,410	قَابَلَ الْمُعَلِّمَ بِبَنَاتَيْنِ	le professeur a rencontré deux filles	3-4-3	4-6-5	2	III	1,880	جَفَّقَتِ الرِّيحُ الْقَمِيصَ	les vents ont séché les chemisiers	4-3-3	4-6-6	4	II	[bi]

Moyenne familiarité 1,388

1,360

N.S. = Nombre de syllabes; N.P. = Nombre de phonèmes; N.G. = Nombre de graphèmes; N.L.I. = Nombre de lettres isolées

* est considéré ici comme un mot, l'ensemble des graphèmes compris entre deux espaces

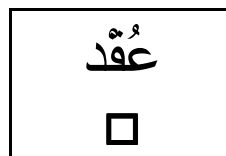
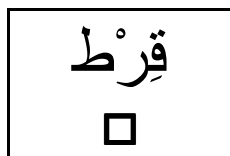
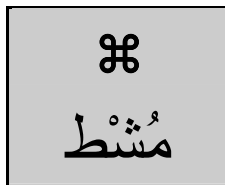
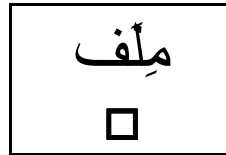
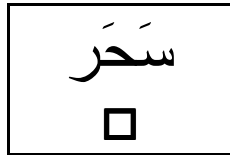
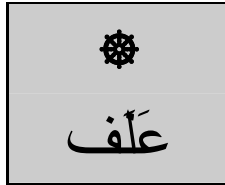
** sans compter la lettre finale

*** réponses moyennes /3 d'une vingtaine d'élèves de 3ème année à un test de familiarité lexicale des verbes et compléments (échelle en trois points)

Annexe 7

Epreuve explicite de détection d'intrus graphophonologique en arabe

Entraînements :





أَشْكَالٌ

عُصْفُورٌ



يُرْمِي



مُشَاطِبٌ

اِنْتَقَلَ



تَخَاطَبَ



اِحْسِنُ

اِحْسِنُ



اِجْعَلْ



😊
تَحْكِي

يَمْضِي
□

يَخْفَى
□

✂
مُؤَلِّف

اِخْتَلَفَ
□

تَخَاصَمَ
□

🕯
اِسْمَع

اِجْتَلَبَ
□

اِتَّبَعَ
□

ثرى

يَهَي

يَقْضِي

مُسَاف

اِبْطَح

تَحَاف

أَجْرَاس

بُرُوس

طُرُوش



تَدْعُو

يَخْشَى



يَمْحُو



مُهْرَج

إِنْدَرَجْ



تَنَاولْ



عَنْب

طُرُقْ



نُقَبْ



أَطْرُقْ

إِحْزِرْ



إِسْحَقْ



Item	Intrus	Traduction	C. G.	S. P.	N.S.	N.P.	N.G.	N.L.I. *	Lié	Traduction	C. G.	S. P.	N.S.	N.P.	N.G.	N.L.I. *	CIBLE	Traduction	C. G.	S. P.	N.S.	N.P.	N.G.	N.L.I. *	Unité phonologique	choix sémantique**
1	عُصْفُورٌ	oiseau	nom	CvCCVC	2	6	5	1	يُرْمِي	tonneau	nom	CvCCVC	2	6	5	1	أَشْكَالٌ	formes	nom	CvCCVC	2	6	5	2	C terminale	45%
2	إِنْتَقِلْ	passee	adjectif	CvCCVCvC	3	8	5	1	تَخَاطَبٌ	entretiens toi avec qq'un	adjectif	CvCVCvC	3	7	5	1	مَشْطَبٌ	rayé	adjectif	CvCvCCvC	3	8	4	0	syllabe	
3	إِجْعَلْ	deviens	verbe	CvCCVC	2	6	4	1	إِجْلِسْ	assied toi	verbe	CvCCVC	2	6	4	1	إِحْبِسْ	emprisonne	verbe	CvCCVC	2	6	4	1	rime	
4	يَحْفَى	cache	verbe	CvCCV	2	5	4	0	يَمْضِي	passer	verbe	CvCCV	2	5	4	0	تُحْكِي	raconter	verbe	CvCCV	2	5	4	0	V terminale	40%
5	تَخَاصَمْ	dispute toi	verbe	CvCVCvC	3	7	5	1	إِخْتَلَفْ	différencie toi	verbe	CvCCVCvC	3	8	5	1	مُؤَلِّفٌ	auteur compositeur	nom	CvCvCCvC	3	8	4	1	syllabe	
6	إِجْلِبْ	ramène	verbe	CvCCVC	2	6	4	1	إِنِّعْ	suis	verbe	CvCCVC	2	6	4	1	إِسْمَعْ	écoute	verbe	CvCCVC	2	6	4	1	rime	90%
7	يَفْضِي	il passe	verbe	CvCCV	2	5	4	0	يَنْهَى	interdire	verbe	CvCCV	2	5	4	0	تَرَى	voir	verbe	CvCV	2	4	3	1	V terminale	80%
8	إِنْبِطِحْ	couche toi à plat ventre	verbe	CvCCVCvC	3	8	5	1	تُحَالِفُ	allie toi avec qq'un	verbe	CvCVCvC	3	7	5	1	مُسَلِّفٌ	prêté, avancé	adjectif	CvCvCCvC	3	8	4	0	syllabe	
9	طُرْبُوشٌ	chapeau	nom	CvCCVC	2	6	5	2	بُرْنُوسٌ	burnous	nom	CvCCVC	2	6	5	2	أَجْرَاسٌ	sonnettes	nom	CvCCVC	2	6	5	3	C terminale	80%
10	يَحْشَى	craindre	verbe	CvCCV	2	5	4	0	يَمْحُو	effacer	verbe	CvCCV	2	5	4	0	تُدْعُو	inviter	verbe	CvCCV	2	5	4	1	V terminale	70%
11	تَتَأَوَّلُ	prendre	verbe	CvCVCvC	3	7	5	2	إِنْدَرَجْ	entre, inscris toi	verbe	CvCCVCvC	3	8	5	3	مُهْرَجٌ	clown	nom	CvCvCCvC	3	8	4	1	syllabe	
12	طُرُقٌ	chemins	nom	CvCvC	2	5	3	1	تُقَبُّ	trous	nom	CvCvC	2	5	3	0	عَنْبٌ	raisin	nom	CvCvC	2	5	3	0	rime	74%
13	إِخْرَازٌ	confisque	verbe	CvCCVC	2	6	4	1	إِسْحَاقٌ	écrase	verbe	CvCCVC	2	6	4	1	أَطْرُقُ	frappe à la porte	verbe	CvCCVC	2	6	4	2	C terminale	60%

C.G. = Catégorie Grammaticale ; S.P. = Structure Phonologique

Moyenne choix sémantique = 67 %

N.S. = Nombre de syllabes; N.P. = Nombre de phonèmes; N.G. = Nombre de graphèmes; N.L.I.= Nombre de lettres isolées

* sans compter la lettre finale

** % moyen de réponses de 20 élèves de 2ème année en faveur de l'intrus

Annexe 8

Epreuve explicite de détection d'intrus grapho-morpho-dérivationnel en arabe



اجْتَمَعَ

احْتَرَسَ



انْتَحَلَ



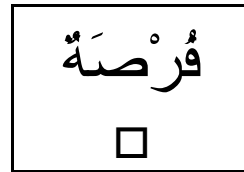
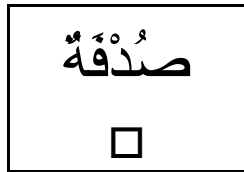
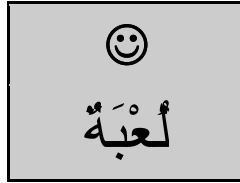
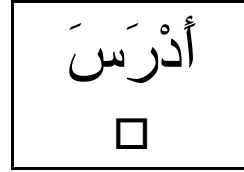
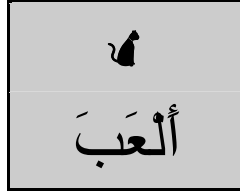
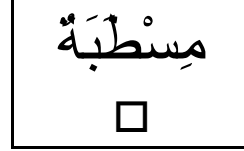
مَكْتَبٌ

مَرْقَأٌ



مَصْنَعٌ






تَاجِرَ

عَالِجَ

جَالِسَ

★
مِكَوَاهَ

مِجْرَقَةَ

مِنْضَدَةَ

→
حُقْنَةَ

أَسْرَةَ

شُعْلَةَ

أَكْثَرَ

أَبْغَى

أَلْقَى

عَادَلَ

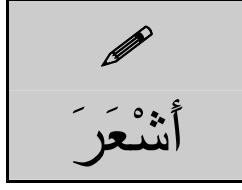
حَاوَرَ

طَالَبَ

أَلْفَهُ

سُرِعَهُ

نُرْهَهُ



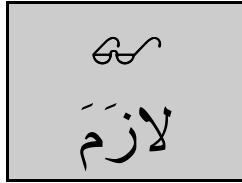
أندَرَ

أخرَجَ



مِشَارٌ

مِفْتَاحٌ



عَابِثٌ

شَاكِسٌ

Item	Intrus	Traduction	N.S.	N.P.	N.G.	N.L.I.*	Familiarité**	Liés	Traduction	N.S.	N.P.	N.G.	N.L.I.*	Familiarité**	Cible	Traduction	N.S.	N.P.	N.G.	N.L.I.*	Type de schème	choix sémantique* **
1	مِسْطَبَةٌ	estrade	4	10	5	0	0,500	مِكْنَسَةٌ	balai	4	10	5	0	1,692	مِرْآةٌ	miroir	3	8	4	2	Instrument	18%
2	أُدْرَكَ	Avoir raison	3	7	4	3	0,760	أُدْرَسَ	faire étudier	3	7	4	3	1,920	أَلْعَبَ	faire jouer	3	7	4	1	Forme IV	50%
3	فُرْصَةٌ	occasion opportunité	3	8	4	1	1,417	صُدْفَةٌ	hasard	3	8	4	1	0,931	لُعْبَةٌ	jouet	3	8	4	0	Substantif	68%
4	عَالَجٌ	soigner, guérir	3	6	4	1	1,560	جَالَسَ	s'asseoir avec qqun	3	6	4	1	1,200	تَأَجَرَ	commercer avec qqun	3	6	4	1	Forme III	36%
5	مِضْدَةٌ	petite table	4	10	4	1	1,440	مِجْرَقَةٌ	pelle	4	10	4	1	0,250	مِكَوَّاةٌ	fer à repasser	3	8	5	2	Instrument	45%
6	أُسْرَةٌ	famille	3	8	4	2	1,565	شُعْلَةٌ	flamme	3	8	4	0	0,652	حَقْنَةٌ	seringue	3	8	4	0	Substantif	73%
7	أَلْفَى	jeter	2	5	4	1	1,435	أَبَكَى	faire pleurer	2	5	4	1	1,087	أَكْتَرَّ	faire augmenter	3	7	4	1	Forme IV	45%
8	حَاوَرَ	discuter	3	6	4	2	0,609	طَالَبَ	demander à qqun	3	6	4	1	1,478	عَادَلَ	faire justice à qqun	3	6	4	2	Forme III	45%
9	نَزْهَةٌ	promenade	3	8	4	1	1,783	سُرْعَةٌ	vitesse	3	8	4	1	1,826	أَلْفَةٌ	familiarité, habitude	3	8	4	1	Substantif	27%
10	أُنْدَرَ	alarmer; alerter	3	7	4	1	0,783	أَخْرَجَ	faire sortir	3	7	4	2	1,783	أَشْعَرَ	faire sentir	3	7	4	1	Forme IV	55%
11	مِثْشَارٌ	scie	3	8	5	1	1,759	مِفْتَاحٌ	clé	3	8	5	1	1,793	مِعْلَاقٌ	porte manteau	3	8	5	1	Instrument	41%
12	شَاكَسَ	taquiner	3	6	4	1	0,520	عَابَثَ	taquiner qqun	3	6	4	1	0,360	لَازَمَ	s'attacher à qqun	3	6	4	2	Forme III	64%

Moyenne familiarité 14,129

Moyenne familiarité 14,973

Moyenne choix sémantique = 47 %

N.S. = Nombre de syllabes; N.P. = Nombre de phonèmes; N.G. = Nombre de graphèmes; N.L.I. = Nombre de lettres isolées

* sans compter la lettre finale

** réponses moyennes /3 d'une vingtaine d'élèves de 3ème année à un test de familiarité lexicale (échelle en trois points)

*** % moyen de réponses de 20 élèves de 2ème année en faveur de l'intrus

Annexe 9

Epreuve explicite de détection d'intrus grapho-morpho-flexionnel en arabe

Entraînements:

⊗
أُخْرَى

أَعْمَى
□

أُولَى
□

⌘
بَقِيَتْ

أَتَيْتُ
□

أَيُّتُ
□



بِرْغُوثَةٌ

خِنْزِيرَةٌ



حَمَامَةٌ



أَعْيَشُ

أَمَرْتُ



أَبْهَتُ



أَفْعَةٌ

دُبَابَةٌ



حِمَارَةٌ





أُدُورُ

أُنْعَتُ



أُمْنَتُ



دَيْنَصُورَةٌ

نَحْلَةٌ



ذَنْبَةٌ



أَنَالُ

أَصْمَتُ



أُمْنَتُ





قِطَّةٌ

تَمْسَاحَةٌ



سُحْقَاءُ



أَجِدُ

أَلْفِتُ



أَقْلِبُ



أَصِلُ

أَسْكُتُ



أَكَلْتُ



رَبَّةٌ

رَبَّةٌ

زَرَاقَةٌ

□

غَزَالَةٌ

□

عَجَلَةٌ

عَجَلَةٌ

فَيْلَةٌ

□

نَاقَةٌ

□

أُزُنٌ

أُزُنٌ

أُذُنٌ

□

أُفْتُ

□

Item	Intrus	Traduction	N.S.	N.P.	N.G.	N.L.I.*	Familiarité**	Lié	Traduction	N.S.	N.P.	N.G.	N.L.I.*	Familiarité**	Cible	Traduction	N.S.	N.P.	N.G.	N.L.I.*	Type de flexion	choix sémantique ***
1	حَمَامَةٌ	pigeon	4	9	5	1	0,868	خِزْيِرَةٌ	truie	4	10	6	2	0,395	بَرُّ عُوْتَةٌ	poux	4	10	6	2	Féminin	61%
2	أَمَرْتُ	j'ai ordonné	3	7	4	2	1,269	أُبْهَتُ	je deviens étonné	3	7	4	1	1,154	أُعِيشُ	je vis	3	6	4	1	Conjugaison verbale	30%
3	دُبَابَةٌ	mouche	4	9	5	2	0,158	حِمَارَةٌ	ânesse	4	9	5	2	0,474	أَفْعَةٌ	vipère	3	8	4	1	Féminin	22%
4	أَمْنْتُ	j'ai fais confiance	3	7	4	1	1,200	أَنْعَتُ	je décris	3	7	4	1	0,160	أُدُورُ	je tourne	3	6	4	2	Conjugaison verbale	74%
5	نَحْلَةٌ	abeille	3	8	4	0	0,079	ذَيْبَةٌ	louve	3	8	4	1	0,632	دَيْئُصُورَةٌ	dinausore	5	11	7	3	Féminin	96%
6	أَمْنْتُ	j'étais à l'abri	3	7	4	1	0,609	أَصْمْتُ	je garde le silence	3	7	4	1	1,565	أَنَالُ	j'obtiens	3	6	4	2	Conjugaison verbale	48%
7	سُلْحَفَاءُ	tortue	4	10	6	1	0,711	تِمْسَاحَةٌ	crocodile	4	10	6	1	0,789	قِطَّةٌ	chatte	3	8	3	0	Féminin	30%
8	أَلْفَتُ	j'ai adapté	3	7	4	1	0,586	أَقْلِتُ	j'échappe	3	7	4	1	0,931	أَجِدُ	je trouve	3	6	3	1	Conjugaison verbale	43%
9	أَكَلْتُ	j'ai mangé	3	7	4	1	1,955	أَسْكُنْتُ	je me tais	3	7	4	1	1,478	أَصِلُ	j'arrive	3	6	3	1	Conjugaison verbale	52%
10	زَرَافَةٌ	giraffe	4	9	5	3	1,500	عَزَالَةٌ	gazelle	4	9	5	2	1,658	دَيْبَةٌ	ours	3	8	3	1	Féminin	61%
11	نَاقَةٌ	chamelle	3	7	4	1		فَيْلَةٌ	éléphante	3	7	4	0	1,816	عِجْلَةٌ	bebis	3	8	4	0	Féminin	48%
12	أَذِنْتُ	j'ai permis	3	7	4	2	1,261	أَلْفَتُ	j'attire	3	7	4	1	1,091	أَزُنُ	je pese	3	6	3	2	Conjugaison verbale	22%

Moyenne familiarité 0,933

1,068

Moyenne choix sémantique = 49 %

N.S. = Nombre de syllabes; N.P. = Nombre de phonèmes; N.G. = Nombre de graphèmes; N.L.I. = Nombre de lettres isolées

* sans compter la lettre finale

** réponses moyennes /3 d'une vingtaine d'élèves de 3ème année à un test de familiarité lexicale (échelle en trois points).

Pour les items "féminin", le score à été calculé à partir de productions d'enfants de primaire à une tâche catégorielle (travaux non publiés de l'équipe EDIPSY)

Annexe 10

Epreuve explicite de détection d'intrus grapho-syntaxique en arabe

Entraînements:



يَعْسِلُ يَدَيْهِ

يَقْفِزُ كَفَّهْدِ



يَقْطِفُ كَرَزِ



يَطْهُو الْقَهْوَةَ

يَهْزُ كَنْفَيْهِ



يُعْنِي كَبْلُوبِ





سَلَفَ كِتَابَهُ

كَسَرَ بَلُورَيْنِ



ضَحَى بِخَرُوفَيْنِ



صَعَدَ السُّلَّمُ

حَكَى لِزَوْجَيْنِ



طَبَخَ لِقَتَيْنِ



عَدَّ دَرَّاجَتَيْنِ

دَهَنَ بِنَائِنِ



طَافَ بِمَنْزَلَيْنِ





دَخَنَ سِيَّجَارَةً

بَاعَ لِفَاقَتَيْنِ



حَكَمَ لِأَخْتَيْهِ



حَفَظَ الْوَلَدُ الْقُرْآنَ

فَرَحَ الطِّقْلُ بِكَلْبَيْنِ



حَفَرَ الرَّجُلُ بئرَيْنِ



تَحَمَّلَ أَخْطَاءَهُ

تَأَلَّمَ لِطِفْلَيْنِ



تَجَنَّبَ لِقَائَيْنِ





اِقْتَطَفَ البُسْتَانِيَّ وَرَدَةً

رَحَبَتْ الأُمُّ بَضِيْعَتَيْنِ



عَبَّأَتِ المَطْرُ بِرُكَّتَيْنِ



خَبَأَ الهَدِيَّةَ

اِخْتَارَ لِتَامِيْنَ



عَضِبَ لِقَوْمِيْنَ



أَطْلَقَ الشَّرْطِيَّ السَّحِيْبَةَ

شَعَرَ المَرِيضُ بِهَمِيْنِ



زَارَ المُسَافِرُ بِلَادِيْنَ





رَاجَعَ تَمَارِينَهُ

تَحَمَّلَ لِقَاحِينَ



فَرِحَ لِأَبُوئِهِ



عَبَدَ الْعَامِلُونَ الطَّرِيقَاتَ

تَبَيَّنَ الْمُرُوضُ بِطَرِيقَيْنِ



تَأَثَّرَ الْبَاحِثُ بِقَوْمَيْنِ



رَاقَصَ الرَّجُلُ الْحَبِيبَةَ

أَسِيفَ الْوَلَدُ لِقَطْنِهِ



قَطَعَ الْجَزَارُ لِسَانَيْنِ



Item	Intrus	Traduction	N.S. /mot*	N.G. /mot*	N.L.I. total**	Forme verbale	Familiarité***	Lié	Traduction	N.S. /mot*	N.G. /mot*	N.L.I. total**	Forme verbale	Familiarité***	Cible	Traduction	N.S. /mot*	N.G. /mot*	N.L.I. total**	Forme verbale	Préposition
1	ضَحَى بِخُرُوفَيْنِ	il a sacrifié deux moutons	2-5	3-7	2	II	1,336	كَسَرَ بِلُورَيْنِ	il a cassé deux vitres	3-4	3-6	2	II	1,480	سَلَفَ كِتَابَهُ	il a empreinté son livre	3-4	3-5	1	II	[bi]
2	حَكَى لِزَوْجَيْنِ	il a raconté à deux époux	2-4	3-6	2	I	1,672	طَبَخَ لِفَتْنَيْنِ	il a cuisiné deux navets	3-3	3-5	0	I	1,738	صَعِدَ السَّلْمَ	il a monté l'escalier	3-3	3-5	1	I	[li]
3	طَافَ بِمَنْزِلَيْنِ	il a tourné autour de deux maisons	2-5	3-7	2	I	1,340	دَهَنَ بِنَائِنِ	il a peint deux bâtiments	3-4	3-6	2	I	1,596	عَدَّ دَرَجَاتَيْنِ	il a compté deux bicyclettes	2-5	2-7	4	I	[bi]
4	حَكَّمَ لِأَخْتَيْهِ	il donnait raison à ses deux sœurs	3-4	3-6	1	I	1,722	بَاعَ لِفَأَقَتَيْنِ	il a vendu deux couvertures	2-5	3-7	2	I	1,182	دَخَّنَ سِجَارَةً	il a fumé une cigarette	3-4	3-6	3	II	[li]
5	فَرِحَ الطِّفْلُ بِكَلْبَيْنِ	l'enfant s'est réjoui de deux chiens	3-2-4	3-5-6	2	I	1,755	حَفَرَ الرَّجُلُ بِئْرَيْنِ	l'homme a creusé deux puits	3-3-3	3-5-5	3	I	1,704	حَفِظَ الْوَلَدُ الْقُرْآنَ	le garçon a appris le coran	3-3-3	3-5-6	5	I	[bi]
6	تَأَلَّمَ لِطِفْلَيْنِ	il a compati à deux garçons	4-4	4-6	1	V	1,828	تَجَنَّبَ لِغَائِنِ	il a évité deux rencontres, au revoir	4-4	4-6	1	V	0,897	تُحَمَّلَ أَخْطَاؤُهُ	il a assumé ses fautes	4-4	4-6	2	V	[li]
7	رَحَّبَتِ الْأُمُّ بِضَيْفَتَيْنِ	la mère a accueilli deux invitées	3-3-5	4-4-7	3	II	1,457	عَبَّأَتِ الْمَطْرُ بِرَكَائِنِ	la pluie a rempli deux étangs	3-4-4	4-5-6	3	II	0,935	إِقْتَطَفَ الْبُسْتَانِيَّ وَرْدَةً	le jardinier a cueilli une rose	4-4-3	5-8-4	6	VIII	[bi]
8	غَضِبَ لِقَوْمَيْنِ	il a soutenu deux peuples	3-4	3-6	1	I	1,065	إِخْتَارَ لِثَامَيْنِ	il a choisi deux voiles	3-4	5-6	3	VIII	1,000	حَبَا الْهَدِيَّةَ	il a caché le cadeau	3-4	3-6	2	II	[li]
9	شَعَرَ الْمَرِيضُ بِهَمَيْنِ	le malade a ressenti deux chagrins, soucis	3-3-4	3-6-5	2	I	1,087	زَارَ الْمُسَافِرُ بِلَادَيْنِ	le voyageur a visité deux pays	2-4-4	3-7-6	5	I	1,955	أَطْلَقَ الشَّرْطِيَّ السَّجِينَةَ	le policier a libéré la prisonnière	3-3-4	4-6-7	4	IV	[bi]
10	فَرِحَ لِأَبْوَيْهِ	il s'est réjoui pour ses parents	3-5	3-6	3	I	1,804	تَحَمَّلَ لِغَائِنِ	il a supporté deux vaccins	4-4	4-6	1	V	1,221	رَاجَعَ تِمَارِيئَهُ	il a révisé ses exercices	3-5	4-7	4	IV	[li]
11	تَأَثَّرَ الْبَاحِثُ بِقَوْمَيْنِ	le chercheur a été affecté par deux peuples	4-3-4	4-6-6	4	V	1,069	تَبَيَّنَ الْمَرُوضُ بِطَرِيقَيْنِ	le dompteur a adopté deux pingouins	3-4-4	4-6-7	4	V	1,026	عَبَّدَ الْعَامِلُونَ الطَّرِيقَاتَ	les ouvriers ont goudronné les rues	3-4-4	3-8-7	6	II	[bi]
12	أَسِيفَ الْوَلَدُ لِطَيْئِهِ	le garçon a regretté ses deux chats	3-3-4	3-5-5	3	I	1,583	قَطَعَ الْجَزَارُ لِلسَّانِنِ	le boucher a découpé deux langues	3-3-4	3-6-6	4	X	1,695	رَاقَصَ الرَّجُلُ الْحَبِيبَةَ	l'homme a dansé avec la bien-aimée	3-3-4	4-5-7	5	IV	[li]

Moyenne familiarité 1,476

1,369

N.S. = Nombre de syllabes; N.P. = Nombre de phonèmes; N.G. = Nombre de graphèmes; N.L.I. = Nombre de lettres isolées

* est considéré ici comme un mot, l'ensemble des graphèmes compris entre deux espaces

** sans compter la lettre finale

*** réponses moyennes /3 d'une vingtaine d'élèves de 3ème année à un test de familiarité lexicale des verbes et compléments (échelle en trois points)

Annexe 11

Epreuve d'identification de la construction morphologique de pseudo-mots
en arabe

Entraînements:

بالبُغَار

«بَقِيَّ أَمَامَ شَيْءٍ مَا» نَقُولُ "BURIKIAN" «بَقِيَّ وَرَاءَ شَيْءٍ مَا» نَقُولُ "BUMEKIAN"	«بَقِيَّ» نَقُولُ "BUKIAN"
«وَضَعَ أَمَامَ شَيْءٍ مَا» نَقُولُ "KARIDIAN" «وَضَعَ وَرَاءَ شَيْءٍ مَا» نَقُولُ "KAMEDIAN"	«وَضَعَ» نَقُولُ "KADIAN"
؟	«وَقَفَّ» نَقُولُ "CESIAN"

كَيْفَ نَقُولُ «وَقَفَّ أَمَامَ شَيْءٍ مَا» بَالْبُغَارِ ؟

CEMESIAN	CERIDIAN	CERISIAN
----------	----------	----------

بِالْفَرَبَّةِ

«قَادَ قَلِيلاً» نَقُولُ "UOSHITA" «قَادَ بِسُرْعَةٍ» نَقُولُ "UOGUTA"	«قَادَ» نَقُولُ "UOTA"
«لَعِبَ قَلِيلاً» نَقُولُ "KAISHIKI" «لَعِبَ بِسُرْعَةٍ» نَقُولُ "KAIGUKI"	«لَعِبَ» نَقُولُ "KAIKI"
؟	«كَبُرَ» نَقُولُ "JEYFO"

كَيْفَ نَقُولُ «كَبُرَ قَلِيلاً» بِالْفَرَبَّةِ ؟

JEYSHIFO	JEYGUFO	JEYSHITO
----------	---------	----------

1. بالكسبة

« ثَمَانِيَةٌ أَسْمَاكٍ » نَقُولُ "XODIKE" « سِتَّةُ أَسْمَاكٍ » نَقُولُ "XODUKO"	« سَمَكَةٌ » نَقُولُ "XOD"
« ثَمَانِيَةٌ فُرُودٍ » نَقُولُ "GOFIKE" « سِتَّةُ فُرُودٍ » نَقُولُ "GOFUKO"	« فِرْدٌ » نَقُولُ "GOF"
؟	« كَلْبٌ » نَقُولُ "RUV"

كَيْفَ نَقُولُ « سِتَّةُ كِلَابٍ » بِالْكَسْبَةِ ؟

RUVUKO	RUFIKE	RUVIKE
--------	--------	--------

2. بالدندز

« " كَأْسٌ "، عِنْدَمَا يَكُونُ هُنَاكَ وَاحِدٌ فَقَطْ » نَقُولُ "TIXKAL" « " كَأْسٌ "، عِنْدَمَا يَكُونُ هُنَاكَ أَلْعَدِيدُ » نَقُولُ "VEXKAL"	« كَأْسٌ » نَقُولُ "XKAL"
« " كِتَابٌ "، عِنْدَمَا يَكُونُ هُنَاكَ وَاحِدٌ فَقَطْ » نَقُولُ "TIBLOMST" « " كِتَابٌ "، عِنْدَمَا يَكُونُ هُنَاكَ أَلْعَدِيدُ » نَقُولُ "VEBLOMST"	« كِتَابٌ » نَقُولُ "BLOMST"
؟	« رِحْلَةٌ » نَقُولُ "REPS"

