



HAL
open science

“ Cul et chemise ”, “ Modes et travaux ”, “ Émilie et Nathan ” : étude des principes gouvernant la coordination par “et” de deux mots en français

Graziella Couasnon

► **To cite this version:**

Graziella Couasnon. “ Cul et chemise ”, “ Modes et travaux ”, “ Émilie et Nathan ” : étude des principes gouvernant la coordination par “et” de deux mots en français. Linguistique. Université Michel de Montaigne - Bordeaux III, 2016. Français. NNT : 2016BOR30022 . tel-01404945

HAL Id: tel-01404945

<https://theses.hal.science/tel-01404945>

Submitted on 29 Nov 2016

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 Bordeaux Montaigne

École Doctorale Montaigne - Humanités (ED 480)

THÈSE DE DOCTORAT EN SCIENCES DU LANGAGE

**« *Cul et chemise* », « *Modes et travaux* »,
« *Émilie et Nathan* » : étude des principes
gouvernant la coordination par «et» de
deux mots en français**

Présentée et soutenue publiquement le 11 juillet 2016 par

Graziella COUASNON

Sous la direction de Laurence Labrune

Membres du jury

Mme Hélène GIRAUDO, Chargée de recherche HDR au CNRS, Université Toulouse Jean Jaurès

M. Haïke JACOBS, Professeur, Université Radboud

Mme Laurence LABRUNE, Professeure, Université Bordeaux Montaigne

M. Frédéric LAMBERT, Professeur, Université Bordeaux Montaigne

M. Vincent RENNERT, Professeur, Université Lumière Lyon 2

Table des matières

Table des matières

Conventions d'écriture.....	8
Abréviations.....	8
Signes conventionnels.....	9
Introduction.....	10
Chapitre 1 : Remarques générales.....	13
1.1. Panorama de la littérature anglo-saxonne.....	14
1.1.1. Cooper et Ross (1975)	15
1.1.2. Pinker et Birdsong (1979).....	17
1.1.3. Wright, Hay et Bent (2002, 2005).....	18
1.1.4. Renner (2006, 2014).....	19
1.1.5. Récapitulatif des facteurs relevés pour les composés binomiaux en anglais.....	21
1.2. Cadre théorique.....	23
1.2.1. Transcription du corpus et décompte syllabique.....	25
1.2.2. Le schwa.....	25
1.2.3. Organisation par traits.....	29
Chapitre 2 : Aspects méthodologiques, protocoles et mises en œuvre.....	33
2.1. Protocole de constitution des corpus.....	34
2.1.1. Pourquoi différents corpus.....	34
2.1.2. Principes généraux et normes méthodologiques de la constitution des corpus.....	36
2.1.3. Protocole de constitution du corpus TOM ET ZOÉ.....	39
2.1.4. Protocole de constitution du corpus TARN ET GARONNE.....	43
2.1.5. Protocole de constitution du corpus ARTS ET MÉTIERS.....	44
2.1.6. Protocole de constitution du corpus CUL ET CHEMISE.....	44
2.1.7. Protocole de constitution du corpus LISSE ET SOYEUX.....	47
2.1.8. Protocole de constitution du corpus VITE ET BIEN.....	49
2.1.9. Présentation récapitulative des corpus.....	50
2.2. Protocole de constitution des questionnaires d'enquête.....	51
2.2.1. Pourquoi différents questionnaires d'enquête.....	52
2.2.2. Principes généraux et normes méthodologiques de la constitution des	

questionnaires d'enquête.....	53
2.2.3. Protocole de constitution du questionnaire d'enquête ÉLISE et ZOÉ	56
2.2.4. Protocole de constitution du questionnaire d'enquête ÉRIC ET AURIC.....	60
2.2.5. Protocole de constitution du questionnaire d'enquête PAUL ET PASCAL	61
2.2.6. Protocole de constitution du questionnaire d'enquête ÉDOUARD ET EDMOND	62
2.2.7. Protocole de constitution du questionnaire d'enquête NAROD ET NANOR.....	63
2.2.8. Protocole de constitution du questionnaire d'enquête LEMOTO ET LEMOTAN	65
2.2.9. Présentation récapitulative des questionnaires d'enquête.....	66
2.3. Exploitation des données.....	67
2.3.1. Remarques générales.....	68
2.3.2. Critères classificatoires.....	71
Chapitre 3 : Facteurs extraphonologiques gouvernant les appariements.....	74
3.1. L'ordre alphabétique.....	75
3.2. Le facteur sériel.....	77
3.3. Le facteur syntaxique.....	80
3.4. Le facteur sémantique.....	81
3.5. Le facteur psycho-affectif.....	83
3.6. La fréquence des mots.....	85
3.7. Le facteur morphologique.....	88
3.8. Le facteur culturel.....	92
Chapitre 4 : Facteurs phonologiques gouvernant les appariements.....	95
4.1. La taille des mots.....	96
4.1.1. Le nombre de syllabes.....	96
4.1.1.1. Longueur relative de M1 et de M2.....	96
4.1.1.2. Nombre de syllabes de M1 et M2.....	102
4.1.1.3. Corrélation entre nombre de syllabes de M1 et nombre de syllabes de M2	110
4.1.1.4. Bilan et analyse des résultats concernant le nombre de syllabes.....	118
4.1.2. Initiale vocalique et nombre de syllabes.....	122
4.1.2.1. Longueur relative -#V_ de M1 et de M2.....	122
4.1.2.2. Nombre de syllabes -#V_ de M1 et M2.....	127
4.1.2.3. Corrélation entre nombre de syllabes -#V_ de M1 et nombre de syllabes de M2.....	136

4.1.2.4. Bilan et analyse des résultats concernant le nombre de syllabes -#V_.....	144
4.2. Début de mots.....	145
4.2.1. Nature du segment initial.....	145
4.2.1.1. Les initiales vocaliques.....	150
4.2.1.1.1. Répartition phonémiques des initiales vocaliques	150
4.2.1.1.2. Répartition par traits distinctifs des initiales vocaliques.....	160
4.2.1.1.3. Bilan et analyse des résultats concernant #V_	164
4.2.1.2. Consonne.....	167
4.2.1.2.1. Répartition phonémique des consonnes initiales	167
4.2.1.2.2. Répartition par traits distinctifs des consonnes initiales.....	175
4.2.1.2.3. Bilan et analyse des résultats concernant #C_	178
4.2.1.3. Glide.....	179
4.2.2. Syllabe initiale.....	180
4.2.2.1. Structure de la syllabe initiale vocalique.....	180
4.2.2.2. Structure de la syllabe initiale consonantique.....	184
4.2.2.3. Structure de la syllabe initiale commençant par un glide.....	191
4.3. Fin de mots.....	192
4.3.1. Nature du segment final.....	192
4.3.1.1. Voyelle.....	196
4.3.1.1.1. Répartition phonémique des finales vocaliques.....	196
4.3.1.1.2. Répartition par traits distinctifs des finales vocaliques.....	207
4.3.1.1.3. Bilan et analyse des résultats concernant les finales vocaliques.....	215
4.3.1.2. Consonne.....	219
4.3.1.2.1. Répartition phonémique des finales consonantiques.....	220
4.3.1.2.2. Répartition par traits distinctifs des finales consonantiques.....	227
4.3.1.2.3. Bilan et analyse des résultats concernant les finales consonantiques. .	237
4.3.1.3. Glide.....	243
4.3.2. Syllabe finale.....	244
4.3.2.1. timbre de la syllabe finale.....	245
4.3.2.1.1. Répartition phonémique du noyau des syllabes finales fermées.....	248
4.3.2.1.2. Répartition par traits distinctifs du noyau des syllabes finales fermées253
4.3.2.1.3. Bilan et analyse des résultats de la répartition des noyaux de syllabes257
finales fermées.....	

4.3.2.2. structure de la syllabe finale.....	261
4.3.2.2.1. Patron de syllabes finales.....	261
4.3.2.2.2. Allongement vocalique	268
4.3.2.2.3. Bilan et analyse des résultats des syllabes finales fermées.....	272
4.3.3. Les phénomènes de liaison.....	274
Conclusion.....	283
Bibliographie.....	287
ANNEXES.....	297

Remerciements

En premier lieu, je tenais à remercier particulièrement les membres de mon jury. Tous, de diverses manières, avaient une importance particulière dans mon parcours de thèse.

Merci à Vincent Renner. Vos travaux rigoureux et inspirants sont un modèle pour moi. Merci de m'avoir fait l'honneur de faire partie de ce jury.

Merci à Hélène Giraud, merci de m'avoir permis de dédramatiser les statistiques. Grâce à toi, j'ai pris la mesure de ces outils fantastiques. Merci de ta patience.

Merci à Frédéric Lambert, le hasard a voulu que nous nous rencontrions un mois avant mon inscription en Sciences du Langage. Vous avez mis un visage sur cette reprise d'études qui alors me terrorisait. Merci de vos encouragements.

Merci à Haike Jacobs. Je ne pouvais concevoir une soutenance de thèse sans ta présence. De nos premiers échanges en colloque à aujourd'hui, merci de ta bienveillance constante.

Merci à Laurence Labrune. Merci pour tous ces moments où vous avez trouvé les mots pour me pousser, mais aussi pour m'encourager et pour me rassurer. J'ai appris beaucoup sur moi tout au long de ce travail de thèse, et pour beaucoup, je vous le dois.

Merci à tous mes anciens enseignants et collègues de l'université Bordeaux 3 / Bordeaux-Montaigne, ainsi qu'aux membres de mon équipe de recherche CLLE-ERSS. Mwata, Claude, Anna, Agnès, Henri, tous m'avaient transmis un peu de vous de la première année de licence jusqu'à aujourd'hui. Un merci particulier à Catherine Mathon, pour tes paroles rassurantes, à Gilles Boyé, pour ta façon toute personnelle d'être un lanceur d'idée avec discrétion et humilité, et à Joan Busquets, merci de ton amitié, et oui, navrée, je fais toujours de la phonologie !

Une pensée pour les doctorants ou jeunes docteurs qui m'ont fait partager leurs expériences et leurs conseils. Shureka, Nancy, Marine, Sandra, Antoine et Angela, mille mercis.

Un grand merci à l'École Doctorale de Bordeaux-Montaigne, merci de votre professionnalisme mais également merci de votre patience, et de votre dévouement tout au long de mon parcours.

Je tenais à remercier aussi les « petites mains » de l'Université, tous ces gens qui de loin en

loin m'ont accompagnée et m'ont soutenue. Je ne peux pas tous les citer mais je tiens particulièrement à remercier le personnel de la bibliothèque Henri Guillemin et toi, Brigitte, pour ta gentillesse et tes conseils. Et merci au Sirtaki. Que serait un étudiant sans un café CROUS ?! Merci à Dominique et à Michael pour vos sourires quotidiens.

Une pensée également pour mes anciens étudiants, cobayes volontaires et enthousiastes. Merci d'avoir joué le jeu.

Enfin, un grand merci à la plus grande équipe de pompoms de la Terre !

Je vais commencer par toi, Pierrick Legrand, super héros de la dernière minute et des situations compliquées. Merci à toi, et à ton collègue Vincent, pour vos conseils précieux et pour l'éclairage scientifique que vous m'avez apporté.

Merci à toi Céline, pour m'avoir fait profiter de ton expérience en matière de correction. Une pensée pour ta douce Emila.

Merci à ma Dream Team, fidèles d'entre les fidèles, du bout du monde, ou de la France, merci de vos conseils, corrections, traductions, mais merci surtout, Anouck, Émilie, Estelle, Philippine et Sébastien, pour vos messages quotidiens, votre soutien et votre amitié indéfectible. Du fond du cœur, et avec toute mon affection, merci.

Un immense merci à mon immense famille. Merci de m'avoir soutenue dans tous les sens du terme dans mes études éternelles. Merci de m'avoir permis d'être libre de mes choix et de les avoir même encouragés. Merci de votre amour inconditionnel et quotidien. Papa, Maman, Hélène, Emmanuelle, Bruno, Cedryc, Romane, Annie, Bernard, Sylvain, Lara, Aïnoa, Maïa, Sébastien, Linda, Samantha, Colleen, Jean Jacques, Anne, Agnès, Samuel, Fleur, Augustine, et Léopold, j'ai de la chance de vous avoir.

A toi Mamie, tu es une femme libre, tolérante, et intelligente, mais aussi une tête dure, une obstinée, une volontaire, merci d'être mon modèle et ma source d'inspiration.

Enfin, pour citer Torlois 2016 « last but not least, idole urbi et orbi des jeunes, des moins jeunes et des animaux de basse cour, parangon de vertu, Porteur de la joie de vivre et du sourire le plus lumineux qu'il puisse être, roi des montagnes et des plaines, des terriers et des collines, dieu des océans et des flaques, air que l'humanité respire, soleil qui nous illumine et nous apporte la joie, la vitamine D et les Ray Ban miroir, grand mamamouchi de la magie expérimentale, archange du moonwalk et surtout, surtout, divinité footballistique », à toi Stéphane, mon amoureux qui est aussi mon mari, merci de me rappeler chaque jour que la vie est une fête.

Conventions d'écriture

Abréviations

Adj= nom adjectivé

Adv = adverbe

N = nom

Vbe = verbe

M1 = terme occupant la position à gauche du coordonnant « et »

M2 = terme occupant la position à droite du coordonnant « et »

M1&M2 : symbolise tout composé coordinatif syndétique par « et »

σ : mot de une syllabe / σσ : mot de deux syllabes / σσσ : mot de trois syllabes etc

C : consonne

V : voyelle

G : glide

V(C)(C).σ# : mot de deux syllabes commençant par une voyelle. V(C)(C).σσ# : mot de trois syllabes commençant par une voyelle

#_ : commence par.

_# : finit par.

#V_ : symbolise une initiale vocalique

#Vorable_ : symbolise une initiale vocalique orale

#Vnasale_ : symbolise une initiale vocalique nasale

#[+ NASAL]- / #[-CONTINU]_ / #[+ VOISE]_ / etc : symbolisent une initiale [+ NASAL] / [-CONTINU] / [+ VOISE] etc

#C_ : symbolise une initiale consonantique.

#_V_ : symbolise la voyelle de la syllabe initiale quand cette dernière ne débute pas par une voyelle.

#_[-NASAL]_ / #_[+ HAUT]_ / #_[+ ARRIÈRE]_ / etc : symbolisent la voyelle de la syllabe initiale [-NASAL] / [+ HAUT] / [+ ARRIÈRE] etc

_V# : symbolise une finale vocalique

_C# : symbolise une finale consonantique.

Binôme : ensemble coordonné par « et » composé de deux éléments « x » et « y », indifféremment de leur position dans la coordonnée

Paire : ordre préférentiel de coordination. Il peut être « x & y » ou « y & x ».

Appariement : ensemble des paires possibles à savoir « x & y » et « y & x ».

Signes conventionnels

MAJUSCULE ET MAJUSCULE = support d'études

un point après V ou au milieu d'une suite de C et V (et accessoirement G) majuscules : indique une frontière syllabique (ex. : CVC.CV)

un mot en petites majuscules entre crochets : indique un trait conceptuel (ex. : [MALE])

un ou plusieurs symboles phonétiques entre deux barres obliques : indique une transcription phonologique (ex. : /t ε z/)

plusieurs chiffres entre deux barres obliques : indique le schéma accentuel du mot (ex. : /102/)

1 : symbolise une syllabe recevant un accent primaire

2 : symbolise une syllabe recevant un accent secondaire

0 : symbolise une syllabe inaccentuée

Introduction

Ce travail de thèse a pour objet l'étude des principes gouvernant la coordination par « et » de deux mots en français. Il a pour but d'observer l'émergence de facteurs actifs dans la sélection d'un ordre préférentiel de coordination binaire directe par « et », hors contexte, selon un angle essentiellement phonologique ; et ainsi de tenter d'en proposer une pré-hiérarchisation en français.

Dans cette étude, nous appréhendons les principes qui déterminent l'ordre des éléments dans des binômes du type «cul et chemise», «modes et travaux» ou «Émilie et Nathan». Il apparaît, en effet, que pour tous ces exemples, l'ordre adopté ici est l'ordre préférentiel en français, et ceci que l'on se base sur le jugement intuitif des locuteurs, ou qu'on l'atteste au moyen de relevés statistiques assis. Il est clair que les locuteurs natifs privilégient souvent spontanément un ordre à un autre dans ce type de formations, et/ou qu'ils jugent plus naturelle que l'autre l'une des deux combinaisons possibles.

Partant de ce constat, la question qui se pose est de savoir quels sont les facteurs qui régissent l'ordre de ces constituants en français. Peut-on assigner la préférence intuitive d'une séquence binomiale coordonnée à des facteurs morphologiques, syntaxiques, sémantiques, culturels, sociologiques, ou autres ? Et qu'en est-il des contraintes phonologiques jusqu'ici largement négligées en français ?

Pour répondre à ces questionnements, nous avons posé deux postulats initiaux :

- Travaillant sur des faits de langue, nous ne pouvons nous restreindre au seul point de vue phonologique. C'est pourquoi, pour examiner la coordination binomiale directe par « et », nous nous devons de proposer une exploration large de ce phénomène en rendant compte entre autres de quelques facteurs extraphonologiques ;
- Cependant, souhaitant néanmoins axer principalement nos travaux sur le rôle des principes phonologiques actifs dans la sélection de composés syndétiques par « et » en français, nous avons mis en place des stratégies afin d'éviter autant que faire se peut ces biais extraphonologiques.

C'est prise dans cette tension *a priori* contradictoire que nous avons construit cette étude.

Nous avons donc choisi de n'examiner que des coordinations binaires directes par « et » hors contexte, car d'une part, les « binômes » comme définis par Malkiel (1959) nous affranchissent de biais discursifs et syntaxiques. D'autre part, il nous a semblé qu'une expression figée telle que « par monts et par vaux » ne peut être étudiée de la même façon que « cul et chemise ». En effet, comment considérer une telle séquence ? Comptabilisons-nous [p a r m ɔ̃ {z}] et [p a r v o] comme des monosyllabes en n'observant que les noms « monts » et « vaux » ou comme des dissyllabes en considérant « par monts » et « par vaux » comme une seule séquence phonologique ? Pour pallier ces incertitudes, nous avons préféré n'étudier que des formes directes coordonnées par « et ».

L'autre argument ayant appuyé plus encore le choix de l'examen de composés binaires directs par « et » est le suivant : comme le montre l'étude de Labbé (2007), à l'oral, « la grande majorité de(s) (...) emplois (de « et ») sont du type "nom et nom" ou "nom adjectif et adjectif" ». Or, même si nous ne travaillons pas directement sur des réalisations orales de binômes coordonnées par « et », il est entendu que nous souhaitons rendre compte d'un phénomène actif dans des actualisations réelles du français contemporain.

Enfin, afin de mettre au jour les principes actifs dans la sélection des formes privilégiées des composés syndétiques, nous avons mis en place un travail guidé par une approche empirique et expérimentale : empirique, du fait de la construction des supports d'étude nous ayant permis de dégager des données, et expérimentale, compte tenue de la vérification systématique de nos données et de nos hypothèses grâce à des tests d'ajustement.

Nous avons donc, dans un premier temps, rassemblé des données statistiquement valides, à partir desquelles nous nous proposons de dégager des principes généraux en vue d'une pré-hiérarchisation formelle dans le cadre d'une approche en termes d'interactions de contraintes.

Organisation de la thèse

La présente thèse se compose de trois parties principales. La première partie aborde le cadre théorique de cette étude, en présentant quelques notions utilisées dans ce travail et en discutant rapidement des différents travaux phonologiques ayant étudié la sélection d'un ordre préférentiel de coordination directe par « et » dans la littérature. La deuxième partie présente la mise en œuvre du recueil et de l'exploitation des données examinées dans nos travaux. Enfin, la troisième partie regroupe les résultats et la discussion finale.

Dans le détail, la première partie débute au chapitre 1. Nous y proposons tout d'abord un panorama de quelques études examinant l'ordre de la coordination binominale binaire en anglais - l'examen phonologique de ce phénomène étant jusqu'alors très peu étudié en français. Puis nous définissons des notions théoriques essentielles à nos travaux et l'utilisation particulière que nous en ferons dans la présente thèse.

La deuxième partie est présentée chapitre 2. Il s'agit d'un chapitre essentiellement méthodologique dans lequel nous rendons compte du recueil et de l'exploitation de nos données à travers la création d'un ensemble composite de supports d'études. Nous y présentons premièrement la mise en place de six corpus, puis celle de six questionnaires d'enquête, et enfin nous exposons les critères classificatoires que nous avons utilisés pour étiqueter et exploiter nos données.

Enfin la dernière partie recouvre les chapitres 3 et 4. Ces chapitres présentent les résultats émergeant de nos supports d'étude et les analyses que nous en proposons.

Dans le chapitre 3, nous exposons les facteurs extraphonologiques dont nous supposons qu'ils ont un rôle actif dans la sélection d'une séquence binominale coordonnée.

Le chapitre 4, quant à lui, détaille les facteurs phonologiques qui pourraient gouverner la coordination par « et » de deux mots en français. Nous l'avons organisé en trois grands pôles.

Le premier rend compte de l'hypothèse de la corrélation entre la taille des mots et la position dans la coordonnée, les deuxième et troisième pôles examinent le rôle des frontières de mots, que ce soit en termes de représentations phonémiques, de traits distinctifs ou de structures de syllabes. Nous y étudions également le rôle de la liaison dans la sélection d'un ordre préférentiel de coordination.

Chapitre 1 : Remarques générales

S'il existe de très nombreux travaux sur la coordination par « et » en français, qu'il s'agisse de ses aspects syntaxique et sémantique (cf. Abeillé (2005), (Lambert (1997), (2001), (2006), Hobaek Haff (1990), ou Labbé (2007) entre autres) ou de ses aspects prosodiques (Mouret, Abeillé, Delais-Roussarie, Marandin, Yoo (2008)), nous n'avons néanmoins pas connaissance d'étude phonologique de la coordination binaire simple par « et » de deux lexèmes pleins en français. En revanche, les recherches sur les constructions binomiales sont légion notamment pour l'anglais et l'allemand. Dans la section à venir, nous proposons un tour d'horizon de différentes études sur la coordination binomiale en anglais ayant initiées notre travail de thèse. Puis, nous nous attacherons à définir quelques notions théoriques essentielles à nos travaux.

1.1. Panorama de la littérature anglo-saxonne

Un certain nombre de chercheurs se sont intéressés à la problématique de la formation binomiale, se concentrant essentiellement sur les constructions de l'anglais et de l'allemand. Depuis Malkiel (1959) qui présentait un aperçu des binômes figés de l'anglais et des principes sémantiques et phonologiques actifs dans leur processus de figement dont la préséance en M1 au plus fort des deux traits conceptuels polarisés et les priorités propres aux structures d'une société jusqu'à Gutsafsson (1976) dont l'étude de corpus de binômes non figés avançait un principe de fréquence confirmé par Fenk-Oczlon (1989) qui concluait sur l'existence d'une corrélation forte entre la fréquence et un ordre binomial privilégié par l'étude de binômes figés, l'ordre binomial des mots est un sujet récurrent.

S'inscrivant dans la continuité de leurs prédécesseurs, des études plus récentes ont émergé. Nous notons spécifiquement celle de Benor et Levy (2006) qui, profitant de l'accessibilité de corpus plus volumineux, proposent quant à eux, un traitement statistique de vingt contraintes relevées dans la littérature spécialisée. Leurs travaux tendent à prouver que les contraintes phonologiques ne contribuent pas aussi significativement à la sélection d'un ordre préférentiel des binômes figés ou non ; que les contraintes sémantiques, métriques ou celles liées à la fréquence.

Enfin, nous relevons particulièrement les travaux de Cooper et Ross (1975), ceux de Pinker et Birdsong (1979), de Wright, Hay et Bent (2002, 2005) et Renner (2014) que nous allons

présenter en détail *infra*.

1.1.1. Cooper et Ross (1975)

L'analyse de Cooper et Ross, à travers l'examen des éléments nominaux placés en première position de la coordination, ce que nous nommons M1 (avant le coordonnant) et de ceux placés après (M2), fait apparaître des facteurs agissant dans la sélection d'un ordre privilégié de coordination en anglais, dont une liste de vingt-deux concepts sémantiques et une première liste de sept facteurs phonologiques différenciant M1 de M2.

Concernant la liste des vingt-deux concepts sémantiques, nous l'exposons telle quelle ; les exemples proposés sont les deux premiers présentés par Cooper et Ross :

1. « [HERE] : *here and there, this and that.* » L'élément spatial le plus proche du locuteur est placé en M1 ;
2. « [NOW] : *now and then, sooner or later.* » L'élément temporel le plus proche du locuteur est placé en M1 ;
3. « [PRESENT GENERATION] : *father and grandfather, son and grandson.* » L'élément le plus contemporain est placé en M1 ;
4. « [ADULT] : *man and boy, men, women and child.* » L'élément le plus adulte est en M1 ;
5. « [MALE] : *man and woman, husband and wife.* » L'élément masculin est en M1 ;
6. « [POSITIVE] : *positive or negative, plus or minus.* » L'élément positif occupe M1 ;
7. « [SINGULAR] : *singular and plural, Mick Jagger and the Rolling Stones.* » L'élément singulier est en M1 ;
8. « [PATRIOTIC] : *cowboys and Indians, United States and Canada.* » L'élément le plus patriotique est en M1 ;
9. « [ANIMATE] : *people and things, person, place or thing.* » L'élément animé occupe M1 ;
10. « [FRIENDLY] : *friend or foe, pro- or anti-labor.* » L'élément le plus amical sélectionne M1 ;
11. « [SOLID] : *land and sea, Army and Navy,* » L'élément le plus solide est M1 ;
12. « [FRONT] : *front and back, front and rear.* » L'élément le plus « devant » est en M1 ;
13. « [AGENTIVE] : *agent and patient, speaker and hearer.* » L'agent sélectionne M1 ;

14. « [POWER SOURCE] : *bow and arrow, sun and moon.* » La source d'énergie occupe M1 ;
15. « [LIVING] : *living or dead, the quick and the dead.* » L'élément vivant est en M1 ;
16. « [AT HOME] : *Aeronautics and Astronautics, Earth and Planetary Science.* » L'élément local sélectionne M1 ;
17. « [GENERAL] : *form and substance, general and particular.* » L'élément le plus général est en M1 ;
18. « [NOMINAL] : *nouns and verbs.* » L'élément nominal est en M1 ;
19. « [COUNT] : *count and mass nouns.* » L'élément comptable favorise M1 ;
20. « [FOOD AND DRINK HIERARCHY] : [FISH] > [MEAT] > [DRINK] > [FRUIT] > [VEGETABLES] > [BAKED GOODS] > [DAIRY PRODUCTS] > [SPICES] : *fish and game, meat and drink.* » Il existe une échelle d'occupation de M1 liée au type d'aliment ou de boisson ;
21. « [DIVINE] : *God and man, church and state.* » Le divin est en M1.
22. « [PLANT] : *plant and animal, flora and fauna.* » L'élément végétal sélectionne en M1.

Notons également le [ME FIRST], concept subsumant les facteurs 1 à 20 et défini par Cooper et Ross comme « First conjuncts refer to those properties which describe the prototypical speaker (whom we will sometimes refer to as « Me »).

Une autre analyse visant à examiner ce que ces traits ont en commun a également mené les deux chercheurs à s'interroger sur l'opposition marqué / non marqué. Cette opposition a notamment été reprise, entre autres, par Benor et Levy (2006) que nous citons précédemment. En effet, il apparaîtrait que cette répartition soit pertinente du fait de la tendance des éléments marqués à se produire dans des contextes marqués, et, inversement, ce qui suggère que l'élément le moins marqué sélectionnerait en priorité M1. Nous observerons cette tendance ultérieurement.

Cette liste de vingt-deux critères sémantiques, bien qu'apparemment massivement validée par Cooper et Ross, ne sera pas retenue telle quelle dans le présent travail. En effet, non seulement, elle souffre d'exceptions, mais en plus certains concepts sont difficilement appréhendables. De plus comme le signalait Renner (2006), « plusieurs concepts ont en outre une portée limitée (...) tandis que d'autres semblent pouvoir être facilement rapprochés ».

Pour toutes ces raisons, nous préférons nous référer à celle proposée par Renner que nous détaillerons page 19.

Parallèlement à la précédente liste de concepts sémantiques actifs dans la composition binomiale, Cooper et Ross relevaient sept facteurs phonologiques différenciant M1 de M2.

Ainsi M2 devrait contenir :

- a) Plus de syllabes ;
- b) Une voyelle plus longue ;
- c) Une initiale plus consonantique ;
- d) Quand les deux lexèmes ont une initiale consonantique, l'initiale de M2 est plus obstruante ;
- e) M2 contient également une voyelle ayant une fréquence de second formant plus basse (du formant le plus haut au plus bas [i] > [e] > [ɛ] > [y] > [ø] > [œ] > [a] > [ɑ] > [ɔ] > [o] > [u]) ;
- f) M2 possède un plus petit nombre de consonnes finales ;
- g) Et enfin quand les deux lexèmes se terminent par une consonne simple, la consonne finale de M2 est moins obstruante.

Il a été reproché à cette étude la taille limitée de ses corpus, et de l'aveu même de ses auteurs, le fait qu'un certain nombre de leurs exemples étaient soumis concomitamment à plusieurs facteurs sémantiques et phonologiques.

1.1.2. Pinker et Birdsong (1979)

A la suite de l'analyse originale de Cooper et Ross que nous venons de présenter, Pinker et Birdsong (1979) ont testé expérimentalement cinq des sept facteurs phonologiques de Cooper et Ross¹ sur des pseudos-expressions binomiales.

Quatre de ces facteurs ont été confirmés. Ainsi, statistiquement il apparaît que dans les expressions binomiales figées en anglais :

1 Les facteurs expérimentés sont :

- a) M2 a plus de syllabes que M1 ;
- b) M2 a une voyelle plus longue que M1 ;
- d) Quand les deux lexèmes ont une initiale consonantique, celle de M2 est plus obstruante ;
- e) M2 contient une voyelle plus aiguë que M1 ;
- f) M2 contient un plus petit nombre de consonnes finales.

- a) M2 a un plus grand nombre de syllabes que M1 ;
- b) Le lexème ayant la voyelle la plus longue est placé en M2 ;
- d) M2 possède une initiale consonantique plus obstruante que M1 ;
- e) M1 aura tendance à être composé de voyelles plus fermées et/ou plus antérieures que M2 pour qui elles seront de fait plus ouvertes et/ou plus postérieures.

Le facteur f) est, quant à lui, disqualifié par Pinker et Birdsong qui, contrairement à ce que Cooper et Ross avaient énoncé, établissent le fait que les mots en M2 possèdent plus de consonnes finales.

D'une manière générale, comme le signale Renner (2006), il est également possible de regrouper ces facteurs en relevant ce qu'ils partagent en commun. En effet, concernant les facteurs b) et e), Birdsong (1995), en considérant les voyelles d'un point de vue acoustique, autorise le rapprochement de ces deux facteurs : les voyelles plus ouvertes et/ou plus antérieures sont plus longues. De plus, le rappelle Renner, le facteur a) est à rapprocher « d'un principe général appelé « contraste court/long » par Ross (1980) et surtout de la « loi du second lourd » de Hagège (1985:184) qui veut que le constituant structurellement le plus simple précède le constituant le plus complexe ».

1.1.3. Wright, Hay et Bent (2002, 2005)

Plus récemment, les recherches de Wright, Hay et Bent (2002, 2005) portent sur la coordination de deux prénoms mixtes en anglais tels que «Romeo and Juliet», «Rhett and Scarlett», «Mickey and Minnie» ou «Fred and Wilma». Grâce à l'étude d'un corpus et à la mise en place de deux expériences dans lesquelles ont été soumis des appariements de prénoms composés d'un prénom masculin et d'un prénom féminin en contexte à une trentaine d'interviewés, elles tendent à prouver que de nombreux facteurs – dont le genre, la phonologie mais aussi la fréquence des prénoms étudiés – jouent un rôle important dans le processus de bonne formation d'items coordonnés.

Ainsi sont actifs dans la coordination de deux prénoms mixtes en anglais :

- a) Les prénoms masculins sélectionnent M1 ;
- b) Les prénoms masculins sont phonologiquement caractérisés pour M1 ;
- c) Les prénoms les plus fréquents occupent M1 ;
- d) Les prénoms monosyllabiques sont plutôt en M1 ;

- e) Les prénoms ayant une initiale plus sonante sont plus en M1 que ceux ayant une initiale occlusive ;
- f) M1 a des finales plus obstruantes que M2 ;
- g) M2 a une finale plus sourde que M1 ;
- h) Les prénoms possédant des voyelles hautes et antérieures sont plutôt en M1.

Wright, Hay et Bent, proposent également l'organisation des contraintes comme suit :

genre > sonorité de la finale > sonorité de l'initiale > voisement de la finale

Cette hiérarchie de contraintes signifie que la sélection de la position M1 ou M2 dans la coordonnée est déterminée d'abord par le genre du mot, puis par la sonorité de la finale, de l'initiale, et enfin par le voisement de la finale.

D'une manière générale, d'un point de vue phonologique, la structure rythmique, la resyllabation, le renforcement du domaine initial, l'amuïssement de la finale et l'allongement des voyelles sont autant de facteurs actifs dans la coordination de deux prénoms en anglais en faveur de M1 ou de M2.

Notre travail s'inspire de Wright *et al.* (2002, 2005) à plusieurs titres :

- Dans la lignée de leurs recherches, nous avons initié la présente étude par l'examen de deux prénoms mixtes. Travailler sur des prénoms nous a permis de nous affranchir un temps d'un maximum de contraintes syntaxiques ou sémantiques ;
- Nous avons également appuyé nos travaux sur un recueil de données hétéroclites. Nos supports d'études sont soit des corpus, soit des expérimentations ;
- Comme elles, nous avons systématiquement validé statistiquement nos résultats ;
- Enfin, convaincue que la sélection d'ordre préférentiel de séquences binomiales coordonnées peut être soumise au jeu d'interaction de contraintes, nous proposons aussi des pré-hiérarchies de contraintes

1.1.4. Renner (2006, 2014)

De récents travaux de Renner (2014) proposent un traitement statistique de douze contraintes phonologiques et extraphonologiques agissant potentiellement sur la parataxe syndétique en anglais (la coordination de binômes), mais aussi sur la parataxe asyndétique (la composition binomiale) et sur les mots-valises ou amalgames. De cette dernière étude, nous retenons

surtout les facteurs opérant sur la coordination binomiale simple en anglais.

Ainsi, opérant sur l'ordre préférentiel des mots en anglais, Renner relève :

- a) Le facteur de [PROXIMITÉ] que nous présenterons ci-après ; ce qui est proche de notre expérience empirique privilégie M1 ;
- b) Le facteur d'[ANTÉRIORITÉ TEMPORELLE] prédit que les éléments faisant référence à un événement ou une action antérieurs occuperont prioritairement M1 ;
- c) Le nombre de syllabes : statistiquement M2 a plus de syllabes que M1 ;
- d) Les initiales consonantiques sont plus complexes en M2 ;
- e) Les initiales consonantiques sont plus obstruantes en M2 ;
- f) Les voyelles sont plus longues en M2 ;
- g) Les finales consonantiques sont plus obstruantes pour M1 dans le cas de coordination équisyllabique dans lesquelles les lexèmes ont des codas simples ;
- h) M1 et M2 sont ordonnés pour respecter l'alternance entre syllabes accentuées et non accentuées ;
- i) Le mot le plus fréquent sélectionne M1.

Selon Renner, le facteur de fréquence lexicale est à rapprocher du facteur c), à savoir le nombre de syllabes. En effet, statistiquement dans des coordinations équisyllabiques, le rôle de la fréquence n'est pas significatif. La fréquence lexicale serait en lien avec la loi de Panini que mentionnent Cooper et Ross (1975:71) ou le « Principe poétique de gradation syllabique » de Jakobson (1963:221) : les mots les plus fréquents tendent à être les plus courts et donc, plus un lexème est fréquent, et court, plus il sera placé en M1.

Revenons à présent sur les facteurs sémantiques proposés par Renner (2006, 2014). Comme nous l'indiquions précédemment, la liste des vingt-deux traits conceptuels proposée par Cooper et Ross nous semblait difficilement accessible. Nous avons donc choisi d'utiliser pour le traitement sémantique de nos supports d'études les six (ou trois facteurs sémantiques) ci-dessous :

1. [PROXIMITÉ SPATIALE] : *ici et là*
2. [PROXIMITÉ TEMPORELLE] : *aujourd'hui et demain*
3. [PROXIMITÉ PSYCHOLOGIQUE] : *bien et mal, mâle et femelle, hommes et machines*
4. [SUPÉRIORITÉ SPATIALE] : *haut et bas*

5. [SUPÉRIORITÉ HIÉRARCHIQUE] : *seigneur et vassal, chef et salarié*

6. [ANTÉRIORITÉ TEMPORELLE] : *causes et effets*

L'avantage premier de ces six concepts généraux que Renner (2014) a unifié en trois ensembles conceptuels ([PROXIMITÉ], [SUPÉRIORITÉ] et [ANTÉRIORITÉ TEMPORELLE]), est leur organisation intuitive.

Le second avantage renvoie à l'opposition marqué / non marqué que nous mentionnions précédemment. Comme évoqué par Cooper et Ross concernant les traits [PROXIMITÉ] subsumant les traits 1 à 20 présentés page 15, si l'élément le plus proche est en M1, il est aussi le plus prototypique, le moins marqué.

1.1.5. Récapitulatif des facteurs relevés pour les composés binomiaux en anglais

Pour finir notre tour d'horizon de la littérature spécialisée au sujet des composés binomiaux en anglais, nous proposons, pour plus de lisibilité, un tableau récapitulatif synthétique des principaux facteurs relevés.

	Cooper et Ross	Pinker et Birdsong	Wright, Hay et Bent	Renner	Autres auteurs, pour l'anglais
Facteur sémantique	22 concepts sémantiques dominants M1 = - marqué M1 = [ME FIRST]			M1 = [PROXIMITÉ] M1 = [ANTÉRIORITÉ TEMPORELLE]	

Facteur de fréquence lexicale			M1 = plus fréquent que M2	Peu significatif pour binômes coordonnés équisyllabiques	Gustafsson (1976), Fenc- Ozelon (1989) et Benor et Levy (2006) : M1 = plus fréquent que M2
Facteur phonologique	M2 = V + longue M2 = #C_ Quand M1 et M2 =#C_, M2 = #C_ + obstruante. M2 = V a une fréquence de second formant plus basse M2 = _ de _C#. Quand M1 et M2 =_C# et coda et coda simple, M2 = _C# - obstruante.	M2 = V + longue M2 = #C_ + obstruante M2 = V + ouverte, + postérieure M2 = + de _C#.	M1 : #_ + sonante > #_ + occlusive M1 = _# + obstruante M2 = _# + sourde M1 = V + fermée, + antérieure	M2 = #C_ plus complexe M2 = #C_ + obstruante M2 = V + longue M1 = _C# + obstruante si longueur M1= longueur M2 et si M1 et M2 ont des codas simples.	
Facteur de genre			M1 = Prénom masculin		

Facteur métrique	M2 = + de syllabes	M2 = + de syllabes	M1 = monosyllabe	M2 a plus de syllabes que M1 M1 et M2 respectent l'alternance entre syllabes accentuées et non accentuées.	Cité par Renner : Jespersen (1905) ² Les séquences « monosyllabes et dissyllabes » sont favorisées car elles permettent alternance accentuelle trochaïque.
Facteur de prototypicalité					Cité par Renner : Kelly, Bock et Keil (1986) ³ : M1 est plus prototypique que M2.

Tableau 1 : Récapitulatif des facteurs relevés pour les composés binomiaux en anglais

L'ensemble des facteurs présentés ici sont à l'origine des critères classificatoires que nous présenterons page 71. Tous sont autant d'hypothèses que nous avons cherché à développer dans le présent travail de thèse.

1.2. Cadre théorique

L'ordre des binômes est un phénomène non catégorique induisant des conflits de contraintes ou de facteurs.

En effet, pour les paires « cul et chemise », « modes et travaux » ou « Émilie et Nathan », en adéquation avec les hypothèses relevées dans la littérature spécialisée, nous pouvons supposer

2 En lien avec la loi de Panini, « Jespersen expliqu(e) dans sa première édition de *Growth and Structure of the English Language* en 1905 que la séquence « monosyllabe-coordonnant-dissyllabe » est privilégiée, car elle permet une alternance accentuelle de type trochaïque »

3 « Kelly, Bock et Keil (1986) démontrent expérimentalement qu'à fréquence et à nombre de syllabes égaux, la différence de typicalité entre les éléments d'un binôme formé en discours influe sur leur ordre linéaire : l'élément prototypique précède généralement l'élément non prototypique »

le poids de plusieurs facteurs :

- La taille des mots : « cul » et « modes » sont des monosyllabes, « chemise et travaux » des dissyllabes (Nous développerons ce point page 96.) ;
- « Cul » « Émilie », tous deux placés en M1, se finissent par une voyelle, tout comme « Nathan » et « travaux » (Ce facteur sera détaillé page 196.) ;
- « Modes » se terminant par une consonne est placé en M1 (cf. page 219) ;
- « Nathan » et « travaux » finissent tous deux par une voyelle non haute (« grave ») (Ce critère est examiné page 207.) ;
- Et « Émilie » débutant par une voyelle occupe M1. (Nous étudierons le rôle des initiales vocaliques page 150.)

Au regard de ces trois exemples, il est tout à fait envisageable que plusieurs principes commandent les mots composant ces formes coordonnées en M1 ou en M2.

De fait, afin d'exploiter au mieux nos résultats, il nous est donc nécessaire de proposer une analyse dans le cadre d'une approche en termes d'interactions de contraintes.

Pour cela, nous avons choisi de nous inscrire à la suite des travaux de Plénat (1996, 1997) (dans ses travaux sur la morpho-phonologie du français, et notamment ses études portant sur les dérivés en -esque). En effet, le cadre théorique proposé par Plénat met en œuvre des modèles par contraintes comme dans la Théorie de l'optimalité, notée à présent OT (Prince et Smolensky (1993)). Cependant, à la différence de OT, le cadre théorique choisi pour le présent travail nous offre la possibilité d'une évaluation hiérarchique et cumulative sans nous contraindre à formaliser strictement nos hiérarchie de principes.

Ce modèle nous permet ainsi d'appréhender plus librement la sélection d'un ordre binomial préférentiel de coordination par « et » de deux noms propres, communs, substantivés ou adjectivés, comme le résultat de conflit entre de contraintes ou de facteurs.

Dans la section suivante, nous délimitons un peu plus notre cadre théorique en définissant certaines notions et l'utilisation que nous en faisons dans cette étude.

1.2.1. Transcription du corpus et décompte syllabique

Nous nous sommes appuyée sur différentes sources afin d'étiqueter nos données selon la variation du français dit « standard » de façon cohérente et homogène pour l'ensemble de nos supports d'études.

Nous avons donc fait le choix de nous conformer principalement à la transcription en API proposée par le TLF⁴, notamment pour les noms communs, adjectifs, adverbes et verbes non tensés.

En effet, le TLF⁴ offre la possibilité de se référer à une transcription standardisée, consensuelle, facilement accessible et relativement récente (la seizième édition date de 1994 et les dernières corrections de 2001 et 2004).

Concernant les réalisations phonologiques des noms propres, ne pouvant consulter la même source à ce sujet, nous avons mis en place deux stratégies différentes.

Pour le corpus TARN ET GARONNE, comme nous l'indiquerons ultérieurement, nous avons fait appel à nos connaissances en tant que locutrice native de la Gironde ainsi qu'à celles d'autres locuteurs natifs des départements de la Dordogne et des Landes, en contactant directement les mairies pour la poignée de communes pour lesquelles nous nourrissions des hésitations.

Plus marginalement, essentiellement pour la centaine de prénoms présents dans nos supports d'études, nous avons suivi les prescriptions proposées par Fouché (1956) et Wioland (1991).

Enfin, nous avons choisi de ne pas réaliser les diérèses en ce qui concerne le traitement des glides. Nous élidons systématiquement le schwa final.

Pour exemple, le mot « bien » est un monosyllabe et non un dissyllabe, tout comme « porte » et « Élise », pour lesquels le schwa final est élidé.

1.2.2. Le schwa

Nous détaillons ci-dessous le traitement spécifique que nous proposons du schwa⁵ dans le cadre de ces travaux.

Le schwa est un problème récurrent de la phonologie du français, au même titre que la liaison que nous examinerons ultérieurement page 274. De fait, un très grand nombre d'études balaient ce phénomène. Nous citerons ici très sommairement Durand (2014), Eychenne

4 Le TLF⁴ ou Trésor de langue française informatisé est la version informatisée en accès libre des 16 volumes (plus un supplément) du TLF (Trésor de la langue française)

5 Dans la lignée de Durand (2014), nous choisissons ce terme de schwa comme préthéorique.

(2006), Scheer (2000), Tranel (1987, 2000), Dell (1985), Martinet (1969), Schane (1968) et plus largement le projet PFC (Phonologie du français contemporain).

De façon très schématique, les théories de la distribution du schwa sont divisées en deux positions opposées :

- Dans la première, le schwa est sous-jacent. Il est éliminé ou non ;
- Dans la seconde, il est épenthétique.

Ces deux considérations extrêmes sont à l'initiative de nombreuses approches théoriques, qu'il s'agisse entre autres, d'analyse générative (décrivant la distribution du schwa par un ensemble de règles ordonnées, Dell (1985), d'analyse syllabique (proposant que tout segment soit intégré dans une syllabe bien formée, Tranel (1987, 2000) ou encore d'analyse de sonorité et de perception (Côté (2000)⁶).

Consciente des très nombreux questionnements que pose la distribution du schwa, nous avons pris des décisions méthodologiques que nous exposons ici.

Ainsi, nous avons éliminé l'ensemble des schwas présents en finale de mots. Et inversement, nous les avons conservés lorsqu'ils s'inscrivent dans la règle dite « règle des trois consonnes » (Grammont (1914), à savoir que le schwa est maintenu pour éviter la formation d'un groupe consonantique de trois consonnes consécutives). Pour exemple, le mot « règlement » /R ɛ ŋ l ə m ɑ̃/ est un mot trisyllabique réalisant systématiquement le schwa afin d'éviter le cluster /g l m/⁷.

Nous pensons utile de signaler qu'aucun prénom du corpus TOM ET ZOÉ ou de ceux présents dans nos questionnaires d'enquête ne possède de schwa interne.

Dans l'ensemble des données de nos supports d'études, seuls les 34 termes que nous présentons ci-dessous questionnent ce phénomène. Nous les proposons sous leurs différentes formes, avec et sans réalisation du ou des schwas internes.

6 CÔTÉ MH., (2000), *Consonant cluster phonotactics: a perceptual approach*. Thèse doctorale, Massachusetts Institute of Technology.

7 Dans les faits, la règle des trois consonnes n'est appliquée que sur cinq mots du corpus : « pamplemousse », « entreprise », « remboursement » « bretelles » et « règlement ». Pour ces cinq termes, nous considérons que le schwa interne est systématiquement réalisé.

CORPUS	_/ə/_	Réalisations phonétiques
CUL ET CHEMISE	<ul style="list-style-type: none"> chemise fenêtres retour passeries venir venues 	<ul style="list-style-type: none"> [ʃmi:z] / [ʃəmi:z] [fənɛtʁ] / [fnɛtʁ] [ʁətu:ʁ] / [ʁtu:ʁ] [pasəʁi] / [pasʁi] [vəni:ʁ] / [vni:ʁ] [vny] / [vəny]
LISSE ET SOYEUX	<ul style="list-style-type: none"> cheveux pamplemousse reconstruction 	<ul style="list-style-type: none"> [ʃəvø] / [ʃvø] [pɑ̃pləmʊs] [ʁəkɔ̃stʁyksjɔ̃] / [ʁkɔ̃stʁyksjɔ̃]
ARTS ET MÉTIERS	<ul style="list-style-type: none"> aménagement citoyenneté demeure devoir entreprise promenade remboursement 	<ul style="list-style-type: none"> [amenaʒmɑ̃] / [amenaʒəmɑ̃] [sitwajɛntɛ] / [sitwajɛnɛtɛ] [dəmœ:ʁ] / [dmœ:ʁ] [dəvwa:ʁ] / [dvwa:ʁ] [ɑ̃tʁəpʁi:z] [promnad] / [promənad] [ʁɑ̃bʊrsəmɑ̃]
VITE ET BIEN	<ul style="list-style-type: none"> avenir bretelle cheveux demande développement enseignement femelle maintenance 	<ul style="list-style-type: none"> [avəni:ʁ] / [avn̩i:ʁ] [brətɛl] [ʃəvø] / [ʃvø] [dəmɑ̃:d] / [dmɑ̃:d] [devlɔpmɑ̃] / [devəlɔpəmɑ̃] / [devəlɔpmɑ̃] / [devlɔpəmɑ̃] / [ɑ̃sɛnəmɑ̃] / [ɑ̃sɛnmɑ̃] [fəmɛl] / [fmɛl] [mɛ̃tənɑ̃:s] / [mɛ̃tnɑ̃:s]

	<ul style="list-style-type: none"> • melon • menace • passerelles • petites • petits • recrutements • règlement • reproduction • semailles • vêtement 	<ul style="list-style-type: none"> • [məlɔ̃] / [mlɔ̃] • [mənɑs] / [mnɑs] • [pasʁɛl] / [pasəʁɛl] • [ptit] / [pətit] • [pti] / [pəti] • [ʁəkʁytəmɑ̃] / [ʁəkʁytmɑ̃] / [ʁkʁytmɑ̃] • [ʁɛgləmɑ̃] • [ʁəpʁɔdyksjɔ̃] / [ʁpʁɔdyksjɔ̃] • [səmɑj] / [smɑj] • [vɛtmɑ̃] / [vɛtəmɑ̃]
--	---	---

Tableau 2 : Mots possédant un schwa interne dans nos corpus

Les différentes formes que nous venons de relever font émerger une question : comment comptabiliser dans nos supports d'étude les mots comprenant un schwa interne ?

D'autant plus que les phénomènes d'actualisation du schwa interne posent un problème de prévisibilité. Dans un même contexte, le schwa peut être réalisé ou non.

S'il est possible de prédire sa réalisation en frontière de morphèmes, à l'intérieur des morphèmes, il est sous-jacent. Si la nature et le nombre de consonnes environnantes peuvent nous renseigner, d'autres facteurs, qu'ils soient morphologiques (devant un suffixe dérivationnel ou un suffixe flexionnel), prosodiques (en fonction du nombre de syllabes encadrant le schwa), dialectaux ou liés ou au registre ou au débit de langue motivent également les réalisations ou non du schwa.

En fait, au-delà même des considérations sur les variations du français, les réalisations du schwa interne relevant d'un phénomène extrêmement complexe, nous préférons exclure les 38 paires⁸ dans lesquelles au moins un élément contient un schwa interne de l'examen des deux champs d'explorations suivants :

⁸ Il y a 38 paires dans lesquelles un des deux mots du binôme possède un schwa interne, et 34 mots possédant un schwa interne. Certains termes comme « cheveux » sont présents dans plusieurs paires.

- Le nombre de syllabes (présentée page 96) ;
- Et le timbre de la syllabe finale (page 245).

1.2.3. Organisation par traits

Dans cette section, nous définirons les traits distinctifs utilisés dans le présent travail de thèse. L'enjeu de notre étude est de mettre au jour les facteurs phonologiques et extraphonologiques gouvernant la sélection d'un ordre préférentiel par « et ». Pour cela, nous avons défini une liste de critères classificatoires qui sont autant d'hypothèses de travail. Parmi ces critères, un certain nombre envisage le rôle des unités phonémiques, que ce soit par timbre de syllabes ou en frontière de mots.

Or, comme nous le présenterons régulièrement, l'examen des répartitions phonémiques en M1 et en M2 nous conduit à nous interroger sur la mise en relation contrastive des phonèmes entre eux. Comme le rappelle Saussure (1916) dans le *Cours de linguistique générale* : « Dans la langue, il n'y a que des différences. » Le système des signes linguistiques se construit par opposition les uns aux autres. L'opposition forme le système. En d'autres termes, nous nous sommes attachée à étudier ce que ces phonèmes ont en commun et ce sur quoi ils divergent. C'est pourquoi, afin de rendre compte de ces oppositions, nous nous sommes référée aux plus petites unités linguistiques contrastives du français, à savoir les traits distinctifs. Dans la lignée de Dell (1985), nous avons choisi l'utilisation de traits binaires.

A cela, plusieurs raisons que nous présentons rapidement ci-dessous :

- Premièrement, les traits binaires marquent une opposition contrastive entre deux valeurs + / -. Tout couple de phonèmes se distingue par au moins la valeur d'un trait ;
- les faisceaux de traits peuvent également rendre compte des niveaux articulatoires et acoustiques ;
- Et surtout, en ce qui concerne nos travaux, à la différence des traits unaires, même si mathématiquement, les traits binaires sont extrêmement redondants, ils permettent de représenter tous les phonèmes avec le même faisceau de traits. Les traits binaires spécifient l'ensemble des classes naturelles par opposition contrastive.

Ce dernier point est déterminant pour nous. En effet, la mise en valeur des facteurs actifs dans la coordination binomiale par « et » passe par une opposition contrastive de l'ensemble des éléments entre eux. Nous ne pouvons comparer que ce qui est comparable.

Il est donc essentiel que nous puissions spécifier l'ensemble des phonèmes du français par l'utilisation des mêmes traits et des mêmes oppositions.

Enfin, le dernier avantage, selon nous, de l'utilisation des traits binaires est que les valeurs non-marquées des traits sont généralement mentionnées par la valeur « - » précédant le trait - exception faite du trait [ANTÉRIEUR] -. Nous manipulons cette généralisation avec prudence, consciente que les traits n'ont qu'une valeur relative (Chomsky et Halle (1968)). Dans les faits, nous utiliserons l'opposition marqué / non marqué dans sa valeur universaliste (Jakobson *et al.* (1952)) et statistique (Clements (2005)). Cependant, cette dichotomie apparemment systématique entre la valeur « + », marquée, et la valeur « - », non marquée, nous a permis de rendre compte d'un certain nombre de phénomènes.

Pour résumer, nous adopterons pour le présent travail le modèle de Dell (1975), même si de nombreux autres modèles ont été proposés depuis. Nous en connaissons les limites. Cependant, la binarité des traits et le fait que chaque phonème soit couvert par le même faisceau de traits, en font une théorie pratique et maniable pour notre relevé statistique. Ce modèle est adapté à notre entreprise.

Nous utiliserons donc la matrice des traits distinctifs que nous proposons dans le tableau à suivre.

Les traits retenus sont :

- Pour les classes majeures :
 - Le trait [± SYLLABIQUE]. Ce trait renvoie à la structure syllabique dans laquelle se trouve le phonème. [+ SYLLABIQUE] désigne un noyau de syllabes, soit en français les voyelles [- SYLLABIQUE], l'ensemble des autres phonèmes ;
 - Le trait [± CONSONANTIQUE]. Ce trait indique l'obstruction du canal vocal. [+ CONSONANTIQUE] désigne une obstruction importante, soit en français les consonnes, [- CONSONANTIQUE], l'ensemble des autres phonèmes.

- Pour le mode d'articulation des consonnes :
 - Le trait [± SONANT]. Ce trait⁹ caractérise l'ouverture du conduit. [+ SONANT] désigne l'ouverture quand la source d'excitation principale est le larynx, soit en français les voyelles, les glides, les liquides et les nasales. [- SONANT] renvoie

9 [± SONANT] est habituellement classé comme un trait de classe majeur dans les modèles géométriques. Il nous a semblé plus pertinent de le traiter avec les traits concernant le mode articuloire des consonnes

- aux phonèmes pour lesquels la source d'excitation est dans la cavité buccale, soit les consonnes occlusives et les fricatives ;
- Le trait [\pm CONTINU]. Ce trait indique les sons que l'on peut prolonger. [-CONTINU] désigne les occlusives orales, [+CONTINU], l'ensemble des autres phonèmes du français ;
 - Le trait [\pm NASAL]. Ce trait renvoie au passage de l'air dans la cavité nasale. [+NASAL] désigne en français les voyelles et consonnes nasales, [-NASAL], l'ensemble des autres phonèmes.
- Pour le lieu d'articulation des consonnes :
 - Le trait [\pm CORONAL]. [+CORONAL] indique les sons produits avec la pointe de la langue ;
 - Le trait [+HAUT] désigne l'articulation avec le dos de la langue ;
 - Le trait [+ANTÉRIEUR] indique une articulation alvéolaire, ou plus en avant ;
 - Le trait [+ARRIÈRE] renvoie aux phonèmes produits avec la masse de la langue en arrière.
 - Pour les voyelles :
 - Les traits [\pm HAUT] et [\pm BAS] indiquent la position de la langue, et désignent les degrés d'aperture ;
 - Le trait [+ARRIÈRE] renvoie aux voyelles postérieures ;
 - Le trait [+ARRONDI] désigne les voyelles et les glides articulés avec une projection des lèvres.
 - Enfin, le trait [\pm REL RET]. Le relâchement retardé permet de distinguer les occlusives (orales et nasales) [-REL RET] de l'ensemble des autres phonèmes du français, en observant le moment où les articulateurs sont relâchés.

Rappelons que le reproche communément fait à cette matrice des traits est son extrême redondance. Cette redondance est un avantage, nous permettant de comparer entre eux l'ensemble des phonèmes du français. Mais c'est également un inconvénient, du fait de sa lourdeur.

C'est pourquoi, dans la présentation des résultats phonologiques à venir, nous réduirons régulièrement l'utilisation de ces traits aux traits pertinents pour décrire les segments étudiés.

	p	b	t	d	k	g	f	v	s	z	ʃ	ʒ	m	n	ɲ	l	r	j	w	ɥ	i	y	u	e	ø	o	ɛ	œ	ɔ	a	ɛ̃	ɔ̃	ã			
[SONANT]	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	[SONANT]	
[SYLLABIQUE]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	[SYLLABIQUE]
[CONSONANTIQUE]	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	[CONSONANTIQUE]
[CONTINU]	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	[CONTINU]
[NASAL]	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	[NASAL]
[HAUT]	-	-	-	-	+	+	-	-	-	-	+	+	-	-	+	+	-	-	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	[HAUT]
[BAS]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	[BAS]
[ARRIÈRE]	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	+	-	+	-	+	-	-	-	+	-	-	+	-	-	+	+	-	+	+	+	[ARRIÈRE]
[ARRONDI]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	+	+	-	+	+	-	+	+	-	-	+	-	-	-	[ARRONDI]
[ANTÉRIEUR]	+	+	+	+	-	-	+	+	+	+	-	-	+	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	[ANTÉRIEUR]
[CORONAL]	-	-	+	+	-	-	-	-	+	+	+	+	-	+	+	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	[CORONAL]
[VOISÉ]	-	+	-	+	-	+	-	+	-	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	[VOISÉ]
[REL RET]	-	-	-	-	-	-	+	+	+	+	+	+	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	[REL RET]

Tableau 3 : Matrice des traits distinctifs inspirée par Dell (1985)

Chapitre 2 : Aspects méthodologiques, protocoles et mises en œuvre

Il n'existait pas, à notre connaissance, pour le français, de données mettant en relation les deux formes binomiales concurrentes de coordination par « et » à leurs effectifs de réalisation. La première étape de ce travail a donc consisté à recueillir et à construire les différents supports d'études que nous présentons dans le chapitre à venir.

Nous appelons « support d'études » deux types différents de données : des corpus et des questionnaires d'enquête.

Nous les avons organisés en deux ensembles distincts ; le premier ensemble est constitué de six corpus, que nous vous présenterons en détail ci-après.

Le second ensemble, quant à lui, se compose d'une série de six questionnaires d'enquête que nous vous exposerons en 2.2. Protocole de constitution des questionnaires d'enquête (page 51).

2.1. Protocole de constitution des corpus

Avant de présenter plus en détail les corpus indépendamment les uns des autres, nous exposerons, dans un premier temps, les raisons pour lesquelles nous avons choisi de construire six objets différents (bien qu'étroitement liés les uns aux autres), puis nous verrons ensuite les considérations générales que ces objets ont en commun.

2.1.1. Pourquoi différents corpus

Compte tenu des travaux existants, sur l'anglais notamment, nous ne pouvons nous extraire des facteurs extraphonologiques décrits par Cooper et Ross (1975), Wright S., Hay J., Bent T. (2005) et par Renner V. (2006) présentés en 1.1. (page 14). C'est un fait, au risque de proposer ici un truisme, la langue est un domaine d'étude pluridimensionnel imposant par essence une visée pluridisciplinaire. Travailler sur un phénomène de langue nous expose à des biais de diverses natures auxquels nous avons été confrontée dès la genèse de notre réflexion.

De cette observation préliminaire, il nous semble évident qu'en plus des principes

phonologiques que nous cherchons à mettre en exergue, d'autres facteurs extraphonologiques actifs dans la parataxe syndétique (ou coordination) sont également à prendre en compte, qu'il s'agisse de facteurs psycho-affectifs, syntaxiques et/ou sémantiques. Nous avons cependant délibérément axé notre démarche selon un point de vue le « plus strictement phonologique possible » en nous interrogeant expressément sur la manière d'extraire et de rendre saillants les facteurs phonologiques de cet ensemble de principes apparemment inextricables.

C'est ce que nous nous sommes attachée à faire notamment en mettant en place le corpus TOM et ZOÉ que nous présenterons page 39. En effet, ce corpus ne collectant que des appariements de prénoms, nous avons pu nous affranchir des facteurs syntaxiques et sémantiques que nous évoquons. Travailler en amont sur des prénoms nous a permis d'asseoir l'existence de facteurs phonologiques agissant dans la coordination binaire de deux prénoms en français, autorisant par là-même de présupposer l'existence de contraintes phonologiques actives dans la coordination binomiale par « et » (Couasnon, 2009). Cependant, l'intérêt que nous portons à la coordination par « et » porte bien plus sur la forme phonologique des mots coordonnés que sur leur étiquetage syntaxique. C'est pourquoi nous avons souhaité vérifier si les facteurs gouvernant la coordination par « et » de deux prénoms sont les mêmes que ceux qui agissent dans la coordination binaire directe par « et » de deux noms communs, de deux adjectifs, de deux adverbes et de deux verbes.

Afin de répondre à ces questionnements, nous avons mis en place cinq nouveaux corpus : les corpus TARN ET GARONNE, ARTS ET MÉTIERS, CUL ET CHEMISE, LISSE ET SOYEUX et VITE ET BIEN. Nous avons ainsi recueilli un total de 3 172 composés coordinatifs syndétiques par « et » que nous allons présenter plus en détail dès la page 39.

Cet ensemble, bien qu'apparemment hétéroclite, est en fait construit graduellement ; les corpus se répondent les uns aux autres dans une volonté de recouvrir au maximum l'ensemble des phénomènes de la coordination binaire directe par « et ».

Ainsi, suite aux prénoms relevés dans le corpus TOM ET ZOÉ, nous avons construit un corpus de noms propres, le corpus TARN ET GARONNE composé de noms toponymiques excluant les facteurs psycho-affectifs que nous évoquons infra. Le corpus ARTS ET MÉTIERS est constitué d'appariements de noms communs et d'adjectifs qui, en désignant des institutions coutumières ou légales, ou des titres de presse, se comportent comme des noms propres, mais également comme des locutions figées, car elles acceptent les six conditions nécessaires du figement que nous présenterons page 44. Le corpus CUL ET CHEMISE, quant à lui, recueille des expressions figées lexicalisées (noms communs, adjectifs, adverbes et verbes coordonnés par « et ») nous permettant de révéler la façon dont s'est fixée la

coordination binaire simple dans le temps. Parallèlement, le corpus LISSE ET SOYEUX rassemble des noms communs, des adjectifs et des verbes tensés dont on peut attester l'actualisation contemporaine. Enfin, le corpus VITE ET BIEN regroupe des co-occurrences par « et » de noms propres, d'adjectifs et d'adverbes statistiquement éprouvés par un test de χ^2 que nous avons par ailleurs étendu à l'ensemble de nos données¹⁰.

De fait, nous avons souhaité asseoir nos travaux sur un ensemble de six corpus autonomes, mais liés, à la fois pluriel et composite, afin de balayer le plus large spectre possible de la coordination binomiale directe par « et ».

2.1.2. Principes généraux et normes méthodologiques de la constitution des corpus

Ces six corpus imbriqués les uns dans les autres partagent ensemble des principes généraux en lien étroit avec la façon dont ils ont été conçus.

La première considération générale est liée au mode de constitution de l'ensemble de nos corpus : pour chacun d'entre eux, nous avons utilisé Internet. A cela, deux motivations essentielles.

La première raison est directement induite par notre volonté de recouvrir l'ensemble de la coordination directe simple par « et ». Travailler sur Internet nous permet d'avoir accès à un échantillon incomparablement plus large que celui qui serait accessible grâce à des entrevues. Le Web met à notre disposition gratuitement et rapidement un flux pléthorique d'informations en perpétuelle évolution sur tout type de sujets qu'il nous serait difficile d'appréhender autrement.

La deuxième raison est, elle, liée aux types d'informations que nous souhaitons recueillir. Nous avons fait le choix de traiter un phénomène de langue d'un point de vue essentiellement phonologique. Notre enjeu principal vise à examiner et à mettre en perspective des hypothèses théoriques avec des productions orales et réelles du français. Pour cela, nous aurions pu procéder à des séries d'entrevues, nous privant de fait du large éventail d'échantillonnage que nous venons d'évoquer.

Nous avons donc préféré utiliser Internet car l'interroger, c'est interroger les Communications médiatisées par ordinateur (ou C.M.O.) décrites par Panckhurst (1997), c'est-à-dire toutes interactions produites sur un ordinateur (qu'il s'agisse de *chats*, de forums, d'email, des

¹⁰ Le test de χ^2 explore les différences entre des données catégorielles. Dans le cadre de nos travaux, il nous a permis de tester les différences entre deux fréquences. Nous développerons page 68 l'utilisation précise que nous avons faite de ce test.

réseaux sociaux ou de blogs) présentées comme des « écrits oralisés » ou des « parlécrit » ayant leurs propres caractéristiques linguistiques. Comme le signalait Anis (2006) : « Il s'agit d'un écrit brut (sans relecture), familier (alors qu'on associe habituellement écrit et formalisme), affectif (expression des sentiments favorisant le relâchement du contrôle), ludique (s'exprimant par la néographie, le jeu de mot), socialisant (dominance de la fonction phatique dans les messages, partage de codes communs). Les caractéristiques formelles de cet écrit sont l'oralité (les effets d'oralité en fait). »

En recueillant des données issues des C.M.O., nous avons eu accès à des productions réelles et oralisées du français, comme nous nous proposons de le faire.

Cependant, il nous faut modérer nos propos. L'utilisation de la Toile comme source suppose des biais inhérents à la nature instable d'Internet.

Ainsi, il n'existe pas ou peu de déclarations de contenus. Il est impossible de contrôler l'ensemble des caractéristiques d'un document, qu'il s'agisse de sa date de création, de son origine ou même de son auteur (dont des renseignements sociologiques pertinents comme son âge, sa situation socio-professionnelle, sa situation géographique, son niveau d'études, et en ce qui nous concerne plus particulièrement, la proximité psycho-affective entretenue avec un prénom ou sa langue première). De la même façon, nous n'avons pas non plus accès aux informations formelles à propos des documents interrogés - leur contexte, leur taille, la langue dans laquelle ils sont rédigés, mais également le type du texte (formel, informel, quel usage...) ou leur répartition sur le Web. Nous avons donc choisi d'effectuer l'ensemble de nos recherches sur les pages « France » afin de contenir au maximum l'absence de déclarations de contenus (même si nous ne pouvons garantir que les auteurs soient exclusivement des locuteurs natifs du français).

Ce manque de renseignements est à corrélérer à un problème plus anecdotique qu'il nous a été néanmoins nécessaire de régler également : le problème de l'orthographe sur Internet. Véronis et Guimier De Neef (2006) l'affirmaient : « L'une des caractéristiques les plus frappantes des NFCE [nouvelles formes de communication écrite] est leur déviance par rapport à l'écrit classique. »

Pour pallier ces phénomènes, nous avons fait le choix de stabiliser nos recherches à une seule forme orthographique (même pour les prénoms) en produisant nos requêtes en majuscule et entre guillemets pour nous dédouaner des problèmes de casses et d'accents, afin de ne rechercher que les formes exactes soumises à interrogation.

Par ailleurs, l'accès à une telle masse de résultats ne peut se faire que grâce à l'utilisation de moteurs de recherche. Dans nos travaux, nous avons choisi de recourir à Google. Nos corpus

ainsi construits sont donc contraints par le choix de l'outil utilisé. L'accessibilité au Web est assujettie au type de requêtes formulées (la requête par mots). L'indexation des moteurs de recherche ne se fait pas sur les résultats, mais sur les pages. La conséquence directe est que toute nouvelle mise à jour est comptabilisée comme étant une nouvelle page. Ce phénomène est accentué par l'afflux constant de nouvelles données sur Internet, qui rend son contenu instable. Il est très souvent difficile d'obtenir exactement le même résultat pour une même requête à quelques heures près.

Cependant, malgré ces « défauts » inhérents à Internet, nous avons fait le choix de l'utiliser ayant pleinement conscience que de tels outils restent des biais. Aucun corpus construit n'est parfait du fait même de sa construction. D'autant plus que tous sont le pendant de la grande qualité d'Internet : le nombre. Certes, il n'a pas de déclarations de contenus et le contenu est instable, mais la masse d'informations arase les biais. Nous avons par ailleurs tenu à vérifier statistiquement l'ensemble de nos données présentes dans les six corpus en testant les appariements par un test de χ^2 . Nous n'avons conservé que les formes coordonnées sélectionnées statistiquement et majoritairement. Ne sont utilisées dans ce travail de thèse que des données pour lesquelles nous avons pu exclure l'hypothèse nulle, que des données statistiquement fiables ($p \leq 0,05$).

La seconde considération générale est, elle, liée à l'écrit. Nous n'avons pas de données orales concernant nos corpus. Comme nous le présentons page 25, nous considérons les réalisations des formes coordonnées étudiées selon la variation du français standard. Il est cependant des cas pour lesquels les réalisations phonologiques peuvent être hésitantes. Afin d'être la plus précise possible, nous avons pris le parti de les exclure de nos corpus.

De la même façon, et pour les mêmes raisons, sans réalisation orale, la question du traitement de la liaison nous a posé problème. En effet, dans les corpus ARTS ET MÉTIERS, CUL ET CHEMISE, LISSE ET SOYEUX, et VITE ET BIEN, 138 appariements sont potentiellement liaisonnants. Dans des cas comme « Mesdames et Messieurs » ou « Fruits et Légumes », rien ne peut nous garantir que l'ensemble des locuteurs ayant sélectionné ces formes coordonnées plutôt que leur inverse réalise systématiquement la liaison. Est-ce la forme /medamemesjø/ ou /medamzemesjø/ qui est retenue ? /fʁɥielegym/ ou /fʁɥizelegym/ ? Considère-t-on M1 dans « Mesdames et Messieurs » comme finissant par une consonne nasale ou par une fricative ? « Fruits » se termine-t-il par une voyelle ou par une consonne ?

La solution la plus radicale eut été d'extraire purement et simplement ces formes

potentiellement liaisonnantes. Elle serait sans doute la moins contestable. Cependant, comme nous avons la possibilité de traiter ces appariements à part, nous avons choisi de les conserver et de les doubler selon leurs deux réalisations phonologiques, avec ou sans liaison. Il est évident que nous ne prendrons pas en compte ces collocations (sous leurs deux formes) dans le cadre de l'étude générale du segment final des lexèmes présentées page 192. Nous avons choisi de les traiter à part dans un chapitre spécifiquement dédié aux phénomènes de liaison (page 274).

Quoi qu'il en soit, nous demeurons absolument convaincue de l'intérêt des ces appariements non seulement du fait de leur nombre, d'une part, mais d'autre part, du fait de la variété de profils phonologiques qu'ils mettent à notre disposition.

2.1.3. Protocole de constitution du corpus TOM ET ZOÉ

Notre premier corpus est constitué du recueil d'un ensemble de 2 500 binômes de prénoms coordonnés par « et » (du type « Tom et Zoé » ou « Emma et Simon ») mettant en corrélation les deux formes concurrentes de deux prénoms (un prénom masculin et un prénom féminin) coordonnés par « et », leur fréquence d'apparition, ainsi que leur rapport numérique.

L'avantage d'un corpus de prénoms est, comme nous l'indiquions précédemment, qu'il permet d'exclure des facteurs extraphonologiques notamment syntaxiques, sémantiques, psycho-affectifs, mais aussi culturels.

Pour ce faire, nous avons créé un support informatisé utilisant l'Internet comme source de données en associant les deux formes concurrentes de coordination des 100 prénoms les plus souvent donnés en 2008 (50 prénoms féminins et 50 prénoms masculins) à leur fréquence d'apparition sur Google.

Ce choix nous a été dicté par les raisons suivantes :

En premier lieu, les recherches sur l'Internet nous permettent une observation effective d'un phénomène en synchronie.

De plus, la sélection des 100 prénoms les plus fréquemment donnés en 2008 permet de balayer un panel représentatif (bien que non exhaustif) des patrons phonologiques des prénoms français. Mais en plus, l'utilisation de prénoms contemporains et fréquents garantit leur présence en grand nombre sur le Net. Ça n'aurait pas été le cas en élisant des prénoms moins actuels ou plus rares.

Enfin, concernant les principes extraphonologiques que nous évoquions dans le précédent paragraphe, l'Internet offre pléthore d'informations gratuites et rapidement accessibles, qui du

fait de leur nombre, nous protège des biais liés à la fréquence ou à la proximité affective que pourrait provoquer un petit nombre d'internautes ; nous le déclarions antérieurement, la masse arase les aspérités. D'autant plus qu'en choisissant de travailler en amont sur la coordination de deux prénoms en français, nous avons mis de côté, dans un premier temps, les principes sémantiques et syntaxiques actifs dans la coordination binomiale. Coordonner deux prénoms, c'est coordonner deux objets syntaxiquement et sémantiquement identiques.

Ainsi, pour constituer le corpus TOM ET ZOÉ, nous avons proposé sur Google les deux patrons de requête ci-dessous :

« PRENOM MASCULIN ET PRENOM FEMININ » - (« PRENOM MASCULIN ET
PRENOM FEMININ » MARIAGE FIANÇAILLES)
« PRENOM FEMININ ET PRENOM MASCULIN » - (« PRENOM FEMININ ET PRENOM
MASCULIN » MARIAGE FIANÇAILLES)

Pour chaque binôme mixte, nous avons associé des formes concurrentes de coordination et leur fréquence, en soustrayant les pages contenant les termes « MARIAGE » et « FIANÇAILLES ». En effet, l'usage prescrit que dans les faire-part de mariage, on place le prénom féminin avant le prénom masculin lorsque l'on cite un couple non marié ; et inversement, lorsque l'on évoque des couples mariés, le prénom masculin doit être suivi du prénom féminin. Ne mesurant pas exactement l'incidence de telles coutumes dans la construction binomiale par « et » et afin de nous en prémunir au maximum, nous avons préféré éliminer en amont les pages évoquant les mariages et les fiançailles.

Nous avons donc soumis 2 500 paires à ce traitement en respectant l'orthographe et la casse proposées par la liste des 100 prénoms les plus fréquemment donnés aux nouveaux nés en 2008.

Avant tous traitements et toutes analyses de ces données, nous les avons épurées une nouvelle fois en retirant le plus possible les appariements pour lesquels nous soupçonnions une influence extraphonologique, dont notamment les facteurs culturels.

En effet, peut-on imaginer que l'on entend plus souvent « Roméo et Juliette » en référence l'œuvre éponyme, par « habitude culturelle » ? Si Shakespeare avait titré « Juliette et Roméo », cette dernière forme nous semblerait-elle plus naturelle et apparaîtrait-elle plus souvent ? Comment ne pas suspecter une contrainte culturelle ? A nouveau, afin de nous prémunir de la possible prégnance d'un tel facteur dans la sélection d'un ordre préférentiel de coordination par « et », lorsqu'un appariement faisait référence à des paires de personnages

publics, réels ou fictifs (personnages réels, titres de films, de chansons, de livres ou marques), nous avons retiré ses deux formes concurrentes de coordination. Nous avons par exemple soustrait les deux formes binomiales coordonnées par « et » pour les prénoms « Théo » et « Inès » initialement présents dans le corpus TOM ET ZOÉ. En effet, « Théo et Inès » est une marque de mobilier pour enfants. Et afin de nous prévenir de tous biais, non seulement de la surnumération des pages Internet liées à cette marque, mais aussi d'effets d'habitude qui feraient que des clients fidèles de ce mobilier pourraient être susceptibles de sélectionner cette forme car elle leur apparaîtrait plus « naturelle », nous avons préféré les supprimer.

De la même façon, afin d'être le plus précise possible, nous avons exclu de nos données tous les termes pour lesquels la réalisation phonologique hésitante était invérifiable. Ainsi, nous avons retiré tous les appariements contenant les prénoms tels que « Dorian » et « Killian », du fait de l'immense hésitation existant quant à leur prononciation en français standard : /k i l j ā / /kiljan/, /d o r j ā , /d o r j a n/.

Enfin, nous avons pris une dernière précaution avant toutes observations, en écartant de nos analyses tous les binômes construits avec le prénom « Marie ». Ce dernier entre très fréquemment dans la constitution de prénoms composés soit dans le second composant (Pierre-Marie, Jean-Marie, Anne-Marie, Claire-Marie, Jacques-Marie, Louis-Marie, Rose-Marie...) ou, massivement, dans le premier (Marie-Adélaïde, Marie-Agnès, Marie-Alice, Marie-Amélie, Marie-Antoinette, Marie-Béatrice, Marie-Chantal, Marie-Christine, Marie-Françoise, Marie-Thérèse...). Et il nous est très difficile de savoir si toutes les occurrences répertoriées du prénom « Marie » en M1 ou en M2 sont des occurrences de « α -Marie », de « Marie- β » ou de « Marie ». Ne pouvant garantir qu'il ne s'agissait pas de prénoms composés, il nous est apparu plus prudent de retirer tous les binômes comportant le prénom « Marie ». Un traitement manuel aurait certes été plus fin, mais très fastidieux et extrêmement « chronophage ».

Ces précautions mises à part, il apparaît que ce corpus ainsi construit, comme tout corpus, est évidemment perfectible. Il présente un certain nombre de zones grises qui seront pour la plupart lissées grâce à la pléthore de résultats que nous offre Internet. Il convient néanmoins d'en prendre conscience et de les présenter.

En effet, du fait de la constitution de ce corpus informatisé, nous ne pouvons être sûre de ne traiter exclusivement la coordination de deux prénoms par « et ». Ce problème est identique à celui que nous venons de mettre en évidence pour les prénoms composés. La requête lancée sur Google ne nous permet pas de distinguer les pages citant des binômes coordonnés de

celles proposant des constructions coordonnées ternaires. « Pierre, Paul et Jacques » pourra donc être, de fait, comptabilisée comme étant une occurrence de la coordination simple de « Paul et Jacques ».

De plus, comme nous le précisons précédemment, il n'existe pas ou peu de déclarations de contenus sur Internet.

Pour ces deux motifs en général, et comme ce corpus en particulier n'est pas de même nature que les autres supports d'études que nous présentons à suivre du fait de sa taille et de la façon dont il est constitué, nous ne validerons jamais un résultat s'il est uniquement présent dans le corpus TOM ET ZOÉ. En effet, non seulement il s'agit du support d'études le moins contrôlé de nos supports d'études, mais c'est également celui qui est composé du plus grand nombre d'appariements. Ainsi, il est tout à fait envisageable que des facteurs extraphonologiques comme le genre, ou la proximité psycho-affective que nous développerons page 83 soient particulièrement actifs dans la sélection d'une forme binomiale coordonnée. Or, ce corpus est conséquent en termes de nombre d'effectifs testés ; ses résultats pourraient influencer l'ensemble de nos observations. C'est pourquoi afin de ne pas biaiser nos résultats totaux, nous ne confirmerons jamais un résultat porté par un seul support d'études, *a fortiori* s'il s'agit de résultats portés par le corpus TOM ET ZOÉ. Car, c'est un fait, nous avons été amenée à prendre un certain nombre de décisions méthodologiques qui ont nécessairement orienté nos travaux. Il demeure que l'atout majeur et premier du corpus TOM ET ZOÉ est « de mettre en place un support d'études rapidement exploitable sur lequel il nous est possible d'asseoir une observation des principes régissant la sélection d'un ordre préférentiel dans la coordination de deux prénoms mixtes et d'en proposer une analyse théorique et formelle » (Couasnon, 2009). Ce travail constitue, à notre connaissance, la seule collection de donnée en français mettant en relation des paires coordonnées de prénoms et leur nombre d'occurrences. Le corpus TOM ET ZOÉ est le premier corpus qui a été élaboré dans le cadre de ce travail. Il nous a permis de mettre en évidence un déterminisme phonologique actif dans la coordination binomiale par « et » et d'émettre des hypothèses. C'est à partir de ces hypothèses que d'autres outils, dont les questionnaires d'enquête que nous présenterons dès la page 51, ont ensuite été développés.

2.1.4. Protocole de constitution du corpus TARN ET GARONNE

Ce deuxième corpus est constitué des noms des départements français et collectivités d'outre-mer ainsi que des noms de l'ensemble des villes et villages d'Aquitaine¹¹ dont les formes mettent en relation deux noms propres coordonnés à l'aide de la conjonction de coordination « et ».

Ainsi ont été collectés les 12 noms de départements (et collectivités) français coordonnés par « et » du type « WALLIS-ET-FUTUNA » et « TARN-ET-GARONNE » ; et les 57 villes et villages d'Aquitaine construits avec « et » dont « LIT ET MIXE », « BRAUD ET SAINT-LOUIS », « COUX ET BIGAROQUE » et « UCHAQ ET PARENTIS ».

Dans les faits, il ne s'agit que de noms de villes et de villages de Dordogne, de Gironde et des Landes. En effet, nous n'avons pas relevé l'existence de villes ou villages coordonnés par « et » dans les départements du Lot-et-Garonne et des Pyrénées-Atlantiques.

Au total, 69 appariements toponymiques prennent part à la constitution du corpus TARN ET GARONNE¹².

La présence de ces noms toponymiques dans notre corpus offre de nombreux avantages.

D'une part, comme nous le présentions au préalable, analyser des noms propres nous permet d'assurer une transition entre l'étude de la coordination de deux prénoms et celle de noms substantivés, adjectivés, de verbes et d'adverbes.

D'autre part, exception faite d'une poignée de villages pour laquelle nous nourrissions quelques hésitations quant à leur prononciation (hésitations qu'il a été aisé de lever en interrogeant des locuteurs natifs de ces départements ou en contactant directement les mairies), en tant que locutrice native de la Gironde, nous en connaissons les réalisations phonétiques.

Enfin, d'un point de vue strictement phonologique, ces villes et villages du Sud-Ouest présentent un large éventail de consonnes finales fixes, dont des clusters en /ks/ ou des finales en /k/ (comme dans les exemples présentés *supra*) plutôt rares en français.

11 Selon la division administrative du territoire antérieure à 2015.

12 Nous avons évidemment « nettoyé » chacun de nos supports d'étude avant de les utiliser tels quels. Comme nous le présenterons en détail en 3.1. L'ordre alphabétique page 75, nous avons notamment confirmé que l'ordre de coordination sélectionné n'était pas imputable à l'ordre alphabétique.

2.1.5. Protocole de constitution du corpus ARTS ET MÉTIERS

Ce troisième corpus rassemble initialement 110 paires constituées de noms d'institutions coutumières ou légales et de titres de magazines.

Sont ainsi réunis, entre autres :

- Des noms d'écoles (« arts et métiers »);
- Des noms d'associations loi 1901 (« enfance et partage ») ;
- Des lois (« informatique et libertés ») ;
- Des labels (« tourisme et handicap ») ;
- Mais aussi des titres de magazines dont des appariements de noms adjectivés du type « jeune et jolie » ou de noms substantivés : « modes et travaux » « art et décoration » ou « air et cosmos ».

Cet ensemble répond aux critères de figement exposés ci-après.

2.1.6. Protocole de constitution du corpus CUL ET CHEMISE

Le quatrième corpus rassemble 49 appariements constitués d'expressions que nous identifions comme figées, du type « cul et chemise », « juge et partie » ou « monts et merveilles ». Nous identifions les expressions comme figées en nous appuyant sur les conditions de figement définies par Gaston Gross.

Selon Gross (1996), les expressions figées doivent répondre à six conditions nécessaires. La séquence « CUL ET CHEMISE » nous servira d'exemple pour illustrer le court rappel à suivre.

