



**HAL**  
open science

# Structures syllabiques et lexicales des langues du monde Données, typologies, tendances universelles et contraintes substantielles

Isabelle Rousset

► **To cite this version:**

Isabelle Rousset. Structures syllabiques et lexicales des langues du monde Données, typologies, tendances universelles et contraintes substantielles. Linguistique. Université Stendhal - Grenoble III, 2004. Français. NNT: . tel-00250154

**HAL Id: tel-00250154**

**<https://theses.hal.science/tel-00250154>**

Submitted on 11 Feb 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 GRENOBLE III – STENDHAL  
U.F.R. DE SCIENCES DU LANGAGE

N° attribué par la bibliothèque  
-----

THÈSE  
pour obtenir le grade de  
DOCTEUR DE L'UNIVERSITÉ GRENOBLE III  
Discipline : Sciences du Langage

Présentée et soutenue publiquement par

Isabelle Rousset  
Le 10 Juin 2004

**Structures syllabiques et lexicales des langues du monde**  
**Données, typologies, tendances universelles et contraintes substantielles**

JURY

Christian Abry	Président du jury
Louis-Jean Boë	Examineur, Directeur de thèse
Barbara Davis	Rapporteur
Didier Demolin	Rapporteur
Jacques Durand	Examineur
Nathalie Vallée	Examineur, Tuteur

Thèse préparée au sein de l'Institut de la Communication Parlée  
CNRS UMR 5009



UNIVERSITÉ GRENOBLE III – STENDHAL  
U.F.R. DE SCIENCES DU LANGAGE

THÈSE  
pour obtenir le grade de  
DOCTEUR DE L'UNIVERSITÉ GRENOBLE III  
Discipline : Sciences du Langage

Présentée et soutenue publiquement par

Isabelle Rousset  
Le 10 Juin 2004

**Structures syllabiques et lexicales des langues du monde**  
**Données, typologies, tendances universelles et contraintes substantielles**

JURY

Christian Abry	Président du jury
Louis-Jean Boë	Examineur, Directeur de thèse
Barbara Davis	Rapporteur
Didier Demolin	Rapporteur
Jacques Durand	Examineur
Nathalie Vallée	Examineur, Tuteur

Thèse préparée au sein de l'Institut de la Communication Parlée  
CNRS UMR 5009

Je tiens tout d'abord à exprimer ma reconnaissance à Louis-Jean Boë qui a bien voulu accepter la direction de ce travail, et dont les conseils ont toujours été prodigués à bon escient.

Merci à Nathalie Vallée pour sa présence tout au long de ce travail, même après l'arrivée de Silouan, ses encouragements et conseils toujours judicieux. Merci aussi pour sa patience et sa rigueur intellectuelle qui m'ont forcée à aller au-delà de mes limites.

Merci à Didier Demolin pour avoir accepté d'être l'un des rapporteurs de ma thèse, m'avoir apporté son aide pour le transcodage du !xóõ ainsi que pour la recherche de dictionnaires découpés en syllabes.

Merci à Barbara Davis d'avoir bien voulu être rapporteur de ma thèse, de m'avoir accordé sa confiance et d'avoir toujours été positive quant à l'avancée de mon travail.

Merci à Jacques Durand qui me fait l'honneur de participer au jury, pour son aide sur la phonologie du français, ses conseils et encouragements répétés.

Merci à Christian Abry pour sa vivacité intellectuelle, l'intérêt qu'il porte à mon travail, ses conseils et toutes nos discussions, que ce soit au niveau de la recherche, de l'enseignement ou de la vie de l'université.

Je tiens aussi à remercier tout particulièrement Ian Maddieson qui nous a fourni 16 langues de la base ULSID, qui fut intéressé par mes résultats et m'encouragea pour la suite lors de notre brève rencontre à Barcelone.

Je tiens également à remercier Peter MacNeilage qui, comme Barbara Davis, m'a accueillie dans son laboratoire à Austin, m'a fait profiter de ses critiques constructives et a toujours manifesté son intérêt pour mes recherches.

Merci à Sophie Wauquier-Gravelines, Ali Trifit, Ashlynn Kinney et Noël N'Guyen avec qui j'ai eu des discussions parfois enflammées mais toujours instructives sur la syllabe.

Je tiens à exprimer ma reconnaissance à Michel Contini et Pascal Perrier qui ont été mes tuteurs respectivement pour l'enseignement et la recherche et ont toujours veillé de loin à mon bien être.

Ma gratitude va également à Jean-Luc Schwartz et à Marc Sato pour le travail accompli en commun lors de la réalisation de l'expérience sur « l'effet LC ». Merci à Jean-Luc d'avoir servi de locuteur et à Marc pour son aide dans le travail statistique des résultats et la rédaction des articles.

Je tiens à remercier Pierre Darlu et Mahé Ben-Ahmed pour leur accueil au sein du groupe OHLL « Congruence » et pour l'aide financière fournie pour la saisie et la mise aux normes des données de Ruhlen.

Merci aussi à Jean-Pierre Chevrot, ainsi qu'à Céline et Aurélie, pour leurs conseils sur les tests statistiques.

Je tiens à féliciter Jean Brun et Sétaré Alirézai pour le travail accompli au niveau de la création et du développement de l'interface MatLab, ainsi que pour leur aide à la compréhension du logiciel et de la programmation.

Je tiens à remercier Alain et Christophe pour toute l'aide informatique qu'ils m'ont apportée, que ce soit sur Mac ou PC, ainsi que Dominique et Marie-Thé pour tous les conseils pratiques et l'organisation du quotidien à l'ICP.

Je souhaite également remercier 'Julie, Claire et Cécile', pour les pauses café, leur soutien constant et l'aide matérielle des derniers moments, Solange, toujours attentive et présente, ainsi qu'Albert, Anne, Aude, Fanny, Jean-Michel, Laurent, Laurent, Lucie, Marie, Marie-Laure, Nicolas, Omran, Sétaré pour les cours de danses orientales, Véronique, les deux Virginie, mais aussi tous les autres doctorants et DEA de l'ICP.

Je pense aussi à Patricia, Cyril, Elisabetta, Jean-Pierre, Jacqueline ainsi que tous les autres doctorants et enseignants du LIDILEM qui m'ont ouverte sur d'autres branches de la linguistique. Un merci tout particulier à Catherine Chevallier qui a été la première à m'accueillir à Grenoble et qui a su m'orienter vers l'ICP.

Merci à Mélanie pour sa présence depuis mon arrivée à Grenoble, les cours de chants révolutionnaires et les parties de fous rires, ainsi qu'à Hélène pour nos discussions syndicales et politiques, les après-midi shopping et sa conception de la vie qu'elle m'ait fait découvrir. Elles savent tout ce que je dois à leur amitié.

Un grand merci à Delphine, pour toutes les choses que nous avons partagées, à Martin, Pierre et tous les autres camarades qui ont su me permettre de penser à autre chose qu'à mon travail et partager avec moi différents moments de détente.

Mille mercis à Monique, Yves et Cécile pour leurs encouragements constants, leur présence tout au long de ce travail, leur confiance en moi et tout ce qui ne peut se dire ici. Merci plus particulièrement à Yves pour m'avoir servi de coach, pour son aide quant à la rédaction et ses nombreuses relectures.

Enfin, je tiens à remercier Stéphane, sans qui je ne serais parvenue au bout de cette thèse, pour son soutien de tous les instants, son aide aussi bien matérielle que morale et surtout pour son amour et sa confiance qui m'aident à avancer.

## TABLE DES MATIERES

<b>Remerciements.....</b>	<b>1</b>
<b>Table des matières.....</b>	<b>2</b>
<b>Liste des figures.....</b>	<b>5</b>
<b>Liste des tables.....</b>	<b>9</b>
<b>Introduction.....</b>	<b>12</b>
<b>Chapitre I : La syllabe, illusion ou réalité.....</b>	<b>20</b>
I.1. Phonologies de la syllabe.....	21
I.1.1. Frontières syllabiques et organisations internes de la syllabe : des phonologies linéaires en quête d'indices phonétiques.....	22
I.1.2. Formalismes quasi-linéaires pour une définition de la syllabe phonologique.....	24
I.1.3. La syllabe : une structure sonore hiérarchisée.....	26
I.1.3.1. Hiérarchie inter-niveaux.....	26
I.1.3.2. Hiérarchie inter-constituants.....	31
I.1.4. La syllabe réinterprétée.....	32
I.1.5. Syllabe et connexionnisme : modélisation de processus de syllabation.....	35
I.2. Descriptions phonétiques.....	36
I.2.1. La recherche d'indices articulatoires et perceptifs.....	37
I.2.2. Les approches motrices.....	43
I.2.2.1. Rythme syllabique et rythme respiratoire.....	43
I.2.2.2. L'oscillation mandibulaire et la théorie « <i>Frame, then Content</i> ».....	44
<b>Chapitre II : ULSID, base de données lexicales et syllabiques.....</b>	<b>50</b>
II.1. De la représentativité de l'échantillon en typologie linguistique.....	51
II.2. UCLA Lexical and Syllabic Inventory Database.....	52
II.3. Les lexiques d'ULSID.....	54
II.3.1. Aspect formel.....	54
II.3.2. Présentation des lexiques d'ULSID.....	57
II.4. Quantification des données d'ULSID.....	75
II.5. Une interface pour l'interrogation d'ULSID.....	77

<b>Chapitre III : Typologies et tendances des structures lexicales et syllabiques.....</b>	<b>84</b>
III.1. De l'importance des recherches en typologie linguistique.....	85
III.2. Organisation structurelle des unités lexicales.....	86
III.2.1. Taille des unités lexicales.....	86
III.2.1.1. Forme canonique : vers une typologie des langues sur le critère du nombre de syllabes par unité lexicale.....	87
III.2.1.2. Une typologie des langues basée sur le nombre de syllabes par unité lexicale.....	89
III.2.1.3. Loi de Menzerath : taille des unités lexicales vs. Taille des syllabes.....	93
III.2.1.4. La typologie confrontée au nombre de phonèmes par syllabe.....	100
III.2.1.5. Taille des unités lexicales basées sur le nombre de phonèmes qu'elles contiennent.....	102
III.2.2. Structuration syllabique des unités lexicales.....	103
III.2.3. Confirmation de la taille des unités lexicales comme critère de classement typologique.....	106
III.2.4. Vers des éléments d'explication des tendances dans l'organisation structurelle des lexiques.....	107
III.3. Organisation des structures syllabiques.....	108
III.3.1. Les types de structures syllabiques.....	109
III.3.2. Fréquence des types syllabiques dans les lexiques.....	110
III.3.3. Prédominance CV ou CVC : 2 types de langues.....	113
III.3.4. Règles implicationnelles des structures syllabiques.....	115
III.3.5. Convergence entre résultats d'études phonétiques et tendances des structures syllabiques.....	119
III.4. Les constituants syllabiques.....	120
III.4.1. Rendement syllabique.....	120
III.4.2. Syllabes de premier plan.....	123
III.4.3. Les noyaux vocaliques « vedettes ».....	125
III.4.4. Tendances consonantiques des structures syllabiques.....	127
III.4.5. Règles implicationnelles affinées.....	128
III.5. Cooccurrences et dépendances intra et intersyllabiques.....	130
III.5.1. Restriction dans l'inventaire syllabique des langues.....	130
III.5.2. Méthodologie pour la recherche de cooccurrences et de dépendances.....	132
III.5.3. Cooccurrences et dépendances entre attaque et noyau d'une même syllabe.....	136

III.5.4. Analyse des constituants successifs, voyelle et consonne, de la rime.....	137
III.5.5. Cooccurrences et dépendances entre attaque et coda d'une même syllabe.....	138
III.5.6. Recherche de l'effet LC de MacNeilage et Davis (2000) aux niveaux intra et intersyllabique.....	139
III.5.7. Discussion.....	141
<b>Chapitre IV : Un corrélat perceptif de « l'effet LC ».....</b>	<b>148</b>
IV.1. Illusions auditives et effet LC : le paradigme des transformations verbales.....	149
IV.2. Méthode.....	151
IV.3. Procédure expérimentale.....	156
IV.4. Résultats.....	157
IV.5. Discussion.....	161
<b>Conclusion et perspectives : vers une modélisation de l'émergence syllabique.....</b>	<b>163</b>
<b>Bibliographie.....</b>	<b>172</b>
<b>Annexe : .....</b>	<b>188</b>
Annexe I : Carte des langues de notre échantillon.....	189
Annexe II : Ouvrages ayant servi pour le transcodage des données.....	190
Annexe III : Gabarits les plus fréquents dans les langues d'ULSID (plus de 2 %)...	191
Annexe IV : Syllabes les plus fréquentes dans les langues d'ULSID.....	193
Annexe V : Ratio entre occurrences observées et attendues pour les structures CV.....	194
Annexe VI : Ratio entre occurrences observées et attendues pour les structures CVC .....	196
Annexe VII : Ratio entre occurrences observées et attendues pour les structures VC.....	198
Annexe VIII : Abréviations utilisées.....	200



## LISTE DES FIGURES

### Introduction

Figure 1 : Représentation schématique de l'organisation interne de la syllabe..... 16

### Chapitre I : La syllabe, illusion ou réalité

Figure I.1 : Notation autosegmentale de [atlas] (atlas) et de [poni] (poni) pour Kahn (1976)..... 27

Figure I.2 : Associations entre segments et positions CV selon Clements et Keyser (1983)..... 27

Figure I.3 : Représentation arborescente de la structure interne de la syllabe en sous-constituants hiérarchiquement organisés. Entre parenthèses, les constituants facultatifs..... 29

Figure I.4 : Représentation des positions terminales de la structure syllabique selon Selkirk (1984, p. 115) (avec O = onset, R = rime, et les chiffres indiquant la position dans le constituant)..... 30

Figure I.5 : Représentation syllabique du mot anglais « beat » combinant structure autosegmentale et métrique (exemple de Goldsmith, 1990)..... 30

Figure I.6 : Représentation de deux types de syllabes (CVC et CCV) selon le formalisme proposé en phonologie du gouvernement..... 32

Figure I.7 : Modèle rythmique proposé par Angoujard..... 32

Figure I.8 : Grille rythmique du terme « ami » (à gauche) et « chaos » (à droite) (Angoujard, 1997)..... 33

Figure I.9 : Grille rythmique du terme « facteur » (Angoujard, 1997)..... 33

Figure I.10 : Le presetting de la langue dans les productions du babillage qui expliquerait les combinaisons CV favorisées, MacNeilage et Davis (2000)..... 46

Figure I.11 : Schémas représentant le mot cible « συλιε » et le lapsus « συιλε », avec inversion des attaques des deux dernières syllabes de l'unité lexicale (Beaulieu, 2001)..... 46

### Chapitre II : ULSID, base de données lexicales et syllabiques

Figure II.1 : Représentation syllabique de l'unité lexicale « ? a . ch i »..... 54

Figure II.2 : Représentation syllabique de l'unité lexicale « ? a ch . x o sh . tlh' oo l »... 55

Figure II.3 : « Grille iconique » symbolisant le triangle vocalique utilisé pour la description des systèmes vocaliques d'ULSID (Vallée, 1994)..... 57

Figure II.4 : Taille des lexiques en nombre d'entrées lexicales..... 75



Figure II.5 : Nombre total de syllabes dans les lexiques de l'échantillon ULSID..... 76

Figure II.6 : Corrélation entre nombre d'items lexicaux et nombre de syllabes pour les 16 langues.....	76
Figure II.7 : Corrélation entre le nombre d'items lexicaux et le nombre de syllabes pour les 11 lexiques de moins de 7 000 entrées.....	77
Figure II.8 : Fenêtre de l'interface MatLab de la base ULSID pour la recherche du nombre de syllabes par unité lexicale.....	79
Figure II.9 : Fenêtre de l'interface MatLab de la base ULSID, résultats de la recherche du nombre de syllabes par unité lexicale pour le finnois. Le chiffre précédant les deux points indique le nombre de syllabe, et celui suivant les deux points indique le nombre d'unités lexicales correspondantes.....	80
Figure II.10 : Fenêtre de l'interface MatLab de la base ULSID (matrice obtenue pour les combinaisons possibles entre les différents lieux consonantiques possibles pour l'attaque et les trois positions de la langue possible pour le noyau en afar).....	80
Figure II.11 : Fenêtre de l'interface MatLab de la base ULSID. Recherche des gabarits lexicaux selon le type de fichier : API, cohorte CV, lieu, mode articulatoire.....	81
Figure II.12 : Fenêtre de l'interface MatLab de la base ULSID pour la sortie graphique des requêtes.....	81
Figure II.13 : Interface MatLab permettant le choix entre la base ULSID et la base de Ruhlen.....	82
Figure II.14 : Interface MatLab permettant l'étude des gabarits (codage des unités lexicales en C et V) de la racine [aja] de Ruhlen en tenant compte du lieu d'articulation des unités.....	83
Figure II.15 : Interface MatLab présentant les gabarits en fonction du lieu d'articulation pour les unités lexicales de la racine [aja].....	83

### **Chapitre III : Typologies et tendances des structures lexicales et syllabiques**

Figure III.1 : Forme canonique des 16 langues de notre échantillon.....	87
Figure III.2 : Pourcentage d'unités lexicales par nombre de syllabes pour l'afar (à gauche) et le finnois (à droite).....	89
Figure III.3 : Typologie des 16 langues d'ULSID, élaborée sur la répartition du lexique en fonction du nombre de syllabes par item lexical.....	90
Figure III.4 : Pourcentage d'unités lexicales par nombre de syllabes pour l'allemand (d'après Menzerath, 1954).....	92
Figure III.5 : Répartition des unités lexicales (en %) par nombre de syllabes pour l'ensemble de nos 16 lexiques.....	93
Figure III.6 : Répartition du nombre de syllabes par unité lexicale en fonction du nombre moyen de phonèmes par syllabe pour l'allemand d'après les données de Menzerath (1954, p. 96).....	94

Figure III.7 : Répartition du nombre moyen de phonèmes par syllabe en fonction du nombre de syllabes par unité lexicale d'après les données de Altmann (1980) sur l'indonésien et l'anglais.....	95
Figure III.8 : Répartition de la durée des syllabes (en millisecondes) en fonction du nombre de syllabes par unité lexicale d'après les données de Altmann (1980) sur la langue bachka.....	95
Figure III.9 : Nombre moyen de syllabes par mots, et nombre moyen de phonèmes par syllabe pour le français et l'anglais (Goldman et al., 1996), l'allemand (Menzerath, 1954) et le gbaya (Monino et Roulon, 1972).....	96
Figure III.10 : Comparaison entre le nombre moyen de phonèmes par syllabe et le nombre de syllabes par unité lexicale pour les 16 langues de notre échantillon.....	97
Figure III.11 : Corrélation entre le nombre de phonèmes par syllabe et le nombre de syllabes par unité lexicale dans les 16 langues de notre échantillon.....	97
Figure III.12 : Corrélation entre le nombre de phonèmes par syllabe et le nombre de syllabes par unité lexicale dans 15 langues de notre échantillon (le !xóõ se démarquant du comportement général des langues est ici exclu).....	98
Figure III.13 : Répartition du nombre moyen de phonèmes par syllabe en fonction du nombre de syllabes par unité lexicale pour les 16 langues de notre échantillon.....	99
Figure III.14 : Répartition du nombre moyen de phonèmes par syllabe en fonction du nombre de syllabes par unité lexicale sur l'ensemble des lexiques d'ULSID.....	99
Figure III.15 : Proportion de syllabes (en pourcentage) par nombre de phonèmes pour les langues du type A.....	100
Figure III.16 : Proportion de syllabes (en pourcentage) par nombre de phonèmes pour les langues du type B.....	101
Figure III.17 : Proportion de syllabes (en pourcentage) par nombre de phonèmes pour les langues du type C.....	101
Figure III.18 : Nombre moyen de phonèmes par unité lexicale pour chacune des 16 langues d'ULSID. ....	102
Figure III.19 : Nombre de types de structures syllabiques différents dans chacune des langues d'ULSID. ....	109
Figures III.20 : Proportions des différents types syllabiques (à gauche), regroupés en fonction de la présence ou non de l'attaque et de la coda (à droite).....	111
Figure III.21 : Structures avec et sans groupements consonantiques dans notre échantillon. ....	112
Figure III.22 : Répartition des types de syllabes ouvertes les plus répandus (en pourcentage sur l'ensemble des syllabes). Les langues majoritairement CVC sont en rouge. ....	114
Figure III.23 : Répartition par langue des types de syllabes fermées les plus répandus	114

(en pourcentage sur l'ensemble des syllabes). Les langues majoritairement CVC sont en rouge.....	
Figure III.24 : Rendement syllabique des langues d'ULSID.....	121
Figure III.25 : Interactions entre attaque et noyau pour les différents types de structures.....	137
<b>Chapitre IV : Un corrélat perceptif de « l'effet LC »</b>	
Figure IV.1 : Cube de Necker : stimulus ambigu et perceptions possibles	150
Figure IV.2 : Stimulus /pa/ utilisé dans notre expérience	153
Figure IV.3 : Stimulus /ta/ utilisé dans notre expérience	153
Figure IV.4 : Stimulus /pi/ utilisé dans notre expérience	154
Figure IV.5 : Stimulus /ti/ utilisé dans notre expérience	154
Figure IV.6 : Stimulus /po/ utilisé dans notre expérience	155
Figure IV.7 : Stimulus /to/ utilisé dans notre expérience	155
Figure IV.8 : Nombre de transformations en fonction de l'effet d'amorçage (ANOVA [F(1,23) = 4,61, p < 0,05])	157
Figure IV.9 : Nombre de transformations vers la forme de base, la forme inversée, une autre forme, en fonction du stimulus de départ.	159
Figure IV.10 : Durée de stabilité des structures :  LC et  CL en fonction de l'effet stimulus (à gauche) et de l'effet d'amorçage (à droite).	160

## LISTE DES TABLES

### Chapitre II : ULSID, base de données lexicales et syllabiques

Table II.1 : Tableau de transcodage des caractères ascii « E », « e » et « @ » pour quatre langues.....	56
Table II.2 : Extrait de fichiers transcrits pour le kannada .....	78
Table II.3 : Extrait de fichiers transcrits pour le kannada (cf. Annexe VIII) .....	79

### Chapitre III : Typologies et tendances des structures lexicales et syllabiques

Table III.1 : Regroupements mis en évidence en fonction de la forme canonique des langues.....	88
Table III.2 : Regroupements mis en évidence en fonction de la forme canonique des langues d'ULSID et de 3 autres langues .....	88
Table III.3 : Nombre de types de gabarits lexicaux pour chacune des langues de notre échantillon, classées selon la typologie en quatre types définie au paragraphe 2.1.....	104
Table III.4 : Structures représentant plus de 2 % des gabarits de l'ensemble des 16 lexiques.....	105
Table III.5 : Récapitulatif des analyses précédentes.....	106
Table III.6 : Les 19 types syllabiques d'ULSID classés en fonction de la complexité de l'attaque.....	110
Table III.7 : Liste et nombre d'occurrences des types de structures syllabiques rencontrées dans les 16 langues d'ULSID (classés par fréquence décroissante). .....	111
Table III.8 : Liste des structures syllabiques d'ULSID classées par type et leur proportion dans le lexique d'une langue donnée. ....	116
Table III.9 : Répartition des voyelles dans les syllabes les plus fréquentes (20 premières de chaque langue) ( ʉ est la voyelle murmurée). ....	123
Table III.10 : Consonnes les plus répandues dans les vingt syllabes les plus fréquentes (en position d'attaque). ....	124
Table III.11 : Pourcentage des voyelles dans les vingt syllabes les plus fréquentes (après regroupement en fonction du timbre de base). ....	126
Table III.12 : Pourcentage de l'inventaire de consonnes pouvant apparaître en attaque et en coda dans les langues d'ULSID (calculé à partir du nombre de syllabes différentes). ....	127
Table III.13 : Nombre de syllabes observées, nombre de syllabes théoriques et rapport syllabique (observé/théorique) pour chacune des langues de notre base de données.....	131

Table III.14 : Matrice de cooccurrences des syllabes CV en afar, avec leur nombre d'apparition dans le lexique. Les matrices sont présentées selon la fréquence décroissante des constituants en position d'attaque (colonnes) et de noyau (syllabique). Le total des totaux correspondant au nombre de syllabes de l'afar, soit 9 896.....	133
Table III.15 : Regroupement des consonnes selon leur lieu d'articulation. ....	134
Table III.16 : Regroupement des voyelles selon leur point d'articulation. ....	134
Table III.17 : Regroupement des consonnes selon leur mode articulatoire.....	135
Table III.18 : Matrice des syllabes CV en afar, après regroupement des consonnes par lieu d'articulation et des voyelles suivant la position antéro-postérieure de la langue (voir annexe VIII pour les abréviations). ....	135
Table III.19 : Ratios entre les occurrences attestées et les occurrences attendues pour les combinaisons C*V pour l'ensemble des syllabes de l'afar (voir annexe VIII pour les abréviations).....	136
Table III.20 : Ratios entre les occurrences attestées et les occurrences attendues pour les combinaisons C*V pour l'ensemble des syllabes de l'ensemble de notre échantillon ( $\chi^2$ significatif, $p < 0.001$ pour chaque colonne).....	136
Table III.21 : Ratios entre les occurrences attestées et les occurrences attendues pour les combinaisons V*C pour l'ensemble des syllabes de l'ensemble de notre échantillon ( $\chi^2$ significatif, $p < 0.001$ pour chaque colonne).....	137
Table III.22 : Ratios entre les occurrences attestées et les occurrences attendues pour les combinaisons C_C(attaque_coda) pour l'ensemble des syllabes de l'ensemble de notre échantillon ( $\chi^2$ significatif, $p < 0.001$ pour chaque colonne). ....	139
Table III.23 : Matrice des cooccurrences consonantiques dans les structures C <sub>1</sub> V .C <sub>2</sub> V en afar.....	140
Table III.24 : Matrice des cooccurrences consonantiques dans les structures C <sub>1</sub> V .C <sub>2</sub> V pour les 10 langues retenues. ....	140
Table III.25 : Résultats sur les modèles CV favorisés dans le babillage de six enfants, les premiers mots de 10 enfants et les mots de 10 langues (MacNeilage et Davis, 2000). Les résultats sont obtenus en calculant les ratios entre fréquence observée de la syllabe et fréquence calculée en fonction de celles des constituants syllabiques. ....	146

## Chapitre IV : Un corrélat perceptif de l'« effet LC »

Table IV.1 : Valeurs des facteurs intrinsèques aux unités pour les stimuli choisis (point d'unicité : nombre de segments nécessaires pour que l'on ne puisse plus confondre un mot avec un autre ; densité de voisinage : nombre d'unités lexicales phonologiquement proches du stimulus).....	152
Table IV.2 : Nombre moyen de transformations et de formes différentes observées en fonction du stimulus (avec écart-types).....	157
Table IV.3 : Nombre moyen et durée de stabilité des transformations liées au stimulus, vers la forme associée ou une autre en fonction du stimulus.....	158
Table IV.4 : Nombre moyen et durée de stabilité (ms.) des transformations vers une forme Labial-Coronal ou vers une forme Coronal-Labial.....	160

## Conclusion

Table C.1 : « Example initial and final vocabularies from the ful model. Blanks in the initial vocabulary indicate words of length 0. It is evident that no regular syllable structure exists in the initial vocabulary, whereas words of the final vocabularies are clearly composed of just a few syllable types. » (Redford et al., 2000, p. 21).....	169
--	-----

# **INTRODUCTION**



Bien que la syllabe constitue, depuis de nombreuses années, un objet de recherche de la phonologie, de la phonétique et de la psycholinguistique, la question de sa nature n'a pas encore trouvé de réponse complète et cohérente aux différents niveaux de l'analyse linguistique. En effet, la syllabe est un objet d'étude complexe tant dans les approches conceptuelles, que dans les méthodes d'analyses des différentes disciplines qui l'ont inscrite dans leur champ d'étude. À partir de plusieurs études typologiques de données lexicales et syllabiques de langues naturelles, nous proposons dans cette thèse des éléments nouveaux qui plaident en faveur de l'existence d'une unité syllabique, structurellement organisée, construite sur des dépendances internes et externes, constitutive des unités lexicales et participant de fait pleinement au fonctionnement du système linguistique.

Si la plupart des courants phonologiques ont considéré la syllabe comme un objet incontournable de la théorie, les approches très variées, du fait qu'elles ne réussissent pas à englober la diversité linguistique, sont toujours très discutées à l'heure actuelle. La difficulté des recherches sur la syllabe, en tant qu'objet d'étude propre, semble résider dans le fait qu'*« on admet généralement sa pertinence mais sans vraiment la définir explicitement »* (Trifit, 2000, p. 9). Cet état de fait résulte probablement de la connaissance intuitive de la syllabe. En effet, comment nier l'existence d'une structuration syllabique des langues lorsqu'à la différence des systèmes d'écriture alphabétique, comme le grec ou le latin, ou consonantique, comme l'écriture arabe, dans lesquelles le symbole graphique transcrit le son, certains systèmes dits logographiques considèrent la syllabe comme unité naturelle de codage pour l'écriture. C'est par exemple le cas des syllabaires bien connus du japonais (hiragana et kata kana), ou d'un certain nombre de langues d'Asie, telles que le tamoul (langue dravidienne), ou le khmer (langue austro-asiatique).

Paradoxalement à la complexité de son organisation et de ses réalisations, la syllabe est souvent rapportée comme une unité intuitivement présente chez les locuteurs, que leur langue soit écrite (alphabétique ou syllabique) ou non. En effet, tout locuteur natif d'une langue, quel que soit son niveau d'instruction, est capable de donner le nombre de syllabes d'un mot ou de trouver plusieurs mots d'un nombre de syllabe donné, alors que ce phénomène n'est pas observable pour le comptage des sons ou des phonèmes d'un mot avec des sujets illettrés. Derwing (1992) précise que les locuteurs de nombreuses langues sont capables non seulement de compter le nombre de syllabes d'un mot, mais qu'ils obtiennent aussi des résultats remarquablement consistants dans la syllabation de mots avec consonne simple ou consonne géminée à l'intervocalique. Cet état de fait sur la délimitation des frontières syllabiques est néanmoins controversé au niveau de sa stabilité : 1) le découpage des groupements consonantiques en position inter-syllabique varie en fonction de la langue étudiée ; 2) de nombreux linguistes considèrent que cette délimitation peut varier entre différents locuteurs d'une même langue. Le phénomène d'ambisyllabité, qui décrit le fait qu'un individu ne sait pas toujours dans quelle syllabe rattacher une consonne (soit en position finale de la syllabe précédente, soit au début de la syllabe suivante), n'apparaît pas avec la même fréquence selon les langues. Ainsi, l'anglais est considéré comme une langue pour laquelle le découpage

syllabique est ambigu, alors que l'espagnol est défini comme ayant une syllabation bien définie et stable (Alvarez, Carrieras et de Vega, 2000). Pourtant, dans la majorité des cas d'unités lexicales ambisyllabiques, s'il existe bien des problèmes de délimitation des frontières, les locuteurs ne ressentent aucune difficulté dans la tâche de comptage des syllabes.

Autre preuve d'une conscience linguistique de la syllabe : la présence de jeux de langue fondés sur l'organisation syllabique. On les retrouve dans les constructions du verlan ou du javanais, formés à partir du français, mais aussi dans un grand nombre d'autres langues (exemple dans Demolin, 1991 ; Hombert, 1986). Bagemihl (1995) présente les cas venant d'un « *ludling* » (jeu de langue) en portugais du Brésil. Ce jeu consiste en la préfixation d'une syllabe [pe] ou [pV] – la voyelle s'identifiant, dans ce cas, à celle de la syllabe suivante – devant chaque syllabe du mot : [menina] « fille » devient [pemepenipena] ou [pemepinipana] selon la règle choisie. Ce type de jeu est répandu dans la plupart des familles de langues. Ainsi, en allemand, on retrouve le même principe avec l'ajout d'une syllabe « bi » pour la création de la « *Bi-Sprache* » décrite par Ringelnatz (1883-1934) :

**Langue-bi**

Ibich habibebi dibich,

Lobittebi, sobi liebib.

Habist aubich dubi mibich

**Allemand**

Ich habe dich

Lotte, so, lieb.

Hast auch du mich

Ce type de jeux de langues peut aussi avoir un rôle social comme c'est le cas dans certaines langues africaines, où les différents niveaux de combinaisons des unités syllabiques correspondent aux différentes catégories sociales qui les utilisent.

La syllabe est aussi recrutée lors du travail sur la forme de la langue dans une recherche poétique, la versification se faisant en comptant le nombre de « pieds » d'un vers (nombre de pics syllabiques). Ainsi, l'exemple suivant, tiré du « *Cid* » (Corneille, 1682) est une tirade composée intégralement en alexandrins, c'est-à-dire en vers comprenant douze pieds ou douze syllabes :

Ô rage ! ô désespoir ! ô vieillesse ennemie !

N'ai-je donc tant vécu que pour cette infamie ?

Et ne suis-je blanchi dans les travaux guerriers

Que pour voir en un jour flétrir tant de lauriers ?

Extrait du « *Cid* » (Corneille, 1682)

Les syllabes, et plus précisément les noyaux syllabiques, jouent un rôle important dans la création du langage poétique, sans que ce phénomène ne soit spécifique aux textes écrits. Des études montrent, en effet, l'utilisation de la syllabe pour la versification dans les langues à

tradition orale, ainsi que dans la structuration universelle des comptines ou de certaines chansons (Arléo, 1997 ; Arléo et Flament, 1988).

Au niveau de la description des langues et de leur organisation interne, la syllabe est souvent associée au traitement des phénomènes prosodiques tels que l'accentuation, les systèmes de tons, le rythme et l'intonation. Dans certaines langues, la syllabe est en effet considérée comme l'élément clé de la description de phénomènes accentuels, quel que soit le principe mis en œuvre pour l'accentuation : dans les langues dites à accent libre, les possibilités pour placer l'accent sont aussi nombreuses qu'il existe de syllabes, alors que dans les langues à accent fixe, l'application des lois régissant la place de l'accent suppose le décompte préalable des syllabes. C'est par exemple le cas bien connu du latin, où l'accent tombe sur la syllabe pénultième si le mot possède moins de trois syllabes, alors que sa position dépend du poids de cette syllabe dans le cas des mots de trois syllabes et plus : si elle est lourde (comme dans le cas d'une voyelle longue), elle prend l'accent ; si elle est légère, c'est la syllabe antépénultième qui le reçoit.

*Exemples (la syllabe accentuée est notée en gras) :*

Mots trisyllabiques et plus :	<b>fide:</b> lis	« fidèle »
	<b>facilis</b>	« facile »
	<b>agricola</b>	« agriculteur »
Mots mono- et dissyllabiques :	<b>do:</b>	« je donne »
	<b>locus</b>	« lieu »

Cependant, la syllabe n'est pas toujours considérée comme étant l'élément majeur de l'organisation rythmique du langage. En effet, traditionnellement, les langues sont divisées en deux types : les « *syllabled-timed* » et les « *stress-timed* » (Pamies Bertrán, 1999 ; Lacheret-Dujour et Beaugendre, 1999). Les premières possèdent une organisation rythmique basée sur un principe de chronométrage syllabique, tandis que les deuxièmes sont plutôt des langues pour lesquelles le rythme se fonde sur un chronométrage accentuel. Dans les secondes, les unités rythmiques se suivent de manière à ce que le laps de temps qui sépare les syllabes soit le plus uniforme possible. Bien que ces deux types d'organisations possibles du rythme diffèrent, ils se réfèrent tous deux au découpage syllabique des langues, même si ce n'est effectivement pas le seul facteur pris en compte.

Enfin, la description des systèmes tonals et de leurs réalisations, dans différentes langues, incite également à penser que la présence d'un niveau syllabique intermédiaire entre le segment et l'unité lexicale permet d'expliquer de manière plus cohérente et logique un nombre important de phénomènes. Selon Creissels (1994), l'apparition des tons modulés dans certaines langues africaines est à la fois plus compréhensible et mieux prédictible si le phénomène est étudié au niveau de la syllabe. Sous cet angle de vue, les tons ne seraient plus associés directement aux segments, mais à l'ensemble des positions segmentales dominées

par le nœud rime (cf. figure 1), qu'ils soient réalisés comme consonne ou voyelle. Cette explication permet de ne pas associer un ton descendant à la voyelle et un ton montant à la consonne lorsque celle-ci ne saurait phonétiquement être le support d'un ton, mais d'associer les tons aux positions et de les faire apparaître, en surface, comme un ton modulé sur la voyelle. Creissels propose, par exemple, pour le terme [kõp.tò] issu de la langue zarma (variété de shonghaï, apparentée aux langues mandées) d'adopter l'explication suivante : le ton modulé de la première syllabe est la réalisation d'un ton bas associé à la première position de la rime et d'un ton haut associé à la deuxième position.

Bien que couramment utilisée pour la description de phénomènes prosodiques, l'existence de la syllabe comme objet ou unité linguistique ainsi que son rôle dans le fonctionnement du langage ne sont toujours pas éclaircis. La phonologie, la phonétique et la psycholinguistique, ont développé des approches diversifiées, parfois complémentaires, qui fournissent des indices sur son existence, sa complexité et améliorent la connaissance de l'objet. Nous les aborderons dans le premier chapitre.

Dans les différents courants phonologiques qui se sont penchés sur la réalité d'un niveau syllabique dans l'organisation des unités sonores des langues, il est néanmoins possible de retrouver des éléments de description analogues. Une des propositions relativement consensuelles en phonologie est l'analyse de la syllabe en deux constituants, l'attaque et la rime, cette dernière étant elle-même composée de deux éléments, le noyau et la coda. Le noyau est le seul élément nécessaire et suffisant pour qu'on puisse parler de syllabe. L'attaque et la coda sont pour leur part facultatives. On obtient ainsi la schématisation suivante de la syllabe phonologique :

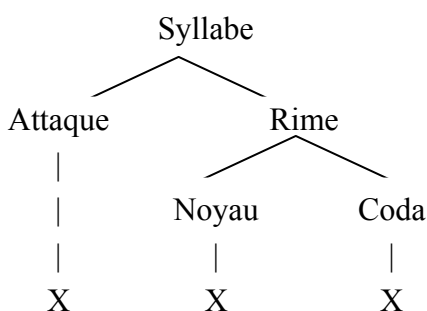


Figure 1 : Représentation schématisée de l'organisation interne de la syllabe

Les différentes positions du squelette (notées X) sont remplies par un ou plusieurs éléments du système phonologique de la langue considérée. La position squelettale associée au noyau est majoritairement remplie par des segments vocaliques ou, plus rarement dans certaines langues, par des types de segments consonantiques (nasales, vibrantes, sifflantes, ...), alors que les positions squelettiques d'attaque et de coda sont toujours associées à des segments consonantiques.

Ce type de formalisme pour représenter la syllabe est le fruit de l'évolution d'un certain nombre de courants phonologiques. Plusieurs des différentes approches s'intéressant à cette unité seront évoquées dans la première partie du chapitre I afin de mieux connaître les principes phonologiques qui semblent régir l'organisation des constituants syllabiques.

Au niveau phonétique, bien des études se sont intéressées aux traces de la syllabe dans le signal acoustique (Krakow, 1999), comme aux niveaux de la motricité et de la perception. Bien que de nombreux résultats expérimentaux aillent dans le sens de l'existence d'une unité syllabique, aucune description phonétique alliant les différents niveaux (acoustique, articulatoire et perceptif) n'a encore été proposée. Nous les présenterons dans la deuxième partie du chapitre I.

Des travaux récents ouvrent cependant de nouvelles perspectives théoriques, c'est le cas de la théorie « *Frame, then Content* » de MacNeilage (1998) qui propose que l'universelle structuration syllabique des langues du monde repose entièrement sur les mouvements alternés et cycliques d'élévation et d'abaissement de la mandibule – le cadre (*the frame*). Cette modulation accordée à l'ouverture et à la fermeture de la cavité buccale correspond à la dimension articulatoire la plus exploitée par les langues pour la réalisation des unités sonores distinctives consonnes et voyelles – le contenu (*the content*) – qui constituent les phases des cycles mandibulaires. La théorie « *Frame, then Content* » attire l'attention sur le rôle prépondérant de la mandibule dans le babillage (Davis et MacNeilage, 1995) mais aussi dans la structuration syllabique du lexique des langues et des protolangues (MacNeilage et Davis, 2000).

Cette approche, basée sur l'hypothèse d'une base articulatoire pour la syllabe, nous a semblé être un cadre intéressant pour situer nos recherches (partie 2 du chapitre I). Ainsi, la comparaison des résultats de plusieurs de nos études avec ceux de MacNeilage et Davis (2000) permettra de valider ou nuancer l'apport de cette théorie.

Mais si la syllabe est une unité linguistique par nature, elle intéresse aussi certains psycholinguistes qui tentent de comprendre la place de cette unité dans le fonctionnement du langage. En effet, la syllabe semble avoir également une réalité psycholinguistique tant au regard des pathologies, de l'acquisition, que des erreurs de langage. Déjà Jakobson (1969) remarquait que la structure CV, présente dans toutes les langues du monde, était à la base des premières productions des enfants, ainsi que des productions restantes chez des patients aphasiques. Depuis, de nombreuses études ont confirmé cette observation (Davis et MacNeilage, 1990 ; MacNeilage et Davis, 1993 ; Davis et MacNeilage, 1994 ; Stefanuto, 1997 ; MacNeilage, 1998 ; MacNeilage et Davis, 2001 ; Davis *et al.*, 2002). Les études sur les lapsus, la plupart respectant la position syllabique des éléments intervertis (Beaulieu, 2001), confirment également le rôle prépondérant de la syllabe dans le fonctionnement du langage. MacNeilage (1998) souligne que les voyelles et les consonnes ne sont interchangeable respectivement qu'avec les voyelles et les consonnes des syllabes adjacentes, ou alors pour les consonnes, avec une autre consonne de la même syllabe : une consonne ne prend jamais la

place d'une voyelle, et vice-versa. Ces résultats confirment bien que le traitement de la parole tient compte d'informations sur la place des segments dans la syllabe et supposent donc une organisation syllabique des langues. Nous reprenons plus en détail cet aspect psycholinguistique de la syllabe pour refermer ce premier chapitre.

Ainsi, l'ensemble de ces recherches laisse penser que la syllabe, unité intuitive forte, se manifeste dans les langues naturelles au travers d'un grand nombre de phénomènes très divers. C'est probablement la raison pour laquelle les approches des différents niveaux de l'analyse linguistique n'ont pas encore réussi à mettre en évidence de définition claire de l'unité syllabique. C'est dans les relations entre phonétique et phonologie qu'il est actuellement possible d'envisager une nouvelle manière d'aborder la syllabe et c'est dans ce type d'approche prometteur que se situe notre étude de la syllabe dans l'organisation lexicale des langues.

Dans la continuité des travaux menés depuis une quinzaine d'années à l'ICP sur les structures sonores des langues du monde (Vallée, 1994 ; Schwartz *et al.*, 1997 ; Vallée *et al.*, 1999 ; Boë *et al.*, 2000; Vallée *et al.*, 2001), nous avons choisi de travailler en premier lieu dans un cadre typologique, à partir de la constitution d'une base de données lexicales et syllabiques (présentée dans le chapitre II) afin de mettre en évidence des grandes tendances et des universaux des structures syllabiques contenues dans les lexiques des langues. Nos analyses typologiques nous ont conduits à postuler un certain nombre de tendances structurelles et organisationnelles qui confirment, prolongent ou remettent en cause des résultats d'études antérieures sur la syllabe (chapitre III). Après avoir établi une typologie des langues en fonction de la structuration syllabique des unités lexicales (nombre de syllabes, de phonèmes, ...), nous nous sommes intéressés à l'organisation interne de la syllabe. La mise en évidence de structures syllabiques extrêmement favorisées, ainsi que de tendances nettes quant à l'apparition des différentes structures dans les langues a conforté l'hypothèse de l'existence de contraintes pesant sur l'organisation syllabique. Nous avons alors cherché d'éventuelles relations entre les différents constituants au sein d'une même syllabe ou de deux syllabes successives.

Nous discutons nos résultats à la lumière de travaux antérieurs relevant aussi bien de la phonologie que de la phonétique ou la psycholinguistique, en se référant à une linguistique « *substance-based* » telle qu'elle a été définie par Liljencrants et Lindblom (1972). Ces derniers font le constat que l'explication des tendances universelles trouvées dans les systèmes sonores des langues du monde est impossible si on reste uniquement au niveau phonologique, comme le recommandent les principes de l'approche structuraliste de la première moitié du XXème siècle. Liljencrants et Lindblom (1972) proposent d'utiliser les résultats obtenus en phonétique pour expliquer certains phénomènes phonologiques, et de se servir de ces derniers pour orienter les recherches phonétiques : « *On the other hand, we must concede that very little research has been consciously undertaken in the fields of phonetics and speech communication to elucidate how language rather than speech is structured. [...] Rather than accept it as a-priori [of the language], we attempt to derive linguistic form as a*

*consequence of various substance-based principles pertaining to the use of spoken language and its biological, sociological, and communicative aspects.* » (Liljencrants et Lindblom, 1972, p. 859).

Depuis, à travers de multiples études, il est apparu de plus en plus difficile, voire impossible, de tenter d'expliquer les tendances générales des structures sonores des langues du monde avec des principes internes de phonologie, orientant les recherches des explications des faits phonologiques dans les aspects articulatoires, acoustiques et psychologiques de la parole (Lindblom *et al.*, 1984 ; Vallée, 1994 ; Schwartz *et al.*, 1997 ; De Boer, 2000). Cette linguistique « *substance-based* » s'est ainsi développée depuis les années 1970, essentiellement autour de la réflexion sur les tendances dans les systèmes phonologiques des langues du monde, à travers une démarche qui considère que l'exploration des données de langues doit être complétée d'études expérimentales et de modélisations s'appuyant sur les caractéristiques articulatoires et perceptives des unités linguistiques étudiées. Cette vision différente des relations entre phonétique et phonologie, nous paraît transposable à la syllabe, et capable d'éclairer l'organisation syllabique des langues. Notre propos n'est pas ici d'analyser le statut linguistique de la syllabe mais son contenu, afin de mettre en évidence les facteurs susceptibles de conditionner des choix ou des restrictions en faisant appel à la substance des unités sonores en contexte immédiat ou proche. Nous prolongerons notre investigation typologique dans le domaine de la substance en montrant (chapitre IV) qu'il existe un corrélat perceptif fort d'une tendance (« *L'effet LC* »). La réalisation d'une expérience de perception nous a permis de tester la stabilité respective de deux patrons consonantiques dont l'un est largement favorisé dans les lexiques par rapport à l'autre : Labial-V-Coronal-V par rapport à Coronal-V-Labial-V.

Enfin, nous terminons notre étude en ouvrant des pistes vers la modélisation de l'émergence syllabique dans les langues (cf. partie Conclusion et Perspectives), perspective importante qui aura pour but de valider de manière objective nos hypothèses en évitant le « *curve-fitting* » décrit par Lindblom (2000) qui prive la typologie linguistique d'explication scientifique.

# **CHAPITRE I**

## **LA SYLLABE : ILLUSION OU RÉALITÉ ?**



À différents niveaux de recherche, aussi bien phonétiques que phonologiques, la syllabe semble être aujourd'hui un élément incontournable pour l'analyse de nombreux phénomènes linguistiques. Nous avons voulu, dans ce chapitre, rassembler différents courants théoriques et travaux expérimentaux qui ont tenté de manière inductive ou déductive de démontrer tout ou partie de l'existence d'une unité syllabique dans le fonctionnement du langage, sans pour autant réussir à la définir complètement dans sa nature et dans son fonctionnement.

## I.1. PHONOLOGIES DE LA SYLLABE

Dans les études phonologiques, la syllabe est souvent considérée comme une unité segmentale ou suprasegmentale. Bien qu'il existe un grand nombre de théories phonologiques de la syllabe, deux courants sont considérés comme fondamentaux :

- le premier considère la chaîne comme une unité composée d'unités phonologiques plus larges que le phonème, organisées essentiellement de manière séquentielle. « *Dans cette optique, la syllabe est donc une séquence de segments chronologiquement ordonnés et délimitée par des frontières* » (Meynadier, 2001, p. 102-103).
- le second se fonde sur une conception *multilinéaire* de la phonologie. Chaque unité possède ses propres caractéristiques tout en étant le domaine de phénomènes phonologiques ou phonétiques particuliers. « *La syllabe est considérée comme une unité phonologique ayant une structure interne et appartenant à la structure hiérarchisée en constituants phonologiquement déterminés* » (Meynadier, 2001, p. 103).

En fait, cette opposition sur le concept de la syllabe s'inscrit dans l'évolution théorique générale de la phonologie et le passage d'une analyse linéaire à une analyse multi-linéaire. Meynadier (2001, p. 107) poursuit : « *Ainsi, parallèlement à l'existence d'une unité phonologique telle que le phonème, la syllabe est indépendamment justifiée par un certain nombre de phénomènes phonologiques généraux. Elle apparaît comme un domaine délimité de régularités phonologiques particulières segmentales et suprasegmentales* ». Cependant, l'ensemble des théories post-généralistes qui se positionnent par rapport à la négation de l'unité syllabique telle qu'elle est développée par Chomsky et Halle et à la domination complète des matrices de traits descriptifs (c'est-à-dire d'éléments constituants des phonèmes) se sont trouvées confrontées à plusieurs problèmes d'ordre général portant atteinte à une définition claire de la syllabe :

- la délimitation de ses frontières ;
- la compréhension de sa structure interne et des liens entre les positions syllabiques et les segments qui s'y actualisent ;
- les relations entre les syllabes dans le cadre de la resyllabation à l'intérieur du mot.

Ces différents points ont été abordés dans les propositions des diverses théories phonologiques qui traitent de la syllabe, même si ces dernières ne leur accordent pas toujours la même importance et le même rôle.

### **I.1.1. Frontières syllabiques et organisations internes de la syllabe : des phonologies linéaires en quête d'indices phonétiques**

Dans ce premier courant d'idées, la syllabe se réduit aux caractéristiques distributionnelles des segments qui la composent. Elle n'a pas d'existence propre et se caractérise par la localisation de ses frontières. Nous verrons que la majeure partie des théories linéaires, telles celles basées sur le principe de sonorité des phonèmes, se borne à décrire les combinaisons possibles sans pour autant réussir à les expliquer.

La pertinence de la syllabe comme unité phonologique est déjà remarquée par Saussure (1916) qui préconise, dans son « *Cours de Linguistique Générale* », d'observer les restrictions dans les combinaisons de segments, plus particulièrement à partir de leurs représentations articulatoires, afin d'établir la définition de la syllabe. L'oreille percevant la division en syllabes de la chaîne parlée, il doit nécessairement exister une différence entre deux types de sons permettant la distinction entre le début et la fin de la syllabe. Pour Saussure, c'est la notion d'ouverture du conduit vocal. Ainsi, la séquence [appa] contient deux [p] différents : le premier est implusif et correspond à un mouvement de fermeture (noté >); le second est explosif et correspond à un mouvement d'ouverture (noté <). Cette opposition existe pour tous les sons à l'exception du [a] : en effet, selon cette hypothèse, plus l'aperture du son est grande, moins la distinction implusif/explosif sera perceptible, jusqu'à devenir inexistante pour le son le plus ouvert, en l'occurrence la voyelle [a].

À partir de cette définition, Saussure définit quatre types de combinaisons possibles permettant de définir la place des éléments dans la syllabe, la frontière syllabique étant marquée par le passage d'une implosion à une explosion (> <), mais pouvant être mouvante en fonction de la rapidité de prononciation ou du degré d'aperture des phonèmes successifs.

La définition de la syllabe à partir d'une catégorisation implusive ou explosive des phonèmes sera reprise par un grand nombre de linguistes, notamment Grammont (1933) qui affine les hypothèses de Saussure grâce à l'évolution de la phonétique expérimentale et au perfectionnement des instruments de mesures, ou encore Kurylowicz (1948), qui tente de déterminer de façon formelle les groupes implusifs et les groupes explosifs.

Malmberg (1974) considèrera la syllabe comme l'élément linguistique qui rend possible la réalisation concrète des phonèmes dans la parole. L'étude de sa réalisation phonétique combinera l'étude des frontières syllabiques à partir de la perception des spectres artificiels qui fournissent les indices de syllabation dans la durée de la transition entre vocoïde et contoïde, avec celle des échelles de sonorité, pour aboutir à la mise en évidence d'une série d'implications sur les types syllabiques attestés.

Ces approches, qui visent à définir selon leurs auteurs la syllabe phonologique, se fondent sur des critères phonétiques (articulatoires ou perceptifs) susceptibles de délimiter les frontières syllabiques. Les mêmes types de critères vont être utilisés pour la proposition de lois régissant les segments et leur apparition aux différentes positions syllabiques.

Les théories du début du siècle apportent effectivement un élément important à la description de la syllabe par la mise en évidence d'une organisation intra-syllabique des segments dans la chaîne parlée basée sur un principe de sonorité. Elles trouvent cependant très vite leurs limites dans la délimitation des frontières de syllabe, beaucoup moins immédiate qu'il n'y paraît dans le cas de groupements consonantiques inter-syllabiques. De plus, les différentes solutions proposées (segment implosif/explosif, spectres artificiels) ne sont pas toujours explicables au niveau phonétique.

En effet, dès Saussure (1916), les frontières syllabiques seules se montrent vite insuffisantes pour déterminer la syllabe comme une unité phonologique. Caractériser les séquences de segments possibles et mettre en lumière des règles qui les régissent constituent des étapes indispensables. Dans cette optique, Saussure introduit la notion d'échelle de classement des segments, en utilisant des données perceptives et articulatoires simplifiées en quatre actes élémentaires (l'expiration, l'articulation buccale, la variation du larynx, la résonance nasale) et en proposant le degré d'ouverture du conduit vocal comme critère d'organisation intra-syllabique (de 0 pour les occlusives à 6 pour les voyelles basses et nasales correspondantes).

La syllabe se construit, selon Saussure, autour de cette échelle à travers la définition du point vocalique, des sonantes et des consonantes. Le premier phonème implosif de la syllabe sera toujours le point vocalique. Les sonants correspondent aux sons pouvant apparaître en position de point vocalique. Les consonants sont ceux qui l'entourent. L'échelle permet ainsi de rendre compte de la sonorité des sons.

Celle-ci sera améliorée à plusieurs reprises par Jespersen (1904), ou Grammont (1933) avec le principe de tension croissante et décroissante selon lequel se placent les segments : entre l'initiale de la syllabe et le point vocalique, la tension (ou ouverture) est croissante, alors qu'entre le point vocalique et la finale, la tension est décroissante.

Ces règles, d'abord définies pour la syllabe phonologique, sont plus ou moins valables pour la syllabe phonétique qui, quant à elle, est spécifique à chaque langue. Dans ces théories, la syllabe phonologique correspond à la syllabe idéale, comme chez Grammont (1933) qui relègue les problèmes de délimitation des frontières à la syllabe phonétique en raison d'une ambisyllabité qui lui est inhérente.

Même si Malmberg (1974) propose une définition différente des frontières de la syllabe, il utilisera aussi les notions d'échelle de sonorité et de tensions définies à partir de la hiérarchie de Jespersen (1904) pour caractériser l'organisation syllabique.

La multiplicité des échelles de sonorité est toujours actuelle, Angoujard (1997) suggérant qu'elles sont encore imparfaites parce que l'on ne connaît pas encore assez bien la description physique des segments que l'on y classe. Ainsi, l'occlusive glottale [ʔ] comme les fricatives glottales [h h̥] sont exclues de la plupart des classements. À partir des travaux de Steriade (1982) qui posent que les traits retenus par les échelles de sonorité sont eux aussi hiérarchisés, Angoujard définit les segments comme un ensemble de traits articulatoires, pour proposer une hiérarchie qui au final reprend en majorité celle des échelles de sonorité précédentes. Sa conclusion est qu'une échelle de sonorité ne saurait tout expliquer et que pour définir les

phénomènes liés à la syllabe, il faut tenir compte d'autres éléments comme le rythme et la courbe prosodique (cf. partie I.1.4.).

La principale difficulté rencontrée par ces théories linéaires est le découpage de la chaîne parlée. La question de la localisation du noyau syllabique, même si sa résolution peut être rendue difficile par l'existence de consonnes syllabiques, est assez vite traitée. La préoccupation principale des phonologues est alors le repérage des frontières entre les segments intervocaliques.

Parmi ces approches, certaines voient la syllabe comme une entité décomposable en trois parties : l'attaque, le noyau et la coda, ces éléments étant relativement autonomes du point de vue segmental. « *La syllabe peut alors être décrite comme la concaténation de ces trois éléments dont le noyau est l'élément minimal nécessaire et suffisant de la syllabe et dont les marges pré-nucléaires (attaque) et post-nucléaires (coda) sont facultatives* » (Meynadier, 2001, p. 109). La frontière n'a alors plus de rôle segmental et découle de la structure linéaire de la syllabe.

Un autre point faible de ces approches est qu'elles ne permettent pas de comprendre pour quelles raisons certaines combinaisons de segments, bien que respectant le principe de sonorité, sont peu fréquentes dans les langues, alors que d'autres le sont davantage. Les échelles de sonorité étant, de plus, propres à chaque langue, il est difficile de tirer des conclusions générales sur la syllabe en tant qu'objet linguistique universel. Afin de surmonter ces interrogations, d'autres théories vont se développer au sein même de la phonologie linéaire.

Ainsi, à la marge de ce courant phonologique, Anderson et Jones (1974) supposent que la syllabe est régie par des relations de subordination entre ces éléments et posent ainsi les bases de la phonologie de dépendance, s'approchant alors d'une conception *multilinéaire* pour la fonction syllabique : le noyau domine les autres constituants, puisque ces derniers ne sont pas obligatoires et ne peuvent exister sans noyau.

### **I.1.2. Formalismes quasi-linéaires pour une définition de la syllabe phonologique**

Déterminer un formalisme pour la syllabe phonologique va contribuer à poser la question des relations qui la lient à ses unités, dorénavant définies comme : attaque (*onset*), noyau (*nucleus*) et coda, et à considérer comme primordiale la définition de la phonotactique.

L'un des principaux initiateurs de l'approche linéaire, Haugen (1956), définit la syllabe comme la plus petite unité de séquences phonémiques récurrentes non contrastive, contrairement au phonème. Toute syllabe se décomposant en attaque, noyau et coda, sa définition nécessite la délimitation des frontières entre attaque et coda, ainsi que la liste des phonèmes ou groupes de phonèmes pouvant apparaître dans chaque position.

Deux formalismes plus ou moins complexes proposés par Haugen ne permettront cependant pas de trancher la question des frontières. L'un des apports de cette approche est la

définition formelle des constituants de la syllabe (attaque, noyau, coda), même si Haugen ne réutilise pas ces derniers dans son approche. Un des problèmes qu'elle pose en revanche est qu'aucune généralisation de la syllabe n'est possible, chaque langue étant caractérisée par un diagramme particulier présentant les structures syllabiques possibles, complété d'un second présentant les instanciations permises pour les phonèmes de la langue.

Une autre approche proposera un formalisme à base d'arbres de dépendance. Développée en partie par Anderson et Jones (1974) et Anderson et Ewen (1987), elle se construit en opposition aux positions prises par Chomsky et Halle dans SPE (*The Sound Pattern of English*, 1968) qui réduisent leur étude de la syllabe à la création d'un trait [syllabique], sans jamais faire référence à une définition explicite de cette unité. La phonologie de dépendance considère aussi bien la structure de la syllabe, que celle des segments, mais aussi les relations entre segments. La syllabe émerge alors directement des segments : on reconnaît le noyau au fait qu'il est la tête (élément dominant) du domaine syllabique ainsi que de la rime. La syllabe est ainsi constituée d'une tête et, s'ils existent, des éléments que cette dernière domine. La tête est donc considérée comme le seul constituant obligatoire à partir duquel se construit la syllabe.

Dans cette étude formelle, comme dans les autres théories linéaires, les frontières se placent en fonction des ruptures dans la courbe de sonorité, mais l'échelle choisie, propre à cette approche, est un continuum et repose sur une redéfinition des segments. En effet, Anderson et Ewen (1987) ne considèrent pas le classement de SPE en classes naturelles pour les consonnes et les voyelles comme étant nécessaire si l'on utilise le traitement en syllabes. La caractérisation des segments ne se fait plus à partir de traits binaires, mais selon des composants unaires, qui sont définis en termes perceptifs. Par exemple pour les voyelles, les composants vocaliques sont au nombre de trois :

- |I| : avant ou aigu ;
- |A| : bas ou compact ;
- |U| : arrondi ou grave ;
- les autres segments étant définis à partir de plusieurs possibilités de combinaisons des trois premiers (avec relation de dépendance ou de domination) :

Exemples :   {|I=>A|} = /ε/           (i domine a)  
                  {|A=>I|} = /æ/           (a domine i)  
                  {|A<=>I|} = /ε/           (i domine a et a domine i)

Le même principe est utilisé pour la définition des systèmes consonantiques, avec la distinction suivante :

- |V| : élément « *relativement périodique* » ;
- |C| : composant « *de réduction périodique d'énergie* ».

Pour la structure syllabique, la hiérarchie suivante est postulée : les segments les plus |V| sont les plus proches du noyau alors que les segments les plus |C| en sont davantage éloignés. La syllabe dépend donc de la représentation des segments. Cette hiérarchie n'est cependant pas toujours généralisable à toutes les langues, car elle doit être adaptée pour tenir compte des particularités phonotactiques de chacune.

Cette théorie connaît actuellement plusieurs tentatives d'approfondissement, avec notamment les travaux de Hulst (1996) et son modèle RcvP (*Radical CV Phonology*) qui accentue les contraintes du modèle de Anderson et Ewen en se fondant sur un patron unique CV et seulement quatre primitives (C, Cv, Vc, V) exemptes de tout contenu phonétique.

L'examen des différentes approches de la syllabe en phonologie linéaire montre des avancées réelles sur le plan de sa description dans un certain nombre de langues avec la caractérisation des relations linéaires entre segments consécutifs dans les groupements consonantiques intra-syllabiques grâce à la mise en évidence du principe de sonorité et/ou des relations de dépendances. Pourtant, l'existence de principes universaux devant être adaptés pour chaque langue ne permet aucune généralisation. Le postulat d'une définition de la syllabe commune à toutes les langues ne paraît pas envisageable si la syllabe n'est considérée qu'en tant que séquence de segments sans tenir compte des inter-relations existant entre ces derniers.

### **I.1.3. La syllabe : une structure sonore hiérarchisée**

#### ***I.1.3.1. Hiérarchie inter-niveaux***

L'impossibilité de décrire des phénomènes suprasegmentaux par la phonologie linéaire, comme la réalisation des systèmes de tons ou le phénomène de gémination dans les langues africaines (cf. Creissels (1994) pour les langues bantoues), a conduit à développer une approche capable de rendre compte de ces phénomènes et de ceux relevant de la morphophonologie. L'information phonologique est alors décrite sur plusieurs niveaux distincts, les relations entre eux se définissant au sein des diverses théories phonologiques, dites *multilinéaires*. Celles-ci remettent fondamentalement en cause la conception de la syllabité de la Phonologie Générative qui réduit au maximum le rôle de la syllabe dans l'organisation et le fonctionnement du langage : la syllabe n'y apparaît plus que dans la définition de règles entre segments, jouant un rôle d'unité intuitive, sans statut linguistique, mais toutefois incontournable dans la description des systèmes linguistiques. La seule considération de la syllabe est la définition d'un trait [+ syllabique] à la place du trait [+ vocalique] apparaissant dans la tradition jakobsonienne.

Kahn (1976) proposera le premier une conception hiérarchique de la syllabe dans un cadre générativiste, avec la volonté de réduire l'ensemble des règles de SPE en un nombre restreint de principes pour délimiter la syllabe en regroupant les phonèmes sous des frontières syllabiques. L'ambisyllabité est traitée en proposant un formalisme permettant de tenir compte des consonnes appartenant à la fois à la coda d'une syllabe et à l'attaque de la syllabe

suivante. La structure syllabique y est appréhendée sans reprendre les notions d'attaque et de rime, pourtant déjà proposées par Fudge (1969).



Figure I.1 : Notation autosegmentale de [ætɫəs] (atlas) et de [poni] (poni) pour Kahn (1976).

Trois principes universels de bonnes formations régissent la structuration des syllabes, auxquels s'ajoutent des règles spécifiques pour chaque langue de manière à les rendre adaptées à leurs particularités syllabiques.

Kahn pose aussi le problème de la séparation entre le débit de parole normal (*connected speech*), l'épellation syllabique (*syllable-to-syllable speech*) et la parole rapide (*faster speech*). Il propose de déterminer, en premier lieu, les règles de syllabation du *syllable-to-syllable speech* transposables selon lui aux autres débits de parole grâce à l'introduction de lignes d'association supplémentaires. Cinq règles régissent alors l'affectation des niveaux supplémentaires lors des différents débits de parole.

À travers la prise en compte du débit, Kahn tente de démontrer la nécessité d'intégrer la variation linguistique dans le traitement de la syllabation des langues. Les plus vives critiques suscitées par cette démarche sont émises par Clements et Keyser (1983), estimant le phénomène de la resyllabation sous-traité. De plus, la démonstration de Kahn ne repose sur aucune structure interne de la syllabe en ne distinguant pas le noyau des éléments marginaux. Clements et Keyser proposent alors d'intercaler un palier CV intermédiaire (le niveau métrique) entre le niveau syllabique et le niveau segmental : « *le squelette* » qui correspond aux positions temporelles C (différent d'un noyau vocalique) ou V (noyau vocalique). Les segments ne sont donc plus intrinsèquement syllabiques, cette éventuelle caractéristique dépendant de leur position dans le squelette. Ce concept de position squelettale permet de rendre compte des segments phonologiques complexes, telles que, par exemple, les géminées (cf. figure I.2) :

Niveau métrique	C	C	C	C	V	V	V	V	C
		(	)	(		(	)		
Niveau segmental	t	t:	t s		a	a:	a	i	
	simple	gémignée	affriquée		brève	longue	diphtongue		
		Consonne				Voyelle			

Figure I.2 : Associations entre segments et positions CV selon Clements et Keyser (1983)

Dans ce formalisme, la syllabe est décrite par une structure à trois niveaux : niveau segmental, niveau des unités de timing (squelette) et niveau syllabique, occupant une place essentielle dans l'analyse du rythme : « *A theory of syllable representation which*

*characterizes the syllable as a three-tiered structure having the formal properties of an autosegmental system. We have claimed that the terminal element of syllable tree are not vowels or consonants themselves, but rather the units of the CV-tier which define positions in the syllable structure that particular consonants and vowels may occupy.* » (Clements et Keiser, 1983, p. 115)

Toutes les critiques de l'approche générativiste quant à l'organisation syllabique ne reprennent pas obligatoirement le niveau squelettal mais définissent toutefois différemment les liens entre les divers niveaux nécessaires à la description de la syllabe, telle la vision hiérarchisée de Fudge (1969). Il y aurait selon lui deux syllabes correspondant à deux niveaux de description (phonétique et phonologique), tout en considérant la syllabe comme un universel phonologique dont l'importance linguistique est l'élément abstrait qui permet de faire des généralisations et de répondre à deux fonctions principales :

- l'accès à la structure prosodique (reprise de Troubetzkoy, 1939) ;
- la mise en évidence des contraintes pesant sur les séquences de phonèmes possibles (règles phonotactiques).

Fudge (1969) fournira une représentation arborescente (qui évoluera avec, par exemple, l'abandon de la position « *Terminaison* » qui correspond à la position syllabique en finale absolue de mot) basée sur des règles et permettant de visualiser les différents phonèmes pouvant apparaître dans les diverses positions de la syllabe. Pour rendre compte de phénomènes irréguliers dans la langue, certaines règles peuvent être violées, obligeant Fudge à placer l'intérêt principal de leur application au niveau sémantique, établissant ainsi une condition entre règles et valeur expressive des mots.