كَيْفَ نَقُولُ « " رِحْلَةٌ "، عِنْدَمَا يَكُونُ هُنَاكَ أَلْعَدِيدُ » بِالْدَنْدُزِ ؟

TIREPS	VEREPS	TISEPS
--------	--------	--------

3. بالككننز

« نَامَ بِمَشَقَّةٍ » نَقُولُ "VOLRIT" « نَامَ كَثِيرًا » نَقُولُ "VALROT"	« نَامَ » نَقُولُ "VLRT"
« أَكَلَ بِمَشَقَّةٍ » نَقُولُ "ZOGZIT" « أَكَلَ كَثِيرًا » نَقُولُ "ZAGZOT"	« أَكَلَ » نَقُولُ "ZGZT"
؟	« حَدَّثَ » نَقُولُ "KPJT"

كَيْفَ نَقُولُ « حَدَّثَ كَثِيرًا » بِالْكَكْنُزِ ؟

KAPJOT	KOPJIT	KAPROT
--------	--------	--------

4. بالزوزنز

« مَعَ الْآبِ » نَقُولُ "AITARENKIN" « لِلْآبِ » نَقُولُ "AITASETZAT"	« الْآبُ » نَقُولُ "AITA"
« مَعَ الْأَخِ » نَقُولُ "ZUMURENKIN" « لِلْأَخِ » نَقُولُ "ZUMUSETZAT"	« الْأَخُ » نَقُولُ "ZUMU"
؟	« الْقِطُّ » نَقُولُ "NEBO"

كَيْفَ نَقُولُ « مَعَ الْقِطِّ » بِالزُّوْنِزِ ؟

NEBOSETZAT	NEBURENKIN	NEBORENKIN
------------	------------	------------

5. بِالْبِرْكَلِ

« أَصْبَحَ قَوِيًّا » نَقُولُ "UNHARBOR" « بَقِيَ قَوِيًّا » نَقُولُ "GEHARBOR"	« قَوِيٌّ » نَقُولُ "HARBOR"
« أَصْبَحَ أَحْمَرًا » نَقُولُ "UNSITBOR" « بَقِيَ أَحْمَرًا » نَقُولُ "GESITBOR"	« أَحْمَرٌ » نَقُولُ "SITBOR"
؟	« جَمِيلٌ » نَقُولُ "RECHLICH"

كَيْفَ نَقُولُ « أَصْبَحَ جَمِيلًا » بِالْبِرْكَلِ ؟

GERECHLICH	UNRECHLICH	GESECHLICH
------------	------------	------------

6. بِالسَّوْمَدِ

« فَعَلَ مَرَّةً أُخْرَى » نَقُولُ "KILIVAKUMI" « فَعَلَ كَثِيرًا » نَقُولُ "KILIVASITA"	« فَعَلَ » نَقُولُ "KILIVA"
« طَارَ مَرَّةً أُخْرَى » نَقُولُ "MAKIKAKUMI" « طَارَ كَثِيرًا » نَقُولُ "MAKIKASITA"	« طَارَ » نَقُولُ "MAKIKA"
؟	« خَاطَ » نَقُولُ "FABASI"

كَيْفَ نَقُولُ « خَاطَ مَرَّةً أُخْرَى » بِالسَّوْمَدِ ؟

FABASASITA	FABASIKUMI	FABASISITA
------------	------------	------------

7. بِالْقَمَكِ

« أَنْتَ بَصَدَدِ الْقَطْعِ » نَقُولُ "KEXYT" « إِنْتَهَيْتَ مِنَ الْقَطْعِ » نَقُولُ "KAXAT"	« الْقَطْعُ » نَقُولُ "KXT"
« أَنْتَ بَصَدَدِ الصُّعُودِ » نَقُولُ "PEMYD" « إِنْتَهَيْتَ مِنَ الصُّعُودِ » نَقُولُ "PAMAD"	« الصُّعُودُ » نَقُولُ "PMD"
؟	« الرَّسْمُ » نَقُولُ "DVB"

كَيْفَ نَقُولُ « أَنْتَ بَصَدَدِ الرَّسْمِ » بِالْقَمَكِ ؟

DEVYB	DAMAB	DAVAB
-------	-------	-------

8. بِالْأَكْرَاوِمِ

« لَا أَحِبُّ » نَقُولُ "CHESIROEM" « أَحِبُّ » نَقُولُ "HASIROEM"	« أَحَبَّ » نَقُولُ "SIROEM"
« لَا أَتَكَلَّمُ » نَقُولُ "CHERANOEM" « أَتَكَلَّمُ » نَقُولُ "HARANOEM"	« تَكَلَّمَ » نَقُولُ "RANOEM"
؟	« فَهِمَّ » نَقُولُ "POSOEM"

كَيْفَ نَقُولُ « لَا أَفْهِمُ » بِالْأَكْرَاوِمِ ؟

HAPOSOEM	HAROSOEM	CHEPOSOEM
----------	----------	-----------

9. بِالنُّوْفِ

« مَا لَا نَسْتَطِيعُ فِعْلَهُ » نَقُولُ "NEMIHUA" « مَا نَسْتَطِيعُ فِعْلَهُ » نَقُولُ "NEMIZTLI"	« فَعَلَ » نَقُولُ "NEMI"
« مَا لَا نَسْتَطِيعُ عِلْمَهُ » نَقُولُ "POXUHUA" « مَا نَسْتَطِيعُ عِلْمَهُ » نَقُولُ "POXUZTLI"	« عِلَّمَ » نَقُولُ "POXU"
؟	« طَلَبَ » نَقُولُ "CEPA"

كَيْفَ نَقُولُ « مَا لَا نَسْتَطِيعُ طَلَبَهُ » بِالنُّوْفِ ؟

CEPAZTLI	CEPIHUA	CEPAHUA
----------	---------	---------

10. بالزئك

« عَكْسُ الشِّرَاءِ » نَقُولُ "DOKAS" « طَرِيقَةُ الشِّرَاءِ » نَقُولُ "DUKIS"	« الشِّرَاءُ » نَقُولُ "DKS"
« عَكْسُ الصُّعُودِ » نَقُولُ "KOGAL" « طَرِيقَةُ الصُّعُودِ » نَقُولُ "KUGIL"	« الصُّعُودُ » نَقُولُ "KGL"
؟	« التَّحْطِيمُ » نَقُولُ "ZTB"

كَيْفَ نَقُولُ « طَرِيقَةُ التَّحْطِيمِ » بِالزَّئِكِ ؟

ZOTAB	ZUKIB	ZUTIB
-------	-------	-------

11. بالبُولِيرِ

« الَّذِي حَفِظَ هُوَ أَنَا » نَقُولُ "MIDOJANNGA" « الَّذِي حَفِظَ هُوَ أَنْتَ » نَقُولُ "ADAJANNGA"	« حَفِظَ » نَقُولُ "JANNGA"
« الَّذِي كَتَبَ هُوَ أَنَا » نَقُولُ "MIDOKINNDU" « الَّذِي كَتَبَ هُوَ أَنْتَ » نَقُولُ "ADAKINNDU"	« كَتَبَ » نَقُولُ "KINNDU"
؟	« عَمِلَ » نَقُولُ "HAALDE"

كَيْفَ نَقُولُ « الَّذِي عَمِلَ هُوَ أَنَا » بِالْبُولِيرِ ؟

ADAHAALDE	MIDOKAALDE	MIDOHAALDE
-----------	------------	------------

12. بالِقَافَانَازِ

« يَكْفُ عَنِ الْقَوْلِ » نَقُولُ "VERZEGEN" « قَالَ مَرَّةً أُخْرَى » نَقُولُ "HERZEGEN"	« قَالَ » نَقُولُ "ZEGEN"
« يَكْفُ عَنِ الْأَخْذِ » نَقُولُ "VERVATTEN" « أَخَذَ مَرَّةً أُخْرَى » نَقُولُ "HERVATTEN"	« أَخَذَ » نَقُولُ "VATTEN"
؟	« سَقَطَ » نَقُولُ "GEDEN"

كَيْفَ نَقُولُ « سَقَطَ مَرَّةً أُخْرَى » بِالِقَافَانَازِ ؟

VERGEDEN	VERZEDEN	HERGEDEN
----------	----------	----------

13. بِالْتَنْفِينِمْ

« الَّذِي حَمَلَ هُوَ أَنَا » نَقُولُ "ZELAT" « الَّذِي حَمَلَ هُوَ أَنْتَ » نَقُولُ "ZOLUT"	« حَمَلَ » نَقُولُ "ZLT"
« الَّذِي تَقَدَّمَ هُوَ أَنَا » نَقُولُ "FEVAP" « الَّذِي تَقَدَّمَ هُوَ أَنْتَ » نَقُولُ "FOVUP"	« تَقَدَّمَ » نَقُولُ "FVP"
؟	« بَدَأَ » نَقُولُ "BRN"

كَيْفَ نَقُولُ « الَّذِي بَدَأَ هُوَ أَنْتَ » بِالْتَنْفِينِمْ ؟

BORUN	BERAN	BOLUN
-------	-------	-------

14. بِالطَّبُولُونُو

« الَّذِي يَمُرُّ هُوَ أَنْتَ » نَقُولُ "MITINI" « الَّذِي يَمُرُّ هُوَ أَنَا » نَقُولُ "MITARI"	« يَمُرُّ » نَقُولُ "MIT"
« الَّذِي يَسْفُطُ هُوَ أَنْتَ » نَقُولُ "CHIRINI" « الَّذِي يَسْفُطُ هُوَ أَنَا » نَقُولُ "CHIRARI"	« يَسْفُطُ » نَقُولُ "CHIR"
؟	« يَطْلُبُ » نَقُولُ "RAV"

كَيْفَ نَقُولُ « الَّذِي يَطْلُبُ هُوَ أَنَا » بِالطَّبُولُونُو ؟

RAVINI	RAVARI	RATINI
--------	--------	--------

15. بِالتَّوَلِيَّتِي

« صَانِعُ الْكُتُبِ » نَقُولُ "KNIHONATCHO" « مَعَارَةُ الْكُتُبِ » نَقُولُ "KNIHORICTVO"	« كُتِبَ » نَقُولُ "KNIHO"
« صَانِعُ الْأَحْدِيَةِ » نَقُولُ "OBUVNATCHO" « مَعَارَةُ الْأَحْدِيَةِ » نَقُولُ "OBUVRICTVO"	« أَحْدِيَةٌ » نَقُولُ "OBUV"
؟	« أَثَاتٌ » نَقُولُ "NABYTOK"

كَيْفَ نَقُولُ « مَعَارَةُ الْأَثَاتِ » بِالتَّوَلِيَّتِي ؟

NABYTOVNATCHO	NABYTOKNATCHO	NABYTOKRICTVO
---------------	---------------	---------------

16. بالأفيون

« "غريب" ، عِنْدَمَا نَتَحَدَّثُ عَنْ وَالدِ » نَقُولُ "SITAKOT" « "غريب" ، عِنْدَمَا نَتَحَدَّثُ عَنْ بِنْتِ » نَقُولُ "SUTIKAT"	« غريب » نَقُولُ "STKT"
« "سريع" ، عِنْدَمَا نَتَحَدَّثُ عَنْ وَالدِ » نَقُولُ "NITASOJ" « "سريع" ، عِنْدَمَا نَتَحَدَّثُ عَنْ بِنْتِ » نَقُولُ "NUTISAJ"	« سريع » نَقُولُ "NTSJ"
؟	« نظيف » نَقُولُ "WKBZ"

كَيْفَ نَقُولُ « "نظيف" ، عِنْدَمَا نَتَحَدَّثُ عَنْ بِنْتِ » بِالْأَفْيُونِ ؟

WIKAKOZ	WUKIBAZ	WIKABOZ
---------	---------	---------

17. بالتسيج

« الَّذِي يَهْتَمُّ بِالْأَسْنَانِ » نَقُولُ "ZEPON" « سِنٌ قَبِيحَةٌ » نَقُولُ "ZUPIN"	« سِنٌ » نَقُولُ "ZPN"
« الَّذِي يَهْتَمُّ بِالرِّجْلِ » نَقُولُ "LETOK" « رَجُلٌ قَبِيحَةٌ » نَقُولُ "LUTIK"	« رَجُلٌ » نَقُولُ "LTK"
؟	« غَابَةٌ » نَقُولُ "DRV"

كَيْفَ نَقُولُ « الَّذِي يَهْتَمُّ بِالْغَابَاتِ » بِالتَّسِيجِ ؟

DURIV	DEROV	DETOV
-------	-------	-------

18. بالقنكن

« "مُضْحِكٌ" ، عِنْدَمَا نَتَحَدَّثُ عَنْ وَالدِ » نَقُولُ "JEJVEGAREK" « "مُضْحِكٌ" ، عِنْدَمَا نَتَحَدَّثُ عَنْ بِنْتِ » نَقُولُ "SWOVEGAREK"	« مُضْحِكٌ » نَقُولُ "VEGAREK"
« "مَرِيضٌ" ، عِنْدَمَا نَتَحَدَّثُ عَنْ وَالدِ » نَقُولُ "JEJKOSZULA" « "مَرِيضٌ" ، عِنْدَمَا نَتَحَدَّثُ عَنْ بِنْتِ » نَقُولُ "SWOKOSZULA"	« مَرِيضٌ » نَقُولُ "KOSZULA"
؟	« فَقِيرٌ » نَقُولُ "SPODNIÉ"

كَيْفَ نَقُولُ « "فَقِيرٌ" ، عِنْدَمَا نَتَحَدَّثُ عَنْ بِنْتِ » بِالْقَنْكَنِ ؟

SWOSPODNIÉ	JEJSPODNIÉ	SWOVPODNIÉ
------------	------------	------------

19. بِالْفُورِكِ

« مَنْ يَسْكُنُ الْمَدِينَةَ » نَقُولُ "RAZDEKAT" « مَنْ يُحِبُّ الْمَدِينَةَ » نَقُولُ "PAJDEKAT"	« الْمَدِينَةُ » نَقُولُ "DEKAT"
« مَنْ يَسْكُنُ الْبَحْرَ » نَقُولُ "RAZVOROVAT" « مَنْ يُحِبُّ الْبَحْرَ » نَقُولُ "PAJVOROVAT"	« الْبَحْرُ » نَقُولُ "VOROVAT"
؟	« الرَّيْفُ » نَقُولُ "SHITZAT"

كَيْفَ نَقُولُ « مَنْ يَسْكُنُ الرَّيْفَ » بِالْفُورِكِ ؟

RAZVOSHITZAT	RAZSHITZAT	PAJSHITZAT
--------------	------------	------------

20. بِالْمَدْرِزِ

« عَلَى السَّرِيرِ » نَقُولُ "BRYTER" « فِي السَّرِيرِ » نَقُولُ "BRETUR"	« السَّرِيرُ » نَقُولُ "BRTR"
« عَلَى الْحَقِيبَةِ » نَقُولُ "KNYPER" « فِي الْحَقِيبَةِ » نَقُولُ "KNEPUR"	« الْحَقِيبَةُ » نَقُولُ "KNPR"
؟	« الْخِرَانَةُ » نَقُولُ "SRKR"

كَيْفَ نَقُولُ « عَلَى الْخِرَانَةِ » بِالْمَدْرِزِ ؟

SRYTER	SREKUR	SRYKER
--------	--------	--------

21. بِالذُّونِلِ

« أَنْتَهَى مِنَ الشِّرَاءِ » نَقُولُ "PYALATETE" « بَصَدَدِ الشِّرَاءِ » نَقُولُ "PYALAKIRA"	« الشِّرَاءُ » نَقُولُ "PYALA"
« أَنْتَهَى مِنَ الْإِعْلَانِ » نَقُولُ "FRANUTETE" « بَصَدَدِ الْإِعْلَانِ » نَقُولُ "FRANUKIRA"	« الْإِعْلَانُ » نَقُولُ "FRANU"
؟	« الْمُسَاعَدَةُ » نَقُولُ "KWANI"

كَيْفَ نَقُولُ « أَنْتَهَى مِنَ الْمُسَاعَدَةِ » بِالذُّونِلِ ؟

KWANITETE	KWANAKIRA	KWANIKIRA
-----------	-----------	-----------

22. بِالْقَمْبَرَا

« رَكَضَ كَثِيرًا » نَقُولُ "MOXIMOA" « عَكْسُ رَكَضَ » نَقُولُ "TIXIMOA"	« رَكَضَ » نَقُولُ "XIMOA"
« أُعْطِيَ كَثِيرًا » نَقُولُ "MOPECHIA" « عَكْسُ أُعْطِيَ » نَقُولُ "TIPECHIA"	« أُعْطِيَ » نَقُولُ "PECHIA"
؟	« دَخَلَ » نَقُولُ "LOTIA"

كَيْفَ نَقُولُ « عَكْسُ دَخَلَ » بِالْقَمْبَرَا ؟

TIXOTIA	TILOTIA	MOLOTIA
---------	---------	---------

23. بِالسَّوْفِرْس

« "يَوْمٌ" ، عِنْدَمَا يَكُونُ هُنَاكَ وَاحِدٌ فَقَطْ » نَقُولُ "DAGEN" « "يَوْمٌ" ، عِنْدَمَا يَكُونُ هُنَاكَ أَلْعَدِيدُ » نَقُولُ "DAGNA"	« "يَوْمٌ" نَقُولُ "DAG"
« "حَقْلٌ" ، عِنْدَمَا يَكُونُ هُنَاكَ وَاحِدٌ فَقَطْ » نَقُولُ "DORDEN" « "حَقْلٌ" ، عِنْدَمَا يَكُونُ هُنَاكَ أَلْعَدِيدُ » نَقُولُ "DORDNA"	« "حَقْلٌ" نَقُولُ "DORD"
؟	« إِمْرَأَةٌ » نَقُولُ "KVIS"

كَيْفَ نَقُولُ « "إِمْرَأَةٌ" ، عِنْدَمَا يَكُونُ هُنَاكَ أَلْعَدِيدُ » بِالسَّوْفِرْس ؟

KVIDNA	KVISEN	KVISNA
--------	--------	--------

24. بِالْبَيْلُوتِ

« قِطَّةٌ صَغِيرَةٌ » نَقُولُ "XAZIPAT" « قِطَّةٌ كَبِيرَةٌ » نَقُولُ "XUZOPIT"	« قِطَّةٌ » نَقُولُ "XZPT"
« شَجَرَةٌ صَغِيرَةٌ » نَقُولُ "GAFITAV" « شَجَرَةٌ كَبِيرَةٌ » نَقُولُ "GUFOTIV"	« شَجَرَةٌ » نَقُولُ "GFTV"
؟	« دَرَاجَةٌ » نَقُولُ "MVLB"

كَيْفَ نَقُولُ « دَرَاجَةٌ صَغِيرَةٌ » بِالْبَيْلُوتِ ؟

MAVILAB	MUVOTAB	MUVOLIB
---------	---------	---------

Annexe 12

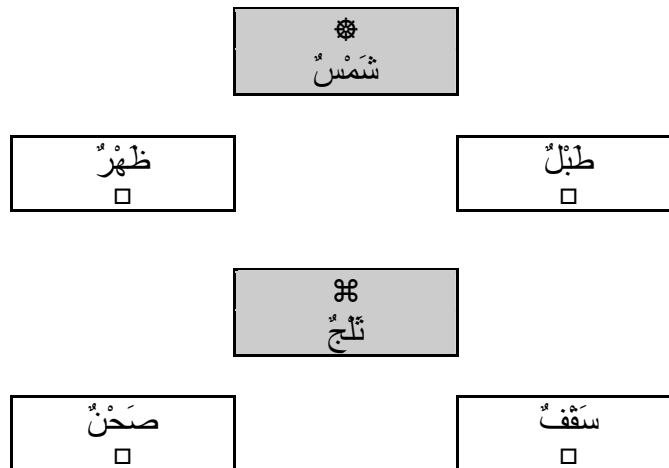
Exemple donné à l'épreuve implicite graphophonologique



Exemple donné aux épreuves implicites grapho-morphologiques



Items d'entraînement graphophonologiques et grapho-morpho-dérivationnels implicites (dans les items de l'épreuve graphophonologique les « tanwin » sont absents, comme dans le reste de l'épreuve)



Items d'entraînement grapho-morpho-flexionnels implicites

سَكَنَ

حَلَمَ

عَطِشَ

تُعَبِّانُ

أَسَدَّ

فَهَّدَ

Exemple donné à l'épreuve implicite grapho-syntaxique

أَطْهَوْا كَعَكَةَ

أَرْتَبُ دَفَاتِرِي

أَقْلَمُ أَظْفِرِي

Items d'entraînement grapho-syntaxiques implicites

حَرَّتْ الْفَلَّاحُ الْأَرْضَ

دَخَلَ الْبَائِعُ الدَّارَ

فَتَحَ الْخَادِمُ الْبَابَ

وَجَدَ الطَّبِيبَ

اِشْتَرَى كُرَّةً

بَاعَ سَيَّارَةً

Annexe 13

Résultats aux tests t (niveau du hasard/6) obtenus aux épreuves implicites de détection d'intrus grapho-morpho-dérivationnel, grapho-morpho-flexionnel et grapho-syntaxique en arabe en fonction du niveau scolaire

Epreuves de détection d'intrus	3 ^e année	4 ^e année	5 ^e année
Implicite grapho-morpho-dérivationnel	5,82 (1,61) t(55)>-1	6,02 (1,72) t (40)<1	5,80 (1,46) t(49)>-1
Implicite grapho-morpho-flexionnel	5,59 (1,97) t(55)=1,56 ns	6,22 (1,29) t(40)=1,09 ns	5,50 (1,79) t(49)=-1,98 ns
Implicite grapho-syntaxique	5,93 (1,41) t(55)>-1	6,59 (1,77) t(40)=2,11p<.05	6,08 (1,60) t(49)<1

Annexe 14

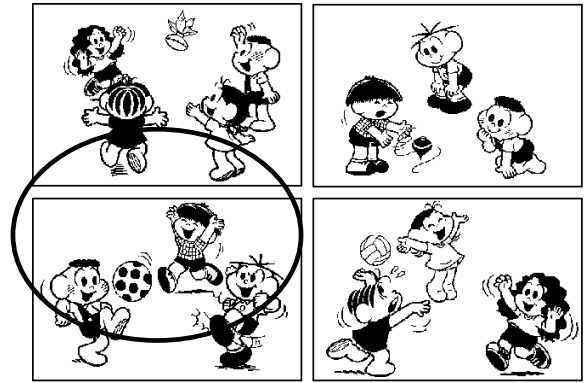
Epreuve de lecture de mots en une minute en portugais

pato	poupa	durex
mato	armadura	milho
nova	moeda	explicação
vela	outono	nascimento
fita	bandeja	aglomerar
loja	palavra	sucesso
medo	chapéu	gramaram
nata	aplicado	rapidez
taco	trevo	bicho
lobo	crê	luxuoso
janela	floresta	rescindido
pipa	globo	pouco
minha	castanha	lençois
saco	projeto	aeronáutica
pacote	atlas	trânsito
garra	ficha	quiosque
caju	acanhado	repugnante
pilha	querido	trajeto
sapato	guitarra	isqueiro
osso	muito	hipócrita
bonito	campo	interpretar
agulha	bruto	advogado
caminhão	concretizar	perseverança
ouvido	tempestade	encasquetar
agora	pingado	atmosfera
tomate	plebéia	coalhada
tijolo	exausto	bravamente
caminho	abusar	marsupiais
acordar	cheiro	vangloriar
costas	garagem	estraçalhar
cozinha	hospedaria	acabrunhado
tamanho	retribuem	excepcional
mel	trouxe	preservada
bicada	azedo	ricochetear
arte	fixa	saguões
isca	chocalho	reverendíssimo

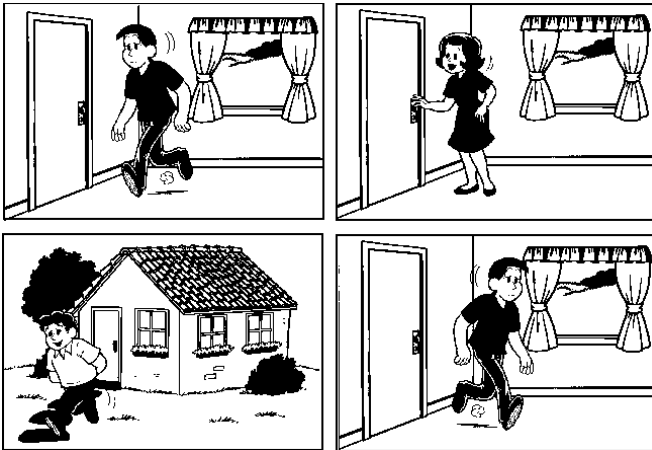
Compreensão de frases escritas em português (Zanella, 2007)

Como no modelo ao lado, apenas uma figura representa melhor a frase escrita abaixo delas. Circule a figura correta.

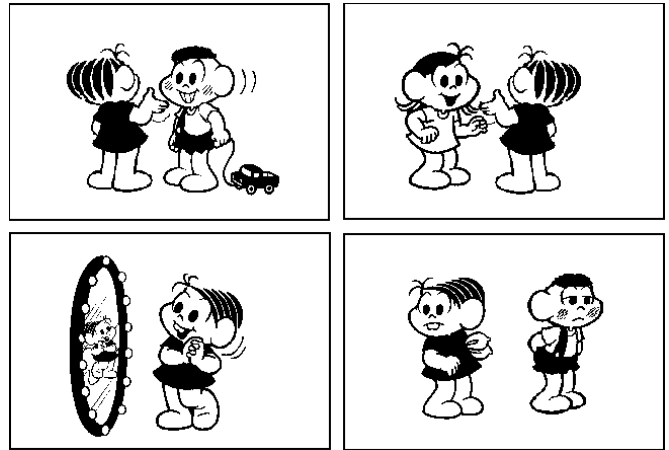
Trabalhe o mais rápido possível.



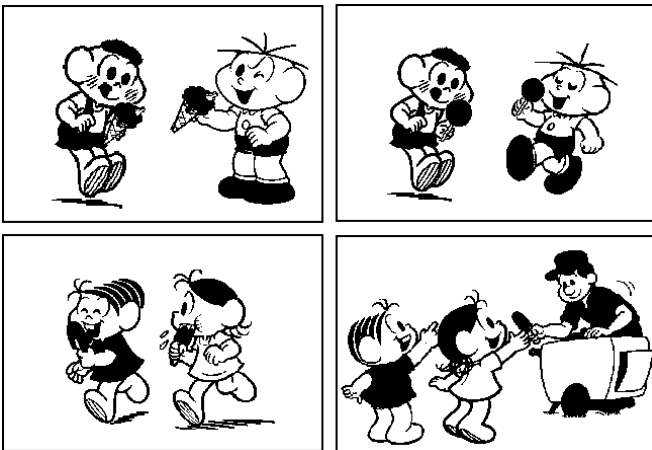
Os meninos jogam bola.



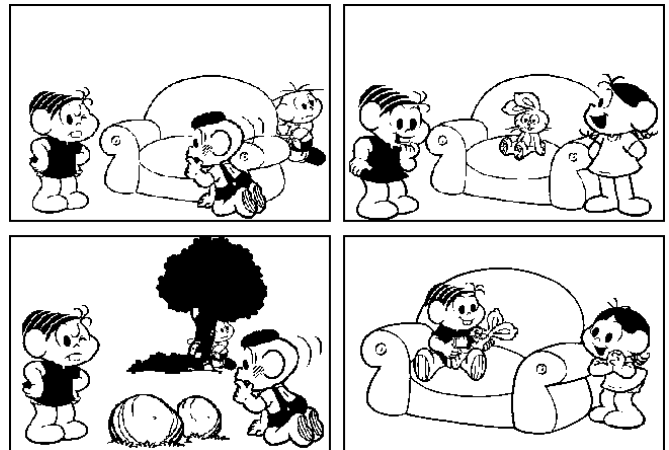
O senhor Pedro vai sair.



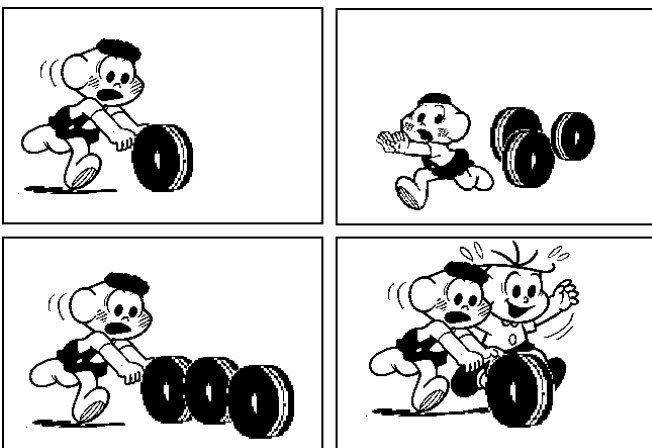
A menina olha para ele.