1. La polylexicalité. Dans une expression figée, chaque terme composant a une existence lexicale qui lui est propre. Dans notre exemple « cul et chemise », les termes « cul » et « chemise » ont des entrées lexicales autonomes dans les dictionnaires ;
2. L'opacité sémantique. Les expressions figées ne relèvent pas de la lecture compositionnelle. Le sens de la séquence ne peut être déduit du sens des mots la composant. Ainsi le sens de l'expression « cul et chemise », comme « être inséparable, lié par une solide amitié » telle qu'elle est définie dans le TLFi (consultation en ligne le 28/03/2015), n'est en aucun cas le produit des éléments autonomes « cul » et « chemise » définis comme « partie du corps humain » et comme « vêtement » ;
3. Le blocage des propriétés transformationnelles. Les expressions figées ne peuvent

subir des transformations syntaxiques. Il nous sera, par exemple, impossible de les recatégoriser, de les permuter ou de modifier leur accord. Ainsi la séquence « culs et chemises » au pluriel renverra aux deux objets lexicaux « culs » et « chemises » coordonnés définis plus haut, et non à l'expression figée « cul et chemise » en tant qu'« intimité entre deux personnes », tout comme la séquence « chemise et cul » ;

4. La non-actualisation des éléments. De la même façon, il nous est impossible d'actualiser les expressions figées en les situant temporellement ou géographiquement. *« être cul et chemise sur une chaise » nous apparaît au mieux bancal ou humoristique, au pire inacceptable ;
5. Le blocage des substitutions synonymiques. Il est également impensable de substituer les composants de la séquence avec un synonyme tout en conservant le sens premier de l'expression figée. Être *« fesse et chemise » n'évoque en rien le rapport intime entre deux individus ;
6. L'impossibilité d'insertions. Enfin, il est invraisemblable d'insérer des éléments entre les deux composants d'une expression figée. *Être « cul-terreux et chemise sale » paraît de fait inadmissible.

Dans ses travaux, Gross n'a pas examiné les structures M1&M2 composées de deux mots coordonnés par « et » ; il nous est néanmoins apparu que des structures comme « CUL ET CHEMISE », « JUGE ET PARTI » ou « MONTS ET MERVEILLES » (tout comme « ENFANCE ET PARTAGE », « ARTS ET MÉTIERS », « MODES ET TRAVAUX » ou « JEUNE ET JOLIE » et l'ensemble du corpus « ARTS ET MÉTIERS » précédemment évoqué) répondaient au moins en partie à la catégorisation qu'il propose. Ces trois séquences sont polylexicales; étant toutes composées de deux mots ayant un sens autonome. La lecture compositionnelle est impossible. Ces séquences en tant qu'expressions figées possèdent un sens figé très éloigné du produit des éléments les composant.

En suivant ces critères, nous avons donc recueilli 49 appariements figés et lexicalisés partant de l'axiome que le figement s'effectue sur les formes les plus usuelles, donc les plus « naturelles ».

Cependant, qu'il s'agisse de noms d'institutions ou de titres de presses ou d'expressions binomiales figées, le figement soulève inévitablement la question du moment où ces appariements ont cessé d'être libres et se sont fixés en tant qu'expressions figées.

Pour reprendre nos précédents exemples, le TLFi (consultation le 28/03/15) assure l'existence

de formes attestées¹³ de la locution « MONTS ET MERVEILLES » en 1812¹⁴ et 1920¹⁵, en 1947¹⁶ et 1949¹⁷ pour « CUL ET CHEMISE » et de 1836¹⁸ pour le binôme « JUGE ET PARTIE ». Quant aux exemples que nous proposons pour les paires des noms d'institutions et de titres de presses, si « ENFANCE ET PARTAGE »¹⁹, « TOURISME ET HANDICAP »²⁰ ET « INFORMATIQUE ET LIBERTÉS »²¹ et « JEUNE ET JOLIE »²² sont à peu près contemporains, le premier numéro de « AIR ET COSMOS » est paru en 1963, « MODES ET TRAVAUX » en 1919 et « ART ET DÉCORATION » en 1897, et l'école des « ARTS ET MÉTIERS » a été créée en 1780.

Bien que cette thèse se donne pour objectif d'étudier la coordination binaire pas « et » dans ses actualisations contemporaines, il ne serait pas judicieux d'exclure des expressions telles que « MONTS ET MERVEILLES », « CUL ET CHEMISE » ou « JUGE ET PARTIE ». En effet, nous pensons que le figement comme procédé qui s'effectue dans le temps sur des formes usuelles et fréquentes nous fournit de précieux indices sur ce qui est phonologiquement acceptable dans la coordination par « et » en français et sur ce qui l'est moins. Nous avons pris en compte ces questions dès la constitution de l'ensemble de nos corpus en faisant en sorte que chaque corpus, bien que traité et analysé conjointement, soit autonome et sécable. Un résultat porté seul par les corpus ARTS ET MÉTIERS et CUL ET CHEMISE sera proposé avec l'ensemble des réserves que nous venons de présenter.

Enfin, l'attestation d'un appariement comme expression figée, qu'elle soit lexicalisée (comme

13 Il est fort probable que le figement soit bien antérieur aux attestations relevées dans le TLFi.

14 « On attend **monts et merveilles** du marquis de Martel, qui s'est vanté de forcer les Algériens à la paix », Étienne de JOUY, (1812), dans *L'Ermite de la Chaussée d'Antin, ou observations sur les mœurs et les usages français au commencement du XIX^{ème} siècle*

15 « Les meilleures, entendant leur mari dire **monts et merveilles** de l'esprit de la duchesse, estimaient que celle-ci était si supérieure au reste des femmes qu'elle s'ennuyait dans leur société, car elles ne savent parler de rien. » Marcel PROUST, (1920), *Le côté de Guermantes*.

16 « Naturellement, avec Langlois, en tout bien tout honneur, ils s'entendaient comme **cul et chemise**. » Jean GIONO, (1947), *Un Roi sans divertissement*.

17 « Ces deux-là, c'est **cul et chemise**, toujours ensemble ou à se courir après » (Jean Paul SARTRE, (1949), *La Mort dans l'âme*.

18 « N'est-ce pas un crime que de donner le jour à de pauvres créatures condamnées par avance à de perpétuelles douleurs? Cependant ma conduite soulève de si graves questions que je ne puis les décider seule ; je suis **juge et partie**. » Honoré de BALZAC, (1836) *Le lys dans la vallée*.

19 Les statuts de l'association « ENFANCE ET PARTAGE » ont été déposés en mai 1977.

20 La loi « INFORMATIQUE ET LIBERTÉ » a été promulguée en 1978.

21 Le label « TOURISME ET HANDICAP » a été mis en place en 2003.

22 La première parution du mensuel « JEUNE ET JOLIE » date de 1987.

dans le corpus CUL ET CHEMISE) ou non lexicalisée (comme dans le corpus ARTS ET MÉTIERS), n'est pas nécessairement synonyme de réalisations statistiquement valides.

Le figement d'une forme ne garantit pas pour autant sa « réussite » ; une forme figée, sans la moindre vérification statistique, ne peut prétendre être la forme binaire coordonnée par « et » majoritairement sélectionnée par les locuteurs natifs du français²³.

2.1.7. Protocole de constitution du corpus LISSE ET SOYEUX

Le cinquième corpus, le corpus LISSE ET SOYEUX, recueille 115 allégations cosmétiques sous la forme d'appariements binaires coordonnés par « et ». Nous appelons « allégations cosmétiques » les assertions présentant ou revendiquant les caractéristiques d'un produit cosmétique, son efficacité ou une ou plusieurs de ses qualités. A la différence de l'étiquetage, ces allégations sont écrites en grandes lettres visibles et saillantes sur l'emballage.

Nous avons recueilli le 22/02/2014 la liste exhaustive des allégations cosmétiques présentes dans les rayons d'une grande surface²⁴ et d'une parapharmacie²⁵. Ces appariements balayaient trois types syntaxiques de coordination par « et » :

23 Prenons l'exemple de «IMAGES ET RÉSEAUX ».

Cet appariement renvoie à une institution fondée en 2005 visant à rassembler les connaissances d'industriels et de chercheurs dans les domaines « des télécoms, de l'internet, de la télévision et de l'image ». Ainsi, les deux termes composant « IMAGES ET RÉSEAUX » possèdent tous deux des sens autonomes : la lecture compositionnelle est impossible. De fait, nous ne pouvons les transformer syntaxiquement ; « image et réseau » au singulier ne renvoyant pas au pôle de compétitivité que nous venons d'introduire mais à quelque chose se référant à une « image » et à un ou du « réseau ». De la même façon et pour le même motif, nous ne pouvons les actualiser, les substituer ou insérer de nouveaux éléments sans en modifier le référent. « Cette image et ce réseau », « réseau et image », « ce nouveau et grand réseau » ne renvoient plus au même objet. Cette séquence « Images et Réseaux », selon la définition proposée par Gaston Gross, est donc une expression figée. Cependant, étudiant les réalisations réelles et contemporaines du français, comme nous le développerons un peu plus tard, nous avons soumis chacune de nos appariements sur Google afin de pouvoir attester de la forme la plus fréquemment utilisée. Nous avons donc soumis les deux requêtes suivantes sur Google le 06/06/2014 : « IMAGES ET RÉSEAUX » et « RÉSEAUX ET IMAGES ». Les recherches ont été effectuées en majuscule afin d'éviter les biais liés à la casse et entre guillemets afin de compiler uniquement et exactement les expressions recherchées. Il apparaît que la forme attendue « IMAGES ET RÉSEAUX » ressort dans seulement 45,77% des pages compilées. ($\chi^2(1) = 143,78$; $p \leq 0,05$). La forme figée n'est donc pas la forme privilégiée.

24 Magasin Auchan Bordeaux Bouliac, rue de la Gabarre 33270 Bouliac.

25 Pharmacie Principale, 10 rue Sainte Catherine, 33000 Bordeaux.

- Des coordinations « Adj et Adj » : « LISSE ET SOYEUX », « NETTOYANTE ET ADOUCISSANTE » ;
- Des coordinations « N et N » : « CÈDRE ET EUCALYPTUS », « COULEUR ET NUTRITION » ;
- Et des coordinations « Vbe et Vbe » : « NOURRIT ET DÉTEND ».

Du fait de sa nature, ce corpus présente plusieurs caractéristiques intéressantes pour notre étude.

Premièrement, il recouvre un large panel des types syntaxiques coordonnés par « et », exception faite des adverbes.

Deuxièmement, les allégations publicitaires sont soumises à une réglementation communautaire qui dépend de l'Autorité européenne de sécurité sanitaire afin de protéger les consommateurs des allégations fausses, ambiguës ou non validées scientifiquement. En France, l'ARPP (anciennement Autorité de régulation professionnelle de la publicité) et le CPP (Conseil paritaire de la publicité), ont proposé en mars 2010, une nouvelle version de la Recommandation « Produits cosmétiques²⁶ » appelant notamment dans l'article 20 1223/2009 « à la mise en place de critères communs » élaborés par la Commission européenne, relatifs aux allégations cosmétiques. Ces critères sont entrés en application dans l'ensemble des pays de l'Union européenne en juillet 2013. Les entreprises cosmétiques se sont donc vues dans l'obligation de réviser leurs allégations et étiquetages, ce qui constitue un immense avantage nous concernant. À la différence des appariements présents dans les deux précédents corpus, cette collection d'allégations publicitaires est tout à fait contemporaine.

Enfin, ces allégations publicitaires, en tant qu'objets de communication ayant une visée ouvertement incitatrice, nous amènent à nous interroger plus précisément sur les contraintes sémantiques. Bien que nourrissant l'ambition d'une étude essentiellement phonologique de la parataxe syndétique, nous savons que des principes extraphonologiques sont actifs. Il nous apparaît mal avisé de les exclure de notre réflexion. Le corpus LISSE ET SOYEUX nous permettra de mener des recherches sur cet aspect.

26 http://www.arpp-pub.org/IMG/pdf/rapport_cosmetique-2.pdf

2.1.8. Protocole de constitution du corpus VITE ET BIEN

Ce sixième corpus, dit corpus VITE ET BIEN, recueille initialement 319 collocations binaires directes coordonnées par « et ». Nous parlons ici de collocations et non de figements comme dans les corpus ARTS ET MÉTIERS et CUL ET CHEMISE. En effet, les binômes relevés et présentés dans ce corpus, bien qu'associant de manière privilégiée deux lexèmes sous une forme plus ou moins libre ne répondent pas à l'ensemble des six critères définis par Gaston Gross. Ainsi nous avons choisi ce terme de « collocation » bien qu'il soit extrêmement polysémique et revête plusieurs acceptions qui ont été largement décrites par Tutin et Grossmann (2003, 2002). Traditionnellement, on en distingue deux : une acception « qualitative » qui définit les collocations à travers leurs fonctionnements syntaxiques et sémantiques, et une autre, plus « quantitative », qui s'appuie sur la fréquence statistiquement remarquable d'association de deux lexèmes.

Nous utilisons, dans le cadre de cette étude, la collocation dans le second sens en tant que combinaison privilégiée de deux constituants au sein d'un groupe linguistique d'un point de vue « quantitatif », c'est-à-dire en relevant des cooccurrences statistiquement testées.

Nous avons donc relevé des collocations recouvrant à nouveau trois types syntaxiques de coordinations directes :

- Des coordinations « Adj et Adj » : « BONNE ET HEUREUSE », « VITE ET BIEN », « CLAIR ET NET », « COURT ET MOYEN » ;
- Des coordinations « Adv et Adv » : « ÇÀ ET LÀ », « ICI ET MAINTENANT » ;
- Et des coordinations « N et N » : « MESDAMES ET MESSIEURS », « ACHAT ET VENTE », « FRUITS ET LÉGUMES », « NUIT ET JOUR ».

Nous avons ensuite statistiquement éprouvé leur fréquence en vérifiant si la forme de la collocation ainsi relevée était la forme privilégiée.

Les quatre exemples « BONNE ET HEUREUSE », « VITE ET BIEN », « ICI ET MAINTENANT » et « ACHAT ET VENTE » nous permettent d'illustrer notre méthodologie de recueil de données.

Nous avons soumis sur Google deux types de requêtes du type « M1 ET M2 » et « M2 ET M1 ». Comme toujours, les requêtes étaient formulées en majuscule afin d'éviter les problèmes liés à la casse, et entre guillemets afin que soit relevée uniquement et exclusivement l'exacte expression recherchée.

- Ainsi, « BONNE ET HEUREUSE » apparaît dans 601 000 occurrences. « HEUREUSE ET BONNE » dans 71 000. « BONNE ET HEUREUSE » est donc la collocation privilégiée dans 89,5 % ($\chi^2(1) = 2090002,98$; $p \leq 0,05$) des appariements associant les lexèmes « bonne » et « heureuse » coordonnés par « et ».

De même :

- « VITE ET BIEN » apparaît dans 466 000 résultats. « BIEN ET VITE » dans 163 000. « VITE ET BIEN » est donc la forme privilégiée dans 74,09 % ($\chi^2(1) = 72980,13$; $p \leq 0,05$) des appariements associant les lexèmes « bien » et « vite » coordonnés par « et » ;
- « ICI ET MAINTENANT » ressort dans 614 000 occurrences, ce qui représente 89,92 % ($\chi^2(1) = 272 600$; $p \leq 0,05$) des formes binaires coordonnées par « et » associant « ici » et « maintenant » ;
- Enfin, « ACHAT ET VENTE » apparaît dans 70,75% ($\chi^2(1) = 349000$; $p \leq 0,05$) des expressions binaires directes par « et » liant le nom « achat » à celui de « vente ».

Compte tenu de l'acceptation du terme collocation que nous retenons pour les besoins de ce travail, il semble manifeste que d'un point de vue statistique, « BONNE ET HEUREUSE », « VITE ET BIEN », « ICI ET MAINTENANT » et « ACHAT ET VENTE » sont des collocations statistiquement remarquables.

2.1.9. Présentation récapitulative des corpus

D'une manière générale, nous n'avons conservé que les formes coordonnées sélectionnées majoritairement et validées statistiquement par un χ^2 , réduisant de près de trois cinquièmes les 3 172 composés coordinatifs initialement recueillis.

Ainsi, comme présenté dans le tableau récapitulatif ci-dessous, l'ensemble des 6 corpus est composé comme suit :

CORPUS	TYPE	NOMBRE
TARN ET GARONNE	Toponymes	69
CUL ET CHEMISE	Expressions figées	48
LISSE ET SOYEUX	Allégations cosmétiques	62
TOM ET ZOÉ	Prénoms	1420
VITE ET BIEN	Collocations remarquables	202
ARTS ET MÉTIERS	Institutions coutumières ou légales.	85
	Titres de magazines.	
TOTAL		1886

Tableau 4 : Les corpus

Grâce à l'ensemble des six corpus que nous venons de présenter, nous avons à notre disposition une collection conséquente de 1 886 composés coordinatifs syndétiques validés statistiquement, nous offrant, comme atout majeur, l'opportunité de manipuler un très grand nombre de données et d'hypothèses.

2.2. Protocole de constitution des questionnaires d'enquête

En complément des six corpus précédemment décrits, la nécessité d'asseoir nos résultats sur un autre type de support d'études a émergé.

Nous avons donc mis en place les six questionnaires d'enquête que nous présentons dans cette partie :

- L'enquête ÉLISE et ZOÉ ;
- L'enquête ÉRIC ET AURIC ;
- L'enquête PAUL ET PASCAL ;
- L'enquête ÉDOUARD ET EDMOND ;
- L'enquête NAROD ET NANOR ;
- Et l'enquête LEMOTO ET LEMOTAN.

Avant de présenter plus en détail ces questionnaires indépendamment les uns des autres, nous exposerons, dans un premier temps, les raisons pour lesquelles nous avons choisi de

construire six objets distincts mais étroitement liés, puis nous verrons les considérations générales et normes méthodologiques que ces objets partagent en commun.

2.2.1. Pourquoi différents questionnaires d'enquête

Le but de cet ensemble de questionnaires, à la différence de l'ensemble des corpus que nous présentions auparavant, n'est pas de collecter un vaste ensemble de données brassant un grand nombre d'appariements et de facteurs, mais au contraire de cibler des hypothèses émergentes pour les confirmer ou les infirmer. C'est pourquoi chaque questionnaire d'enquête circonscrit un très petit nombre d'inférences. Ces six objets, bien qu'autonomes, se répondent également les uns les autres selon deux axes : le premier renseigne des hypothèses plutôt extraphonologiques, le second, des hypothèses exclusivement phonologiques.

Ainsi concernant les hypothèses plutôt extraphonologiques : le questionnaire d'enquête ÉLISE et ZOÉ examine, entre autres, les questions du genre et des contraintes « culturelles » - faisant suite à l'enquête ÉRIC ET AURIC , explorant la question de la fréquence -, mais il introduit également la question du statut des voyelles initiales plus largement appréhendée dans le questionnaire d'enquête PAUL ET PASCAL qui sonde des critères liant la taille des constituants à la sélection d'une place privilégiée en M1 ou en M2.

Concernant les hypothèses exclusivement phonologiques : les trois derniers questionnaires d'enquête parcourent le type de segments finals des constituants (pseudo-prénoms). Par conséquent, le questionnaire d'enquête ÉDOUARD ET EDMOND observe des constituants à finales consonantiques coordonnés à des constituants à finales vocaliques, observations également développées dans l'enquête NAROD ET NANOR qui met, par ailleurs, en relation des pseudo-mots ayant des finales consonantiques de différents types. Enfin, le questionnaire d'enquête LEMOTO ET LEMOTAN examine la coordination par « et » de deux pseudo-mots à finales vocaliques.

De fait, nous avons souhaité asseoir nos travaux sur un ensemble de six questionnaires autonomes mais liés, afin de recouvrir un échantillonnage d'hypothèses phonologiques et extraphonologiques remarquables dont nous recherchions à infirmer ou à confirmer la validité.

Enfin, bien que ces six questionnaires d'enquête mettent en œuvre des prénoms (ou des pseudo-prénoms), nous considérons que leurs résultats sont valables au-delà des seuls prénoms. En effet, ces derniers, tels que nous les avons utilisés dans les questionnaires d'enquête, sont observés comme de « purs » objets phonologiques, exempts au maximum

d'autres faits langagiers. Rappelons en effet que l'avantage d'une collection de prénoms est qu'elle nous permet de nous dédouaner des facteurs syntaxiques et sémantiques qui pourraient être actifs dans la sélection d'un ordre préférentiel de coordination. On peut donc supposer que les phénomènes phonologiques rendus visibles grâce aux prénoms, ou aux pseudo-prénoms, sont valables pour d'autres mots que ces derniers.

2.2.2. Principes généraux et normes méthodologiques de la constitution des questionnaires d'enquête

Les six questionnaires d'enquête partagent ensemble des principes généraux répondant à une motivation fondamentale : celle d'interroger des locuteurs identifiés.

En effet, l'objection que l'on pourrait faire aux corpus présentés précédemment est le manque d'informations concernant les émetteurs. Comme nous l'indiquions, il n'y a pas de déclarations de contenu sur Internet. Vérifier l'intégralité des caractéristiques des données que nous avons relevées est impossible. Les dates de création, l'origine des documents, les renseignements sociologiques concernant les auteurs de ces pages (leur âge, leur sexe, leur langue maternelle, leur niveau d'études, leur situation géographique) : tout ceci nous est inconnu. Les situations d'énonciation sont complètement opaques. Et quand bien même, la masse des résultats arase les biais.

Nous avons donc souhaité tester certaines de nos hypothèses auprès de locuteurs identifiés grâce à des expériences nous permettant de contrôler plusieurs facteurs. Nous considérons ces expérimentations comme une sorte de mise à l'épreuve des précédents corpus. Il s'agit d'étalonner nos données. Si l'ensemble de nos corpus et expérimentations confirme un échantillon d'hypothèses, alors nous pouvons les affirmer sans la moindre hésitation possible.

Nous avons donc mis en place une série de questionnaires d'enquête que nous présenterons en détail dans les parties à venir, enquêtes dans lesquelles nous avons associé les deux formes concurrentielles de binômes coordonnées par « et » à leur fréquence de réalisation, c'est-à-dire en observant statistiquement si la forme « M1&M2 » est plus ou moins fréquente que la forme « M2&M1 ».

Ces expérimentations ont toutes été construites sur le même modèle, en respectant les mêmes contraintes et exigences. Il s'agit de questionnaires d'enquête standardisés et systématisés proposant les mêmes tâches à chaque sujet testé, à savoir la sélection, pour chaque paire de prénoms ou de pseudo-mots, de l'une des deux formes binomiales coordonnées par « et » préalablement suggérées.

Pour exemple, pour illustrer nos propos, nous proposons le binôme « TOM / JERRY » sous ses deux formes possibles :

Tom et Jerry

Jerry et Tom

Le fait de tester des prénoms ou des pseudo-prénoms nous permet d'évacuer rapidement tous biais syntaxiques ou sémantiques liés. Pour les mêmes raisons, nous avons proposé des appariements bruts, hors contexte. En effet, même si les prénoms ne sont pas de « purs » objets phonologiques (à la différence des pseudo-mots que nous nous proposerons de tester), ils ne sont pas sémantiquement contraints. Il ne s'agit que d'objets référentiels désignant des individus, pour lesquels il est impossible d'assigner une signification. Tout comme le fait de travailler hors contexte nous permet d'éviter l'influence de procédés syntaxiques comme la focalisation par exemple.

Ainsi chaque questionnaire répond aux mêmes normes méthodologiques.

Les interviewés sont tous et sans exception des locuteurs natifs du français. Il s'agit de locuteurs témoins naïfs n'ayant absolument aucune connaissance de l'objet de nos travaux.

En outre, si à travers nos corpus internet, il était possible de tester des centaines d'appariements, ce n'est pas le cas auprès de nos locuteurs témoins. Pour ne pas risquer la lassitude des personnes interrogées, compte tenu de la répétition du même format de questions semi-fermées, nous avons restreint l'ensemble de nos expérimentations à un maximum de quarante-trois appariements par expérience. Du fait de ce nombre, il a été absolument nécessaire de sélectionner des appariements pertinents ne ciblant qu'une seule hypothèse. Pour la même raison, même si les recommandations traditionnelles en sociolinguistiques (Calvet et Dumont, 1999) préconisent une durée maximale d'une heure pour des passations de questionnaires en une salle de classe), aucun test n'a excédé le temps imparti maximum de cinq minutes, car comme l'indique Boukous (1999), « les sujets sont plus enclins à répondre quand le questionnaire est court et précis ».

De surcroît, « l'ordre des questions peut aussi influencer sur la réponse du sujet qui a tendance à interpréter les questions selon leur position dans le questionnaire » (Boukous, 1999). Consciente que l'écrit peut biaiser les résultats de nos enquêtes, nous avons systématiquement distribué au groupe de locuteurs interrogés deux types de questionnaires en prenant soin d'inverser l'ordre des items proposés de haut en bas et de gauche à droite.

Ainsi par exemple si le questionnaire A de l'enquête LEMOTO ET LEMOTAN débutait par :

Akin et Ako				Ako et Akin
Amassa et Amassan				Amassan et Amassa

Et finissait par :

Lémoto et Lémotan				Lémotan et Lémoto
Massarou et Massaron				Massaron et Massarou

Le questionnaire B débutait par :

Massaron et Massarou				Massarou et Massaron
Lémotan et Lémoto				Lémoto et Lémotan

Et se terminait par :

Amassan et Amassa				Amassa et Amassan
Ako et Akin				Akin et Ako

Nous avons proposé nos enquêtes à un nombre conséquent d'interviewés pour qu'une validation statistique soit possible.

Les groupes de locuteurs témoins étaient composés de 50 à 105 individus.

Sur les bases des observations relevées lors du questionnaire d'enquête préliminaire ÉLISE et ZOÉ que nous développerons en suivant en 2.2.3 (page 56), nous avons construit un patron de questionnaires commun aux questionnaires d'enquête ÉRIC ET AURIC, PAUL ET PASCAL, ÉDOUARD ET EDMOND, NAROD ET NANOR et LEMOTO ET LEMOTAN.

D'une manière générale, ces cinq questionnaires répondent tous à deux exigences communes :

- Une exigence de représentativité : chaque expérience ne teste qu'une seule hypothèse à la fois ;
- Une exigence de neutralisation. Pour chacun des appariements testés, nous avons veillé à mettre en place une neutralisation culturelle, sémantique et syntaxique.

Pour ce faire, nous avons testé des binômes composés de prénoms non mixtes, ou de logatomes présentés comme des « prénoms étrangers », hors contexte et hors figement culturel.

En effet, afin d'éviter l'empreinte de « figement » ou de « collocation » culturels, du type « ROMÉO ET JULIETTE », « TOM ET JERRY » ou « PAUL ET VIRGINIE », nous avons systématiquement recherché au préalable les appariements testés sur Google, sous leurs deux formes concurrentielles, en évacuant en amont les paires sur lesquelles nous nourrissions quelques hésitations. Par ailleurs, nous méfiant de l'usage et des conventions sociales, nous n'avons proposé que des binômes non mixtes (« masculin et masculin », « féminin et féminin ») afin de nous affranchir des usages coutumiers.

Au travers de ces cinq enquêtes, nous avons surtout souhaité neutraliser les composés coordinatifs syndétiques un à un afin que le seul critère de variabilité entre les deux composants d'un binôme soit l'hypothèse éprouvée par l'enquête en place.

Enfin, nous avons également pris soin de poser la question filtre suivante :

« L'un de vos proches porte-t-il l'un de ces prénoms ? : OUI NON
si oui, le(s)quel(s) »

Le degré de proximité étant laissé à la libre appréciation des locuteurs témoins.

Nous avons retiré avant analyse toutes les paires contenant au moins l'un des prénoms cités, afin de nous dédouaner des contraintes psycho-affectives. Nous développerons ce point page 83.

Les interviewés étaient interrogés en même temps, en un même lieu, une salle de cours.

Ils répondaient à la question suivante : « Pour chaque ligne, sélectionnez les paires qui vous semblent les plus naturelles. ».

Les feuilles d'enquêtes sont proposées en annexe.

2.2.3. Protocole de constitution du questionnaire d'enquête ÉLISE et ZOÉ

Le questionnaire d'enquête ÉLISE ET ZOÉ (présenté en annexe) est un questionnaire composé de 43 questions dichotomiques. Il fut soumis le 10 novembre 2010 en salle de cours à 105 étudiants en première année de licence.

Ont été soumises les deux formes coordonnées par « et » des binômes ci-dessous :

AMÉLIE / ILÉNA	ILÉNA / ÉMILIE	ROMÉO / JULIETTE
ANNA / EMMA	JADE / TOM	ROMÉO / TOM
AUDREY / ENZO	JEANNE / SERGE	ROMÉO / VALENTIN
AUDREY / HUGO	LAURENCE / GUILLAUME	STÉPHANE / BAPTISTE
CHARLOTTE / BAPTISTE	LOÏC / CLÉMENCE	STÉPHANE / CLÉMENCE
CHLOÉ / ZOÉ	LOU / TOM	STÉPHANE / LAURENCE
CLÉMENCE / ALEX	LOUIS / LOU	STÉPHANE / STÉPHANIE
ÉLISA / ZOÉ	LOUISON / LOUISA	THOMAS / ÉLISA
ÉLISE / LISA	LUC / CHLOÉ	THOMAS / LISA
ÉLISE / ZOÉ	LUC / JULES	TOM / THOMAS
ENZO / EMMA	MATHILDE / DYLAN	VALENTIN / VICTOR
HÉLÉNA / ANNA	MATHIS / MAXIME	VICTOR / ALEX
HÉLÉNA / HÉLÈNE	MAXIME / MARINE	VICTOR / BAPTISTE
HÉLÉNA / NICOLAS	PAUL / LUC	ZOÉ / LISA
HÉLÈNE / NICOLAS		

Sur ces 43 appariements de prénoms coordonnés, 22 sont des appariements mixtes et 21 des appariements non mixtes répartis comme suit : 11 binômes de prénoms féminins et 10 binômes de prénoms masculins.

Les hypothèses testées lors de ce questionnaire étaient au nombre de quatre, deux d'entre elles cherchaient à valider des intuitions extraphonologiques, les deux autres des hypothèses phonologiques.

Concernant les hypothèses extraphonologiques, évoquées page 40, nous avons eu très tôt le pressentiment de l'existence d'une « contrainte culturelle » active qui fixerait certains appariements binomiaux dans un ordre coordonné plutôt que sa forme contraire. Nous avons donc souhaité mettre à l'épreuve cette intuition en soumettant volontairement à notre échantillon six paires culturellement marquées. En effet, les six paires de binômes ci-dessous renvoient tous à des paires de héros fictifs, plus ou moins célèbres, pour lesquelles la forme coordonnée par « et » a été fixée par le titre de l'œuvre dont ils sont issus. Nous pensons que plus l'œuvre est célèbre, plus le binôme est marqué.

Ainsi étaient proposés, du plus marqué au moins marqué : « ROMÉO ET JULIETTE »,

« JEANNE ET SERGE ²⁷ », « TOM ET THOMAS ²⁸ », « HÉLÈNE ET NICOLAS ²⁹ », « LOU ET LOUIS ³⁰ » et « TOM ET LOU ³¹ ».

La seconde hypothèse extraphonologique que nous avons testée dans ce questionnaire d'enquête est à mettre en lien avec le postulat de départ de l'article de Wright, Hay et Bent (2005) que nous commentons page 18 : l'idée d'une « conspiration » du genre, à savoir que la sélection d'une position de la coordonnée plutôt que l'autre pourrait être imputée au genre porté par les prénoms. La question posée est : est-ce que le genre des prénoms a une incidence sur le choix d'une forme coordonnée ?

Pour cela, nous avons proposé aux 105 locuteurs interrogés dans le cadre de cette expérience 20 appariements binomiaux mixtes dont « MARINE / MAXIME », « LOUIS / LOU » et « CHARLOTTE / BAPTISTE », pour ne présenter que quelques exemples.

En ce qui concerne les deux hypothèses phonologiques que nous désirions tester, elles sont toutes deux en lien avec les frontières vocaliques des prénoms.

La première cherche à vérifier s'il existe une tendance à éviter la séquence /e e/ que formerait la finale de M1, ou l'initiale de M2, juxtaposée à la conjonction de coordination « et ».

Afin de tester cette hypothèse, nous avons proposé des appariements dont l'un des prénoms soit commence par /e/ comme « ÉLISE », soit se termine par /e/ comme « CHLOÉ ». Nous avons soumis les dix paires suivantes :

La deuxième hypothèse phonologique est à associer à l'hypothèse étudiée grâce au questionnaire d'enquête PAUL ET PASCAL liant la taille du mot à sa place dans « M1&M2 ». Dans le présent questionnaire, nous souhaitons examiner le statut des voyelles initiales, à savoir : si l'on exclut du décompte métrique les voyelles initiales, les mots dont le patron phonologique le permet se comportent-ils comme des mots plus courts ou conservent-ils leurs caractéristiques premières en ce qui concerne la sélection d'un ordre préférentiel de coordination ?

27 « Jeanne et Serge » sont les héros de la série animée éponyme créée par Shizuo Koizumi, diffusée pour la première fois en France en 1987 et ayant fait l'objet de nombreuses rediffusions jusqu'en 2007.

28 « Tom et Thomas » est le titre d'un film de Esmé Lammers datant de 2002.

29 « Hélène et Nicolas » sont les héros d'une série télévisée française « Hélène et les garçons » créée en 1992 ayant fait l'objet de nombreuses séries dérivées (« Le miracle de l'amour », « Les vacances de l'amour », « Les mystères de l'amour ») encore diffusées à ce jour.

30 « Lou et Louis » est le titre d'une chanson de 1996 chantée par Maurane.

31 « Tom et Lou » sont les héros d'un livre jeunesse de Sophie Faudais : « Tom et Lou, les jumeaux à la maternité » publié en janvier 2007

Pour répondre à cette interrogation, nous avons construit les huit binômes suivants :

« THOMAS / ÉLISA »,	« ÉLISE / ZOÉ »,	« AUDREY / ENZO »,
« THOMAS / LISA »,	« ÉLISE / LISA »,	et « HÉLÈNE / NICOLAS ».
« HÉLÉNA / HÉLÈNE »,	« ENZO / EMMA »,	

On notera que certains appariements permettent de tester deux des hypothèses présentées (« THOMAS / ÉLISA », « ÉLISE / ZOÉ », « ÉLISE / LISA », « HÉLÈNE / NICOLAS »).

Concernant les 13 derniers binômes présentés ci-dessous, ils n'ont pas été utilisés à proprement dit pour tester l'une des quatre hypothèses que nous venons d'évoquer. Il s'agit d'appariements témoins qui nous ont permis d'observer et d'étalonner nos résultats de manière plus précise, pour les binômes proposés ci-dessus.

AMÉLIE / ILÉNA,	MATHIS / MAXIME	STÉPHANE / BAPTISTE
ANNA / EMMA	PAUL / LUC	VALENTIN / VICTOR
HÉLÉNA / ANNA	ROMÉO / TOM	VICTOR / ALEX
LOUISON / LOUISA	ROMÉO / VALENTIN	VICTOR / BAPTISTE
LUC / JULES		

Pour exemple, « AMÉLIE / ILÉNA » répond au binôme « ILÉNA / ÉMILIE », « ANNA / EMMA » à « ENZO / EMMA » et à « HÉLÉNA / HÉLÈNE » et « MAXIME / MATHIS » étalonne le binôme « MAXIME / MARINE ».

Il convient de préciser que nous considérons ce questionnaire d'enquête comme un questionnaire « pilote », dans le sens où il nous a amenée à prendre de nouvelles normes méthodologiques pour les questionnaires d'enquête suivants.

Nous avons en premier lieu révisé la longueur du questionnaire en tentant de le réduire au moins de moitié. En effet ÉLISE ET ZOÉ est un questionnaire relativement long (43 appariements) et nous avons craint que la répétition d'un trop grand nombre de questions dichotomiques standardisées lasse les interviewés.

D'autre part, nous avons décidé de neutraliser au maximum les prochains binômes soumis à interrogation afin de n'expérimenter qu'une seule et unique hypothèse à la fois. En effet, si nous prenons l'exemple du binôme « HÉLÈNE / NICOLAS », il peut relever des quatre hypothèses testées :

- L'influence du genre liée aux contraintes culturelles ;
- L'évitement de /e e/ ;
- Le statut des segments vocaliques à l'initiale (et en frontière de mots, celui de la finale consonantique) ;
- Et enfin, le poids de la fréquence des prénoms.

La difficulté est bien évidemment de déterminer, dans ces conditions, le facteur gouvernant le plus expressément l'ordre de coordination majoritairement privilégié par notre échantillon.

2.2.4. Protocole de constitution du questionnaire d'enquête ÉRIC ET AURIC

Le questionnaire d'enquête ÉRIC ET AURIC (présenté en annexe) est un questionnaire composé de 11 questions dichotomiques. Il fut soumis le 12 novembre 2011 en salle de cours à 50 étudiants en première année de licence. Comme nous l'indiquons dans le paragraphe précédent, nous avons tenté de neutraliser au maximum nos appariements afin de ne tester qu'une seule hypothèse à la fois. Dans le cadre du questionnaire d'enquête ÉRIC ET AURIC, nous avons cherché à examiner une hypothèse de fréquence qui lierait la fréquence d'un mot à sa position dans la coordination.

Pour ce faire, nous avons construit les 11 binômes suivants :

ALAIN / ODIN	CLEO / CHLOE	LOIC / LORIC
AMAEL / ANAEL	ELARA / HELANA	LUC / ZACH
AUDE / AUBE	EMILA / EMILIE	NOELIE / NOEMIE
AURIC / ERIC	JULIENNE / FABIENNE	

A la différence de Cooper et Ross (1975) présenté en 1.1.1.(page 15), nous avons essayé de nous extraire des biais syntaxiques et sémantiques en travaillant des assemblages de prénoms hors contexte. Ces prénoms étaient non mixtes pour éviter l'influence du « genre » dans la coordination. Facteur que nous avons préalablement éprouvé dans le questionnaire d'enquête ÉLISE ET ZOÉ.

Ainsi, pour ne tester que la fréquence, les appariements proposés étaient donc composés de deux prénoms équisyllabiques et possédaient les mêmes types d'initiales et de finales. Seule la fréquence différait.

Nous avons sélectionné des prénoms très fréquents (Alain, Émilie, Fabienne...) et des

prénoms peu fréquents (Odin, Emila, Julienne...).

Pour exemple, nous avons testé le binôme « ALAIN / ODIN ». Les deux prénoms sont des prénoms masculins, tous deux des termes dissyllabiques, commençant tous les deux par une voyelle orale et se terminant par une voyelle nasale. Seule leur fréquence diffère significativement : selon l'INSEE³², Alain comptait 452 409 individus en France début 2010 et Odin, 24.

2.2.5. Protocole de constitution du questionnaire d'enquête PAUL ET PASCAL

Le questionnaire d'enquête PAUL ET PASCAL (présenté en annexe) est un questionnaire composé de 9 questions dichotomiques. Il fut soumis le 12 décembre 2011 en salle de cours à 50 étudiants en première année de licence. Dans le cadre du questionnaire d'enquête PAUL ET PASCAL, nous avons cherché à questionner l'existence d'un facteur liant la taille des lexèmes à la sélection d'une position privilégiée dans la coordination binaire par « et ».

Pour ce faire, nous avons construit les 9 binômes suivants :

AGATHA / ASSIA	LOU / LILOU	PASCAL / PAUL
KIM / CORALINE	LUC / LOIC	SAM / SALEM
LINE / LAURINE	LUDOVIC / LUC	VIC / SOAZIC

Afin de ne tester que l'influence de la taille des lexèmes, nous n'avons étudié que des appariements neutralisés dans lesquels les lexèmes appariés sont non équisyllabiques. Nous avons donc sélectionné des prénoms monosyllabiques (Paul, Luc, Sam...) et des prénoms polysyllabiques (Pascal, Ludovic, Salem...).

Pour exemple, nous avons testé le binôme « LOU / LILOU ».

Il s'agit de deux prénoms féminins, commençant par la même consonne initiale /l/ et se terminant par la même voyelle /u/. « LOU » et « LILOU » ont approximativement de la même fréquence : selon les données INSEE dont nous disposons, 24 993 personnes vivaient en France début 2010 portait le prénom de « LOU », 24 610 celui de « LILOU ». La seule variable est leur longueur syllabique : « LOU » est un prénom monosyllabique, « LILOU » un prénom dissyllabique.

32 Ces données ont été obtenues sur <http://meilleursprenoms.com/> grâce au concours précieux de Stéphanie Rapoport

2.2.6. Protocole de constitution du questionnaire d'enquête ÉDOUARD ET EDMOND

Le questionnaire d'enquête ÉDOUARD ET EDMOND (présenté en annexe) est un questionnaire composé de 24 questions dichotomiques. Il fut soumis le 12 décembre 2012 en salle de cours à 94 étudiants en première année de licence.

Dans le cadre du questionnaire d'enquête ÉDOUARD ET EDMOND, nous avons cherché à examiner la possible action des finales consonantiques en lien avec un procédé de resyllabation avec la conjonction de coordination (nous développerons plus précisément ce point page 238) dans la sélection d'un ordre de composés binomiales par « et ».

Pour ce faire, nous avons construit les 24 binômes suivants :

ADEL / ANIS	ELYA / ELIETTE	LOIC / LUC
ALANE / ADILE	EUDORE / EUTROPE	MAXO / MAXEL
ALANE / AGNAN	FLORIDE / FLORICE	MERIL / MERIC
AMARINE / AMARANDE	FLORIMON / FLORIBERT	MERINE / MERYL
ANATOLINE / ANATOLIE	FRANCESCO / FRANCISQUE	NELE / NEIGE
APOLLINAIRE / APOLINARIE	GRACIELLE / GRACIETTE	PASCAL / PAUL
EDMOND / EDOUARD	LEONINE / LEONIDE	THEODULE / THEOTIME
ELI / ELLIS	LILIANE / ELIANE	TIMON / THIBALT

Afin de ne tester que le facteur concernant les finales consonantiques, nous avons observé 10 appariements mettant en concurrence un lexème ayant une finale consonantique avec un lexème ayant une finale non consonantique ($_[\pm \text{ CONSONANTIQUE}]\#$ et $_[- \pm \text{ CONSONANTIQUE}]\#$) et 11 appariements mettant en concurrence deux prénoms ayant des finales consonantiques différentes. Trois binômes supplémentaires « PAUL / PASCAL », « LUC / LOÏC » ET « LILIANE / ELIANE » ont été retestés comme appariements témoins. Ils avaient été préalablement soumis lors du questionnaire d'enquête PAUL ET PASCAL.

Nous avons donc sélectionné des prénoms rares ayant une syllabe finale fermée (MAXEL / APOLLINAIRE...) et des prénoms rares ayant une syllabe finale ouverte (MAXO / APOLINARIE...).

Pour exemple, nous avons testé le binôme « Elya / Elyette ».

Il s'agit de deux prénoms féminins commençant par la même initiale vocalique /e/ et ayant, dans la composition de leur syllabe finale, la même suite /lj/. Elya et Elyette ont approximativement la même fréquence : selon les données INSEE dont nous disposons, 698 personnes en France début 2010 portaient le prénom de Elya, 369, celui de Elyette. La seule

variable est le type de segment final: Elya se termine par une finale vocalique, et Elyette, une finale consonantique.

De la même façon, nous avons proposé les deux formes concurrentielles coordonnées de « Léonine / Léonide ».

Il s'agit de deux prénoms féminins trisyllabiques commençant par la même syllabe initiale /le/ et partageant la même attaque et le même noyau de leur syllabe finale : /ni/. Selon les données INSEE, début 2010, 13 individus en France portait le prénom Léonine, 84, celui de Léonide. La seule variable entre ces deux prénoms est la nature de leur coda finale : les deux prénoms se terminent par des consonnes sonores apicales, Léonine par une consonne nasale, Léonide par une consonne occlusive.

2.2.7. Protocole de constitution du questionnaire d'enquête NAROD ET NANOR

Le questionnaire d'enquête NAROD ET NANOR (présenté en annexe) est un questionnaire composé de 28 questions fermées. Il fut soumis le 21 janvier 2013 en salle de cours à 53 étudiants en première année de licence.

Nous avons mis en place de nouvelles normes méthodologiques en ne proposant plus de questions dichotomiques. En effet, étant donné les facteurs variables dans les deux prochains questionnaires d'enquête à venir (l'enquête NAROD ET NANOR et l'enquête LEMOTO ET LEMOTAN), il nous a semblé qu'une question dichotomique ne laissait pas la possibilité à nos interviewés de faire part de leurs doutes. Ainsi, nous avons pris le parti de poser une question fermée en mettant à la disposition de notre échantillon trois items différents : « M1 ET M2 », « M2 ET M1 » et « NSP » (Ne Se Prononce pas).

De plus, compte tenu des difficultés que nous rencontrons dans la création d'appariements pertinents neutralisant au maximum l'ensemble des paramètres agissant dans la coordination binomiale par « et », exceptions faites des hypothèse examinées, nous avons fait le choix de ne plus utiliser des prénoms mais des pseudo-prénoms que nous avons introduits auprès notre échantillon comme étant des prénoms étrangers. Dans le cadre du questionnaire d'enquête NAROD ET NANOR, nous avons présenté nos appariements comme des binômes de prénoms arméniens. Dans les faits, la plupart des pseudo-prénoms proposés sont très librement inspirés de prénoms ou de termes arméniens (et japonais en ce qui concerne le questionnaire d'enquête LEMOTO ET LEMOTAN). Nous avons d'ailleurs intégré quelques « véritables » prénoms dans les listes que nous soumettions afin de rendre notre propos plus

vraisemblable. Dans le but d'éviter les hésitations liées à la prononciation de ces pseudo-prénoms, nous informions en début de questionnaire le fait que l'orthographe avait été francisée et que chaque lettre devait être prononcée. Et afin de lever le moindre doute, nous avons lu l'ensemble des appariements proposés³³.

L'hypothèse évaluée par le questionnaire d'enquête NAROD ET NANOR est en lien avec le questionnaire d'enquête ÉDOUARD ET EDMOND. Nous avons cherché à observer les finales consonantiques des pseudo-mots testés.

Pour ce faire, nous avons construit les 28 binômes suivants :

ANTRANIG / ANDRANIK	AVÉDIG / AVÉTIK	LOUSSVART / LOUSSNART
ARARAD / ARARAT	AVÉTIK / AVÉTISS	MANNIG / MARAL
ARÈG / AVÈV	AVÈV / ARÈK	MARAL / MARO
ARÈK / ARÈG	AZAT / AZAD	MARO / MANNIG
ARÈVCHAT / ARÈVCHAD	BARSAME / BARSÈG	NANÉ / NANOR
ARMÉNOUCHE / ARMINÉ	CHIRAG / CHIRAZ	NAROD / NANOR
ARPINÉE / ARPÉNIG	ENDZAG / ENTZAK	SÉVANE / SÉDA
ARSÈNE / ARTÈME	KARPIG / KARPISS	YÉRAME / YÉZAZ
ASBÈB / ASPÈT	LOUSSINE / LOUSSIG	ZÉPOU / ZÉVART
AVAG / AVAK		

Afin de ne tester que l'influence des finales consonantiques des segments testés, nous avons étudié des appariements neutralisés pour lesquels nous faisons varier, pour l'essentiel des paires présentées, la qualité des consonnes finales, à savoir leur lieu d'articulation, leur mode articulaire, leur voisement et aussi leur organe articulaire, exceptions faites de cinq paires reprenant l'hypothèse vérifiée dans le questionnaire d'enquête ÉDOUARD ET EDMOND, à savoir l'opposition entre syllabe finale ouverte et syllabe finale fermée.

Pour exemple, nous avons testé le binôme *Azat / Azad*.

Du fait de l'utilisation de pseudo-mots, le genre et la fréquence des termes sont neutralisés, ainsi que les problèmes d'ordre sémantique ou syntaxique. *Azat / Azad* sont tous deux des segments dissyllabiques, ayant pour seule variation la qualité de leur coda finale du point de vue de leur voisement. *Azat* se termine par une occlusive alvéodentale sourde, *Azad* par une occlusive alvéodentale sonore.

³³ Pour exemple :

/ãtɾanig/ /ãtɾanik/, /ararad/ /ararat/, /arɛg/ /avɛv/, /arɛk/ /arɛg/
 /arɛvʃat/ /arɛvʃad/, /ar menuʃ/ /ar mine/, /ar pine/ /ar penig/
 /arɛn/ /artɛm/, /asbɛb/ /aspɛt/, /avag/ /avak/, /avedig/ /avetik/,
 /avetik/ /avetis/,...

De la même façon, considérons le binôme Arèg / Avèv.

Il est également neutralisé, seules les codas finales varient selon leur mode articuloire et leur lieu d'articulation : Arèg se termine par une consonne occlusive vélaire sonore, Avèv par une consonne fricative labiodentale sonore.

2.2.8. Protocole de constitution du questionnaire d'enquête LEMOTO ET LEMOTAN

Le questionnaire d'enquête LEMOTO ET LEMOTAN (présenté en annexe) est un questionnaire composé de 25 questions fermées. Il fut soumis le 18 décembre 2013 en salle de cours à 86 étudiants en première année de licence.

Dans le cadre du questionnaire d'enquête LEMOTO ET LEMOTAN, nous avons cherché à examiner un facteur liant la qualité des segments finaux préalablement testé dans les questionnaires d'enquête ÉDOUARD ET EDMOND et NAROD ET NANOR, à savoir l'hypothèse de l'existence d'une contrainte active dans la sélection d'un ordre préférentiel des composants liant la qualité du segment vocalique final à une position dans le segment binaire coordonné.

Une nouvelle fois, nous avons présenté des pseudo-prénoms que nous avons introduits auprès de notre échantillon comme étant des prénoms japonais.

Pour ce faire, nous avons construit les 25 binômes suivants :

AKIN / AKO	DANZO / DANZON	HAROUTO / HAROUTAN
AMASSA / AMASSAN	DENBE / DENBON	HAZE / HAZIN
AMOU / AMON	EIKO / EIKON	HIDEKO / HIDEKI
ASAFOUMI / ASAFOUMON	EMIKIN / EMIKO	KAGEHAN / KAGEHO
BYAKOUREN / BYAKOUYA	FOUYOUKA / FOUYOUKI	KAHO / KAHIN
CHIEKA / CHIEKO	FUSAHE / FUSUHAN	KOTAN / KOTO
CHIHARIN / CHIHARU	GIKAN / GIKO	LEMOTO / LEMOTAN
DAIGEN / DAIGO	HANAHE / HANAHI	MASSAROU / MASSARON
DAISSAKOU / DAISSAKAN		

Pour éviter toutes confusions inutiles, nous informions nos interviewés que l'orthographe avait été francisée, que chaque lettre devait être prononcée. De plus, nous avons lu l'ensemble des trois items soumis au test (« M1&M2 », « M2&M1 » et NSP) avant le début de la passation. La lecture du test a été particulièrement insistante et appuyée dans le cadre de cette expérimentation de manière à éviter toute confusion possible entre un segment « voyelle +

consonne nasale » et un segment « voyelle nasale ».

Afin de ne tester que l'influence des finales vocaliques des segments testés, nous n'avons étudié que des appariements neutralisés pour lesquels nous faisons varier la qualité des voyelles finales.

Pour exemple, nous avons testé le binôme *Amassa /Amassan*.

A nouveau, du fait de l'utilisation de pseudo-mots, le genre, la fréquence des termes, ainsi que les problèmes d'ordre sémantique ou syntaxique, sont neutralisés. *Amassa /Amassan* sont tous deux des segments trissyllabiques, ayant pour seule variation la qualité de leur voyelle finale. *Amassa* se termine par une voyelle orale, *Amassan* par une voyelle nasale.

De la même façon, considérons le binôme *Chieka / Chieko* :

Le binôme est également neutralisé, seule la hauteur des voyelles finales varie, *Chieka* se terminant par une voyelle plus basse que *Chieko*.

2.2.9. Présentation récapitulative des questionnaires d'enquête

Au travers de cet ensemble de questionnaires d'enquête que nous venons de présenter, nous avons souhaité neutraliser les composés coordinatifs syndétiques afin que le seul critère de variabilité entre les deux composants d'un appariement soit le facteur éprouvé par le questionnaire d'enquête en place recueillant ainsi 138 nouveaux binômes. Pour chacun d'entre eux, nous avons relevé la paire préférentielle c'est-à-dire celle ayant été majoritairement sélectionnée par nos interviewés. Puis, à nouveau, comme nous l'avons fait pour les corpus, nous avons étendu un test de χ^2 à l'ensemble des formes coordonnées binomiales précédemment sélectionnées afin de ne conserver que celles éprouvées statistiquement, réduisant de plus de moitié les 138 composés coordinatifs initialement recueillis.

Ainsi, comme présentés dans le tableau récapitulatif ci-dessous, l'ensemble des 6 questionnaires d'enquête est composé de 60 paires réparties comme suit :

QUESTIONNAIRES D'ENQUÊTE	FACTEURS TESTÉS	NOMBRE
ÉLISE ET ZOÉ	Facteur culturel	20
	Genre	
	Évitement de /e e/	
	Statut de la voyelle initiale	
ÉRIC ET AURIC	Fréquence	8
ÉDOUARD ET EDMOND	_ $[\alpha \text{ cons}]$ \# et _ $[-\alpha \text{ cons}]$ \#	12
NAROD ET NANOR	Type de finales consonantiques	7
LEMOTO ET LEMOTAN	Timbre du segment vocalique final	5
PAUL ET PASCAL	Taille	8
TOTAL		60

Tableau 5 : Collection des questionnaires d'enquête

Le travail de recueil de données que nous achevons ici s'est révélé long et fastidieux.

Les corpus que nous présentions en 2.1. (à partir de la page 34) supposent un très gros travail de tri et de traitement des résultats, *a posteriori*, du fait de la pléthore de données mises à notre disposition sur Internet.

En revanche, la mise en place des questionnaires d'enquête nécessitait la préconception (*a priori*) de contraintes fortes ainsi que l'instauration de normes protocolaires strictes afin de nous autoriser à révéler, pour chacun d'entre eux, un unique facteur, en éliminant au maximum tout biais qu'il soit phonologique, morphologique, sémantique, syntaxique, psychologique ou sociologique.

Quoi qu'il en soit, nous pensons que ces deux types de données constituées à partir de principes et de méthodes totalement différents sont complémentaires, et offrent une base empirique et variée recouvrant un ensemble représentatif des formes binomiales coordonnées directement par « et ».

2.3. Exploitation des données

Une fois franchi le travail de compilation que nous venons de présenter, nous avons mis en place l'étiquetage et la classification des données retenues. Pour ce faire, nous avons établi un certain nombre de critères qui constituent la trame de l'étiquetage proposé en suivant.

Ces critères sont de diverses origines.

Nombre d'entre eux sont directement inspirés de la littérature scientifique existante que nous présentons page 14. D'autres sont nés de notre intuition de locutrice native du français. Enfin, d'autres trouvent leur source dans la construction empirique des questionnaires d'enquête. Quoi qu'il en soit, tous questionnent les facteurs gouvernant la sélection d'un ordre préférentiel de coordination binaire directe par « et » en français.

Cet étiquetage est fondamental dans notre travail de thèse. Il délimite nos champs d'investigation. Chaque critère constitue un champs d'exploration préalable que nous nous appliquerons à démontrer statistiquement dans les sections à venir.

2.3.1. Remarques générales

La collection des corpus et des questionnaires d'enquête décrite ci-dessus nous permet d'avoir à notre disposition 1 946 formes binomiales coordonnées par « et ».

Chaque composé syndétique par « et » collecté ou expérimenté a fait l'objet d'un étiquetage numérisé sous format Excel.

A titre d'illustration, voici ci-après une reproduction des tableaux ainsi obtenus pour trois appariements, « ANAÏS / ADRIEN », « VITE / BIEN » ET « US / COUTUMES ». Les étiquettes classificatoires que nous avons retenues afin de permettre une l'exploitation raisonnée de nos données sont exposées en détail en 2.3.2. Critères classificatoires.

La reproduction du fichier Excel comportant l'ensemble de nos données ne sera pas fournie dans la version papier de cette thèse, compte tenu de la taille et de la lourdeur du fichier.

Il est cependant librement accessible en ligne.

Une fois la numérisation et l'étiquetage des 1 946 appariements mises en place, nous avons traité automatiquement et statistiquement, l'ensemble des contraintes phonologiques et extraphonologiques qui agissent sur la coordination en français. Pour ce faire, nous avons à nouveau utilisé comme test d'ajustement un test de χ^2 , la taille de nos échantillons s'y prêtant. Cependant, si dans la sélection des paires dont l'ordre de coordination est privilégié, nous avons été restrictive avec un risque d'erreur de 5%, nous l'avons été un peu moins dans l'examen des facteurs à suivre ($p \leq 0,1$).

A cela, deux raisons : la première concerne le recueil des formes de coordination privilégiées. En effet, afin d'avoir la possibilité de faire émerger des facteurs actifs dans la sélection d'un ordre coordonné, il nous a semblé pertinent de recueillir des paires statistiquement remarquables pour lesquelles la répartition des effectifs en « x et y » ou « y et x » fait l'objet d'un déséquilibre significatif.

Concernant l'examen des facteurs phonologiques et extraphonologiques, l'enjeu est différent. Nous considérons que l'ordre des coordonnées fait l'objet d'un jeu d'interactions de contraintes. Dès lors, nous envisageons que les facteurs en concurrence n'ont, ni le même rôle, ni la même force. C'est pourquoi nous retenons des tendances statistiques dans le but de faire émerger des facteurs peut être moins évidents mais néanmoins actifs dans la sélection d'un ordre préférentiel d'un binôme par «et ».

Dans le paragraphe à venir, nous présenterons l'ensemble des critères de classement utilisés pour l'examen des facteurs actifs dans la sélection d'un ordre préférentiel de coordination ainsi que les résultats de ce traitement statistique.

2.3.2. Critères classificatoires

L'ensemble des paires relevées et statistiquement validées ont été organisées en répondant aux vingt-neuf critères classificatoires suivants :

1. L'ordre alphabétique est-il respecté ? M1 commence t-il alphabétiquement avant M2 ?
Ou inversement ? ($A \rightarrow Z, Z \rightarrow A$)
2. La fréquence des mots est-elle en jeu ? M1 est-il statistiquement plus fréquent que M2 ? Ou inversement ?

3. Dans le cadre de coordination genrée notamment dans le corpus « Tom et Zoé », le genre a-t-il une incidence sur la sélection d'un ordre préférentiel de coordination ? M1 est-il plutôt masculin ? Plutôt féminin ? Qu'en est-il de M2 ?
4. Combien de syllabes M1 et M2 contiennent-ils ? M1 en contient-il plus que M2 ? Ou inversement ?
5. De combien de syllabes s'agit-il ? Une, deux, trois, quatre et plus ?
6. La voyelle initiale entre-t-elle dans le décompte des syllabes de M ? A combien tombe le nombre de syllabes de M1 ou de M2, en extrayant du décompte syllabique l'ensemble des segments #V_ ?
7. Avec ce nouveau décompte, M1 est-il plus court, plus long ou de même longueur syllabique que M2 ?
8. Quelle est la catégorie grammaticale de M1 ? De M2 ?
9. Quel est le patron syllabique de la première syllabe de M1 ? De M2 ?
10. Les syllabes initiales de M1 ont-elles le même poids que celles de M2 ?
11. Quel est le patron syllabique de la dernière syllabe de M1 ? De M2 ?
12. Les syllabes finales de M1 ont-elles le même poids que celles de M2 ?
13. Quel est le noyau de la première syllabe de M1 ? De M2 ?
14. Quel est le noyau de la dernière syllabe de M1 ? De M2 ?
15. Quelle est la nature du segment initial de M1 ? M2 ?
16. Si ce segment est une voyelle, quelle est-elle ?
17. Quelle est son ouverture ? Sa hauteur ? Est-elle écartée ? Arrondie ? Est-ce une voyelle orale ? Nasale ? Est-ce un /e/ ?
18. Si le segment initial de M1 ou de M2 est une consonne, quelle est-elle ?
19. Quel est son voisement ? Est-ce une consonne nasale ? Arrière ? Arrondie ? Est-ce un cluster consonantique, ou une consonne simple ?
20. Le segment initial de M1 ou de M2 est-il un glide ?
21. Si le segment initial est un glide, est-il arrière ? Arrondi ?
22. Quelle est la nature du segment final de M1 ? De M2 ?
23. Est-ce une voyelle ? De quel type ?
24. Est-ce une consonne ? De quel type ? Est-ce un cluster ?
25. Est-ce un glide ? De quel type ?
26. Le segment initial de M1, quel qu'il soit, interagit-il en réponse avec le segment initial de M2 ?
27. Le segment final de M1, quel qu'il soit, interagit-il en réponse avec le segment final de

M2 ?

28. Pour tous les appariements autres que les prénoms, des traits sémantiques sont-ils en jeu ?

29. Quels sont ces traits : [PROXIMITÉ SPATIALE] ? [PROXIMITÉ TEMPORELLE] ? [PROXIMITÉ PSYCHOLOGIQUE] ? [SUPÉRIORITÉ SPATIALE] ? [SUPÉRIORITÉ HIÉRARCHIQUE] ? [ANTÉRIORITÉ TEMPORELLE] ?

Une fois l'étiquetage effectué, nous avons traité statistiquement la réponse de chaque question soulevée par ces hypothèses, en les organisant selon deux grands axes décrits ci-après : les interrogations relevant de critères extraphonologiques et celles relevant de critères phonologiques.

Chapitre 3 : Facteurs extraphonologiques gouvernant les appariements

Même si notre visée principale est d'étudier les contraintes phonologiques actives dans la coordination de deux mots par « et » en français, travailler sur un phénomène de langue ne nous permet pas de nous émanciper d'autres types de contraintes. Ainsi, en proposant une série de critères extraphonologiques, nous avons cherché à démontrer dans quelle mesure ces derniers peuvent avoir une incidence dans la sélection d'un ordre préférentiel de coordonnés syndétiques. Les critères extraphonologiques que nous avons pris en considération sont l'ordre alphabétique, le facteur syntaxique, le facteur sémantique, le facteur psycho-affectif, la fréquence des mots et le facteur culturel.

3.1. L'ordre alphabétique

Il s'agit ici de déterminer si l'ordre alphabétique joue un rôle dans la sélection d'un ordre de coordination privilégié : l'ordre alphabétique est-il respecté ? M1 commence-t-il alphabétiquement avant M2 ? Ou inversement ?

La moitié de nos supports d'études est issue de sources écrites, et même si, comme nous l'indiquons page 34, ces sources Internet fonctionnent comme un écrit oralisé, nous pouvons néanmoins soupçonner l'influence de l'ordre alphabétique dans le choix d'un ordre coordonné, a fortiori en ce qui concerne le corpus TARN ET GARONNE. En effet, cette collection de toponymes coordonnés par « et » est le recueil d'un ensemble de villes et de villages antérieurement réunis, dont le nouveau nom coordonné par « et » marque la nouvelle identité. Il est donc tout à fait envisageable que la sélection de ce nouveau nom ait été faite par des facteurs extraphonologiques, qu'il s'agisse de la taille, de la renommée d'un toponyme ou tout simplement du respect de l'ordre alphabétique.

Par ailleurs, Benor et Levy (2006) le signalaient : « It is possible that binomial ordering is influenced by the alphabetic location of the first letter of each word. We expected that this would have more of an effect in the ordering of names, such as business partners or joint authors. » Nous pouvons donc supposer l'existence d'une corrélation entre l'ordre alphabétique et un ordre binomial privilégié.

Pour répondre à cette interrogation, nous avons relevé les six corpus composant nos supports d'études en mettant de côté les questionnaires d'enquête. En effet, ces derniers ont été

construits avec la volonté de n'interroger qu'une seule hypothèse à la fois. La question de l'ordre alphabétique avait donc été neutralisée en amont.

Nombre de paires		A→Z	Z→A
69	TARN ET GARONNE % TOTAL =	37 53,62	32 46,38
48	CUL ET CHEMISE % TOTAL =	25 52,08	23 47,92
62	LISSE ET SOYEUX % TOTAL =	32 51,61	30 48,39
1420	TOM ET ZOÉ % TOTAL =	726 51,13	694 48,87
202	VITE ET BIEN % TOTAL =	94 46,53	108 53,47
85	ARTS ET MÉTIERS % TOTAL =	52 61,18	33 38,82
1886	% TOTAL =	966 51,22	920 48,78

Tableau 6 : Pourcentage des paires respectant ou ne respectant pas l'ordre alphabétique dans la coordination

Dans le détail, concernant le corpus TARN ET GARONNE pour lequel nous suspicions particulièrement une incidence de l'ordre alphabétique dans la sélection d'un ordre préférentiel de coordination, il apparaît que 46,38 % des appariements ne respectent pas l'ordre alphabétique ($\chi^2(1) = 0,36$; $p \leq 0,55$), tout comme dans les corpus CUL ET CHEMISE (47,92 %, ($\chi^2(1) = 0,08$; $p \leq 0,77$), LISSE ET SOYEUX (48,39% ($\chi^2(1) = 0,06$; $p \leq 0,80$)), TOM ET ZOÉ (48,87 %, ($\chi^2(1) = 0,72$; $p \leq 0,40$) et VITE ET BIEN, (53,47 %, ($\chi^2(1) = 0,97$; $p \leq 0,32$)). Notons que dans ces cinq corpus, la répartition n'est pas significative. Nous ne pouvons affirmer l'existence d'une corrélation entre un ordre alphabétique croissant et la sélection d'un ordre préférentiel de coordination.

Le corpus ARTS ET MÉTIERS, quant à lui, fait apparaître un déséquilibre significatif : dans 61,18% des cas, l'ordre alphabétique est respecté ($\chi^2(1) = 4,25$; $p \leq 0,04$), il apparaît donc ici que le respect de l'ordre alphabétique influence la sélection des formes coordonnées.

Cependant, comme le montre le tableau ci-dessus, sur 1 886 paires testées, 966 respectent l'ordre alphabétique (51,22%), 920 ne le respectent pas (48,78%) ($\chi^2(1) = 1,12$; $p \leq 0,29$). La valeur asymptotique est supérieure à .01. La distribution n'est pas donc statistiquement

valide. Il nous est impossible d'extraire l'hypothèse nulle. Nous en concluons donc que pour l'ensemble de nos supports d'études, l'ordre alphabétique n'a pas d'incidence directe sur l'ordre de coordination.

3.2. Le facteur sériel

Dans ce paragraphe, il s'agit d'évaluer le poids du facteur sériel sur la sélection d'un ordre préférentiel de composés syndétiques par « et ». Nous entendons par « facteur sériel » un facteur qui permettrait de déduire l'ordre des binômes à partir d'un modèle existant créant des effets de séries.

Dans le cadre de l'étude de cette hypothèse, nous avons relevé 142 paires concernées par ce facteur dans l'ensemble de nos supports d'études, exception faite des questionnaires d'enquête et des corpus TOM ET ZOÉ et TARN et GARONNE. En effet, ces derniers recueillent des paires de noms propres dont la liste est fermée. Le corpus TOM ET ZOÉ apparie notamment les 100 prénoms les plus fréquemment donnés en 2008. Ainsi, même s'il existe des effets de séries pour ces derniers, nous pensons qu'il s'agit de « faux » étant donné qu'ils sont directement imputables à la façon dont nous avons construit ces supports d'études. Dans le tableau ci-dessous, nous présentons une illustration³⁴ de quelques paires tombant, selon nous, sous l'effet du facteur sériel que nous nous proposons de démontrer dans cette section.

THÈMES	PARADIGMES	
	M1	M2
ART(S)		<ul style="list-style-type: none"> • essai • décoration • gastronomie • manufactures • métiers • vie
COURT	<ul style="list-style-type: none"> • haut 	<ul style="list-style-type: none"> • concis • long • moyen

³⁴ Les 142 paires présentant un effet de séries sont compilées en annexe page Erreur : source de la référence non trouvée.

HYDRATE		<ul style="list-style-type: none"> • adoucit • nourrit • illumine • rafraichit
ICI		<ul style="list-style-type: none"> • là • maintenant • ailleurs

Tableau 7 : Exemples d'appariements sériels et position dans la coordination

Ainsi, pour exemple, le mot ART(S) fonctionne comme un thème qui est décliné dans nos supports d'études dans les paires suivantes : « art et essai », « art et décoration », « art et gastronomie », « art et manufactures », « arts et métiers » et « art et vie ». Notons que pour l'ensemble des paires dans lequel ART(S) est présent, il occupe systématiquement M1. C'est ce que nous nommons « facteur sériel ». ART(S) serait un thème présent en M1 et l'ensemble des déclinaisons proposées en M2 fonctionneraient comme des cases paradigmatiques.

Notons également que nous avons relevé ce comportement particulier dans des composés syndétiques de noms propres (ART(S)), d'adjectifs (COURT dans « court et moyen » et « court et long »), de verbes tensés (HYDRATE) et d'adverbes (ICI).

Nous avons donc relevé deux types de données de l'observation de ces 142 paires. Premièrement, le respect ou non de l'ordre induit par la série et, deuxièmement, la répartition des paradigmes en M1 et en M2. Nous présentons les résultats de ces observations dans le tableau ci-dessous :

FACTEUR SÉRIEL	+ SÉRIE	- SÉRIE	PARADIGMES	
			M1	M2
142	123	19	27	115
% TOTAL =	86,62	13,38	19,99	80,99

Tableau 8 : Pourcentage des paires sérielles respectant ou ne respectant pas l'ordre imposé par leur série, et position dans la coordination des paradigmes

Les 52 thèmes que nous avons relevés dans nos supports d'études respectent massivement des séries fixant un ordre comme modèle à suivre dans 86,62% des cas ($\chi^2(1) = 76,17$; $p \leq 0,01$). Il apparaît également que les paradigmes occupent très majoritairement M2 dans 80,99% des

142 appariements relevés ($\chi^2(1) = 54,54 ; p \leq 0,01$). Étant donné que le déséquilibre entre les répartitions présentées ci-dessus est statistiquement significatif, nous affirmons qu'un facteur de série induit un ordre préférentiel de coordination et que ce même facteur favorise les paradigmes en M2 et par conséquent le mot thème en M1.

Comment analyser de tels résultats ?

Selon nous, ces résultats étaient attendus. Depuis Saussure, il est communément entendu que la langue est un système. Il n'est pas surprenant de penser que des composés binomiaux se construisent en suivant un modèle préexistant ou saillant.

Cependant, comment expliquer la sélection du thème en M1 ? Ces 52 mots sont-ils de meilleurs candidats pour occuper cette position ? Et pour quelles raisons ? Est-ce lié à des facteurs purement sériels et automatiques faisant qu'on reproduit un ordre existant, ou pouvons-nous justifier leurs présences en M1 pour des raisons sémantiques en considérant notamment l'opposition Thème / Rhème ?

Est-il possible d'y voir un facteur morphologique ?

Nous avons volontairement utilisé dans ce paragraphe un vocabulaire lié aux travaux de la morphologie flexionnelle. Pouvons-nous entrevoir dans la sélection les déclinaisons paradigmatiques que nous venons de présenter un fonctionnement proche des cases flexionnelles ?

De plus, nous supposons l'existence d'effet de fréquences d'occurrences qui, comme le rappelle Bybee (2005), peuvent avoir un effet d'enracinement. « À force d'être utilisés, les items fréquents voient leur force lexicale augmenter et les items qui ont une grande force lexicale sont plus facilement accessibles que les autres et résistent mieux au changement analogique. » Le facteur sériel illustre, selon nous, ce phénomène. Plus la série s'installe, plus la fréquence d'occurrences est importante, plus le lien entre le thème et l'une des deux positions de la coordonnée devrait être fort.

Enfin, pour revenir à des caractérisations de nature plus phonologique, pourquoi ne pas supposer que ces effets de séries soient imputables à des facteurs phonologiques ? En effet, sur les 52 mots thèmes relevés, 23 sont des mots monosyllabiques et 14 commencent par une initiale vocalique. Est-il possible que ces différents facteurs phonologiques influent sur l'ordre

des coordonnés ? Nous développerons en détail ces points dans les paragraphes à venir et notamment dans le Chapitre 4 : Facteurs phonologiques gouvernant les appariements en prenant soin de prendre en compte l'influence notable du facteur sériel que nous avons relevé.

3.3. Le facteur syntaxique

Il s'agit ici d'évaluer le rôle que la syntaxe joue dans la sélection d'un ordre préférentiel de coordination binaire par « et ».

Comme nous l'évoquions en 1.1. Panorama de la littérature anglo-saxonne, il existe de très nombreux travaux portant sur les aspects syntaxiques et sémantiques de la coordination par « et » en français³⁵, notamment en ce qui concerne les coordinations binaires et n-aires³⁶. Ces travaux nous ont permis de délimiter notre champ de recherche en tentant d'exclure en amont tous facteurs syntaxiques actifs dans la sélection des composés syndétiques par « et ». C'est ainsi que nous entendons étudier les binômes directs coordonnés par « et ». En effet, comme le rappelle Hobaek Haff (1990) « ce qui caractérise les coordonnants et qui les distingue des subordinants, c'est cependant qu'ils ont pour rôle d'unir des éléments de même fonction ou de même statut syntaxique. Si les termes reliés ne se trouvent pas hiérarchiquement sur le même niveau, il n'y a pas coordination ».

M1 sera donc systématiquement de même catégorie syntaxique et de même niveau hiérarchique que M2.

En mettant en relation directe par « et » deux éléments syntaxiquement identiques « Nom et Nom », « Adjectif et Adjectif », « Verbe et Verbe » et « Adverbe et Adverbe », nous avons tenté de neutraliser au maximum les facteurs syntaxiques.

Nous demeurons néanmoins consciente que nous ne pouvons totalement nier l'existence de facteurs syntaxiques et discursifs.

En effet, bien que travaillant sur des appariements hors contexte, il est tout à fait envisageable que l'environnement contextuel influe sur l'ordre des mots. Nous pensons en particulier au thème du discours et à la reprise anaphorique. Pour illustrer nos propos, reprenons l'exemple initial de la paire « STÉPHANE ET GRAZIELLA » :

- « Graziella et Stéphane vont se marier, il n'a pas acheté son costume. »
- « Stéphane et Graziella vont se marier, elle n'a pas trouvé sa robe. »

35 Notons à nouveau ici les travaux d'Abeillé (2005), Mouret et al (2008) et Lambert (1991, 2001, 2006).

36 Les coordinations n-aires définissent des coordinations de plus de deux éléments.

Dans deux exemples, nous pouvons imputer la sélection de M2 à la proximité du pronom anaphorique.

Cependant, dans nos supports d'études, ces facteurs sont, selon nous, arasés statistiquement. La sélection de paires significatives, mais également la taille de nos échantillons, nous préservent de ces biais.

3.4. Le facteur sémantique

Dans ce paragraphe, nous nous proposons de déterminer le rôle joué par les facteurs sémantiques relevés dans la littérature spécialisée notamment par Cooper et Ross (1975) et Renner (2006) dans la sélection d'un ordre de composés syndétiques privilégié.

Pour ce faire, nous avons examiné les supports d'études pour lesquels les facteurs sémantiques n'étaient pas neutralisés : à savoir les corpus CUL ET CHEMISE, LISSE ET SOYEUX, VITE ET BIEN et ARTS ET MÉTIERS.

Pour l'étiquetage sémantique des 397 paires retenues, nous avons utilisé les six principes sémantiques utilisés par Renner : la [PROXIMITÉ SPATIALE], la [PROXIMITÉ TEMPORELLE], la [PROXIMITÉ PSYCHOLOGIQUE], la [SUPÉRIORITÉ SPATIALE], la [SUPÉRIORITÉ HIÉRARCHIQUE] et l'[ANTÉRIORITÉ TEMPORELLE]. Ces six traits peuvent être regroupés en trois grands principes :

- Le principe de proximité plaçant en M1 le mot pour lequel le locuteur idéal (dans son acception chomskyenne) a le plus d'empathie, plus de proximité ;
- Le principe de supériorité qui, comme le propose Renner, « prédit que tout composant entrant dans une relation sémantique de supériorité par rapport un autre, sera placé en première position dans un composé coordinatif » ;
- Et le principe d'antériorité plaçant en M1 l'élément de la coordonnée le plus antérieur³⁷.

37 LAMBERT (2001) réinterprète ce principe d'antériorité par la relation binaire ouvert/fermé dont la temporalité et l'antériorité découlent. « Nous avons pu constater que 'et' est le support d'un clivage thème/rhème qui réintroduit une dissymétrie fonctionnelle (et systématiquement binaire) entre les éléments antérieurs et les éléments postérieurs à lui. Un examen plus approfondi du lien entre les deux côtés de cette dissymétrie a permis alors de constater que ceux-ci entretenaient en fait un lien de complémentarité, et contraignant à réinterpréter les éléments antérieurs comme « ouverts » et posant les éléments subséquents

Grâce à cette répartition, Renner a démontré qu'en anglais contemporain, l'ordre des composés syndétiques était impacté par les contraintes de proximité et d'antériorité. Par manque de données, le principe de supériorité n'avait pas été validé.

Concernant nos travaux sur la coordination binomiale directe par « et », cette répartition par traits sémantiques est uniquement un point de vue accessoire. Notre enjeu n'est pas de déterminer si tel trait sémantique est plus saillant qu'un autre, même si nous proposerons en annexe le tableau A 1 présentant la répartition par traits sémantiques et par corpus des paires obéissant aux facteurs sémantiques³⁸. Il s'agit surtout pour nous d'évaluer dans quelle mesure les paires constituant nos supports d'études obéissent à des facteurs sémantiques. Nous avons donc observé 397 composés syndétiques par « et » en vérifiant si l'occupation en M1 des termes étaient impactée par l'un des trois (six) concepts définis par Renner.

Dans le tableau ci-dessous, nous avons compilé et mis en perspective, de façon binaire, les paires obéissant à ces concepts sémantiques et celles n'y obéissant pas.

Afin de faciliter la lecture de nos résultats, nous avons noté en caractères gras les résultats pour lesquels le χ^2 a une valeur-p inférieure à .05, c'est-à-dire l'ensemble des résultats pour lesquels la répartition des effectifs révèle un déséquilibre significatif, ceux pour lesquels avec un risque d'erreur faible de 5% nous pouvons statistiquement affirmer un résultat.

comme « fermés ». Et est donc loin d'être indifférent à la successivité linéaire. On ne sera pas étonné, dans ces conditions, qu'il puisse rencontrer la successivité temporelle, par l'intermédiaire de l'iconicité. On a finalement montré que la relation ouvert/fermé était la source d'un niveau supérieur de complétude qui permettait de construire à la fois des « étapes » discursives et des supports de cohérence qui ne sont sans doute pas sans effet sur l'élaboration de la temporalité discursive. »

38 Ce tableau fait apparaître que sur les 397 paires examinées, les facteurs de proximité et d'antériorité sont actifs dans près d'un quart des coordonnées (respectivement 25,94% et 25,19%) comme le facteur de supériorité dans une moindre mesure (16,37%).

TOTAL		CONCEPTS SÉMANTIQUES	
		ACTIFS	NON ACTIFS
48	CUL ET CHEMISE % TOTAL =	32 66,67	17 33,33
62	LISSE ET SOYEUX % TOTAL =	46 74,19	16 25,81
202	VITE ET BIEN % TOTAL =	143 70,44	59 29,56
85	ARTS ET MÉTIERS % TOTAL =	47 55,29	38 44,71
397	% TOTAL =	268 67,51	129 32,49

Tableau 9 : Répartition des paires impactées et non impactées par des facteurs sémantiques

Nous constatons cinq résultats statistiquement significatifs. Le seul corpus pour lequel on ne peut déterminer si les facteurs sémantiques sont opérants ou non est le corpus ARTS ET MÉTIERS ($\chi^2(1) = 0,95$; $p \leq 0,3$). Pour tous les autres, et d'une manière générale, l'ordre de coordination est massivement impacté par les facteurs sémantiques proposés par Renner ($\chi^2(1) = 48,67$; $p \leq 0,00000000001$).

Nous affirmons donc l'existence d'un facteur sémantique actif dans la sélection d'un ordre de coordination privilégié.

Selon nous, ce facteur sémantique est à rapprocher d'autres facteurs présentés dans cette partie, à savoir le facteur psycho-affectif et celui lié à la fréquence des mots que nous détaillons en suivant.

3.5. Le facteur psycho-affectif

Comme nous l'exposons dans nos travaux antérieurs (Couasnon, (2009, 2012)), nos interrogations sur la coordination binomiale par « et » sont nées d'une simple observation anecdotique : dans notre famille, lorsque l'on évoque notre couple, c'est la forme coordonnée « Graziella et Stéphane » qui est sélectionnée. *A contrario*, dans la famille de notre conjoint, c'est la forme « Stéphane et Graziella » qui ressort systématiquement. Il apparaît intuitivement que plus le degré de proximité relationnelle entretenue avec l'un des individus est fort, plus le prénom de ce dernier est choisi en tête de la coordination ; les proches de Stéphane le citent

en premier, et inversement pour ceux de Graziella. On peut donc supposer que cette notion d'affect liant un prénom à M1 peut se manifester au-delà du cadre relationnel familial ou amical. Il est tout à fait envisageable de la même façon un facteur psycho-affectif de groupe ou culturel.

Pour répondre à ces interrogations préliminaires, nous avons mis en place, dans Couasnon (2009), une expérimentation visant à éprouver ce facteur. Pour cela, nous avons sélectionné deux nouvelles paires de prénoms en plus de la paire « Stéphane » / « Graziella » : « Lara » / « Sylvain » et « Hélène » / « Bruno ». Pour chacune de ces paires, nous avons observé les réalisations spontanées de quatre groupes d'interviewés différents :

- GL correspond au cercle familial de « Lara »
- GSy à celui de « Sylvain »
- GH à celui de « Hélène »
- GB à celui de « Bruno »

Ces groupes étaient très homogènes en termes d'âge, de sexe et de catégories socio-professionnelles. De même, les informateurs étaient tous des locuteurs natifs du français. La seule variable était le degré de proximité psycho-affective existant entre les groupes et les prénoms « Lara », « Sylvain », « Hélène » et « Bruno ».

L'idée de cette expérience était de déterminer l'influence que pourrait avoir un prénom fortement marqué en termes de proximité psycho-affective dans la sélection d'une paire coordonnée par « et » : les locuteurs placent-ils en M1 les prénoms portés par des individus qui leur sont proches ?

Pour ce faire, nous avons observé, à l'insu de nos informateurs, les appariements de l'ensemble de leurs réalisations spontanées, en nous fixant un minimum de dix réalisations par informateurs.

Les résultats relevés étaient massifs : sur un total de 127 appariements, les locuteurs placent systématiquement le prénom de l'individu qui leur est proche en M1 ($\chi^2(1) = 127$; $p \leq 0,01$). Nous affirmons donc l'existence d'un facteur psycho-affectif liant le prénom de la personne dont le locuteur est le plus proche à M1.

Nous analysons ce facteur comme suit :

« Dès qu'un prénom fait référence à un individu appartenant au cercle privé du locuteur, la distanciation entre le prénom et l'individu qu'il représente devient quasiment impossible. Le

prénom n'est plus un hyperonyme, il est actualisé, il fait référence. Le prénom renvoie alors à un individu précis et perd de fait son statut d'objet de désignation uniquement composé de sons articulés successivement. Il devient alors un matériau complexe liant des sons aux propriétés propres à l'individu auquel il renvoie. Le prénom est alors « chargé » de sens et devient socialement connoté. ». (Couasnon, 2009).

Selon nous, ce facteur est à rapprocher des facteurs sémantiques que nous détaillons en 3.4. Le facteur sémantique, notamment du facteur de [PROXIMITÉ] proposé par Renner et inspiré du [Me First] de Cooper et Ross. Ce qui est proche du locuteur est prioritairement placé en M1. M1 est occupé par ce qui est proche du « je, ici, et maintenant » décrit comme procédé d'appropriation par Benveniste. M1 est sélectionné par ce qui est proche de notre expérience empirique, de notre expérience individuelle : il est non-marqué.

En fait, en adéquation avec les résultats rencontrés dans la littérature spécialisée, il semblerait que la répartition entre marqué / non-marqué plaçant les éléments les moins marqués en M1, puisse, comme l'affirmaient Benor et Levy (2006), être observée pour le facteur psycho-affectif.

Il apparaît également qu'au regard des résultats massifs que nous avons obtenus, il existe une prégnance du facteur psycho-affectif dans la sélection d'une paire coordonnée par « et » sur tout autre facteur, qu'il soit extraphonologique ou phonologique.