Pour de nombreux autres linguistes, la syllabe est une structure hiérarchisée de constituants sub-syllabiques plus complexe que celle proposée par Clements et Keyser (1983). Pike et Pike (1947) ont été les précurseurs de cette analyse en constituants immédiats avec hiérarchisation de ces derniers. Leur approche est, à l'origine, plutôt descriptive, et motivée par le fait qu'il est plus économique de décrire les langues en syllabes plutôt qu'en phonèmes, comme par exemple pour le mazateco (famille Oto-Manguean, groupe Mazatecan). Selon eux, le principal problème est de déterminer la structure de la syllabe et les distributions possibles des segments dans cette structure. Les syllabes sont considérées non comme une succession linéaire de segments, mais comme des « *couches superposées de briques* », liées entre elles par une structure. La syllabe phonétique fondée sur des critères physiques et acoustiques est distinguée de la syllabe phonologique qui provient de critères distributionnels et sert pour la résolution de problèmes phonologiques comme le placement de l'accent. Leurs travaux évoquent la notion de gabarit syllabique avec des positions pouvant être occupées par un ou plusieurs segments. Mais les post-SPE retiendront surtout de ces études la comparaison de la syllabe à une structure syntaxique.



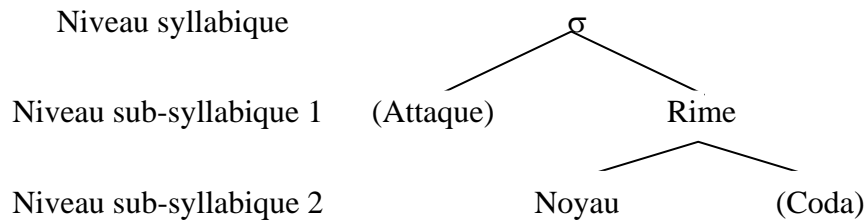


Figure I.3 : Représentation arborescente de la structure interne de la syllabe en sous-constituants hiérarchiquement organisés. Entre parenthèses, les constituants facultatifs.

En fait, les approches métriques considèrent la syllabe comme un gabarit phonologique abstrait et préexistant qui permet de rendre compte des contraintes phonotactiques entre ses constituants. La formation de ces derniers repose sur les principes suivants : plus deux éléments sont liés phonologiquement, plus ils cherchent à se regrouper en un même constituant. Inversement, deux éléments apparaissant dans des constituants différents manifestent une certaine indépendance.

La principale différence entre les diverses approches de ce courant réside dans l'analyse de la sous-division de la rime. Cette dernière doit son apparition à des observations qui lient plus fortement le noyau à la coda que le noyau à l'attaque. La première explication proposée se fonde sur le patron de sonorité qui permet de distinguer deux parties dans la structure syllabique. Selon Klein (1993), la manière d'intégrer linéairement les segments en serait responsable : l'attaque répondrait à un principe reposant sur l'échelle de consonanticité alors que la rime s'appuierait sur celle de sonorité. Cette dichotomie prend aussi du sens dans la détermination du poids des syllabes, car seule la composition de la rime y joue un rôle. Certaines évolutions diachroniques soutiennent aussi une dépendance plus importante entre le noyau et la coda qu'entre l'attaque et le noyau. Nous confirmerons cette observation dans notre partie résultats (chapitre III.3.) avec l'analyse de combinaisons syllabiques favorisées qui mettent elles aussi en évidence un lien plus fort entre le noyau et la coda d'une syllabe qu'entre le noyau et l'attaque.

L'adoption d'une structure hiérarchisée pour la syllabe permet non seulement de rendre compte du fonctionnement syllabique à partir des relations entre constituants, mais aussi de certaines restrictions d'occurrence des traits distinctifs traitées au niveau des constituants de la syllabe et non pas à celui des segments. La structure hiérarchisée est utilisée dans de nombreuses études pour décrire l'organisation syllabique d'une ou plusieurs langues. Ainsi Selkirk (1984) reprend, pour l'anglais, le schéma d'arbre à branchement et tente de lier la notion d'échelle de sonorité à celle de gabarit, en postulant que chaque langue présente au niveau lexical un ensemble de gabarits syllabiques possibles avec des attaques maximisées tant qu'elles ne vont pas à l'encontre des contraintes pesant sur les gabarits. De la sorte, l'affectation d'un segment se fait majoritairement en faveur de l'attaque si deux choix (attaque ou coda) sont possibles. La position des segments dans le mot peut modifier ce principe d'affectation : [ps] ou [sl] sont des attaques possibles en français à l'initiale de mot,

alors qu'en interne de mot, ils se trouvent dans deux syllabes consécutives. Exemples : [psi . ko . lo . ʒi] vs. [kap . syl] et [slav] vs. [is . lãd].

Dans sa démonstration, Selkirk propose d'affecter les différents segments aux positions squelettiques terminales des gabarits syllabiques en fonction des capacités de ces dernières à accueillir des segments dotés d'un indice de sonorité donné (cf. figure I.4). Les conditions exactes d'association dépendent de chaque langue. Ainsi, à partir de la représentation des positions terminales et de règles différentes, sa démonstration tient compte de la phonotactique de chaque langue.

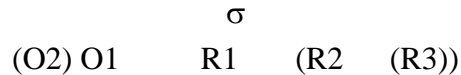


Figure I.4 : Représentation des positions terminales de la structure syllabique selon Selkirk (1984, p. 115) (avec O = onset, R = rime, et les chiffres indiquant la position dans le constituant)

Selkirk introduit un indice de sonorité, IS entre 0,5 pour [p, t, k] et 10 pour [a], permettant l'affectation de certains phonèmes aux différentes positions. En plus du principe de « *Sonority Sequencing Generalization* » (toute syllabe comporte un pic de sonorité précédé et suivi d'une séquence de segments ayant des valeurs de sonorité progressivement décroissantes), Selkirk introduit dans sa démarche la « *Minimum Sonority Difference* » (entre deux positions syllabiques on observe une différence de sonorité dont le minimum est défini au niveau des positions terminales). Les travaux de Selkirk ont été construits sur l'anglais et l'espagnol, mais des problèmes se posent, comme le souligne Angoujard (1997, p. 31), lorsque la démonstration est appliquée à d'autres langues comme l'arabe.

Les théories autosegmentale et métrique ont très vite été perçues comme complémentaires, incitant les linguistes à combiner les avantages des deux théories. Il en résulte des propositions d'une structure constituante hiérarchique hybride comme celle mentionnée dans la figure I.5 :

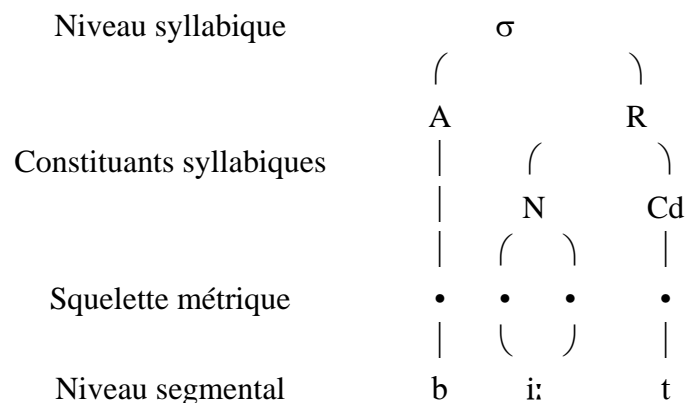


Figure I.5 : Représentation syllabique du mot anglais « beat » combinant structure autosegmentale et métrique (exemple de Goldsmith, 1990)

Le squelette répond alors à des positions internes à la syllabe, et non plus seulement à des positions temporelles et générales. « *Cette approche permet de rendre compte à la fois des relations entre les unités segmentales et leur poids métrique par la ligne temporelle de la base (squelette), et des restrictions distributionnelles des segments par la constituance structurelle de la syllabe* » (Meynadier, 2001, p. 121). D'autres approches vont alors émerger, telles que celle de Encrevé (1998) où les constituants sont traités comme des paliers parallèles indépendants ou autosegments. Dans les approches *multilinéaires*, la syllabe est considérée comme un gabarit métrique et phonotactique sous-jacent.

### **1.1.3.2. Hiérarchie inter-constituants**

La phonologie du gouvernement de Kaye, Lowenstamm et Vergnaud (1988) de type *Principes et Paramètres* (Principes de la Grammaire Universelle (GU)) fondée sur une approche morphophonologique, considère la syllabe dès le niveau lexical et les contextes morphologiques comme les réanalyseurs syllabiques. Un des principaux outils fourni par cette théorie est la notion de positions vides non-exprimées, qui s'ajoute aux indices de marque et à une représentation squelettale.

Kaye, Lowenstamm et Vergnaud (1990) proposent une évolution de leur théorie dans laquelle la syllabe n'est plus un élément important, lui substituant une organisation en terme de domaines de gouvernement, impliquant que tous les constituants syllabiques sont maximalelement binaires. Ils redéfinissent une organisation interne à la syllabe à partir des caractéristiques phonétiques des segments, en restant dans un cadre descriptif, utilisées pour affecter les segments aux positions squelettale. Les segments gouverneurs doivent être charmés et les éléments gouvernés sont sans charme. Tout élément vocalique est charmé positivement et tout élément consonantique l'est négativement, avec entre les deux des éléments neutres (glides, liquides, nasales). Seuls les éléments charmés positivement peuvent apparaître dans les positions nucléaires. Un gouvernement inter-constituant fondamentalement différent du précédent sera défini par plusieurs principes, dont certains issus de la théorie de Gouvernement Propre en syntaxe (Chomsky, 1987).

Charette (1990) propose une évolution en remplaçant le principe du Gouvernement Propre (qui exclut tout autre domaine de gouvernement) par une condition de licenciement : pour qu'il y ait une relation de gouvernement entre une tête non-nucléaire A et son complément B, A doit être licencié par le noyau précédent. La théorie sera radicalisée par Lowenstamm (1996) qui propose de passer à un cadre plus strict en posant que toute structure syllabique, quelle qu'elle soit, se réduit à une forme CV. Dès lors, toute syllabe fermée, avec groupement consonantique, ou avec géminée, doit être réinterprété en fonction du cadre CV (cf. figure I.6) :

Syllabe CVC				Syllabe avec attaque branchante			
A	N	A	N	A	N	A	N
C	V	C	Ø	C	Ø	C	V
t	a	t		t		r	a

Figure I.6 : Représentation de deux types de syllabes (CVC et CCV) selon le formalisme proposé en phonologie du gouvernement.

L'utilisation des catégories vides apporte une vision différente des possibilités de représentations syllabiques. Mais le problème des groupes consonantiques super-lourd (VCCC), comme il en existe dans des langues telles que le tachelhit (langue berbère de la famille afro-asiatique), semble nécessiter une nouvelle évolution de la théorie.

#### I.1.4. La syllabe réinterprétée

En 1997, Angoujard constate que si une théorie phonologique de la syllabe n'a pas encore été élaborée, « *ce n'est peut-être pas parce que nous n'avons pas compris ce qu'est la syllabe, mais tout simplement parce que cet objet n'existe pas, qu'il n'est, au mieux, que l'effet variable des interactions entre plusieurs éléments élémentaires* » (Angoujard, 1997, p. 24). Ainsi, il propose de remplacer la notion de syllabe par la rythmicité en considérant que toute unité lexicale est décomposable selon un modèle « *composé d'un creux initial, d'un sommet (noté par la présence d'un X de niveau 2) et d'une troisième position (de type « creux ») facultative (cf. figure 7). La conformité des séquences de positions à ce modèle rythmique impliquera donc :*

- *l'obligation, pour toute séquence, de commencer par un creux ;*
- *l'impossibilité de voir se succéder deux sommets (la position suivant un sommet sera soit une position 3 soit une position de type 1) ;*
- *l'impossibilité de voir se succéder plus de deux creux rythmiques »* (Angoujard, 1997, p. 80).

X	X X	(X)
1	2	3

Figure I.7 : Modèle rythmique proposé par Angoujard.

Cependant, si le nombre de positions est supérieur à quatre, le recours à l'échelle de sonorité et à l'association segments/position est nécessaire afin de choisir la bonne grille rythmique parmi les différentes possibilités offertes par le modèle. Les principes de l'échelle de sonorité définie par Angoujard sont tirées des caractéristiques phonétiques des segments (cf. partie 1.1.2.).

Afin de permettre la représentation de tous les types de « syllabes », Angoujard précise que des positions vides peuvent apparaître (cf. figures I.8 et I.9), même dans des positions de sommet. Mais que dans ce cas, une voyelle facultative telles [i] ou [ə] disparaîtra. Le modèle rythmique ne permet cependant pas à lui seul de prédire quelles positions pourront rester vides ou devront être remplies par [i] ou [ə] pour les sommets et [ʔ] pour les creux. La théorie suggère que, pour chaque langue, l'étude des courbes prosodiques permet de définir les structures privilégiées et de prédire le remplissage des positions vides.



Figure I.8 : Grille rythmique du terme « ami » (à gauche) et « chaos » (à droite)  
(Angoujard, 1997).

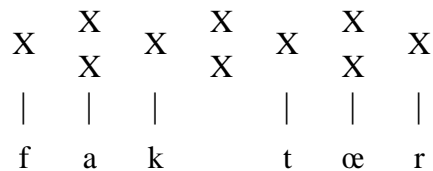


Figure I.9 : Grille rythmique du terme « facteur » (Angoujard, 1997).

Ces courbes sont vues comme des objets discontinus qui comprennent une position initiale, un sommet et une position finale facultative, avec la possibilité d'apparition de positions intermédiaires. Les positions initiales et finales sont associées aux positions rythmiques 1 et 2 s'il n'y a pas de positions rythmiques vides. L'affectation des segments aux positions rythmiques se fait alors en fonction de la courbe prosodique : « *un segment ne peut être associé qu'à une position rythmique elle-même associée à une position de courbe (nous dirons, à une position rythmique liée)* » (Angoujard, 1997, p. 112). Cette condition permet de rendre compte des possibilités pour une langue de posséder des positions 3 non-vides. L'association de la courbe prosodique et du modèle rythmique permet, selon Angoujard, d'expliquer certains problèmes phonologiques connus comme la gémation ou les quadrilitères du sémitique.

Pour Angoujard, « *l'abandon de [l'objet syllabe], ou, plus exactement, sa réinterprétation au travers de l'interaction de trois objets élémentaires (la grille rythmique, le segment et la courbe prosodique), ouvre la voie à une analyse univoque et basée sur des contraintes* » (Angoujard, 1997, p. 197). Cette approche apporte une nouvelle vision des possibilités d'analyse de l'objet syllabique au travers d'éléments qui ne sont plus uniquement les segments. Elle ne permet en revanche pas de comprendre pourquoi, outre par la hiérarchisation des segments, les langues favorisent certaines séquences par rapport à

d'autres, avec des caractéristiques semblables, que ce soit dans la sonorité des segments, au niveau rythmique comme au niveau prosodique.

Les approches multilinéaires ont permis de mettre en évidence des relations particulières entre les différents constituants de la syllabe, par exemple en illustrant par la création du nœud rime le fait que les liens entre le noyau et la coda semblent plus étroits que ceux entre l'attaque et le noyau. Pourtant, si les syllabes d'un grand nombre de langues sont de mieux en mieux décrites et comprises, il ne semble pas encore possible de définir des lois universelles s'appliquant à l'organisation syllabique quelle que soit la langue étudiée.

La théorie de l'Optimalité (« *Optimality Theory* », OT), avec la possibilité de violer les contraintes et de les faire évoluer, se donne les capacités de rendre compte à la fois de l'universalité de certains phénomènes linguistiques, et de la variation qui existe entre les langues du monde. Évolution des plus récente de la phonologie générative, OT (Prince et Smolensky, 1995) propose de remplacer les notions de dérivation et d'opération pour ne retenir qu'un ensemble de contraintes hiérarchisées entre elles selon un ordre spécifique à chaque langue et appartenant directement à la grammaire universelle et permettant de mettre en évidence les représentations. Le choix de règlement des conflits apparaissant lors de l'application de contraintes contradictoires permet d'expliquer les différences qui existent entre les langues.

Les contraintes proposées sont de deux sortes : contrainte de fidélité et contrainte de marque. La syllabe *y* est interprétée comme le résultat :

- 1) de contraintes de marque qui vont permettre de définir les conditions de bonne formation des éléments, ainsi que l'information sur sa valeur marquée ou non. Ainsi, quatre contraintes permettent de « *poser CV comme la syllabe canonique : une syllabe doit posséder une attaque et un noyau mais pas de coda, et aucun de ses constituants ne peut être complexe.* » (Durand et Lyche, 2001, p. 131).
- 2) de contraintes de fidélité qui seront du type suivant :
  - « *MAXIO Chaque élément de base a un correspondant dans la représentation de surface (l'élision est interdite).*
  - « *DEPIO Chaque élément de la représentation de surface a un correspondant dans la base (l'épenthèse est interdite).* » (Durand et Lyche, 2001, p. 130).

La possibilité de tenir compte de propriétés phonétiques des éléments dans le choix des contraintes pourrait être une voie de développement prometteuse pour expliquer la phonologie de la syllabe. Cependant, la plupart des théories phonologiques privilégient les faits observés sans explications recherchées dans la substance, même si certaines des avancées obtenues en phonétique s'intègrent aujourd'hui plus largement dans la recherche d'explications des contraintes et des règles mises en évidence dans l'organisation syllabique des langues (cf. chapitre III).

### I.1.5. Syllabe et connexionnisme : modélisation de processus de syllabation

Au vu des problèmes non traités par les théories de la syllabe, qu'elles soient linéaires ou non, ainsi que des résultats récents en neurologie et en cognition, certains linguistes ont proposé une nouvelle approche de l'organisation syllabique : la phonologie connexionniste. Ce courant propose des modèles subsymboliques où le processus calculatoire est supporté par la diffusion d'énergies et l'auto-stabilisation dynamique des systèmes. Dans ce cadre, et en s'inscrivant dans la continuité de la phonologie autosegmentale, Goldsmith (1993) développe la phonologie harmonique fondée sur trois niveaux essentiels d'analyse comme :

- le niveau M, morphophonologique : celui auquel les morphèmes sont phonologiquement spécifiés ;
- le niveau W, phonologique : où les expressions sont structurées en syllabes et en mots bien formés, mais avec un minimum d'information phonologique redondante ;
- le niveau P phonétique : un niveau de description phonétique large qui sert d'interface avec les dispositifs articulatoires et acoustiques périphériques.

Chaque niveau possède ses propres processus de syllabation et l'application libre et non-ordonnée de ces derniers permet d'obtenir un maximum harmonique, livrant la bonne formation d'une représentation.

Les modèles connexionnistes se fondent sur des réseaux dits neuromimétiques à cause de leur similarité avec les réseaux utilisés par les traitements neurobiologiques. Un des modèles proposés est le DLM (*Modèle Linéaire Dynamique*, Goldsmith et Larson, 1990 ; Goldsmith, 1993 ; Laks, 1995), qui reprend l'idée d'échelle de sonorité, mais considère la sonorité d'un segment comme une variable dépendante de la sonorité des segments gauche et droit. Le modèle fonctionne en associant les autosegments à des unités squelettales (neurones formels) et en calculant la sonorité dérivée des segments par la formule suivante :

$$k_i^{t+1} = [\text{sonorité inhérente de } k_i] + \alpha \cdot k_{i+1}^t + \beta \cdot k_{i-1}^t$$

Les valeurs de  $\alpha$  et de  $\beta$  sont identifiées lorsqu'est atteint un état d'équilibre, révélant la syllabation correcte. Selon Goldsmith, les paramètres sont les mêmes pour toutes les langues, alors que Laks (1995) montre qu'ils peuvent varier. La sonorité inhérente aux segments est calculée selon un produit de vecteurs dont le premier spécifie le segment et le second la valeur des différents traits binaires pour la langue étudiée. Lorsque l'état d'équilibre est atteint, les sonorités dérivées calculées permettent d'obtenir la syllabation cherchée. Le maximum local de sonorité dérivée correspond à un pic ou noyau de syllabe et un minimum de sonorité dérivée correspond à une attaque.

La critique formulée par Klein (1993) reproche au DLM de ne pas permettre d'exprimer les relations existant entre les constituants de la syllabe, puisque cette dernière découle

uniquement des interactions segmentales. Le DLM ne permet donc pas de traiter des problèmes de diérèse ou de synérèse. Pour y remédier, Klein propose deux « *Principes d'Intégration Linéaire Segmentaux* » basés sur la séparation des représentations articulatoires et perceptives. Il postule que les voyelles maximisent leur représentation perceptive, alors que les consonnes maximisent leur représentation articulatoire. La syllabe est alors définie à un niveau suprasegmental comme l'interface entre les deux représentations, les segments étant définis en termes de poids articulatoire et perceptif.

Ces théories semblent ouvrir la voie à des avancées prometteuses en lien avec l'ensemble des recherches sur la cognition en général, et la place qu'y occupe le langage en particulier. Les évolutions de la phonologie connexionniste semblent aller dans le sens d'un rapprochement avec les recherches actuelles en phonétique, psycholinguistique, ainsi qu'avec les résultats récents en acquisition et en pathologie du langage.

Si les avancées des recherches phonologiques montrent la nécessité de mieux connaître la syllabe en se basant sur les différentes règles qui contraignent l'organisation des segments, il est pour le moment difficile d'expliquer, uniquement avec les principes des phonologies actuelles, les réalisations et généralités observées à travers les langues du monde. Au-delà d'une démarche purement descriptive, les démonstrations des différentes phonologies semblent, en effet, ne pas être encore en mesure de définir l'existence d'une unité syllabique à partir des données de langue, ni d'expliquer l'origine de l'organisation syllabique de ces dernières. Il semble qu'il faille aller chercher ces explications dans des disciplines connexes. Cette ouverture à d'autres éléments qu'aux données observées de la langue semble constituer à l'heure actuelle un point important pour le développement futur des théories phonologiques en ce qui concerne la syllabe, comme c'est le cas, depuis une trentaine d'années, pour les systèmes sonores des langues du monde (pour les systèmes vocaliques : Liljencrants et Lindblom, 1972 ; Lindblom, 1986 ; Vallée, 1994 ; Schwartz *et al.*, 1997 ; Lindblom, 2000 ; plus récemment pour les systèmes consonantiques : Vallée *et al.*, 2002 ; Boë *et al.*, à paraître 2004).

## I.2. DESCRIPTIONS PHONETIQUES

En balayant la plupart des courants théoriques de la phonologie, nous constatons que la syllabe est une unité qui a été beaucoup étudiée par les phonologues, mais sans qu'un consensus sur sa définition n'arrive véritablement à émerger. Les phonéticiens se sont aussi beaucoup interrogés sur la nature de la syllabe. Mais au lieu de prendre pour point de départ l'existence ou non d'une unité syllabique, ils recherchent les éléments présents dans la parole qui pourraient aller dans le sens d'une définition physique de la syllabe.

Plusieurs voies de recherche existent et se différencient par les méthodes et moyens mis en œuvre. Certaines études tentent de démontrer l'existence de la syllabe à travers des corrélats perceptifs pour la plupart reliés à des indices articulatoires. À ce type de travaux



des phonéticiens s'ajoute aussi l'apport des psycholinguistes afin de déterminer si la syllabe est ou non une unité pertinente dans les représentations mentales du langage. Une autre approche aujourd'hui très exploitée est la recherche d'une base motrice à la syllabe, notamment au travers de la théorie « *Frame, then Content* » (MacNeilage, 1998).

### **I.2.1. Recherche d'indices articulatoires et perceptifs**

Contrairement au phonème, la syllabe est une unité qui possède une réalité consciente chez les locuteurs. Par conséquent, le phonéticien devrait trouver des indices permettant de la retrouver dans le signal acoustique. Ainsi, de nombreuses études ont porté et portent encore sur la recherche de ces indices syllabiques potentiellement présents dans le flux de parole et qui permettent le découpage syllabique des mots et des énoncés. D'autres s'intéressent plutôt à la perception des différents types syllabiques au cours du développement langagier, dans les différentes phases d'acquisition de la parole.

Kawasaki (1982) montre ainsi qu'un certain nombre de contraintes phonotactiques peuvent s'expliquer en grande partie par des phénomènes auditori-acoustiques, fournissant une confirmation supplémentaire de l'hypothèse selon laquelle l'auditeur joue un rôle significatif dans la détermination des systèmes sonores et l'évolution des changements phonologiques. Son étude porte par exemple sur la recherche d'explications perceptives pour les contraintes phonotactiques pesant sur l'organisation des segments dans les syllabes des langues du monde : « *The present study proposes a hypothesis that they [some universally attested sequential constraints] are, to a large extent, acoustically and auditorily based. One determinant, the magnitude of acoustic modulation in a given sound sequence, assumed to be directly related to its perceptual saliency, should affect the viability of the sequence. Another determinant, the degree of acoustic difference among various sound sequences, should be inversely related to the likelihood of perceptual confusion and hence predict the sequence which are less susceptible to merger than others* » (Kawasaki, 1982, p. 4).

Son étude porte sur deux facteurs acoustiques qui sont posés comme déterminants pour connaître les séquences de sons favorisées ou non : la magnitude de la modulation acoustique dans une séquence et le degré de différences acoustiques entre des séquences. Le manque de modulation acoustique expliquerait la faible fréquence des groupements consonantiques, vocaliques et des groupements d'obstruantes avec le même mode d'articulation aussi bien que celle des groupements 'nasale + liquide' en opposition aux groupements 'obstruente + liquide'. L'hypothèse est testée expérimentalement par des applications à certaines contraintes séquentielles de nature plus spécifique.

Les auditeurs, ou les détecteurs de signaux, doivent préférer des signaux bien modulés acoustiquement, pour qu'ils soient perceptuellement saillants par rapport aux signaux avec des petits changements. Par conséquent, les séquences sonores universellement favorisées devraient être celles qui présentent de grands changements acoustiques. Une autre hypothèse est proposée par Kawasaki (1982) : si des séquences de deux sons ou plus partagent certaines

caractéristiques acoustiques, ils peuvent être perceptuellement similaires. Sans tenir compte de la somme des changements acoustiques dans de telles séquences, ils peuvent être confondus et, par conséquent, susceptibles de fusion. Ainsi, les séquences de sons favorisées devraient non seulement présenter un contraste acoustique maximal entre ses constituants, mais aussi être suffisamment différentes des autres séquences de sons.

Les résultats acoustiques de Kawasaki (1982) suggèrent que les groupements ‘occlusives + [r]’ sont phonotactiquement moins sévèrement contraints que les groupements ‘occlusives + [l]’ parce que les premiers sont mieux modulés et mieux contrastés entre eux que les seconds. Une légère préférence pour les groupements ‘occlusives alvéolaires + [r]’ peut être due à leurs importants changements dans les fréquences formantiques. La non préférence des groupements ‘occlusive alvéolaire + [l]’, cependant, ne peut être expliquée en faisant référence aux mêmes facteurs. Il semble que cette contrainte particulière soit attribuable aux similarités spectrales de ces groupements avec les groupements ‘occlusive + sonante’, spécialement les groupements ‘vélaire + [l]’ et ‘occlusive alvéolaire + [w]’. Ces résultats prédisent aussi la rareté des séquences ‘liquide + voyelle d’arrière’, fait qui n’est pas immédiatement confirmé par les données phonologiques.

Selon Kawasaki, cette approche fournit une voie possible pour expliquer les tendances universelles comme l’évitement de groupements longs de consonnes et de voyelles, les modulations acoustiques étant relativement faibles dans ces séquences. La même explication est proposée pour la tendance à éliminer les séquences de consonnes produites avec le même mode d’articulation, la rareté des groupes ‘nasale + liquide’ par rapport aux groupes ‘obstruente + liquide’, ainsi que la rareté des nasales postérieures en début de mot.

D’autres recherches se sont davantage intéressées à la mise en évidence d’indices révélateurs des différentes positions consonantiques des syllabes, et ce pour identifier les mécanismes perceptifs en jeu dans les délimitations de syllabes. Redford et Diehl (1999) tentent de comprendre les différences observées entre la perception des consonnes en position initiale et en finale de syllabe, et proposent d’expliquer la prédominance des syllabes avec consonne initiale grâce à des propriétés de perception ou de production. Pour cela, deux expériences d’identification de consonnes sont organisées sur l’anglais-américain. Les mesures sont effectuées sur : la durée, l’amplitude, la mesure de fréquence moyenne, et deux mesures de fréquence en relation avec les transition sur F2. Les résultats obtenus leur permettent de conclure que l’avantage perceptif remarqué pour les consonnes initiales semble bien avoir une base dans la production : les consonnes en attaque de syllabe sont plus longues et plus fortes.

Cette étude vérifie trois prédictions :

- 1) les consonnes initiales sont mieux identifiées que les consonnes finales ;
- 2) cet avantage perceptif existe quelles que soient les consonnes, les voyelles et les locuteurs ;
- 3) les consonnes post-vocaliques précédant une voyelle sont mieux identifiées que celles se trouvant devant une autre consonne.

L'expérience de Redford et Diehl met en évidence qu'au niveau des propriétés acoustiques, l'amplitude contribue plus à la reconnaissance de la position consonantique que la durée. Les mesures de fréquences, notamment les transitions en F2, contribuent aussi fortement à l'identification des consonnes. Le contexte ne semble pas être responsable de l'avantage perceptif observé pour les consonnes initiales, mais accentue le désavantage de certaines consonnes en position finale. Ce qui rejoint bien l'observation faite par les recherches en typologie que seul un sous-ensemble de l'inventaire des consonnes apparaît en position finale (cf. partie III.4.4) et les travaux des phonologues sur la lénition dans les langues, beaucoup plus fréquente en finale qu'en initiale de syllabe (Colloque du GDR « Phonologies », 1999).

L'étude de Redford et Diehl montre bien l'existence d'une base articulaire aux différences de perception observées entre les consonnes en position initiale et celles en finale de syllabes. Mais il faudrait tester ces résultats sur des locuteurs d'autres langues, et vérifier si l'avantage attribué aux consonnes en initiale de syllabe ne se limite pas aux consonnes en initiale de mots afin de pouvoir considérer ces résultats comme représentatifs de l'organisation syllabique des langues du monde.

Ces résultats vont dans le même sens que ceux de Fougeron et Keating (1997) qui montrent que pour les consonnes en initiale de domaines prosodiques, les contacts entre langue et palais sont plus importants et plus appuyés : « *At least some tongue blade and body articulations are more extreme in domain-initial position.* » (Fougeron et Keating, 1997, p. 3739). Par ailleurs, différentes études (Krakow, 1999) mettent en évidence qu'en attaque de syllabe, la force articulaire est plus grande, le geste plus précis, et que les segments montrent une plus grande stabilité, une résistance plus importante à la coarticulation.

D'autres travaux se sont focalisés sur la recherche de contraintes perceptives pouvant expliquer l'organisation syllabique. En effet, s'il est évident à travers la littérature que certains types de changements phonétiques et syllabiques correspondent directement à des phases de transition entre les gestes articulaires, il est intéressant de mieux comprendre le concept de stabilité du phasage relatif des mouvements des articulateurs et de prouver sa validité. Dans ce but, Tuller et Kelso (1990, 1991) proposent plusieurs expériences qui tentent de montrer que les changements linéaires articulaires sont perçus de façon discontinue : « *exploit articulatory nonlinearities, that is, phase transitions, as a tool for identifying collective variables for speech production. [...] If the nonspecific influence of a continuously varying control parameter results in a change of pattern, then one can identify collective variable and determine their dynamics through an analysis of stability. Furthermore, if a candidate collective variable and its dynamics could be found, they might also constitute relevant information for speech perception.* » (Tuller et Kelso, 1991, p. 502).

Sur ce point, Tuller et Kelso (1991) ont mené une expérience avec les syllabes /ip/ et /pi/, en multipliant le débit de parole de départ par 4 et en l'augmentant jusqu'à 6. Un calcul de phasage entre les mouvements pharyngaux et ceux des lèvres a été réalisé pour chaque syllabe. Ils notent que la syllabe /pi/ garde à peu près le même phasage quelle que soit la

vitesse de débit. Pour la syllabe /ip/, par contre, quand le débit augmente, le phasage change pour rejoindre celui de /pi/. Tuller et Kelso (1991) attribuent le changement de phasage des syllabes VC à un manque de stabilité articulatoire. Leurs résultats montrent que le continuum linéaire des changements acoustiques peut avoir des conséquences discontinues sur la perception. « *When an articulatory phase transition was observed, it correspond precisely to a perceived change in syllable structure.* » (Tuller et Kelso, 1990, p. 447) Les conséquences acoustiques d'un changement du phasage articulatoire sont par conséquent suffisantes pour changer la perception de la structure de la syllabe. Cette étude constitue donc un argument important concernant la relative stabilité de CV par rapport à VC, et rejoint une étude précédente (Tuller et Kelso, 1990) qui affirmaient que VC se transforme en CV alors que l'inverse ne se produit jamais et donc que CV est intrinsèquement plus stable.

Ces résultats mettent en évidence la plus grande stabilité perceptive de la structure CV par rapport à la structure VC, et le fait que cette différence relève des caractéristiques articulatoires des deux formes. Tuller et Kelso (1991) concluent que des études menées sur la production devraient permettre de comprendre cette stabilité des articulateurs qui facilite la perception des stimuli plus stables.

Dans un cadre plutôt psycholinguistique, de nombreuses études se sont penchées sur la question de la pertinence de l'unité syllabique dans la recherche des représentations mentales du langage. Certaines se sont concentrées sur les capacités de segmentation des unités syllabiques par des locuteurs adultes, ou même des nouveaux-nés. Enfin, d'autres études ont porté aussi sur les processus d'extraction d'unités lexicales du flux de parole.

La recherche des unités perceptives qui rendent possible la segmentation du signal de parole représente une des questions centrales de la psycholinguistique. Plusieurs candidats possibles ont par conséquent été étudiés, même si tous n'ont pas revêtu la même importance. En effet, le phonème a traditionnellement été considéré comme l'unité de base de la perception de la parole. Pourtant, les difficultés rencontrées pour retrouver les informations sur un phonème précis dans le signal acoustique ont amené un certain nombre de chercheurs à abandonner cette unité et à choisir la syllabe comme unité de base de la perception : « *Considered as the smallest pronounceable unit, a popular choice for 'the primary perceptual unit', and now used as an acoustical representation for robust speech recognition, the syllable is of importance across a wide spectrum of speech research fields.* » (Goslin *et al.*, 1999).

Pour comprendre la place de la syllabe dans la segmentation et la reconnaissance de parole, Mehler *et al.* (1981) ont réalisé une expérience portant sur la reconnaissance de séquences cibles de phonèmes qui correspondent à une séquence sonore présente dans le mot entendu, mais n'appartiennent pas forcément à la même syllabe. On trouvera par exemple les cibles /pa/ et /pal/ avec les mots /palace/ et /palmier/. Les séquences sonores des deux cibles apparaissent toutes deux dans les deux mots, mais elles correspondent chacune à la première syllabe d'un seul des deux mots. L'hypothèse proposée est que si la syllabe est l'unité de

segmentation de la parole, les sujets devraient être plus rapides à détecter la cible correspondant à la structure syllabique du mot entendu. Les résultats valident cette solution puisque les temps de réaction des sujets semblent bien être liés à l'adéquation ou non de la cible et de la première syllabe du mot entendu, alors que deux autres portant sur la longueur de la cible sont rejetées. Pour compléter les résultats obtenus, une deuxième expérience est montée avec les cibles /a/ et /al/. Les résultats obtenus vont dans le même sens que ceux de la première expérience : « *the results reported here provide evidence for syllabic segmentation of speech* » (Mehler *et al.*, 1981, p. 304).

L'approche de la syllabe comme unité fondamentale de perception de la parole est en accord avec des analyses menées sur des nouveaux-nés et portant sur les capacités de reconnaissance de la parole. En effet, dès la naissance, les nouveaux-nés possèdent des capacités à discriminer un certain nombre de phénomènes linguistiques, et différents travaux montrent que la syllabe est l'un d'entre eux. Plusieurs études portent sur les capacités des nourrissons, dès 2 ou 3 jours, à retrouver une information syllabique dans des listes de mots présentés auditivement afin de détecter leur compétence dans la discrimination de divers types d'unités lexicales. La majorité de ces expériences utilisent la mesure de l'amplitude de la succion non-nutritive pour mettre en évidence les réactions des enfants devant les stimuli présentés.

Ainsi, Bijeljac-Babic, Bertoncini et Mehler (1993) ont mené trois expériences pour observer si des enfants âgés de 4 jours discriminent les énoncés multisyllabiques sur la base du nombre de syllabes ou du nombre de phonèmes. La première expérience montre que ces enfants séparent les énoncés composés de 2 ou 3 syllabes CV. La deuxième expérience a été réalisée afin de vérifier si la durée en nombre de phonèmes (4 ou 6) des stimuli dissyllabiques présentés aux enfants influent sur les résultats. La discrimination des deux types de stimuli ne semble pas affecter par ce critère, confirmant les résultats de la troisième expérience qui observe la reconnaissance d'unités dissyllabiques possédant 4 ou 6 phonèmes. Les résultats ne donnent aucune preuve que les nouveaux-nés soient sensibles à un tel changement dans le nombre des constituants de la syllabe. Les résultats provenant de ces trois expériences soutiennent donc l'idée que les nouveaux-nés utilisent l'unité syllabique pour séparer des listes de mots, mais sans tenir compte de la composition interne de la syllabe.

Cette hypothèse est confortée par les résultats de Ooijen *et al.* (1997) sur la distinction par des nouveaux-nés âgés de 2 jours des syllabes faibles et fortes. L'expérience a été menée sur deux groupes de vingt enfants français qui ont chacun entendu deux listes d'énoncés : la première, commune aux deux groupes, est une liste de mots dissyllabiques alternant des syllabes faible (possédant une voyelle réduite non-accentuée) et forte (possédant une voyelle pleine portant l'accent primaire ou secondaire) ; la seconde est soit une liste de mots monosyllabiques forts, soit une liste de mots dissyllabiques ayant deux syllabes fortes. Les résultats montrent que les nouveaux-nés discriminent mieux les mots dissyllabiques des mots monosyllabiques que les mots dissyllabiques entre eux, même si le poids des syllabes varie. Ces résultats vont bien dans le sens d'une discrimination des stimuli en fonction du nombre

de syllabes, et conforte l'hypothèse de l'existence d'une unité syllabique dans la représentation mentale du langage pour les nouveaux-nés.

Une troisième étude supporte cette hypothèse. Bertocini *et al.* (1995) vérifient si dans une langue comme le japonais où l'unité rythmique est la more (Labrune, 2001, 2003) la syllabe joue un rôle identique de discrimination des mots. Les listes de stimuli diffèrent en fonction du nombre d'unités syllabiques, ou du nombre d'unités subsyllabiques comme la more. Deux groupes d'enfants âgés de 3 jours sont testés, chacun d'entre eux entendant deux listes de stimuli, le premier avec des mots dissyllabiques *vs.* trisyllabiques, le second avec une alternance entre mots dissyllabiques de 2 ou 3 mores. Les résultats montrent que les mots trisyllabiques sont mieux distingués des dissyllabiques, comme les bimoriques des trimoriques. À notre connaissance, aucune analyse ne compare la reconnaissance respective des deux unités que sont la syllabe et la more.

Une autre série d'études (Mattys et Jusczyk, 2001) porte sur la reconnaissance, par des enfants de 8 mois et demi, de séquences CVC lorsqu'elles se trouvent au sein d'un même mot (exemple « *dice* »), ou à la frontière entre deux mots (exemple : « *cold ice* »). Pour cette série d'expériences, le protocole utilisé n'est pas celui de la succion non-nutritive, mais le paradigme d'orientation contrôlée de la tête du nouveau-né vers ce qu'il reconnaît. Les résultats obtenus montrent que les nouveaux-nés ne traitent pas les deux structures de la même manière et que si la structure CVC est bien reconnue à ce stade de développement, la structure avec voyelle initiale ne sera segmentée correctement qu'à partir de 16 mois : « *The other noticeable finding in the present study is that infants have considerably more difficulty segmenting vowel-initial than consonant-initial words from fluent speech.* » (Mattys et Jusczyk, 2001, p. 19).

L'analyse de ces différentes expériences permet de formuler l'hypothèse selon laquelle la syllabe, et notamment la structure CV, serait une unité particulièrement importante durant la première étape de l'acquisition de la parole, indépendamment de la langue et du rythme syllabique de cette dernière.

L'ensemble de ces résultats tend à prouver que la syllabe est une unité fondamentale pour la perception de la parole, au moins dans la phase d'acquisition du langage, bien que sa nature et sa structure organisationnelle ne soient pas encore complètement élucidées. En effet, des recherches montrent qu'il existe des variations de segmentation pour les sujets quand la syllabation donne des groupements consonantiques. Ainsi, Goslin *et al.* (1999) ont mené une expérience afin de connaître le degré de consistance des locuteurs français dans le placement des frontières syllabiques dans un ensemble composé d'une ou de plusieurs consonnes. Les résultats montrent une variation certaine dans le placement des frontières par des auditeurs français, mais les différences ne sont pas équivalentes selon la position syllabique qu'occupent les consonnes étudiées. Ainsi, Goslin *et al.* (1999) considèrent que leurs résultats confirment l'existence d'un principe maximisant l'attaque (Maximal Onset Principle, notamment développé par Selkirk, 1984, cf. partie I.1.3.1.), c'est-à-dire que l'attaque de la

seconde syllabe est bien mieux détectée que la coda de la première, ce qui rejoint les travaux de Redford et Diehl (1999). Mais la segmentation semble aussi liée au noyau apparaissant dans la syllabe, les voyelles ouvertes attirant les syllabes fermées et vice-versa.

L'ensemble des études sur la perception auditive de la syllabe, notamment à travers la perception de ses frontières, ont permis de poser plusieurs hypothèses quant aux facteurs articulatoires ou acoustiques à prendre en considération pour définir l'unité syllabique, sans pour autant réussir à offrir un cadre global d'explication.

## I.2.2. Les approches motrices

### I.2.2.1. Rythme syllabique et rythme respiratoire

Il serait difficile d'évoquer ce champ de recherche sans mentionner la théorie de Stetson (1951) qui place la syllabe comme unité de base de la parole, unité à part entière constituée d'un assemblage de segments individuels et qui doit pouvoir être définie de façon cohérente quelle que soit la langue étudiée. « *When the language signals of different languages are compared, the discussion must strictly be phonetic. There are no consistent meanings and syllables compared are in effect nonsense syllables treated as physiological processes. The syllables must be independent of the meanings, although meanings gave rise to the syllables in each language in question.* » (Stetson, 1951, p. 6).

Dans son étude, Stetson analyse les mécanismes articulatoires toujours présents lors de la production de parole et, plus spécifiquement, de syllabes. Il remarque que la pression du flux d'air phonatoire varie au cours de la parole : la contraction des muscles intercostaux entraîne des augmentations de la pression qui permettent la production des voyelles. Stetson considère que la pression du flux d'air phonatoire peut être activée ou arrêtée par l'activité de ces muscles respiratoires, ou par la fermeture du conduit vocal nécessaire à la production d'une consonne. Ainsi, une syllabe sera définie, pour Stetson, comme une simple pression d'air et ses limites seront l'augmentation et l'arrêt du flux d'air phonatoire : « *the chest muscles (intercostals) produce a separate pulse of pressure for each syllable* » (Stetson, 1951, p. 3).

L'hypothèse de Stetson a été infirmée par un certain nombre de travaux dont celui de Ladefoged (1967) qui montre qu'il n'y a pas adéquation entre le nombre de syllabes et le nombre de pics d'expiration de l'air phonatoire dans un énoncé : plusieurs syllabes peuvent apparaître dans un seul pic expiratoire. La critique de la théorie de Stetson par Ladefoged se fonde sur une réfutation de ses hypothèses grâce à l'analyse de nouvelles données, mais ne remet pas fondamentalement en cause la position générale de Stetson en ce qui concerne les relations sons/syllabes dans la production du langage. C'est pourtant, selon Redford (1999), le principal problème posé par les travaux de ce dernier. En effet, plus que la non-adéquation entre les rythmes respiratoire et syllabique, la définition de la syllabe comme unité totalement indépendante des sons semble improbable. Pour Stetson, « *In the individuality of the syllable the sound is secondary; syllables are possible without sound.* » (Stetson, 1951, p. 33). Pourtant, si la fonction principale du langage est bien la communication, l'utilisation du canal

oral est prépondérante, et il est difficile d'envisager, dans ce contexte, que les sons ne soient qu'accessoires.

Les résultats de Stetson sont aussi remis en cause au niveau de la séparation des aspects acoustiques de la parole et des syllabes, ce qui affaiblit les relations possibles entre la syllabe et la pression du flux phonatoire. Malgré ces critiques, l'approche de Stetson apporte une vision nouvelle de la syllabe. Ainsi, pour Redford (1999), « *The real virtue of Stetson's theory was that the syllable movement was assumed to be slower than the movements that define segmental articulation. In this way, Stetson was able to derive a suprasegmental unit of sound organization from the hierarchical organization of the speech system. The problem, however, was that the movement chosen was so far removed from the local acoustic structure of speech that Stetson was unable to define syllable boundaries or internal structural attributes of syllables, which, as the phonological theories emphasize, are the main obstacles to defining syllable.* » (Redford, 1999, p. 20). Et pour répondre à ce problème, Redford s'inscrit dans le prolongement de MacNeilage (1998) à propos de la théorie « *Frame, then Content* », en confirmant, à travers plusieurs expériences, l'hypothèse de l'oscillation mandibulaire comme mouvement de base pour expliquer l'organisation syllabique des langues.

#### **1.2.2.2. L'oscillation mandibulaire et la théorie « *Frame, then Content* »**

La théorie « *Frame, then Content* » (Davis et MacNeilage, 1995 ; MacNeilage, 1998) propose que l'universelle structuration syllabique des langues du monde repose entièrement sur les mouvements alternés et cycliques d'élévation et d'abaissement de la mandibule « *the frame* » (le cadre). Cette modulation accordée à l'ouverture et à la fermeture de la cavité buccale correspond à la dimension articulatoire la plus exploitée par les langues pour la réalisation des unités sonores distinctives : consonnes et voyelles « *The content* » (le contenu) constituants les phases des cycles mandibulaires. Cette théorie attire l'attention sur le rôle prépondérant de la mandibule dans le babillage mais aussi dans la structuration syllabique du lexique des langues : « *syllable-structure frames are represented in some way during the production process independent of segmental content elements* ». (MacNeilage, 1998, p. 499).

La théorie « *Frame, then Content* » s'inscrit dans une perspective phylogénétique qui cherche à comprendre comment la capacité de parole a évolué chez l'espèce humaine. En effet, « *d'autres espèces animales partagent avec l'homme le fait d'utiliser à des fins communicatives des séquences d'ouverture et de fermeture du conduit vocal, mais nous seuls ajoutons à ce mouvement une modulation articulatoire : la spécificité résiderait donc dans cette seconde articulation du mouvement, ce que MacNeilage appelle le contenu.* » (Vilain, 2000, p. 73). Ainsi, selon MacNeilage (1998), chez les primates non-humains, on n'observe presque jamais de communications par cyclicités mandibulaires accompagnées de phonations. En revanche, on trouve de nombreux exemples de communications visuo-faciales basées sur la mandibule. Ces exemples sont observés dans de très nombreuses situations sociales et impliquent tous une communication. Il s'agit donc de comportements très similaires à la parole humaine, que MacNeilage considère comme des pistes d'une évolution phylogénétique



de la parole, des sortes d'intermédiaires entre ingestion et parole. En effet, pour MacNeilage (1998), cette alternance cyclique de mouvements d'ouverture et de fermeture de la cavité buccale dans la parole proviendrait de l'évolution du processus de mastication des aliments.

Dans un cadre hypothétique de récapitulation (Haeckel, 1879), la théorie « *Frame, then Content* » inscrit également le développement du contrôle de la parole chez le bébé humain selon les mêmes étapes que le développement de ce contrôle chez l'espèce humaine. Les travaux sur lesquels s'appuie la théorie mettent en évidence les grandes étapes de l'acquisition du langage, et les aptitudes liées à la mâchoire qui s'y rapportent. À cinq mois, les bébés sont capables de jeux de mâchoires, même si aucune vocalisation n'y est encore rattachée. Le babillage canonique, qui apparaît vers 7 mois, est produit sous forme de suites répétitives de proto-syllabes CVCV, et donc d'ouvertures/fermetures de la cavité buccale accompagnées de phonation (Davis et MacNeilage, 1995). C'est l'apparition du cadre porteur de la parole (*the Frame*), qui est considéré comme dominant jusqu'à l'apparition du contenu segmental (*the Content*), au stade des premiers mots vers l'âge de 18 mois. L'observation des combinaisons CV présentes dans le babillage (Davis et MacNeilage, 1995, MacNeilage et Davis, 2000) montre que les segments consonantiques et vocaliques produits ont généralement le même lieu d'articulation. Cette observation va dans le sens de la prédominance pure du cadre, les autres articulateurs étant alors passifs et uniquement portés par la mandibule. La configuration du conduit vocal utilisée pour la production de la closure est maintenue pour la production du segment vocalique : la position de la langue lors de la phase d'ouverture reste la même dans la phase de fermeture, ce qu'ils appellent le « *preseting* » de la langue. Les combinaisons favorisées observées dans le babillage correspondent aux productions liées à ce prépositionnement (« *presetting* ») de la langue (cf. figure I.10). L'apparition du contrôle des autres articulateurs permettra de mettre en place le contenu (pour une évolution des degrés de liberté articulaire, cf. Vilain, 2001).

Le « *preseting* » de la langue rejoint donc un autre concept repris par la théorie « *Frame, then Content* » : la recherche d'un coût minimal de production, qui vise à limiter la mise en mouvement des autres articulateurs du conduit vocal pour privilégier le mouvement d'élévation et d'abaissement de la mâchoire qui posséderait un coût minimal (principe d'économie articulaire).

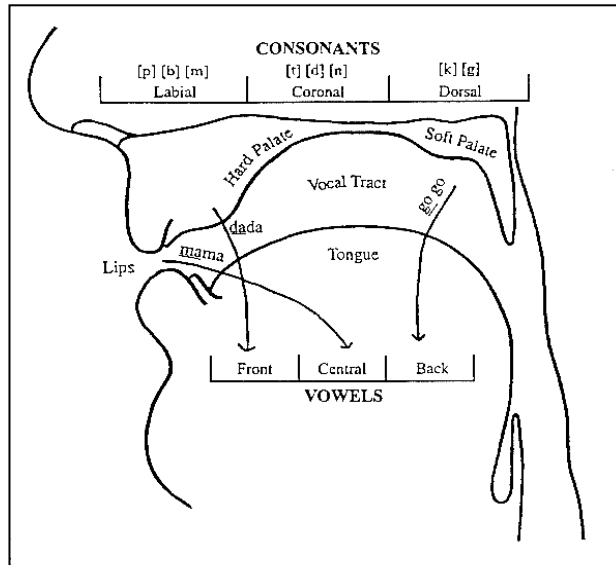


Figure I.10 : Le *preseting* de la langue dans les productions du babillage qui expliquerait les combinaisons CV favorisées, MacNeilage et Davis (2000).

MacNeilage et Davis (2000) ont postulé la théorie « *Frame, then Content* » à partir de données provenant essentiellement du babillage et des premiers mots d'enfants, mais ils ont ensuite cherché à la vérifier sur des données de langues ou de la phylogénèse. Ainsi, les combinaisons favorisées dans les productions des enfants sont retrouvées dans les études des lexiques de 10 langues naturelles, ainsi que dans les 27 racines proposées par Ruhlen (1997) comme reconstructions provenant d'une langue mère universelle.

Ces résultats sont aussi confirmés par des études portant sur les erreurs de production réalisées par les adultes (lapses, contrepèteries, ...). Les interactions ont le plus souvent lieu entre segment occupant la même position dans la structure syllabique (MacNeilage, 1998). La plupart des lapsus tiennent donc compte de la position syllabique des éléments intervertis, comme le montre l'étude de Beaulieu (2001) à propos de 417 lapsus français. Les contraintes de positions syllabiques sont respectées dans 98,5 % des lapsus ayant lieu entre deux mots, et dans 95 % des cas, lorsqu'il a lieu à l'intérieur d'un même mot, comme dans l'exemple suivant :

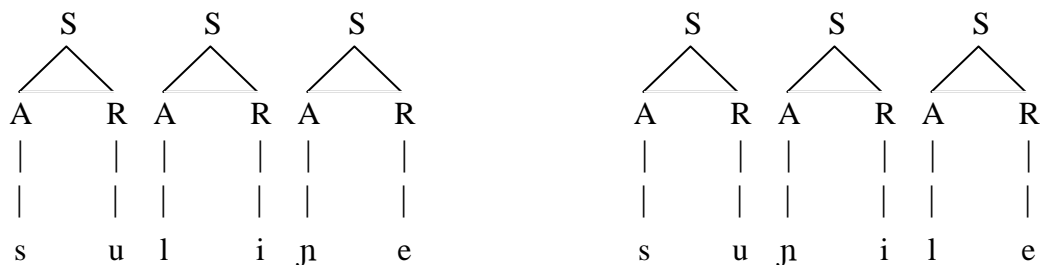


Figure I.11 : Schémas représentant le mot cible « *suli je* » et le lapsus « *su jile* », avec inversion des attaques des deux dernières syllabes de l'unité lexicale (Beaulieu, 2001).

MacNeilage (1998) appuie sa théorie sur l'hypothèse d'une évolution progressive de deux systèmes corticaux primates qui ont permis le développement des contrôles moteurs nécessaires pour produire les cadres et les contenus :

- (1) un système intrinsèque médian lié à la motivation, incluant le cortex « *cingulaire* » antérieur et l'aire motrice supplémentaire, pour le comportement « *self-generated* », jadis responsable du contrôle de la vocalisation ancestrale et maintenant responsable des cadres ;
- (2) un système extrinsèque latéral, incluant l'aire de Broca et ses environs, et l'aire de Wernicke, spécialisée dans la réponse aux stimuli externes (et dans la capacité émergente d'apprentissage vocal) et qui est en plus responsable du contenu.

Le développement de la théorie « *Frame, then Content* » se réalise autour de l'importance des phases d'ouverture et de fermeture de la cavité buccale dans le fonctionnement du langage humain. Un certain nombre de travaux en phonétique portent sur la description des mouvements de la mandibule. Ils prépareront et conforteront la réflexion sur les relations existant entre la syllabe et le cycle mandibulaire. On peut citer dans ce cadre les études de Redford qui conclut : « *the jaw cycle provides a mechanical and temporal constraint on segmental articulation. It was argued that this constraint accounts for the patterns of segment sequencing and relative segment duration normally associated with syllables.* » (Redford, 1999, p. 123).

Dans les études antérieures sur l'importance de la mandibule dans la production du langage, on peut citer Lindblom (1983) et Keating (1983) qui ont analysé les influences du mouvement mandibulaire dans la coarticulation au sein d'énoncés respectivement suédois et anglais, mais leurs résultats sont contradictoires. Pour Lindblom, la hauteur de mâchoire nécessaire à la production de voyelles influe la hauteur d'ouverture des consonnes qui l'entourent. En revanche, pour Keating, la hauteur de mâchoire nécessaire à la production des consonnes influence celle des voyelles.

En 1994, Keating, Lindblom, Lubker et Kreiman décident de travailler ensemble pour tenter de résoudre cette contradiction. Deux hypothèses sont alors avancées : soit le contexte segmental a un effet important sur la hauteur de mâchoire des segments adjacents (influence relative) ; soit certains segments varient contextuellement plus que d'autres (variabilité relative). Une autre possibilité est envisagée : Lindblom et Keating ont travaillé sur deux langues différentes, qui ont peut-être chacune une organisation différente. Afin de répondre à ces questions, des données de hauteur de mâchoire sont collectées pour différentes combinaisons de consonnes et de voyelles, et les segments sont comparés dans tous les contextes. Un premier résultat indique que l'anglais et le suédois ont un comportement semblable : le rang d'apparition des consonnes par hauteur d'ouverture est quasiment identique dans les deux langues, avec des différences mineures en ce qui concerne les voyelles.

Les résultats obtenus sont plus proches de ceux de Lindblom (1983) que de ceux de Keating (1983). Le modèle qui émerge clairement est que l'effet des voyelles sur les consonnes est plus grand que l'effet des consonnes sur les voyelles. En effet, bien que ce ne soit pas le cas pour tous les sujets de l'étude, l'effet des voyelles sur les consonnes est hautement significatif tandis que l'effet des consonnes sur les voyelles ne l'est pas. Ils constatent également que les consonnes peuvent accommoder leur hauteur de mâchoire en fonction de leur voisin vocalique. Ces résultats renforcent aussi fortement le propos de Lindblom (1983) qui montre que les consonnes diffèrent dans leur propension de coarticulation. Une distinction existant entre la variabilité globale (Keating, 1983) et la variabilité attribuable au contexte (Lindblom, 1983) est aussi mise en évidence.

Ces résultats, portant sur les liens existant entre consonnes et voyelles et insistant sur l'adaptabilité des premières en fonction de l'abaissement de la mâchoire nécessaire à la production des secondes, vont dans le sens d'une organisation des segments autour d'un noyau vocalique. En effet, les études menées par Redford (1999) montrent que les segments adaptent leur degré d'ouverture du conduit vocal en fonction de la position qu'ils occupent dans la syllabe et donc ainsi par rapport à leur contexte segmental.

Ainsi, l'hypothèse formulée par Redford (1999) selon laquelle si l'articulateur, dont les mouvements sont responsables de la structuration syllabique est la mâchoire, les mouvements de montée et d'abaissement de la mandibule doivent contraindre l'articulation des segments de façon à produire des patrons phonétiques et phonologiques qui forment la base de la perception de la syllabe.

En travaillant à partir d'un corpus de 42 monosyllabes issues du russe, Redford (1999) mesure, pour l'ensemble des stimuli, la durée de la syllabe et de ses constituants, la hauteur de la mâchoire pour chaque segment et pour chaque cycle. « *The results [...] provided evidence for the view that a single syllable corresponded to a single cycle jaw, defined by two consecutive points of maximal jaw closure* » (Redford, 1999, p. 116). Cette étude, en effet, montre qu'il existe bien une corrélation entre la hiérarchie de sonorité et la cyclicité du mouvement mandibulaire : la nature du segment a une influence sur la position de la mâchoire, et en même temps, le cycle mandibulaire produit une contrainte mécanique forte sur l'articulation des segments : [l] dans [lba] ou [bla] est articulé avec une position plus haute dans la première que dans la seconde.

L'ensemble des recherches concernant la base articulo-mandibulaire de la syllabe converge vers le mouvement mandibulaire. La théorie « *Frame, then Content* » nous a semblé être un cadre intéressant pour situer nos analyses de la structure syllabique. En effet, l'observation des différentes structures syllabiques présentes dans les langues du monde, et notamment dans notre corpus, nous permet de mettre en évidence des combinaisons favorisées qui rejoignent, renforcent et prolongent celles obtenues par MacNeilage et Davis (2000).

Les différentes études, qu'elles soient inscrites dans un cadre purement perceptif ou acoustico-perceptif, ou dans un cadre plutôt moteur, ne permettent pas de conclure de façon

unanime sur le rôle de la syllabe dans le fonctionnement du langage. La piste de la théorie « *Frame, then Content* » semble prometteuse pour comprendre l'origine de l'universelle structuration syllabique des lexiques et des énoncés, ainsi que l'organisation des structures syllabiques, tout comme les études articulatoires et perceptives ouvrent des perspectives sur le rôle de la syllabe comme unité linguistique dans le fonctionnement du langage.

Il ne fait cependant plus aucun doute que c'est dans un cadre théorique liant les aspects articulatoires et perceptifs de la parole, ainsi que ces aspects aux descriptions phonologiques que la nature et le fonctionnement de la syllabe pourront être élucidés et expliqués. Une théorie telle celle de « *La perception pour le contrôle de l'action* » de Schwartz *et al.* (2002) pourrait inscrire la syllabe dans son propos. Il s'agit d'une théorie qui tente de lier les aspects productif et perceptif de la parole en considérant la perception de la parole comme l'ensemble des processus perceptifs (auditifs, visuels) fournissant un ensemble de représentations permettant le contrôle de ses propres actions et le suivi des actions de l'autre.

Nous constatons donc que la recherche d'explications pour les tendances générales concernant les systèmes sonores des langues du monde dans les propriétés physiologiques de la parole apporte des avancées importantes pour la compréhension du fonctionnement du langage humain. Il semble, par conséquent, intéressant d'utiliser la même approche pour tenter de mieux comprendre l'unité syllabique. L'étude des faits de langues à partir de l'établissement de typologies devrait permettre de mettre en évidence des tendances générales sur l'organisation syllabique des langues et ainsi permettre de poser des hypothèses quant aux mécanismes de production ou de perception qui y jouent un rôle. Le principal enjeu de cette approche « *substance-based* » est, à terme, la mise en évidence des relations existant entre le système linguistique et des composantes structurelles basées sur les contraintes articulatoires, acoustiques et perceptives des sons de parole (*cf.* par exemple, Lindblom *et al.*, 1984 ; Schwartz *et al.*, 1997 ; Vallée *et al.*, 1999 ; Lindblom, 2000 ; De Boer, 2000).

## **CHAPITRE II**

**ULSID,**

**BASE DE DONNÉES LEXICALES ET SYLLABIQUE**

L'ensemble des travaux en typologie linguistique, associés à la recherche d'universaux, a montré que la constitution de bases de données est une étape indispensable et fondamentale dans la compréhension du rôle des unités linguistiques dans le fonctionnement du langage humain (Greenberg, 1966 ; Chomsky & Halle, 1968 ; Maddieson, 1984, 1991). Après avoir posé le problème de la constitution d'un échantillon linguistique représentatif, nous présentons, dans cette partie, l'historique des données collectées à UCLA (Université Californie Los Angeles), à partir desquelles nous avons mené notre recherche, et précisons leur contenu. Nous spécifions également les choix réalisés lors de la constitution d'une interface informatique permettant le traitement automatique des données lexicales. Cette interface a été programmée de manière à constituer une banque de données extensible : dans le cadre de ce travail, deux autres langues ont été adjointes aux données de UCLA.

## II.1. DE LA REPRESENTATIVITE DE L'ECHANTILLON EN TYPOLOGIE LINGUISTIQUE

Les travaux typologiques en lien avec la recherche d'universaux du langage initiée par Greenberg dans les années 60 ont montré que l'étude d'un large échantillon de langues est nécessaire pour pouvoir envisager une généralisation des résultats. En effet, comme dans toute étude fondée sur l'expérience, cet objectif nécessite d'étudier un échantillon suffisamment représentatif des individus. De même, l'étude d'un grand nombre de langues est incontournable lorsque la recherche consiste à délimiter les éventails de variations possibles et les limites de ces dernières. Chomsky & Halle eux-mêmes, qui dans leur ouvrage « *Aspect of the theory of syntax* » (1965) se contentent d'une étude approfondie de l'anglais pour généraliser sa théorie, se sert de plus d'une centaine d'autres langues lorsqu'il étudie la phonologie (*The Sound Pattern of English*, 1968).

Plus l'échantillon est vaste, plus les résultats qu'il permet d'obtenir sont généralisables et apportent des informations sur l'organisation générale du langage humain. L'échantillon le plus vaste qui soit devrait contenir toutes les langues du monde passées, présentes et futures. Cette hypothèse est irréaliste pour plusieurs raisons : 1) certaines langues se sont éteintes sans avoir été répertoriées, ou avec des informations incomplètes ; 2) de nouvelles langues ou variétés de langue vont apparaître et offrir d'autres possibilités de variations qu'il sera peut être impossible de prévoir ; 3) trop de langues sont actuellement parlées dans le monde pour que l'on puisse toutes les étudier individuellement (plus de 5 000 langues selon les estimations de Hagège (2000)).

Le problème réside donc dans la définition ou l'estimation de la représentativité d'un échantillon des langues parlées à la surface du globe. Souvent l'inventaire est limité à des langues parlées aujourd'hui, augmenté d'un petit nombre de langues mortes pour lesquelles il existe suffisamment de données écrites. Il s'appuie sur deux hypothèses : les langues représentent toutes le même niveau d'évolution ; le nombre de langues parlées aujourd'hui est suffisamment important pour inclure toutes les variations possibles du langage humain. Se pose alors la question de la constitution d'un échantillon représentatif : 1) les langues doivent

provenir d'un large inventaire de familles linguistiques, puisque, par définition, les membres d'une même famille sont les héritiers d'un ancêtre commun et il ne serait donc pas possible de trancher aisément entre des traits reçus de la parenté commune et ceux qui relèvent d'universaux du langage ; 2) il convient aussi d'écartier les biais géographiques, car les langues en contact s'empruntent mutuellement des caractères linguistiques et la communauté de traits spécifiques pourrait ne pas être due à un universal. Pour obtenir un échantillon réellement représentatif de la diversité des langues du monde, Croft (1990) propose : « *The possibility of genetic and areal sources for cross-linguistic similarities requires a wide distribution for the typological sample. The essential idea is to take a classification of languages by genetic affiliation and geographical location, and to use that to develop an unbiased sample. This has turned out to be more controversial topic than the size of sample issue. First, it should be pointed out that the areal-genetic distribution problem is different for small samples than large ones. In small samples, the problem is the possibility of excluding from consideration large linguistic families or large areas which may be of potential importance for the phenomenon in question. In the large sample, the problem is how to avoid having subsets of the sample "clustered" together areally and/or genetically* » (Croft, 1990, p. 22).

Les données de langues sur lesquelles s'appuient notre recherche (et que nous décrivons très précisément ci-après) peuvent paraître limitées (une quinzaine de langues) au regard de la taille des inventaires phonologiques, tels ceux d'UPSID (Maddieson, 1984 : 317 langues ; Maddieson et Precoda, 1989 : 451 langues) qui ont permis de sérieuses avancées dans la compréhension de l'organisation des systèmes sonores, à travers la mise en évidence de nombreuses tendances générales sur l'organisation des systèmes vocaliques et consonantiques des langues du monde (Maddieson, 1986 ; Laver, 1980 ; Vallée, 1994 ; Vallée *et al.*, 1999 ; Vallée *et al.*, 2001) et le développement de modèle de prédiction de ces tendances comme celui de Schwartz *et al.* (1997) ou encore celui de Boer (2000).

## II.2. UCLA LEXICAL AND SYLLABIC INVENTORY DATABASE

Ainsi baptisée par notre équipe, ULSID a été élaborée dans le but de servir à la recherche d'universaux sur l'organisation syllabique des langues du monde. Riche au départ de 32 langues, elle respecte un certain nombre de critères de représentativité définis par Maddieson (1986, 1991) pour UPSID. Les deux premiers portent sur la recherche d'une diversité aux niveaux :

- 1- génétique : les langues incluses dans la base doivent être génétiquement indépendantes sur la base du principe que deux langues de souche commune doivent être séparées de leur ancêtre commun d'au moins 1500 ans ;
- 2- géographique : les langues doivent offrir un éventail le plus large possible de provenances géographiques.

Par contre, pour Maddieson, « *the size of extant populations of speakers of languages is an accident of political and social history that is quite irrelevant to the questions relating to*



*the structure of human languages* » (Maddieson, 1986, p. 158). Le nombre de locuteurs ne constitue donc pas un critère de figuration dans une banque de données pour la recherche d'universaux des systèmes sonores. Au contraire, les appartenances génétique et aérale sont les critères de sélection des langues. Ainsi, la base ULSID est constituée de langues issues des cinq continents et de plusieurs familles linguistiques différentes.

Les trente-deux langues sélectionnées par Maddieson et Precoda (1992) sont les suivantes (les seize en caractères gras sont celles implémentée à l'ICP dans le cadre de ce travail) :

- Asie* : chinois standard, tibétain, coréen, **thaï**, jeh, **wa**, **nyah kur**, **sora**, **kannada**, darai
- Amériques* : comanche, **totonaque**, **quechua**, **navaho**, **kwakw'ala**, **yup'ik**, shipibo, pirahã
- Afrique* : **!xóõ**, **ngizim**, maninka, **gbaya**, **kanouri**, igbo, **afar**
- Europe* : turc, polonais, **finnois**
- Pacifique* : fasu, hawaïen, kadazan, rotokas

Les données d'ULSID se présentent sous la forme de lexiques ou de dictionnaires contenant l'information sur le découpage syllabique des unités. « *Generally the sources are printed or computer-readable dictionaries or which provide a phonetic or phonemic transcription of entries, or employ an orthography that is interpretable in such terms* » (Maddieson, 1993, p. 1). Le choix des entrées lexicales, présentes ou non dans les fichiers d'ULSID, se fait selon les critères suivants : « *Recent loan words, especially those of wide international currency relating to modern technological, political or cultural concepts ('telephone', 'democracy', 'football'), have been excluded wherever recognizable* » (Maddieson, 1993, p. 1). Le découpage des lexiques en syllabes a pour sa part été réalisé soit à partir de lexiques syllabés édités, soit à partir d'informateurs natifs de cette langue, sachant que « *l'"ignorance" des sujets parlants en matière de phonétique constitue un meilleur guide de la réalité phonologique que la "science" du spécialiste* » (Sapir, 1953, p. 181).