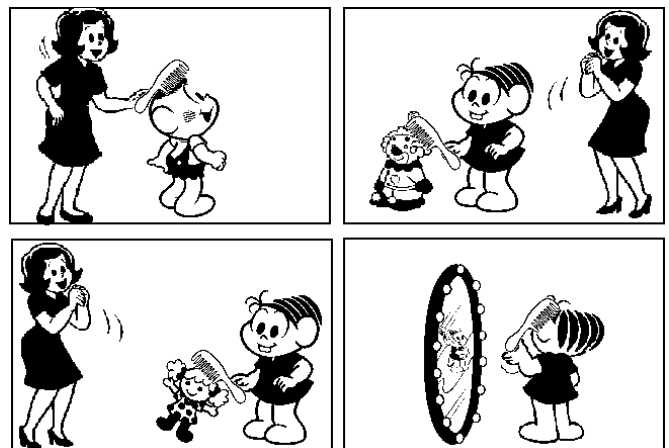
Elas chuparão sorvetes.



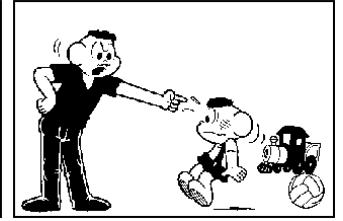
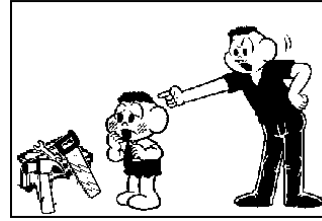
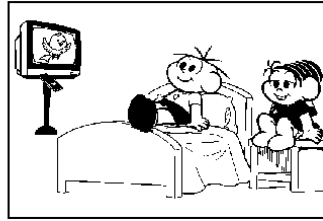
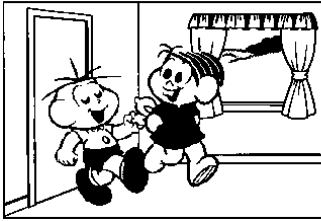
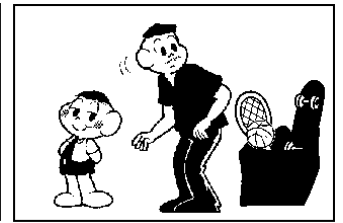
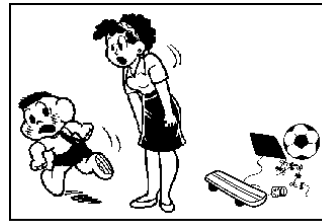
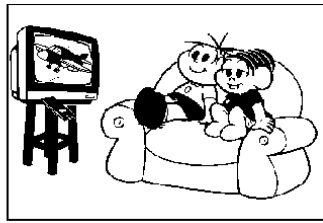
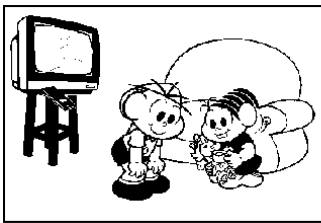
-Mas, onde está o coelhinho que eu deixei no sofá?



Os pneus são empurrados pelo menino.

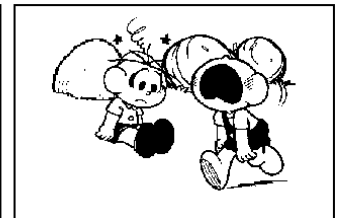
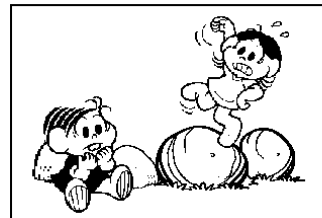
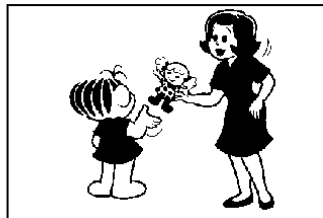
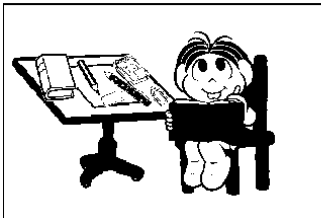
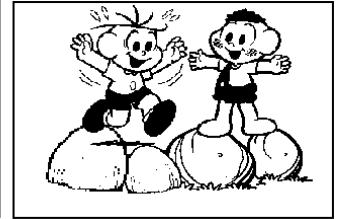
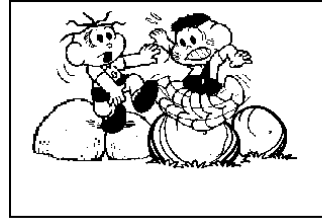
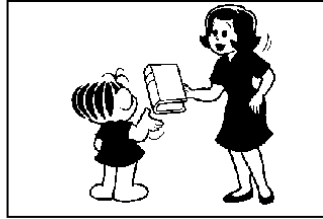
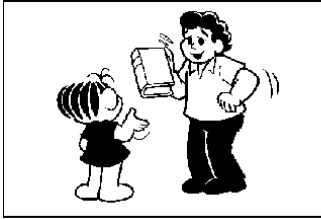


Mamãe vê a Mônica penteando os cabelos dele.



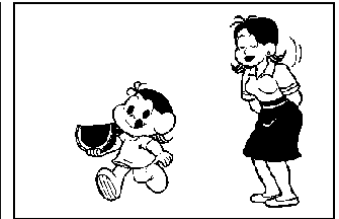
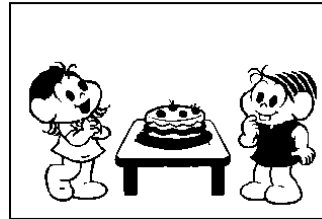
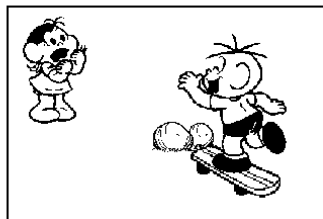
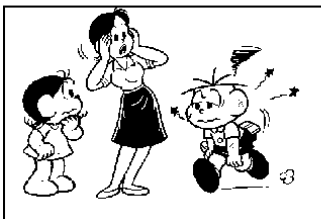
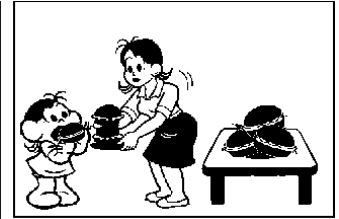
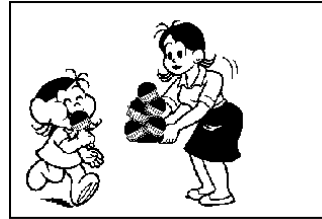
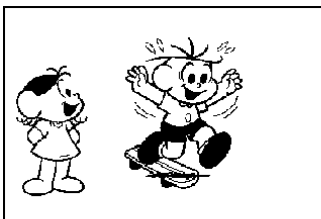
Depois de entrarem em casa, as crianças assistem televisão sentadas no sofá.

Ele não guardou os seus brinquedos e foi repreendido por seu pai.



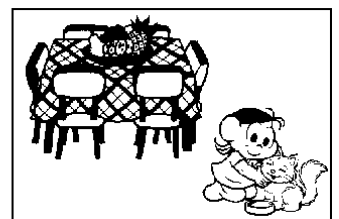
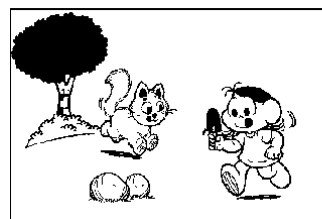
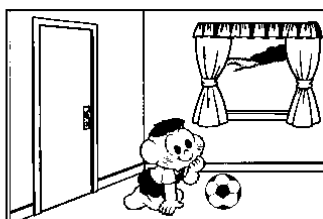
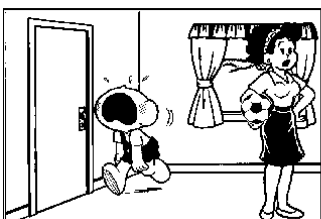
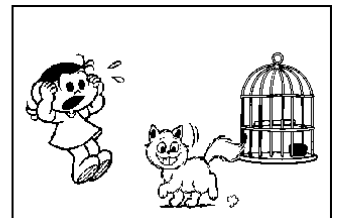
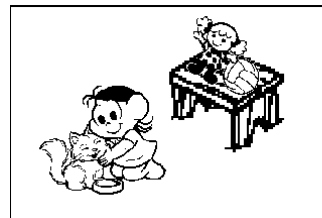
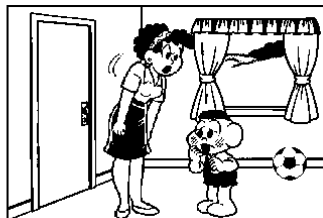
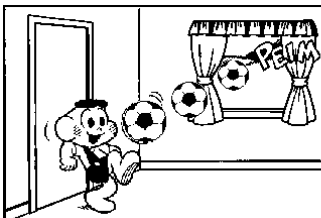
- Vá estudar! Me diz mamãe.

Eles caíram das pedras.



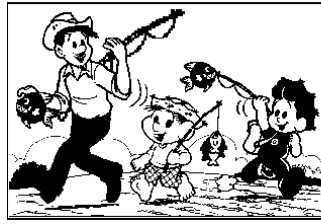
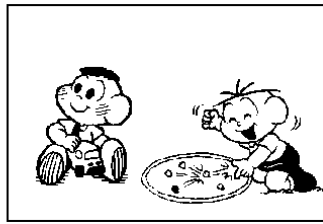
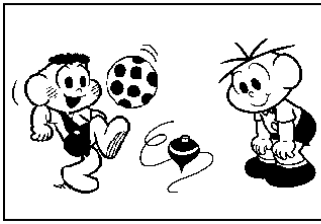
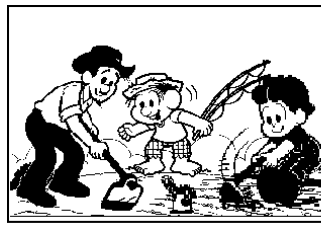
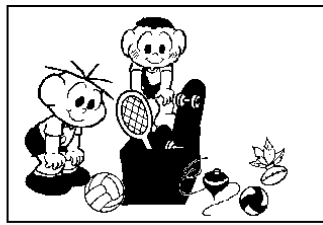
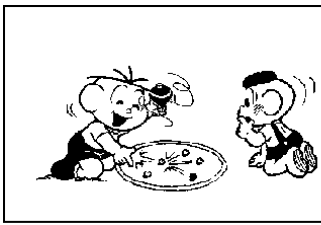
- O menino caiu?

Ela come os docinhos que a mamãe fez.



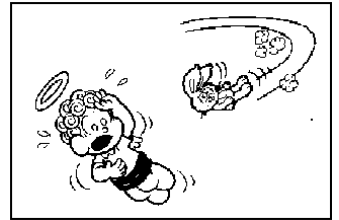
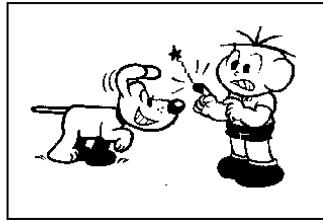
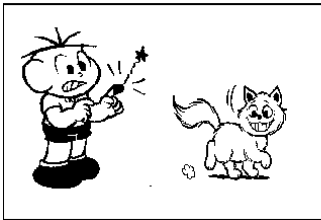
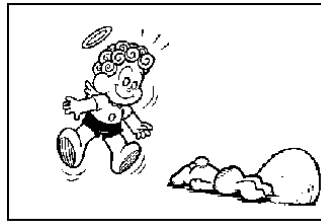
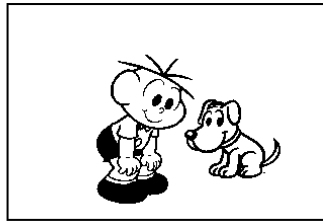
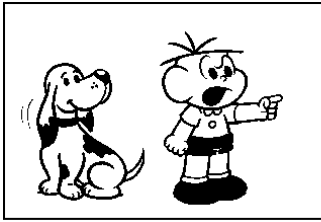
Cascão chutou sua bola muito forte e quebrou a janela. Sua mãe brigou e o puniu tirando-lhe a bola.

Depois de jantar, a menina dá leite para o gatinho



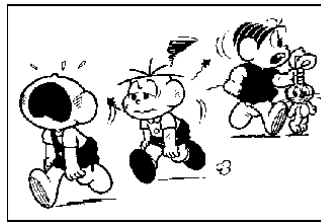
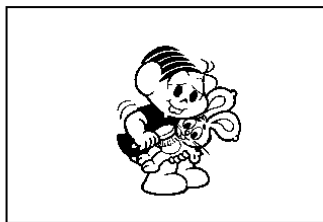
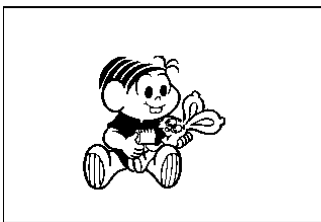
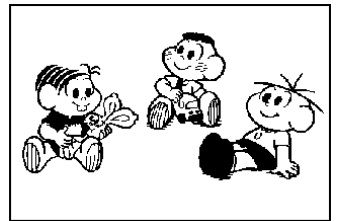
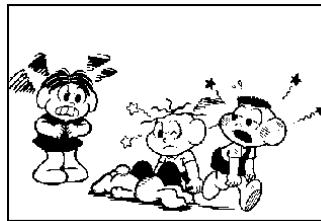
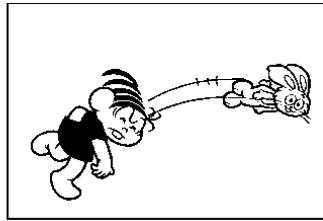
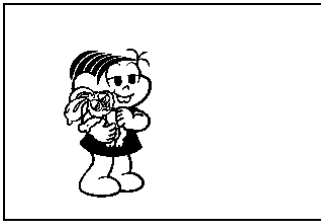
- Ganhei um pião no jogo de bolinhas de gude!

Eles procuram minhocas para irem pescar no lago.



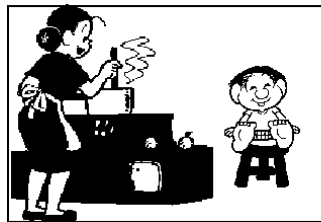
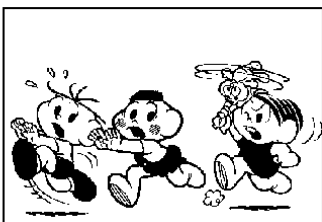
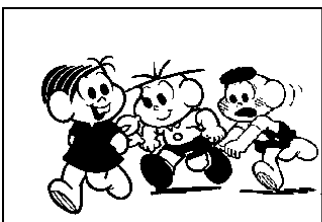
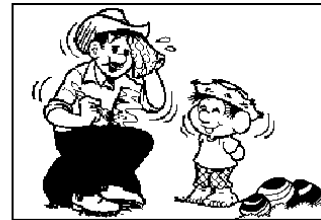
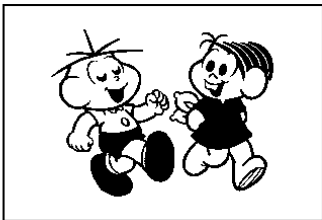
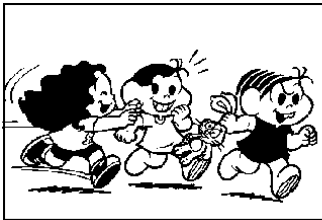
O meu dedo foi mordido por esse cão.

O anjinho será acordado pelo coelhinho.



O coelhinho é arremessado pela menina.

Os meninos apanharão.



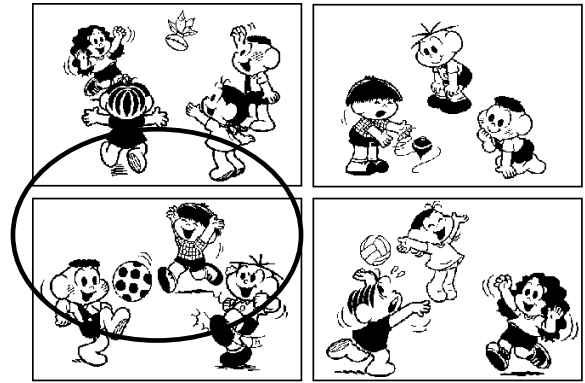
A Mônica é seguida pelos garotos.

Chico esperou sua mãe preparar o almoço e o levou para seu pai que trabalhava no campo. Ele sentou-se numa pedra e comeu satisfeito.

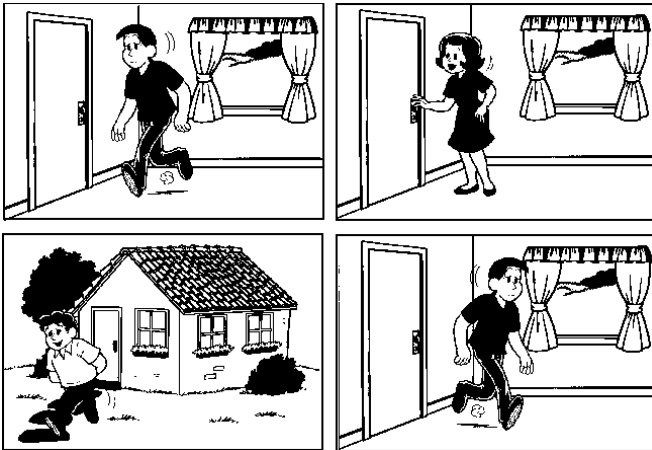
Compreensão de frases escritas em português (Zanella, 2007)

Como no modelo ao lado, apenas uma figura representa melhor a frase escrita abaixo delas. Circule a figura correta.

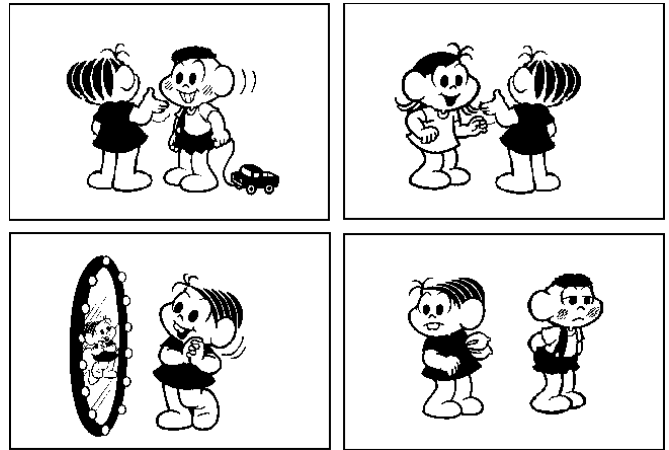
Trabalhe o mais rápido possível.



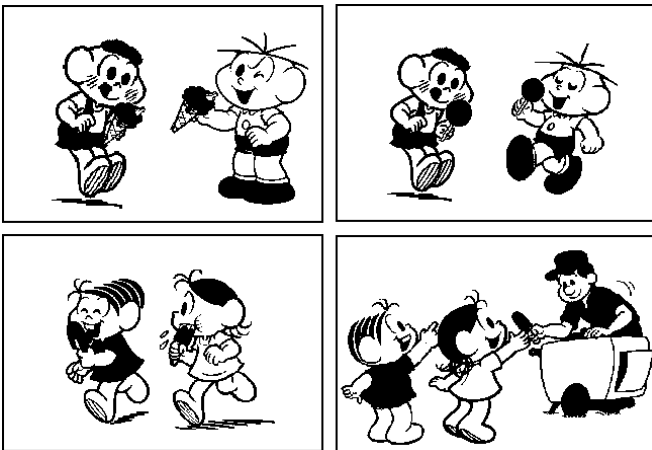
Os meninos jogam bola.



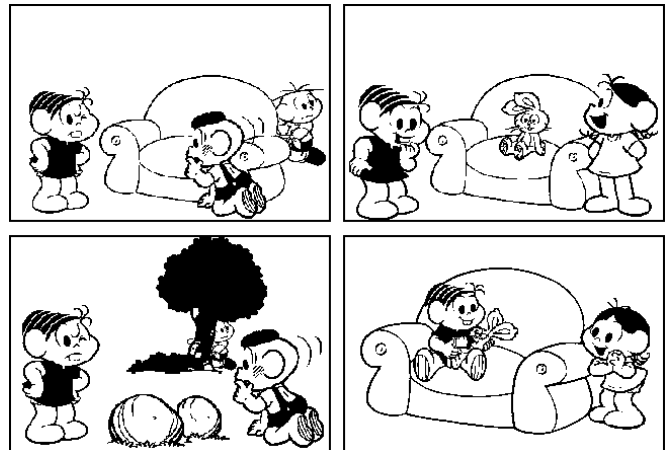
O senhor Pedro vai sair.



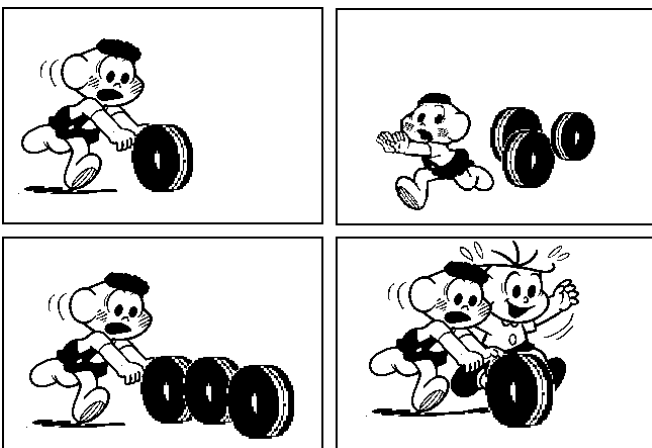
A menina olha para ele.



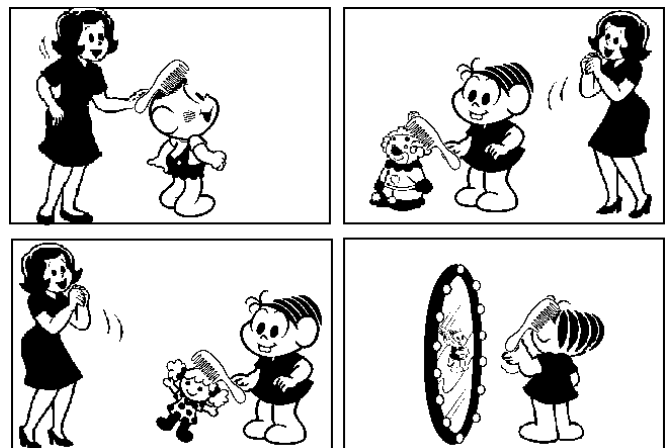
Elas chuparão sorvetes.



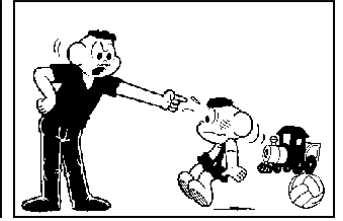
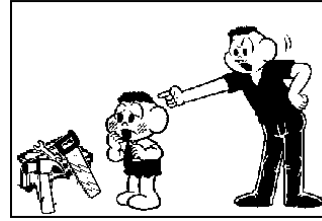
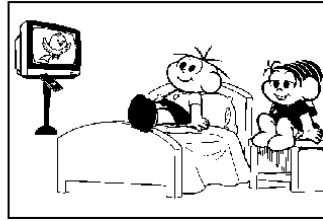
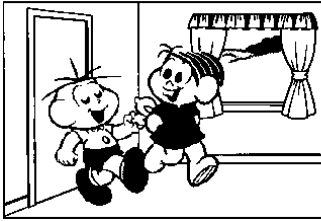
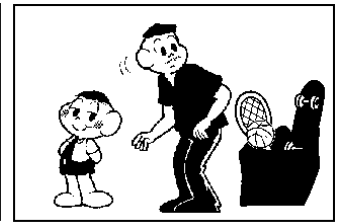
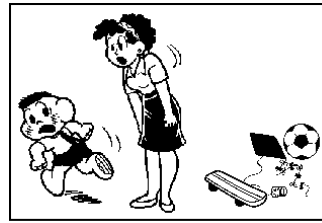
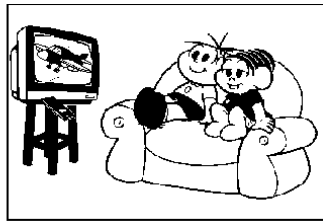
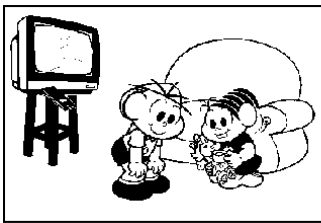
-Mas, onde está o coelhinho que eu deixei no sofá?



Os pneus são empurrados pelo menino.

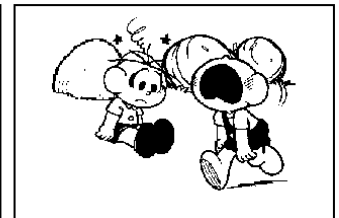
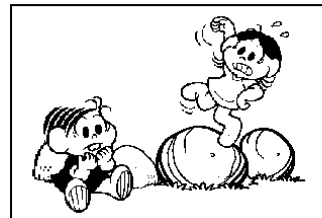
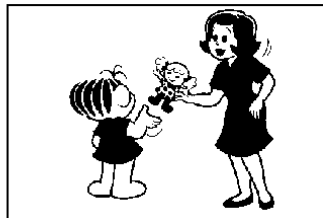
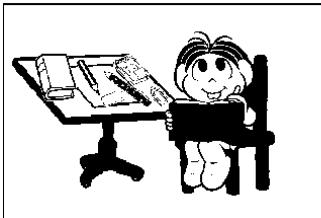
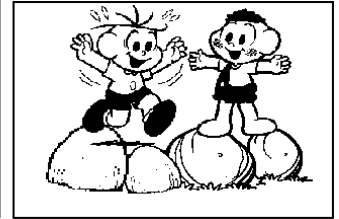
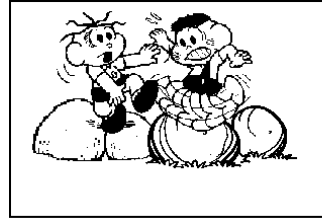
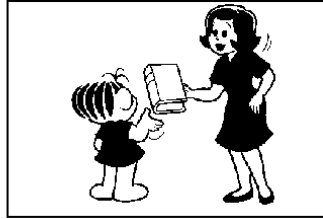
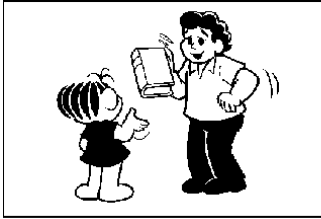


Mamãe vê a Mônica penteando os cabelos dele.



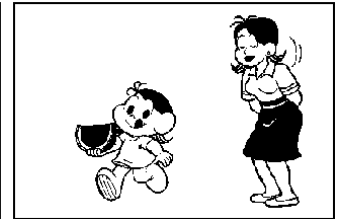
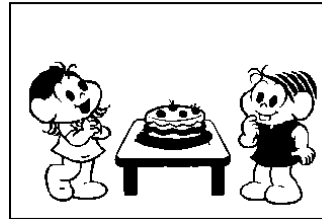
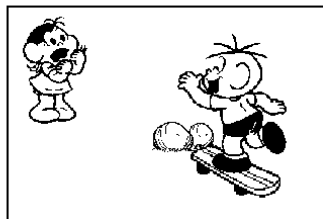
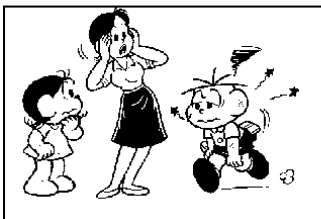
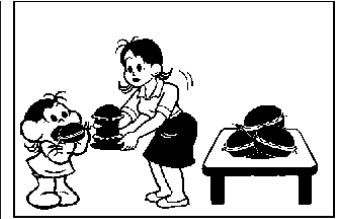
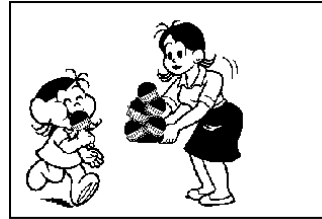
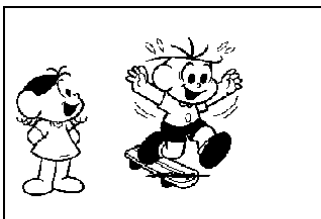
Depois de entrarem em casa, as crianças assistem televisão sentadas no sofá.

Ele não guardou os seus brinquedos e foi repreendido por seu pai.