3.6. La fréquence des mots

Reprenons l'observation que nous faisons sur les prénoms « Stéphane » et « Graziella » : le nombre d'individus portant le prénom « Stéphane » en France est nettement plus important que celui de ceux portant le prénom « Graziella ». On peut donc aussi se demander si la fréquence relative des prénoms ne joue pas un rôle dans la sélection d'un ordre préférentiel des appariements. S'il en est ainsi, « Stéphane et Graziella » devrait être privilégié. En effet, l'INSEE dans les premiers mois de l'année 2010 comptabilisait en France 297 627 individus portant le prénom « Stéphane » pour 7 147 « Graziella »³⁹. Le prénom « Stéphane » était donc, à cette période de test, 97,6 % fois plus fréquent en France que celui de « Graziella ». En suivant l'hypothèse que nous développerons à venir, liant la fréquence et le degré proximité

³⁹ Ces données sont disponibles sur <http://meilleursprenoms.com/> grâce au précieux concours de Stéphanie Rapoport

relationnelle à M1, l'appariement « Stéphane et Graziella » devrait apparaître dans les mêmes proportions. Et c'est le cas. Une recherche sommaire sur Google fait ressortir la forme « Stéphane et Graziella » dans plus de 94% des appariements rencontrés ($\chi^2(1) = 0,32$; $p \leq 0,05$). La fréquence du prénom « Stéphane » le lie fortement à M1.

Comme nous venons de le voir, les locuteurs interviewés placent prioritairement le prénom porté par l'individu qui leur est le plus proche en M1. De plus, il semble exister une relation étroite entre la fréquence d'un prénom et sa position dans la coordination binomiale par « et ». Il y aurait donc une prégnance du facteur de proximité relationnelle liant la « valeur affective » et la fréquence d'un prénom à la première position de la coordination.

Par extension, nous avons souhaité étendre ces principes de facteurs psycho-affectifs et de fréquence, que nous pensons étroitement liés, à l'ensemble de nos supports d'études en nous demandant si la fréquence des mots détermine l'ordre de coordination, en tentant de répondre aux deux questions suivantes : M1 est-il statistiquement plus fréquent que M2 ? Ou inversement, M2 est-il plus fréquent que M1 ?

Pour cela, nous avons relevé les différents supports d'études répondant à cette question, à savoir cinq corpus, (CUL ET CHEMISE, LISSE ET SOYEUX, VITE ET BIEN, TOM ET ZOÉ et ARTS ET MÉTIERS) et le questionnaire d'enquête « ÉRIC ET AURIC » auquel nous porterons une attention particulière étant donné qu'il examine précisément la question de la corrélation entre la fréquence d'un mot et la sélection d'une position dans la coordination. Les autres questionnaires d'enquête étaient neutralisés en amont sur cette hypothèse et le corpus TARN ET GARONNE relevant des toponymes coordonnés par « et » n'est pas pertinent sur cette question.

Nous nous proposons donc d'observer les résultats que nous obtenons pour l'ensemble des supports d'études examinant l'hypothèse d'un lien entre la fréquence des mots et la sélection d'une place privilégiée dans les composés asyndétiques ci-dessous :

TOTAL		FRÉQUENCE	
		M1>M2	M2>M1
48	CUL ET CHEMISE % TOTAL =	28 58,33	20 41,67
62	LISSE ET SOYEUX % TOTAL =	41 66,13	21 33,87
1360	TOM ET ZOÉ % TOTAL =	722 53,09	638 46,91
202	VITE ET BIEN % TOTAL =	103 50,99	99 49,01
85	ARTS ET MÉTIERS % TOTAL =	56 65,88	29 34,12
6	ÉRIC ET AURIC % TOTAL =	6 100	0 0
1763	% TOTAL =	956 54,23	807 45,77

Tableau 10 : Fréquence des mots et position dans la coordination pour l'ensemble des supports d'études

Pour plus de lisibilité, à nouveau, sont notés en caractères gras les résultats pour lesquels la signification asymptotique est inférieure à .05.

Dans le tableau ci-dessus, dans le détail, cinq résultats statistiquement significatifs émergent, associant majoritairement le mot le plus fréquent à la position M1. Dans le questionnaire ÉRIC ET AURIC traitant spécifiquement cette question en tentant de neutraliser tout autre facteur qu'il soit phonologique ou extraphonologique⁴⁰, les résultats sont sans appel : 100% des mots en M1 sont les plus fréquents.

D'une manière générale, il apparaît que sur 1 763 paires testées sur le champs d'exploration liant la fréquence des mots et une position privilégiée de la coordination, le mot le plus fréquent sélectionne M1 dans 54,23% des cas ($\chi^2(1) = 12,59$; $p \leq 0,0003$).

Nous affirmons donc qu'en adéquation avec les résultats relevés dans la littérature spécialisée par Wright, Hay et Bent (2005), par Gustafsson (1976), par Fenc-Ozelon (1989) et par Benor et Levy (2006), la fréquence est un facteur déterminant dans la sélection d'un ordre préférentiel de coordination par « et » : les mots les plus fréquents sélectionnent prioritairement M1.

⁴⁰ Pour ce faire, comme nous le présentons en 2.2.4. Protocole de constitution du questionnaire d'enquête ÉRIC ET AURIC, nous n'avons traité que des binômes de « mêmes profils phonologiques », équisyllabiques, et possédant les mêmes types d'initiales et de finales. Seule la fréquence des prénoms différait.

Comment analyser un tel résultat ?

Conformément à ce que proposaient Cooper et Ross en attribuant ce principe à Panini, nous pensons que le facteur de fréquence liant le mot le plus fréquent à M1 est à rapprocher du « Principe poétique de gradation syllabique » de Jakobson, ou de la loi Zipf⁴¹ qui veut que les mots les plus fréquents tendent à être les plus courts.

On peut également rendre compte de ces résultats en invoquant à nouveau l'opposition marqué / non-marqué. Rappelons que selon Troubetzkoy cité par Carvalho J. (1997), « la marque expliquerait pour partie, la fréquence des phonèmes : « Des deux termes d'une opposition privative le terme non-marqué apparaît plus souvent dans le discours suivi que le terme marqué. » ».

Ainsi en adéquation avec ce que nous relevions dans la littérature spécialisée, les mots les plus fréquents étant les moins marqués, ils sélectionnent prioritairement la position la moins marquée, à savoir M1.

De fait, ce principe apparemment extraphonologique pourrait, en fait, relever de facteurs phonologiques tels que la longueur syllabique que nous développerons ultérieurement en 4.1.1. Le nombre de syllabes.

3.7. Le facteur morphologique

Dans cette section, il s'agit de déterminer si le lien que nous soupçonnions dans des études antérieures (Couasnon, 2009) associant un facteur morphologique à la sélection d'un ordre de prénoms privilégiés est vérifiable et également applicable à des composés syndétiques en « et ». Nous avons en effet remarqué que pour quelques paires de prénoms présents dans le corpus TOM ET ZOÉ, M1 et M2 pouvaient être construits sur la base de deux règles classiques de la morphologie dérivationnelle (Apothéloz, 2002) : la troncation et la suffixation.

Pour exemple, examinons quelques paires relevées dans le tableau ci-après.

Il apparaît que partant d'un « mot-source », nous obtenons un « mot dérivé » grâce à l'application de trois types de traitements dérivationnels :

⁴¹Zipf, G. K. (1974). « La Psychobiologie du langage. Une Introduction à la philologie dynamique ». Paris : Retz-C.E.P.L.

- Une troncation simple dans laquelle le point de troncation est le dernier segment phonologique ([t o ma]→[t o m]⁴². Il est possible de justifier le passage de la forme longue à la forme courte par l'apocope du segment [a]) ;
- Une troncation et une suffixation ([ʒ y l j ɛ̃] →[ʒ y l j ɛ t]. Une des propositions justifiant le passage de la forme courte à la forme longue serait une apocope de la nasalisation suivie d'une suffixation en [t]) ;
- Et enfin une suffixation simple ([v w a l]→ [v w a l j e]), le passage de la forme courte à la forme longue s'effectuerait par une suffixation en [j e]).

Le traitement dérivationnel proposé n'est qu'un point de vue accessoire de nos travaux. En aucun cas, nous ne nous interrogeons sur l'existence d'une « forme de base », d'une « forme source » ou d'une « forme dérivée ». Nous retenons surtout l'idée que les procédés ainsi présentés nous permettent de motiver l'hypothèse d'un lien morphologique entre les composants des paires examinées dans le tableau à venir, faisant émerger une opposition entre forme courte (notée FC) et forme longue (FL).

Ainsi, afin d'observer si la répartition en M1 et en M2 de mots morphologiquement liés est impactée par l'opposition FC/FL, nous avons recueilli dans nos supports d'études l'ensemble des paires pour lesquelles au moins un des deux procédés dérivationnels, à savoir la troncation ou la suffixation, est appliqué. Nous relevons un total de seize paires réparties dans les corpus TOM ET ZOÉ et VITE ET BIEN et dans le questionnaire d'enquête ÉLISE ET ZOÉ.

Dans le tableau ci-dessous, nous avons noté en caractères gras les résultats pour lesquels le χ^2 présente un seuil de signification asymptotique inférieur à .05. Pour les binômes opposant une forme longue à une forme courte morphologiquement liées, nos supports d'études fournissent les résultats suivants :

42 Notons une ouverture du [ɔ] de « Tom » caractéristique des syllabes finales fermées.

SUPPORTS D'ÉTUDES	TOTAL	«FC ET FL»	«FL ET FC»	Total	%	
					«FC ET FL»	«FL ET FC»
TOM ET ZOÉ	PAUL ET PAULINE	382	198	580	65,86	34,14
	LOUIS ET LOUISE	3453	971	4424	78,05	21,95
	JULIEN ET JULES	1214	737	1951	37,78	62,22
	JULIE ET JULIEN	1961	1501	3462	56,64	43,36
	JULES ET JULIE	3825	256	4081	93,73	6,27
	CLÉMENT ET CLÉMENCE	841	128	969	86,79	13,21
	JULIETTE ET JULES	1214	737	1951	37,78	62,22
VITE ET BIEN	VILLE ET VILLAGE	265000	25300	290300	91,28	8,72
	ROUTES ET AUTOROUTES	366000	25500	391500	93,49	6,51
	PARENTS ET GRANDS-PARENTS	264000	33700	297700	88,68	11,32
	VOILES ET VOILIERS	12000000	1970000	13970000	85,9	14,1
ÉLISE ET ZOÉ	TOM ET THOMAS	77	28	105	73,33	26,67
	HÉLÈNE ET HÉLÉNA	89	16	105	84,76	15,24
	STÉPHANE ET STÉPHANIE	80	25	105	76,19	23,81
TOTAUX	TOTAL TOKEN =	12908136	2059097	14967233	86,24	13,76
	TOTAL TYPE =	14	2	16	87,5	12,5

Tableau 11 : Facteur morphologique et position dans la coordination

Nous sommes consciente que le nombre d'appariements étudié est trop peu élevé pour traiter de manière approfondie l'hypothèse de l'existence d'une corrélation entre des facteurs morphologiques et la sélection d'un ordre préférentiel de coordination binomiale. De surcroît, nos résultats portent essentiellement sur des paires de prénoms. Néanmoins, nos premiers résultats nous semblent suffisamment notables pour que nous procédions malgré tout à un examen, même sommaire, de ces données.

En effet, tous les résultats présentés dans le tableau ci-dessus sont statistiquement significatifs. Il apparaît que sur seize paires de mots morphologiquement liés, en termes de type⁴³, quatorze, soit 87,5% d'entre elles, sélectionnent FC en M1 et FL en M2 ($\chi^2(1) = 9$; $p \leq 0,002$). En termes de token, le résultat est très homogène avec 86,24% des paires favorisant FC en M1 et FL en M2 ($\chi^2(1)$; $p \leq 0,01$). Les seules exceptions relevées positionnent toutes deux le prénom « Jules » en M2. De fait, il nous faudra examiner ultérieurement ce que terme a de particulier.

Quoi qu'il en soit, nous pouvons affirmer qu'il existe un facteur liant la forme courte de deux mots morphologiquement liés à M1 et sa forme longue à M2. Ce facteur est selon nous l'indice de l'existence d'un facteur morphologique.

⁴³ Nous reviendrons plus précisément sur la distinction entre type et token distinguant les objets en général de leur nombre de réalisations réelles page 174.

Comment analyser un tel résultat ?

En adéquation avec l'analyse que nous proposons dans Couasnon (2009), en formalisant l'hypothèse liant FC à M1 et FL à M2, nous pouvons proposer comme modèle de construction des paires morphologiquement liées le schéma suivant :

« FC ET FL »

De la même façon, nous pouvons également schématiser les opérations de la morphologie dérivationnelle que nous opérons sur nos paires comme suit :

- Pour la troncation : FL + apocope → FC.
- Pour la troncation suivie d'une suffixation : (FL + apocope) + suffixe → FC.
- Et pour la suffixation : FC + suffixe → FL.

La formalisation du modèle de coordination de binômes morphologiquement liés du type «FC ET FL» pourrait être ainsi réduite (au sens mathématique du terme) :

« FC ET (FC+ suffixe) »

« (FL+ apocope) ET FL »

Ces schémas de composés syndétiques de deux mots morphologiquement liés nous renvoient aux phénomènes de redoublement partiel des noms hypocoristiques en français comme « Totor » ou « Nanar » décrits par Plénat (1999) dans lesquels la seconde syllabe est une copie partielle de la première :

$\sigma\alpha + (\sigma\alpha + \text{suffixe})$

Il ne s'agit là que d'une réflexion hypothétique. En effet, ne disposant que de peu de données, même si nous affirmons que pour des binômes morphologiques liés FC est en M1 et FL en M2, à ce stade de nos travaux, il nous est impossible d'asseoir une corrélation entre facteur morphologique et position de la coordonnée.

En effet, ce que nous présentons dans ce paragraphe comme une caractérisation morphologique pourrait en fait être un facteur phonologique lié à la taille des mots. Il est totalement envisageable de considérer les mêmes résultats en termes de longueur syllabique. Nous reviendrons abondamment sur cette hypothèse page 96.

3.8. Le facteur culturel

Dans le cadre de coordinations genrées _à savoir quand l'un des termes est au masculin et l'autre au féminin, que nous nommons paires, composés ou binômes mixtes_ notamment dans le corpus TOM ET ZOÉ, le genre a-t-il une incidence sur la sélection d'un ordre préférentiel de coordination ? M1 est-il plutôt masculin ? Plutôt féminin ? Qu'en est-il de M2 ?

Il s'agit de répondre à cette question de manière statistique, même si nous sommes convaincue que la question du genre n'est pas seulement une question sociologique. Pour examiner en détail l'hypothèse corrélant le genre des mots à une position privilégiée de la coordination, nous avons relevé 1 536 paires des différents supports d'études considérant cette question, en mettant de côté le corpus TARN ET GARONNE non concerné par cette hypothèse, et les questionnaires d'enquête. En effet, ces derniers sont neutralisés sur la question du genre. Ils ne sont composés que de paires non mixtes.

TOTAL		Masc ET Fem	Fem ET Masc
14	CUL ET CHEMISE % TOTAL =	14 100	0 0
15	LISSE ET SOYEUX % TOTAL =	3 20	12 80
1420	TOM ET ZOÉ % TOTAL =	674 47,46	746 52,54
62	VITE ET BIEN % TOTAL =	34 54,84	28 45,16
42	ARTS ET MÉTIERS % TOTAL =	19 54,76	23 45,24
1553	% TOTAL =	744 47,91	809 52,09

Tableau 12 : Genre et position dans la coordination pour l'ensemble des supports d'études

Dans le tableau ci-dessus, afin d'en faciliter la lecture, nous avons noté en caractères gras les résultats pour lesquels le χ^2 a une valeur-p de .05 et en italique ceux pour lesquels il est compris entre $0,05 < p \leq 0,1$.

Ainsi, par supports d'études, trois résultats statistiquement significatifs et une tendance statistique apparaissent.

Dans le détail, dans le corpus TOM ET ZOÉ, contrairement à ce qu'avaient relevé Wright Hay et Bent (2005) qui indiquaient que «Male names tend to be characterized by First

Position Phonology, while female names tend to be characterized by Second Position Phonology (...) Results from our name-ordering experiments provide further evidence that this gendered phonology plays a role in ordering preferences.»), les prénoms féminins occupent majoritairement M1 dans 52,54 % des 1 420 paires testées sur cette question ($\chi^2(1) = 3,65$; $p \leq 0,05$).

Nous constatons les mêmes résultats dans le corpus LISSE ET SOYEUX, dans lequel, dans 80 % des paires mixtes examinées M1 est un féminin ($\chi^2(1) = 5,04$; $p \leq 0,02$). Inversement, dans les composés syndétiques mixtes présents dans le corpus CUL ET CHEMISE, les masculins occupent systématiquement M1 ($\chi^2(1) = 14$; $p \leq 0,0001$).

Notons particulièrement les résultats des corpus CUL ET CHEMISE et LISSE ET SOYEUX en relevant ce qu'ils ont de commun et de dissemblable.

Ces deux corpus proposent une quinzaine de composés syndétiques par « et » mixtes en « N&N » et « Adj&Adj ». Le corpus CUL ET CHEMISE, recueillant des binômes figés place systématiquement les mots masculins en M1. Inversement, dans le corpus LISSE ET SOYEUX, compilant des allégations cosmétiques récentes, les mots féminins occupent très massivement cette première position.

Comment analyser une telle différence ?

Nous ne nous risquons pas à proposer une distinction entre des figements « datés » et des collocations « récentes », ni même à évoquer le caractère marketing des binômes LISSE ET SOYEUX visant à séduire la « ménagère » responsable des achats du foyer. Ces hypothèses seraient à étudier ultérieurement.

Quoi qu'il en soit, dans l'ensemble de nos supports d'études, en opposition aux résultats avancés dans la littérature spécialisée, les mots féminins ont statistiquement tendance à occuper M1 dans 52,09% ($\chi^2(1) = 2,72$; $p \leq 0,1$).

Nous avons souhaité observer les mêmes résultats en les confrontant avec ceux que nous obtenons sans le corpus TOM ET ZOÉ, étant donné que du fait de sa taille et de la façon dont il a été constitué, ce corpus est le moins contrôlé de nos supports d'études.

Ainsi, en ce qui concerne la répartition en M1 et M2 par genre des paires mixtes présentes dans nos supports d'études sans les données relevées dans le corpus TOM ET ZOÉ, nous

obtenons les résultats suivants :

TOTAL-TOM ET ZOÉ		Masc→Fem	Fem→Masc
14	CUL ET CHEMISE % TOTAL =	14 100	0 0
15	LISSE ET SOYEUX % TOTAL =	3 20	12 80
62	VITE ET BIEN % TOTAL =	34 54,84	28 45,16
42	ARTS ET MÉTIERS % TOTAL =	19 54,76	23 45,24
133	% TOTAL =	70 52,63	63 47,37

Tableau 13 : Genre et position dans la coordination pour l'ensemble des supports d'études sans TOM ET ZOÉ

Dans le tableau ci-dessus, nous constatons que pour l'ensemble de nos données sans le corpus TOM ET ZOÉ, la répartition par genre en M1 et en M2 ne présente pas de déséquilibre statistiquement significatif ($\chi^2(1) = 0,37$; $p \leq 0,54$).

En l'état, nous ne pouvons affirmer l'existence d'un facteur culturel assignant une position privilégiée aux mots masculins ou féminins.

En fait, d'une manière générale, comme nous l'avons démontré tout au long de cette partie, nous affirmons l'existence de facteurs extraphonologiques actifs dans la sélection d'un ordre préférentiel de coordination par « et ». Cependant, nous sommes convaincue que certains facteurs apparemment extraphonologiques, (nous pensons particulièrement aux facteurs sériels, aux facteurs de fréquence et aux facteurs morphologiques) pourraient en fait relever de questions phonologiques que nous nous proposons de détailler dans le chapitre à venir.

Chapitre 4 : Facteurs phonologiques gouvernant les appariements

Les facteurs phonologiques que nous avons retenus pour l'analyse de nos données sont organisés selon trois grands axes.

Dans un premier temps, nous observerons la longueur syllabique relative de M1 et M2 avant de confronter la taille de M1 par rapport à celle M2. Puis, nous observerons les débuts des mots en M1 et en M2 en examinant tout d'abord les segments initiaux par types, puis par traits avant d'observer les syllabes initiales. Enfin, nous proposerons le même traitement pour les segments finaux.

4.1. La taille des mots

4.1.1. Le nombre de syllabes

4.1.1.1. Longueur relative de M1 et de M2

Au travers de ce critère, nous nous proposons d'examiner la question de savoir combien de syllabes M1 et M2 contiennent. Cette question inclut également celle de savoir si M1 en contient plus que M2 ou inversement. Nous avons cherché à déterminer si le nombre de syllabes de M, autrement dit sa taille, joue un rôle dans la sélection préférentielle d'une position dans la coordination.

Pour cela, nous avons analysé différents supports d'études. Nous avons toutefois mis de côté les questionnaires d'enquête étant donné que pour la plupart d'entre eux, la question de la taille était neutralisée, exception faite des questionnaires d'enquête ÉLISE ET ZOÉ et PAUL ET PASCAL qui traitaient précisément la question de la corrélation pouvant exister entre la taille d'un mot et la sélection d'une position dans la coordination. Rappelons également que nous avons retiré de l'examen du nombre de syllabes toutes les paires dont l'un des mots contient un schwa interne.

Nous avons donc testé 1 877 paires, desquelles nous avons ensuite extrait 926 paires équisyllabiques ne nous permettant pas de déterminer un rôle actif ou non de la taille.

Les résultats apparaissent dans le tableau ci-dessous.

Comme le montre le tableau, sur 951 binômes non équisyllabiques testés, dans 511 formes

coordonnées non équisyllabiques, le nombre de syllabes de M1 est inférieur à celui de M2 (53,73%) et inversement, pour 440 binômes, le nombre de syllabes de M1 est inférieur à celui de M2 (46,27%). Comme nous l'indiquons supra, nous avons noté en italiques les résultats pour lesquels le χ^2 est compris entre $0,05 < p \leq 0,1$, et en caractères gras ceux pour lesquels $p \leq 0,05$. Ces deux seuils rendent compte de la différence entre des résultats statistiquement significatifs pour lesquels la valeur-p est inférieure à .05 et des tendances possédant une valeur asymptotique comprise entre .05 et .10.

TOTAL	$n\sigma M1 < n\sigma M2$	$n\sigma M2 < n\sigma M1$
951		
M1	511 53,73	
M2		440 46,27
Total :	951 100,00%	

Tableau 14 : Longueur syllabique et position dans la coordination

Ce résultat nous permet d'affirmer l'existence d'une propension générale corrélatant la taille des mots à leur position dans la coordination. Les mots les plus courts ont statistiquement plus tendance à occuper majoritairement M1 ($\chi^2 (1) = 5,3$; $p \leq 0,02$).

Nous avons mené le même test, indépendamment, support d'études par support d'études. Car conformément à ce que nous annonçons page 33, bien qu'étroitement liés à l'ensemble dont ils dépendent (corpus ou questionnaires d'enquête), ils sont autonomes et donc aisément exploitables.

TARN ET GARONNE	$n\sigma M1 < n\sigma M2$	$n\sigma M2 < n\sigma M1$
41		
M1	27 65,85	
M2		14 34,15
Total :	41 100,00%	

Tableau 14a : Longueur syllabique et position dans la coordination dans le corpus TARN ET GARONNE

CUL ET CHEMISE 19	$n\sigma M1 < n\sigma M2$	$n\sigma M2 < n\sigma M1$
M1	14 73,68	
M2		5 26,32
Total :	19 100,00%	

Tableau 14b : Longueur syllabique et position dans la coordination dans CUL ET CHEMISE

LISSE ET SOYEUX 39	$n\sigma M1 < n\sigma M2$	$n\sigma M2 < n\sigma M1$
M1	28 71,79	
M2		11 28,21
Total :	39 100,00%	

Tableau 14c : Longueur syllabique et position dans la coordination dans LISSE ET SOYEUX

TOM ET ZOÉ 671	$n\sigma M1 < n\sigma M2$	$n\sigma M2 < n\sigma M1$
M1	316 47,09	
M2		355 52,91
Total :	671 100,00%	

Tableau 14d : Longueur syllabique et position dans la coordination dans TOM ET ZOÉ

VITE ET BIEN 110	$n\sigma M1 < n\sigma M2$	$n\sigma M2 < n\sigma M1$
M1	63 57,27	
M2		47 42,73
Total :	110 100,00%	

Tableau 14e : Longueur syllabique et position dans la coordination dans VITE ET BIEN

ARTS ET MÉTIERS	$n\sigma M1 < n\sigma M2$	$n\sigma M2 < n\sigma M1$
53		
M1	47 88,68	
M2		6 11,32
Total :	53 100,00%	

Tableau 14f : Longueur syllabique et position dans la coordination dans ARTS ET MÉTIERS

ÉLISE ET ZOÉ	$n\sigma M1 < n\sigma M2$	$n\sigma M2 < n\sigma M1$
10		
M1	8 80	
M2		2 20
Total :	10 100,00%	

Tableau 14g : Longueur syllabique et position dans la coordination dans ÉLISE ET ZOÉ

PAUL ET PASCAL	$n\sigma M1 < n\sigma M2$	$n\sigma M2 < n\sigma M1$
8		
M1	8 100	
M2		0 0
Total :	8 100,00%	

Tableau 14h : Longueur syllabique et position dans la coordination dans PAUL ET PASCAL

Les tableaux présentent des résultats tout à fait intéressants. Des contrastes importants apparaissent dans nos supports d'étude.

Dans le détail, même si nous proposons ci-dessous un relevé général de ces résultats, nous nous intéressons plus particulièrement à trois d'entre eux, à savoir ceux présentés dans les questionnaires d'enquêtes PAUL ET PASCAL et ÉLISE ET ZOÉ, ainsi que le résultat

contradictoire relevé dans TOM ET ZOÉ.

PAUL ET PASCAL et ÉLISE ET ZOÉ sont les questionnaires d'enquête qui ont été mis en œuvre pour tester précisément l'hypothèse d'une corrélation entre la taille des mots et la position dans la coordination. C'est pourquoi nous pensons qu'ils méritent une attention particulière. Or, nous remarquons que pour ces deux questionnaires d'enquête, les résultats sont massifs : les mots les plus courts sélectionnent M1 prioritairement dans 100 % des paires dans PAUL ET PASCAL et dans 80% des occurrences testées dans ÉLISE ET ZOÉ.

Certes, la taille de l'échantillon de PAUL ET PASCAL est trop réduite pour que le test de khi deux soit probant. Il n'en reste pas moins que les résultats produits par ce dernier sont à mettre en rapport avec les résultats obtenus par le questionnaire d'enquête ÉLISE ET ZOÉ : dans 80% des binômes non équisyllabiques testés, M1 est plus court que M2 ($\chi^2 (1) = 8,33$; $p \leq 0,003$). Notons également que les résultats présentés sont en adéquation avec ceux relevés dans les autres supports d'études.

Ainsi, dans le corpus ARTS ET MÉTIERS, sur 53 binômes non équisyllabiques testés, le nombre de syllabes de M1 est inférieur à celui de M2 dans 88,68% des cas, $\chi^2 (1) = 30,77$; $p \leq 0.000\ 000\ 03$).

De la même façon, dans le corpus CUL ET CHEMISE, le mot le plus court occupe M1 dans 73,78 % des paires non équisyllabiques testées ($\chi^2 (1) = 4,26$; $p \leq 0,04$).

Concernant le corpus LISSE ET SOYEUX, sur 39 paires non équisyllabiques, le nombre de syllabes de M1 est inférieur à celui de M2 dans 71,79% des appariements testés, ($\chi^2 (1) = 7,41$; $p \leq 0,01$).

Le corpus TARN ET GARONNE suit encore la même distribution : 65,85% des occurrences non équisyllabiques placent prioritairement le mot le plus court en M1 ($\chi^2 (1) = 4,12$; $p \leq 0,04$).

VITE ET BIEN présente encore la même répartition : dans 57,27% des paires, le mot le plus court est en M1 ($\chi^2 (1) = 2,33$; $p \leq 0,13$). Notons que ce résultat n'est pas statistiquement validé par un Khi Deux. Néanmoins sa valeur asymptotique est juste au-dessus du seuil de .1. Rappelons que la valeur-p est graduelle, ce qui signifie qu'avec une valeur de 0,13, même si nous ne pouvons rejeter l'hypothèse nulle, le « non-rejet » n'est pas très fort. D'autant plus, et c'est selon nous l'argument le plus fort, ce résultat demeure en adéquation avec l'ensemble des résultats relevés sur cette hypothèse.

Enfin, le corpus TOM ET ZOÉ présente des résultats contraires à ceux des précédents

supports d'études. En effet, une très légère majorité des binômes non équisyllabiques expérimentés (52,91%) semble favoriser le mot le plus long en M1.

Cependant, d'une part, la répartition des mots en fonction de leur taille n'est pas statistiquement valide ($\chi^2(1) = 2,33 ; p \leq 0,13$). Et d'autre part, comme nous l'indiquons en 2.1.3. Protocole de constitution du corpus TOM ET ZOÉ, le corpus TOM ET ZOÉ est, du fait de même de sa nature hétérogène, le support d'études pour lequel il nous est le plus difficile de contrôler les données. Il est donc très probable que d'autres facteurs phonologiques ou extraphonologiques interfèrent dans la sélection d'une position dans la coordination binaire par « et ».

A cause sa taille importante, TOM ET ZOÉ est donc susceptible d'introduire un biais dans les résultats. Il est donc intéressant de présenter une nouvelle répartition qui exclut les 671 M1 et M2 de TOM ET ZOÉ sur la question de la taille des mots.⁴⁴

Ainsi, nous obtenons la nouvelle répartition présentée dans le tableau ci-dessous :

TOTAL-TOM ET ZOÉ 280	$n\sigma M1 < n\sigma M2$	$n\sigma M2 < n\sigma M1$
M1	195 69,64	
M2		85 30,36
Total :	280 100,00%	

Tableau 15 : Longueur syllabique et position dans la coordination sans le corpus TOM ET ZOÉ

Comme le montre le tableau ci-dessus, dans 280 binômes non équisyllabiques testés, le nombre de syllabes de M1 est inférieur à celui de M2 69,64% dans 195 paires, et inversement pour 85 paires, le nombre de syllabes de M1 est inférieur à celui de M2 (30,36%) ($\chi^2(1) = 43,21 ; p \leq 0,000\ 000\ 000\ 04$).

Au regard de ces éléments, il existe une très forte présomption contre l'hypothèse nulle.

Nous pouvons donc affirmer l'existence d'une corrélation forte liant la taille des mots à leur position dans la coordination. Les mots les plus courts ont statistiquement tendance à occuper préférentiellement la position M1.

⁴⁴ D'une manière générale, pour l'ensemble de nos travaux, nous proposerons systématiquement la confrontation des résultats pour l'ensemble des données et ceux présentés sans les paires issues de TOM ET ZOÉ.

4.1.1.2. Nombre de syllabes de M1 et M2

Nous venons de voir que les mots les plus courts occuperaient préférentiellement M1. Nous nous attacherons dans cette section à aller plus loin en examinant plus précisément le nombre de syllabes de M1 et de M2.

Pour cela, dans un premier temps, nous nous sommes interrogée sur le nombre de syllabes de M1 et M2 : combien de syllabes M1 et M2 possèdent-ils ? Une, deux, trois, quatre et plus ? Compte tenu du peu de mots de quatre, cinq, et six syllabes, nous avons initialement choisi de les regrouper en une seule catégorie : les mots de quatre syllabes et plus.

Les supports d'études pertinents pour l'examen de cette question sont les mêmes que ceux précédemment utilisés pour déterminer le lien existant entre longueur syllabique et position dans la coordination, à savoir la collection complète de six corpus et les deux questionnaires d'enquête ÉLISE ET ZOÉ et PAUL ET PASCAL.

Dans la mesure où ce que nous souhaitons rendre visible est la comparaison entre le nombre de n syllabes en M1 et en M2, il convient de considérer comme ratio pertinent non pas le rapport entre le nombre de syllabes sur le nombre total d'appariements, mais le nombre de mots de n syllabes en M1, puis en M2 sur le nombre total de mots en n syllabes. Les résultats proposés dans le tableau ci-dessous présentent le pourcentage des appariements en fonction de leur nombre de syllabes.

TOTAL 1877	nombre de σ			
	σ .	$\sigma.\sigma$.	$\sigma.\sigma.\sigma$.	$\sigma.\sigma.\sigma.\sigma$. et +
M1	304 55,88	1171 49,87	377 46,37	25 51,02
M2	240 44,12	1177 50,13	436 53,63	24 48,98
Total :	544 100,00%	2348 100,00%	813 100,00%	49 100,00%

Tableau 16 : Nombre de syllabes et position dans la coordination

Les 55,88 % de mots monosyllabiques en M1 ne correspondent pas à 55,88% d'occurrences sur les 1 877 paires testées, mais à 55,88% de l'ensemble des mots monosyllabiques rencontrés au sein des 1 877 binômes, c'est-à-dire 304 mots en M1, sur un total de 544 mots de une syllabe.

Ainsi, comme le montre le tableau ci-dessus, nous pouvons affirmer que les mots

monosyllabiques occupent majoritairement M1 ($\chi^2(1) = (7,53 ; p \leq 0,01)$).

Il est intéressant de constater qu'il n'en va pas de même avec les mots plus longs. Aucun résultat statistiquement valide ne se dégage des répartitions des dissyllabes, des trisyllabes et des mots de quatre syllabes et plus ($p > 0,1$). Nous relevons cependant qu'en termes de score (et non de validation statistique), nous observons une distribution graduelle décroissante de l'occupation de M1 dans laquelle il apparaîtrait que plus un mot est court, plus il aurait une prédisposition à sélectionner la première place dans la coordination. 55,88% des mots monosyllabiques favorisent la position M1, pour 49,75 % des termes dissyllabiques et 49,87% des mots trisyllabiques. Seuls les mots de quatre syllabes et plus, dont nous signalions le faible effectif (seulement 49 termes), semblent résister à cette tendance. En effet, une faible majorité des mots de quatre syllabes et plus (51,02%) occupe plutôt M1.

Afin de scruter plus encore cette distribution graduelle décroissante, il nous a semblé pertinent de présenter nos résultats, indépendamment, support d'études par support d'études.

TARN ET GARONNE 69	nombre de σ			
	σ .	$\sigma.\sigma$.	$\sigma.\sigma.\sigma$.	$\sigma.\sigma.\sigma.\sigma$. et +
M1	19 61,29	31 50,82	17 39,53	2 66,67
M2	12 38,71	30 49,18	26 60,47	1 33,33
Total :	31 100,00%	61 100,00%	43 100,00%	3 100,00%

Tableau 17a : Nombre de syllabes et position dans la coordination pour TARN ET GARONNE

CUL ET CHEMISE 42	nombre de σ			
	σ .	$\sigma.\sigma$.	$\sigma.\sigma.\sigma$.	$\sigma.\sigma.\sigma.\sigma$. et +
M1	35 55,56	6 37,5	1 25	0 0
M2	28 44,44	10 62,5	3 75	1 100
Total :	63 100,00%	16 100,00%	4 100,00%	1 100,00%

Tableau 17b : Nombre de syllabes et position dans la coordination pour CUL ET CHEMISE

LISSE ET SOYEUX	nombre de σ			
	σ .	$\sigma.\sigma$.	$\sigma.\sigma.\sigma$.	$\sigma.\sigma.\sigma.\sigma$. et +
59	17	30	<u>11</u>	<u>1</u>
M1	58,62	57,69	35,48	16,67
M2	12	22	<u>20</u>	<u>5</u>
	41,38	42,31	<u>64,52</u>	<u>83,33</u>
Total :	29	52	31	6
	100,00%	100,00%	100,00%	100,00%

Tableau 17c : Nombre de syllabes et position dans la coordination pour LISSE ET SOYEUX

TOM ET ZOÉ	nombre de σ			
	σ .	$\sigma.\sigma$.	$\sigma.\sigma.\sigma$.	$\sigma.\sigma.\sigma.\sigma$. et +
1420	116	998	306	
M1	46,22	50,71	49,28	
M2	135	970	315	
	53,78	49,29	50,72	
Total :	251	1968	621	0
	100,00%	100,00%	100,00%	100,00%

Tableau 17d : Nombre de syllabes et position dans la coordination pour TOM ET ZOÉ

VITE ET BIEN	nombre de σ			
	σ .	$\sigma.\sigma$.	$\sigma.\sigma.\sigma$.	$\sigma.\sigma.\sigma.\sigma$. et +
184	66	67	33	18
M1	60,55	42,95	42,31	72
M2	43	89	45	7
	39,45	57,05	57,69	28
Total :	109	156	78	25
	100,00%	100,00%	100,00%	100,00%

Tableau 17e : Nombre de syllabes et position dans la coordination pour VITE ET BIEN

ARTS ET MÉTIERS	nombre de σ			
	σ .	$\sigma.\sigma$.	$\sigma.\sigma.\sigma$.	$\sigma.\sigma.\sigma.\sigma$. et +
75	39	27	5	<u>4</u>
M1	84,78	39,13	23,81	28,57
M2	7	<u>42</u>	16	<u>10</u>
	15,22	<u>60,87</u>	76,19	<u>71,43</u>
Total :	46	69	21	14
	100,00%	100,00%	100,00%	100,00%

Tableau 17f : Nombre de syllabes et position dans la coordination pour ARTS ET MÉTIERS

ÉLISE ET ZOÉ	nombre de σ			
	σ .	$\sigma.\sigma$.	$\sigma.\sigma.\sigma$.	$\sigma.\sigma.\sigma.\sigma$. et +
20				
M1	5 62,5	11 55	4 33,33	
M2	3 37,5	9 55	8 33,33	
Total :	8 100,00%	20 100,00%	12 100,00%	0 100,00%

Tableau 17g : Nombre de syllabes et position dans la coordination pour ÉLISE ET ZOÉ

PAUL ET PASCAL	nombre de σ			
	σ .	$\sigma.\sigma$.	$\sigma.\sigma.\sigma$.	$\sigma.\sigma.\sigma.\sigma$. et +
8				
M1	7 100	<u>1</u> 16,76	<i>0</i> 0	
M2	0 0	<u>5</u> <u>83,33</u>	<i>3</i> <i>100</i>	
Total :	7 100,00%	6 100,00%	3 100,00%	0 100,00%

Tableau 17h : Nombre de syllabes et position dans la coordination pour PAUL ET PASCAL

Comme précédemment, nous avons noté en caractères gras les résultats pour lesquels le χ^2 est de $p \leq 0,05$ et en italiques ceux pour lesquels il est compris entre $0,05 < p \leq 0,1$. Nous mentionnons également, en les soulignant, les résultats pour lesquels la valeur-p est comprise en $0,1 < p \leq 0,14$. Il est bien entendu que nous ne les présenterons jamais comme des résultats statistiquement probants. Nous sommes néanmoins convaincue que certains d'entre eux méritent une attention plus particulière.

Dans le détail, les corpus ARTS ET MÉTIERS, VITE ET BIEN et le questionnaire d'enquête PAUL ET PASCAL révèlent des répartitions statistiquement valides ($p \leq 0,05$).

Plus précisément, par support d'études, nous dégageons les résultats suivants :

PAUL ET PASCAL nous procure à nouveau des résultats massifs : 100% des mots monosyllabiques semblent prédisposés à occuper M1⁴⁵. Même si ce résultat n'est pas statistiquement aussi fort que nous le souhaiterions, il convient de préciser que ce

⁴⁵ La présente formulation paraît ambiguë. Elle tient compte de fait que eu égard à la taille de l'échantillon testé (8 paires), le test de Khi Deux ne peut être probant.

questionnaire d'enquête a été spécifiquement mis en œuvre pour examiner le rapport entre la taille des mots et leur position dans la coordonnée.

Concernant les résultats obtenus pour les mots de plus de deux syllabes, même s'ils ne sont pas statistiquement valides ($p > 0,1$), nous notons néanmoins qu'ils sont récurrents dans la quasi totalité de nos supports d'études.

Ainsi nous retrouvons la dispersion graduelle décroissante que nous évoquions précédemment. Pour le questionnaire d'enquête PAUL ET PASCAL, 100% des mots monosyllabiques favorisent la position M1, pour 16,67 % des termes dissyllabiques et aucun mot trisyllabique.

Dans le corpus ARTS ET MÉTIERS, on observe clairement que les mots monosyllabiques occupent prioritairement M1. 84,78% des monosyllabes testés dans ce corpus sélectionnent la première position dans la coordonnée ($\chi^2 (1) = 22,26$; $p \leq 0,000\ 002$). Inversement, les mots de trois syllabes sont 76,19 % à préférer M2 ($\chi^2 (1) = 5,76$; $p \leq 0,02$). Par ailleurs, il existe une tendance statistique ($\chi^2 (1) = 3,26$; $p \leq 0,07$) plaçant majoritairement les dissyllabes, dans 60,87% des occurrences testant cette hypothèse, en M2.

A nouveau, nous retrouvons une dispersion graduelle décroissante. Pour le corpus ARTS ET MÉTIERS, 84,78% des mots monosyllabiques favorisent la position M1, pour 38,13 % des termes dissyllabiques et 23,81% des mots trisyllabiques. A noter, les mots de quatre syllabes et plus ne respectent pas tout à fait le tracé décroissant de la répartition des mots les plus courts aux mots les plus longs en M1 : en effet, sur les 14 termes testés, 28,57% des mots de quatre syllabes et plus occupent prioritairement M1, soit un peu moins de 4,46 % de plus que les trisyllabes.

Le questionnaire d'enquête ÉLISE ET ZOÉ, même s'il ne fournit aucun résultat confirmé statistiquement, présente encore une fois la même répartition graduelle : 62,5% des monosyllabes favorisent la position M1, pour 55 % des mots dissyllabiques et 33% des termes trisyllabiques.

Concernant le corpus VITE ET BIEN, nous relevons trois répartitions significatives : les monosyllabes occupent majoritairement (60,55%) la première partie de la coordination ($\chi^2 (1) = 4,85$; $p \leq 0,03$), tout comme les mots de quatre syllabes et plus dans 72% des mots examinés sur cette hypothèse ($\chi^2 (1) = 4,84$; $p \leq 0,03$). Inversement, les mots de deux syllabes sélectionnent la position opposée, M2, dans 57,05% des occurrences testées ($\chi^2 (1) =$

3,1 ; $p \leq 0,03$).

Encore une fois, en faisant abstraction des résultats relevés pour les mots de quatre syllabes et plus, nous notons une dispersion graduelle décroissante des monosyllabes aux trisyllabes : 60,55% des mots monosyllabiques favorisent la position M1, pour 42,95 % des termes dissyllabiques et 42,31% des mots trisyllabiques. .

Le corpus TARN ET GARONNE ne fait pas apparaître de résultats pour lesquels le test de Khi Deux est probant. Cependant, même sans validation statistique, nous observons qu'en termes de score, TARN ET GARONNE suit également la gradation décroissante en M1 relevée dans les autres supports d'étude : 61,29% des mots monosyllabiques favorisent la position M1, pour 50,82 % des termes dissyllabiques et 39,53 % des mots trisyllabiques. À l'opposé, les trois mots de quatre syllabes et plus, quant à eux, favorisent prioritairement M1 dans 66,67% des cas.

Les corpus LISSE ET SOYEUX et CUL ET CHEMISE suivent à nouveau la répartition graduelle décroissante précédemment relevée. A savoir, pour le corpus LISSE ET SOYEUX, 58,62% des mots monosyllabiques favorisent la position M1, pour 57,69 % des termes dissyllabiques, 35,48% des mots trisyllabiques et 16,67% des six mots testés de quatre syllabes et plus, et pour le corpus CUL ET CHEMISE, 55,56,% des mots monosyllabiques sélectionnent prioritairement la première position de la coordination M1, pour 37,5 % des termes dissyllabiques, 25% des mots trisyllabiques et 0% du seul mot testé de quatre syllabes et plus ($p > 0,1$).

D'une manière générale, sur l'ensemble de ces sept supports d'études, nous pouvons nous accorder sur trois points : le premier, c'est qu'étant donné le peu d'effectifs par type de mots, seuls des résultats très forts sont statistiquement valides. Le deuxième, c'est que la catégorie des mots de quatre syllabes et plus semble hétéroclite. Les résultats relevés ne semblent pas cohérents d'un support d'études à l'autre. (A nouveau, nous attirons l'attention sur le très petit nombre de mots de quatre syllabes et plus, et sur cette catégorie « fourre-tout » sur laquelle nous reviendrons plus en détail.) Enfin, et cela nous semble le plus pertinent, dans tous les support d'études qui viennent d'être discutés, l'occupation de M1 suit une gradation décroissante des monosyllabes aux trisyllabes : plus un mot est court, plus il semble favoriser M1.

Le corpus TOM ET ZOÉ présente encore une fois des résultats en contradiction avec ceux des

précédents support d'études. En effet, une minorité de 46,22% des mots monosyllabiques favorisent la position M1, pour une majorité hésitante de 50,71 % des termes dissyllabiques et 49,28% des mots trisyllabiques.

Cependant, ces résultats sont à nouveau à nuancer. En effet, d'une part, les résultats sont bien moins massifs que ceux présentés dans les sept autres supports d'études testés sur la question du nombre de syllabes. Ils semblent révélateurs d'une grande hésitation. D'autre part, à la différence des supports d'études détaillés *supra*, malgré le nombre considérable de binômes coordonnées par « et » expérimentés dans TOM ET ZOÉ, la répartition des mots en fonction de leur nombre de syllabes n'est pas statistiquement valide : elle n'est pas significative. En effet, dans ce corpus, 251 monosyllabes ont été éprouvés ($\chi^2(1) = 1,44$; $p \leq 0,2$). Puis, nous avons testé 1 968 mots dissyllabiques ($\chi^2(1) = 0,4$; $p \leq 0,5$). Pour finir, nous avons observé la répartition de 621 mots de trois syllabes ($\chi^2(1) = 0,13$; $p \leq 0,7$).

Enfin, et toujours, le corpus TOM ET ZOÉ est, du fait de même de sa nature, le support d'études le moins contrôlé de l'ensemble de nos données.

Compte tenu de ces trois objections, nous avons, encore une fois, choisi d'exclure le corpus TOM ET ZOÉ du tableau étudiant les liens entre la position dans la coordination et le nombre de syllabes. Nous nous proposons de nommer ce nouveau total excluant les formes issues de TOM ET ZOÉ : « TOTAL-TOM ET ZOÉ ».

Ainsi, nous obtenons la nouvelle répartition présentée dans le tableau ci-dessous :

TOTAL-TOM ET ZOÉ	nombre de σ			
	σ .	$\sigma.\sigma$.	$\sigma.\sigma.\sigma$.	$\sigma.\sigma.\sigma.\sigma$. et +
457				
M1	188 64,16	173 45,53	71 36,98	25 51,02
M2	105 35,84	207 54,47	121 63,02	24 48,98
Total :	293 100,00%	380 100,00%	192 100,00%	49 100,00%

Tableau 18 : Nombre de syllabes et position dans la coordination pour l'ensemble des supports d'études sans le corpus TOM ET ZOÉ

Comme le montre le tableau ci-dessus, nous pouvons statistiquement affirmer deux résultats liés à ce que nous présentons ci-dessus. Premièrement, dans les 293 monosyllabes testés, la majorité (64,16%) occupent prioritairement M1 ($\chi^2(1) = 23,51$; $p \leq 0,000\ 001$). Et

deuxièmement, seuls 71 des mots de trois syllabes, soit 36,98% ($\chi^2 (1) = 13,02 ; p \leq 0,0003$), préfèrent M1, alors qu'ils ont 63,01 à préférer M2.

Nous pouvons donc affirmer que les mots monosyllabiques sélectionnent prioritairement M1 contrairement aux termes trisyllabiques.

Observons maintenant les autres résultats présentés dans ce tableau.

Sur 380 mots dissyllabiques, 173 ont tendance à privilégier M1 soit 45,53% et 207, soit 54,47% M2 ($\chi^2 (1) = 3,04 ; p \leq 0,08$). La tendance statistique que nous relevons ici peut indiquer deux phénomènes⁴⁶ :

- Soit d'autres facteurs phonologiques ou extraphonologiques obscurcissent ce résultat ;
- Soit ce facteur n'est pas aussi fort que les autres facteurs liant la taille des mots à une position privilégiée dans la coordination.

Comment comprendre ce résultat ?

Selon nous, et au regard des éléments dont nous disposons au travers de ce tableau ainsi que ceux relevés *supra* pour chaque support d'études, si nous déterminons poussivement que les mots dissyllabiques sélectionnent prioritairement M1, nous constatons néanmoins que dans l'ensemble de nos données, les mots dissyllabiques sélectionnent un peu moins M1 que les mots monosyllabiques, et un peu plus que les termes trisyllabiques, conformément à dispersion graduelle décroissante que nous observions précédemment. Il apparaîtrait donc, avec toutes les précautions nécessaires, une disposition à ce que, pour des mots composés de une à trois syllabes, plus un mot est court, plus il occupera M1.

Enfin, concernant les mots de quatre syllabes et plus, sur 57 mots éprouvés, 29 occupent M1 (soit 50,88%) pour 28, M2 (soit 49,12%) ($\chi^2 (1) = 0,01 ; p \leq 0,93$). Nous ne pouvons rejeter l'hypothèse nulle, et donc il nous est impossible d'affirmer quoi que ce soit à propos de l'existence d'un lien entre les mots de quatre syllabes et plus et ou l'une ou l'autre position de la coordination.

Nous pouvons imputer cette absence de résultats à deux phénomènes extérieurs à l'hypothèse formulée : d'une part, nous n'avons testé qu'assez peu d'effectifs, compte tenu de la façon dont nous avons recueilli nos données. Et d'autre part, cette catégorie de quatre syllabes et plus est, comme nous l'indiquions plus haut, une catégorie un peu « fourre-tout ». Nous avons traité de la même façon les mots de quatre syllabes comme ceux de cinq ou de six syllabes en

46 Nous écartons de fait la taille de l'échantillon testé.

préjugant du fait que comme il y avait peu d'effectifs, il aurait été difficile d'obtenir une significativité statistique pour chacune de ces catégories. Nous espérons qu'en les regroupant, l'augmentation numérique ferait apparaître des résultats, mais il est possible que dans le détail ce regroupement obscurcisse nos résultats.

Jusqu'à présent, nous avons obtenu des résultats assez clairs corrélant la taille des mots à la sélection d'une place dans la coordination.

- Premièrement, comme présenté dans le tableau page 102, les mots en M1 sont plus courts dans 72,63% des appariements non équisyllabiques testés que les mots en M2 ;
- Deuxièmement, les mots monosyllabiques occupent majoritairement M1 dans 64,16% des termes monosyllabiques éprouvés ;
- Inversement, les mots trisyllabiques favorisent M2 dans 63,02% des mots de trois syllabes présents dans nos supports d'études ;
- Tout comme les mots de deux syllabes pour lesquels nous relevons une tendance à occuper M2 dans 54,47% des cas ;
- Enfin, pour les mots contenant de une à trois syllabes, il apparaîtrait une disposition dans laquelle plus un mot est court, plus il occuperait M1 de façon graduelle.

4.1.1.3. Corrélation entre nombre de syllabes de M1 et nombre de syllabes de M2

Les résultats précédents sont tous considérés indépendamment de ce qui se passe dans la reste de la coordination. En effet, nous avons observé la répartition en M1 et en M2 des mots en fonction de leur taille de manière absolue. Il est maintenant nécessaire d'approfondir ces résultats en les confrontant aux données présentes dans l'ensemble de la coordonnée. A savoir : nous savons à présent que les mots monosyllabiques sélectionnent prioritairement M1, mais précisément, dans le détail, nous ne connaissons pas la répartition des mêmes mots monosyllabiques lorsqu'ils sont confrontés à des mots de deux, trois, quatre, cinq ou six syllabes.

C'est ce que nous nous proposons à présent d'observer.

Les supports d'études traitant de cette question sont les mêmes que ceux précédemment

utilisés pour déterminer le lien existant entre le nombre de syllabes et la position dans la coordination, à savoir la collection complète de six corpus et les questionnaires d'enquête ÉLISE ET ZOÉ et PAUL ET PASCAL. Nous avons donc testé 1 877 appariements.

Dans la mesure où ce que nous souhaitons rendre visible est la répartition des mots de x syllabes par rapport à ceux de y syllabes, en fonction de leur position dans la coordination, il nous était nécessaire de considérer comme ratio pertinent les appariements « x syllabes et y syllabes », puis « y syllabes et x syllabes » sur le nombre total d'appariements ayant les mêmes spécificités, c'est-à-dire d'être composé d'un mot de x syllabes et d'un mot de y syllabes.

A nouveau, nous avons noté en italique les résultats pour lesquels le χ^2 est compris entre $0,05 < p \leq 0,1$, et en caractères gras ceux pour lesquels $p \leq 0,05$. Les cases en gris clair désignent les effectifs des appariements équisyllabiques. Pour éviter de surcharger ce tableau, nous avons délibérément laissé des cases vides correspondant aux types de confrontations inexistantes dans nos supports d'études. Seules sont notées les cases correspondant aux types de confrontations pour lesquelles il y a des effectifs, même s'il arrive que cet effectif soit de zéro.

TOTAL M1	M2						TOTAL M1=
	σ	$\sigma\sigma$	$\sigma\sigma\sigma$	$\sigma\sigma\sigma\sigma$	$\sigma\sigma\sigma\sigma\sigma$	$\sigma\sigma\sigma\sigma\sigma\sigma$	
σ %	94 38,54	177 61,46	26 42,62	7 100			304
$\sigma\sigma$ %	111 38,54	744	310 55,56	5 41,67	1 50		1171
$\sigma\sigma\sigma$ %	35 57,38	248 44,44	88	6 35,29			377
$\sigma\sigma\sigma\sigma$ %	0 0	7 58,33	11 64,71	3	1 100	1 50	23
$\sigma\sigma\sigma\sigma\sigma$ %		1 50	0 0				1
$\sigma\sigma\sigma\sigma\sigma\sigma$ %			1 50				1
TOTAL M2 =	240	1177	436	21	2	1	

Tableau 19 : Confrontation par nombre de syllabes et position dans la coordination

Comme le montre le tableau ci dessus, sur les 1 877 appariements testés, nous observons deux

résultats statistiquement remarquables allant dans le sens des résultats que nous présentions précédemment, à savoir que les mots les courts ont plus tendance à occuper M1.

Nous pouvons affirmer que la distribution de « σ & $\sigma\sigma$ » est statistiquement significative ($\chi^2(1) = 15,13$; $p \leq 0,0001$). Les mots monosyllabiques sont majoritairement placés en M1 dans 61,46% des cas lorsqu'ils sont confrontés à des mots dissyllabiques, tout comme les mots dissyllabiques qui sélectionnent prioritairement M1 (dans 55,56 % des occurrences) quand ils sont en présence de termes trisyllabiques ($\chi^2(1) = 6,89,29$; $p \leq 0,008$).

Considérons plus en détail les autres résultats apparaissant dans ce tableau.

Mettons dès à présent de côté les binômes totalisant moins de dix occurrences, qui, bien que présentant pour certains des résultats massifs, ne peuvent être analysés. Sont concernées les paires associant des mots de une syllabe à des mots de quatre syllabes, des mots de six syllabes à des termes trisyllabiques, des mots des quatre syllabes à des mots de six syllabes, des mots dissyllabiques à des mots de cinq syllabes et des mots de quatre syllabes à des mots de cinq syllabes. En effet, ces hypothèses étant testées sur des échantillons de moins de dix mots, le test de Khi Deux ne peut être probant. Nous ne le relèverons pas.

Les trois derniers résultats présentent des résultats non statistiquement confirmés en contradiction avec ceux précédemment confirmés. En effet, il apparaîtrait que les mots de quatre syllabes sélectionnent majoritairement M1 dans 64,71% des cas lorsqu'ils sont confrontés à des mots trisyllabiques ($\chi^2(1) = 1,47$; $p \leq 0,2$) et dans 58,33 % des cas lorsqu'ils sont opposés à des mots dissyllabiques ($\chi^2(1) = 0,33$; $p \leq 0,6$). Les mots de trois syllabes, quant à eux, sélectionnent prioritairement M1 dans 57,38 % des paires quand ils sont en présence de monosyllabes ($\chi^2(1) = 0,33$; $p \leq 0,2$). Ces résultats n'ont aucune valeur statistiquement parlant. Il en demeure qu'en termes de score, leur apparente contradiction avec les résultats précédemment mis au jour nous interroge.

A nouveau afin d'asseoir plus explicitement les deux résultats pertinents que nous avons relevés, il nous a semblé intéressant de proposer nos résultats, indépendamment, support d'études par support d'études.

Les tableaux ci-dessus présentent l'ensemble des résultats traitant de la confrontation de M1 et M2 par nombre de syllabes pour chaque support d'études.

TARN ET GARONNE M1	M2						TOTAL M1=
	σ	σσ	σσσ	σσσσ	σσσσσ	σσσσσσ	
σ %	9 12,45	7 87,5	3 60				19
σσ %	1	17	13				31
σσσ %	2	6	9			0	17
σσσσ %						1	1
σσσσσ %						100	0
σσσσσσ %			1	0			1
			100	0			
TOTAL M2 =	12	30	26	0	0	1	

Tableau 20a : Confrontation par nombre de syllabes et position dans TARN ET GARONNE

CUL ET CHEMISE M1	M2						TOTAL M1=
	σ	σσ	σσσ	σσσσ	σσσσσ	σσσσσσ	
σ %	25 17,65	9 75		1 100			35
σσ %	3	0	3				6
σσσ %		1					1
σσσσ %	0						0
σσσσσ %	0						0
σσσσσσ %							0
TOTAL M2 =	28	10	3	1	0	0	

Tableau 20c : Confrontation par nombre de syllabes et position dans CUL ET CHEMISE

LISSE ET SOYEUX M1	M2						TOTAL M1=
	σ	σσ	σσσ	σσσσ	σσσσσ	σσσσσσ	
σ %	4 41,67	7 58,33	5 62,5	1 100			17
σσ %	5	13	10	2			30
σσσ %	3	2	4	2			11
σσσσ %							1
σσσσσ %	0	0	1				1
σσσσσσ %	0	0	33,33				0
σσσσσσσ %		0					0
σσσσσσσσ %							0
TOTAL M2 =	12	22	20	5	0	0	

Tableau 20b : Confrontation par nombre de syllabes et position dans LISSE ET SOYEUX

TOM ET ZOÉ M1	M2						TOTAL M1=
	σ	σσ	σσσ	σσσσ	σσσσσ	σσσσσσ	
σ %	19 49,17	92 658	5 251				116
σσ %	89	658	251				998
σσσ %	27	220	59				306
σσσσσ %	84,38	46,71					0
σσσσσσ %							0
σσσσσσσ %							0
σσσσσσσσ %							0
TOTAL M2 =	135	970	315	0	0	0	

Tableau 20d : Confrontation par nombre de syllabes et position dans TOM ET ZOÉ

VITE ET BIEN M1	M2						TOTAL M1=
	σ	σσ	σσσ	σσσσ	σσσσσ	σσσσσσ	
σ %	28 27,91	31 72,09	7 70				66
σσ %	12 27,91	36	18 54,55	1 14,29	0 0		67
σσσ %	3 30	15 45,55	12 72,73	3 27,27			33
σσσσ %		6 85,71	8 72,73	2 27,27	1 100		17
σσσσσ %		1 100		0 0			1
σσσσσσ %							0
TOTAL M2 =	43	89	45	6	1	0	

Tableau 20e : Confrontation par nombre de syllabes et position dans VITE ET BIEN

ARTS ET MÉTIERS M1	M2						TOTAL M1=
	σ	σσ	σσσ	σσσσ	σσσσσ	σσσσσσ	
σ %	6 3,85	25 96,15	3 100	5 100			39
σσ %	1 3,85	14	9 81,82	2 66,67	1 100		27
σσσ %	0 0	2 18,18	2 81,82	1 33,33			5
σσσσ %	0 0	1 33,33	2 66,67	1			4
σσσσσ %		0 0					0
σσσσσσ %							0
TOTAL M2 =	7	42	16	9	1	0	

Tableau 20g : Confrontation par nombre de syllabes et position dans ARTS ET MÉTIERS

ÉLISE ET ZOÉ M1	M2						TOTAL M1=
	σ	σσ	σσσ	σσσσ	σσσσσ	σσσσσσ	
σ %	3 100	1 100	1 100				5
σσ %	0 0	6 71,43	5 71,43				11
σσσ %	0 0	2 28,57	2				4
σσσσ %							0
σσσσσ %							0
σσσσσσ %							0
TOTAL M2 =	3	9	8	0	0	0	

Tableau 20f : Confrontation par nombre de syllabes et position dans ÉLISE ET ZOÉ

PAUL ET PASCAL M1	M2						TOTAL M1=
	σ	σσ	σσσ	σσσσ	σσσσσ	σσσσσσ	
σ %		5 100	2 100				7
σσ %	0 0		1 100				1
σσσ %	0 0	0 0					0
σσσσ %							0
σσσσσ %							0
σσσσσσ %							0
TOTAL M2 =	0	5	3	0	0	0	

Tableau 20h : Confrontation par nombre de syllabes et position dans PAUL ET PASCAL

Dans le détail, cinq supports d'études ont des résultats statistiquement valides ($p \leq 0,05$) : les corpus VITE ET BIEN, ARTS ET MÉTIERS, TOM ET ZOÉ et LISSE ET SOYEUX. Le corpus CUL ET CHEMISE, quant à lui, fournit une tendance statistique ($0,05 < p \leq 0,1$).

Le corpus ARTS ET MÉTIERS fait apparaître deux résultats massifs statistiquement significatifs : les monosyllabes quand ils sont mis en relation avec des mots dissyllabiques occupent M1 dans 96,15% des 26 appariements testés sur ce critère ($\chi^2(1) = 22,15$; $p \leq 0,000003$). Et les mots de deux syllabes privilégient très nettement M1 lorsqu'ils sont confrontés à des mots trisyllabiques dans 81,82% des appariements ($\chi^2(1) = 4,45$; $p \leq 0,03$). Les autres résultats bien que non statistiquement validés, du fait sans aucun doute du faible nombre d'effectifs (aucun critère n'est testé sur plus de 13 appariements), font néanmoins émerger une constante : lorsque deux mots non équisyllabiques sont mis en présence, le plus court occupe plutôt M1.

Dans le corpus VITE ET BIEN, nous relevons à nouveau un résultat fort validé statistiquement : ici encore, les mots monosyllabiques opposés à des mots dissyllabiques occupent M1 dans 72,09% des 43 appariements testés sur ce critère ($\chi^2(1) = 8,4$; $p \leq 0,003$). De plus, les monosyllabes confrontés à des mots trisyllabiques favorisent M1 dans 70 % des appariements testés ($\chi^2(1) = 1,6$; $p \leq 0,2$). Comme précédemment, même si ces résultats ne sont pas aussi solides que nous le souhaiterions, ils sont en adéquation avec ceux que nous observions précédemment. Les autres résultats traitant de faibles effectifs ne peuvent être validés par un test de Khi Deux. Notons néanmoins qu'ils vont tous encore (à la seule exception des mots de quatre syllabes) dans le sens de la constante que nous relevions : lorsque deux mots non équisyllabiques sont mis en présence, le plus court occupe plutôt M1. Enfin, nous attirons également l'attention sur les résultats obtenus pour les mots de trois syllabes lorsqu'ils sont confrontés à des dissyllabes : l'examen de nos données met au jour une hésitation. VITE ET BIEN ne nous permet pas de déterminer si c'est la forme « $\sigma\sigma$ & $\sigma\sigma\sigma$ » ou la forme « $\sigma\sigma\sigma$ & $\sigma\sigma$ » qui sera privilégiée.

Concernant le corpus LISSE ET SOYEUX, nous constatons une répartition significative favorisant les mots dissyllabiques en M1 dans 83,33% des 12 données testées lorsqu'ils sont opposés à des mots de trois syllabes ($\chi^2(1) = 5,33$; $p \leq 0,02$). Les autres résultats, bien que non statistiquement validés sans aucun doute à cause du faible nombre d'effectifs (au maximum onze appariements), révèlent, encore une fois, le fait que lorsque deux mots non

équisyllabiques sont opposés dans une coordination directe par « et », le plus court occupe plutôt M1. Il est très probable qu'un plus grand nombre d'effectifs éprouvés sur ces critères nous aurait permis de rejeter l'hypothèse nulle et donc de valider statistiquement des résultats.

Le corpus CUL ET CHEMISE fait également apparaître une tendance statistique plaçant massivement les monosyllabes en M1 quand ils sont confrontés à des mots de deux syllabes dans 75 % des 12 appariements traitant ce critère ($\chi^2(1) = 3$; $p \leq 0,08$). Les autres résultats traitant de faibles effectifs (moins de quatre appariements par critères), bien que non statistiquement validés, vont encore dans le sens d'une constante plaçant le mot le plus court plutôt en M1 lorsque deux mots non équisyllabiques sont mis en présence.

Le corpus TARN ET GARONNE, et les questionnaires d'enquête ÉLISE ET ZOÉ et PAUL ET PASCAL, du fait du faible nombre d'effectifs en jeu (moins de 19 pour le corpus TARN ET GARONNE, moins de sept pour ÉLISE ET ZOÉ et moins de cinq appariements testés pour PAUL ET PASCAL), n'apportent aucun résultat strictement valide d'un point de vue statistique. Relevons tout de même le fait que dans les questionnaires d'enquête PAUL ET PASCAL et ÉLISE ET ZOÉ, les paires « σ & $\sigma\sigma$ » et « σ & $\sigma\sigma\sigma$ » sont systématiquement préférées à leur contraire.

Enfin le corpus TOM ET ZOÉ présente encore une fois des résultats en contradiction avec ceux des précédents supports d'études. En effet, une minorité de 15,63% des mots monosyllabiques favorise la position M1 lorsqu'ils sont opposés à des mots trisyllabiques ($\chi^2(1) = 15,13$; $p \leq 0,0001$).

Cependant, nous nuancions à nouveau ces résultats. Le corpus TOM ET ZOÉ met régulièrement au jour des résultats opposés à ce que nous relevons dans l'ensemble des supports d'étude. De plus, il est d'une nature différente des autres supports d'études, induisant ainsi de l'hétérogénéité. Pour ces raisons, nous pensons pertinent de présenter les résultats des confrontations par nombre de syllabes en excluant TOM ET ZOÉ.

Nous obtenons alors la nouvelle répartition présentée dans le tableau ci-après :

TOTAL-TOM ET ZOÉ M1	M2						TOTAL M1=
	σ	σσ	σσσ	σσσσ	σσσσσ	σσσσσσ	
σ %	75 20,56	85 79,44	21 72,41	7 100			188
σσ %	22 20,56	86 32,18	59 67,82	5 41,67	1 50		173
σσσ %	8 27,59	28 32,18	29 35,29	6 35,29			71
σσσσ %	0 0	7 58,33	11 64,71	3 35,29	1 100	1 50	23
σσσσσ %		1 50	0 0				1
σσσσσσ %			1 50				1
TOTAL M2 =	105	207	121	21	2	1	

Tableau 21 : Total des confrontations du nombre de syllabes de M1 et M2 sans le corpus TOM ET ZOÉ

Comme le montre le tableau ci-dessus, nous pouvons statistiquement affirmer trois résultats liés à ce que nous présentions auparavant.

Les mots monosyllabiques opposés à des mots dissyllabiques occupent massivement M1 dans 79,44 % des 107 appariements testés sur ce critère ($\chi^2(1) = 37,09$; $p < 0,000\ 000\ 001$).

De plus, les monosyllabes mis en présence de termes de trois syllabes favorisent M1 dans 72,41% des appariements traités sur ce critère ($\chi^2(1) = 5,83$; $p < 0,02$).

Conformément à ce que nous présentions précédemment, les dissyllabiques se placent majoritairement en M1 dans 67,82 % des appariements testés, lorsqu'ils sont confrontés à des mots trisyllabiques ($\chi^2(1) = 11,05$; $p < 0,000\ 9$).

Nous observons également que les paires « σ & $\sigma\sigma\sigma$ » sont systématiquement préférées à leur contraire. Bien que ne pouvant valider ce résultat statistiquement, il semble très probable qu'un effectif d'appariements testés plus important (seuls sept appariements relevaient de ce critère) nous permettrait d'asseoir ce résultat.

On remarquera que les mots trisyllabiques associés à des mots de quatre syllabes n'occupent M1 que dans 35,29% des 17 appariements testés, ce qui paraît contredire le principe de la sélection de la position M1 par le mot le plus court. Cependant, ce résultat inattendu n'est pas statistiquement significatif ($\chi^2(1) = 1,47$; $p \leq 0,2$).

Enfin, les appariements associant des mots de quatre syllabes à des mots de cinq syllabes et des termes de quatre syllabes à des mots de six syllabes ne peuvent faire l'objet d'une conclusion compte tenu de la taille des échantillons testés.

4.1.1.4. Bilan et analyse des résultats concernant le nombre de syllabes

Au regard de l'ensemble des résultats relevés sur la question d'une corrélation entre la taille des mots et la sélection d'une place dans la coordination, il apparaît que :

- Dans 69,64% des appariements non équisyllabiques testés, les mots en M1 sont plus courts que ceux en M2 ;
- Nous observons également que dans l'absolu :
 - les monosyllabes sélectionnent majoritairement M1 dans 64,16% des termes monosyllabiques éprouvés ;
 - Les mots trisyllabiques privilégient M2 dans 63,02% des mots de trois syllabes présents dans nos supports d'études ;
 - 54,47 % des mots de deux syllabes présents dans nos supports d'étude ont tendance à sélectionner M2 ;
 - Ces trois résultats, pour les mots contenant de une à trois syllabes, mettent au jour une disposition graduelle dans laquelle plus un mot est court, plus il occuperait M1.
- Les précédents résultats sont confortés par l'observation relative du nombre de syllabes :
 - En effet, la confrontation relative par nombre de syllabes des mots en M1 et en M2 fait apparaître que lorsque sont opposés des termes monosyllabes à des mots de deux, trois ou quatre syllabes, sur un total de 143 paires examinées, dans 79,02% des cas, la forme privilégiée est « σ & n σ »⁴⁷ ($\chi^2(1) = 48,17$; $p < 0,000\ 000\ 000\ 004$) ;
 - D'une manière générale, nous pouvons affirmer que pour les mots contenant de une à trois syllabes, plus un mot est court, plus il occupe M1. Les mots monosyllabiques occuperaient plus M1 que les mots dissyllabiques qui seraient eux-mêmes plus présents en M1 que les mots trisyllabiques.

Il est utile de rappeler que compte tenu de la méthodologie que nous avons mise en place pour collecter nos données, si les paires issues des questionnaires d'enquête sont neutralisées et ne traitent qu'un seul facteur à la fois, ce n'est absolument pas le cas de celles extraites des

⁴⁷ $n > 1$

corpus. C'est pourquoi les résultats que nous présentons ici sont d'autant plus massifs. Ils émergent, malgré tout, et très nettement, de données pouvant être l'objet d'un conflit de contraintes.

Pour illustrer nos propos, revenons sur les résultats que nous relevons pour les paires « σ & n σ »⁴⁸. Sur les 143 paires étudiées, 130 pourraient potentiellement être influencées par d'autres facteurs. Pourtant, nous relevons un résultat massif : la forme privilégiée est « σ & n σ »⁴⁹ dans 79,02% des cas. De fait, la corrélation entre la longueur syllabique et la place dans le binôme coordonné est un critère fort.

Face à la prégnance d'un tel résultat, nous avons décortiqué les 30 paires ne sélectionnant pas la forme « σ & n σ »⁵⁰. Nous les avons observées en fonction des facteurs que nous développerons à venir (pages 145 et 192), à savoir : leurs segments en frontière de mots - par phonèmes et par traits-, ainsi que leurs structures syllabiques en initiale et en finale.

Or, alors que pour l'ensemble de nos supports d'étude, ces examens révèlent l'existence d'un certain nombre de facteurs actifs dans la sélection d'un ordre préférentiel de coordination, pour ces 30 mots en question, aucun résultat n'est probant.

A priori, nous ne pouvons imputer la sélection de la forme « n σ & σ » aux facteurs phonologiques étudiés.

Nous nous sommes donc penchée sur les facteurs extraphonologiques que nous développons page 74. Le résultat le plus saillant de cet examen, même s'il révèle que 64% des formes « n σ & σ » présentes dans nos supports d'étude sont impactées par des facteurs sémantiques, n'est cependant pas statistiquement significatif ($\chi^2(1) = 1,96$; $p < 0,16$).

En fait, ici encore, nous ne relevons aucun résultat.

Cette absence de résultat pourrait indiquer que nous n'avons pas examiné les bons critères. Il reste néanmoins plus probable que l'échantillon de 30 mots divisé par le nombre d'hypothèses testées soit trop restreint pour faire émerger des résultats.

Quoi qu'il en soit, il en demeure qu'il existe un lien entre le nombre de syllabes de mots et une place privilégiée dans la coordonnée.

Comment expliquer de ces résultats ?

Les travaux de la littérature spécialisée présentés page 14 rendent également compte de ce

48 n>1

49 idem

50 idem

facteur de taille. Pour ne citer qu'eux : Pinker et Birdsong (1979), Cooper et Ross (1975) , Wright *et al.* (2002,2005), Renner (2006) et Benor et Levy (2006), tous relèvent l'existence d'une corrélation entre la taille des mots et le positionnement dans le binôme coordonné. Pour tous, comme dans nos travaux, plus le mot est court, plus il occupe M1.

L'explication sans doute la plus vraisemblable de ce phénomène, en tout cas la plus souvent donnée, est le lien que Cooper et Ross établissent entre la taille des mots et leur fréquence. Selon un principe⁵¹ qu'ils attribuent à Panini, les mots les plus fréquents sont les plus courts. Taille et fréquence des mots sont liées. Les mots fréquents sont courts. Les mots courts occupent M1, tout comme les mots fréquents.

En fait, il semblerait que le facteur de taille ne soit pas un facteur exclusivement phonologique. Nous pourrions même envisager qu'il ne s'agisse pas, à proprement dit, d'un facteur autonome. Le facteur de taille pourrait n'être qu'un artefact du facteur de fréquence.

De fait, nous pouvons justifier l'occupation des mots les plus courts en M1 de la même façon que pour les mots les plus fréquents. Ainsi, comme nous le signalons page 85, l'opposition marqué / non marqué peut proposer une explication vraisemblable à l'assignation préférentielle des mots en M1. Les mots fréquents sont moins marqués, et nous savons que les éléments les moins marqués sélectionnent prioritairement la position la moins marquée.

De façon un peu syllogistique, nous pensons qu'il est possible d'associer la valeur non marquée à la taille des mots. Les mots les plus courts seraient non marqués, comme les mots les plus fréquents.

Si l'opposition marqué / non marqué motive vraisemblablement la présence des mots les plus courts en M1, nous pensons qu'un principe prosodique peut également justifier leur présence à cette place.

Les chercheurs (citons Di Cristo (2000), Jun et Fougeron (2000), Delais-Roussarie (2000)) s'accordent sur le fait que le système accentuel du français est caractérisé par un accent primaire en syllabe finale accentuable.

Mais, si la caractérisation dans les systèmes accentuels du français de l'accent primaire final est consensuelle, ce n'est pas le cas ni du pied, ni de l'accent secondaire. Si l'accent primaire final est fixe, l'accent secondaire, en initiale de mots est optionnel.

Des corrélats acoustiques tels que la montée précoce de la fréquence fondamentale en initiale de mots ont permis de caractériser l'accent secondaire. Ainsi, grâce à ces indices acoustiques,

51 Ce principe est également connu comme loi de Zipf (cf. note de bas de page, page 88) ou « Principe poétique de gradation syllabique », Jakobson (1963).

Goad et Buckley (2006) proposent que le système accentuel du français, bien que donnant la priorité au pied iambique, possède deux pieds : un pied trochaïque en frontière gauche du mot et un pied iambique, en frontière droite.

Le patron accentuel des formes coordonnées, pour lesquelles nous avons relevé des résultats statistiquement forts, pourrait donc être schématisé comme suit :⁵²

- « σ & σσ » : /2001/
- « σ & σσσ » : /20201/
- « σσ & σσσ » : ou /200201/

Au regard de ces schémas, nous émettons l'hypothèse que dans la lignée des travaux de Jespersen (1905) cité par Renner qui avançait, pour l'anglais, que la séquence « monosyllabe-coordonnant-disyllabe » est privilégiée, car elle permet une alternance accentuelle de type trochaïque : il est probable que l'ordre des séquences binomiales coordonnées en français tende également à respecter une alternance accentuelle.

Cependant, il est entendu que n'examinant pas les productions orales de nos données, nous ne pouvons proposer un relevé précis de leurs patrons accentuels. A plus forte raison, comme les binômes que nous observons sont présentés hors contexte, il nous est impossible de prédire où tombent les accents.

Pour exemple, les trois paires « bonne et due », « bonne et heureuse » et « petites et moyennes » nous questionnent particulièrement : quel est leur patron accentuel ? Est-ce : « bonne et DUE » ou « bonne et due FORME » ? Considérons-nous « bonne et heuREUSE » ou « bonne et heureuse anNÉE » ? « Petites et moyENnes » ou « petites et moyennes entrePRISe »?⁵³

Compte tenu de la nature de ce présent travail de thèse, bien que convaincue de l'intérêt de l'hypothèse que nous avançons ici, il nous est impossible de la valider. Une étude des réalisations orales en contexte des formes préférentielles des séquences binaires coordonnées par « et » devrait apporter des réponses à ces questionnements.

52 Rappelons que pour nos schémas /1/ représente l'accent primaire, /2/, l'accent secondaire, /0/, les syllabes inaccentuées. Notons que l'accent ne tombera jamais sur la conjonction de coordination, et que deux accents ne peuvent être consécutifs.

53 Pour ces exemples, nous ne proposons que la représentation de l'accent primaire que nous mentionnons ici en majuscule.

4.1.2. Initiale vocalique et nombre de syllabes

4.1.2.1. Longueur relative -#V_ de M1 et de M2

Il s'agit ici d'examiner en détail la longueur de mots commençant par une initiale vocalique.

La question que nous nous sommes posée est la suivante : la voyelle initiale entre-t-elle dans le décompte des syllabes de M ? A combien tombe le nombre de syllabes de M1 ou de M2, en extrayant du décompte syllabique l'ensemble des syllabes initiales commençant par une voyelle #V_ ?

En effet, des travaux ont montré (Plénat, (1994, 2009), Villoing (2012)) que les syllabes initiales dépourvues d'attaque pourraient ne pas entrer dans le décompte syllabique. A savoir, par exemple, que les mots dissyllabiques commençant par une voyelle initiale se comporteraient comme des monosyllabes. Plénat parle d'une « extramétrie des voyelles initiales ».

Nous avons cherché à déterminer si ce phénomène déjà connu était actif dans la coordination binaire de deux mots par « et ».

Pour cela, nous avons relevé les différents supports d'études pouvant apporter une réponse à cette question, en l'occurrence les questionnaires d'enquête ÉLISE ET ZOÉ et PAUL ET PASCAL qui considèrent précisément l'hypothèse de la corrélation pouvant exister entre la taille d'un mot dont on a extrait du décompte syllabique la voyelle initiale et la sélection d'une position dans la coordination et la collection de six corpus.

Le nouveau décompte syllabique a été obtenu en suivant la méthode décrite en 1.2.1. Transcription du corpus et décompte syllabique. Pour chaque mot possédant une syllabe initiale dépourvue d'attaque, nous avons retiré cette syllabe du nombre total de syllabes du mot. Ainsi, pour reprendre un précédent exemple, selon ce nouveau décompte « Élise » n'est pas un dissyllabe mais un monosyllabe.

A noter, dès à présent que pour les deux premiers examens de cette section, à savoir la corrélation entre position dans la coordination et longueur syllabique -#V_, et la corrélation entre position dans la coordination et nombre de syllabes -#V_, nous avons comptabilisé des termes ayant pour patron syllabique V(C)(C)# ou V(C)(C)σ# comme des monosyllabes. Ainsi « Élise » et « or » seront tous deux des mots monosyllabiques. La distinction entre V(C)(C)# et V(C)(C)σ# sera introduite lors de la confrontation du nombre de syllabes -#V_ entre M1 et

M2.

Nous avons donc appliqué ce nouveau décompte syllabique à 1 877 appariements que nous avons testés desquels nous avons extrait 940 paires équisyllabiques ne nous permettant pas de déterminer un rôle actif ou non de l'extramétrie des voyelles initiales⁵⁴. Notons que sont considérées comme paires équisyllabiques, les binômes associant des mots $n\sigma$ à d'autres mots $n\sigma$, mais également ceux possédant $n\sigma$ aux termes $V(C)(C).n-1\sigma$ ⁵⁵.

Ainsi, comme le montre le tableau suivant, sur 937 paires non équisyllabiques testées, dans 570 paires non équisyllabiques $-#V_$, le nombre de syllabes $-#V_$ de M1 est inférieur à celui de M2 (60,83%) et inversement pour 367 binômes, le nombre de syllabes $-#V_$ de M1 est inférieur à celui de M2 (39,17%). A nouveau, nous avons noté en caractères gras les résultats pour lesquels le χ^2 est de $p \leq 0,05$. Dans le tableau ci-dessous : $\chi^2(1) = 43,98$; $p \leq 0,000\ 000\ 000\ 3$.

TOTAL	(-) #V	
	$n\sigma M1 < n\sigma M2$	$n\sigma M2 < n\sigma M1$
937		
M1	570 60,83	
M2		367 39,17
Total :	937 100,00%	

Tableau 22 : Position et longueur syllabique $-#V_$

Au regard de ces éléments, nous pouvons rejeter l'hypothèse nulle et affirmer l'existence d'une propension corrélant la taille des mots $-#V_$ à leur position dans la coordination. Selon le nouveau décompte syllabique excluant les syllabes commençant par une voyelle, les mots les plus courts ont statistiquement plus tendance à occuper majoritairement M1.

Afin d'asseoir encore plus explicitement ce premier résultat fort, il nous a semblé pertinent de présenter nos résultats, indépendamment, support d'études par support d'études.

54 Nous noterons l'extramétrie des voyelles initiales par $-#V_$ qui signale le retrait des syllabes initiales vocaliques du décompte syllabique.

55 « n » symbolise ici la variable du nombre de syllabes.

TARN ET GARONNE	(-) #V	
	nσM1<nσM2	nσM2<nσM1
43		
M1	29 67,44	
M2		14 32,56
Total :	43 100,00%	

Tableau 22a : Position et longueur syllabique -#V_ dans TARN ET GARONNE

CUL ET CHEMISE	(-) #V	
	nσM1<nσM2	nσM2<nσM1
16		
M1	14 87,5	
M2		2 12,5
Total :	16 100,00%	

Tableau 22b : Position et longueur syllabique -#V_ dans CUL ET CHEMISE

LISSE ET SOYEUX	(-) #V	
	nσM1<nσM2	nσM2<nσM1
36		
M1	27 75	
M2		9 25
Total :	36 100,00%	

Tableau 22c : Position et longueur syllabique -#V_ dans LISSE ET SOYEUX

TOM ET ZOÉ	(-) #V	
	nσM1<nσM2	nσM2<nσM1
663		
M1	371 55,96	
M2		292 44,04
Total :	663 100,00%	

Tableau 22d : Position et longueur syllabique -#V_ dans TOM ET ZOÉ

VITE ET BIEN	(-) #V	
	nσM1<nσM2	nσM2<nσM1
107		
M1	64 59,81	
M2		43 40,19
Total :	107 100,00%	

Tableau 22e : Position et longueur syllabique -#V_ dans VITE ET BIEN

ARTS ET MÉTIERS	(-) #V	
	nσM1<nσM2	nσM2<nσM1
52		
M1	46 88,46	
M2		6 11,54
Total :	52 100,00%	

Tableau 22f : Position et longueur syllabique -#V_ dans ARTS ET MÉTIERS

ÉLISE ET ZOÉ	(-) #V	
	nσM1<nσM2	nσM2<nσM1
12		
M1	11 91,67	
M2		1 8,33
Total :	12 100,00%	

Tableau 22g : Position et longueur syllabique -#V_ dans ÉLISE ET ZOÉ

PAUL ET PASCAL	(-) #V	
	nσM1<nσM2	nσM2<nσM1
8		
M1	8 100	
M2		0 0
Total :	8 100,00%	

Tableau 22h : Position et longueur syllabique -#V_ dans PAUL ET PASCAL

Les tableaux ci-dessus présentent la répartition des mots en fonction du décompte syllabique tenant compte de l'extramétrie des syllabes initiales commençant par une voyelle.

