

Thèse de doctorat
de l'Université Sorbonne Paris Cité
Préparée à l'Université Paris Diderot
Ecole Doctorale Sciences du Langage (ED 132)
Équipe d'accueil CLILLAC-ARP (EA 3967)

**ÉLABORATION DE CRITÈRES PROSODIQUES
POUR UNE ÉVALUATION SEMI-AUTOMATIQUE
DES APPRENANTS FRANCOPHONES DE L'ANGLAIS**

Par **Évelyne CAUVIN**

Thèse de doctorat de linguistique anglaise

Sous la direction de Nicolas Ballier

Présentée et soutenue publiquement à Paris le 4 décembre 2017

Président du jury : Hanote, Sylvie / Professeur des Universités / Université de Poitiers

Rapporteurs : Herment, Sophie / Professeur des universités / Aix-Marseille Université

Hirst, Daniel / Directeur de recherche émérite CNRS / Aix-Marseille Université

Examineur : Horgues, Céline / Maître de conférences / Université Sorbonne Nouvelle - Paris 3

Directeur de thèse : Ballier, Nicolas / Professeur des universités / Université Paris Diderot - Paris 7

Titre : Élaboration de critères prosodiques pour une évaluation semi-automatique des apprenants francophones de l'anglais

Résumé :

L'objectif de cette thèse est de modéliser l'interlangue prosodique des apprenants francophones de l'anglais afin de pouvoir élaborer des critères utilisables pour une évaluation semi-automatique de leur niveau prosodique. Le domaine évaluatif requiert la plus grande rigueur dans la mise en place de ses critères pour aboutir à la validité, la fiabilité, la faisabilité et l'équité maximales, alors que la prosodie anglaise de la langue cible se caractérise par son extrême variabilité. Aussi, peu d'études se sont engagées dans l'évaluation de la prosodie, qui représente une réelle gageure. Pour relever ce défi, une stratégie particulière a été mise en place pour élaborer une méthodologie permettant d'atteindre l'objectif fixé, en lecture.

L'approche choisie repose sur la symbiose permanente qu'entretient la prosodie avec le monde dans lequel évolue le locuteur. Cette méthodologie, ou « profilage », est destinée à sélectionner par inclusion ou exclusion les éléments analysés tant au niveau perceptif qu'acoustique. Le profilage des réalisations sur l'axe syntagmatique permet de sélectionner les locuteurs natifs servant de modèles, et celui basé sur le phénomène d'emphase rend possible un ciblage de leurs réalisations les plus pertinentes à modéliser sur l'axe paradigmatic. Conformément à cette méthodologie d'investigation nouvelle et aux résultats perceptifs et acoustiques concordants pour la langue cible, les réalisations des apprenants francophones du corpus *Longdale-Charliphonia* sont analysés acoustiquement. Le classement automatique à partir des variables prosodiques (acoustiques et perceptives) est confronté à celui d'experts évaluant par perception classique.

Les travaux de cette thèse aboutissent essentiellement à :

- Une modélisation de la prosodie anglaise non native par grilles évaluatives critériées s'appuyant sur critères distinctifs natifs et non natifs issus de variables temporelles (vitesse d'élocution avec ou sans pauses), de registre et de mélodie, ainsi que de rythme,
- À partir de ces variables, une évaluation semi-automatisée de 15 apprenants représentatifs du corpus par classement et notation,
- Une correspondance des résultats de l'évaluation traditionnelle avec celle semi-automatique évoluant entre 56,83% et 59,74% dans une catégorisation des apprenants en 3 niveaux de maîtrise, en fonction du profilage d'experts évaluateurs.

Mots clefs : évaluation semi-automatique / traits critériés / apprenants francophones / prosodie anglaise / intonation / rythme / vitesse de lecture / profilage / emphase / saillance

Title : Devising prosodic criteria for a semi-automatic assessment of Francophone learners of English

Abstract :

The aim of our study is to modelise the prosodic interlanguage of Francophone learners of English in order to provide useful criteria for a semi-automatic assessment of their prosodic level in English. Learner assessment is a field that requires to be very rigorous and fair when setting up criteria that ensure validity, reliability, feasibility and equality, whereas English prosody is highly variable. Hence, few studies have carried out research in assessing prosody because it represents a real challenge. To address this issue, a specific strategy has been devised to elaborate a methodology that would ensure assessing a reading task successfully.