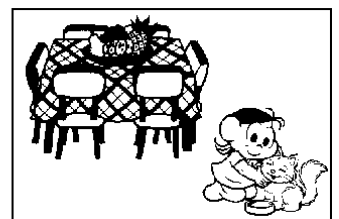
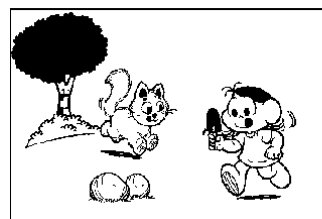
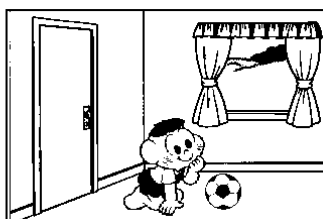
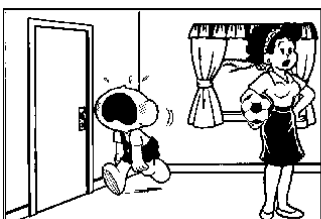
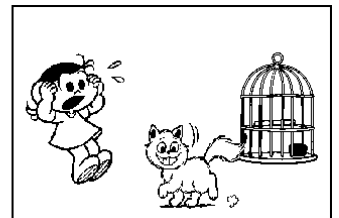
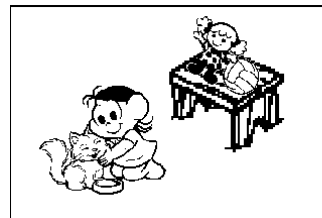
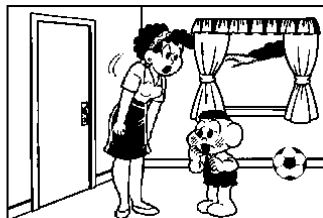
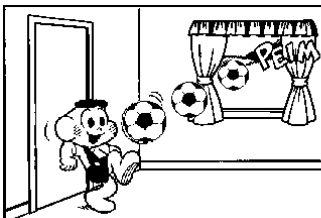
- Vá estudar! Me diz mamãe.

Eles caíram das pedras.



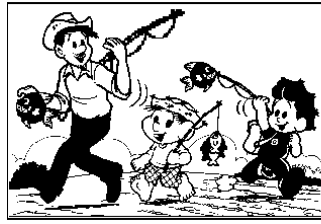
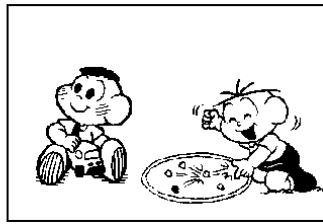
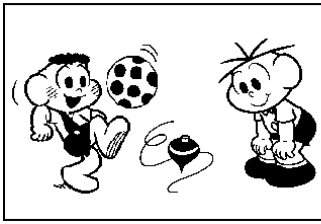
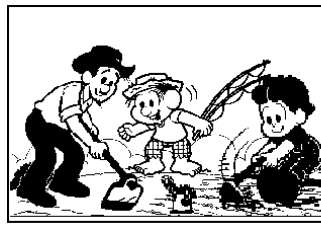
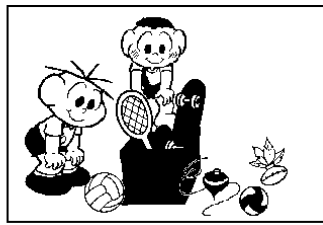
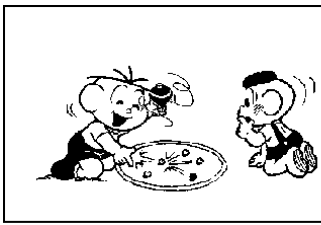
- O menino caiu?

Ela come os docinhos que a mamãe fez.



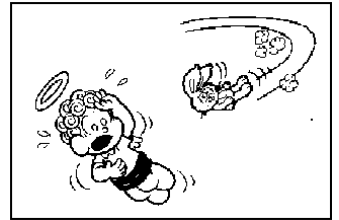
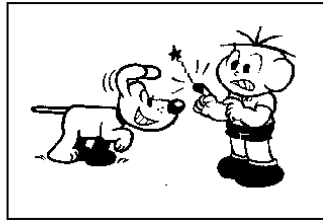
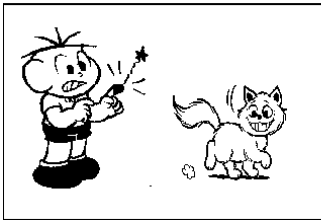
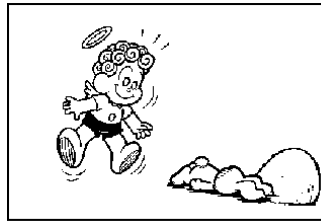
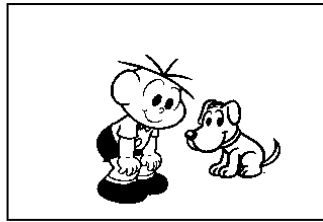
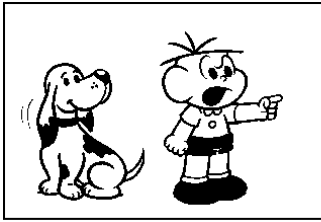
Cascão chutou sua bola muito forte e quebrou a janela. Sua mãe brigou e o puniu tirando-lhe a bola.

Depois de jantar, a menina dá leite para o gatinho



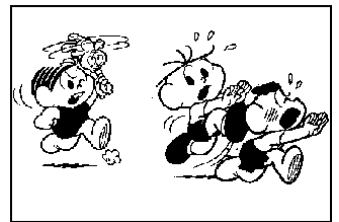
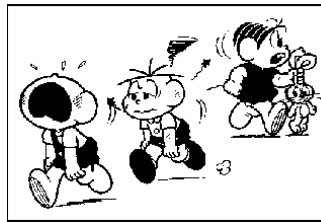
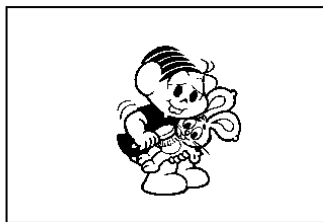
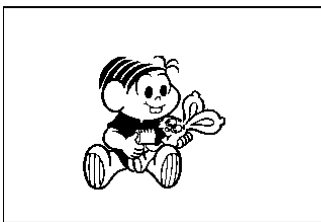
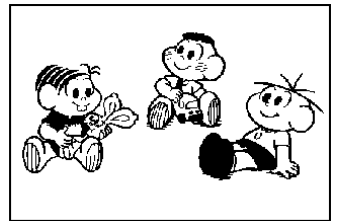
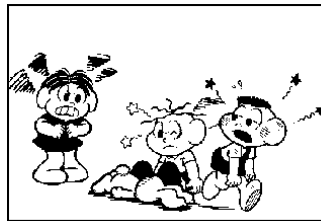
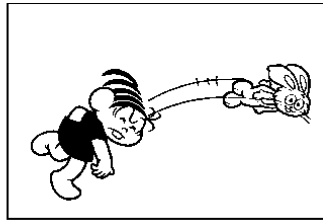
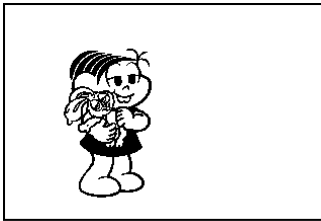
- Ganhei um pião no jogo de bolinhas de gude!

Eles procuram minhocas para irem pescar no lago.



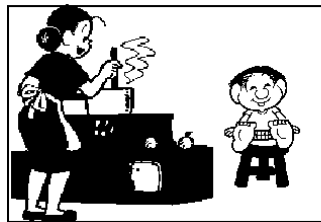
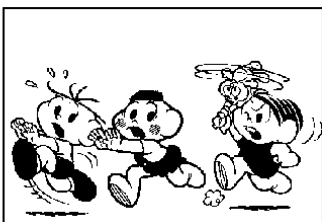
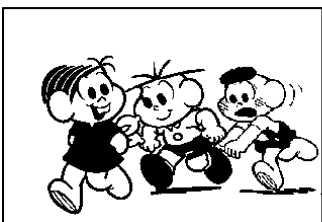
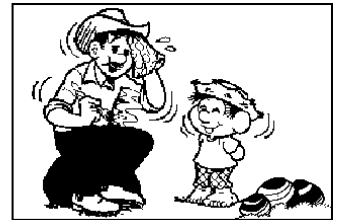
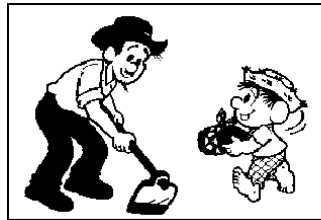
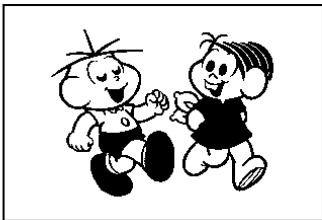
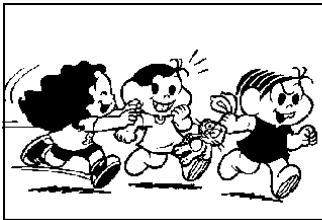
O meu dedo foi mordido por esse cão.

O anjinho será acordado pelo coelhinho.



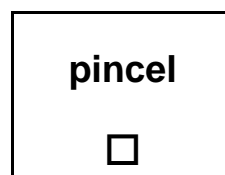
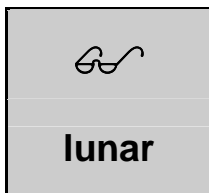
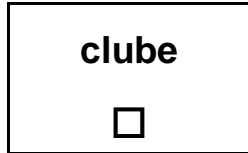
O coelhinho é arremessado pela menina.

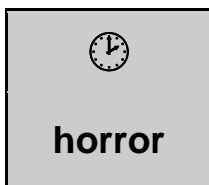
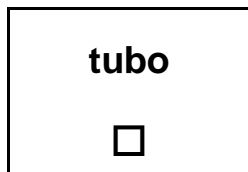
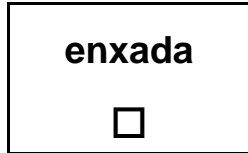
Os meninos apanharão.



A Mônica é seguida pelos garotos.

Chico esperou sua mãe preparar o almoço e o levou para seu pai que trabalhava no campo. Ele sentou-se numa pedra e comeu satisfeito.







viola

semente



panela



portão

museu



salão



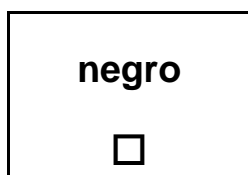
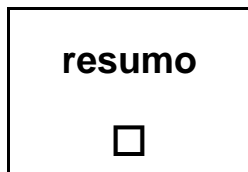
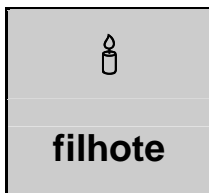
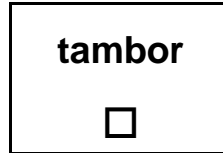
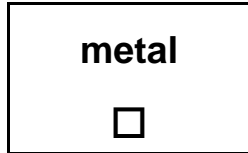
feroz

clamor



cuscuz





Item	Unité phonologique	Cible	traduction	C.G.	N.S.	N.P.	N.G.	Familiarité **	Lié	traduction	C.G.	N.S.	N.P.	N.G.	Familiarité **	Intrus	traduction	C.G.	N.S.	N.P.	N.G.	Familiarité **
1	V terminale	MENTE	menthe	n.f	2	4	5	60	CLUBE	club	n.m	2	5	5	64	DISCO	disque	n m	2	5	5	113
2	C terminale	LUNAR	lunaire	adj	2	5	5	17	PASTOR	berger	n.m	2	6	6	46	XADREZ	jeu d'échec	n m	2	6	6	44
3	rime	FOGÃO	cuisinière (objet)	n.m	2	5	5	44	MAMÃO	papaye	n.m	2	5	5	52	PINCEL	pinceau	n m	2	5	6	26
4	syllabe	CORRIDA	course	n.f	3	6	7	76	ENXADA	houe	n.f	3	5	6	41	ALFACE	laitue	n.f	3	6	6	59
5	V terminale	CHAVE	clé	n.f	2	4	5	116	MONTE	mont	n.m	2	4	5	119	TUBO	tube	n m	2	4	4	70
6	C terminale	HORROR	horreur	n.m	2	5	6	31	BAZAR	bazar	n.m	2	5	5	9	TELEX	minitel	n m	2	5	5	13
7	syllabe	VIOLA	petite guitare	n.f	3	5	5	35	PANELA	poil	n.f	3	6	6	76	SEMENTE	ciment	n.f	3	6	7	62
8	rime	PORTÃO	grande porte	n.m	2	6	6	56	SALÃO	salon	n.m	2	5	5	47	MUSEU	musée	n m	2	5	5	71
9	C terminale	FEROZ	feroce	adj	2	5	5	50	CUSCUZ	couscous	n.m	2	6	6	7	CLAMOR*	clameur	n m	2	6	6	2
10	rime	CASAL	couple	n.m	2	5	5	33	METAL	metal	n.m	2	5	5	40	TAMBOR	tambour	n m	2	5	6	45
11	syllabe	FILHOTE	chiot	n.m/f	3	6	7	67	TAPETE	tapis	n.m	3	6	6	64	RESUMO*	résumé	n m	3	6	6	56
12	V terminale	SÍTIO	ferme	n.m	2	5	5	117	NEGRO	noir	n.m	2	5	5	135	CORDA	corda	n.f	2	5	5	127

Moyenne familiarité 59

58

57

n.m = nom masculin; n.f = nom féminin, adj = adjectif

* graphème commun avec la cible mais ayant une autre prononciation

** Fréquence cumulée calculée à partir des valeurs de la base de donnée ADB (Pinheiro, 1987) issus des manuels scolaires de 2ème, 3ème et 4ème année



descolorir

deslizar



destorcer



chaveiro

cinzeiro



pandeiro



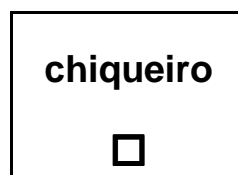
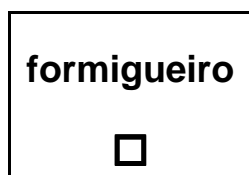
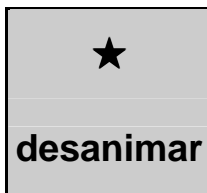
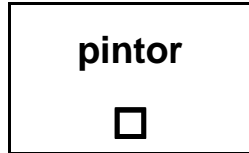
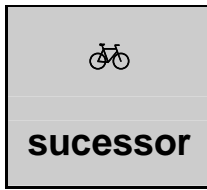
reflorir

recomendar



reconhecer







instrutor

major



feitor



reaparecer

recuperar



reproduzir



desabafar

desatar



destinar





passageiro

açougueiro



escoteiro



amargor

vigor



temor



revender

reformatar



relatar



Item	Affixe	Cible	traduction	N.S.	N.P.	N.G.	Familiarité*	Lié	traduction	N.S.	N.P.	N.G.	Familiarité**	Intrus	traduction	N.S.	N.P.	N.G.	Familiarité**	choix sémantique***
1	DES-	DESCOLORIR	décolorer	4	9	9	1	DESTORCER	détordre	3	9	9	6	DESLIZAR	glisser	3	8	8	10	36%
2	-EIRO	CHAVEIRO	porte clé	3	8	7	21	CINZEIRO	cendrier	3	8	8	9	PANDEIRO	tambourin	3	8	7	14	71%
3	RE-	REFLORIR	refleurir	3	8	8	2	RECONHECER	reconnaître	4	10	9	32	RECOMENDAR	recommander	4	10	9	5	79%
4	-OR	SUCCESSOR	successeur	3	8	7	110	PINTOR	peintre	2	6	5	72	DOUTOR	docteur	2	6	6	103	32%
5	DES-	DESANIMAR	décourager	4	9	9	7	DESOBEDECER	desobéir	5	11	11	8	DESPEDIR	dire au revoir	3	8	8	8	43%
6	-EIRO	GALINHEIRO	poulailler	4	10	9	75	FORMIGUEIRO	fourmilière	4	11	10	40	CHIQUEIRO	porcherie	3	9	8	23	25%
7	-OR	INSTRUTOR	instructeur	3	9	8	8	FEITOR	celui qui fait	2	6	6	6	MAJOR	majeur	2	5	5	2	68%
8	RE-	REAPARECER	réapparaître	5	10	10	2	REPRODUZIR	reproduire	4	10	10	13	RECUPERAR	reécuperer	4	9	9	22	39%
9	DES-	DESABAFAR	desétouffer	4	9	9	3	DESATAR	délier	3	7	7	3	DESTINAR	destiner	3	8	8	2	39%
10	-EIRO	PASSAGEIRO	passager	5	4	7	1	AÇOUGUEIRO	boucher	4	10	8	22	ESCOTEIRO	scout	4	9	9	3	14%
11	-OR	AMARGOR	amer	3	7	7	1	TEMOR	crainte	2	5	5	9	VIGOR	vigueur	2	5	5	5	79%
12	RE-	REVENDER	revendre	3	7	6	1	REFORMAR	reformer	3	9	9	2	RELATAR	relater	3	7	7	9	54%

Moyenne familiarité 19

19

17

N.S. = Nombre de syllabes; N.P. = Nombre de phonèmes; N.G. = Nombre de graphèmes

Moyenne choix sémantique = 48 %

** Fréquence cumulée calculée à partir des valeurs de la base de donnée ADB (Pinheiro, 1987) issus des manuels scolaires de 2ème, 3ème et 4ème année

*** % moyen de réponses de 20 élèves de 2ème année en faveur de l'intrus



diretora

garota



poeta



fabriquei

amarei



esperei



mineira

viúva



atleta





festejam

respiram



mandaram



latina

astronauta



operária



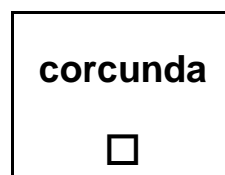
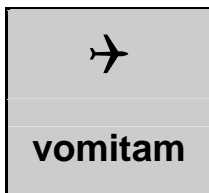
trabalhei

encontrei



andarei







verifiquei

espirrarei



exagerei



moderna

egoísta



nervosa



adicionam

provocaram



consideram



Item	Flexion	Cible	traduction	N.S.	N.P.	N.G.	Familiarité**	Lié	traduction	N.S.	N.P.	N.G.	Familiarité**	Intrus	traduction	N.S.	N.P.	N.G.	Familiarité**	choix sémantique***
1	Féminin	DIRETORA	directrice	4	8	8	37	GAROTA	gamine	3	6	6	80	POETA	poète	3	5	5	89	43%
2	Conjugaison passé	FABRIQUEI	j'ai fabriqué	3	8	9	63	ESPEREI	j'ai attendu	3	7	7	76	AMAREI	j'aimerai	3	6	6	64	43%
3	Féminin	MINEIRA	habitante du minas	3	7	7	116	VIÚVA	veuve	3	5	5	28	ATLETA	athlète	3	6	6	27	68%
4	Conjugaison présent	FESTEJAM	ils fêtent	3	8	8	20	RESPIRAM	ils respirent	3	8	8	55	MANDARAM	ils ont envoyé	3	7	8	52	32%
5	Féminin	LATINA	latine	3	6	6	11	OPERÁRIA	ouvrière	4	8	8	12	ASTRONAUTA	astronaute	4	10	10	23	71%
6	Conjugaison passé	TRABALHEI	j'ai travaillé	3	9	9	304	ENCONTREI	j'ai trouvé	3	7	9	310	ANDAREI	je marcherai	3	6	7	339	46%
7	Féminin	BAILARINA	balerine	4	9	9	52	SECRETARIA	secrétaire	4	10	10	13	CARIOCA	habitant de rio de janeiro	4	7	7	7	39%
8	Conjugaison présent	VOMITAM	ils vomissent	3	7	7	1	MELHORAM	ils améliorent	3	8	8	58	MARCARAM	ils ont marqué	3	8	8	57	75%
9	Féminin	CABRITA	chèvre	3	7	7	9	CANTORA	chanteuse	3	6	7	19	CORCUNDA	bossu	3	7	8	5	36%
10	Conjugaison passé	VERIFIQUEI	j'ai vérifié	4	9	10	91	EXAGEREI	j'ai exagéré	4	8	8	11	ESPIRRAREI	je éternuerai	4	9	10	10	64%
11	Féminin	MODERNA	moderne	3	7	7	41	NERVOSA	nerveuse	3	7	7	22	EGOÍSTA	égoïste	4	7	7	16	50%
12	Conjugaison présent	ADICIONAM	ils additionnent	3	9	9	31	CONSIDERAM	ils considèrent	4	8	9	24	PROVOCARAM	ils ont provoqué	4	10	10	24	61%

Moyenne familiarité 65

59

59

N.S. = Nombre de syllabes; N.P. = Nombre de phonèmes; N.G. = Nombre de graphèmes

Moyenne choix sémantique = 49 %

** Fréquence cumulée calculée à partir des valeurs de la base de donnée ADB (Pinheiro, 1987) issus des manuels scolaires de 2ème, 3ème et 4ème année

*** % de réponses de 20 élèves de 2ème année en faveur de l'intrus

Item	Structure syntaxique	Cible	traduction (approximative)	N.S. /mot	N.G. /mot	Familiarité**	Lié	traduction (approximative)	N.S. /mot	N.G. /mot	Familiarité**	Intrus	traduction (approximative)	N.S. /mot	N.G. /mot	Familiarité**	choix sémantique***
1	Ter + participe passé	Talita tem saído.	Talita est sortie.	3-1-3	6-3-5	299	Carina tem mentido.	Carina a menti	3-1-3	6-3-7	6	Joana tem marido.	Joana a un mari	3-1-3	5-3-6	71	46%
2	Estar + gérondif	O carro esta passando.	la voiture passe	1-2-1-3	1-5-4-8	338	O cão esta correndo	le chien court	1-1-1-3	1-3-4-8	145	A lago esta fundo	le lac est profond	1-2-1-2	1-4-4-5	260	89%
3	Ter + participe passé	Aqui tem doído sem parar.	A cet endroit (du corp), il y a de la douleur sans arrêt	2-1-3-1-2	4-3-5-3-5	6	Aqui tem vendido bastante	Ici, on vend beaucou	2-1-3-3	4-3-7-8	213	Aqui tem tecido com brilho .	ici, il y a du tissu avec des brillants	2-1-3-1-2	4-3-6-3-6	195	61%
4	Estar + gérondif	O violão esta tocando.	on joue de la guitare	1-2-1-3	1-6-4-7	78	O caminhão esta batendo	le camion est accidenté	1-2-1-3	1-6-4-7	213	O buraco esta profundo	le trou est profond	1-2-1-3	1-6-4-7	22	68%
5	Ter + participe passé	A galinha tem posto muito	La poule a pondu beaucoup	1-3-1-2-2	1-7-3-5-5	128	O aluno tem escrito mau.	l'élève écrit mal	1-3-1-3-1	1-5-3-7-3	461	O remédio tem efeito ruim.	le médicament a un effet mauvais	1-3-1-3-1	1-7-3-6-4	27	32%
6	Ter + participe passé	Tenho corrido no patio.	J'ai couru dans la cours	2-3-1-2	5-7-2-5	145	tenho perdido no jogo.	j'ai perdu le jeu	2-3-1-2	5-7-2-4	76	tenho zumbido no ouvido.	J'ai un bourdonnement dans l'oreille	2-3-1-3	5-7-2-6	7	71%
7	Estar + gérondif	Hoje, ele esta dormindo.	aujourd'hui il dort	2-2-1-3	4-3-4-8	286	Hoje, ele esta hesitando	aujourd'hui il hesite	2-2-1-3	4-3-4-8	2	Hoje, ele esta vagabundo	aujourd'hui, il est un vagabond	2-2-1-3	4-3-4-8	4	46%
8	Ter + participe passé	Tem entupido na pia.	L'évier est bouché	1-4-1-2	3-8-2-3	2	Tem florido na varanda.	il a fleuri dans la véranda	1-3-1-3	3-7-2-7	1	Tem liquido na seringa.	Il y a du liquide dans la seringue	1-3-1-3	3-7-2-7	145	39%
9	Estar + gérondif	Ele esta sorrindo.	il sourit	2-1-3	3-4-8	33	Ee esta demorando	il prend du temps	2-1-3	3-4-8	9	Ele esta estupendo	il est parfait	2-1-3	3-4-8	1	64%
10	Ter + participe passé	Ele tem andando.	Il a marché	2-1-3	3-3-6	339	Ele tem rido	il a ri	2-1-2	3-3-4	87	Ele tem medo	il a peur	2-1-2	3-3-4	605	61%
11	Estar + gérondif	Isto esta chegando	cela arrive	2-1-3	4-4-8	338	Isto esta cabendo	cela va (un vêtement, chaussure)	2-1-3	4-4-8	7	Isto esta horrendo	cela est affreux	2-1-3	4-4-8	1	39%
12	Ter + participe passé	Tem aquecido neste verão.	il fait chaud cet été.	1-4-2-2	3-7-5-5	21	Tem chovido neste pais.	Il a plu dans ce pays	1-3-2-2	3-7-5-4	70	Tem bandido neste bairro.	Il y a un bandit dans ce quartier	1-3-2-2	3-7-5-6	12	75%

Moyenne familiarité 168

Moyenne familiarité 108

Moyenne familiarité 113

N.S. = Nombre de syllabes; N.P. = Nombre de phonèmes; N.G. = Nombre de graphèmes

Moyenne choix sémantique = 58 %


** Fréquence cumulée calculée à partir des valeurs de la base de donnée ADB (Pinheiro, 1987) issus des manuels scolaires de 2ème, 3ème et 4ème année


*** % moyen de réponses de 20 élèves de 2ème année en faveur de l'intrus

Annexe 20

Epreuve explicite de détection d'intrus graphophonologique en portugais

Entraînements:

 boi	
pai <input type="checkbox"/>	mão <input type="checkbox"/>

 frango	
bala <input type="checkbox"/>	jogo <input type="checkbox"/>



pavão

lençol



limão



redonda

escada



estante



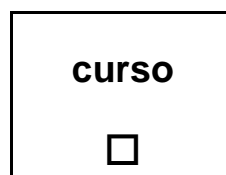
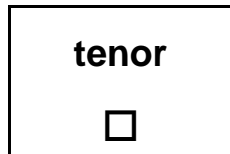
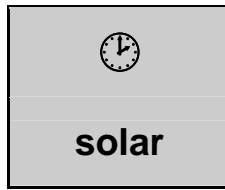
manga

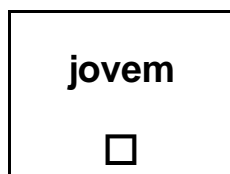
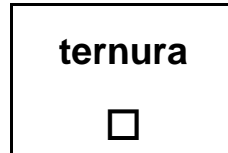
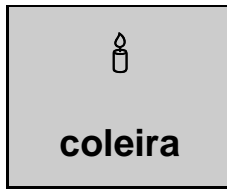
lenço



pilha









chicote

parente



balanço



sagaz

pavor



capuz



tigre

choque



chumbo



Item	Unité phonologique	Cible	traduction	C.G.	N.S.	N.P.	N.G.	Familiarité**	Lié	traduction	C.G.	N.S.	N.P.	N.G.	Familiarité*	Intrus	traduction	C.G.	N.S.	N.P.	N.G.	Familiarité*
1	rime	PAVÃO	paon	n.m	2	5	5	36	LIMÃO	citron	n.m	2	5	5	39	LENÇOL	drap	n m	2	6	6	72
2	syllabe	REDONDA	ronde	n.f	3	6	7	54	ESCADA	escalier	n.f	3	6	6	57	ESTANTE	étagère	n.f	3	6	7	61
3	V terminale	MANGA	mangue	n.f	2	4	5	45	PILHA	pile	n.f	2	5	5	45	LENÇO	mouchoir	n m	2	4	5	45
4	C terminale	SOLAR	solaire	adj	2	5	5	96	TENOR	tenor	n.m	2	5	5	5	PIREX	pirex	n m	2	5	5	6
5	rime	ONTEM	hier	adv	2	4	5	315	NUVEM	nuage	n.f	2	5	5	177	GALHO	branche	n m	2	5	5	136
6	V terminale	BAILE	bal	n.m	2	5	5	33	POSTE	poste	n.m	2	5	5	38	CURSO	cours	n m	2	5	5	38
7	syllabe	COLEIRA	collier	n.f	3	7	7	9	TERNURA	tendresse	n.f	3	7	7	7	SERPENTE	serpent	n.f	3	7	8	4
8	C terminale	AZAR	chance	adj	2	4	4	17	NARIZ*	nez	n.m	2	5	5	247	CALOR	chaleur	n m	2	5	5	306
9	rime	MARGEM	marge	n.f	2	6	6	112	JOVEM	jeune	n.m/f	2	5	5	85	VINHO	vin	n m	2	5	5	69
10	syllabe	CHICOTE	fouet	n.m	3	6	7	21	PARENTE	parent	n.m/f	3	6	7	25	BALANÇO*	balançoire	n m	3	6	7	21
11	C terminale	SAGAZ	perspicace	adj	2	5	5	3	CAPUZ	capuche	n.m	2	5	5	15	PAVOR	terreur	n m	2	5	5	25
12	V terminale	TIGRE	tigre	n.m	2	5	5	18	CHOQUE	choc	n.m	2	4	6	22	CHUMBO	plomb	n m	2	4	6	26