Sur les huit supports d'études testés, sept présentent des résultats statistiquement valides ($p \leq 0,05$). Le huitième, PAUL ET PASCAL, ne peut être validé par un Khi Deux. Son échantillon est trop petit pour que le test de significativité soit applicable.

Dans le détail, PAUL ET PASCAL et ÉLISE ET ZOÉ sont les questionnaires d'enquête qui ont été mis en œuvre pour tester l'hypothèse générale d'un lien entre la position dans la coordination et la taille des mots (avec ou sans décompte des syllabes initiales débutant par une voyelle). Or, nous remarquons que pour ces deux questionnaires, les résultats sont massifs : selon le nouveau décompte syllabique excluant les syllabes initiales vocaliques, les mots les plus courts sélectionnent M1 prioritairement.

Dans ÉLISE ET ZOÉ, le mot le plus court -#V_ favorise prioritairement M1 dans 91,67% des paires non équissyllabiques -#V_ testées ($\chi^2(1) = 8,33$; $p \leq 0,003$). Le questionnaire d'enquête PAUL ET PASCAL, quant à lui, montre que M1 est systématiquement plus court que M2. Bien que ce résultat ne soit pas statistiquement aussi fort que nous le souhaiterions, étant donné que PAUL ET PASCAL teste spécifiquement l'hypothèse d'une relation entre la position dans la coordination et le nombre de syllabes des mots _et que de fait, tout autre facteur a été neutralisé en amont_, nous pensons qu'il est pertinent de le conserver.

Les corpus CUL ET CHEMISE, ARTS ET MÉTIERS et LISSE ET SOYEUX font tous trois apparaître des résultats conséquents.

En effet, le corpus CUL ET CHEMISE atteste un résultat fort dans lequel le mot le plus court, selon le nouveau décompte syllabique tenant compte de l'extramétrie des syllabes initiales vocaliques, occupe M1 dans 87,5 % des binômes non équissyllabiques testés ($\chi^2(1) = 9$; $p \leq 0,002$).

Dans le corpus ARTS ET MÉTIERS, sur 52 binômes non équissyllabiques -#V_ testés, pour 46 formes coordonnées, le nombre de syllabes -#V_ de M1 est inférieur à celui de M2, soit dans 88,46% des cas ($\chi^2(1) = 30,77$; $p \leq 0,000\ 000\ 03$).

Concernant le corpus LISSE ET SOYEUX, sur 36 binômes non équissyllabiques -#V_, dans 27 paires, le nombre de syllabes -#V_ de M1 est inférieur à celui de M2 soit 75% des paires testées ($\chi^2(1) = 9$; $p \leq 0,003$).

De la même façon, les corpus TARN ET GARONNE et VITE ET BIEN révèlent des

répartitions similaires, à savoir que respectivement 67,44% et 59,81% des occurrences non équisyllabiques -#V_ testées placent prioritairement le mot le plus court -#V_ en M1, ($\chi^2(1) = 5,23$; $p \leq 0,02$ et $\chi^2(1) = 4,12$; $p \leq 0,04$).

Enfin, le corpus TOM ET ZOÉ présente des résultats bien moins massifs que ceux relevés dans les précédents support d'études. En effet, selon le nouveau décompte syllabique tenant compte de l'extramétrie des syllabes initiales commençant par une voyelle, une légère majorité des appariements non équisyllabiques expérimentés (55,96%) semblent favoriser le mot le plus long en M1 ($\chi^2(1) = 9,41$; $p \leq 0,002$).

Cependant, il est important de nuancer ces résultats.

D'une part, les résultats sont bien moins conséquents que ceux présentés dans les sept autres supports d'études testés sur la question de la taille -#V_. Et d'autre part, compte tenu de la nature particulière du corpus TOM ET ZOÉ, il est probable que d'autres facteurs extraphonologiques interfèrent dans la sélection d'une position dans la coordination binaire par « et ».

En raison de ces objections, nous pensons qu'il est pertinent de présenter une nouvelle répartition qui exclut les données fournies par TOM ET ZOÉ sur la question de la taille des mots selon le nouveau décompte syllabique tenant compte de l'extramétrie des initiales vocaliques. Nous nommons ce total « TOTAL-TOM ET ZOÉ ».

Ainsi, nous obtenons la nouvelle répartition présentée ci-dessous :

TOTAL-TOM ET ZOÉ	(-) #V	
	nσM1<nσM2	nσM2<nσM1
274		
M1	199 72,63	
M2		75 27,37
Total :	274 100,00%	

Tableau 23 : Position et longueur syllabique -#V_ dans TOTAL-TOM ET ZOÉ

Comme le montre le tableau ci-dessus, selon le décompte syllabique tenant compte de l'extramétrie des syllabes initiales commençant par une voyelle, pour 274 binômes non équisyllabiques-#V_ testés, dans 199 paires, le nombre de syllabes -#V_ de M1 est inférieur à

celui de M2 (72,63%), et inversement, pour 78 occurrences, le nombre de syllabes -#V_ de M1 est inférieur à celui de M2 (27,37%) : $\chi^2(1) = 56,12$; $p \leq 0,000\ 000\ 000\ 000\ 1$).

Au regard de ces éléments, il existe une très forte présomption contre l'hypothèse nulle. Nous avançons donc l'existence d'une corrélation forte liant la taille -#V_ des mots à leur position dans la coordination. Selon le décompte syllabique tenant compte de l'extramétricité des syllabes initiales commençant par une voyelle, les mots les plus courts ont statistiquement plus tendance à occuper majoritairement M1.

4.1.2.2. Nombre de syllabes -#V_ de M1 et M2

Nous avons souhaité approfondir ce résultat en nous interrogeant plus précisément sur le nombre de syllabes -#V_ de M1 et de M2.

Combien de syllabes M1 et M2 possèdent-ils ? Une⁵⁶, deux, trois, quatre et plus⁵⁷ ?

Les supports d'études nous permettant d'examiner ces hypothèses sont les mêmes que ceux utilisés précédemment pour déterminer le lien existant entre la position dans la coordination et la longueur syllabique -#V_, à savoir la collection complète des six corpus et les questionnaires d'enquête ÉLISE ET ZOÉ et PAUL ET PASCAL.

Dans les deux sections à venir, nous utilisons le décompte syllabique et les termes suivants :

	nombre de σ -#V_			
Patrons syllabiques	V(C)(C)# V(C)(C) σ # ou σ #	V(C)(C) $\sigma\sigma$ # ou $\sigma\sigma$ #	V(C)(C) $\sigma\sigma\sigma$ # ou $\sigma\sigma\sigma$ #	V(C)(C) $\sigma\sigma\sigma\sigma$ et +# ou $\sigma\sigma\sigma\sigma$ et +#
Dénominations	monosyllabe	dissyllabe	trisyllabe	mot de quatre syllabes et+

Tableau 24 : Décompte syllabique -#V_

Les résultats proposés dans le tableau ci-dessous présentent la répartition en M1 et en M2 des mots en fonction de leur nombre de syllabes -#V_.

56 Comme précisé dans le tableau ci-dessus, la catégorie des monosyllabes rassemble, pour cet examen du nombre de syllabes -#V_, des termes ayant pour patron syllabique V(C)(C)#, V(C)(C) σ # ou σ #.

57 Comme précédemment, compte tenu du peu de mots de quatre, cinq, et six syllabes, nous avons choisi de les regrouper, dans un premier temps, en une seule catégorie, celle des mots de quatre syllabes et plus.

	nombre de σ			
	V(C)(C)# V(C)(C) σ # ou σ #	V(C)(C) $\sigma\sigma$ # ou $\sigma\sigma$ #	V(C)(C) $\sigma\sigma\sigma$ # ou $\sigma\sigma\sigma$ #	V(C)(C) $\sigma\sigma\sigma\sigma$ et +# ou $\sigma\sigma\sigma\sigma$ et+#
TOTAL 1877				
M1	549 57,91	1123 48,74	191 40,81	14 41,18
M2	399 42,09	1181 51,26	277 59,19	20 58,82
Total :	948 100,00%	2304 100,00%	468 100,00%	34 100,00%

Tableau 25 : Position et nombre de syllabes -#V_

Ainsi comme le montre le tableau *supra*, selon le décompte syllabique tenant compte de l'extramétrie des syllabes initiales commençant par une voyelle, nous pouvons affirmer deux résultats : les monosyllabes occupent majoritairement M1 pour 57,91% des monosyllabes testés ($\chi^2(1) = 23,73$; $p \leq 0,000\ 001$) et, de façon inversement proportionnelle, les mots trisyllabiques semblent favoriser prioritairement M2 dans 59,19% des termes trisyllabiques testés ($\chi^2(1) = 15,8$; $p \leq 0,000\ 07$)

Concernant les mots de deux syllabes et ceux de quatre syllabes et plus, selon le décompte tenant compte de l'extramétrie des syllabes initiales sans attaque, nous ne pouvons asseoir statistiquement nos résultats ($p > 0,1$). Cependant, il est intéressant de relever ici encore les contours d'une distribution graduelle décroissante de l'occupation de M1 dans laquelle il apparaîtrait que plus un mot est court -#V_, plus il aurait une prédisposition à sélectionner la première place dans la coordination.

Afin de scruter plus encore la distribution graduelle décroissante de l'occupation de M1 que nous relevons ci-dessus, il nous a semblé pertinent de présenter nos résultats, indépendamment, support d'études par support d'études.

TARN ET GARONNE	nombre de σ			
	V(C)(C)# V(C)(C)σ# ou σ#	V(C)(C)σσ# ou σσ#	V(C)(C)σσσ# ou σσσ#	V(C)(C)σσσσ et +# ou σσσσ et+#
69	24	28	15	2
M1	64,86	45,16	41,67	66,67
M2	13	34	21	1
	35,14	54,84	58,33	33,33
Total :	37 100,00%	62 100,00%	36 100,00%	3 100,00%

Tableau 26a : Position et nombre de syllabes -#V_ dans TARN ET GARONNE

LISSE ET SOYEUX	nombre de σ			
	V(C)(C)# V(C)(C)σ# ou σ#	V(C)(C)σσ# ou σσ#	V(C)(C)σσσ# ou σσσ#	V(C)(C)σσσσ et +# ou σσσσ et+#
59	24	29	5	1
M1	64,86	50,88	26,32	20
M2	13	28	14	4
	35,14	49,12	73,68	80
Total :	37 100,00%	57 100,00%	19 100,00%	5 100,00%

Tableau 26c : Position et nombre de syllabes -#V_ dans LISSE ET SOYEUX

CUL ET CHEMISE	nombre de σ			
	V(C)(C)# V(C)(C)σ# ou σ#	V(C)(C)σσ# ou σσ#	V(C)(C)σσσ# ou σσσ#	V(C)(C)σσσσ et +# ou σσσσ et+#
42	39	2	1	0
M1	57,35	18,18	25	0
M2	29	9	3	1
	42,65	81,82	75	100
Total :	68 100,00%	11 100,00%	4 100,00%	1 100,00%

Tableau 26b : Position et nombre de syllabes -#V_ dans CUL ET CHEMISE

TOM ET ZOÉ	nombre de σ			
	V(C)(C)# V(C)(C)σ# ou σ#	V(C)(C)σσ# ou σσ#	V(C)(C)σσσ# ou σσσ#	V(C)(C)σσσσ et +# ou σσσσ et+#
1420	318	972	130	
M1	53,63	50,21	41,8	
M2	275	964	181	
	46,37	49,79	58,2	
Total :	593 100,00%	1936 100,00%	311 100,00%	0 100,00%

Tableau 26d : Position et nombre de syllabes -#V_ dans TOM ET ZOÉ

	nombre de σ			
	V(C)(C)# V(C)(C) σ # ou σ #	V(C)(C) $\sigma\sigma$ # ou $\sigma\sigma$ #	V(C)(C) $\sigma\sigma\sigma$ # ou $\sigma\sigma\sigma$ #	V(C)(C) $\sigma\sigma\sigma\sigma$ et +# ou $\sigma\sigma\sigma\sigma$ et+#
VITE ET BIEN 184	77 59,23	64 41,83	32 47,06	11 64,71
M1	53 40,77	89 58,17	36 52,94	6 35,29
M2	130 100,00%	153 100,00%	68 100,00%	17 100,00%
Total :				

Tableau 26e : Position et nombre de syllabes -#V_ dans VITE ET BIEN

	nombre de σ			
	V(C)(C)# V(C)(C) σ # ou σ #	V(C)(C) $\sigma\sigma$ # ou $\sigma\sigma$ #	V(C)(C) $\sigma\sigma\sigma$ # ou $\sigma\sigma\sigma$ #	V(C)(C) $\sigma\sigma\sigma\sigma$ et +# ou $\sigma\sigma\sigma\sigma$ et+#
ÉLISE ET ZOÉ 20	11 73,33	8 42,11	1 16,67	
M1	4 26,67	11 57,89	5 83,33	
M2	15 100,00%	19 100,00%	6 100,00%	0 100,00%
Total :				

Tableau 26g : Position et nombre de syllabes -#V_ dans ÉLISE ET ZOÉ

	nombre de σ			
	V(C)(C)# V(C)(C) σ # ou σ #	V(C)(C) $\sigma\sigma$ # ou $\sigma\sigma$ #	V(C)(C) $\sigma\sigma\sigma$ # ou $\sigma\sigma\sigma$ #	V(C)(C) $\sigma\sigma\sigma\sigma$ et +# ou $\sigma\sigma\sigma\sigma$ et+#
ARTS ET MÉTIERS 75	48 80	20 33,33	7 31,82	0 0
M1	12 20	40 66,67	15 68,18	8 100
M2	60 100,00%	60 100,00%	22 100,00%	8 100,00%
Total :				

Tableau 26f : Position et nombre de syllabes -#V_ dans ARTS ET MÉTIERS

	nombre de σ			
	V(C)(C)# V(C)(C) σ # ou σ #	V(C)(C) $\sigma\sigma$ # ou $\sigma\sigma$ #	V(C)(C) $\sigma\sigma\sigma$ # ou $\sigma\sigma\sigma$ #	V(C)(C) $\sigma\sigma\sigma\sigma$ et +# ou $\sigma\sigma\sigma\sigma$ et+#
PAUL ET PASCAL 8	8 100	0 0	0 0	
M1	0 0	6 100	2 100	
M2	8 100,00%	6 100,00%	2 100,00%	0 100,00%
Total :				

Tableau 26h : Position et nombre de syllabes -#V_ dans PAUL ET PASCAL

Comme précédemment, nous avons noté en caractères gras les résultats pour lesquels le χ^2 est de $p \leq 0,05$ et en italiques ceux pour lesquels, il est compris entre $0,05 < p \leq 0,1$. De même, pour ne alourdir la présentation de nos résultats, nous ne présenterons, en général, que ceux obtenus pour la sélection de M1.

Les tableaux *supra* mettent au jour les résultats relevés pour chaque support d'études, sur le traitement de la question du nombre de syllabes -#V_ en rapport avec la sélection d'une position dans la coordonnée.

Dans le détail, tous les supports d'études, exception faite du questionnaire d'enquête PAUL ET PASCAL, révèlent au moins une dispersion statistiquement valide ($p \leq 0,05$).

Dans le détail, observons tout d'abord les questionnaires d'enquête.

PAUL ET PASCAL et ÉLISE ET ZOÉ proposent tous deux des binômes permettant l'examen spécifique de la longueur des mots. Du fait de la neutralisation des autres facteurs non testés, ils font apparaître des résultats forts. Ainsi, dans PAUL ET PASCAL, 100% des mots monosyllabiques occupent M1 et inversement, les mots de deux et trois syllabes privilégient systématiquement M2. Aucun de ces deux résultats ne peut prétendre à une validation statistique. La valeur théorique des échantillons testés étant inférieure à cinq, le χ^2 n'est pas probant. Nous sommes néanmoins convaincue que les résultats issus de PAUL ET PASCAL traitant spécifiquement ces hypothèses sont pertinents.

Le questionnaire d'enquête ÉLISE ET ZOÉ met au jour une tendance statistique : 73,33% des monosyllabes favorisent la position M1 ($\chi^2 (1) = 3,27 ; p \leq 0,07$). Notons également la même répartition graduelle que nous observions page 118 : 73,33% des monosyllabes occupent M1 pour 42,11 % des mots dissyllabiques et 16,67% des termes trisyllabiques.

Dans le corpus ARTS ET MÉTIERS, on observe nettement deux résultats conséquents : les mots monosyllabiques occupent prioritairement M1 (80% des monosyllabes testés dans ce corpus sélectionnent la première position dans la coordonnée ($\chi^2 (1) = 21,6 ; p \leq 0,000 003$), inversement ceux de deux syllabes privilégient M2 dans 66,67 % des cas ($\chi^2 (1) = 6,67 ; p \leq 0,009$). Par ailleurs, nous relevons une tendance statistique ($\chi^2 (1) = 2,91 ; p \leq 0,08$) plaçant une minorité de mots de trois syllabes (31,82%) en M1.

A nouveau, nous retrouvons une dispersion graduelle décroissante. Pour le corpus ARTS ET MÉTIERS, 80% des mots monosyllabiques favorisent la position M1, pour 33,33 % des termes dissyllabiques, 31,82% des mots trisyllabiques et 0% des mots de quatre syllabes et

plus.

Dans le corpus CUL ET CHEMISE, nous relevons un résultat fort pour lequel il existe une très forte présomption contre l'hypothèse nulle ($\chi^2(1) = 5,44$; $p \leq 0,02$). En effet, les mots de deux syllabes semblent éviter majoritairement M1, ils ne sont que 18,18% à sélectionner cette position de la coordination.

Concernant le corpus VITE ET BIEN, nous relevons deux répartitions significatives : les mots monosyllabiques sont 59,23 % à occuper majoritairement (60,14%) la première partie de la coordination ($\chi^2(1) = 4,43$; $p \leq 0,04$), Inversement, les mots de deux syllabes évitent plutôt d'occuper M1 ; ils ne sont que 41,83% à préférer cette position ($\chi^2(1) = 4,08$; $p \leq 0,04$).

Notons également les résultats des mots de trois syllabes et ceux de quatre syllabes et plus -#V_. Si aucun des deux n'est statistiquement valide ($p > 0,1$), nous remarquons le fait que ni les mots trisyllabiques, ni les mots de quatre syllabes et plus, ne respectent la dispersion graduelle décroissante qui semblait pourtant se confirmer. En effet, les mots trisyllabiques occupent M1 dans 47,06% des mots trisyllabiques testés, soit un peu plus que les mots de deux syllabes. Quant aux mots de quatre syllabes et plus -#V_, dans le corpus VITE ET BIEN, tout comme nous le verrons ci-dessous dans le corpus TARN ET GARONNE, favorisent pour la première fois dans nos supports d'études M1 dans 64,71%. Certes, ces résultats ne sont pas fiables d'un point de vue statistique, mais nous nous attacherons néanmoins à les examiner en détail lors de la confrontation par nombre de syllabes -#V_ de M1 et M2. Pour ce qui est des résultats des trisyllabes -#V_, nous remarquons, tout de même, que comme dans l'ensemble de nos supports d'études, les mots de trois syllabes sélectionnent plutôt majoritairement M2.

Le corpus LISSE ET SOYEUX, quant à lui, suit scrupuleusement la dispersion graduelle décroissante précédemment relevée. A savoir, pour le corpus LISSE ET SOYEUX, 64,86% des mots monosyllabiques ont tendance à favoriser la position M1 ($\chi^2(1) = 3,27$; $p \leq 0,07$), pour 50,88 % des termes dissyllabiques ($\chi^2(1) = 0,02$; $p \leq 0,89$), 26,32% des trisyllabes ($\chi^2(1) = 4,26$; $p \leq 0,03$), et 20% des mots de quatre syllabes et plus testés ($\chi^2(1) = 1,8$; $p \leq 0,18$). Deux des résultats présentés sont validés statistiquement : les monosyllabes présentent une tendance statistique, les mots de trois syllabes et plus, une répartition significative.

Le corpus TARN ET GARONNE adopte lui-aussi les mêmes contours que les autres supports

d'études, à savoir que 64,86% des monosyllabes favorisent la position M1 ($\chi^2(1) = 3,27$; $p \leq 0,07$), pour 45,16 % des termes dissyllabiques ($\chi^2(1) = 0,58$; $p \leq 0,45$) et 41,67 % des trisyllabes ($\chi^2(1) = 1$; $p \leq 0,317$). Notons que seuls les monosyllabes présentent une dispersion significative.

Enfin, les trois mots de quatre syllabes et plus testés dans TARN ET GARONNE favorisent prioritairement M1 dans 66,67% des cas ($\chi^2(1) = 0,33$; $p \leq 0,56$). Intuitivement, nous imputons ce résultat accidentel au faible effectif éprouvé ; néanmoins, comme nous le annoncions précédemment, nous observerons les mots de quatre et plus -#V_ plus précisément lors de l'examen du nombre de syllabes de M1 et M2.

D'une manière générale, sur l'ensemble de ces supports d'études, nous pouvons nous accorder sur trois faits : le premier, c'est qu'étant donné le peu d'effectifs par type de mots, seuls des résultats très forts sont statistiquement valides. Nous sommes convaincue qu'un plus grand nombre d'effectifs confirmerait plus aisément les résultats en question. Le deuxième, c'est que la catégorie des mots de quatre syllabes et plus est ici encore hétéroclite. Les résultats relevés ne semblent pas cohérents d'un support d'études à l'autre. Enfin, ici encore, conformément aux résultats que nous avons relevé sur la corrélation entre la taille des mots (sans le décompte extramétrique des initiales dépourvues d'attaques) et leur position dans la coordination syndétique par « et », dans tous nos précédents support d'études, l'occupation de M1 suit une gradation décroissante des monosyllabes -#V_ aux trisyllabes -#V_ : plus un mot est court, plus il semble favoriser M1.

Enfin, le corpus TOM ET ZOÉ présente des résultats en adéquation avec ceux des précédents support d'études. En effet, à nouveau, nous retrouvons une dispersion graduelle décroissante dans laquelle une faible majorité de 53,63% des monosyllabes ont tendance à favoriser la position M1 ($\chi^2(1) = 3,12$; $p \leq 0,07$), pour une majorité hésitante de 50,21 % des termes dissyllabiques ($\chi^2(1) = 0,03$; $p \leq 0,85$) et seulement 41,8% des mots trisyllabiques ($\chi^2(1) = 8,36$; $p \leq 0,003$). Les résultats relevés pour les monosyllabes font apparaître une tendance statistique, ceux des trisyllabes, un résultat significatif.

Il convient cependant de nuancer ici encore ces résultats.

En effet, les résultats sont bien moins massifs que ceux présentés dans les autres supports d'études testés sur la question du nombre de syllabes. Ils semblent révélateurs d'une très grande hésitation. D'autant plus que malgré le nombre important de binômes coordonnées par

« et » testés dans ce corpus, la répartition des mots en fonction de leur nombre de syllabes n'est pas statistiquement significative pour les 1 936 mots de deux syllabes testant cette hypothèse.

Enfin, compte tenu de la nature du corpus, TOM ET ZOÉ est, d'une part, le support d'études le moins contrôlé de l'ensemble de nos données, et d'autre part, le corpus le plus imposant en termes de taille. Il est tout à fait envisageable que les résultats y émergeant biaisent l'ensemble de nos résultats.

Motivée par ces objections, il nous a semblé pertinent de proposer un traitement de la question concernant la corrélation entre une position dans la coordination et le nombre de syllabes excluant les données issues dans TOM ET ZOÉ.

C'est pourquoi dans le tableau à venir, nous présentons la répartition des mots en M1 et en M2 en fonction du décompte syllabique tenant compte de l'extramétrie des syllabes initiales sans attaque pour un ensemble de données excluant les résultats de TOM ET ZOÉ.

Nous nommons ici encore ce nouvel ensemble : « TOTAL-TOM ET ZOÉ ».

TOTAL-TOM ET ZOÉ	nombre de σ			
	V(C)(C)# V(C)(C) σ # ou σ #	V(C)(C) $\sigma\sigma$ # ou $\sigma\sigma$ #	V(C)(C) $\sigma\sigma\sigma$ # ou $\sigma\sigma\sigma$ #	V(C)(C) $\sigma\sigma\sigma\sigma$ et +# ou $\sigma\sigma\sigma\sigma$ et +#
457	231	151	61	14
M1	65,07	41,03	38,85	41,18
M2	124	217	96	20
	34,93	58,97	61,15	58,82
Total :	355	368	157	34
	100,00%	100,00%	100,00%	100,00%

Tableau 27 : Position et nombre de syllabes -#V_ TOTAL-TOM ET ZOÉ

Comme le montre le tableau ci-dessus, nous pouvons statistiquement affirmer trois résultats liés à ce que nous présentions lors de l'examen du nombre de syllabes -#V_ pour chaque support d'études. Les 355 monosyllabes testés occupent largement M1 pour 65,07% d'entre eux ($\chi^2(1) = 32,25$; $p \leq 0,000\ 000\ 01$). Sur 368 dissyllabes examinés, seuls 41,03% sélectionnent prioritairement M1 ($\chi^2(1) = 11,84$; $p \leq 0,000\ 6$). Enfin, seuls 61 des mots de trois syllabes sur les 157 testés, soit 38,85% ($\chi^2(1) = 7,8$; $p \leq 0,005$), préfèrent M1.

Nous pouvons donc affirmer que le décompte syllabique -#V_ met au jour une dispersion

graduelle décroissante _ des monosyllabes aux trisyllabes _, dans laquelle plus un mot est court, plus il sélectionne prioritairement M1. Nous notons également que seuls les mots monosyllabiques privilégient largement M1. Les dissyllabes et les trisyllabes favorisent plutôt M2

Enfin, concernant les mots de quatre syllabes et plus, sur 34 mots éprouvés, 14 occupent M1 (soit 41,18%) ($\chi^2(1) = 1,06$; $p \leq 0,30$). Nous n'obtenons pas de significativité statistique, et donc il nous est impossible d'affirmer quoi que ce soit à propos de l'existence d'un lien entre les mots de quatre syllabes et plus et ou l'une ou l'autre position de la coordination.

Jusqu'à présent, nous avons obtenu des résultats assez forts associant la taille -#V_ des mots à la sélection d'une place dans la coordination :

- Les mots en M1 sont plus courts que les mots en M2 dans 72,63% des appariements non équisyllabiques -#V_ testés ;
- L'examen du nombre de syllabes -#V_ montre que les monosyllabes occupent majoritairement M1 dans 65,07% des termes monosyllabiques éprouvés ;
- En revanche, les mots de deux et trois syllabes favorisent M2 : dans 58,97% des mots dissyllabiques et 61,15% des trisyllabes présents dans nos supports d'études.
- Enfin, comme le montrent les diagrammes ci-dessous, pour les mots contenant de une à trois syllabes selon le décompte métrique excluant les initiales dépourvues d'attaques, il apparaîtrait deux dispersions graduelles, déductibles l'une de l'autre, dans lesquelles plus un mot est court, plus il occuperait M1, et plus un mot est long, plus il favorise M2.⁵⁸

58 Ces deux dispersions sont volontairement redondantes, à ce stade de la présentation de nos travaux, nous avons arbitrairement choisi de présenter nos résultats selon la sélection de M1. Cependant, nous gardons à l'esprit, et cela sera discuté ultérieurement, que la sélection dans la coordination peut également se faire par non M1, par M2, par non M2 ou par rapport à « et ».

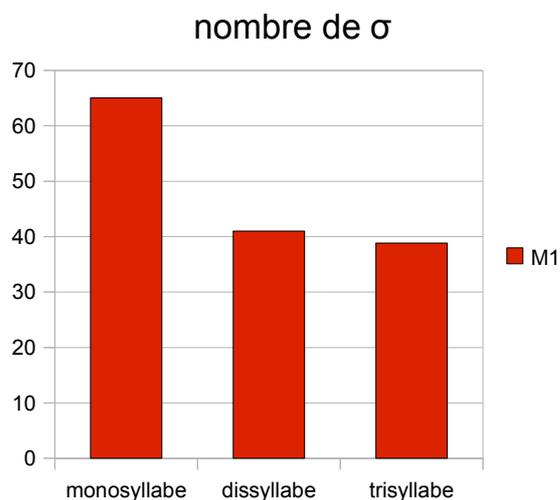


Diagramme 1 a : Pourcentage d'occupation de M1 par nombre de syllabes -#V_ (p ≤ 0,05)

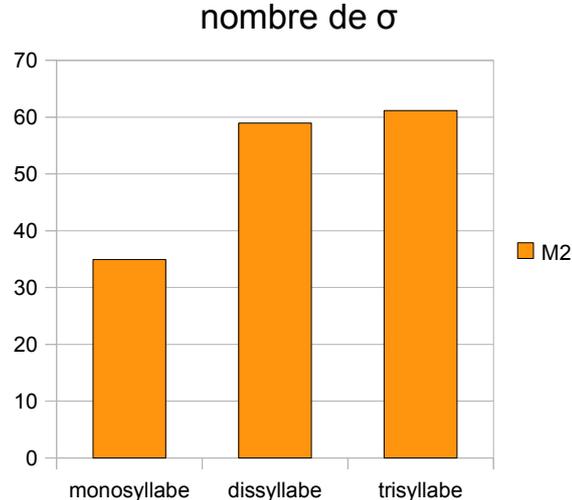


Diagramme 1 b : Pourcentage d'occupation de M2 par nombre de syllabes -#V_ (p ≤ 0,05)

4.1.2.3. Corrélation entre nombre de syllabes -#V_ de M1 et nombre de syllabes de M2

Les résultats précédents ont tous été relevés indépendamment de ce qui se passe dans le reste de la coordination. En effet, nous avons examiné la nombre de syllabes -#V_ dans l'absolu. Il convient à présent d'approfondir ces résultats. C'est pourquoi, dans la présente section, nous nous proposons d'observer le nombre de syllabes -#V_ de M1 relativement au nombre de syllabes -#V_ de M2.

Les supports d'études traitant de cette interrogation sont les mêmes que ceux précédemment utilisés pour déterminer le lien existant entre le nombre de syllabes selon le décompte syllabique excluant les syllabes initiales commençant par une voyelle et la position dans la coordination, à savoir la collection complète de six corpus et les questionnaires d'enquête ÉLISE ET ZOÉ et PAUL ET PASCAL. 1 877 binômes ont été testés.

A nouveau, nous avons noté en italique les résultats pour lesquels le χ^2 est compris entre $0,05 < p \leq 0,1$, et en caractères gras ceux pour lesquels $p \leq 0,05$. Nous avons signalé en clair les effectifs des paires équisyllabiques -#V_. Afin d'éviter de surcharger les tableaux à venir, nous avons délibérément laissé des cases vides correspondant aux types de confrontations

absentes de nos supports d'études. Les cases les plus sombres, quant à elles, reprennent pour des raisons purement pratiques les confrontations par nombre de syllabes M1 et M2 sans le décompte extramétrique. Nous les avons commentées précédemment, et ne le referons pas dans cette section.

Consciente de la difficulté à lire les résultats ainsi présentés, nous nous attacherons à relever et détailler chaque résultat pertinent en suivant.

Enfin, nous avons révisé l'organisation du décompte syllabique :

- Ainsi les monosyllabes tels qu'ils étaient comptabilisés jusqu'à présent sont désormais répartis en trois catégories : celle des « vrais » monosyllabes commençant par une attaque, notés σ , celle des « vrais » monosyllabes commençant par une voyelle notés $V(C)(C)\#$ et celle de monosyllabes $\#V_$, à savoir la catégorie des mots ayant pour patron $V(C)(C)\sigma\#$;
- De la même façon, les dissyllabes seront représentés par deux catégories : les dissyllabes commençant par une attaque notés $\sigma\sigma$, et les dissyllabes $\#V_$ répondant au patron $V(C)(C)\sigma\sigma\#$;
- Les trisyllabes commençant par une consonne seront notés $\sigma\sigma\sigma$, les trisyllabes $\#V_$ par le patron $V(C)(C)\sigma\sigma\sigma\#$;
- Enfin la catégorie des mots de quatre syllabes et plus a été subdivisée en catégories des quatre, cinq et six syllabes, notées respectivement $\sigma\sigma\sigma\sigma$ et $V(C)(C)\sigma\sigma\sigma\sigma\#$, $\sigma\sigma\sigma\sigma\sigma$ et $V(C)(C)\sigma\sigma\sigma\sigma\sigma\#$, et $\sigma\sigma\sigma\sigma\sigma\sigma$.

TOTAL		M2										
M1		σ	σσ	σσσ	σσσσ	σσσσσ	σσσσσσ	(-)#V_				
								V(C)(C)#	V(C)(C)σ#	V(C)(C)σσ#	V(C)(C)σσσ#	V(C)(C)σσσσ#
σ		94	177	26	7			4	21	15		
%			61,46	42,62	100			40	42,86	55,56		
σσ		111	744	310	5	1		1	77	88	0	
%		38,54		55,56	41,67	50		10	38,69	41,31	0	
σσσ		35	248	88	6				14	13	0	
%		57,38	44,44		64,71			0	38,89	25,49	0	
σσσσ		0	7	11	3	1	1	0	1	0	0	
%		0	58,33	35,29		100	50	0	100	0	0	
σσσσσ			1	0							0	0
%			50	0							0	0
σσσσσσ				1								
%				50								
(-)#V_	V(C)(C)#	6	9	1	3				2	2		
	%	60	90	100	100				50	66,67		
	V(C)(C)σ#	28	122	22	0			2	27	27	2	
	%	57,14	61,31	61,11	0			50		54	66,67	
	V(C)(C)σσ#	12	125	38	1			1	23	33	0	1
%	44,44	58,69	74,51	100			33,33	46		0	50	
V(C)(C)σσσ#		2	8	3	1			1	1	3		
%		100	100	100	100			33,33	100			
V(C)(C)σσσσ#					1				1			
%					100				50			

Tableau 28 : Position et nombre de syllabes -#V_

Le tableau ci-dessus fournit quatre résultats statistiquement remarquables confortant globalement les résultats que nous présentions précédemment ; à savoir que les mots les courts -#V_ ont plus tendance à occuper M1.

Nous pouvons donc affirmer que premièrement les mots dissyllabiques V(C)(C)σσ# sont majoritairement placés en M1 dans 74,51 % des cas lorsqu'ils sont confrontés à des trisyllabiques commençant par une attaque (σσσ) ($\chi^2(1) = 12,25$; $p \leq 0,0005$).

Les monosyllabes V(C)(C)σ# occupent prioritairement M1 dans 61,31% des cas lorsqu'ils sont opposés à des dissyllabiques commençant par une attaque (σσ) ($\chi^2(1) = 10,18$; $p \leq 0,001$).

Les mots V(C)(C)# sont massivement placés en M1 dans 90 % des cas lorsqu'ils sont confrontés à des dissyllabiques commençant par une attaque (σσ) ($\chi^2(1) = 6,4$; $p \leq 0,01$).

Enfin, les mots de deux syllabes V(C)(C)σσσ# sélectionnent prioritairement M1 (58,69 %) quand ils sont en présence de termes dissyllabiques commençant par une attaque (σσ) ($\chi^2(1) = 6,43$; $p \leq 0,01$).

Considérons plus en détail les autres résultats apparaissant dans ce tableau.

Mettons dès à présent de côté les dix paires totalisant moins de cinq occurrences, qui, bien que présentant des résultats massifs, ne peuvent en aucun cas faire l'objet d'examen

approfondis.⁵⁹ Le test de Khi Deux ne peut être probant sur de si petits échantillons.

Les cinq derniers résultats, bien que non validés par le test de Khi Deux ($p > 0,1$), sont en adéquation avec les résultats précédemment relevés.

En effet, il apparaîtrait que :

- Les mots en $V(C)(C)\#$ sélectionnent majoritairement M1 dans 60% des cas lorsqu'ils sont confrontés à des monosyllabes commençant par une consonne (σ) ;
- Les dissyllabes $V(C)(C)\sigma\sigma\#$ occupent M1 dans 57,14 % des cas lorsqu'ils sont opposés à σ ;
- σ sélectionnent prioritairement M1 (55,56 %) quand ils sont en présence de dissyllabes $V(C)(C)\sigma\sigma\#$;
- $V(C)(C)\#$ préfèrent majoritairement M1 dans 61,11% des cas lorsqu'ils sont confrontés à des trisyllabes commençant par une attaque ($\sigma\sigma\sigma$) ;
- Et enfin les monosyllabes $V(C)(C)\#$ occupent M1 dans 54 % des cas lorsqu'ils sont opposés à des mots dissyllabiques $V(C)(C)\sigma\sigma\#$.

Globalement⁶⁰, les mots les plus courts selon le décompte extramétrique sélectionnent majoritairement M1.

A nouveau afin d'asseoir plus explicitement les résultats prénants que nous avons relevés, il nous a semblé pertinent de présenter nos résultats, indépendamment, support d'études par support d'études.

Les tableaux ci-après présentent les résultats de chaque support d'études traitant de la confrontation de M1 et M2 par nombre de syllabes $\#V_.$

Dans le détail, un seul support d'études fournit des résultats statistiquement pertinents : le corpus TOM ET ZOÉ.

Afin de ne pas opacifier nos propos, nous ne présenterons que les résultats pour lesquels la distribution est majoritaire et statiquement pertinente. La répartition étant binaire, il ne nous paraît pas nécessaire de préciser des résultats aisément déductibles les uns des autres.

59 A savoir $\sigma\sigma$ vs. $V(C)(C)\sigma\sigma\sigma\#$, $\sigma\sigma\sigma$ vs. $V(C)(C)\#$, $\sigma\sigma\sigma\sigma$ vs. $V(C)(C)\#$, vs. $V(C)(C)\sigma\#$, vs. $V(C)(C)\sigma\sigma\#$, et vs. $V(C)(C)\sigma\sigma\sigma\#$, $\sigma\sigma\sigma\sigma\sigma$ vs. $V(C)(C)\sigma\sigma\sigma\#$, $V(C)(C)\sigma\sigma\sigma\sigma\#$, $V(C)(C)\#$ vs. $V(C)(C)\sigma\#$ et vs. $V(C)(C)\sigma\sigma\#$.

60 Nous nuancions nos propos par ce terme. En effet, une nouvelle hypothèse semble émerger. Nous la développerons à venir. Les résultats des trisyllabes $V(C)(C)\sigma\sigma\sigma\#$ favorisant M1 confrontés à des trisyllabes commençant par une consonne ($\sigma\sigma\sigma$) et ceux de $V(C)(C)\#$ préférant M1 en présence des monosyllabes débutant par une attaque (σ), nous laisse à penser que la qualité du segment initial peut être un critère déterminant dans la sélection d'une place dans la coordination. Nous examinerons ce phénomène en détail dès la page 145.

TARN ET GARONNE		M2										
M1	σ	σσ	σσσ	σσσσ	σσσσσ	σσσσσσ	(-)#V					
							V(C)(C)#	V(C)(C)σ#	V(C)(C)σσ#	V(C)(C)σσσ#	V(C)(C)σσσσ#	
σ	9	7	3									
%	87,5	60					0					
σσ	1	17	13					1	2			
%	12,45	68,42					0	50	66,67			
σσσ	2	6	9							0		
%	40	31,58					0	0				
σσσσ										1		
%										100		
σσσσσ												
%												
(-)#V_	V(C)(C)#	2	1									
	%	100	100									
	V(C)(C)σ#		1	4								
	%		50	100								
	V(C)(C)σσ#		1	1								
%		33,33	100									
V(C)(C)σσσ#												
%												
V(C)(C)σσσσ#												
%												

Tableau 29a : Position et nombre de syllabes -#V_ dans TARN ET GARONNE

LISSE ET SOYEUX		M2										
M1	σ	σσ	σσσ	σσσσ	σσσσσ	σσσσσσ	(-)#V					
							V(C)(C)#	V(C)(C)σ#	V(C)(C)σσ#	V(C)(C)σσσ#	V(C)(C)σσσσ#	
σ	4	7	5	1								
%	58,33	62,5	100				1	1				
σσ	5	13	10	2								
%	41,67	83,33	100				100	25	33,33			
σσσ	3	2	4	2								
%	37,5	16,67	66,67					0				
σσσσ	0	0	1									
%	0	0	33,33									
σσσσσ												
%												
σσσσσσ												
%												
(-)#V_	V(C)(C)#											
	%	0	0								0	
	V(C)(C)σ#	3	3	2							1	
	%	66,67	75	100								
	V(C)(C)σσ#		2									3
%		66,67							100			
V(C)(C)σσσ#												
%												
V(C)(C)σσσσ#												
%												

Tableau 29c : Position et nombre de syllabes -#V_ dans LISSE ET SOYEUX

CUL ET CHEMISE		M2										
M1	σ	σσ	σσσ	σσσσ	σσσσσ	σσσσσσ	(-)#V					
							V(C)(C)#	V(C)(C)σ#	V(C)(C)σσ#	V(C)(C)σσσ#	V(C)(C)σσσσ#	
σ	25	9		1								
%	75	100		100			3	1				
σσ	3	0	3									
%	17,65	75					0	0				
σσσ		1										
%		25					0	0				
σσσσ	0											
%	0											
σσσσσ												
%												
(-)#V_	V(C)(C)#	2	1	1								
	%	40	100	100								
	V(C)(C)σ#	3	3	1								
	%	75	100	100								
	V(C)(C)σσ#											
%												
V(C)(C)σσσ#												
%												
V(C)(C)σσσσ#												
%												

Tableau 29b : Position et nombre de syllabes -#V_ dans CUL ET CHEMISE

TOM ET ZOÉ		M2											
M1	σ	σσ	σσσ	σσσσ	σσσσσ	σσσσσσ	(-)#V						
							V(C)(C)#	V(C)(C)σ#	V(C)(C)σσ#	V(C)(C)σσσ#	V(C)(C)σσσσ#		
σ	19	92	5										
%	59,83	15,63									16	12	
σσ	89	658	251								45,71	52,17	
%	49,17	53,29									72	83	
σσσ		220	59								40,91	41,29	
%	84,38	46,71									13	13	
σσσσ											59,09	28,26	
%													
σσσσσ													
%													
σσσσσσ													
%													
(-)#V_	V(C)(C)#												
	%												
	V(C)(C)σ#	19	104	9									
	%	54,29	59,09	40,91								19	21
	V(C)(C)σσ#	11	118	33								47,5	23
%	47,83	58,71	71,74										
V(C)(C)σσσ#													
%													
V(C)(C)σσσσ#													
%													

Tableau 29d : Position et nombre de syllabes -#V_ dans TOM ET ZOÉ

VITE ET BIEN		M2											
MI		σ	σσ	σσσ	σσσσ	σσσσσ	σσσσσσ	(-)#V					
								V(C)(C)#	V(C)(C)σ#	V(C)(C)σσ#	V(C)(C)σσσ#	V(C)(C)σσσσ#	
σ		27	38	6	1			1	3	1			
%			76	85,71	50			50	50	50			
σσ		12	39	19	2	1			2	2			
%		24	78	52,78	25	50		0	25	50	0		
σσσ		1	17	17	3				1	0	0		
%		14,29	47,22	47,22	20				100	0	0		
σσσσ		1	6	12	1	1			1				
%		50	75	80	100	100			100		0		
σσσσσ			1		1							0	
%			50		100							0	
σσσσσσ													
%													
(-)#V_	V(C)(C)#	1	2							1			
	%	50	100						0	100			
	V(C)(C)σ#	3	6					2	4	2			
	%	50	75	0	0			100	50	50	0		
	V(C)(C)σσ#	1	2	3				0	2	4		1	
%	50	50	100					50	100	100	50		
V(C)(C)σσσ#		1	6	1				1	1	3			
%		100	100	100				100	100	100			
V(C)(C)σσσσ#					1				1				
%					100				50				

Tableau 29e : Position et nombre de syllabes -#V_ dans VITE ET BIEN

ARTS ET MÉTIERS		M2										
MI		σ	σσ	σσσ	σσσσ	σσσσσ	σσσσσσ	(-)#V				
								V(C)(C)#	V(C)(C)σ#	V(C)(C)σσ#	V(C)(C)σσσ#	V(C)(C)σσσσ#
σ		6	25	3	5					1		
%			96,15	100	100			0		100		
σσ		1	14	9	2	1			1			
%		3,85	47,62	81,82	66,67	100		0	25	0	0	
σσσ		0	2	2	1						0	
%		0	18,18	22,22	33,33				0		0	
σσσσ		0	1	2	1							
%		0	33,33	66,67	100			0		0	0	
σσσσσ			0									
%			0									0
σσσσσσ												
%												
(-)#V_	V(C)(C)#	1	5		3				2	1		
	%	100	100		100				100	100		
	V(C)(C)σ#		3	5				0	2	1	2	
	%		75	100					0	33,33	100	
	V(C)(C)σσ#	1	1		1				0	2	2	
%	0	100		100				0	66,67	100		
V(C)(C)σσσ#		1	2	2	1				0			
%		100	100	100	100				0			
V(C)(C)σσσσ#												
%												

Tableau 29f : Position et nombre de syllabes -#V_ dans ARTS ET MÉTIERS

ÉLISE ET ZOË		M2										
MI		σ	σσ	σσσ	σσσσ	σσσσσ	σσσσσσ	(-)#V				
								V(C)(C)#	V(C)(C)σ#	V(C)(C)σσ#	V(C)(C)σσσ#	V(C)(C)σσσσ#
σ		3	1	1								
%		100	100	100								
σσ		0	6	5								
%		0	71,43	71,43						0	0	
σσσ		0	2	2								
%		0	28,57	28,57						0	0	
σσσσ												
%												
σσσσσ												
%												
σσσσσσ												
%												
(-)#V_	V(C)(C)#											
	%											
	V(C)(C)σ#		2	1						1	2	
	%		100	100						0	100	
	V(C)(C)σσ#		1	1							1	
%		100	100							0		
V(C)(C)σσσ#												
%												
V(C)(C)σσσσ#												
%												

Tableau 29g : Position et nombre de syllabes -#V_ dans ÉLISE ET ZOË

PAUL ET PASCAL		M2										
MI		σ	σσ	σσσ	σσσσ	σσσσσ	σσσσσσ	(-)#V				
								V(C)(C)#	V(C)(C)σ#	V(C)(C)σσ#	V(C)(C)σσσ#	V(C)(C)σσσσ#
σ			5	2								
%			100	100								
σσ		0	1	1								
%		0	100	100								
σσσ		0	0									
%		0	0									
σσσσ												
%												
σσσσσ												
%												
σσσσσσ												
%												
(-)#V_	V(C)(C)#											
	%											
	V(C)(C)σ#										1	
	%										100	
	V(C)(C)σσ#											
%										0		
V(C)(C)σσσ#												
%												
V(C)(C)σσσσ#												
%												

Tableau 29h : Position et nombre de syllabes -#V_ dans PAUL ET PASCAL

Examinons dans le détail les résultats émergeant dans ces tableaux support d'études par support d'études.

Les corpus TARN ET GARONNE, CUL ET CHEMISE, LISSE ET SOYEUX, et VITE ET BIEN, ainsi que les questionnaires d'enquête ÉLISE ET ZOÉ et PAUL ET PASCAL ne permettent pas l'examen d'hypothèse sur un échantillon supérieur à dix occurrences. Le test d'ajustement ne peut être probant, nous avons donc choisi de ne pas les relever. Notons que ce phénomène est imputable à la division des échantillons par le nombre d'hypothèses testées. Plus les hypothèses testées sont nombreuses, plus il est difficile d'observer des résultats.

Enfin, le corpus TOM ET ZOÉ présente trois résultats pertinents. Nous pouvons affirmer premièrement que :

- Les dissyllabes du type $V(C)(C)\sigma\sigma\#$ favorisent M1 dans 71,74% des binômes lorsqu'ils sont confrontés à des trisyllabes commençant par une attaque $\sigma\sigma\sigma$ ($\chi^2(1) = 8,7$; $p \leq 0,003$) ;
- Les mots $V(C)(C)\sigma\#$ sont prioritairement placés en M1 dans 59,09% des cas lorsqu'ils sont confrontés à des dissyllabiques commençant par une attaque $\sigma\sigma$ ($\chi^2(1) = 5,82$; $p \leq 0,02$)
- Enfin, les mots de deux syllabes du type $V(C)(C)\sigma\sigma\#$ sélectionnent prioritairement M1 (58,71%) quand ils sont en présence de termes dissyllabiques commençant par une attaque ($\sigma\sigma$) ($\chi^2(1) = 6,09$; $p \leq 0,01$).

En dehors de toutes considérations statistiques, nous constatons que d'une manière générale, les résultats par support d'études corroborent les résultats que nous avons précédemment constatés, même en ce qui concerne le corpus TOM ET ZOÉ qui avait pourtant présenté jusqu'ici des résultats systématiquement opposés à ce que nous relevions dans l'ensemble des supports d'études.

De plus, il apparaît que nos supports d'études, pris indépendamment les uns des autres, n'ont globalement pas assez d'effectifs par hypothèses testées pour nous fournir des résultats fiables, raison pour laquelle il nous est nécessaire d'examiner les résultats pour l'ensemble de nos données. Cependant, compte tenu de la nature différente du corpus TOM ET ZOÉ induisant de l'hétérogénéité, et malgré l'exceptionnelle convergence des résultats que nous constatons ici, nous avons néanmoins choisi de l'exclure.

Nous observerons donc la répartition relative du nombre de syllabes des mots pour un nouvel ensemble que nous nommons « TOTAL-TOM ET ZOÉ ».

TOTAL -TOM ET ZOË		M2										
M1		σ	σσ	σσσ	σσσσ	σσσσσ	σσσσσσ	(-)#V				
								V(C)(C)#	V(C)(C)σ#	V(C)(C)σσ#	V(C)(C)σσσ#	V(C)(C)σσσσ#
σ		75	85	21	7			4	5	3		
%			79,44	72,41	100			40	35,71	75		
σσ		22	199,18	59	5	1		1	5	5	0	
%		20,56		67,82	41,67	50		10	21,74	41,67	0	
σσσ		8	28	29	6			0	1	0	0	
%		27,59	32,18		35,29			0	7,14	0	0	
σσσσ		0	7	11	3	1	1	0	1	0	0	
%		0	58,33	64,71		100	50	0	100	0	0	
σσσσσ			1	0							0	0
%			50	0							0	0
σσσσσσ				1								
%				50								
(-)#V_	V(C)(C)#	6	9	1	3				2	2		
	%	60	90	100	100				50	66,67		
	V(C)(C)σ#	9	18	13	0			2	8	6	2	
	%	64,29	78,26	92,86	0			50		60	66,67	
	V(C)(C)σσ#	1	7	5	1			1	4	10	0	1
%	25	58,33	100	100			33,33	40		0	50	
V(C)(C)σσσ#		2	8	3	1			1	1	3		
%		100	100	100	100			33,33	100			
V(C)(C)σσσσ#					1				1			
%					100				50			

Tableau 30 : Position et nombre de syllabes -#V_ dans TOTAL-TOM ET ZOË

Comme le montre le tableau ci-dessus, nous pouvons statistiquement affirmer trois résultats forts liés à ce que nous présentions auparavant.

Les monosyllabes du type V(C)(C)σ# opposés à des trisyllabes σσσ occupent M1 dans 92,86% des 14 appariements testés sur ce critère ($\chi^2(1) = 10,29$; $p < 0,001$).

Parallèlement, nous notons que V(C)(C)σ# mis en présence de termes de deux syllabes commençant par une attaque σσ favorisent également M1 dans 78,26% des appariements traités sur ce critère ($\chi^2(1) = 7,25$; $p < 0,006$).

Enfin, les mots ayant pour patron V(C)(C)# montrent une occupation importante de 90% de M1 lorsqu'ils sont opposés à des mots de deux syllabes σσ ($\chi^2(1) = 6,4$; $p < 0,01$).

Concernant les autres résultats, nous avons donc choisi de ne pas relever les résultats observés sur moins de dix occurrences par hypothèse car non seulement ces effectifs très faibles ne permettent pas d'obtenir des résultats catégoriques, mais étant donnée que la valeur de p est supérieure à 0,1, nous ne pouvons rejeter l'hypothèse nulle et donc ne pouvons ni infirmer ni confirmer les hypothèses formulées.

Revenons néanmoins sur trois résultats, qui, bien que ne présentant pas de validation statistique aussi forte que nous le souhaitions ($p > 0,1$), mettent au jour des résultats en adéquation avec ce que nous avons observés jusqu'à présent.

Ainsi les monosyllabes $V(C)(C).\sigma\#$ mis en présence de termes monosyllabiques commençant par une attaque (σ) sélectionnent M1 dans 64,29 %. Les mots ayant pour patron $V(C)(C)\#$ confrontés aux mots monosyllabiques (σ) favorisent M1 dans 60% des cas. Et enfin, les mots dissyllabiques du type $V(C)(C).\sigma\sigma\#$ opposés à des termes dissyllabiques commençant par une attaque ($\sigma\sigma$) sélectionnent M1 dans 58,33% des occurrences testées sur cette hypothèse.

4.1.2.4. Bilan et analyse des résultats concernant le nombre de syllabes -#V_

Au regard de l'ensemble des résultats que nous avons proposé sur la question d'une corrélation entre la taille des mots pour laquelle nous avons pris en compte le décompte extramétrique des syllabes initiales sans attaque et la sélection d'une place dans la coordination, il apparaît que :

- En prenant en compte l'extramétricité des initiales vocaliques, dans 74,09% des paires non équisyllabiques testées, les mots en M1 sont plus courts que les mots en M2 ;
- En ce qui concerne les mots contenant de une à trois syllabes selon le décompte métrique excluant les initiales dépourvues d'attaques, il apparaîtrait deux dispersions graduelles, déductibles l'une de l'autre : plus un mot est court, plus il occuperait M1, et plus un mot est long, plus il favorise M2. Les monosyllabes -#V_ favorisent M1, les dissyllabes -#V_ et les trisyllabes -#V_, M2 ;
- De plus, ce résultat est corroboré par la confrontation du nombre de syllabes en M1 par rapport à ceux de M2 : pour les mots ayant un patron $V(C)(C)\#$ associés à des dissyllabes $\sigma\sigma$, ainsi que pour les monosyllabes $V(C)(C).\sigma\#$ confrontés respectivement à des trisyllabes $\sigma\sigma\sigma$ et à des dissyllabiques $\sigma\sigma$, nous affirmons que plus un mot est court, plus il occupe M1.

Comment comprendre et interpréter ces résultats ?

Dans l'ensemble, les résultats relevés dans cette section corroborent l'hypothèse de l'extramétrie des voyelles initiales telle que Plénat (1994 / 2009) la définit. Les mots commençant par une syllabe dépourvue d'attaques se comportent comme les mots les plus courts et tombent très vraisemblablement sous les principes que nous décrivions page 122.

Nous pouvons néanmoins aborder ce phénomène en examinant une nouvelle hypothèse : comme le montrent les résultats obtenus pour les monosyllabes $V(C)(C)\sigma\#$ mis en présence des mots dissyllabiques $\sigma\sigma$, il est possible de supposer que la qualité du segment initial peut être un critère déterminant dans la sélection d'une place dans la coordination.

Nous développerons cette hypothèse concernant les segments initiaux dans la section à suivre.

4.2. Début de mots

4.2.1. Nature du segment initial

Dans cette section, nous nous proposons d'examiner le critère définitoire soulevé par la question : quelle est la nature du segment initial de M1 et M2 ? Les initiales vocaliques favorisent-elles l'une des deux positions ? Et qu'en est-il des mots possédant des initiales consonantiques et de ceux débutant par une semi-consonne ?

Cette question fait écho à l'hypothèse émergente que nous avons relevée en examinant la corrélation entre la sélection d'une place dans la coordination et la taille des mots $\#V_$. Les résultats laissent à penser que la qualité du segment initial pourrait être un critère déterminant dans la sélection d'une place dans la coordination.

Afin d'explorer plus précisément ce nouveau postulat, tout en répondant au questionnement général qu'est la nature des segments initiaux de M1 et de M2, nous avons relevé les segments initiaux de M1 puis de M2 de sept supports d'études, à savoir, la collection des six corpus et le questionnaire d'enquête « ÉLISE et ZOÉ ». Les cinq autres questionnaires d'enquête ont été mis de côté car ils n'étaient pas interrogés sur l'hypothèse d'un segment initial actif dans la sélection d'un ordre préférentiel de coordination. Chaque paire était neutralisée sur ces questions ; leur segment initial était identique.

Étant donné que nos supports d'études sont autonomes, il nous a semblé pertinent, dans un premier temps, de présenter nos résultats, indépendamment, support d'études par support d'études.

TARN ET GARONNE 69	#V_	#C_	#G_
M1	11	57	1
%	73,33%	46,72%	###
M2	4	65	0
%	26,67%	53,28%	0,00%
Total :	15	122	1
	100,00%	100,00%	100,00%

Tableau 31a : Répartition des segments initiaux dans TARN ET GARONNE

CUL ET CHEMISE 42	#V_	#C_	#G_
M1	9	33	
%	69,23%	46,48%	
M2	4	38	
%	30,77%	53,53%	
Total :	13	71	0
	100,00%	100,00%	100,00%

Tableau 31b : Répartition des segments initiaux dans CUL ET CHEMISE

LISSE ET SOYEUX 62	#V_	#C_	#G_
M1	14	48	0
%	53,85%	49,48%	0,00%
M2	12	49	1
%	46,15%	50,52%	###
Total :	26	97	1
	100,00%	100,00%	100,00%

Tableau 31c : Répartition des segments initiaux dans LISSE ET SOYEUX

TOM ET ZOÉ 1420	#V_	#C_	#G_
M1	386	1030	4
%	56,76%	47,82%	66,67%
M2	294	1124	2
%	43,24%	52,18%	33,33%
Total :	680	2154	6
	100,00%	100,00%	100,00%

Tableau 31d : Répartition des segments initiaux dans TOM ET ZOÉ

VITE ET BIEN 202	#V_	#C_	#G_
M1	31	153	0
%	51,67%	50,00%	0,00%
M2	29	153	1
%	48,33%	50,00%	100,00%
Total :	60	306	1
	100,00%	100,00%	100,00%

Tableau 31e : Répartition des segments initiaux dans VITE ET BIEN

ARTS ET MÉTIERS 85	#V_	#C_	#G_
M1	31	54	
%	70,45%	42,86%	
M2	13	72	
%	29,55%	57,14%	
Total :	44	126	0
	100,00%	100,00%	100,00%

Tableau 31f : Répartition des segments initiaux dans ARTS ET MÉTIERS

TOTAL 1900	#V_									
	NASAL		HAUT		BAS		ARRIÈRE		ARRONDI	
	+	-	+	-	+	-	+	-	+	-
M1	66	431	36	461	320	177	303	194	65	432
	65,35%	56,49%	37,50%	60,03%	58,50%	55,84%	60,00%	54,04%	56,03%	57,75%
M2	35	332	60	307	227	140	202	165	51	316
	34,65%	43,51%	62,50%	39,97%	41,50%	44,16%	40,00%	45,96%	43,97%	42,25%
Total :	101	763	96	768	547	317	505	359	116	748
	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

Tableau 31g : Répartition des segments initiaux dans ÉLISE ET ZOÉ

Comme précédemment, nous avons noté en caractères gras les résultats pour lesquels le χ^2 est de $p \leq 0,05$ et en italiques ceux pour lesquels, il est compris entre $0,05 < p \leq 0,1$.

Les tableaux ci-dessus présentent les résultats de chaque support d'études traitant l'hypothèse de la répartition dans une des deux positions dans la coordonnée en fonction de la qualité du segment initial.

Dans le détail, les corpus TOM ET ZOÉ et ARTS ET MÉTIERS révèlent au moins une dispersion statistiquement valide ($p \leq 0,05$). TARN ET GARONNE, quant à lui, nous permet de dégager une tendance statistique ($0,05 < p \leq 0,1$).

Observons à présent les résultats relevés par support d'études.

Le corpus TOM ET ZOÉ, du fait du grand nombre d'effectifs le composant, met au jour deux répartitions significatives : les mots commençant par une initiale vocalique occupent majoritairement M1 dans 56,76% des mots à initiale vocalique ($\chi^2 (1) = 12,45 ; p \leq 0,0004$).

Inversement, les mots débutant par initiale consonantique sélectionnent prioritairement M2 dans 52,18% des occurrences testés sur cette hypothèse ($\chi^2 (1) = 4,1 ; p \leq 0,04$).

La dispersion des mots commençant par un glide du fait du très faible effectif testé (6 mots) n'est pas statistiquement valide ($p > 0,1$). Il nous est donc difficile de la commenter.

Dans le corpus ARTS ET MÉTIERS, on observe nettement une répartition significative conséquente : les mots à initiale vocalique (70,45% des mots #V_ testés dans ce corpus) sélectionnent très majoritairement la première position dans la coordonnée ($\chi^2 (1) = 7,36 ; p \leq 0,006$).

De plus, nous observons que les résultats des mots ayant une initiale consonantique, bien que ne montrant pas de validation statistique aussi forte que nous le souhaiterions ($\chi^2 (1) = 2,57 ; p \leq 0,11$), respectent néanmoins les contours de la prédisposition que nous relevions dans le précédent examen de TOM ET ZOÉ. En fait, seulement 42,86% des occurrences commençant par une consonne se placent prioritairement en M1. La valeur asymptotique de ce résultat est très proche de la valeur seuil de .1, De fait, au regard de ces deux motifs, nous pensons qu'il est pertinent et mérite notre attention. Nous nous attacherons, bien entendu, à le vérifier en le confrontant avec les résultats relevés dans les autres supports d'étude.

Aucun des résultats présents dans corpus CUL ET CHEMISE n'est validé par un test de Khi Deux. Ils corroborent néanmoins les résultats relevés dans les autres supports d'étude. En

effet, les mots à initiale vocalique sélectionnent majoritairement M1 dans 69,23 % des mots débutant par une voyelle ($\chi^2 (1) = 1,92 ; p \leq 0,16$). La répartition des mots commençant par une consonne respecte également ce que nous notions auparavant, à savoir que les mots à initiale consonantique évitent M1 : ils ne sont que 46,48% à occuper cette position ($\chi^2 (1) = 0,35 ; p \leq 0,55$).

Dans le TARN ET GARONNE, 73,33% des mots en #V_ ont tendance à M1, conformément à ceux que nous relevions auparavant ($\chi^2 (1) = 3,27 ; p \leq 0,07$).

Enfin, en ce qui concerne les supports d'études ÉLISE ET ZOÉ, LISSE ET SOYEUX et VITE ET BIEN, si aucun d'entre eux ne présente des résultats statistiquement valides ($p > 0,1$), nous relevons toutefois que tous sans exception respectent les contours de ce que nous observions précédemment. A savoir que les mots à initiale vocalique sélectionnent prioritairement M1 (respectivement dans 69,23%, 53,85% et 51,67%). Quant aux mots commençant par un glide, les résultats ne peuvent être pertinents compte tenu du trop faible effectif (Un seul mot pour LISSE ET SOYEUX et VITE ET BIEN).

A travers l'examen de ces sept supports d'études, nous voyons apparaître une prédisposition à ce que les mots débutant par une initiale vocalique occupent prioritairement M1. Inversement, il semblerait que les mots possédant une initiale consonantique occupent plutôt M2, bien que la dispersion statistique, compte tenu, sans doute, du faible effectif dégagé par hypothèses, ne nous permette pas toujours de dégager des résultats aussi probants que nous le souhaiterions.

Étant donné que le nombre de données testées ne nous permet pas systématiquement d'asseoir nos résultats par une significativité statistique, et que, parallèlement, l'ensemble de nos supports d'études reprennent tous et sans exception aucune la même répartition par segments initiaux pour M1 et M2, il nous a semblé pertinent d'étudier ces résultats dans leur ensemble pour le total des supports d'études examinés.

Nous avons donc observé les résultats fournis par l'ensemble des binômes testant l'hypothèse d'un lien entre une position privilégiée dans la coordination et la qualité du segment initial. Comme précédemment, nous avons noté en caractères gras les résultats pour lesquels le χ^2 est inférieur à .05.

Ainsi, dans le tableau suivant, sur 1 900 paires testées soit 3 763 termes, 851 mots commencent par une initiale vocalique représentant 22,6 % du total des mots testés, 2 903 termes débutent par une consonne soit une majorité de 77,1 % des termes testés, pour seulement neuf mots possédant un glide comme segment initial (0,23%) ($\chi^2(2)$; $p \leq 0,00000001$). Nous trouvons ici confirmation du fait que les mots du français commencent plus par une consonne que par une voyelle ou par un glide.

TOTAL 1900	#V_	#C_	#G_
M1	491	1386	5
%	57,70%	47,74%	55,56%
M2	360	1517	4
%	42,30%	52,26%	44,44%
Total :	851	2903	9
	100,00%	100,00%	100,00%

Tableau 32 : Répartition des segments initiaux pour TOTAL

Dans le détail de la répartition par segment entre M1 et M2, au regard des éléments présentés dans le tableau ci-dessus, nous pouvons affirmer l'existence de fortes présomptions contre l'hypothèse nulle validant statistiquement deux hypothèses corrélant la qualité des segments initiaux à une place dans la coordination. Il apparaît que les mots commençant par une initiale vocalique sélectionnent prioritairement M1 dans 57,70% des mots commençant par une voyelle ($\chi^2(1) = 20,17$; $p \leq 0,00001$).

Inversement les mots commençant par une consonne occupent majoritairement M2 dans 52,26% des mots débutant par une initiale consonantique ($\chi^2(1) = 5,91$; $p \leq 0,02$).

Nous ne pouvons commenter les résultats relevés pour les mots dont l'initiale est un glide, l'effectif de dix mots au total est trop faible pour que la répartition présentée ci-dessus soit valide.

Comme le montre le graphique ci-dessous, nous constatons que les mots commençant par une voyelle privilégient la première position dans la coordination, et que ceux débutant par une consonne sélectionnent à l'inverse M2.

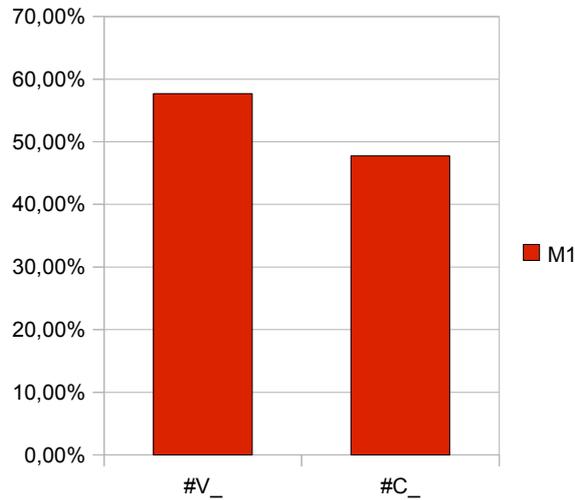


Diagramme 2 : Pourcentage d'occupation de M1 par segments initiaux

Au regard des résultats que nous venons de présenter, il apparaît qu'il existe une corrélation forte entre la qualité des segments initiaux et la sélection d'une position dans la coordination. Nous avons souhaité approfondir ce résultat en observant plus en détail les initiales vocaliques, les initiales consonantiques, puis, dans la mesure du possible, les mots débutant par un glide.

4.2.1.1. Les initiales vocaliques

4.2.1.1.1. Répartition phonémiques des initiales vocaliques

Dans le présent paragraphe, nous nous proposons d'étudier l'influence du timbre de la voyelle initiale des mots dans la sélection de la position M1 ou M2 des séquences binomiales coordonnées.

Pour ce faire, nous utiliserons deux ensembles de supports d'études.

Pour l'examen des voyelles initiales orales (ou #V_{orale}_⁶¹), nous compilerons les six corpus et le questionnaire d'enquête ÉLISE ET ZOÉ. Les autres questionnaires d'enquête ont été mis de côté, ne traitant pas la question du type d'initiale, cette variable a été neutralisée en amont pour chaque appariement.

Pour l'examen des voyelles initiales nasales (ou #V_{nasale}_), nous ne recueillerons que la collection des six corpus.

⁶¹ La forme #V_{orale}_ désigne les voyelles initiales orales par opposition à #V_{nasale}_ qui indique les voyelles initiales nasales. Nous les nommerons également #V[-NASALE]_ et #V[+NASALE]_.

Afin d'explorer plus précisément la répartition des initiales vocaliques dans les séquences binomiales coordonnées, nous nous attacherons, après un examen général de la répartition des voyelles initiales, à observer le type phonémique des initiales vocaliques orales de M1 et de M2, puis nous examinerons celui des initiales vocaliques nasales.

Dans un premier temps, nous nous proposons d'approfondir les résultats présentés *supra* en proposant deux tableaux observant la répartition des initiales vocaliques orales. Dans le premier tableau, nommé TOTAL-TOM ET ZOÉ, nous recueillons la dispersion des 148 mots possédant une initiale vocalique orale de l'ensemble des supports d'études sans ceux du corpus TOM ET ZOÉ. Dans le second, que nous nommons TOTAL, nous regroupons l'ensemble des 763 termes commençant par une voyelle orale. Il nous apparaît, en effet, que compte tenu de la taille et du mode de constitution du corpus TOM ET ZOÉ, il est préférable d'être en mesure de vérifier qu'il ne biaise pas nos résultats en mettant en avant d'autres facteurs.

Ainsi, comme le montrent les tableaux suivants, les deux résultats sont sensiblement identiques. Nous notons clairement que pour les deux totaux mettent au jour la même préférence à ce que les mots commençant par une voyelle orale occupent prioritairement M1, dans 60,14% des cas dans TOTAL-TOM ET ZOÉ ($\chi^2 (1) = 6,08 ; p \leq 0,001$), et 56,49% des mots #Vocale_ pour TOTAL ($\chi^2 (1) = 12,85 ; p \leq 0,0003$).

TOTAL-TOM ET ZOÉ 480	#Vocale_
M1	89 60,00%
M2	59 40,00%
Total :	148 100,00%

Tableau 33 : Répartition des voyelles orales initiales pour l'ensemble des supports d'études sans TOM ET ZOÉ

TOTAL 1900	#Vocale_
M1	431 56,47%
M2	322 43,53%
Total :	753 100,00%

Tableau 34 : Répartition des voyelles orales initiales pour l'ensemble des supports d'études

La confrontation de ces deux résultats homogènes indique également qu'en ce qui concerne les hypothèses liant les initiales vocaliques orales à la sélection d'une position dans la coordination, le corpus TOM ET ZOÉ présente des résultats semblables à ceux présents dans l'ensemble des supports d'études testés à ce sujet, ce qui laisse à penser que ces facteurs sont plutôt prégnants par rapport à ceux que nous relevions au sujet de la taille des mots par exemple. Il nous faudra le confirmer.

Intéressons-nous à présent à la répartition phonémique de voyelles initiales en M1 et en M2. Comme nous l'avons préalablement indiqué, nos supports d'études sont autonomes. Ce recueil de données éclectiques est l'une des richesses de ce travail de thèse. Grâce à l'ensemble de données dont nous disposons, il nous est possible de relever un vaste panel de composés syndétiques. C'est pourquoi il nous semble pertinent de présenter nos résultats support d'études par support d'études.

Cependant, il apparaît que plus les hypothèses à tester sont nombreuses, plus les effectifs examinés sont dispersés. En conséquence de ceci, il nous est très difficile d'obtenir des résultats solides, statistiquement valides de l'observation des voyelles orales initiales de M1 et de M2 par supports d'études.

En effet, dans le détail, comme présentés en annexe (des tableaux A 2a à A 2g), par hypothèses, un maximum de cinq effectifs sont testés pour les corpus TARN ET GARONNE et CUL ET CHEMISE, ils sont moins de dix pour le questionnaire d'enquête ÉLISE ET ZOÉ et moins de 15 pour les corpus VITE ET BIEN, ARTS ET MÉTIERS, et LISSE ET SOYEUX.

De l'observation des voyelles orales initiales par support d'études, nous ne pouvons extraire que quatre résultats tous issus du corpus TOM ET ZOÉ.

Ainsi, dans le corpus TOM ET ZOÉ, sur les 156 mots testés sur l'initiale vocalique /e/, 96 sélectionnent M1, soit 61,54% des mots débutant par /e/ ($\chi^2(1) = 8,31$; $p \leq 0,003$), tout comme les 30 mots commençant par /o/ qui occupent majoritairement M1 dans 70% des mots examinant cette hypothèse ($\chi^2(1) = 4,8$; $p \leq 0,03$), et les termes ayant pour initiale la voyelle orale /a/ qui favorisent M1 dans 57,83 % des mots examinés débutant par /a/ ($\chi^2(1) = 7,67$; $p \leq 0,005$).

Inversement, il apparaît que les mots commençant par la voyelle orale /i/ sélectionnent massivement M2, dans 92 % des occurrences testées à ce sujet ($\chi^2(1) = 17,64$; $p \leq 0,00003$).

Dans le recueil des résultats de chaque support d'études étudié individuellement, les mots commençant par /e/, par /a/, et par /o/ occupent majoritairement M1, alors qu'inversement, ceux débutant par /i/ privilégient plutôt M2.

Il semble également que le petit nombre d'effectifs testés soit un frein à l'observation de résultats statistiquement significatifs.

Compte tenu de ce dernier point, nous avons souhaité asseoir plus fortement les précédents résultats en observant les initiales vocaliques orales pour l'ensemble des supports d'études, augmentant ainsi le nombre d'effectifs par hypothèses testées. Cependant, à plusieurs reprises, nous avons signalé que le corpus TOM ET ZOÉ est, du fait de sa constitution, le support d'études le moins contrôlé de nos supports d'études. Nous l'avons d'ailleurs supprimé de l'observation générale de certaines hypothèses, tant il apparaissait que d'autres facteurs plus prégnants que ce que nous relevions biaisaient nos résultats.

C'est pourquoi nous nous proposons d'observer à présent la confrontation des répartition phonémiques des initiales vocaliques orales relevées dans TOTAL – à savoir les 763 mots commençant par une voyelle orale recueillis dans l'ensemble des supports d'étude _, et dans TOTAL-TOM ET ZOÉ – c'est-à-dire les 148 mots relevés dans l'ensemble des supports d'études sans les données issus de TOM ET ZOÉ.

Ici encore, nous avons noté en caractères gras les résultats pour lesquels le χ^2 est de $p \leq 0,05$.

Ainsi, comme le montrent les tableaux suivants, nous pouvons observer la répartition de M1 et de M2 par type d'initiale vocalique orale suivante :

-/e/ TOTAL-TOM ET ZOÉ 160	#V									
	NASAL		HAUT		BAS		ARRIÈRE		ARRONDI	
	+	-	+	-	+	-	+	-	+	-
M1	22 61,11%	72 58,06%	19 61,29%	75 58,14%	68 57,63%	26 61,90%	57 58,76%	37 58,73%	29 63,04%	65 57,02%
M2	14 38,89%	52 41,94%	12 38,71%	54 41,86%	50 42,37%	16 38,10%	40 41,24%	26 41,27%	17 36,96%	49 42,98%
Total :	36 100,00%	124 100,00%	31 100,00%	129 100,00%	118 100,00%	42 100,00%	97 100,00%	63 100,00%	46 100,00%	114 100,00%

Tableau 35 : Répartition phonémique des initiales vocaliques orales en M1 et en M2 dans TOTAL-TOM ET ZOÉ

TOTAL 753	#Vocale										
	i	e	ɛ	a	ɑ	y	ø	œ	u	o	ɔ
M1	14 31,11%	113 62,78%	33 51,56%	206 57,22%	0	22 43,14%	0 0,00%	1 33,33%	0	28 70,00%	14 73,68%
M2	31 68,89%	67 37,22%	31 48,44%	154 42,78%	0	29 56,86%	1 100,00%	2 66,67%	0	12 30,00%	5 26,32%
Total :	45 100,00%	180 100,00%	64 100,00%	360 100,00%	0 100,00%	51 100,00%	1 100,00%	3 100,00%	0 100,00%	40 100,00%	19 100,00%

Tableau 36 : Répartition phonémique des initiales vocaliques orales en M1 et en M2 dans TOTAL

Les deux tableaux *supra* mettent au jour sept résultats significatifs de la répartition phonémique des mots possédant une initiale vocalique orale ; cinq proviennent de TOTAL et les deux derniers de TOTAL-TOM ET ZOÉ. Notons dès à présent qu'un seul de ces résultats privilégie la position M2. Nous reviendrons sur ce résultat un peu plus tard.

Dans les paragraphes à venir, nous nous attacherons à confronter les deux résultats relevés pour les cinq phonèmes pour lesquels nous observons des résultats, à savoir : /i/, /e/, /a/, /o/ et /ɔ/.

Dans le détail, les mots en #/i/_ sont 68,89% à occuper massivement M2 dans TOTAL pour seulement 40% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 6,42$; $p \leq 0,01$ et $\chi^2(1) = 0,8$; $p \leq 0,4$). Ces résultats sont disparates tant dans leur score que dans leur validation statistique. TOTAL met au jour une répartition significative, ce qui n'est pas le cas de TOTAL-TOM ET ZOÉ. Ces différences nous indiquent que le résultat fourni par TOTAL n'est en fait porté que par le corpus TOM ET ZOÉ. En effet, sur 45 items testés, 25 proviennent de TOM ET ZOÉ. De ce fait, nous ne conserverons pas ce résultat.

Inversement, les mots commençant par /e/ sélectionnent prioritairement M1 dans 62,78% des mots ayant pour initiale /e/ dans TOTAL. Et ils sont 70,83% à occuper la même position dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 11,76$; $p \leq 0,0006$ et $\chi^2(1) = 4,17$; $p \leq 0,04$). Les deux totaux font émerger deux répartitions significatives. Les résultats sont sensiblement identiques, il apparaît qu'il existe une corrélation entre la position M1 et l'initiale vocalique orale /e/.

Les mots en #/a/_ privilégient M1 dans 57,22% des cas dans TOTAL ($\chi^2(1) = 751$; $p \leq 0,006$), et ils sont 53,19% à occuper la même position dans TOTAL-TOM ET ZOÉ ($\chi^2(1) =$

0,19 ; $p \leq 0,66$). Ces résultats sont *a priori* semblables, seule leur validation statistique diffère. TOTAL présente un résultat significatif, pas TOTAL-TOM ET ZOÉ. Cette différence est selon nous la marque du fait que le résultat de TOTAL est en fait majoritairement porté par TOM ET ZOÉ. En effet, sur les 360 mots en #/a/_ TOTAL, 313 proviennent de TOTAL-TOM ET ZOÉ. Ne résultant que d'un seul corpus, nous ne pouvons conserver ce résultat.

Dans le tableau TOTAL, les mots en #/o/_ privilégient majoritairement M1 dans plus de 70% des occurrences testant cette hypothèse, tout comme dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2 (1) = 6,4$; $p \leq 0,01$ et $\chi^2 (1) = 1,6$; $p \leq 0,20$). Les résultats étant absolument identiques en termes de score, nous pourrions être tentée de les conserver. D'autant plus que le petit échantillon de TOTAL-TOM ET ZOÉ (dix mots) peut justifier de façon assez convaincante l'absence de validation statistique. Néanmoins, dans la mesure où ce résultat est à nouveau porté par un seul corpus, *a fortiori* TOM ET ZOÉ, nous ne pouvons le confirmer. Rappelons en effet, que le corpus TOM ET ZOÉ est un corpus très vaste, essentiel à la présente thèse dans le sens où il est à l'origine de la constitution de nos questionnaires d'enquête, mais faillible du fait de sa nature.

Enfin, les mots commençant par le voyelle initiale orale /ɔ/ sont 73,68% à sélectionner M1 dans TOTAL, et TOTAL-TOM ET ZOÉ ($\chi^2 (1) = 2,13$; $p \leq 0,04$). Ces résultats absolument identiques sont portés par 19 termes issus de l'ensemble des corpus examinés _ exception faite de TOM ET ZOÉ. De plus, ils présentent une répartition significative. C'est pourquoi nous affirmons que les mots débutant par /ɔ/ privilégient M1.

Au regard de l'examen des initiales vocaliques orales, nous affirmons que :

- Les mots en #Vocale sélectionnent M1 dans 60,14% des occurrences relevées ;
- Tout comme les mots en #/e/_ et #/ɔ/_ qui privilégient massivement cette même position.

Après avoir examiné la répartition phonémique des voyelles initiales orales, nous proposons le même examen pour les initiales vocaliques nasales. Rappelons que pour cet examen, nous recueillons des paires issues de la seule collection des six corpus.

Dans un premier temps, nous nous proposons donc d'explorer la répartition générale des initiales vocaliques nasales.

Pour cela, nous présentons ci-dessous pour les prochaines hypothèses deux tableaux observant la répartition des initiales vocaliques nasales. Le premier tableau regroupe les 36 mots de l'ensemble des supports d'études, sans ceux du corpus TOM ET ZOÉ possédant une initiale vocalique nasale. Ce total se nomme à nouveau TOTAL-TOM ET ZOÉ. Le second, nommé TOTAL, recueille l'ensemble des 101 termes commençant par une #Vnasale_ présents dans les six corpus.

Ainsi, comme le montrent les tableaux suivants, les deux résultats sont sensiblement identiques. Nous notons clairement que pour les deux résultats, il existe une préférence à ce que les mots commençant par une voyelle nasale occupent prioritairement M1 dans 61,11% des cas dans le TOTAL-TOM ET ZOÉ ($\chi^2(1) = 1,78 ; p \leq 0,18$), et 65,35 % des mots #Vnasale_ pour le TOTAL ($\chi^2(1) = 9,51 ; p \leq 0,002$).

La seule différence notable provient de la validation statistique de ces mêmes résultats.

En effet, les 36 mots possédant une initiale vocalique nasale ne permettent pas d'établir un résultat statistique aussi fort que le total des 101 mots testés sur cette même hypothèse : nous ne pouvons exclure l'hypothèse nulle dans TOTAL-TOM ET ZOÉ ($\chi^2(1) = 1,78 ; p \leq 0,18$), alors que TOTAL met au jour une répartition significative.

TOTAL -TOM ET ZOÉ 490	#C_
M1 %	356 47,37%
M2 %	393 52,63%
Total :	749 100,00%

Tableau 37 : Répartition des voyelles nasales initiales pour l'ensemble des supports d'études sans TOM ET ZOÉ

TOTAL 1900	#C_
M1 %	1386 47,74%
M2 %	1517 52,26%
Total :	2903 100,00%

Tableau 38 : Répartition des voyelles nasales initiales pour l'ensemble des supports d'études

Bien que ces deux résultats soient homogènes, il apparaît que le résultat présenté dans TOTAL n'est en fait porté que par TOTAL-TOM ET ZOÉ. Nous ne pouvons pas le conserver.

Nous ne pouvons ni confirmer ni infirmer l'existence d'une relation liant les voyelles initiales nasales à une position préférentielle dans les composés syndétiques par « et ».

L'examen phonémique des voyelles initiales nasales pourrait apporter un nouvel éclairage à cette hypothèse.

C'est pourquoi nous nous proposons à présent d'examiner dans le détail la répartition phonémique en M1 et en M2 des initiales vocaliques nasales.

Comme nous l'avons proposé précédemment, nous présentons nos résultats, indépendamment, pour chaque support d'études impacté par l'hypothèse d'une corrélation entre une voyelle initiale nasale et la sélection d'une position dans la coordonnée. Nous les fournissons en l'état en annexe (des tableaux A 3a à A 3g).

Cependant, en raison de la dispersion du nombre de termes testés par support d'études et par hypothèse, il est difficile de relever des résultats.

En effet, dans le détail, les corpus TARN ET GARONNE, CUL ET CHEMISE, LISSE ET SOYEUX et ARTS ET MÉTIERS ne proposent qu'un maximum de sept mots (respectivement trois, quatre, un et sept mots) commençant par une voyelle initiale nasale. Seuls les corpus TOM ET ZOÉ et VITE ET BIEN avec un nombre de 65 et 21 mots à initiale vocalique nasale sont susceptibles de fournir un résultat statistiquement valide.

En fait, de l'observation des voyelles initiales nasales par support d'études, nous ne pouvons extraire que deux résultats, qui sont détaillés ci-dessous.

Dans le corpus TOM ET ZOÉ, nous relevons un résultat statistiquement significatif : les mots commençant par la voyelle nasale /ã/ sélectionnent majoritairement M1, dans 67,69 % des occurrences testées à ce sujet ($\chi^2(1) = 8,14$; $p \leq 0,004$).

Le corpus VITE ET BIEN, quant à lui, fait émerger une répartition significative : les termes ayant pour initiale la voyelle nasale /ẽ/ privilégient très massivement M1 dans 81,82 % des mots examinés débutant par /ẽ/ ($\chi^2(1) = 4,45$; $p \leq 0,03$).

Il apparaît donc que dans le recueil des résultats de chaque support d'études étudié individuellement, les mots commençant par /ã/ et par /ẽ/ occuperaient majoritairement M1.

Cependant nous avançons ces résultats avec prudence.