The approach relies upon the constant symbiosis between prosody and a speaker's subjective response to their environment. Our methodology, also known as « profiling », first aims at selecting relevant native perceived and acoustic prosodic features that will optimize assessment criteria by using their degree of emphasis and creating speakers' prosodic profiles. Then, using the *Longdale-Charliphonia* corpus, the learner's productions are analysed acoustically. The automatic classification of the learners based on acoustic or perception prosodic variables is then submitted to expert aural assessment which assesses the learner evaluation criteria.

This study achieves:

- A modelisation of non-native English prosody based on assessment grids that rely upon features of both native and non-native speakers of English, namely, speech rate – with or without the inclusion of pauses, register, melody and rhythm,
- A semi-automatic evaluation of 15 representative learners based on the above modelisation – ranking and marking,
- A comparison of the semi-automatic results with those of experts' auditory assessment; correspondence between the two varies from 56.83% to 59.74% when categorising the learners into three prosodic proficiency groups.

Keywords : semi-automatic assessment / criterial features / French-speaking learners / English prosody / intonation / rhythm / speech rate / profiling / emphasis / saliency

« Caminante, son tus huellas
el camino y nada más.
Caminante, no hay camino,
se hace camino al andar. »

Antonio Machado, *Soledades* (1936)

Voyageur, ce ne sont que les traces de tes pas
qui donnent vie au chemin, sans plus.
Voyageur, il n'existe pas de chemin,
c'est en le foulant que tu le crées.
Traduction É. Cauvin

Je souhaite dédier ce travail de recherche à deux personnes qui m'ont ardemment encouragée dans cette voie : Messieurs Lionel Guierre et Georges Boulakia. Je pense les avoir profondément déçus en ne répondant pas à leur attente il y a de cela trente ans. Je me devais de réparer cette faute.

Mon parcours dans l'enseignement à tous les niveaux depuis tant d'années, et maintenant plus que jamais, m'a montré le lien fondamental que devait préserver tout enseignant entre sa pratique de terrain et sa recherche dans la compréhension des phénomènes en jeu. L'étude suprasegmentale de l'anglais est donc tout naturellement devenue pour moi un passage obligé.

Remerciements

En tout premier lieu, je tiens à remercier très sincèrement mon directeur de recherche, Monsieur Nicolas Ballier, sans lequel cette thèse n'aurait jamais vu le jour. Il a suivi ses convictions en m'offrant les conditions optimales pour la mener à bien. Son soutien indéfectible m'a permis de réaliser un projet longtemps retardé.

J'exprime ma plus profonde gratitude à Madame Sophie Herment, référente de thèse, qui, par ses encouragements, sa patience et la justesse de sa vision, m'a guidée à des moments charnières de la réalisation de ce travail.

Je remercie les membres du jury de me faire le grand honneur de lire et d'évaluer la présente thèse : Mesdames Sylvie Hanote et Céline Horgues, Monsieur Daniel Hirst.