Moyenne familiarité 63

64

67

n.m = nom masculin; n.f = nom féminin, adj = adjectif

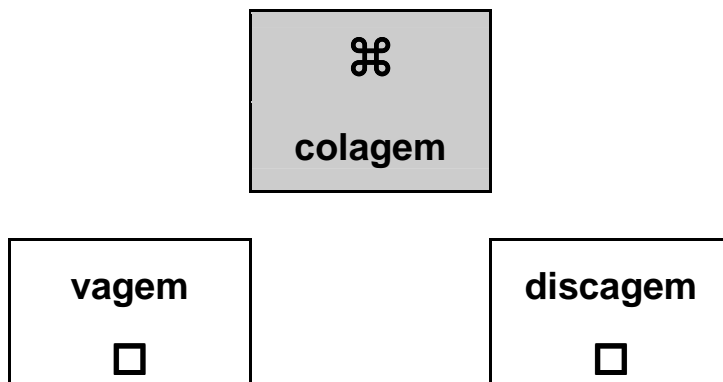
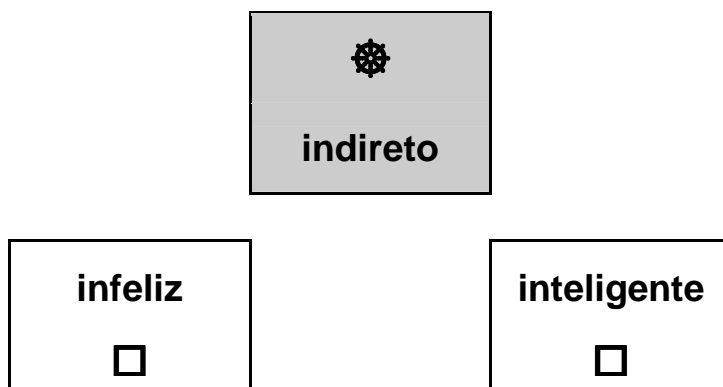
* graphème commun avec la cible mais ayant une autre prononciation

** Fréquence cumulée calculée à partir des valeurs de la base de donnée ADB (Pinheiro, 1987) issus des manuels scolaires de 2ème, 3ème et 4ème année

Annexe 21

Epreuve explicite de détection d'intrus grapho-morpho-dérivationnel en portugais

Entraînements :







reabrir

reler



regar



desconhecer

desafiar



desrespeitar




letrado

faqueiro



cheiro




reescrever

relacionar

reflorestar


protetor

autor

inventor


desatar

descuidar

desperdiçar



bagageiro

banheiro



picadeiro



replantar

reservar



relembrar



fedor

frescor



furor



Item	Affixe	Cible	traduction	N.S.	N.P.	N.G.	Familiarité*	Lié	traduction	N.S.	N.P.	N.G.	Familiarité**	Intrus	traduction	N.S.	N.P.	N.G.	Familiarité*	choix sémantique***
1	-EIRO	JARDINEIRO	jardinier	4	10	10	16	GUERREIRO	guerrier	3	9	8	35	PIONEIRO	pionnier	4	8	8	4	25%
2	DES-	DESARMAR	desarmer	3	8	8	1	DESCONFIAR	ne pas faire confiance	4	10	10	17	DESPERTAR	réveiller	3	9	9	19	54%
3	-OR	CORREDOR	couloir	3	8	7	30	REGADOR	arrosoir	3	7	7	16	COMPUTADOR	ordinateur	4	10	9	4	43%
4	RE-	REABRIR	rouvrir	3	7	7	2	RELER	relire	2	5	5	6	REGAR	arroser	2	5	5	8	79%
5	-EIRO	LETREIRO	panneau	3	8	8	2	FAQUEIRO	boîte à couteau	3	8	7	6	CHEIRO	odeur	2	6	5	151	61%
6	DES-	DESCONHECER	ne pas reconnaître	4	11	10	4	DESRESPEITAR	ne pas respecter	4	12	12	5	DESAFIAR	defier	4	8	8	2	29%
7	RE-	REESCREVER	réécrire	4	10	10	3	REFLORESTAR	reboiser	4	11	11	8	RELACIONAR	mettre en relation	4	10	10	8	11%
8	-OR	PROTETOR	protecteur	3	8	8	12	INVENTOR	inventeur	3	8	6	29	AUTOR	auteur	2	5	5	140	43%
9	DES-	DESATAR	détacher	3	7	7	3	DESCUIDAR	ne pas faire attention	3	9	9	2	DESPERDIÇAR	dépenser	4	11	11	1	68%
10	-EIRO	BAGAGEIRO	coffre à bagage	4	9	9	2	BANHEIRO	salle de bain	3	8	7	53	PICADEIRO	piste de cirque	4	9	9	12	54%
11	RE-	REPLANTAR	replanter	3	9	8	2	RELEMBRAR	resouvenir	3	9	8	4	RESERVAR	réserver	3	8	8	7	64%
12	-OR	FEDOR	puanteur	2	7	7	1	FRESCOR	fraicheur	2	7	7	1	FUROR	fureur	2	5	5	1	61%

Moyenne familiarité 7

15

30

N.S. = Nombre de syllabes; N.P. = Nombre de phonèmes; N.G. = Nombre de graphèmes

Moyenne choix sémantique = 49 %


** Fréquence cumulée calculée à partir des valeurs de la base de donnée ADB (Pinheiro, 1987) issus des manuels scolaires de 2ème, 3ème et 4ème année

*** % moyen de réponses de 20 élèves de 2ème année en faveur de l'intrus

Annexe 22


Epreuve explicite de détection d'instrus grapho-morpho-flexionnel en portugais

Entraînements :


gigante

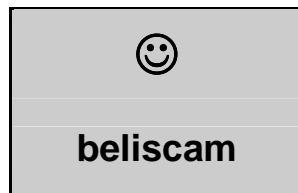
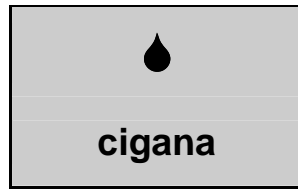
docente

mestre


corremos

veremos

queremos





enfeitei

testarei



numerei



louca

pirata



romana



transpiram

devoram



riscaram





delicada

confusa



peralta



atravessei

despertarei



cooperei



holandesa

coelha



caipora





impressionam

convidaram



recuperam



solteira

camarada



dançarina



adorei

procurei



pegarei



Item	Flexion	Cible	traduction	N.S.	N.P.	N.G.	Familiarité **	Lié	traduction	N.S.	N.P.	N.G.	Familiarité **	Intrus	traduction	N.S.	N.P.	N.G.	Familiarité **	choix sémantique***
1	Féminin	CIGANA	tsigane	3	6	6	9	FILHA	fille	2	5	5	154	GUARDA	garde	2	6	6	164	46%
2	Conjugaison présent	BELISCAM	ils pincent	3	8	8	4	SEGURAM	ils tiennent	3	7	7	37	SONHARAM	ils ont rêvé	3	8	8	39	68%
3	Féminin	NOIVA	fiancée	2	5	5	21	AUTORA	auteur	3	6	6	88	ARTISTA	artiste	3	7	7	78	39%
4	Conjugaison passé	ENFEITEI	j'ai arrangé	3	7	8	51	NUMEREI	j'ai énuméré	3	7	7	10	TESTAREI	je testerai	3	8	8	11	29%
5	Féminin	LOUCA	folle	2	5	5	36	ROMANA	romaine	3	6	6	24	PIRATA	pirate	3	6	6	28	68%
6	Conjugaison présent	TRANSPIRAM	ils respirent	3	9	10	6	DEVORAM	ils dévorent	3	7	7	6	RISCARAM	ils ont risqué	3	8	8	8	57%
7	Féminin	DELICADA	délicate	4	8	8	30	CONFUSA	confuse	3	6	7	9	PERALTA	coquin	3	7	7	12	43%
8	Conjugaison passé	ATRAVESSEI	j'ai traversé	4	9	10	116	COOPEREI	j'ai coopéré	4	8	8	17	DESPERTAREI	je me réveillerai	4	10	11	19	79%
9	Féminin	HOLANDESA	holandaise	4	7	9	21	COELHA	lapine	3	6	6	81	CAIPORA	personnage folklorique	3	7	7	31	36%
10	Conjugaison présent	IMPRESSIONAM	ils impressionnent	3	10	12	4	RECUPERAM	ils récupèrent	4	9	9	22	CONVIDARAM	ils ont invité	4	9	10	25	54%
11	Féminin	SOLTEIRA	célibataire	3	8	8	11	DANÇARINA	danseuse	4	8	9	10	CAMARADA	camarade	4	8	8	9	61%
12	Conjugaison passé	ADOREI	j'ai adoré	3	6	6	10	PROCUREI	j'ai cherché	3	8	8	150	PEGAREI	je prendrai	3	7	7	160	39%

Moyenne familiarité 27

51

49

N.S. = Nombre de syllabes; N.P. = Nombre de phonèmes; N.G. = Nombre de graphèmes

Moyenne choix sémantique = 49 %

** Fréquence cumulée calculée à partir des valeurs de la base de donnée ADB (Pinheiro, 1987) issus des manuels scolaires de 2ème, 3ème et 4ème année

*** % de réponses de 20 élèves de 2ème année en faveur de l'intrus

Annexe 23

Epreuve explicite de détection d'intrus grapho-syntaxique en portugais

Entraînements :



ela tinha comido

ela tinha aprendido

ela tinha apelido



ele estava pulando

ele estava horrendo

ele estava jogando



**o paciente tem visto seu
médico**

tenho dito para parar



tenho jeito para dançar



agora, você esta caindo

agora, voce esta vendo



agora, você esta lindo



**tem ardido de verdade na
garganta**

**tem cozido de carne na
geladeira**



**tem nascido de tudo na
floresta**





o balão esta subindo

o buraco esta redondo



o objeto esta partindo



tenho relido para a prova

tenho resumido para a prova



tenho comprimido para a gripe



o furo esta crescendo

o fogo esta brando



o frio esta vindo





**o cientista tem
descoberto planetas**

no interior tem feito frio



no bolo tem confeito rosa



**o musico tem composto
sobre o carnaval**

**o time tem respeito
sobre o adversario**



**o jornal tem descrito
sobre o escandalo**



o copo esta quebrando

o nariz esta doendo



o jardim esta fecundo





**tenho aberto devagar
esta porta**

**tenho direito de dormir
esta tarde**

**Tenho previsto de sair
esta noite**



João esta saindo

João esta imundo

João esta errando



ela tem sofrido de noite

ela tem saído de noite

ela tem vestido de noite

Item	Structure syntaxique	Cible	traduction (approximative)	N.S. /mot	N.G. /mot	Familiarité**	Lié	traduction (approximative)	N.S. /mot	N.G. /mot	Familiarité**	Intrus	traduction (approximative)	N.S. /mot	N.G. /mot	Familiarité**	choix sémantique***
1	Ter + participe passé	O paciente tem visto seu médico.	le patient rend visite au médecin	1-3-1-2-2-3	1-8-3-5-3-6	919	O professor tem dito para estudar	le professeur a dit d'étudier	1-3-1-2-2-3	1-9-3-4-4-7	570	O menino tem jeito para escrever.	L'enfant a une habileté pour écrire	1-3-1-2-2-3	1-6-3-5-4-8	307	46%
2	Estar + gérondif	Agora, você esta caindo	maintenant, tu es en train de tomber	3-2-1-3	5-4-4-6	153	agora, você esta vendo	aujourd'hui vous voyez	3-2-1-2	5-4-4-5	213	agora, você esta lindo	aujourd'hui tu es joli	3-2-1-2	5-4-4-5	230	57%
3	Ter + participe passé	Tem ardido de verdade na garganta.	La gorge est vraiment enflammée	1-3-1-3-1-3	3-3-7-2-5	5	Tem nascido de tudo na floresta.	Il né de tout dans la forêt	1-3-1-2-1-3	3-7-2-4-2-8	96	Tem cozido de carne na geladeira.	il y a un plat de viande dans le frigo	1-3-1-2-1-4	3-6-2-5-2-9	16	46%
4	Estar + gérondif	O balão esta subindo	le ballon s'envole	1-2-1-3	1-5-4-7	144	O objeto esta partindo	l'objet est en traine de se casser	1-2-1-3	1-5-4-7	230	O buraco esta redondo	le trou est rond	1-2-1-3	1-5-4-7	36	46%
5	Ter + participe passé	Tenho relido para a prova.	J'ai relue pour l'examen	2-3-2-1-2	5-6-4-1-5	6	Tenho revendido para a loja.	J'ai revendu à un autre magasin	2-4-2-1-2	5-9-4-1-4	11	Tenho comprimido para a gripe.	Il y a un médicament pour la grippe	2-4-2-1-2	5-10-4-1-5	16	64%
6	Estar + gérondif	O furo esta crescendo.	le trou s'agrandi	1-2-1-3	1-4-4-9	104	o frio esta vindo	le froid vient	1-2-1-3	1-4-4-9	81	o fogo esta brando	le feu est faible	1-2-1-3	1-4-4-9	18	86%
7	Ter + participe passé	O cientista tem descoberto planetas.	Le scientifique a découvert des planètes	1-4-1-4-3	1-9-3-10-8	375	No interior tem feito frio.	A l'intérieur (du pays) il a fait froid	1-3-1-2-2	2-8-3-5-4	333	No bolo tem confeito rosa.	Dans le gateau il y a des dragées roses	1-2-1-3-2	2-4-3-8-4		71%
8	Ter + participe passé	O musico tem composto sobre o carnaval.	Le musicien a composé sur le carnaval	1-3-1-3-2-1-3	1-6-3-8-5-1-8	11	O jornal tem descrito sobre o escandalo.	le journal a décrit le scandale	1-2-1-3-2-1-4	1-6-3-8-5-1-9	42	O time tem respeito sobre o adversario.	L'équipe a du respect pour l'adversaire	1-2-1-3-2-1-4	1-4-3-8-5-1-10	168	32%
9	Estar + gérondif	O copo esta quebrando.	le verre se casse	1-2-1-3	1-4-4-9	19	O nariz esta doendo	le nez est douloureux	1-2-1-3	1-4-4-9	6	o jardim esta fecundo	le jardin est fécond	1-2-1-3	1-4-4-9	2	71%
10	Ter + participe passé	Tenho aberto devagar esta porta.	j'ai ouvert lentement cette porte	2-3-3-2-2	5-6-7-4-5	109	Tenho previsto de dormir esta tarde	j'ai prévu de dormir cette après-midi	2-3-1-2-2-2	5-8-2-6-4-5	4	Tenho direito de sair esta noite	j'ai le droit de sortir cette nuit	2-3-1-2-2-2	5-7-2-4-4-5	257	39%
11	Estar + gérondif	João esta saindo.	jean est en train de sortir	1-1-3	4-4-6	299	João esta errando	jean se trompe	1-1-3	4-4-6	7	João esta imundo	jean est immonde	1-1-3	4-4-6	1	39%
12	Ter + participe passé	Ela tem sofrido de noite	Elle a eu mal pendant la nuit	2-1-3-1-2	3-3-7-2-5	19	Ela tem saido de noite	Elle est sorti de nuit	2-1-3-1-2	3-3-5-2-5	299	Ela tem vestido de noite	Elle a une robe de soirée	2-1-3-1-2	3-3-7-2-5	209	75%

Moyenne familiarité 180

Moyenne familiarité 158

Moyenne familiarité 115

N.S. = Nombre de syllabes; N.P. = Nombre de phonèmes; N.G. = Nombre de graphèmes

Moyenne choix sémantique = 56 %

** Fréquence cumulée calculée à partir des valeurs de la base de donnée ADB (Pinheiro, 1987) issus des manuels scolaires de 2ème, 3ème et 4ème année

*** % moyen de réponses de 20 élèves de 2ème année en faveur de l'intrus

Annexe 24

Epreuve d'indentification de la construction morphologique
de pseudo-mots en portugais

Entraînements :

Em bogar

“ficar” se diz “BUKIAN”	“ficar diante de qualquer coisa” se diz “BURIKIAN” “ficar atrás de qualquer coisa” se diz “BUMEKIAN”
“por” se diz “KADIAN”	“por diante de qualquer coisa” se diz “KARIDIAN” “por atrás de qualquer coisa” se diz “KAMEDIAN”
“colocar” se diz “CESIAN”	?

Como se diz « colocar diante de qualquer coisa » em bogar ?

CERISIAN	CERIDIAN	CEMESIAN
----------	----------	----------

Em valvanês

“dormir” se diz “UOTA”	“comer pouco” se diz “UOSHITA” “fazer dormir” se diz “UOGUTA”
“rir” se diz “KAIKI”	“rir pouco” se diz “KAISHIKI” “fazer rir” se diz “KAIGUKI”
“crescer” “JEYFO”	?

Como se diz « fazer crescer » em bogar?

JEYSHITO	JEYGUFO	JEYSHIFO
----------	---------	----------

1. Em casuco

“peixe” se diz “XOD”	“oito peixes” se diz “XODIKE” “seis peixes” se diz “XODUKO”
“macaco” se diz “GOF”	“oito macacos” se diz “GOFIKE” “seis macacos” se diz “GOFUKO”
“cachorro” se diz “RUV”	?

Como se diz « oito cachorros » em casuco?

RUVIKE	RUFIKE	RUVUKO
--------	--------	--------

2. Em dindozem

“bola” se diz “XKAL”	“bola, só tem uma” se diz “TIXKAL” “bola, tem várias” se diz “VEXKAL”
“livro” se diz “BLOMST”	“livro, só tem um” se diz “TIBLOMST” “livro, tem vários” se diz “VEBLOMST”
“viagem” se diz “REPS”	?

Como se diz « viagem, tem várias » em dindozem ?

TISEPS	VEREPS	TIREPS
--------	--------	--------

3. Em calanguêz

“dormir” se diz VLRT”	“dormir mal” se diz “VOLRIT” “dormir muito” se diz “VALROT”
“comer” se diz “ ZGZT”	“comer mal” se diz “ZOGZIT” “comer muito” se diz “ZAGZOT”
“falar” se diz “KPJT ”	?

Como se diz « falar muito » em calanguês?

KAPROT	KOPJIT	KAPJOT
--------	--------	--------

4. Em zunini

“o pai” se diz “AITA”	“com o pai” se diz “AITARENKIN” “para o pai” se diz “AITASETZAT”
“o irmão” se diz “ZUMU”	“com o irmão” se diz “ZUMURENKIN” “para o irmão” se diz “ZUMUSETZAT”
“o gato” se diz “NEBO”	?

Como se diz « com o gato » em zunini ?

NEBORENKIN	NEBURENKIN	NEBOSETZAT
------------	------------	------------

5. Em braquitês

“forte” se diz “HARBOR”	“ficar forte” se diz “UNHARBOR” “ser forte” se diz “GEHARBOR”
“vermelho” se diz “SITBOR”	“ficar vermelho” se diz “UNSITBOR” “ser vermelho” se diz “GESITBOR”
“bonito” se diz “RECHLICH”	?

Como se diz « ficar bonito » em braquitês?

GESECHLICH	UNRECHLICH	GERECHLICH
------------	------------	------------

6. Em swapadi

“fazer” se diz “KILIVA”	“fazer de novo” se diz “KILIVAKUMI” “fazer muito” se diz “KILIVASITA”
“voar” se diz “MAKIKA”	“voar de novo” se diz “MAKIKAKUMI” “voar muito” se diz “MAKIKASITA”
“costurar” se diz “FABASI”	?

Como se diz « costurar de novo » em swapadi ?

FABASISITA	FABASIKUMI	FABASASITA
------------	------------	------------

7. Em gamir

“cortar” se diz “KXT”	“você está cortando” se diz “KEXYT” “você terminou de cortar” se diz “KAXAT”
“subir” se diz “PMD”	“você está subindo” se diz “PEMYD” “você terminou de subir” se diz “PAMAD”
“notar” se diz “DVB”	?

Como se diz « você está notando » em gamir ?

DAVAB	DAMAB	DEVYB
-------	-------	-------

8. Em bacabroem

“amar” se diz “SIROEM”	“eu não amo” se diz “CHESIROEM” “eu amo” se diz “HASIROEM”
“falar” se diz “RANOEM”	“eu não falo” se diz “CHERANOEM” “eu falo” se diz “HARANOEM”
“compreender” se diz “POSOEM”	?

Como se diz « eu não compreendo » em bacabroem?

CHEPOSOEM	HAROSOEM	HAPOSOEM
-----------	----------	----------

9. Em natufi

“fazer” se diz “NEMI”	“que não se pode fazer” se diz “NEMIHUA” “que se pode fazer” se diz “NEMIZTLI”
“conhecer” se diz “POXU”	“que não se pode conhecer” se diz “POXUHUA” “que se pode conhecer” se diz “POXUZTLI”
“pedir” se diz “CEPA”	?

Como se diz « que se pode pedir » em natufi?

CEPAHUA	CEPIHUA	CEPAZTLI
---------	---------	----------

10. Em zetuke

“comprar” se diz “DKS”	“o contrário de comprar” se diz “DOKAS” “a ação de comprar” se diz “DUKIS”
“subir” se diz “KGL”	“o contrário de subir” se diz “KOGAL” “a ação de subir” se diz “KUGIL”
“partir” se diz “ZTB”	?

Como se diz « a ação de partir » em zetuke ?

ZUTIB	ZUKIB	ZOTAB
-------	-------	-------

11. Em poaliir

“aprender” se diz “JANNGA”	“quem aprende sou eu” se diz “MIDOJANNGA” “quem aprende é você” se diz “ADAJANNGA”
“escrever” se diz “KINNDU”	“quem escreve sou eu” se diz “MIDOKINNDU” “quem escreve é você” se diz “ADAKINNDU”
“trabalhar” se diz “HAALDE”	?

Como se diz « quem trabalha sou eu » em poaliir ?

ADAHAALDE	MIDOKAALDE	MIDOHAALDE
-----------	------------	------------

12. Em gavanês

“falar” se diz “ZEGEN”	“parar de falar” se diz “VERZEGEN” “falar de novo” se diz “HERZEGEN”
“levar” se diz “VATTEN”	“parar de levar” se diz “VERVATTEN” “levar de novo” se diz “HERVATTEN”
“cair” se diz “GEDEN”	?

Como se diz « cair de novo » em gavanês?

HERGEDEN	VERZEDEN	VERGEDEN
----------	----------	----------

13. Em tanvinês

“carregar” se diz “ZLT”	“quem carrega sou eu” se diz “ZELAT” “quem carrega é ele” se diz “ZOLUT”
“avançar” se diz “FVP”	“quem avança sou eu” se diz “FEVAP” “quem avança é ele” se diz “FOVUP”
“começar” se diz “BRN”	?

Como se diz « quem começa é ele » em tanvinês?

BOLUN	BERAN	BORUN
--------------	--------------	--------------

14. Em tepolano

“passar” se diz “MIT”	“quem passa é você” se diz “MITINI” “quem passa é ele” se diz “MITARI”
“cair” se diz “CHIR”	“quem cai é você” se diz “CHIRINI” “quem cai é ele” se diz “CHIRARI”
“pedir” se diz “RAV”	?

Como se diz « quem pede é ele » em tepolano?

RATINI	RAVARI	RAVINI
---------------	---------------	---------------

15. Em turiti

“livros” se diz “KNIHO”	“aquele que fabrica os livros” se diz “KNIHONATCHO” “loja de livros” se diz “KNIHORICTVO”
“sapatos” se diz “OBUV”	“aquele que fabrica sapatos ” se diz “OBUVNATCHO” “loja de sapatos” se diz “OBUVRICTVO”
“móveis” se diz “NABYTOK”	?

Como se diz « loja de móveis » em turiti ?

NABYTOKRICTVO	NABYTOKNATCHO	NABYTOVNATCHO
----------------------	----------------------	----------------------

16. Em afivonês

“estranho” se diz “STKT”	“estranho para um menino” se diz “SITAKOT” “estranho para uma menina” se diz “SUTIKAT”
“rápido” se diz “NTSJ”	“rápido para um menino” se diz “NITASOJ” “rápido para uma menina” se diz “NUTISAJ”
“limpo” se diz “WKBZ”	?

Como se diz « limpo para uma menina » em afivonês ?

WIKABOZ	WUKIBAZ	WIKAKOZ
---------	---------	---------

17. Em tasige

“dente” se diz “ZPN”	“aquele que cuida dos dentes” se diz “ZEPON” “um dente feio” se diz “ZUPIN”
“mão” se diz “LTK”	“aquele que cuida das mãos” se diz “LETOK” “uma mão feia” se diz “LUTIK”
“floresta” se diz “DRV”	?

Como se diz « aquele que cuida da floresta » em tasige?

DETOV	DEROV	DURIV
-------	-------	-------

18. Em vankenês

“engraçado” se diz “VEGAREK”	“engraçado para um menino” se diz “JEJVEGAREK” “engraçado para uma menina” se diz “SWOVEGAREK”
“doente” se diz “KOSZULA”	“doente para um menino” se diz “JEJKOSZULA” “doente para uma menina” se diz “SWOKOSZULA”
“pobre” se diz “SPODNIE”	?

Como se diz « pobre para uma menina » em vankenês ?

SWOVPODNIE	JEJSPODNIE	SWOSPODNIE
------------	------------	------------

19. Em voroquês

“cidade” se diz “DEKAT”	“quem mora na cidade” se diz “RAZDEKAT” “quem ama a cidade” se diz “PAJDEKAT”
“mar” se diz “VOROVAT”	“quem mora no mar” se diz “RAZVOROVAT” “quem ama o mar” se diz “PAJVOROVAT”
“campo” se diz “SHITZAT”	?

Como se diz « quem mora no campo » em voroquês ?

PAJSHITZAT	RAZSHITZAT	RAZVOSHITZAT
------------	------------	--------------

20. Em ulamadês

“cama” se diz “BRTR”	“sobre uma cama” se diz “BRYTER” “na cama” se diz “BRETUR”
“mala” se diz “KNPR”	“sobre a mala” se diz “KNYPER” “na mala” se diz “KNEPUR”
“armário” se diz “SRKR”	?

Como se diz « sobre o armário » em ulamadês ?

SRYKER	SREKUR	SRYTER
--------	--------	--------

21. Em dunês

“ele compra” se diz “PYALA”	“ele acabou de comprar” se diz “PYALATETE” “ele vai comprar mais tarde” se diz “PYALAKIRA”
“ele sobe” se diz “FRANU”	“ele acabou de subir” se diz “FRANUTETE” “ele vai subir mais tarde” se diz “FRANUKIRA”
“ele ajuda” se diz “KWANI”	?

Como se diz « ele acabou de ajudar » em dunês ?

KWANIKIRA	KWANAKIRA	KWANITETE
-----------	-----------	-----------

22. Em xambara

“rolar” se diz “XIMOA”	“rolar demais” se diz “MOXIMOA” “fazer o contrário de rolar” se diz “TIXIMOA”
“rodar” se diz “PECHIA”	“rodar demais” se diz “MOPECHIA” “fazer o contrário de rodar” se diz “TIPECHIA”
“emprestar” se diz “LOTIA”	?

Como se diz « fazer o contrário de emprestar » em bambara ?

MOLOTIA	TILOTIA	TIXOTIA
---------	---------	---------

23. Em soforês

“dia” se diz “DAG”	“dia, tem só um” se diz “DAGEN” “dia, tem vários” se diz “DAGNA”
“campo” se diz “DORD”	“campo, tem só um” se diz “DORDEN” “campo, tem vários” se diz “DORDNA”
“mulher” se diz “KVIS”	?

Como se diz « mulher, tem várias » em soforês?

KVISNA	KVISEN	KVIDNA
--------	--------	--------

24. Em bibolês

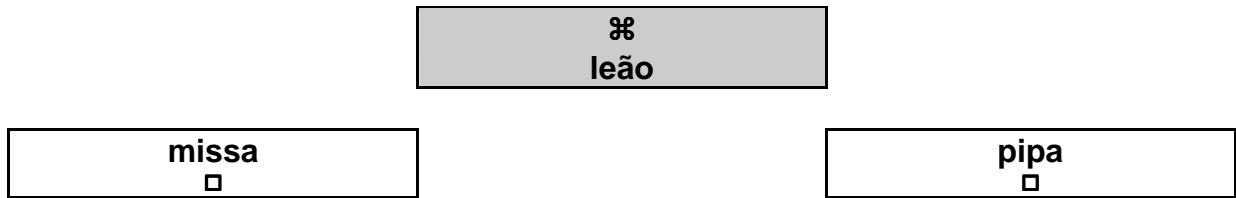
“gato” se diz “XZPT”	“gato pequeno” se diz “XAZIPAT” “gato grande” se diz “XUZOPIT”
“árvore” “GFTV”	“árvore pequeno” se diz “GAFITAV” “árvore grande” se diz “GUFOTIV”
“bicicleta” se diz “MVLB”	?

Como se diz « bicicleta pequena » em bibolês ?