En effet, non seulement le nombre restreint de noms testés est un frein à l'observation de résultats solides, mais en plus, nous nous devons de prendre en compte le fait que ces résultats

ne sont portés, dans les deux cas, que par un seul support d'études, à savoir les corpus TOM ET ZOÉ et VITE ET BIEN. C'est pourquoi il convient de les valider ultérieurement.

Ainsi, afin de consolider les précédents résultats, nous observerons l'ensemble des supports d'études augmentant ainsi le nombre d'effectifs par hypothèses testées dans le tableau nommé TOTAL. Nous nous proposons également d'observer deux nouveaux totaux : le premier TOTAL-TOM ET ZOÉ exclut comme précédemment les données fournies par TOM ET ZOÉ, le second TOTAL-VITE ET BIEN, celles issues du corpus VITE ET BIEN.

Nous allons donc présenter trois tableaux observant la répartition des initiales vocaliques nasales. Dans le premier, représentant les 36 mots possédant une initiale vocalique nasale de l'ensemble des supports d'études sans ceux du corpus TOM ET ZOÉ, nous prêterons une attention particulière à la répartition de la voyelle /ã/. Dans le deuxième donnant la répartition des voyelles initiales nasales en M1 et M2 des 80 mots composant l'ensemble des données de nos supports d'études sans celles de VITE ET BIEN, nous observerons spécifiquement les mots commençant par /ẽ/. Enfin le troisième tableau nous fournira l'ensemble de la répartition des initiales vocaliques nasales pour tous les supports d'études examinés pour cette hypothèse.

TOTAL M	ẽ																
	p	t	k	b	d	g	f	s	ʃ	v	z	ʒ	l	m	n	ɲ	ŋ
M	41	2	38	20	19	4	17	6	10	27	0	6	20	9	7		23
	3,9%	47%	4,2%	3,6%	3,0%	3,3%	4,9%	2,0%	3,0%	4,5%	0,0%	3,3%	3,3%	3,3%	3,4%		4,4%
M2	38	3	40	26	30	1	22	5	10	31	2	14	19	4	14		30
	4,0%	32%	3,2%	4,0%	4,2%	0,6%	6,6%	5,4%	4,1%	3,4%	0,0%	4,6%	4,2%	4,4%	4,1%		5,6%
tot	79	4	78	46	49	5	39	11	20	58	2	20	39	11	21	0	53
	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%

Tableau 39 : Répartition des voyelles nasales initiales pour l'ensemble des supports d'études sans TOM ET ZOÉ

TOTAL M1	ẽ																
	p	t	k	b	d	g	f	s	ʃ	v	z	ʒ	l	m	n	ɲ	ŋ
M1	109	63	67	58	58	29	26	60	3	99	12	63	104	29	58		5
	5,6%	32%	4,4%	4,2%	4,4%	4,3%	3,0%	3,3%	0,0%	3,3%	4,0%	4,3%	4,3%	3,0%	3,0%		4,5%
M2	67	75	65	73	56	35	40	9	3	95	13	61	208	29	78		7
	4,3%	4,7%	3,9%	3,7%	3,5%	3,6%	4,2%	4,3%	0,0%	4,0%	3,0%	3,6%	3,7%	3,0%	4,4%		5,4%
tot	176	138	132	131	114	64	66	69	4	194	25	124	412	58	136	0	12
	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%

Tableau 40 : Répartition des voyelles nasales initiales pour l'ensemble des supports d'études sans VITE ET BIEN

TOTAL 241	HC																	
	SONANT		CONTINU		NASAL		HAUT		ARRIERE		ANTERIEUR		CORONAL		VOISE		REL.RET	
	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-
M1	559	846	921	454	308	1097	342	1063	253	1152	1016	399	643	762	878	527	613	792
	48,52%	47,29%	47,33%	40,64%	50,00%	47,18%	45,84%	48,49%	46,00%	48,18%	48,81%	45,34%	47,42%	48,08%	46,11%	50,82%	46,09%	49,16%
M2	593	945	1025	511	308	1228	404	1132	297	1239	1055	401	713	823	1026	510	717	819
	51,48%	52,71%	52,67%	51,36%	50,00%	52,42%	54,16%	51,57%	54,00%	51,82%	51,19%	54,66%	52,58%	51,92%	53,49%	49,18%	53,91%	50,84%
Total :	1152	1791	1946	965	616	2325	746	2195	550	2391	2071	800	1356	1585	1904	1037	1330	1611
	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

Tableau 41 : Répartition des voyelles nasales initiales pour l'ensemble des supports d'études

Ainsi, comme le montrent les tableaux ci-dessus, concernant les mots ayant pour initiale la voyelle nasale /ã/, les résultats sont assez proches, nous notons une même inclination plaçant les mots commençant par la voyelle nasale /ã/ plutôt en M1, dans 57,89% des cas dans TOTAL-TOM ET ZOÉ ($\chi^2(1) = 1,78$; $p \leq 0,2$), et 65,48% pour le TOTAL ($\chi^2(1) = 8,05$; $p \leq 0,005$). Les résultats semblent identiques, mais ils ne jouissent pas de la même validation statistique. TOTAL annonce une répartition significative, ce qui n'est pas le cas de TOTAL-TOM ET ZOÉ. Cette différence nous indique que ce résultat est seul porté par TOM ET ZOÉ. C'est pourquoi nous ne le conservons pas.

Confrontant à présent les résultats relevés dans TOTAL et dans TOTAL-VITE ET BIEN. Concernant les mots ayant pour initiale /ẽ/, il apparaît que dans le TOTAL-VITE ET BIEN, les mots commençant par /ẽ/ ne parviennent à sélectionner une position privilégiée. Inversement, dans TOTAL, les mots commençant par #/ẽ/ ont tendance à sélectionner majoritairement M1 ($\chi^2(1) = 3,27$; $p \leq 0,07$). Si l'on considère le petit nombre d'effectifs testés (seulement 15 au total), et l'absence de résultats relevée TOTAL-VITE ET BIEN, nous ne pouvons confirmer la sélection d'une position dans la coordination par les mots en #/ẽ/.

Dans l'examen de la répartition par type des mots commençant par une initiale vocalique nasale en M1 et en M2, nous ne pouvons extraire de résultat.

4.2.1.1.2. Répartition par traits distinctifs des initiales vocaliques

La question que nous nous posons à présent est de savoir pourquoi seuls les mots en #/e/_ et #/ɔ/_ occupent plutôt M1. Pourquoi ne relevons-nous pas de lien entre une position préférentielle de la séquence binomiale coordonnée par « et » et les mots débutant par une voyelle nasale ? Que partagent /e/ et /ɔ/ qui les différencie tant des autres initiales vocaliques ?

Pour répondre à ces questionnement, nous nous proposons d'observer la répartition des initiales vocaliques orales et nasales par traits binaires distinctifs en commençant par présenter cet examen pour chaque support d'étude avant de commenter la confrontation de ces données pour l'ensemble de mots examinant le rôle des initiales vocaliques. Pour cela, nous utilisons la matrice des traits distinctifs de Dell (1975) comme présentée en 1.2.3. Organisation par traits page 29. Les résultats de la répartition des traits distinctifs des initiales vocaliques par support d'études sont accessibles en annexe (tableaux A 4a à A 4g).

Dans le détail, dans les six corpus et le questionnaire d'enquête ÉLISE ET ZOÉ, nous relevons douze résultats statistiquement significatifs notés en caractères gras et une tendance statistique mentionnée en italique. Notons dès à présent que quasiment tous les résultats que nous présentons - à une exception près - mettent en avant une corrélation entre un trait distinctif et la position M1. Ces treize résultats sont fournis par trois corpus : CUL ET CHEMISE, TOM ET ZOÉ, et ARTS ET MÉTIERS. Observons les résultats relevés pour ces trois corpus.

Le corpus CUL ET CHEMISE fait émerger un résultat dans lesquels les mots #[- NASAL]_ sélectionnent massivement, dans 83,33% des cas, M1 ($\chi^2(1) = 5,33 ; p \leq 0,02$).

Dans le corpus TOM ET ZOÉ, nous recueillons sept résultats dont deux apparemment contradictoires.

Concernant les mots possédant une initiale vocalique [\pm NASAL], nous relevons deux résultats. Les mots en #[+ NASAL]_ sont 67,69% à occuper M1 ($\chi^2(1) = 8,14 ; p \leq 0,004$), ceux #[- NASAL]_ sont 55,61% dans la même position ($\chi^2(1) = 7,74 ; p \leq 0,005$). Ce résultat semble signifier que le trait [\pm NASAL] n'est pas un facteur pertinent dans la sélection d'un ordre préférentiel de coordination. Nous reviendrons dessus en les confrontant avec ceux obtenus pour l'ensemble des supports d'études.

Les mots débutant par une voyelle [\pm HAUT] mettent tous deux au jour des résultats. Les

mots #[+ HAUT]_ sélectionnent majoritairement M2 dans 73,85% des occurrences testant cette hypothèse ($\chi^2(1) = 14,78$; $p \leq 0,0001$). Inversement, ceux #[- HAUT]_ occupent M1 dans 60% des cas ($\chi^2(1) = 24,6$; $p \leq 0,000001$).

Les mots débutant par une voyelle [+ BAS] sont 58,74% à préférer M1 ($\chi^2(1) = 13,11$; $p \leq 0,0003$), tout comme ceux commençant par une voyelle [+ ARRIÈRE] ou [- ARRONDI] qui sont respectivement 60,29% et 57,38% à sélectionner cette même position ($\chi^2(1) = 17,29$; $p \leq 0,00003$ et $\chi^2(1) = 13,28$; $p \leq 0,0003$).

Enfin, le corpus ARTS ET MÉTIERS révèle quatre résultats significatifs et une tendance.

Conformément à ce que nous observions dans les précédents corpus, les mots #[- NASAL]_ sont 66,67% à occuper M1, pour 73,81% des mots #[- HAUT]_, 73,08% des mots #[+ ARRIÈRE]_ et 67,57% des mots commençant par une voyelle [- ARRONDI] (respectivement $\chi^2(1) = 4,33$; $p \leq 0,04$, $\chi^2(1) = 9,52$; $p \leq 0,002$, $\chi^2(1) = 5,54$; $p \leq 0,02$ et $\chi^2(1) = 4,57$; $p \leq 0,03$).

Les mots commençant par une voyelle [+ BAS], quant à eux, ont tendance à occuper M1 : ils sont 66,67% à préférer cette position ($\chi^2(1) = 3,67$; $p \leq 0,06$).

Au regard des éléments collectés de la répartition en M1 et en M2 des mots commençant par une voyelle en fonction de leur trait distinctif, nous relevons sept résultats : les mots possédant une voyelle [\pm NASAL] et/ou [-HAUT] et/ou [+ BAS] et/ou [+ ARRIÈRE] et/ou [- ARRONDI] sélectionnent majoritairement M1. Inversement les mots #[+HAUT]- ont plutôt tendance à occuper M2.

Deux résultats sont apparemment contradictoires : les traits [\pm NASAL] occupent tous deux M1. De fait, soit les traits [\pm NASAL] ne jouent pas de rôle dans la sélection d'un ordre préférentiel de coordination, soit d'autres facteurs impactent les résultats que nous observons, et les résultats émergeant pour l'une des deux valeurs du trait [\pm NASAL] sont biaisés.

Afin de clarifier ce point, il nous faut les mettre en perspective.

C'est pourquoi dans le paragraphe à venir, nous nous attacherons à confronter les résultats relevés pour chaque support d'étude à ceux de deux totaux. Le premier total (TOTAL) regroupe l'ensemble des données interrogeant le rapport entre initiale vocalique et position dans la séquence binomiale coordonnée pour les six corpus et le questionnaire d'enquête ÉLISE ET ZOÉ. Le second total, pour les raisons que nous avons déjà évoquées, examine le même sans les données issues du corpus TOM ET ZOÉ. Nous le nommons TOTAL-TOM ET ZOÉ.

TOTAL 1900	#V_									
	NASAL		HAUT		BAS		ARRIÈRE		ARRONDI	
	+	-	+	-	+	-	+	-	+	-
M1	66 65,35%	431 56,49%	36 37,50%	461 60,03%	320 58,50%	177 55,84%	303 60,00%	194 54,04%	65 56,03%	432 57,75%
M2	35 34,65%	332 43,51%	60 62,50%	307 39,97%	227 41,50%	140 44,16%	202 40,00%	165 45,96%	51 43,97%	316 42,25%
Total :	101 100,00%	763 100,00%	96 100,00%	768 100,00%	547 100,00%	317 100,00%	505 100,00%	359 100,00%	116 100,00%	748 100,00%

Tableau 42 : Répartition par traits des mots en #V_ pour TOTAL

TOTAL -TOM ET ZOÉ 480	#V_									
	NASAL		HAUT		BAS		ARRIÈRE		ARRONDI	
	+	-	+	-	+	-	+	-	+	-
M1	22 61,11%	89 60,14%	19 61,29%	92 60,13%	68 57,63%	43 65,15%	57 58,76%	54 62,07%	29 63,04%	82 59,42%
M2	14 38,89%	59 39,86%	12 38,71%	61 39,87%	50 42,37%	23 34,85%	40 41,24%	33 37,93%	17 36,96%	56 40,58%
Total :	36 100,00%	148 100,00%	31 100,00%	153 100,00%	118 100,00%	66 100,00%	97 100,00%	87 100,00%	46 100,00%	138 100,00%

Tableau 43 : Répartition par traits des mots en #V_ pour TOTAL-TOM ET ZOÉ

Les deux tableaux ci-dessus font apparaître treize résultats significatifs et trois tendances.

Dans le détail, concernant les traits [\pm NASAL], nous relevons trois répartitions significatives. Ainsi, les mots #[- NASAL]_ sont 56,49 % dans TOTAL et 60,14% dans TOTAL-TOM ET ZOÉ à occuper majoritairement M1 (respectivement $\chi^2(1) = 12,85$; $p \leq 0,0003$ et $\chi^2(1) = 6,08$; $p \leq 0,01$). Les résultats sont homogènes dans les deux totaux, nous affirmons donc l'existence d'un lien entre M1 et les mots #[- NASAL]_.

Les mots #[+ NASAL]_ sont, quant à eux, 65,37% dans TOTAL et 61,11% dans TOTAL-TOM ET ZOÉ à favoriser la position M1 (respectivement $\chi^2(1) = 9,51$; $p \leq 0,002$ et $\chi^2(1) = 1,78$; $p \leq 0,18$). Ces deux résultats sont *a priori* très proches, exception faite de leur validation statistique ; TOTAL présente un résultat, pas TOTAL-TOM ET ZOÉ. Cette différence nous indique que ce résultat est uniquement porté par TOM ET ZOÉ. En effet, sur 101 mots #[+ NASAL]_, 65 proviennent du corpus TOM ET ZOÉ. Attendu que nous ne validerons jamais un résultat fourni par un seul corpus, nous ne conservons pas ce résultat.

La répartition des mots dont l'initiale vocalique porte les traits [\pm HAUT] met au jour trois résultats.

Les mots #[-HAUT]_ sont 60,03 % dans TOTAL et 60,13% dans TOTAL-TOM ET ZOÉ à favoriser préférentiellement M1 (respectivement $\chi^2(1) = 30,88$; $p \leq 0,000\ 000\ 03$ et $\chi^2(1) = 6,28$; $p \leq 0,01$). Les résultats sont identiques. Il existe une corrélation entre M1 et les mots #[-HAUT]_.

Quant aux mots en #[+ HAUT]_, TOTAL révèle qu'ils sont 62,5% à occuper M2 pour 38,71% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 6$; $p \leq 0,01$ et $\chi^2(1) = 1,58$; $p \leq 0,2$). Les résultats sont disparates, et en termes de score, et en termes de validation statistique. Nous ne les conserverons pas.

Nous observons également trois répartitions significatives et une tendance statistique pour les traits [\pm BAS] en initiale vocalique de mots.

Les mots #[+ BAS]_ sont 58, 5% à sélectionner M1 dans TOTAL, pour 57,63% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 15,81$; $p \leq 0,000\ 07$ et $\chi^2(1) = 2,75$; $p \leq 0,09$). Les résultats sont très proches en termes de score. Cependant, nous notons une différence de validation statistique. TOTAL met au jour un résultat, TOTAL-TOM ET ZOÉ, une tendance. Convaincue que des phénomènes de conflits de contraintes peuvent obscurcir nos résultats, nous pensons qu'il est nécessaire de prêter une attention particulière aux tendances statistiques. C'est pourquoi nous avançons avec précaution que les mots #[+ BAS]_ ont tendance à occuper M1.

Les mots #[- BAS]_, quant à eux, sont 55,84% dans TOTAL et 65,15% dans TOTAL-TOM ET ZOÉ à occuper majoritairement M1 (respectivement $\chi^2(1) = 4,32$; $p \leq 0,04$ et $\chi^2(1) = 6,06$; $p \leq 0,01$). Ces deux résultats mettent tous deux en évidence la propension des mots #[-BAS]_ à sélectionner prioritairement M1, nous affirmons donc ce résultat.

Concernant le trait [+ ARRIÈRE], les mots #[+ ARRIÈRE]_ sont 60% dans TOTAL à sélectionner M1, pour 58,76% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 20,2$; $p \leq 0,000\ 1$ et $\chi^2(1) = 2,75$; $p \leq 0,08$). Exception faite des résultats du test d'ajustement, ces résultats sont très proches, TOTAL met au jour une répartition significative, TOTAL-TOM ET ZOÉ, une tendance. Comme précédemment et pour les mêmes raisons, nous conservons ces résultats.

Les mots #[- ARRONDI]_ sont 57,75 % dans TOTAL à sélection M1, pour 59,42% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 17,99$; $p \leq 0,000\ 02$ et $\chi^2(1) = 4,9$; $p \leq$

0,03). Ces résultats sont identiques, nous affirmons donc qu'il existe une corrélation entre les mots #[- ARRONDI]_ et M1.

Enfin, TOTAL-TOM ET ZOÉ fait émerger deux nouveaux résultats en apparence contradiction avec les deux derniers résultats exposés. Tout comme les mots #[+ ARRIÈRE]_, les mots #[- ARRIÈRE]_ sont 62, 07% à occuper préférentiellement M1 ($\chi^2 (1) = 2,98$; $p \leq 0,07$).

Les mots #[+ ARRONDI]_, comme les mots #[- ARRONDI]_, ont tendance à sélectionner M1 dans 63,04 % des occurrences testant cette hypothèse ($\chi^2 (1) = 2,98$; $p \leq 0,07$).

Nous nous attacherons à commenter ces deux résultats dans la section suivante.

4.2.1.1.3. Bilan et analyse des résultats concernant #V_

Au regard de l'examen de la répartition des mots possédant une initiale vocalique en M1 et M2, nous affirmons que :

- Les mots commençant par une initiale vocalique sélectionnent majoritairement M1 pour 60,14% des mots testant cette hypothèse.
- La répartition phonémique montre que les mots débutant par les initiales vocaliques /e/ et /ɔ/ occupent prioritairement M1.
- Les mots ayant pour initiale une voyelle [- NASAL] (/i, e, ε, a, α, ɔ, o, u, y, ø, œ, ə/), et/ou [- HAUT] (/e, ø, ε, œ, ɔ, a, ē, ã, õ/), et/ou [- BAS] (/i, y, u, e, ø, o/), et/ou [- ARRONDI] (/i, e, ε, a, ã, ē/), et/ou [- ARRIÈRE] (/i, y, e, ø, ε, œ, ē/) sont majoritairement placés en M1 ;
- Tout comme les mots possédant une initiale vocalique [+ BAS] (/ε, œ, ɔ, a, ē, ã, õ/), et/ou [+ ARRIÈRE] (/u, o, ɔ, a, õ, ã/), et/ou [+ ARRONDI] (/y, u, o, ɔ, ø, œ, õ/) qui mettent en avant la même tendance ;
- Concernant les traits [± BAS] et/ou [± ARRIÈRE] et/ou [± ARRONDI] les résultats plaçant les deux valeurs des traits en M1 semblent a priori contradictoires ; notons cependant que :
 - Les mots #[- BAS]_ sélectionnent plus M1 que les mots #[+ BAS]_ : le premier résultat met au jour une répartition significative, le second, une tendance ;

- Pour les mêmes raisons, les mots #[- ARRIÈRE]_ occupent plus M1 que les mots #[+ ARRIÈRE]_ ;
- Et les mots #[- ARRONDI]_ favorisent plus en M1 que les mots #[+ ARRONDI]_

Comment comprendre et analyser les résultats compilés ci-dessus ?

La préférence des mots en #V_ à occuper M1 est vraisemblablement imputable à l'évitement du hiatus. Ce principe, noté en Théorie de l'Optimalité *VV, prédit que la séquence de deux voyelles consécutives est interdite.

Cependant, si *VV rend aisément compte de la présence des mots possédant une initiale vocalique en M1, elle n'explique pas pourquoi les mots commençant par /e/ et /ɔ/, et ceux dont l'initiale vocalique est soit [- NASAL], soit [-HAUT], soit [- BAS], soit [- ARRONDI] , soit [- ARRIÈRE], sélectionnent majoritairement M1.

Notons que les traits [- NASAL], [-HAUT], [- BAS], [- ARRONDI], [- ARRIÈRE] constituent le faisceau de traits distinctifs de la voyelle /e/. En fait, pour les présents résultats, plus que d'observer les traits, il nous faut avant tout expliquer l'occupation privilégiée des mots possédant une initiale vocalique /e/ ou /ɔ/.

Afin de motiver la présence préférentielle des mots #/e/_ en M1, nous pouvons invoquer le Principe du contour obligatoire dit OCP (Goldsmith (1976)). Ce principe est un principe de dissimilation qui prohibe la présence adjacente de deux éléments identiques. Nous reviendrons dessus page 217 pour l'observation des voyelles finales.

Il apparaît que les mots débutant par /e/ seraient en fait rejetés en M1 afin de ne pas produire la séquence /ee/ avec la conjonction de coordination « et ».

Ici encore, si le principe du contour obligatoire rend vraisemblablement compte de l'occupation préférentielle des mots #/e/_ en M1, il ne motive en rien celle des mots #/ɔ/_.

Afin d'expliquer ce résultat, nous avons observé en détail les 19 mots commençant par /ɔ/ dans TOTAL-TOM ET ZOÉ. Or, sur les 14 termes occupant M1, trois sont des réalisations de « HOMME ». Consciente du rôle actif du facteur sériel que nous avons exposé en 3.2. Le facteur sériel, nous avons donc choisi d'examiner la même répartition des mots #/ɔ/_ en

termes de type, et non de token⁶². Les résultats relevés dans cette nouvelle répartition montrent que nous pouvons imputer la présence en M1 des mots #/ɔ/_ aux multiples réalisations du mot « HOMME ». En effet, si en termes de score les résultats obtenus révèlent encore une préférence à ce que les mots #/ɔ/_ occupent M1 – ils sont 68,75% à sélectionner cette position-, le résultat n'est pas confirmé par le test de khi deux ($\chi^2(1) = 2,25$; $p \leq 0,13$). Il semblerait que ce résultat soit un faux résultat.

Enfin, compte tenu des résultats massifs dégagés pour les mots commençant par /e/, il nous a semblé intéressant d'observer la répartition par traits distinctifs en M1 et en M2 des mots commençant par une initiale vocalique qui n'est pas /e/ dans le but de faire émerger de nouveaux résultats.

-/e/ TOTAL-TOM ET ZOÉ 160	#V									
	NASAL		HAUT		BAS		ARRIÈRE		ARRONDI	
	+	-	+	-	+	-	+	-	+	-
M1	22 61,11%	72 58,06%	19 61,29%	75 58,14%	68 57,63%	26 61,90%	57 58,76%	37 58,73%	29 63,04%	65 57,02%
M2	14 38,89%	52 41,94%	12 38,71%	54 41,86%	50 42,37%	16 38,10%	40 41,24%	26 41,27%	17 36,96%	49 42,98%
Total :	36 100,00%	124 100,00%	31 100,00%	129 100,00%	118 100,00%	42 100,00%	97 100,00%	63 100,00%	46 100,00%	114 100,00%

Tableau 44 : Répartition par traits des mots en #V_(-/e/) pour TOTAL-TOM ET ZOÉ

Le tableau ci-dessus ne révèle que cinq tendances statistiques ($0,5 < p \leq 0,1$) notées en italique : les mots en [- NASAL], et/ou [- HAUT], et/ou [+ BAS], et/ou [+ ARRONDI], et/ou [+ ARRIÈRE] auraient tendance à sélectionner M1. Ces cinq traits constituent le faisceau de traits distinctifs de /ɔ/. Sachant que les résultats des mots #/ɔ/_ sont galvaudés par la surnumération du mot « HOMME », nous pensons que ces cinq traits tombent sous le même biais.

Cependant, ces résultats (ou cette absence de résultats) nous confortent dans l'idée que d'une part, les résultats relevés pour le faisceau de traits [- NASAL], [- HAUT], [- BAS], [- ARRONDI], [- ARRIÈRE] sont essentiellement dus à /e/, et d'autre part, si l'on peut supposer le rôle du principe *VV positionnant préférentiellement les mots commençant par une voyelle en M1, l'évitement de la séquence /e e/ est assurément plus actif.

⁶² Nous reviendrons abondamment sur ces deux méthodes de comptabilisations différentes page 174

Nous pensons donc que dans le cadre de la coordination par « et », de deux mots en français, l'action d'OCP est plus forte que celle de *VV.

4.2.1.2. Consonne

4.2.1.2.1. Répartition phonémique des consonnes initiales

Comme nous l'indiquions précédemment page 149 lors de l'examen global de la répartition des mots en M1 et en M2 en fonction de leur segment initial, s'il apparaît de façon assez solide et comme nous l'avons exposé dans le détail *supra* que les mots commençant par une voyelle sélectionnent prioritairement M1, inversement, les mots débutant par une consonne privilégient M2.

Dans cette section, nous nous proposons donc de relever les initiales consonantiques cette fois-ci, de M1, puis de M2, en nous appuyant sur la collection des six corpus et sur le questionnaire d'enquête ÉLISE ET ZOÉ. Les autres questionnaires d'enquêtes sont, ici encore, exclus de ce relevé, ne traitant pas la question du type d'initiale, cette variable était neutralisée en amont pour chaque appariement.

Afin d'explorer plus précisément la répartition des initiales consonantiques entre M1 et M2, nous nous attacherons dans un premier temps à étudier la répartition générale des initiales consonantiques pour chaque support d'études. Puis, nous examinerons en détail la répartition phonémique des consonnes initiales en M1 et en M2.

Les tableaux ambitionnant d'établir une corrélation entre une position dans la coordonnée et les initiales consonantiques sont proposés en annexe du tableau A 5a au tableau A 5g.

Le premier constat que nous faisons de ces tableaux est que les mots commençant par une consonne sont extrêmement nombreux. Dans les supports d'études visant la répartition des mots en fonction de leurs segments initiaux, 851 mots débutent par une voyelle, pour 2 903 par une consonne et pour seulement neuf mots par un glide. Normalement de tels effectifs devraient nous permettre de valider statistiquement aisément nos résultats.

Pourtant, force est de constater que nous ne relevons qu'un seul résultat significatif. Dans le corpus TOM ET ZOÉ, les mots débutant par une consonne (#C_) occupent majoritairement M2 dans 52,18 % des occurrences testant cette hypothèse ($\chi^2(1) = 4,1$; $p \leq 0,04$).

En fait, de l'observation par support d'études de la répartition en M1 et M2 des mots commençant par une consonne, nous constatons tout d'abord que :

- Malgré le grand nombre d'effectifs testés, la répartition des mots possédant une initiale consonantique est très hésitante ;
- Cependant, un seul et unique corpus met au jour un résultat dans lequel les mots en #C_ privilégient M2.

Il convient d'asseoir plus solidement ces constatations. D'autant plus que le résultat le plus catégorique extrait de ces tableaux n'est porté que par un seul support d'études, le corpus TOM ET ZOÉ dont nous indiquions à plusieurs reprises le particularisme. En effet, compte tenu de la façon dont il a été constitué d'autres facteurs phonologiques et/ou extraphonologiques pourraient intervenir et biaiser le résultat que nous venons de relever.

Afin d'écarter les quelques réserves dont nous avons toujours fait part, nous pensons que la comparaison des résultats pour l'ensemble des mots commençant par une consonne avec ceux obtenus pour le même ensemble auquel on aurait soustrait ceux recueillis dans le corpus TOM ET ZOÉ, nous permettra de lever nos doutes et de valider plus fermement des résultats.

Pour cela nous proposons les deux tableaux ci-dessous. Ceux-ci observent la répartition des initiales consonantiques. Le premier tableau, nommé TOTAL-TOM ET ZOÉ, présente les 749 mots en #C_ de l'ensemble des supports d'études sans ceux du corpus TOM ET ZOÉ. Le second, nommé TOTAL, recueille le total des 2 903 termes de l'ensemble des supports d'études commençant par une consonne. Les données présentées dans TOTAL sont extraites du tableau 32 page 149.

TOTAL -TOM ET ZOÉ	#C_
490	
M1	356
%	47,37%
M2	393
%	52,63%
Total :	749
	100,00%

**Tableau 45 : Répartition des consonnes
pour TOTAL-TOM ET ZOÉ**

TOTAL	#C_
1900	
M1	1386
%	47,74%
M2	1517
%	52,26%
Total :	2903
	100,00%

**Tableau 46 : Répartition des consonnes
pour TOTAL**

Ainsi, comme le montrent les tableaux ci-dessus, les deux résultats sont sensiblement identiques. Nous notons clairement qu'il existe une légère préférence à ce que les mots commençant par une consonne occupent prioritairement M2 (dans 52,63% des cas dans

TOTAL-TOM ET ZOÉ, et 52,26% des mots à initiale consonantique pour TOTAL).

Les résultats sont en termes de score, extrêmement homogènes. La seule différence notable provient de la validation statistique de ces mêmes résultats.

En effet, les 749 mots possédant une initiale consonantique de l'ensemble TOTAL-TOM ET ZOÉ ne permettent pas d'établir un résultat statistique aussi fort que le TOTAL des 2 903 mots testés sur cette même hypothèse. Nous ne pouvons exclure l'hypothèse nulle dans le premier cas ($\chi^2 (1) = 1,83$; $p \leq 0,17$), alors que les résultats fournis par l'ensemble des supports d'études sans restriction produisent une répartition significative ($\chi^2 (1) = 5,91$; $p \leq 0,01$).

Il nous est difficile d'incriminer le nombre d'effectifs testés pour justifier le fait que nous n'obtenions pas de résultats statistiques aussi forts que nous les souhaiterions. En effet que l'on considère le TOTAL-TOM ET ZOÉ et ses 749 mots à initiale consonantique examinés ou les 2 903 termes du tableau TOTAL, le nombre de mots éprouvés est très largement supérieur à ce qu'il est nécessaire de tester pour affirmer des résultats solides. Malgré le grand nombre d'effectifs testés, la répartition en M1 et M2 pour les mots commençant par une consonne est beaucoup moins saillante que celle que nous obtenions, par exemple, pour les mots à initiale vocalique.

Que faire de ces deux résultats ?

Si, de prime abord, les résultats énoncés par TOTAL nous laissaient supposer l'existence d'une corrélation entre la position M2 et les mots débutant par une consonne, les résultats émis par TOTAL-TOM ET ZOÉ nous indiquent qu'ils n'étaient en fait portés que par le corpus TOM ET ZOÉ. Pour les raisons déjà évoquées, nous ne pouvons conserver ce résultat.

Néanmoins, il est envisageable que, si nous ne pouvons assigner une position préférentielle aux mots en #C_, l'examen phonémique des initiales consonantiques mettent au jour de nouveaux résultats.

C'est pourquoi nous nous proposons à présent d'examiner dans le détail la répartition phonémique en M1 et en M2 des initiales consonantiques.

Comme nous l'avons proposé précédemment, nous présentons nos résultats, indépendamment, pour chaque support d'études impacté par l'hypothèse d'une corrélation entre les mots à initiale consonantique et la sélection d'une position dans la coordonnée. Nous les fournissons en l'état en annexe des tableaux A 6a à A 6g.

A nouveau, en raison entre autre de la dispersion du nombre de termes testés par support d'études et par hypothèses, ainsi que du fait des hésitations à sélectionner une position préférentielle pour les mots commençant par une consonne que nous relevions précédemment, il est difficile de relever des résultats statistiquement valides de l'observation des consonnes initiales de M1 et de M2 support d'études par support d'études.

En effet, dans le détail, les corpus CUL ET CHEMISE et ÉLISE ET ZOÉ ne proposent qu'un maximum de neuf mots (respectivement neuf et cinq mots) par hypothèses testées. Nous ne commenterons pas leurs résultats, les faibles effectifs éprouvés rendant ces commentaires hasardeux.

Les corpus VITE ET BIEN et LISSE ET SOYEUX et ARTS ET MÉTIERS, bien que testant un nombre suffisant d'items, ne produisent aucun résultat statistiquement valide. De surcroît, d'un corpus à l'autre, les résultats sont hétérogènes. Nous n'y voyons émerger aucun contour d'une dispersion par type de consonnes initiales. Le seul constat qu'il nous est possible de faire sur les résultats produits par ces trois corpus, c'est qu'ils sont hésitants. Les commenter plus en détail serait imprudent.

Seuls les corpus TOM ET ZOÉ et TARN ET GARONNE fournissent des résultats statistiquement valides. De l'observation des types de consonnes initiales par support d'études, nous extrayons six répartitions significatives et trois tendances.

Ainsi, le corpus TARN ET GARONNE met au jour deux tendances statistiques. Les mots #/m/_ privilégient massivement M2 dans 70% des mots examinés sur cette hypothèse ($\chi^2(1) = 3,2$; $p \leq 0,07$). Inversement, les mots #/s/_ sélectionnent dans 68,18% des occurrences M1 ($\chi^2(1) = 2,91$; $p \leq 0,09$).

Dans le corpus TOM ET ZOÉ, nous relevons six résultats statistiquement significatifs et une tendance.

Les mots commençant par /b/, par /f/, par /v/, par /z/ et par /n/ sélectionnent tous majoritairement M2, dans 61,04%, 70,37%, 66,67%, 57,87% et 65,98% des occurrences testées sur ces hypothèses (respectivement $\chi^2(1) = 3,75$; $p \leq 0,05$, $\chi^2(1) = 4,48$; $p \leq 0,03$, $\chi^2(1) = 10,67$; $p \leq 0,001$, $\chi^2(1) = 6,3$; $p \leq 0,01$ et $\chi^2(1) = 9,91$; $p \leq 0,002$).

Inversement, les mots débutant par /m/ privilégient M1 dans 55,58% des cas ($\chi^2(1) = 5,14$; $p \leq 0,02$), tout comme les mots #/p/_ qui ont tendance à occuper cette même position dans 58,12% des cas ($\chi^2(1) = 3,09$; $p \leq 0,08$).

Nous avançons ces résultats avec prudence.

En effet, ceux-ci ne sont essentiellement portés que par le corpus TOM ET ZOÉ. En raison du principe de précaution que nous avons précédemment invoqué, même si les résultats issus de TOM et ZOÉ sont établis statistiquement, nous préférons les confirmer plus avant

Ainsi, afin d'attester plus fermement les résultats présentés *supra*, nous nous proposons de les observer pour l'ensemble des supports d'études augmentant ainsi le nombre d'effectifs par hypothèses testées, mais également pour l'ensemble des supports d'études dont on aura extrait les données du corpus TOM ET ZOÉ

Nous proposons donc ci-dessous deux tableaux observant la répartition des initiales consonantiques. Dans le premier, nommé TOTAL-TOM ET ZOÉ, sont regroupés les 787 mots possédant une initiale consonantique dans l'ensemble des supports d'études sans ceux du corpus TOM ET ZOÉ. Nous prêterons une attention particulière à la répartition des consonnes initiales /b/, /f/, /v/, /ʒ/, /n/, /p/, /m/ et /f/. Le second tableau donne la répartition des mots débutant par une consonne initiale en M1 et M2 des 2941⁶³ mots composant l'effectif TOTAL, nous reviendrons sur les résultats des consonnes initiales que nous venons d'évoquer et nous y observerons spécifiquement les mots commençant par /s/.

63 Nous comptons 2 941 mots au lieu des 2 903 présentés précédemment dans le tableau 32, car nous comptabilisons ici les mots possédant un schwa interne. cf. page 25.

TOTAL-TOM ET ZOË 787	#C_																	
	p	t	k	b	d	g	f	s	ʃ	v	z	ʒ	l	m	n	ɲ	ɥ	ʀ
M1	41 51,90%	21 47,73%	38 48,72%	28 51,85%	19 38,78%	4 33,33%	17 43,59%	65 52,85%	10 50,00%	27 46,55%	0 0,00%	16 53,33%	20 51,28%	29 38,16%	17 54,84%			23 43,40%
M2	38 48,10%	23 52,27%	40 51,28%	26 48,15%	30 61,22%	8 66,67%	22 56,41%	58 47,15%	10 50,00%	31 53,45%	2 100,00%	14 46,67%	19 48,72%	47 61,84%	14 45,16%			30 56,60%
Total :	79 100,00%	44 100,00%	78 100,00%	54 100,00%	49 100,00%	12 100,00%	39 100,00%	123 100,00%	20 100,00%	58 100,00%	2 100,00%	30 100,00%	39 100,00%	76 100,00%	31 100,00%	0 100,00%	0 100,00%	53 100,00%

Tableau 47 : Répartition phonémique des initiales consonantiques pour TOTAL-TOM ET ZOË

TOTAL 2941	#C_																	
	p	t	k	b	d	g	f	s	ʃ	v	z	ʒ	l	m	n	ɲ	ɥ	ʀ
M1	109 55,61%	83 52,53%	167 47,44%	58 44,27%	38 40,43%	29 45,31%	25 37,88%	120 54,79%	23 50,00%	59 38,31%	12 48,00%	123 43,31%	194 48,26%	258 52,87%	50 39,06%			57 42,54%
M2	87 44,39%	75 47,47%	185 52,56%	73 55,73%	56 59,57%	35 54,69%	41 62,12%	99 45,21%	23 50,00%	95 61,69%	13 52,00%	161 56,69%	208 51,74%	230 47,13%	78 60,94%			77 57,46%
Total :	196 100,00%	158 100,00%	352 100,00%	131 100,00%	94 100,00%	64 100,00%	66 100,00%	219 100,00%	46 100,00%	154 100,00%	25 100,00%	284 100,00%	402 100,00%	488 100,00%	128 100,00%	0 100,00%	0 100,00%	134 100,00%

Tableau 48 : Répartition phonémique des initiales consonantiques pour TOTAL

Ainsi, comme le montre la comparaison des tableaux ci-dessus, concernant les mots possédant une consonne initiale, nous relevons cinq résultats et une tendance statistique.

Dans le détail, les mots $\#/f/_$ sont 62,12% à occuper M2 dans TOTAL, pour 56,41% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 3,88$; $p \leq 0,05$ et $\chi^2(1) = 0,64$; $p \leq 0,42$). Ces résultats semblent proches, seule leur validation statistique diffère. TOTAL met au jour une répartition significative, ce qui n'est pas le cas de TOTAL-TOM ET ZOÉ. Cette différence nous indique que le résultat assignant les mots $\#/f/_$ en M2 est essentiellement porté par le corpus TOM ET ZOÉ. En effet, sur 66 items testés dans TOTAL, 49 proviennent de TOTAL-TOM ET ZOÉ. De fait, nous ne pouvons conserver ce résultat.

De la même façon, les résultats pour les mots $\#/v/_$ présentent tous deux une prédisposition à placer les mots commençant par la consonne $/v/$ plutôt en M2 dans 61,69% des 154 mots testés sur cette hypothèse dans TOTAL ($\chi^2(1) = 8,42$; $p \leq 0,004$), pour 53,45% des 58 mots commençant par $\#/v/_$ dans TOTAL-TOM ET ZOÉ ($\chi^2(1) = 0,28$; $p \leq 0,60$). A nouveau, les deux résultats ne prétendent pas à la même validation statistique, ce qui signifie que les résultats sont biaisés par TOM ET ZOÉ. Nous ne les confirmons pas.

Les mots $\#/z/_$, quant à eux, présentent des résultats totalement divergents : 56,69% des termes commençant par $/z/$ sélectionnent M2 dans TOTAL, pour 53,33% en M1 dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 5,08$; $p \leq 0,02$ et $\chi^2(1) = 0,13$; $p \leq 0,72$). Nous ne pouvons les commenter.

Les résultats obtenus au sujet de l'hypothèse d'une corrélation entre une position dans la coordonnée et les mots commençant par $/n/$ font également apparaître des données disparates de la comparaison des tableaux TOTAL et TOTAL-TOM ET ZOÉ. En effet, si les mots en $\#/n/_$ dans TOTAL, comme dans ceux précédemment relevés page 170, occupent majoritairement M2 dans 60,94% des mots commençant par $\#/n/_$ ($\chi^2(1) = 6,13$; $p \leq 0,01$), ils sont, en revanche, 54,84% à préférer M1 dans le tableau TOTAL-TOM ET ZOÉ ($\chi^2(1) = 0,29$; $p \leq 0,59$).

Bien que le résultat proposé par le tableau TOTAL sur cette hypothèse rende compte d'une répartition statistique, il n'est néanmoins pas fiable. En effet, sur 128 mots $\#/n/_$ testés dans TOTAL, 97 proviennent de TOM ET ZOÉ. En fait, l'hypothèse liant les mots en $\#/n/_$ à M2 n'est réellement portée que par le corpus TOM ET ZOÉ.

Or, ce corpus constitué par la mise en relation de 100 prénoms pour l'ensemble de leurs

appariements mixtes possibles, n'est en fait composé que de trois prénoms éprouvant l'hypothèse d'une corrélation entre la position dans la coordination et les mots commençant par /n/ : « Nathan », « Nicolas » et « Noémie ». Pour reprendre la distinction proposée par Pierce (1931)⁶⁴, même si les « tokens » éprouvés sont au nombre conséquent de 128, il n'y a que trois « types » examinant les mots ayant pour initiale /n/. Difficile de ne pas supposer la possible influence d'autres facteurs tels que le genre ou le type de finale.

C'est pourquoi, comme précédemment, nous avons choisi de ne pas conserver ce résultat.

Concernant les mots ayant pour initiale /r/, les résultats sont homogènes dans les deux tableaux, mais peu solides statistiquement parlant. En effet, il apparaît que dans le TOTAL-TOM ET ZOÉ, les mots commençant par /r/ sélectionnent M2 dans 56,6% des cas, pour 57,46% dans TOTAL (respectivement ($\chi^2(1) = 0,92$; $p \leq 0,34$) et ($\chi^2(1) = 2,99$; $p \leq 0,08$). Si TOTAL fait émerger une tendance statistique, ce n'est pas le cas de TOTAL-TOM ET ZOÉ. De fait, nous ne pouvons affirmer la sélection d'une position dans la coordination par les mots en #/r/_.

Les résultats obtenus au sujet de l'hypothèse d'une corrélation entre une position dans la coordonnée et les mots commençant par /m/ font apparaître des données semblables entre les deux tableaux. En effet, les mots en #/m/_ dans les résultats de TOTAL occupent plutôt M1 dans 52,87% des mots commençant par #/m/_, tout comme dans le tableau TOTAL-TOM ET ZOÉ où ils sont 61,84% à occuper la même position (respectivement $\chi^2(1) = 1,61$; $p \leq 0,20$ et $\chi^2(1) = 4,26$; $p \leq 0,04$). A nouveau, les deux totaux ne présentent pas la même validation statistique. TOTAL-TOM ET ZOÉ met au jour un résultat, ce qui n'est pas le cas de TOTAL. Selon nous cette différence marque un biais porté par TOM ET ZOÉ. En effet, si TOTAL ne fait émerger de résultats pour les mots commençant par /m/, c'est parce que 435 mots sur les 488 testés proviennent de TOM ET ZOÉ. TOM ET ZOÉ a obscurci un résultat pourtant bien saillant de l'examen des six autres supports d'études. C'est pourquoi nous attestons le résultat relevé dans TOTAL-TOM ET ZOÉ et affirmons que les mots #/m/_ occupent majoritairement M2.

64 Peirce C. S., 1931-1935, *Collected Papers*, Cambridge, Harvard University Press.

La distinction entre type/token établit la différenciation entre deux façons de comptabiliser et/ou d'identifier des objets. Ainsi, on distinguera le général, ou « type », c'est-à-dire le genre d'un objet, du particulier nommé « token » correspondant au nombre concret d'instances ou d'occurrences d'un objet.

Dans la répartition phonémique des mots commençant par une initiale consonantique en M1 et M2, nous n'affirmons qu'un seul résultat : les mots #/m/_ privilégient M2.

4.2.1.2.2. Répartition par traits distinctifs des consonnes initiales

Dans cette section, nous nous sommes attachée à révéler la répartition des initiales consonantiques par traits distinctifs. En nous employant à observer ce que les consonnes ont de semblable ou de divergeant, et en scrutant précisément ce que la consonne /m/ a de particulier, nous souhaitons faire émerger des hypothèses jusqu'alors opaques.

Pour ce faire, dans un premier temps, nous relèverons les résultats présentés par chaque support d'études, indépendamment les uns des autres. Puis, nous confronterons les résultats retenus à l'ensemble de nos données.

Les résultats relevés pour chacun de nos supports d'étude interrogeant cette question, à savoir ici encore les six corpus et le questionnaire d'enquête ÉLISE ET ZOÉ, sont présentés en annexe dans les tableaux A 7a à A 7g.

L'observation par support d'étude de la répartition en M1 et en M2 des initiales consonantiques en fonction de leur trait distinctif fait apparaître treize répartitions significatives et cinq tendances statistiques assignant tous un trait distinctif à M2.

Dans le corpus TARN ET GARONNE, les mots #[+ SONANT] sont 63,41% à avoir tendance à occuper majoritairement M2 ($\chi^2(1) = 2,95$; $p \leq 0,08$).

Le corpus CUL ET CHEMISE fait émerger deux tendances.

77,78% des mots #[+ NASAL]_ ont tendance à sélectionner M2 ($\chi^2(1) = 2,78$; $p \leq 0,09$), tout comme les mots possédant une initiale consonantique [+ REL RET] dans 64,29% des occurrences testées ($\chi^2(1) = 3,43$; $p \leq 0,06$).

Le corpus TOM ET ZOÉ révèle huit répartitions significatives.

En apparence contradiction avec les trois tendances précédemment relevées, les mots commençant par une consonne [- SONANT] ou [- NASAL] ou [- REL RET] privilégient également M2 dans 53,37%, 53,31% et 54,97% des mots testant ces hypothèses (respectivement $\chi^2(1) = 5,46$; $p \leq 0,02$, $\chi^2(1) = 7,22$; $p \leq 0,007$ et $\chi^2(1) = 9,54$; $p \leq 0,002$). Les mots #[+ CONTINU]_, #[+ HAUT]_, #[- ANTÉRIEUR]_, #[+ CORONAL]_ et #[+ VOISÉ]_ sélectionnent eux aussi M2 dans 52,75%, 54,79%, 55,17%, 53,34% et 53,67% des

mots examinant ces hypothèses (respectivement $\chi^2(1) = 4,45$; $p \leq 0,03$, $\chi^2(1) = 5,55$; $p \leq 0,02$, $\chi^2(1) = 7,34$; $p \leq 0,007$, $\chi^2(1) = 4,54$; $p \leq 0,03$ et $\chi^2(1) = 8,07$; $p \leq 0,005$).

Le corpus ARTS ET MÉTIERS met en avant cinq résultats et une tendance statistique.

Conformément à ce que nous relevions dans le corpus TOM ET ZOÉ, les mots dont l'initiale consonantique est [- SONANT] ou [- NASAL] ou [+ HAUT] ou [- ANTÉRIEUR] privilégient M2 dans 59,80%, 59,63%, 80% et 73,08% des occurrences testant ces hypothèses (respectivement $\chi^2(1) = 3,92$; $p \leq 0,05$, $\chi^2(1) = 4,05$; $p \leq 0,04$, $\chi^2(1) = 7,2$; $p \leq 0,01$ et $\chi^2(1) = 5,54$; $p \leq 0,02$).

Inversement, en contradiction avec ce que nous observions dans TOM ET ZOÉ, les mots #[+ ARRIÈRE]_ occupent également M2 dans 75% des cas ($\chi^2(1) = 4$; $p \leq 0,05$).

Enfin, émerge un nouveau résultat : les mots #[- ARRIÈRE]_ ont tendance à sélectionner majoritairement M2 dans 60,87% des cas ($\chi^2(1) = 3,26$; $p \leq 0,07$).

L'examen de la répartition en M1 et en M2 des initiales consonantiques par traits distinctifs met au jour des résultats semblant corrélés l'assignation d'un trait à une position préférentielle d'une séquence binaire coordonnée. Cependant trois résultats semblent *a priori* contradictoires : les mots commençant par une consonne soit [\pm SONANT], soit [\pm NASAL], soit [\pm VOISE] privilégient tous la position M2. C'est pourquoi afin de faire la lumière sur ces apparentes incohérences, nous proposons dans les deux tableaux à suivre l'observation des traits distinctifs pour les deux totaux TOTAL et TOTAL-TOM ET ZOÉ.

Pour rappel comme précédemment, nous avons mentionné les résultats significatifs en caractères gras. Les tendances sont indiquées en italique.

TOTAL 2941	#C_																	
	SONANT		CONTINU		NASAL		HAUT		ARRIERE		ANTERIEUR		CORONAL		VOISE		REL RET	
	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-
M1	559 48,52%	846 47,29%	921 47,33%	484 48,64%	308 50,00%	1097 47,18%	342 45,84%	1063 48,43%	253 46,00%	<i>1152</i> 48,18%	1006 48,81%	399 45,34%	<i>643</i> 47,42%	762 48,08%	878 46,11%	527 50,82%	613 46,09%	792 49,16%
M2	593 51,48%	943 52,71%	1025 52,67%	511 51,36%	308 50,00%	1228 52,82%	404 54,16%	1132 51,57%	297 54,00%	<i>1239</i> 51,82%	1055 51,19%	481 54,66%	<i>713</i> 52,58%	823 51,92%	1026 53,89%	510 49,18%	717 53,91%	819 50,84%
Total :	1152 100,00%	1789 100,00%	1946 100,00%	995 100,00%	616 100,00%	2325 100,00%	746 100,00%	2195 100,00%	550 100,00%	2391 100,00%	2061 100,00%	880 100,00%	1356 100,00%	1585 100,00%	1904 100,00%	1037 100,00%	1330 100,00%	1611 100,00%

Tableau 49 : Répartition par traits distinctifs des initiales consonantiques dans TOTAL

TOTAL-TOM ET ZOÉ 787	#C_																	
	SONANT		CONTINU		NASAL		HAUT		ARRIERE		ANTERIEUR		CORONAL		VOISE		REL RET	
	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-
M1	89 44,72%	286 48,64%	224 47,56%	151 47,78%	46 42,99%	329 48,38%	68 48,57%	307 47,45%	65 45,45%	310 48,14%	284 47,81%	91 47,15%	168 49,70%	207 46,10%	183 45,30%	192 50,13%	178 48,90%	197 46,57%
M2	110 55,28%	302 51,36%	247 52,44%	165 52,22%	61 57,01%	351 51,62%	72 51,43%	340 52,55%	78 54,55%	334 51,86%	310 51,19%	102 52,85%	170 50,30%	242 53,90%	221 54,70%	191 49,87%	186 51,10%	226 53,43%
Total :	199 100,00%	588 100,00%	471 100,00%	316 100,00%	107 100,00%	680 100,00%	140 100,00%	647 100,00%	143 100,00%	644 100,00%	594 100,00%	193 100,00%	338 100,00%	449 100,00%	404 100,00%	383 100,00%	364 100,00%	423 100,00%

Tableau 50 : Répartition par traits distinctifs des initiales consonantiques dans TOTAL-TOM ET ZOÉ

Les deux totaux présentés ci-dessus sont visiblement très disparates. Aucune des sept répartitions significatives et des trois tendances statistiques émergeant de TOTAL n'est confirmée par TOTAL-TOM ET ZOÉ.

Or, comme précédemment lorsque TOTAL présente un résultat non validé par TOTAL-TOM ET ZOÉ, cela signifie que TOM ET ZOÉ biaise les résultats de TOTAL. De fait, nous ne retiendrons aucun résultat émergeant du tableau 49.

Le tableau TOTAL-TOM ET ZOÉ met néanmoins au jour deux tendances statistiques.

Les mots #[- CORONAL]_ ont tendance à sélectionner majoritairement M2 dans 53,9% des 449 mots #[- CORONAL]_ examinés ($\chi^2(1) = 2,73$; $p \leq 0,10$). De la même façon, les mots possédant une initiale consonantique [+ VOISÉ] ont également tendance à favoriser M2 dans 54,7% des occurrences testant cette hypothèse ($\chi^2(1) = 3,57$; $p \leq 0,06$)

4.2.1.2.3. Bilan et analyse des résultats concernant #C_

Au regard de l'examen détaillé des initiales consonantiques, nous relevons trois résultats :

- Les mots #/m/_ occupent majoritairement M2 ;
- De plus les mots #[- CORONAL]_ (/p, b, k, g, f, v, m, ŋ, R/), ainsi que les mots #[+ VOISÉ]_ (/b, d, g, v, z, ʒ, l, m, n, ɲ, R/), ont tendance à sélectionner la même position.

Comment expliquer ces résultats ?

Notons tout d'abord que les traits [- CORONAL] et [+ VOISÉ] spécifient tous deux la consonne /m/. De fait, nous pensons que la tendance plaçant les mots commençant par une consonne [- CORONAL] et/ou [+ VOISÉ] n'est que l'artefact de l'assignation privilégiée de mots commençant par /m/ en M2.

La question est donc de savoir pourquoi les mots #/m/_ occupent majoritairement M2.

Selon nous, les efforts articulatoires en initiale de mots peuvent rendre compte de la présence préférentielle des initiales /m/ en M2.

En effet, comme le signale Fougeron (1998), la phase initiale demande plus d'effort articulatoire que les segments en finale. De fait, les segments en position initiale sont souvent

préservés alors que ceux qui sont en position plus faible peuvent être supprimés. Wright *et al.* (2005) associent ce même principe au principe de distribution asymétrique des phonèmes (Martinet (1955)) : les segments en position forte, c'est-à-dire en initiale de mots, sont plus susceptibles d'être des obstruantes, et ceux en position faible auront plus tendance à être des sonantes.

Selon ce principe, la consonne initiale /m/, sonante, a plus tendance à occuper la position faible de la séquence binomiale coordonnée, M2.

Fortes des résultats que nous venons de présenter, nous pensons que la corrélation entre une position privilégiée de la coordonnée par «et» et les mots possédant une initiale consonantique en termes de type et de traits n'est pas une corrélation forte. Elle est tout au moins beaucoup moins forte que celle que nous affirmions précédemment entre la position M1 et les mots commençant par une voyelle.

4.2.1.3. Glide

Comme nous l'indiquions page 149 en examinant la répartition des mots en M1 et en M2 en fonction de leur segment initial, s'il apparaît de façon assez solide que les mots commençant par une voyelle sélectionnent prioritairement M1, nous ne pouvons rien affirmer concernant les mots débutant par un glide.

En effet, sur les 1 900 appariements que nous avons testés au sujet d'une possible corrélation entre le segment initial et une position privilégiée de la coordonnée (c'est-à-dire ceux relevés dans la collection des six corpus et dans le questionnaire d'enquête ÉLISE ET ZOÉ), seuls neuf mots au total commencent par un glide.

Il nous est complètement impossible d'extraire des résultats fiables. Et même si intuitivement et au regard des données chiffrées, il semblerait que les mots commençant par un glide sélectionneraient plutôt M1, il ne s'agit là que de notre intuition de locutrice native du français, des recherches ultérieures devront le confirmer.

4.2.2. Syllabe initiale

Nous questionner sur le lien existant entre le type de segment initial et la sélection d'une forme coordonnée nous a mené à nous interroger sur la structure de la syllabe initiale:

En effet, si nous venons de définir une corrélation entre les initiales vocaliques et M1, comment ne pas supposer que la structure de la syllabe initiale pourrait avoir une incidence dans la sélection d'une forme coordonnée plutôt que son contraire.

C'est ce que nous nous proposons de relever dans les sections à venir.

Pour cela, nous utiliserons la collection des six corpus et le questionnaire d'enquêtes ÉLISE ET ZOÉ, les autres questionnaires étant neutralisé sur cette question.

Rappelons que comme nous l'indiquons page 25 dans la section 1.2.2. Le schwa, nous avons retiré de l'examen de ces données, les 38 paires dans lesquelles un schwa interne est potentiellement réalisé.

Ainsi, afin d'explorer en détail les patrons de syllabes initiales, nous nous attardons d'abord sur les syllabes à initiales vocaliques, puis sur les syllabes possédant une initiale consonantique, et enfin, nous évoquerons les mots dont la première syllabe débutent par un glide.

4.2.2.1. Structure de la syllabe initiale vocalique

Dans cette partie, nous proposons un examen détaillé des différents patrons syllabiques en début de mots commençant par une voyelle présents dans nos supports d'études, à savoir #V, #VC et #VCC. Pour ce faire, nous avons utilisé l'ensemble des six corpus ainsi que le questionnaire d'enquête ÉLISE ET ZOÉ, les autres questionnaires d'enquête étaient neutralisés en amont sur la question du type d'initiales.

Nos supports d'études étant autonomes, il nous a semblé pertinent de présenter nos résultats support d'études par support d'études.

Dans le composite de tableau ci-dessous, nous avons comme précédemment et comme pour l'ensemble de nos résultats, noté en caractères gras les résultats pour lesquels le χ^2 est de $p \leq 0,05$ et en italiques ceux pour lesquels, il est compris entre $0,05 < p \leq 0,1$. De plus, afin de les rendre visiblement accessibles, nous avons souligné les résultats pour lesquels la valeur-p était très légèrement au dessus de 0,1, nous les commenterons à suivre.

TOTAL	type #σ		
	#V	#VC	#VCC
M1	420	70	1
%	56,07%	71,43%	25,00%
M2	329	28	3
%	43,93%	28,57%	75,00%
Total :	749	98	4
	100,00%	100,00%	100,00%

**Tableau 51b : Répartition par syllabes
vocaliques initiales
dans TARN ET GARONNE**

TARN ET GARONNE 69	type #σ								
	#CGVC	#CGV	#CV	#CVC	#CCV	#CCVC	#CCVCC	#CCVCC	#CVG
M1	0		35	15,00	3	1	3		
%	0,00%		44,30%	0,56	60,00%	100,00%	50,00%		
M2	4		44	12,00	2	0	3		
%	100,00%		55,70%	0,44	40,00%	0,00%	50,00%		
Total :	4	0	79	27	5	1	6	0	0
	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

**Tableau 51b : Répartition par syllabes
vocaliques initiales
dans CUL ET CHEMISE**

CUL ET CHEMISE 42	type #σ								
	#CGVC	#CGV	#CV	#CVC	#CCV	#CCVC	#CCVCC	#CCVCC	#CVG
M1	2	1	14	12	2	1	1	0	
%	100,00%	33,33%	51,85%	38,71	40,00%	100,00%	50,00%		
M2	0	2	13	19	3	0	1	0	0
%	0,00%	66,67%	48,15%	61,29	60,00%	0,00%	50,00%		
Total :	2	3	27	31	5	1	2	0	0
	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

**Tableau 51c : Répartition par syllabes
vocaliques initiales
dans LISSE ET SOYEUX**

VITE ET BIEN 202	type #σ		
	#V	#VC	#VCC
M1	25	6	0
%	52,08%	54,55%	0,00%
M2	23	5	1
%	47,92%	45,45%	100,00%
Total :	48	11	1
	100,00%	100,00%	100,00%

**Tableau 51d : Répartition par syllabes
vocaliques initiales dans TOM ET ZOÉ**

VITE ET BIEN 202	type #σ		
	#V	#VC	#VCC
M1	25	6	0
%	52,08%	54,55%	0,00%
M2	23	5	1
%	47,92%	45,45%	100,00%
Total :	48	11	1
	100,00%	100,00%	100,00%

**Tableau 51e : Répartition par syllabes
vocaliques initiales dans VITE ET BIEN**

LISSE ET SOYEUX 62	type #σ								
	#CGVC	#CGV	#CV	#CVC	#CCV	#CCVC	#CCVCC	#CCVCC	#CVG
M1	0	0	32	11,00	2		3		
%	0,00%	0,00%	47,06%	0,73	28,57%		75,00%		
M2	2	1	36	4,00	5		1		
%	100,00%	100,00%	52,94%	0,27	71,43%		25,00%		
Total :	2	1	68	15	7	0	4	0	0
	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

**Tableau 51f : Répartition par syllabes
vocaliques initiales dans ARTS ET
MÉTIER**

TOM ET ZOÉ 1420	type #σ								
	#CGVC	#CGV	#CV	#CVC	#CCV	#CCVC	#CCVCC	#CCVCC	#CVG
M1	32		741	180,00	77				
%	60,38%		47,35%	0,49	46,39%				
M2	21		824	190,00	89				
%	39,62%		52,65%	0,51	53,61%				
Total :	53	0	1565	370	166	0	0	0	0
	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

**Tableau 51g : Répartition par syllabes
vocaliques initiales dans ÉLISE ET ZOÉ**

Les tableaux ci-dessus présentent les résultats de chaque support d'études relevés sur la question de la répartition par type de syllabes vocaliques initiales dans une des deux positions dans la coordonnée.

Dans le détail, nous observons fait trois dispersions statistiquement valides ($p \leq 0,05$) concernant #V._ et #VC._ et une tendance statistique.

Les corpus TARN ET GARONNE, CUL ET CHEMISE, LISSE ET SOYEUX, VITE ET BIEN, et le questionnaire d'enquête ÉLISE ET ZOÉ ne présentent aucun résultat. Cette absence de résultats est selon nous directement imputable à la dispersion par nombre de patrons syllabiques. En effet, nous avons organisé la répartition des syllabes initiales par type selon treize patrons syllabiques : trois pour les syllabes débutant par une voyelle, neuf pour ceux débutant par une consonne et un pour les mots commençant par un glide. Rappelons de plus que les mots commençant par une initiale vocalique ne représentent, avec 851 termes en #V._, que 22,6% de l'ensemble des mots testés sur les types de syllabes initiales. En conséquence de ceci, les hypothèses examinées sont particulièrement réparties. Raison pour laquelle, nous ne pouvons extraire de résultats pour les corpus TARN ET GARONNE, LISSE ET SOYEUX et le questionnaire d'enquête ÉLISE ET ZOÉ (respectivement testés au maximum par hypothèses sur 15, 26 et 13 mots).

Au contraire, le corpus TOM ET ZOÉ, fort du grand nombre d'items dont il est composé, fait apparaître deux résultats statistiquement significatifs dans lesquels les mots en #V._ occupent M1 dans 55,39 % des cas ($\chi^2 (1) = 7,23 ; p \leq 0,007$), tout comme ceux en #VC._ qui sélectionnent la même position dans 71,19% des occurrences testées sur cette hypothèse ($\chi^2 (1) = 10,59 ; p \leq 0,001$).

Le corpus ARTS ET MÉTIERS met au jour une tendance statistique et un résultat significatif. Sur les 20 mots en #VC._ testés dans ce corpus 75% sélectionnent majoritairement M1 ($\chi^2 (1) = 5 ; p \leq 0,02$). De la même façon, les mots en #V ont également tendance à occuper la première position de la séquence binomiale coordonnée dans 66,67% des occurrences testant cette hypothèse ($\chi^2 (1) = 2,67 ; p \leq 0,10$.)

De manière générale, de l'examen de chaque support d'études étudié indépendamment les uns des autres des mots commençant par une syllabe vocalique, nous observons donc, que les mots en #V._ et #VC._ occupent majoritairement M1.

Comme nous l'indiquions *supra*, comme le petit nombre d'items testés semble avoir été vraisemblablement un frein à l'observation de résultats solides, nous avons souhaité asseoir plus fortement nos précédents résultats en observant la syllabe en début de mots des termes commençant par #V_ pour l'ensemble des supports d'études.

Nous allons donc présenter deux tableaux observant la répartition des syllabes vocaliques en début de mots, par patrons. Le premier tableau, nommé TOTAL-TOM ET ZOÉ, livre la répartition en M1 et M2 des syllabes en #V_ des 171 mots composant l'ensemble des données de nos supports d'études sans celles issues du corpus TOM ET ZOÉ. Dans le second tableau, TOTAL, nous présenterons l'ensemble de la répartition par patron des syllabes en #V_ pour tous les supports d'études en confrontant spécifiquement les trois précédents résultats à ceux de TOTAL-TOM ET ZOÉ.

TOTAL-TOM ET ZOÉ	type #σ_		
	#V	#VC	#VCC
M1	76	28	1
%	59,38%	71,79%	25,00%
M2	52	11	3
%	40,58%	27,50%	60,00%
Total :	128	39	4
	100,00%	100,00%	100,00%

Tableau 52 : Répartition par patrons des syllabes vocaliques initiales pour TOTAL-TOM ET ZOÉ

TOTAL	type #σ_		
	#V	#VC	#VCC
M1	420	70	1
%	56,07%	71,43%	25,00%
M2	329	28	3
%	43,93%	28,57%	75,00%
Total :	749	98	4
	100,00%	100,00%	100,00%

Tableau 53 : Répartition par patrons des syllabes vocaliques initiales pour l'ensemble des supports d'études

D'une manière générale, les deux tableaux font apparaître des résultats très similaires en adéquation avec nos précédents résultats.

Revenons sur les deux patrons pour lesquels nous observons des résultats de l'examen de nos supports d'études pris indépendamment les uns des autres, à savoir les mots en #V._ et ceux en #VC._⁶⁵.

Concernant les mots en #V._, les résultats sont hétérogènes.

Les mots en #V._ sont 56,07% dans TOTAL et 59,38% dans TOTAL-TOM ET ZOÉ à sélectionner majoritairement M1 (respectivement $\chi^2(1) = 11,06$; $p \leq 0,00088$ et $\chi^2(1) = 4,5$; $p \leq 0,03$). Ces deux résultats sont identiques. Nous affirmons donc que les mots en #V._ sélectionnent majoritairement M1.

De la même façon, nous relevons également de la confrontation des deux tableaux pour les mots en #VC._ des résultats très proches, statistiquement significatifs. En effet, dans TOTAL-TOM ET ZOÉ, 71,79% des 39 mots en #VC._ privilégient massivement M1 ($\chi^2(1) = 48,5$; $p \leq 0,006$) et ils sont 71,43% à sélectionner la même position dans TOTAL ($\chi^2(1) = 18$; $p \leq 0,00002$). Compte tenu de ces deux résultats marqués, nous attestons que les mots en #VC._ occupent prioritairement M1

D'une manière générale, de la répartition en M1 et M2 des syllabes vocaliques initiales par patrons, en adéquation avec les résultats présentés page 164 attestant de l'existence d'une corrélation forte entre les mots débutant par une initiale vocalique et M1, nous affirmons donc que:

- les mots en #V._ occupent majoritairement M1,
- tout comme les mots en #VC._ privilégiant massivement cette position.

4.4.2..2. Structure de la syllabe initiale consonantique

Dans cette section, nous examinerons les huit patrons des syllabes initiales commençant par une consonne présents dans nos supports d'études, à savoir #CV._, #CVC._, #CVCC._, #CCV._, #CCVC._, #CGV._, #CVG._, #CGVC._.

Pour ce faire, nous avons à nouveau utilisé l'ensemble des six corpus et le questionnaire d'enquête ÉLISE ET ZOÉ, les autres questionnaires d'enquêtes n'explorant pas cette hypothèse étaient neutralisés en amont.

Nos supports d'études étant autonomes, il nous a semblé pertinent de présenter nos résultats

⁶⁵ Nous ne commenterons pas les résultats relevés pour les mots en #VCC._, testés sur seulement quatre items, les scores de 25% et 75% ne peuvent en aucun cas être considérés comme des résultats probants.

support d'études par support d'études comme nous les proposons ci-dessous.

TARN ET GARONNE 69	type #σ								
	#CGVC	#CGV	#CV	#CVC	#CCV	#CCVC	#CVCC	#CCVCC	#CVG
M1	0		35	15,00	3	1	3		
%	0,00%		44,30%	0,56	60,00%	100,00%	50,00%		
M2	4		44	12,00	2	0	3		
%	100,00%		55,70%	0,44	40,00%	0,00%	50,00%		
Total :	4	0	79	27	5	1	6	0	0
	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

Tableau 54 a : Répartition par patrons des syllabes initiales consonantiques pour TARN ET GARONNE

CUL ET CHEMISE 42	type #σ								
	#CGVC	#CGV	#CV	#CVC	#CCV	#CCVC	#CVCC	#CCVCC	#CVG
M1	2	1	14	12	2	1	1	0	
%	100,00%	33,33%	51,85%	38,71	40,00%	100,00%	50,00%		
M2	0	2	13	19	3	0	1	0	0
%	0,00%	66,67%	48,15%	61,29	60,00%	0,00%	50,00%		
Total :	2	3	27	31	5	1	2	0	0
	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

Tableau 54 b : Répartition par patrons des syllabes initiales consonantiques pour CUL ET CHEMISE

LISSE ET SOYEUX 62	type #σ								
	#CGVC	#CGV	#CV	#CVC	#CCV	#CCVC	#CVCC	#CCVCC	#CVG
M1	0	0	32	11,00	2		3		
%	0,00%	0,00%	47,06%	0,73	28,57%		75,00%		
M2	2	1	36	4,00	5		1		
%	100,00%	100,00%	52,94%	0,27	71,43%		25,00%		
Total :	2	1	68	15	7	0	4	0	0
	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

Tableau 54 c : Répartition par patrons des syllabes initiales consonantiques pour LISSE ET SOYEUX

	type #σ								
TOM ET ZOÉ 1420	#CGVC	#CGV	#CV	#CVC	#CCV	#CCVC	#CVCC	#CCVCC	#CVG
M1	32		741	180,00	77				
%	60,38%		47,35%	0,49	46,39%				
M2	21		824	190,00	89				
%	39,62%		52,65%	0,51	53,61%				
Total :	53	0	1565	370	166	0	0	0	0
	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

**Tableau 54 d : Répartition par patrons des syllabes initiales consonantiques pour
TOM ET ZOÉ**

	type #σ								
VITE ET BIEN 202	#CGVC	#CGV	#CV	#CVC	#CCV	#CCVC	#CVCC	#CCVCC	#CVG
M1	4	5	74	42	18	6	4		
%	66,67%	31,25%	46,84%	56,00	51,43%	54,55%	80,00%		
M2	2	11	84	33	17	5	1		
%	33,33%	68,75%	53,16%	44,00	48,57%	45,45%	20,00%		
Total :	6	16	158	75	35	11	5	0	0
	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

**Tableau 54 e : Répartition par patrons des syllabes initiales consonantiques pour
VITE ET BIEN**

	#CGVC	#CGV	#CV	#CVC	#CCV	#CCVC	#CVCC	#CCVCC	#CVG
ARTS ET MÉTIERS 85									
M1	8	4	23	10,00	1	6	2		
%	100,00%	66,67%	34,33%	0,45	8,33%	75,00%	66,67%		
M2	0	2	44	12,00	11	2	1		
%	0,00%	33,33%	65,67%	0,55	91,67%	25,00%	33,33%		
Total :	8	6	67	22	12	8	3	0	0
	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

**Tableau 54 f : Répartition par patrons des syllabes initiales consonantiques pour ARTS
ET MÉTIERS**

ÉLISE ET ZOÉ 20	type #σ ₋								
	#CGVC	#CGV	#CV	#CVC	#CCV	#CCVC	#CVCC	#CCVCC	#CVG
M1		0	3	7	1		0		
%		0,00%	20,00%	100,00%	33,33%		0,00%		
M2		1	12	0	2		1		
%		100,00%	80,00%	0,00%	66,67%		100,00%		
Total :	0	1	15	7	3	0	1	0	0
	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

Tableau 54 g : Répartition par patrons des syllabes initiales consonantiques pour ÉLISE ET ZOÉ

L'ensemble des tableaux ci-dessus présente les résultats de chaque support d'études relevés sur la question de la répartition par type de syllabes initiales consonantiques dans l'une des deux positions de la séquence binomiale coordonnée.

Dans le détail, nous relevons quatre résultats statistiquement significatifs et une tendance statistique. Ces résultats concernent les mots ayant comme patrons de syllabes initiales : #CV._, #CVC._ et #CCV._, les cinq autres hypothèses ne font apparaître de résultats solides.

Plus précisément, support d'études par support d'études.

Les corpus TARN ET GARONNE, CUL ET CHEMISE, VITE ET BIEN ne présentent aucun résultat. Deux raisons à cette absence de résultats : d'une part certaines structures sont sous-représentées, nous pensons particulièrement aux mots en #CVCC._, #CCVC._, et #CVG._. (Ils sont un maximum de 21 mots testés par hypothèse pour l'ensemble des supports d'études). Dans ce cas, l'absence de résultats est selon nous directement imputable à la dispersion par nombre de patrons syllabiques.

D'autre part, en ce qui concerne les autres hypothèses, nous pensons que cette même absence est le signe d'une hésitation. Rappelons que les résultats précédemment obtenus que nous présentons page 178, ne permettent pas de déterminer l'existence d'un lien entre les mots débutant par une consonne et la sélection d'un ordre préférentiel de coordination binaire directe par « et ». Nous parvenons toutefois à voir émerger des résultats.

Dans le corpus ARTS ET MÉTIERS, nous relevons deux résultats.

Ainsi, les mots en #CV._ occupent majoritairement M2 dans 65,67% des termes examinés sur cette hypothèse ($\chi^2(1) = 6,5$; $p \leq 0,01$), tout comme les mots en #CCV._ qui sélectionnent

massivement cette même position M2 dans 91,67% des cas ($\chi^2(1) = 8,33$; $p \leq 0,003$).

Le questionnaire d'enquête ÉLISE ET ZOÉ met au jour une répartition significative.

Les mots en #CV._ sélectionnent très majoritairement M2 dans 80% des cas ($\chi^2(1) = 5,4$; $p \leq 0,02$).

Nous ne relevons pas les résultats des mots en #CVC._ qui bien que systématiques, ne peuvent en aucun cas être validés par un test de Khi Deux probant. L'échantillon testé (sept mots) est beaucoup trop restreint pour prétendre à une validation statistique.

Le corpus TOM ET ZOÉ, malgré le grand nombre d'items dont il est composé, ne rend compte que d'un seul et unique résultat : sur les 1565 mots en #CV._ testés, 52,65% d'entre eux favoriseraient M2 ($\chi^2(1) = 4,4$; $p \leq 0,04$).

Enfin, le corpus LISSE ET SOYEUX révèle une tendance statistique.

Les mots en #CVC._ sont 73,33% à occuper préférentiellement M1. ($\chi^2(1) = 3,27$; $p \leq 0,07$).

Ce dernier résultat est à l'opposé de ce que l'ensemble des autres résultats semble avancer. Il conviendra de le vérifier ultérieurement.

De manière générale, de l'examen de chaque support d'études étudié indépendamment les uns des autres des mots commençant par une syllabe consonantique, nous constatons, que les mots en #CV._ et #CCV._ occuperaient majoritairement M2, contrairement à ceux en #CVC._ qui sélectionneraient M1.

Nous avons souhaité asseoir plus fortement nos précédents résultats en observant les syllabes initiales des mots commençant par #C._ pour l'ensemble des supports d'études.

Pour cela, nous nous proposons de présenter deux tableaux observant la répartition des syllabes consonantiques initiales par patrons. Comme précédemment ce choix de deux tableaux comparatifs trouve sa justification dans la constitution même du corpus TOM ET ZOÉ. Ce corpus, en examinant la fréquence de réalisations d'appariements de prénoms non mixtes sur le Web, est non seulement le moins contrôlé de nos supports d'études, mais il est également celui qui propose le plus grand nombre de données. De plus, constitué d'appariements construits à partir des 100 paires de prénoms les plus fréquemment donnés en 2008, nous pouvons envisager qu'un seul et même prénom, par ses multiples réalisations en termes de tokens, galvaude les résultats. Afin d'éviter que TOM ET ZOÉ ne biaise nos résultats, nous proposons systématiquement de confronter l'ensemble de nos résultats

observés avec, et sans ses données.

Ainsi, dans le premier total, ici encore nommé TOTAL-TOM ET ZOÉ, nous mettons en avant la répartition en M1 et M2 des syllabes en #C_ des 749 mots composant l'ensemble des données de nos supports d'études sans celles issues du corpus TOM ET ZOÉ.

Dans le second tableau, nommé TOTAL, nous présenterons l'ensemble de la répartition par patron des syllabes initiales en #C_ pour tous les supports d'étude.

TOTAL-TOM ET ZOÉ 749	type #σ_								
	#CGVC	#CGV	#CV	#CVC	#CCV	#CCVC	#CVCC	#CCVCC	#CVG
M1	14	10	181	97	27	14	13	0	
%	63,64%	37,04%	43,72%	54,80%	40,30%	66,67%	61,90%		
M2	8	17	233	80	40	7	8	0	
%	39,13%	58,62%	56,28%	44,09%	60,56%	33,33%	38,10%		
Total :	22	27	414	177	67	21	21	0	0
	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

Tableau 55 : Répartition par patrons des syllabes initiales consonantiques pour TOTAL-TOM ET ZOÉ

TOTAL-TOM ET ZOÉ 179	Vocale #						
	i	/E/	/A/	y	/œ/	u	/O/
M1	21	23	15	1	2	3	13
%	46,67%	42,59%	37,50%	50,00%	40,00%	60,00%	46,43%
M2	24	31	<u>25</u>	1	3	2	15
%	53,33%	57,41%	62,50%	50,00%	60,00%	40,00%	53,57%
Total :	45	54	40	2	5	5	28
	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

Tableau 56 : Répartition par patrons des syllabes initiales consonantiques pour TOTAL

De prime abord, la confrontation des deux tableaux fait apparaître trois répartitions significatives.

Revenons sur les trois patrons pour lesquels nous observons des résultats de l'examen de nos supports d'études pris indépendamment les uns des autres, à savoir les mots en #CV._ tout d'abord et ceux en #CVC._, et #CCV._ à suivre⁶⁶.

⁶⁶ Nous ne commenterons pas les résultats relevés pour les mots en #CVCC._, #CCVC._, et #CVG._, testés sur seulement 21 items et sans validation statistique, ils ne peuvent en aucun cas être considérés comme des

Concernant les mots en #CV._, les résultats sont hétérogènes et significatifs.

Dans TOTAL-TOM ET ZOÉ, les mots en #CV._ occupent en M2 dans 56,28% des mots testés sur cette hypothèse ($\chi^2(1) = 6,53$; $p \leq 0,01$), tout comme dans TOTAL, où 53,41% des mots en #CV._ ($\chi^2(1) = 9,21$; $p \leq 0,00002$) favorisent M2.

Nous affirmons donc que les mots en en #CV._ sélectionnent majoritairement M2.

Les mots en #CCV._ ont, quant à eux, tendance à occuper M2 dans 55,73% des cas dans TOTAL, pour 59,7% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 2,68$; $p \leq 0,10$ et $\chi^2(1) = 2,52$; $p \leq 0,11$). Les résultats relevés sont extrêmement proches tant en termes de scores que de validations statistiques. A une nuance près : si TOTAL avance une tendance, ce n'est pas le cas de TOTAL-TOM ET ZOÉ. La valeur asymptotique juste au-deçà du seuil de .1 nous laisse à penser qu'il existe vraisemblablement un lien entre les mots en #CCV._ et M2, et intuitivement, nous serions tentée de l'avancer. Cependant, afin d'être cohérente et rigoureuse dans notre démarche, nous ne conservons pas ce résultat.

La confrontation des deux totaux révèle que les mots en #CVC._ favorisent la position M1 dans 50,64% des cas dans TOTAL, pour 54,8% TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 0,09$; $p \leq 0,72$ et $\chi^2(1) = 1,63$; $p \leq 0,20$). Les deux résultats, tant par leur score que par leur validation statistique, font montre d'une incapacité à sélectionner une position préférentielle.

Enfin, les mots en #CGVC._ occupent M1 dans 61,33% des cas dans TOTAL pour 63,64% TOTAL – TOM ET ZOÉ (respectivement $\chi^2(1) = 3,85$; $p \leq 0,05$ et $\chi^2(1) = 1,64$; $p \leq 0,20$). Les résultats sont *a priori* semblables. Nous notons cependant que TOTAL met au jour une répartition significative, ce qui n'est pas le cas de TOTAL-TOM ET ZOÉ. Cette différence de validation statistique signale un biais porté par le corpus TOM ET ZOÉ. En effet, sur le 75 mots en #CGVC._ examinés dans TOTAL, 53 sont issus du corpus TOM ET ZOÉ. Ce résultat n'est en fait validé que par un seul corpus, nous ne pouvons le conserver.

D'une manière générale, la répartition en M1 et M2 des syllabes initiales consonantiques par patrons, en adéquation avec les résultats relevés page 178, peine à attester de l'existence d'une corrélation entre les mots débutant par une initiale consonantique et l'une des deux positions dans la coordonnée. Nous pouvons néanmoins affirmer que les mots en #CV._

résultats probants.

occupent majoritairement M2.

4.2.2.3. Structure de la syllabe initiale commençant par un glide

Comme nous l'indiquions page 149, s'il apparaît de façon assez solide que les mots en #V_ occupent prioritairement M1, nous ne pouvons rien affirmer concernant les mots débutant par un glide. A cela une raison majeure, seuls neuf mots commencent par un glide dans l'ensemble de nos supports d'études.

Nous proposons à titre informatif, le tableau résumant la répartition de mots commençant par un glide par patron, à savoir pour nos données, #GV_ ci-dessous. Cependant, il nous est complètement impossible d'extraire des résultats fiables. Des recherches ultérieures devront s'attacher à observer la répartition en M1 et en M2 des mots commençant par un glide.

TOTAL 1331	Vale#						
	i	e/	A/	y	0/	u	/O/
M1	21 4,67%	56 4,20%	34 4,23%	1 5,00%	16 4,00%	3 4,00%	107 4,64%
M2	24 5,33%	60 5,22%	35 5,77%	1 5,00%	24 6,00%	2 4,00%	12 5,14%
total:	45 100,0%	16 100,0%	59 100,0%	2 100,0%	40 100,0%	5 100,0%	119 100,0%

Tableau 57 : Répartition par patrons des syllabes initiales commençant par un glide pour l'ensemble des supports d'études

4.3. Fin de mots

4.3.1. Nature du segment final

Dans cette section, nous nous proposons d'examiner le critère définitoire soulevé par la question : quelle est la nature du segment final de M1 et M2 ? Les finales vocaliques favorisent-elles l'une des deux positions ? Et qu'en est-il des mots possédant une finale consonantique et de ceux se terminant par une semi-consonne ?

Afin d'étudier ces questions, nous avons observé les finales de M1 puis de M2 de différents supports d'études en excluant les formes liaisonnantes que nous traiterons à part en 4.3.3. Les phénomènes de liaison page 274. Nous présentons ci-après les différents supports d'études retenus en fonction des hypothèses éprouvées.

Ainsi, pour collecter les données concernant les voyelles finales, nous avons relevé l'ensemble de nos supports d'études à l'exception du questionnaire d'enquête PAUL ET PASCAL. En effet, ce dernier cherchant à faire émerger le rôle du nombre de syllabes dans la sélection d'un ordre préférentiel des composés syndétiques par « et », les appariements proposés étaient neutralisés sur la question de la finale. Au total, pour examiner l'hypothèse d'une corrélation liant $_V\#$ à l'une ou l'autre des deux positions de la coordonnée, nous avons recueilli 1 962 mots.

De la même façon, pour collecter les données concernant les consonnes finales, nous avons observé l'ensemble des six corpus et les questionnaires d'enquête ÉLISE ET ZOÉ, ÉDOUARD ET EDMOND et NAROD ET NANOR interrogeant tous trois la corrélation entre les finales consonantiques et la sélection d'une des deux positions dans la coordination. Nous prêterons une attention particulière aux résultats fournis par le corpus ÉDOUARD ET EDMOND testant spécifiquement l'opposition entre finales vocaliques et finales consonantiques. A nouveau, nous avons écarté les trois derniers questionnaires d'enquête car ils n'étaient pas interrogés sur l'hypothèse d'un segment final consonantique actif dans la sélection d'un ordre préférentiel de coordination. Ce facteur avait été neutralisé en amont. Au total, 1 580 mots ont été testés sur l'hypothèse corrélant $_C\#$ à l'une des deux positions de la coordonnée.

Enfin, pour collecter les données concernant les mots ayant pour finale un glide, nous avons étudié l'ensemble des six corpus. Les questionnaires d'enquête n'interrogeant pas cette hypothèse, ils étaient tous neutralisés en amont. Au total, 46 mots en $_G\#$ ont été examinés afin de déterminer l'existence d'un rôle actif des finales $_G\#$ dans la sélection de M1 ou de

M2 dans la coordonnée.