Sur les 16 langues de l'échantillon de Maddieson et Precoda (1992), 14 ont été implémentées à l'ICP après un travail conséquent d'harmonisation du formalisme des lexiques, qui n'a pu être achevé à temps pour deux d'entre elles : le totonaque (langue amérindienne) et le gbaya (langue nigéro-congolaise). Nous en invoquons les raisons page 56. L'étude des structures lexicales et syllabiques que nous proposons dans le cadre de cette thèse porte donc sur les langues suivantes : afar, finnois, kannada, kanouri, kwakw'ala, navaho, ngizim, nyah kur, quechua, sora, thaï, wa, yup'ik, !xóõ, que nous avons complétées par le suédois (Berlitz, 1981) et le français (base BDLex-Syll, Pérennou et Calmès, 2002).

Notre corpus est ainsi constitué d'un ensemble de 16 langues regroupant 12 familles, 16 sous-familles, réparties sur 4 continents (cf. annexe I). L'échantillon retenu remplit donc bien les conditions de diversité génétique et géographique. Le respect de ces deux critères est nécessaire à la représentativité de la base de données, mais ce ne sont pas les seuls existant.

En effet, au regard des systèmes vocaliques et consonantiques des différentes langues de notre échantillon, la diversité est aussi présente au niveau des systèmes phonologiques. Nous livrons ci-après l'ensemble de ces informations, de manière détaillée, pour chacune des langues.

## II.3. LES LEXIQUES D'ULSID

### II.3.1. Aspect formel

L'échantillon constitué à UCLA par Maddieson et Precoda (1992) comporte les lexiques de langues sous forme de fichier texte en code ASCII avec les conventions suivantes pour le découpage syllabique :

- un item lexical par ligne ;
- les syllabes sont séparées par des points ;
- les constituants de la syllabe (attaque, noyau, coda) sont séparés par des blancs graphiques (espace).

Illustration ci-dessous avec les dix premières entrées du lexique du navaho :

? a  
 ? a .ch i t  
 ? a .ch inin h  
 ? a .ch inin sh .t a h  
 ? a .ch o  
 ? a .ch o ?  
 ? a ch .x o sh .tlh' oo l  
 ? a .ch' anan h  
 ? a .ch' a .h a .y a h  
 ? a .ch' e .? e .d anan ?

Dans cet extrait, nous pouvons identifier une unité lexicale monosyllabique, cinq dissyllabiques, quatre trisyllabiques et une quadrisyllabique. Le deuxième item lexical « ? a .ch i » s'analyse phonologiquement de la manière suivante (avec  $\sigma$  = syllabe, A = attaque, R = rime, N = noyau) :

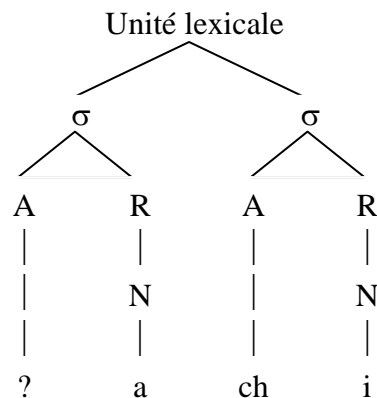


Figure II.1 : Représentation syllabique de l'unité lexicale « ? a . ch i ».

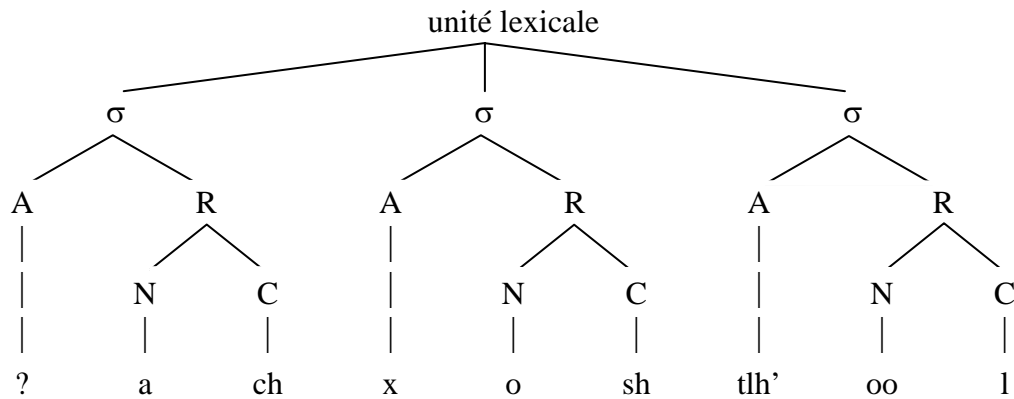


Figure II.2 : Représentation syllabique de l'unité lexicale « ? a ch . x o sh . tlh' oo l ».

L'implantation des données à l'ICP s'est trouvée confrontée à trois types de problèmes majeurs pour pouvoir exploiter les données.

- 1- Le codage ASCII de chaque lexique a été plus ou moins adapté aux particularités phonologiques des langues et même si les lexiques se présentent sous le même formalisme, il existe des différences dans le codage des constituants de la syllabe.
- 2- Ensuite, le codage hétérogène de segment identique mais dans des langues différentes a nécessité l'harmonisation des données pour envisager l'automatisation de l'interrogation de la base. Un nouveau codage des données, que nous avons appelé la phase de transcodage, a été réalisé en utilisant les caractères de l'Alphabet Phonétique International (API) révisé de 1996 (Handbook of the IPA, 1999).
- 3- Enfin, il a fallu repérer et traiter les erreurs de saisie, essentiellement des omissions ou rajouts d'espaces. Par exemple, dans certains cas, les voyelles placées en début de mot étaient précédées d'un espace. Cette frappe n'étant pas toujours observée, la différence apparemment minime entre les items précédés ou non d'un caractère blanc est en réalité véritablement problématique si l'on veut pouvoir effectuer un traitement automatique des données. Il a, par conséquent, été nécessaire de procéder à une phase appelée de « *nettoyage des lexiques* », après avoir identifié l'ensemble des irrégularités dans le format des données, et ce sans perdre d'information sur le découpage syllabique. En effet, un espace peut être un bruit, mais rappelons que c'est aussi le séparateur des constituants syllabiques.

Pour le transcodage des données, la première étape a consisté en une recherche des caractères utilisés pour le codage de chaque voyelle et chaque consonne de chaque lexique, donc à retrouver une logique de codage et ce en retournant aux sources phonétiques et phonologiques des langues. Il a donc fallu définir exactement la variété de langue dont nous disposions et par là même son système phonologique et ses particularités. Une des difficultés rencontrées réside dans le fait que, selon les descriptions et les auteurs, une même langue peut avoir plusieurs appellations. De plus, les informations phonologiques sont parfois délivrées

avec une transcription propre à l'auteur et souvent peu ou pas décrite. C'est à travers ces différentes sources (cf. annexe II) associées à la recherche de la logique de codage que les systèmes sonores des 14 langues de l'échantillon de Maddieson et Precoda (1992) ont pu être identifiés. Les deux exemples ci-dessous illustrent bien la problématique rencontrée :

- la séquence /a:/ est codée de différentes manières selon les langues. Dans la majorité des lexiques, comme en kannada, elle est transcrite avec la séquence « aa », ainsi les premières syllabes dans « k a. j e » et « k aa. k e » se distinguent bien. Par contre, en nyah kur, le même segment est codé par la séquence « a a », avec un blanc graphique entre les deux symboles. Ainsi, les items lexicaux « b a. b i » et « b a a. c i » sont tous deux composés de deux syllabes de deux constituants chacune, alors qu'une autre entrée « b a. a. c i » est composée de trois syllabes de respectivement deux, un et deux segments. L'espace qui marque normalement la séparation entre deux items différents bruite le codage lorsqu'il est utilisé entre les composants d'une même unité sonore. Ce problème a aussi été rencontré dans le cas des consonnes affriquées.
- différentes voyelles d'avant non-arrondies d'aperture moyenne sont codées par des symboles identiques : en effet, si l'on observe le tableau ci-dessous, on comprend que la simple étude des systèmes phonologiques n'aurait pas permis de retrouver, avec certitude, toutes les correspondances et qu'il nous a fallu, dans certains cas, identifier directement la transcription d'unités lexicales afin de choisir entre les différentes possibilités.

Codage ASCII fourni	Codage API proposé après investigation			
	ngizim	nyah kur	thaï	yup'ik
E	ə	ɛ	æ	
e	e	e	e	ə
@		ə	ə	

Table II.1 : Tableau de transcodage des caractères ascii « E », « e » et « @ » pour quatre langues.

Deux langues sont toujours en cours de transcodage et n'ont donc pas pu être à temps intégrées à notre étude :

- Pour le totonaque, il n'a pas encore été possible d'établir la correspondance complète entre le codage ASCII et les divers systèmes phonologiques trouvés selon les variétés.
- Pour la seizième langue fournie par Ian Maddieson, le gbaya, l'étude de l'ouvrage *Phonologie du Gbaya Kara'Bodoe de Ndongue Bongowen : (région de Bouar, république Centrafricaine)* de Monino et Roulon (1972), ainsi qu'un entretien particulier avec Mme Roulon-Doko, nous a permis d'établir que notre fichier n'était pas exploitable en l'état. En effet, le dictionnaire dont il est tiré a considéré comme entrées potentielles les unités lexicales de la langue, mais aussi dans certains cas des parties d'énoncés voire des phrases entières, ce qui le rend inexploitable pour notre recherche.

Nous avons donc opté pour l'intégration du français et du suédois dans les données, langues pour lesquelles nous avons pu disposer de dictionnaires syllabés (cf. références page 53). À l'heure actuelle, notre corpus comporte 16 langues. Après avoir présenté plus en détail les langues d'ULSID et leurs systèmes phonologiques, nous précisons la masse de données que représente notre échantillon (cf. partie II.4.). Nous présenterons à la suite les différentes interfaces mises en place pour leur analyse automatique (cf. partie II.5.).

### II.3.2. Présentation des lexiques d'ULSID

Nous avons choisi de présenter, par ordre alphabétique des langues, les systèmes phonologiques que nous avons obtenus au terme des étapes de transcodage pour les différentes langues de notre échantillon sous la forme suivante (sources bibliographiques en annexe II) :

- les consonnes sont placées selon le tableau conventionnel de l'API en fonction de leur lieu et mode d'articulation ;
- les voyelles sont disposées sur une « grille iconique » symbolisant un triangle vocalique (cf. figure 1) représentant les différentes qualités vocaliques en fonction « du point le plus haut de la langue dans la cavité buccale, du degré de cette ouverture dite interne, et de la position des lèvres (étirées ou protruses) » (Vallée, 1994, p. 49).

Pour chacune des dimensions articulatoires secondaires recrutées dans une langue (nasalité, longueur, ...), une grille supplémentaire précisera les timbres attestés.

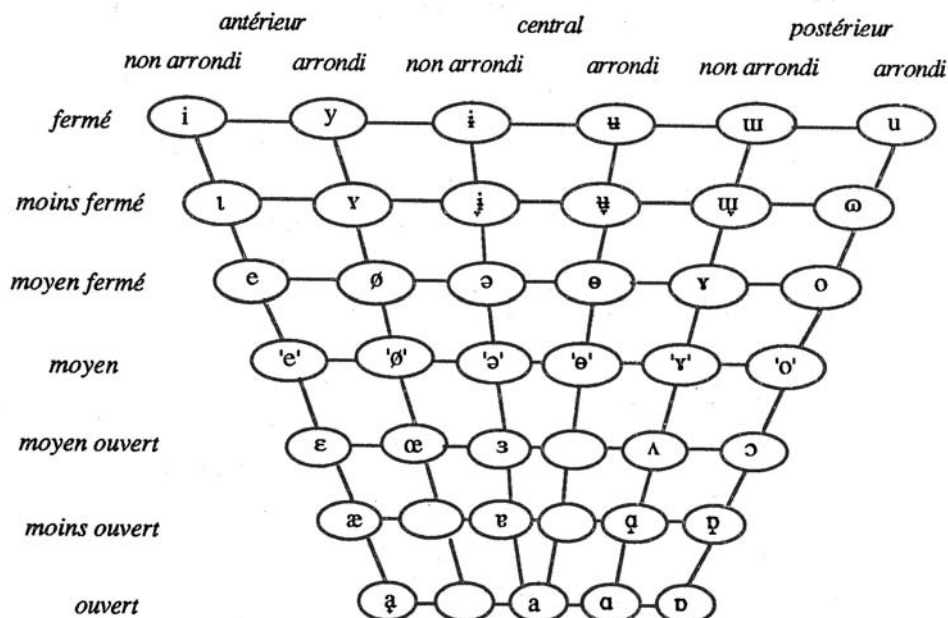


Figure II.3 : « Grille iconique » symbolisant le triangle vocalique utilisé pour la description des systèmes vocaliques d'ULSID (Vallée, 1994).

**Langue :** Afar

**Famille et sous-famille :** Afro-asiatique, couchitique.

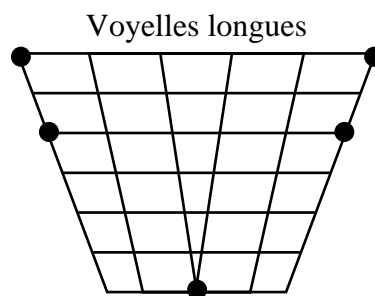
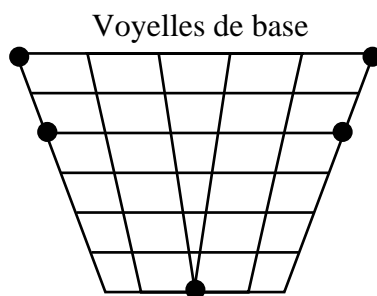
**Situation géographique :** Corne orientale de l’Afrique (Nord de l’Ethiopie, Erythrée, Djibouti).

**Nombre de locuteurs :** 1 500 000, dont plus de 900 000 en Ethiopie, 300 000 à Djibouti et 300 000 en Erythrée.

**Système consonantique :** 17 consonnes.

AFAR		Bilabiale	Labio-dentale	Coronale	Rétroflexe	Vélaire	Pharyngale	Glottale	Labio-vélaire
Plosive	Non-voisée	p		t		k			
	Voisée	b		d	ɖ	g			
Nasale		m		n					
Vibrante				r					
Fricative	Non-voisée		f	s			ħ	h	
	Voisée						ʕ		
Approximante Latéral-approximante									w
				l					

**Système vocalique :** 10 voyelles



**Langue :** Finnois

**Famille et sous-famille :** Ouralo-altaïque, finno-ougrienne.

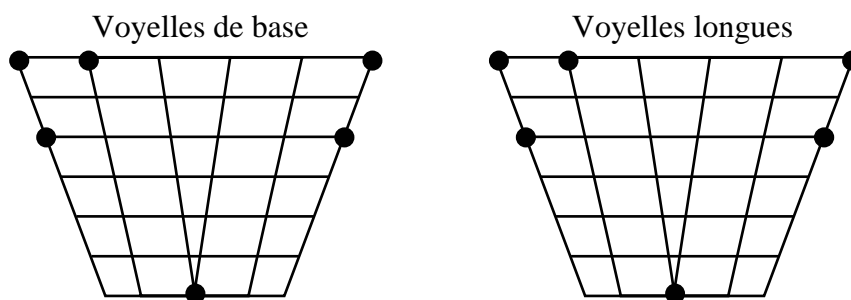
**Situation géographique :** Finlande, Suède (dans le comté de Norrbotten), Russie (Carélie), Estonie, Norvège (dans la région de Finnmark), ainsi que par de petites communautés aux Etats-Unis et au Canada.

**Nombre de locuteurs :** 6 000 000 dont 4 700 000 en Finlande, 300 000 en Suède.

**Système consonantique :** 17 consonnes.

FINNOIS		Bilabiale	Labio-dentale	Coronale	Palatale	Vélaire	Glottale
Plosive	Non-voisée	p		t, t'		k	
	Voisée	b		d		g	
Nasale		m		n			
Vibrante				r			
Fricative	Non-voisée		f	s			h
	Voisée		v				
Affriquée	Non-voisée			ts			
Approximante Latéral-approximante					j		
				l			

**Système vocalique :** 12 voyelles + 20 diphtongues



Diphtongues:

iy, ie, iu, yi, ya, yo, ei, ey, eu, ai, ay, ao, au, oi, oy, ou, ui, uy, ue, uo

**Langue :** Français (variété de France)

**Famille et sous-famille :** Indo-européenne, romane.

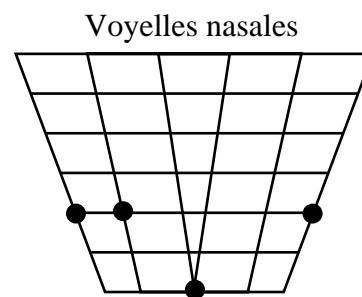
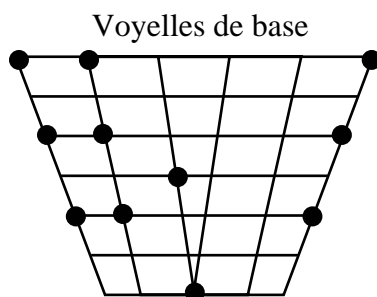
**Situation géographique :** France, Amérique du Nord, Afrique, Océanie, Asie du Sud-Est.

**Nombre de locuteurs :** environ 80 000 000 dont 60 000 000 en France.

**Système consonantique :** 21 consonnes.

FRANÇAIS		Bilabiale	Labio-dentale	Alvéolaire — Coronale Post-alvéolaire		Palatale	Vélaire	Uvulaire	Labio-palatale	Labio-vélaire
Plosive	Non-voisée	p		t			k			
	Voisée	b		d			g			
Nasale		m		n		ɲ	ŋ			
Fricative	Non-voisée		f	s	ʃ					
	Voisée		v	z	ʒ			ʁ		
Approximante Latéral- approximante						j			ɥ	w
				l						

**Système vocalique :** 15 voyelles





**Langue :** Kannada

**Famille et sous-famille :** Dravidienne, sud-dravidienne.

**Situation géographique :** Inde, dans les Etats suivants : Karnataka, Andhra Pradesh, Tamil Nadu, Maharashtra.

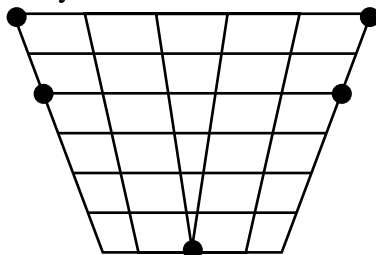
**Nombre de locuteurs :** 35 000 000.

**Système consonantique :** 27 consonnes.

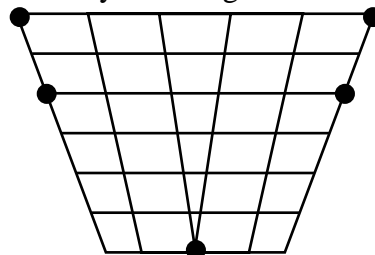
KANNADA		Bilabiale	Labio-dentale	Alvéolaire Post-alvéolaire	Coronale	Rétroflexe	Palatale	Vélaire	Glottale
Plosive	Non-voisée	p		t t <sup>h</sup>		ʈ	c	k	
	Voisée	b		d d <sup>h</sup>		ɖ	ʃ	g	
Nasale		m		n		ɳ	ɲ	ŋ	
Vibrante				r		ɽ			
Fricative	Non-voisée			s	ʃ	ʂ			h
	Voisée		v						
Approximante Latéral- approximante							j		
				l		ɭ			

**Système vocalique :** 16 voyelles + 2 diphtongues

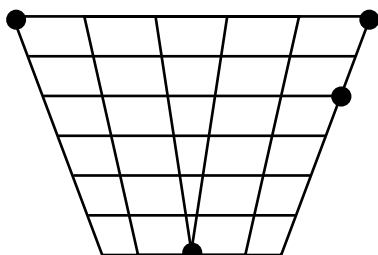
Voyelles de base



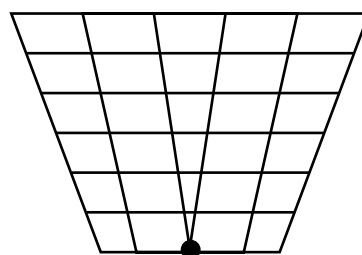
Voyelles longues



Voyelles nasales



Voyelles nasales longues



Diphtongues  
ai, au

**Langue :** Kanouri

**Famille et sous-famille :** Nilo-saharienne (ou nilo-tchadienne), saharienne.

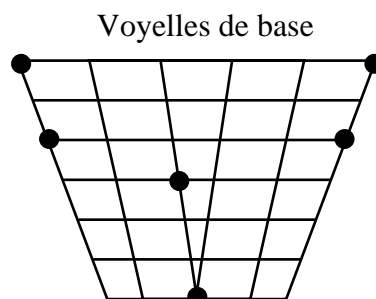
**Situation géographique :** Nord-Est du Nigeria (Etats de Bournou, Yobé, Jigawa), Est du Niger, Tchad, Soudan, Cameroun.

**Nombre de locuteurs :** 4 000 000 dont plus de 3 000 000 au Nigeria.

**Système consonantique :** 22 consonnes.

KANOURI		Bilabiale	Labio-dentale	Coronale		Palatale	Vélaire	Glottale	Labio-vélaire
				Alvéolaire	Post-alvéolaire				
Plosive	Non-voisée	p		t			k	ʔ	
	Voisée	b		d			g		
Nasale		m		n		ɲ	ŋ		
Battue				r					
Fricative	Non-voisée		f	s	ʃ			h	
	Voisée			z	ʒ				
Affriquée	Non-voisée			tʃ					
Approximante Latéral- approximante						j			w
				l					

**Système vocalique :** 6 voyelles



**Langue :** Kwakw'ala

**Famille et sous-famille :** Nord-américaine, almosane.

**Situation géographique :** Ouest du Canada (Colombie britannique).

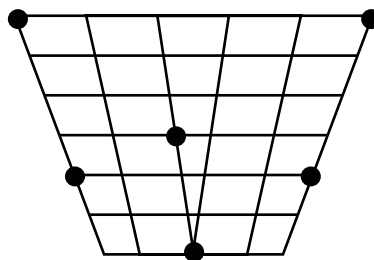
**Nombre de locuteurs :** 250.

**Système consonantique :** 43 consonnes.

KWAKW'ALA		Bilabiale	Coronale	Palatale	Vélaire	Uvulaire	Glottale	Labio-vélaire
Plosive	Non-voisée	p p'	t t'		k k' k <sup>w</sup> k <sup>w</sup> '	q q' q <sup>w</sup> q <sup>w</sup> '	ʔ	
	Voisée	b	d		g g <sup>w</sup>	ɢ ɢ <sup>w</sup>		
Nasale		m m̃	n ñ					
Battue			r					
Fricative	Non-voisée		s		x x <sup>w</sup>	χ χ <sup>w</sup>	h	
Latérale fricative	Non-voisée		ʃ					
Affriquée	Non-voisée		ts tʃ ts' tʃ'					
	Voisée		dz dʒ					
Approximante Latéral-approximante				j j				w w̃
			l l̃					

**Système vocalique :** 6 voyelles

Voyelles de base



**Langue :** Navaho

**Famille et sous-famille :** Na-déné, athabaskane.

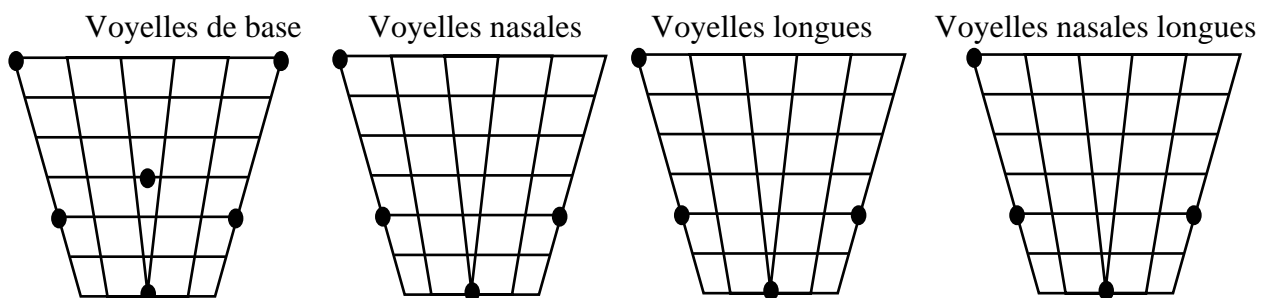
**Situation géographique :** Etats-Unis (Arizona, Nouveau Mexique, Utah, Colorado, Texas, Oklahoma, Kansas).

**Nombre de locuteurs :** 150 000.

**Système consonantique :** 38 consonnes.

NAVAHO		Bilabiale	Coronale		Palatale	Vélaire	Glottale	Labio-vélaire
			Alvéolaire	Post-alvéolaire				
<b>Plosive</b>	<b>Non-voisée</b>	p	t t' t <sup>h</sup>			k k' k <sup>h</sup> k <sup>w</sup>	ʔ	
	<b>Voisée</b>							
<b>Nasale</b>		m	n n'					
<b>Battue</b>			r					
<b>Fricative</b>	<b>Non-voisée</b>		s s'	ʃ	x		h h <sup>w</sup>	
	<b>Voisée</b>		z z'	ʒ	ɣ			
<b>Latérale fricative</b>	<b>Non-voisée</b>		ɬ ɬ'					
<b>Affriquée</b>	<b>Non-voisée</b>		ts ts' ts <sup>h</sup>	tʃ tʃ' tʃ <sup>h</sup>				
			tɬ tɬ' tɬ <sup>h</sup>					
<b>Approximante Latéral-approximante</b>					j			w
			l l'					

**Système vocalique :** 18 voyelles + 4 diphtongues et 4 triptongues



Diphtongues : eɪ, ai, aɔ, ɔi

Triptongues : aai, aii, eii, ɔɔi

**Langue :** Ngizim

**Famille et sous-famille :** Afro-asiatique, tchadique.

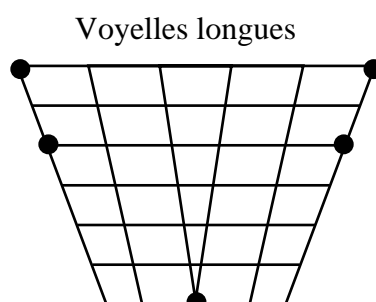
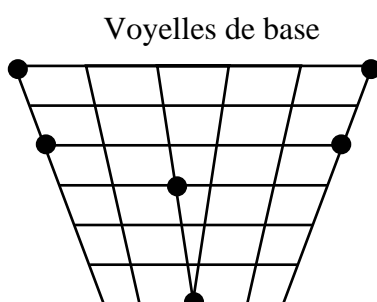
**Situation géographique :** Nord-est du Nigeria (Etat de Yobé).

**Nombre de locuteurs :** 80 000.

**Système consonantique :** 37 consonnes.

NGIZIM		Bilabiale	Labio-dentale	Coronale		Palatale	Vélaire	Glottale	Labio-vélaire
				Alvéolaire	Post-alvéolaire				
<b>Plosive</b>	<b>Non-voisée</b>	p		t		c	k kk k <sup>w</sup>		
	<b>Voisée</b>	b		d		ɟ	g g <sup>w</sup>		
<b>Implosive</b>	<b>Voisée</b>	ɓ		ɗ					
<b>Prénasalisée</b>	<b>Voisée</b>	mb		nd			ŋg ŋg <sup>w</sup>		
<b>Nasale</b>		m		n		ɲ	ŋ		
<b>Vibrante</b>				r					
<b>Battue</b>				r					
<b>Fricative</b>	<b>Non-voisée</b>		f	s	ʃ			h	
	<b>Voisée</b>		v	z	ʒ				
<b>Latérale fricative</b>	<b>Non-voisée</b>			ɬ					
	<b>Voisée</b>			ɮ					
<b>Affriquée</b>	<b>Non-voisée</b>			kʃ					
<b>Approximante Latéral-approximante</b>						j			w
				l					

**Système vocalique :** 11 voyelles et 3 diphtongues



Diphtongues  
au, ai, eu

**Langue :** Nyah Kur

**Famille et sous-famille :** Austro-asiatique, mön.

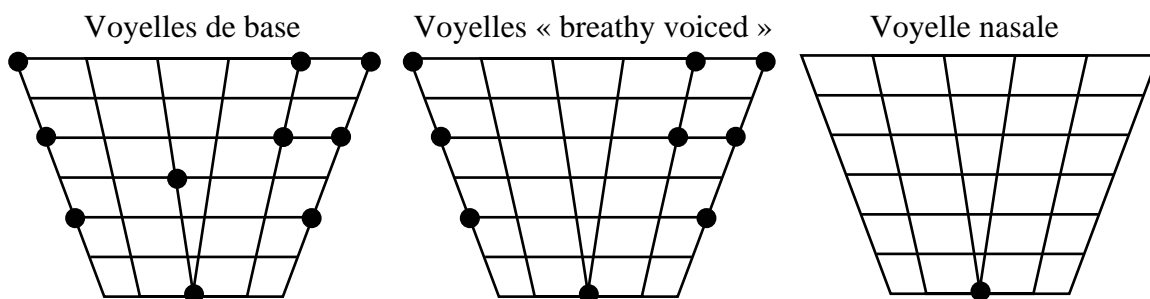
**Situation géographique :** Centre de la Thaïlande.

**Nombre de locuteurs :** 300 à 10 000.

**Système consonantique :** 30 consonnes.

NYAH KUR		Bilabiale	Labio-dentale	Coronale	Palatale	Vélaire	Glottale	Labio-vélaire
Plosive	Non-voisée	p p <sup>h</sup>		t t <sup>h</sup>	c c <sup>h</sup>	k k <sup>w</sup> k <sup>h</sup>	ʔ	
	Voisée	b		d				
Nasale		m m̰		n n̰	ɲ	ŋ		
Vibrante				r r̰				
Fricative	Non-voisée		f	s			h	ʌ
Approximante Latéral-approximante				ɹ	j j̰			w
				l l̰				

**Système vocalique :** 20 voyelles + 3 diphtongues



Diphtongues : iə, uə, ʊə

**Langue :** Quechua (variété de Quito). Il existe plusieurs variétés de quechua entre lesquelles l'intercompréhension est difficile, voire impossible.

**Famille et sous-famille :** Sud-américaine, andine.

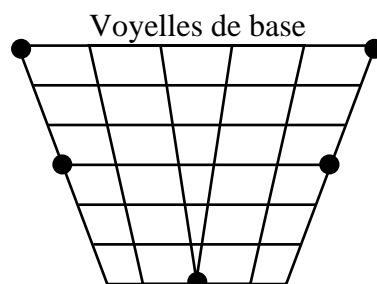
**Situation géographique :** Amérique du Sud (Pérou, Bolivie, Equateur, Argentine, Chili, Colombie).

**Nombre de locuteurs :** environ 10 000 000 si l'on comprend toutes les variétés.

**Système consonantique :** 33 consonnes.

QUECHUA		Bilabiale	Labio-dentale	Coronale		Rétroflexe	Palatale	Vélaire	Uvulaire	Glottale	Labio-vélaire
				Alvéolaire	Post-alvéolaire						
Plosive	Non-voisée	p p <sup>h</sup> p'		t t <sup>h</sup> t'			c	k k <sup>h</sup> k'	q q <sup>h</sup> q'		
	Voisée	b		d				g			
Nasale		m		n			ɲ				
Battue				r							
Fricative	Non-voisée		f	s	ʃ			x			h
	Voisée					ʒ					
Affriquée	Non-voisée				tʃ tʃ <sup>h</sup> tʃ'						
Approximante Latéral-approximante							j				w
					l		ʎ				

**Système vocalique :** 5 voyelles



**Langue :** Sora

**Famille et sous-famille :** Austro-asiatique, moundâ.

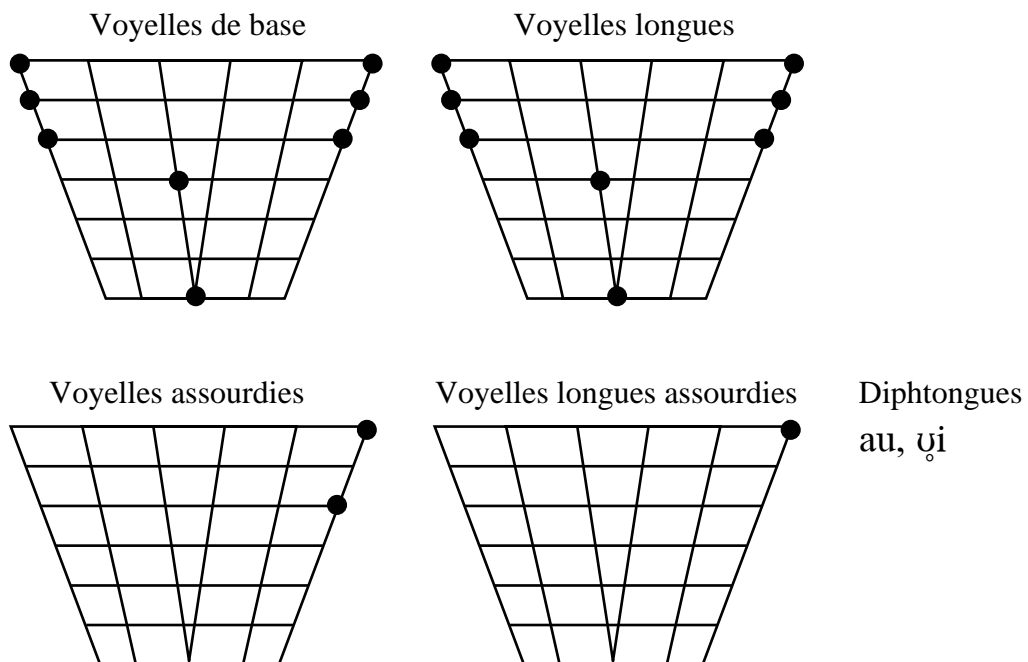
**Situation géographique :** Inde (Etats de l'Orissa, Andhra Pradesh, Madhya Pradesh, Bihar, Tamil Nadu, Bengale occidental, Assam).

**Nombre de locuteurs :** 300 000.

**Système consonantique :** 51 consonnes.

SORA		Bilabiale	Coronale	Rétroflexe	Palatale	Vélaire	Glottale
Plosive	Non-voisée	p p <sup>h</sup> pp	t t'			k k <sup>h</sup> k' kk	ʔ
	Voisée	b b' bb ḅ	d d' dd ḍ			g g' gg	
Nasale		m m' mm	n n' nn		ɲ n'	ŋ ŋŋ ŋ	
Vibrante			r r' ṛ				
Fricative	Non-voisée		s s' ss				
	Voisée		z	ẓ			
Affriquée	Voisée		dj dj' dj̣				
Approximante Latéral- approximante					j j' jj j̣		
			l l' ll ḷ				

**Système vocalique :** 19 voyelles + 2 diphtongues





**Langue :** Suédois (variété de Stockholm)

**Famille et sous-famille :** Indo-européenne, germanique du nord.

**Situation géographique :** Suède, Finlande.

**Nombre de locuteurs :** 9 000 000 dont 8 000 000 en Suède.

La langue étant en cours de transcription, le système phonologique utilisé dans le lexique dont nous disposons n'est pas encore mis en évidence.

**Langue :** Thaï

**Famille et sous-famille :** Kam-thaï, kadaï.

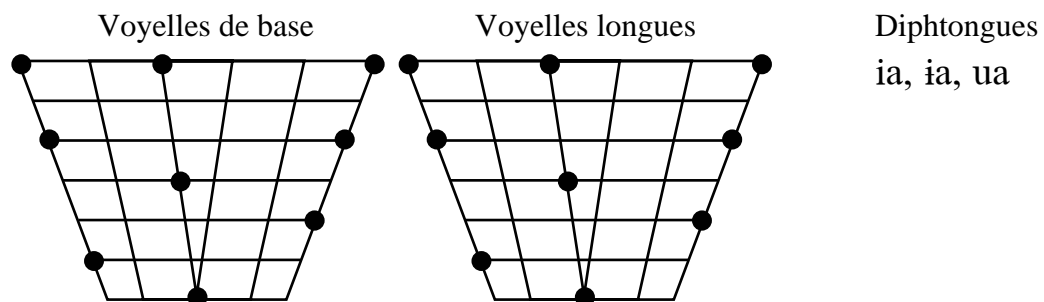
**Situation géographique :** Centre de la Thaïlande, Birmanie, Chine, Vietnam, Laos.

**Nombre de locuteurs :** 30 000 000.

**Système consonantique :** 22 consonnes.

THAÏ		Bilabiale	Labio-dentale	Coronale	Palatale	Vélaire	Glottale	Labio-vélaire
Plosive	Non-voisée	p p <sup>h</sup>		t t <sup>h</sup>	c c <sup>h</sup>	k k <sup>h</sup>	ʔ	
	Voisée	b		d		g		
Nasale		m		n		ŋ		
Battue				r				
Fricative	Non-voisée		f	s			h	
Approximante Latéral- approximante					j			w
				l				

**Système vocalique :** 18 voyelles + 3 diphtongues



**Langue :** Wa

**Famille et sous-famille :** Austro-asiatique, palaung.

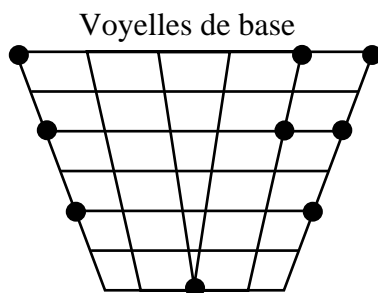
**Situation géographique :** Birmanie (Etat Shan), Chine (Sud du Yunnan).

**Nombre de locuteurs :** 530 000 dont 350 000 en Birmanie.

**Système consonantique :** 37 consonnes.

WA		Bilabiale	Labio-dentale	Coronale		Vélaire	Glottale
				Alvéolaire	Post-alvéolaire		
Plosive	Non-voisée	p p <sup>h</sup>		t t <sup>h</sup>		k k <sup>h</sup>	ʔ
	Voisée	b b̥		d d̥		g g̥	
Prénasalisée	Non-voisée					ŋk	
Nasale		m m̥		n n̥ ɲ ɲ̥		ŋ	
Vibrante				r̥			
Fricative	Non-voisée		f	s			h
	Voisée		v v̥		ʃ ʃ̥		
Affriquée	Non-voisée			tɕ tɕ <sup>h</sup>			
	Voisée			ts <sup>h</sup> ts̥			
Latéral-approximante				l l̥			

**Système vocalique :** 9 voyelles + 14 diphtongues et 4 triptongues



Diphtongues :

ɣi, iɛ, ia, io, iu, au, ai,  
au, oi, oi, ui, ui, ua, uo

Triptongues :

iuo, iau, aui, uai

**Langue :** Yup'ik (variété d'Alaska central) ; Trois variétés : le yup'ik central d'Alaska, le yup'ik de l'île St Lawrence et le yup'ik du Pacifique. Le territoire couvert par ces trois yup'ik comprend l'Alaska et le nord-est de la Sibérie.

**Famille et sous-famille :** Eskimo-aléoute, eskimo.

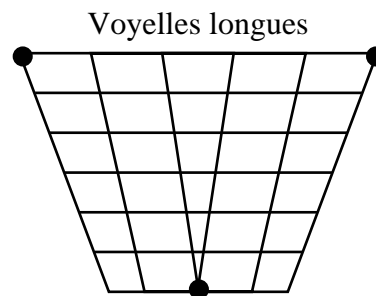
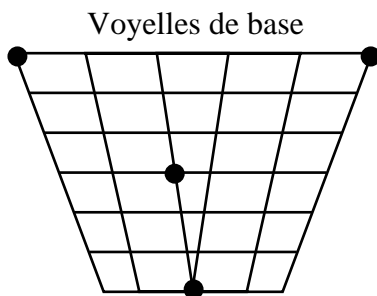
**Situation géographique :** Etat-Unis (Alaska), Russie (Nord-Est de la Sibérie).

**Nombre de locuteurs :** 11 000 dont 10 000 en Alaska.

**Système consonantique :** 40 consonnes.

YUP'IK		Bilabiale	Labio-dentale	Coronale	Palatale	Vélaire	Uvulaire	Glottale
Plosive	Non-voisée	p pp		t tt	c cc	k kk	q q	ʔ
	Voisée							
Nasale		m m̥		n n̥		ŋ ŋ		
Fricative	Non-voisée		f	s ss		x x <sup>w</sup> xx	χ χ <sup>w</sup> χχ	
	Voisée		v vv	z zz		ɣ ɣɣ	ʁ ʁ <sup>w</sup> ʁʁ	
Approximante					j jj			
Latéral-approximante				l ll l̥ ll̥				

**Système vocalique :** 7 voyelles + 3 diphtongues



Diphtongues :  
ai, au, ui

**Langue :** !Xóõ

**Famille et sous-famille :** Khoïsan, khoïsan-septentrional.

**Situation géographique :** Botswana (désert du Kalahari), Namibie.

**Nombre de locuteurs :** 3 000 à 4 000 (la très grande majorité au Botswana).

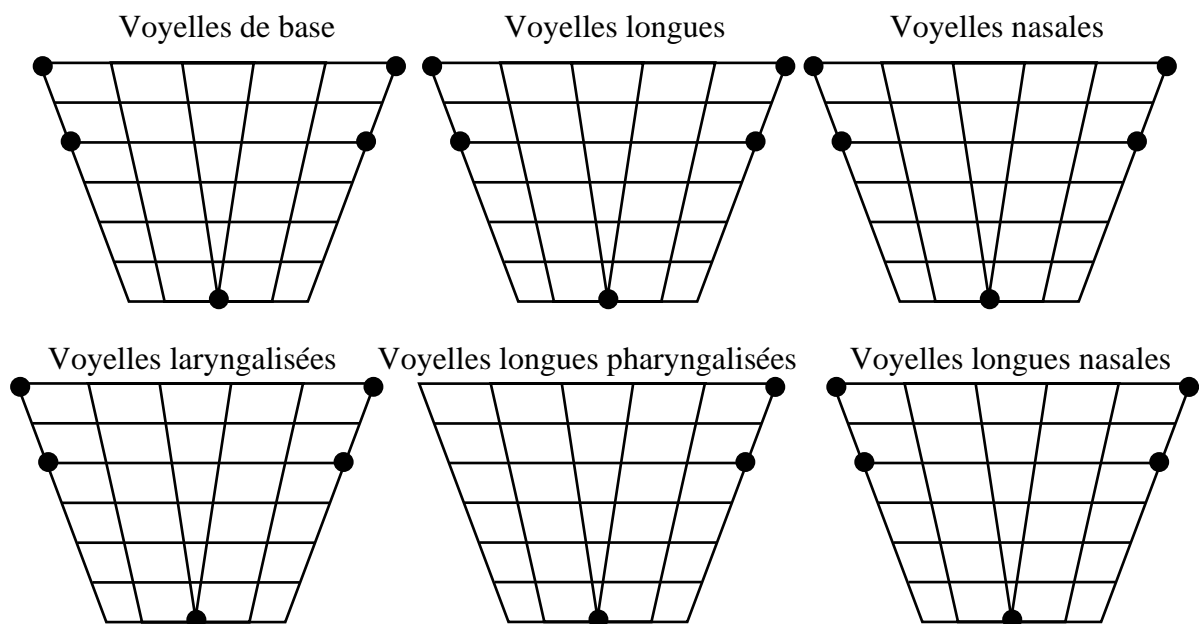
**Système consonantique :** 46 consonnes et 83 clicks.

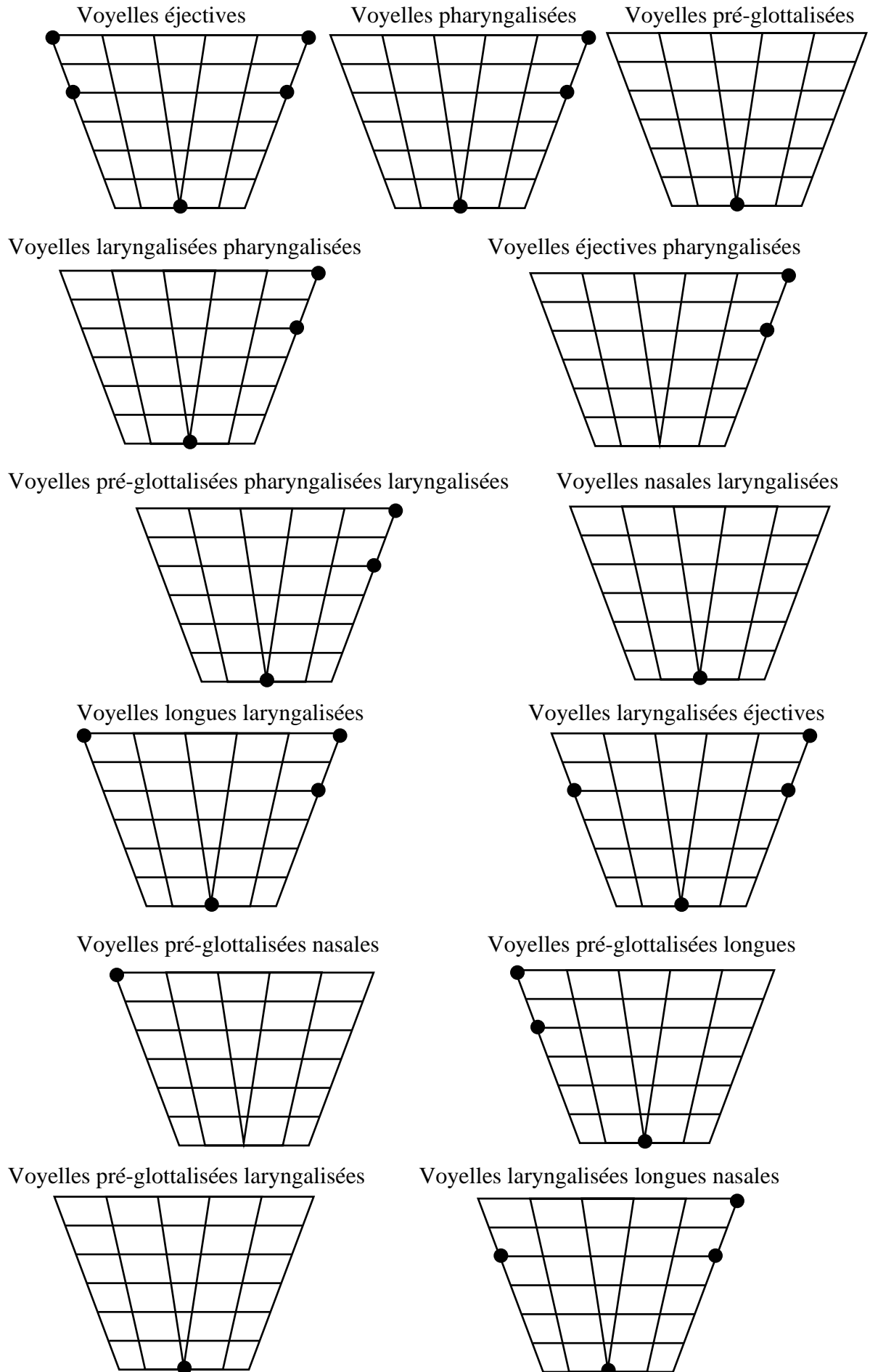
!Xoo		Bilabiale	Coronale	Palatale	Vélaire	Uvulaire	Glottale	Labio-vélaire
Plosive	Non-voisée	p p <sup>h</sup>	t̚ t̚ <sup>h</sup> t̚'		k k <sup>h</sup> k'	q q <sup>h</sup> q'	ʔ	
	Voisée	b	d̚		g	ɠ		
Nasale		m m̚	n̚ n̚	ɲ				
Vibrante			r					
Fricative	Non-voisée		s		x		h	
Affriquée	Non-voisée		ts ts <sup>h</sup> ts'					
	Voisée		dʒ					
Approximante Latéral-approximante				j				w
			l					

Consonnes avec doubles plosions : dts<sup>h</sup>, dt<sup>h</sup>, gk<sup>h</sup>, Gq<sup>h</sup>, p'kx', t'kx', ts'kx', dt'kx', dts'kx', dt<sup>x</sup>, t<sup>x</sup>, gk<sup>x</sup>, dts<sup>hx</sup>

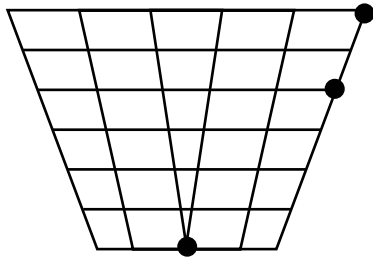
5 clicks (⊙, ‡, †, ‡, ‡) et 78 différentes combinaisons avec les articulations secondaires

**Système vocalique :** 74 voyelles + 71 diphtongues

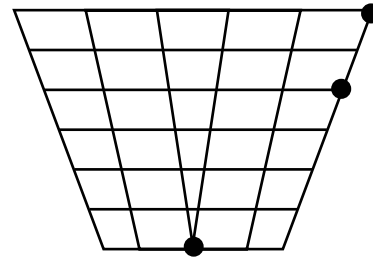




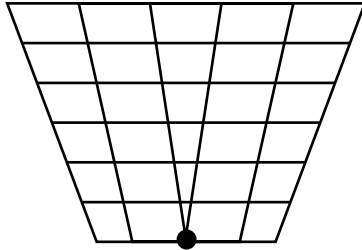
Voyelles pharyngalisées nasales longues



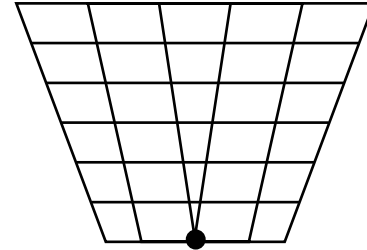
Voyelles pharyngalisées laryngalisées longues



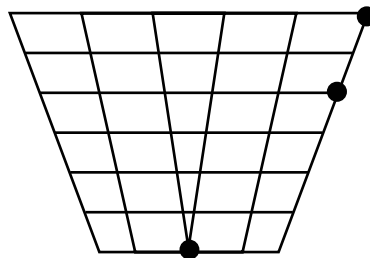
Voyelles laryngalisées pharyngalisées éjectives



Voyelles pré-glottalisées nasales longues



Voyelles pharyngalisées laryngalisées nasales longues



**Diphthongues :**

De base : ae, ai, ao, au, ia, oa, oe, ou, ua, ue, ui

Pharyngalisées : u<sup>ɸ</sup>i, u<sup>ɸ</sup>e, u<sup>ɸ</sup>a, ou<sup>ɸ</sup>, oi<sup>ɸ</sup>, oa<sup>ɸ</sup>, au<sup>ɸ</sup>, ao<sup>ɸ</sup>, ai<sup>ɸ</sup>, ae<sup>ɸ</sup>

Laryngalisées : u̠i, u̠e, u̠a, o̠e, o̠a, a̠u, a̠o, a̠i, a̠e

Pharyngalisées nasales : u̠<sup>ɸ</sup>i, u̠<sup>ɸ</sup>e, u̠<sup>ɸ</sup>a, o̠<sup>ɸ</sup>u, o̠<sup>ɸ</sup>a, a̠<sup>ɸ</sup>u, a̠<sup>ɸ</sup>o, a̠<sup>ɸ</sup>i, a̠<sup>ɸ</sup>e

Pharyngalisées laryngalisées : u̠<sup>ɸ</sup>i, u̠<sup>ɸ</sup>e, u̠<sup>ɸ</sup>a, ou̠<sup>ɸ</sup>, oa̠<sup>ɸ</sup>, au̠<sup>ɸ</sup>, ao̠<sup>ɸ</sup>, ai̠<sup>ɸ</sup>, ae̠<sup>ɸ</sup>

Laryngalisées nasales : ũ̠a, õ̠u, õ̠a, ã̠u, ã̠o, ã̠i

Pharyngalisées nasales laryngalisées : ũ̠<sup>ɸ</sup>i, ũ̠<sup>ɸ</sup>a, õ̠<sup>ɸ</sup>u, ã̠<sup>ɸ</sup>u, ã̠<sup>ɸ</sup>o, ã̠<sup>ɸ</sup>i

Nasales : ũ̠i, ũ̠a, õ̠i, õ̠a, ẽ̠i, ã̠u, ã̠o, ã̠i, ã̠e

Pré-glottalisées : 'au

Pré-glottalisées laryngalisées : 'a̠u

Ejectives : ae'

## II.4. QUANTIFICATION DES DONNEES D'ULSID

Les travaux d'harmonisation et de transcription des différents lexiques de l'échantillon effectués, une description détaillée de la réalité des données qui le compose s'impose. À cette étape, un certain nombre d'éléments quantitatifs et des liens pouvant exister entre eux et qui pourraient constituer un biais à notre analyse des structures lexicales et syllabiques ont été mesurés.

Les premières observations effectuées portent sur le nombre d'unités lexicales et de syllabes composant les dictionnaires de nos langues. Ainsi, les 16 lexiques de notre échantillon correspondent à 247 252 syllabes qui résultent du découpage de 94 535 items lexicaux, avec une moyenne de 15 453 syllabes par langue.

Nous relevons une grande diversité dans la taille respective des différents lexiques (figure II.4) puisque le nombre d'unités lexicales varie de 1 989 pour le ngizim à 12 181 pour le français, la moyenne se situant à 5 908 items lexicaux.

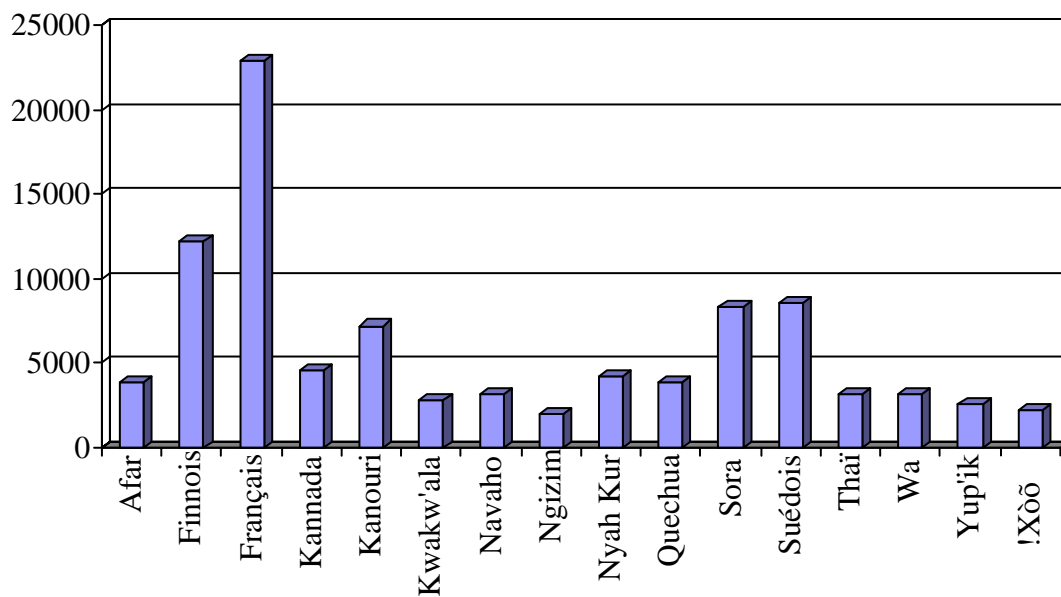


Figure II.4 : Taille des lexiques en nombre d'entrées lexicales.

Pour le nombre de syllabes contenues dans les différents lexiques (figure II.4), une grande diversité est également observée avec un nombre moyen de syllabes de 15 453, un minimum de 3 180 pour la wa et un maximum de 60 994 pour le français. On peut se douter que les langues possédant les plus grands lexiques sont celles qui présentent le plus grand nombre de syllabes. Mais, l'existence d'une relation directe entre ces deux données n'apparaît pas comme évidente si l'on compare les langues deux à deux. Ainsi, le kannada et le nyah kur, qui possèdent des lexiques de taille comparable (respectivement 4 559 et 4 188 unités lexicales), ont un nombre de syllabes très différent (respectivement 12 126 et 5 503).

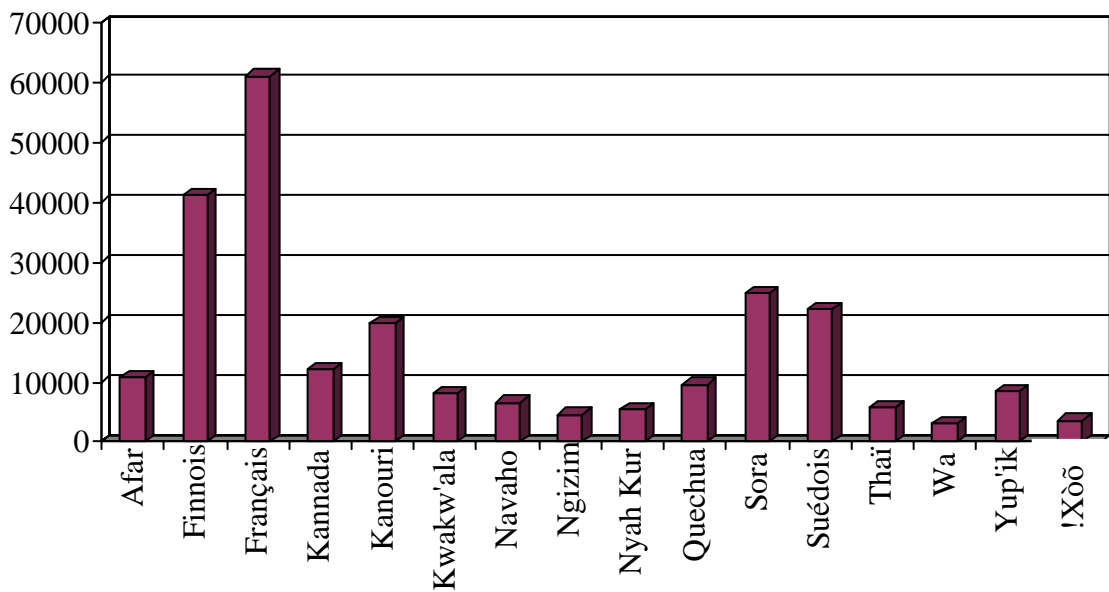


Figure II.5 : Nombre total de syllabes dans les lexiques de l'échantillon ULSID.

Il paraît donc important de chercher à mieux comprendre les relations qui lient la taille des lexiques au nombre de leurs syllabes. Cette vérification est d'autant plus nécessaire que la diversité qui existe au niveau du nombre d'unités lexicales contenues dans les différents dictionnaires de nos langues peut influencer certains résultats. Ainsi, il semble impératif de vérifier que la grande taille d'un lexique n'influence pas les résultats globaux et que la confusion entre une particularité propre à une langue pour laquelle nous possédons un grand lexique et une tendance générale à toutes nos langues soit écartée.

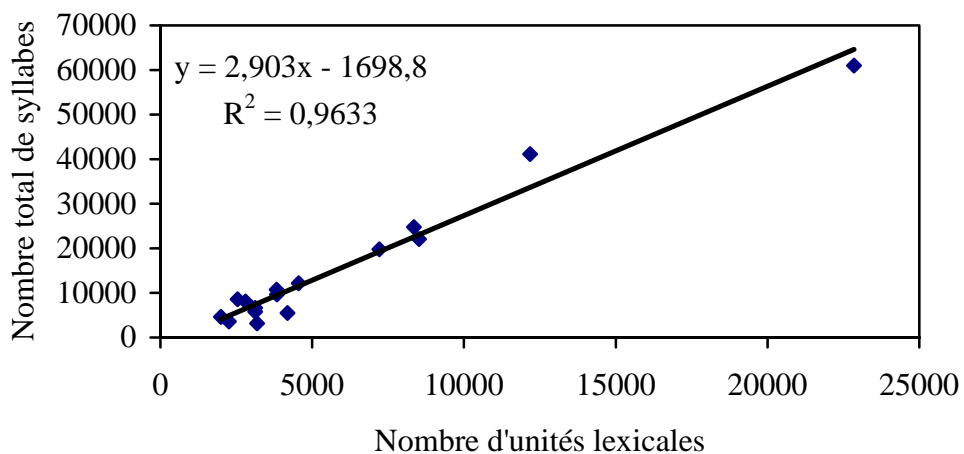


Figure II.6 : Corrélation entre nombre d'items lexicaux et nombre de syllabes pour les 16 langues.



La corrélation entre la taille des lexiques et le nombre total de syllabes pour toutes les langues retenues (cf. figure II.6) révèle que le nombre total de syllabes d'une langue semble être fortement lié au nombre d'unités lexicales de cette langue. En effet, le coefficient de corrélation est supérieur à 0,96. Pourtant, ce résultat va à l'encontre de l'observation concernant le nombre de syllabes et le nombre d'items lexicaux pour le kannada et le nyah kur. Une observation plus fine de la figure 3 permet de remarquer que les lexiques de plus de 5 000 unités "tirent" la corrélation. Nous avons, par conséquent, décidé d'écarter les cinq plus grands lexiques (ceux de plus de 7 000 unités lexicales) pour recalculer la corrélation. On remarque alors que les deux facteurs étudiés ne sont plus corrélés (cf. figure II.7). En effet, le nombre de syllabes d'une langue est globalement indépendant de la taille du lexique (coefficient de corrélation inférieur à 0,37). Les unités syllabiques de nos langues pourront donc être étudiées sans que nous tenions compte de la taille des différents lexiques, tout en gardant à l'esprit que les cinq lexiques les plus importants (finnois, français, kanouri, sora et suédois) possèdent une influence non négligeable. Nous contrôlerons les résultats obtenus sur les tendances par rapport aux structures favorisées de ces cinq lexiques.

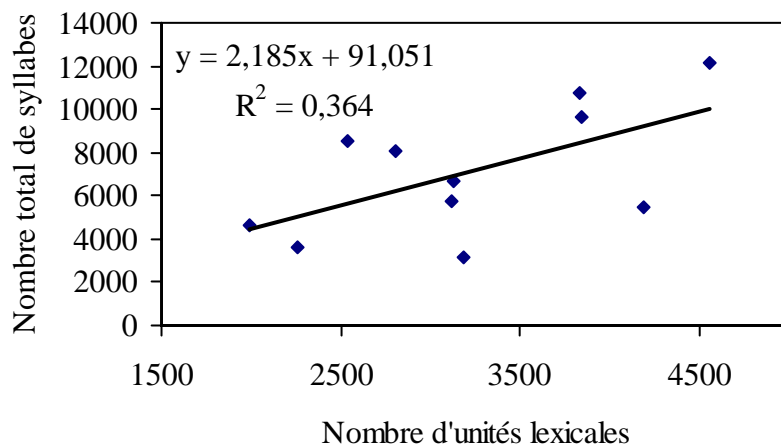


Figure II.7 : Corrélation entre le nombre d'items lexicaux et le nombre de syllabes pour les 11 lexiques de moins de 7 000 entrées.

## II.5. UNE INTERFACE POUR L'INTERROGATION D'ULSID

Après une première implantation des données sous environnement Hypercard (logiciel Macintosh), le choix d'un nouveau logiciel pour la création d'une deuxième interface d'interrogation de notre base de données a été décidé dans le but de répondre aux limites d'HyperCard. Le logiciel retenu n'est pas en soi un outil conçu pour le traitement de banques de données, mais ses possibilités de traitement de fichiers textes, de sorties statistiques et de graphisme interactif ont permis de créer une interface totalement adaptée au traitement des centaines de milliers de données et à la recherche de grandes tendances dans un temps optimisé.

Cette interface est conçue en deux plates-formes distinctes : la première pour le traitement des fichiers de lexiques au préalable de l'analyse (formatage et harmonisation des données), la seconde pour l'interrogation du contenu linguistique de ces lexiques à partir d'un certain nombre de requêtes prédéfinies (Brun, 2001 ; Alirezai, 2002).

La première partie de cette interface répond au travail de mise en forme avant traitement des données qui s'élabore selon trois étapes :

- 1) la correction de la structure du lexique : phase qui consiste à nettoyer le fichier contenant le lexique d'origine avec les fautes de frappe, les espaces insérés en début de mot, etc., ainsi qu'à homogénéiser la structure du lexique en considérant le contexte des séparateurs de syllabes et de constituants syllabiques, etc. Par exemple, la position des espaces par rapport au point séparateur de syllabe a nécessité une analyse particulière afin de ne pas représenter un obstacle à un traitement automatique identique pour tous les fichiers.
- 2) la transcription des fichiers avec l'Alphabet Phonétique International, suivie d'une transcription en C (consonnes) et V (voyelles) des items lexicaux : cette phase de transcodage des données a été possible grâce à l'élaboration au préalable de tables de correspondance ASCII-API considérant tous les segments relevés dans le lexique. Ainsi retranscrits sous forme CV, les lexiques ont pu livrer les types de cohortes et de gabarits les plus présents dans nos données (cf. table II.2).

Fichier ASCII	Fichier API	Fichier CV
a.k.a.l.u	a.k.a.l.u	V.CV.CV
a.k.k.a.n	a.k.k.ã	VC.CV
a.k.k.a.r.e	a.k.k.a.r.e	VC.CV.CV
a.k.k.i	a.k.k.i	VC.CV
a.k.k.u	a.k.k.u	VC.CV
a.g.a.t.e	a.g.a.t.e	V.CV.CV
a.g.a.t.y.a	a.g.a.t.j.a	V.CVC.CV
a.g.a.l.a	a.g.a.l.a	V.CV.CV
a.g.a.l.u	a.g.a.l.u	V.CV.CV
a.g.a.s.e	a.g.a.s.e	V.CV.CV
a.g.a.r.l.u	a.g.a.ɭ.u	V.CV.CV
a.g.a.r.l.i	a.g.a.ɭ.i	V.CV.CV
a.g.u.r.l.i.k.a.a.s.u	a.g.u.ɭ.i.k.a:.s.u	V.CV.CV.CV.CV

Table II.2 : Extrait de fichiers transcrits pour le kannada

- 3) Le codage d'informations sur la description au niveau segmental en fonction du lieu ou du mode d'articulation pour les consonnes et de la position avant-arrière de la langue pour les voyelles (cf. annexe VIII pour les abréviations utilisées par la suite) : étape nécessaire pour la mise en évidence de tendances liées à la nature même des constituants syllabiques (cf. table II.3).

Fichier API	Fichier lieu	Fichier Mode
a . k a . l u	Ce . Ve Ce . Co Po	Ce . Pl Ce . Ap Po
a k . k ã	Ce Ve . Ve Ce	Ce Pl . Pl Ce
a k . k a . r e	Ce Ve . Ve Ce . Co An	Ce Pl . Pl Ce . Vi An
a k . k i	Ce Ve . Ve An	Ce Pl . Pl An
a k . k u	Ce Ve . Ve Po	Ce Pl . Pl Po
a . g a . t e	Ce . Ve Ce . Co An	Ce . Pl Ce . Pl An
a . g a t . j a	Ce . Ve Ce Co . Pa Ce	Ce . Pl Ce Pl . Ap Ce
a . g a . l a	Ce . Ve Ce . Co Ce	Ce . Pl Ce . Ap Ce
a . g a . l u	Ce . Ve Ce . Co Po	Ce . Pl Ce . Ap Po
a . g a . s e	Ce . Ve Ce . Co An	Ce . Pl Ce . Fr An
a . g a . l u	Ce . Ve Ce . Co Po	Ce . Pl Ce . Ap Po
a . g u . l i . k a : . s u	Ce . Ve Po . Co An . Ve Ce . Co Po	Ce . Pl Po . Ap An . Pl Ce . Fr Po

Table II.3 : Extrait de fichiers transcrits pour le kannada (cf. Annexe VIII).

Une fois les fichiers formatés et transcrits, la deuxième partie de l'interface permet de choisir la forme des données à explorer. Les possibilités d'interrogation de la base de données sont multiples, en voici quelques unes des plus représentatives :

- La recherche du nombre de syllabes par item lexical (cf. figure II.8 et II.9), qui permet d'obtenir, pour chaque langue, le nombre d'unités lexicales composées d'une, de deux ou plusieurs syllabes.