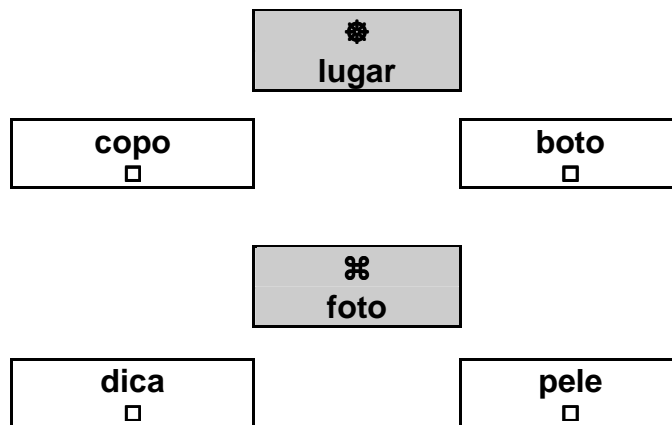
MUVOLIB	MUVOTAB	MAVILAB
---------	---------	---------

Annexe 25

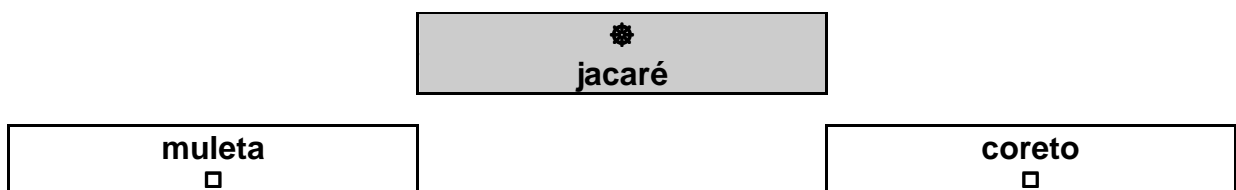
Exemple donné à l'épreuve implicite graphophonologique



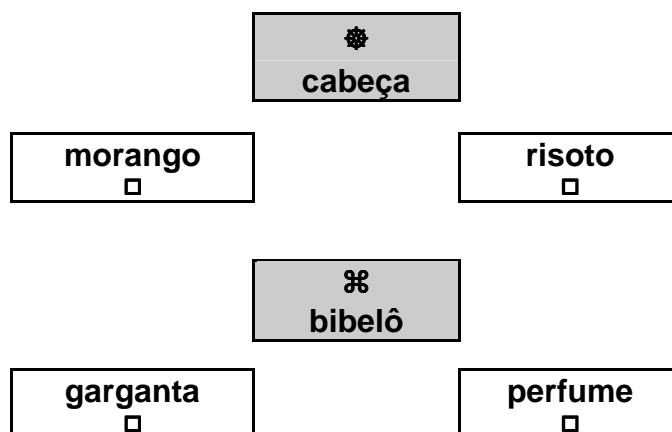
Items d'entraînement graphophonologiques implicites



Exemple donné aux épreuves implicites grapho-morphologiques



Items d'entraînement grapho-morphologiques implicites



Exemple donné à l'épreuve implicite grapho-syntaxique


nós vamos brincar de roda

este pássaro é comum

o almoço está pronto

Items d'entraînement grapho-syntaxiques implicites


o pato vai nadar

minha jarra quebrou

ganhei um presente

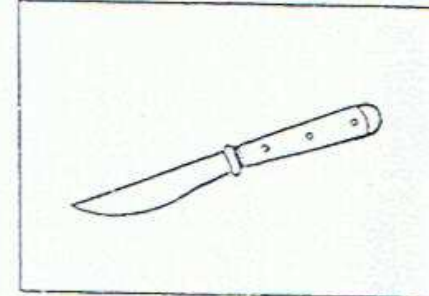
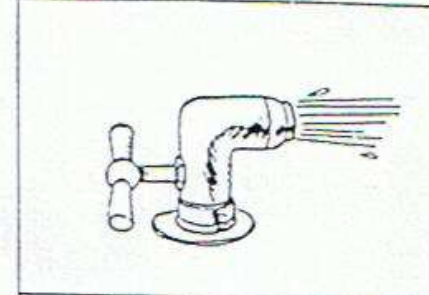
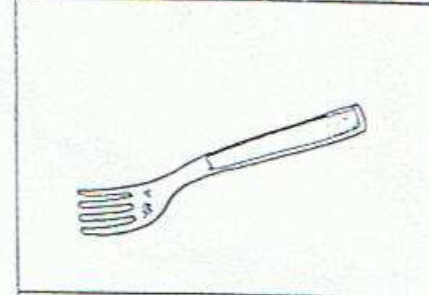
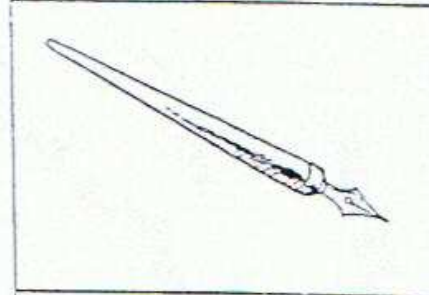
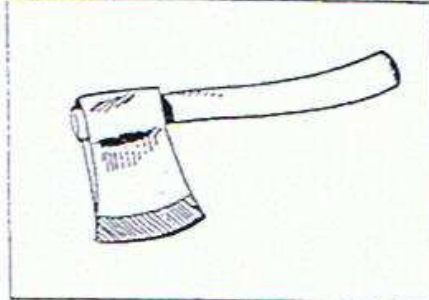
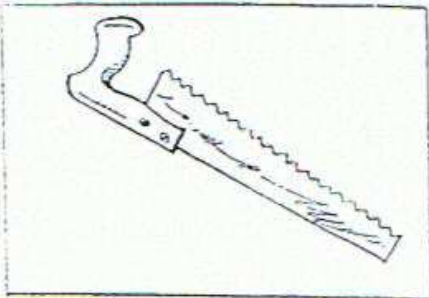

o gato dormiu no mato

a bala é de cereja

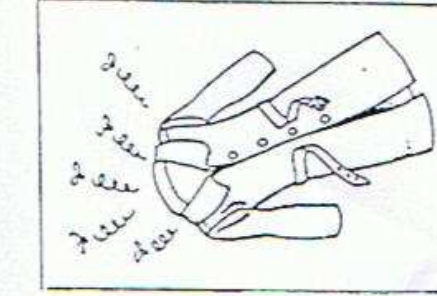
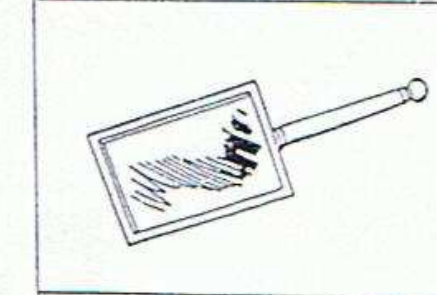
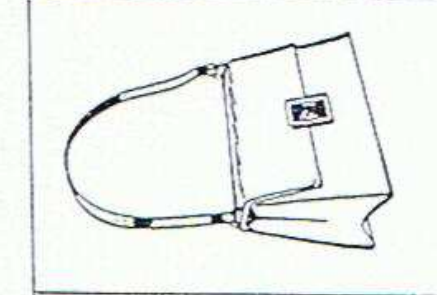
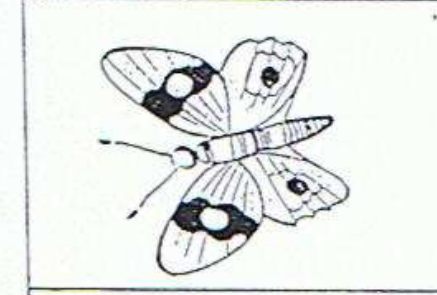
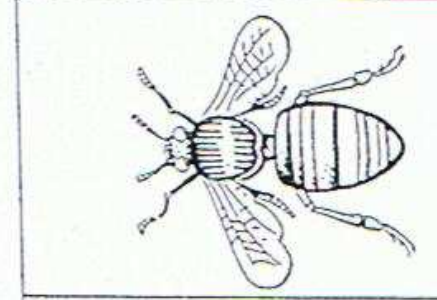
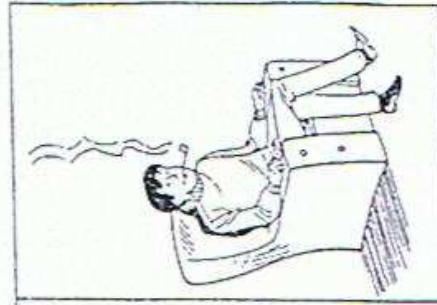
a sandália é de mulher

Annexe 26
Exemples d'item du test de Vocabulaire passif en français
(Deltour & Hupkens, 1980)

« couteau »



« sac »



Annexe 27

Test de lecture de mots en une minute en français (Khomsi, 1999)

1 il	37 chorale	73 plongeant
2 un	38 bruit	74 parent
3 le	39 monsieur	75 idéal
4 lui	40 agir	76 obéir
5 nu	41 cousin	77 griller
6 os	42 unir	78 travail
7 et	43 naïf	79 yacht
8 fil	44 curieux	80 animal
9 bol	45 venir	81 surprise
10 tous	46 départ	82 doigt
11 mur	47 écho	83 sommeil
12 sac	48 devient	84 réunir
13 clé	49 mardi	85 briller
14 est	50 drôle	86 guêpe
15 rue	51 histoire	87 feuillage
16 petit	52 glacé	88 gymnastique
17 camion	53 derrière	89 fauteuil
18 nom	54 coloquinte	90 obtenir
19 acide	55 stylo	91 jusque
20 film	56 magasin	92 expliquer
21 grand	57 cueillir	93 magnifique
22 mars	58 compagnon	94 spécial
23 jaune	59 chronomètre	95 fenouil
24 parc	60 continue	96 gourmand
25 sept	61 soleil	97 chirurgien
26 bloc	62 pharmacie	98 prévenir
27 faim	63 enveloppe	99 assiette
28 compter	64 éléphant	100 dangereux
29 poisson	65 trésor	101 observer
30 herbe	66 choc	102 obscurité
31 pied	67 fier	103 descendre
32 chemin	68 nouveau	104 installeront
33 joie	69 expert	105 inquiétude
34 phare	70 sportif	106 gastronomie
35 ciel	71 merveilleux	107 urbanisme
36 coup	72 odeur	108 anthropophage

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

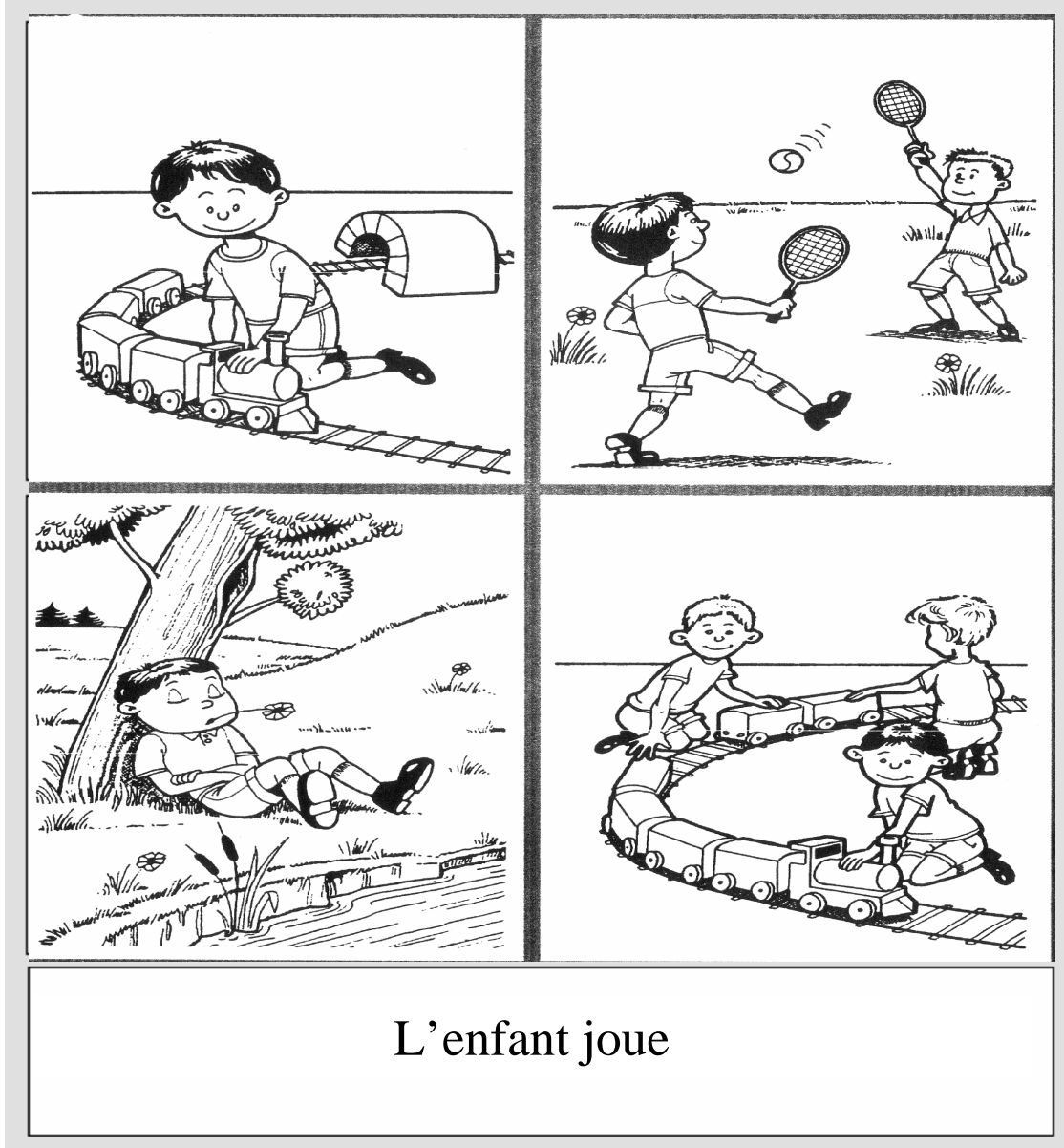
.....

.....

.....

Vous allez lire une
petite phrase et retrouver
l'image qui va avec.

Modèle



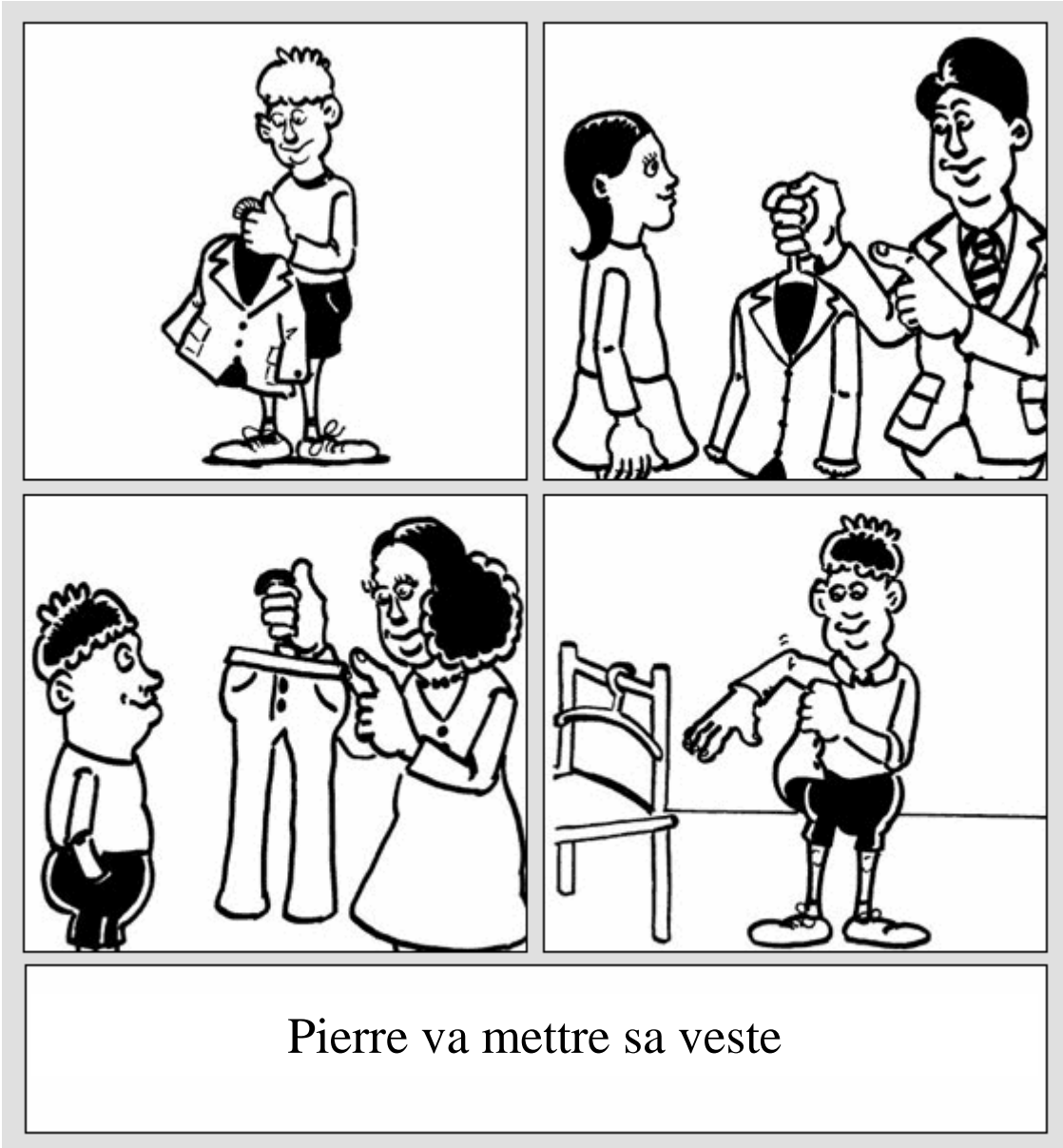
L'enfant joue

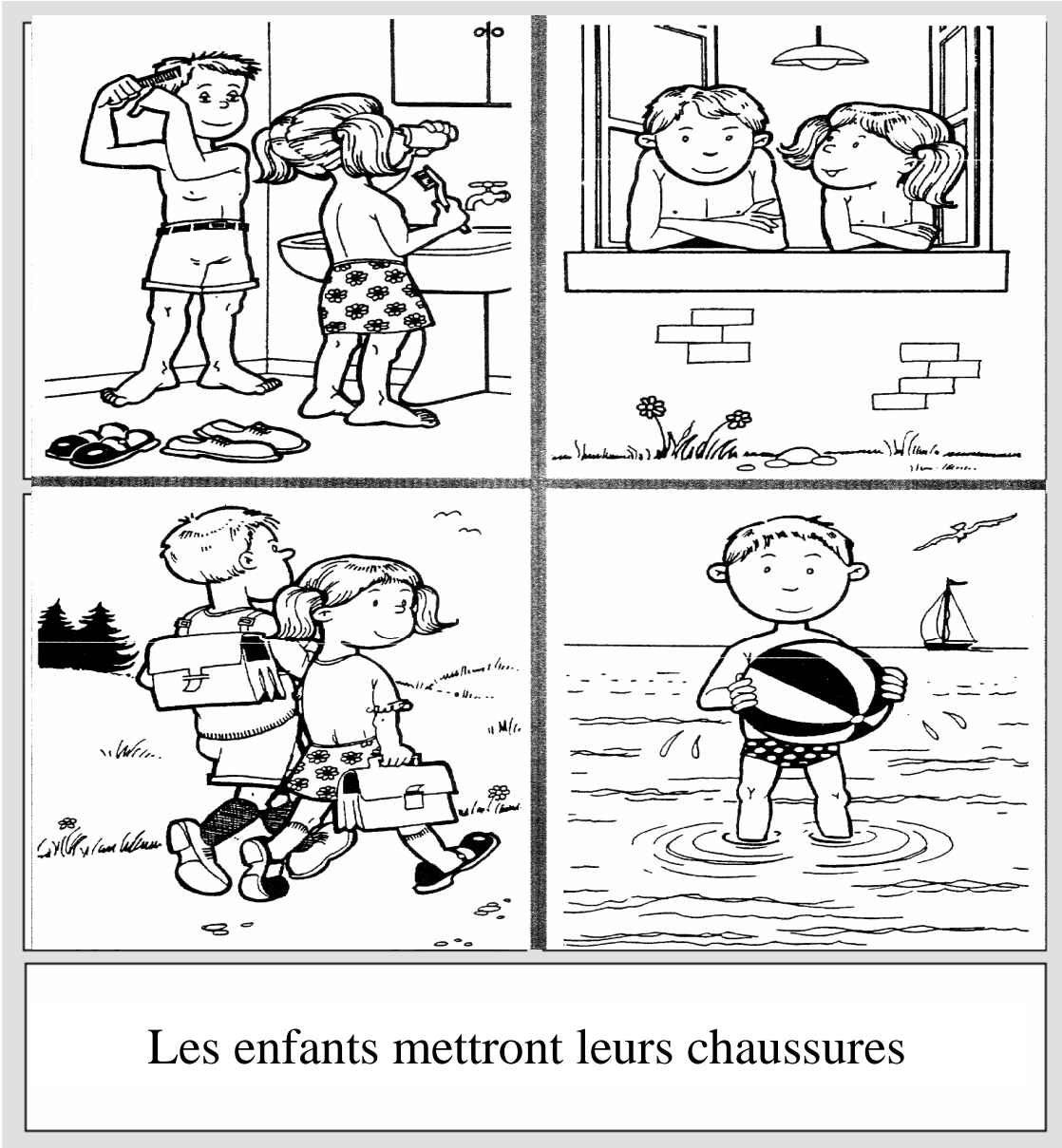
Comme dans le modèle, il n'y a qu'une seule image qui va avec ce qui est écrit : entoure-la. Travaille le plus vite possible.

Essai

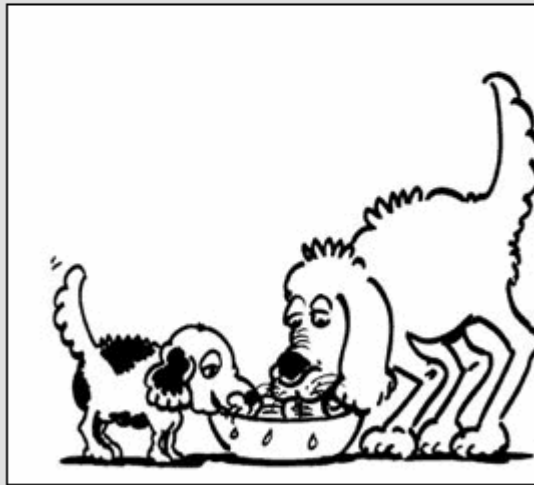
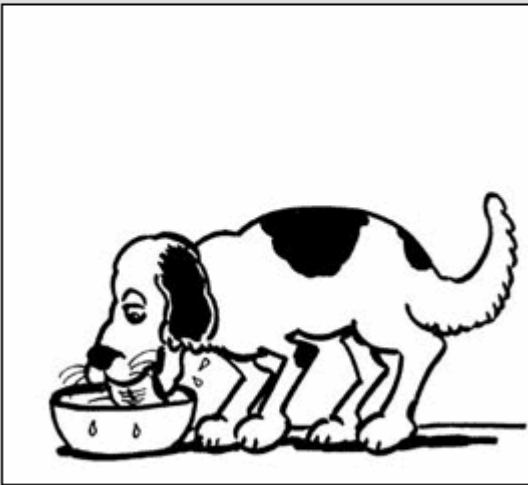
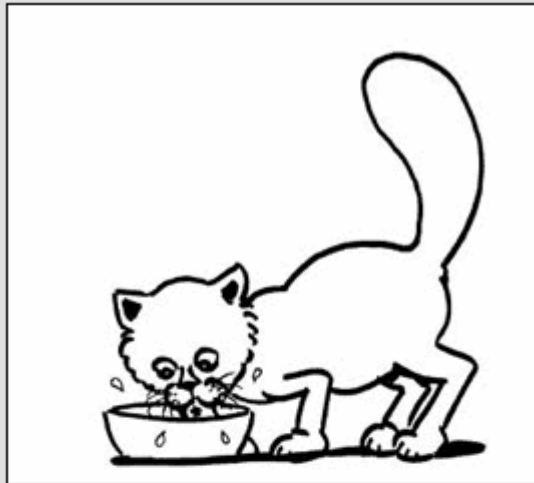
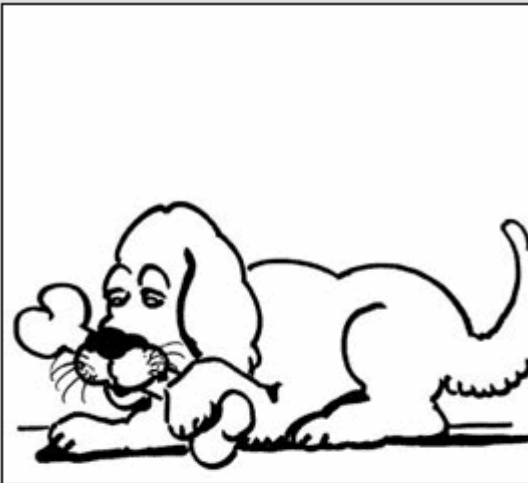


Qui est cette fille ?