Ensuite, dans un premier temps, nous avons observé les résultats de façon générale par type de segment final pour l'ensemble des supports d'études. Dans une deuxième étape, nous nous attacherons spécifiquement à l'étude des voyelles finales en proposant un relevé par phonèmes de leur répartition en M1 et en M2, puis un examen par traits distinctifs. La troisième étape nous amènera à explorer les finales consonantiques, par phonèmes puis par traits. Dans une quatrième étape, nous étudierons la répartition des mots se terminant par un glide en M1 et en M2. Enfin, nous examinerons la répartition par traits de l'ensemble des segments finaux précédemment détaillés avant d'examiner les patrons des syllabes finales.

Nous nous proposons donc d'examiner tout d'abord la répartition en M1 et M2 par type de finales. Comme nous l'indiquions ci-dessus, les supports d'études examinés dépendent des hypothèses testées. C'est pourquoi il n'y a pas de correspondance du nombre de paires compilées questionnant les hypothèses d'une position de la coordonnée privilégiée pour les mots en *_V#*, en *_C#*, et en *_G#*. Cependant, étant donné que nous présentons chaque hypothèse indépendamment les unes des autres en fonction de leur répartition en M1 et en M2 et non en fonction de leur rapport avec les nombres de paires testées, cela n'a aucune incidence sur nos résultats.

Ainsi, dans les deux tableaux à venir, nous observons la répartition en M1 et en M2 des mots en *_V#*, en *_C#*, et en *_G#*. Le premier tableau nommé TOTAL présente cette répartition pour l'ensemble des paires examinées sur ces questions dans nos supports d'études. Nous confronterons les résultats de TOTAL à ceux que nous obtenons sans les données fournis par le corpus TOM ET ZOÉ dans le second tableau nommé TOTAL-TOM ET ZOÉ. En effet, le corpus TOM ET ZOÉ est moins contrôlé que les autres supports d'études. Il est également le plus important en nombre de paires. C'est pourquoi, par mesure de précaution, afin de nous prémunir du biais possible d'autres facteurs, nous nous sommes laissé la liberté d'observer nos résultats indépendamment de ceux recueillis dans ce corpus.

Pour faciliter la lecture de nos résultats, nous avons noté en caractères gras les résultats pour lesquels la valeur-p est inférieure à .05, c'est-à-dire ceux pour lesquels nous relevons une répartition statistiquement significative.

TOTAL	_V#	_C#	_G#
M1	0 46,53%	869 55,00%	12 26,09%
M2	0 53,47%	711 45,00%	34 73,91%
Total :	0 100,00%	1580 100,00%	46 100,00%

Tableau 58 : Répartition des segments finaux pour TOTAL

TOTAL-TOM ET ZOÉ	_V#	_C#	_G#
M1	120 40,82%	251 56,53%	3 30,00%
M2	174 59,18%	193 43,47%	7 70,00%
Total :	294 100,00%	444 100,00%	10 100,00%

Tableau 59 : Répartition des segments finaux pour TOTAL-TOM ET ZOÉ

Dans le détail de la répartition par segment entre M1 et M2, au regard des éléments présentés dans les deux tableaux ci-dessus, nous pouvons affirmer l'existence de deux très fortes présomptions contre l'hypothèse nulle.

En effet, il apparaît que les mots possédant une finale consonantique sélectionnent prioritairement M1 dans 55 % des mots en _C# dans TOTAL et 56,53% des cas dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 15,8$; $p \leq 0,0001$ et $\chi^2(1) = 7,58$; $p \leq 0,006$).

Inversement, les mots en _V# ont statistiquement plus tendance à occuper M2. Ils sont 53,47% à sélectionner cette position dans TOTAL et 59,18% des occurrences dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 9,43$; $p \leq 0,002$ et $\chi^2(1) = 9,92$; $p \leq 0,002$).

Concernant les résultats relevés pour les glides, même si la validation statistique n'est pas aussi forte que nous le souhaiterions, la répartition dans TOTAL et TOTAL-TOM ET ZOÉ est homogène. Les mots en _G# privilégient massivement M2 : ils sont 73,91% à favoriser cette position dans TOTAL, et 70% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 10,52$; $p \leq 0,001$ et $\chi^2(1) = 1,6$; $p \leq 0,21$). Exception faite de leur validation statistique, la répartition

des mots en _G# est très homogène. TOTAL présente une répartition significative, ce qui n'est pas le cas de TOTAL-TOM ET ZOÉ. Or, nous sommes convaincue que la faiblesse statistique du résultat relevé dans TOTAL-TOM ET ZOÉ est directement imputable au faible nombre d'occurrences testées (seulement dix mots dans TOTAL-TOM ET ZOÉ pour 46 dans TOTAL). Nous avançons cependant ce résultat avec précaution et reviendrons en détail sur cette hypothèse page 243.

Les diagrammes ci-dessous illustrent les résultats statistiquement significatifs relevés dans TOTAL-TOM ET ZOÉ. Il s'agit de la répartition des mots en fonction de la qualité de leur segment final en M1 et en M2 ; à savoir la répartition des mots en M1 et en M2 par finale consonantique et par finale vocalique. Nourrissant des hésitations quant à la validité statistique des résultats relevés pour les mots se terminant par un glide, nous ne les présenterons pas dans les diagrammes ci-dessous.

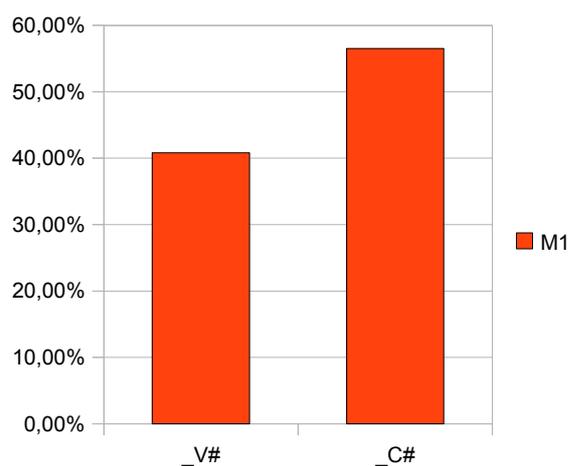


Diagramme 3a : Pourcentage d'occupation de M1 par segment final

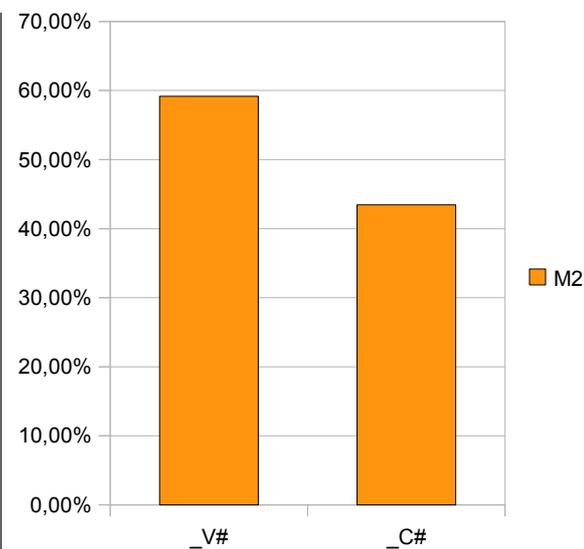


Diagramme 3b : Pourcentage d'occupation de M2 par segment final

En observant ces deux diagrammes, nous remarquons que les mots possédant une coda consonantique comme segment final semblent favoriser la première position dans la coordination. Inversement, les mots n'en possédant pas, occupent majoritairement M2.

Au regard des résultats que nous venons de présenter, il apparaît que le fait que M se termine par une consonne ou une voyelle est un facteur actif dans la sélection d'un ordre préférentiel de coordination. Nous supposons que cela relève en réalité de structures syllabiques. Nous y

reviendrons ultérieurement page 261. Mais avant, nous nous proposons tout d'abord d'examiner si la qualité de la voyelle ou de la consonne finale joue un rôle dans le choix de l'ordre d'une séquence binomiale coordonnée.

4.3.1.1. Voyelle

L'intérêt que nous portons à la répartition des finales vocaliques est double. D'une part, nous pensons qu'il est pertinent de rapprocher leurs résultats à ceux que nous présentions en 4.2.1.1.3. Bilan et analyse des résultats concernant #V_ pour les mots possédant une initiale vocalique. En effet, si les mots en _V# favorisent majoritairement M2, ceux en #V_ ont un comportement exactement opposé, dans des proportions similaires, en sélectionnant prioritairement M1.

Pour schématiser, dans des binômes où l'un des deux termes possède une frontière vocalique, « #V_ & _ » et « _ & _V# » sont majoritairement sélectionnés. Or, il apparaît que ce que ces deux schémas ont en commun, c'est l'évitement d'une séquence VV qui pourrait potentiellement se former avec la conjonction « et ». Cet évitement de l'hiatus est un phénomène récurrent dans la littérature spécialisée. En effet, comme le rappelle Tranel (2000) en présentant quelques caractéristiques de la Théorie de l'Optimalité, la séquence VV est manifestement défavorisée « dans les inventaires séquentiels asymétriques attestés dans les langues du monde. » Pour cette raison, comme le suppose Tranel, même si « une explication plus fondamentale est sans aucun doute à la source de cette généralisation », il existe une contrainte universelle qui bannit les séquences *VV.

Ainsi, dans un cadre OT, il apparaîtrait que les formes « #V_ & _ » et « _ & _V# » pourraient être des effets de la contrainte universelle *VV. Nous reviendrons ultérieurement sur cette analyse au regard des éléments observés dans la section à venir par l'examen précis de la répartition en M1 et M2 des mots possédant une voyelle finale.

Pour ce faire, nous présenterons tout d'abord les finales vocaliques en relevant la répartition de leur phonème en M1 et en M2 pour les voyelles orales, puis pour les voyelles nasales, par supports d'études, et pour le total des paires possédant une finale vocalique. Enfin, nous proposerons une organisation de ces phonèmes relevés en fonction de leurs traits distinctifs.

4.3.1.1.1. Répartition phonémique des finales vocaliques

Nous nous proposons d'explorer la répartition par phonème des mots finissant par une voyelle orale pour chaque support d'études examinant cette hypothèse, à savoir les six corpus et les

questionnaires d'enquête, exception faite du questionnaire PAUL ET PASCAL n'interrogeant pas cette hypothèse - il était neutralisé en amont. Les résultats sont en annexe du tableau A 8a au tableau A 8k.

Le premier constat que nous formulons de l'examen phonémique des mots possédant une finale vocalique orale, est que même si les mots se terminant par une voyelle orale sont nombreux dans nos supports d'études (1 351 au total), il nous est néanmoins difficile d'obtenir des résultats statistiquement solides de leur observation par support d'études.

Dans le détail, le résultat le plus lisible de l'examen de nos supports d'études observés indépendamment les uns des autres est l'absence de résultats. Les mots possédant une finale vocalique orale ne semblent pas privilégier de position particulière dans la coordination, exception faite d'un résultat relevé dans le corpus TOM ET ZOÉ. En effet, ce dernier fait émerger une tendance statistique ($\chi^2(1) = 2,77$; $p \leq 0,09$) : les mots en $_/i/\#$ sont 54,47% à sélectionner majoritairement M1. Il nous faudra confirmer ce résultat ultérieurement.

Attardons-nous à présent sur les résultats émergeant des questionnaires d'enquête. Nous y relevons quelques données pour lesquelles la valeur-p est inférieure à .1. Mais compte tenu de la très petite taille des échantillons présentés (moins de treize mots par hypothèse), il serait hasardeux d'avancer une validation statistique.

Notons cependant deux points essentiels.

D'une part, trois de ces questionnaires (à savoir NAROD ET NANOR, LEMOTO ET LEMOTAN et ÉDOUARD ET EDMOND) ont été spécifiquement mis en place pour étudier les hypothèses corrélant le type de segment final à une position privilégiée dans la coordination. Comme nous le présentons en 2.2. Protocole de constitution des questionnaires d'enquête, les appariements nous permettant de sélectionner les paires canoniques de coordination ne sont éprouvés que sur une seule hypothèse à la fois. Tout autre facteur est neutralisé.

D'autre part, rappelons ce que nous indiquions page 71. Les paires composant les questionnaires d'enquête ont été sélectionnées par une double validation statistique. Dans un premier temps, pour constituer nos supports d'études, nous retenons les paires présentant l'ordre de mots préférentiel des locuteurs interrogés. Puis, une fois nos supports d'études constitués, nous observons l'ensemble de ces paires en proposant un nouveau traitement statistique pour chaque hypothèse examinée. Or, si du fait de la taille des échantillons testés, les résultats des questionnaires d'enquête que nous présentons ci-dessous ne peuvent prétendre à une validation statistique pour la seconde étape, chaque paire retenue a néanmoins

été soumise et sélectionnée de façon restrictive ($p \leq 0,05$), par une cinquantaine de locuteurs natifs du français. C'est pourquoi, nous pensons qu'il est pertinent de commenter leurs résultats.

Ainsi, dans le détail, par questionnaire d'enquête, nous relevons les résultats suivants.

Le questionnaire d'enquête NAROD ET NANOR, présenté page 63, est le questionnaire que nous avons mis en place pour explorer l'hypothèse établissant une relation entre le type de phonèmes en finale de mot (voyelles et consonnes confondues) et l'assignation d'une position dans la coordonnée. Ce support d'études présente des résultats massifs. L'ensemble des mots possédant une finale vocalique, à savoir pour questionnaire d'enquête, les mots se terminant en $_ /e/\#$, en $_ /a/\#$ et en $_ /o/\#$, sélectionnent systématiquement M2. Il est évident que le nombre de types testés (trois paires pour les mots en $_ /e/\#$, deux pour les finales en $/o/$ et une seule pour les mots en $_ /a/\#$) nous interdit toute validation statistique. Rappelons néanmoins que cet ordre préférentiel plaçant ces six mots ayant une finale vocalique en M2 a été sélectionné par 53 locuteurs natifs du français de façon massive et statistique. Dans une sélection proposant « x & y », « y & x » et « NSP », environ 70% des locuteurs interviewés ont positionné les six paires traitant du type de finales vocaliques en M2.⁶⁷ Au regard de cette considération, nous pensons que les résultats recueillis dans les questionnaires d'enquête sont pertinents et nécessitent notre attention particulière. Ainsi, dans le questionnaire d'enquête NAROD ET NANOR, les mots possédant une finale vocalique en $/e/$, $/a/$ et $/o/$ occupent M2. Notons également que les mots en $_ /e/\#$ sélectionnent également M2 dans les mêmes proportions dans le questionnaire d'enquête ÉLISE ET ZOÉ.

De la même façon, le questionnaire d'enquête ÉDOUARD ET EDMOND examine précisément, comme nous le signalions en 2.2.6. Protocole de constitution du questionnaire d'enquête ÉDOUARD ET EDMOND, les segments finaux en s'attardant spécifiquement sur les binômes opposant des finales [+ CONS] à des finales [- CONS].

67 Dans le détail :

- La forme « NANOR ET NANÉ » a été sélectionnée par 73,67% des locuteurs interviewés ($\chi^2 (2) = 44,19$; $p \leq 0,0000000002$)
- « ARPÉNIG ET ARPINÉE » : 66% ($\chi^2 (2) = 34,68$; $p \leq 0,00000002$),
- « ARMÉNOUCHE ET ARMONÉE » : 75,5% ($\chi^2 (2) = 45,5$; $p \leq 0,0000000001$)
- « MARAL ET MARO » : 66% ($\chi^2 (2) = 34,68$; $p \leq 0,0000002$),
- « MANNIG ET MARO » : 67,9% ($\chi^2 (2) = 36,72$; $p \leq 0,0000001$),
- et « SÉVANE ET SÉDA » : 75,5% ($\chi^2 (2) = 47,13$; $p \leq 0,0000000001$)

Ainsi, nous observons de l'opposition binaire des mots à finale vocalique aux mots à finale consonantique que les mots en $_V\#$ occupent systématiquement, dans 100% des paires canoniques retenues, M2. Par phonèmes, dans le questionnaire d'enquête ÉDOUARD ET EDMOND, les mots en $_/a/\#$, $_/o/\#$ et $_/i/\#$ privilégient M2.

Enfin, le questionnaire d'enquête LEMOTO ET LEMOTAN étudie à l'aide de pseudo-prénoms (cf. page 65) la sélection d'une forme privilégiée de coordination en fonction du type de finale vocalique des binômes. Les 25 binômes testés proposaient d'observer les oppositions binaires des finales vocaliques par traits distinctifs. De fait, seules cinq paires ont été retenues.⁶⁸ Trois d'entre elles permettent d'examiner l'opposition [+ NASAL] / [- NASAL]. Nous les commenterons plus tard. Les deux dernières traitent des oppositions [+ HAUT] / [- HAUT] et [+ ARRONDI]/[- ARRONDI]. C'est pourquoi, en mettant en relation deux mots possédant une finale vocalique, il est normal (voire tautologique) d'observer des mots à finale vocalique dans les deux positions. Notons cependant deux résultats intéressants : en adéquation avec les résultats relevés dans NAROD ET NANOR, et ÉDOUARD ET EDMOND (mais également dans ÉLISE ET ZOÉ) dans le questionnaire d'enquête LEMOTO ET LEMOTAN, les mots en $_/e/\#$ et les mots en $_/a/\#$, confrontés à d'autres mots possédant une finale vocalique, sélectionnent systématiquement M2.

Du recueil par support d'études, concernant l'existence d'une corrélation entre la sélection d'une position dans la coordonnée et les phonèmes de voyelles orales finales, nous relevons :

- Une tendance statistique : les mots en $_/i/\#$ ont tendance à sélectionner M2 ;
- Des pistes de réflexion : les questionnaires d'enquête suggèrent que dans des conditions d'expérimentation dans lesquels tout facteur externe a été neutralisé, les mots en $_/e/\#$, en $_/a/\#$ et en $_/o/\#$ se placent en M2 ;
- Et surtout une absence de résultats notable.

En raison de ce dernier point et de notre volonté d'asseoir plus fortement les résultats et pistes de réflexion précédents, nous avons choisi d'observer les finales vocaliques orales pour

68 Les paires retenues pour le questionnaire d'enquête LEMOTO ET LEMOTAN sont : « CHIEKO ET CHIEKA » (63,6%, $\chi^2(2) = 45,91$; $p \leq 0,0000000001$), « DANZO ET DANZON » (59,3% $\chi^2(2) = 38,81$; $p \leq 0,000000003$), « HANAHI ET HANAHE » (66,7% $\chi^2(2) = 54,07$; $p \leq 0,00000000002$), « HAROUTO ET HAROUTAN » (59,3% $\chi^2(2) = 42,86$; $p \leq 0,0000000004$) et « LEMOTO ET LEMOTAN » (59,3% $\chi^2(2) = 38,81$; $p \leq 0,000000003$).

l'ensemble des supports d'études.

Pour cela, souhaitant que le regroupement de données nous permette de faire émerger des résultats statistiquement solides, nous avons produit deux tableaux observant la répartition en M1 et en M2 par phonèmes des mots en V[- NASAL]# : le premier présentant les 179 mots possédant une finale vocalique orale de l'ensemble des supports d'études sans ceux issus du corpus TOM ET ZOÉ, et le second, le total 1 351 termes se terminant par une finale vocalique orale.

Comme précédemment, nous avons préféré proposer une confrontation de ces deux résultats afin de pallier les biais possibles portés par le corpus TOM ET ZOÉ. De l'examen de ces deux nouvelles répartitions, nous nous attarderons spécifiquement sur la place privilégiée des mots en /i/# que TOM ET ZOÉ positionne majoritairement en M1, et celles en M2 des mots en /e/#, en /a/# et en /o/# repérées de l'examen attentif des questionnaires d'enquête.

TOTAL-TOM ET ZOÉ 179	_Vocale #										
	i	e	ɛ	a	ɑ	y	ø	œ	u	o	ɔ
M1	21 46,67%	21 42,86%	2 40,00%	15 38,46%	0 0,00%	1 50,00%	2 50,00%	0 0,00%	3 60,00%	13 46,43%	
M2	24 53,33%	28 57,14%	3 60,00%	24 61,54%	1 100,00%	1 50,00%	2 50,00%	1 100,00%	2 40,00%	15 53,57%	
Total :	45 100,00%	49 100,00%	5 100,00%	39 100,00%	1 100,00%	2 100,00%	4 100,00%	1 100,00%	5 100,00%	28 100,00%	

Tableau 60 : Répartition par phonèmes des finales vocaliques orales pour l'ensemble des supports d'études sans TOM ET ZOÉ

TOTAL 1351	_Vocale #										
	i	e	ɛ	a	ɑ	y	ø	œ	u	o	ɔ
M1	210 53,57%	54 48,65%	2 40,00%	264 47,31%	0 0,00%	1 50,00%	16 41,03%	0 0,00%	11 47,83%	107 48,86%	
M2	182 46,43%	57 51,35%	3 60,00%	294 52,69%	1 100,00%	1 50,00%	23 58,97%	1 100,00%	12 52,17%	112 51,14%	
Total :	392 100,00%	111 100,00%	5 100,00%	558 100,00%	1 100,00%	2 100,00%	39 100,00%	1 100,00%	23 100,00%	219 100,00%	

Tableau 61 : Répartition par phonèmes des finales vocaliques orales pour l'ensemble des supports d'études

Au regard des données présentées dans les deux tableaux ci-dessus, nous ne relevons aucun résultat. Le résultat plaçant les mots en /i/# majoritairement en M1 n'est visible que dans le corpus TOM ET ZOÉ ; il nous est impossible de confirmer cette hypothèse. Quant aux pistes relevées dans les questionnaires d'enquête, même si intuitivement en tant que locutrice native du français nous sommes convaincue de leur intérêt, nous ne pouvons les affirmer.

Face à cette absence de résultats, nous proposons une nouvelle organisation des phonèmes en neutralisant le degré d'aperture des voyelles moyennes, à savoir les paires /e/~ε/, /ø/~œ/ et /ɔ/~o/. Nous avons également neutralisé le degré d'antériorité de la paire /ɑ/~a/ étant donné que cette opposition est rarement réalisée (nous ne relevons qu'une réalisation unique de /ɑ/# pour l'ensemble de nos supports d'études). Pour ce faire, et compte tenu de la loi de distribution complémentaire des voyelles moyennes en syllabe accentuée, nous avons utilisé les archiphonèmes /E/, /ɛ/, /O/ et /A/.

Cependant, comme le montrent à nouveau les tableaux ci-dessous, cela ne permet pas de faire émerger l'assignation d'une place privilégiée en fonction du type phonémique de la voyelle orale finale, exception faite d'un résultat concernant les mots en /A/# qu'il nous faut commenter.

TOTAL-TOM ET ZOÉ 179	Vocale #						
	i	/E/	/A/	y	/ɛ/	u	/O/
M1	21 46,67%	23 42,59%	15 37,50%	1 50,00%	2 40,00%	3 60,00%	13 46,43%
M2	24 53,33%	31 57,41%	25 62,50%	1 50,00%	3 60,00%	2 40,00%	15 53,57%
Total :	45 100,00%	54 100,00%	40 100,00%	2 100,00%	5 100,00%	5 100,00%	28 100,00%

Tableau 62 : Répartition des archiphonèmes des finales vocaliques orales pour l'ensemble des supports d'études sans TOM ET ZOÉ

TOTAL 1351	Vocale #						
	i	/E/	/A/	y	/ɛ/	u	/O/
M1	21 46,67%	56 48,28%	264 47,23%	1 50,00%	16 40,00%	3 60,00%	107 48,86%
M2	24 53,33%	60 51,72%	295 52,77%	1 50,00%	24 60,00%	2 40,00%	112 51,14%
Total :	45 100,00%	116 100,00%	559 100,00%	2 100,00%	40 100,00%	5 100,00%	219 100,00%

Tableau 63 : Répartition des archiphonèmes des finales vocaliques orales pour l'ensemble des supports d'études

La confrontation de ces deux tableaux fait apparaître un seul résultat : dans TOTAL-TOM ET ZOÉ, les mots en $_ /A/\#$ sont 62,5% à sélectionner M2. Bien que la validation statistique ne soit pas aussi forte que nous le souhaiterions ($\chi^2 (1) = 2,5 ; p \leq 0,11$), nous avançons néanmoins ce résultat pour deux raisons.

Tout d'abord, la valeur-p que nous présentons ici est très légèrement supérieure à .01. Or ce seuil de .01 est pris pour référence afin de définir une présomption plus ou moins forte contre l'hypothèse nulle. Cette présomption est en faveur de l'hypothèse éprouvée (quand $p < 0,1$), ou en sa défaveur (quand $p > 0,1$). Cependant, ne pas rejeter l'hypothèse nulle ne signifie pas pour autant que l'on rejette nécessairement l'hypothèse testée. Cela signifie seulement que l'on ne peut rien conclure à son propos : ni l'infirmier, ni la confirmer. La valeur de p est graduelle, on rejette plus ou moins l'hypothèse nulle. Or la valeur-p de 0,11 montre bien que si l'on ne peut pas rejeter l'hypothèse nulle, son « non-rejet » est beaucoup moins fort que celui que nous rencontrons pour la distribution des autres finales vocaliques orales. Rappelons également que ce résultat plaçant majoritairement en M2 les mots en $_ /A/\#$ est en adéquation avec l'ensemble des résultats que nous retenons de l'observation des questionnaires d'enquête ÉDOUARD ET EDMOND, LEMOTO ET LEMOTAN et NAROD ET NANOR. C'est pourquoi qu'il existe une corrélation entre les mots en $_ /A/\#$ et M2.

En fait, au regard des éléments interrogeant l'existence d'une corrélation entre l'assignation d'une position privilégié et les phonèmes de finales vocaliques orales, nous n'affirmons qu'un seul résultat : les mots se terminant par $_ /A/\#$ occupent majoritairement M2.

Gardons également à l'esprit les résultats obtenus dans les questionnaires d'enquête : dans des coordinations binomiales neutralisées pour lesquels nous interrogeons spécifiquement les relations entre les positions de la coordonnée et le type phonémique de la finale, les mots en $_ /e/\#$ et en $_ /o/\#$ sélectionnent massivement M2.

Comment analyser ses résultats ?

Le principe d'évitement de l'hiatus que nous évoquions en début de section semble traiter efficacement ces phénomènes : les mots se terminant par $_ /A/\#$, $_ /e/\#$ et en $_ /o/\#$ occupent M2, ou évitent M1 afin de ne pas violer la contrainte *VV. De plus, il nous est possible de rapprocher le résultat des mots en $_ /e/\#$ privilégiant M2 de ceux que nous commentons dans la section 4.2.1.1.3. Bilan et analyse des résultats concernant #V_ pour les mots en #/e/_ en M1. Dans les deux cas, nous y voyons une tendance à éviter la séquence /e e/.

Néanmoins, la question que nous nous posons est : pourquoi seules trois voyelles orales finales suivent ce contour ? Que partagent-elles en commun ? Qu'ont-elles de spécifique les distinguant des autres voyelles finales orales pour lesquelles nous n'avons pu extraire de résultats tangibles ?

La répartition par traits distinctifs de ces trois voyelles devrait répondre à ces questions. Nous reviendrons dessus après avoir étudié la dispersion par phonème en M1 et en M2 des mots se terminant par une voyelle nasale, lors de l'examen par traits distinctifs que nous mènerons pour l'ensemble des mots finissant par une finale vocalique.

Nous examinons à présent la répartition par phonème des mots finissant par une voyelle nasale. Pour cela, nous avons utilisé les supports d'études suivants : les six corpus et les questionnaires d'enquête, exceptions faites des questionnaires PAUL ET PASCAL et NAROD ET NANOR. En effet, ces deux questionnaires explorent l'existence de principes actifs liés respectivement à la longueur syllabique et au type de finales consonantique dans l'assignation d'une position préférentielle. Le rôle des finales dans la sélection d'un ordre de coordination était donc neutralisé en amont. Les résultats sont présentés les uns indépendamment des autres en annexe du tableau A 9a au tableau A 9j.

De la répartition en M1 et en M2 des mots finissant par une voyelle nasale par support d'études, nous notons trois résultats statistiquement significatifs ($p \leq 0,05$) et une tendance statistique ($p \leq 0,1$).

Dans le détail, par support d'études, nous relevons les résultats suivants.

Le corpus TARN ET GARONNE fait apparaître une tendance statistique : les mots en $_/̃/#$ occupent systématiquement M2 ($\chi^2(1) = 3,8$; $p \leq 0,08$). Ce résultat peut sembler solide, nous ne le retenons cependant pas. En effet, l'échantillon testé est bien trop petit (trois mots) pour pouvoir asseoir un quelconque résultat.

Dans le corpus TOM ET ZOÉ, nous relevons trois résultats : les mots en $_/̃/#$ sélectionnent M1 dans 65,14% des cas ($\chi^2(1) = 9,99$; $p \leq 0,001$) alors qu'inversement, les mots en $_/̃̃/#$ et $_/̃̃̃/#$ semblent favoriser plutôt M2 dans respectivement 66,41% et 64,45% des occurrences testant ces deux hypothèses ($\chi^2(1) = 15,11$; $p \leq 0,0001$ et $\chi^2(1) = 21,39$; $p \leq 0,000004$).

Enfin, les questionnaires d'enquête, et notamment ÉDOUARD ET EDMOND et LEMOTO ET LEMOTAN qui explorent particulièrement le rôle des finales vocaliques nasales, ne fournissent pas de résultats proprement dits. Étant relevés sur des échantillons de moins de cinq mots, ils ne peuvent prétendre à une validation statistique. Néanmoins, comme ces deux

questionnaires traitent précisément du rôle des finales vocaliques, nous nous devons de les relever : les paires examinées dans ÉDOUARD ET EDMOND opposent des mots ayant une finale vocalique à des mots possédant une finale consonantique, et celles du questionnaire LEMOTO ET LEMOTAN confrontent des finales vocaliques orales à des finales vocaliques nasales. Ainsi, il apparaît que dans les questionnaires d'enquête LEMOTO ET LEMOTAN, et ÉDOUARD ET EDMOND, les mots en $_{/}\tilde{\text{o}}/\#$, tout comme ceux en $_{/}\tilde{\text{a}}/\#$, occupent systématiquement M2.

Les résultats par support d'études que nous venons de relever semblent assez solides. Cependant, comme nous l'indiquions précédemment par mesure de précaution, nous ne conserverons pas de résultats portés par un seul corpus. Raison pour laquelle nous nous proposons d'observer les données de l'ensemble de nos supports d'études questionnant la relation entre position dans la coordonnée et les types phonémiques des finales vocaliques nasales. Nous avons à nouveau choisi de présenter ces résultats en confrontant deux totaux dans deux tableaux distincts. Dans le premier tableau, nous avons sélectionné les 115 mots possédant une finale vocalique nasale dans l'ensemble des supports d'études interrogés sur cette hypothèse sans les résultats du corpus TOM ET ZOÉ. Nous nommerons ce total : TOTAL-TOM ET ZOÉ. De son observation, nous explorerons très précisément la répartition en M1 et en M2 des mots en $_{/}\tilde{\text{o}}/\#$, en $_{/}\tilde{\text{a}}/\#$ et en $_{/}\tilde{\text{e}}/\#$ pour lesquels nous relevions des résultats assez forts dans le corpus TOM ET ZOÉ. Dans le second tableau, nous mettons à disposition l'ensemble des résultats de nos supports d'études questionnant le rôle des finales vocaliques nasales, à savoir 611 mots. Nous nommons ce nouveau total : TOTAL.

Comme précédemment, dans les deux tableaux ci-dessous, nous avons mis en valeur les résultats saillants en notant en caractère gras ceux pour lesquels nous relevions une dispersion significative, c'est-à-dire ceux pour lesquels la valeur-p est inférieure à .05.

TOTAL-TOM ET ZOÉ 115	Vnasale#			
	$\tilde{\text{e}}$	$\tilde{\text{œ}}$	$\tilde{\text{o}}$	$\tilde{\text{a}}$
M1	15 44,12%		15 32,61%	12 34,29%
M2	19 55,88%		31 67,39%	23 65,71%
Total :	34 100,00%		46 100,00%	35 100,00%

Tableau 63 : Répartition par phonèmes des finales vocaliques nasales pour l'ensemble des supports d'études sans TOM ET ZOÉ

TOTAL 611	Vnasale#			
	ē	œ	ō	ā
M1	106 36,55%		86 55,48%	56 33,73%
M2	184 63,45%		70 44,52%	110 66,27%
Total :	290 100,00%	0 100,00%	156 100,00%	166 100,00%

Tableau 64 : Répartition par phonèmes des finales vocaliques nasales pour l'ensemble des supports d'études

De la confrontation des deux tableaux ci-dessous, nous relevons trois résultats pour lesquels la dispersion est significative.

Dans le détail, revenons sur les résultats que nous retenions de l'observation de la répartition phonémique des mots se terminant par une voyelle nasale par support d'études.

Concernant les mots en $_{/}\tilde{o}/\#$: dans le corpus TOM ET ZOÉ, nous notons que les mots en $_{/}\tilde{o}/\#$ occupaient majoritairement M1. Or, en confrontant les résultats des deux totaux, nous constatons deux résultats différents. Dans TOTAL, les mots en $_{/}\tilde{o}/\#$ sélectionnent effectivement M1 dans 55,48% des cas. Cependant, la distribution n'est pas significative ($\chi^2(1) = 1,86$; $p \leq 0,17$). Inversement, TOTAL-TOM ET ZOÉ fait émerger un résultat significatif ($\chi^2(1) = 5,57$ $p \leq 0,016$) dans lequel les mots se terminant par $_{/}\tilde{o}/\#$ se placent très majoritairement en M2, dans 67,39% des cas.

Comment comprendre et expliquer une telle différence ?

Nous sommes convaincue que les résultats relevés dans TOTAL sont induits par ceux du corpus TOM ET ZOÉ. En effet, pour l'ensemble de nos supports d'études, 155 mots se terminent par $_{/}\tilde{o}/\#$. Or sur, ces 155 mots, 109 proviennent du corpus TOM ET ZOÉ. Nous avons donc examiné en détail TOM ET ZOÉ en collectant les 78 mots en $_{/}\tilde{o}/\#$ sélectionnant M1, et les 38 qui préféraient M2. Or pour l'ensemble de ces 109 réalisations, nous n'avons en fait recueilli que trois types (en M1 et en M2) : « SIMON », « MARION » et « MANON ». Le fait que cette hypothèse ne soit effectivement testée que sur trois mots est une raison supplémentaire pour ne pas conserver les résultats du corpus TOM ET ZOÉ.

Ainsi, nous affirmons donc, conformément à ce que nous relevions dans TOTAL-TOM ET ZOÉ, que les mots en $_{/}\tilde{o}/\#$ occupent majoritairement M2.

Au sujet des mots se terminant par la voyelle nasale / $\tilde{\epsilon}$ /, les résultats observés dans les deux tableaux ont la même répartition : dans TOTAL, les mots en $_/\tilde{\epsilon}/\#$ sélectionnent majoritairement M2 dans 63,45% des 290 mots testés. Ils sont 55,88% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 20,98$ $p \leq 0,000005$ et $\chi^2(1) = 0,47$ $p \leq 0,49$). Les résultats sont très proches. Leur seule différence réellement notable est leur différence de validation statistique. Nous pourrions être tentée de valider ce résultat. Cependant, sur les 290 mots testés dans TOTAL, 256 proviennent de TOM ET ZOÉ.

Nous avons vérifié dans le détail les mots en $_/\tilde{\epsilon}/\#$ présents dans le corpus TOM ET ZOÉ. A nouveau, sur 290 tokens, seuls neuf types sont réellement testés en M1 et en M2 (ADRIEN, BASTIEN, BENJAMIN, CORENTIN, DAMIEN, JULIEN, QUENTIN, ROMAIN et VALENTIN). Il apparaît donc, que non seulement le résultat plaçant les mots en $_/\tilde{\epsilon}/\#$ majoritairement en M2 n'est porté que par un seul corpus, mais que de surcroît, il n'est comptabilisé pour l'immense majorité des appariements que sur neuf prénoms. Pour ces deux raisons, nous ne conserverons donc pas ce résultat.

Enfin, concernant les mots en $_/\tilde{\alpha}/\#$, la confrontation des deux tableaux fait émerger deux résultats extrêmement similaires : les mots en $_/\tilde{\alpha}/\#$ occupent très majoritairement M2 dans 66,27% des cas dans TOTAL, pour 65,71% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 17,57$ $p \leq 0,00002$ et $\chi^2(1) = 3,46$ $p \leq 0,06$). Compte tenu que ce résultat apparaît dans l'ensemble de nos données, comptabilisées avec et sans le corpus TOM ET ZOÉ, nous affirmons donc que les mots en $_/\tilde{\alpha}/\#$ sélectionnent prioritairement M2.

Au regard de l'examen de la répartition en M1 et en M2 des types phonémiques des finales vocaliques nasales, nous affirmons que les mots en $_/\tilde{\omega}/\#$, tout comme les mots en $_/\tilde{\alpha}/\#$, occupent généralement M2.

Que penser de ces résultats ?

A nouveau, nous pensons que la contrainte *VV est active dans la sélection d'un ordre préférentiel de coordination. De plus, ces derniers résultats sont à rapprocher de ceux que nous relevions pour les voyelles finales orales. En effet, à condition d'accepter la neutralisation des voyelles moyennes et de /A/, nous constatons que les voyelles orales et leurs voyelles nasales correspondantes, en finales de mots, occupent les mêmes positions dans la coordonnée. Ainsi, exception faite pour les finales en $_/\tilde{e}/\#$ pour lesquelles nous

soupçonnons un comportement particulier eu égard à la présence de la conjonction de coordination « et », /A/~/*ã*/ et /O/~/*õ*/ en finale sélectionnent M2.

La question que nous nous posons à présent est de savoir pourquoi les mots en *_/A/#*, en *_/e/#* et en *_/o/#*, tout comme ceux en *_/*ã*/#* et en *_/*õ*/#* occupent plutôt M2. En plus du principe d'évitement de l'hiatus, que partagent-elles en commun, qu'est-ce qui les pousse en M2 ?

C'est pourquoi, nous nous proposons d'observer la répartition des finales vocaliques orales et nasales par traits binaires distinctifs en commençant par présenter cet examen pour chaque support d'études, avant de commenter la confrontation de ces données pour l'ensemble des mots examinant le rôle des finales vocaliques.

4.3.1.1.2. Répartition par traits distinctifs des finales vocaliques

Les résultats de la répartition des traits distinctifs des finales vocaliques par support d'études sont accessibles en annexe du tableau A 10a au tableau A 10k.

Dans le détail, dans les six corpus et le questionnaire d'enquête ÉLISE ET ZOÉ, nous relevons neuf résultats statistiquement significatifs notés en caractères gras et cinq tendances statistiques mentionnées en italique. Notons dès à présent que tous les résultats que nous présentons mettent en avant une corrélation entre un trait distinctif et la position M2. Nous examinerons les résultats des questionnaires d'enquête ci-après. En effet, les résultats que nous y observons étant calculés sur des échantillons de moins de dix mots, ils ne peuvent être validés statistiquement. Nous estimons néanmoins que pour les motifs que nous exposons page 197, nous nous devons de les relever et de les commenter.

Les corpus CUL ET CHEMISE, LISSE ET SOYEUX, et VITE ET BIEN, bien que mesurés sur des échantillons suffisants (de respectivement 30, 39 et 87 mots), ne fournissent pas de résultats.

Dans le corpus TARN ET GARONNE, nous relevons quatre résultats et une tendance statistique liant un trait binaire de la finale vocalique à M2. Par ordre décroissant : 76% des mots se terminant par une voyelle en [- HAUT] et en [+ BAS] occupent M2 (respectivement $\chi^2(1) = 6,08$; $p \leq 0,01$) et $\chi^2(1) = 6,76$; $p \leq 0,01$), tout comme 75% des mots *_[+ ARRIÈRE]#* et *_[+ NASAL]#* (respectivement $\chi^2(1) = 6$ $p \leq 0,01$ et $\chi^2(1) = 5$ $p \leq 0,03$). Enfin, 65,63% des 32 mots ayant pour finale vocalique une voyelle [-ARRONDI] ont

tendance à favoriser M2 ($\chi^2(1) = 3,13$; $p \leq 0,08$).

L'examen du corpus TOM ET ZOÉ fait également émerger quatre résultats significatifs et une tendance. Du résultat le plus saillant au moins saillant : 58,47% des mots se terminant par une voyelle [+ NASAL] occupent majoritairement M2 ($\chi^2(1) = 14,23$; $p \leq 0,0002$), tout comme les mots _[+ BAS]# dans 55,17% des 1 015 occurrences testées ($\chi^2(1) = 10,86$; $p \leq 0,01$). Les mots _[- HAUT]# sélectionnent la même position dans 54,26% des cas ($\chi^2(1) = 9,46$; $p \leq 0,002$), et 53,92% des mots se terminant par une voyelle [- ARRONDI] ($\chi^2(1) = 8,07$; $p \leq 0,004$). Enfin, 53,29% des 700 mots du corpus TOM ET ZOÉ possédant une finale vocalique [- ARRIÈRE] ont tendance à favoriser M2 ($\chi^2(1) = 3,02$; $p \leq 0,08$). Ces résultats, exceptions faites de ceux relevés pour les voyelles finales [- ARRIÈRE], sont en adéquation avec ceux relevés dans TARN ET GARONNE, même s'ils peuvent sembler moins massifs. Nous pensons, en outre, que cette différence n'est qu'un artefact de la taille des échantillons testés : 43 mots pour TARN ET GARONNE, pour 1 868 dans TOM ET ZOÉ.

Dans le corpus ARTS ET MÉTIERS, en adéquation avec les résultats observés dans le corpus TOM ET ZOÉ pour les mots _[- ARRIÈRE]#, nous relevons une tendance statistique plaçant 68,18% des 22 mots se terminant par une voyelle [- ARRIÈRE] en M2 ($\chi^2(1) = 2,91$; $p \leq 0,09$).

Enfin, dans le questionnaire d'enquête ÉLISE ET ZOÉ, nous dégageons un résultat significatif et trois tendances statistiques. Suivant les contours de ce que nous observons dans les corpus TARN ET GARONNE et TOM ET ZOÉ, les mots _[- ARRONDI]# sélectionnent majoritairement M2 dans 73,68% des 19 mots testés sur cette hypothèse ($\chi^2(1) = 4,26$; $p \leq 0,04$). Nous extrayons également deux tendances statistiques : les mots _[- HAUT]#, comme ceux _[- NASAL]# occupent prioritairement M2 dans 72,22% et 68,18% des cas (respectivement $\chi^2(1) = 3,56$; $p \leq 0,06$ et $\chi^2(1) = 2,91$; $p \leq 0,09$), tout comme 85,71% des sept mots _[- ARRIÈRE]# qui sélectionnent cette même position ($\chi^2(1) = 3,57$; $p \leq 0,06$). Compte tenu de la taille de l'échantillon éprouvant cette dernière hypothèse, bien que suivant le contour de résultats déjà relevés, nous ne pouvons la considérer comme un résultat.

Du recueil des données interrogeant l'assignation d'une position dans la coordonnée en fonction des traits distinctifs de la finale vocalique dans les corpus et le questionnaire d'enquête ÉLISE ET ZOÉ, nous constatons trois résultats récurrents : les mots ayant une

finale vocalique [- HAUT], [+ BAS] et [- ARRONDI] sélectionnent majoritairement M2. Nous relevons également deux résultats apparemment contradictoires pour les mots finissant par une voyelle [\pm ARRIÈRE] et [\pm NASAL]. Ces deux derniers résultats pourraient nous laisser penser que les traits distinctifs des voyelles finales [\pm ARRIÈRE] et [\pm NASAL] ne sont pas des facteurs actifs dans la sélection d'un ordre préférentiel de coordination. Cependant, selon nous, seuls les données pour les mots $_{[\pm \text{ARRIÈRE}]\#}$, fournissent des résultats réellement opposés. En effet, il n'y a rien de contradictoire à ce que les mots finissant par une voyelle [\pm NASAL] occupent préférentiellement M2. Cela signifie simplement que d'une manière générale, les mots se terminant par une voyelle nasale ou orale sélectionneraient plutôt M2.

En ce qui concerne à présent les résultats dégagés des questionnaires d'enquête ; nous ne prétendons pas une validation statistique étant donnée la taille réduite des échantillons testés. Néanmoins, nous pensons que du fait de leurs constitutions, les résultats massifs émergeant des questionnaires d'enquête méritent notre plus grande attention. En effet, les binômes soumis aux interviewés n'interrogeaient qu'une seule hypothèse à la fois. Tout autre facteur était neutralisé. Or, cette neutralisation nous permet de rendre visible des facteurs, ou des nuances, peut être moins évidents que ce que nous relevons des corpus, mais tout aussi pertinents. Nous avons donc retenu des résultats dans ÉDOUARD ET EDMOND, NAROD ET NANOR et LEMOTO ET LEMOTAN.

Rappelons que :

- Le premier questionnaire examine l'opposition finale vocalique / finale consonantique :
[- CONSONANTIQUE] vs [+ CONSONANTIQUE] ;
- Le deuxième, l'opposition finale consonantique / finale vocalique orale ;
[- CONSONANTIQUE] vs. $\left(\begin{array}{c} + \text{ CONSONANTIQUE} \\ - \text{ NASAL} \end{array} \right)$;
- Et le dernier, l'opposition finale vocalique orale / finale vocalique nasale.
 $\left(\begin{array}{c} + \text{ CONSONANTIQUE} \\ - \text{ NASAL} \end{array} \right)$ vs. $\left(\begin{array}{c} + \text{ CONSONANTIQUE} \\ + \text{ NASAL} \end{array} \right)$.

Nous avons relevé les résultats selon deux critères précis. Premièrement, la valeur-p est inférieure à .10. Ce critère n'est qu'un critère de sélection, il ne s'agit en aucun cas d'une validation statistique. L'échantillon est beaucoup trop restreint pour présenter des résultats

significatifs d'un point de vue statistique.

Deuxièmement, les résultats présentent la sélection systématique d'une position dans la coordonnée, M2 pour les résultats que nous détaillons ci-dessous.

Ainsi, dans ÉDOUARD ET EDMOND, les mots $_{[+ \text{NASAL}]\#}$, ceux $_{[- \text{NASAL}]\#}$, $_{[- \text{HAUT}]\#}$, $_{[+ \text{BAS}]\#}$, $_{[- \text{BAS}]\#}$, $_{[+ \text{ARRIÈRE}]\#}$, $_{[+ \text{ARRONDI}]\#}$ et ceux $_{[- \text{ARRONDI}]\#}$, occupent systématiquement M2. Rappelons que ce questionnaire mettait en évidence les corrélations entre M2 et les voyelles finales /i/, /o/, /a/, /õ/ et /ã/.

Dans NAROD ET NANOR, les mots possédant des finales vocaliques $[- \text{NASAL}]$, $[- \text{HAUT}]$, $[- \text{BAS}]$, $[+ \text{ARRIÈRE}]$, $[- \text{ARRIÈRE}]$ et ceux $_{[- \text{ARRONDI}]\#}$ sélectionnent dans 100% des cas M2. Ce questionnaire révélait les corrélations entre M2 et les finales vocaliques /o/, /a/ et /e/.

Enfin, dans le questionnaire d'enquête LEMOTO ET LEMOTAN, les mots $_{[+ \text{NASAL}]\#}$, tout comme ceux $_{[- \text{HAUT}]\#}$, $_{[+ \text{BAS}]\#}$ et $_{[- \text{BAS}]\#}$ sont systématiquement placés en M2. Ce questionnaire faisait émerger les corrélations entre M2 et les finales vocaliques /a/, /e/, /õ/ et /ã/. D'une manière générale, les résultats relevés dans les questionnaires d'enquête pour ce qui est des traits $[\pm \text{NASAL}]$ et $[- \text{HAUT}]$ suivent les contours de ceux que nous observions dans les corpus. Concernant les traits des finales vocaliques $[\pm \text{BAS}]$, $[\pm \text{ARRIÈRE}]$ et $[\pm \text{ARRONDI}]$, nous notons des résultats apparemment contradictoires. Un examen de ces traits distinctifs pour l'ensemble des données devrait clarifier ces points. Pour cette raison, nous proposons de confronter, dans les deux tableaux à venir, les résultats produits par deux différents totaux.

Le résultat présenté dans le tableau TOTAL traite l'hypothèse d'une corrélation entre l'une des deux positions de la coordonnée en fonction des traits distinctifs des voyelles finales pour l'ensemble des 1 962 mots interrogeant cette hypothèse. Nous y observerons particulièrement les traits des finales vocaliques $[\pm \text{BAS}]$ et $[\pm \text{ARRONDI}]$ pour lesquels nous avons relevé des résultats contradictoires.

Dans le tableau TOTAL-TOM ET ZOÉ, nous proposons les mêmes résultats sans ceux produits par le corpus TOM ET ZOÉ, soit l'étude de 294 mots examinant le lien entre la position dans la coordonnée et les traits des finales vocaliques. Dans ce dernier total, nous prêterons une attention particulière aux mots dont les finales vocaliques ont pour traits

distinctifs finales vocaliques [- ARRIÈRE]. En effet, l'assignation en M2 des finales vocaliques possédant ce trait est essentiellement portée par TOM ET ZOÉ et par un total de dix mots dans les questionnaires d'enquête ÉLISE ET ZOÉ et NAROD ET NANOR. Il nous faut asseoir plus solidement ce résultat, car comme nous l'avons indiqué à maintes reprises, nous nourrissons quelques hésitations quant aux résultats fournis par le corpus TOM ET ZOÉ. Comme nous le proposons précédemment, nous avons conservé les mêmes codes de présentation de nos tableaux, à savoir que sont notés en caractères gras les résultats pour lesquels la répartition est statistiquement significative, et en italiques ceux pour lesquels nous relevons une tendance.

TOTAL 1962	_V#									
	NASAL		HAUT		BAS		ARRIÈRE		ARRONDI	
	+	-	+	-	+	-	+	-	+	-
M1	248 40,59%	665 49,22%	222 53,24%	691 44,72%	514 43,71%	399 50,76%	524 46,70%	389 46,31%	221 50,34%	692 45,44%
M2	363 59,41%	686 50,78%	195 46,76%	854 55,28%	662 56,29%	387 49,24%	598 53,30%	451 53,69%	218 49,66%	831 54,56%
Total :	611 100,00%	1351 100,00%	417 100,00%	1545 100,00%	1176 100,00%	786 100,00%	1122 100,00%	840 100,00%	439 100,00%	1523 100,00%

Tableau 64 : Répartition par traits des mots possédant une finale vocalique pour l'ensemble de nos supports d'études

TOTAL-TOM ET ZOÉ 294	_V#									
	NASAL		HAUT		BAS		ARRIÈRE		ARRONDI	
	+	-	+	-	+	-	+	-	+	-
M1	42 36,52%	78 43,58%	25 48,08%	95 39,26%	59 36,65%	61 45,86%	58 37,66%	62 44,29%	34 39,53%	86 41,35%
M2	73 63,48%	<i>101</i> <i>56,42%</i>	27 51,92%	147 60,74%	102 63,35%	72 54,14%	96 62,34%	78 55,71%	52 60,47%	122 58,65%
Total :	115 100,00%	179 100,00%	52 100,00%	242 100,00%	161 100,00%	133 100,00%	154 100,00%	140 100,00%	86 100,00%	208 100,00%

Tableau 65 : Répartition par traits des mots possédant une finale vocalique pour l'ensemble de nos supports d'études sans TOM ET ZOÉ

De la confrontation des tableaux présentant la répartition par traits, nous dégagons, au premier examen, trois résultats consensuels statistiquement significatifs.

Ainsi, les mots $_{[+ \text{NASAL}]\#}$ occupent majoritairement M2 dans 59,41% des cas dans TOTAL, et 63,48% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 21,64$; $p \leq 0,000003$ et $\chi^2(1) = 8,36$; $p \leq 0,004$), tout comme les mots $_{[- \text{HAUT}]\#}$ sélectionnant la même position dans 55,28% dans TOTAL, et 60,74% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 17,2$; $p \leq 0,00003$ et $\chi^2(1) = 11,17$; $p \leq 0,001$), et ceux $_{[+ \text{BAS}]\#}$ sont 56,29% à favoriser M2 dans TOTAL pour 63,35% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 18,63$; $p \leq 0,00002$ et $\chi^2(1) = 11,48$; $p \leq 0,001$).

Il existe une corrélation entre M2 et les mots possédant une finale vocalique $_{[+ \text{NASAL}]\#}$, $_{[- \text{HAUT}]\#}$ et/ou $_{[+ \text{BAS}]\#}$.

De la confrontation des deux tableaux, nous parvenons à extraire deux nouveaux résultats : un résultat statistiquement significatif et une tendance.

D'une part, concernant les mots $_{[\pm \text{ARRIÈRE}]\#}$, les résultats sont de prime abord contradictoires. En effet, le tableau TOTAL positionne en M2 non seulement les mots $_{[+ \text{ARRIÈRE}]\#}$ dans 53,3% des 1 122 mots éprouvés sur cette hypothèse, mais également ceux $_{[- \text{ARRIÈRE}]\#}$ dans 53,69 % des 840 mots $_{[- \text{ARRIÈRE}]\#}$ présents dans nos supports d'études (respectivement $\chi^2(1) = 4,88$; $p \leq 0,03$ et $\chi^2(1) = 4,58$; $p \leq 0,03$). Or, ces résultats ne sont pas maintenus dans TOTAL-TOM ET ZOÉ. Dans ce nouveau total, seuls les mots $_{[+ \text{ARRIÈRE}]\#}$ font apparaître une dispersion significative : les mots $_{[+ \text{ARRIÈRE}]\#}$ occupent majoritairement M2 dans 62,34% des 154 mots testés sur cette hypothèse ($\chi^2(1) = 9,38$; $p \leq 0,002$).

Que retenir de ce résultat ?

En observant attentivement les résultats recueillis pour les mots en $_{[- \text{ARRIÈRE}]\#}$, nous constatons que sur les 840 mots testés dans le tableau TOTAL, 700 sont issus du corpus TOM ET ZOÉ. Étant donné, que le résultat plaçant en M2 les mots $_{[- \text{ARRIÈRE}]\#}$ ne sont portés que par le corpus TOM ET ZOÉ, nous ne pouvons le conserver. De l'examen des mots $_{[\pm \text{ARRIÈRE}]\#}$, nous affirmons que seuls les mots $_{[+ \text{ARRIÈRE}]\#}$ occupent majoritairement M2.

D'autre part, le tableau TOTAL-TOM ET ZOÉ fait émerger une tendance statistique dans laquelle les mots $_{[- \text{NASAL}]\#}$ auraient tendance à favoriser M2 dans 56,42% des cas ($\chi^2(1)$

= 2,96 ; $p \leq 0,09$). A la différence des résultats précédents, les résultats retenus pour les mots $_{[\pm \text{NASAL}]\#}$ ne sont pas contradictoires. Ils abondent dans le sens du facteur que nous présentions page 194, plaçant les mots se terminant par une voyelle (qu'elle soit orale ou nasale) en M2 dans 59,18% des occurrences éprouvant cette hypothèse. Nous attestons donc que les mots $_{[- \text{NASAL}]\#}$ ont tendance à occuper M2.

Enfin, concernant les traits des finales vocaliques $[\pm \text{ARRONDI}]$ pour lesquels nous observons des résultats contradictoires, la confrontation des tableaux TOTAL et TOTAL-TOM ET ZOÉ ne nous permet pas de les exploiter en l'état. En effet, le tableau TOTAL-TOM ET ZOÉ met en avant deux résultats plaçant les mots possédant une finale $[\pm \text{ARRONDI}]$ en M2. Or, au regard de l'ensemble des résultats présentés, cela attire notre attention ; soit ce trait n'est pas un facteur dans la sélection d'un ordre de coordination privilégié, soit le traitement des finales vocaliques $[\pm \text{NASAL}]$ opacifie nos résultats. Afin de répondre à ce questionnement, nous proposons ci-dessous deux tableaux. Par économie de traitement, nous ne soumettons que les résultats produits par l'ensemble des données sans celles extraites du TOM ET ZOÉ. Ce seul total à lui seul nous permet de préciser nos résultats. Ainsi, le premier tableau présente la répartition par traits distinctifs des voyelles finales orales et le second, la répartition par traits distinctifs des finales vocaliques nasales.

TOTAL-TOM ET ZOÉ 179	V#									
	NASAL		HAUT		BAS		ARRIÈRE		ARRONDI	
	+	-	+	-	+	-	+	-	+	-
M1		78 43,58%	25 48,08%	53 41,73%	17 36,96%	61 45,86%	31 42,47%	47 44,34%	19 47,50%	59 42,45%
M2		101 56,42%	27 51,92%	74 58,27%	29 63,04%	72 54,14%	42 57,53%	59 55,66%	21 52,50%	80 57,55%
Total :		179 100,00%	52 100,00%	127 100,00%	46 100,00%	133 100,00%	73 100,00%	106 100,00%	40 100,00%	139 100,00%

Tableau 66 : Répartition par traits des mots se terminant par une finale vocalique orale pour l'ensemble de nos supports d'études sans TOM ET ZOÉ

TOTAL-TOM ET ZOÉ	V #									
	NASAL		HAUT		BAS		ARRIÈRE		ARRONDI	
115	+	-	+	-	+	-	+	-	+	-
M1	42 36,52%		0	42 36,52%	42 36,52%	0	27 33,33%	15 44,12%	15 32,61%	27 39,13%
M2	73 63,48%		0	73 63,48%	73 63,48%	0	54 66,67%	19 55,88%	31 67,39%	42 60,87%
Total :	115 100,00%		0 100,00%	115 100,00%	115 100,00%	0 100,00%	81 100,00%	34 100,00%	46 100,00%	69 100,00%

Tableau 67 : Répartition par traits des mots se terminant par une finale vocalique nasale pour l'ensemble de nos supports d'études sans TOM ET ZOÉ

Conformément à ce que nous relevions précédemment, la répartition des mots possédant une finale vocalique [\pm NASAL], [- HAUT] et [+ BAS] est en adéquation avec ce que nous observons de la répartition des mots se finissant par une finale vocalique [\pm NASAL]. En effet, concernant les finales vocaliques orales, les mots possédant des finales [- NASAL], [- HAUT] ou/et [+ BAS] ont tendance à occuper M2 dans 56,42%, 58,27% et 63,04% des occurrences testées sur ces hypothèses (respectivement $\chi^2(1) = 2,96$ $p \leq 0,09$, $\chi^2(1) = 3,47$ $p \leq 0,06$ et $\chi^2(1) = 3,13$ $p \leq 0,08$). De la même façon, concernant les finales vocaliques nasales, les mots possédant des finales [+ NASAL], [- HAUT] ou/et [+ BAS] sélectionnent prioritairement M2 dans 63,48,% des occurrences testées sur ces trois hypothèses ($\chi^2(1) = 8,36$; $p \leq 0,004$).

Concernant les mots ayant pour finales vocaliques une voyelle [+ ARRIÈRE], l'observation de ces deux nouveaux tableaux nous permet d'assigner la position M2 pour l'essentiel aux mots possédant une finale vocalique $_ [+ NASAL] \#$. En effet, les mots ayant une finale vocalique nasale [+ ARRIÈRE] sont 66,67% à occuper M2 ($\chi^2(1) = 9$; $p \leq 0,003$).

Les mots se terminant par une voyelle orale [+ ARRIÈRE] suivent les mêmes contours, ils sont 57,53% à occuper prioritairement M2 ($\chi^2(1) = 1,66$ $p \leq 0,20$). Cependant, leur répartition n'étant pas statistiquement significative, il nous est impossible de conserver ce résultat.

Enfin, les deux répartitions en fonction du trait de nasalité des finales vocaliques nous permettent de clarifier les résultats que nous rencontrons pour les mots possédant des voyelles finales [\pm ARRondi]. Concernant les mots se terminant par une voyelle nasale, il

apparaît que les mots ayant des voyelles finales orales [- ARRONDI] ont tendance à favoriser prioritairement M2 dans 57,55% des cas ($\chi^2(1) = 3,17$ $p \leq 0,07$).

Au sujet des mots finissant par une voyelle nasale, les résultats sont moins lisibles.

En effet, si nous notons que les mots possédant une finale vocalique nasale [+ ARRONDI] sélectionnent majoritairement M2 dans 67,39% des cas ($\chi^2(1) = 5,57$ $p \leq 0,02$), nous relevons aussi que ceux [- ARRONDI] ont également tendance à occuper la même position dans 60,87% des cas ($\chi^2(1) = 3,26$ $p \leq 0,07$).

Comment expliquer ce résultat *a priori* contradictoire ? Rappelons dans un premier temps que les seules finales vocaliques nasales pour lesquelles nous avons relevé des répartitions significatives sont les voyelles /ã/ et /õ/. Or, si ces voyelles partagent l'ensemble de leurs traits distinctifs, elles se différencient par leur trait d'arrondissement : /ã/ est [- ARRONDI], et /õ/, [+ ARRONDI]. Il n'est donc pas surprenant au regard de nos résultats d'observer l'assignation de M2 pour les mots possédant une finale vocalique nasale en [+ ARRONDI] et [- ARRONDI]. De prime abord, nous en déduisons qu'en ce qui concerne la corrélation entre M2 en fonction des traits distinctifs des finales vocaliques nasales, le trait d'arrondissement n'est pas réellement actif. D'autant plus que nous constatons que les résultats présentés ci-dessus n'ont pas la même validation statistique. Les mots possédant une finale vocalique nasale [+ ARRONDI] fournissent un résultat, ceux [- ARRONDI], une tendance. Cette nuance nous permet d'affirmer que les mots possédant une finale vocalique nasale [+ ARRONDI] sélectionnent M2, comme ceux [- ARRONDI] dans une moindre mesure. Par extension, nous pouvons donc affirmer que les mots en _/õ/# occupent plus M2 que ceux en _/ã/#.

4.3.1.1.3. Bilan et analyse des résultats concernant les finales vocaliques

Au regard des résultats observés, de l'examen des mots se terminant par une finale vocalique, nous affirmons que :

- D'une manière générale, les mots possédant une finale vocalique occupent majoritairement, dans 59,18% des cas, M2 ;
- Les mots possédant une finale vocalique nasale sélectionnent plus massivement M2 (63,48% des cas) que les mots finissant par une finale vocalique orale, pour lesquels il ne s'agit que d'une tendance (56,42% des occurrences testées sur cette hypothèse).

- Dans le détail, pour les mots possédant une finale vocalique orale :
 - Les mots $_ /A/\#$ ont tendance à occuper M2 ;
 - Notons également les résultats relevés dans les questionnaires d'enquête prouvant que les mots en $_ /e/\#$, et en $_ /o/\#$ privilégient également cette même position ;
 - Enfin, les mots se terminant par une voyelle orale en termes de trait [- NASAL] et qui sont soit [- HAUT] (/e ø o ε œ ɔ a/), soit [+ BAS] (/ε œ ɔ a/) soit [- ARRONDI] (/i e ε/) ont tendance à être placés en M2.
- Concernant les mots possédant une finale vocalique nasale :
 - Les mots en $_ /ĩ/\#$ et en $_ /ã/\#$ sélectionnent massivement M2 ;
 - Tout comme les mots se terminant par une voyelle nasale en terme de trait [+ NASAL] et qui sont soit [- HAUT] (/ẽ ã õ/), soit [+ BAS] (/ẽ ã õ/), soit [+ ARRIÈRE] (/ã õ/), soit [+ ARRONDI] (/õ/);
 - Les mots [- ARRONDI] (/ẽ ã/) ont également tendance à occuper cette même position M2.
- D'une manière générale, pour l'ensemble des mots se terminant par une voyelle, il existe une corrélation entre M2 et les traits [+ NASAL], [- NASAL], [- HAUT], [+ BAS], ou/et [+ ARRIÈRE].

Comment expliquer ces résultats ?

On peut rendre compte de la préférence des mots se terminant par une finale vocalique pour la position M2 en invoquant deux contraintes bien connues en Théorie de l'Optimalité : la contrainte *VV, présentée précédemment, qui sanctionne les suites de deux voyelles, et la contrainte de non-marque *CODA, imposant des syllabes ouvertes.

*VV : deux voyelles ne peuvent se suivre.

*Coda (aussi connue sous le nom de « NoCoda ») : les syllabes n'ont pas de coda.

Le fait de repousser la finale $_ /e/\#$ le plus loin possible de la conjonction de coordination afin d'éviter de produire une séquence /e e/ relève d'une instanciation du principe dit « de contour

obligatoire » (Obligatory Contour Principle, Goldsmith (1976)) initialement proposé par Leben (1973)⁶⁹. Ce principe est un principe de dissimilation qui proscrit la répétition d'un même segment. La théorie OT le formalise sous la forme d'une contrainte dite « OCP » (McCarthy 1986).

OCP : pas de segments identiques adjacents.

Cependant, si la préférence des mots en $_ /e/\#$ pour la position M2 semble motivée par la conjonction de ces trois contraintes, cette conjonction n'explique pas pourquoi les seules autres finales vocaliques favorisant prioritairement M2 sont $/A/, /ã/, /O/$ et $/õ/$.

Quels sont les autres facteurs en jeu ? De même, comment expliquer également que pour ces quatre finales que l'on peut organiser selon une opposition $[- \text{NASAL}] \sim [+ \text{NASAL}]$ ($/A/ \sim /ã/$ et $/O/ \sim /õ/$), la sélection de M2 est statistiquement plus forte pour les mots finissant par une finale vocalique spécifiée $[+ \text{NASAL}]$ que pour ceux spécifiés par une voyelle finale $[- \text{NASAL}]$?

Nous pouvons tenter de trouver une explication, ou du moins, une formalisation à ceci en invoquant les sous-hiérarchies universelles de contraintes, telles que proposées par Prince et Smolensky (1993 : 129). Celle qui nous intéresse ici est la HIÉRARCHIE DES NOYAUX (également connue sous le nom « Hiérarchie des Pics », *Peak Hierarchy*), que nous reproduisons sous la forme que leur donne Clements (1997 : 296) :

HIÉRARCHIE DES NOYAUX :

$*N/t \gg *N/d \gg *N/s \gg *N/z \gg *N/n \gg *N/r \gg *N/i \gg *N/a$

Cette hiérarchie formalise le fait que comme le rappelait Plénat (1999), « la syllabation d'une chaîne de segments donnée est, d'une façon ou d'une autre, fonction de la sonorité relative des éléments de cette chaîne ». Dans la lignée de l'échelle de sonorité de Grammont (1933)⁷⁰, plus un segment phonologique est haut sur l'échelle de sonorité, plus il est susceptible d'être

69 Leben, W. R. (1973). *Suprasegmental phonology* (Doctoral dissertation, Massachusetts Institute of Technology). Dans sa thèse de doctorat Leben relatait des phénomènes tonals condamnant l'apparition de deux éléments identiques adjacents sur le même niveau de représentation. Ce phénomène a ensuite été étendu aux segments puis aux traits par McCarthy (1986). «Au niveau mélodique, deux éléments identiques et adjacents sont prohibés.»

70 Une échelle de sonorité inspirée par Grammont 1933 est proposée en annexe.

noyau de syllabes. Les voyelles ouvertes sont donc favorisées en noyau.

La conjonction des contraintes *CODA, *VV et de la HIÉRARCHIE DES NOYAUX conduit à une hypothèse vraisemblable expliquant la position en M2 des voyelles finales /A/ et /O/. En effet, d'une part, en rejetant en M2 les mots possédant pour voyelle finale /A/ ou /O/, nous ne transgressons pas *VV. D'autre part, comme indiqué par la HIÉRARCHIE DES NOYAUX /A/ et /O/ sont de meilleurs candidats pour être noyaux de syllabes que les voyelles /i/, /y/ ou /u/, de même que /E/, puisque ces trois voyelles sont [– HAUT].

Concernant la position des voyelles nasales $_{/}\tilde{/}\#$ et $_{/}\tilde{\alpha}/\#$ en M2, nous proposons ici quelques pistes de réflexion, qui demanderont à être approfondies lors de travaux futurs. Si notre précédente hypothèse motive la présence en M2 des voyelles nasales, elle n'explique pas pourquoi ces voyelles favorisent plus M2 que leurs voyelles orales correspondantes.

On pourrait supposer l'action d'un paramètre de phonétique articulatoire. En effet, les voyelles nasales sont moins ouvertes que les voyelles orales du fait de l'abaissement du voile du palais nécessaire à leur réalisation, ainsi que le suggèrent Delvaux, Metens, et Soquet (2002) : « le corrélat acoustique de l'abaissement du voile se situe essentiellement au niveau de l'intensité : les nasalisées (et *a fortiori* les nasales) accusent une perte d'énergie généralisée par rapport aux orales. » On en revient donc à la hiérarchie des noyaux, qui, est, comme nous le disions plus haut, ni plus ni moins qu'une adaptation OT de la hiérarchie de sonorité. Au vu des résultats de notre étude et si l'on admet donc que les voyelles nasales sont moins ouvertes que les orales correspondantes, nous devons alors placer les voyelles nasales plus bas dans la hiérarchie des noyaux que leurs correspondantes orales, comme suit :

$$*N/i, *N/y, *N/u \gg *N/A, *N/O, *N/E \gg *N/ \tilde{\alpha}, *N/ \tilde{\omega}$$

Ce dernier paramètre n'est pas sans évoquer la loi (ou hiérarchie) dite « du second lourd » proposée par Hagège (1985, 244-246). Ce principe selon lequel dans un binôme, le second terme est plus lourd et plus complexe que le premier, peut également rendre compte de l'assignation préférentielle en M2 des mots en $_{/}\tilde{/}\#$ et $_{/}\tilde{\alpha}/\#$. En effet, la complexité du second segment se manifeste en termes de nombres de syllabes, de nombres de consonnes, de consonnes ou de voyelles plus arrières, ou de phonèmes de plus basse fréquence ; le second segment a plus de syllabes que le premier, plus de consonnes, etc. Or, les voyelles les plus

basses en termes de fréquence sont /ɔ̃/ et /ɑ̃/⁷¹. Il semble vraisemblable que conformément aux observations de la loi du second lourd, les mots se terminant par ces voyelles nasales soient rejetés en seconde position du binôme, tout comme les mots en _/A/# et _/O/# dans une moindre mesure⁷².

Enfin, pour revenir à une caractérisation de nature plus phonologique, nous rappellerons que les voyelles nasales peuvent être considérées comme des voyelles marquées par rapport aux voyelles orales. C'est possiblement ce caractère marqué qui fait préférer la position M2 aux mots terminés par une telle voyelle nasale.

En effet, si l'on entend l'opposition marqué / non marqué comme proposée par Jakobson *et al.* (1952), c'est-à-dire en considérant les marques de corrélation, les voyelles nasales sont marquées du fait de leur nasalisation. D'autant plus que d'une manière générale, la distribution des voyelles dans l'ensemble des langues du monde le confirme : « il n'existe aucun système phonémique qui compte plus de voyelles nasales que de voyelles orales. » Durand (2005).

Or, comme le relevaient Benor et Levy (2006) (cf page 16), il existe une tendance à ce que l'élément le moins marqué tende à occuper la position la moins marquée, et inversement. De fait, les mots les plus marqués sélectionneraient plutôt M2 qui est la position la plus marquée, ce qui, au regard de nos résultats, justifie que les mots possédant une finale vocalique nasale, étant plus marqués que ceux se terminant par une finale vocalique orale, occupent plus expressément M2.

4.3.1.2. Consonne

L'intérêt que nous portons à l'examen des finales consonantiques est double.

L'observation des finales consonantiques permet, d'une part, d'examiner plus avant les phénomènes observés dans les sections précédentes. C'est pourquoi, nous prêterons une attention particulière aux mots se terminant par des segments [± SONANT], [± CONTINU], [± NASAL], [± HAUT], [± ARRIÈRE], [± VOISÉ]. Ces traits étant partagés avec les voyelles, il nous paraît pertinent d'en proposer une analyse globale.

D'autre part, comme nous l'indiquons page 195, il existe une corrélation entre la qualité du

71 Cf Léon (1992) page 83. Nous proposons le classement de la voyelle la plus aiguë à la plus grave suivant :

/i/ >> /e/ >> /ɛ/ >> /y/ >> /a/ >> /ø/ >> /œ/ >> /ɑ/ >> /ə/ >> /ɔ/ >> /o/ >> /u/ >> /ẽ/ >> /œ̃/ >> /ã/ >> /ɔ̃/

72 En effet, /A/ et /O/ étant plus haut en termes de fréquence que /ɔ̃/ et /ɑ̃/, il semble attendu qu'ils soient un peu moins enclins à occuper M2.

segment final et la sélection d'une forme coordonnée syndétique par « et » : les mots se terminant par une coda occupent majoritairement M1. A présent, l'enjeu de cette section est de révéler les segments activant cette sélection et les traits qu'ils ont en commun.

Nous nous proposons donc d'examiner la répartition par phonème des mots finissant par une consonne pour chaque support d'études examinant cette hypothèse. Pour cela, nous exploiterons les six corpus et les questionnaires d'enquête ÉLISE ET ZOÉ, ÉDOUARD ET EDMOND et NAROD ET NANOR. Ces trois derniers supports explorent précisément la corrélation entre les finales consonantiques et la sélection d'une des deux positions dans la coordination. Nous prêterons une attention particulière aux résultats qu'ils fournissent.

Les autres questionnaires d'enquête, n'interrogeant pas cette nouvelle hypothèse, avaient été neutralisés sur cette question. Nous ne les utiliserons pas.

Au total, nous avons étudié la répartition en M1 et en M2 de 1 580 mots finissant par une consonne.

Dans une première étape, nous présenterons un relevé phonémique des finales consonantiques en fonction de leur position dans la coordonnée. Nous observerons ces répartitions pour chaque support d'études examiné indépendamment les uns des autres, avant de commenter ces mêmes répartitions pour l'ensemble de nos données interrogeant le rôle des finales consonantiques. Puis, nous étudierons ce que ces consonnes partagent en commun en proposant leur répartition par traits distinctifs, d'abord par support d'études, puis pour l'ensemble de nos données. Enfin, nous effectuerons une analyse des résultats concernant les finales consonantiques.

4.3.1.2.1. Répartition phonémique des finales consonantiques

Dans cette section, nous étudions la répartition générale des finales consonantiques pour chaque support d'études. Les neuf tableaux A 11 sont proposés en annexe.

De l'examen des corpus observant la répartition en M1 et en M2 des mots possédant une finale consonantique par type phonémique, nous relevons quatre résultats statistiquement significatifs et cinq tendances. Ces résultats sont respectivement mentionnés en caractères gras quand leur répartition est significative, et en italique lorsque nous relevons des tendances. Rappelons que nous avons fait le choix d'étendre le relevé de nos résultats aux résultats pour lesquels la valeur-p du χ^2 est de .10. En effet, même si certains critères peuvent sembler

moins massifs que ceux pour lesquels nous relevons une répartition statistiquement significative, ils n'en sont pas moins pertinents. Travaillant dans le cadre d'une approche en termes d'interactions de contraintes, nous sommes convaincue que les paires que nous examinons peuvent être sujettes aux jeux de divers facteurs. Les tendances statistiques ne sont parfois que des tendances du fait de l'action d'autres critères. C'est pourquoi nous pensons qu'il est nécessaire de les relever et de les commenter.

Dans le détail, tous les corpus, exception faite du corpus CUL ET CHEMISE, permettent l'examen de résultats. De plus, les résultats relevés montrent tous l'assignation en M1 de segments phonémiques des finales consonantiques.

Dans le corpus TARN ET GARONNE, nous relevons un résultat et une tendance statistique. Tout d'abord, les mots en $_ /z/\#$ sélectionnent systématiquement M1 ($\chi^2(1) = 4$; $p \leq 0,05$). Même si ce résultat semble apparemment catégorique, nous ne le conservons pas. Calculé sur un échantillon de quatre mots, il ne peut prétendre à une validation statistique. Nous relevons cependant de ce corpus que les mots en $_ /s/\#$ ont tendance à occuper M1 dans 70,59% des occurrences se terminant par $/s/$ ($\chi^2(1) = 2,88$; $p \leq 0,09$).

Dans le corpus LISSE ET SOYEUX, les mots en $_ /t/\#$ ont tendance à sélectionner M1 dans 77,78% des cas ($\chi^2(1) = 2,78$; $p \leq 0,09$). Cependant, à nouveau, nous ne conserverons pas ce résultat. En effet, sa valeur théorique étant inférieure à cinq, le test de Khi Deux n'est pas applicable.

Le corpus TOM ET ZOÉ, quant à lui, fait apparaître deux résultats significatifs et une tendance : les mots en $_ /d/\#$ et en $_ /n/\#$ occupent M1 dans 68,42% et 58,67 % des occurrences testant ces hypothèses (respectivement $\chi^2(1) = 7,74$; $p \leq 0,01$ et $\chi^2(1) = 10,4$; $p \leq 0,001$), tout comme les mots en $_ /k/\#$ dont la répartition montre une tendance à sélectionner cette même position dans 68% des cas ($\chi^2(1) = 3,24$; $p \leq 0,07$).

Le corpus VITE ET BIEN révèle un résultat et une tendance statistique : les mots en $_ /k/\#$, en adéquation avec la tendance observée dans le corpus TOM ET ZOÉ, favorisent M1 dans 75 % des occurrences testées ($\chi^2(1) = 5$; $p \leq 0,03$). De la même façon, les mots se terminant par la consonne $/m/$ semblent avoir massivement tendance à occuper M1 dans 83,33% des cas ($\chi^2(1) = 2,67$; $p \leq 0,10$). Cependant, nous ne conservons pas ce dernier résultat. L'échantillon est

trop petit (à peine six mots) pour que le test de Khi Deux soit applicable.

Enfin, dans le corpus ARTS ET MÉTIERS, nous relevons une tendance statistique. Les mots en *_/n/#* semblent occuper systématiquement M1 ($\chi^2(1) = 3 ; p \leq 0,08$). Cependant, n'étant testée que sur trois mots, nous ne pouvons affirmer cette hypothèse.

Attardons-nous à présent sur les résultats émergeant des questionnaires d'enquête ÉLISE ET ZOÉ, ÉDOUARD ET EDMOND et NAROD ET NANOR. Nous y relevons quelques données pour lesquelles la valeur-p est inférieure à .1 que nous ne pouvons valider par un Khi Deux, compte tenu de la petite taille des échantillons testés. Nous présentons néanmoins leurs résultats. En effet, d'une part, ces trois questionnaires explorent précisément l'hypothèse examinée dans cette section, à savoir si la qualité du segment final est un facteur actif dans la sélection d'un ordre binomial coordonné. De fait, exception faite d'ÉLISE ET ZOÉ, pour tout autre questionnaire d'enquête, les facteurs autres que l'hypothèse testée ont été neutralisés. D'autre part, chaque paire retenue a été statistiquement sélectionnée en amont ($p \leq 0.05$) par plus de cinquante interviewés.

Pour ces motifs, nous pensons que l'examen attentif de leurs résultats est pertinent.

Dans le détail, pour le questionnaire d'enquête ÉLISE ET ZOÉ, les mots en *_/m/#* sélectionnent systématiquement M1. Cependant, l'examen attentif des quatre paires⁷³ en question révèle que le prénom « TOM » est présent dans trois binômes. En termes de type (par opposition au décompte en token déjà évoqué), au total dans ce questionnaire d'enquête, nous ne testons pas quatre mots, mais seulement deux. De plus, deux des paires relevées peuvent être induites par d'autres facteurs : « TOM ET THOMAS » peut être un artefact du facteur morphologique que nous décrivions page 88 et « TOM ET ROMÉO » et à nouveau « TOM ET THOMAS » pourraient être gouvernés par le facteur de taille dont nous avons démontré la prégnance page 118. Pour ces raisons, nous ne conservons pas ce résultat.

Dans le questionnaire d'enquête ÉDOUARD ET EDMOND, il apparaît que les mots en *_/R/#* occupent systématiquement M1. Ce résultat est *a priori* intéressant. Cependant, les trois paires éprouvées sur cette hypothèse opposent systématiquement une finale vocalique à une finale consonantique. Nous avons démontré que les mots se terminant par une consonne occupent majoritairement M1. Comment ne pas supposer qu'au-delà d'une potentielle sélection en tête

⁷³ Les paires retenues sur l'hypothèse liant les mots en *_/m/#* à M1 sont : « TOM ET THOMAS », « TOM ET LOU », « TOM ET ROMÉO » et « MAXIME ET MARINE ».

de coordonnées des mots en $_ /r/\#$, l'assignation en M1 que nous relevons ici est simplement due au caractère consonantique de $/r/$? Ne pouvant le préciser, nous ne conservons pas ce résultat.

Enfin, dans le questionnaire d'enquête NAROD ET NANOR, les mots en $_ /g/\#$ sélectionnent systématiquement M1. Cependant, pour les mêmes raisons que précédemment, étant donné que les mots en $_ /g/\#$ sont toujours confrontés à des mots se terminant par une voyelle, nous ne conservons pas ce résultat.

Au regard de l'examen par support d'études de la répartition phonémique des finales consonantiques en M1 et en M2, nous relevons trois résultats : les mots en $_ /s/\#$, en $_ /k/\#$, en $_ /n/\#$ et en $_ /d/\#$ sélectionnent majoritairement M1.

Nous constatons également que ces résultats ne sont souvent portés que par un seul support d'études. Il nous faut les asseoir plus solidement.

C'est pourquoi, nous les comparerons dans les paragraphes à suivre à ceux obtenus pour l'ensemble des données testant l'hypothèse de la corrélation entre une position dans la coordonnée et les phonèmes des finales consonantiques. Nous appellerons cet ensemble de données « TOTAL ».

Rappelons que le corpus TOM ET ZOÉ est, du fait de sa constitution, le support d'études le moins contrôlé de nos supports d'études. Bien que présentant d'immenses avantages dont sa taille et les neutralisations syntaxiques et sémantiques liées à l'examen des prénoms, ce corpus recueilli sur Internet ne possède pas de déclarations de contenus. Plus que tout autre, il peut être assujéti à d'autres facteurs dont les facteurs psycho-affectifs que nous présentions page 83. C'est pourquoi, sans remettre en question ses qualités indéniables, nous le manipulons avec précaution.

Nous proposerons donc les résultats de l'ensemble de nos données sans ceux fournis par le corpus TOM ET ZOÉ. Nous nommerons ce nouvel ensemble « TOTAL-TOM ET ZOÉ ». Dans ce dernier, nous porterons une attention particulière à la répartition en M1 et en M1 des mots en $_ /k/\#$, en $_ /n/\#$ et en $_ /d/\#$. En effet, comme nous le précisons ci-dessus, ces résultats étant essentiellement présents dans le corpus TOM ET ZOÉ, il nous faut les confirmer.

TOTAL 1788	_C#																	
	p	t	k	b	d	g	f	s	ʃ	v	z	ʒ	l	m	n	ɲ	ŋ	ʁ
M1	0 0,00%	66 50,77%	54 65,06%	2 66,67%	44 69,84%	4 80,00%	1 50,00%	107 50,71%	3 50,00%	1 20,00%	36 49,32%	12 60,00%	109 55,61%	68 50,75%	225 57,99%	2 50,00%	1 50,00%	134 52,96%
M2	2 100,00%	64 49,23%	29 34,94%	1 33,33%	19 30,16%	1 20,00%	1 50,00%	104 49,29%	3 50,00%	4 80,00%	37 50,68%	8 40,00%	87 44,39%	66 49,25%	163 42,01%	2 50,00%	1 50,00%	119 47,04%
Total :	2 100,00%	130 100,00%	83 100,00%	3 100,00%	63 100,00%	5 100,00%	2 100,00%	211 100,00%	6 100,00%	5 100,00%	73 100,00%	20 100,00%	196 100,00%	134 100,00%	388 100,00%	4 100,00%	2 100,00%	253 100,00%

Tableau 68 : Répartition phonémiques des finales consonantiques pour l'ensemble des supports d'études

TOTAL-TOM ET ZOÉ 368	_C#																	
	p	t	k	b	d	g	f	s	ʃ	v	z	ʒ	l	m	n	ɲ	ŋ	ʁ
M1	0 0,00%	21 60,00%	37 63,79%	2 66,67%	5 83,33%	4 80,00%	1 50,00%	29 56,86%	3 50,00%	1 20,00%	9 56,25%	12 60,00%	33 57,89%	13 65,00%	22 52,38%	2 50,00%	1 50,00%	56 50,91%
M2	2 100,00%	14 40,00%	21 36,21%	1 33,33%	7 16,67%	1 20,00%	1 50,00%	22 43,14%	3 50,00%	4 80,00%	7 43,75%	8 40,00%	24 42,11%	7 35,00%	20 47,62%	2 50,00%	1 50,00%	54 49,09%
Total :	2 100,00%	35 100,00%	58 100,00%	3 100,00%	6 100,00%	5 100,00%	2 100,00%	51 100,00%	6 100,00%	5 100,00%	16 100,00%	20 100,00%	57 100,00%	20 100,00%	42 100,00%	4 100,00%	2 100,00%	110 100,00%

Tableau 69 : Répartition phonémiques des finales consonantiques pour l'ensemble des supports d'études sans TOM ET ZOÉ

La confrontation des tableaux ci-dessus fait apparaître quatre résultats pour lesquels la répartition est statistiquement significative. Nous les avons mentionnés en caractères gras. Nous relevons également une tendance statistique notée en italique.