Cette thèse est un travail d'équipe, comme le prouve le nombre important de tests perceptifs, et conséquemment de volontaires qui ont accepté d'être enregistrés ou de participer à ces tests comme auditeurs. Je les remercie du fond du cœur. Que ceux, nombreux, qui ne trouvent pas leur nom parmi ceux de cette liste me pardonnent. Je remercie nommément Mesdames et Messieurs Reynald Adjahi, Maëlle Amand, Myriam Annacer, Nicolas Ballier, Nadia Bacor, Jiewon Baek, Grant Raymond Barrett, Róisín Barry, James Batchelor, Sylvie I. Bebohi, Carmen Berechet, Pierre Berrezig, Cécile Bianchi, Elizabeth Blot, Caroline Boclet, Mariana Bonnouvrier, Agnès Botz, Caroline Bouzon, Hélène Bresson, Neil James Broadbent, Georges Boulakia, Bernadette Bourmorcq, Jane Brothwell, Françoise Brown, Hélène Bultelle, Christopher Bushell King, Daniel Cahill, Alan Casavant, Joy Chan, Aurélie Ceccaldi, Clara Cohen, Alain Corbière, Stephen Dadds, Shihong Dai, Christine Degeorge, Élodie De Wever, Alain Diana, Jenny Dixon, Nicole Dunham, Valentin Dupont, Emmanuelle Durieux, Elizabeth Dutertre, Gaëlle El Ghezal, Kambiz Elhami, Marie Elven, Antoine Ertlé, Christelle Exare, Emmanuel Ferragne, Kevin Ferragut, Kevin Ferris, Pierre Fournier, Bernard Galtier, Joëlle Garault, Arnaud Gazeaux, Joy Gelardi, Hélène Ginard, William Gleeson, Pascale Goutereaux, Mark Gray, Alice Hall, Trevor Harris, Michael Harvey, Sophie Herment, Daniel Hirst, Joséphine Hodgson-Ancel, Daniel Hoffmann, Jeffrey Hopes, Céline Horgues, Shena R. House, Ibrahim Ismaeil, Kathryn Jones, Jean-Paul Jullien, Christopher Kent, Stephan Kraitsowits, Soufiane Lamchaouri, Marie Laniel, Morgan Le Nôtre, Sébastien Le Pipec,

Nicolas Lecointre, Brian Lowrey, Chantal Maciborko, Didier Madelaine, Tim Mahrt, Céline Mansanti, Declan McCavana, Susan Moore Mauroux, Dominique Morel, Benjamin Newman, David Nowell-Smith, Donnacha O’Ceallaigh, Martin O’Connor, Martin Parsons, Michelle Phillips, Patricia Picard, Andrea Poret Gilbert, Anne Poulain, Michelle Preddy, Keith Randall, Émile Rat, Oriana Reid-Collins, Patricia Rousseau, Kate Roussel, Megan Roussel, Sohrab Salami, Aline Salvat, Nathalie Saudo-Welby, Steven Schaefer, Diane Schmitt, Martine Schuwer, Michelle Selliez, Jessica Shore, Isabelle Sicard, Emma Simmons, Frédérique Spill, Jessica Stark, Adam Stephenson, Grace Stockton-Bliard, Corine Tachtiris, Anne Talbot, Ann Tardy, Henry Tavernier, Maria Thomas, Anne Tortel, Marc Umdenstock, Carla Vale, Ana Vioque, Jane Weston, Laurence Whiteside, Shona Whyte, Heidi Wild, Shân Williams, Emilia Wilton-Godberfforde, Aeysha Wing, John Richard Wisdom, Lei Yang, Iyonna Naomi Zortman, ainsi que les personnes dont je n’ai pu retrouver le nom de famille : Angelina, Bronagh, Cathryn, Daniel, Eleanor, Elinor, José, Laurence, Naomi, Nicole, Ned, Philip, Steve, William.

Je remercie en outre tous ceux qui m’ont soutenue en me prodiguant des conseils d’ordre technique à des moments cruciaux de ce travail de recherche, et plus particulièrement Maëlle Amand, Ismaïl Benali, Georges Boulakia, Alain Diana, Emmanuel Ferragne, Philippe Martin, Yves Morda, Aline Salvat et Éric Salvat.

Abréviations, sigles et acronymes

Un rappel des 28 variables créées figure en annexe ainsi que des tableaux d'équivalence des lecteurs natifs et non natifs et des expérimentations mises en œuvre.