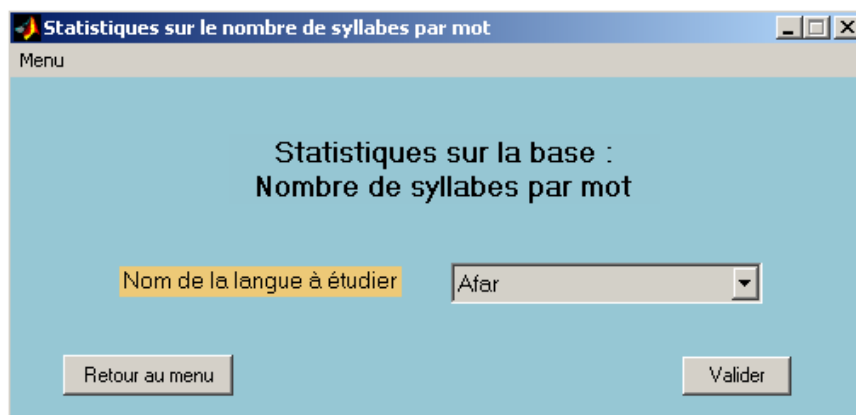


Figure II.8 : Fenêtre de l'interface MatLab de la base ULSID pour la recherche du nombre de syllabes par unité lexicale.

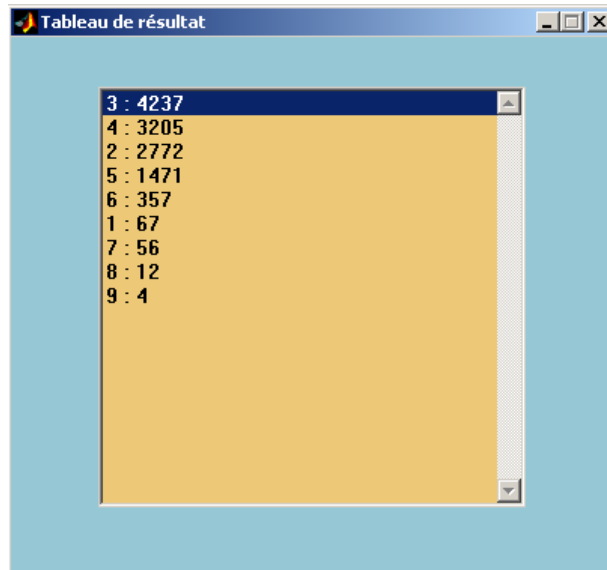


Figure II.9 : Fenêtre de l'interface MatLab de la base ULSID, résultats de la recherche du nombre de syllabes par unité lexicale pour le finnois. Le chiffre précédant les deux points indique le nombre de syllabe, et celui suivant les deux points indique le nombre d'unités lexicales correspondantes.

- Le calcul de matrices d'occurrences (cf. figure II.10) correspondant à l'analyse des combinaisons existant dans chaque langue entre l'attaque et le noyau d'une même syllabe, le noyau et la coda ou l'attaque et la coda.

		Nom de l'Attaque							
		Co	Bi	Ve	Ph	Lde	Pa	Lve	Gl
Nom du Noyau	Ce	1872	820	555	496	128	186	158	81
	An	1980	368	382	375	116	94	98	39
	Po	1057	365	373	193	71	35	18	36

Figure II.10 : Fenêtre de l'interface MatLab de la base ULSID (matrice obtenue pour les combinaisons possibles entre les différents lieux consonantiques possibles pour l'attaque et les trois positions de la langue possibles pour le noyau en afar).

- Le nombre et la forme des différents gabarits lexicaux pour chaque langue en fonction du type de fichiers transcodés (API, lieu, mode d'articulation).

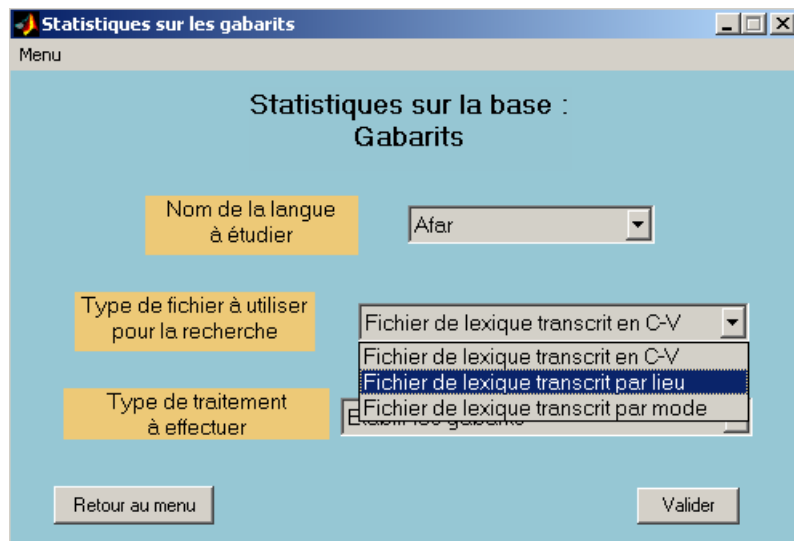
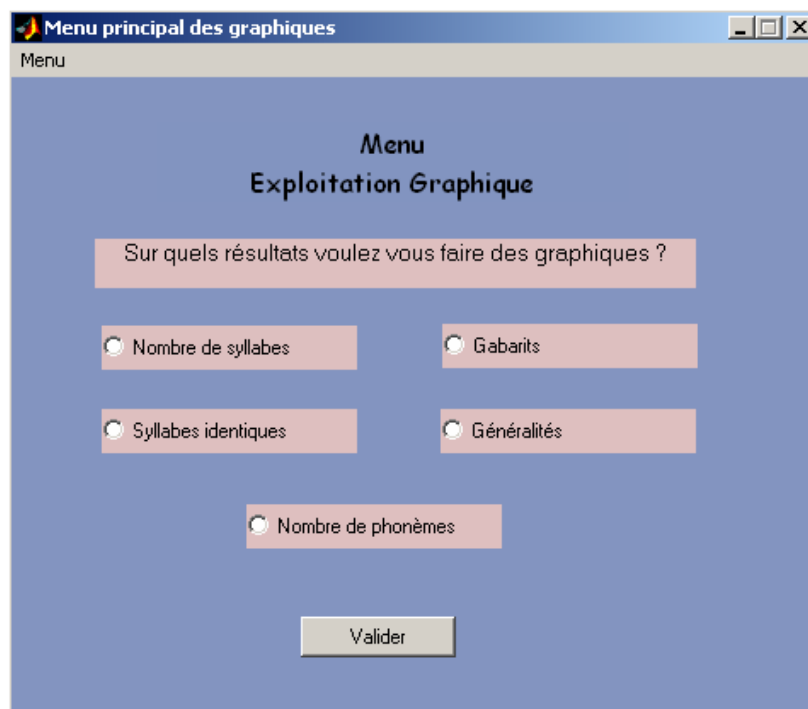


Figure II.11 : Fenêtre de l'interface MatLab de la base ULSID. Recherche des gabarits lexicaux selon le type de fichier : API, cohorte CV, lieu, mode articuloire.

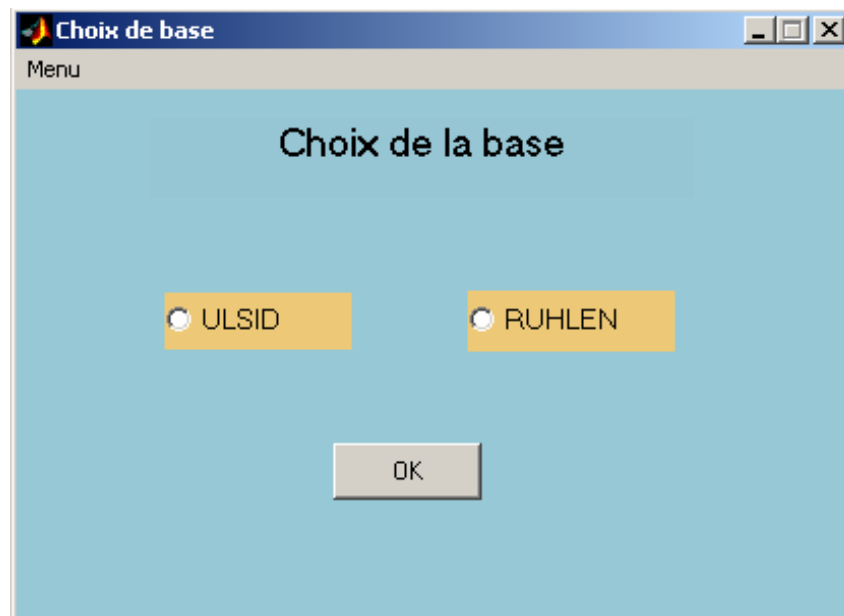
Précisons que chacun des modules présentés ci-dessus offre deux possibilités d'affichage des résultats : 1) les données chiffrées des requêtes ; 2) les représentations graphiques correspondantes.



*Figure II.12 : Fenêtre de l'interface MatLab de la base ULSID pour la sortie graphique des requêtes.*

L'ensemble des traitements sur les lexiques d'entrées codés en ASCII, ainsi que la totalité des résultats que nous présentons en détail dans le chapitre suivant, ont été obtenus grâce à cette interface MatLab, qui a été conçue extensible, tant au niveau des données d'entrées (lexiques de langues syllabés) qu'au niveau des possibilités de requêtes.

Dans le cadre de la vérification des résultats présentés par MacNeilage et Davis (2000), à propos d'une tendance des structures lexicales des langues du monde qu'ils retrouvent dans les 27 racines universelles reconstruites proposées par Ruhlen (1997), nous avons complété l'interface ULSID en implémentant, sur les mêmes principes, les données fournies en annexe de l'ouvrage de Ruhlen (1997) « *L'origine des langues* ». Il s'agit de données collectées dans le but de montrer l'existence de racines universelles et de démontrer que toutes les langues parlées ou éteintes à la surface du globe sont issues d'une langue mère unique. Sans entrer dans quelque considération que ce soit à propos de la thèse de Ruhlen (1997), nous avons harmonisé l'ensemble des données lexicales issues de langues et de proto-langues et procédé à leur exploitation sur la même base que les données d'ULSID.



*Figure II.13 : Interface MatLab permettant le choix entre la base ULSID et la base de données deRuhlen.*

Le traitement se fait de manière semblable pour les deux bases (ULSID et Ruhlen), même si certaines requêtes ne sont pas applicables à la deuxième du fait que ces données ne soient pas découpées en syllabes. Une des analyses à mener est la recherche portant sur les structures gabaritiques possibles pour les unités lexicales appartenant à la base de Ruhlen.

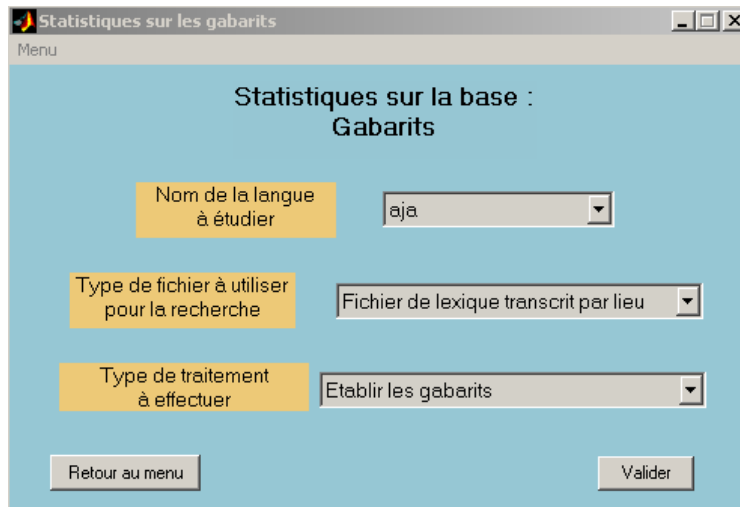


Figure II.14 : Interface MatLab permettant l'étude des gabarits (codage des unités lexicales en C et V) de la racine [aja] de Ruhlen en tenant compte du lieu d'articulation des unités.

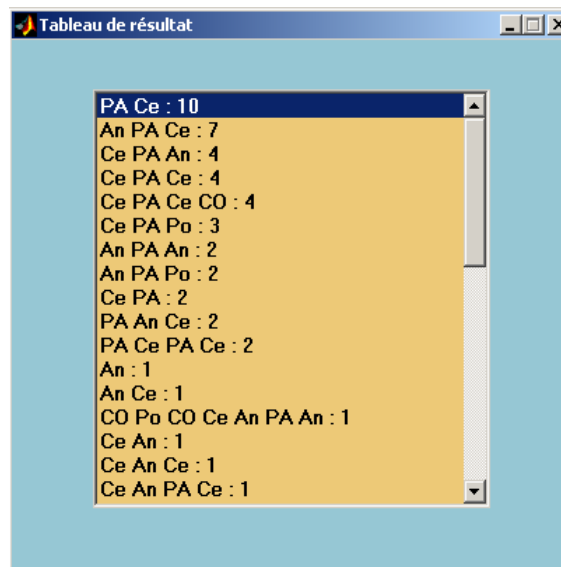


Figure II.15 : Interface MatLab présentant les gabarits en fonction du lieu d'articulation pour les unités lexicales de la racine [aja].

Nous présentons, dans le chapitre suivant, l'ensemble des résultats typologiques obtenus ainsi que les analyses menées sur les données pour mettre en évidence de certaines grandes tendances des structures lexicales et syllabiques.

## **CHAPITRE III**

# **TYPOLOGIES ET TENDANCES DES STRUCTURES LEXICALES ET SYLLABIQUES**



### III.1. DE L'IMPORTANCE DES RECHERCHES EN TYPOLOGIE LINGUISTIQUE

À partir du moment où des ressemblances ont été mises en évidence entre les langues, il est naturel que les linguistes aient cherché à comprendre les raisons de ces similitudes. Beaucoup d'études ont porté sur ce sujet, aboutissant à des avancées certaines comme dans le cas des universaux phonologiques, sans pour autant arriver à un consensus total puisque les similitudes observées peuvent concerner un grand nombre de faits de langue qui restent néanmoins débattus dans la communauté linguistique. De plus, les explications avancées varient, elles aussi, en fonction des théories adoptées.

Ainsi, Ruhlen (1997, p. 22) aborde cette question de la façon suivante : « Comment expliquer l'existence de formes semblables ayant le même sens dans diverses langues ? En fait, les différentes explications possibles sont les mêmes que pour les ressemblances entre des protéines, des animaux ou des religions. [...] Quel que soit le domaine - animaux, langues, religions, il n'importe -, on peut expliquer les ressemblances par trois mécanismes en tout et pour tout : la convergence, l'emprunt et l'origine commune. Personne n'a proposé de façon convaincante une quatrième explication ».

Comrie (1980) considère quant à lui que les similitudes entre deux langues peuvent avoir quatre origines possibles : le fait que les langues sont génétiquement apparentées et ont hérité ce caractère de leur ancêtre commun (1) ; les contacts géographiques entre les deux langues (emprunt) (2) ; l'existence d'un universel du langage humain (3) ; le hasard (4).

Ces différentes explications possibles n'ont pas toutes la même portée. En effet, si les facteurs génétiques et géographiques sont reconnus par tous et les procédés assez bien décrits, les autres sont beaucoup plus flous. Ainsi, Ruhlen appelle « *convergence* » le processus qui fait que des objets différents en arrivent à se ressembler accidentellement ou sous l'effet d'un facteur extérieur à la langue, recouvrant en partie les deux derniers aspects développés par Comrie, si l'on considère que les universaux du langage peuvent être dus à des facteurs non-linguistiques au sens saussurien du terme. Mais, ce dernier fait n'est actuellement pas accepté par tous les linguistes.

De ces quatre aspects énoncés, seuls deux ont donné naissance à deux branches de la linguistique, fondées sur des principes et des objectifs différents : l'approche génétique en linguistique est basée sur la recherche d'une origine commune entre des groupes de langues, tandis que l'approche typologique se développe en lien avec la recherche d'universaux du langage dans le but de comprendre les mécanismes qui sous-tendent le fonctionnement du langage humain.

Hjelmslev (1966) écrit : « une typologie linguistique exhaustive est la tâche la plus grande et la plus importante qui s'offre à la linguistique. [...] En fin de compte, sa tâche est de répondre à la question : quelles structures linguistiques sont possibles ? Et pourquoi telles structures sont possibles quand d'autres ne le sont pas ? Ce faisant, elle doit, plus qu'aucune autre espèce de linguistique, s'approcher de ce qu'on pourrait appeler le problème de la nature

du langage. Et en dernier ressort, elle s'avèrera supérieure à la linguistique génétique ; seule en effet, elle peut permettre de comprendre les lois générales selon lesquelles les langues changent, et les possibilités de changement que comporte un type donné. C'est seulement par la typologie que la linguistique s'élève à des points de vue tout à fait généraux et devient une science » (Hjelmslev, 1966, p. 128-129). Ainsi, la typologie linguistique s'efforce de découvrir, de manière orthogonale aux niveaux génétique et géographique, des régularités structurelles qui permettent de les classer dans un même type, considéré comme l'un des systèmes possibles satisfaisant aux conditions générales du fonctionnement du langage humain.

La typologie va connaître un tournant important avec la proposition de Greenberg à la conférence de Dobbs Ferry en 1961 qui l'associe, de façon incontournable, à la recherche des universaux du langage : « *Le travail typologique de Greenberg (1963) change radicalement de perspectives : il ne s'agit plus de classer pour classer mais de découvrir les universaux du langage* » (Coyaud, 1980, p. 14).

La recherche des universaux est représentée par deux théories distinctes qui se développent en parallèle. La première, en rapport avec la typologie des langues, est dans la mouvance de Greenberg (Greenberg, 1966). En 1967, Chomsky tient, à Austin, une autre conférence sur les universaux dans la théorie linguistique (Bach & Harms, 1968). Ces deux théories, bien qu'utilisant des méthodes différentes, poursuivent un but commun ainsi défini par Bach et Harms (1968), qui suivent la théorie chomskyenne : « *What is linguistic theory itself if not the attempt to discover what is common to all languages, what is essential in the notion 'natural language', what are the limits within which languages can vary, what are the (universal !) terms by means of which this variation can be described ?* » (Bach & Harms, 1968, p. vi). Ce propos est rejoint par celui de Comrie, reprenant la théorie de Greenberg : « *Language universals research is concerned with finding those properties that are common to all languages* » (Comrie, 1980, p. 30).

C'est dans cette optique générale de la recherche des universaux du langage humain que se situe notre étude des 16 langues de la base ULSID, tant au niveau de la structuration syllabique des unités lexicales qu'à celui de l'organisation interne de la syllabe.

## **III.2. ORGANISATION STRUCTURELLE DES UNITES LEXICALES**

### **III.2.1. Taille des unités lexicales**

L'étude de la structuration des unités lexicales en syllabes, nous a conduit à examiner, dans chaque langue, le nombre de syllabes dans ces unités, leur position ainsi que les contraintes pouvant peser sur la concaténation de deux types syllabiques successifs. Un ensemble de grandes tendances a ainsi été mis en évidence auquel nous tentons d'apporter des éléments d'explication.

### III.2.1.1. Forme canonique : vers une typologie des langues sur le critère du nombre de syllabes par unité lexicale

L'une des premières analyses réalisées à partir de la base ULSID a consisté à calculer ce que nous avons appelé la forme canonique (FC) des langues selon la formulation de Monino et Roulon (1972), c'est-à-dire le rapport entre le nombre total de syllabes pour l'ensemble d'un lexique donné et le nombre d'entrées lexicales de ce dernier (figure III.1). Ce calcul correspond en fait au nombre moyen de syllabes que contiennent les unités lexicales d'une langue. Monino et Roulon (1972) obtiennent, pour le gbara (langue africaine de la famille Niger-Congo), une forme canonique de 1,92. Il est intéressant de comparer ce résultat aux formes canoniques des langues de notre échantillon (cf. figure III.1). On observe ainsi que, pour ces dernières, la forme canonique varie entre 1 pour le wa et 3,38 pour le finnois, avec une moyenne pour l'ensemble de notre corpus de 2,42. Il semble donc que le nombre moyen de syllabes par unité lexicale tourne majoritairement autour de 2 et que les langues composant l'échantillon retenu présentent des lexiques globalement dissyllabiques. L'étude globale de la forme canonique n'apporte que peu d'informations supplémentaires, alors qu'une observation plus fine laisse apparaître des regroupements possibles.

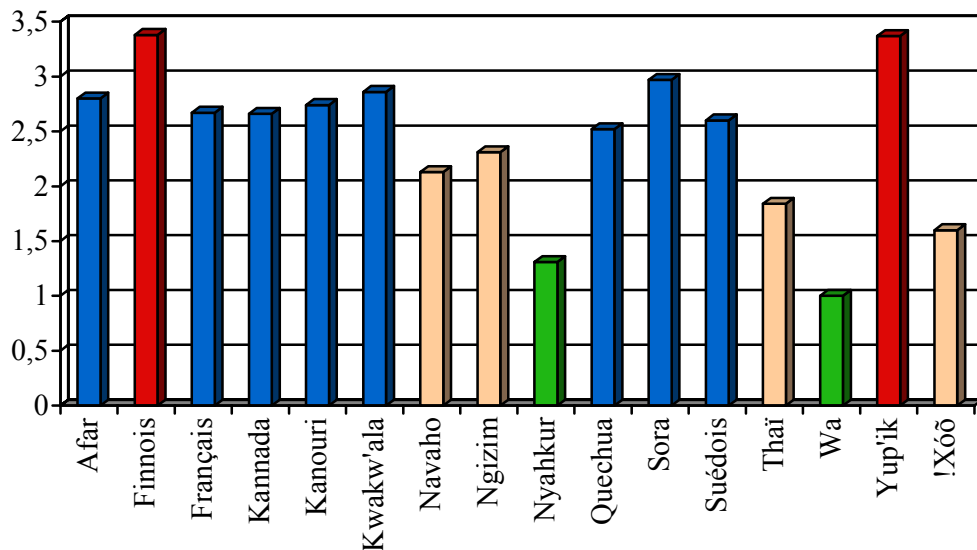


Figure III.1 : Forme canonique des 16 langues de notre échantillon.

Ainsi, l'analyse de la figure 5 permet de remarquer que les 16 langues d'ULSID se répartissent en 4 types assez nets (cf. table III.1) dont un, le troisième, est largement dominant :

Valeurs de la Forme Canonique	Nombre de langues	Langues	% d'ULSID
$FC \leq 1,5$	2	nyah kur, wa	12,5 %
$1,5 < FC \leq 2,5$	4	navaho, ngizim, thaï et !xóǝ	25 %
$2,5 < FC \leq 3$	8	afar, français, kannada, kanouri, kwakw'ala, quechua, sora et suédois	50 %
$3 < FC$	2	finnois et yup'ik	12,5 %

Table III.1 : Regroupements mis en évidence en fonction de la forme canonique des langues

La répartition des langues en fonction de la forme canonique est relativement inégale : les langues ayant une forme canonique comprise entre 1,5 et 3 représentent 75 % des langues d'ULSID et le troisième groupe (FC entre 2,5 et 3) contient même à lui seul la moitié de celles-ci. Il ressort que le nombre moyen de syllabes par unité lexicale le plus répandu est compris entre 2,5 et 3. Cette tendance des items lexicaux de notre échantillon est confirmée par des études portant sur d'autres langues. Ainsi, Goldman *et al.* (1996) présentent des données pour le français et l'anglais provenant de deux bases lexicales. Pour ces deux langues, la forme canonique est respectivement de 2,83 et 2,62. Elles renforcent par conséquent le groupe 3, comme l'allemand pour lequel les données proposées par Menzerath (1954) nous ont permis de calculer une forme canonique de 2,63. Le gbaya, pour sa part, rejoint le deuxième type puisque les calculs de Monino et Roulon (1972) donnent une forme canonique de 1,92. Le tableau III.2 tient compte de ces quatre langues supplémentaires dans la répartition en quatre groupes. Les deux types les mieux représentés sont donc toujours ceux pour lesquels la forme canonique se situe autour de 2,5.

Valeurs de la Forme Canonique	Nombre de langues	Langues	% sur 19 langues
$FC \leq 1,5$	2	nyah kur, wa	10,53 %
$1,5 < FC \leq 2,5$	5	navaho, ngizim, thaï, !xóǝ et gbaya	26,32 %
$2,5 < FC \leq 3$	10	afar, français, kannada, kanouri, kwakw'ala, quechua, sora, suédois, anglais et allemand	52,62 %
$3 < FC$	2	finnois et yup'ik	10,53 %

Table III.2 : Regroupements mis en évidence en fonction de la forme canonique des langues d'ULSID et de 3 autres langues

Cette distribution montre des préférences dans la formation des unités lexicales en fonction des langues. Afin de pouvoir étudier plus finement les différentes possibilités, nous avons décidé d'observer la répartition des items lexicaux en fonction du nombre de syllabes pour chacune des 16 langues. On obtient alors des graphiques tels ceux pour l'afar et le finnois (cf. figure III.2) :

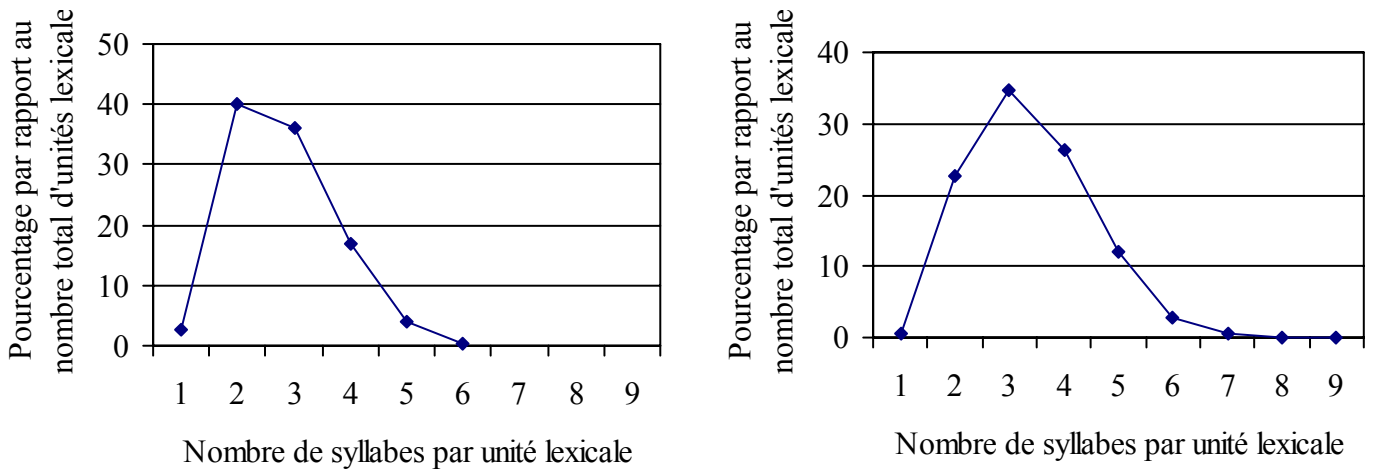


Figure III.2 : Pourcentage d'unités lexicales par nombre de syllabes pour l'afar (à gauche) et le finnois (à droite).

### III.2.1.2. Une typologie des langues basée sur le nombre de syllabes par unité lexicale

La comparaison de la distribution des unités lexicales par nombre de syllabes permet de mettre en évidence, à une exception près, la pertinence du classement obtenu à partir de la forme canonique. En effet, si on excepte le français qui appartient au groupe 4 avec une forme canonique de 2,67, seule une redéfinition des limites de catégories est nécessaire pour obtenir un classement cohérent selon les deux conditions (forme canonique et nombre de syllabes par unité lexicale). Ainsi, nous proposons la séparation entre les groupes 2 et 3 avec une forme canonique de 2,3 au lieu de 2,5 et la frontière des groupes 3 et 4 à partir d'une valeur de 2,9 au lieu de 3. Les types obtenus respectent alors les deux critères observés : la forme canonique d'une langue et la répartition des différentes structures des unités lexicales en fonction du nombre de syllabes. Nous proposons alors la typologie des langues suivante, basée sur la distribution du lexique en fonction du nombre de syllabes par unité lexicale :

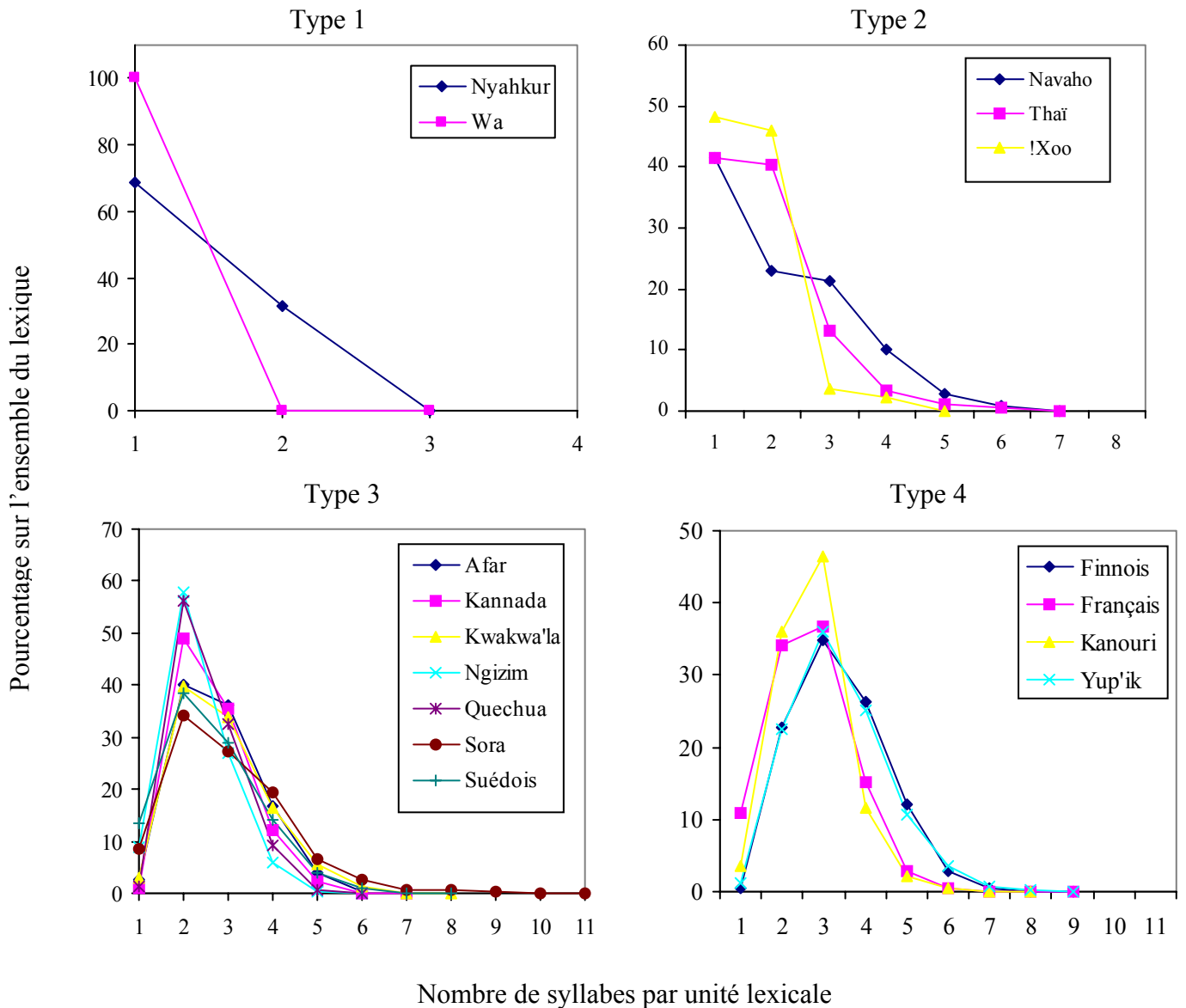


Figure III.3 : Typologie des 16 langues d'ULSID, élaborée sur la répartition du lexique en fonction du nombre de syllabes par item lexical.

- Le type 1 (cf. figure III.3) rassemble les langues totalement ou très majoritairement monosyllabiques (forme canonique proche ou égale à 1) : le wa (100 % de monosyllabes), et le nyah kur (70 % d'unités monosyllabiques et 30 % de dissyllabes). Notons que ces langues ne possèdent pas d'items lexicaux de plus de 2 syllabes, ce facteur caractérisant les langues du type 1 par rapport à celles du type 2 défini ci-après.

Toutes deux attestées en Asie, le wa étant parlé en Chine et le nyah kur au centre de la Thaïlande, ces deux langues sont issues de la famille austro-asiatique, mais appartiennent à deux branches différentes. Le nyah kur est une langue môn et le wa est issu de la branche palaung.

- Le type 2 regroupe les langues qui présentent une distribution des unités lexicales telle que 40 % au moins d'entre elles sont monosyllabiques, entre 20 et 40 % sont dissyllabiques, et entre 10 et 20 % trisyllabiques (navaho, thaï et !xóǝ de forme canonique comprise entre 1,5 et 2,2).

Signalons qu'il s'agit de trois langues génétiquement et géographiquement éloignées : le thaï (famille kam-thaï) est parlé en Asie, le navaho (famille na-déné) est attesté aux États-Unis et le !xóǝ (famille khoisan) est parlé dans le désert du Kalahari (sud de l'Afrique).

- Dans le type 3 ont été classées les langues qui présentent une majorité d'items dissyllabiques, avec moins de 14 % d'unités monosyllabiques et environ un tiers d'entrées lexicales trisyllabiques. Il s'agit du type le plus répandu dans les langues de notre échantillon (formes canoniques entre 2,3 et 3). Quelle que soit la taille du lexique (de 1 989 entrées pour le ngizim à 8 515 pour le suédois), la superposition de ces distributions est remarquable.

Notons également qu'il n'y a pas de lien entre les origines géographiques et génétiques des langues du type 3, puisqu'elles se répartissent en 6 familles, 7 sous-familles et qu'elles apparaissent sur 4 continents différents.

- Le type 4 regroupe le finnois, le français, le kanouri et le yup'ik. Ces langues possèdent très peu d'unités monosyllabiques dans leur lexique, mais ont une majorité d'unités trisyllabiques, 20 à 40 % de dissyllabes, 25% de quadrisyllabes, et de 2 (kanouri) à 12 % (finnois) de quinquasyllabes (formes canoniques supérieures à 3, sauf pour le kanouri avec 2,34).

Ces quatre langues sont issues de quatre familles différentes (ouralo-altaïque, indo-européenne, nilo-saharienne et eskimo-aléoute) et sont parlées dans des régions géographiquement très éloignées (Europe du Nord, France, Afrique Centrale et Alaska).

L'analyse des données de Menzerath (1954) permet de renforcer le type 4, la répartition des unités lexicales de l'allemand étant très proche des quatre langues d'ULSID qui favorise les tri-syllabiques (figure III.4). Les unités lexicales de l'allemand peuvent contenir jusqu'à 9 syllabes, même si ce type d'items ne totalise que 0,03 % des représentants lexicaux, mais les unités trisyllabiques sont majoritaires avec une fréquence d'apparition de plus de 34 %, et les dissyllabes sont toujours très présentes avec plus de 30 %.

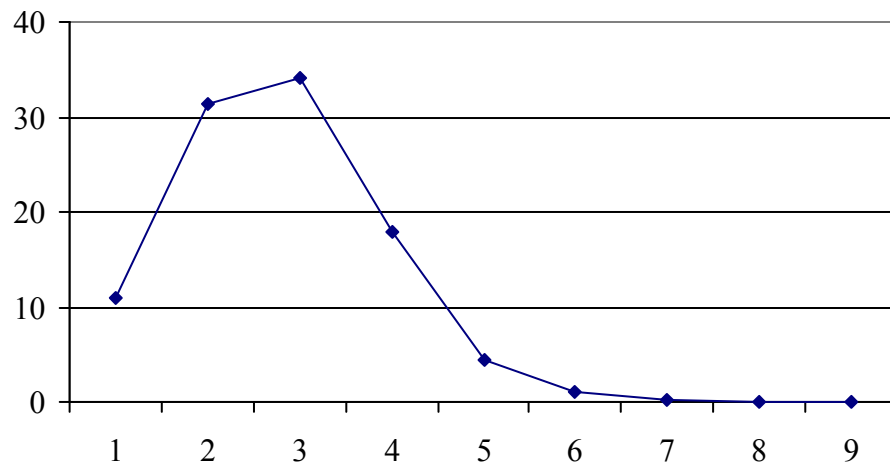


Figure III.4 : Pourcentages d'unités lexicales par nombre de syllabes pour l'allemand (d'après Menzerath, 1954)

La répartition des langues est toutefois inégale dans les quatre types mis en évidence. Comme nous l'avons observé pour le classement des langues à partir de la forme canonique, le groupe 3 est de loin le plus représenté, avec 7 langues sur 16, soit plus de 43 % des langues de notre échantillon. Il semble donc que le type d'organisation le plus fréquent dans les langues soit celui qui privilégie les unités dissyllabiques : les structures à deux syllabes, présentes en quantité non-négligeable dans les types 2 et 4, représentent le type d'unités lexicales le plus répandu dans les langues de notre corpus.

Les unités trisyllabiques sont, elles aussi, très présentes dans les lexiques de cet échantillon puisqu'elles correspondent à plus de 30 % des structures pour ces langues. Les unités d'une seule syllabe apparaissent presque exclusivement dans les langues majoritairement monosyllabiques : leur présence est faible dans les langues des types 3 et 4 (moins de 14 %) et bien qu'elles soient fortement représentées dans les langues des types 1 et 2, elles ne sont que peu représentatives au niveau de l'échantillon total.

En ce qui concerne la quantité d'unités de plus de trois syllabes dans les lexiques, on observe que si les éléments de quatre syllabes sont encore assez fréquents (sauf dans les langues du type 1), les unités de cinq syllabes ne dépassent les 10 % du lexique que dans deux langues : le finnois et le yup'ik (deux langues du type 4). Les unités de six syllabes apparaissent encore dans 13 langues sur 16 même si elles ne représentent jamais plus de 3,7 % des unités lexicales d'une même langue. Les items lexicaux de sept syllabes et plus ne sont présents que dans 10 langues sur 16, mais dans des proportions minimales puisqu'ils n'atteignent jamais 1 % du lexique.



Quelle distribution des unités lexicales est alors observable pour l'ensemble des lexiques ? On remarque des unités dissyllabiques largement majoritaires, les trisyllabiques et monosyllabiques étant encore relativement présentes (respectivement 30 et 15 %). Le fait que 3 des 4 langues majoritairement trisyllabiques possèdent les lexiques les plus importants (finnois, français et kanouri) influe légèrement sur les résultats puisque sans elles, les fréquences d'apparition des unités tri- et monosyllabiques sont inversées (respectivement 23 et 24 %), mais les items dissyllabiques restent toujours majoritaires avec juste un pourcentage légèrement plus important (37 %). Une organisation générale favorisant les unités dissyllabiques puis les items à 1 ou 3 syllabes, en fonction du type de langue, émerge donc de nos données (cf. figure III.5).

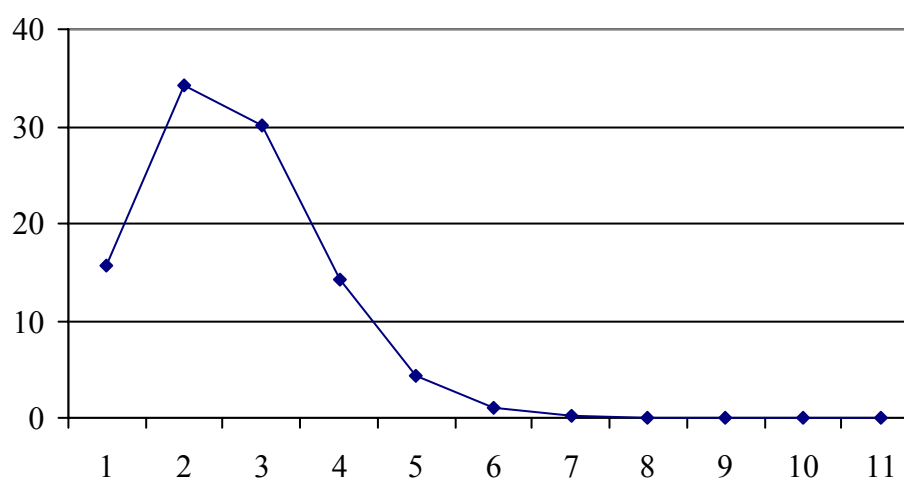


Figure III.5 : Répartition des unités lexicales (en %) par nombre de syllabes pour l'ensemble de nos 16 lexiques.

L'analyse des données recueillies sur la taille des unités lexicales, que ce soit langue par langue ou globalement, montre qu'un patron majoritaire émerge. Le nombre de syllabes par item lexical semble en effet se structurer autour d'un modèle privilégiant les unités dissyllabiques au détriment de toutes les autres. Il semble par conséquent nécessaire de mieux connaître l'influence d'autres facteurs qui pourrait influencer sur la structuration syllabique des lexiques.

### III.2.1.3. Loi de Menzerath : taille des unités lexicales vs. taille des syllabes

Nos premiers résultats révèlent que le nombre de syllabes des unités lexicales d'une langue semble fortement contraint puisqu'on note un lien entre la fréquence d'apparition des items lexicaux et le nombre de syllabes qui les composent. Des études plus anciennes ont porté sur l'existence ou non de relations entre le nombre moyen de syllabes par unité lexicale et diverses autres données de structuration de ces unités. Ainsi, à partir de son analyse des

items lexicaux de l'allemand, Menzerath (1954) propose une loi liant le nombre moyen de syllabes des unités lexicales à la taille des syllabes qui les composent : « *Die relative Lautzahl nimmt mit steigender Silbenzahl ab, oder mit anderer Formel gesagt : je mehr Silben ein Wort hat, um so (relative) kürzer (lautärmer) ist es.* » (Menzerath, 1954, p. 100).

Menzerath (1954, p. 96) propose d'observer la répartition du nombre de phonèmes par syllabes en fonction du nombre de syllabes par unité lexicale. La représentation des données de l'allemand fournies par Menzerath illustre bien sa loi (cf. figure III.6). En effet, le nombre de phonèmes par syllabe a tendance à diminuer lorsque le nombre de syllabes par unité lexicale augmente. Ainsi, ce sont les unités lexicales composées d'une seule syllabe qui contiennent le plus de segments. Le nombre de phonèmes par syllabes connaît une diminution très nette lorsque le nombre de syllabes par unité lexicale est compris entre 2 et 4 et semble se stabiliser autour de 2,3 lorsque les items lexicaux possèdent plus de 4 syllabes.

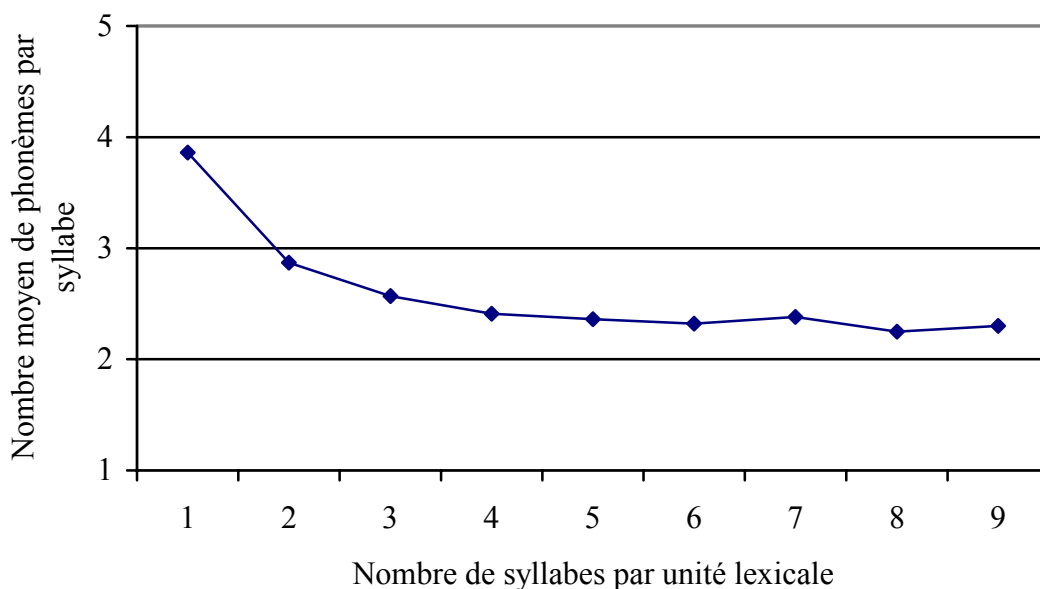


Figure III.6 : Répartition du nombre de syllabes par unité lexicale en fonction du nombre moyen de phonèmes par syllabe pour l'allemand d'après les données de Menzerath (1954, p. 96).

La loi mise en évidence par Menzerath (1954) sera reprise et reformulée, notamment dans le cadre de la linguistique quantitative. Ainsi, Altmann (1980) propose une version générale de la loi : « *The longer a language construct the shorter its components (constituents).* » (Altmann, 1980, p.1), et fournit des données provenant de langues diverses appuyant son existence. Celles-ci portent sur le nombre de phonèmes par syllabe pour 13 444 morphèmes indonésiens et 15 465 010 unités lexicales anglaises (Roberts, 1965), ainsi que sur la durée en millisecondes de 488 items lexicaux du bachka (langue germanique parlée en Yougoslavie). Les figures III.7 et III.8 représentent les résultats obtenus par Altmann dans cette étude.

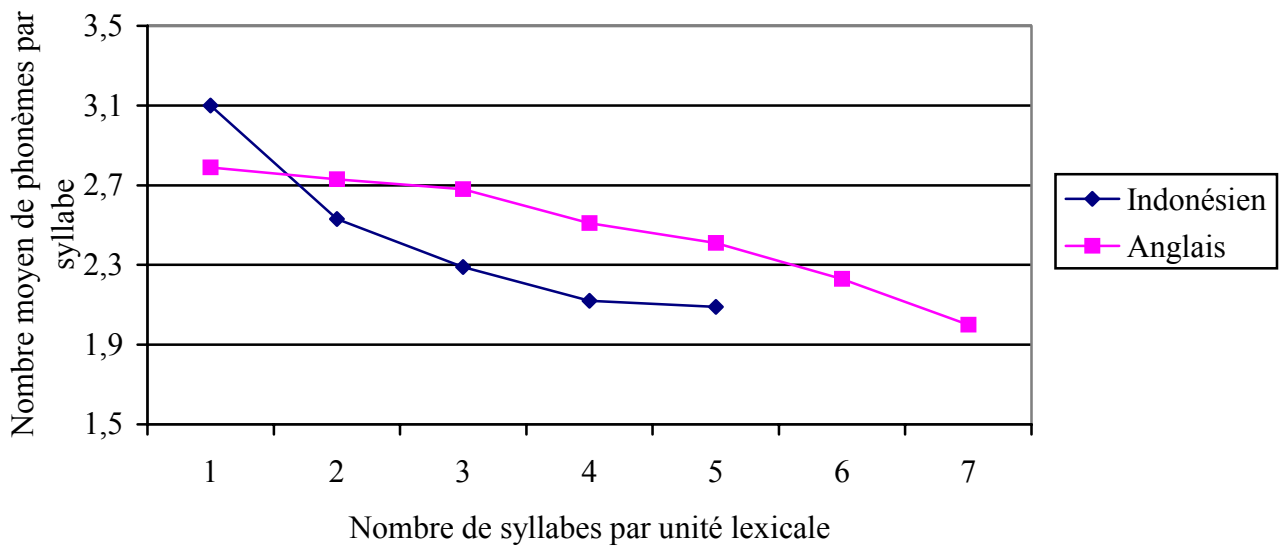


Figure III.7 : Répartition du nombre moyen de phonèmes par syllabe en fonction du nombre de syllabes par unité lexicale d'après les données de Altmann (1980) sur l'indonésien et l'anglais.

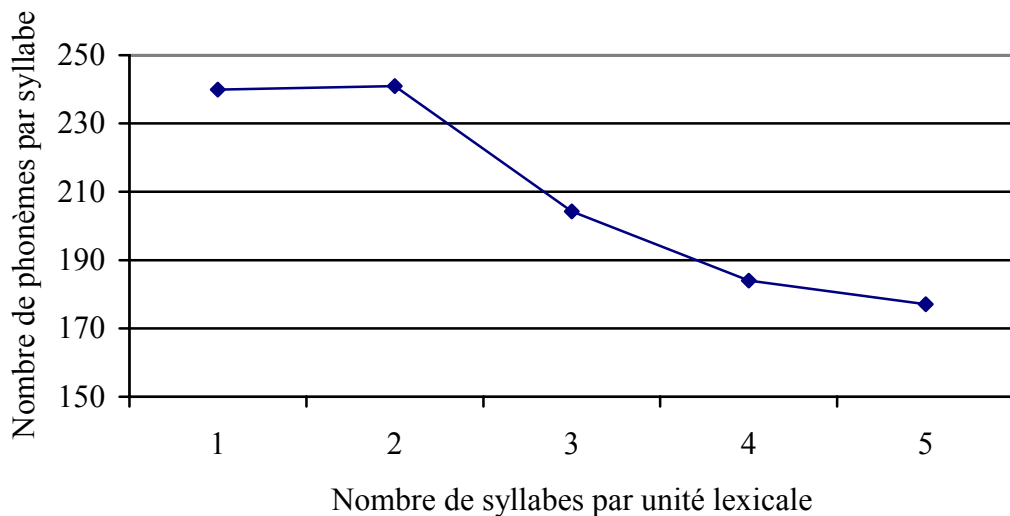
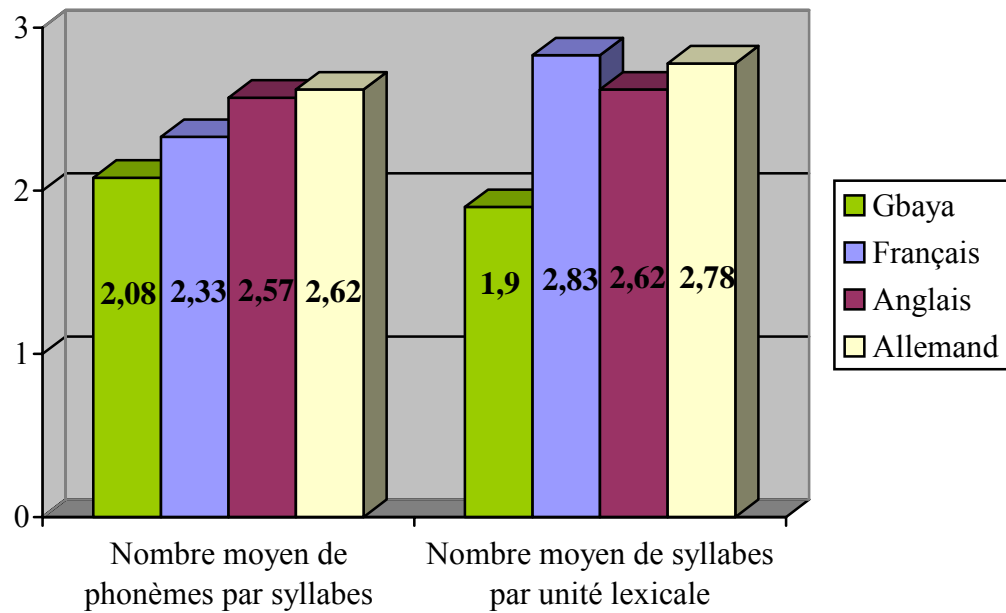


Figure III.8 : Répartition de la durée des syllabes (en millisecondes) en fonction du nombre de syllabes par unité lexicale d'après les données de Altmann (1980) sur la langue bachka.

D'autres études fournissent la possibilité de comparer, disons cette fois de manière générale, le nombre moyen de phonèmes par syllabe pour une langue au nombre moyen de syllabes par unité lexicale. C'est par exemple le cas des travaux de Goldman *et al.* (1996) qui portent sur le français et l'anglais, et une nouvelle fois les données de Menzerath (1954) à propos de l'allemand.



*Figure III.9 : Nombre moyen de syllabes par mot, et nombre moyen de phonèmes par syllabe pour le français et l'anglais (Goldman et al., 1996), l'allemand (Menzerath, 1954) et le gbaya (Monino et Roulon, 1972).*

Les résultats de Altmann (1980) et Menzerath (1954) mettent respectivement en évidence que l'anglais et l'allemand suivent la loi proposée par Menzerath, mais l'étude de la distribution du nombre moyen de phonèmes par syllabe au sein des langues, comparée à la distribution du nombre moyen de syllabes par unité lexicale, ne semble pas confirmer la pertinence de cette loi. Si elle était valable, l'allure croissante pour les données de gauche devrait se trouver décroissante pour celles de droite. Il semble dès lors que cette loi doit se limiter à l'analyse des relations existant entre les deux critères étudiés au sein d'une même langue. Afin de vérifier cette hypothèse, nous avons décidé d'étudier les langues de notre échantillon selon les deux optiques : comparaison des nombres moyens entre les différentes langues, et observation au sein d'une même langue.

La figure III.10 confirme l'hypothèse selon laquelle la loi de Menzerath n'est pas pertinente pour la comparaison, entre différentes langues, des distributions des nombres moyens de phonèmes par syllabe et de syllabes par unité lexicale. Pourtant, une tendance allant dans le sens de la loi semble se dessiner. Il nous a donc paru intéressant de vérifier s'il existe une corrélation statistique entre les deux critères observés dans le cadre de la loi de Menzerath (cf. figure III.10).

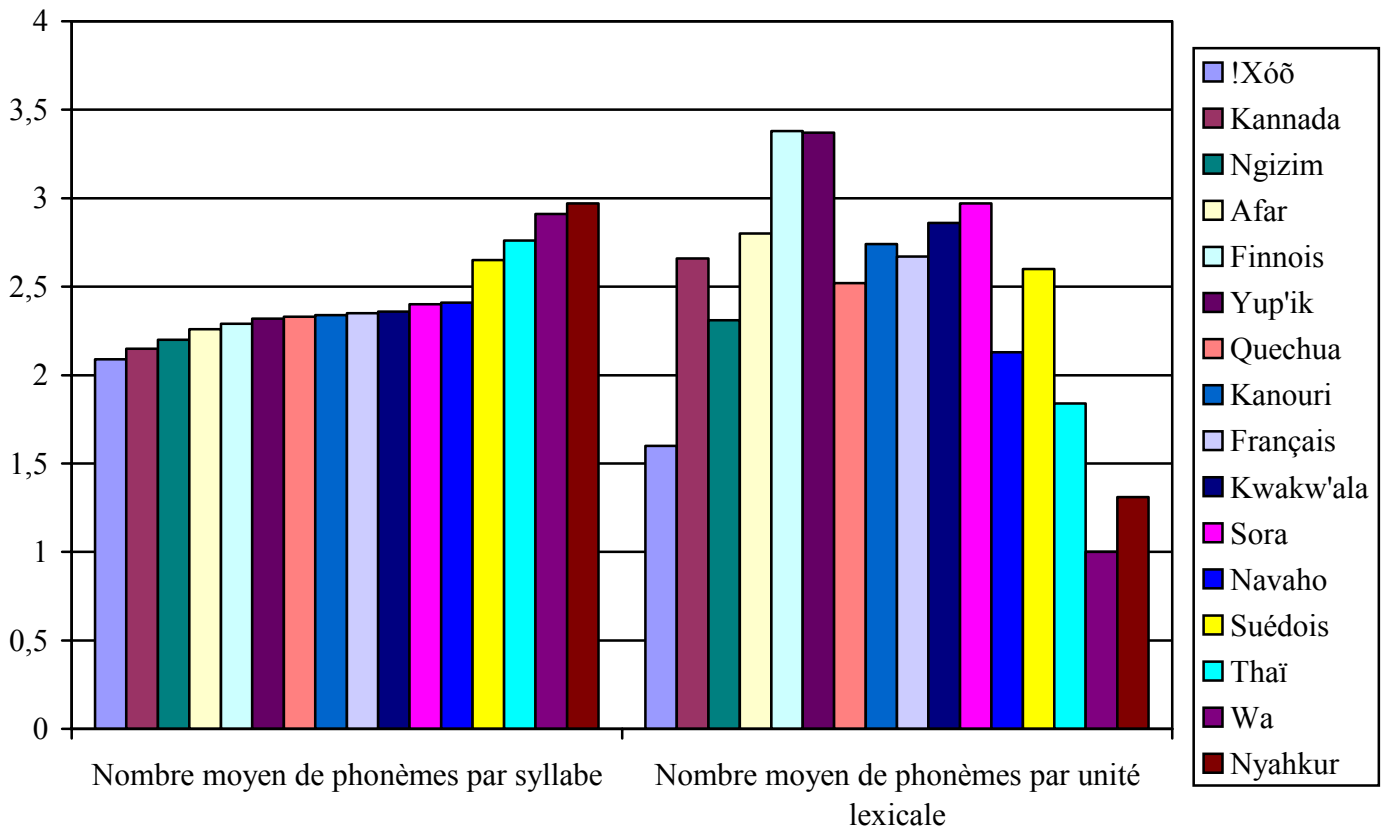


Figure III.10 : Comparaison entre le nombre moyen de phonèmes par syllabe et le nombre de syllabes par unité lexicale pour les 16 langues de notre échantillon.

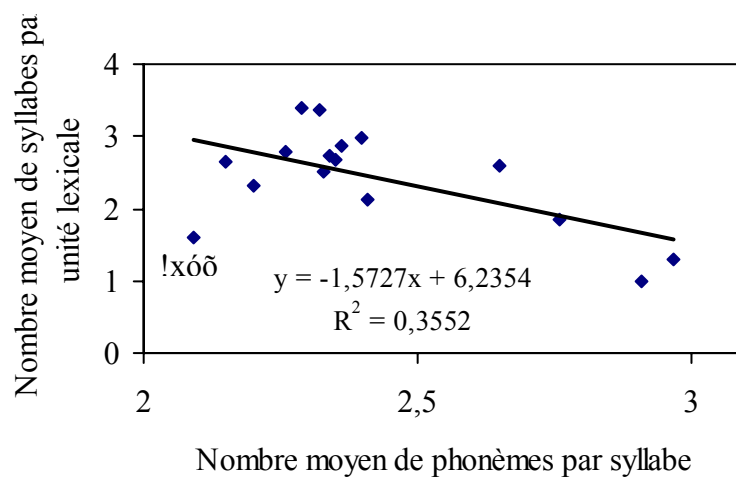
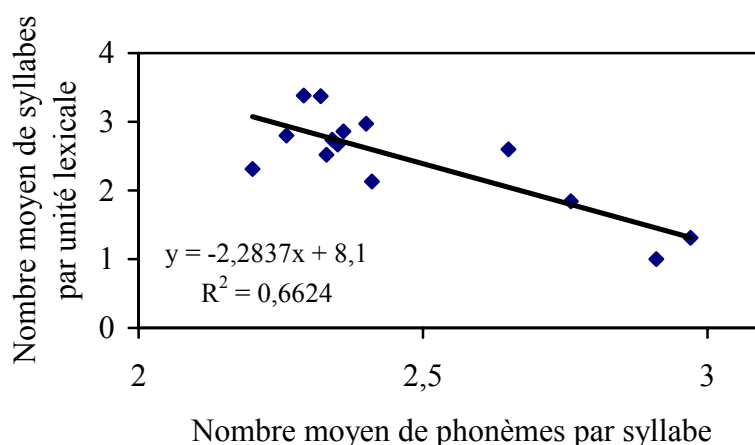


Figure III.11 : Corrélation entre le nombre de phonèmes par syllabe et le nombre de syllabes par unité lexicale dans les 16 langues de notre échantillon.

Le calcul de la corrélation entre le nombre moyen de phonèmes par syllabe et le nombre moyen de syllabes par unité lexicale donne un coefficient de 0,36. Par conséquent, même si une tendance allant dans le sens de la loi de Menzerath apparaît, elle ne semble pas évidente quand on observe statistiquement les données. Le !xóǎ paraît, cependant, à lui seul, "tirer" la corrélation vers le bas, et affaiblir la tendance. Rappelons la particularité de cette langue à disposer d'un inventaire phonologique très important, avec un grand nombre de clicks (5 clicks simples et 78 avec dimensions articulatoires supplémentaires), de consonnes à doubles plosions (9) et de dimensions secondaires pour les voyelles (23 combinaisons de traits supplémentaires possibles). Cette présence importante de phonèmes complexes influe certainement sur le nombre moyen de phonèmes par syllabe (2,09 pour cette langue), sans que cela ne joue sur le nombre moyen de syllabes par mots, assez proche de 1 dans cette langue. En excluant le !xóǎ du calcul de la corrélation (figure III.12), on observe alors que la relation entre le nombre moyen de phonèmes par syllabe et le nombre moyen de syllabes par unité lexicale est plus nette. Nous pouvons donc en conclure qu'il existe une tendance à suivre la loi observée par Menzerath, même si elle ne semble pas avoir de valeur universelle.



*Figure III.12 : Corrélation entre le nombre de phonèmes par syllabe et le nombre de syllabes par unité lexicale dans 15 langues de notre échantillon (le !xóǎ se démarquant du comportement général des langues est ici exclu).*

Le nombre de phonèmes par syllabe est, semble-t-il, à relier au nombre de syllabes par unité lexicale (figure III.12). Nous devrions donc retrouver dans les langues d'ULSID une tendance à la diminution du nombre de phonèmes par syllabes lorsque le nombre de syllabes par unité lexicale augmente. L'étude de la répartition du nombre de phonèmes par syllabes en fonction du nombre de syllabes par unité lexicale peut permettre de vérifier cette hypothèse.

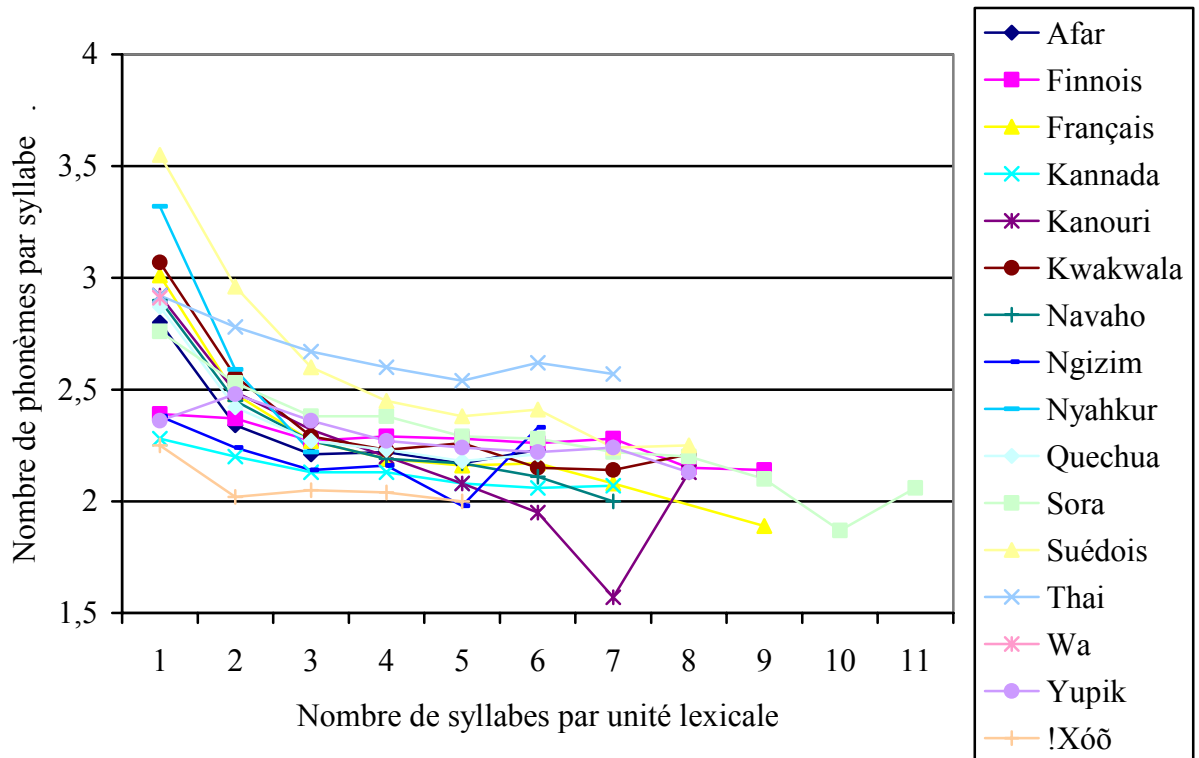


Figure III.13 : Répartition du nombre moyen de phonèmes par syllabe en fonction du nombre de syllables par unité lexicale pour les 16 langues de notre échantillon.

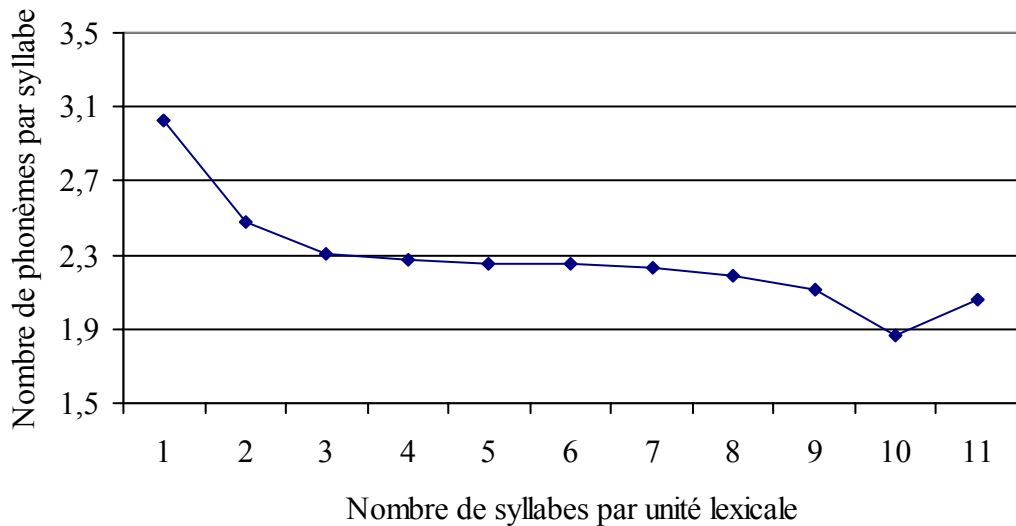


Figure III.14 : Répartition du nombre moyen de phonèmes par syllabe en fonction du nombre de syllables par unité lexicale sur l'ensemble des lexiques d'ULSID.

L'analyse des figures III.13 et III.14 révèle que les 15 langues possédant les unités lexicales de plus d'une syllabe respectent globalement la loi de Menzerath. La diminution du nombre de phonèmes par syllabe est très nette pour les unités contenant moins de cinq syllables (de 3,03 à 2,27) et semble se stabiliser par la suite pour la majorité des langues

(autour de 2,15). Ces résultats vont dans le sens d'une organisation lexicale des langues défavorisant les items trop longs, que ce soit en réduisant le nombre de syllabes par unité ou en diminuant le nombre de phonèmes à l'intérieur des syllabes.

#### III.2.1.4. La typologie confrontée au nombre de phonèmes par syllabe

La vérification de la loi de Menzerath sur les langues d'ULSID permet de confirmer une organisation autour du nombre de phonèmes par syllabe lié au nombre de syllabes par unité lexicale. Or, l'analyse de la répartition du nombre de syllabes par unité lexicale nous a permis de dégager quatre types de langues (cf. paragraphe III.2.1.2.). Il nous a dès lors paru intéressant de comparer cette typologie avec les résultats sur le nombre de phonèmes par syllabe. Le regroupement de nos 16 langues en fonction du nombre de phonèmes par syllabe permet de dégager trois types d'organisations (A, B et C) assez dissemblables les uns des autres. Comme les deux critères sont liés, d'où proviennent les différences ?

Les langues des types 1 et 2 de la typologie précédente sont majoritairement ou fortement monosyllabiques. Elles devraient donc, d'après la loi de Menzerath, contenir une forte proportion de syllabes constituées d'un nombre élevé de phonèmes. Or, si cette observation est vraie pour les deux langues majoritairement monosyllabiques (le nyahkur et le wa), seul le thaï (type 2) appartient au type A caractérisé par une distribution du nombre de phonèmes par syllabes présentant un pic à 3 (figure III.15). Le navaho et le !xóǝ, langues possédant pourtant des taux d'unités monosyllabiques de plus de 40 % et que notre typologie associait au type 2, appartiennent au type B décrit ci-dessous, où le pic de la distribution est clairement situé au niveau de 2 phonèmes par syllabes (cf. figure III.16).