Dans le détail, nous reviendrons prioritairement sur les précédents résultats que nous avons relevés de l'examen phonémiques des finales consonantiques par support d'études, à savoir la répartition des mots en */s/#*, */k/#*, en */d/#*, et en */n/#*.

Ainsi, comme le montre la comparaison des tableaux ci-dessus, les résultats relevés pour les mots en */s/#* dans le corpus TARN ET GARONNE ne sont pas confirmés : les mots en */s/#* sont 50,71% dans TOTAL, et 56,86% dans TOTAL-TOM ET ZOÉ, à sélectionner M1. Cependant, aucune de ces deux répartitions n'exclut l'hypothèse nulle (respectivement $\chi^2(1) = 0,04$; $p \leq 0,84$ et $\chi^2(1) = 0,96$; $p \leq 0,33$). De fait, nous ne pouvons conserver ce résultat.

Concernant les mots en */k/#*, la confrontation des deux totaux corrobore les résultats fournis par TOM ET ZOÉ et VITE ET BIEN. En effet, les mots en */k/#* sont 65,06% dans TOTAL et 63,79% dans TOTAL-TOM ET ZOÉ à préférer M1 (respectivement $\chi^2(1) = 7,53$; $p \leq 0,01$ et $\chi^2(1) = 4,41$; $p \leq 0,04$). Nous affirmons donc l'existence d'une corrélation liant les mots en */k/#* à la position M1.

La confrontation des résultats émis pour les deux totaux pour les mots en */d/#* fait apparaître deux répartitions en M1 suivant *a priori* les mêmes contours. Les mots en */d/#* occupent M1 dans 69,84% des cas dans TOTAL et 83,33% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 9,92$; $p \leq 0,002$ et $\chi^2(1) = 2,67$; $p \leq 0,10$). Si les deux résultats semblent indiquer la même répartition, la validation statistique est différente dans les deux totaux. TOTAL met en avant une distribution significative, TOTAL-TOM ET ZOÉ, une tendance. Il apparaît également que cette tendance est calculée sur un échantillon de six items. Nous ne pouvons donc pas appliquer le test de Khi Deux pour les mots en */d/#* dans TOTAL-TOM ET ZOÉ.

De plus, si les résultats de TOTAL semblent massifs, leur examen détaillé révèle que sur les 63 paires étudiées, 57 proviennent du corpus TOM ET ZOÉ. Or, ces 57 paires ne mobilisent en fait, en termes de type, que deux prénoms en */d/#* : « JADE » et « MATHILDE ». Nous ne pouvons pas asseoir une hypothèse sur un échantillon aussi petit. D'autant plus que ces deux prénoms pourraient être influencés par d'autres facteurs. « MATHILDE » se termine par une coda branchante si l'on considère comme nous l'indiquions page 25 que le e muet en

finale de mots n'est pas présent au niveau sous-jacent. Soupçonnant un rôle actif en fonction de la structure des syllabes finales, nous examinerons précisément ce point en comparant entre autre les mots en $_VC\#$ et en $_VCC\#$ en 4.3.2.2. structure de la syllabe finale page 261. Quant au prénom « JADE », en tant que prénom monosyllabique, il tombe également sous le facteur de taille que nous présentions page 96 liant les mots les plus courts à M1.

Pour tous ces motifs, nous ne conservons pas les résultats liant M1 aux mots en $_d/\#$.

L'observation des deux tableaux fait également apparaître que les mots en $_n/\#$ occupent majoritairement M1 dans 57,99% des cas dans TOTAL et 52,38% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 9,91$; $p \leq 0,002$ et $\chi^2(1) = 0,1$; $p \leq 0,76$). Ces résultats semblent confirmer ceux relevés dans TOM ET ZOÉ. Nous notons cependant que la signification asymptotique du résultat dans TOTAL-TOM ET ZOÉ est nettement supérieure à .10. N'excluant pas l'hypothèse nulle, la répartition des mots en $_n/\#$ dans TOTAL-TOM ET ZOÉ n'est pas significative. De plus, il apparaît que sur les 388 mots en $_n/\#$ retenus dans TOTAL, 346 proviennent du corpus TOM ET ZOÉ. De fait, n'étant porté que par un seul corpus, nous ne conservons pas le résultat assignant les mots en $_n/\#$ en M1.

Enfin, TOTAL fait émerger un résultat pour les mots en $_l/\#$. Ces derniers sélectionnent prioritairement M1 dans 55,61% des 196 mots testés sur cette hypothèse ($\chi^2(1) = 2,47$; $p \leq 0,11$). Certes, ce résultat n'est pas statistiquement aussi fort que nous le souhaiterions. Rappelons que la valeur-p de .10 est prise pour référence afin de définir, en-deçà, une présomption plus ou moins forte contre l'hypothèse nulle en faveur de l'hypothèse éprouvée, et au dessous, l'impossibilité de la rejeter. Or, ne pas rejeter de l'hypothèse nulle ne signifie pas pour autant que l'on rejette nécessairement l'hypothèse testée. Cela signifie seulement qu'on ne peut rien conclure à son propos : ni l'infirmer, ni la confirmer. De plus, la signification asymptotique est graduelle. De fait, avec une valeur-p de 0,11, si on ne peut pas rejeter l'hypothèse nulle, son « non-rejet » n'est pas réellement ferme. En termes d'interactions de contraintes, nous pensons que c'est un résultat qu'il nous faut interpréter.

Cependant, en le confrontant notamment à ceux relevés dans TOTAL-TOM ET ZOÉ dans lequel les mots en $_l/\#$ occupent M1 dans 57,89% des cas ($\chi^2(1) = 1,42$; $p \leq 0,23$). Exception faite de la validation statistique, ce résultat est en adéquation avec le précédent résultat. TOTAL-TOM ET ZOÉ ne confirme pas l'assignation en M1 des mots en $_l/\#$. Nous ne pouvons conserver ce résultat.

De la répartition phonémique des finales consonantiques pour l'ensemble de nos supports d'études, au regard des résultats relevés, seul un résultat est confirmé. Il existe une corrélation entre M1 et les mots en $_ /k/\#$. Ces derniers favorisent majoritairement la première position dans la coordonnée.

Ce qui semble étonnant, c'est que si l'assignation en M1 des mots se terminant par une consonne est établie, la répartition phonémique ne fait émerger qu'un seul résultat pour les mots en $_ /k/\#$.⁷⁴

Pourquoi seules les consonnes finales /k/ font apparaître une corrélation avec l'une des deux positions dans la coordonnée ? Qu'ont-elles de singulier ? Nous apporterons une attention particulière à ces questionnements en 4.3.1.2.3. Bilan et analyse des résultats concernant les finales consonantiques page 237 .

4.3.1.2.2. Répartition par traits distinctifs des finales consonantiques

Nous interrogeant sur ce que les finales consonantiques possèdent de commun et de dissemblable, nous proposons dans cette section une répartition par traits distinctifs en M1 et en M2 de celles-ci.

Dans un premier temps, nous relèverons les résultats présentés par chaque support d'études, indépendamment les uns des autres. Puis, nous les confronterons aux résultats retenus de l'ensemble de nos données.

Rappelons que les supports d'études examinant l'hypothèse d'une corrélation entre les traits distinctifs des finales consonantiques et l'une des deux positions dans la parataxe syndétique sont les mêmes que précédemment, à savoir les six corpus et les questionnaires d'enquête ÉLISE ET ZOÉ, ÉDOUARD ET EDMOND et NAROD ET NANOR.

Les neuf tableaux proposant la répartition par traits des finales consonantiques sont en annexe du tableau A 12a au tableau A 12i.

De l'examen des corpus observant la répartition par traits distinctifs des finales consonantiques, nous relevons vingt-huit résultats statistiquement significatifs, mentionnés en caractère gras, et quinze tendances, notées en italique. Nous présentons ci-après les quinze résultats et les sept tendances provenant de l'ensemble des six corpus. Les treize résultats et neuf tendances issus des questionnaires d'enquête seront commentés ultérieurement.

⁷⁴ Les mots se terminant par une consonne occupent M1 dans 56,53% des cas ($\chi^2(1) = 7,58$; $p \leq 0,006$)

Dans le détail, comme précédemment, tous les corpus, exception faite du corpus CUL ET CHEMISE, fournissent des résultats ou tendances. Chacun d'entre eux révèle l'assignation en M1 de traits distinctifs.

Dans le corpus TARN ET GARONNE, nous relevons cinq répartitions significatives. Les mots se terminant par une consonne [+ CORONAL] occupent majoritairement M1 dans 68,09% des cas ($\chi^2(1) = 6,15$; $p \leq 0,01$). De la même façon, les mots possédant une finale [- ARRIÈRE] sont 66% à privilégier cette même position ($\chi^2(1) = 5,12$; $p \leq 0,02$). Les mots _[- ANTÉRIEUR]# sélectionnent également M1 dans 65,96% des occurrences testant l'hypothèse d'une corrélation entre la finale consonantique [- ANTÉRIEUR] et l'une des deux positions dans la coordonnée ($\chi^2(1) = 4,79$; $p \leq 0,03$). 64% des mots _[- VOISÉ]# favorisent M1 ($\chi^2(1) = 3,92$; $p \leq 0,05$). Enfin, les mots ayant une finale consonantique [- SONANT]# sont 63,79% à être prioritairement placés en M1 ($\chi^2(1) = 4,41$; $p \leq 0,04$).

Notons également deux résultats qui, bien que non validés statistiquement, méritent néanmoins notre attention. Les mots possédant une finale consonantique [- NASAL] ou [- HAUT] semblent être plutôt positionnés en M1 dans 58,09 % et 59,09 % des mots testant ces hypothèses (respectivement $\chi^2(1) = 2,45$; $p \leq 0,12$ et $\chi^2(1) = 2,18$; $p \leq 0,14$). La valeur-p de ces deux résultats est juste en-deçà du seuil de 0.10. Nous sommes néanmoins convaincue de leur intérêt. La faiblesse de leur validation statistique est, selon nous, un artefact de l'interaction des nombreux facteurs actifs dans la sélection de composés syndétiques. Quoiqu'il en soit, nous ne les validerons qu'après les avoir confrontés aux résultats relevés pour l'ensemble de nos supports d'études.

Le corpus LISSE ET SOYEUX fait émerger un nouveau résultat : les mots se terminant par une consonne [- CONTINU] occupent majoritairement M1 dans 76,92% des cas ($\chi^2(1) = 3,77$; $p \leq 0,05$). Calculé sur seulement treize mots, ce résultat peut paraître fragile. Néanmoins, étant donné que la valeur théorique du test de Khi Deux mis en place est supérieure à 5, le test est applicable. Nous confirmons ce résultat.

Notons également la répartition en M1 et M2 des consonnes _[- REL RET]#. Il apparaît que les mots ayant une finale consonantique [- REL RET] semblent sélectionner M1 dans 65,22% des occurrences ($\chi^2(1) = 2,13$; $p \leq 0,14$). A nouveau, ce résultat ne peut prétendre à une validation statistique, sa valeur-p étant juste au dessus de .10. Cependant, comme précédemment, convaincue du jeu de l'interaction de contraintes, nous conservons ce résultat.

Le corpus TOM ET ZOÉ, quant à lui, met en valeur neuf répartitions significatives et cinq tendances. Les mots se terminant par une consonne [+ SONANT] occupent majoritairement M1 dans 55,53% des cas ($\chi^2(1) = 9,06$; $p \leq 0,003$). 56,36% des mots ayant une finale consonantique [+ REL RET] sélectionnent la même position ($\chi^2(1) = 10,3$; $p \leq 0,001$). Ceux possédant une consonne finale [+ VOISÉ] sont 55,84% à privilégier M1 ($\chi^2(1) = 11,68$; $p \leq 0,001$), de la même façon que les mots [+ CORONAL]# sont majoritairement placés la même position dans 54,8% ($\chi^2(1) = 7,87$; $p \leq 0,01$).

Les dix autres résultats fournis par le corpus TOM ET ZOÉ sont apparemment ambigus. En effet, ces résultats présentent des paires en [\pm TRAIT] sélectionnant tous la même position : M1.

Les mots se terminant par une consonne [\pm NASAL] occupent majoritairement M1. 56,69% des mots possédant une finale consonantique [+ NASAL] et 53,25% de ceux en [- NASAL] sélectionnent cette position (respectivement $\chi^2(1) = 6,82$; $p \leq 0,01$ et $\chi^2(1) = 2,86$; $p \leq 0,09$).

De même, les mots se terminant par une consonne [\pm ARRIÈRE] privilégient prioritairement M1. Ceux en [- ARRIÈRE] sont 54,03 % à occuper M1, pour 56,55% des mots se terminant par une finale consonantique [+ ARRIÈRE] (respectivement $\chi^2(1) = 6,29$; $p \leq 0,01$ et $\chi^2(1) = 2,88$; $p \leq 0,09$).

Les mots finissant par une consonne [\pm CONTINU] favorisent la même position M1. Les mots se terminant par une consonne [+ CONTINU] sont 53,91% à sélectionner M1, pour 57,06% des mots possédant une finale consonantique [- CONTINU] (respectivement $\chi^2(1) = 5,87$; $p \leq 0,02$ et $\chi^2(1) = 3,53$; $p \leq 0,06$).

Les mots se terminant par une consonne [\pm ANTÉRIEUR] sélectionnent majoritairement M1. 54,03% des mots possédant une finale consonantique [+ ANTÉRIEUR] occupent cette position pour 56,55% de ceux en [- ANTÉRIEUR] (respectivement $\chi^2(1) = 6,29$; $p \leq 0,01$ et $\chi^2(1) = 2,88$; $p \leq 0,09$).

Enfin, les mots ayant pour finale consonantique une finale [\pm HAUT] favorisent plutôt M1. Les mots se terminant par une consonne [- HAUT] sont 54,1% à sélectionner M1, pour 68% des mots possédant une finale consonantique [+ HAUT] (respectivement $\chi^2(1) = 7,45$; $p \leq 0,01$ et $\chi^2(1) = 3,24$; $p \leq 0,07$).

Le fait que ces paires de traits aient la même répartition pourrait nous indiquer que ces traits ne sont pas vraiment actifs dans la sélection d'une forme coordonnée. Notons cependant que

par paires de traits distinctifs, les validations statistiques n'ont pas le même poids. Pour l'ensemble des dix paires, le premier résultat fait état d'une répartition statistiquement significative ($p \leq 0,05$), le second, une tendance ($p \leq 0,10$). Cette différence laisserait entendre que les premiers mettent en valeur des facteurs plus prégnants et donc peut être plus actifs que les seconds. Un examen détaillé de ces résultats confrontés à ceux de l'ensemble des données exploitant les finales consonantiques nous permettra d'asseoir cette hypothèse.

Concernant le corpus VITE ET BIEN, la répartition des mots possédant une finale consonantique fait émerger une tendance statistique : les mots $_{[+ \text{HAUT}] \#}$ ont tendance à occuper M1 dans 68 % des mots testés sur cette hypothèse ($\chi^2 (1) = 3,24 ; p \leq 0,07$). Nous notons également un résultat pour lequel nous n'avons pas de validation statistique. Cependant, sa signification asymptotique étant juste au dessus du seuil .10, pour les raisons que nous avons précédemment évoquées, il paraît pertinent de le conserver. Ainsi les mots se terminant par une consonne $[- \text{CONTINU}]$ sont 62,86% à sembler favoriser M1 ($\chi^2 (1) = 2,31 ; p \leq 0,13$).

Enfin, le corpus ARTS ET MÉTIERS fait émerger une tendance statistique : conformément aux résultats que nous relevons dans le corpus TARN ET GARONNE et TOM ET ZOÉ, les mots en $[+ \text{CORONAL}]$ sont 68% à sélectionner M1 ($\chi^2 (1) = 3,24 ; p \leq 0,07$).

Revenons à présent sur les résultats issus des questionnaires d'enquête. Pour les raisons que nous évoquons précédemment page 222, nous ne pouvons les conserver. Même si les questionnaires d'enquête ÉLISE ET ZOÉ et ÉDOUARD ET EDMOND font apparaître des résultats statistiquement établis, nous pensons que les présenter ici revient à présenter un relevé artificiel de traits distinctifs. En effet, d'une part, dans le questionnaire ÉLISE ET ZOÉ, sur 17 prénoms se terminant par une finale consonantique en termes de tokens, nous n'examinons en fait que 12 types : le prénom « TOM » est présent dans trois paires, « HÉLÈNE » et « ÉLISE », dans deux. Examiner les traits des finales consonantiques du questionnaire d'enquête ÉLISE ET ZOÉ reviendrait en fait à examiner des traits des finales consonantiques biaisés par la surnumération de ces trois prénoms.⁷⁵ D'autre part, ÉDOUARD ET EDMOND et NAROD ET NANOR, en particulier, ont été construits afin d'explorer le rôle des finales consonantiques. Pour cela, nous avons confronté des paires présentant une

75 Nous avons songé exclure ces trois prénoms et proposer une nouvelle répartition en M1 et en M2 des finales consonantiques du questionnaire ÉLISE ET ZOÉ. Statistiquement, cela n'apportait rien. Les échantillons devenaient trop petits pour prétendre à une validation statistique.

finale consonantique et une finale vocalique, et des paires opposant deux consonnes finales différentes. Or, les seules paires pour lesquelles les interviewés ont massivement sélectionné une forme privilégiée, sont celles opposant les finales vocaliques aux finales consonantiques. Nous n'avons pu retenir les autres paires à l'exception de « NAROD ET NANOR », « GRACIELLE ET GRACIETTE » et « MERYL ET MERINE ». De fait, le seul résultat que produisent réellement ces deux questionnaires, porte sur le rôle du trait [+ CONSONANTIQUE] comme présenté page 195.

Nous notons cependant que la non-sélection de paires privilégiées pour les paires « _C# & _C# », nous semble être l'indice de ce que nous relevions précédemment de l'examen phonémique des finales consonantiques. Plus qu'une sélection par type de finales consonantiques, il semblerait que ce soit le trait [+ CONSONANTIQUE] qui prévale.

Du recueil par support d'études, concernant l'existence d'une corrélation entre la sélection d'une position dans la coordonnée et les traits des finales consonantiques, il apparaît que :

- Les mots se terminant par une finale consonantique privilégient majoritairement M1 ;
- Les mots possédant une finale consonantique [+ CORONAL] occupent majoritairement M1, tout comme les mots en [- ARRIÈRE] et ceux en [- CONTINU] ;
- Les mots ayant pour finale une consonne [± ANTÉRIEUR], [± VOISÉ], [± SONANT], [± NASAL], [± HAUT] ou [± REL RET] sélectionnent tous prioritairement M1.

En raison de ce dernier point et de son apparente ambiguïté, ainsi que de notre volonté de confirmer les résultats précédemment relevés, nous avons choisi d'observer les finales consonantiques pour l'ensemble des supports d'études. A nouveau, nous présenterons deux totaux différents. Le premier nommé TOTAL explore la totalité des 1 580 mots interrogeant le rôle des finales consonantiques présents dans notre corpus. Le second, nommé TOTAL-TOM ET ZOÉ, inventorie le même total sans les données issues du corpus TOM ET ZOÉ, soit 444 mots finissant par une finale consonantique.

Dans TOTAL, nous examinerons précisément les mots possédant une finale consonantique [+ CORONAL], [- ARRIÈRE] ou [- CONTINU] afin de confirmer les résultats précédemment. Dans TOTAL-TOM ET ZOÉ, nous interrogerons spécifiquement les cas ambigus que nous relevions, à savoir les mots se terminant par une consonne [± ANTÉRIEUR], [± VOISÉ], [± SONANT], [± NASAL], [± HAUT] ou [± REL RET].

TOTAL 1788	C#																	
	SONANT		CONTINU		NASAL		HAUT		ARRIERE		ANTERIEURE		CORONAL		VOISE		REL RET	
	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-
M1	539	330	699	170	296	573	76	793	193	676	659	210	604	265	638	231	403	466
	55,17%	54,73%	54,02%	59,44%	56,06%	54,47%	63,33%	54,32%	56,27%	54,65%	54,60%	56,30%	55,36%	54,19%	55,67%	53,23%	52,61%	57,25%
M2	438	273	595	116	232	479	44	667	150	561	548	163	487	224	508	203	363	348
	44,83%	45,27%	45,98%	40,56%	43,94%	45,53%	36,67%	45,68%	43,73%	45,35%	45,40%	43,70%	44,64%	45,81%	44,33%	46,77%	47,39%	42,75%
Total :	977	603	1294	286	528	1052	120	1460	343	1237	1207	373	1091	489	1146	434	766	814
	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

Tableau 70 : Répartition par traits distinctifs des finales consonantiques pour l'ensemble des supports d'études

TOTAL-TOM ET ZOÉ 368	C#																	
	SONANT		CONTINU		NASAL		HAUT		ARRIERE		ANTERIEURE		CORONAL		VOISE		REL RET	
	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-
M1	127	124	182	69	38	213	59	192	98	153	136	115	136	115	160	91	144	107
	54,04%	59,33%	54,33%	63,30%	55,88%	56,65%	62,11%	55,01%	56,00%	56,88%	56,90%	56,10%	57,38%	55,56%	55,17%	59,09%	53,93%	60,45%
M2	108	85	153	40	30	163	36	157	77	116	103	90	101	92	130	63	123	70
	45,96%	40,67%	45,67%	36,70%	44,12%	43,35%	37,89%	44,99%	44,00%	43,12%	43,10%	43,90%	42,62%	44,44%	44,83%	40,91%	46,07%	39,55%
Total :	235	209	335	109	68	376	95	349	175	269	239	205	237	207	290	154	267	177
	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

Tableau 71 : Répartition par traits distinctifs des finales consonantiques pour l'ensemble des supports d'études sans TOTAL-TOM ET ZOÉ

La confrontation des deux tableaux fait apparaître des résultats pour chaque trait distinctif.

Ainsi, concernant le trait [\pm SONANT], nous relevons trois résultats.

Dans TOTAL, 55,17% de mots possédant une finale consonantique [+ SONANT] occupent M1 pour 54,04% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 10,44$; $p \leq 0,001$ et $\chi^2(1) = 1,54$; $p \leq 0,22$). Ces résultats sont *a priori* semblables. Seule leur validation statistique diverge. TOTAL présente un résultat, ce qui n'est pas le cas de TOTAL-TOM ET ZOÉ. Cette différence est la preuve que le résultat énoncé dans TOTAL n'est essentiellement porté que par TOM ET ZOÉ. En effet, sur les 977 mots [+ SONANT] de TOTAL, 742 proviennent de TOM ET ZOÉ. Ayant choisi de ne pas valider de résultats issus d'un seul support d'études, pour les motifs que nous avons déjà invoqués, nous ne conservons pas ce résultat.

Un autre résultat au sujet du trait [\pm SONANT] émerge également de la confrontation des deux totaux. Dans TOTAL, 54,73% de mots finissant par une consonne [- SONANT] sélectionnent M1 pour 59,33% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 5,39$; $p \leq 0,02$ et $\chi^2(1) = 7,28$; $p \leq 0,01$). Ces résultats sont homogènes. Il apparaît qu'il existe une corrélation entre les mots se terminant par une finale consonantique [- SONANT] et M1.

La confrontation des deux totaux fait également émerger trois résultats concernant le trait [\pm CONTINU].

Dans TOTAL, 54,02% de mots finissant par une consonne [+ CONTINU] sélectionnent M1 pour 54,33% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 8,36$; $p \leq 0,004$ et $\chi^2(1) = 2,51$; $p \leq 0,11$). Ces résultats sont très semblables. Seule leur validation statistique diverge. TOTAL met en avant une répartition significative alors que la valeur-p de TOTAL-TOM ET ZOÉ est juste en-deçà du seuil de .10. Même si ce dernier mérite notre attention du fait du caractère graduel de la valeur-p, il est trop fragile. Ne pouvant le contrôler par d'autres tests, nous préférons ne pas le conserver en l'état. De fait, le résultat assignant M1 aux consonnes finales [+ CONTINU] n'étant porté que par le corpus TOM ET ZOÉ, nous ne pouvons l'affirmer.

Un second résultat lié au trait [\pm CONTINU] apparaît également. Conformément à ce que nous relevions dans LISSE ET SOYEUX, TOM ET ZOÉ et VITE ET BIEN, dans TOTAL, 59,44% de mots possédant une finale consonantique [- CONTINU] occupent M1 pour 63,3% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 10,2$; $p \leq 0,001$ et $\chi^2(1) = 7,72$; $p \leq 0,01$). Ces résultats sont homogènes. Il apparaît qu'il existe une corrélation entre les mots se terminant par une finale consonantique [- CONTINU] et M1.

Concernant le trait [\pm NASAL], nous relevons à nouveau trois résultats statistiquement significatifs de l'examen des deux totaux.

Dans TOTAL, 56,06% de mots possédant une finale consonantique [+ NASAL] occupent M1 pour 55,88% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 7,76$; $p \leq 0,005$ et $\chi^2(1) = 0,94$; $p \leq 0,33$). Les résultats sont très sensiblement identiques, exception faite de leur validation statistique. Les résultats présentés par TOTAL-TOM ET ZOÉ n'excluent pas l'hypothèse nulle. De fait, il apparaît que l'assignation en M1 des finales consonantiques [+ NASAL] est essentiellement portée par TOM ET ZOÉ ; nous ne conservons pas ce résultat.

Notons également la répartition des mots se terminant par une consonne [- NASAL]. Dans TOTAL, 54,47% de mots finissant par une consonne [- NASAL] sélectionnent M1 pour 56,65% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 8,4$; $p \leq 0,004$ et $\chi^2(1) = 6,65$; $p \leq 0,01$). Ces résultats sont homogènes. Il existe une corrélation entre les mots se terminant par une finale consonantique [- NASAL] et M1.

La confrontation des deux totaux fait également émerger trois résultats et une tendance statistique concernant le trait [\pm HAUT].

Dans TOTAL, 63,33% de mots finissant par une consonne [+ HAUT] sélectionnent M1 pour 62,11% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 8,53$; $p \leq 0,003$ et $\chi^2(1) = 5,57$; $p \leq 0,02$). Ces résultats sont identiques. Les mots se terminant par une consonne [+ HAUT] occupent M1.

Il apparaît également que les mots possédant une finale consonantique [- HAUT] sont 54,32% dans TOTAL et 51,01% dans TOTAL-TOM ET ZOÉ à occuper majoritairement M1 (respectivement $\chi^2(1) = 10,87$; $p \leq 0,001$ et $\chi^2(1) = 3,51$; $p \leq 0,06$). Ces résultats sont homogènes, exception faite de leur validation statistique. TOTAL présente un résultat, TOTAL-TOM ET ZOÉ, une tendance statistique. Les mots se terminant par une finale consonantique [- HAUT] sont prioritairement placés en M1.

Concernant le trait [\pm ARRIÈRE], nous relevons à nouveau trois résultats statistiquement significatifs de l'examen des tableaux TOTAL et TOTAL-TOM ET ZOÉ.

Dans TOTAL, 56,27% des mots finissant par une consonne [+ ARRIÈRE] sélectionnent M1 pour 56% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 5,39$; $p \leq 0,02$ et $\chi^2(1) = 2,52$; $p \leq 0,11$). Exception faite de leur validation, les deux tableaux mettent en valeur les mêmes résultats. Si TOTAL avance une répartition significative, la valeur-p de TOTAL-TOM ET ZOÉ est juste au-dessus du seuil de validation de .10. Comme nous l'indiquons

précédemment, même si ce dernier résultat est, selon nous, l'artefact d'une interaction de facteurs, du fait de sa fragilité, nous préférons ne pas le conserver. De fait, le résultat assignant M1 aux consonnes finales [+ ARRIÈRE] n'est essentiellement porté que par le corpus TOM ET ZOÉ. Nous ne le confirmons pas.

Nous notons également que conformément à ce que nous relevions dans TARN ET GARONNE et TOM ET ZOÉ, dans TOTAL, 54,65% des mots possédant une finale consonantique [- ARRIÈRE] occupent M1 pour 56,88% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 10,69$; $p \leq 0,001$ et $\chi^2(1) = 5,09$; $p \leq 0,02$). Ces résultats sont homogènes. Il apparaît qu'il existe une corrélation entre les mots se terminant par une finale consonantique [- ARRIÈRE] et M1.

La confrontation des deux totaux fait également émerger trois résultats significatifs et une tendance statistique concernant le trait [\pm ANTÉRIEUR].

Dans TOTAL, 54,6% des mots finissant par une consonne [+ ANTÉRIEUR] sélectionnent M1 pour 56,9% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 10,21$; $p \leq 0,001$ et $\chi^2(1) = 4,56$; $p \leq 0,03$). Ces résultats sont identiques. Les mots se terminant par une consonne [+ ANTÉRIEUR] occupent M1.

Il apparaît également que les mots possédant une finale consonantique [- ANTÉRIEUR] sont 56,3% dans TOTAL et 56,1% dans TOTAL-TOM ET ZOÉ à occuper majoritairement M1 (respectivement $\chi^2(1) = 5,92$; $p \leq 0,01$ et $\chi^2(1) = 3,05$; $p \leq 0,08$). Ces résultats sont identiques. Seule leur validation statistique diverge. TOTAL présente un résultat, TOTAL-TOM ET ZOÉ, une tendance statistique. Ces deux résultats sont néanmoins solides. Nous affirmons donc que les mots se terminant par une finale consonantique [- ANTÉRIEUR] sont prioritairement placés en M1.

Concernant le trait [\pm CORONAL], nous relevons à nouveau deux résultats significatifs et une tendance statistique de l'examen des tableaux TOTAL et TOTAL-TOM ET ZOÉ.

Dans TOTAL, 54,19% de mots finissant par une consonne [- CORONAL] ont tendance à sélectionner M1 pour 55,56% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 3,44$; $p \leq 0,06$ et $\chi^2(1) = 2,56$; $p \leq 0,11$). Exception faite de leur validation statistique, ces résultats sont très proches. TOTAL met en avant une tendance statistique alors que la valeur-p de TOTAL-TOM ET ZOÉ est juste en-deçà du seuil de .10. Ce dernier résultat est très fragile. D'autre part, la confrontation des deux totaux indique que l'hypothèse liant M1 aux consonnes finales [+ CORONAL] n'est en fait portée que par le corpus TOM ET ZOÉ. C'est

pourquoi nous ne pouvons la conserver.

Un second résultat lié au trait [\pm CORONAL] apparaît également. Conformément à ce que nous relevons dans TARN ET GARONNE, TOM ET ZOÉ et ARTS ET MÉTIERS, dans TOTAL, 55,36% des mots possédant une finale consonantique [+ CORONAL] occupent M1 pour 57,38% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 12,55$; $p \leq 0,0004$ et $\chi^2(1) = 5,17$; $p \leq 0,02$). Ces résultats sont très proches. Il existe donc une corrélation entre les mots se terminant par une finale consonantique [+ CORONAL] et M1.

La confrontation des deux totaux fait également émerger deux dispersions significatives et une tendance statistique [\pm VOISÉ].

Dans TOTAL, 55,67% des mots finissant par une consonne [+ VOISÉ] sélectionnent M1 pour 55,17% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 14,75$; $p \leq 0,001$ et $\chi^2(1) = 3,1$; $p \leq 0,08$). Ces résultats sont identiques. Seule leur validation statistique diverge légèrement. TOTAL présente une répartition significative, TOTAL-TOM ET ZOÉ, une tendance statistique. Il apparaît donc que les mots se terminant par une consonne [+ VOISÉ] occupent M1.

Notons également un résultat émergent dans TOTAL-TOM ET ZOÉ : les mots possédant une finale consonantique [- VOISÉ] sont 59,09% à occuper majoritairement M1 ($\chi^2(1) = 5,09$; $p \leq 0,02$). Nous affirmons donc que les mots se terminant par une finale consonantique [- VOISÉ] sont prioritairement placés en M1.

Enfin, nous relevons deux derniers résultats statistiquement significatifs concernant les mots possédant par une finale consonantique [- REL RET].

Dans TOTAL, 57,25% de mots finissant par une consonne [- REL RET] sélectionnent M1 pour 60,45% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 17,11$; $p \leq 0,00004$ et $\chi^2(1) = 7,73$; $p \leq 0,01$). Ces résultats étant très proches, nous affirmons donc que les mots se terminant par une consonne [- REL RET] occupent M1.

4.3.1.2.3. Bilan et analyse des résultats concernant les finales consonantiques

Au regard des résultats observés, de l'examen des mots se terminant par une finale consonantique, nous affirmons que :

- D'une manière générale, les mots possédant une finale consonantique occupent majoritairement, dans 56,53% des cas, M1 ;
- Dans le détail, l'examen des consonnes finales révèle que la répartition phonémique ne fait apparaître qu'un seul résultat : les mots en /k/# sélectionnent prioritairement M1.
- La répartition par traits distinctifs des finales consonantiques met en évidence le fait que :
 - Les mots se terminant par une consonne [- SONANT] (/p t k b d g f s ʃ v z ʒ/), et/ou [- CONTINU] (/p t k b d g/), et/ou [- NASAL] (/p t k b d g f s ʃ v z ʒ l R/), et ou [± HAUT], et/ou [- ARRIÈRE] (/p t b d f s ʃ v z ʒ l m n ɲ/), et/ou [± ANTÉRIEUR], et/ou [+ CORONAL] (/t d s ʃ z ʒ l n ɲ/), et/ou [± VOISÉ] et/ou [- REL RET] (/p t k b d m n ɲ ʁ/) ont tendance à être placés en M1 ;
 - Les mots se terminant par une consonne [+ HAUT] (/k g ʃ ʒ ɲ ʁ/) sélectionnent plus massivement M1 que ceux possédant une consonne finale [- HAUT] (/p t b d f s v z l m n R /);
 - Les mots possédant une finale consonantique [+ ANTÉRIEUR] (/p t b d f s v z l m n /) semblent plus occuper M1 que ceux se terminant par une consonne [- ANTÉRIEUR] (/ k g ʃ ʒ l ɲ R ʁ/);
 - Enfin que les mots ayant une finale consonantique [- VOISÉ] (/p t k f s ʃ l m /) sont plus souvent placés en M1 que ceux possédant une consonne finale [+ VOISÉ] (/ b d g v z ʒ l m n ɲ r R ʁ/).

Comment expliquer l'ensemble des résultats que nous venons de présenter ?

Plus que tout autre résultat, l'examen des finales consonantiques met au jour une forte cohésion générale entre les mots se terminant par une finale consonantique et la position M1. Nous pouvons supposer que cette corrélation est imputable à un principe fréquemment observé dans l'organisation des segments phonologiques en syllabes (Labrune 2005) : le

principe de maximisation des attaques. Ce principe est exprimé en Théorie de l'Optimalité par Onset >> *Coda.

Onset : toute syllabe possède une attaque.

*Coda : les syllabes n'ont pas de codas.

La maximisation des attaques formalise la tendance à ce qu'une consonne finale se syllabe comme attaque de la voyelle suivante.

Ainsi, si l'on observe nos résultats au regard de ce principe général, en plaçant prioritairement les mots se terminant par une syllabe fermée en M1, d'une part, nous ne transgressons pas *Coda, et d'autre part, comme indiqué par la maximisation des attaques, nous favorisons la syllabation de « ET ». En effet, la coda de la syllabe finale fermée devient attaque d'une nouvelle syllabe dont le noyau est /e/.

Cependant, si la préférence des mots se terminant par une consonne pour M1 semble motivée par ce principe, il n'explique pas pourquoi l'examen phonémique des finales consonantiques révèle que seuls les mots en . _/k/# favorisent M1.

Quels sont les autres facteurs en jeu ? De même, comment interpréter les corrélations entre M1 et les finales consonantiques [- SONANT], [- CONTINU], [- NASAL] [± HAUT], [- ARRIÈRE], [± ANTÉRIEUR], [+ CORONAL], [± VOISÉ] ou [- REL RET] dégagées de l'examen des finales consonantiques par traits ? Enfin, comment expliquer que la sélection de M1 soit statistiquement plus forte pour les mots finissant par une finale consonantique spécifiée [+ HAUT], [+ ANTÉRIEUR] ou [- VOISÉ] que pour ceux spécifiés [- HAUT], [- ANTÉRIEUR] ou [- VOISÉ] ?

Concernant la répartition par traits distinctifs des finales consonantiques, on peut rendre compte de la préférence des mots se terminant par une consonne pour la position M1 en invoquant à nouveau l'opposition marqué / non marqué présentée précédemment. Cette opposition tend à placer l'élément le moins marqué dans la position la moins marquée, à savoir M1.

Nous retenons ici les définitions fonctionnelles universelles de la marque de Jakobson *et al.* (1952) et de Clements (2007) proposant qu'« au sein de toute classe de sons dans laquelle un trait est potentiellement distinctif, la valeur marquée de ce trait est celle qui est absente de certaines langues, et la valeur non-marquée est celle qui est présente dans toutes les langues ».

La valeur la moins fréquente sera considérée comme marquée.

Ainsi, reprenant les critères de Clements, les consonnes [- SONANT], [- CONTINU], [- NASAL], [+ CORONAL]⁷⁶, [- ARRIÈRE], ou [- REL RET] sont moins marquées.

Dans la lignée de Renner (2006) citant Kelly, Bock et Keil (1986), « à fréquence et à nombre de syllabes égaux, la différence de typicalité entre les éléments d'un binôme formé en discours influe sur leur ordre linéaire : l'élément prototypique précède généralement l'élément non prototypique ». Il est donc tout à fait envisageable que le caractère non marqué de ces traits puisse déterminer l'occupation préférentielle des mots se terminant par une consonne en M1.

Cependant, si on peut expliquer la prédisposition des mots se terminant par une finale consonantique [- SONANT], [- CONTINU], [- NASAL], [+ CORONAL], [- ARRIÈRE], [- VOISÉ] ou [- REL RET] à occuper M1, par la préférence des éléments moins marqués à occuper une position non marquée, cette opposition n'explique pas pourquoi les mots possédant une finale consonantique [± HAUT], [± ANTÉRIEUR], ou [± VOISÉ] favorisent tous, qu'ils soient marqués ou non, cette même position. D'autant plus que ces résultats semblent a priori contradictoires.

A cela, nous trouvons deux explications potentielles :

- Les valeurs [± HAUT], [± ANTÉRIEUR], ou [± VOISÉ] ne sont pas actives dans la sélection de l'ordre préférentiel de coordination de deux mots par « et » ;
- Ou alors, d'autres facteurs interagissent et obscurcissent nos résultats.

Notons tout d'abord qu'il nous faut nuancer ces résultats.

Certes, les mots se terminant par une consonne [± HAUT], [± ANTÉRIEUR], ou [± VOISÉ] occupent majoritairement M1. Cependant, comme nous l'indiquons plus haut, les mots en $_ [+ HAUT] \#$ sélectionnent plus massivement M1 que ceux possédant une consonne finale [- HAUT], tout comme les mots $_ [+ ANTÉRIEUR] \#$ par rapport aux $_ [- ANTÉRIEUR] \#$ et les mots se terminant par une finale consonantique [- VOISÉ] par rapport à ceux se terminant par une consonne [+ VOISÉ].

⁷⁶ Le trait de lieu [CORONAL] est fréquemment non marqué par opposition aux traits [DORSAL] et [LABIAL] (Paradis et Prunet (1991), Clements (2007), Yamaguchi (2015)). Comme nous l'indiquons page 29, nous considérons que la valeur marquée de ce trait unaire [CORONAL] correspond au trait binaire [+ CORONAL] que nous utilisons. Ainsi, même si ce trait porte la valeur « + », le trait [+ CORONAL] est le trait non marqué.

En effet, les résultats associant les finales consonantiques [-ANTÉRIEUR], [-HAUT] et [+VOISÉ] à la première position de la coordination sont des tendances ($0,05 < p \leq 0,1$), alors que ceux liant les finales [-ANTÉRIEUR], [+HAUT] et [+VOISÉ] à M1 sont des résultats statistiquement significatifs ($p \leq 0,05$).

Nous pensons que les tendances relevées pour les valeurs [-HAUT], [-ANTÉRIEUR], et [+VOISÉ] ne sont finalement que des artefacts du principe de maximisation des attaques. En effet, les mots se terminant par une consonne possédant ces valeurs-ci ont plus tendance à être en tête de coordonnée, car ils se terminent par des consonnes et favorisent la syllabation.

Nous prêterons donc une attention plus particulière aux valeurs [+HAUT], [+ANTÉRIEUR] et [-VOISÉ].

Concernant les valeurs [+ANTÉRIEUR] et [-VOISÉ], nous invoquons à nouveau la conjonction du principe de maximisation des attaques et la valeur asymétrique marqué / non marqué. Nous considérons que les consonnes [-VOISÉ] sont moins marquées que les consonnes [+VOISÉ], tout comme les [+ANTÉRIEUR] par rapport aux [-ANTÉRIEUR].⁷⁷ Les finales consonantiques [+ANTÉRIEUR] et [-VOISÉ] favorisent donc la syllabation de « ET » en respectant la contrainte Onset évoquée plus haut. De plus, étant moins marquées, elles occupent prioritairement la position la moins marquée : M1.

Cependant, à nouveau, cette conjonction de principes, même si elle motive vraisemblablement la présence en M1 des mots se terminant par une finale consonantique [+ANTÉRIEUR] ou [-VOISÉ], n'explique en rien les raisons pour lesquelles les mots possédant une consonne finale [+HAUT] dont ceux en $_ /k/\#$ sélectionnent également M1.

Il est important de noter que sur les 95 mots testés possédant une finale consonantique [+HAUT], plus de 61% d'entre eux se terminent par $/k/$. Les résultats associant les mots se terminant par une consonne finale [+HAUT] à M1 ne sont en fait portés que par les mots en $_ /k/\#$. En fait, pour rendre compte de la préférence des mots se terminant par une consonne [+HAUT] à occuper M1, il faut rendre compte de la préférence des mots en $_ /k/\#$ pour cette même position.

Pour tenter de trouver une explication à ceci, nous pouvons supposer l'action des hiérarchies des attaques formalisant les aptitudes des consonnes à occuper les fonctions d'attaque ou de coda, à savoir la HIÉRARCHIE DES MARGES (ou Margin Hierarchy) et la HIÉRARCHIE

⁷⁷ Comme nous l'indiquons page 29, rappelons que nous considérons que le trait [ANTÉRIEUR] comme l'opposé du trait [\pm POSTÉRIEUR] définit par Clements.

DES CODAS. Nous reproduisons ci-dessous la forme que leur donne Plénat (1999) :

HIÉRARCHIE DES MARGES

*N/a >> *N/i >> *N/r >> *N/n *N/z >> *N/s >> *N/d >> *N/t

HIÉRARCHIE DES CODAS

*C/t >> *C/d >> *C/s >> *C/z >> *C/n >> *C/r >> *C/i >> *C/a

La HIÉRARCHIE DES MARGES suggère le principe selon lequel, dans une gradation strictement opposée à celle de l'échelle de sonorité de Grammont (1933) (présentée en annexe), plus un segment phonologique est bas sur l'échelle de sonorité, plus il peut prétendre être marge de syllabe.

La HIÉRARCHIE DES CODAS, quant à elle, formalise le fait que les phonèmes [+SONANT] sont plus aptes à occuper la fonction de codas que les consonnes [-SONANT].

Comme le rappelle Plénat (1999), « il est préférable qu'une attaque soit momentanée plutôt que continue et obstruante plutôt que sonante ; toutes choses égales par ailleurs, il vaut mieux qu'une coda soit une sonante plutôt qu'une obstruante et une continue plutôt qu'une momentanée ».

Ainsi, les consonnes occlusives, placées très bas dans la HIÉRARCHIE DES MARGES et très haut dans la HIÉRARCHIE DES CODAS, seront de « bonnes » attaques, et de « mauvaises » codas. Inversement le /R/, placé très haut dans la HIÉRARCHIE DES MARGES et très bas dans la HIÉRARCHIE DES CODAS, fera une « meilleure » coda.

Cette conjonction de hiérarchies est illustrée par deux résultats.

Les mots se terminant par l'occlusive sourde /k/, étant potentiellement de « mauvaises » codas mais de « bonnes » attaques, tombent aisément sous les effets conjoints de la HIÉRARCHIE DES MARGES, de la HIÉRARCHIE DES CODAS et du principe de maximisation des attaques. En occupant M1, les codas /k/ deviennent attaques d'une nouvelle syllabe dont /e/ (« ET ») est le noyau.

Inversement, concernant les mots se terminant par /R/, il apparaît que *Coda et le principe de maximisation des attaques entrent en conflit avec la HIÉRARCHIE DES CODAS. Les mots en /R/# semblent plus prompts à violer *Coda. Selon nous, ces tensions obscurcissent la lisibilité d'une assignation préférentielle des mots en /R/# pour l'une des deux positions de la

coordonnée. L'observation des 110 mots $_ /R/\#^{78}$ ne nous a, en effet, pas permis de faire émerger une position privilégiée dans la coordination.

Cependant, si la conjonction des hiérarchies présentées ci-dessus et le principe de maximisation des attaques proposent une motivation vraisemblable à ce que les mots en $_ /k/\#$ privilégient M1, il n'explique en rien pourquoi seuls les mots se terminant par la finale consonantique $/k/$ occupent M1.

Pourquoi, par phonèmes, les autres finales occlusives ne fournissent-elles pas les mêmes résultats ?

On pourrait supposer l'action d'un paramètre lié à la fréquence de réalisation des phonèmes en français.

D'une manière générale, la fréquence des codas relevées en finale de mots dans nos supports d'études est conforme à celle proposée pour le français par Wioland (1985), à savoir par ordre décroissant : les liquides, les occlusives sourdes, les consonnes nasales, les fricatives, les occlusives sonores et les glides. En finale de mots, il y a proportionnellement plus de codas liquides ou rhotiques que de codas occlusives, et plus de codas occlusives sourdes que de codas occlusives sonores.

Ainsi, dans TOTAL-TOM ET ZOÉ, 167 mots se terminent par $/R/$ ou $/l/$, pour 95 par une occlusive sourde ($_ /p/\# = 2$ mots, $_ /t/\#, 35$ et $_ /k/\#, 58$) et 14 par une occlusive sonore ($_ /b/\# = 3$ mots, $_ /d/\#, 6$ et $_ /g/\#, 5$).

De fait, comme attendu compte tenu de la fréquence des phonèmes en position de coda, nous ne relevons pas de résultats pour les consonnes occlusives sonores par manque de données éprouvées. Le faible effectif des mots en $_ /p/\#$ ne nous permet pas non plus d'extraire une répartition privilégiée.

En revanche, les mots se terminant $_ /k/\#$ ou en $_ /t/\#$ étant relativement fréquents en français (et dans nos supports d'études), nous sommes en mesure de dégager des résultats.

Ainsi, concernant l'examen des mots en $_ /t/\#$, même si nous ne relevons pas de répartitions statistiquement significatives, nous observons que conformément aux prescriptions conjointes des HIÉRARCHIE DES MARGES, de la HIÉRARCHIE DES CODAS et de la maximisation des attaques, 60% des mots en $_ /t/\#$ occupent M1 ($\chi^2(1) = 1,4$; $p \leq 0,24$). Ce résultat est d'autant plus pertinent que l'observation détaillée des onze paires semblant violer cette conjonction de contraintes révèle que deux d'entre elles coordonnent en fait deux mots en

78 Dans TOTAL-TOM ET ZOÉ, 50,91% des 110 mots $_ /R/\#$ sélectionnent M1, pour 49,09%, M2 ($p < 0,1$).

_/t/# (« HYDRATANTE ET APAISANTE » et « HYDRATANTE ET ADOUCISSANTE »). Ces deux formes exclues, en adéquation avec les résultats que nous présentions ci-haut, les mots en _/t/# ont tendance à sélectionner M1 dans 67,86 % des occurrences ($\chi^2(1) = 3,57$; $p \leq 0,06$)⁷⁹. La conjonction de la HIÉRARCHIE DES MARGES, de la HIÉRARCHIE DES CODAS et du principe de maximisation des attaques semble donc proposer une motivation vraisemblable à ce que les mots se terminant par une occlusive non voisée occupent préférentiellement M1.

4.3.1.3. Glide

Revenons à présent sur les résultats collectés pour les mots se terminant par un glide.

Comme nous le présentions page 192, nous avons dégagé de l'ensemble des six corpus examinés sur cette question, deux résultats *a priori* semblables :

- Dans TOTAL⁸⁰, 73,91% des 46 mots en _G# présents dans nos corpus occupent massivement M2 ($\chi^2(1) = 10,52$; $p \leq 0,001$) ;
- Dans TOTAL-TOM ET ZOÉ⁸¹, sur 10 mots en _G# testés, 70 % sélectionnent préférentiellement M2 ($\chi^2(1) = 1,6$; $p \leq 0,21$).

La répartition est homogène, pas la validation statistique.

Que penser de ces résultats ?

Il apparaît que dans les faits, la préférence des mots en _G# à occuper M2 n'est portée que par TOM ET ZOÉ. Or, si les 36 paires observées dans ce corpus en termes de tokens semblent cohérentes et suffisantes pour nous permettre de dégager des résultats, en termes de type, les résultats proviennent d'un seul prénom : « CAMILLE ». Il paraît impensable de conserver ces résultats, même si intuitivement, et en adéquation avec l'ensemble des résultats dégagés de la répartition des segments finaux, nous pensons que les mots se terminant par un glide occuperaient plutôt M2.

En effet, les mots en _/j/# (seul glide à occuper la position de coda) sont des phonèmes [+

79 À nouveau, nous sommes convaincue que les tendances statistiques signalent le jeu d'interaction de facteurs.

80 TOTAL recueille l'ensemble des données examinant l'hypothèse d'une corrélation entre l'une des deux positions de la coordonnée et les mots se terminant par un glide.

81 TOTAL-TOM ET ZOÉ regroupe le même ensemble que TOTAL duquel on a extrait les données fournies par le corpus TOM ET ZOÉ. Ce dernier étant le corpus le moins contrôlé de nos supports d'études, nous avons systématiquement proposé de confronter ses résultats émergents à l'ensemble des résultats.

SONANT], [- CONSONANTIQUE], [+ CONTINU], [+ HAUT], [+ VOISÉ] et [+ REL RET]. Au regard de l'opposition marqué / non marqué et de la HIÉRARCHIE DES CODAS que nous présentions précédemment, nous constatons que non seulement ces traits sont non marqués, mais en plus, /j/ est assez bas dans la HIÉRARCHIE DES CODAS présentée page 241. La conjonction de cet ensemble de principes favoriserait la présence des mots se terminant par un glide en M2.

Néanmoins, compte tenu de la petite taille de l'échantillon testé et de l'absence de résultats significatifs, nous ne pouvons nous risquer en conjecture. Nous envisageons un examen ultérieur sur de plus vastes données afin de vérifier nos hypothèses à ce propos.

4.3.2. Syllabe finale

L'intérêt que nous portons à l'examen des syllabes finales est triple.

Tout d'abord, l'opposition marqué / non marqué et le principe de maximisation des attaques dont nous soupçonnons les effets dans la sélection d'un ordre préférentiel de coordination nous mènent à nous interroger sur la structure de la syllabe finale. En effet, la maximisation des attaques Onset >> *Coda propose que la structure de la syllabe non marquée soit CV. Or si CV est non marqué, peut-on imaginer que les syllabes finales en M2, en position marquée, aient un *pattern* plus marqué que CV ?

De plus, nous questionner sur le lien existant entre le type de segment final et la sélection d'une forme coordonnée nous conduit également à observer le timbre des syllabes finales et précisément de celles qui violent *Coda. Quel est le timbre des finales se terminant par une consonne ou par un glide ? Peut-on imaginer un effet de la HIÉRARCHIE DES NOYAUX semblable à celui que nous relevions page 217 ? Le timbre nucléique est-il actif dans la sélection d'un ordre préférentiel de coordination binomial ?

Enfin, la conjonction du principe *VV et du principe de maximisation des attaques a-t-elle une incidence sur la réalisation ou non des consonnes de liaisons ? Les paires potentiellement liaisonnantes que nous avons retenues tombent-elles sous les mêmes principes que les paires non liaisonnantes ? Dans quelle mesure la liaison est-elle un paramètre dans la sélection d'un ordre préférentiel de coordination ?

Pour répondre à ces questionnements, nous nous proposons d'explorer dans les paragraphes à suivre la répartition en M1 et en M2 du timbre des noyaux des syllabes finales fermées, puis nous examinerons plus précisément la structure des syllabes finales ; enfin nous relèverons la

répartition des formes liaisonnantes en M1 et en M2.

4.3.2.1. timbre de la syllabe finale

Dans ce paragraphe, nous présentons la répartition en M1 et en M2 du noyau vocalique des syllabes finales fermées, à savoir le noyau des mots se terminant par une consonne ou un glide.

Pour cela, nous exploitons les six corpus et les questionnaires d'enquête ÉLISE ET ZOÉ, ÉDOUARD ET EDMOND et NAROD ET NANOR. Les autres questionnaires d'enquête ont été mis de côté, ne traitant pas la question du type de finale, les appariements testés étaient neutralisés sur cette question. Au total, sur l'hypothèse d'une corrélation entre le timbre du noyau des syllabes finales fermées et l'une des deux positions de la coordonnée, nous examinons 1 629 items.

Observons tout d'abord les répartitions en M1 et en M2 du noyau vocalique des syllabes finales fermées.

Dans les deux tableaux ci-dessous, nous comparons deux totaux.

Le total obtenu pour l'ensemble des 1 629 items possédant une syllabe finale fermée est présenté dans TOTAL.

Dans TOTAL-TOM ET ZOÉ, nous proposons le même ensemble amputé des paires issues de TOM ET ZOÉ, soit un total de 457 mots. En effet, comme nous le rappelons régulièrement, à cause de sa taille et de la méthodologie mise en place pour recueillir ses données, le corpus TOM ET ZOÉ est le moins contrôlé de nos supports d'études. D'autres facteurs que ceux testés pourraient biaiser nos résultats. C'est pourquoi, bien que possédant des qualités indéniables nous ayant permis entre autres de mettre en place nos questionnaires d'enquêtes, nous effectuons systématiquement une vérification supplémentaire des résultats en émergeant. Nous avons par ailleurs étendu cette vérification à l'ensemble des résultats que nous relevons : d'une manière générale, aucun résultat n'est conservé s'il est porté par un unique corpus et s'il n'est pas confirmé par l'ensemble des composants étudiés.

Ainsi, les deux tableaux ci-dessous présentent la répartition en M1 et en M2 des noyaux des syllabes finales fermées, d'une manière générale et en fonction de leur trait de nasalité.

A nouveau, nous avons noté en caractères gras les résultats excluant l'hypothèse nulle pour lesquels nous relevons une répartition statistiquement significative, c'est-à-dire ceux pour lesquels la valeur du Khi Deux est inférieure à .05.

TOTAL 1629	_V_#		
	_ NASAL	+ NASAL	TOTAL
M1	824 54,57%	58 48,74%	882 54,14%
M2	686 45,43%	61 51,26%	747 45,86%
Total :	1510 100,00%	119 100,00%	1629 100,00%

Tableau 72 : répartition en _V_# pour l'ensemble des supports d'études

TOTAL-TOM ET ZOÉ 457	_V_#		
	_ NASAL	+ NASAL	TOTAL
M1	230 55,83%	25 55,56%	255 55,80%
M2	182 44,17%	20 44,44%	202 44,20%
Total :	412 100,00%	45 100,00%	457 100,00%

Tableau 73 : répartition en _V_# pour l'ensemble des supports d'études sans TOM ET ZOÉ

La confrontation des totaux fait apparaître deux résultats extrêmement liés : les mots possédant pour noyau de syllabes finales fermées une voyelle [- NASAL], et ceux qui, d'une manière générale, se terminent par une syllabe fermée, occupent majoritairement M1 dans respectivement 55,83 % et 55,80% des items testés sur ces hypothèses ($\chi^2(1) = 10,52$; $p \leq 0,001$ et $\chi^2(1) = 10,52$; $p \leq 0,001$)).

Comment analyser ces résultats ?

Le premier résultat, concernant les noyaux vocaliques des syllabes finales fermées, corrobore ce que nous relevions pages 237. En effet, comme le rapporte Labrune (2005), en français, toute syllabe s'organise autour d'un noyau vocalique⁸².

Compte tenu de ce principe, il est attendu que les syllabes fermées possèdent un noyau vocalique et donc, par extension, et conformément à nos précédent résultats, les mots

⁸² « Toute syllabe s'organise autour d'une voyelle », comme le rappelle l'auteure, exception faite de « l'interjection *psst*, par exemple, qui forme une syllabe sans voyelle ». Il est utile de préciser que les syllabes sans voyelle sont absentes des supports d'études que nous utilisons.

possédant un noyau vocalique de syllabes finales fermées suivent la même répartition que les mots se terminant par une finale consonantique⁸³. Cependant, à ce stade de nos travaux, si nous pouvons justifier la présence en M1 des mots en *_V_#*, nous ne savons si leur présence est uniquement induite par la finale consonantique, par la structure de la syllabe finale ou si le timbre vocalique a une incidence dans la sélection d'un ordre préférentiel de coordination. Nous tenterons d'apporter des réponses à cette question en 4.3.2.1.3. Bilan et analyse des résultats de la répartition des noyaux de syllabes finales fermées page 257.

Nous notons également que les noyaux vocaliques [- NASAL] sélectionnent majoritairement M1. Selon nous, cette prédisposition à placer en M1 les mots dont le noyau vocalique des syllabes finales fermées est [- NASAL] se réclame à nouveau de l'opposition marqué / non marqué présentée précédemment. En effet, cette opposition favorise l'élément le moins marqué dans la position la moins marquée. Or, suivant les définitions universalistes de la marque de Jakobson *et al* (1952) et de Clements (2007), l'inventaire des langues du monde révèle que si toutes les langues possèdent des voyelles [- NASAL], celles possédant des voyelles [+ NASAL] sont plus rares. De plus, aucune langue dont le système vocalique autorise les voyelles [+ NASAL] n'interdit les voyelles [- NASAL]. La réciproque n'est pas vraie. Les voyelles [+ NASAL] sont donc marquées par rapport aux voyelles [- NASAL]. Cette asymétrie rend vraisemblablement compte de la présence des mots possédant un noyau vocalique [- NASAL] de syllabes finales fermées en M1, étant donnée que cette position est la moins marquée.

En fait, la question que nous nous posons ici est de savoir si le type, par phonème ou par trait, de ces noyaux de syllabes finales fermées agit dans la sélection d'un ordre préférentiel de coordination.

Pour cela, nous recueillerons tout d'abord les répartitions phonémiques des noyaux vocaliques oraux, puis des noyaux vocaliques [+ NASAL]. Nous examinerons, dans une deuxième étape, ces mêmes répartitions en fonction des traits distinctifs des noyaux vocaliques des syllabes finales fermées que nous avons relevés. Enfin, nous proposerons un bilan et une analyse des résultats.

83 Nous ne considérons pas ici les mots se terminant par un glide. Nous n'avons pu relever de résultats les concernant. De plus, leur faible nombre (seulement dix mots) ne modifie en rien les présents résultats.

4.3.2.1.1. Répartition phonémique du noyau des syllabes finales fermées

Afin d'explorer plus précisément la répartition par timbre vocalique des noyaux des syllabes finales fermées entre M1 et M2, nous nous sommes penchée, tout d'abord, sur l'observation des voyelles [- NASAL] que nous nommerons dans ce paragraphe *_Vorale_#*. L'examen phonémique des noyaux vocaliques [+ NASAL] des syllabes finales fermées, nommés ici *_Vnasale_#* sera exposé ultérieurement.

Nous nous proposons donc d'explorer la répartition générale des mots *_Vorale_#* pour chaque support d'études. Comme précédemment les répartitions statistiquement significatives sont mentionnées en caractères gras, et les tendances, en italique.

Le tableau présenté en annexe tableau A 13 ne révèle que très peu de données mettant au jour une corrélation avec l'une des deux positions de la coordonnée et *_Vorale_#*. Nous ne relevons en effet que trois résultats et deux tendances statistiques. Or, sur ces cinq hypothèses, deux seulement peuvent être retenues ; les trois autres étant calculées sur moins de cinq items. Il apparaît donc de l'examen de la répartition phonémique en M1 et en M2 de *_Vorale_#*, seul le corpus TOM ET ZOÉ nous permet de dégager deux résultats : les mots en *_/i/_#* et en *_/y/_#* occupent majoritairement M1 dans respectivement 56,33% et 62,5% des occurrences testant ces hypothèses ($p \leq 0,05$).

Il est entendu que bien que ces résultats soient statistiquement fiables, n'étant portés que par un seul corpus, nous ne pouvons les valider en l'état. C'est pourquoi dans le paragraphe à venir, nous étudierons la répartition en M1 et en M2 des mots en *_Vorale_#* pour l'ensemble de nos supports d'études. Nous présenterons ces résultats dans le tableau récapitulatif TOTAL.

De plus, afin d'asseoir les hypothèses relevées dans TOM ET ZOÉ, à savoir le positionnement privilégié en M1 des mots *_/i/_#* et en *_/y/_#*, nous confronterons TOTAL à un nouvel ensemble nommé TOTAL-TOM ET ZOÉ.

TOTAL-TOM ET ZOÉ inventorie la répartition de l'ensemble des mots en *_Vorale_#* sans ceux fournis par le corpus TOM ET ZOÉ. La confrontation de ces deux totaux nous offre la possibilité d'examiner un vaste panel de composés coordinatifs par « et » tout en nous dédouanant autant que possible d'autres facteurs dont TOM ET ZOÉ pourrait être le vecteur.

TOTAL 1510	_Vorale_ #											
	i	e	ε	a	ɑ	y	ø	œ	ə	u	o	ɔ
M1	329 56,92%		144 52,55%	163 53,62%	1 100,00%	52 59,77%	0 0,00%	12 50,00%		10 62,50%	17 34,69%	96 55,17%
M2	249 43,08%		130 47,45%	141 46,38%	0 0,00%	35 40,23%	3 100,00%	12 50,00%		6 37,50%	32 65,31%	78 44,83%
Total :	578 100,00%	0 100,00%	274 100,00%	304 100,00%	1 100,00%	87 100,00%	3 100,00%	24 100,00%	0 100,00%	16 100,00%	49 100,00%	174 100,00%

Tableau 74 : Répartition phonémique de _Vorale_ # pour TOTAL

TOTAL-TOM ET ZOË 412	_Vorale_ #											
	i	e	ε	a	ɑ	y	ø	œ	ə	u	o	ɔ
M1	52 <i>59,09%</i>		61 64,21%	55 50,00%	1 100,00%	12 52,17%	0 0,00%	12 50,00%		10 62,50%	3 20,00%	24 64,86%
M2	36 40,91%		34 35,79%	55 50,00%	0 0,00%	11 47,83%	3 100,00%	12 50,00%		6 37,50%	12 80,00%	13 35,14%
Total :	88 100,00%	0 100,00%	95 100,00%	110 100,00%	1 100,00%	23 100,00%	3 100,00%	24 100,00%	0 100,00%	16 100,00%	15 100,00%	37 100,00%

Tableau 75 : Répartition phonémique de _Vorale_ # pour TOTAL-TOM ET ZOË

Les deux tableaux ci-dessus font apparaître quatre résultats notés en caractères gras et cinq tendances statistiques (en italique) qu'il nous faut commenter.

Ainsi, concernant les mots en *_/ø/_#*, nous relevons deux tendances statistiques concernant leur répartition privilégiée en M2. Cependant les deux répartitions n'étant mesurées que sur trois items, nous ne pouvons conserver ce résultat.

Revenons à présent sur les résultats émergeant du corpus TOM ET ZOË, à savoir la prédisposition des mots en *_/i/_#* et en *_/y/_#* à occuper M1. Les deux tableaux mettent au jour un résultat statistiquement significatif et deux tendances.

En effet, il apparaît que les mots en *_/i/_#* sont 56,92% dans TOTAL, et 59,02% dans TOTAL-TOM ET ZOË à occuper majoritairement M1 (respectivement $\chi^2(1) = 11,07$; $p \leq 0,001$ et $\chi^2(1) = 2,91$; $p \leq 0,09$). Ces résultats sont identiques. Seule leur validation statistique diverge. TOTAL présente un résultat, tandis que TOTAL-TOM ET ZOË avance

une tendance statistique. Ils demeurent néanmoins solides. C'est pourquoi nous affirmons l'existence d'une corrélation entre M1 et les mots en $_/i/_#$.

Les mots en $_/y/_#$, quant à eux, sont 59,77% dans TOTAL, pour 52,17% dans TOTAL-TOM ET ZOÉ à sélectionner préférentiellement M1 (respectivement $\chi^2(1) = 3,32$; $p \leq 0,07$ et $\chi^2(1) = 0,04$; $p \leq 0,83$). Exception faite de leur validation statistique, ces résultats semblent *a priori* semblables : TOTAL nous permet de dégager une tendance statistique, ce qui n'est pas le cas de TOTAL-TOM ET ZOÉ. Cette différence notable illustre le fait que TOM ET ZOÉ porte seul ce résultat. En effet, sur les 87 mots en $_/i/_#$ de TOTAL, 64 proviennent de TOM ET ZOÉ. La confrontation des deux totaux au sujet de l'hypothèse d'une corrélation entre M1 et les mots en $_/i/_#$ révèle la fragilité de résultats pourtant apparemment identiques. De fait, ayant choisi de ne pas valider de résultats issus d'un unique support d'études, nous ne les conservons pas.

La confrontation des deux totaux ci-dessus met également au jour trois nouvelles hypothèses concernant l'assignation préférentielle dans l'une des deux positions de la coordonnée des mots en $_/o/_#$, en $_/ε/_#$ et en $_/ɔ/_#$. Les deux dernières hypothèses révèlent une assignation en M1, comme précédemment, la première, une assignation en M2.

Concernant les mots en $_/o/_#$, nous relevons deux résultats.

Dans TOTAL, 65,31% de mots en $_/o/_#$ sélectionnent M2, pour 80% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 4,59$; $p \leq 0,03$ et $\chi^2(1) = 5,4$; $p \leq 0,02$). Ces résultats sont homogènes. Il apparaît qu'il existe une corrélation entre les mots en $_/o/_#$ et M2.

D'autre part, le tableau TOTAL-TOM ET ZOÉ fait émerger deux nouvelles hypothèses.

Les mots en $_/ε/_#$ privilégient majoritairement M1 dans 64,21% des cas ($\chi^2(1) = 7,67$ $p \leq 0,01$). Ce résultat étant extrait de l'ensemble de nos données sans celles fournies par le corpus TOM ET ZOÉ, nous attestons donc l'existence d'un lien plaçant majoritairement les mots en $_/ε/_#$ en M1.

Enfin, apparaît dans TOTAL-TOM ET ZOÉ une tendance statistique dans laquelle les mots en $_/ɔ/_#$ ont tendance à favoriser M1 dans 64,86% des cas ($\chi^2(1) = 3,27$ $p \leq 0,07$). Ce résultat n'est pas statistiquement aussi fort que le précédent. Rappelons néanmoins que convaincue du

jeu d'interactions de facteurs ou de principes dans la sélection d'un ordre préférentiel de composés syndétiques, il est vraisemblable que la faiblesse de cette validation statistique ne soit que l'artefact d'un conflit de contraintes. C'est pourquoi, même s'il ne semble pas aussi solide que nous le souhaiterions, nous conservons ce résultat et affirmons qu'il existe une tendance plaçant prioritairement en M1, les mots en $_ / \text{ɔ} / _ \#$.

Au regard de la confrontation de TOTAL et TOTAL-TOM ET ZOÉ concernant l'hypothèse d'un ordre préférentiel de coordination en lien avec le type phonémique des noyaux vocaliques oraux des syllabes finales fermées, il apparaît que :

- Les mots en $_ / \text{ɛ} / _ \#$ sélectionnent plutôt M1 ;
- Tout comme ceux en $_ / \text{i} / _ \#$ ou en $_ / \text{ɔ} / _ \#$ pour lesquels nous constatons la même tendance ;
- En revanche, les mots en $_ / \text{o} / _ \#$ occupent majoritairement M2.

Nous commenterons ces résultats en 4.2.1.2.3. Bilan et analyse des résultats concernant #C_ page 178 et ceux que nous relevons concernant la répartition phonémique en M1 et en M2, de $_ \text{Vnasale} _ \#$ à venir .

Dans un premier temps, étant donné que nous avons construit un ensemble de supports d'études volontairement hétéroclites afin de balayer le plus largement possible la coordination binaire directe de deux mots par « et », nous avons souhaité présenter les résultats, par support d'études, indépendamment les uns des autres. (Le tableau A 14 est présenté en annexe) Cependant, force est de constater que le relevé individuel par support d'études ne permet de dégager aucun résultat. A cela, selon nous, plusieurs motifs peuvent être avancés :

- Soit, $_ \text{Vnasale} _ \#$ n'est pas un facteur actif dans la sélection d'un ordre préférentiel de coordination ;
- Soit, $_ \text{Vnasale} _ \#$ est un facteur actif, mais il entre en conflit de contraintes : d'autres facteurs opacifient la visibilité des résultats ;
- Ou alors, les échantillons testés sur ces hypothèses sont trop limités pour rendre compte de résultats saillants.

Il devrait nous être aisé de vérifier cette dernière hypothèse en augmentant le nombre d'items testés. Pour cela, nous nous proposons d'observer la répartition en M1 et en M2 des mots en $_ \text{Vnasale} _ \#$ pour l'ensemble des supports d'études. A nouveau, pour des raisons plusieurs fois

évoquées, nous présenterons deux totaux :

- TOTAL recueille 119 mots en Vnasale_# de l'ensemble des supports d'études examinant l'hypothèse d'une corrélation entre type phonémique du noyau de la syllabe finale fermée et position dans le composé syndétique ;
- TOTAL-TOM ET ZOÉ compile les mêmes données sans celles extraites du corpus TOM ET ZOÉ, soit au total 45 mots.

TOTAL 119	<u>Vnasale_#</u>			
	<u>ɛ̃</u>	<u>œ̃</u>	<u>ɔ̃</u>	<u>ɑ̃</u>
M1	3 100,00%	0	0 0,00%	55 49,55%
M2	0 0,00%	0	5 100,00%	56 50,45%
Total :	3 100,00%	0 100,00%	5 100,00%	111 100,00%

Tableau 76 : Répartition phonémique de Vnasale_# pour TOTAL

TOTAL-TOM ET ZOÉ 45	<u>Vnasale_#</u>			
	<u>ɛ̃</u>	<u>œ̃</u>	<u>ɔ̃</u>	<u>ɑ̃</u>
M1	3 100,00%		0 0,00%	22 59,46%
M2	0 0,00%		5 100,00%	15 40,54%
Total :	3 100,00%	0 100,00%	5 100,00%	37 100,00%

Tableau 77 : Répartition phonémique de Vnasale_# pour TOTAL-TOM ET ZOÉ

Les deux totaux ci-dessus font apparaître deux résultats et deux tendances statistiques que nous ne pouvons commenter. En effet, leur valeur théorique est inférieure à cinq : il nous est impossible de proposer une validation statistique.

Que penser de ces résultats ?

A nouveau, nous pouvons invoquer les trois motivations dont nous faisons part plus haut. Il est possible que d'un point de vue statistique, trop peu d'items soient testés pour nous

permettre de dégager des résultats.

D'autre part, d'un point de vue théorique, à ce stade de nos travaux, nous pouvons présumer que soit le facteur liant les phonèmes de *_Vnasale_#* à une position préférentielle dans la coordonnée n'est pas actif, soit il est obscurci par un conflit de contraintes.

Nous pouvons également supposer que l'hypothèse posant l'assignation de M1 ou de M2 des mots en *_Vnasale_#* ne se fait pas comme nous le pensions en fonction de leur types phonémique, mais en fonction de leur répartition par trait distinctif.

Nous commenterons l'ensemble de ces justifications ainsi que l'absence de résultats notables concernant la répartition phonémique en M1 et en M2 de *_Vnasale_#* en 4.3.2.1.3. Bilan et analyse des résultats de la répartition des noyaux de syllabes finales fermées page 257.

4.3.2.1.2. Répartition par traits distinctifs du noyau des syllabes finales fermées

Nous interrogeant sur ce que les noyaux de syllabes finales fermées partagent de commun ou ont de dissemblable, nous proposons dans cette section une répartition par traits distinctifs en M1 et en M2 de ces derniers.

Dans un premier temps, nous relèverons donc les résultats présentés par chaque support d'études, indépendamment les uns des autres. Puis, nous confronterons les résultats retenus à l'ensemble de nos données.

Rappelons que les supports d'études examinant l'hypothèse d'une corrélation entre les traits distinctifs des noyaux vocaliques [\pm NASAL] des syllabes finales fermées et l'une des deux positions de la coordonnée sont les mêmes que précédemment, à savoir les six corpus et les questionnaires d'enquête ÉLISE ET ZOÉ, ÉDOUARD ET EDMOND et NAROD ET NANOR.

Nous présentons en annexe de la répartition des traits distinctifs de *_V_#* pour chaque support d'études dans le tableau A 15.

Ce tableau nous permet de dégager dix résultats et huit tendances.

Dans le corpus TARN ET GARONNE, nous relevons un résultat statistiquement significatif et une tendance. Les deux assignent une position privilégiée en M1. Ainsi, 66,67% des mots *_-ARRIÈRE_#* sélectionnent M1 ($\chi^2(1) = 4,33$ $p \leq 0,03$). De la même façon, 60,26% des mots *_-ARRONDI_#* ont tendance à occuper la même position ($\chi^2(1) = 3,28$ $p \leq 0,07$).

Le corpus TOM ET ZOÉ révèle quant à lui cinq résultats : les mots *_-NASAL_#*, *_+HAUT_#*, *_-BAS_#*, *_-ARRIÈRE_#* et *_-ARRONDI_#* occupent majoritairement M1

dans respectivement 54,1%, 57,22%, 56,29%, 54,57% et 53,47% des occurrences éprouvées sur ces hypothèses ($p \leq 0,05$)⁸⁴. Les deux derniers résultats corroborent ceux déjà relevés dans le corpus TARN ET GARONNE.

Concernant le corpus ARTS ET MÉTIERS, même si nous relevons une tendance statistique concernant les mots $_{[+ \text{NASAL}] \#}$, nous ne la conserverons pas. En effet, ce résultat étant calculé sur moins de quatre items, il nous est impossible de le valider.

Dans le questionnaire d'enquête ÉLISE ET ZOÉ, deux résultats et une tendance émergent. Nous ne retiendrons cependant pas leurs résultats des mots $_{[- \text{BAS}] \#}$ et $_{[+ \text{ARRONDI}] \#}$. En effet, leur valeur théorique étant très nettement inférieure à cinq, nous ne pouvons les confirmer.

En revanche, ÉLISE ET ZOÉ met au jour le fait qu'en adéquation avec les résultats relevés dans TOM ET ZOÉ, les mots $_{[- \text{NASAL}] \#}$ sont 76,47% à occuper préférentiellement M1 ($\chi^2 (1) = 4,76 p \leq 0,03$).

Nous dégageons également du questionnaire d'enquête ÉDOUARD ET EDMOND une répartition significative et deux tendances. A nouveau, compte tenu du faible nombre d'items éprouvés, nous ne conserverons pas le résultat relevé pour les mots $_{[+ \text{NASAL}] \#}$.

Notons cependant que conformément à ce que nous relevions dans TOM ET ZOÉ et ÉLISE ET ZOÉ, les mots $_{[- \text{NASAL}] \#}$ occupent massivement M1 dans 75% des occurrences éprouvant cette hypothèse ($\chi^2 (1) = 4 p \leq 0,05$). De la même façon, les mots en $_{[- \text{ARRIÈRE}] \#}$ ont tendance à favoriser M1 ($\chi^2 (1) = 3,36 p \leq 0,06$). Même si ce dernier résultat n'est pas aussi statistiquement solide que nous le souhaiterions, il corrobore néanmoins les résultats retenus dans TARN ET GARONNE et TOM ET ZOÉ.

Enfin, le questionnaire d'enquête NAROD ET NANOR fait émerger deux résultats et deux tendances concernant les mots en $_{[- \text{NASAL}] \#}$, $_{[+ \text{HAUT}] \#}$, $_{[- \text{BAS}] \#}$ et $_{[- \text{ARRONDI}] \#}$. Étant tout quatre calculés sur de trop petits échantillons, nous ne les conserverons pas.

84 Respectivement pour chacun des résultats observés, les valeurs des Khi Deux sont : $\chi^2 (1) = 7,38 p \leq 0,01$, $\chi^2 (1) = 11,55 p \leq 0,001$, $\chi^2 (1) = 9,31 p \leq 0,002$, $\chi^2 (1) = 6,12 p \leq 0,01$ et $\chi^2 (1) = 4,51 p \leq 0,03$

L'examen de la répartition par trait distinctif de $_V_\#$ en M1 et en M2 révèle que les mots $_[-\text{NASAL}]_\#$, $_[+\text{HAUT}]_\#$, $_[-\text{BAS}]_\#$, $_[-\text{ARRIÈRE}]_\#$ et $_[-\text{ARRONDI}]_\#$ occupent majoritairement M1.

Attendu que nous ne validerons jamais un résultat porté par un seul support d'études, nous proposons ci-après de confronter les résultats ainsi dégagés à deux ensembles de données. Le premier ensemble TOTAL regroupe l'ensemble des mots $_V_\#$ présents dans nos supports d'études. Le second TOTAL-TOM ET ZOÉ recueille les mêmes données sans celles fournies par TOM ET ZOÉ, pour les motivations que nous avons déjà développées infra.

TOTAL 1629	$_V_\#$									
	NASAL		HAUT		BAS		ARRIÈRE		ARRONDI	
	+	-	+	-	+	-	+	-	+	-
M1	58 48,74%	824 54,57%	391 57,42%	491 51,79%	474 52,90%	408 55,66%	341 51,75%	540 55,73%	187 52,23%	695 54,68%
M2	61 51,26%	686 45,43%	290 42,58%	457 48,21%	422 47,10%	325 44,34%	318 48,25%	429 44,27%	171 47,77%	576 45,32%
Total :	119 100,00%	1510 100,00%	681 100,00%	948 100,00%	896 100,00%	733 100,00%	659 100,00%	969 100,00%	358 100,00%	1271 100,00%

Tableau 78 : Répartition en M1 et en M2 des traits distinctifs de $_V_\#$ pour TOTAL

TOTAL-TOM ET ZOÉ 457	$_V_\#$									
	NASAL		HAUT		BAS		ARRIÈRE		ARRONDI	
	+	-	+	-	+	-	+	-	+	-
M1	25 55,56%	230 55,83%	74 58,27%	181 54,85%	178 57,05%	77 53,10%	114 51,82%	140 59,32%	61 49,59%	194 58,08%
M2	20 44,44%	182 44,17%	53 41,73%	149 45,15%	134 42,95%	68 46,90%	106 48,18%	96 40,68%	62 50,41%	140 41,92%
Total :	45 100,00%	412 100,00%	127 100,00%	330 100,00%	312 100,00%	145 100,00%	220 100,00%	236 100,00%	123 100,00%	334 100,00%

Tableau 79 : Répartition en M1 et en M2 des traits distinctifs de $_V_\#$ pour TOTAL-TOM ET ZOÉ

Les deux tableaux ci-dessus mettent au jour neuf répartitions significatives et trois tendances. Ainsi, les mots $_[-\text{NASAL}]_\#$ occupent majoritairement M1 dans 54,57% des cas dans TOTAL, et 55,57% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 12,61$ $p \leq 0,0004$ et $\chi^2(1) = 5,59$ $p \leq 0,02$). Ces résultats sont homogènes, nous affirmons donc l'existence d'une corrélation entre M1 et les mots $_[-\text{NASAL}]_\#$.

Les mots $[- \text{ARRIÈRE}]_{\#}$ sont 55,73% à sélectionner M1 dans TOTAL, pour 59,32% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 12,72$ $p \leq 0,0004$ et $\chi^2(1) = 8,2$ $p \leq 0,004$). Ces résultats sont identiques, nous confirmons donc que les mots en $[- \text{ARRIÈRE}]_{\#}$ favorisent préférentiellement M1.

De la même façon, les mots $[- \text{ARRONDI}]_{\#}$ sont 54,68% dans TOTAL, et 58,08% dans TOTAL-TOM ET ZOÉ, à favoriser majoritairement M1 (respectivement $\chi^2(1) = 11,14$ $p \leq 0,001$ et $\chi^2(1) = 8,73$ $p \leq 0,003$). Ces résultats étant similaires, nous affirmons donc qu'il existe une corrélation entre M1 et les mots $[- \text{ARRONDI}]_{\#}$.

Enfin, concernant les mots $[\pm \text{HAUT}]_{\#}$ et $[\pm \text{BAS}]_{\#}$, les résultats émergents sont de prime abord contradictoires.

Concernant les mots $[\pm \text{BAS}]_{\#}$, la confrontation des deux totaux révèle une tendance et deux résultats.

Les mots $[+ \text{BAS}]_{\#}$ sont 52,9% dans TOTAL, et 57,05% dans TOTAL-TOM ET ZOÉ, à occuper prioritairement M1 (respectivement $\chi^2(1) = 3,02$ $p \leq 0,08$ et $\chi^2(1) = 6,21$ $p \leq 0,01$). Les deux résultats sont très proches, seule leur validation statistique diffère. TOTAL met au jour une tendance statistique, TOTAL-TOM ET ZOÉ, un résultat. Nous affirmons donc que les mots $[+ \text{BAS}]_{\#}$ privilégient M1.

En revanche, concernant les mots $[- \text{BAS}]_{\#}$, si nous relevons une répartition statistique dans TOTAL, il n'en est rien dans TOTAL-TOM ET ZOÉ. Le résultat plaçant dans 55,66% en M1 les mots $[- \text{BAS}]_{\#}$ ($\chi^2(1) = 9,4$ $p \leq 0,002$) semble, en fait, n'être porté que par le corpus TOM ET ZOÉ. En effet, sur 733 mots $[- \text{BAS}]_{\#}$ examinés dans TOTAL, 588 proviennent du corpus TOM ET ZOÉ. Étant donné que ce résultat n'est pas validé par l'ensemble des corpus, nous ne le conservons pas.

Concernant les mots $[\pm \text{HAUT}]_{\#}$, les deux totaux permettent de dégager un résultat et deux tendances. En effet, le tableau TOTAL, 57,42% des mots $[+ \text{HAUT}]_{\#}$ occupent majoritairement M1 ($\chi^2(1) = 14,98$ $p \leq 0,0001$). En adéquation avec ce résultat, dans TOTAL-TOM ET ZOÉ, 58,27% $[+ \text{HAUT}]_{\#}$ ont tendance à favoriser M1 ($\chi^2(1) = 3,47$ $p \leq 0,06$). TOTAL met au jour une répartition statistiquement significative et TOTAL-TOM ET ZOÉ, une tendance. Exception faite de leur validation statistique, ces résultats sont très homogènes. Nous affirmons donc qu'il existe une corrélation entre les mots $[+ \text{HAUT}]_{\#}$ et

M1.

Cependant, TOTAL-TOM ET ZOÉ fait émerger une tendance *a priori* en contradiction avec le précédent résultat concernant les mots $[- \text{HAUT}]_{\#}$. Les mots $[- \text{HAUT}]_{\#}$ ont tendance à être placés en M1 dans 54,85% des cas ($\chi^2 (1) = 3,1$ $p \leq 0,08$). Nous commenterons en détail ces résultats dans la section à venir.

4.3.2.1.3. Bilan et analyse des résultats de la répartition des noyaux de syllabes finales fermées

Au regard de l'examen de la répartition des noyaux vocaliques des syllabes finales fermées, il apparaît que :

- Les mots dont la syllabe finale fermée possède un noyau vocalique oral se placent plutôt en M1 ;
- ceux dont la syllabe finale fermée possède un noyau vocalique nasal ne permettent de dégager aucun résultat.

- D'un point de vue phonémique, nous relevons que :
 - Les mots en $[_{\varepsilon}]_{\#}$ sélectionnent plutôt M1 ;
 - Tout comme ceux en $[_i]_{\#}$ ou en $[_\text{ɔ}]_{\#}$ pour lesquels nous constatons la même tendance ;
 - En revanche, les mots en $[_\text{o}]_{\#}$ occupent majoritairement M2.