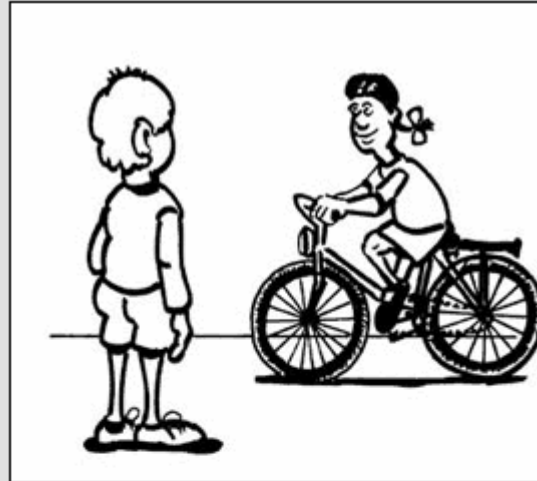
Les enfants mettront leurs chaussures



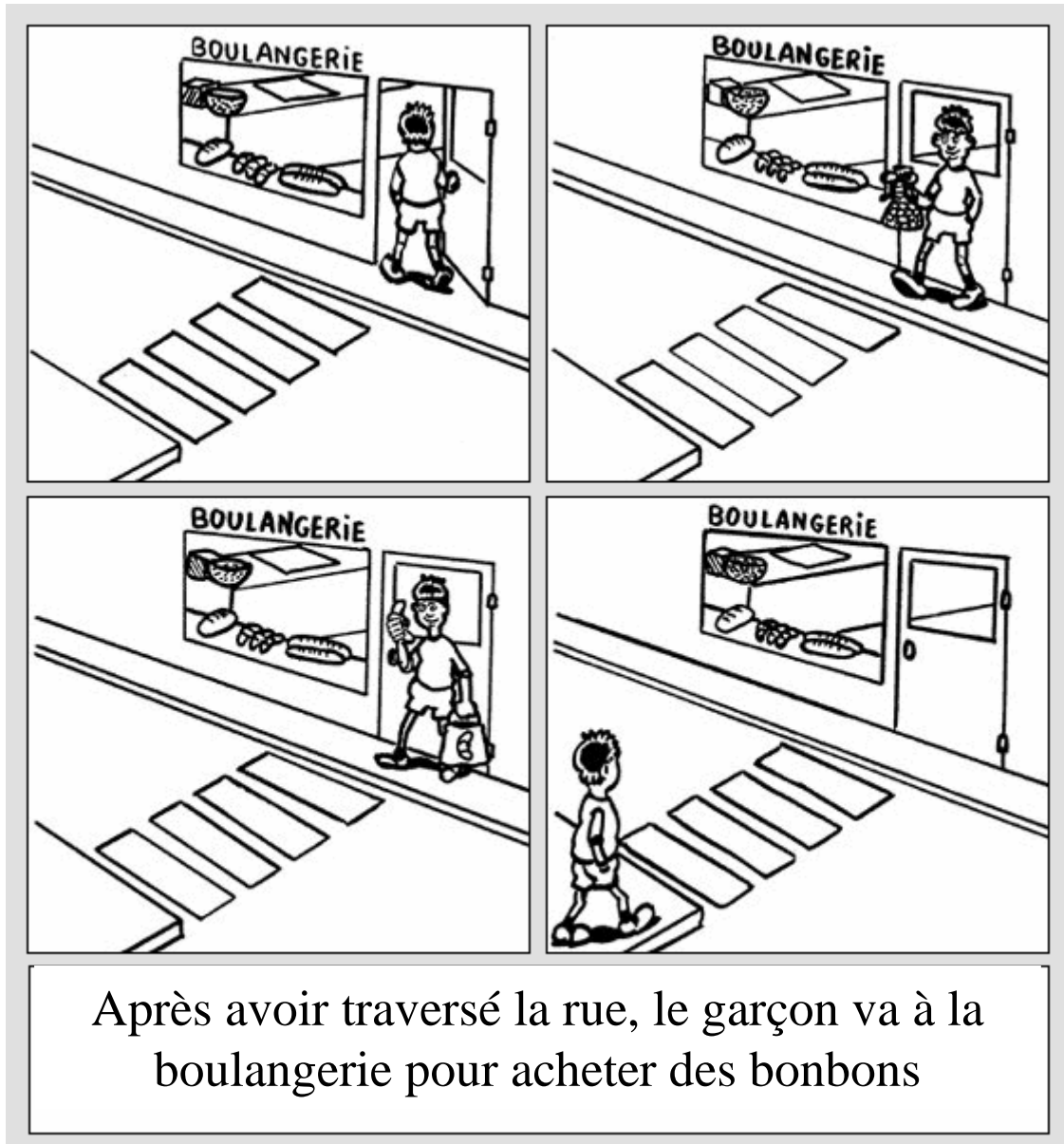
Le chien boit son lait

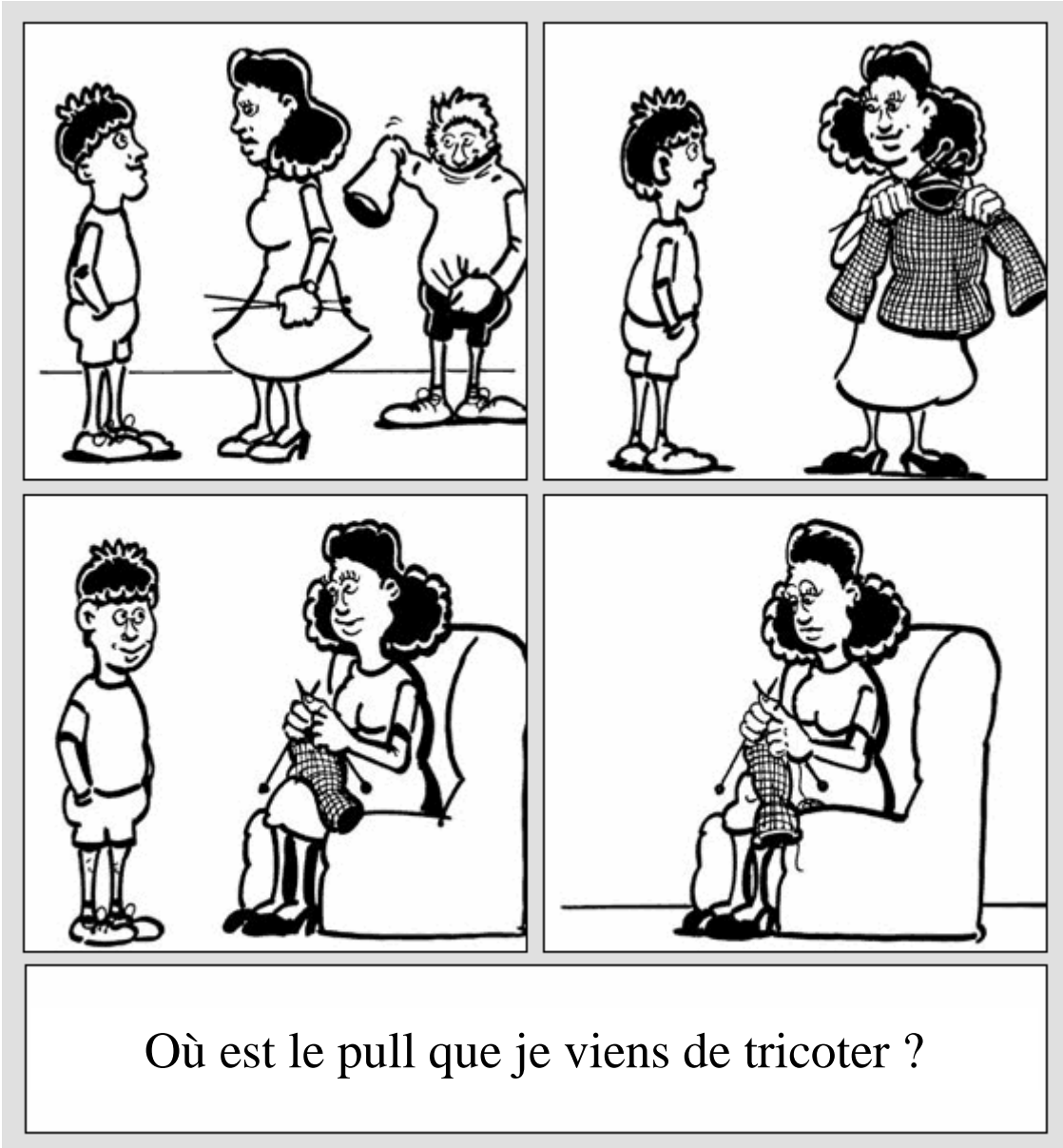


Je mange les cerises que maman cueille



Marie est poussée par Pierre

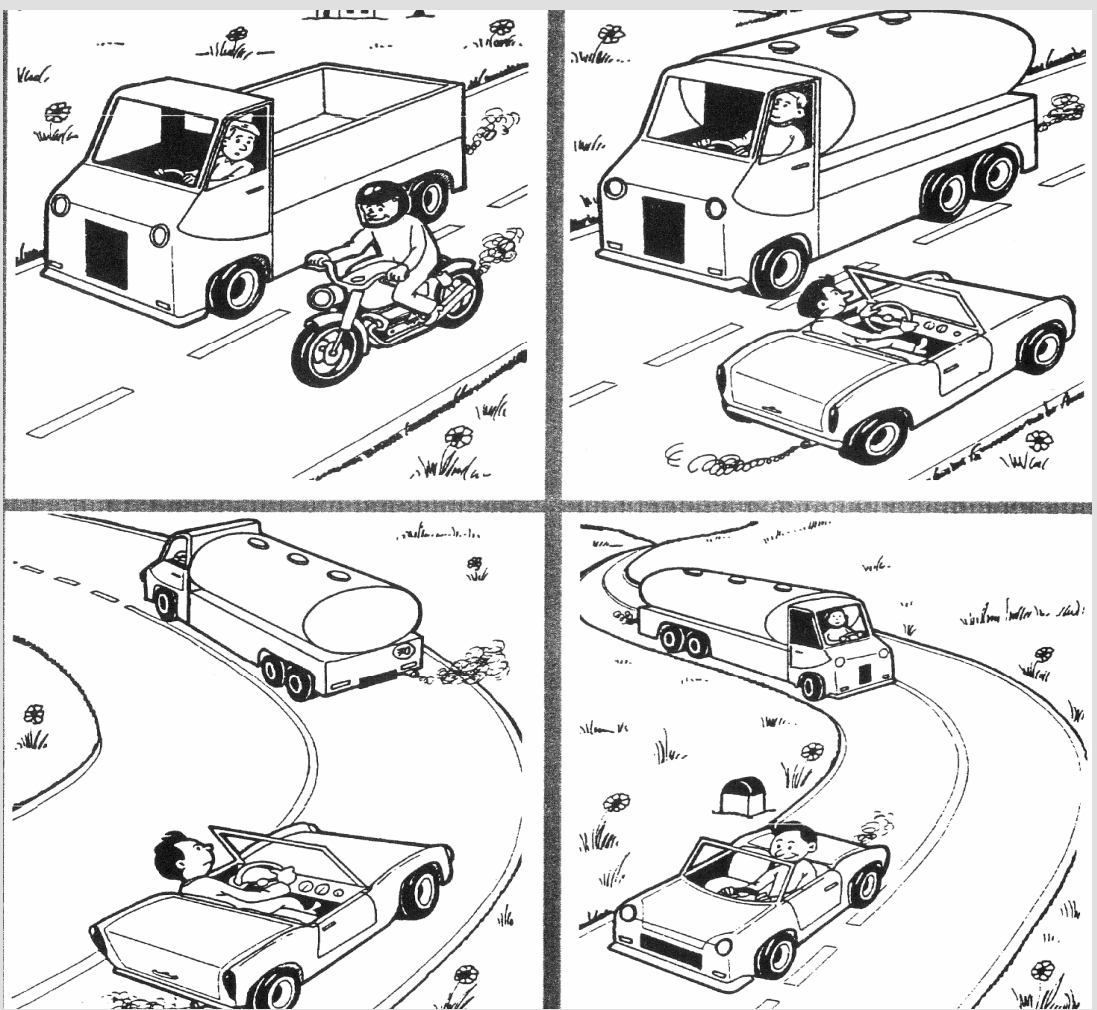




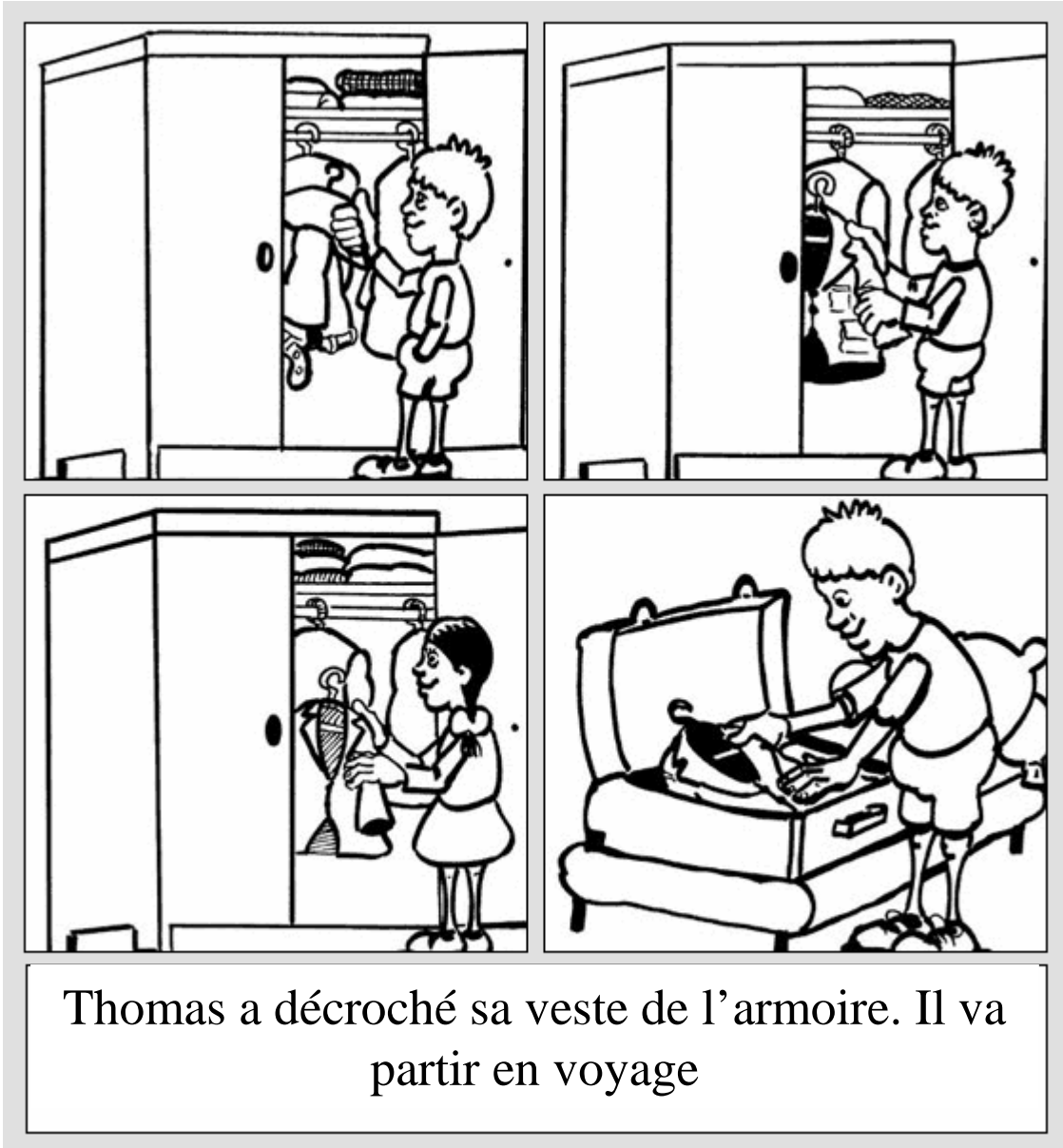


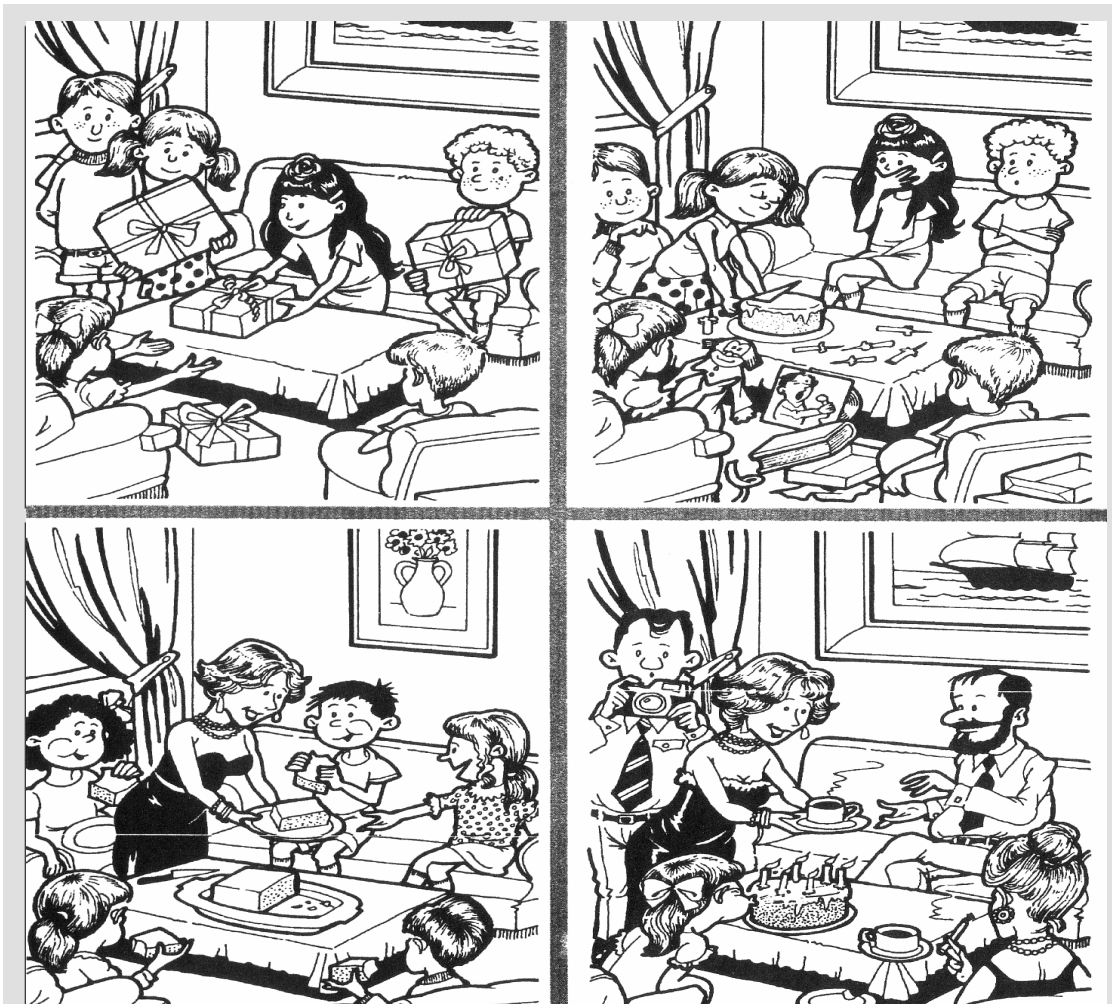
Le chat dont j'ai tiré la queue m'a griffé





Le camion est suivi par la voiture





La fille, dont c'est l'anniversaire, a invité ses amis. Ils vont manger un gâteau



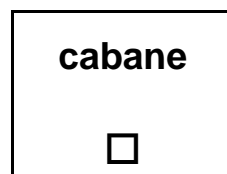
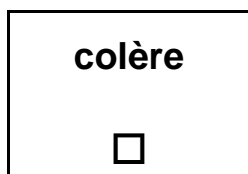
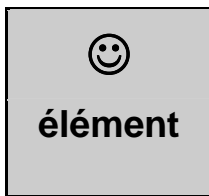
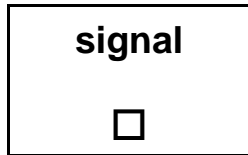
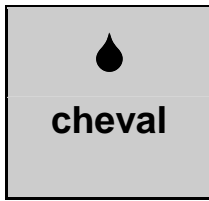
J'aimerais bien aller dehors !

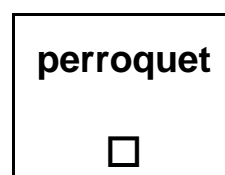
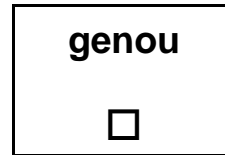
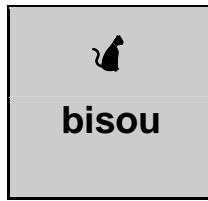


Nicolas n'en a pas envie, mais il va se brosser les dents. Maman lui en a donné l'ordre



Après avoir gonflé le bateau et joué dans l'eau,
Pierre et Marie rentrent chez eux







ballon

cheveu



mouton



époque

boutique



fontaine



mascara

invité



opéra





pizza

lama



loto



grimace

fortune



surface



modèle

pirate



pétale



Item	Unité phonologique	Cible	C.G.	N.S.	N.P.	N.G.	Familiarité**	Lié	C.G.	N.S.	N.P.	N.G.	Familiarité**	Intrus	C.G.	N.S.	N.P.	N.G.	Familiarité**				
1	rime	cheval	n.m	2	5	6	65,35	signal	n.m	2	5	6	55,29	moteur	n.m	2	5	6	58,70				
2	syllabe	élément	n.m	3	5	7	59,32	monument	n.m	3	6	8	51,03	magazine	n.m	3	7	8	54,74				
3	C terminale	histoire	n.f	2	6	8	69,91	colère	n.f	2	5	6	61,70	cabane	n.f	2	5	6	58,94				
4	V terminale	bisou	n.m	2	4	5	47,73	genou	n.m	2	4	5	59,13	gamin	n.m	2	4	5	55,30				
5	rime	tirelire	n.f	2	6	8	50,61	navire	n.m	2	5	6	59,00	bagage	n.m	2	5	6	54,67				
6	syllabe	ustensile	n.m	3	7	9	49,61	domicile	n.m	3	7	8	50,74	perroquet	n.m	3	6	9	56,74				
7	V terminale	ballon	n.m	2	4	6	62,73	mouton	n.m	2	4	6	61,47	cheveu	n.m	2	4	6	62,93				
8	C terminale	époque	n.f	2	4	6	57,64	boutique	n.f	2	5	8	55,89	fontaine	n.f	2	5	8	54,41				
9	syllabe	mascara	n.m	3	7	7	24,18	opéra	n.m	3	5	5	48,26	invité	n.m	3	5	6	53,20				
10	V terminale	pizza	n.f	2	4	5	45,90	lama	n.m	2	4	4	46,36	loto	n.m	2	4	4	51,02				
11	rime	grimace	n.f	2	6	7	52,88	surface	n.f	2	6	7	58,11	fortune	n.f	2	6	7	55,02				
12	C terminale	modèle	n.m	2	5	6	59,38	pétale	n.m	2	5	6	54,17	pirate	n.m	2	5	6	57,60				
Moyenne familiarité							53,77								55,10								56,11

C.G. = Catégorie grammaticale; N.S. = Nombre de syllabes; N.P. = Nombre de phonèmes; N.G. = Nombre de graphèmes

n.m = nom masculin; n.f = nom féminin, adj = adjectif

** Fréquence issue de la base de donnée MANULEX (Lété, Sprenger-Charolles & Colé, 2003) calculée à partir des manuels scolaires du CP au CM2



revoir

refuser



replonger



déformer

dérouler



déchirer



voleur

facteur



coiffeur





caissier

jardinier



menuisier



relire

renifler



refermer



fraîcheur

hauteur



humeur





dégeler

décrire



défaire



beurrier

dentier



papier




refleurir

recopier




repérer




longueur

couleur

douceur


déconnecter

déboutonner

dégringoler


saladier

courrier

fichier

Item	Affixe	Cible	N.S.	N.P.	N.G.	Familiarité**	Lié	N.S.	N.P.	N.G.	Familiarité**	Intrus	N.S.	N.P.	N.G.	Familiarité**
1	re-	revoir	2	6	6	56,46	replonger	3	7	9	43,93	refuser	3	6	7	58,07
2	dé-	déformer	3	7	8	46,93	dérouler	3	6	8	56,07	déchirer	3	6	8	56,17
3	-eur	voleur	2	5	6	60,39	coiffeur	2	6	8	53,20	facteur	2	6	7	57,43
4	-ier	caissier	2	6	8		jardinier	3	8	9	57,53	menuisier	3	8	9	52,82
5	re-	relire	2	5	6	58,48	refermer	3	7	8	57,06	renifler	3	7	8	52,23
6	-eur	fraîcheur	2	6	9	49,85	hauteur	2	5	7	57,87	humeur	2	4	6	55,57
7	dé-	dégeler	2	6	7	40,37	défaire	2	5	7	51,71	décrire	2	6	7	59,59
8	-ier	beurrier	2	5	8	17,15	dentier	2	5	7	42,52	papier	2	5	6	64,48
9	re-	refleurir	3	8	9	40,34	recopier	3	7	8	65,11	repérer	3	6	7	58,54
10	-eur	longueur	2	5	8	56,42	douceur	2	5	7	54,17	couleur	2	5	7	66,38
11	dé-	déconnecter	4	9	11	34,95	déboutonner	4	8	11	43,17	dégringoler	4	9	11	51,38
12	-ier	saladier	3	7	8	51,89	fichier	2	5	7	36,69	courrier	2	5	8	54,84

Moyenne familiarité 46,66

51,59

57,29

N.S. = Nombre de syllabes; N.P. = Nombre de phonèmes; N.G. = Nombre de graphèmes

** Fréquence issue de la base de donnée MANULEX (Lété, Sprenger-Charolles & Colé, 2003) calculée à partir des manuels scolaires du CP au CM2



nous tremblons

nous suivrons



nous couvrons



petite

touriste



cliente



vous travaillez

vous poursuivrez



vous découvrez





muette

patte



chatte



nous commençons

nous illustrons



nous abattons



courte

prête



vaste





vous participez

vous enregistrez



vous apparaîtrez



forte

cuite



triste



nous attendons

nous permettrons



nous rencontrons





idiot

fleuriste



servant



vous cueillez

vous pleurez



vous lirez



puissante

inquiète



infecte



Item	Flexion	Cible	C.G.*	N.S.	N.P.	N.G.	Familiarité**	Lié	C.G.*	N.S.	N.P.	N.G.	Familiarité**	Intrus	C.G.*	N.S.	N.P.	N.G.	Familiarité**					
1	Conjugaison verbale	nous tremblons	g1	2	6	9	60,29	nous couvrons	g3	2	5	8	61,03	nous suivrons	g3	2	6	8	66,63					
2	féminin	courte	adj	1	4	6	61,58	prête	adj	1	4	5	60,35	vaste	adj	1	4	5	55,52					
3	Conjugaison verbale	vous travaillez	g1	3	8	10	63,41	vous découvrez	g3	3	7	9	63,97	vous poursuivrez	g3	3	9	11	58,68					
4	féminin	forte	adj	1	4	5	61,96	cuite	adj	1	4	5	53,80	triste	adj	1	5	6	61,45					
5	Conjugaison verbale	nous commençons	g1	3	6	10	67,00	nous illustrons	g1	3	7	10	57,26	nous abattons	g3	3	6	9	56,59					
6	féminin	puissante	adj	2	6	9	58,63	inquiète	adj	2	5	8	59,12	infecte	adj	2	5	7	56,82					
7	Conjugaison verbale	vous participez	g1	4	9	10	55,80	vous enregistrez	g1	4	9	11	52,80	vous apparaîtrez	g3	4	8	11	62,08					
8	féminin	petite	n	2	5	6	75,52	cliente	n	2	5	7	59,00	touriste	n	2	6	8	51,94					
9	Conjugaison verbale	nous attendons	g3	3	5	9	66,84	nous rencontrons	g1	3	7	11	63,33	nous permettrons	g3	3	8	11	63,82					
10	féminin	muette	adj	1	4	6	57,23	chatte	n	1	3	6	69,79	patte	n	1	3	5	65,57					
11	Conjugaison verbale	vous cueillez	g3	2	5	8	60,63	vous pleurez	g1	2	5	7	63,54	lirez	g3	2	4	5	70,70					
12	féminin	idiote	n	2	5	6	51,01	servante	n	2	6	8	49,69	fleuriste	n	2	7	9	48,26					
							Moyenne familiarité	61,66								59,47								59,84

C.G. = Catégorie grammaticale; N.S. = Nombre de syllabes; N.P. = Nombre de phonèmes; N.G. = Nombre de graphèmes

* Dans cette colonne, est également précisé le groupe des verbes conjugués

** Fréquence issue de la base de donnée MANULEX (Lété, Sprenger-Charolles & Colé, 2003) calculée à partir des manuels scolaires du CP au CM2

Les valeurs données sont celles des verbes à l'infinif et des noms masculins



le chat pêche des poissons

la neige casse des sapins



la pluie tombe des nuages



il a oublié de vendre

il a besoin de marcher



il a fini de manger



il est tranquille

il est amusant



il est arrivé





il trouve des billets

il rentre des vacances



il chante des chansons



il est nouveau

il est resté



il est poli



il vend des voitures

il jette des chapeaux



il parle des copains





il a ronflé

il a peur



il a fui



**le garçon attrape des
balles**

**le gardien carresse des
chats**



le voleur profite des gens



il a peint

il a ri



il a honte





**il choisit des couleurs
magnifiques**

**il décide des travaux
importants**



**il explique des projets
inutiles**



il est étrange

il est malin



il est tombé



il entend des paroles

il surgit des buissons



il traverse des jardins



Item	Structure syntaxique étudiée	Cible	N.S. /mot	N.G. /mot	Familiarité* *	Lié	N.S. /mot	N.G. /mot	Familiarité**	Intrus	N.S. /mot	N.G. /mot	Familiarité* *
1	Transivité (D/I)	le chat pêche des poissons	1-1-1-1-2	2-4-5-3-8	64,21	la neige casse des sapins	1-1-1-1-2	2-5-5-3-6	62,16	la pluie tombe des nuages	1-1-1-1-2	2-5-5-3-6	64,82
2	Avoir (passé composé)	il a oublié de vendre	1-1-3-1-2	2-1-6-2-6	63,35	il a fini de manger	1-1-2-1-2	2-1-4-2-6	67,28	il a besoin de marcher	1-1-2-1-2	2-1-6-2-7	64,12
3	Etre (V+attribut)	il est tranquille	1-1-2	2-3-10	58,16	il est amusant	1-1-3	2-3-8	56,21	il est arrivé	1-1-3	2-3-6	60,51
4	Transivité (D/I)	il trouve des billets	1-2-1-2	2-6-3-7	64,76	il chante des chansons	1-1-1-2	2-6-3-8	60,97	il rentre des vacances	1-1-1-3	2-6-3-8	64,03
5	Etre (V+attribut)	il est nouveau	1-1-2	2-3-7	61,51	il est poli	1-1-2	2-3-4	53,15	il est resté	1-1-2	2-3-5	55,83
6	Transivité (D/I)	il vend des voitures	1-1-1-2	2-4-3-8	64,31	il jette des chapeaux	1-1-1-2	2-5-3-8	63,80	il parle des copains	1-2-1-2	2-6-3-7	64,54
7	Avoir (passé composé)	il a ronflé	1-1-2	2-1-6	70,70	il a fui	1-1-1	2-1-3	57,86	il a peur	1-1-1	2-1-4	67,27
8	Transivité (D/I)	le garçon attrape des balles	1-2-2-1-1	2-6-7-3-6	63,39	le gardien caresse des chats	1-2-2-1-1	2-7-8-3-5	61,83	le voleur profite des gens	1-2-2-1-1	2-6-7-3-4	60,86
9	Avoir (passé composé)	il a peint	1-1-1	2-1-5	59,83	il a ri	1-1-1	2-1-2	65,91	il a honte	1-1-1	2-1-5	54,55
10	Transivité (D/I)	il choisit des couleurs magnifiques	1-2-1-2-3	2-7-3-8-11	63,97	il explique des projets inutiles	1-2-1-2-3	2-8-3-7-8	57,77	il décide des travaux importants	1-2-1-2-3	2-6-3-7-10	63,42
11	Etre (V+attribut)	il est étrange	1-1-2	2-3-7	59,11	il est malin	1-1-2	2-3-5	56,32	il est tombé	1-1-2	2-3-5	57,54
12	Transivité (D/I)	il entend des paroles	1-2-1-2	2-6-3-7	62,50	il traverse des jardins	1-2-1-2	2-8-3-7	64,55	il surgit des buissons	1-2-1-2	2-6-3-8	55,78