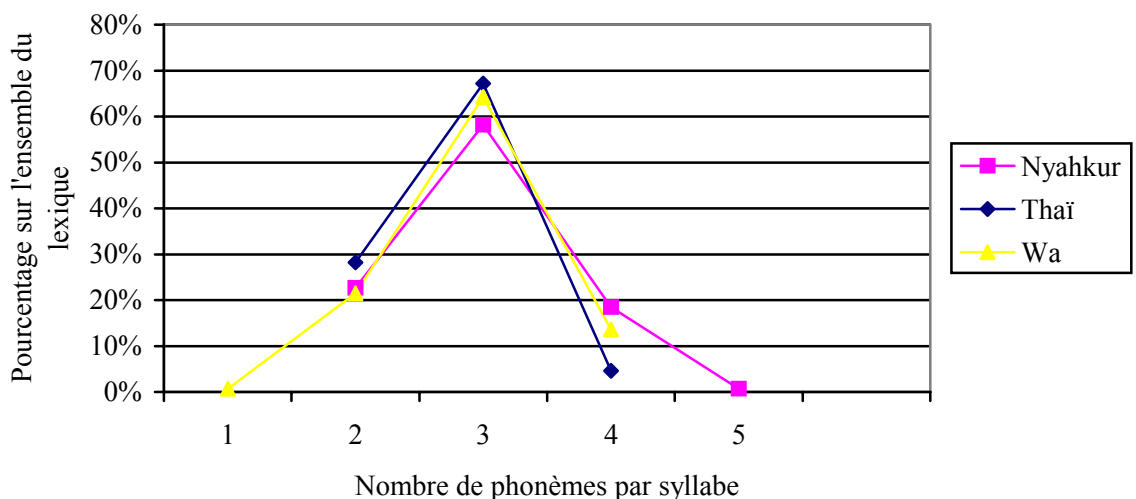


Figure III.15 : Proportion de syllabes (en pourcentage) par nombre de phonèmes pour les langues du type A.

Le type B contient, en plus des deux langues fortement monosyllabiques, la majorité des langues dissyllabiques ainsi que trois des quatre langues majoritairement trisyllabiques. Il



semble donc, comme pour le nombre de syllabes par unité lexicale, que l'organisation interne des syllabes se fasse autour d'un patron comprenant majoritairement deux phonèmes. C'est de très loin le type le plus répandu, puisqu'il contient 10 langues sur 16, soit plus de 60 % de notre échantillon.

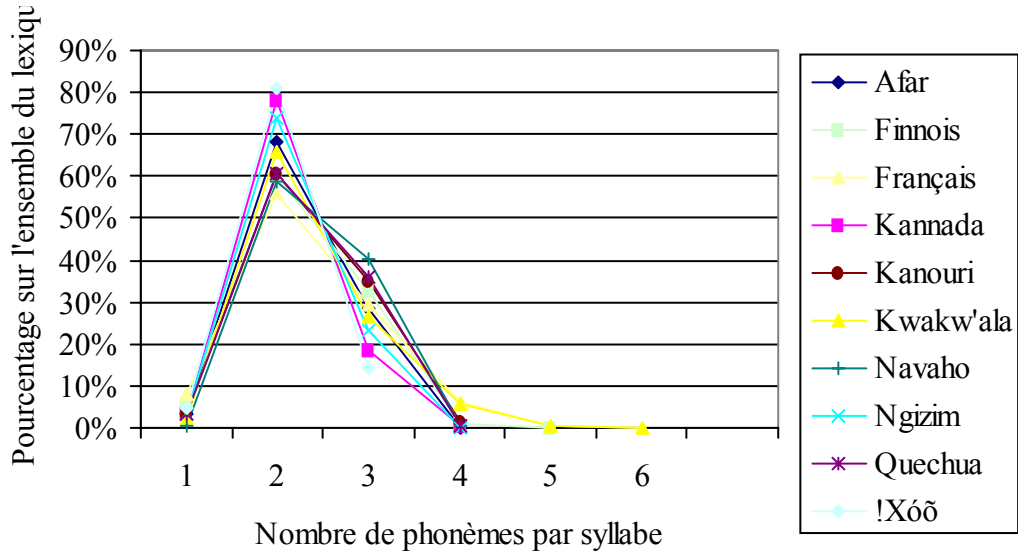


Figure III.16 : Proportion de syllabes (en pourcentage) par nombre de phonèmes pour les langues du type B.

Le troisième type (type C) contient, pour sa part, deux langues dissyllabiques et une trisyllabique (cf. figure III.17). Pour ces trois langues, les pourcentages de syllabes de 2 et de 3 phonèmes sont très proches et représentent entre 79 et 92,4% de l'ensemble des syllabes. Il est intéressant de remarquer que, bien que possédant des taux d'apparition très proches, les syllabes de 2 phonèmes sont toujours les plus fréquentes.

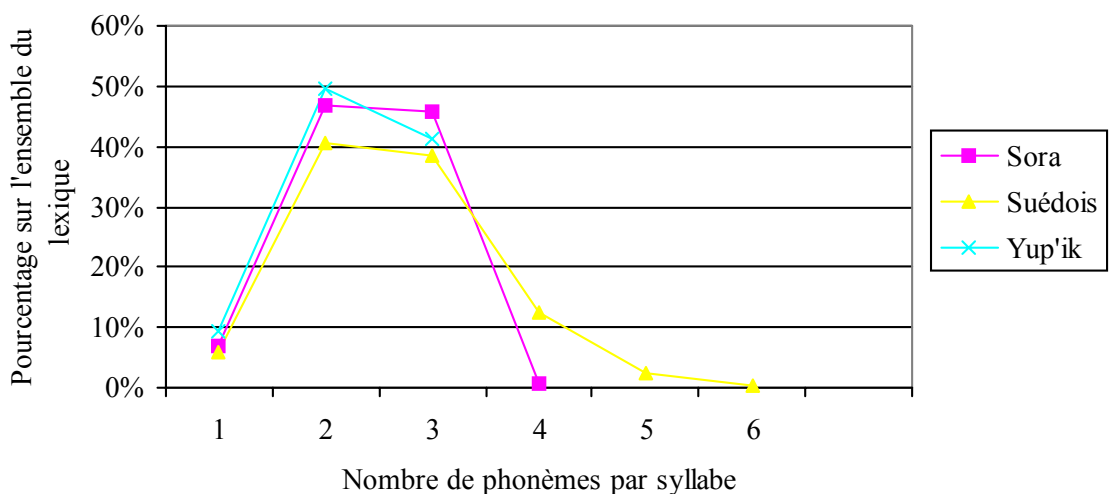


Figure III.17 : Proportion de syllabes (en pourcentage) par nombre de phonèmes pour les langues du type C.

Cette deuxième typologie fait apparaître un patron d'organisation lexicale majoritaire lié au nombre de phonèmes qui composent les syllabes en fonction du nombre de syllabes dans l'unité lexicale. Hormis pour les langues totalement ou fortement monosyllabiques, les langues paraissent privilégier les syllabes de 2 phonèmes tout comme elles favorisent les unités lexicales de 2 syllabes. Il est aussi intéressant de noter que les syllabes de plus de trois phonèmes sont très peu fréquentes et ce quel que soit le type de langues (entre 10 et 20 % pour trois langues, moins de 8 % pour les 13 autres, et même moins de 1 % pour dix d'entre elles). Il semble, par conséquent, que l'organisation des syllabes en phonèmes soit, elle aussi, soumise à certaines contraintes qui mériteraient d'être explicitées.

### III.2.1.5. Taille des unités lexicales basée sur le nombre de phonèmes qu'elles contiennent

Les résultats précédents montrent que l'organisation des unités lexicales en syllabes et de ces dernières en phonèmes ne semble pas due au hasard et que certaines tendances générales émergent. La répartition des phonèmes dans les unités lexicales devrait donc découler de l'application par les langues de ces tendances. Une des implications de la loi de Menzerath (1954) concerne la limitation de la longueur totale du mot dans les langues, même si les constituants qui le composent varient. Dans le prolongement de cette implication, nous nous sommes intéressés au nombre total de phonèmes par item lexical.

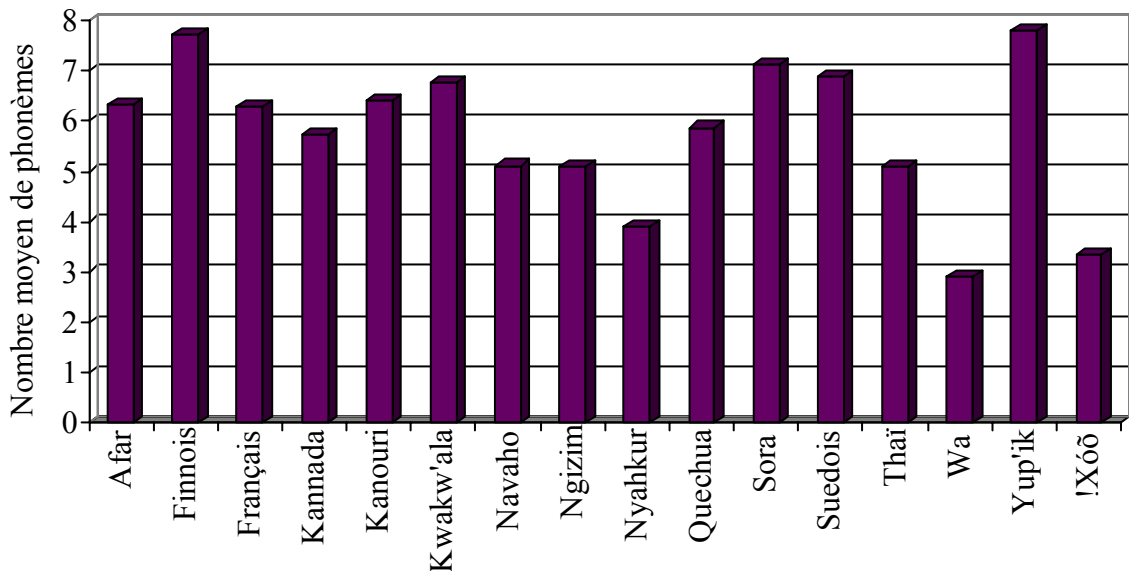


Figure III.18 : Nombre moyen de phonèmes par unité lexicale pour chacune des 16 langues d'ULSID.

Le nombre moyen de phonèmes par unité lexicale varie entre 2,91 pour le wa et 7,81 pour le yup'ik. Les cinq langues majoritairement ou fortement monosyllabiques (types 1 et 2) sont en dessous de 5,5. Les langues des types 3 et 4 (majoritairement di- et trisyllabiques) se répartissent indifféremment entre 5,09 et 7,91. Ce regroupement entre deux comportements permet de formuler plusieurs hypothèses quant à l'organisation des unités lexicales en phonèmes et en syllabes. Ainsi, le fait que le nombre moyen de phonèmes par item lexical soit inférieur à 8 renforce la loi de Menzerath : les langues acceptant un grand nombre de syllabes par unité lexicale pondèrent la longueur de ces dernières par la longueur des syllabes, limitant de la sorte la taille des items. Pour ce qui est du plus faible nombre de phonèmes par item lexical dans les langues majoritairement ou fortement monosyllabiques, il permet de supposer qu'indépendamment de la taille de l'unité lexicale, la taille des syllabes est, elle aussi, contrainte. L'application de la loi de Menzerath prédit des syllabes plus longues lorsque les unités lexicales sont plus petites, mais on remarque que pour le wa, langue strictement monosyllabique, le nombre moyen de phonèmes par unité lexicale ne dépasse pas 3. Rappelons que, d'une manière générale, les syllabes de plus de trois segments sont très minoritaires dans les langues de notre échantillon (moins de 1 % pour 10 langues sur 16 et moins de 20 % pour les 6 autres).

### **III.2.2. Structuration syllabique des unités lexicales**

Les différents résultats sur la taille des unités lexicales montrent qu'il est possible de dégager des tendances organisationnelles au niveau lexical. À un autre niveau, il semble intéressant de rechercher si des tendances existent au niveau de la concaténation des syllabes qui forment les items lexicaux : existe-t-il, au niveau de la structure syllabique des lexiques, des patrons lexicaux favorisés ? Les langues favorisent-elles certaines successions de structures syllabiques ou utilisent-elles les très nombreuses possibilités de combinaisons qui s'offrent à elles ?

Afin de répondre à ces questions, nous avons codé les différents lexiques de nos 16 langues en deux constituants C (consonne) et V (noyau vocalique), et obtenu ainsi la liste des gabarits lexicaux possibles (sorte de squelette lexical) pour reprendre une terminologie de la théorie phonologique (cf. partie I.1.3.).

La première remarque porte sur la variabilité observée quant au nombre de gabarits attestés dans les langues (cf. table III.3). Il évolue en effet entre 6 pour le wa et 1 208 pour le suédois, avec une moyenne de 285 pour les 16 langues.

Type	Langue	Nombre de structures lexicales
1	nyahkur	19
	wa	6
2	navaho	73
	thaï	119
	!xóǝ	25
3	afar	119
	kannada	89
	kwakw'ala	189
	ngizim	67
	quechua	113
	sora	495
4	suédois	1208
	finnois	583
	français	949
	kanouri	210
	yup'ik	293

*Table III.3 : Nombre de types de gabarits lexicaux pour chacune des langues de notre échantillon, classées selon la typologie en quatre types définie au paragraphe 2.1.*

La comparaison du nombre de gabarits lexicaux de chaque langue en fonction du type auquel elle appartient (rappelons que le type définit la répartition du lexique en fonction du nombre de syllabes par unité lexicale) montre que les regroupements proposés pour notre première typologie restent assez cohérents. En effet, les langues majoritairement monosyllabiques (types 1 et 2), ont le nombre de gabarits le plus faible (compris entre 6 et 119), possèdent des items lexicaux moins longs et par conséquent auront moins de possibilités de combinaisons différentes que les langues di- et trisyllabiques, qui possèdent un potentiel gabaritique plus élevé (supérieur à 110 pour 9 langues sur 10).

La présence d'une aussi grande diversité au niveau du nombre des gabarits nous a incité à nous pencher sur les formes des gabarits les plus fréquents dans les différentes langues de l'échantillon étudié. Pour cela, nous avons établi pour chaque type de langue, la liste de tous les gabarits représentant plus de 2 % des unités lexicales, en les classant par langue en fonction du nombre de syllabes (pour plus de détails, voir tableau annexe III).

Il est intéressant de remarquer que les mêmes gabarits apparaissent dans la majorité des langues. Ainsi, la structure monosyllabique CVC est le gabarit le plus fréquent pour les langues des types 1 et 2 (entre 30 et 60 % des unités), à l'exception du !xóǝ, pour laquelle ce type de gabarit se situe au troisième rang par ordre de fréquence décroissante (13 % des items). Pour ce qui est des langues des groupes 3 et 4, les gabarits les plus fréquents sont les combinaisons des deux structures syllabiques CV et CVC, avec une légère préférence pour

CV.CV (présent dans les gabarits de plus de 2 % dans 10 langues sur 11, cf. tableau annexe III).

Il ressort que les langues ont des préférences structurelles pour leurs gabarits lexicaux et donc qu'elles ne se servent pas de toutes les possibilités qui leur sont offertes par le nombre de combinaisons de structures syllabiques possibles pour former des unités lexicales. L'étude des gabarits les plus fréquents dans l'ensemble des lexiques le montre, puisque 41 % des gabarits présents dans l'ensemble des lexiques se répartissent en seulement 7 structures (cf. table III.4).

Gabarits	Fréquence d'apparition dans l'ensemble des lexiques
CVC	8,80 %
CV.CV	7,79 %
CV.CVC	6,85 %
CV.CV.CV	5,74 %
CVC.CV	5,44 %
CVC.CVC	4,60 %
CV	2,62 %

Table III.4 : Structures représentant plus de 2 % des gabarits de l'ensemble des 16 lexiques.

Cette analyse des gabarits les plus fréquents dans les lexiques permet de mieux connaître l'organisation syllabique des unités. La première remarque porte sur la place des unités monosyllabiques, en première et en septième positions des gabarits les plus fréquents bien qu'elles ne représentent qu'environ 15 % des unités lexicales de nos lexiques. Cette forte fréquence des éléments monosyllabiques est due à la présence de cinq langues favorisant ce type d'items, ainsi qu'à la forte proportion de ces structures dans les items monosyllabiques des autres langues de notre échantillon. La diversité possible des unités monosyllabiques étant moins grande que pour les structures plus longues, il n'est donc pas étonnant que les structures CVC et CV se retrouvent parmi les gabarits les plus fréquents sur l'ensemble de nos lexiques.

Le gabarit favori des unités monosyllabiques est la structure fermée CVC (plus de 34 % des structures dans 15 langues sur 16), qui est la structure lexicale la plus fréquente dans 10 langues. Le recours majoritaire à la structure fermée dans les unités monosyllabiques renforce les conclusions obtenues avec l'application de la loi de Menzerath (cf. partie III.2.1.3.).

Autre remarque : la faible présence des groupements consonantiques dans ces sept types les plus fréquents. Ils n'apparaissent jamais à l'intérieur d'une syllabe et dans seulement deux structures au niveau inter-syllabique. Il sera donc intéressant de cerner de manière plus précise quelle place occupent les groupements consonantiques dans l'organisation syllabique des langues, que ce soit au niveau inter ou intra-syllabique.

### III.2.3. Confirmation de la taille des unités lexicales comme critère de classement typologique

Les différentes analyses menées sur les unités lexicales des langues montrent une organisation structurale autour de tendances qui ne confortent qu'en partie la typologie des langues proposée à la partie III.2.1.2. Le classement des langues effectué à partir du nombre de syllabes par unité lexicale, s'il nous apporte des informations importantes sur la structuration syllabique des lexiques, ne permet pas de rendre compte finement d'autres facteurs que nous avons étudiés. Une redéfinition des types afin d'obtenir un classement qui tienne compte de la majorité des données dont nous disposons sur l'organisation syllabique des unités lexicales semble nécessaire.

Le critère qui semble le plus pertinent pour classer les langues dans des types en fonction de l'organisation syllabique semble être la longueur dominante au niveau des unités lexicales, sans tenir compte de la fréquence des différentes structures.

Dominance	Langue	Forme canonique	Nombre de phonèmes par unité lexicale	Nombre de phonèmes par syllabe	Gabarits les plus fréquents
1 et 2 monosyllabique	nyah kur	FC < 2,2	Nb < 5,13	Nb > 2,65 (2 exceptions)	CVC CV CV.CV
	wa				
	navaho				
	thaï				
	!xóõ				
3 dissyllabique	afar	2,2 ≤ FC FC < 3	5,10 ≤ Nb Nb < 7,2	2,35 ≤ Nb Nb ≤ 2,65	CV.CV CV.CVC CVC.CV CVC.CVC
	kannada				
	kwakw'ala				
	ngizim				
	quechua				
	sora				
	suédois				
4 trisyllabique	finnois	3 ≤ FC	7,2 ≤ Nb (2 exceptions)	Nb < 2,35	
	français				
	kanouri				
	yup'ik				

Table III.5 : Récapitulatif des analyses précédentes.

La table III.5 montre bien que le critère retenu pour l'établissement de notre typologie (répartition des unités lexicales en fonction de leur nombre de syllabes) est pertinent puisque l'ensemble des autres analyses que nous avons effectuées sur nos 16 langues confirment ce classement en trois grands groupes (après regroupement des types 1 et 2, pour lesquels les unités majoritaires sont les monosyllabes).

### III.2.4. Vers des éléments d'explications des tendances dans l'organisation structurale des lexiques

L'étude de la structuration des unités lexicales en syllabes ainsi que des syllabes en phonèmes dans nos seize langues montre qu'une organisation majoritaire émerge des données. L'existence d'un patron privilégiant les unités lexicales et syllabiques contenant deux constituants se retrouve à travers un certain nombre d'études portant sur divers aspects du langage humain.

Tout d'abord, au niveau des unités lexicales, l'organisation majoritairement dissyllabique des langues peut être reliée à l'existence de la boucle articulatoire, ou boucle phonologique (Baddeley, 1990). Cette dernière renvoie à la mémoire de travail qui intervient dans l'encodage de la longueur phonologique. Cette boucle reliant différentes aires du cerveau ne permet d'encoder qu'une durée phonologique de 1,6 secondes à la fois, ce qui pourrait expliquer la désaffection que connaissent les unités de plus de trois syllabes, puisque « *la longueur de l'empan mnésique est profondément liée à la durée de prononciation des items à mémoriser [...]. Ce phénomène est connu sous le nom d'effet longueur de mot (worth length) et est traditionnellement associé au processus de contrôle articulatoire.* » (Sato, 2001, p. 11).

De plus, l'analyse commune de données provenant d'études sur l'acquisition des contrôles moteurs nécessaires à la mise en place de la parole et d'études sur l'acquisition des contrôles moteurs des segments du corps nécessaires à la mise en place du mot chez l'enfant montre l'existence d'un « rendez-vous développemental » entre ces deux phénomènes au niveau de l'acquisition de la coarticulation, qui constitue le contrôle de base de la parole, en parallèle avec celle des premiers mots de l'enfant. De la même façon qu'on peut dégager le « pas de la parole », c'est-à-dire le rythme syllabique à 3 Hz, Ducey et Abry (2004) ont cherché « le pas du signe ». Ils utilisent pour cela l'unité métrique qu'est le pied, qui fournit un gabarit pour intégrer la syllabe dans ce pas du signe, constitué d'un « stroke », c'est-à-dire une action commençant par une phase d'initiation d'un geste de pointage jusqu'au bras et index tendus. En effet lors des premiers mots d'enfants, ces mots « prosodiques » varient d'une à deux syllabes.

Ainsi, on peut intégrer « deux cycles de la mandibule (le babillage canonique peut se faire à environ 3 Hz, soit 300-350 ms) dans un pied (1.5 Hz, soit 600-700 ms) » (Ducey et Abry, 2004, p. 4). En d'autres termes, Ducey et Abry (2004) remarquent et soulignent : « *Ce que*

*l'enfant entendant ou sourd finit par faire avec la mandibule (cadre de la parole), il l'intègre dans le bras qui pointe, l'index de la main tendu, soit dans le cadre du signe, en coordination avec la prosodie vocale : autrement dit le mot* » (Ducey et Abry, 2004, p. 4). Il semble donc qu'au niveau de l'acquisition des premiers mots, les unités dissyllabiques bénéficient d'une préférence du fait de leur lien avec le geste de pointage. Or, MacNeilage (1998) postulant que les contraintes pesant sur l'ontogenèse influent fortement sur les structures présentes dans les langues naturelles, ce phénomène peut renforcer les résultats déjà existants validant l'hypothèse d'une organisation des unités lexicales autour d'un patron non universel mais très fréquent d'une longueur relativement constante, et correspondant à la taille d'un dissyllabe CV.CV. La dominance des patrons structurels lexicaux CVC dans les langues totalement ou majoritairement monosyllabiques trouve peut-être là aussi son explication.

Les observations à propos de la diminution de la taille des constituants quand leur nombre augmente au sein des unités du langage est en accord avec les études de Redford qui postule : « *Although the relative duration of the cycle may change depending on the amplitude of the cycle, the change is moderated by the fact that the jaw appears to have a preferred oscillation frequency of about 4 cycles per second (e.g., Ohala, 1975; Nelson, Perkell, Westbury, 1984). This preferred oscillation frequency may provide a further temporal constraint on the articulations segments within the cycle so that segment duration will be inversely correlated with the number of segments articulated within a single syllable.* » (Redford, 1999, p. 32). Cette affirmation rejoint la loi proposée par Menzerath et vérifiée, pour notre échantillon, au niveau du nombre de phonèmes par syllabe en fonction du nombre de syllabes par unité lexicale.

La mise en relation de l'ensemble de ces résultats permet de mieux comprendre la part des différentes contraintes pesant sur l'organisation syllabique des langues. Ainsi, les données de l'ontogenèse et les résultats récents portant sur l'acquisition des contrôles moteurs (cf. partie I.2.) confortent la voie ouverte par MacNeilage et Davis (2000) quant à la position prédominante du cycle mandibulaire dans l'organisation des structures lexicales et syllabiques. De plus, les investigations sur la boucle articulatoire, visant dans un premier temps à mieux comprendre le rôle des différentes aires du cerveau dans le cadre de la parole, permettent de justifier les hypothèses faites quant à la taille des unités.

### **III.3. ORGANISATION DES STRUCTURES SYLLABIQUES**

Un de nos précédents résultats montre que les variations du nombre moyen de phonèmes par syllabe sont peu importantes entre les 16 langues d'ULSID (min : 2,09 pour le !xóǝ ; max : 2,97 pour le nyah kur). Nous nous proposons dans cette partie d'examiner comment sont organisées les syllabes, dans les différentes langues de notre échantillon, au niveau de leurs constituants consonantiques et vocaliques. Pour ce faire, nous avons réalisé le regroupement de toutes les syllabes en fonction du caractère vocalique (V) ou consonantique



(C) des différents constituants. Les représentations des unités lexicales par leurs structures syllabiques en C et V constituent ce que nous appellerons désormais les cohortes syllabiques.

### III.3.1. Les types de structures syllabiques

La recherche des différentes cohortes présentes dans les langues de notre échantillon nous a permis de relever un nombre de types syllabiques fortement limité : de 4 types différents pour le thaï et le yup'ik à 19 structures différentes pour le suédois. Le nombre moyen est de 9 types par langues, ce qui est très faible quand on le compare au nombre de syllabes différentes répertoriées dans les langues : en moyenne 15 453 syllabes par langue, et une variation comprise entre 3 180 pour le wa et 60 994 pour le français.

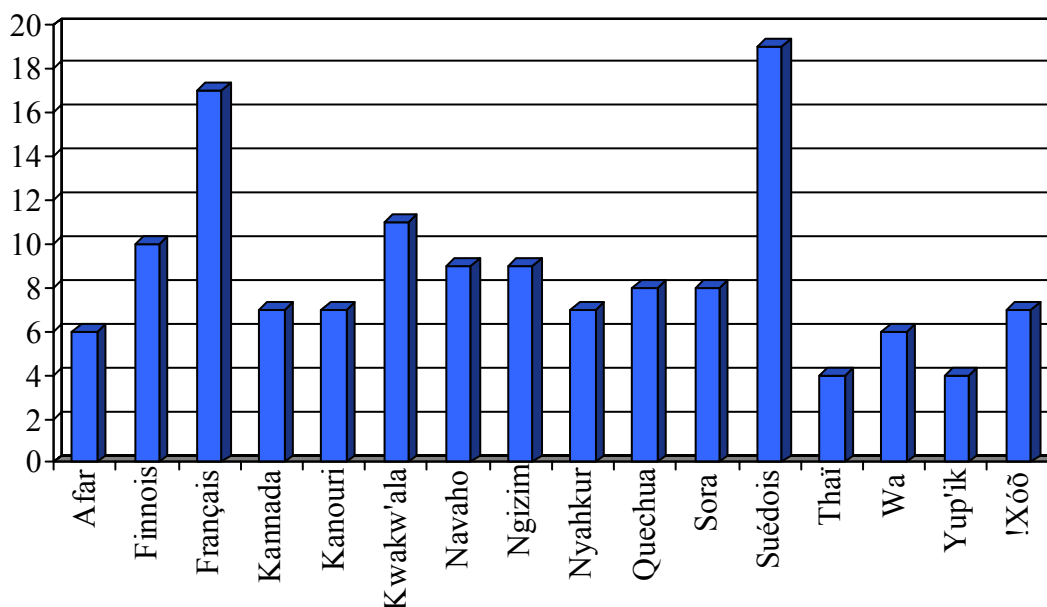


Figure III.19 : Nombre de types de structures syllabiques différents dans chacune des langues d'ULSID.

Il est important de préciser que le nombre de types de structures syllabiques d'une langue n'est corrélé ni à la taille des lexiques ( $r^2 = 0,10$ ), ni au nombre de syllabes ( $r^2 = 0,0002$ ), ni au nombre de phonèmes ( $r^2 = 0,005$ ) de la langue concernée. Les trois langues d'ULSID qui possèdent 7 structures syllabiques différentes illustrent bien cette tendance. Ainsi, au niveau de la taille du lexique, le kannada et le nyah kur, par rapport au kanouri, possèdent un nombre d'unités lexicales assez différent, respectivement 4 559, 4 188 et 7 211. Le kanouri possède presque deux fois plus d'items lexicaux que les deux autres langues pour le même nombre de types de structures syllabiques. On constate aussi que le nyah kur possède beaucoup moins de syllabes différentes (5 503) que le kanouri (19 000) et le kannada (plus de 12 000). En ce qui

concerne le nombre de phonèmes, il varie pour ces trois langues entre 28 et 54, et n'influe pas non plus sur le nombre de types syllabiques.

La totalité des syllabes de notre corpus, pour les 16 lexiques, livrent un éventail de 19 cohortes possibles, c'est-à-dire 19 types de structures syllabiques. Signalons au passage que les 12 types de l'indonésien présentés par Zörnig et Altmann (1993) se retrouvent dans notre liste (cf. table III.6) :

V	CV	CCV	CCCV	C
VC	CVC	CCVC	CCVC	CC
VCC	CVCC	CCVCC	CCCVCC	
VCCC	CVCCC	CCVCCC	CCCVCCC	
	CVCCCC			
<b>Attaque vide</b>	<b>Attaque simple</b>	<b>Attaque complexe</b>		<b>C syllabique</b>

*Table III.6 : Les 19 types syllabiques d'ULSID classés en fonction de la complexité de l'attaque.*

Les différents types sont analysables en fonction des degrés de complexité de leur structure. Ils peuvent contenir jusqu'à 3 positions consonantiques en attaque et 4 en coda. Les syllabes fermées présentent une diversité beaucoup plus importante : 13 types contre 4 pour les syllabes ouvertes (les types C et CC mis à part). Cette disparité est explicable puisqu'une syllabe ouverte n'ayant par définition pas de coda, seule la complexité de l'attaque peut varier. Or, le nombre de consonnes en attaque de syllabe ne dépassant pas trois d'après notre corpus, il n'est pas possible de trouver plus de quatre types de syllabes ouvertes. Les types fermés sont plus nombreux puisque la coda peut contenir jusqu'à quatre positions consonantiques, et que la combinaison avec les différentes attaques existantes donne à 16 structures possibles parmi lesquelles nos langues en ont choisi 13.

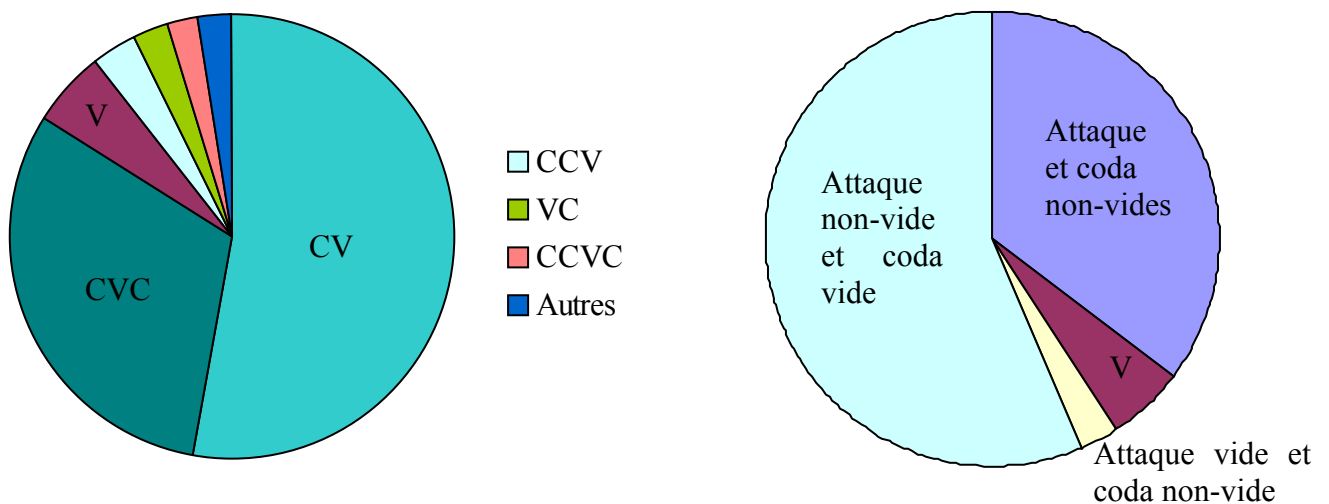
Une observation plus fine des types syllabiques attestés dans les 16 langues de l'échantillon montre des tendances fortes dans l'apparition des différentes structures dans les lexiques qui laissent à penser qu'un certain nombre de facteurs contraignent l'organisation syllabique des langues.

### **III.3.2. Fréquences des types syllabiques dans les lexiques**

Le nombre d'occurrences de chaque type syllabique dans l'ensemble des lexiques permet de mettre en évidence des structures plus favorisées.

Type	0/000	Type	0/000
CV	5299	CVCCC	9
CVC	3119	CCCVC	8
V	539	CCCVCC	2
CCV	338	C	2
VC	254	CCVCCC	1
CCVC	214	VCCC	0,7
CVCC	164	CVCCCC	0,2
VCC	21	CC	0,2
CCVCC	20	CCCVCCC	0,04
CCCV	9		

Table III.7 : Liste et nombre d'occurrences des types de structures syllabiques rencontrées dans les 16 langues d'ULSID (classés par fréquence décroissante).



Figures III.20 : Proportions des différentes types syllabiques (à gauche), regroupés en fonction de la présence ou non de l'attaque et de la coda (à droite).

L'observation de la fréquence des types syllabiques permet de faire émerger les deux types dominants : CV et CVC qui totalisent à eux seuls près de 85 % des syllabes de nos langues. Les structures CV constituant à elles seules plus de 50 %, elles représentent le type syllabique le plus recruté par les langues. Cette nette domination de la structure CV confirme les résultats de toutes les études antérieures qui posent CV comme structure universelle (MacNeilage, 1998).

Des études précédentes sur différentes langues livrent des proportions comparables aux nôtres. Ainsi, Molinu (1999) montre que dans un parler roman d'Italie du Sud, 2 structures

(CV et CVC) représentent 86 % des syllabes. Des résultats analogues sont observés par Goldman (1996) : en français, 4 structures (CV, VC, CVC, CCV) représentent 91 % des syllabes ; en anglais, 4 structures (CV, CVC, CCV, VC) représentent 76,6 % des syllabes.

Si on observe ces données en fonction de la présence ou de l'absence de consonnes en attaque et en coda, on obtient un classement encore plus net, dans lequel les structures avec attaque pleine sont largement favorisées par rapport aux autres types syllabiques avec ou sans coda (figure III.20 à droite). Ce résultat peut être mis en relation avec le « *Maximal Onset principle* », souvent mis en avant dans les études en phonologie (ex. Selkirk, 1982, 1984) selon lequel les attaques, dans la délimitation des frontières syllabiques en interne d'unité lexicale, récupèrent les positions consonantiques.

Une autre tendance très forte qui se dégage de cette typologie est que les groupements consonantiques intra-syllabiques sont nettement défavorisés : on ne les rencontre que dans 7,86 % des quelques 160 500 syllabes de notre corpus. Si on se penche sur la place de ces groupements consonantiques dans la syllabe, nous constatons que la plupart (67 %) occupent la position initiale, contre 33 % en finale de syllabe. Ces résultats sont renforcés par ceux de Zerling (2000) et de Molinu (1999) qui montrent que les groupements consonantiques sont défavorisés et que lorsqu'ils existent, ils se trouvent plus fréquemment en attaque qu'en coda (Molinu, 1999).

Avec groupement consonantique

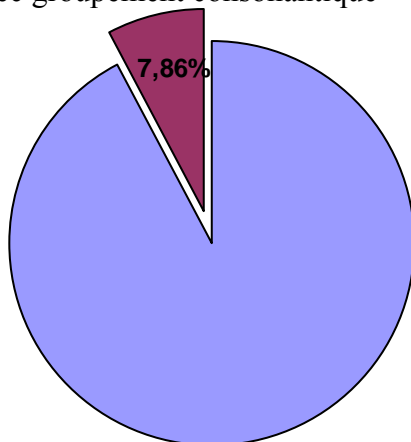


Figure III.21 : Structures avec et sans groupements consonantiques dans notre échantillon.

La tendance pour les groupements consonantiques à apparaître majoritairement en attaque est connue depuis longtemps, et un certain nombre de théories phonologiques s'appuient sur cette observation pour l'établissement de règles permettant de déterminer les frontières syllabiques (Selkirk, 1982, 1984). Les études actuelles en phonétique montrent que ce phénomène peut être lié à des contraintes sur le nombre et la position des consonnes dans le cycle mandibulaire (Redford, 1999), la phase d'abaissement étant plus longue que la

remontée, elle laisserait plus de place à l'articulation consonantique en position initiale de syllabe.

Nous avons donc mis en évidence que les différents types n'apparaissent pas de manière équitable dans les lexiques, et que certains sont d'ailleurs très favorisés au détriment d'autres qui ont des taux d'occurrence quasi nuls. La fréquence des structures semble inversement liée à leur complexité. En effet, une syllabe semble avoir d'autant plus de chances d'apparaître que sa structure sera proche de celle des syllabes de base CV ou CVC, présentes dans toutes nos langues. Ainsi, les syllabes avec attaque vide, groupement consonantique ou consonnes syllabiques sont nettement moins fréquentes dans les langues de notre échantillon.

### **III.3.3. Prédominance CV et CVC : 2 types de langues**

CV et CVC étant les deux seules structures que l'on rencontre dans toutes les langues de la base, il a paru intéressant d'examiner leur répartition dans les lexiques en fonction de la prédominance d'une de ces deux structures dans la langue. En d'autres termes, nous avons étudié la fréquence respective des différentes structures ouvertes et fermées de notre échantillon, en lien avec la présence majoritaire de CV ou CVC (cf. figures III.22 et III.23).

Alors que 10 langues sur 12 totalisent 60 % ou plus de leurs syllabes dans le type CV et que cette structure est favorisée pour 11 de nos lexiques (cf. figure III.22), le wa, le nyah kur, le thaï (toutes trois issues de la famille austro-asiatique mais de branches différentes), le sora et le suédois possèdent majoritairement des syllabes de type CVC. De manière plus générale, ces cinq lexiques favorisent, contrairement aux autres, les syllabes fermées, c'est-à-dire à coda pleine (cf. figure III.23). Il est intéressant de remarquer que ces lexiques proviennent de 5 des 6 langues à tons de notre échantillon, même si le nombre de tons recrutés n'est pas le même pour chacune d'entre elles. En effet, ce chiffre varie de 2 pour le suédois à 5 pour le thaï, les autres langues étant à un niveau intermédiaire. Cette utilisation de la prosodie semble influencer sur le recrutement des unités syllabiques, comme cela a été montré par des études précédentes (Hombert, Ohala et Ewan, 1979).

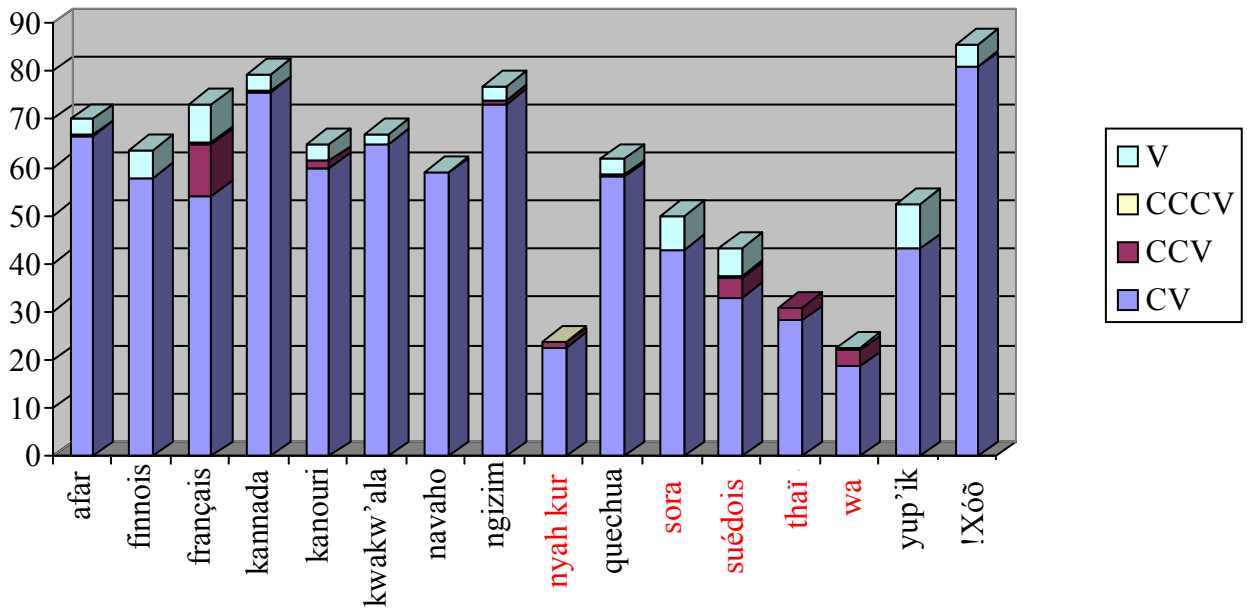


Figure III.22 : Répartition des types de syllabes ouvertes les plus répandus (en pourcentage sur l'ensemble des syllabes). Les langues majoritairement CVC sont en rouge.

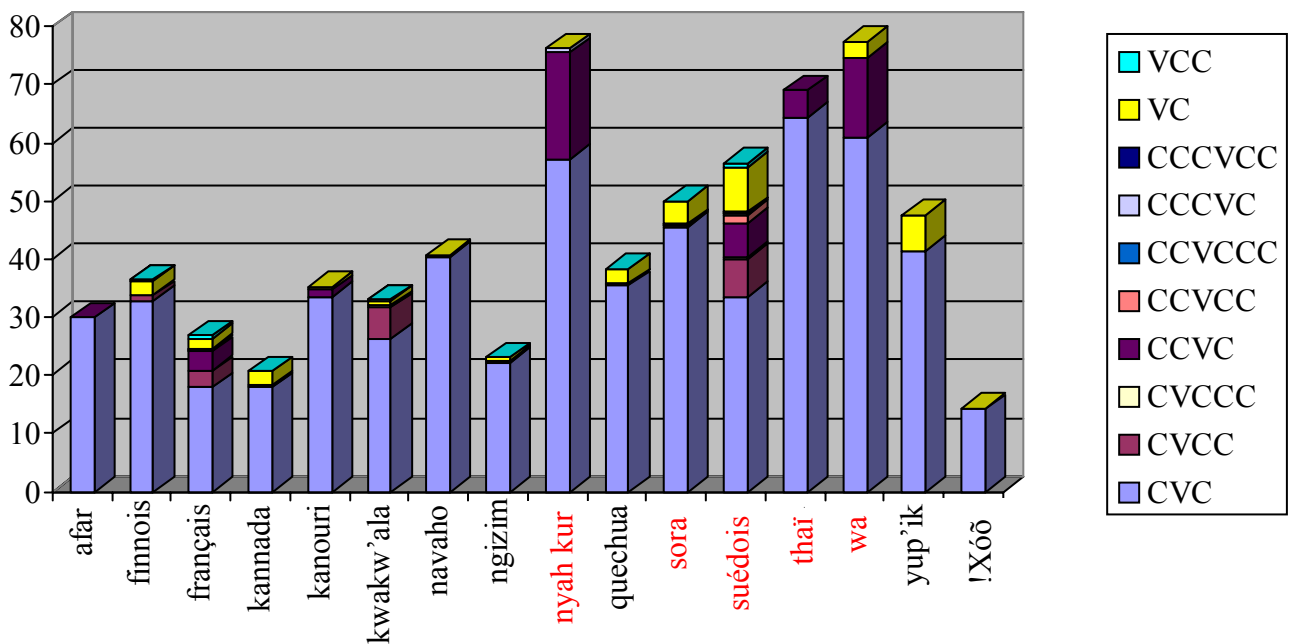


Figure III.23 : Répartition par langue des types de syllabes fermées les plus répandus (en pourcentage sur l'ensemble des syllabes). Les langues majoritairement CVC sont en rouge.

L'observation de la répartition des structures ouvertes et fermées dans les langues permet de remarquer que les séquences fermées sont plus fréquentes dans les langues possédant CVC comme syllabe majoritaire et que les séquences ouvertes le sont dans les langues pour

lesquelles la structure majoritaire est CV. La mise en évidence d'une telle organisation au niveau de la répartition des structures syllabiques des langues laisse supposer l'existence de contraintes régissant l'apparition des types syllabiques dans les langues, construites sur un patron de base dominant.

### III.3.4. Règles implicationnelles des structures syllabiques

L'étude des inventaires des structures syllabiques possibles dans un échantillon composé de 11 langues a permis à Juliette Blevins (1995-1996) de formuler des lois implicationnelles régissant l'apparition des différents types de cohortes : « *if clusters of  $n$  Cs are possible syllable-initially, then clusters of  $n-1$  Cs are also possible syllable-initially, and if clusters of  $n$  Cs are possible syllable- finally, then clusters of  $n-1$  Cs are also possible finally. [...] De plus, « if a language does not allow syllables consisting solely of V, then it does not allow any V-initial syllables » (Blevins, 1995-1996, p. 217). Ces règles se vérifient effectivement si l'on observe les inventaires syllabiques présentés dans différentes autres études. Ainsi, les types syllabiques du suédois livrés dans le travail de Robert Bannert (1998), permettent de proposer les implications suivantes :*

$$\begin{aligned} \text{CCCVCC} &\Rightarrow \text{CCVCC} \Rightarrow \text{CVCC} \Rightarrow \text{CVCC} \Rightarrow \text{VCC} \\ \text{CVCCC} &\Rightarrow \text{CVCC} \Rightarrow \text{CVC} \Rightarrow \text{CV} \\ \text{VCC} &\Rightarrow \text{VC} \Rightarrow \text{V} \end{aligned}$$

Les données de Molinu (1999) sur les syllabes dans un parler roman de l'Italie du Sud montrent que ces lois ne semblent pas régir uniquement l'apparition ou non des structures plus ou moins complexes, mais aussi leur fréquence d'occurrences. « *Au fur et à mesure que la syllabe se complexifie, son taux d'occurrence se réduit, devenant déjà négligeable pour toutes les sous-productions à partir du 7<sup>ème</sup> rang* » (Molinu, 1999, p. 152). Il paraît par conséquent incontournable d'examiner nos langues à la lumière de ces lois implicationnelles.

	afar	finnois	français	kannada	kanouri	kwakw'ala	navaho	ngizim
<b>CV</b>	<b>63,55 %</b>	<b>57,74 %</b>	<b>54,17 %</b>	<b>75,52 %</b>	<b>59,96 %</b>	<b>64,83 %</b>	<b>58,94 %</b>	<b>73,13 %</b>
<b>CCV</b>	0,11 %	0,12 %	10,82 %	0,38 %	1,40 %	0,01 %	0,08 %	0,74 %
<b>CCCV</b>		0,005 %	0,25 %					
<b>V</b>	<b>3,15 %</b>	<b>5,75 %</b>	<b>7,94 %</b>	<b>3,43 %</b>	<b>3,43 %</b>	<b>2,17 %</b>	<b>0,11 %</b>	<b>2,95 %</b>
<b>VC</b>	4,65 %	2,63 %	1,92 %	2,45 %	0,53 %	0,87 %	0,02 %	0,63 %
<b>VCC</b>		0,04 %	0,41 %	0,03 %		0,16 %		0,04 %
<b>VCCC</b>			0,005 %			0,04 %		
<b>CVC</b>	<b>28,52 %</b>	<b>32,69 %</b>	<b>17,92 %</b>	<b>17,91 %</b>	<b>33,42 %</b>	<b>26,18 %</b>	<b>40,30 %</b>	<b>22,20 %</b>
<b>CVCC</b>		0,96 %	2,70 %	0,27 %		5,44 %	0,06 %	0,04 %
<b>CVCCC</b>			0,12 %		0,03 %	0,27 %		
<b>CVCCCC</b>			0,002 %			0,02 %		
<b>CCVC</b>	<b>0,02 %</b>	<b>0,06 %</b>	<b>3,27 %</b>		<b>1,24 %</b>		<b>0,23 %</b>	<b>0,15 %</b>
<b>CCVCC</b>		0,01 %	0,32 %					
<b>CCVCCC</b>			0,01 %					
<b>CCCVC</b>			<b>0,09 %</b>					
<b>CCCVCC</b>			<b>0,03 %</b>					
<b>CCCVCCC</b>								
<b>C</b>			0,02 %			0,01 %	0,26 %	0,11 %
<b>CC</b>							0,02 %	

	nyah kur	quechua	sora	suédois	thaï	wa	yup'ik	!xóõ
<b>CV</b>	<b>22,64 %</b>	<b>58,05 %</b>	<b>42,86 %</b>	<b>33,05 %</b>	<b>28,23 %</b>	<b>18,71 %</b>	<b>43,13 %</b>	<b>80,89 %</b>
<b>CCV</b>	1,09 %	0,48 %	0,06 %	4,16 %	2,75 %	3,33 %		0,03 %
<b>CCCV</b>	0,05 %			0,30 %				
<b>V</b>		<b>3,18 %</b>	<b>7,08 %</b>	<b>5,91 %</b>		<b>0,69 %</b>	<b>9,22 %</b>	<b>4,73 %</b>
<b>VC</b>	0,04 %	2,55 %	3,87 %	7,60 %		2,74 %	6,37 %	0,08 %
<b>VCC</b>		0,02 %	0,15 %	0,82 %				
<b>VCCC</b>				0,05 %				
<b>CVC</b>	<b>57,11 %</b>	<b>35,42 %</b>	<b>45,44 %</b>	<b>33,56 %</b>	<b>64,41 %</b>	<b>60,94 %</b>	<b>41,28 %</b>	<b>14,14 %</b>
<b>CVCC</b>		0,1 %	0,50 %	6,33 %				
<b>CVCCC</b>				0,56 %				
<b>CVCCCC</b>				0,01 %				
<b>CCVC</b>	<b>18,39 %</b>	<b>0,2 %</b>	<b>0,04 %</b>	<b>5,65 %</b>	<b>4,61 %</b>	<b>13,58 %</b>		
<b>CCVCC</b>				1,34 %				
<b>CCVCCC</b>				0,08 %				
<b>CCCVC</b>	<b>0,67 %</b>			<b>0,45 %</b>				
<b>CCCVCC</b>				0,11 %				
<b>CCCVCCC</b>				0,005 %				
<b>C</b>				0,05 %				0,08 %
<b>CC</b>				0,005 %				0,06 %

Table III.8 : Liste des structures syllabiques d'ULSID classées par type et leur proportion dans le lexique d'une langue donnée.

L'observation des trois premières lignes de la table III.8. nous apprend que les structures CCCV ne peuvent apparaître que lorsque CCV est attestée. De même, cette dernière n'est possible que si CV existe dans la langue. Si on observe les autres types (à partir de la troisième ligne), on note qu'il en va de même pour les structures à attaque complexe et coda simple. D'après les langues de notre échantillon, il est par conséquent possible d'établir très précisément la règle suivante :



**R1 : Si une structure syllabique avec attaque complexe de n consonnes est attestée dans une langue, elle implique les structures avec attaques de n-i consonnes (i=1 à n-1) ;**

Ce qui revient à formuler, pour les types de syllabes des treize langues de notre échantillon, les implications suivantes :

$$\begin{aligned} \text{CCCV} &\Rightarrow \text{CCV} \Rightarrow \text{CV} \\ \text{CCCVC} &\Rightarrow \text{CCVC} \Rightarrow \text{CVC} \end{aligned}$$

L'étude des lignes 4 à 7 montre que les structures avec attaque vide sont soumises à une loi similaire. En effet, excepté le cas du nyah kur, les onze langues comportant une forme de type VC ou VCC attestées possèdent aussi la structure V. Nous pouvons donc définir la règle suivante :

**R2 : Les structures syllabiques avec attaque vide impliquent la présence de la structure V dans la langue.**

Pour le nyah kur, nous proposons, vu le faible taux d'apparition de la forme VC (0,04 %, soit 2 occurrences sur 5 503 syllabes) que ce cas ne remet pas en question la loi proposée.

À présent, l'étude des structures comportant une coda complexe fait remarquer que, comme dans le cas des attaques, la présence d'une forme complexe implique celle de la structure simple associée. Ainsi, une règle analogue à celle des attaques peut être énoncée :

**R3 : Si une structure syllabique avec coda complexe de n consonnes est attestée dans une langue, elle implique les structures avec coda de n-i consonnes (i=1 à n-1).**

Les différentes structures avec coda complexe, présentes dans notre base de données, respecteront donc les implications suivantes :

$$\begin{aligned} \text{CVCCCC} &\Rightarrow \text{CVCCC} \Rightarrow \text{CVCC} \Rightarrow \text{CVC} \\ \text{VCCC} &\Rightarrow \text{VCC} \Rightarrow \text{VC} \Rightarrow \text{V} \end{aligned}$$

Les trois règles précédentes permettent de définir des conditions d'apparition de la majorité des types de structures syllabiques des langues de notre corpus. Mais pour certaines, d'autres règles sont nécessaires. Ainsi, pour les langues acceptant les consonnes syllabiques, il semble y avoir une corrélation entre la présence de CC et celle de C. Cette règle pourrait être construite sur le modèle des précédentes en déclarant :

**R4 : Si le groupement syllabique CC apparaît dans une langue, il implique la présence d'une consonne syllabique C.**

Pour ce qui est de la structure CCVCC, elle met en jeu des éléments obéissant à deux des règles, et on arrive par conséquent à l'implication suivante :

$$\text{CCVCC} \Rightarrow (\text{CCVC et CVCC}) \Rightarrow \text{CVC}$$

Les règles d'implication ne sont pas les seules informations que nous tirons de la table 3. En effet, les fréquences d'apparition des différentes structures dans les langues semblent elles aussi soumises à des règles strictes. À part deux exceptions au niveau des formes V en afar (V : 3,5 % et VC : 4,65 %), en suédois (V : 5,91 % et VC : 7,6 %) et en wa (V : 0,69 % et VC : 2,74 %) et pour des proportions très faibles dans les trois cas, nous pouvons remarquer que plus la complexité de l'attaque ou de la coda augmente, plus la fréquence de la structure diminue. Nous constatons, l'ordre de fréquence suivante :

$$\begin{aligned} & \text{CCCV} \lll \text{CCV} \ll \text{CV} \\ & \text{CCVC} \lll \text{CCVC} \ll \text{CVC} \\ & \text{CVCCCC} \lll \text{CVCCC} \ll \text{CVCC} < \text{CVC} \\ & \text{VCCC} \lll \text{VCC} \ll \text{VC} < \text{V} \\ & \text{CCVCC} \ll (\text{CCVC et CVCC}) \\ & \text{CC} \ll \text{C} \end{aligned}$$

(avec < pour inférieur, << pour bien inférieur et <<< pour très largement inférieur).

Ce qui permet de postuler la règle suivante :

**R5 : La fréquence d'apparition d'une structure syllabique diminue lorsque la complexité de son attaque et/ou de sa coda augmente.**

La complexité se mesurant par le nombre d'éléments constitutifs de l'attaque et/ou de la coda, cette loi permet de situer, par exemple, le rang d'apparition d'une forme telle que CCVCC qui possède à la fois une attaque et une coda complexe, et ne qui pourra donc apparaître que rarement dans les langues du monde.

L'ensemble des lois implicationnelles validées sur nos langues renforcent celles mises en évidence par Blevins (1996-1997) sur un échantillon n'ayant qu'une seule langue en commun avec le nôtre. Il apparaît, à travers ce résultat, que les groupements consonantiques intra-syllabiques sont soumis à des contraintes organisationnelles. Une étude plus approfondie sur la nature phonétique des différents constituants des attaques et des codas sera nécessaire afin de cerner davantage les restrictions qui semblent peser sur ces structures (cf. chapitre III.4.3).

### III.3.5. Convergence entre résultats d'études phonétiques et tendances des structures syllabiques

La connaissance du faible nombre de structures possibles (19 types) pour décrire les 247 252 syllabes de notre corpus, et l'étude des différents types de structures syllabiques des langues de notre base de données nous a permis de montrer que leur caractère plus ou moins fréquent ne doit probablement rien au hasard, un certain nombre de règles implicationnelles pouvant être postulées. Ces observations peuvent, au moins en partie, s'expliquer grâce à des résultats obtenus à propos des caractéristiques articulatoires et perceptives des types syllabiques apparaissant de façon majoritaire.

Ainsi, la préférence nette des structures avec attaque pleine et coda vide rejoint les études phonétiques qui montrent qu'il existe des différences entre les deux positions consonantiques, que ce soit au niveau de la production ou de la perception. Les consonnes seraient plus intenses et plus longues en attaque, et au niveau perceptif, les consonnes d'attaque résisteraient mieux au bruit (Redford et Diehl, 1999).

De plus, l'apparition majoritaire de structures simples (sans groupements consonantiques) peut être reliée aux résultats obtenus en phonétique. Ainsi, la plupart des études montrant l'existence d'une asymétrie entre les consonnes d'attaque et de coda remarquent aussi que les groupements de consonnes sont moins stables au niveau perceptif que les consonnes seules, ce qui pourrait expliquer qu'ils ne soient que peu recrutés par les langues (Fougeron et Keating, 1997 ; Redford et Diehl, 1999).

L'ensemble de ces résultats rejoint l'idée de l'intégration de la syllabe dans le cycle mandibulaire, avec une place assignée pour chacun de ses constituants. Le temps disponible pour la phase d'ouverture étant plus long que celui accordé pour la fermeture de la cavité buccale, le locuteur peut mieux articuler les consonnes en position initiale de syllabe et les rendre ainsi plus perceptibles et plus stables que les consonnes apparaissant en position finale (Fougeron et Keating, 1997 ; MacNeilage, 1998 ; Redford, 1999).

La présence majoritaire des types syllabiques avec attaque pleine et coda vide est aussi reliée aux résultats obtenus en perception quant à la stabilité des structures CV et VC (Tuller et Kelso, 1990, 1991), ainsi qu'aux études portant sur la reconnaissance des types syllabiques par des nouveaux-nés (Bijeljac-Babic *et al.*, 1993 ; Bertoncini *et al.*, 1995 ; Ooijen *et al.*, 1997 ; Mattys et Jusczyk, 2001). En effet, l'ensemble de ces travaux montre que les structures avec attaque pleine sont perçues plus tôt et sont plus stables chez l'adulte, alors que la segmentation des structures avec voyelle initiale (attaque vide) est plus tardive et constitue une forme moins stable dans une tâche de perception.

Les contraintes de production et de perception convergent pour faire de la syllabe CV (attaque pleine et coda vide), la syllabe canonique, c'est-à-dire la première à apparaître au stade du babillage (6-7 mois) dans le développement du langage, présente dans toutes les

langues du monde (même dans celles majoritairement CVC, cf. table III.8), la plus stable lors de l'écoute d'un stimulus ambigu ou dans le bruit, et celle qui colle au mieux au cycle d'abaissement et de remontée de la mâchoire, correspondant au cycle mandibulaire.

### III.4. LES CONSTITUANTS SYLLABIQUES

L'analyse du nombre de phonèmes par syllabes ainsi que des types syllabiques et de leur fréquence dans les langues d'ULSID montre que l'organisation structurelle des syllabes est probablement soumise à un certain nombre de contraintes. Dans le prolongement des tendances et régularités pesant sur la structure de la syllabe, nous avons observé la nature des éléments constituants des syllabes afin de rechercher d'éventuels comportements tendanciels des lexiques à ce niveau de l'organisation syllabique.

#### III.4.1. Rendement syllabique

Le calcul de la forme canonique des langues (cf. partie III.2.1.1) montre que la majorité utilise des items lexicaux composés de plus d'une syllabe. Il a semblé utile d'examiner dans quelle proportion les langues exploitent plusieurs fois la même syllabe, en la combinant avec d'autres pour construire des unités lexicales. Afin de vérifier ce qu'il en est exactement, nous avons calculé le rendement syllabique de chaque langue, c'est-à-dire le rapport entre le nombre total de syllabes obtenues par le découpage d'un lexique donné et le nombre de syllabes différentes comptabilisées après regroupement des syllabes identiques au niveau de la nature des segments qui la composent (cf. figure III.2.4).

Ainsi, un rendement syllabique (RS) égal à 1 signifie que chaque syllabe n'apparaît qu'une seule et unique fois dans le corpus. Plus le RS augmente, plus la fréquence d'utilisation d'une même syllabe dans les unités lexicales augmente. Par exemple, la syllabe « t a m » n'apparaît qu'une seule fois en navaho, en suédois, en wa, en !xóǒ et en yup'ik, mais entre 4 et 9 fois pour 8 langues (afar, finnois, français, kannada, ngizim, nyah kur, quechua et thaï), 30 fois en kanouri et 104 fois en sora, alors qu'elle est totalement absente du kwakw'ala. On peut alors en déduire que le rendement de cette syllabe est différent selon les langues.

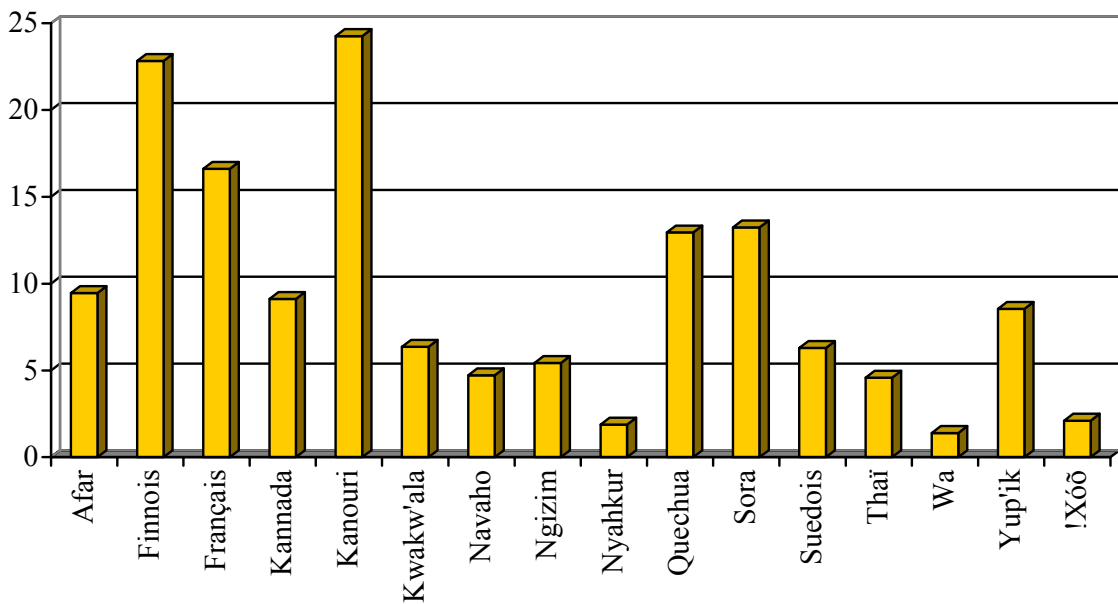


Figure III.24 : Rendement syllabique des langues d'ULSID

L'analyse du rendement syllabique montre que l'on peut retrouver un classement de nos langues en différents types, et que ce dernier rejoint la typologie des 16 langues sur le critère du nombre de syllabes par unité lexicale (cf. partie III.2.1.2).

- Les deux langues totalement ou majoritairement monosyllabiques du groupe 1, le wa et le nyah kur, ont un rendement syllabique inférieur à 2 ;
- Les langues du type 2, c'est-à-dire celles possédant plus d'unités lexicales de plus d'une syllabe ont tendance à plus réutiliser leurs syllabes, et ont un rendement syllabique compris entre 2 et 5 ;
- Les langues du troisième type, qui ont des unités lexicales majoritairement dissyllabiques, se caractérisent par un rendement syllabique compris entre 5 et 14, et réutilisent donc davantage les mêmes syllabes dans la construction de différentes unités lexicales ;
- Pour les langues du type 4, majoritairement trisyllabiques, trois d'entre elles possèdent un rendement syllabique très élevé (entre 16 et 24), mais la quatrième, le yup'ik, possède un rendement syllabique plus faible, quoi que non négligeable (8,55).

Si l'on observe en premier les langues des types 1 et 2, on observe que le faible taux de réutilisation des syllabes est lié avec le caractère majoritairement monosyllabique de ces langues. On peut même considérer que si une syllabe équivaut à une unité lexicale, le rapport entre le nombre total de syllabes et le nombre total d'items lexicaux devrait valoir 1. La présence de structures de plus d'une syllabe permet d'expliquer les valeurs de rendement

syllabique supérieur à 1 pour les langues en possédant, mais même pour le wa, seule langue totalement monosyllabique de notre échantillon, le rendement syllabique est supérieur à 1 (RS = 1,39). Ce phénomène peut s'expliquer par l'étude du rôle de la prosodie dans les différentes langues des types 1 et 2. Ainsi, le wa, le thaï et le navaho sont des langues à tons, ce qui signifie qu'une différenciation des unités lexicales se fera en fonction de la hauteur de tons associée à la syllabe. Les tons ne sont pas notés dans nos lexiques. Ce phénomène permet de comprendre comment les 3 180 unités lexicales du wa sont réalisées par 3 180 unités syllabiques, mais que seulement 2 288 syllabes différentes sont recensées dans le lexique. Certaines syllabes apparaissent plusieurs fois de façon identique au niveau de la nature des segments qui la compose, mais en tant qu'éléments lexicaux différents, avec des sens différents pour chacune de ces réalisations.

À contrario, il semble naturel que ce soit pour les langues des types 3 et 4 que les syllabes soient le plus souvent réutilisées. En effet, le sens d'un élément lexical ne repose pas sur une seule syllabe, mais sur l'association de plusieurs d'entre elles et par conséquent, si le contexte change, l'unité lexicale le fait aussi, même si une de ses syllabes reste identique. Ainsi, « d a. d e n » sera différent de « d a. d i » même si la syllabe « d a » est présente dans les deux cas. Dès lors, plus le nombre de syllabes pouvant composer une unité lexicale augmente, plus le nombre d'associations possibles croît, et la fréquence d'utilisation des syllabes évolue dans le même sens.

L'étude du rendement syllabique permet de mieux comprendre comment les langues utilisent leurs syllabes et quelles sont leurs stratégies au niveau de la construction des unités lexicales. La réutilisation de syllabes dans la construction du lexique permet aux langues de limiter leur inventaire syllabique. Le calcul du rendement syllabique permet de mettre le doigt sur un comportement des langues : celles-ci augmentent d'autant plus la réutilisation de syllabes que la longueur de leurs unités lexicales est grande. Visiblement au niveau des constituants syllabiques, les langues n'utilisent pas tout leur potentiel, c'est-à-dire l'ensemble des combinaisons possibles entre V et C dans les différentes structures syllabiques, calculables à partir des inventaires phonologiques de voyelles et de consonnes.

Dans la partie suivante, nous analysons en détail la nature des constituants des syllabes à plus fort rendement. Les différentes études qui suivent nécessitent la connaissance précise du système phonologique de chacune des langues observées. Or, la transcription entre les données saisie à partir du dictionnaire Berlitz (1981) et le système phonologique du suédois n'a pu être réalisé en raison de problèmes rencontrés dans la distinction de certaines voyelles. Ils n'ont pas été tous résolus. Les analyses suivantes ne porteront par conséquent que sur 15 langues.

### III.4.2. Syllabes de premier plan

Afin de déterminer s'il existe des tendances dans la nature des constituants syllabiques, nous avons extrait des données les 20 syllabes les plus fréquentes dans chacune des langues de notre échantillon (voir liste en annexe IV). Nous avons ainsi obtenu 300 combinaisons de consonnes et de voyelles que nous étudions séparément pour identifier les segments les plus recrutés par les langues.

La première remarque concerne le nombre peu élevé de voyelles différentes qui apparaissent dans ces 300 syllabes. En effet, seuls 19 noyaux vocaliques différents se dégagent dans des proportions toutefois loin d'être équitables (cf. table III.9).

Voyelles (API)	Nombre d'occurrences
a	119
i / ɪ	46
u	33
ə	30
e / ε	25
ɔ / o	12
longues	23
dont a	15
nasales	6
diphthongues	5
u̥	1

Table III.9: Répartition des voyelles dans les syllabes les plus fréquentes (20 premières de chaque langue) (u̥ est la voyelle murmurée).