AAVE : African American Vernacular English

ACM : Analyse en Composantes Multiples

ACP : Analyse en Composantes Principales

ADELEX : (projet de recherche) Assessing and Developing Lexical Competence through the Internet (Université de Grenade)

AFC : Analyse Factorielle par Correspondances

ALTE : Association of language testers in Europe

ANN-TON-N1 : première expérimentation portant sur l'annotation tonale (syllabique) chez les natifs (Chapitre III)

ANN-TON-N2 : deuxième expérimentation portant sur l'annotation tonale (syllabique) chez les natifs (Chapitre III)

B1 : (niveau de maîtrise de la langue selon les critères du Cadre Européen Commun de Référence pour les Langues du Conseil de l'Europe) B1 correspond au niveau « indépendant seuil ».

B2 : (niveau de maîtrise de la langue selon les critères du Cadre Européen Commun de Référence pour les Langues du Conseil de l'Europe) B2 correspond au niveau « indépendant avancé ».

C1 : (niveau de maîtrise de la langue selon les critères du Cadre Européen Commun de Référence pour les Langues du Conseil de l'Europe) utilisateur « expérimenté autonome ».

CAED : Cambridge American English Dictionary

CBED : Cambridge British English Dictionary

CECRL ou **CECR** : Cadre européen commun de référence pour les Langues : Apprendre, Enseigner, Évaluer

CEFR : Common European Framework of Reference for Languages: Learning, Teaching, Assessment

COR-AC-N1 : expérimentation portant sur le premier recueil de corrélats (acoustiques) chez les natifs (Chapitre II)

COR-AC-N2 : expérimentation portant sur le deuxième recueil de corrélats (acoustiques) chez les natifs (Chapitre III)

COR-AC-NN : expérimentation portant sur le recueil de corrélats acoustiques chez les non natifs (Chapitre IV)

COR-DIV-N3 : expérimentation portant sur le troisième recueil de corrélats (de nature diverse) chez les natifs (Chapitre III)

CRIT-LEC-NN : expérimentation de création de critères distinctifs en lecture chez les non natifs (Chapitre IV)

D1 : (catégorie d'apprenants du) premier décile (en vitesse de lecture globale)

D9 : (catégorie d'apprenants du) neuvième décile (en vitesse de lecture globale)

DFL : Dictionnaire de français Larousse

EDRSAM : Encyclopédie de Diderot et d'Alembert

EEG : électroencéphalographie

EPP : English Profile Programme

EXP-LEC-N : expérimentation portant sur l'expression des natifs enregistrés (Chapitre II)

FIDIS : Consortium Future of Identity in the Information Society

GDT : Le grand dictionnaire terminologique de l'Office québécois de la langue française

IP (pluriel **IPs**) : syntagme intonatif (*intonation phrase*)

IRM : imagerie par résonance magnétique

L1 : langue maternelle

L2 : langue étrangère (éventuellement langue seconde)

LANSAD : LANGues pour Spécialistes d'Autres Disciplines

LDOCE : Longman Dictionary of Contemporary English

LeaP corpus : corpus résultant du projet Learning Prosody in a Foreign Language

LFP : Lexical Frequency Profile

LONGDALE : Longitudinal Database of Learner English

LSWE corpus : Longman Spoken and Written English Corpus

Méd : (catégorie d'apprenants de la) médiane (en vitesse de lecture globale)

MEG : magnétoencéphalographie

MMD : Macmillan Dictionary

MWD : Merriam-Webster's Online Dictionary

NPIs : negatively-oriented polarity-sensitive items (classe des composés de any selon Huddleston et Pullum, 2002 : 823)

OED : Oxford English Dictionary

PER-EMPH-N : expérimentation portant sur la perception de l'emphase chez les natifs (Chapitre III)

PERL : (langage de programmation de scripts) Practical Extraction and Reporting Language

PER-LEC-N : expérimentation portant sur la perception en lecture chez les natifs (Chapitre II)

PER-LEC-NN : expérimentation portant sur la perception en lecture chez les non natifs (Chapitre V)

PER-LEX-NN : expérimentation portant sur la perception des patrons d'accentuation lexicale réalisés dans les polysyllabes par les non natifs (Chapitre IV)

PER-PROM-N : expérimentation portant sur la perception de proéminences chez les natifs (Chapitre III)

PER-THEM-NN : expérimentation portant sur la perception d'une accentuation du pronom them chez les non natifs (Chapitre IV)