- L'examen de la répartition par traits distinctifs de $[_V]_{\#}$ en M1 et en M2 révèle que :
 - Les mots $[_{-} \text{NASAL}]_{\#}$, $[_{+} \text{BAS}]_{\#}$, $[_{-} \text{ARRIÈRE}]_{\#}$ et $[_{-} \text{ARRONDI}]_{\#}$ privilégient plutôt M1 ;
 - Tout comme les mots $[_{\pm} \text{HAUT}]_{\#}$ pour lesquels nous constatons la même tendance.
 - A noter également que les mots $[_{+} \text{HAUT}]_{\#}$ semblent occuper légèrement plus fréquemment M1 que les mots $[_{-} \text{HAUT}]_{\#}$. En effet, ils sont respectivement 58,27% et 54,85% à sélectionner cette position.

Les résultats concernant les mots $[_{\pm} \text{HAUT}]_{\#}$, ainsi que les mots en $[_i]_{\#}$ ou en $[_\text{ɔ}]_{\#}$ ne bénéficient pas de la même validation statistique que les autres. Si l'ensemble des corrélations

entre $_V_\#$ et l'une des deux positions de la coordonnée affirment des résultats significatifs, celles des mots en $_/ɔ/_\#$, en $_/i/_\#$ et $_[\pm \text{HAUT}]_\#$ présentent des tendances. Nous proposons néanmoins ces tendances, car consciente des conflits de contraintes en œuvre dans la sélection d'un ordre préférentiel de coordination par « et », nous pensons que la faiblesse statistique de ces résultats n'est que l'artefact de l'interaction d'autres facteurs. Assurément, il existe des corrélations entre M1 et les mots en $_/ɔ/_\#$ en $_/i/_\#$ et $_[\pm \text{HAUT}]_\#$, mais ces corrélations sont moins remarquables, moins évidentes.

Que penser de l'examen de la répartition des noyaux de syllabes fermées ?

Nous notons tout d'abord que les résultats ci-dessus renforcent ceux que nous relevions précédemment. Les mots possédant un noyau de syllabe finale fermée suivent la même répartition que ceux se terminant par une consonne. En effet, comme nous l'indiquions page 246, toute syllabe s'organise autour d'un noyau. Il est donc naturel que $_V_\#$ soit plutôt assignée en M1, comme $_C_\#$.

Ces résultats étant liés, il semble légitime de supposer que les principes agissant dans la sélection d'un ordre préférentiel des mots en $_C_\#$ soient les mêmes pour $_V_\#$: même causes, mêmes conséquences.

C'est pourquoi, nous pensons que comme exposé page 238, le principe de MAXIMISATION DES ATTAQUES favorise l'occupation en M1 des mots en $_V_\#$.

Cependant, si ce principe motive la préférence des mots en $_V_\#$ à occuper M1, il n'explique en rien les corrélations entre cette position et $_[- \text{NASAL}]_\#$, $_ [+ \text{BAS}]_\#$, $_ [- \text{ARRIÈRE}]_\#$ et $_ [- \text{ARRONDI}]_\#$ ou $_ [\pm \text{HAUT}]_\#$, ni pourquoi, si les mots en $_/ɛ/_\#$, $_/i/_\#$ ou en $_/ɔ/_\#$ sélectionnent plutôt M1, ceux en $_/o/_\#$ sont rejetés en M2.

L'opposition marqué / non marqué peut à nouveau rendre compte de façon vraisemblable de ces phénomènes. Rappelons que dans la littérature spécialisée (Benor et Levy (2006) et Kelly, Bock et Keil (1986), entre autres), cette opposition semble prédire que les éléments non marqués ont tendance à occuper la position non marquée. L'élément le plus prototypique, le moins marqué, précède l'élément le plus marqué.

Or, conformément aux définitions de la marque que nous évoquions précédemment (Clements (2007)), $_ [- \text{NASAL}]_\#$, $_ [- \text{ARRIÈRE}]_\#$ et $_ [- \text{ARRONDI}]_\#$ sont non marqués. De même, les voyelles $/ɛ/$ et $/i/$, toutes deux $\left(\begin{array}{c} - \text{ARRIÈRE} \\ - \text{ARRONDI} \end{array} \right)$, sont moins marquées que $/o/$ $\left(\begin{array}{c} + \text{ARRIÈRE} \\ + \text{ARRONDI} \end{array} \right)$.

Cependant si le recours à l'opposition marqué / non marqué fournit une motivation vraisemblable à l'assignation en M1 des mots $[- \text{NASAL}]_{\#}$, $[+ \text{BAS}]_{\#}$ ⁸⁵, $[- \text{ARRIÈRE}]_{\#}$, $[- \text{ARRONDI}]_{\#}$, $[_{\varepsilon}/_{\#}$ et $[_i]/_{\#}$ et de celle en M2 des mots en $[_o]/_{\#}$, il ne justifie en rien la préférence qu'ont les mots $[_{\pm} \text{HAUT}]_{\#}$ et $[_\text{ɔ}]/_{\#}$ à occuper M1.

Concernant la tendance plaçant les mots en $[_\text{ɔ}]/_{\#}$ en M1, il nous faut observer deux phénomènes :

- La neutralisation des voyelles moyennes, ou loi de position ;
- Et les effets de la loi du second lourd que nous évoquons page 218.

Revenons tout d'abord sur la loi de position.

La loi de position telle qu'elle est présentée par Fouché (1956) prescrit que le degré d'aperture des voyelles moyennes dépend de leur position. A savoir, les voyelles moyennes sont fermées en syllabe ouverte et ouvertes en finale fermée accentuable. D'une manière générale, le relevé de nos supports d'études corrobore ce principe : nous n'observons, par exemple, aucun noyau vocalique /e/ en syllabe finale fermée.

Nous aurions pu proposer ici l'examen de la distribution des archiphonèmes /E/, /Ø/ et /O/, leur degré d'aperture étant prévisible. Il nous est néanmoins apparu que dans le cadre de nos travaux, la neutralisation des voyelles moyennes en raison des très nombreuses exceptions à la loi de position, ne nous permet qu'une évaluation rudimentaire de la répartition de ces voyelles notamment en ce qui concerne l'examen de /O/ comme noyau de syllabe fermée. Rappelons que l'archiphonème /O/ en syllabe entravée se réalise en [o] devant /z/ et en [ɔ] devant /R, g, ɲ/. /O/ ne nous permet donc pas d'observer la tendance plaçant les mots en $[_\text{ɔ}]/_{\#}$ en M1, et la préférence de ceux en $[_o]/_{\#}$ à occuper M2.

Cependant, les lois de position nous poussent à réévaluer l'examen de la répartition des mots en $[_\text{ɔ}]/_{\#}$. En fait, nous pensons que plutôt que de justifier la présence des mots en $[_\text{ɔ}]/_{\#}$ en M1, il nous faut surtout les comparer aux mots en $[_o]/_{\#}$ qui eux, favorisent très majoritairement M2.

Pour cela, nous invoquons à nouveau la loi du second lourd. Pour rappel, ce principe rend compte de l'assignation en M2 du mot le plus lourd, le plus complexe. Or, selon des paramètres de phonétique articulatoire, /ɔ/ est plus aiguë (F1 ≈ 528 Hz) que /o/ (F1 ≈ 428

85 Si l'on considère les deux degrés extrêmes de constriction des voyelles du français, on peut envisager que le trait [+ BAS] dans son opposition privative à [+ HAUT], est non marqué.

Hz)⁸⁶. /o/ possédant une fréquence plus basse, est donc un noyau de syllabe plus complexe que /ɔ/. Cette différence motive selon nous l'occupation préférentielle des mots en $_ / \text{ɔ} / _ \#$ en M2, et la tendance statistique à ce que les mots en $_ / \text{ɔ} / _ \#$ les précèdent en occupant plutôt M1.

Pour expliquer les résultats *a priori* contradictoires relevés pour les mots $_ [\pm \text{HAUT}] _ \#$, nous présentons trois causes potentielles :

- a) soit il s'agit d'un problème lié à l'hypothèse posée : l'hypothèse liant l'une des deux positions dans la coordination à $_ [\pm \text{HAUT}] _ \#$ n'est pas une bonne hypothèse ;
- b) soit les facteurs $_ [\pm \text{HAUT}] _ \#$ ne sont pas des facteurs actifs dans la sélection d'une forme binomiale préférentielle ;
- c) ou alors le facteur $_ [\pm \text{HAUT}] _ \#$ est actif dans la sélection d'un ordre coordonné, mais ce facteur est obscurci par un conflit de contraintes.

Les traits $_ [\pm \text{HAUT}] _ \#$ indiquant la position de la langue dans la bouche sont utilisés pour représenter le degré d'aperture de $_ V _ \#$. Or, le degré d'aperture des voyelles ne permet en français que des oppositions graduelles (exception faite de deux degrés extrêmes de constriction). De plus, compte tenu des effets des lois de position sur les voyelles moyennes, nous pensons que les hypothèses observant la répartition en M1 ou en M2 en fonction du degré d'aperture de noyaux de syllabes finales fermées sont inadéquates. Selon nous, les résultats relevés pour les $_ [\pm \text{HAUT}] _ \#$ étayent, en fait, les résultats que nous présentions précédemment à savoir que les mots en $_ V _ \#$, comme ceux en $_ C \#$ sélectionnent plutôt M1.

Enfin, il semble intéressant de comparer l'occupation en M1 et en M2 des noyaux vocaliques des syllabes finales entravées à celle des noyaux vocaliques des syllabes finales ouvertes présentée page 215.

Il apparaît que si les mots en $_ V \#$ sélectionnent généralement M2, ceux en $_ V _ \#$ préfèrent M1. D'une façon générale, les mots en $_ V \#$ sont marqués, ceux en $_ V _ \#$ non marqués.

Notons une similitude : les mots en $_ / \text{o} / _ \#$, comme ceux en $_ / \text{o} / \#$ occupent majoritairement M2.

Exception faite des mots en $/ \text{e} / \#$ pour lesquels il nous faut invoquer le principe du contour obligatoire, les résultats concernant les noyaux de syllabes finales, qu'il s'agisse de syllabes

⁸⁶ Les fréquences des premiers formants de ces deux voyelles sont extraites de Georgeton et al (2012) reprenant les travaux de Gendrot et Adda Decker (2005).

ouvertes ou fermées, tombent tous sous les effets de deux principes : l'opposition marqué / non marqué, et la loi du second lourd.

4.3.2.2. structure de la syllabe finale

L'intérêt que nous portons à la répartition en M1 et en M2 des mots en fonction de leur *pattern* de syllabes finales est triple.

Les résultats observant les frontières droites des mots font apparaître deux principes récurrents : l'opposition marqué / non marqué, et en lien avec celle-ci, la loi du second lourd.

Si comme nous le présenterons plus en détail page 272, la forme non marquée est CV, peut-on imaginer que l'ordre préférentiel de coordination par « et » privilégiera le maintien de cette structure syllabique ?

D'autre part, conformément aux prescriptions de la loi du second lourd, peut-on supposer un poids de syllabes actif dans la composition syndétique par « et » ?

Enfin, quel rôle joue l'allongement vocalique, qu'il soit déclenché par une consonne allongante ou lié au timbre de la voyelle ? Est-il conforme aux effets de la loi du second lourd ?

Afin de répondre à ces questionnements, dans la section à venir, nous examinerons tout d'abord la répartition en M1 et en M2 des mots en fonction de la structure de leur syllabe finale, puis nous observerons la dispersion des syllabes finales possédant une voyelle allongée, enfin nous proposerons une analyse de nos relevés.

4.3.2.2.1. Patron de syllabes finales

Afin d'étudier le rôle des syllabes finales dans la sélection d'un ordre préférentiel de coordination binaire par « et », nous avons étiqueté l'ensemble de nos supports d'études en fonction de leur patron de syllabes finales, exception faite de PAUL ET PASCAL. En effet, ce questionnaire d'enquêtes évalue les questions de la taille des mots relativement à leur position dans la coordination, il est donc neutralisé sur tout autre facteur, dont le type de finales syllabiques.

Au total, nous avons relevé 1 799 paires.

L'étiquetage en fonction des patrons de syllabes finales s'est effectué comme suit :

- Pour les mots en *_V#* :
 - *_V#*
 - *_GV#*
 - *_CGV#*
 - *_CV#*
 - *_CCV#*

- Pour les mots en *_C#* :
 - *_(C)GVC(C)#*
 - *_VC#*
 - *_VCC#*
 - *_CVC#*
 - *_CCVC#*
 - *_CVCC#*
 - *_CCVCC#*

- Pour les mots en *_G#* :
 - *_VG#*
 - *_CVG#*

Les répartitions par supports d'études des patrons de syllabes finales en M1 puis en M2 sont présentées en annexe dans les tableaux 16a et 16b. Dans ces deux tableaux, nous avons mentionné en caractères gras les résultats significatifs et en italique, les tendances.

Dans le détail, nous relevons deux résultats significatifs et cinq tendances statistiques. D'une manière générale, l'observation sommaire des deux tableaux fait apparaître deux orientations : globalement, à une exception près que nous présenterons à venir, les mots en *_C#* occupent M1, ceux en *_V#* et en *_G#*, M2.

Plus précisément, support d'études par support d'études, nous observons les résultats suivants :

Dans le corpus TARN ET GARONNE, 71,88 % des mots en *_CV#* sélectionnent massivement M2 ($\chi^2(1) = 6,13$; $p \leq 0,01$).

Le corpus TOM ET ZOÉ, quant à lui, met au jour quatre résultats significatifs et une tendance statistique.

Ainsi les mots en $_CCV\#$, en $_CVG\#$, et en $_CV\#$ occupent majoritairement M2 dans respectivement 100%, 75 % et 53,82% des occurrences éprouvées sur ces trois hypothèses ($\chi^2(1) = 15$; $p \leq 0,0001$, $\chi^2(1) = 9$; $p \leq 0,003$, $\chi^2(1) = 7,25$; $p \leq 0,01$).

Inversement, ceux en $_CVCC\#$ sont 66,25% à favoriser M1 ($\chi^2(1) = 8,78$; $p \leq 0,003$), tout comme ceux en $_(C)GVC(C)\#$ sont 56,25 % à avoir tendance à sélectionner la même position ($\chi^2(1) = 2,75$; $p \leq 0,10$).

Dans le corpus VITE ET BIEN, nous relevons deux tendances statistiques.

Nous ne conserverons pas celle concernant les mots en $_GV\#$. La tendance les plaçant en M2 étant calculée sur trois mots, le test de Khi Deux n'est pas probant.

Concernant les mots en $_CV\#$, à la différence des résultats émergents de chaque support d'études examiné individuellement, le corpus VITE ET BIEN fait apparaître une tendance à ce qu'ils occupent majoritairement M1, dans 62,26 % des cas ($\chi^2(1) = 3,19$; $p \leq 0,07$) ;

Deux tendances semblent émerger du corpus ARTS ET MÉTIERS. Nous ne conservons pas celle liant les mots en $_VC\#$ en M1. Calculé sur un échantillon de cinq mots, le test de Khi Deux n'est pas concluant.

Nous relevons néanmoins que conformément aux résultats que nous relevions dans les corpus TARN ET GARONNE et TOM ET ZOÉ, 66,25% des mots en $_CV\#$ ont tendance à favoriser M1 ($\chi^2(1) = 3,33$; $p \leq 0,07$).

Enfin, dans le questionnaire d'enquêtes ÉLISE ET ZOÉ, nous recueillons un résultat significatif assignant 85,71 % des mots en $_CVC\#$ en M1 ($\chi^2(1) = 7,14$; $p \leq 0,01$.)

L'examen de la structure de la syllabe finale par support d'études met au jour des résultats pour les mots en $_(C)CGVC(C)\#$, en $_CVG\#$, en $_CV\#$, en $_CCV\#$, en $_CVCC\#$, en $VC\#$ et en $_CVC\#$. Seul le résultat associant M2 au final ayant pour structure $_CV\#$ est récurrent. Pour tous les autres, ayant fait le choix de ne présenter aucun résultat porté par un unique support d'études, nous devons mettre en place une vérification.

Pour cela, dans le paragraphe à suivre, nous nous proposons de comparer ces hypothèses avec les résultats produits par l'ensemble de nos supports d'études.

Comme précédemment, nous présentons deux totaux. Le premier TOTAL envisage conjointement les 1 799 paires de l'ensemble de nos supports d'études. Le second, TOTAL -TOM ET ZOÉ, parcourt le même ensemble sans les paires extraites du corpus TOM ET ZOÉ, soit 379 paires. Dans ce dernier, nous prêtons une attention particulière aux résultats portés par le corpus TOM ET ZOÉ, à savoir ceux concernant les mots en *_(C)CGVC(C)#*, en *_CVG#*, en *_CV#*, en *_CCV#*, et en *_CVCC#*.

Dans les tableaux présentés page suivante, nous avons mentionné les répartitions statistiquement significatives en caractère gras et les tendances en italique. Nous avons également souligné deux résultats. Leur valeur-p étant proche du seuil de .10 que nous avons fixé, il nous a semblé pertinent de les signaler au moins à titre indicatif.

	type_σ#													
	(C)GVC(C)#	CGV#	GV#	VG#	CVG#	V#	VC#	VCC#	CV#	CVC#	CCV#	CCVC#	CVCC#	CCVCC#
TOTAL 1799														
M1	123	130	0		12	95	100	3	664	564	5	12	71	0
%	55,41%	50,78%	0,00%		26,09%	49,74%	64,94%	50,00%	45,82%	52,76%	16,13%	57,14%	62,28%	
M2	99	126	6		34	96	54	3	785	505	26	9	43	0
%	44,59%	49,22%	100,00%		73,91%	50,26%	35,06%	50,00%	54,18%	47,24%	83,87%	42,86%	37,50%	
Total :	222	256	6		46	191	154	6	1449	1069	31	21	114	
	100,00%	100,00%	100,00%		100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	

Tableau 80 : Répartition en M1 et en M2 des structures de syllabes finales pour l'ensemble des supports d'études

	type_σ#													
	(C)GVC(C)#	CGV#	GV#	VG#	CVG#	V#	VC#	VCC#	CV#	CVC#	CCV#	CCVC#	CVCC#	CCVCC#
TOTAL-TOM ET ZOË 379														
M1	24	16	0		3	3	16	3	89	172	5	12	16	
%	52,17%	40,00%	0,00%		30,00%	27,27%	80,00%	50,00%	43,63%	54,78%	31,25%	57,14%	51,61%	
M2	22	24	6		7	8	4	3	115	142	11	9	15	
%	47,83%	60,00%	100,00%		70,00%	72,73%	20,00%	50,00%	56,37%	45,22%	68,75%	42,86%	48,39%	
Total :	46	40	6		10	11	20	6	204	314	16	21	31	
	100,00%	100,00%	100,00%		100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	

Tableau 81 : Répartition en M1 et en M2 des structures de syllabes finales pour l'ensemble des supports d'études sans TOM ET ZOË

La confrontation des deux totaux fait émerger six répartitions significatives et trois tendances. Dans le détail, nous présentons un à un les couples de résultats de gauche à droite dans les tableaux. Nous ne conservons pas les résultats relevés pour les mots en *_GV#* qui bien qu'identiques dans les totaux, ne sont examinés que sur six mots. Le calcul de Khi Deux ne peut être probant.

Concernant l'hypothèse assignant les mots en *_CVG#* en M2 relevée dans le corpus TOM ET ZOÉ, la confrontation des deux totaux rapporte que 73,91% d'entre eux sélectionnent M2 dans TOTAL, pour 70% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 10,52$; $p \leq 0,001$ et $\chi^2(1) = 1,6$; $p \leq 0,21$). Ces résultats sont *a priori* identiques. Notons cependant que leur validation diverge. TOTAL présente une répartition significative, ce qui n'est pas le cas de TOTAL-TOM ET ZOÉ. Cette différence nous indique que le résultat observé dans TOTAL n'est en fait porté que par le corpus TOM ET ZOÉ. En effet, sur 46 mots testés, 36 proviennent du corpus TOM ET ZOÉ. Même si intuitivement nous supposons qu'il est possible d'imputer l'absence de validation statistique à la petite taille de l'échantillon examiné dans TOTAL-TOM ET ZOÉ (seulement dix mots), nous ne pouvons conserver ce résultat.

Les mots en *_VC#*, pour lesquels nous relevons des résultats dans le corpus ARTS ET MÉTIERS, sont 64,94% à occuper M1 dans TOTAL, pour 80% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 13,74$; $p \leq 0,0002$ et $\chi^2(1) = 7,2$; $p \leq 0,01$). La répartition des mots en *_VC#* suit la même inclinaison vers M1 dans les deux totaux. Les résultats sont tous deux significatifs. Nous affirmons l'existence d'une corrélation entre M1 et les mots en *_VC#*.

Concernant les mots en *_CV#* pour lesquels nous observons des résultats plutôt consensuels dans TARN ET GARONNE, TOM ET ZOÉ et VITE ET BIEN, ils sont 54,18% à favoriser M2 dans TOTAL, pour 56,37% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 10,1$; $p \leq 0,001$ et $\chi^2(1) = 3,31$; $p \leq 0,07$). Ces résultats sont extrêmement proches, exception faite de leur validation statistique. TOTAL met au jour un résultat significatif, TOTAL-TOM ET ZOÉ, une tendance. Cette tendance statistique est selon nous l'indice d'un conflit de contraintes dans la sélection d'un ordre préférentiel de coordination binaire par «et». Ainsi, bien que ces résultats ne soient pas aussi forts que nous le souhaiterions, nous sommes convaincue que les mots en *_CV#* sélectionnent majoritairement M2.

Pour les mots en *_CVC#*, pour lesquels nous remarquons des résultats dans le questionnaire

d'enquête ÉLISE ET ZOÉ, nous relevons deux tendances statistiques. Les mots en $_CVC\#$ sont 52,76% à avoir tendance à favoriser M1 dans TOTAL, pour 54,78% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 3,26$; $p \leq 0,07$ et $\chi^2(1) = 2,87$; $p \leq 0,09$). Ces résultats identiques, bien que ne présentant pas de répartition significative aussi forte que nous le souhaiterions, mettent au jour une corrélation entre les mots en $_CVC\#$ et la position M1.

Les mots en $_CCV\#$ pour lesquels nous observons des résultats dans le corpus TOM ET ZOÉ, sont 83,87% à occuper M2 dans TOTAL, pour 68,75% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 14,23$; $p \leq 0,0002$ et $\chi^2(1) = 2,25$; $p \leq 0,13$). Les résultats présentés dans les deux tableaux suivent les mêmes contours en assignant majoritairement M2 aux mots se terminant par une syllabe du type CCV. Cependant, les validations statistiques sont inégales. Si TOTAL met au jour une répartition significative, TOTAL-TOM ET ZOÉ présente un résultat dans laquelle la valeur-p est juste au-dessus du seuil de .10. La différence de validation statistique nous indique habituellement l'influence du corpus TOM ET ZOÉ. Ce n'est pas le cas ici : sur 31 mots testés dans TOTAL, seuls 15 sont issus de TOM ET ZOÉ.

Que faire de ce résultat ? Comme nous l'avons indiqué précédemment la signification asymptotique est graduelle. De fait, avec une valeur-p de 0,13, si on ne peut pas rejeter l'hypothèse nulle, son « non-rejet » n'est pas catégorique. Cela signifie qu'il y a moins de 13% que le résultat soit dû au hasard. La faiblesse statistique de ce résultat peut être le résultat d'interactions de contraintes, mais elle peut être également due à la petite taille de l'échantillon testé (seulement 16 mots). C'est pourquoi ; avec toutes les précautions nécessaires, nous présentons le résultat assignant les mots en $_CCV\#$ plutôt en M2.

Avec la même prudence et pour les mêmes raisons, nous notons également que les mots en $_V\#$ occupent plutôt M2 dans 72,73 % des 11 occurrences examinant cette hypothèse ($\chi^2(1) = 2,27$; $p \leq 0,13$).

Enfin, concernant les mots en $_CVCC\#$, s'ils sont 62,28% à sélectionner M1 dans TOTAL, ils ne sont plus que 51,61 % à favoriser cette même position dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 6,28$; $p \leq 0,01$ et $\chi^2(1) = 0,03$; $p \leq 0,86$). TOTAL présente un résultat significatif, ce qui n'est pas le cas de TOTAL-TOM ET ZOÉ. De plus, sur les 114 mots testés dans TOTAL, 83 proviennent du corpus TOM ET ZOÉ. Il apparaît que ce résultat n'est porté que par le corpus TOM ET ZOÉ, nous ne le conservons pas.

Au regard de l'examen de la répartition en M1 et en M2 des mots en fonction du *pattern* de leur syllabe finale, nous relevons que :

- Les mots en *_C#* occupent M1 ;
- Ceux en *_V#* plutôt M2.

- Dans le détail, concernant les mots se terminant par une consonne :
 - Les mots en *_VC#* sélectionnent M1 ;
 - Tout comme ceux en *_CVC#* pour lesquels on relève la même tendance.

- Concernant les mots se terminant par une voyelle :
 - Les mots en *_CV#* ont tendance à favoriser M2 ;
 - Tout comme ceux en *_CCV#* et *_V#* pour lesquels les résultats sembleraient indiquer une préférence à occuper M2.

Nous développerons ces résultats en tentant d'apporter une analyse motivant l'assignation d'une position préférentielle dans la coordonnée de ces patrons syllabiques en 4.3.2.2.3. Bilan et analyse des résultats des syllabes finales fermées page 272.

4.3.2.2.2. Allongement vocalique

Dans ce paragraphe, nous nous intéressons aux phénomènes d'allongement vocalique réguliers. Ces phénomènes décrits notamment par Tranel (1987) et Béchade (1992) s'opèrent en syllabe fermée.

Ils sont de deux types : soit l'allongement est déclenché par une consonne finale allongeante, soit l'allongement vocalique est lié au timbre de la voyelle.

Ainsi, tout noyau vocalique de syllabe finale entravée par /v/, /z/, /ʒ/, /R/ ou /vR/ est allongé, de même que tout noyau vocalique de syllabe finale fermée en *_/ø/_#*, en *_/o/_#*, en *_/ɑ/_#* ou en *_V[+NASAL]_#*.

L'intérêt que nous portons aux phénomènes d'allongement vocalique est en lien avec nos interrogations sur les patrons de syllabes finales. L'opposition marqué / non marqué, tout comme la loi du second lourd partagent toutes deux la tendance à repousser en seconde position l'élément le plus complexe. Les effets de ces principes sont-ils effectifs sur le poids de syllabe ?

A travers l'examen de la répartition en M1 et en M2 des mots en fonction de l'allongement vocalique de leur syllabe finale fermée, nous souhaitons examiner l'hypothèse d'un rôle du poids de syllabes actif dans la sélection préférentielle d'un ordre de coordination.

Afin de traiter cette question, nous avons recueilli l'ensemble des données présentes dans nos supports d'études dans lesquels au moins un des deux termes se termine par une syllabe dont le noyau est une voyelle longue.

Nous en proposons l'étiquetage suivant :

- Pour l'allongement déclenché par une consonne allongeante :
 - $_V | /v_R/\#$;
 - $_V | /v, z, ʒ, R/$;
 - Et la somme de ces patrons de syllabes finales allongeantes notée $_V: C(C)\#$.

- Pour l'allongement lié au timbre du noyau de syllabe finale entravée :
 - $_V[+NASAL]C\#$;
 - $_ /ø, o, a/ C\#$.
 - Ainsi que la somme de ces patrons de syllabes finales allongeantes notée $_V : | C\#$

- Et enfin en comparaison :
 - $_V[+NASAL]\#$;
 - $_ /ø, o, a/\#$
 - et $_C | /R, l/\#$.

Dans les deux tableaux page suivante, nous mettons en avant leur répartition en M1 et en M2. Le premier tableau examine l'ensemble de mots $_V: \#$ présent dans nos supports d'études. Nous le nommons TOTAL. Le second, pour les raisons que nous avons déjà évoquées, met à disposition le même ensemble sans les paires extraites du corpus TOM ET ZOÉ. Nous nommons ce total TOTAL -TOM ET ZOÉ.

A titre informatif, nous proposons également la répartition en M1 et en M2 des mots en $_V: \#$ par support d'études en annexe dans le tableau A 17. Cependant, compte tenu du très faible nombre de résultats apparents, nous ne le détaillerons pas ici.

Pour l'ensemble des tableaux présentés *infra*, les répartitions significatives ($p \leq 0,05$) sont à nouveau notées en caractères gras et les tendances ($p \leq 0,10$) en italique.

TOTAL 1799	type σ#						-V[+ NASAL]#	_oəɑ#	_C R, l #
	_V v+r#	_V RZ3v#	_V : C(C)#	-V[+ NASAL]C#	oəɑ/C#	_V :/C#			
M1	0	183	183	37	3	40	248	387	21
%	0,00%	52,14%	51,99%	56,92%	18,75%	49,38%	40,26%	47,43%	44,68%
M2	1	168	169	28	13	41	368	429	26
%	100,00%	47,86%	48,01%	43,08%	81,25%	50,62%	59,74%	52,57%	55,32%
Total :	1	351	352	65	16	81	616	816	47
	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

Tableau 82 : Répartition en M1 et en M2 des mots en _V:# pour TOTAL

TOTAL-TOM ET ZOÉ 329	type σ#						-V[+ NASAL]#	_oəɑ#	_C R, l #
	_V v+r#	_V RZ3v#	_V : C(C)#	-V[+ NASAL]C#	oəɑ/C#	_V :/C#			
M1	0	78	78	18	3	21	42	30	7
%	0,00%	51,66%	51,32%	81,82%	18,75%	55,26%	36,52%	42,25%	46,67%
M2	1	73	74	4	13	17	73	41	8
%	100,00%	48,34%	48,68%	18,18%	81,25%	44,74%	63,48%	57,75%	53,33%
Total :	1	151	152	22	16	38	115	71	15
	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

Tableau 83 : Répartition en M1 et en M2 des mots en _V:# pour TOTAL-TOM ET ZOÉ

De la confrontation de ces deux tableaux, nous ne dégagons que cinq répartitions significatives. Deux d'entre elles corroborent des résultats démontrés précédemment page 215 à savoir que les mots possédant une finale vocalique nasale occupent majoritairement M2.

Les trois autres résultats notables mettent au jour des corrélations entre l'une des deux positions dans la coordonnée et l'allongement vocalique lorsque ce dernier est déclenché par le timbre du noyau de la syllabe finale.

Notons qu'aucun résultat saillant ne fait apparaître un lien entre la sélection d'un ordre préférentiel de coordination et les syllabes finales dont l'allongement vocalique est provoqué par la présence d'une coda allongeante.

Ainsi, par ordre de présentation dans les tableaux, nous relevons les résultats suivants.

Les mots en $_V[+ \text{NASAL}]C\#$ sont 56,96% à occuper M1 dans TOTAL, pour 81,82% dans TOTAL-TOM ET ZOÉ (respectivement $\chi^2(1) = 1,25$; $p \leq 0,26$ et $\chi^2(1) = 8,91$; $p \leq 0,003$). Même si ces résultats indiquent tous deux une préférence à ce que les mots en $_V[+ \text{NASAL}]C\#$ sélectionnent M1, ils sont différents en termes de score et de validation statistique. En effet, seuls les résultats présentés dans TOTAL-TOM ET ZOÉ sont significatifs.

Comment comprendre de telles différences ?

La confrontation de ces deux résultats rend compte du biais porté par le corpus TOM ET ZOÉ. En effet, sur 65 items testés dans TOTAL, 43 proviennent du corpus TOM ET ZOÉ. Or, l'examen attentif de ces 43 tokens révèle qu'ils ne sont en fait construits que sur trois types « CLÉMENCE », « MAXENCE » et « ALEXANDRE ». En comparaison TOTAL-TOM ET ZOÉ met à notre disposition 22 tokens pour 18 types. Le résultat présenté par TOTAL n'est donc pas fiable. Nous affirmons donc, qu'en adéquation avec les résultats produits par TOTAL-TOM ET ZOÉ, les mots en $_V[+ \text{NASAL}]C\#$ sélectionnent majoritairement M1.

Les mots en $_ / \emptyset, o, \alpha / | C\#$, quant à eux, sont 81,25% à privilégier massivement M2 dans TOTAL et TOTAL-TOM ET ZOÉ ($\chi^2(1) = 6,25$; $p \leq 0,01$). Nous attestons de l'existence d'une corrélation forte entre les mots en $_ / \emptyset, o, \alpha / | C\#$ et M2.

4.3.2.2.3. Bilan et analyse des résultats des syllabes finales fermées

Au regard de l'examen de la répartition en M1 et en M2 des mots en fonction de la structure de leur syllabe finale, nous relevons que :

- Les mots en *_C#* occupent M1 ;
- Ceux en *_V#* plutôt M2.

- Dans le détail, concernant les mots se terminant par une consonne, nous constatons que :
 - Les mots en *_VC#* sélectionnent M1 ;
 - Tout comme ceux en *_CVC#* pour lesquels on relève la même tendance.

- Les mots se terminant par une voyelle révèlent que :
 - Les mots en *_CV#* ont tendance à favoriser M2 ;
 - Tout comme ceux en *_CCV#* et *_V#* pour lesquels les résultats sembleraient indiquer une préférence à occuper M2.

- Concernant les effets de l'allongement vocalique, l'examen de nos données fait apparaître une différence entre l'allongement vocalique déclenché par une consonne allongeante, et celui lié au timbre des noyaux des syllabes finales fermées :
 - Nous ne relevons pas de résultats de l'examen de l'allongement vocalique provoqué par la présence d'une coda allongeante ;
 - Nous remarquons que les deux résultats assignant une position préférentielle dans la coordination examinent des mots dont le timbre du noyau de la syllabe finale fermée est allongeant.

- Dans le détail, nous observons que :
 - Les mots en *_V[+ NASAL]C#* occupent préférentiellement M1 ;
 - Inversement, ceux en *_ /ø, o, α / C#* sélectionnent massivement M2.

Comment comprendre et analyser de tels résultats ?

En mettant en place deux étiquetages observant les *pattern* et l'allongement vocalique des syllabes finales, et en comparant les répartitions en M1 et en M2 des syllabes lourdes (C^nVV)

ou C^nVCC^{n87}) à celles des syllabes légères (C^nV), nous souhaitons avant tout examiner l'hypothèse d'un poids de syllabes actif dans la sélection d'un ordre préférentiel de coordination par « et ». Notre idée initiale est que plus la rime de la syllabe finale est complexe, plus les mots tombent sous les effets de la loi du second lourd présentée page 218. Conséquemment à ces considérations, les mots en $_CV\#$, en $_V[+NASAL]C\#$, en $_VC\#$ et en $_CVC\#$ devraient logiquement être rejetés en M2.

Or, aucun des résultats présentés ici ne confirme l'action du poids de la syllabe finale. Les syllabes les plus lourdes n'occupent pas forcément de positions préférentielles : les mots en $_CV\#$ occupent M2, ceux en $_V[+NASAL]C\#$, en $_VC\#$ et en $_CVC\#$, M1.

Cependant, si les résultats relevés de l'examen des structures des syllabes finales ne nous permettent pas de confirmer l'hypothèse que nous avançons, ils valident néanmoins des résultats déjà observés, à savoir l'occupation préférentielle des mots en $_C\#$ en M1 et celle des mots en $_V\#$ en M2.

Dans le détail, nous pouvons rendre compte de la préférence des mots en $_V[+NASAL]C\#$, en $_VC\#$ et $_CVC\#$ à occuper M1 en invoquant le principe de Maximisation des Attaques que nous présentions page 238. La coda finale en M1 se ressyllabe avec le segment « et » formant une nouvelle syllabe qui respecte le schéma CV : Onset >> *Coda. Notons que le relevé par types de structures syllabiques finales examiné page 261 montre que d'une manière générale, en termes de score et non de validation statistique, tous les *pattern* se terminant par $_C\#$ favorisent M1.

Concernant, les mots en $_V\#$, il apparaît (à nouveau en termes de score et non de validation statistique) que tous les patrons de syllabes finales ouvertes sélectionnent plutôt M2, respectant le principe d'évitement de *VV que nous évoquions page 216. Les mots en $_CV\#$, $_CV\#$, $_CCV\#$ et $_V\#$ sont rejetés en M2 afin d'éviter l'hiatus.

D'une manière générale, les résultats relevés dans cette section sont tous vraisemblablement motivés par l'opposition marqué / non marqué.

En effet, les finales consonantiques, [-SONANT], sont moins marquées (Clements 2007) que les finales vocaliques. Elles semblent donc favoriser la position la moins marquée, M1. Inversement, les finales [+SONANT] marquées occupent plutôt la position marquée, M2.

87 L'indice ⁿ mentionne un nombre minimal de un élément.

Conformément à ce que nous observions précédemment, les termes non marqués tendent à occuper la position non marquée. L'élément le plus prototypique, le moins marqué, précède l'élément le plus marqué.

4.3.3. Les phénomènes de liaison

Dans nos travaux, nous nous devions de proposer un aperçu des phénomènes de liaisons.

En effet, la liaison agissant sur les frontières de mots, il nous était impossible de ne pas observer ces manifestations agissant potentiellement en finale de mots entre M1 et « et ».

D'autant plus, que comme les voyelles nasales et le schwa, c'est l'un des problèmes les plus récurrents et les plus célèbres de la phonologie du français.

Dans un premier temps, nous exposerons donc un tour d'horizon sommaire de quelques théories explorant la liaison en français, avant d'exploiter nos données concernant ces phénomènes.

Pour rappel, d'un point de vue diachronique, la liaison est une survivance de l'époque où toutes les consonnes finales étaient prononcées en français.

D'un point de vue synchronique, la littérature spécialisée est prolifique. Citons entre autres les travaux Tranel (2010), Côté (2005), Bybee (2005), Encrevé (1988), Klausenburger (1984)⁸⁸, Schane (1967), et Delattre (1966).

D'une façon très schématique, on peut remarquer deux grandes orientations définitives ou explicatives ancrées sur le statut de la consonne de liaison (nommé ici CL) :

- Les explications ou justifications pour lesquelles la réalisation contextuelle de la consonne finale latente est mise en avant ;
- Et celles pour lesquelles, inversement, on invoque plutôt des phénomènes d'élision de la consonne finale en contexte de non-liaison.

Ainsi, lorsque la consonne finale est latente, nous relevons trois types d'approches :

- Deux approches explicatives :
 - L'approche par troncation (Schane (1967)) : CL est latente, elle fait partie de la représentation lexicale du mot. En contexte non liaisonnant, les consonnes de liaison subissent une règle d'élision ; $C \rightarrow 0 \mid _C, \#$. La forme liaisonnante est la

⁸⁸ Klausenburger, J. (1984). *French liaison and linguistic theory* (Vol. 10). Franz Steiner Verlag Wiesbaden GmbH.

forme par défaut ;

- L'approche OT (Tranel (2010)) : la forme liaisonnante est l'aboutissement de la hiérarchie suivante :

DEP (C), MAX (C), MAX (V) >> *VV >> GENRE, DEP (L)⁸⁹.

- Une approche descriptive : Delattre (1966) propose une importante typologie de la liaison encore utilisée aujourd'hui⁹⁰. La liaison est la réalisation d'une consonne finale normalement muette lorsque cette dernière est en contact d'un mot commençant par une voyelle. Delattre distingue trois contextes interdisant ou prescrivant la réalisation de la liaison : les liaisons obligatoires, les liaisons facultatives et les liaisons interdites.

Lorsque CL n'est pas latente, nous relevons les traitements des phénomènes de liaisons suivants :

- Les approches par insertion, dans lesquelles CL ne fait pas partie de la représentation lexicale du mot. La liaison y est décrite comme un phénomène d'épenthèse _ on ajoute un segment consonantique entre deux phonèmes vocaliques _ ;
- Les approches supplétives : les mots liaisonnants possèdent deux représentations lexicales ;
- Les approches dans le cadre de la phonologie non-linéaire (Encrevé (1988)) dans lesquelles CL est flottante. La liaison est un phénomène de repositionnement : soit CL se rattache au mot en occupant une attaque vide, soit elle est élidée.

Dans le cadre de nos travaux, nous ne nous interrogerons pas sur le statut de CL. Nous observerons les formes avec et sans réalisation de la liaison des paires potentiellement liaisonnantes présentes dans nos supports d'études. Nous comparerons leurs résultats avec ceux que nous avons relevés pour l'ensemble de nos supports d'études.

En effet, questionner l'ordre préférentiel d'un binôme coordonné nous interroge sur la sélection de deux⁹¹ *output* potentiels : « x&y » et « y&x ».

89 DEP sont des contraintes de FIDÉLITÉ qui s'opposent à l'insertion d'un segment C pour les consonnes épenthétiques et L pour la consonne latente.

MAX ou MAXIMALITÉ s'oppose à la suppression d'un segment V vocalique ou C consonantique.

GENRE est une contrainte hypothétique réclamant un accord en genre.

90 La typologie de la liaison de Delattre est encore utilisée aujourd'hui aux exceptions près de l'évolution de la langue et de l'utilisation des termes plus descriptifs et moins normatifs de Catégorique, Variable et Erratique préconisée par le projet de Phonologie du Français Contemporain (PFC).

91 Dans les faits, ces deux *output* potentiels excluent les questions du schwa et de la liaison. Il faut y ajouter des *output* considérant la réalisation ou non des schwas et ceux celles des formes liaisonnantes ou non.

Au regard des relevés des résultats précédents, notamment en ce qui concerne les frontières vocaliques (initiales et finales), il apparaît que la contrainte *VV motive la sélection d'un ordre préférentiel de coordination : quand « x » finit par une consonne et « y » par une voyelle, l'ordre de coordination retenu est « x&y », car il ne transgresse pas *VV.

Qu'en est-il pour les formes liaisonnantes ?

Si « x » possède une consonne finale et « y » se termine par $_V\{CL\}\#$, quelles stratégies se mettront en place ? La forme « x&y » sera-t-elle à nouveau retenue, favorisant la resyllabation de la consonne fixe ? Ou est-ce la forme « y et x » qui sera privilégiée, signalant une prédominance de la consonne de liaison ?

En fait la question à laquelle nous nous proposons de répondre ici est : la réalisation de la liaison influe-t-elle ou n'influe-t-elle pas sur la sélection d'un ordre de coordination binomiale par « et » ?

Afin de répondre à ces interrogations, nous avons observé les formes préférentielles de 138 paires potentiellement liaisonnantes issues des corpus « TARN ET GARONNE », « CUL ET CHEMISE », « LISSE ET SOYEUX », « VITE ET BIEN » et « ARTS ET MÉTIERS ». Avant de débiter à proprement dit notre analyse, il nous faut prendre en compte trois considérations :

- Comme nous l'indiquions en 2.1. Protocole de constitution des corpus page 34, ces cinq corpus présentent tous, à divers degrés de figement, des expressions collocatives dans le sens statistique du terme. Or, comme le rappelle Bybee (2005), les « collocations de mots fréquemment utilisées ont des représentations très fortes dans la mémoire » provoquant des effets de fréquences impliqués dans la réalisation ou non de la liaison. Les expressions à forte cohésion syntaxique réalisent potentiellement la liaison plus que les autres. De ce point de vue là, les 138 paires que nous avons retenues devraient réaliser la liaison ;
- Dans les contextes étudiés dans le cadre de nos travaux, c'est-à-dire « N et N », « Adj et Adj » et « Adv et Adv », les liaisons pour reprendre la typologie de Delattre (1966), sont facultatives, ou variables. Le locuteur les réalise, ou pas ;
- Enfin, la consonne de liaison est {z} pour 137 paires ⁹² sur 138, et {z} est la marque morphologique du pluriel dans 136 binômes. Ce qui signifie, en d'autres termes, que les deux mots de ces binôme sont affectés par une même marque, et peuvent donc

⁹² La seule exception est {t} dans « TANT ET PLUS ».

potentiellement réaliser la même liaison. CL étant la même, l'examen des finales consonantiques de M1 comparativement à celles de M2 ne peut être concluant.

En fait, dans nos supports d'études⁹³, les réalisations de la liaison et leurs effets dans la sélection d'un ordre préférentiel binaire par « et » sont opaques. Nous ne pouvons que supposer les réalisations de la liaison. Il nous est impossible de les observer directement.

Étant donné que le phénomène que nous souhaitons analyser n'est pas directement accessible, nous devons mettre en place une démonstration par contraposée.

Pour étudier le rôle de la liaison dans la sélection d'un ordre préférentiel de coordination, nous devons observer les résultats des formes potentiellement liaisonnantes quand CL n'est pas réalisée.

Nous posons donc l'implication suivante : si la liaison est active, alors les résultats observés sur les formes non liaisonnées devraient être différents des résultats phonologiques que nous relevions précédemment. Ce qui induit que si les résultats sont identiques, alors la réalisation de la liaison n'a pas de rôle prégnant dans la sélection d'une forme coordonnée.

Cependant, cette démonstration a ses limites. La preuve apagogique fonctionne dans un rapport binaire du type : « si A alors B. Or, non B, donc A ». Or, nous travaillons en interaction de contraintes ce qui signifie qu'il existe une différence entre les deux implications ci-dessus.

En effet, si les résultats sans CL font émerger les mêmes facteurs phonologiques que nos précédents relevés, alors la liaison ne motive pas (ou très peu) l'ordre du composé syndétique par « et ». Mais, si les résultats sans CL sont différents, on ne peut assurer que seule la liaison est responsable de cette divergence. D'autres facteurs, dont les facteurs extraphonologiques présentés page 74 pourraient affecter la forme des paires coordonnées. Il nous faudra prendre en compte cet élément lors de l'analyse de nos résultats.

Ainsi, pour l'examen du rôle de la liaison dans la sélection d'un ordre préférentiel de coordination, nous étudions les segments en frontière de mots et la taille de M1 et M2 des 138 formes liaisonnantes quand CL est non réalisée (nous nommerons désormais ces paires CLNR). L'ensemble des résultats détaillant les répartitions en M1 et en M2 par phonème et par trait distinctif des segments en frontière de mots sont présentés en annexe du tableau A 18

⁹³ Consciente des limites de nos supports d'études sur les questions de liaison, nous sommes convaincue de la nécessité à l'avenir de mettre en place une expérimentation observant les réalisations orales de formes liaisonnantes étudiées.

au tableau A 29. Nous ne proposons ici que les résultats généraux, à savoir la répartition par taille des mots en M1 et en M2 et par type de finales.

Pour l'ensemble des tableaux présentés dans cette section, ci-dessous, ou en annexe, nous mentionnons à nouveau les résultats significatifs ($p \leq 0,05$) en caractères gras. Sont notés en italique ceux pour lesquels nous relevons une tendance ($p \leq 0,1$). Pour information, nous indiquons également les résultats pour lesquels la valeur-p est juste au-dessus du seuil de .1 ($0,11 \leq p \leq 0,149$).

Dans le tableau ci-dessous, nous proposons la répartition en M1 et en M2 des mots en fonction de leur nombre de syllabes.

Ces données ne sont présentées que sur 125 paires. Comme précédemment, nous avons préféré exclure les mots possédant un schwa interne de ce décompte, étant donné que seule la mise en place d'une expérimentation sur l'oral nous renseignerait précisément sur leur réalisation ou non.

CLNR	nombre de σ			
	σ .	$\sigma.\sigma$.	$\sigma.\sigma.\sigma$.	$\sigma.\sigma.\sigma.\sigma$. et +
M1	62 69,66	49 42,98%	7 21,21%	7 50,00%
M2	27 30,34%	65 57,02%	26 78,79%	7 50,00%
Total :	89 100,00%	114 100,00%	33 100,00%	14 100,00%

Tableau 84 : Répartition en M1 et en M2 de CLNR en nombre de syllabes

Le tableau ci-dessus met au jour une répartition en adéquation avec celle que nous présentions page 108.

Les monosyllabes sont 69,66% à privilégier M1 ($\chi^2(1) = 13,76$; $p \leq 0,002$). Inversement, les mots de trois syllabes occupent très massivement M2, dans 78,79% des occurrences testées M2 ($\chi^2(1) = 10,94$; $p \leq 0,001$).

Concernant les facteurs liant la taille des mots à une position préférentielle dans la coordonnée, les résultats de CLNR sont identiques à ceux de TOTAL-TOM ET ZOÉ ce qui pourrait indiquer qu'il n'existe pas d'effets de liaison.

Cependant, comme nous l'indiquions en 4.1. La taille des mots, les facteurs corrélant le nombre de syllabes à une place privilégiée dans les composés syndétiques sont des principes très actifs dans la sélection d'un ordre de coordination. En termes de hiérarchisation, nous pensons qu'ils sont très haut placés. Il est possible que le rôle de liaison soit masqué par la prégnance de ces facteurs.

C'est pourquoi, dans le tableau à venir, nous nous proposons d'examiner les frontières finales des mots. Notre propos est le suivant : si la liaison est un phénomène actif dans la sélection d'un ordre préférentiel de coordination, alors la réalisation de la consonne latente devrait mettre en place une stratégie efficace d'évitement du segment VV.

Ainsi, pour reprendre la démonstration que nous proposons ci-avant, dans des paires où « x » finit par une consonne, et « y » par $_V\{CL\}\#$, les stratégies pour ne pas transgresser *VV n'ont plus lieu d'être. La répartition en M1 et en M2 de $_V\#$ ne serait, dans ce cas, motivée par l'évitement de l'hiatus, et nous devrions observer une répartition des segments finaux différente de celle relevée pour TOTAL-TOM ET ZOÉ.

	$_C\#$	$_V\#$	$_G\#$
M1	89 57,42%	47 41,23%	2 28,57%
M2	66 42,58%	67 58,77%	5 71,43%
Total :	155 100,00%	114 100,00%	7 100,00%

Tableau 85 : Répartition des finales vocaliques pour CLNR

Le tableau ci-dessus fait apparaître des résultats semblables à ceux de TOTAL-TOM ET ZOÉ présentés précédemment en 4.3.1.1. Voyelle. Seules leurs validations statistiques diffèrent : TOTAL-TOM ET ZOÉ met au jour un résultat significatif, nous n'observons ici qu'une tendance plaçant majoritairement les mots $_V\#$ en M2 dans 58,77% des occurrences examinées sur cette hypothèse ($\chi^2(1) = 3,51 ; p \leq 0,06$).

Cette différence de validation statistique pourrait-elle être la trace d'un effet de liaison ?

Notons que les résultats proposés ci-dessus sont des résultats dans l'absolu, ils observent la

répartition des mots possédant une finale vocalique sans considération pour le second élément du binôme. Nous avons donc observé les 114 mots en _V# de CLNR en fonction du terme qui leur est associé.

Or, l'examen de ces paires révèle que 25 paires opposent deux mots possédant une finale vocalique. Nous les avons donc retirées de la répartition des finales vocaliques suivante :

	_V#
M1	22 34,38%
M2	42 65,63%
Total :	64 100,00%

Tableau 86 : Répartition des finales vocaliques pour CLNR sans « _V# & _V# »

Ainsi, conformément aux résultats constatés dans TOTAL-TOM ET ZOÉ, les mots se terminant par une voyelle sont 65,63 % à privilégier M2 ($\chi^2(1) = 6,25 ; p \leq 0,01$).

Cependant, cela ne nous renseigne pas sur le rôle de la liaison dans la sélection d'un ordre préférentiel de coordination. Nous pouvons seulement affirmer que ces résultats sont les indices d'une préférence à resyllaber « et » par la consonne finale fixe plutôt qu'avec la consonne de liaison.

C'est pourquoi, afin de tenter de répondre à notre questionnement initial, nous explorons plus en détail les douze tableaux proposés en annexe.

L'examen rapide des données en annexe ne met au jour que très peu de résultats - seulement cinq répartitions significatives et neuf tendances.

Dans le tableau suivant, nous les comparons avec ceux que nous retenons de l'ensemble des supports d'études sans le corpus TOM ET ZOÉ. Cette comparaison doit mettre au jour leurs similitudes et leurs dissemblances.

Trois types de différences pourraient émerger :

- CLNR ferait émerger un résultat absent de TOTAL-TOM ET ZOÉ ;
- CLNR présenterait des résultats dont la répartition en M1 et en M2 est différente de

celle relevée dans TOTAL-TOM ET ZOÉ ;

- CLNR révélerait des facteurs binaires opposés à ceux relevés dans TOTAL-TOM ET ZOÉ.

Dans les faits, la comparaison des résultats ne rend compte que de la troisième différence : CLNR révèle quelques facteurs binaires opposés à ceux relevés précédemment. Nous avons spécifié ces cas par le caractère « ≠ ».

FACTEURS	CLNR	TOTAL-TOM ET ZOÉ	
#/m/_	M2	M2	
_V[+ NASAL]#	<i>M2</i>	<i>M2</i>	
_V[- BAS]#	M2	_V[+ BAS]# M2	≠
_V[- ARRIÈRE]#	M2	_V[+ ARRIÈRE]# M2	≠
_V[- ARRONDI]#	<i>M2</i>	M2	
_C[- HAUT]#	M1	<i>M1</i>	
_C[- NASAL]#	<i>M1</i>	M1	
_C[- CORONAL]#	<i>M1</i>	_C[+ CORONAL]# M1	≠
_C[- REL RET]#	<i>M1</i>	M1	
V[- HAUT]#	<i>M1</i>	<i>M1</i>	
V[+ BAS]#	<i>M1</i>	_V[- BAS]_# M1	≠
V[+ ARRIÈRE]#	<i>M1</i>	_V[- ARRIÈRE]_# M1	≠
V[+ ARRONDI]#	<i>M1</i>	_V[- ARRONDI]_# M1	≠
_ /ɔ/ _#	M1	M1	

Tableau 87 : confrontation des résultats de CLNR et de TOTAL-TOM ET ZOÉ

D'une manière générale, les répartitions en M1 et en M2 par facteurs étudiés sont identiques. Notons également que si l'on observe les résultats significatifs, seuls deux facteurs diffèrent : ceux relevés pour les mots **_V[- ARRIÈRE]#** et ceux pour les termes **_V[- BAS]#**.

Or, nous avons remarqué page 164 une différence de résultats entre les finales vocaliques orales et les finales vocaliques nasales : les finales vocaliques orales pour lesquelles nous relevions une assignation préférentielle en M2 sont **_ [+ BAS]#** et/ou **_ [+ ARRIÈRE]#**. Les finales vocaliques nasales favorisant M2 sont **_ [- BAS]#** et/ou **_ [- ARRIÈRE]#**. Il est important de rappeler que plus de 60 % des finales vocaliques présentes dans TOTAL-TOM

ET ZOÉ sont des finales vocaliques orales. Dans l'ensemble de nos supports d'études, les corrélations entre M2 et les facteurs $_{[+ \text{BAS}]\#}$ et/ou $_{[+ \text{ARRIÈRE}]\#}$ sont portées par $_{\text{V}[- \text{NASAL}]\#}$.

Il apparaît donc, que les résultats relevés dans CLNR concernant les mots $_{\text{V}[- \text{ARRIÈRE}]\#}$ $_{\text{V}[- \text{BAS}]\#}$ ne sont pas en contradiction avec ceux relevés précédemment.

En ce qui concerne à présent les neuf tendances statistiques dont CLNR rend compte : quatre d'entre elles sont opposées à des résultats présents dans TOTAL-TOM ET ZOÉ.

Rappelons tout d'abord que sont comparées ici des tendances par rapport à des répartitions significatifs. Les facteurs, en fonction de leur validation statistique, n'ont pas la même incidence sur la sélection d'un ordre préférentiel de coordination.

D'autre part, nous sommes convaincue que l'observation des tendances est l'artefact de conflit de contraintes. Ces tendances sont les indices de l'action d'autres principes gouvernant la coordination par « et ». Nous pensons particulièrement aux facteurs extraphonologiques présentés page 74.

La question que nous nous posons ici est de savoir si les effets de la liaison, en opacifiant ces quatre résultats, participent également à ce conflit de contraintes.

En fait, l'examen détaillé de paires CLNR ne nous permet pas de répondre à cette question. Il nous est impossible de confirmer ou d'infirmer l'existence du rôle de la liaison dans la sélection d'un ordre préférentiel de coordination.

En revanche, le relevé que nous venons d'effectuer nous permet d'affirmer que si la liaison est un facteur actif, c'est un facteur moindre. Les facteurs présentés dans l'ensemble de ce travail de thèse sont plus prégnants que les effets de liaison.

Conclusion

Au terme de cette thèse, nous souhaitons présenter une synthèse générale de notre travail.

A travers l'étude des principes gouvernant la coordination par « et » de deux mots en français, nous souhaitons rendre de compte des facteurs actifs dans la sélection de l'ordre préférentiel de séquences binomiales coordonnées par « et » en observant ces facteurs selon un angle essentiellement phonologique.

Pour ce faire, dans le chapitre 2, en accord avec les récents travaux observant la coordination directe de deux mots en anglais - nous pensons précisément à Wright *et al.* (2005, 2002), Renner (2008,2006) et Benor et Levy (2006) - nous avons construit des supports d'études nous permettant de relever des données statistiquement validées par un test d'ajustement.

Ce que nous appelons support d'études est en fait un ensemble hétéroclite de corpus et de questionnaires d'enquête que nous avons mis en place afin de recouvrir le panel le plus représentatif possible des composés syndétiques directs par « et ». Le principe premier de nos supports d'études est que chaque composant, bien qu'autonome, demeure étroitement lié aux autres. Ainsi, les six corpus s'articulent les uns avec les autres : TOM ET ZOÉ est un corpus de prénoms coordonnées par « et », TARN ET GARONNE relève des toponymes, CUL ET CHEMISE, des expressions figées, ARTS ET MÉTIERS, des noms d'institutions, d'écoles, VITE ET BIEN, des collocations dans le sens statistique du terme et LISSE ET SOYEUX des allégations publicitaires.

Parallèlement, les questionnaires d'enquête découlent de cet ensemble. Ils sont tous les six issus de l'exploitation des premiers résultats fournis par le corpus TOM ET ZOÉ. C'est pourquoi, bien qu'ayant longuement hésité à présenter les résultats émis par TOM ET ZOÉ compte tenu des défauts inhérents à sa constitution, nous avons fait le choix de le présenter avec toutes les précautions nécessaires. En effet, ce dernier est non seulement à l'origine de notre travail, mais il reste à notre connaissance le seul corpus mettant en relation de prénoms mixtes directement coordonnées par « et » en français.

Au demeurant, la construction de cet ensemble de supports d'études caractérise notre travail.

Une fois affranchi ces aspects méthodologiques, nous avons organisé l'exploitation de nos données en définissant une liste de critères classificatoires directement inspirés des résultats de la littérature spécialisée concernant la coordination ou la juxtaposition de deux mots en anglais.

Ces critères nous ont permis de mettre au jour deux types de résultats : des résultats extraphonologiques et des résultats phonologiques.

Ainsi, dans le chapitre 3, nous avons révélé la prégnance de facteurs extraphonologiques tels que le facteur sériel, le facteur morphologique, les facteurs sémantiques, les facteurs psycho-affectifs et le facteur associant la fréquence des mots à leur position dans la séquence binomiale coordonnée.

Si les facteurs sériels et les facteurs morphologiques peuvent être l'artefact de facteurs phonologiques comme le facteur de taille, les facteurs sémantiques, les facteurs psycho-affectifs et les facteurs liés à la fréquence des mots ne laissent aucun doute quant à leur prévalence. Ils sont tous trois particulièrement actifs dans la sélection d'un ordre préférentiel de coordination.

De plus, nous sommes convaincue que les facteurs sémantiques, les facteurs psycho-affectifs et la fréquence des mots ne relèvent en fait que d'un seul phénomène : l'opposition marqué / non marqué.

En effet, si l'on considère que la forme prototypique, coutumière est la forme non marquée, il apparaît que les mots les plus fréquents, les plus proches sémantiquement et affectivement sont non marqués. Conformément à ce que relevaient Benor et Levy (2006), l'élément non marqué sélectionne prioritairement M1.

Dans le chapitre 4, nous avons mis au jour un certain nombre de facteurs phonologiques :

- Les mots les plus courts sélectionnent M1.

- Les mots commençant par une voyelle occupent majoritairement M1 ;
- Les mots commençant par la voyelle /e/ évitent M2 ;
- Les mots commençant par /m/ se placent en M2.

- D'une manière générale, les mots possédant une finale vocalique occupent majoritairement M2 ;
- Les mots possédant une finale vocalique nasale sélectionnent plus massivement M2 que les mots finissant par une finale vocalique orale ;
- Les mots $_ /A/\#$ ont tendance à occuper M2 ;
- Les mots en $_ /e/\#$, et en $_ /o/\#$ privilégient également cette même position ;
- Les mots en $_ /ɔ̃/\#$ et en $_ /ã/\#$ sélectionnent massivement M2 ;

- Les mots possédant une finale consonantique occupent majoritairement M1 ;
- Les mots en *_/k/#* sélectionnent prioritairement M1.

- Concernant la structure syllabique en finale de mots :
 - Les mots en *_VC#* sélectionnent M1 ;
 - Tout comme ceux en *_CVC#*.
- Les mots se terminant par une voyelle révèlent que :
 - Les mots en *_CV#* ont tendance à favoriser M2 ;
 - Tout comme ceux en *_CCV#* et *_V#* pour lesquels les résultats sembleraient indiquer une préférence à occuper M2.

- Concernant les effets de l’allongement vocalique, nous n’observons pas de résultats de l’examen des mots assujettis à un allongement vocalique provoqué par une coda allongeante ;
- Les mots en *_V[+ NASAL]C#* occupent préférentiellement M1 ;
- Inversement, ceux en *_/ø, o, α/ | C#* sélectionnent massivement M2.

- Nous ne relevons pas d’effets de liaison.

Notons que comme le signalaient Wright *et al.* (2005), la situation des segments finaux est plus claire que celle des segments initiaux.

D’autre part, il est possible d’organiser ces principes dans la sélection d’un ordre préférentiel de coordination en deux grands axes :

- Les facteurs dans l’absolu ;
- Les facteurs par rapport à « et ».

Nous nommons « facteur dans l’absolu » tous les facteurs actifs que nous avons relevés qui ne sont pas directement liés à « et ». Nous pensons particulièrement au facteur de taille, aux facteurs extraphonologiques, mais également à ceux positionnant les voyelles finales nasales en M2. Ce sont les facteurs les plus actifs dans la sélection d’un ordre préférentiel de coordination, ce sont ceux qui ont mis au jour les résultats les plus forts.

Selon nous, tous ces facteurs, ancrés à gauche, sont instanciés par l’opposition marqué / non

marqué ou la loi du second lourd.

Les facteurs par rapport à « et » sont des facteurs purement phonologiques. L'évitement de *VV, OCP, la maximisation des attaques : tous opèrent par rapport à « et ». Cependant, ils sont moins actifs que les précédents principes.

D'une manière générale, nous pensons que des facteurs phonologiques sont actifs dans la sélection d'un ordre préférentiel de coordination binaire par « et », mais ils semblent de moindre force que les facteurs extraphonologiques.

Afin de les explorer plus en détail, il serait pertinent de mettre en perspective les phénomènes de coordination binaire par « et » en français avec ceux de l'espagnol et de l'italien proposant une alternance « y / e » et « e / ed ».

Enfin, il nous semble que nos travaux devraient être complétés par une étude de terrain observant les réalisations réelles des binômes coordonnés par « et ».

Bibliographie

- ABEILLE A. (2005). « Les syntagmes conjoints et leurs fonctions syntaxiques » dans *Langages*, 39e année, n°160, p. 42-66.
- ADDA-DECKER M. (2006). « De la reconnaissance automatique de la parole à l'analyse linguistique de corpus oraux » dans *JEP*, 389-400.
- ANIS J. (1999). « Chats et usages graphiques » dans *Internet, communication et langue française*, Paris, Hermès Science Publications.
- ANIS J. (2006). « Communication électronique scripturale et formes langagières » <http://rhrt.edel.univ-poitiers.fr/document.php?id=547>
- APOTHELOZ D. (2002). *La Construction du lexique français : Principes de morphologie dérivationnelle*, Paris, édition Ophrys, p. 117-127.
- BECHADE H. D. (1992). *Phonétique et morphologie du français moderne et contemporain*, Presses universitaires de France-PUF.
- BECKMAN J. N. (1997). « Positional faithfulness, positional neutralisation and Shona vowel harmony » dans *Phonology*, 14(01), 1-46.
- BENOR S. B. & LEVY R. (2006). « The chicken or the egg? A probabilistic analysis of English binomials » dans *Language*, p. 233-278.
- BENVENISTE E., (1970). « L'appareil formel de l'énonciation », dans *Langages*, vol. 5, n°17, p. 12-18
- BIRDSONG D. (1995). « Iconicity, Markedness, and Processing Constraints in Frozen Locutions » dans *Syntactic Iconicity and Linguistic Freezes. The Human Dimension*, Berlin, M. E. Landsberg (ed.), p. 31-45.
- BRANDÃO DE CARVALHO J. (1997). « Primitives et naturalité » dans *Langages*, 31e année, n° 125. *Nouvelles phonologies*, p. 14- 34.
- BRANDÃO DE CARVALHO J., NGUYEN N. & WAUQUIER S. (2010). *Comprendre la phonologie*. Paris : Presses universitaires de France.
- BYBEE J. L. & HOPPER P. J. (eds). (2001). *Frequency and the emergence of linguistic structure* (vol. 45). John Benjamins Publishing.

- BYBEE J. (2005). « La liaison : Effets de fréquence et constructions » dans *Langages*, 159, 24-37.
- CALABRESE A. (1995). « A Constraint-Based Theory of Phonological Markedness and Simplification Procedures » dans *Linguistic Inquiry*, 26(3), 373-463.
- CHOMSKY N., HALLE, M. (1968). *The sound pattern of English*. New York : Harper & Row.
- CLEMENTS G. N. (1997). « Berber Syllabification: Derivations or Constraints? », dans I. Roca (éd.), *Derivations and Constraints in Phonology*, Oxford, Clarendon Press, p. 289-330.
- CLEMENTS G. N. (2003). « Feature economy as a phonological universal » dans *15th International Congress of Phonetic Sciences*, Barcelone.
- CLEMENTS G. N. (2007). « L'Évitement de la marque : Vers un nouveau modèle des inventaires phonologiques » dans E. Delais-Roussarie & L. Labrune (éds), *Des sons et des sens : Données et modèles en phonologie et morphologie*, 25-47. Paris & Londres : Hermès.
- CLEMENTS G. N. (2009). « The role of features in phonological inventories » dans *Contemporary views on architecture and representations in phonological theory*, p. 19-68.
- COOPER W. & ROSS J. (1975). « Word order », dans *Papers from the parasession on functionalism*, Chicago, Chicago Linguistic Club, p. 63-111.
- CÔTÉ M. H. (2005). « Le statut lexical des consonnes de liaison », dans *Langages*, (2), p. 66-78.
- CÔTÉ M. H. (2010). « La longueur vocalique devant consonne allongeante en contexte final et dérivé en français laurentien » dans Carmen LeBlanc, France Martineau & Yves Frenette, éd. *Vues sur les français d'ici*. Québec : Presses de l'Université Laval, 49-75.
- COUASNON G. (2009). *Étude des principes phonologiques gouvernant la coordination de deux prénoms en français*, Mémoire de Travaux d'étude et de recherches.
- COUASNON G. (2012). « Une étude des principes gouvernant la coordination par « et » de deux prénoms en français » dans *SHS Web of Conferences* (vol. 1, p. 1455-1469). EDP Sciences.
- CRISTO A. D. (2000). « Vers une modélisation de l'accentuation du français » dans *Journal of French Language Studies*, 10(01), 27-44.

- CRISTO A. D. (2003). « De la métrique et du rythme de la parole ordinaire : L'exemple du français » dans *Semen. Revue de sémio-linguistique des textes et discours* (16).
- DELAIS-ROUSSARIE E. (1999). « Accentuation et réalisation des clitiques en français » dans *Cahiers de Grammaire*, 24, 17-37.
- DELAIS-ROUSSARIE, E. (2000). « Vers une nouvelle approche de la structure prosodique » dans *Langue française*, 126, 92-112.
- DELATTRE P. (1938). « L'accent final en français : Accent d'intensité, accent de hauteur, accent de durée » dans *The French Review*, 12(2), 141-145.
- DELATTRE P. (1944). « L'aperture et la syllabation phonétique » dans *The French Review*, 17(5), 281-285.
- DELATTRE P. (1959). « Rapports entre la durée vocalique, le timbre et la structure syllabique en français » dans *The French Review*, 32(6), 547-552.
- DELATTRE P. (1965). « La nasalité vocalique en français et en anglais » dans *The French Review*, 39(1), 92-109.
- DELATTRE P. (1966). *Studies in French and comparative phonetics: Selected papers in French and English* (vol. 18). Walter de Gruyter GmbH & Co. KG.
- DELATTRE P. (1968). « Divergences entre nasalités vocalique et consonantique en français » dans *Word*, 24(1-3), 64-72.
- DELL F., HIRST D. & VERGNAUD J. R. (éds). (1984). *Forme sonore du langage* (vol. 1411). Herman-Éditeurs des Sciences et des Arts.
- DELL F. (1985). *Les Règles et les sons : Introduction à la phonologie générative*. Paris : Hermann.
- DELVAUX V., METENS T. & SOQUET A. (2002). « Propriétés acoustiques et articulatoires des voyelles nasales du français » dans *XXIVe Journées d'étude sur la parole, Nancy, 1*, 348-352.
- DURAND M. (1953). « De la formation des voyelles nasales » dans *Studia Linguistica*, 7(1 - 2), 33-53.
- DURAND J. (2005). « Les primitives phonologiques : Des traits distinctifs aux éléments » dans *Phonologie et phonétique: forme et substance*. Paris: Hermès, 63-93.

- DURAND J. (2014). « À la recherche du schwa : Données, méthodes et théories » dans *SHS Web of Conferences* (vol. 8, p. 23-43). EDP Sciences.
- DURAND J., LYCHE C. (2008). « French liaison in the light of corpus data » dans *Journal of French Language Studies*, 18-1.
- ENCREVE P. (1988). *La Liaison avec et sans enchaînement*, Paris, Le Seuil.
- EYCHENNE J. (2006). *Aspects de la phonologie du schwa dans le français contemporain. Optimalité, visibilité prosodique, gradience*, Doctoral dissertation, PhD thesis, Université de Toulouse-Le Mirail, France.
- FENK-OCZLON G. (1989). *Word frequency and word order in freezes*, dans *Linguistics*, 27.517-56.
- FÉRY C. & VAN DE VIJVER R. (2003). *The Syllable in Optimality Theory*, Cambridge, Cambridge University Press.
- FOUCHÉ P. (1956). *Traité de prononciation française*, Paris, Klincksieck.
- FOUGERON C. (1998). *Variations articulatoires en début de constituants prosodiques de différents niveaux en français* (Doctoral dissertation, Université Paris).
- GADET F. (1994). « La genèse du concept de marque » dans *Cahiers de l'ILSL*, 5, p.81-92.
- GEERTS T. (2011). « Ch (e) va : Schwa français en syllabe initiale » dans *Langue française*, (1), p. 39-54.
- GEORGETON L., PAILLERAU N., LANDRON S., GAO J. & KAMIYAMA T. (2012, juin). « Analyse formantique des voyelles orales du français en contexte isolé : A la recherche d'une référence pour les apprenants de FLE » dans *Conférence conjointe JEP-TALN-RECITAL 2012* (p. 145-152).
- GOAD H., & BUCKLEY, M. (2006). « Prosodic structure in child French: Evidence for the Foot » dans *Catalan Journal of Linguistics*, 5(1), p.109-142.
- GRAMMONT M. (1914). *Traité pratique de prononciation française*, Paris, Delagrave, éd. 1963.
- GRAMMONT M. (1933). *Traité de phonétique*, Paris, Delagrave.
- GROSS G. (1996). *Les Expressions figées en français*, Paris, Ophrys.
- GUSTAFSSON M. (1976). « The frequency and frozenness of some English binomials », dans *Neuphilologische Mitteilungen*, 75, 663-77.

- HAGÈGE C. (1985). *L'Homme de paroles : Contribution linguistique aux sciences humaines*, Fayard, Le temps des sciences, Paris, p. 244-249.
- HATHOUT. N., NAMER F., PLÉNAT M. & TANGUY L. (2009). « La collecte et l'utilisation des données en morphologie » dans *Aperçus de morphologie du français*. Saint-Denis, Presses universitaires de Vincennes.
- HAYES B. (1995). *Metrical stress theory: Principles and case studies*, University of Chicago Press.
- HERDAN G. (1960). *TYPE _ TOKEN mathematics – a textbook of mathematical linguistics*, La Hague, Mouton et Co's Gravenhage.
- HILAIRE-DEBOVE G. & KEHOE M. (2004). « Acquisition des consonnes finales (codas) chez les enfants francophones : Des universaux aux spécificités de la langue maternelle » dans *Actes de la 25e Journée d'Études sur la Parole*, 265-268.
- HOBÆK HAF M. (1990). « Coordonnants et éléments coordonnés : Une étude sur la coordination en français moderne » dans *L'information grammaticale*, 46(1), 17-21.
- JAKOBSON R., FANT G. & HALLE M. (1952). *Preliminaries to Speech Analysis*. Cambridge, MA : MIT Press.
- JUN, S. A., & FOUGERON, C. (2000). « A phonological model of French intonation » dans *Intonation*, Springer Netherlands, p. 209-242.
- KILGARIFF A. & GREFFENSTETTE G. (2003). « Introduction to the special issue on the Web as corpus », dans *Computational Linguistics* 29-3 : 333-347.
- LABBÉ D. (2007). « Coordination et subordination en français oral » dans *Les Nouvelles Journées de l'ERLA, n° 4, Coordination/subordination dans le texte de spécialité*, L'Harmattan, p. 161-182.
- LABRUNE L. (2005). « Autour de la syllabe : Les constituants prosodiques mineurs en phonologie », dans *Phonétique et phonologie, approches contemporaines*, Nguyen N., Wauquiers S., Durand J. (éds), Paris, Hermes Science Publications, Lavoisier, p. 95-116.
- LABRUNE L. (2006). « Patterns of phonemic preferences in Japanese non-headed binary compounds: what waa-puro, are-kore and mecha-kucha have in common », *Gengo Kenkyuu, Journal of the Linguistic Society of Japan*, 129.
- LABRUNE L. (2007). « Morphologie des emprunts abrégés en japonais : Aspects prosodiques et mélodiques », dans *Des sons et des sens. Données et modèles en phonologie et*

- en morphologie*, Delais-Roussarie E., Labrune L. (éds), Paris, Hermes Science Publications, Lavoisier, p. 173-197.
- LACHERET A., LYCHE C. & TCHOBANOV A. (2011). « Schwa et position initiale revisités : L'éclairage de la prosodie en phonologie du français contemporain » dans *Langue française*, (1), p. 137-158.
- LAKS B. (1997). *Phonologie accentuelle : Métrique, autosegmentalité et constituance*. CNRS Editions.
- LAMBERT F. (1991). « Observations sur la coordination par et en français », dans *Cahiers de Grammaire*, n° 16, Toulouse, Université de Toulouse le Mirail ERSS-CNRS, p. 75-102.
- LAMBERT F. (2001). « Le et temporel est-il temporel » dans *Cahiers de grammaire*, n° 26, Toulouse, Université de Toulouse le Mirail ERSS-CNRS, p. 143-163.
- LAMBERT F. (2006). « Pourquoi met-on une virgule devant et ? », dans *Cahiers de grammaire*, n° 30, Toulouse, Université de Toulouse le Mirail ERSS-CNRS, p. 205-218.
- LÉON P. (1992). *Phonétisme et prononciations du français avec des travaux d'application et leurs corrigés*, Paris, Nathan.
- LOMBARDI L. (2001). *Segmental Phonology in Optimality Theory, Constraints and Representations*, Cambridge, Cambridge University Press.
- LYCHE C. (2005). « Des règles aux contraintes : Quelques aspects de la théorie de l'optimalité » dans *Phonétique et phonologie, approches contemporaines*, Nguyen N., Wauquiers S., Durand J. (éds), Paris, Hermes Science Publications, Lavoisier, p. 209-237.
- MAGNE B. (2005). « Expressions figées et fonctionnements des contraintes dans La vie mode d'emploi : L'exemple des couples », dans *SEMEN* n°19, l'ordre des mots. URL : <http://semen.revues.org/2106>
- MALKIEL Y. (1959). « Studies in irreversible binomials », dans *Lingua* 8.
- MARTIN P. (2004). « Intonation de la phrase dans les langues romanes : L'exception du français » dans *Langue française* (1), 36-55.
- MCCARTHY J. (1986). « OCP Effects: Geminata and Antigeminata » dans *Linguistic Inquiry* 17, 207-263
- MCCARTHY J. (2004). *Optimality Theory In Phonology*, Oxford, Blackwell Publishing.

- McCARTHY J. & PRINCE A. (1997). « L'émergence du non-marqué. L'optimalité en morphologie prosodique » dans *Langages*, vol. 31, n° 125, p. 55-99.
- MEYNADIER Y. (2001). « La syllabe phonétique et phonologique : Une introduction » dans *Travaux Interdisciplinaires du Laboratoire Parole et Langage d'Aix-en-Provence (TIPA)*, n° 20, 91-148.
- MONTREUIL J. P. (2000). « Inventaires et contraintes consonantiques » dans *Langue française*, (126), 73-91.
- MORIN Y. C. (1986). « La loi de position ou de l'explication en phonologie historique » dans *Revue québécoise de linguistique*, 15(2), 199-231.
- MORIN Y. C. (2003). « Remarks on prenominal liaison consonants in French » dans S. Ploch (ed.), *Living on the Edge. 28 Papers in Honour of Jonathan Kaye*, Berlin : Mouton de Gruyter.
- MOURET F., ABEILLE A., DELAIS-ROUSSARIE E., MARANDIN J. M. & YOO H. (2008). « Aspects prosodiques des constructions coordonnées du français » dans *Actes des XXVIIIe Journées d'études sur la Parole*, Avignon.
- MULLER C. (1993). *Initiation aux méthodes de la statistique linguistique*, Genève, Honoré Champion Editeur, coll. « Unichamp ».
- NGUYEN N., WAUQUIER-GRAVELINES S. & DURAND J. (2005). *Phonologie et phonétique, forme et substance*, Paris, Hermes Science Publications, Lavoisier.
- PANCKHURST R. (1997). « La communication médiatisée par ordinateur ou la communication médiée par ordinateur ? » dans *Terminologies nouvelles*, 17, p. 56-58.
- PARADIS C., PRUNET J.-F. (1991). *The Special Status of Coronals: Internal and External Evidence*, Academic Press, Inc., San Diego.
- PLÉNAT M. (1994). L'« extramétrie » des voyelles initiales. *French Generative Phonology: Retrospective and Perspective*, Salford, University of Salford, AFLS & ESRI, 239-257.
- PLÉNAT M. (1996). « De l'interaction des contraintes : Une étude de cas », dans *Current Trends in Phonology: Models and Methods*. Salford : ESRI, J. Durand & B. Laks (éds), University of Salford, vol. 2, p. 585-615.
- PLÉNAT M. (1997). « Quelques thèmes de recherche actuels en morphophonologie française », dans *Cahiers de lexicologie*, n°77, p. 27-62

- PLÉNAT M. (1999). « Prolégomènes à une étude variationniste des hypocoristiques à redoublement en français », dans *Cahiers de grammaire* n° 24 (235 p.), p. 183-219.
- PLÉNAT M. (2008). « La liaison « obligatoire » avec et sans enchaînement », dans *Congrès Mondial de Linguistique Française*, EDP Sciences, p. 149.
- PLÉNAT M. (2009). « Les contraintes de taille », *Aperçus de morphologie du français*, dans Fradin B., Kerleroux F. et Plénat M. (éds), Saint-Denis, Presses universitaires de Vincennes.
- PLÉNAT M. & SOLARES HUERTA P. (2006). « Domi, Seb, Flo et toute la famille », dans *Cahiers de la Grammaire*, n° 30, p. 331-357.
- PINKER S. & BIRDSOONG D. (1979). « Speakers sensitivity to rules of frozen word order », dans *Journal of Verbal Learning and Verbal behaviour* 18, n° 4, p. 497-508.
- PRINCE A. & SMOLENSKY P. (1993). *Optimality Theory, Constraint Interaction in Generative Grammar*, Ms, Rutgers Un. & Un. du Colorado.
- RENNER V. (2006). « Les composés coordinatifs en anglais contemporain », Thèse de doctorat.
- RENNER V. (2008). « On the semantics of English coordinate compounds » dans *English Studies*, 89(5), 606-613. doi:10.1080/00138380802011982.
- RENNER V. (2014). « A Study if Element Ordering in English Coordinate Lexical Items », dans *English Studies*, DOI : 10.1080/0013838X.2014.908039.
- RENNER V. & FERNÁNDEZ-DOMÍNGUEZ J. (2011). « Coordinate compounding in English and Spanish » dans *Poznań Studies in Contemporary Linguistics*, 47(4), 873-883. doi:10.2478/psicl-2011-0041.
- RIDOUANE R., MEYNADIER Y., & FOUGERON C. (2011). « La syllabe: objet théorique et réalité physique », dans *Faits de langue*, n° 37, 225-246.
- RIZZOLO O. (2002). « Du leurre phonétique des voyelles moyennes en français et du divorce entre Licenciement et Licenciement pour gouverner », Unpublished PhD thesis, Université de Nice-Sophia Antipolis.
- SAUZET P. (2004). « La singularité phonologique du français » dans *Langue française* (1), 14-35.
- SCHANE S. (1967). « L'élision et la liaison en français » dans *Langages*, 8.

- SCHEER T. (2000). « L'immunité de schwa en début de mot » dans *Langue française*, n° 126, p. 113-126.
- SPENCE N. C. (1988). « Loi de position ou durée vocalique ? » dans *Revue québécoise de linguistique*, 17(1), 223-234.
- TANGUY L. & HATHOUT N. (2002). « Webaffix : Un outil d'acquisition morphologique dérivationnelle à partir du Web », dans *Actes de Traitement Automatique des Langues 02*, Nancy, INRIA, p. 245-254.
- TUTIN A., GROSSMANN F. (2002). « Collocations régulières et irrégulières : Esquisse de typologie du phénomène collocatif » dans *Revue française de linguistique appliquée*, 7(1), 7-25.
- TUTIN A., GROSSMANN F. (2003). *Les collocations : Analyse et traitement*. Amsterdam, De Werelt.
- TRANEL B. (1987). *The sounds of French: An introduction*. Cambridge university press.
- TRANEL B. (1987). « French schwa and nonlinear phonology » dans *Linguistics*, 25(5), 845-866.
- TRANEL B. (2000). « Aspects de la phonologie du français et la théorie de l'optimalité », dans *Langue Française*, n° 126, p. 39-72.
- TROUBETZKOY N. S. (1970). *Principes de phonologie*, Paris : Klincksieck.
- VAISSIÈRE J. (2000). « Changements de sons et changements prosodiques : Du latin au français » Version corrigée le 24 septembre 2001. dans *Revue Parole*, 15(16), 53-87.
- VÉRONIS J. & GUIMIER DE NEEF E. (2006). « Le traitement des nouvelles formes de communication écrite » dans SABAH G. (éd.), *Compréhension automatique des langues et interaction* (p. 227-248). Paris: Hermès Science.
- VILLOING F. (2012). « Contraintes de taille dans les mots composés : Quand la phonologie entre en concurrence avec les contraintes morphologiques » dans *SHS Web of Conferences* (vol. 1, p. 1425-1440). EDP Sciences.
- WIOLAND F. (1985). *Les structures syllabiques du français : Fréquence et distribution des phonèmes consonantiques, Contraintes idiomatiques dans les séquences consonantiques*. Slatkine-Champion, Genève-Paris.
- WIOLAND F. (1991). *Prononcer les mots du français : Des sons et des rythmes*, Paris, Hachette.

WRIGHT S., HAY J. & BENT T. (2002). « Fred and Wilma. A phonological conspiracy », dans *Gender and Linguistic Practice*, Stanford CA, éd. Benor S., Rose M., Sharma D., Sweetland J. et Zhang G., p. 175-91.

WRIGHT S., HAY J. & BENT T. (2005). « Ladies First? Phonology, Frequency, and the Naming Conspiracy », dans *Linguistics*, vol. 43, n° 3, p. 531-561.

YAMAGUCHI N. (2007). « Acquisition des traits phonologiques : Éclairage par les données », dans *Actes des IXe Rencontres Jeunes Chercheurs de l'Ecole Doctorale 268 « Langage et Langue »*, 80-83.

YAMAGUCHI N. (2015). « L'acquisition phonologique, de Jakobson aux modèles fréquentiels », dans *Langages*, (2), 31-50.