On constate que seules 6 voyelles ont une fréquence d'apparition supérieure à 4 %, et que la voyelle centrale /a/ est très largement favorisée. Elle représente à elle seule près de 40 % des 300 noyaux vocaliques étudiés, et si on lui ajoute la nasale correspondante /ã/, cette fréquence atteint 44 %. Au second rang avec un peu plus de 15 % des noyaux vocaliques, on retrouve la voyelle /i/, ensuite /u/ et /ə/ avec une fréquence d'environ 10 % chacun. Il semble donc que les voyelles les plus souvent recrutées au sein des syllabes les plus fréquentes de l'échantillon correspondent aux voyelles extrêmes du triangle vocalique, qui se trouvent être les voyelles les plus fréquentes dans les systèmes phonologiques des langues du monde (Vallée, 1994). Plus de la moitié des noyaux vocaliques sont occupés par une voyelle centrale.

Observant à présent la nature des consonnes qui figurent en attaque des vingt syllabes les plus fréquentes de chaque langue, un constat identique apparaît. En effet, certaines consonnes sont sur-représentées au détriment d'autres qui n'apparaissent pas.

Consonnes	Nombre d'occurrences	Consonnes	Nombre d'occurrences
t / ṭ	34	d	9
k	31	b	8
l / ḷ	25	j	7
m	21	r	7
s	16	retroflexes	7
n / ṇ	16	w	6
aspirées	13	affriquées	6
g	11	p	5
r	10	attaques complexes	5
?	9	autres consonnes (moins de 1,5 %)	33

Table III.10: Consonnes les plus répandues dans les vingt syllabes les plus fréquentes (en position d'attaque).

Les vingt premières syllabes de chaque langue possèdent majoritairement une attaque constituée de plosives sourdes vélaires et coronales [k t], devant la latérale coronale [l], les nasales coronales et labiales [n m], et la plosive sonore vélaire [g]. Ensuite, apparaissent la fricative coronale [s], la plosive sonore coronale [d] et la vibrante coronale [r].

Le lieu d'articulation coronal est donc largement privilégié dans la position d'attaque des syllabes les plus fréquentes des langues de notre échantillon. Cette prédilection pour les coronales est à rapprocher de la fréquence de ce type de consonnes dans les systèmes des langues du monde. Ainsi, les études sur UPSID<sub>317</sub> et d'UPSID<sub>451</sub> (Vallée *et al.*, 1998, 1999) ont montré qu'elles représentent 44.5 % des consonnes des systèmes phonologiques. De plus, lorsqu'il n'y a qu'un lieu d'articulation, développé par les systèmes dans les catégories plosives, fricatives, nasales ou autre, il s'agit de consonnes coronales (Vallée *et al.*, 2002).

Cette première approche laisse à penser que les syllabes les plus fréquentes dans les langues de notre échantillon sont la combinaison des unités sonores les plus fréquentes. Or, les études sur l'organisation structurelle des syllabes montrent que ces dernières connaissent des contraintes de formation propres. Il semble nécessaire de généraliser l'étude de la nature des constituants syllabiques afin de vérifier si à ce niveau, les contraintes observées par les langues ont pour origine les segments ou l'unité syllabique. Dans le premier cas, l'organisation syllabique doit se contenter d'appliquer les contraintes d'apparition des segments. Dans le second, les syllabes, tout en tenant compte des contraintes pesant sur les segments, ajoutent leurs propres règles de combinaison. Afin de vérifier cette hypothèse sur l'échantillon retenu, il est apparu nécessaire d'analyser la composition des différents constituants syllabiques, en débutant par le noyau vocalique.



### III.4.3. Les noyaux vocaliques « vedettes »

L'analyse des fréquences d'apparition des différentes voyelles au sein des syllabes de l'échantillon met en évidence, comme pour les 300 syllabes les plus fréquentes, une présence très dominante de la voyelle centrale ouverte /a/ (avec ou sans dimension secondaire : nasalité, quantité, ...) dont l'articulation est réalisée par l'abaissement maximum de la mandibule. Elle représente plus de 35 % de l'ensemble des noyaux vocaliques et est la voyelle la plus fréquente dans 11 langues sur 15. Dans les quatre exceptions, le français, le navaho, le nyah kur et le wa, elle apparaît en deuxième position. Ces langues privilégient toutes des voyelles différentes, /e/ pour le français, /i/ pour le navaho, /ə/ pour le nyah kur et /o/ pour le wa.

Pour les voyelles de rang inférieur, /i/ occupe le plus souvent la deuxième position en représentant presque 17 % des noyaux vocaliques. Au troisième rang se trouve /e/, mais la présence majoritaire de cette voyelle dans le lexique du français, qui, rappelons le, est le lexique le plus important de notre base de données, influe sur ce résultat. Ainsi, en ôtant les résultats provenant du français, /e/ passe en quatrième position et /u/ remonte à la troisième avec, dans ce cas, plus de 14 % des noyaux vocaliques.

Une autre remarque que l'on peut faire à partir de la répartition des différentes voyelles en position de noyaux vocaliques dans les langues de notre échantillon, concerne la présence de la voyelle centrale /ə/ qui représente plus de 6 % de l'ensemble des noyaux vocaliques alors qu'elle n'est présente que dans 6 langues de notre échantillon. Dans les 4 langues où elle est fortement présente (kanouri, nyah kur, sora et yup'ik), cette voyelle centrale représente entre 11,73 et 21,20 % des noyaux vocaliques, alors que dans les deux autres langues (français et thaï), elle ne représente qu'entre 1,39 et 4,52 des centres de syllabes.

	Afâr	Finnôis	Français	Kannada	Kanouri	Kwakw'ala	Navaho	Ngizim	Nyah kur	Quechua	Sora	Thaï	Wa	Yup'ik	X66	Total
<b>a</b>	44,49	32,16	25,32	39,01	34,13	37,35	35,41	45,45	19,83	49,32	32,23	53,65	12,52	37,04	38,14	<b>32,99</b>
<b>i</b>	19,16	16,65	16,25	15,35	21,09	18,78	38,98	9,05	6,31	25,42	13,35	11,42	7,17	17,89	9,15	<b>16,98</b>
<b>e</b>	14,80	13,85	27,93	14,69	4,66	31,49	11,43	2,46	9,69	0,13	12,15	3,43	9,21		7,47	<b>15,44</b>
<b>u</b>	10,69	10,59	2,95	23,45	11,57	9,71		22,02	10,76	24,98	15,34	6,18	8,90	29,68	17,78	<b>11,05</b>
<b>o</b>	10,86	9,73	16,69	6,66	8,51	2,67	11,98	1,91	16,34	0,14	9,48	15,18	16,38		13,74	<b>10,70</b>
<b>ə</b>			4,52		20,05				21,10		17,43	1,39		11,73		<b>6,20</b>
<b>y</b>		4,01	5,01				2,20	4,50								<b>2,09</b>
<b>œ</b>			1,33													<b>0,36</b>
<b>ɤ</b>									8,74				8,55			<b>0,33</b>
<b>Diphthongues</b>		13,00		0,86				14,60	7,25		0,01	8,75	37,26	3,67	13,72	<b>3,87</b>

Table III.11 : Pourcentage des voyelles dans les vingt syllabes les plus fréquentes (après regroupement en fonction du timbre de base).

Comme pour les 300 syllabes les plus fréquentes, les trois voyelles privilégiées dans les lexiques de notre base de données se trouvent donc bien être celles qui sont les plus fréquentes dans les langues du monde.

Nous nous sommes donc intéressés à la répartition de ces voyelles au sein de différentes structures syllabiques observées dans les langues. Au niveau des structures avec attaque, on ne remarque aucune restriction particulière. Toutes les voyelles peuvent apparaître dans tous les types syllabiques, en respectant globalement les fréquences observées sur l'ensemble des syllabes.

Le cas des syllabes ayant une attaque vide justifie une attention spécifique. Ce type de structure est présent de manière significative dans 12 langues sur 15. On y retrouve les mêmes tendances que pour l'ensemble des autres syllabes au niveau de l'apparition des voyelles. En effet, /a/ est dominant dans 11 langues sur 12 et représente plus de 30 % des noyaux vocaliques. La seule langue faisant exception est le wa, /a/ n'apparaissant qu'en deuxième position après /o/.

/a/ constitue aussi le noyau le plus « populaire » des structures syllabiques présentant attaque et coda vides. En effet, sur les douze langues pour lesquelles ce type de syllabes n'est pas marginal, dix privilégient /a/. Le sora lui préfère la voyelle d'avant /i/ et la voyelle centrale /ə/ dans ses syllabes de type V. Notons que /ə/ est d'ailleurs toujours fréquent. Même s'il ne peut être considéré comme favorisé par les langues de notre échantillon puisqu'il n'existe que dans quatre langues sur les treize qui possèdent ce type de structures syllabiques. /i/, toujours présent dans les langues, apparaît le plus souvent au deuxième rang des occurrences après /a/.

On observe également que les diphtongues peuvent être présentes dans des syllabes avec attaque vide, et ce dans la plupart des langues possédant cette structure syllabique, mais il faut noter leur faible taux d'occurrence au sein de chaque langue, par exemple, au maximum 3,81 % pour /ai/ en finnois.

Il semble alors que, même si la répartition des voyelles dépend en majeure partie de leur fréquence dans les systèmes phonologiques des langues du monde, certaines contraintes apparaissent lorsque l'on étudie la place des voyelles dans les différentes structures. L'observation du positionnement des différentes voyelles en fonction des consonnes apparaissant en attaque et en coda sera donc nécessaire pour faire apparaître d'éventuelles contraintes au niveau des cooccurrences.

#### III.4.4. Tendances consonantiques des structures syllabiques

Au niveau des consonnes, l'analyse doit porter sur les possibilités d'apparition en fonction de la position dans la syllabe. La table III.12 livre le pourcentage de segments de l'inventaire consonantique pouvant apparaître en attaque et en coda et ce pour nos 15 langues.

Langues	Pourcentage de consonnes pouvant apparaître en		
	attaque	coda	Plus de 10 occurrences en coda
afar	100 %	100 %	94,44 %
finnois	100 %	82,35 %	53,20 %
français	95,24 %	90,48 %	90,48 %
kannada	100 %	92,59 %	77,78 %
kanouri	95,45 %	77,27 %	50 %
kwakw'ala	97,62 %	78,57 %	45,24 %
navaho	94,59 %	37,84 %	27,03 %
ngizim	97,3 %	70,27 %	40,54 %
nyah kur	86,67 %	60 %	50 %
quechua	96,97 %	66,67 %	36,36 %
sora	92 %	62 %	38 %
thaï	95,45 %	50 %	40,91 %
wa	97,37 %	21,05 %	21,05 %
yup'ik	97,56 %	65,85 %	41,46 %
!xóǝ	100 %	3,08 %	3,08 %

Table III.12 : Pourcentage de l'inventaire de consonnes pouvant apparaître en attaque et en coda dans les langues d'ULSID (calculé à partir du nombre de syllabes différentes).

Pour ce qui concerne la position d'attaque syllabique, on constate qu'il n'y a que peu de disparité entre les langues, et que la majorité des consonnes de la langue considérée peut apparaître à cette position (toujours plus de 86,67 %). Pour ce qui est des coda, on remarque que le comportement varie beaucoup plus d'une langue à l'autre. Ainsi, en afar, toutes les consonnes peuvent apparaître en coda, alors qu'en wa seulement 21,05 % d'entre elles peuvent occuper cette position. Cette proportion descend même à 3 % pour le !xóǝ. Pour la plupart des autres langues, il faut faire la distinction entre les consonnes qui n'apparaissent que marginalement en coda et celles dont la présence est régulière. Il est alors possible d'observer que, pour 11 langues, les pourcentages de l'inventaire consonantique pouvant apparaître en coda sont compris entre 20 et 54 % si l'on se contente des consonnes fréquentes en coda, et c'est encore le cas pour 7 langues si l'on observe toutes les consonnes pouvant apparaître dans cette position.

L'étude des consonnes apparaissant majoritairement en coda, permet de remarquer qu'il s'agit en majorité de consonnes simples. Les consonnes possédant une articulation supplémentaire (aspirées, murmurées...) ne sont que très peu fréquentes dans cette position. [p t k ʔ m n ŋ s l] sont de loin les coda les plus fréquentes. Nous remarquons qu'il s'agit de plosives sourdes [p t k ʔ], de nasale [m n ŋ] et de coronales [t n s l] que l'on retrouve dans les productions du babillage (Locke, 1983), mais aussi qui sont les plus répandues dans les langues du monde (Vallée *et al.*, 2002).

Comme pour les voyelles, les consonnes les plus fréquemment recrutées correspondent aux segments les plus fréquents dans les systèmes phonologiques des langues du monde. Il existe donc, comme pour les voyelles, des tendances syllabiques qui peuvent être expliquées par l'existence de tendances au niveau des systèmes sonores des langues du monde (Vallée *et al.* 1998 ; Vallée *et al.*, 2002).

Pourtant, la présence de différences significatives entre le nombre de consonnes pouvant apparaître en attaque et celui des consonnes pouvant apparaître en coda montre que des contraintes sont observables à un autre niveau qui ne peut s'expliquer par les caractéristiques intrinsèques aux segments ou au système phonologique. Il est donc naturel de postuler que des contraintes lient la position dans la structure syllabique et la nature des segments, les structures étant elles aussi soumises à des règles déjà observées et que nous pouvons affiner.

### III.4.5. Règles implicationnelles affinées

Nous avons observé précédemment que les syllabes V sont, de manière générale, peu fréquentes dans les différentes langues de notre base de données (4,5 %), mais que leur existence est nécessaire à la présence de structures avec attaque vide et coda pleine dans le lexique. Il semble alors intéressant de vérifier que cette tendance reste vraie lorsque la nature phonétique du noyau est prise en considération. Il faut dès lors étudier, langue par langue, quelles voyelles peuvent apparaître dans les structures avec attaque vide et parmi elles, celles qui ne sont pas présentes sous la forme syllabique V. On ne trouve que de très rares

exceptions, dans des langues où la règle implicationnelle est respectée par les autres voyelles. Nous pouvons, par conséquent, affiner la tendance :

**R2 : Les structures syllabiques avec attaque vide impliquent la présence de la structure V dans la langue**

et proposer la règle suivante (avec L = lexique d'une langue donnée) :

**R2b : La syllabe  $V_iC$  (avec  $V_i$  appartient à L) implique la présence de la syllabe  $V_i$  dans L.**

À cette tendance s'en ajoute une autre sur les fréquences d'apparition des différentes structures :

**R5 : La fréquence d'apparition d'une structure syllabique diminue lorsque la complexité de son attaque et/ou de sa coda augmente**

qui se vérifie aussi sur nos données à propos des syllabes sans attaque. Ainsi, on aura la règle suivante :

**R5b : La fréquence d'apparition d'une structure syllabique sans attaque diminue lorsque la complexité de sa coda augmente**

Ces différentes règles implicationnelles peuvent être illustrées par les exemples suivants :

En finnois : /u/ (55 occurrences) < /ur/ (12 occurrences) < /urt/ (2 occurrences)

En kwakw'ala : /i/ (58 occurrences) < /iʔ/ (5 occurrences) < /iʔs/ (1 occurrence)

En sora : /a/ (254 occurrences) < /an/ (53 occurrences) < /and/ (5 occurrences)

Les règles implicationnelles mises en évidence au niveau structurel sont respectées au niveau de la nature des segments dans le cas des syllabes à attaque vide. Il est donc intéressant d'observer ce qu'il en est pour les règles R1 et R3 observées respectivement pour les attaques et les coda et complétées par l'affirmation R5 portant sur leur fréquence d'apparition.

L'étude des syllabes apparaissant dans notre échantillon permet de valider ces règles implicationnelles au niveau de la nature des segments qui les composent. Il est alors possible de les généraliser à la nature des consonnes impliquées dans les attaques ou les coda. C'est, par exemple, le cas dans les syllabes suivantes pour les attaques :

En finnois : /s p r/  $\Rightarrow$  (/s p/ et /p r/)  $\Rightarrow$  (/s/ /p/ /r/)

/s p r/ < (/s p/ et /p r/)  $\ll$  (/s/ /p/ /r/)

En nyah kur : /ɲ c<sup>h</sup> r/  $\Rightarrow$  (/ɲ c<sup>h</sup>/ et /c<sup>h</sup> r/)  $\Rightarrow$  (/ɲ/ /c<sup>h</sup>/ /r/)

/ɲ c<sup>h</sup> r/ < (/ɲ c<sup>h</sup>/ et /c<sup>h</sup> r/)  $\ll$  (/ɲ/ /c<sup>h</sup>/ /r/)

Et dans les suivantes pour les coda :

En français :  $/s t \mathfrak{R}/ \Rightarrow (/s t/ \text{ et } /t \mathfrak{R}/) \Rightarrow (/s/ /t/ / \mathfrak{R}/)$   
 $/s t \mathfrak{R}/ < (/s t/ \text{ et } /t \mathfrak{R}/) \ll (/s/ /t/ / \mathfrak{R}/)$

En kwakw'ala :  $/\mathfrak{?} \chi^w s/ \Rightarrow (/ \mathfrak{?} \chi^w/ \text{ et } / \chi^w s/) \Rightarrow (/ \mathfrak{?}/ / \chi^w/ /s/)$   
 $/\mathfrak{?} \chi^w s/ < (/ \mathfrak{?} \chi^w/ \text{ et } / \chi^w s/) \ll (/ \mathfrak{?}/ / \chi^w/ /s/)$

Le même type de vérification a été faite pour l'assertion suivante :

**R4 : Si le groupement extrasyllabique CC apparaît dans une langue, il implique la présence d'une consonne extrasyllabique (C).**

Et comme pour les autres, elle se trouve vérifiée au niveau de la nature des segments, comme le montrent les exemples suivants :

En navaho :  $/m \mathfrak{?}/ \Rightarrow /m/$   
 $/m \mathfrak{?}/ < /m/$

En !xóõ:  $/\mathfrak{?} n/ \Rightarrow (/ \mathfrak{?}/ \text{ et } /n/)$   
 $/\mathfrak{?} n/ < (/ \mathfrak{?}/ \text{ et } /n/)$

Comme pour l'organisation des types syllabiques, les constituants syllabiques sont soumis à certaines tendances. Nous avons observé que voyelles et consonnes sont toutes les deux soumises au même type de règles implicationnelles au niveau de la complexité de la structure, il semble donc important de vérifier, à ce stade, si leurs combinaisons connaissent également des restrictions.

### III.5. COOCCURRENCES ET DEPENDANCES INTRA ET INTERSYLLABIQUES

Les segments qui apparaissent le plus fréquemment se trouvent être les plus fréquents dans les systèmes phonologiques des langues du monde. De plus, aux contraintes pesant sur l'apparition des différents types syllabiques dans les langues semblent s'en ajouter d'autres sur les consonnes pouvant apparaître en attaque ou en coda. Il semble donc intéressant de rechercher d'éventuelles cooccurrences ou dépendances entre les éléments consonantiques et vocaliques des syllabes. Ce qui revient à poser la question suivante : existe-t-il des combinaisons favorisées au niveau intra- et inter-syllabique, et si oui, lesquelles et pourquoi ?

#### III.5.1. Restriction dans l'inventaire syllabique des langues

Le calcul du rendement syllabique a permis d'observer que les langues réutilisent les syllabes plutôt que de puiser dans l'ensemble des combinaisons possibles. Nous proposons, dans cette partie, de prolonger cet aspect en étudiant le lien existant entre le nombre de syllabes attestées (apparaissant dans le corpus) et le nombre de syllabes théoriques, c'est-à-dire toutes les combinaisons possibles de voyelles et de consonnes de chaque langue.

Nous avons calculé, pour chaque langue, le rapport syllabique, défini comme le nombre de syllabes effectivement rencontrées sur le nombre de syllabes théoriques. Le résultat (cf. table III.13) varie entre 0 (si aucune syllabe n'existait dans la langue), et 1 (si le nombre de syllabes différentes de la langue est égal au nombre de syllabes possibles).

Plus une langue possède de phonèmes, plus le nombre de combinaisons est important. Or, ce calcul exprime clairement qu'en n'utilisant pas tout leur potentiel syllabique, les langues restreignent leur inventaire de syllabes. Nous montrons que la proportion de restrictions diffère entre les structures syllabiques les plus fréquentes (cf. table III.13) et proposons, à la suite de cette partie, plusieurs études et analyses des cooccurrences et dépendances entre les constituants d'une même syllabe, mais aussi entre syllabes consécutives de l'unité lexicale.

Langues	CV			CVC		
	Syllabes attestées	Syllabes possibles	Rapport syllabique	Syllabes attestées	Syllabes possibles	Rapport syllabique
<b>afar</b>	167	180	0,92	871	3240	0,27
<b>finnois</b>	314	544	0,58	1157	9248	0,13
<b>français</b>	208	378	0,55	1285	7938	0,16
<b>kannada</b>	258	486	0,53	913	13122	0,07
<b>kanouri</b>	114	132	0,86	569	2904	0,20
<b>kwakw'ala</b>	218	252	0,87	719	10584	0,07
<b>navaho</b>	351	888	0,40	1027	32856	0,03
<b>ngizim</b>	304	518	0,59	492	19166	0,03
<b>nyahkur</b>	101	720	0,14	1106	21600	0,05
<b>quechua</b>	97	165	0,59	571	5445	0,10
<b>sora</b>	287	1050	0,27	1393	52500	0,03
<b>thai</b>	196	462	0,42	921	10164	0,09
<b>wa</b>	373	1026	0,36	1405	38988	0,04
<b>yup'ik</b>	175	410	0,43	740	16810	0,04
<b>!xóõ</b>	1274	18850	0,07	408	2450500	0,0002

Table III.13 : Nombre de syllabes observées, nombre de syllabes théoriques et rapport syllabique (observé/théorique) pour chacune des langues de notre base de données.

Dans le cas des syllabes CV, on remarque que pour huit langues sur quinze, le rapport syllabique est supérieur à 0,5. Cinq d'entre elles utilisent entre 50 et 60 % des possibilités théoriques tandis que, dans les trois autres, presque toutes les réalisations possibles de syllabes sont utilisées, puisque le rapport syllabique dépasse les 85 % atteignant même 92 % pour l'afar.

Pour les six autres langues, le rapport syllabique est inférieur à 0,5, c'est-à-dire que ces langues utilisent moins de 50 % des possibilités théoriques pour les syllabes CV (combinaisons de n voyelles et p consonnes de la langue). Ainsi, le navaho, le thai, le wa et le

yup'ik n'utilisent environ que 40 % des possibilités. Le nyah kur et le sora ont un rapport syllabique encore plus faible puisque ces langues n'utilisent que de 15 à 27 % des syllabes théoriques. Le !xóð, pour sa part, est une langue qui ne réalise qu'une petite minorité des syllabes possibles, puisque le rapport syllabique vaut 0,07.

Ces résultats renforcent l'idée qu'il existe des contraintes de combinaisons pour les syllabes CV (MacNeilage et Davis, 2000). En effet, si l'inventaire de l'afar se rapproche du nombre théorique, le nyah kur, le navaho ou le !xóð paraissent mettre en jeu des règles phonotactiques plus strictes qui expliqueraient le taux moins important de réalisations syllabiques rencontrées par rapport au nombre théorique.

Quant aux rapports syllabiques des langues pour les syllabes de type CVC, ils sont encore plus faibles. Dans aucune langue les syllabes observées ne correspondent à plus de 30 % des syllabes théoriques. On peut donc supposer que toutes les langues de notre échantillon sont peut-être soumises à des restrictions rendant optimales ou non les combinaisons de C et de V, ou de C et de C, pour la création de syllabes CVC.

Cependant, nous constatons des différences de comportement entre les langues de notre échantillon. En effet, le rapport syllabique concernant le navaho approche les 0,3, ce qui laisse supposer des restrictions fortes pour cette langue. La majorité des langues (onze sur quinze) ont un rapport syllabique inférieur à 0,1, ce qui signifie qu'elles n'utilisent qu'entre 5 et 10 % des syllabes théoriques. Les restrictions dans les syllabes CVC semblent donc être plus importantes que pour les structures CV.

L'étude des autres types de syllabes fait apparaître des résultats similaires : plus le nombre de constituants assemblés dans une même syllabe augmente, plus le nombre de syllabes théoriques croît. Or nous avons vu que le nombre de syllabes observées dans les structures autres que CV et CVC est très faible par rapport à ces deux structures dominantes. Il semble donc nécessaire d'étudier plus en détail, au niveau de la nature des constituants syllabiques (attaque, noyau et coda), les différentes combinaisons relevées dans notre échantillon afin de mettre en évidence des tendances qui pourraient relever de contraintes substantielles.

### **III.5.2. Méthodologie pour la recherche de cooccurrences et de dépendances**

Toutes les syllabes possibles ne sont donc pas attestées dans les langues. Il doit donc exister des regroupements consonne-voyelle, ou consonne-consonne, qui ne sont pas réalisés dans les langues. Afin de les mettre en évidence, nous avons établi automatiquement des matrices permettant de les faire émerger. En voici un exemple pour l'afar (cf. table 14).



	t	l	b	s	d	r	m	k	ʔ	n	g	ɖ	j	ħ	f	w	h	p	Total
<b>a</b>	219	326	231	237	294	235	353	263	296	212	212	153	144	104	173	139	70		<b>3661</b>
<b>i</b>	99	120	182	191	145	166	85	120	118	99	122	67	96	64	37	27	27		<b>1765</b>
<b>e</b>	237	64	139	168	72	104	57	67	90	56	55	31	50	32	54	61	9	3	<b>1349</b>
<b>o</b>	147	82	81	56	58	80	50	87	46	54	43	40	37	25	23	8	13		<b>930</b>
<b>u</b>	81	81	57	53	78	71	80	96	50	87	31	29	36	35	10	8	14		<b>897</b>
<b>a:</b>	27	70	48	38	53	47	71	39	41	41	42	36	15	24	13	19	11		<b>635</b>
<b>o:</b>	8	32	10	16	15	11	6	16	7	14	19	7	7	7	2	2	7		<b>186</b>
<b>e:</b>	9	15	17	16	24	14	7	7	7	10	0	18	6	10	2	7	3		<b>172</b>
<b>i:</b>	5	8	23	12	19	16	9	15	4	8	10	19	4	10	1	3			<b>166</b>
<b>u:</b>	11	24	9	10	12	7	10	9	8	10	4	13	2	4			2		<b>135</b>
<b>Total</b>	<b>843</b>	<b>822</b>	<b>797</b>	<b>797</b>	<b>770</b>	<b>751</b>	<b>728</b>	<b>719</b>	<b>667</b>	<b>591</b>	<b>538</b>	<b>413</b>	<b>397</b>	<b>315</b>	<b>315</b>	<b>274</b>	<b>156</b>	<b>3</b>	<b>9896</b>

Table III.14 : Matrice de cooccurrences des syllabes CV en afar, avec leur nombre d'apparition dans le lexique. Les matrices sont présentées selon la fréquence décroissante des constituants en position d'attaque (colonnes) et de noyau (ligne). Le total des totaux correspond au nombre de syllabes de l'afar, soit 9 896.

Il s'est avéré que l'élaboration de telles matrices ne permettait pas de mettre directement en évidence des tendances fortes entre consonne et voyelle, par exemple. Les relations semblent plutôt apparaître entre un type de consonne particulier (fonction du lieu ou du mode d'articulation) à un type de voyelle (selon le timbre vocalique, le degré d'aperture ou la position de la langue). Il est par conséquent nécessaire dans notre démarche d'effectuer des regroupements (équivalences typologiques) entre consonnes ou voyelles ayant des caractéristiques phonétiques communes afin d'obtenir des résultats plus consistants.

À la suite de travaux antérieurs (Kawasaki, 1982 ; Janson, 1986 ; MacNeilage et Davis, 2000) qui mettent en évidence des liens entre le lieu d'articulation de la consonne et la position antéro-postérieure de la langue pour les voyelles dans les syllabes de type CV, nous avons effectué un certain nombre de regroupements. Les tables suivantes présentent nos choix d'équivalences typologiques pour les consonnes et les voyelles (cf. tables III.15 et III.16). Dans le regroupement des consonnes, nous avons basé notre réflexion sur les travaux de Keating (1990), qui précisent que, dans une même langue, les oppositions distinctives entre consonnes coronales de même mode articuloire sont extrêmement rares, à l'exception des affriquées.

Lieu d'articulation	Consonne (API)
bilabial	p b m
labio-dental	f v
coronal	t d n s z ʃ ʒ ʧ ʤ l r r tʃ ts tʃ tʃ dj dz dʒ dʒ t d ŋ ʃ r l
palatal	c ɲ j j ʎ
vélaire	k g ŋ ɣ ŋg x
uvulaire	q ɢ ʁ ʁ
pharyngal	ʕ h
glottal	ʔ h h <sup>w</sup>
labio-vélaire	w ɱ w̥
labio-palatale	ɥ

Table III.15 : Regroupement des consonnes selon leur lieu d'articulation.

Traits vocaliques	Voyelles (API)
antérieur non arrondi	i i e ε æ
Central	a ə i
postérieur arrondi	u u o o
antérieur arrondi	y œ ø
postérieur non arrondi	ɯ ɣ

Table III.16 : Regroupement des voyelles selon leur point d'articulation.

Nous avons aussi décidé de réunir les consonnes en catégories, en fonction de leur mode d'articulation (cf. table III.17).

Mode d'articulation	Consonne (API)
plosif	p b t t̥ d d̥ c k g q ɢ ʔ
fricatif	f s ʃ ʒ h ɦ ɸ χ
nasal	m n ŋ ɲ ŋ
vibrant	r r̥ ɾ
approximant	j w ɥ
affriqué	dʒ dz dʒ dʒ̥ tʃ ts tʃ tʃ̥

Table III.17 : Regroupement des consonnes selon leur mode articuloire.

Deux types de matrices sont alors réalisées. Les premières présentent, en tenant compte des équivalences typologiques, les occurrences observées pour chaque combinaison entre une consonne et une voyelle, par exemple dans le cas des structures CV de l'afar (cf. table III.18). Ce type de matrices permet de connaître les lieux d'articulation les plus fréquemment utilisés dans les langues, mais ne permettent cependant pas de mettre en évidence des combinaisons favorisées. Pour résoudre ce problème, nous avons pondéré les occurrences des structures en fonction des fréquences respectives des segments vocaliques et consonantiques qui la composent. Par exemple, la fréquence de la syllabe /ta/ est divisée par le produit des fréquences de /t/ et de /a/ dans le lexique. Les ratios alors obtenus permettent de mettre en évidence si une combinaison est ou non favorisée par rapport aux segments qui la composent. Ainsi, un ratio supérieur à 1 voudra dire que la combinaison est favorisée, un ratio inférieur à 1 que la structure est défavorisée, et un ratio égal à 1 ne permet pas de conclure. On obtient alors un deuxième type de matrices (cf. table III.19) desquelles nous avons tiré un ensemble de tendances au niveau des cooccurrences entre constituants syllabiques.

Occurrences	Bi	Lde	Co	Pa	Ve	Ph	Lve	Gl	Total
<b>An</b>	368	116	1980	94	382	375	98	39	<b>3452</b>
<b>Ce</b>	820	128	1872	186	555	496	158	81	<b>4296</b>
<b>Po</b>	365	71	1057	35	373	193	18	36	<b>2148</b>
<b>Total</b>	<b>1553</b>	<b>315</b>	<b>4909</b>	<b>315</b>	<b>1310</b>	<b>1064</b>	<b>274</b>	<b>156</b>	<b>9896</b>

Table III.18 : Matrice des syllabes CV en afar, après regroupement des consonnes par lieu d'articulation et des voyelles suivant la position antéro-postérieure de la langue (voir annexe VIII pour les abréviations).

Ratio	Bi	Lde	Co	Pa	Ve	Ph	Lve	Gl
<b>An</b>	0,68	1,06	<b>1,16</b>	0,86	0,84	1,01	1,03	0,72
<b>Ce</b>	<b>1,22</b>	0,94	0,88	1,36	0,98	1,07	1,33	1,20
<b>Po</b>	1,08	1,04	0,99	0,51	<b>1,31</b>	0,84	0,30	1,06

Table III.19 : Ratios entre les occurrences attestées et les occurrences attendues pour les combinaisons C\*V pour l'ensemble des syllabes de l'afar (voir annexe VIII pour les abréviations).

### III.5.3. Cooccurrences et dépendances entre attaque et noyau d'une même syllabe

Nous avons montré dans la partie III.5.1, avec le calcul du rapport syllabique, que pratiquement toutes les voyelles et la majorité des consonnes d'une langue peuvent apparaître dans une syllabe de type CV. Pourtant, l'analyse des matrices de cooccurrences entre les constituants de ce type de syllabe pour chaque langue, montre qu'il existe des combinaisons largement favorisées. Celles-ci figurent dans la table III.20, dans laquelle nous avons choisi de regrouper l'ensemble des autres combinaisons minoritaires.

Ratio	Co	Bi	Ve	Autres
<b>An</b>	<b>1,09</b>	0,94	1,05	1,04
<b>Ce</b>	0,87	<b>1,10</b>	0,70	1,05
<b>Po</b>	1,01	0,93	<b>1,32</b>	0,93
<b>Autres</b>	1,09	1,06	1,10	0,73

Table III.20 : Ratios entre les occurrences attestées et les occurrences attendues pour les combinaisons C\*V pour l'ensemble des syllabes de l'ensemble de notre échantillon ( $\chi^2$  significatif,  $p < 0.001$  pour chaque colonne).

L'étude des ratios pour l'ensemble des syllabes de notre échantillon nous permet de mettre en évidence trois types de combinaisons favorisées :

- les consonnes labiales avec les voyelles centrales ;
- les consonnes coronales avec les voyelles antérieures non-arrondies ;
- les consonnes vélares avec les voyelles postérieures arrondies.

Le regroupement des autres types de consonnes fait émerger une préférence nette pour les voyelles centrales, mais qui est essentiellement due à la forte proportion de ce type de voyelle dans les langues de notre échantillon (de 12.5 % à 50 % des noyaux selon les langues) (cf. partie III.4.2).

Ces résultats portent sur l'ensemble des types syllabiques possédant, en intra-structure, la combinaison CV, avec C en position d'attaque : CV, CVC, CCV, CCVC, ... Le détail par type de structures permet de conclure sur la non-existence de particularités propres à certaines structures. Pour cela, nous avons calculé les ratios entre les occurrences observées et attestées des combinaisons CV pour deux types de structures : CV et CVC. Aucun changement significatif n'est observable entre les résultats portant sur les structures ouvertes (à coda vide) ou fermées (avec coda) ou sur l'ensemble des types syllabiques (cf. figure III.25).

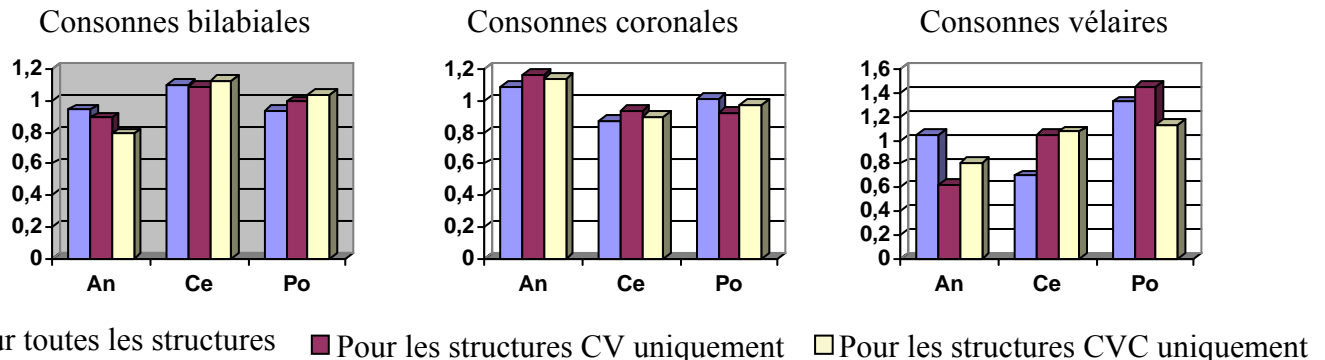


Figure III.25 : Interactions entre attaque et noyau pour les différents types de structures.

Il semble bien, par conséquent que, quel que soit le type de structure syllabique, les langues présentent une tendance forte à favoriser des combinaisons attaque\*noyau qui minimisent les déplacements articulatoires entre les gestes consonantique et vocalique. L'économie de geste articulatoire serait donc un des principes organisateurs de l'unité syllabique. Nous discuterons ce propos dans la partie III.5.7 en comparant nos résultats à ceux d'études antérieures.

### III.5.4. Analyse des constituants successifs, voyelle et consonne, de la rime

L'étude des relations entre le noyau et la coda devrait permettre de savoir si ce principe s'applique, comme le dit MacNeilage (1998), à toute l'organisation syllabique ou si certains constituants sont soumis à des contraintes plus fortes voire à d'autres contraintes.

Ratio	An	Ce	Po	Autres
<b>Co</b>	<b>1,11</b>	1,00	0,89	0,91
<b>Bi</b>	0,53	<b>1,32</b>	1,18	0,71
<b>Ve</b>	0,92	0,86	<b>1,22</b>	1,35
<b>Autres</b>	1,01	0,92	1,06	1,17

Table III.21 : Ratios entre les occurrences attestées et les occurrences attendues pour les combinaisons V\*C pour l'ensemble des syllabes de l'ensemble de notre échantillon ( $\chi^2$  significatif,  $p < 0.001$  pour chaque colonne).

L'analyse des ratios existant entre occurrences observées et occurrences attendues en fonction de la fréquence des constituants montre que, comme pour les structures CV, des combinaisons sont favorisées entre noyau et consonne en coda :

- les voyelles antérieures non-arrondies et les consonnes coronales ;
- les voyelles centrales et les consonnes labiales ;
- les voyelles postérieures arrondies et les consonnes vélares.

Les tendances qui émergent sont les mêmes que celles obtenues dans les structures CV. Il semble que les contraintes qui pèsent sur les combinaisons entre voyelle et consonne dans le cadre de la rime syllabique sont de nature à minimiser les mouvements des articulateurs.

Comme pour les structures CV, nous avons analysé séparément les divers types de structures syllabiques, et nos résultats ne montrent pas de différences majeures en fonction de la présence ou non d'une attaque (VC, CVC, VCC, CVCC, ...).

Si l'on s'intéresse maintenant aux ratios en eux-mêmes, on constate que le lien entre voyelle et consonne semble plus important entre noyau et coda qu'entre attaque et noyau. Cette observation soutient l'hypothèse faite par de nombreux phonologues quant à l'existence d'une rime, niveau intermédiaire entre le noyau et la coda, et la syllabe elle-même, établie sur des observations linguistiques en diachronie qui montrent que le noyau et la coda ont une évolution plus liée que celles entre l'attaque et le noyau (cf. théories phonologiques, partie I.1).

### **III.5.5. Cooccurrences et dépendances entre attaque et coda d'une même syllabe**

L'analyse des autres structures syllabiques fermées de type CVC s'est montrée intéressante pour les relations entre la nature phonétique de la consonne en attaque et les caractéristiques de celle en position de coda. En effet, le pourcentage de consonnes pouvant apparaître en attaque comparé à celui des consonnes se rencontrant en coda montre que, même s'il existe une grande disparité entre les langues, les contraintes sur la position de coda sont globalement bien plus importantes que celles sur la position d'attaque (cf. parties III.4.4. et III.5.1).

Les matrices d'occurrences pour les syllabes de types CVC, en considérant le lieu d'articulation des consonnes de l'attaque et de la coda, font apparaître des relations moins favorisées que d'autres. Comme dans le cas des cooccurrences consonne\*voyelle, les résultats ont été pondérés en fonction des fréquences d'apparition des consonnes en attaque et en coda. La comparaison entre les matrices d'occurrences et les matrices pondérées montre que les consonnes coronales, qui sont pourtant très majoritaires dans toutes nos langues, en attaque comme en coda, sont défavorisées dans les structures de type CVC lorsque attaque et coda partagent le même lieu d'articulation (cf. table III.22). La même situation est observée pour toutes les autres consonnes. Si le principe d'économie articulatoire ne peut être invoqué pour

expliquer ce résultat, d'autres pistes, notamment perceptives, seront discutées dans la partie III.5.7.

<b>Ratio</b>	<b>Bi</b>	<b>Co</b>	<b>Ve</b>
<b>Bi</b>	<b>0,45</b>	1,18	1,04
<b>Co</b>	1,05	<b>0,95</b>	1,15
<b>Ve</b>	0,97	1,19	<b>0,49</b>

Table III.22 : Ratios entre les occurrences attestées et les occurrences attendues pour les combinaisons C\_C (attaque\_coda) pour l'ensemble des syllabes de l'ensemble de notre échantillon ( $\chi^2$  significatif,  $p < 0.001$  pour chaque colonne).

Concernant la fréquence des structures  $C_1VC_2$  dominantes dans les langues d'ULSID, dans lesquelles  $C_1$  et  $C_2$  n'ont pas le même lieu d'articulation, il est intéressant de remarquer qu'il existe plusieurs combinaisons favorisées :

- consonnes bilabiales et consonnes coronales ;
- consonnes coronales et consonnes bilabiales ;
- consonnes coronales et consonnes vélares ;
- consonnes vélares et consonnes coronales.

Pourtant, si nous comparons le nombre d'occurrences de ces différentes combinaisons dans les langues observées, nous constatons que les résultats sont variables. En effet, la structure labial-coronal est beaucoup plus fréquente dans les langues que le patron inverse coronal-labial. La première combinaison est majoritaire dans 9 langues sur 13, avec un ratio (labiale\*v\*coronale) / (coronale\*v\*labiale) de 1,44. Pour ce qui est des structures avec consonnes coronales et consonnes vélares, les résultats sont plus mitigés puisque les combinaisons labial-vélaire ne sont favorisées que dans 7 langues sur 13.

Il semble donc bien que des contraintes pèsent sur les consonnes des structures CVC. Ces résultats permettent de vérifier l'hypothèse de MacNeilage et Davis (2000) sur l'apparition, avec les premiers mots du langage chez les enfants, d'un « effet LC » prédominant qui implique la plus forte fréquence des formes commençant par une consonne labiale et suivie d'une consonne coronale au détriment de la forme inverse.

### **III.5.6. Recherche de l'« effet LC » de MacNeilage et Davis (2000) aux niveaux intra et intersyllabique**

MacNeilage et Davis (2000) mettent en évidence l'existence d'un « effet LC » qui, bien qu'absent du babillage, apparaît avec les premiers mots et se retrouve dans l'étude de 10 langues naturelles. Dans leur analyse, MacNeilage et Davis ne s'intéressent qu'aux structures dissyllabiques dans lesquelles les consonnes labiales et coronales sont les attaques respectives des deux syllabes consécutives.

Ayant validé l'existence de l'« *effet LC* » en intra-syllabique, c'est-à-dire entre attaque et coda d'une même syllabe, il nous a semblé intéressant de vérifier son existence dans les dissyllabes provenant des langues de notre échantillon. Dans ce but, nous avons comparé la répartition des consonnes aux niveaux intra- et inter-syllabiques. Les résultats portent sur des items dissyllabiques de type CV .CV contenus dans les lexiques de 10 langues sur les 15 de notre échantillon, dans lesquels les occurrences de cette structure lexicale sont supérieures à 10 entrées : afar, finnois, français, kannada, kwakw'ala, navaho, ngizim, quechua, sora, yup'ik.

Pour 8 des 10 langues (cf. l'exemple de l'afar, tables III .23 et III.24), le nombre de séquences labiale\*V .coronale\*V est effectivement plus important que le nombre de formes coronale\*V .labiale\*V (test binomial,  $p < 0.005$ , pour 7 d'entre elles) avec un rapport de 1.73. La table III.24 permet, quant à elle, de constater que sur le total des unités lexicales CV .CV, les syllabes labiale\*V .coronale\*V sont cette fois nettement plus fréquentes que les séquences inverses coronale\*V .labiale\*V : le ratio s'élève à 2.39.

	<b>Labiale</b>	<b>Coronale</b>	<b>Vélaire</b>
<b>Labiale</b>	4	<b>30</b>	10
<b>Coronale</b>	<b>52</b>	79	39
<b>Vélaire</b>	10	27	0

Table III.23 : Matrice des cooccurrences consonantiques dans les structures  $C_1 V . C_2 V$  en afar.

	<b>Labiale</b>	<b>Coronale</b>	<b>Vélaire</b>
<b>Labiale</b>	65	<b>241</b>	83
<b>Coronale</b>	<b>577</b>	1005	582
<b>Vélaire</b>	104	253	75

Table III.24 : Matrice des cooccurrences consonantiques dans les structures  $C_1 V . C_2 V$  pour les 10 langues retenues.

Nos résultats montrent que sur le total des unités lexicales CV .CV, les structures labiale-voyelle-coronale sont nettement plus fréquentes que les structures coronale-voyelle-labiale. Le ratio LC / CL est de 2.39, ce qui correspond à celui de MacNeilage et Davis (2000) (ratio : 2.55).

MacNeilage et Davis (2000) ont, de plus, vérifié leurs résultats sur les 27 racines fournies par Ruhlen (1997). L'implantation de la base de données lexicales provenant de langues et de proto-langues fournie par Ruhlen (1997) (cf. chapitre II, partie 5), nous permet de vérifier si les patrons syllabiques favorisés dans les langues du monde se retrouvent dans les 2 825 formes et proto-formes phonétiques correspondant à tous les signifiés des 27 racines universelles. L'analyse des gabarits montre la présence forte de « *l'effet LC* » observée par MacNeilage et Davis (2000) dans les 27 racines principales. Le ratio (Labial-



Coronal)/(Coronal-Labial) est de 10,25. Ce premier résultat sera complété par une étude semblable en séparant les langues des proto-langues ainsi qu'une recherche des autres combinaisons observées comme étant favorisées dans les langues du monde.

### III.5.7. Discussion

Nos résultats peuvent être comparés avec ceux obtenus lors d'études antérieures sur les structures CV des langues du monde. Ainsi, Kawasaki (1982) conduit son analyse afin de trouver des explications possibles pour les contraintes phonotactiques générales mises en évidence par les analyses de grands échantillons de langues. Kawasaki reprend les données sur l'absence ou la rareté de certaines combinaisons dans les descriptions phonologiques de différentes bases de données (tirées de Crothers *et al.*, 1979) afin d'en déduire les structures favorisées.

Un grand nombre de propositions ont été apportées, avant Kawasaki, pour expliquer la présence de contraintes séquentielles universelles. Ainsi, Hooper (1976) propose des hiérarchies basées sur la sonorité/la force ; Diver (1975) avance des contraintes séquentielles des gestes articulatoires ; Saporta (1955) ou Cutting (1975) suggèrent un degré de contraste perceptif nécessaire entre des phonèmes successifs. Cependant, la majorité de ces tentatives ne permettent pas d'expliquer l'ensemble des universaux observés.

Kawasaki (1982) reprend l'hypothèse de Cutting selon laquelle les séquences sonores favorisées les plus fréquentes sont celles où les sons qui les constituent sont maximale-ment différents, et teste la validité de cette hypothèse non pas en termes de traits distinctifs comme Cutting, mais en termes de paramètres acoustiques : « *The present study proposes a hypothesis that they [some universally attested sequential constraints] are, to a large extent, acoustically and auditorily based. One determinant, the magnitude of acoustic modulation in a given sound sequence, assumed to be directly related to its perceptual saliency, should affect the viability of the sequence. Another determinant, the degree of acoustic difference among various sound sequences, should be inversely related to the likelihood of perceptual confusion and hence predict the sequence which are less susceptible to merger than others* » (Kawasaki, 1982, p. 4).

Son hypothèse est testée expérimentalement sur les contraintes séquentielles suivantes :

- plosive + liquide ;
- plosive + glide ;
- consonne + voyelle ;
- l'opposition entre CV et VC.

L'hypothèse acoustique de Kawasaki (1982) (cf. chapitre I, partie 2.1.) fournirait une voie possible pour expliquer les tendances universelles qui suppriment les groupements longs de consonnes et de voyelles, les modulations acoustiques étant relativement faibles dans ces séquences. La même explication est avancée pour la tendance à éliminer les séquences de

consonnes produites avec le même mode articulatoire, la rareté des groupes nasale+liquide par rapport aux groupes obstruente+liquide, ainsi que la rareté des nasales postérieures en début de mot.

La thèse de Kawasaki montre qu'un grand nombre de contraintes phonotactiques peuvent s'expliquer en grande partie par des phénomènes auditori-acoustiques, fournissant une confirmation supplémentaire du rôle significatif de l'auditeur dans la détermination des structures phonologiques et l'évolution des changements. Cependant, même s'il s'agit d'une étude qui ne se penche pas sur l'aspect articulatoire, sa conclusion n'exclut pas la possibilité que certains universaux phonotactiques soient articulatoirement déterminés.

Janson (1986) propose une étude des séquences C + V présentes dans un corpus de textes de plusieurs langues. Bien que le matériel étudié soit différent de celui de Kawasaki (1982), il considère que les séquences CV sont déterminées par les propriétés articulatoires et perceptives des sons qui les composent : « *there are good reasons to believe that these two kinds of facts should usually be well correlated, and both kinds have been used in earlier typological investigations. [...] Thus the lack of correspondence in the results is a matter of real concern.* » (Janson, 1986, p. 181).

La méthode utilisée par Janson consiste à comparer la fréquence d'une séquence CV aux fréquences respectives de la consonne et de la voyelle qui la composent. Toute variation du ratio (séquence/sons) permet de mettre en évidence le caractère favorisé ou non des séquences. Pour traiter automatiquement les données, les langues choisies ont toutes un dictionnaire informatisé et une graphie peu différente du système phonologique : finnois, turc, latin et letton. Le setswana a été ajouté à partir de données personnelles. La proximité géographique de la plupart de ces langues peut avoir des conséquences sur les similitudes observées, mais, pour Janson, il est néanmoins raisonnable de croire qu'une tendance très nette obtenue à partir de ces 5 langues peut indiquer des faits universels.

Les systèmes phonologiques des cinq langues étant différents, des regroupements au niveau des unités sonores ont été effectués par Janson pour rendre la comparaison possible. Les voyelles sont regroupées dans 4 classes : I (voyelles antérieures non arrondies), A (voyelles centrales), U (voyelles postérieures arrondies) et Y (voyelles antérieures arrondies). Pour chaque syllabe, Janson calcule pour chaque séquence la fréquence observée, ainsi que la fréquence attendue (fréquence de la consonne \* fréquence de la voyelle). Les consonnes sont regroupées dans les classes suivantes : dentale/alvéolaire, vélaire, bilabiale et liquide.

Janson tire les conclusions suivantes par classe de consonnes :

Séquences favorisées :

- consonnes dentales/alvéolaires + voyelles antérieures ;
- consonnes vélaire + voyelles postérieures ;
- consonnes occlusives + voyelles postérieures ;
- consonnes liquides + voyelles antérieures ;

Séquences défavorisées :

- consonnes vélares + voyelles antérieures ;
- consonnes liquides + voyelles postérieures.

Les trois consonnes occlusives sourdes /p t k/ montrent une préférence pour une association avec les voyelles postérieures, comme les consonnes occlusives sourdes aspirées du setswana. Les résultats de Janson (1986) sont en contradiction avec ceux de Kawasaki (1982). Il pense que les généralisations faites par Kawasaki sont trop rapides et vont plus loin que ne le permettent les résultats. Quoiqu'il en soit, il note que les combinaisons favorisées, de son corpus, sont celles qui nécessitent un déplacement moindre des articulateurs entre le geste consonantique et le geste vocalique. Mais Janson, lui-même, ne rejette pas l'idée que les capacités perceptives sont aussi importantes que les capacités de production, et que des contraintes contradictoires à ces deux niveaux peuvent interagir dans l'organisation syllabique des langues du monde.

À partir des données d'ULSID, Maddieson (1992) propose une vérification des résultats de Janson, en tentant de déterminer les rôles de la minimisation de l'effort articulo-moteur et de la maximisation du contraste acoustico-perceptif, deux principes qui structureraient les systèmes sonores.

Maddieson reprend les données, la méthodologie et les conclusions de Janson (1986) afin de les analyser dans une optique qui prend en compte l'ampleur et la consistance des déviations entre les résultats observés et les résultats attendus. Bien que l'interaction des voyelles soit significative dans une analyse des variances, l'étude au cas par cas des différentes combinaisons montre qu'aucune n'est significative au critère de  $p < 0,05$  (t-test). Afin de savoir si l'impossibilité de conclure de façon nette vient de l'absence de facteurs phonétiques dans les langues ou de biais dans l'établissement et l'analyse des données, Maddieson reprend les résultats provenant d'une étude plus large. Ses données proviennent de lexiques plutôt que de textes et son échantillon respecte des critères de diversité génétique et géographique. Les cinq langues choisies par Maddieson possèdent un inventaire de segments limité et une phonotactique simple afin de pouvoir facilement inclure toutes les syllabes dans les comptes.

Les données lexicales sont analysées en syllabes CV. Pour chaque consonne, Maddieson calcule sa fréquence ainsi que la fréquence des voyelles apparaissant avec elle (/i/ apparaît dans 16 % des syllabes commençant par un /p/). Il cherche aussi la fréquence d'apparition totale de chaque voyelle (/i/ apparaît dans 19 % des syllabes), et calcule alors la déviation pour la syllabe /pi/ : soustraction de la fréquence totale à la fréquence avec /p/. Il obtient donc une déviation de -3. Cette méthode est analogue à celle de Janson. Après une comparaison entre la moyenne en fonction du nombre de langues et celle en fonction du nombre de segments, Maddieson décide de garder la seconde qui offre un plus grand nombre de points de données. Pour les consonnes labiales, il ne trouve pas de préférence marquée avec les voyelles arrondies, mais plutôt avec la voyelle centrale basse /a/. En ce qui concerne les consonnes

coronales, la déviation avec les voyelles d'avant est très légèrement positive (comme cela était prévisible du fait du principe de minimisation des coûts articulatoires). Pour les vélares, les prédictions articulatoires et acoustiques sont contradictoires, et ne peuvent être confirmées puisque les déviations varient en fonction des langues.

Cette étude ne met donc en évidence aucune raison de donner la préférence à une explication plutôt qu'à une autre (ni articulatoire, ni acoustique). En revanche, deux résultats sont particulièrement saillants. Tout d'abord, les glides et les articulations secondaires tendent à être évités avec les voyelles proches (\*[wu], \*[ji]). Ensuite, les consonnes vélares semblent largement défavorisées avec les voyelles hautes d'avant. Ces résultats peuvent respectivement être attribués à des facteurs auditifs et articulatoires. Mais dans l'ensemble, les déviations étant faibles, la fréquence des combinaisons CV peut être retrouvée à partir des fréquences des segments qui la composent. Maddieson suppose que les langues ayant fait un choix de segments et déterminé leur fréquence relative, elles ne restreignent pas davantage, pour la plupart, les combinaisons dans lesquelles les segments peuvent entrer. Cela rejoindrait, à un niveau syllabique, la tendance des langues à faire « *une utilisation maximale des traits distinctifs disponibles* » à un niveau segmental, plus connue sous l'appellation de « Principe du MUAFF » (*Maximal Use of Available Feature*) de Ohala (1980).

Maddieson (1993) propose alors une analyse de 27 autres langues, qui rappellent le, figurent dans les données d'ULSID, et qui a permis de livrer :

- (1) le nombre d'occurrences de chaque attaque, chaque noyau et chaque coda ;
- (2) le nombre d'occurrences de chaque combinaison attaque/noyau et noyau/coda ;
- (3) le nombre d'occurrences des syllabes entières.

Cette étude des séquences de consonnes, focalisée sur la sonorité, contrairement aux précédentes, s'intéresse à la fréquence des séquences CV afin de cerner l'objet des préférences présidant à la structuration sonore des langues. Maddieson ne propose pas de dichotomie entre structures possibles et structures inexistantes, il place les syllabes CV sur un continuum allant des plus acceptables aux moins acceptables, y compris celles interdites par la langue. L'hypothèse de Maddieson étant que les lois imposées aux syllabes CV (type le plus fréquent) doivent être les mêmes que pour les autres types, tels que CVC, CCV ..., il se contente alors d'étudier les structures CV.

L'étude d'un échantillon plus large que celui de la première étude montre une forte influence du principe d'économie articulatoire, mais aussi de facteurs acoustiques. Maddieson (1993) n'exclut pas non plus l'existence de facteurs cognitifs, mais sans plus de précision.

L'économie articulatoire façonne les séquences avec consonnes dorsales (vélaire, uvulaire et palatale). Dans un grand nombre de langues, les voyelles d'arrière sont privilégiées avec ces consonnes. Il existe d'autres cas pour lesquels une assimilation de lieu entre la voyelle et la consonne est observée. Ainsi, Maddieson cite les cas du chinois standard et du polonais, langues dans lesquelles seules les consonnes palatales ou palatalisées peuvent apparaître devant /i/.

Dans la méthodologie de Maddieson (1993), les probabilités de C et de V sont multipliées afin de donner celles de la syllabe CV, comme si les segments qui la composent étaient indépendants. Les valeurs attendues et observées sont comparées pour chaque syllabe. Les résultats (observés comme pourcentages de l'attendu) pour /Ki/ et /Ku/ dans les seize langues (avec K = k plus les autres consonnes vélaires de la langue) sont analysés en comparaison avec les fréquences de /K/ et des voyelles concernées. Pour /Ku/, la préférence n'est pas uniforme (entre 202 et 100) pour les douze premières langues, et pour les quatre dernières, le pourcentage est même inférieur à 100. Pour l'association consonne vélaire et voyelle d'avant, la majorité des langues ne la favorise pas. Pour le chinois standard et le polonais, toutes les combinaisons seront possibles, mais elles seront moins exploitées. La moyenne pour /Ku/ est de 115 %, mais elle n'est pas significativement différente de 100 (t-test). Par contre pour /Ki/, la moyenne de 70 % est significativement différente de 100. Comme les consonnes vélaires sont produites à partir d'un mouvement de la langue et que la différence entre les voyelles d'arrière et d'avant dépend aussi de la langue, il est naturel de dire, pour Maddieson, que ce résultat résulte de la tendance à minimiser les mouvements articulatoires. Contrairement aux consonnes vélaires, pour les coronales (alvéolaire ou dentale) le corps de la langue ne réalise pas directement l'occlusion, laissant supposer que les coronales vont être les plus indépendantes des voyelles. Maddieson présente les résultats statistiques pour /Ti/ et /Tu/ et remarque qu'il n'y a aucune combinaison favorisée ou défavorisée en fonction de la voyelle (aucune des deux moyennes n'est significativement différente de 100).

Maddieson (1992) affirme que les biais de fréquence pour certaines syllabes CV ne démontrent pas « *a general strategy to reduce articulatory trajectories between adjacent segments* » (Maddieson, 1992, p. 4), mais révèlent plutôt l'influence d'une tendance plus limitée à réduire les mouvements du corps de la langue. D'un autre côté, le changement articulatoire minimum entre la consonne et la voyelle pourrait résulter d'un contrôle minimal. Dans les langues de notre échantillon, les séquences avec glide et voyelle correspondante (par exemple, /ji/ et /wu/) sont également défavorisées et /ju/ et /wi/ sont favorisées. Ces résultats seront discutés plus tard, mais ils vont à l'encontre de l'argument d'une recherche d'économie articulatoire, car /ji/ et /wu/ ne nécessitent que peu de mouvement. Lorsque le critère de distinctivité perceptive est rempli, le patron syllabique basé sur la simplicité articulatoire sera fonctionnel.

D'autres résultats de l'étude inter-langues semblent plutôt devoir être attribués aux propriétés inhérentes de certains segments qui font d'eux les meilleurs, ou les moins bon, pour la construction de structures plus larges. C'est pour vérifier cette hypothèse que Maddieson (1993) rappelle les travaux sur la théorie de la dispersion (Lindblom et Liljencrants, 1972) et ceux sur la théorie quantique (Stevens, 1972). Dans les données de Maddieson, la voyelle centrale basse /a/ est prédominante par rapport aux autres voyelles (très nettement pour 7 langues). Maddieson (1993) remarque que /a/ est beaucoup plus populaire que ne le prédisent ces théories. Il semble raisonnable de proposer que ces préférences viennent du fait que /a/

présente une plus grande amplitude que n'importe quelle autre voyelle, et que cela est dû à la proximité de F1 et F2 associée à une position extrême des articulateurs : « *the absence of any attenuation resulting from a narrow outlet* » (Maddieson, 1993, p. 6). La forte présence du /a/ semble davantage liée à ses propriétés inhérentes qu'à des relations entre voyelles et consonnes dans la syllabe. Ces résultats rejoignent ceux obtenus à partir des langues d'ULSID, analysées dans ce travail (cf. partie III.4.3).

L'ensemble de ces travaux soutient en fait l'idée de l'existence d'une influence commune de l'économie articulatoire et de la distinctivité acoustique dans la préférence de certaines séquences, plutôt que la domination de l'une sur l'autre. Ces processus sont évoqués lors d'études de grandes quantités de données tirées d'échantillons de langues convenablement structurés pour fournir suffisamment de données indépendantes : des tendances générales n'émergeront qu'à partir de larges échantillons et les déviations particulières des langues par rapport aux préférences générales peuvent être assez marquées. À côté des résultats de préférences partielles, des préférences pour des types de segments particuliers apparaissent alors, suggérant que les explications des données phonologiques peuvent inclure une évaluation des avantages et des désavantages inhérents à un segment individuel, aussi bien que les transitions entre segments dans la parole. De plus, l'utilisation partielle du potentiel syllabique observée pourrait venir de facteurs cognitifs quant à l'efficacité du stockage et de l'accès aux connaissances linguistiques.

MacNeilage et Davis (2000) proposent, dans le cadre de la théorie « *Frame then Content* », une étude des structures CV favorisées dans 10 langues à la suite de la mise en évidence de 3 séquences favorisées dans le babillage et les premiers mots des enfants (Davis et MacNeilage, 1995). Les combinaisons favorisées qu'ils remarquent sont les suivantes (cf. table III.25) :

- consonnes labiales avec voyelles centrales ;
- consonnes coronales avec voyelles d'avant ;
- consonnes dorsales avec voyelles d'arrière.

	<b>C Labiale + V centrale</b>	<b>C Coronale + V d'avant</b>	<b>C Dorsale + V d'arrière</b>
<b>Babillage</b>	1,34	1,28	1,22
<b>Premiers mots</b>	1,29	1,48	1,39
<b>Langues</b>	1,10	1,16	1,27

Table III.25 : Résultats sur les modèles CV favorisées dans le babillage de six enfants, les premiers mots de 10 enfants et les mots de 10 langues (MacNeilage et Davis, 2000). Les résultats sont obtenus en calculant les ratios entre fréquence observée de la syllabe et fréquence calculée en fonction de celles des constituants syllabiques.

Leurs résultats vont dans le sens de l'existence d'un coût minimal de production, qui vise à limiter la mise en mouvement des autres articulateurs pour privilégier le mouvement d'ouverture et de fermeture du conduit vocal grâce à la mâchoire qui posséderait un coût minimal de production. Il s'agit d'une hypothèse qui rejoint celle, bien connue, de l'économie articulatoire (cf. section 2.2.2 du chapitre I pour plus de détails).

Nos résultats sont en accord avec ceux de MacNeilage et Davis (2000), mais seulement partiellement avec ceux des études précédentes. Les combinaisons favorisées dans nos langues soutiennent donc l'hypothèse de la recherche d'une économie articulatoire entre l'attaque et le noyau d'une syllabe.

Nos résultats sur l'analyse des relations entre consonnes en position d'attaque et de coda dans une même syllabe montre que si des combinaisons favorisées existent, une tendance forte à défavoriser les consonnes de même lieu d'articulation est aussi présente (/pat/ vs. /tat/ ou /pap/), alors que les structures CVC avec attaque et coda de même mode articulatoire sont possibles. Nous en avons conclu que le principe de parcimonie pour le trait de lieu dans la syllabe CVC ne fonctionne pas. Ce résultat, s'il va dans le sens d'un coût minimal de production pour la syllabe (initiée par un cycle de l'oscillation mandibulaire), avancé pour la prédominance du patron labial-V-coronal-V par MacNeilage et Davis (2000), peut tout aussi bien être l'illustration de la recherche d'une certaine efficacité acoustique au niveau intra-syllabique et donc d'un contraste perceptif nécessaire entre le début et la fin de la syllabe. Les asymétries phonologiques, également observées dans nos langues entre attaque et coda, sont souvent avancées dans la littérature comme le résultat des propriétés des systèmes auditif (Lindblom, 2000) et articulatoire (Krakow, 1999 ; Sato, 2002). Si l'explication des syllabes CVC défavorisées peut être la conséquence de la recherche d'une certaine efficacité acoustique, on pourrait aussi rechercher des explications relatives à une demande de rééquilibrage de l'avantage perceptif de l'attaque sur la coda (Redford et Dhiel, 1999) ou à la réalisation de deux tâches identiques successives qui serait esquivée, car plus difficile à réaliser dans le même cadre mandibulaire (MacNeilage, Davis, Kinney et Matyear, 2000).

Enfin, l'observation des unités lexicales CV.CV, nous a permis de retrouver l'« effet LC » mis en évidence par MacNeilage et Davis (2000). Ce résultat s'ajoute à celui que nous avons obtenu pour les structures intra-syllabiques CVC et renforce ceux de MacNeilage et Davis (2000) puisque seul le finnois est commun à nos échantillons. La confirmation des résultats de MacNeilage et Davis (2000) sur l'« effet LC » ouvre sans doute la piste d'une organisation qui serait, cette fois, inter-syllabique au niveau lexical.

Si on rapproche ces résultats des propositions de la phonologie métrique, qui considère que, plus deux éléments sont liés phonologiquement plus ils cherchent à se regrouper en un même constituant (cf. chapitre I.1.), et si l'« effet LC » est une propriété de l'organisation phonologique des langues, valable aussi bien en interne de syllabe qu'entre deux syllabes consécutives, il semble qu'une redéfinition des domaines d'application des liens phonologiques soit nécessaire au niveau de l'unité lexicale.

## **CHAPITRE IV**

### **UN CORRÉLAT PERCEPTIF DE « L'EFFET LC »**



## IV.1. ILLUSIONS AUDITIVES ET EFFET LC : LE PARADIGME DES TRANSFORMATIONS VERBALES

Nous avons vu dans le chapitre précédent (cf. partie III.5.6.), que l'analyse des structures dissyllabiques de notre corpus confirme et étend les observations de MacNeilage et Davis (2000) qui montrent que les unités CVCV commencent majoritairement par une consonne labiale suivie d'une consonne coronale par rapport à la combinaison inverse. Ils expliquent que ce type de structure est favorisé car plus simple à produire. Pourtant, nos résultats sur les syllabes CVC avec attaque et coda non vides montrent que le principe d'économie articulaire ne semble pas être le seul facteur intervenant dans l'organisation syllabique des langues du monde. La faible proportion de syllabes avec attaque et coda de même lieu d'articulation (cf. partie III.3.5.5.) semble plutôt appuyer l'hypothèse de l'existence d'un facteur perceptif : la recherche d'une efficacité acoustique et donc d'un maximum de distinctivité entre le début et la fin de la syllabe.

Concernant les patrons lexicaux favorisés, une propriété perceptive peut-elle expliquer le fait que des langues favorisent les structures commençant par une consonne labiale suivie d'une consonne coronale ? Et comment le mettre expérimentalement parlant en évidence ?

Des études portant sur des illusions auditives ont montré leur capacité à mettre en lumière certains processus sous-jacents de la perception de la parole et apporter des éléments de compréhension du fonctionnement du langage (Kaminska, Pool et Mayer, 2000 ; Shoaf et Pitt, 2002).