PR : (dictionnaire) Le Petit Robert

PRO-LEC-N : expérimentation portant sur la production de lecture à voix haute par les natifs (Chapitre II)

Q1 : (catégorie d'apprenants du) premier quartile (en vitesse de lecture globale)

Q9 : (catégorie d'apprenants du) neuvième quartile (en vitesse de lecture globale)

REP-PIAN-NN : expérimentation de repérage des apprenants « pianistes » (article « Intonational phrasing as a potential indicator for establishing prosodic learner profiles », 2013)

SAMPA : (alphabet, jeu de caractères) Speech Assessment Methods Phonetic Alphabet

SMC : ouvrage sur la Stylistique de Calas

STAT-N1 : première expérimentation par analyse statistique chez les natifs (Chapitre II)

STAT-NN1 : première expérimentation par analyse statistique chez les non natifs (Chapitre IV)

STAT-NN2 : deuxième expérimentation par analyse statistique chez les non natifs (Chapitre IV)

STAT-NN3 : troisième expérimentation par analyse statistique chez les non natifs (Chapitre IV)

STAT-NN4 : quatrième expérimentation par analyse statistique chez les non natifs (Chapitre V)

STAT-NN5 : cinquième expérimentation par analyse statistique chez les non natifs (Chapitre V)

STAT-NN6 : sixième expérimentation par analyse statistique chez les non natifs (Chapitre V)

TLFI : (dictionnaire) Le Trésor de la Langue Française Informatisé

ToBI : (système de transcription et d'annotation prosodique) Tones and Break Indices

UWL : University Word List

Sommaire

Remerciements	5
Abréviations, sigles et acronymes	7
Sommaire	11
Table des illustrations.....	20
Introduction	40
CHAPITRE I – Prosodie anglaise et évaluation prosodique d’apprenants	65
CHAPITRE II – Profils et profilage : mode d’organisation hiérarchique de la prosodie des anglophones natifs.....	116
CHAPITRE III – Emphase : mode d’organisation paradigmatique des réalisations prosodiques des anglophones natifs	222
CHAPITRE IV – Expérimentation de création de profils et de critères d’évaluation prosodiques d’apprenants.....	379
CHAPITRE V — Vérification de la fiabilité du système d’évaluation prototypique semi-automatique par variables.....	492
Conclusion.....	532
Bibliographie.....	542
Annexes.....	555
Index.....	582

Table des matières

Remerciements	5
Abréviations, sigles et acronymes	7
Sommaire	11
Table des illustrations.....	20
1 Figures	20
2 Tableaux	32
Introduction	40
1 Résumé de l'Introduction	40
2 Présentation générale et justification de ce travail de recherche	40
3 Spécificité de la prosodie.....	42
3.1 Concepts prosodiques définitoires	42
3.2 Fonctionnement prosodique.....	48
3.3 L'organisation prosodique : une conjonction spécifique de corrélats	51
4 Enjeux de l'évaluation prosodique	55
4.1 Concepts définitoires fondamentaux en évaluation	55
4.2 Objectifs fonctionnels de l'évaluation	59
4.3 Phases de développement d'un test évaluatif.....	62
5 Problématique : Cadrer la variabilité et repérer des tendances	63
CHAPITRE I – Prosodie anglaise et évaluation prosodique d'apprenants	65
1 Résumé et introduction du Chapitre I.....	65
2 Organisation prosodique.....	66
2.1 L'organisation temporelle de la prosodie	66
2.2 L'organisation mélodique	74
2.3 L'organisation rythmique.....	89
2.4 Synthèse de l'organisation prosodique	101