Moyenne familiarité 62,98

60,65

61,11

** Fréquence moyenne pour l'ensemble des mots de la phrase calculée à partir de la base de donnée MANULEX (Lété, Sprenger-Charolles & Colé, 2003)

Annexe 33


Epreuve explicite de détection d'intrus graphophonologique en français

Entraînements :


maison

blouson

château


sol

fête

ville



difficile

imbécile



intérêt



cadi

vélo



midi



bonjour

silence



tambour





tempête

visite



girafe



caméra

insolite



angora



villa

schéma



coucher





vitrine

mesure



racine



échelle

parole



police



vendredi

canapé



bigoudi





canal

métal



désir



antenne

musique



machine



chemin

lapin



tapis



Item	Unité phonologique	Cible	C.G.	N.S.	N.P.	N.G.	Familiarité**	Lié	C.G.	N.S.	N.P.	N.G.	Familiarité**	Intrus	C.G.	N.S.	N.P.	N.G.	Familiarité**				
1	syllabe	difficile	adj	3	7	9	61,59	imbécile	n.m/f	3	6	8	51,98	intérêt	n.m	3	5	7	51,63				
2	V terminale	cadi	n.m	2	4	4	39,54	midi	n.m	2	4	4	59,14	vélo	n.m	2	4	4	61,93				
3	rime	bonjour	n.m	2	5	7	61,50	tambour	n.m	2	5	7	57,91	silence	n.m	2	5	7	61,27				
4	C terminale	tempête	n.f	2	5	7	59,38	visite	n.f	2	5	6	59,16	girafe	n.f	2	5	6	58,05				
5	syllabe	caméra	n.f	3	6	6	53,25	angora	adj	3	5	6	41,80	infini	adj	3	5	6	47,40				
6	V terminale	villa	n.f	2	4	5	50,77	schéma	n.m	2	4	6	50,80	coucher	n.m	2	4	7	51,65				
7	rime	vitrine	n.f	2	6	7	53,58	racine	n.f	2	5	6	52,26	mesure	n.f	2	5	6	56,23				
8	C terminale	échelle	n.f	2	4	7	57,90	parole	n.f	2	5	6	61,23	police	n.f	2	5	6	59,06				
9	syllabe	vendredi	n.m	3	7	8	55,51	étourdi	n.m	3	6	7	49,81	acrobate	n.m	3	7	8	53,21				
10	rime	canal	n.m	2	5	5	51,71	métal	n.m	2	5	5	56,47	désir	n.m	2	5	5	53,05				
11	C terminale	antenne	n.f	2	4	7	53,82	machine	n.f	2	5	7	60,97	musique	n.f	2	5	7	62,61				
12	V terminale	chemin	n.m	2	4	6	64,22	lapin	n.m	2	4	5	63,75	tapis	n.m	2	4	5	59,85				
Moyenne familiarité							55,23								55,44								56,33

C.G. = Catégorie grammaticale; N.S. = Nombre de syllabes; N.P. = Nombre de phonèmes; N.G. = Nombre de graphèmes

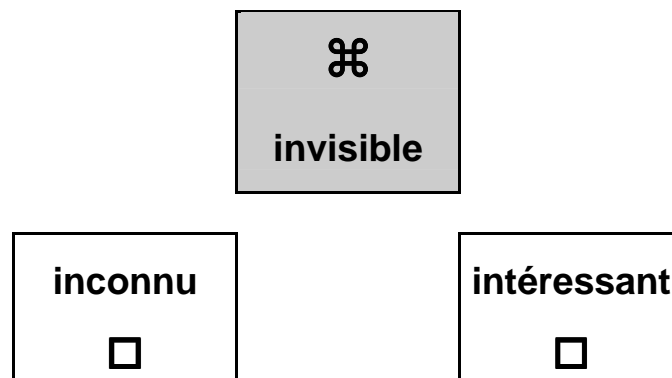
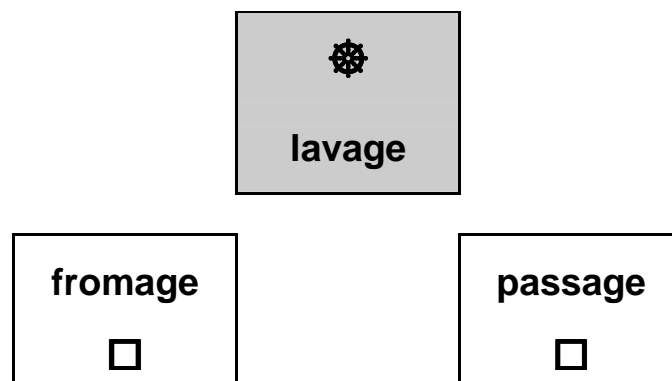
n.m = nom masculin; n.f = nom féminin, adj = adjectif

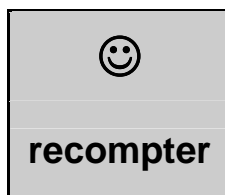
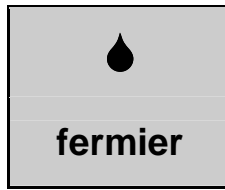
** Fréquence issue de la base de donnée MANULEX (Lété, Sprenger-Charolles & Colé, 2003) calculée à partir des manuels scolaires du CP au CM2

Annexe 34

Epreuve explicite de détection d'intrus grapho-morpho-dérivationnel en français

Entraînements :







dévisser

détacher



déguiser



ressortir

replacer



reculer



vacancier

cavalier



couturier





débrancher

développer



décharger



nageur

chanteur



chauffeur



redonner

regarder



retrouver





cendrier

boîtier



cahier



grandeur

vapeur



lenteur



déboucher

déplacer



décider



Item	Affixe	Cible	N.S.	N.P.	N.G.	Familiarité**	Lié	N.S.	N.P.	N.G.	Familiarité**	Intrus	N.S.	N.P.	N.G.	Familiarité**					
1	-ier	fermier	2	6	7	57,83	cuisinier	3	8	9	57,28	cordonnier	3	8	10	50,02					
2	re-	recompter	3	6	6	56,46	revenir	3	7	7	66,39	recevoir	3	8	8	63,32					
3	-eur	pêcheur	2	5	7	59,02	coureur	2	5	7	55,08	seigneur	2	5	8	56,38					
4	dé-	dévisser	3	6	8	42,22	détacher	3	6	8	56,10	déguiser	3	6	8	55,64					
5	re-	ressortir	3	8	9	53,08	replacer	3	7	8	55,63	reculer	3	6	7	57,75					
6	-ier	vacancier	3	7	9	40,30	couturier	3	7	9	49,49	cavalier	3	7	8	54,19					
7	dé-	débrancher	3	7	10	44,62	décharger	3	7	9	46,83	développer	3	7	10	53,51					
8	-eur	nageur	2	5	6	50,95	chanteur	2	5	8	54,83	chauffeur	2	5	9	56,43					
9	re-	redonner	3	6	8	49,42	retrouver	3	7	9	67,66	regarder	3	7	8	70,32					
10	-ier	cendrier	3	6	8	42,26	boîtier	2	6	7	36,63	cahier	2	4	6	61,14					
11	-eur	grandeur	2	6	8	47,68	lenteur	2	5	7	46,79	vapeur	2	5	6	54,48					
12	dé-	déboucher	3	6	9	50,19	déplacer	3	7	8	58,87	décider	3	6	7	64,76					
						Moyenne familiarité	49,50							54,30							58,16


C.G. = Catégorie grammaticale; N.S. = Nombre de syllabes; N.P. = Nombre de phonèmes; N.G. = Nombre de graphèmes

** Fréquence issue de la base de donnée MANULEX (Lété, Sprenger-Charolles & Colé, 2003) calculée à partir des manuels scolaires du CP au CM2

Annexe 35

Epreuve explicite de détection d'intrus grapho-morpho-flexionnel en français

Entraînements :


froide

malade

grande


je crie

je mis

je lis



haute

brute



sainte



vous cassez

vous livrez



vous vivrez



morte

juste



plate





nous sonnons

nous barrons



nous croirons



complète

adroite



injuste



vous défendez

vous conduirez



vous soupirez





gratuite

absente



adulte



nous laissons

nous boirons



nous tirons



glissante

ouverte



honnête





vous confondez

vous étirez



vous inscrirez



interdite

égoïste



avocate



nous rejoignons

nous retirons



nous produirons



Item	Flexion	Cible	C.G.*	N.S.	N.P.	N.G.	Familiarité**	Lié	C.G.*	N.S.	N.P.	N.G.	Familiarité**	Intrus	C.G.*	N.S.	N.P.	N.G.	Familiarité**
2	Conjugaison verbale	vous cassez	g1	2	4	6	61,28	vous livrez	g1	2	5	6	56,11	vous vivrez	g3	2	5	6	66,58
10	Conjugaison verbale	vous confondez	g3	3	6	9	55,16	vous étirez	g1	3	5	6	54,13	vous inscrirez	g3	3	7	9	55,02
Item	Conjugaison verbale	vous défendez	g3	3	6	8	56,89	vous soupirez	g1	3	6	8	56,47	vous conduirez	g3	3	7	9	61,00
1	féminin	haute	adj	1	3	5	62,75	sainte	adj	1	3	6	45,48	brute	adj	1	4	5	45,29
3	féminin	morte	adj	1	4	5	56,42	plate	adj	1	4	5	58,17	juste	adj	1	4	5	60,57
4	Conjugaison verbale	nous sonnons	g1	2	4	7	59,68	nous barrons	g1	2	4	7	58,37	nous croirons	g3	2	6	8	66,53
5	féminin	complète	adj	2	6	8	61,91	adroite	adj	2	6	7	53,79	injuste	adj	2	5	7	45,59
7	féminin	gratuite	adj	2	7	8	52,23	absente	adj	2	5	7	51,59	adulte	adj	2	5	6	50,35
8	Conjugaison verbale	nous laissons	g1	2	4	8	67,03	nous tirons	g1	2	4	6	65,51	nous boirons	g3	2	5	7	64,20
9	féminin	glissante	adj	2	6	9	50,97	ouverte	adj	2	5	7	58,35	honnête	adj	2	4	7	51,95
11	féminin	interdite	adj	3	7	9	49,89	avocate	n	3	6	7	51,11	égoïste	n	3	6	7	44,23
12	Conjugaison verbale	nous rejoignons	g3	3	7	10	58,77	nous retirons	g1	3	6	8	58,73	nous produirons	g3	3	8	10	58,29

Moyenne familiarité 57,75

55,65

55,80

C.G. = Catégorie grammaticale; N.S. = Nombre de syllabes; N.P. = Nombre de phonèmes; N.G. = Nombre de graphèmes

* Dans cette colonne, est également précisé le groupe des verbes conjugués

** Fréquence issue de la base de donnée MANULEX (Lété, Sprenger-Charolles & Colé, 2003) calculée à partir des manuels scolaires du CP au CM2

Les valeurs données sont celles des verbes à l'infinitif et des noms masculins

Annexe 36

Epreuve explicite de détection d'intrus grapho-syntaxique en français

Entraînements :



tu mets tes chaussures

tu lis un livre



tu vas au parc



tu as couru

tu as faim



tu as mangé





il met des chaussettes

il beurre des tartines



il joue des cymbales



il a menti au voisin

il a joué au ballon



Il a droit au dessert



**les musiciens gagnent
des chaises**

**les étudiants sortent
des salles**



**les amoureux cueillent
des fleurs**





il est bizarre

il est âgé



il est allé



il raconte des blagues

il revient des courses



il reprend des pommes



**il a essayé de
comprendre**

il a envie de dormir



il a prévu de sortir





il lance des jouets

il vend des poules



il souffre des jambes



il est aveugle

il est entré



il est énorme



il lave des pantalons

**il cherche des
camarades**



il doute des résultats





il a tenté de voler

il a raison de refuser



il a choisi de continuer



**les sorciers ramassent
des araignées**

**les soldats triomphent
des ennemis**



**les messieurs délivrent
des animaux**



il est étonnant

il est agréable



il est apparu



Item	Structure syntaxique étudiée	Cible	N.S. /mot	N.G. /mot	Familiarité **	Lié	N.S. /mot	N.G. /mot	Familiarité **	Intrus	N.S. /mot	N.G. /mot	Familiarité **
1	Transitivité (D/I)	il met des chaussettes	1-1-1-2	2-3-3-11	63,39	il beurre des tartines	1-1-1-2	2-6-3-8	49,74	il joue des cymbales	1-1-1-2	2-4-3-8	56,92
2	Avoir (passé composé)	il a menti au voisin	1-1-2-1-2	2-1-5-2-6	58,86	il a joué au ballon	1-1-2-1-2	2-1-4-2-6	66,07	Il a droit au dessert	1-1-1-1-2	2-1-5-2-7	58,03
3	Transitivité (D/I)	les musiciens gagnent des chaises	1-3-1-1-1	3-9-7-3-7	59,71	les amoureux cueillent des fleurs	1-3-1-1-1	3-8-9-3-6	58,33	les étudiants sortent des salles	1-3-1-1-1	3-9-7-3-6	65,56
4	Etre (V+attribut)	il est bizarre	1-1-2	2-3-7	58,14	il est âgé	1-1-2	2-3-3	52,28	il est allé	1-1-2	2-3-4	60,19
5	Transitivité (D/I)	il raconte des blagues	1-2-1-1	2-7-3-7	58,61	il reprend des pommes	1-2-1-1	2-7-3-6	64,09	il revient des courses	1-2-1-2	2-7-3-7	64,78
6	Avoir (passé composé)	il a essayé de comprendre	1-1-3-1-3	2-1-6-2-10	66,19	il a prévu de sortir	1-1-2-1-2	2-1-5-2-6	63,63	il a envie de dormir	1-1-2-1-2	2-1-5-2-6	63,59
7	Transitivité (D/I)	il lance des jouets	1-1-1-2	2-5-3-6	61,37	il vend des poules	1-1-1-1	2-4-3-6	62,31	il souffre des jambes	1-2-1-1	2-7-3-6	59,17
8	Etre (V+attribut)	il est aveugle	1-1-2	2-3-7	55,43	il est énorme	1-1-2	2-3-6	60,98	il est entré	1-1-2	2-3-5	53,84
9	Transitivité (D/I)	il lave des pantalons	1-1-1-3	2-4-3-9	59,95	il cherche des camarades	1-2-1-3	2-7-3-9	65,29	il doute des résultats	1-1-1-3	2-5-3-9	56,46
10	Avoir (passé composé)	il a tenté de voler	1-1-2-1-2	1-2-5-2-5	61,05	il a choisi de continuer	1-1-2-1-3	2-1-6-2-9	62,29	il a raison de refuser	1-1-2-1-3	2-1-6-2-7	61,90
11	Transitivité (D/I)	les sorciers ramassent des araignées	1-2-2-1-3	3-7-9-3-9	59,89	les messieurs délivrent des animaux	1-1-3-1-3	3-8-9-3-7	57,65	les soldats triomphent des ennemis	1-2-2-1-3	3-7-10-3-7	54,76
12	Etre (V+attribut)	il est étonnant	1-1-3	2-3-8	52,18	il est agréable	1-1-3	2-3-8	57,03	il est apparu	1-1-3	2-3-6	41,10

Moyenne familiarité 59,56

59,97

58,02

** Fréquence moyenne pour l'ensemble des mots de la phrase calculée à partir de la base de donnée MANULEX (Lété, Sprenger-Charolles & Colé, 2003)

Annexe 37
Epreuve grapho-morphologique implicite de
plausibilité lexicale écrite en français

Qu'est ce qui ressemble le plus à un mot français

dépunir <input type="checkbox"/>	déconvir <input type="checkbox"/>
jupore <input type="checkbox"/>	vasier <input type="checkbox"/>
socalmer <input type="checkbox"/>	calsomer <input type="checkbox"/>
piscinier <input type="checkbox"/>	réponsore <input type="checkbox"/>
tosomber <input type="checkbox"/>	sotomber <input type="checkbox"/>
gondriake <input type="checkbox"/>	gonkadier <input type="checkbox"/>
vafirier <input type="checkbox"/>	visagier <input type="checkbox"/>

sovoler <input type="checkbox"/>	découner <input type="checkbox"/>
branchore <input type="checkbox"/>	chobrane <input type="checkbox"/>
vécondir <input type="checkbox"/>	dévoncir <input type="checkbox"/>
dridinane <input type="checkbox"/>	nindadier <input type="checkbox"/>
désaner <input type="checkbox"/>	somanger <input type="checkbox"/>
dégenper <input type="checkbox"/>	gendéper <input type="checkbox"/>
tartore <input type="checkbox"/>	fruvier <input type="checkbox"/>
sopasser <input type="checkbox"/>	dédonner <input type="checkbox"/>

pralome <input type="checkbox"/>	lampore <input type="checkbox"/>
-------------------------------------	-------------------------------------

débouver <input type="checkbox"/>	désauver <input type="checkbox"/>
--------------------------------------	--------------------------------------

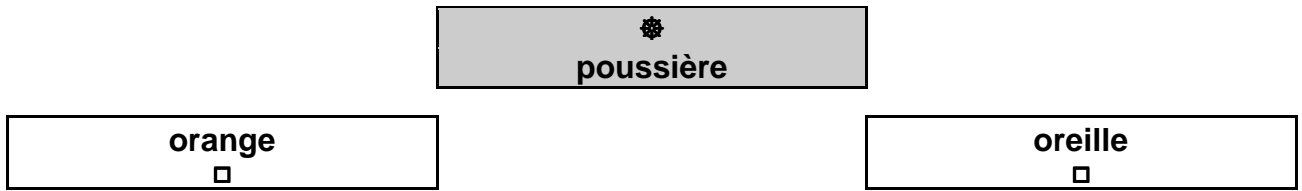
robier <input type="checkbox"/>	ropier <input type="checkbox"/>
------------------------------------	------------------------------------

bripier <input type="checkbox"/>	portore <input type="checkbox"/>
-------------------------------------	-------------------------------------

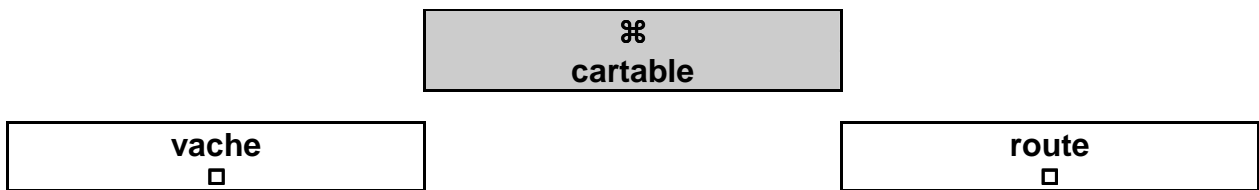
débouger <input type="checkbox"/>	sodanser <input type="checkbox"/>
--------------------------------------	--------------------------------------

Annexe 38

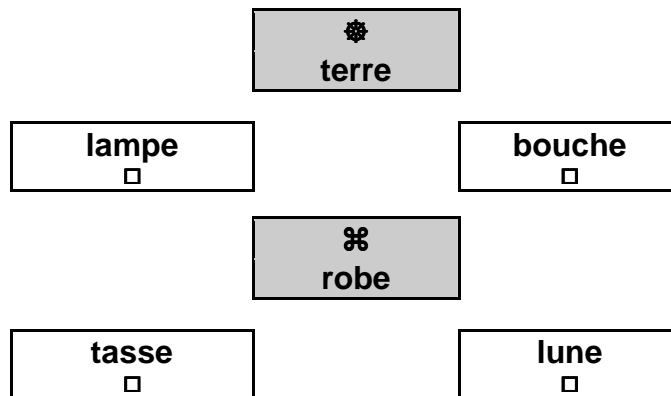
Exemple donné à l'épreuve implicite graphophonologique



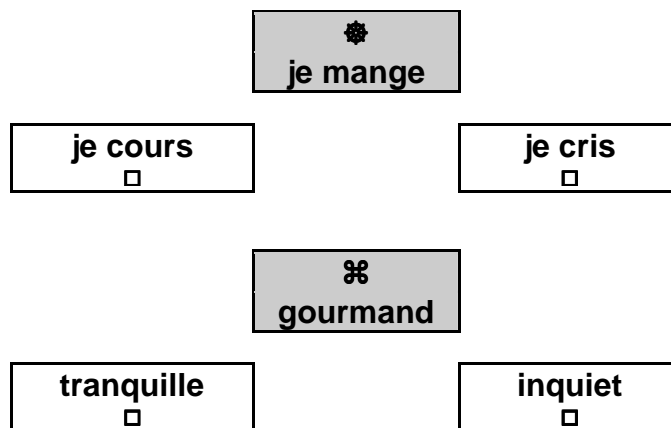
Exemple donné aux épreuves implicites grapho-morphologiques



Items d'entraînement graphophonologiques et grapho-morpho-dérivationnels implicites



Items d'entraînement grapho-morpho-flexionnels implicites



Exemple donné à l'épreuve implicite grapho-syntaxique

⊛
j'achète des bonbons

je prends un bain
□

je lis un livre
□

Items d'entraînement grapho-syntaxiques implicites

⊛
il regarde un film

il prépare une surprise
□

il conduit une voiture
□

⊛
je cherche un crayon

je plante des fleurs
□

je porte des robes
□

Annexe 39

Résultats aux tests t (niveau du hasard/6) obtenus aux épreuves implicites graphophonologique, grapho-morpho-dérivationnel, grapho-syntaxique et à l'épreuve explicite grapho-morpho-flexionnel en français en fonction du niveau scolaire et de la langue première

	Arabophones		
Epreuves de détection d'intrus	Niv A	Niv B	Niv C
Implicite grapho-morpho-dérivationnel	6,15 (1,76) t(39)= 0,54; ns	6,24 (1,75) t(36)= 0,84; ns	5,83 (1,68) t(41)= -0,64; ns
Implicite grapho-syntaxique	5,88 (1,73) t(39)= -0,46; ns	5,84 (1,77) t(36)= -0,56; ns	5,38 (1,53) t(41)= -2,62; p<.01
Explicite grapho-morpho-flexionnel	6,61 (1,88) t(40)=2,07; p<.05	7,42(1,61) t(41)=5,17; p<.001	7,67 (2,00) t(41)=5,41; p<.001
	Lusophones		
Epreuves de détection d'intrus	Niv A	Niv B	Niv C
Implicite grapho-morpho-dérivationnel	5,54 (1,97) t(38)=-1,46; ns	5,71 (1,76) t(20)=-0,74; ns	5,04 (1,55) t(22)=-2,96; p<.01
Implicite grapho-syntaxique	5,72 (1,73) t(38)=-1,02; ns	5,86 (1,96) t(20)=-0,33; ns	5,65 (1,90) t(22)=-0,88; ns
Explicite grapho-morpho-flexionnel	5,53 (1,78) t(37)=-1,64; ns	5,95 (2,01) t(20)=-0,11; ns	7,35 (1,75) t(22)=3,70; p<.001

Annexe 40

Coefficients de corrélation entre les différents scores et le niveau de compréhension de phrases chez les arabophones et les lusophones de niveau A, B et C

Chez les arabophones

	Détection d'intrus en arabe L1					Ident. Const. Morpho pseudo-mots	Lecture mots L1	Comp. phrases L1	Détection d'intrus en français L2					Plausi.	Voca.	Lecture mots FL2
	Imp G-phono.	Imp G-synt	Exp G-morpho-dériv	Exp G-morpho-flex	Exp G-synt.				Imp G-phono.	Imp G-morpho-flex.	Exp G-morpho-dériv.	Exp G-morpho-flex.	Exp G-syntax.			
Niveau A (N = 32)	.25		.11	-.10	-.14	.33 ^t	.33 ^t	.29	.30	.25	.10	.34 ^t	.21	.11	.30	.68***
Niveau B (N = 32)	-.45**		.21	.05	.16	.11	.06	.56***			.50**	-.31		.06	.48**	.30
Niveau C (N = 41)			.37*	.34*	.41**	.35*	.32*	.15	.22		.26	.39*	.30 ^t	.55***	.59***	.65***

*** r significatif à $p < .001$, ** à $p < .01$ * à $p < .05$ ^t tendanciel à $p < .07$

Chez les lusophones

	Détection d'intrus en arabe L1					Ident. Const. Morpho pseudo-mots	Lecture mots L1	Comp. phrases L1	Détection d'intrus en français L2					Plausi.	Voca.	Lecture mots FL2
	Imp G-phono.	Imp G-morpho-flex.	Imp G-synt	Exp G-morpho-dériv	Exp G-morpho-flex				Exp G-synt.	Imp G-phono.	Imp G-morpho-flex.	Exp G-morpho-dériv.	Exp G-morpho-flex.			
Niveau A (N = 31)	.21		-.11	.10	.11	.13	.09	.18	-.02	.16	.20			.09	.44*	.18
Niveau B (N = 20)				.23	.29	-.24	.50*	.31	.38	-.04		-.15		-.07	-.17	.20
Niveau C (N = 19)	.15	.48	.42	.37			.70***	.07	.40	-.03	.49*	.13	.41	.21	.05	-.01

*** r significatif à $p < .001$, ** à $p < .01$ * à $p < .05$

RESUME

Cette recherche interlangue a pour premier objectif d'analyser l'évolution des connaissances graphophonologiques, grapho-morphologiques et grapho-syntaxiques, implicites et explicites et la contribution de ces connaissances à la lecture en arabe vocalisé et en portugais langue première. Les résultats indiquent que les élèves de langue arabe se focalisent davantage sur la structure morpho-dérivationnelle en racine/schème que sur le marquage flexionnel, alors que les élèves de langue portugaise développent, non seulement une sensibilité phonologique à la rime et à la syllabe mais aussi de meilleures connaissances grapho-morpho-flexionnelles que grapho-morpho-dérivationnelles. Par ailleurs, si les connaissances morphologiques jouent un rôle sur la reconnaissance des mots écrits quelle que soit la langue, on remarque que ce sont celles les plus décontextualisées qui interviennent en arabe. La contribution des connaissances grapho-morpho-dérivationnelles à la compréhension de phrases en arabe souligne là encore l'importance de la morphologie dans cette langue écrite. Enfin, la compréhension en lecture apparaît plus dépendante des processus d'automatisation en portugais, du fait de l'opacité relative de son orthographe.

Le deuxième objectif de ce travail est d'étudier le rôle des acquisitions de la lecture en langue première (arabe vs. portugais) sur l'apprentissage de la lecture en français langue seconde. Les comparaisons ont permis de conclure que le statut privilégié de la morphologie en arabe peut inciter les apprenants à porter attention à la structure interne des mots français pour les lire. En revanche, les similitudes entre le portugais et le français facilitent le développement de la sensibilité graphophonologique et amplifient la symétrie entre les habiletés de lecture de mots en L1 et en L2. Les profils de compréhension en français L2 et leur évolution sont également très différents entre arabophones et lusophones.

MOTS CLES : Etude interlangue arabe vs. portugais - Français langue seconde - Apprentissage de la lecture - Reconnaissance des mots écrits - Compréhension en lecture - Connaissances implicites / explicites- Morphologie.

TITLE: Linguistic features and learning to read in first language and French as a second language: Developmental and Crosslinguistic perspective between Arabic and Portuguese.

ABSTRACT

First, this crosslinguistic study aimed to investigate graphophonological, grapho-morphological and grapho-syntactical implicit and explicit knowledge development and the contribution of this knowledge to reading in vocalised Arabic and Portuguese mother tongue. Results indicate that Arabic pupils focus on root/pattern morpho-derivational structure, more than on inflected cues, whereas Portuguese pupils show a phonological sensitivity to rime and syllable but also better knowledge about inflexional than derivational morphology. Moreover, whatever mother tongue we studied, morphological knowledge plays a role on written words recognition, when more decontextualized knowledge occurs mainly in Arabic. The grapho-morpho-derivational knowledge contribution in Arabic reading comprehension accentuates the significance of morphology in this written language. Lastly, reading comprehension seems to be more dependent from automatised processes in Portuguese because of its relatively irregular orthography.

The second aim of this research was to study the role of reading acquisition in mother tongue (Arabic vs. Portuguese) on learning to read in French as a second language. Comparisons allow to conclude that the special status of morphology in Arabic can conduct learners to pay attention on internal structure of French words in order to read it. In the other hand, similarity between Portuguese and French facilitates the development of graphophonological sensitivity and emphasises the symmetry between first and second language reading words skills. French second language reading comprehension profiles and their development are also very different between Arabic and Portuguese speakers.

KEYWORDS: Crosslinguistic study Arabic vs. Portuguese - French as second language - Learning to read - Written word recognition - Reading comprehension - Implicit and explicit knowledge - Morphology.

DISCIPLINE : Psychologie

Université Rennes 2
Centre de Recherches en Psychologie, Cognition, Communication EA 1285
Laboratoire de Psychologie du Développement
Place du Recteur Henri Le Moal – CS 24307
35043 Rennes cedex