Les transformations verbales sont un type particulier d'illusions auditives qui apparaissent lors de l'écoute répétée et continue d'un stimulus. Dans ce paradigme expérimental, le sujet soumis à l'expérience perçoit tout d'abord la forme de base du stimulus, puis au bout d'un certain temps, un changement perceptif se produit chez celui-ci qui croit entendre alors des mots ou des pseudo-mots plus ou moins en rapport avec le stimulus de départ (Sato, 2001 ; Sato et Schwartz, 2003). Ce processus de transformation persiste pendant toute la durée de la répétition correspondant à différents changements perceptifs successifs d'une forme vers une autre, incluant la forme de départ. Ce type d'expérience est une extension aux images auditives de celles menées dans le cadre de la perception visuelle de figures ambiguës, telles que le cube de Necker (Chambers et Reisberg, 1985) : « *Le caractère d'ambiguïté globale est caractérisé par le fait qu'un objet, et pas seulement ses parties, peut être interprété de deux manières.* » (Sato, 2001, p. 23). Au niveau visuel, cet aspect va se caractériser par deux configurations possibles (cf. figure IV.1). Pour le perceptif, la caractéristique observée est une resyllabation du flux de parole vers une autre forme possible du stimulus. Ainsi, la répétition du mot « life » fournit un flux sonore compatible avec deux types de segmentation : « life » ou « fly ».

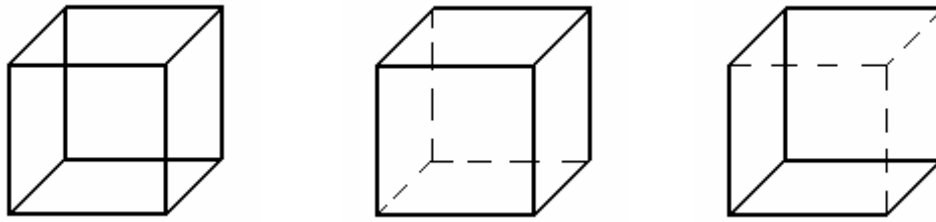


Figure IV.1 : Cube de Necker : stimulus ambigu et perceptions possibles

Le paradigme des transformations verbales nous a semblé particulièrement intéressant pour notre problématique car ces illusions auditives sont généralement considérées comme étant le reflet de processus mentaux liés à l'organisation perceptive humaine et donc à l'interprétation de la parole. Les études portant spécifiquement sur les mécanismes liés aux transformations verbales considèrent que deux fonctions sont particulièrement actives lors de la réinterprétation du signal auditif lorsque les transformations ont lieu vers des non-mots : celle de satiété et celle de changement de critère (Warren & Meyers, 1987 ; MacKay, Wulf *et al.*, 1993 ; Kaminska, Pool & Mayer, 2000). On considère, en effet, que l'écoute répétée d'un stimulus entraîne l'activation de sa représentation mentale jusqu'à satiété. Le critère de changement, utilisé pour catégoriser les stimuli, change alors brusquement et l'activation d'une nouvelle représentation est obtenue.

De précédentes études sur ces illusions ont analysé les différents types de transformations observées, ainsi que leurs contextes d'apparition. Il semble que les variations dans la forme des réponses diffèrent en fonction d'un certain nombre de facteurs liés au stimulus de départ :

- la densité du voisinage lexical (correspondant au nombre d'unités lexicales phonologiquement proches du stimulus répété) ;
- la fréquence d'usage ;
- le statut lexical du mot.

Les influences de ces différents facteurs sont variables. Ainsi, l'augmentation de la densité lexicale entraîne une plus grande variabilité au niveau du nombre et de la forme des transformations, alors que les deux autres facteurs ont une influence conjointe : l'utilisation comme stimulus d'un non-mot entraînera l'apparition d'un nombre moins important de transformations que si l'on choisit un mot. De surcroît, plus ce dernier est fréquent, plus le nombre de transformations observées augmente. D'après Shoaf et Pitt (2002), ce dernier effet est majoritairement dû à la stabilité du stimulus de départ qui, pour sa part, serait liée à un mécanisme de rétablissement qui dépendrait, quant à lui, du statut lexical des stimuli employés. Un tel phénomène rejoint les observations de Ditzinger, Tuller et Kelso (1997) à propos d'une dynamique des transformations verbales qui semble favoriser le couplage par paires des formes émergentes. En effet, bien qu'on obtienne généralement différents types de réponses, la perception de tous les sujets est majoritairement dominée par deux formes, comprenant systématiquement le stimulus de départ.

L'observation et l'analyse des formes obtenues dans le cadre de l'écoute répétée et continue d'un stimulus donné semblent donc fournir des informations sur l'organisation des représentations lexicales et sub-lexicales grâce aux variations de la stabilité perceptive des transformations. L'expérience que nous avons réalisée permet d'explorer cette question à travers le test de la validité perceptive d'un modèle organisationnel spécifique des lexiques, correspondant à « l'effet LC » introduit par MacNeilage et Davis (2000). Ce phénomène concerne la fréquence relativement élevée des structures CVCV commençant par une consonne labiale suivie d'une consonne coronale au détriment de la combinaison inverse. Cet effet LC, apparaissant selon MacNeilage et Davis (2000) avec les premiers mots des enfants, se retrouve dans leur étude d'une dizaine de langues, ainsi que dans 10 des langues de notre échantillon pour les structures CV.CV. Nous montrons qu'il est également présent dans les structures CVC (cf. partie III.5.6.).

Afin de tester « l'effet LC » dans le paradigme des transformations verbales, nous avons choisi de travailler à partir de trois paires d'unités dissyllabiques réversibles, formées autour des deux patrons consonantiques impliqués : labial-coronal et coronal-labial. En nous appuyant sur les résultats de Ditzinger, Tuller et Kelso (1997), l'hypothèse que nous formulons est que la distribution des transformations obtenues doit s'organiser autour d'un couplage par paires entre les deux formes possibles des stimuli répétés. Si tel est le cas, l'analyse des asymétries perceptives attendues entre les deux formes observées, [pV.tV] et [tV.pV], doit permettre de tester la stabilité perceptive de « l'effet LC » dans une tâche de transformations verbales.

## IV.2. METHODE

Pour tester notre hypothèse concernant la stabilité de formes perceptives dissyllabiques présentant un des deux patrons consonantiques suivants, labial-coronal et coronal-labial, nous avons sélectionné trois paires de stimuli réversibles de structure  $C_1V_1.C_2V_1$ , avec  $C = /p t/$  et  $V = /i a o/$ . L'analyse des différents facteurs intrinsèques aux stimuli susceptibles d'influer sur la tâche de transformations verbales a été effectuée à l'aide de données provenant de deux bases de données lexicales du français : VoCoLex (Dufour *et al.*, 2002) et BDLex-Syll (Pérennou et Calmès, 2002).

Ainsi, pour les trois paires de stimuli obtenues (/pa.ta/, /ta.pa/, /po.to/, /to.po/, /pi.ti/, /ti.pi/), on relève les valeurs suivantes pour les facteurs intrinsèques des stimuli utilisés (cf. table IV.1).

Stimulus	Point d'unicité	Densité de voisinage	Fréquence d'usage	Fréquence lexicale
<b>/pa.ta/</b>	5	71	0	19
<b>/ta.pa/</b>	5	55	2	10
<b>/po.to/</b>	5	39	9	6
<b>/to.po/</b>	5	30	1	11
<b>/pi.ti/</b>	5	22	0	7
<b>/ti.pi/</b>	5	29	0	11

*Table IV.1 : Valeurs des facteurs intrinsèques aux unités pour les stimuli choisis (point d'unicité : nombre de segments nécessaires pour que l'on ne puisse plus confondre un mot avec un autre ; densité de voisinage : nombre d'unités lexicales phonologiquement proches du stimulus)*

On peut remarquer que le point d'unicité est identique pour tous nos stimuli, que leur densité de voisinage est assez proche, avec pour chaque paire des ratios variants de 0,76 à 1,30, et que les fréquences d'usage de tous nos stimuli sont également proches et relativement basses (comprises entre 0 et 9). La fréquence lexicale (définie comme le nombre d'entrées lexicales possédant une structure CV.CV identique à celle du stimulus à n'importe quelle position dans le mot) est quasi-identique pour les stimuli /pV.tV/ et /tV.pV/, avec une moyenne à 10,67. Si l'on regarde les ratios en fonction de la voyelle, les différences sont légèrement plus marquées avec des valeurs variant entre 0,55 et 1,9.

Une fois le choix des stimuli terminé, les six séquences CV nécessaires, c'est-à-dire /pa/, /ta/, /po/, /to/, /pi/, /ti/, ont été enregistrées individuellement par un phonéticien entraîné, puis numérisées sur 16 bits à une fréquence d'échantillonnage de 48 kHz. Les caractéristiques acoustiques des six combinaisons CV (cf. figures IV.2 à IV.7) ont été contrôlées de manière à ce que les différences entre les structures avec /p/ et celles avec /t/ soient minimales pour les valeurs de la fréquence fondamentale, des trois premiers formants et de l'intensité (moins de 1 Hz d'écart entre les fréquences fondamentales de la partie /pV/ du dissyllabe et de la partie /tV/).

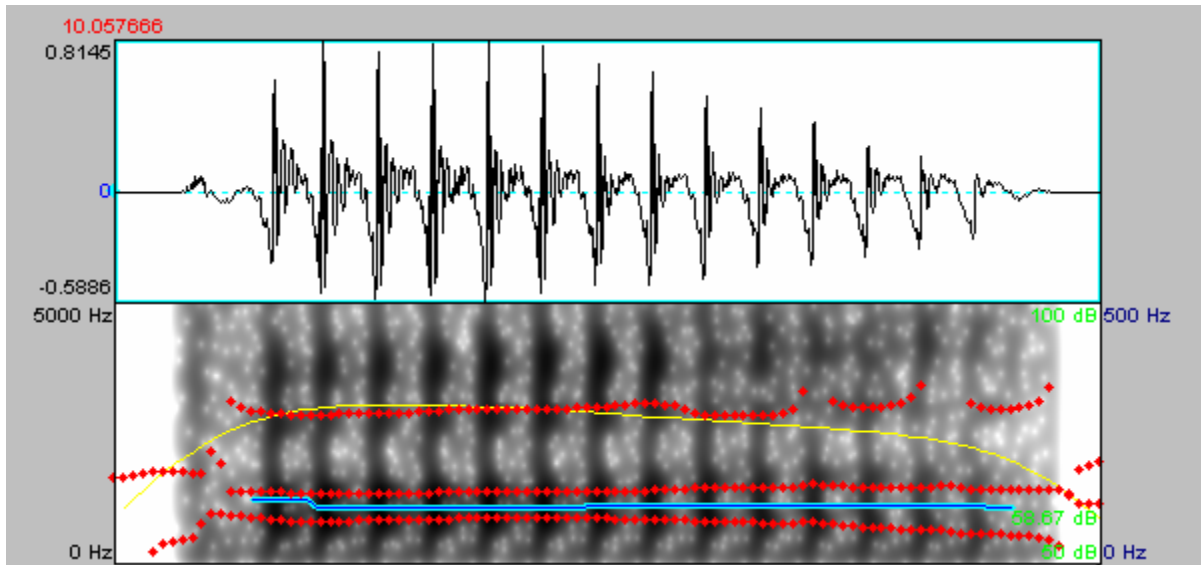


Figure IV.2 : Stimulus /pa/ utilisé dans notre expérience

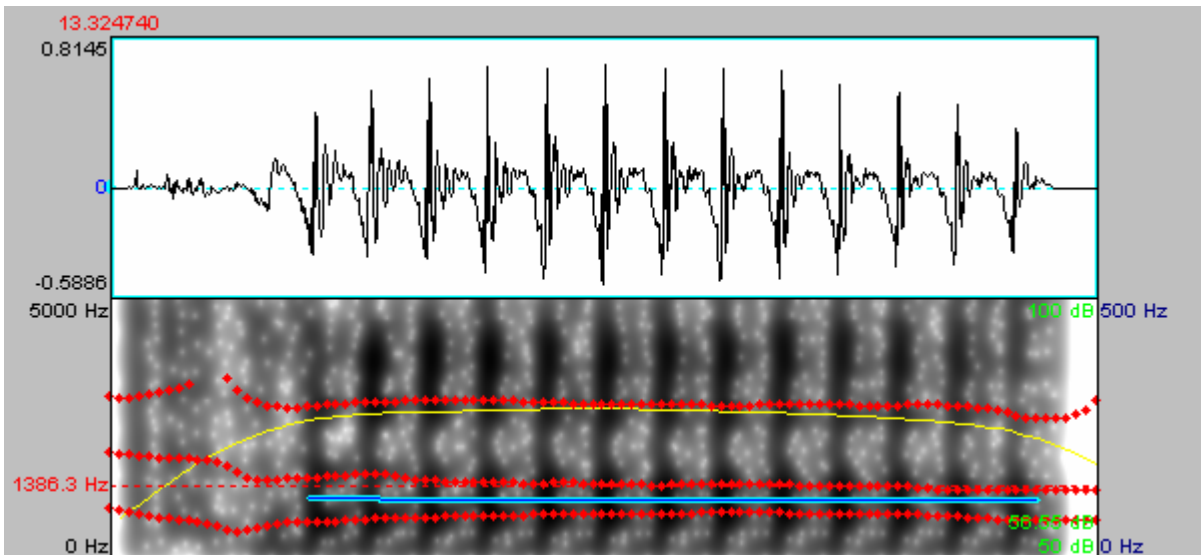
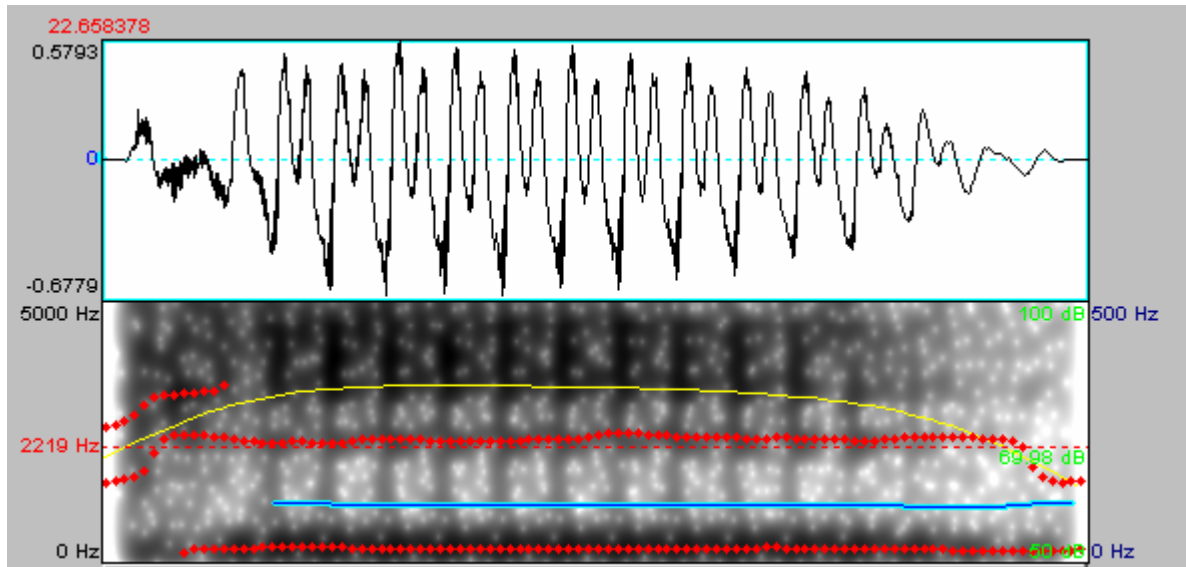


Figure IV.3 : Stimulus /ta/ utilisé dans notre expérience



igure IV.4 : Stimulus /pi/ utilisé dans notre expérience

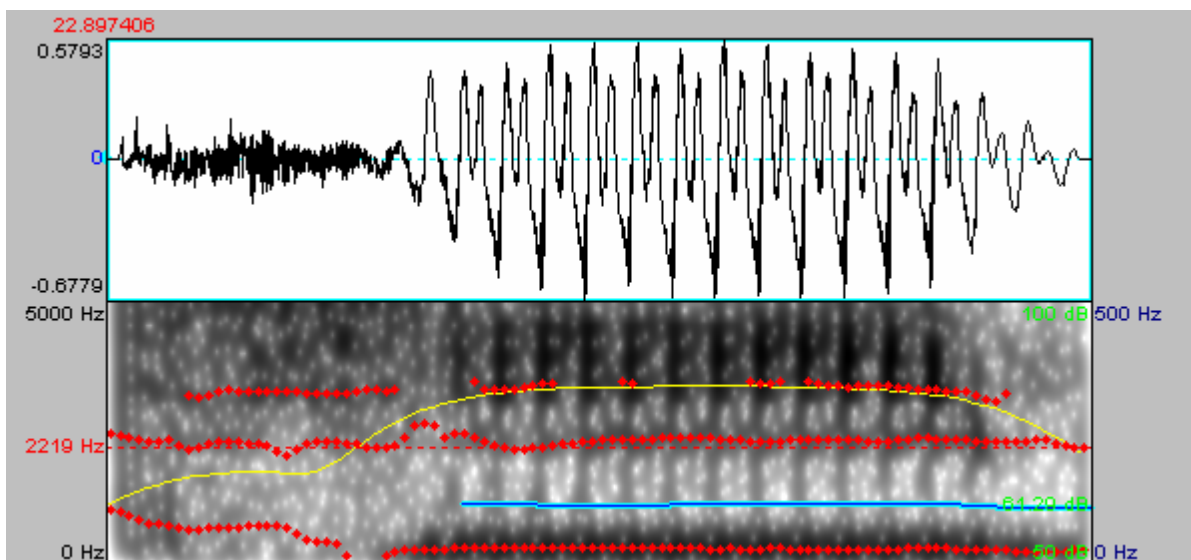


Figure IV.5 : Stimulus /ti/ utilisé dans notre expérience

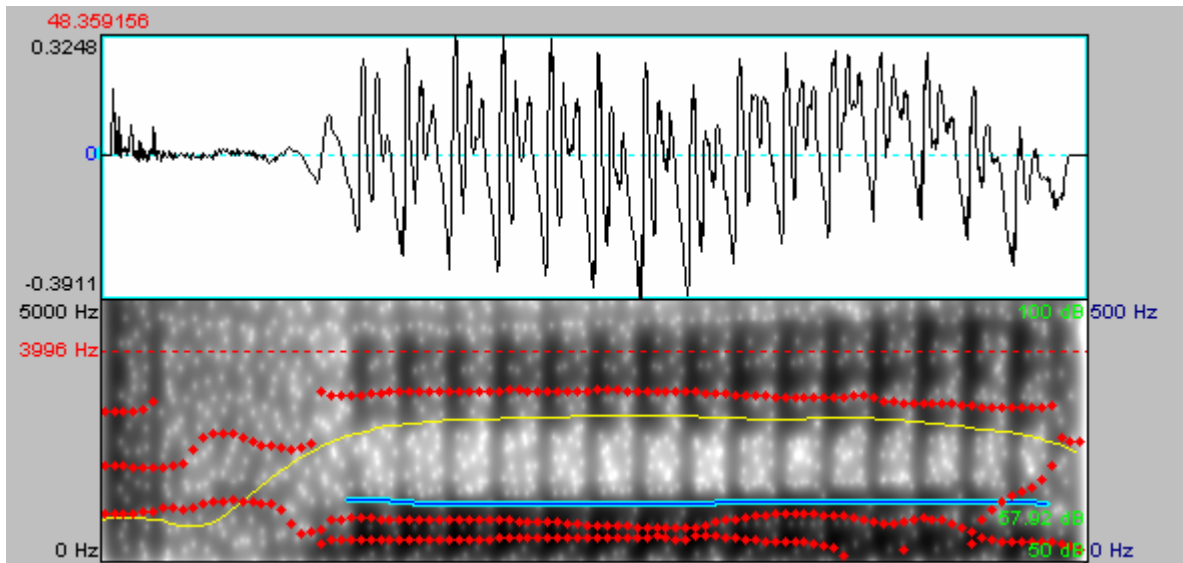


Figure IV.6 : Stimulus /po/ utilisé dans notre expérience

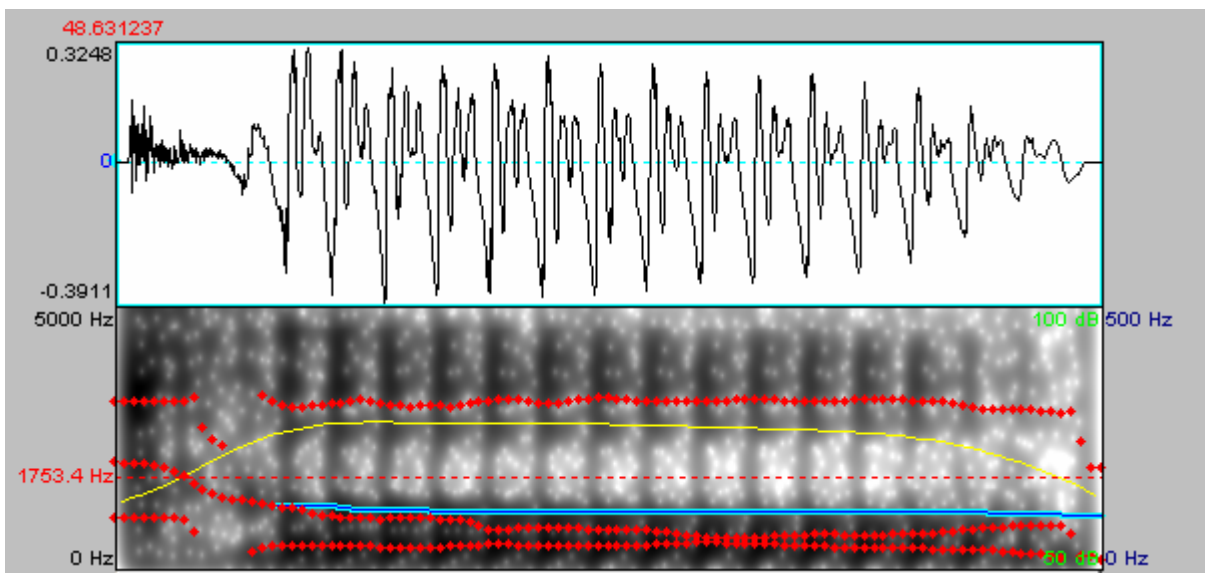


Figure IV.7 : Stimulus /to/ utilisé dans notre expérience

Les structures dissyllabiques des stimuli ont été construites par la concaténation, dans un même fichier-son, des deux syllabes CV appropriées. La durée de l'intervalle entre les deux séquences CV a été fixée à 100 msec, rendant la durée totale des structures dissyllabiques comprise entre 503 et 531 msec, avec une moyenne de 520 msec. Chaque forme CV.CV ainsi obtenue a été concaténée 300 fois de suite dans un fichier audio afin d'obtenir un stimulus d'une durée moyenne de 157 msec (min : 152 msec ; max : 161 msec).

### IV.3. PROCEDURE EXPERIMENTALE

Notre expérience s'est déroulée avec 24 sujets, 17 femmes et 7 hommes, âgés de 18 à 48 ans, tous locuteurs natifs du français sans problème auditif ou articuloire particulier.

Dans un premier temps, il était expliqué par l'expérimentateur, à chacun des sujets, le principe des transformations verbales de la manière suivante : « *il arrive que la perception prolongée d'une séquence verbale entraîne chez le locuteur une perception différente de cette séquence. Par exemple, je vais répéter rapidement le mot « life », vous demander de m'écouter attentivement et de me signaler si vous percevez une transformation de ce mot lors de cette répétition* ». Tous les sujets ont perçu un passage vers le mot « fly ». Ensuite, l'expérimentateur expliquait le déroulement pratique de l'expérience : l'écoute, à travers un casque audio, d'un stimulus auditif passé en boucle, avec la consigne d'être attentif aux changements de l'énoncé répété et de prononcer à haute voix, dans le microphone placé devant les sujets, toutes les transformations perçues. Il était aussi précisé que ces dernières peuvent être n'importe quel changement auditif qui intervient lors de l'écoute de la bande sonore, qu'il soit très perceptible ou plus subtil, et que ce soit vers un mot ou vers un non-mot (mot sans sens). Il a aussi été mentionné aux 24 sujets qu'il n'existe pas de bonnes ou de mauvaises réponses, et qu'il est tout à fait possible de ne percevoir aucun changement et que dans ce cas, il ne faut rien dire.

Une fois la mise en condition du sujet effectuée, un pré-test a été réalisé avec un stimulus prévu à cet effet à partir des séquences /pa/ et /to/. Le but de cet essai est de vérifier la bonne compréhension de la consigne, ainsi que la bonne qualité des niveaux d'écoute et d'enregistrement. Pour chaque sujet, le test comprenant l'écoute des six stimuli dans un ordre semi-aléatoire (les ordres présentant deux fois de suite une combinaison formée à partir de la même voyelle sont écartés). Des pauses ont été accordées entre chaque groupe de trois écoutes, amenant la durée totale du test à 30 minutes.

L'enregistrement des réponses données par les sujets nous a permis d'obtenir pour chaque sujet et chaque stimulus, un fichier son contenant les formes prononcées. Le logiciel Praat nous a donné les moyens d'étiqueter les différentes formes perçues par les sujets et de retrouver leur nombre total d'apparitions, ainsi que leurs durées respectives (temps écoulé entre la prononciation de cette forme et la prononciation de la suivante).



## IV.4. RESULTATS

Les analyses que nous avons menées portent sur les différences existant entre chaque paire de stimuli (*effet stimulus*, c'est-à-dire les formes ayant des patrons consonantiques inverses, mais partageant la même voyelle), ainsi qu'entre chaque forme dissyllabique possible (*effet d'amorçage*, c'est-à-dire dépendant de l'ordre des consonnes et non de la voyelle). À partir de toutes les données recueillies, plusieurs types d'analyses statistiques ont pu être réalisées.

Au niveau d'une analyse globale des résultats recueillis, deux types de variables semblent intéressantes : le nombre moyen de transformations et le nombre moyen de formes différentes (cf. table IV.2).

Stimulus	Transformations		Formes	
	Moyenne	Ecart-type	Moyenne	Ecart-type
/pa.ta/	26,58	16,11	3,83	2,35
/ta.pa/	22,67	14,06	4,21	1,93
/pi.ti/	24,04	12,26	4,33	1,97
/ti.pi/	23,87	13,21	4,29	2,14
/po.to/	29,17	16,45	5,04	3,62
/to.po/	26,46	12,60	4,79	2,78

Table IV.2 : Nombre moyen de transformations et de formes différentes observées en fonction du stimulus (avec écart-types).

À propos du nombre moyen de transformations, le calcul d'une ANOVA à doubles facteurs à mesures répétées montre un effet significatif de l'effet d'amorçage [ $F(1,23) = 4,61$ ,  $MSe = 184,51$ ,  $p < 0,05$ ], avec les stimuli /pV.tV/ montrant en moyenne 2,26 fois plus de transformations que les stimuli /tV.pV/ (figure 4), alors que l'effet stimulus est marginal sur le nombre de formes modifiées [ $F(2,46) = 3,11$ ,  $MSe = 203,67$ ,  $p < 0,06$ ]. Le nombre de transformations varie entre 47,91 et 55,63 pour chacune des trois paires de stimuli réversibles. L'interaction entre les deux conditions n'est pas significative [ $F(2,46) = 0,60$ ,  $MSe = 43,97$ ,  $p > 0,05$ ].

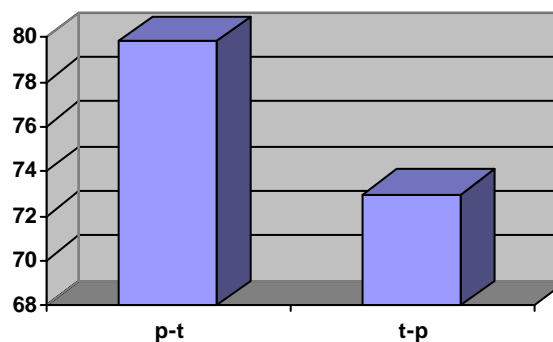


Figure IV.8 : Nombre de transformations en fonction de l'effet d'amorçage (ANOVA [ $F(1,23) = 4,61$ ,  $p < 0,05$ ])

Au niveau du nombre moyen de formes observées, le calcul d'une ANOVA à doubles facteurs à mesures répétées ne montre aucun effet, ni pour l'effet stimulus [ $F(2,46) = 1,83$ ,  $MSe = 10,02$ ], ni pour l'effet d'amorçage [ $F(1,23) = 0,01$ ,  $MSe = 0,03$ ], ni pour leur interaction [ $F(2,46) = 0,40$ ,  $MSe = 1,22$ ].

Les différentes analyses effectuées au niveau global ne permettent pas de conclure sur la stabilité des différentes formes proposées dans notre expérience. Un niveau supplémentaire d'analyses statistiques a donc été appliqué sur nos données afin de tester de possibles groupes perceptifs à travers les transformations observées. De manière à obtenir une mesure plus précise de la stabilité perceptive des stimuli, toutes les transformations ont été classées en fonction de leur adéquation (cf. table IV.3) :

- au stimulus de départ (transformation vers le stimulus de base),
- à la forme inverse du stimulus (transformation vers la forme associée)
- à une autre transformation (transformation vers une autre forme).

Stimulus	Type de transformations					
	Transformations vers le stimulus de base		Transformations vers la forme associée		Autres transformations	
	Nombre	Stabilité	Nombre	Stabilité	Nombre	Stabilité
/pa.ta/	9,96	62,88	8,04	39,95	8,58	48,04
/ta.pa/	7,42	43,39	8,42	56,34	6,83	50,64
/pi.ti/	9,79	57,20	8,17	56,65	6,08	44,31
/ti.pi/	8,04	50,87	8,83	52,85	7,00	54,11
/po.to/	11,46	67,83	8,87	46,37	8,83	43,70
/to.po/	9,25	46,41	9,75	65,97	7,46	45,52
<b>Total</b>	<b>55,92</b>	<b>328,58</b>	<b>52,08</b>	<b>318,13</b>	<b>44,78</b>	<b>286,32</b>

Table IV.3 : Nombre moyen et durée de stabilité des transformations liées au stimulus, vers la forme associée ou une autre en fonction du stimulus.

La première remarque que l'on peut faire en étudiant ces résultats est que la distribution des formes observées démontre que l'organisation principale des transformations perceptives est celle d'un couplage par paires entre la forme de base et la forme inverse (cf. figure IV.9). Ce type de transformations représente plus de 70 % des transformations observées.

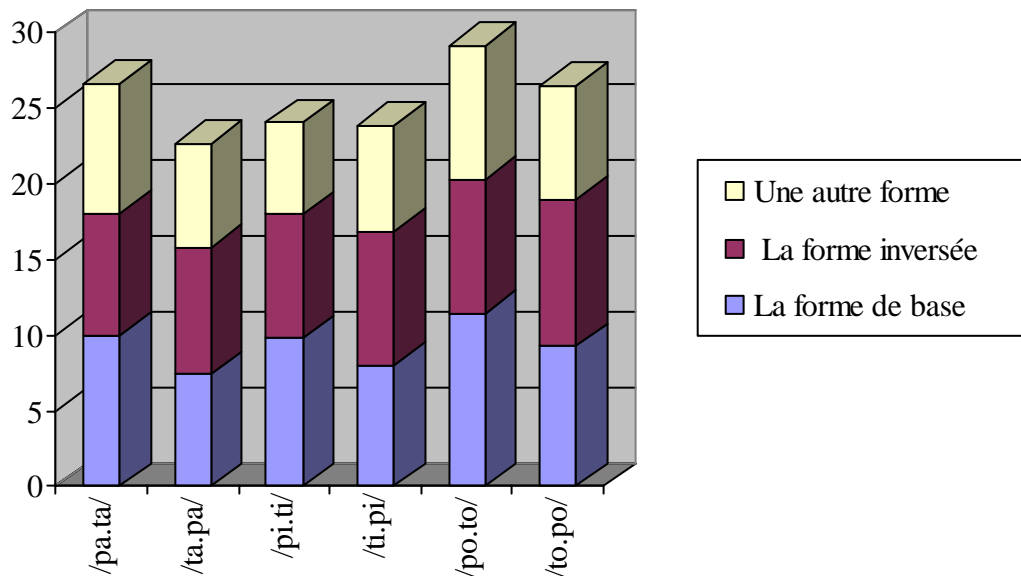


Figure IV.9 : Nombre de transformations vers la forme de base, la forme inversée, une autre forme, en fonction du stimulus de départ.

Le calcul d'une ANOVA à double-facteurs à mesures répétées, avec le nombre de transformations vers une forme donnée (c'est-à-dire la forme de base, la forme associée ou une autre forme) comme variable dépendante, montre un effet significatif de l'effet d'amorçage pour les transformations vers le percept de base [ $F(1,23) = 9,82$ ,  $MSe = 169$ ,  $p < 0,01$ ]. On remarque en effet que les stimuli /pV.tV/ reviennent vers leur forme de base en moyenne 2,71 fois plus que les stimuli inverses. Par contre, aucun effet significatif n'est visible pour les transformations vers une autre forme, que ce soit la forme couplée ou une autre. Ces résultats expliquent, en partie, la présence d'un effet d'amorçage au niveau de l'analyse globale des données, puisqu'on a vu que les stimuli /pV.tV/ connaissent en moyenne 2,26 fois plus de transformations que les stimuli /tV.pV/.

Pour ce qui est de l'effet stimulus, un effet marginal est observable sur le nombre de transformations vers la forme de base [ $F(2,46) = 2,60$ ,  $MSe = 39,17$ ,  $p < 0,09$ ], alors qu'aucun effet ne se remarque pour les deux autres types de transformations. De même, l'interaction entre les deux conditions (effet d'amorçage et effet stimulus) n'est jamais significative, quel que soit le type de transformation étudié.

La troisième analyse que nous avons menée sur les données recueillies concerne la validité perceptive de « l'effet LC ». Pour cela, les transformations vers la forme de base ou vers la forme associée ont été classées en fonction de leur adéquation avec le modèle Labial-Coronal ou avec le patron consonantique inverse (cf. table IV.4).

Stimulus	Transformations vers une forme LC		Transformations vers une forme CL	
	Nombre	Stabilité	Nombre	Stabilité
/pa.ta/	9,96	62,88	8,04	39,95
/ta.pa/	8,42	56,34	7,42	43,39
/pi.ti/	9,79	57,20	8,17	56,65
/ti.pi/	8,83	52,85	8,04	50,87
/po.to/	11,46	67,83	8,87	46,37
/to.po/	9,75	65,97	9,25	46,41
<b>Total</b>	<b>58,21</b>	<b>363,07</b>	<b>49,79</b>	<b>283,64</b>

Table IV.4 : Nombre moyen et durée de stabilité (ms.) des transformations vers une forme Labial-Coronal ou vers une forme Coronal-Labial.

On remarque que le nombre de transformations vers une forme LC est toujours plus important, quel que soit le type de stimuli. Le calcul d'un test-t unilatéral sur l'ensemble de nos données montre une asymétrie significative entre les deux structures [ $t(1,143) = 3,40$ ,  $p < 0,001$ ], avec en moyenne 1,4 fois plus de transformations vers les formes Labial-Coronal que vers les patrons inverses.

Pour compléter les informations à propos de cette asymétrie perceptive, nous nous sommes intéressés aux durées de stabilités respectives des deux patrons observés (cf. figure IV.10). Comme pour le nombre de transformations, on remarque que quel que soit le stimuli, la durée perceptive des transformations /pV.tV/ est toujours supérieure à celle des transformations /tV.pV/. Le calcul d'un test-t unilatéral montre une asymétrie significative entre les deux formes [ $t(1,143) = 3,73$ ,  $p < 0,005$ ], avec une durée de stabilité moyenne supérieure de 13,24 s. en faveur des formes de type LC. Pour affiner ces résultats, nous avons étudié les taux de transformation (/pV.tV/ - /tV.pV/) en tenant compte de l'effet stimulus. Des différences importantes apparaissent pour les stimuli possédant les voyelles /a/ (17,94 s.) et /o/ (20,51 s.) mais elles sont beaucoup moindres pour les stimuli contenant /i/ (1,26 s.). Le calcul d'un test-t unilatéral avec correction de Bonferroni permet de vérifier la significativité de ces trois différences, et montre l'existence d'un effet pour /a/ [ $t(1,47) = 2,90$ ,  $p < 0,05$ ] et /o/ [ $t(1,47) = 3,39$ ,  $p < 0,005$ ], mais pas pour /i/ [ $t(1,47) = 0,21$ ].

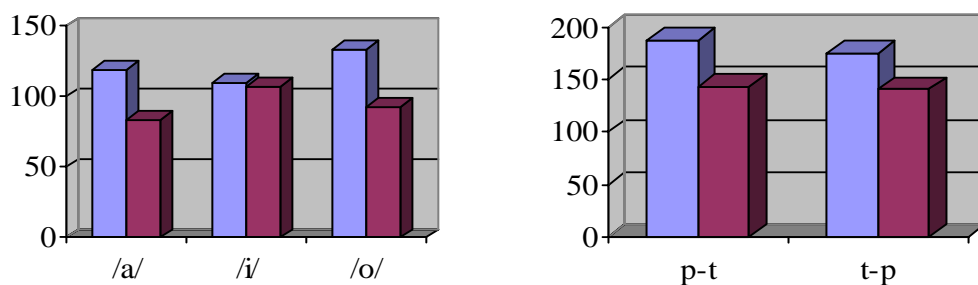


Figure IV.10 : Durée de stabilité des structures : ■ LC et ■ CL en fonction de l'effet stimulus (à gauche) et de l'effet d'amorçage (à droite).

## IV.5. DISCUSSION

Un premier résultat intéressant de cette expérience est celui de la dynamique observée au niveau des changements perceptifs. Les données montrent que, pour tous les sujets, l'organisation principale des transformations est toujours celle d'un couplage par paires entre les deux formes possibles de chaque stimulus réversible (c'est-à-dire que les transformations correspondent au stimulus de départ ou à sa forme syllabique associée). Ces deux patrons attirent en moyenne 71 % du nombre total de transformations et 69 % de la durée totale du temps de l'expérience. Ce mécanisme de couplage par paires est pratiquement similaire à celui précédemment souligné par les travaux de Ditzinger, Tuller et Kelso (1997) qui montrent que, malgré l'observation d'un certain nombre d'alternatives différentes, toutes les perceptions des sujets sont dominées par seulement deux formes, comprenant toujours la forme de base. Par conséquent, nos résultats et les leurs convergent vers un mécanisme de couplage par paires impliquant toujours le percept de départ. Ce mécanisme de couplage par paires est intéressant pour valider la mesure de la durée de stabilité perceptive. En effet, nous observons une corrélation positive entre le nombre de transitions et la durée de la stabilité perceptive des formes dominantes.

Dans notre expérience l'analyse des transformations et de la durée de stabilité perceptive livre bien une corrélation de « l'effet LC ». En effet, quel que soit le stimulus ou le paramètre mesuré, la préférence perceptive pour les formes /pV.tV/ par rapport aux formes /tV.pV/ est nettement observée. Ce biais asymétrique ne peut pas être expliqué par des facteurs tels que la densité du voisinage lexical ou la fréquence lexicale des stimuli. Comme souligné précédemment, un grand nombre de voisins devrait augmenter le nombre de candidats lexicalement possibles, ce qui entraînerait un plus grand nombre et une plus grande étendue des transformations (MacKay, Wulf et Abrams, 1993). Étant donné l'importance de la valeur de la densité du voisinage lexical moyenne pour les stimuli /pV.tV/ par rapport aux stimuli /tV.pV/ (ratio de 1,16), l'effet de la densité du voisinage lexical pourrait paraître en accord avec le plus grand nombre de transformations relevées pour les stimuli /pV.tV/.

Cependant, l'importance des transformations pour les stimuli /pV.tV/ est principalement due à la tendance pour ces stimuli de « switcher » vers leur forme de départ qui n'apparaît pas alors comme un mécanisme compétitif. De plus, l'effet de la densité du voisinage lexical ne peut expliquer la préférence perceptive pour les formes /pV.tV/ par rapport aux formes /tV.pV/. Par conséquent, la densité du voisinage lexical ne fournit vraisemblablement pas l'explication des asymétries présentes. Pas plus que l'interprétation concurrente qui viendrait de la fréquence lexicale des stimuli CV.CV dans le lexique des sujets. En effet, cette fréquence lexicale pourrait biaiser les transformations vers des entrées lexicales plus fréquentes. Pourtant, nous observons un plus grand nombre d'entrées lexicales pour /to.po/ que pour /po.to/, un résultat qui n'est pas en accord avec les préférences perceptives observées dans les transformations de /po.to/ (ratio de 0,50).

Rappelons que « *l'effet LC* » est une tendance dans l'organisation phonologique et lexicale des langues à favoriser un ordre consonantique dans les structures CV.CV aussi bien que CVC : Labial-Coronal plutôt que l'inverse Coronal-Labial, effet que l'on retrouve en production au stade de l'acquisition des premiers mots, bien qu'absent dans le babillage (MacNeilage et Davis, 2000 ; Davis et MacNeilage, 2003). Ces différentes études ont supposé « *l'effet LC* » comme la conséquence de propriétés articulatoires de l'appareil de parole : les consonnes labiales résulteraient d'un mouvement basique de la mandibule, activée seule, alors qu'un mouvement additionnel de la langue est nécessaire à la recherche d'une position avant de la langue pour une consonne coronale. Comment alors réconcilier cette explication articulatoire et la validité perceptive de l'effet LC dans une tâche de transformation verbale ? Une des solutions envisageables pourrait provenir de théories sur la perception du langage qui, bien que discutées pour certaines, postulent une base de représentations articulatoires ou gestuelles (Lieberman et Mattingly, 1985 ; MacNeilage, 1998 ; Schwartz *et al.*, 2002). Selon ces théories, la plus forte cohésion articulatoire des structures syllabiques LabialV-CoronalV, en comparaison aux formes syllabiques CoronalV-LabialV, pourrait expliquer la stabilité/attractivité perceptive plus importante des formes /pV.tV/ et ainsi valider perceptivement l'effet LC dans la tâche de transformation verbale. En d'autres mots, les contraintes de stabilité provenant du système moteur pourraient participer à la représentation mentale des formes phonologiques pendant l'écoute répétée de parole. Et puisqu'il est avéré que les transformations verbales apparaissent aussi bien lors de la perception que lors de la production répétée d'unités de la parole (Reisberg *et al.*, 1989 ; Shoaf et Pitt ; 2002 ; Sato et Schwartz, 2003), l'effet de transformation verbale pourrait alors être un élément central d'évaluation de cette proposition et donc d'investigation sur des propriétés structurales du système langagier.

Il sera donc intéressant de réutiliser, dans le prolongement de cette thèse, le paradigme des transformations verbales afin de mieux comprendre le fonctionnement des langues et en particulier l'apparition d'autres types de patrons consonantiques tels que Labial-Vélaire par rapport à Vélaire-Labial, chacun d'entre eux apparaissant en position favorisée dans environ la moitié des langues d'ULSID (respectivement 7 et 6 langues sur les 13 attestant des vélaires en coda).

Si notre étude sur les structures lexicales des langues (cf. chapitre III.3.5.5.) confirme bien que « *l'effet LC* » est un modèle d'organisation lexicale, les résultats obtenus sur la stabilité des formes LC avec un paradigme des transformations verbales montrent que cet effet est sans doute, de manière plus générale, un modèle d'organisation du langage.

**CONCLUSION ET PERSPECTIVES :**  
**VERS LA MODELISATION DE L'EMERGENCE**  
**SYLLABIQUE**

Toute unité lexicale, quelle que soit la langue, est construite sur une concaténation d'un ou plusieurs types syllabiques existant dans cette langue. Cependant, comme nous l'avons montré dans le premier chapitre, la nature de l'unité syllabique est encore actuellement l'objet d'un débat linguistique, tant au niveau phonologique, que phonétique et psycholinguistique. Depuis Saussure, phonologues et phonéticiens ont très souvent fait l'hypothèse que l'existence de l'unité syllabique ne pourrait être démontrée que par l'identification de ses frontières.

Du point de vue phonétique, les études pour délimiter la syllabe ont surtout été menées dans l'exploration des limites des phénomènes de coarticulation et sur la recherche de caractéristiques articulatoires et perceptives qui seraient plus spécifiques des attaques syllabiques et d'autres qui seraient des indices révélateurs des codas.

Parmi les différentes phonologies qui visent à identifier la nature des représentations phonologiques de la syllabe, certaines ont utilisé les résultats d'études phonétiques pour expliquer les asymétries phonologiques, le repérage des frontières, ou pour établir des échelles de sonorité, et de manière plus générale, les phonologies ont cherché à démontrer, dans l'élaboration des modèles phonologiques, le rôle organisationnel de la syllabe. Mais toutes ces études, tant au niveau de la substance, que de la forme, n'ont pas encore mis au grand jour d'éléments pertinents qui pourraient appuyer l'existence phonétique, phonologique et psychologique d'une unité syllabique.

Pourtant, sans disposer d'une définition de la syllabe cohérente à plusieurs niveaux de la description linguistique, l'ensemble des résultats du chapitre III montre qu'il est tout à fait légitime de penser qu'une organisation syllabique structure les unités lexicales. Ainsi, les tendances dans la longueur des items lexicaux et les schèmes structurels (gabarits) des différents lexiques, les fréquences plus élevées de certaines combinaisons inter- et intra-langue aux niveaux syllabique et lexical, la mise en évidence de cooccurrences favorisées entre consonne et voyelle ou entre consonnes de la même syllabe ou de deux syllabes consécutives, donnent de fortes raisons pour laisser penser que l'organisation syllabique des lexiques ne soit pas à considérer comme le fruit du hasard. De même, l'inventaire des consonnes en attaque bien plus large que celui des consonnes en coda, des types de consonnes plus fréquents en position initiale qu'en finale de syllabe, des tendances implicationnelles dans la complexité des structures qui restent vraies si on considère la nature des consonnes constituants de ces structures, expriment probablement des contraintes fortes de position intra-syllabique.

Une fois de plus, nul doute que l'établissement de typologies et la recherche de grandes tendances trouvent leur place dans la construction et la réflexion de la théorie phonologique. Mais cette étape, bien que fondamentale, ne peut rester telle qu'elle et doit s'inscrire dans une démarche qui tente « d'ancrer la forme dans la substance » pour reprendre les termes de Laks et Plénat (1993). Dans cette optique, l'exploration des données de langues est complémentaire aux études théoriques et expérimentales qui cherchent à montrer l'existence d'une unité syllabique 1) à travers la recherche de son origine dans le fondement du langage humain ; 2) à travers les mécanismes de production et de perception de la parole.



Nous avons analysé certains de nos résultats à la lumière de la théorie « *Frame, then Content* » (MacNeilage, 1998) qui définit le cadre syllabique comme un simple cycle de l'oscillation mandibulaire, le début et la fin de chaque syllabe étant déterminés par un maximum de fermeture (position haute de la mandibule). Le cycle mandibulaire fournit la contrainte temporelle de la syllabe (période de ¼ de seconde environ) qui se traduit par une constante de durée (Greenberg, 1998), et la contrainte mécanique de mode, les articulations vocaliques et consonantiques s'inscrivant respectivement dans les phases d'ouverture et de fermeture de la cavité buccale. La théorie « *Frame, then Content* » répond du même coup à deux grands principes de la phonologie de la syllabe (Redford, 1999 ; Redford *et al.*, 2000) :

1) Au principe de l'attaque maximale (*Maximum Onset Principle*) selon lequel les attaques pleines sont préférées aux attaques vides et aux codas pleines, les attaques complexes sont préférées aux codas complexes, ce qui pourrait s'expliquer par l'asymétrie de phasage du cycle mandibulaire, celui-ci présentant une phase d'ouverture plus lente, plus ample, si elle est comparée à la phase de fermeture ;

2) Au principe de sonorité (*Sonority Sequencing Principle*), qui préconise que la sonorité des segments (définie à travers le mode articulatoire) augmente à mesure que l'on se rapproche du noyau de la syllabe, les segments se conformant au mouvement mandibulaire, ce dernier générant le séquençage et la hiérarchie de sonorité.

La confirmation des résultats de MacNeilage et Davis (2000) sur les séquences de type Labial-V-Coronal, où Labial est une caractéristique de la position initiale de la syllabe 1 et Coronal un trait de l'attaque de la syllabe 2 (l'effet LC), ouvre sans doute la piste d'une organisation, cette fois inter-syllabique, au niveau lexical. MacNeilage et Davis (2000) évoquent le principe de la simplicité articulatoire pour expliquer ce patron lexical favorisé, considérant un coût minimum de la mâchoire dans les mouvements d'ouverture/fermeture du conduit buccal.

En utilisant le paradigme des transformations verbales, l'étude expérimentale que nous proposons dans le chapitre IV témoigne de l'existence d'un « effet LC » au niveau perceptif. Nos résultats montrent que dans une tâche d'écoute de stimuli auditifs répétés à l'identique [p V t V] vs. [t V p V] / V ∈ [ i a o ] (répétitions construites sur des stimuli parfaitement contrôlés au niveau de la durée et de la structure acoustique, idem pour leur fréquence dans le lexique du français), l'existence d'une perception active fait « *switcher* » la perception des sujets français en faveur du patron syllabique Labial-V.Coronal-V, quel que soit le timbre de la voyelle et quel que soit le statut lexical du stimuli répété (mot vs. non-mot).

Nos résultats typologiques montrent que les syllabes CVC présentent aussi cette tendance avec Labial comme le lieu d'articulation de l'attaque et Coronal comme caractéristique de la coda. Toujours dans les structures syllabiques CVC, nous avons montré que le principe de parcimonie des traits pour la syllabe ne fonctionne pas : les syllabes possédant le même lieu d'articulation pour l'attaque et la coda sont nettement défavorisées dans les lexiques d'ULSID. Ce résultat peut venir renforcer le concept de coût minimal de production pour la

syllabe LC, réalisée par un seul geste mandibulaire, mais il va sans doute aussi dans le sens de la recherche d'une certaine efficacité acoustique au niveau intra-syllabique. Donc un contraste perceptif peut-être nécessaire entre le début et la fin de la syllabe.

D'autres investigations des données d'ULSID pourraient compléter ces observations, mais aussi d'autres expériences perceptives sur d'autres patrons lexicaux et syllabiques favorisés, comme Vélaire-V-Coronal ou Coronal-V-Vélaire. Cela permettrait d'affiner nos hypothèses sur l'existence de ce corrélat perceptif de « *l'effet LC* » (cf. Chapitre IV, section 5).

Quoiqu'il en soit, on sent bien que les explications sont à rechercher dans l'esprit de ce qui est plaidé par Lindblom *et al.* (1984) à savoir, l'optimisation des séquences de segments à partir des propriétés des systèmes humains de production et de perception de parole, c'est-à-dire en cherchant à vérifier l'hypothèse que les types de syllabes les plus répandus dans les langues du monde sont certainement plus fonctionnels que d'autres (au niveau articuloire et perceptif).

Mais de quelle manière éviter les hypothèses implicites sur l'organisation syllabique des lexiques ? Probablement en se rendant capables d'élaborer un modèle de l'émergence syllabique à l'image des modèles de prédiction des systèmes vocaliques développés depuis une trentaine d'années par les phonéticiens (Liljencrants et Lindblom, 1972 ; Lindblom, 1986 ; Schwartz *et al.*, 1997 ; de Boer, 2000).

Plusieurs travaux ont été menés en ce sens sans fournir d'éléments déterminants généralisables sur les facteurs susceptibles de conditionner des choix ou des restrictions en faisant appel à la substance des unités sonores en contexte immédiat ou proche. Nous retiendrons les travaux de Lindblom *et al.* (1984), le modèle ICP de Berrah *et al.* (1995) et celui de Redford *et al.* (2000). Les deux premiers sont totalement ancrés dans la substance, le troisième fait appel, presque essentiellement, à des propriétés formelles des structures lexicales et syllabiques.

Lindblom *et al.* (1984) proposent une simulation de l'émergence de structures phonologiques à l'aide d'un modèle auto-organisationnel. Ils visent ainsi à montrer que les syllabes obéissent à un tel principe qui caractérise certaines productions des systèmes vivants. Ils limitent leur recherche à un type particulier d'articulation : les transitions entre fermeture (stop-like) et ouverture du conduit vocal (vowel-like), qui correspondent aux syllabes CV, les plus répandues dans les langues du monde. Leur hypothèse de départ est que le signal phonétique comporte les caractéristiques qui tendent à faciliter production et perception. Ils posent ainsi les conditions de performances hypothétiques suivantes : (1) discriminabilité sensorielle, (2) préférence pour des articulations moins extrêmes, (3) distance perceptive, (4) saillance perceptive. Les conditions (2) et (4) sont des caractéristiques intra-syllabiques alors que (1) et (3) sont systémiques ; (1) et (2) renvoient aux conditions de performance pour le locuteur tandis que (3) et (4) renvoient à celles de l'auditeur.

Le modèle de Berrah (1995) est un modèle de prédiction des structures syllabiques CV, établi dans le prolongement des propositions de Lindblom *et al.* (1984) sur le concept de la négociation entre efficacité acoustique, coût articulatoire, et distinctivité systémique (Berrah, 1995 ; Vallée *et al.* 1999) et l'hypothèse que les syllabes les plus répandues dans les langues du monde sont celles dont les signaux phonétiques évoluent de manière à faciliter leur production et leur perception.

Pour mesurer ces propriétés substantielles sur les syllabes CV les plus fréquentes, Berrah *et al.* ont calculé l'efficacité acoustique intra-syllabique en évaluant la distance acoustique entre le spectre initial de la consonne et le spectre final de la voyelle pour chaque combinaison de C et de V tirées d'un échantillon déterminé au préalable et contenant quatre consonnes occlusives (bilabiales /p b/, alvéolaires /t d/, vélares /k g/ et uvulaires /q G/) et cinq voyelles /i 'e' a 'o' u/. Berrah fait l'hypothèse qu'une syllabe sera d'autant plus saillante perceptivement que les fréquences de résonances de la consonne et de la voyelle seront différentes. Par exemple, /gu/ est moins saillante que /gi/. L'efficacité acoustique inter-syllabique, correspondant à la distance perceptive entre syllabes prises deux à deux, est aussi considérée et permettra de classer, entre elles, toutes les paires de transitions CV possibles à partir de l'inventaire des segments retenus.

Berrah (1995) propose de calculer le coût articulatoire intra-syllabique des combinaisons CV possibles en utilisant l'équation de Lindblom et Engstrand (1989). Celui-ci est obtenu en calculant la distance entre les paramètres articulatoires des consonnes et des voyelles impliquées dans les cohortes. Le poids des articulateurs est pondéré en fonction de son habileté, de sa célérité et de sa capacité de discrimination entre plusieurs voyelles.

L'efficacité acoustique, le coût articulatoire, l'efficacité globale ont été calculés pour l'ensemble des syllabes choisies. En reprenant les principes de la prédiction des systèmes vocaliques (Liljencrants et Lindblom, 1972 ; Schwartz *et al.*, 1997), l'énergie totale du système de syllabes est calculée à partir de l'efficacité globale des structures mais aussi de la distance séparant les syllabes les unes des autres. Un système est d'autant plus stable que son énergie est faible.

Les syllabes ont été simulées dans un espace 3-D, semblable à l'espace vocalique F1/F2/F3, en utilisant l'analyse multidimensionnelle de Kruskal (1977). Les prototypes articulatoire-acoustiques des syllabes ont été construits à partir des prototypes vocaliques de Vallée (1994) et des équations de locus des consonnes [b d g G].

Le lexique qui émerge des simulations avec l'énergie la plus faible est /ba da ga bi di gi bu du gu/.

Le modèle de Berrah (1995) a permis de générer un système constitué de syllabes comptant parmi les plus fréquentes dans les langues du monde, en écartant les syllabes construites sur des constituants moins répandus dans les langues (tels la plosive uvulaire). Cependant le modèle n'est pas suffisamment élaboré pour prédire 1) d'autres types de

structure syllabique ; 2) des unités lexicales plus longues (structures di- ou trisyllabiques). Par ailleurs le modèle nécessite plusieurs améliorations telles que, par exemple, l'intégration de caractéristiques spectrales des segments comme la focalisation pour les voyelles (cf. Scwhartz et al, 1997). Ce travail de modélisation a néanmoins le mérite de s'être attaché à définir un espace physique de simulation des objets syllabiques et d'avoir ajusté un système linguistique basé sur une optimisation des processus de production et de perception de la parole, à partir de la minimisation des coûts articulatoires et de la maximisation de la saillance perceptive.

Toujours dans le cadre de l'émergence de la structure syllabique des langues, plus récemment le modèle ESS « *Emergent Syllable Systems model* », élaboré par Redford *et al.* (2000), permet de simuler l'émergence des systèmes syllabiques à partir d'un ensemble de contraintes fonctionnelles sur les systèmes de syllabes. Les contraintes sont basées sur des faits typologiques et certaines d'entre-elles dérivées de principes phonétiques comme la distinctivité perceptive et la facilité articulatoire.

ESS est construit à partir de six contraintes réparties en 2 catégories :

- 1) Trois contraintes sur la formation des mots du vocabulaire rendent compte des combinaisons de sons favorisées et défavorisées au niveau du lexical :

**C1 : Contrainte d'efficacité sur la production des mots selon le Principe du Moindre Effort.** Les mots courts sont préférés aux mots longs. C1 pénalise donc les vocabulaires avec des mots longs ;

**C2 : Contrainte d'intelligibilité du sens des mots selon le Principe de Non-Homophonie.** Pour rendre la communication efficace, le plus souvent, ce sont des séquences sonores différentes qui nomment des objets différents. C2 pénalise les vocabulaires avec des mots identiques ;

**C3 : Contrainte de mémorisation des types syllabiques.** La réutilisation des patrons syllabiques dans les langues permet de minimiser le nombre de syllabes différentes et donc de réduire le coût de stockage en mémoire en facilitant également leur acquisition. C3 pénalise les vocabulaires avec un large nombre de syllabes différentes.

- 2) Trois contraintes s'appliquent à la structure interne des mots :

**C4 : Contrainte sur le type syllabique selon le Principe du Maximum Onset Principle.** Les onsets syllabiques pleins sont préférés aux codas. C4 pénalise les types CVC et VC ;

**C5 : Contrainte sur la taille des onsets et des codas.** Les consonnes simples sont préférées aux groupements de plusieurs consonnes. C5 pénalise les clusters ou les hiatus ;

**C6 : Contrainte sur les sons successifs selon le Principe de Sonorité.** La sonorité des segments est croissante depuis le début de la syllabe jusqu'au noyau alors

qu'elle est décroissante depuis le noyau vers la fin de la syllabe. C6 pénalise les séquences de segments avec une position mandibulaire identique ou différente qui ne rendrait pas compte du principe de sonorité.

Ci-dessous figure un exemple de vocabulaire initial (généralisé aléatoirement) et final (prédit par le modèle ESS à six contraintes C1-C6). Les lignes vides dans le vocabulaire initial correspondent à des mots de longueur zéro.

Initial Vocabulary		Final Vocabulary	
nial	ska	kiliaku	puliasu
lpnssu	klplptai	pitaa	sitilsu
p	lppnntustnu	kukin	saliaku
tpltintit	aiik	situn	pata
aiaautk	iapansk	kupipsu	pat
nlasanusiua		tilianu	lipu
itlliiluau		tinianu	kapulsu
ut	kplk	pisiaku	kapun
p	nsininikaps	tilu	piliatu
alsnsu	ktptailiklt	pitiltu	putu
s	s	kituaku	kupipsu
u	uia	pitu	pitu
aipnaspstta	p	kuniaku	kitiaki
kisikaspap	iinaptapsu	latip	kupunku
	sntikn	tatilsu	silu
iit	upnppiipknu	kuliatu	ninuanu
stiaitsk	ku	kupun	kap
pisun	iiapistitpp	nipultu	pitin
katppaktllt	k	tiliatu	tatua
tnunptnikls	stiaitsk	titilsu	kilia
tl	uupsnklpkuu	ninianu	kuti
tis	nik	pitin	kuti
kstn	kiskpp	kunus	titilsu
saip	iapp	pip	tatuanu
utsikalliap	n	sati	kunu

*Table C.1 : « Example of initial and final vocabularies from the full model. Blanks in the initial vocabulary indicate words of length 0. It is evident that no regular syllable structure exists in the initial vocabulary, whereas words of the final vocabularies are clearly composed of just a few syllable types. » (Redford et al., 2000, p. 21).*

Contrairement au modèle de Berrah (1995), le modèle ESS est basé sur des contraintes de type formel. La seule donnée orientée substance provient de la position mandibulaire des segments que ce soit pour les contraintes liées au principe de sonorité et à celui de l'attaque maximale. Les autres contraintes, bien que faisant référence aux principes phonétiques de distinctivité et d'économie, sont en réalité établies sur des faits typologiques.

Le résultat des simulations est plus proche de l'organisation syllabique d'une langue naturelle, tant au niveau des structures lexicales, que des structures syllabiques. Mais si le modèle ESS est capable de simuler des lexiques à partir d'un lexique initial généré aléatoirement, il ne modélise pas l'émergence de la syllabe (ce que prétend faire le modèle de Berrah) et n'est donc pas adapté pour simuler l'émergence du langage.

Bien sûr, le critère méthodologique fondamental d'évaluation des modèles théoriques de prédiction des structures sonores (phonologique ou syllabique) est la comparaison des simulations avec les données typologiques d'un large échantillon de langues naturelles. Mais selon la théorie « *Frame, then Content* » la syllabe a sa place dans une phonologie qui rendrait compte non seulement du développement des structures sonores de la langue maternelle, mais aussi de l'émergence du langage du point de vue phylogénétique. Les études récentes sur les capacités du nouveau-né à extraire des syllabes (Bijeljic-Babic *et al.*, 1993 ; Bertocini *et al.*, 1995 ; Mattys *et al.*, 2001), comme celles sur les productions du babillage (Davis et MacNeilage, 1995 ; Davis *et al.*, 1999) avec des tentatives de modélisation (Sherkane *et al.*, 2003) ou sur le développement des contrôles moteurs des organes de la parole (Lalevée *et al.*, 2003), ou encore les travaux s'inscrivant dans la voie de Lindblom *et al.* (1984) comme ceux de Schwartz *et al.* (2002) – même s'ils ne traitent pas directement la syllabe – sur les stratégies de parole adulte, montrent que le cadre syllabique est aujourd'hui un élément linguistique incontournable pour comprendre le développement et le fonctionnement du langage.

L'élaboration d'un modèle de simulation de l'émergence lexicale, basé sur les propriétés des systèmes syllabiques, ne pourra se faire sans intégrer les bases de ces réflexions théoriques dans une optique fondamentalement dynamique et évolutive, pour être capable de simuler l'émergence par apprentissage dans l'interaction (langue de l'environnement) d'un langage adaptatif. Un tel projet ne peut se faire qu'en rassemblant des spécialistes des sciences cognitives, sciences du langage, de la linguistique et anthropologie linguistique, de la neuropsychologie du développement, de la modélisation robotique, de la modélisation des structures sonores. L'ICP s'est lancé tout récemment dans un tel défi à travers la réponse à l'appel d'offre 2004 de l'Action Concertée Systèmes complexes en Sciences Humaines et Sociales avec le projet « *Pati, papa ?* » Modélisation de l'émergence d'un langage articulé dans une société d'agents sensori-moteurs en interaction (Resp. J.L. Schwartz)

Dans ce projet, que ce soit pour expérimenter la complexité des lexiques et de leur émergence, pour rechercher les attracteurs syllabiques et lexicaux, ou pour évaluer les simulations de l'émergence de processus langagiers entre agents communicants, les données de langues naturelles dans leur diversité (génétique et géographique, phonétique et phonologique), et l'élaboration de typologies des structures lexicales et syllabiques constituent une étape fondamentale. C'est aussi toute l'orientation que nous avons voulu donner à cette thèse.

Allier toutes ces forces, c'est se donner les moyens, à terme, d'apporter des réponses aux questions suivantes : comment les structures syllabiques interviennent dans la morphogénèse des lexiques ? Comment émergent la diversité dans les syllabes et la complexité dans les lexiques ? Qu'est-ce qui façonne les formes fréquentes et particulières des structures syllabiques et lexicales ? Quels sont les principes systémiques (auto-organisation, non-linéarité) et les niveaux hiérarchiques dans la structuration ? Quels sont les principes dynamiques oeuvrant à différent niveau de structuration : intra et inter syllabique ? Existe-t-il des patrons syllabiques et lexicaux stables ? De quelle manière, à un moment donné, un état de langue va sélectionner des formes stables ? Est-ce que les processus syllabiques sont indépendants d'une finalité globale lexicale ?

L'objectif ambitieux de ces recherches, en ayant pour finalité la compréhension de l'émergence, du développement et du fonctionnement du langage humain, permettra par la même occasion, en se penchant sur les processus et mécanismes des systèmes syllabiques impliqués dans la construction des lexiques des langues naturelles, de trancher sur la question des « *hasards* » du lexique et de l'arbitraire linguistique, en montrant que les syllabes et les lexiques ne répondent pas seulement à un pur formalisme, mais aussi à des contraintes cognitives liées aux systèmes de production et de perception de parole.

## **BIBLIOGRAPHIE**



- (1981), *Ordbok fransk-svensk / Dictionnaire suédois-français*, Berlitz Publishing Company Ltd, Oxford.
- (1999), *Handbook of the International Phonetic Association : A guide to use of the IPA*, Cambridge University Press, Cambridge.
- (1999), *1er colloque international du GDR-1954 Phonologie, "Lenition and Fortition"*, Nice.
- ALIREZAI A., (2003), *Extension d'une plate-forme MATLAB/PC pour le traitement et l'analyse statistique de données lexicales et syllabiques des langues du monde*, Rapport de stage IUT, Université Stendhal.
- ALTMANN G., (1980), « Prolegomena to Menzerath's Law », *Glottometrika*, Vol. 2, pp. 1-10.
- ALVAREZ C. J., CARREIRAS M. & DE VEGA M., (2000), « Syllable-frequency effect in visual word recognition : Evidence of sequential-type processing », *Psicologica*, Vol. 21, pp. 341-374.
- ANDERSON J. M. & JONES C., (1974), « Three Theses concerning Phonological Representation », *Journal of Linguistic*, Vol. 10, pp. 1-26.
- ANDERSON J. M. & EWEN C., (1987), *Principles of Dependency Phonology*, Cambridge University Press, Cambridge.
- ANGOIJARD J.-P., (1997), *Théorie de la syllabe. Rythme et qualité*, CNRS Editions, Gap.
- APILUCK T., (1994), « Perhaps, the tones are in the consonants? », *Mon-Khmer Studies*, Vol. 23, pp. 11-41.
- ARLEO A. & FLAMENT B., (1988), « " Une poule sur un mur "... : rythme et mélodie d'une comptine à partir d'une analyse mingographique », *Le français moderne*, Vol. 56 n° 1/2, pp. 33-59.
- ARLEO A., (1997), « Counting-Out and the Search for Universals », *Journal of American Folklore*, Vol. 110 n° 438, pp. 391-407.
- BACH E. & HARMS R., (1968), *Universals in Linguistic Theory*, Holt, Rinehart and Winston, New York.
- BADDELEY A., (1990), *Human Memory. Theory and Practice*, Laurence Erlbaum Associates, London.
- BAGEMIHL B., (1995), « Language Games and Related Areas », *The Handbook of Phonological Theory*, Ed. by Goldsmith J.A., Blackwell Publishers, Oxford, pp. 697-712.

- BANNERT R., (1998), « Two thousand and one syllables in spoken Standard Swedish : aspects of syllabification », *Phonum*, Vol. 6, pp. 51-81.
- BEAULIEU K., (2001), « La structure interne de la syllabe : ce qu'en disent les lapsus », *Colloque des Etudiants en Sciences du Langage (CESLa)*, Université de Québec à Montréal, Montréal.
- BERRAH R., (1994), *L'émergence des structures sonores : les syllabes CV*, Mémoire de DEA, Institut National Polytechnique de Grenoble.
- BERRAH R., BOË L.-J. & SCHWARTZ J.-L., (1995), « Emergent syllable using articulatory and acoustic principles », *XIIIth International Congress of Phonetic Sciences*, Vol. 1/4, pp. 396-399.
- BERTONCINI J., FLOCCIA C., NAZZI T. & MEHLER J., (1995), « Morae and syllables : rhythmical basis of speech representations in neonates », *Language and speech*, Vol. 38 (4), pp. 311-329.
- BIJELJAC-BABIC R., BERTONCINI J. & MEHLER J., (1993), « How Do 4-Day-Old Infants Categorize Multisyllabic Utterances ? », *Developmental Psychology*, Vol. 4, pp. 711-721.
- BLEVINS J., (1995-96), « The syllable in Phonological Theory », *The Handbook of Phonological Theory*, Ed. by Goldsmith J.A., Blackwell Publishers, Oxford.
- BOË L.-J., SCHWARTZ J.-L. & VALLÉE N., (2000), « Les tendances des structures phonologiques : le poids de la forme sur la substance. Typologie, modélisation et ontogenèse », *La parole : des modèles cognitifs aux machines communicantes*, Ed. by Escudier P. et Schwartz J.L., Hermès, Paris, pp. 283-323.
- BOË L.-J., ABRY C., CATHIARD M., SCHWARTZ J. L., BADIN P. & VALLÉE N., (à paraître en 2004), « Comment les exceptions des handicaps révèlent les universaux phonologiques bimodaux : contraintes audiovisuelles des systèmes consonantiques des langues du monde ? », *Faits de Langues*, Vol. 23, Les exceptions, Orphys.
- BOPP F., (1866), *Grammaire comparée des langues indo-européennes comprenant le sanscrit, le zend, l'arménien, le grec, le latin, le lithuanien, l'ancien slave, le gothique et l'allemand*, Imprimerie Nationale, Paris.
- BRIGHT W., (1970), « Phonological Rules in Literary and Colloquial Kannada », *American Oriental Society*, Vol. 90 (1).
- BRUN J., (2002), *Création d'une interface MATLAB pour l'exploitation d'une base de données lexicale et syllabique*, Rapport de Stage, IUT 2 - Grenoble.

- CHAMBERS D. & REISBERG D., (1985), « Can mental images be ambiguous? », *Journal of Experimental Psychology : Human Perception and Performance*, Vol. 3, pp. 317-328.
- CHARETTE M., (1990), « Licence to Govern », *Phonology*, Vol. 7 (2), pp. 233-253.
- CHOMSKY N., (1965), *Aspects of the Theory of Syntax*, Massachusetts Institute of Technology, Cambridge.
- CHOMSKY N. & HALLE M., (1968), *The Sound Pattern of English*, Harper Row, New York.
- CHOMSKY N., (1987), *La nouvelle syntaxe, concepts et conséquences de la théorie du liage et du gouvernement*, Editions du Seuil, Paris.
- CLEMENTS G. & KEYSER S., (1983), *CV Phonology - A generative theory of the syllable*, The MIT Press, Cambridge.
- COMRIE B., (1980), *Language Universals and Linguistic Typology. Syntax and Morphology*, Basil Blackwell, Oxford.
- CORNEILLE P., (1682), *Le Cid*, Augustin Courbe, Paris.
- COYAUD M., (1980), *Essais de typologie linguistique. Morpho-syntaxe et sémantique*, Jean Favard, Saint-Sulpice-de-Favières.
- CREISSELS D., (1994), *Aperçu sur les structures phonologiques des langues négro-africaines*, Ellug Université Stendhal, Grenoble.
- CROFT W., (1990), *Typology and Universals*, Cambridge University Press, Cambridge.
- CROTHERS J., LORENTZ J., SHERMAN D. & VIHMAN M., (1979), *Handbook of phonological data from a sample of the world's languages*, Stanford.
- CUTTING J., (1975), « Predicting initial cluster frequencies by phonetic difference », *Status Report on Speech Research (Haskins Laboratories)*, SR-42/43, pp. 233-239.
- DART S. N., (1987), « A Bibliography of X-Ray Studies of Speech », *UCLA Working Papers in Phonetics*, Vol. 66, pp. 1-97.
- DAVIS B. & MACNEILAGE P. F., (1990), « Acquisition of correct vowel production: a quantitative case study », *Journal of Speech and Hearing Research*, Vol. 33, pp. 16-27.
- DAVIS B. & MACNEILAGE P. F., (1994), « Organization of babbling: a case study », *Language and speech*, Vol. 37 (4), pp. 341-355.
- DAVIS B. & MACNEILAGE P. F., (1995), « The articulatory basis of babbling », *Journal of Speech and Hearing Research*, Vol. 38, pp. 1199-1211.

- DAVIS B., MACNEILAGE P. F. & MATYEAR C., (1999), « Intrasyllabic Patterns in Babbling and Early Speech », *Proceedings of the 13th International Congress of Phonetic Sciences*, Vol. 4, San Francisco.
- DAVIS B., MACNEILAGE P. F. & MATYEAR C. L., (2002), « Acquisition of Serial Complexity in Speech Production: A Comparison of Phonetic and Phonological Approaches to First Word Production », *Phonetica*, Vol. 59(2-3), pp. 75-107.
- DAVIS B. & MACNEILAGE P. F., (2003), « Universal Intrasyllabic Patterns in Early Acquisition », *Proceedings of the 15th International Congress of Phonetic Sciences*, pp. 379-382, Barcelona.
- DE BOER B., (2000), « Self-organization in vowel systems », *Journal of phonetics*, Vol. 28, pp. 441-465.
- DEJONG K. & MCDONOUGH J., (1993), « Tone and Tonogenesis in Navaho », *UCLA Working Papers in Phonetics*, Vol. 84, pp. 169-181.
- DEMOLIN D., (1991), « L'analyse des segments, de la syllabe et des tons dans un jeu de langue mangbetu », *Langages*, Vol. 101, pp. 30-50.
- DERWING B. L., (1992), « A 'Pause-Break' task for eliciting syllable boundary judgments from literate and illiterate speakers preliminary results for five diverse languages », *Language and Speech*, Vol. 35 (1, 2), pp. 219-235.
- DITZINGER T., TULLER B. & KELSO J. A. S., (1997), « Temporal patterning in an auditory illusion: The verbal transformation effect », *Biological Cybernetics*, Vol. 77, pp. 23-30.
- DUCEY V. & ABRY C., (2004), « Le cadre de la parole et le cadre du signe : un rendez-vous développemental », *XXVèmes Journées d'Etude de la Parole*, Fès, Maroc.
- DUFOUR S., PEEREMAN R., PALLIER C. & RADEAU M., (2002), « VoCoLex : A lexical database on phonological similarity between French words », *L'Année Psychologique*, Vol. 102, pp. 725-746.
- DURAND J. & LYCHE C., (2001), « Des règles aux contraintes en phonologie générative », *Revue québécoise de linguistique*, Vol. 30, pp. 91-154.
- EHRET C., (1995), *Reconstructing Proto-Afroasiatic (Proto-afroasian). Vowels, Tone, Consonants and Vocabulary*, University of California Press.
- ENCREVÉ P., (1998), *La liaison sans enchaînement. Phonologie tridimensionnelle et usages du français*, Edition du Seuil, Paris.
- ENGTRAND O., (1986), « Durational Correlates of Quantity and Sentence Stress: a Cross-Language Study of Swedish, Finnish and Czech », *UCLA Working Papers in Phonetics*, Vol. 63, pp. 1-25.

- FANT G., KRUCKENBERG A., ANDERSON R., ABELIN A., ALLWOOD J. & LINDBLAD P., (1999), « Syllable and word prominence in Swedish », *Gothenburg papers in theoretical linguistics*, Vol. 81, pp. 57-60.
- FOUGERON C. & KEATING P. A., (1997), « Articulatory strengthening at edges of prosodic domains », *Journal of the Acoustical Society of America*, Vol. 101, pp. 3728-3740.
- FUDGE E. C., (1969), « Syllables », *Journal of Linguistics*, Vol. 5, pp. 253-286.
- GEDNEY W. J., (1967), « Linguistics in South East Asia: Thailand and Laos », *Current Trends in Linguistics*, Vol. 2, pp. 782-814.
- GOLDMAN J.-P., CONTENT A. & FRAUENFELDER U. H., (1996), « Comparaison des structures syllabiques en français et en anglais », *XXIèmes Journées d'Etude sur la Parole*, pp. 119-122, Avignon.
- GOLDSMITH J., (1990), *Autosegmental and Metrical Phonology*, Blackwell Publisher, Oxford.
- GOLDSMITH J. A. & LARSON G., (1990), « Local Modeling and Syllabification », *Papers from the 26th Annual Regional Meeting of the Chicago Linguistic Society*, Part Two, Ed. by Ziolkowski K.M., Deaton K. et Noske, K., Chicago Linguistics Society, Chicago, pp. 129-142.
- GOLDSMITH J. A., (1993), « La phonologie harmonique », *De Natura Sonorum*, Ed. by Laks B. et Plénat M., Presses universitaires de Vincennes, Saint-Denis, pp. 221-269.
- GOSLIN J., CONTENT A. & FRAUENFELDER U., (1999), « Syllable Segmentation : are humans consistent ? », *Proceedings of Eurospeech*, Budapest.
- GRAMMONT M., (1933), *Traité de phonétique*, Delagrave, Paris.
- GREENBERG J. H., (1957), *Essays in Linguistics*, University of Chicago Press, Chicago.
- GREENBERG J. H., (1966), « Some universals of Grammar with Particular reference to the Order of meaningful Elements », *Universals of Language*, Ed. by Greenberg J.H., The MIT Press, Cambridge, pp. 73-113.
- GREENBERG S., (1998), « A syllable-centric framework for the evolution of spoken language », *Behavioral and brain sciences*, Vol. 21, pp. 518.
- HAECKEL E., (1879), *The Evolution of Man*, C. Keegan Paul and Co, Londres.
- HAGÈGE C., (2000), *Halte à la mort des langues*, Editions Odile Jacob, Paris.

- HAUGEN E., (1956), « The Syllable in Linguistic Description », *For Roman Jakobson*, Ed. by Morris H. et Benveniste E., Mouton and Co, The Hague.
- HJELMSLEV L., (1966), *Le Langage : une introduction*, Editions de Minuit, Paris.
- HOMBERT J.-M., OHALA J. J. & EWAN W. G., (1979), « Phonetic explanations for the development of tones », *Language*, Vol. 55, pp. 37-58.
- HOMBERT J.-M., (1986), « Word Games : Some Implications for Analysis of Tone and Other Phonological Constructs », *Experimental Phonology*, Ed. by. Ohala J.J et Jaeger J.J., Academic Press, Orlando, pp. 175-186.
- HOOPER J., (1976), *An introduction to natural generative phonology*, Academic Press, New York.
- HUFFMAN M., (1985), « Mesures of Phonation Type in Hmong, *UCLA Working Papers in Phonetics*, Vol. 61, pp. 1-25.
- HULST H. V. D., (1996), « The Segment Syllable Connection », *Current trends in phonology models and methods*, Ed. by. Laks B. et Durand J., European Studies Research Institute, Salford, Manchester, pp. 333-361.
- JACKSON M., LADEFOGED P., HUFFMAN M. & ANTOÑANZAS-BARROSO N., (1985), « Measure of Spectral Tilt », *UCLA Working Papers in Phonetics*, Vol. 61, pp. 72-78.
- JAKOBSON R., FANT G. & HALLE M., (1952), *Preliminaries to Speech Analysis. The Distinctive Features and Their Correlates*, MIT Press, Cambridge, Mass.
- JAKOBSON R., (1963), « Les études typologiques et leur contribution à la linguistique historique comparée », *Essais de linguistique générale*, Editions de Minuit, Paris.
- JAKOBSON R., (1969), *Langage enfantin et aphasie*, Editions de Minuit, Paris.
- JANSON T., (1986), « Cross-linguistic trends in the frequency of CV sequences », *Phonology Yearbook*, Vol. 3, pp. 179-195.
- JESPERSEN O., (1904), *Lehrbuch der Phonetik*, Teubner, Leipzig.
- JOHNSON K., (1993), « Acoustic and Auditory Analyses of Xhosa Clicks and Pulmonics », *UCLA Working Papers in Phonetics*, Vol. 83.
- KAHN D., (1976), *Syllable-based Generalizations in English Phonology*, Indiana University Linguistics Club, Bloomington.

- KAMINSKA Z., POOL M. & MAYER P., (2000), « Verbal Transformation: Habituation or Spreading Activation? », *Brain and Language*, Vol. 71, pp. 285-298.
- KAWASAKI H., (1982), *An Acoustical Basis for Universal Constraints on Sound Sequences*, Thèse, University of California.
- KAYE J. D., LOWENSTAMM J. & VERGNAUD J. R., (1988), « La structure interne des éléments phonologiques : une théorie du charme et du gouvernement », *Recherches linguistiques de Vincennes*, Vol. 17, pp. 193-231.
- KAYE J. D., LOWENSTAMM J. & VERGNAUD J.-R., (1990), « Constituent structure and government in phonology », *Phonology Yearbook*, 7, pp. 193-231.
- KEATING P., (1983), « Phonetic and Phonological Representation of Stop Consonant Voicing », *UCLA Working Papers in Phonetics*, Vol. 57, pp. 26-38.
- KEATING P. A., (1990), « Coronal places of articulation », *UCLA working Papers in Phonetics*, Vol. 74, pp. 35-60.
- KEATING P. A., LINDBLOM B., LUBKER J. & KREIMAN J., (1994), « Variability in jaw height for segments in English and Swedish VCVs », *Journal of Phonetics*, Vol. 22, pp. 407-422.
- KEATING P. A., (1995), « Segmental phonology and non-segmental phonetics », *Proceedings of the 13th international congress of phonetic sciences*, Vol. 3, pp. 26-32.
- KLEIN M., (1993), « La syllabe comme interface de la production et de la perception phoniques », *De Natura Sonorum : essais de phonologie*, Ed. by Laks B. et Plénat M., Presses Universitaires de Vincennes, Saint-Denis, pp. 99-141.
- KÖHLEN O., LADEFOGED P., SNYMAN J., TRAILL T. & VOSSEN R., (1989), « The Symbol for Clicks », *UCLA Working Papers in Phonetics*, Vol. 72.
- KRAKOW R. A., (1999), « Physiological organization of syllables : a review », *Journal of Phonetics*, Vol. 27, pp. 23-54.
- KRAUSS M. E., (1973), « Eskimo-Aleut », *Current Trends in Linguistics*, Vol. 10, pp. 798-902.
- KRUSKAL, J. B., (1977), « Multidimensional Scaling and Other Methods for Discovering Structure », *Statistical Methods for Digital Computers*, Vol. III of the Wiley-Interscience series, Mathematical Methods for Digital Computers Ed. by Kurt E., Ralston A. et Wilf H.S., Wiley, New York, pp. 296-339.
- KURYLOWICZ J., (1948), « Contribution à la théorie de la syllabe », *Bulletin de la société polonaise de linguistique*, Vol. 8, pp. 80-115.
- LABRUNE L., (2001), « Structure de la syllabe japonaise », *Faits de Langues japonais-coréen*, Ed. by Blin R. et Tamba I., Ophrys, Paris, pp. 111-122.

- LACHERET-DUJOUR A. & BEAUGENDRE F., (1999), *La prosodie du français*, CNRS Editions, Paris.
- LADEFOGED P., (1967), *Three areas of experimental phonetics*, Oxford, London.
- LADEFOGED P. & TRAILL T., (1980), « The Phonetic Inadequacy of Phonological Specifications of Clicks », *UCLA Working Papers in Phonetics*, Vol. 49.
- LADEFOGED P. & TRAILL T., (1980), « Instrumental Phonetic Fieldwork », *UCLA Working Papers in Phonetics*, Vol. 49.
- LADEFOGED P. & MADDIESON I., (1996), *The Sounds of World's Languages*, Blackwell Publishers Ltd, Oxford.
- LAKS B., (1995), « A Connexionist Account of French Syllabification », *Lingua*, Vol. 95, pp. 51-76.
- LAKS B. & PLÉNAT M. (Eds.), *De natura sonorum : essais de phonologie*, Presses Universitaires de Vincennes, Saint-Denis.
- LALÉVÉE C., (2003), « Development of Speech Frame Control: A Longitudinal Study of the Oral/Nasal Control », *15th International Congress of Phonetic Sciences*, pp. 2003-2006, Barcelona.
- LALÉVÉE C., (2003), *Développement du contrôle de la production de parole, du cadre protosyllabique vers la syllabe : suivi audiovisuel de 2 enfants de 6 à 12 mois*, Mémoire de DEA, Université Stendhal - Grenoble 3.
- LAVER J., (1980), *The phonetic description of voice quality*, Cambridge University Press, Cambridge.
- LAZARD G., (2000), « Que cherchent les chercheurs ? », *Bulletin de la Société de Linguistique de Paris*, Vol. XCV, fasc. 1, pp. 445-460.
- LIBERMAN A. M. & G. M. I., (1985), « The Motor Theory of Speech Perception Revised », *Cognition*, Vol. 21, pp. 1-36.
- LILJENCRANTS J. & LINDBLOM B., (1972), « Numerical Simulation of Vowel Quality Systems : the Role of Perceptual Contrast », *Language*, Vol. 48, pp. 839-862.
- LINCOLN N. J., (1985), « Vowel, consonant, and resonant in Kwakwala », *Faculté des lettres et sciences humaines de Nice, Annales* 52.
- LINDAU M., (1982), « Phonetic Differences in Glottalic Consonants », *UCLA Working Papers in Phonetics*, Vol. 54, pp. 66-80.
- LINDBLOM B., (1975), « Experiments in Sound Structure », *8e Congrès International des Sciences Phonétiques*, Leeds.
- LINDBLOM B., (1983), « On the Teleological Nature of Speech Processes », *Speech Communication*, Vol. 2, pp. 155-158.



- LINDBLOM B., MACNEILAGE P. F. & STUDDERT-KENNEDY M., (1984), « Self-organizing Processes and the Explanation of Phonological Universals », *Explanations of language universals* Ed. by Butterworth B., Comrie, B. et Dahl, O., Mouton, pp. 181-203.
- LINDBLOM B., (1986), « Phonetic Universals in Vowel Systems », *Experimental Phonology*, Ed. by Ohala J.J., Academic Press, Orlando, pp. 13-44.
- LINDBLOM B., (1990), « Models of phonetic variation and selection », *PERLIS (University of Stokholm)*, Vol. 11, pp. 65-100.
- LINDBLOM B., (2000), « Developmental origins of adult phonology : The interplay between phonetic emergents and the evolutionary adaptations of sound patterns », *Phonetica*, Vol. 57, pp. 297-314.
- LINDBLOM B. & ENGSTRAND O., (1989), « In what sense is speech quantal ? », *Journal of Phonetics*, Vol.17, pp. 107-121.
- LOWENSTAMM J., (1996), « CV as the only Syllable Type », *Current trends in phonology models and methods*, ED. by Durand J. et Laks B., European Studies Research Institute, Salford (Manchester), pp. 419-441.
- MACKAY D. G., WULF G., YIN C. & ABRAMS L., (1993), « Relations between word perception and production: New theory and data on the verbal transformation effect », *Journal of Memory and Language*, Vol. 32, pp. 624-646.
- MACNEILAGE P. F. & DAVIS B., (1993), « Motor explanations of babbling and early speech patterns », *Changes in Speech and Face Processing in Infancy: A Glimpse at Developmental Mechanisms of Cognition*, Ed. by Boysson-Bardie B., Schonen S., Jusczyk P., MacNeilage P. et Morton J., Kluwer, Dordrecht, pp. 341-352.
- MACNEILAGE P. F., (1998), « The Frame/Content theory of evolution of speech production », *Behavioral and brain sciences*, Vol. 21, pp. 499-546.
- MACNEILAGE P. F. & DAVIS B., (2000), « On the origin of internal structure of word forms », *Sciences*, Vol. 288.
- MACNEILAGE P. F. & DAVIS B., (2001), « The role of rhythmic cyclicities in infant action development », *Developmental Science*, Vol. 4:1, pp. 79-83.
- MADDIESON I., (1984), *Patterns of sounds*, Cambridge University Press, Cambridge.
- MADDIESON I., (1986), « The Size and structure of Phonological Inventories : Analysis of UPSID », *Experimental Phonology*, Ed. by Ohala J.J., Academic Press, New York, pp. 105-123.
- MADDIESON I. & PRECODA K., (1989), « Updating UPSID », *UCLA Working Papers in Phonetics*, Vol. 74, pp. 104-111.

- MADDIESON I., (1991), « Investigating linguistic universals », *Proceedings of the 12th International Congress of Phonetic Sciences*, Vol. 1, pp. 346-354, Aix-en-Provence.
- MADDIESON I. & PRECODA K., (1992), « Syllable structure and phonetic models », *Phonology*, Vol. 9, pp. 45-60.
- MADDIESON I., (1993), « The structure of segment sequences », *UCLA Working Papers in Phonetics*, Vol. 83, pp. 1-8.
- MALHERBE M., (1993), *Les langages de l'humanité. Une encyclopédie des 3000 langues parlées dans le monde*, Seghers, Paris.
- MALMBERG B., (1974), *Manuel de phonétique générale : introduction à l'analyse scientifique de l'expression du langage*, Picard A. et J., Paris.
- MATTYS S. L. & JUSCZYK P. W., (2001), « Do Infants Segment Words or Recurring Contiguous Patterns? », *Journal of Experimental Psychology : Human Perception and Performance*, Vol. 27 (3), pp. 644-655.
- MCDONOUGH J., (1993a), « The Phonological Representation of Laterals », *UCLA Working Papers in Phonetics*, 83, pp. 19-30.
- MCDONOUGH J. & AUSTIN-GARRISON M., (1994), « Vowel Enhancement and Dispersion in the Vowel Space of Western Navaho: a study of Traditional Navajo Speakers », *UCLA Working Papers in Phonetics*, 87, pp. 93-104.
- MCDONOUGH J. & LADEFOGED P., (1993b), « Navajo Stops », *UCLA Working Papers in Phonetics*, Vol. 84, pp. 151-164.
- MCDONOUGH J., LADEFOGED P. & GEORGE H., (1993c), « Navajo Vowels and Phonetic Universal Tendencies », *UCLA Working Papers in Phonetics*, Vol. 84, pp. 143-150.
- MEHLER J., DOMMARGUES J. Y., FRAUENFELDER U. & SEGUI J., (1981), « The syllable's role in speech segmentation », *Journal of Verbal Learning and Verbal Behavior*, Vol. 20, pp. 298-305.
- MEILLET A. & COHEN M., (1952), *Les langues du monde*, CNRS Paris VII, Paris.
- MENZERATH P., (1954), *Die Architektonik des deutsch Wortschatzes*, Dummler, Bonn.
- MEYNADIER Y., (2001), « La syllabe phonétique et phonologique : une introduction », *Travaux Interdisciplinaires du Laboratoire Parole et Langage*, Vol. 20, pp. 91-148.
- MÉTOZ L., (2001), *L'hypothèse des "racines mondiales" de Merritt Ruhlen : analyse méthodologique et évaluation statistique*, Mémoire de maîtrise, Université Stendhal.
- MÉTOZ L., VALLÉE N., ROUSSET I., BOË L.-J. & BESSIÈRE P., (2002), « Des formes phonétiques aux proto-formes de la langue originelle. La théorie de Merritt Ruhlen à l'épreuve des probabilités », *XXIVèmes Journées d'Etude sur la Parole*, Nancy.

- MÉTOZ L., VALLÉE N., ROUSSET I., BOË L.-J. & BESSIÈRE P., (2003), « L'hypothèse des racines universelles de Merritt Ruhlen. Evaluation statistique et analyse méthodologique », *Les universaux sonores*, Presses Universitaires de Rennes, à paraître.
- MOLINU L., (1998), *La syllabe en Sarde*, Thèse de Doctorat, Université Stendhal - Grenoble 3.
- MOLINU L. & ROMANO A., (1999), « La syllabe dans un parler roman de l'Italie du Sud », *Journée d'Etudes Linguistiques "SyllabeS"*, AAI, pp. 148-153, Nantes.
- MONINO Y. & ROULON-DOKO P., (1972), *Phonologie du Gbaya Kara'Bodoe de Ndongue Bongowen : région de Bouar, république Centrafricaine*, Bibliothèque du SELAF, Paris.
- NOEL P. & DELAFOSSE P. D. M., (1923), *Petit manuel français-kanouri*, Librairie orientaliste P. Geuthner, Paris.
- OHALA J. J., (1980), « The application of phonological universals in speech pathology », *Speech and Language, advances in Basic Research and Practice*, Vol. 3, Ed. by Lass N.J., Academic Press, New York, pp. 75-94.
- OHALA J., (1990), « There is no interface between phonology and phonetics », *Journal of phonetics*, Vol. 18 : 2, pp. 153-171.
- OOIJEN B. V., BERTONCINI J., SANSAVINI A. & MEHLER J., (1997), « Do weak syllables count for newborns? », *Journal of the Acoustical Society of America*, Vol. 102 (6), pp. 3735-3741.
- PAMIES BERTRAN A., (1999), « Prosodic Typology: On the Dichotomy between Stress-Timed and Syllable-Timed Languages », *Language Design*, Vol. 2, pp. 103-130.
- PARKER G. J., (1969), *Ayacucho Quechua Grammar and Dictionary*, Mouton, The Hague, Paris.
- PÉRENNOU G. & DE CALMÈS M., (2002), « Ressources lexicales BDLex-v2.1.2 », *ELRA/ELDA*.
- PIKE K. L. & PIKE E. V., (1947), « Immediate Constituents of Mazateco Syllables », *International journal of American Linguistics*, Vol. 13, pp. 78-91.
- PRINCE A. & SMOLENSKY P., (1993), *Optimality Theory: Constraint Interaction in Generative Grammar*, Technical Report 2, Center for Cognitive Science, Rutgers

- University.
- RAMANUJAN A. K. & MASICA C., (1969), « Toward a Phonological Typology of the Indian Linguistic Area », *Current Trends in Linguistics*, Vol. 5, pp. 543-577.
- REDFORD M. A. & DIEHL R. L., (1999), « The relative perceptual distinctiveness of initial and final consonants in CVC syllables », *Journal of the Acoustical Society of America*, Vol. 106 (3), pp. 1555- 1565.
- REDFORD M. A., (1999), *An Articulatory Basis for the Syllable*, PhD thesis, The University of Texas.
- REISBERG B., FRANSSSEN E., SCLAN S., KLUGER A. & FERRIS S., (1989), « Stage specific incidence of potentially remediable behavioral symptoms in aging and Alzheimer disease », *Bulletin of Clinical Neurosciences*, Vol. 54, pp. 95-112.
- RINGE D. A., (1996), « The Mathematics of “Amerind” », *Diachronica*, Vol. XIII : 1, pp. 135-154.
- ROBERTS A. H., (1965), *A statistical linguistic analysis of American English*, Mouton, The Hague.
- RUHLEN M., (1987), *A Guide to the World's Languages*, Stanford University Press, Californie, Stanford.
- RUHLEN M., (1997), *L'origine des langues*, Belin, Saint-Germain-du-Puy.
- SALA M. & VINTILA-RADULESCU I., (1982), *Les langues du monde. Petite Encyclopédie*, Les Belles Lettres, Paris.
- SANDS B. E., (1990), « Some of the Acoustical Characteristics of Xhosa Clicks », *UCLA Working Papers in Phonetics*, Vol. 74.
- SANDS B. E., (1991), « Evidence for Clicks Features: Acoustic Characteristics of Xhosa clicks », *UCLA Working Papers in Phonetics*, Vol. 80, pp. 6-37.
- SAPIR E. D., (1953), *Le langage. Introduction à l'étude de la parole*, Payot, Paris.
- SAPORTA S., (1955), « Frequency of consonant clusters », *Language*, Vol. 31, pp. 25-30.
- SATO M., (2001), *La nature articulatoire de la boucle articulatoire*, Mémoire de DEA, Institut National Polytechnique de Grenoble.
- SATO M., SCHWARTZ J.-L., CATHIARD M.-A. & ABRY C., (2002), « Les contraintes de contrôle articulatoire intra-syllabique dans la mémoire de travail verbale », *XXIVèmes Journées d'Études sur la Parole*, pp. 245-248, Nancy, France.
- SATO M. & SCHWARTZ J.-L., (2003), « Linking Speech, Imagery and Memory processes: Intrasyllabic Articulatory Control Constraints in Verbal Transformations », *15th*

- International Congress of Phonetic Sciences*, pp. 435-438, Barcelona.
- SAUSSURE F. D., (1916), *Cours de linguistique générale*, Payot, Paris.
- SCHLEICHER A., (1863), *Die darwinische Theorie und die Sprachwissenschaft*, Böhlau, Weimar.
- SCHUH R. G., (1981), « A Dictionary of Ngizim », *Linguistics*, Vol. 99.
- SCHWARTZ J.-L., BOË L.-J., VALLÉE N. & ABRY C., (1997), « The Dispersion-Focalisation Theory of vowel Systems », *Journal of Phonetics*, Vol. 25, pp. 255-286.
- SCHWARTZ J. L., ABRY C., BOË L. J. & CATHIARD M. A., (2002), « Phonology in a theory of perception-for-action-control », *Phonetics, Phonology and Cognition*, Ed. by Laks B. et Durand J., Oxford University Press, Oxford, pp. 244-280.
- SELKIRK E. O., (1982), « The syllable », *The structure of phonological representations*, Ed. by Van der Hulst et Smith, University of Massachusetts, Amherst.
- SELKIRK E., (1984), « On the major Class Features and Syllable Theory », *Language and Sound Structure*, Ed. by Aronoff M. et Oerhle R.T., MIT Press, Cambridge.
- SERKHANE J. E., (2003), « Simulating Vocal Imitation in Infants: Using a Growth Articulatory Model and Speech Robotics », *15th International Congress of Phonetic Sciences*, Barcelona.
- SHINEN D. C., (1962), « Notes on Saint Lawrence Island Yupik phonology », *Science in Alaska : Proceedings of the 12th Alaska Science Conference*, Vol. 16-17, Ann Arbor : Edwards Brothers.
- SHIVANAND V., (1994), *Studies in Kannada sociolinguistics*, Kannada Chair Publication, Varanasi.
- SHOAF L. S. & PITT M. A., (2002), « Does node stability underlie the Verbal Transformation Effect? A Test of Node Structure Theory », *Perception & Psychophysics*, Vol. 64, pp. 795-803.
- STEFANUTO M., (1997), *Apport de la neurocognition comparée, développementale et pathologique dans l'étude des tendances universelles phonologiques*, Mémoire de DEA, Université Stendhal - Grenoble 3.
- STERIADE D., (1982), *Greek Prosodies and the Nature of Syllabification*, Phd Thesis, MIT.
- STETSON, (1951), *Motor phonetics : A study of speech movements in action*, North-Holland Publishing Compagny, Amsterdam.
- STEVENS K. N., (1972), « The Quantal Nature of Speech : Evidence from Articulatory-Acoustic Data », *Human Communication : a Unified View*, Ed. by Denes P.B. , , McGraw-Hill, New York pp. 51-66.
- STEVENS K. N., (1989), « On the quantal nature of speech », *Journal of phonetics*, Vol. 17,

- pp. 3-45.
- SVANTESSON J.-O., (2001), « Phonology of a southern Swedish idiolect », *Working Papers*, Vol. 49, pp. 156-159.
- THONGKUM T. L., (1987), « Phonation Types in Mon-Kmer Languages », *UCLA Working Papers in Phonetics*, Vol. 67, pp. 29-48.
- THONGKUM T. L., (1987), « Another look at the register distinction in Mon », *UCLA Working Papers in Phonetics*, Vol. 67, pp. 132-165.
- TRAILL A., (1985), *Phonetic and phonological studies of !Xóo Bushman*, Buske H., Hamburg.
- TRAILL A., (1994), *A !Xóo dictionary*, Köppe R., Köln.
- TRIFIT A., (2000), *Syllabe : architecture et modèles*, Mémoire de DEA, Université Paris X - Nanterre.
- TROUBETZKOY N. S., (1939), « Grundzüge des Phonologie », *Travaux du Cercle Linguistique de Prague*, Vol. 7, traduit en français par Cantineau J. en 1970 sous le titre : « Principes de phonologie », Klincksieck, Paris.
- TULLER B. & KELSO J. A. S., (1990), « Phase transitions in speech production and their perceptual consequences », *Attention and Performance XIII*, Ed. by Jeannerod M., Erlbaum, Hillsdale, pp. 429-452.
- TULLER B. & KELSO J. A. S., (1991), « The Production and Perception of Syllable Structure », *Journal of Speech and Hearing Research*, Vol. 34 (3), pp. 501-508.
- VALLÉE N., BOË L.-J., SCHWARTZ J.-L., BADIN P. & ABRY C., (2002), « The weight of substance in phonological structure tendencies of the world's languages », *ZAS Papers in Linguistics*, pp. 145-168, Berlin.
- VALLÉE N., (1994), *Systèmes vocaliques : de la typologie aux prédictions*, Thèse de Doctorat, Université Stendhal - Grenoble 3.
- VALLÉE N., BOË L.-J. & STEFANUTO M., (1998), « Les systèmes consonantiques : des tendances universelles à l'ontogenèse », *XXIIèmes Journées d'Etude sur la Parole*, pp. 241-244, Martigny, Suisse.
- VALLÉE N., BOË L.-J. & STEFANUTO M., (1999), « Typologies phonologiques et tendances universelles. Approche substantialiste », *LINX*, Numéro Spécial, pp. 18-34.
- VALLÉE N., ROUSSET I. & BOË L.-J., (2001), « Des lexiques aux syllabes des langues du monde. Typologies, tendances et organisations structurelles », *LINX*, Université Paris

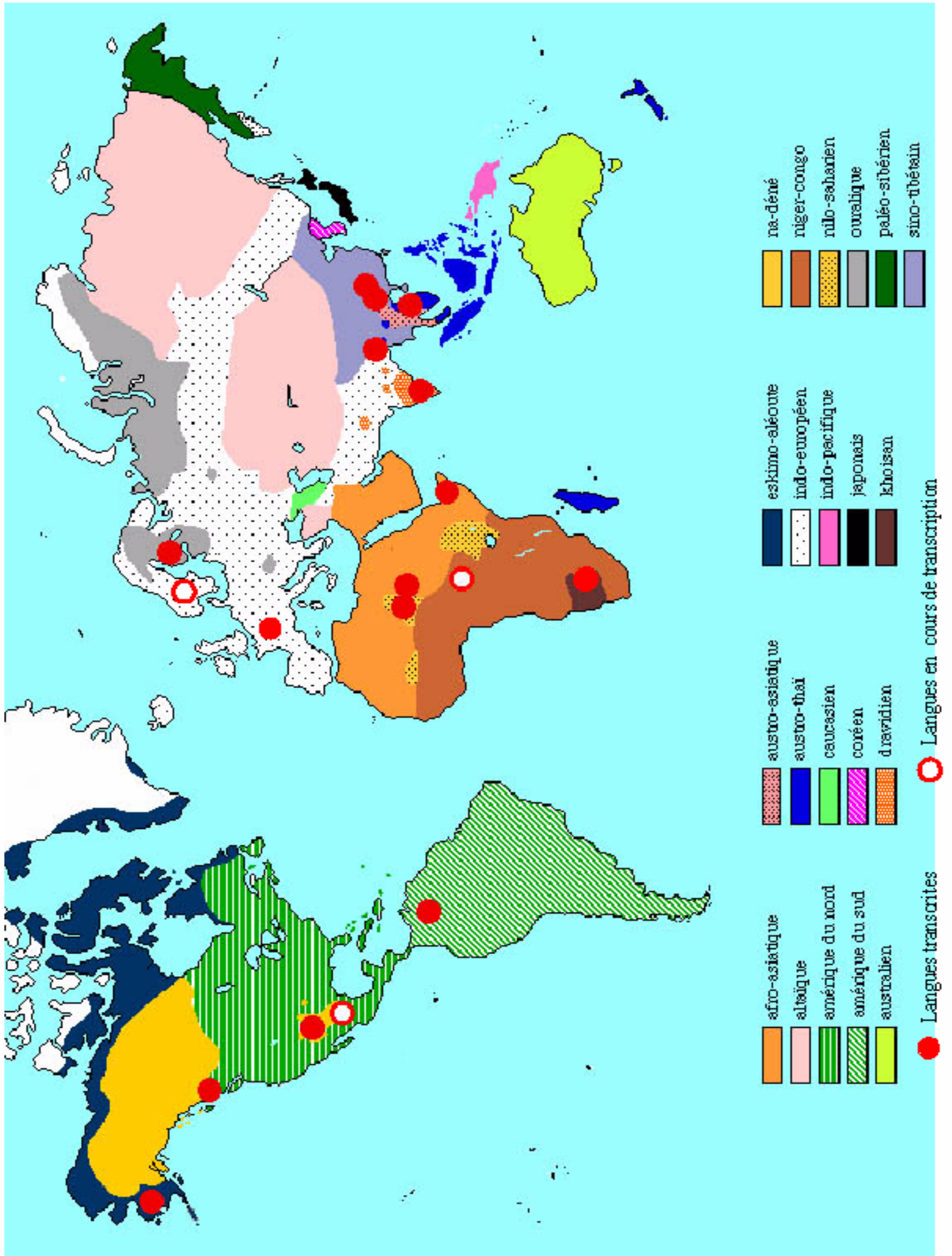
---

X - Nanterre, Numéro spécial.

- VILAIN A., (2000), *Apports de la modélisation des degrés de liberté articulatoires à l'étude de la coarticulation et du développement de la parole*, Thèse de Doctorat, Université Stendhal - Grenoble 3.
- WALTER H., (1975), *La dynamique des phonèmes dans le lexique français contemporain*, Thèse de Doctorat, Université René Descartes - Paris V.
- WARREN M. R. & MEYERS D. M., (1987), « Effects of listening to repeated syllables : Category boundary shifts versus verbal transformation », *Journal of Phonetics*, Vol. 15, pp. 169-181.
- WEBER D. J., (1989), « A Grammar of Huallaga (Huánuco) Quechua », *Linguistics*, Vol. 112.
- ZERLING J.-P., (2000), « Structure syllabique et morphologique des mots à caractères onomatopéique et répétitif en français », *Travaux de l'Institut de Phonétique de Strasbourg*, Vol. 30, pp. 115-162.
- ZÖRNIG P. & ALTMANN G., (1993), « A model for the distribution of syllable types », *Glottometrika*, Vol. 14, pp. 190-196.

## **ANNEXES**





Afar	Ehret (1995)
Finnois	Engstrand (1986)
Français	Walter (1975)
Kannada	Shivanand (1994) Bright (1970)
Kanouri	Noel (1923)
Kwakw'ala	Lincoln (1985)
Navaho	Dejong (1993) McDonough (1993a) McDonough (1993b) McDonough (1993b) McDonough (1994)
Ngizim	Schuh (1981)
Nyah kur	Thongkum (1987a) Thongkum (1987b)
Quechua	Parker (1969) Weber (1989)
Sora	Ramanujan (1969)
Suédois	Engstrand (1986) Fant (1999) Ordbok Dictionnaire (1994) Svantesson (2001)
Thaï	Gedney (1967) Apiluck (1994)
Wa	Thongkum (1987a) Thongkum (1987b) Huffman (1985)
Yup'ik	Krauss (1973) Shinen (1962)
!Xóõ	Johnson (1993) Köhlen et al.(1989) Ladefoged (1980a) Ladefoged et al.(1980b) Sands (1990) Sands (1991) Traill (1985) Traill (1994)

Ouvrages ayant servi pour toutes les langues :

Meillet et Cohen (1952)  
Ladefoged, et Maddieson (1996)  
Malherbe (1993)  
Ruhlen (1987)  
Sala et Vintila-Radulescu (1982)  
Dart (1987)

Types	Langues	Éléments monosyllabiques		Éléments dissyllabiques		Éléments trisyllabiques		Éléments quadrisyllabiques	
1 et 2	<b>Navaho</b>	CVC	36,6 %	CV.CVC	11 %	CV.CV.CV	7,7 %	CV.CV.CV.CV	4,1 %
		CV	4,4 %	CV.CV	6,2 %	CV.CV.CVC	5,1 %		
				CVC.CVC	3,9 %	CVC.CV.CV	3 %		
						CV.CVC.CVC	2,6 %		
	<b>Nyahkur</b>	CVC	41 %	CV.CVC	23,5 %				
		CCVC	22,9 %	CVC.CVC	5 %				
		CV	2,6 %						
	<b>Thaï</b>	CVC	31,9 %	CVC.CVC	17 %	CVC.CV.CVC	2,8 %		
		CV	6,2 %	CV.CVC	10 %	CV.CV.CVC	2,4 %		
		CCVC	2,8 %	CVC.CV	4,7 %	CVC.CVC.CVC	2,1 %		
				CV.CV	2,8 %				
	<b>Wa</b>	CVC	60,9 %						
		CV	18,7 %						
		CCVC	13,6 %						
		CCV	3,3 %						
		VC	2,7 %						
	<b>!Xoo</b>	CV	34 %	CV.CV	32,5 %	CV.CV.CV	2,7 %		
		CVC	13 %	CVC.CV	5,9 %				
				CV.V	5,7 %				
3	<b>Afar</b>	CVC	2 %	CV.CV	11,7 %	CV.CV.CV	9,7 %	CV.CV.CV.CV	2,7 %
				CVC.CV	9,8 %	CVC.CV.CV	7,3 %	CVC.CV.CV.CV	2 %
				CV.CVC	8,3 %	CV.CVC.CV	4,9 %		
				CVC.CVC	4,6 %	VC.CV.CV	3,6 %		
				VC.CV	2,3 %	CVC.CVC.CV	2,9 %		
	<b>Kannada</b>			CV.CV	21,5 %	CV.CV.CV	15,3 %	CV.CV.CV.CV	4,5 %
				CVC.CV	20,4 %	CVC.CV.CV	7,3 %		
				VC.CV	3 %	CV.CVC.CV	4,2 %		
				V.CV	2,7 %	V.CV.CV	3,1 %		
	<b>Kwakw'ala</b>			CV.CVC	10 %	CV.CV.CV	13,1 %	CV.CV.CV.CV	5,2 %
				CV.CV	10 %	CVC.CV.CV	4,1 %		
				CVC.CVC	7,2 %	CV.CV.CVC	4,1 %		
				CVC.CV	3,8 %	CV.CVC.CVC	2,1 %		
				CVCC.CVC	2,4 %				

	<b>Ngizim</b>	CV 5,2 % CVC 2,7 %	CV.CV 26,3 % CVC.CV 13,7 % CV.CVC 9 % CVC.CVC 3 % V.CV 2,2 %	CV.CV.CV 13,7 % CVC.CV.CV 3,2 % CV.CV.CVC 2,7 % CV.CVC.CV 2,4 %	
	<b>Quechua</b>		CV.CVC 13,5 % CV.CV 13,5 % CVC.CV 13 % CVC.CVC 9,4 %	CV.CV.CV 7,8 % CV.CV.CVC 5,9 % CVC.CV.CV 3,6 % CVC.CV.CVC 2,8 % CV.CVC.CV 2,5 %	CV.CV.CV.CVC 2,1 %
	<b>Sora</b>	CVC 6,4 %	CVC.CVC 10,4 % CV.CVC 9,3 % CVC.CV 3,3 % CV.CV 3,1 % V.CVC 2,4 % VC.CVC 2,4 %	CV.CV.CVC 6,4 % CVC.CV.CVC 4,6 % CV.CVC.CVC 2,2 % V.CV.CVC 2 %	CV.CVC.CV.CVC 3,7 % CV.CV.CV.CVC 2,7 % CVC.CV.CV.CVC 2 %
	<b>Suédois</b>	CVC 5 % CVCC 2,6 % CCVC 2,5 %	CVC.CVC 5 % CVC.CV 3,7 % CV.CVC 3,6 % CV.CV 3,5 %		
4	<b>Finnois</b>		CVC.CV 7,9 % CV.CV 6,5 % CV.CVC 2,4 % CVC.CVC 2,2 %	CV.CVC.CV 5,8 % CV.CV.CV 4,9 % CVC.CV.CV 3,6 % CVC.CVC.CV 3,3 % CV.CV.CVC 2,9 % CVC.CV.V 2,2 % CVC.CV.CVC 2,2 %	CV.CV.CV.CV 2,3 % CV.CV.CVC.CV 2,3 %
	<b>Français</b>	CVC 3,7 %	CV.CVC 6,9 % CV.CV 6,6 % CCV.CV 2,2 %	CV.CV.CV 6,8 % CV.CV.CVC 4,1 % V.CV.CV 2,5 %	CV.CV.CV.CV 2,3 %
	<b>Kanouri</b>	CVC 2,3 %	CVC.CV 9,6 % CV.CV 9 % CVC.CVC 7,4 % CV.CVC 5,3 %	CV.CV.CV 10,3 % CV.CV.CVC 7,1 % CV.CVC.CV 5,4 % CV.CVC.CVC 5,1 % CVC.CV.CV 4,7 % CVC.CV.CVC 3,3 %	CV.CV.CV.CV 2,4 %
	<b>Yup'ik</b>		CV.CVC 7,3 % CVC.CVC 5,5 % V.CVC 3 % VC.CVC 2,6 %	CV.CVC.CVC 6,6 % CV.CV.CVC 5,3 % V.CV.CVC 3,5 % V.CVC.CVC 2,4 % CVC.CV.CVC 2,3 % CVC.CVC.CVC 2 %	

Afar	Finnois	Français	Kannada	Kanouri	Kwakw'ala	Navaho	Ngizim	Nyahkur	Quechua	Sora	Thaï	Wa	Yup'ik	!Xöö
te	t̃a	te	ka	te	la	ʔa	tu	ke	pa	i	sa	ku	tuq	li
ba	a	a	ra	ɲin	ʔid	mi	ku	c <sup>h</sup> e	na	ɲen	ka:n	lai	a	ba
ma	neŋ	de	ku	ta	ʔa	ti	ga	pə	ka	ə	na:	lok	ci	le
a	li	e	ru	ʒin	ʔi	pi	wa	te	qa	da:	t <sup>h</sup> i:	tɕaŋ	u	la
ta	ka	ʒe	su	ka	ju	ma	ma	k <sup>h</sup> e	ra	den	k <sup>h</sup> wa:m	tɕe	ka	ni
da	t̃a:	ka	lu	la	sa	hi	a	ce	ku	bən	maj	daŋ	ta	ma
se	va	se	ɖi	wu	ma	ha	ba	le	ta	ren	t <sup>h</sup> a	ka	qa	bu
ka	ʒa	ne	ge	a	ta	ma:	ka	t <sup>h</sup> e	wa	le	ma	kɔŋ	ki	lo
la	t̃o	ē	ɖu	ma	na	ʔi	da	re	a	a	ra	kɔn	na	be
ʒa	t̃i	le	ri	wa	se	ta	ta	p <sup>h</sup> e	ja	ne:	kra	la	tu	bi
na	ja	sjō	gu	ge	a	ki:	ɖu	kəl	tʃij	djen	wan	li	li	na
ja	si	ā	t̃u	ke	da	mi:	gu	mə	tʃa	len	pra	preʔ	uq	lu
to	ma	ti	ma	ra	li	kɔ	ra	ʔe	ma	ma	k <sup>h</sup> on	twi	i	t̃e
sa	ki	si	ta	sə	wa	si	ju	c <sup>h</sup> yʔ	ʒa	te	ni:	tɕai	ni	mi
ra	ko	ko	ga	li	qa	ʔi:	na	c <sup>h</sup> ak	sa	ra	ʔa:	tɕiat	ca	na
li	ri	mā	ke	ku	qwa	ʔat	ja	ʔe	tu	ra:	p <sup>h</sup> a:	vai	ku	sa
le	t̃u	kō	li	i	ŋa	t̃i	ge	c <sup>h</sup> ec	jaj	ga	p <sup>h</sup> u:	dɛh	qu	a
ga	ku	ʒe	lu	tʃin	gi	tʃi	du	pər	ru	ta:	k <sup>h</sup> a:	dʒon	ta:	i
ri	t̃e	ma	ɖa	ba	k'a	tʃi	ru	təm	ti	la:	na:m	gah	ʒa:	se
ɖa	sa	ʒa	pa	na	ʒa	hɔ	mu	tər	ri	ta	si:	ke	ma	ā

Les 20 syllabes les plus fréquentes pour les 15 langues d'ULSID

**Ratios entre occurrences observées et attendues  
pour les structures CV par lieu**

<b>Afar</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>An</b>	1,16	0,68	0,84
<b>Ce</b>	0,88	1,22	0,98
<b>Po</b>	0,99	1,08	1,31

<b>Nyah Kur</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>An</b>	1,08	0,74	0,62
<b>Ce</b>	0,84	0,99	1,48
<b>Po</b>	1,09	1,19	0,63

<b>Finnois</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>An</b>	1,13	0,91	0,85
<b>Ce</b>	0,90	1,06	0,86
<b>Po</b>	1,05	0,95	1,35

<b>Quechua</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>An</b>	1,27	1,01	0,73
<b>Ce</b>	0,89	1,04	0,93
<b>Po</b>	0,94	0,92	1,42

<b>Français</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>An</b>	1,21	0,83	0,31
<b>Ce</b>	0,82	1,23	1,27
<b>Po</b>	0,74	1,08	2,02

<b>Sora</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>An</b>	1,16	0,70	0,74
<b>Ce</b>	0,97	1,11	1,02
<b>Po</b>	0,91	1,06	1,21

<b>Kannada</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>An</b>	1,14	0,77	0,85
<b>Ce</b>	0,91	1,22	0,99
<b>Po</b>	1,00	0,94	1,12

<b>Thaï</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>An</b>	1,07	0,99	0,71
<b>Ce</b>	0,94	0,94	1,08
<b>Po</b>	1,01	1,04	1,22

<b>Kanouri</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>An</b>	1,00	0,72	1,07
<b>Ce</b>	1,06	1,13	0,96
<b>Po</b>	0,84	1,01	1,02

<b>Wa</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>An</b>	0,99	1,01	0,91
<b>Ce</b>	1,00	1,05	0,92
<b>Po</b>	1,01	0,98	1,08

<b>Kwakw'ala</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>An</b>	0,84	0,76	1,00
<b>Ce</b>	1,03	1,08	1,07
<b>Po</b>	1,08	0,97	0,63

<b>Yup'ik</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>An</b>	0,96	1,35	0,94
<b>Ce</b>	0,81	1,19	1,11
<b>Po</b>	1,27	0,44	0,86

<b>Navaho</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>An</b>	1,15	1,45	0,87
<b>Ce</b>	0,80	0,68	0,82
<b>Po</b>	0,98	0,14	1,98

<b>!Xóõ</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>An</b>	0,83	1,11	1,07
<b>Ce</b>	0,68	0,68	1,47
<b>Po</b>	2,50	2,23	0,03

<b>Ngizim</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>An</b>	0,93	1,00	0,60
<b>Ce</b>	0,97	1,10	1,00
<b>Po</b>	1,09	0,76	1,22

**Ratios entre occurrences observées et attendues  
pour les structures CV par mode**

<b>Afar</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Ce</b>	0,96	0,97	1,23
<b>An</b>	0,92	1,13	0,78
<b>Po</b>	1,21	0,86	0,88

<b>Nyah Kur</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Ce</b>	1,15	0,77	0,66
<b>An</b>	0,94	1,05	1,22
<b>Po</b>	0,88	1,18	1,24

<b>Finnois</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Ce</b>	0,81	1,05	1,38
<b>An</b>	0,99	1,09	0,86
<b>Po</b>	1,31	0,68	0,75

<b>Quechua</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Ce</b>	0,94	1,19	0,86
<b>An</b>	0,96	0,79	1,12
<b>Po</b>	1,15	0,84	1,15

<b>Français</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Ce</b>	0,97	0,99	1,18
<b>An</b>	0,90	1,11	0,99
<b>Po</b>	1,20	0,80	0,81

<b>Sora</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Ce</b>	0,97	1,18	0,91
<b>An</b>	1,12	0,73	1,00
<b>Po</b>	0,95	0,92	1,17

<b>Kannada</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Ce</b>	0,99	1,14	1,05
<b>An</b>	0,97	1,10	0,72
<b>Po</b>	1,03	0,72	1,24

<b>Thaï</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Ce</b>	0,97	1,11	0,99
<b>An</b>	1,07	0,81	1,05
<b>Po</b>	1,03	0,89	1,02

<b>Kanouri</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Ce</b>	1,26	0,83	0,85
<b>An</b>	0,51	1,51	1,44
<b>Po</b>	0,92	0,80	0,85

<b>Wa</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Ce</b>	1,04	0,81	1,11
<b>An</b>	0,94	1,17	1,07
<b>Po</b>	1,00	1,07	1,14

<b>Kwakw'ala</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Ce</b>	0,95	1,10	1,01
<b>An</b>	1,22	0,77	0,87
<b>Po</b>	0,89	0,81	1,18

<b>Yup'ik</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Ce</b>	0,90	1,06	1,27
<b>An</b>	0,93	1,37	0,82
<b>Po</b>	1,19	0,52	0,86

<b>Navaho</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Ce</b>	0,90	1,03	1,18
<b>An</b>	1,20	0,85	1,05
<b>Po</b>	0,83	1,30	0,22

<b>!Xóõ</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Ce</b>	0,98	1,44	1,01
<b>An</b>	1,40	2,25	1,87
<b>Po</b>	1,07	0,27	0,65

<b>Ngizim</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Ce</b>	0,97	1,09	1,07
<b>An</b>	0,88	1,14	0,71
<b>Po</b>	1,15	0,70	0,93

**Ratios entre occurrences observées et attendues  
pour les structures CVC par lieu**

<b>Afar</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>Co</b>	0,77	1,24	1,13
<b>Bi</b>	1,29	0,52	1,21
<b>Ve</b>	1,34	0,97	0,51

<b>Nyah Kur</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>Co</b>	0,91	1,00	0,96
<b>Bi</b>	1,30	0,65	0,67
<b>Ve</b>	1,29	0,76	0,49

<b>Finnois</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>Co</b>	1,02	0,99	0,97
<b>Bi</b>	1,01	0,23	1,84
<b>Ve</b>	1,06	1,14	0,67

<b>Quechua</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>Co</b>	0,89	1,38	1,16
<b>Bi</b>	1,06	0,37	1,31
<b>Ve</b>	1,06	0,87	0,84

<b>Français</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>Co</b>	0,94	1,01	1,08
<b>Bi</b>	1,30	0,44	1,07
<b>Ve</b>	1,25	0,67	0,27

<b>Sora</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>Co</b>	0,93	1,16	1,09
<b>Bi</b>	1,03	0,52	1,29
<b>Ve</b>	1,23	0,78	0,40

<b>Kannada</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>Co</b>	0,73	1,16	1,26
<b>Bi</b>	1,54	0,26	0,88
<b>Ve</b>	1,47	0,88	0,35

<b>Thaï</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>Co</b>	0,85	1,03	1,21
<b>Bi</b>	1,39	0,39	0,86
<b>Ve</b>	1,22	1,02	0,38

<b>Kanouri</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>Co</b>	0,79	1,16	1,26
<b>Bi</b>	1,52	0,32	0,44
<b>Ve</b>	1,40	0,95	0,29

<b>Wa</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>Co</b>	1,11	0,90	0,85
<b>Bi</b>	1,38	0,66	0,72
<b>Ve</b>	1,23	0,86	0,61

<b>Kwakw'ala</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>Co</b>	0,73	1,11	1,21
<b>Bi</b>	1,41	0,37	0,65
<b>Ve</b>	1,38	0,99	0,78

<b>Yup'ik</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>Co</b>	0,61	1,36	1,34
<b>Bi</b>	0,70	1,96	2,18
<b>Ve</b>	0,86	1,67	0,65

<b>Navaho</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>Co</b>	0,91	1,08	1,04
<b>Bi</b>	0,00	0,00	0,00
<b>Ve</b>	1,53	0,00	0,00

<b>!Xóõ</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>Co</b>	0,88	3,24	1,00
<b>Bi</b>	1,05	0,00	1,02
<b>Ve</b>	0,00	0,00	0,00

<b>Ngizim</b>	<b>Co</b>	<b>Bi</b>	<b>Ve</b>
<b>Co</b>	0,76	1,32	1,23
<b>Bi</b>	1,35	0,31	1,04
<b>Ve</b>	1,16	1,06	0,40



**Ratios entre occurrences observées et attendues  
pour les structures CVC par mode**

<b>Afar</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Pl</b>	1,10	0,92	0,88
<b>Fr</b>	0,92	0,80	1,14
<b>Na</b>	0,79	1,25	1,31

<b>Nyah Kur</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Pl</b>	0,98	1,45	0,63
<b>Fr</b>	1,08	0,39	0,78
<b>Na</b>	0,98	1,10	0,75

<b>Finnois</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Pl</b>	0,89	1,01	0,87
<b>Fr</b>	1,05	0,85	0,79
<b>Na</b>	0,83	0,72	1,72

<b>Quechua</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Pl</b>	0,92	1,52	0,60
<b>Fr</b>	1,11	0,83	0,87
<b>Na</b>	1,03	1,32	1,31

<b>Français</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Pl</b>	0,83	1,07	0,99
<b>Fr</b>	1,05	0,95	0,98
<b>Na</b>	0,86	0,98	1,19

<b>Sora</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Pl</b>	1,04	1,18	0,87
<b>Fr</b>	0,84	4,58	0,69
<b>Na</b>	0,93	0,72	1,07

<b>Kannada</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Pl</b>	1,01	0,95	0,86
<b>Fr</b>	0,96	1,68	0,74
<b>Na</b>	0,96	1,02	1,15

<b>Thaï</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Pl</b>	0,87	1,27	0,88
<b>Fr</b>	1,84	0,00	0,00
<b>Na</b>	0,97	0,98	1,08

<b>Kanouri</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Pl</b>	0,91	0,93	0,45
<b>Fr</b>	1,25	0,34	1,10
<b>Na</b>	0,81	1,08	1,12

<b>Wa</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Pl</b>	0,80	1,24	1,34
<b>Fr</b>	0,89	1,19	0,81
<b>Na</b>	0,80	1,20	1,24

<b>Kwakw'ala</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Pl</b>	1,25	0,49	0,71
<b>Fr</b>	0,86	0,74	1,44
<b>Na</b>	0,85	1,94	0,98

<b>Yup'ik</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Pl</b>	1,15	0,95	0,77
<b>Fr</b>	0,81	0,95	1,27
<b>Na</b>	0,87	1,42	1,00

<b>Navaho</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Pl</b>	0,87	1,01	0,93
<b>Fr</b>	1,04	1,03	1,03
<b>Na</b>	1,10	1,04	0,00

<b>!Xóõ</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Pl</b>	1,40	0,74	0,67
<b>Fr</b>	0,00	0,00	0,00
<b>Na</b>	0,85	1,10	1,12

<b>Ngizim</b>	<b>Pl</b>	<b>Fr</b>	<b>Na</b>
<b>Pl</b>	0,85	1,03	1,16
<b>Fr</b>	1,17	0,60	1,42
<b>Na</b>	0,95	0,91	1,12

**Ratios entre occurrences observées et attendues  
pour les structures VC par lieu**

<b>Afar</b>	<b>An</b>	<b>Ce</b>	<b>Po</b>
<b>Co</b>	0,98	1,02	0,99
<b>Bi</b>	0,90	1,00	1,12
<b>Ve</b>	0,83	0,93	1,38

<b>Nyah Kur</b>	<b>An</b>	<b>Ce</b>	<b>Po</b>
<b>Co</b>	0,72	1,10	0,85
<b>Bi</b>	0,76	1,02	0,91
<b>Ve</b>	0,61	0,93	1,37

<b>Finnois</b>	<b>An</b>	<b>Ce</b>	<b>Po</b>
<b>Co</b>	1,06	0,96	1,00
<b>Bi</b>	0,62	1,35	1,18
<b>Ve</b>	0,69	1,10	0,75

<b>Quechua</b>	<b>An</b>	<b>Ce</b>	<b>Po</b>
<b>Co</b>	1,15	0,85	1,15
<b>Bi</b>	0,95	1,00	1,08
<b>Ve</b>	1,26	0,65	1,49

<b>Français</b>	<b>An</b>	<b>Ce</b>	<b>Po</b>
<b>Co</b>	0,97	1,21	0,82
<b>Bi</b>	0,71	1,13	1,65
<b>Ve</b>	1,45	0,65	0,73

<b>Sora</b>	<b>An</b>	<b>Ce</b>	<b>Po</b>
<b>Co</b>	1,02	1,07	0,83
<b>Bi</b>	0,65	0,97	1,37
<b>Ve</b>	1,21	0,73	1,39

<b>Kannada</b>	<b>An</b>	<b>Ce</b>	<b>Po</b>
<b>Co</b>	0,94	1,01	1,03
<b>Bi</b>	0,98	0,91	1,17
<b>Ve</b>	1,16	0,99	0,90

<b>Thaï</b>	<b>An</b>	<b>Ce</b>	<b>Po</b>
<b>Co</b>	1,34	0,91	1,03
<b>Bi</b>	0,75	1,10	1,05
<b>Ve</b>	1,08	0,73	1,38

<b>Kanouri</b>	<b>An</b>	<b>Ce</b>	<b>Po</b>
<b>Co</b>	1,30	0,77	0,99
<b>Bi</b>	0,21	1,74	0,57
<b>Ve</b>	0,13	1,31	2,28

<b>Wa</b>	<b>An</b>	<b>Ce</b>	<b>Po</b>
<b>Co</b>	1,10	1,13	1,20
<b>Bi</b>	0,93	1,32	1,53
<b>Ve</b>	0,52	1,07	0,97

<b>Kwakw'ala</b>	<b>An</b>	<b>Ce</b>	<b>Po</b>
<b>Co</b>	1,28	0,93	0,79
<b>Bi</b>	0,08	1,48	0,27
<b>Ve</b>	1,01	0,86	1,69

<b>Yup'ik</b>	<b>An</b>	<b>Ce</b>	<b>Po</b>
<b>Co</b>	1,21	1,00	0,84
<b>Bi</b>	1,30	1,18	0,71
<b>Ve</b>	1,49	1,04	0,67

<b>Navaho</b>	<b>An</b>	<b>Ce</b>	<b>Po</b>
<b>Co</b>	1,08	0,89	0,99
<b>Bi</b>	0,00	0,00	0,00
<b>Ve</b>	0,00	0,00	6,96

<b>!Xóõ</b>	<b>An</b>	<b>Ce</b>	<b>Po</b>
<b>Co</b>	0,00	0,00	0,00
<b>Bi</b>	0,00	0,86	1,15
<b>Ve</b>	2,42	1,57	0,39

<b>Ngizim</b>	<b>An</b>	<b>Ce</b>	<b>Po</b>
<b>Co</b>	1,36	0,96	1,10
<b>Bi</b>	0,58	1,04	0,95
<b>Ve</b>	0,61	1,05	0,89

**Ratios entre occurrences observées et attendues  
pour les structures VC par mode**

<b>Afar</b>	<b>Ce</b>	<b>An</b>	<b>Po</b>
<b>Pl</b>	1,07	0,83	1,04
<b>Fr</b>	1,04	0,91	1,03
<b>Na</b>	0,91	1,17	0,99

<b>Nyah Kur</b>	<b>Ce</b>	<b>An</b>	<b>Po</b>
<b>Pl</b>	1,08	0,88	1,11
<b>Fr</b>	1,29	0,83	1,50
<b>Na</b>	1,10	0,90	0,95

<b>Finnois</b>	<b>Ce</b>	<b>An</b>	<b>Po</b>
<b>Pl</b>	0,80	0,92	0,98
<b>Fr</b>	0,59	0,92	1,33
<b>Na</b>	1,49	0,88	0,79

<b>Quechua</b>	<b>Ce</b>	<b>An</b>	<b>Po</b>
<b>Pl</b>	1,13	1,02	0,66
<b>Fr</b>	0,84	1,08	1,28
<b>Na</b>	0,85	1,15	1,15

<b>Français</b>	<b>Ce</b>	<b>An</b>	<b>Po</b>
<b>Pl</b>	1,13	0,99	0,92
<b>Fr</b>	0,99	0,91	1,04
<b>Na</b>	1,21	0,58	1,52

<b>Sora</b>	<b>Ce</b>	<b>An</b>	<b>Po</b>
<b>Pl</b>	0,79	1,27	1,21
<b>Fr</b>	0,78	1,23	1,27
<b>Na</b>	1,15	0,80	0,87

<b>Kannada</b>	<b>Ce</b>	<b>An</b>	<b>Po</b>
<b>Pl</b>	0,95	1,06	1,02
<b>Fr</b>	1,13	0,91	0,87
<b>Na</b>	1,02	1,00	0,95

<b>Thaï</b>	<b>Ce</b>	<b>An</b>	<b>Po</b>
<b>Pl</b>	0,81	1,27	1,46
<b>Fr</b>	0,00	0,00	7,04
<b>Na</b>	0,93	1,08	0,97

<b>Kanouri</b>	<b>Ce</b>	<b>An</b>	<b>Po</b>
<b>Pl</b>	1,36	0,38	1,44
<b>Fr</b>	1,34	0,37	1,54
<b>Na</b>	0,75	1,52	0,45

<b>Wa</b>	<b>Ce</b>	<b>An</b>	<b>Po</b>
<b>Pl</b>	1,04	1,00	1,03
<b>Fr</b>	0,52	1,45	0,60
<b>Na</b>	1,16	0,81	1,14

<b>Kwakw'ala</b>	<b>Ce</b>	<b>An</b>	<b>Po</b>
<b>Pl</b>	0,58	1,99	1,27
<b>Fr</b>	0,92	0,88	1,65
<b>Na</b>	1,52	0,05	0,15

<b>Yup'ik</b>	<b>Ce</b>	<b>An</b>	<b>Po</b>
<b>Pl</b>	1,36	0,86	0,67
<b>Fr</b>	0,54	1,16	1,44
<b>Na</b>	0,8539	1,0854	1,1195

<b>Navaho</b>	<b>Ce</b>	<b>An</b>	<b>Po</b>
<b>Pl</b>	0,99	1,02	0,95
<b>Fr</b>	0,96	1,04	1,06
<b>Na</b>	1,32	0,85	0,31

<b>!Xóõ</b>	<b>Ce</b>	<b>An</b>	<b>Po</b>
<b>Pl</b>	1,00	0,81	1,20
<b>Fr</b>	0,00	0,00	0,00
<b>Na</b>	1,00	1,07	0,93

<b>Ngizim</b>	<b>Ce</b>	<b>An</b>	<b>Po</b>
<b>Pl</b>	1,14	1,01	0,82
<b>Fr</b>	0,44	1,00	1,33
<b>Na</b>	0,95	1,02	0,92

---

## Consonnes

---

Pour les lieux d'articulation :

bilabial	Bi
labiodental	Lde
coronal	Co
palatal	Pa
vélaire	Ve
labio-vélaire	Lve
uvulaire	Uv
pharyngal	Ph
glottal	Gl

Pour les modes d'articulation :

plosif	Pl
fricatif	Fr
nasal	Na
approximant	Ap
affriqu�	Af
vibrant/battu	Vi

## Voyelles

---

Pour la position avant-arri re de la langue

ant�rieur	An
central	Ce
post�rieur	Po

---

## RÉSUMÉ

Les structures internes de la syllabe et leur combinaison pour former des items lexicaux sont analysées à partir de la constitution d'une base de données informatisée contenant les lexiques de 16 langues découpés en syllabes. L'élaboration de typologies, la recherche de tendances et cooccurrences entre constituants syllabiques et entre syllabes consécutives d'une unité lexicale constituent la majeure partie de cette étude. Les résultats sont analysés à la lumière de travaux antérieurs pouvant fournir une base explicative des tendances, notamment ceux s'inscrivant dans le cadre d'une linguistique orientée substance, et plus particulièrement la théorie « *Frame, then Content* » (MacNeilage, 1998). Utilisant un paradigme expérimental original, nous montrons qu'il existe un corrélat perceptif de l'effet lexical Labial-Coronal. L'ensemble du travail apporte des éléments incontournables dans une perspective de modélisation de l'émergence des structures syllabiques et de la complexité des lexiques.

**Mots clés :** syllabes, typologie des langues, tendances universelles, phonétique/phonologie, linguistique orientée substance, théorie « *Frame, then Content* », modélisation

---

## ABSTRACT

Syllable internal structures and their combinations building lexical items will be analysed in constituting a computerized data base of 16 language lexicons cut up in syllables. Drawing up typologies, we have found tendencies and cooccurrences between syllable components and between two consecutive syllables in the same lexical item. The results was analysed in the light of previous works especially in the field of a substance-based linguistic, and more particularly those in keeping with the "*Frame, then Content*" Theory (MacNeilage, 1998). With an original paradigm testing the asymmetries in verbal transformations we have found a perceptual correlate of the "Labial-Coronal" Effect and interpret these findings in line with previous models of speech perception and our statistical studies on universal organizational syllabic patterns. Globally the thesis gives lot of data and results which will be took into account to build a model of syllable structure emergence and lexical complexity.

**Keywords:** syllables, language typology, universal tendencies, phonetic/phonology, substance-based linguistic, "*Frame, then Content*" theory, modelisation