3	Principes évaluatifs.....	102
3.1	Concept d’interlangue.....	102
3.2	Descripteurs prosodiques du <i>CECRL</i>	104
3.3	Criterial features (critères d’évaluation) adaptés à la prosodie anglaise.....	107
4	Conditions expérimentales de cette recherche.....	109
4.1	Corpus <i>Longdale-Charliphonia</i>	109
4.2	Parole lue.....	111
5	Démarche et organisation générale de la thèse.....	111
5.1	Choix méthodologiques pour remédier à la difficulté de modéliser l’interlangue prosodique.....	112
5.2	Questions de recherche, hypothèses et structure de la thèse.....	114
CHAPITRE II – Profils et profilage : mode d’organisation hiérarchique de la prosodie des anglophones natifs.....		116
1	Résumé du Chapitre II et introduction.....	116
2	Profil et profilage : définitions propres au domaine.....	118
2.1	Analyse lexicographique du mot « profil ».....	118
2.2	Étude générale de la notion de profil et profilage.....	119
2.3	Bilan et discussion.....	128
3	Profils personnalisés : centrage sur l’individu en vue d’une identification ou d’une remédiation.....	129
3.1	Profilage médico-légal (<i>offender profiling</i> et <i>forensic profiling</i>).....	129
3.2	Approche pionnière des profils linguistiques (Crystal 1982).....	136
3.3	Approche pionnière des profils prosodiques (Crystal 1982).....	141
3.4	Synthèse.....	147
4	Profils de groupes : catégorisation natifs-non natifs.....	148
4.1	Introduction : corpus d’apprenants oraux.....	148
4.2	Profilage prosodique multidimensionnel : une approche pionnière de l’interlangue (Gut 2009).....	149

4.3	Profilage du rythme par comparaison de métriques : l'importance de la parcimonie du modèle (Tortel 2009).....	163
4.4	Profilage prosodique des anglicistes francophones : réalisations et accent français (Horgues 2010).....	168
4.5	Profilage suprasegmental par analyse automatique (Ferragne 2013)	169
4.6	Profilage intonatif par étude multidimensionnelle quantitative et qualitative (Herment <i>et al.</i> 2014)	172
4.7	Profilage du rythme par durée syllabique (Ballier, Martin, Amand 2016).....	175
4.8	Profilage évaluatif de la qualité lexicale (Laufer et Nation 1995).....	181
5	Bilan des phénomènes de profils et de profilage.....	186
6	Profilage de la norme native en lecture	187
6.1	Profilage prosodique qualitatif par test perceptif (expérimentation PER-LEC-N) 188	
6.2	Analyse des commentaires de lecteurs sur leur prestation (expérimentation EXP-LEC-N).....	191
6.3	Profilage des lecteurs par analyse acoustique (expérimentation COR-AC-N1)..	195
6.4	Profilage des lecteurs par analyse statistique (expérimentation STAT-N1).....	204
7	Synthèse : création d'outils pour analyser l'interlangue	213
7.1	Profils pour établir des modèles d'évaluation prosodique.....	213
7.2	Critères acoustiques pour modèle et norme évaluatifs	214
7.3	Validation de l'Hypothèse 1	214
7.4	Potentiel et limites des méthodes de profilage de la littérature pour ma recherche 215	
7.5	Détermination de ma méthode de profilage.....	216
CHAPITRE III – Emphase : mode d'organisation paradigmatique des réalisations prosodiques des anglophones natifs		222
1	Résumé du Chapitre III.....	222
2	Phénomène de l'emphase	224
2.1	Définitions de l'emphase	224

2.2	Méthodologie d’investigation perceptive et acoustique de l’emphase	240
2.3	Mes choix sur les corrélats perceptifs et acoustiques de l’emphase	247
2.4	Synthèse	248
3	Test perceptif de détection et hiérarchisation de l’emphase chez les anglophones natifs (expérimentation PER-EMPH-N)	248
3.1	Rappel des objectifs	248
3.2	Profilage des lecteurs	249
3.3	Profilage des auditeurs	250
3.4	Profilage de l’expérimentation.....	250
3.5	Classement et résultat des données perceptives.....	253
3.6	Bilan de l’étude perceptive de l’emphase (expérimentation PER-EMPH-N)	261
4	Étude acoustique des marqueurs d’emphase principaux chez les anglophones natifs (expérimentation COR-AC-N2).....	263
4.1	L’exclamative en <i>How</i>	263
4.2	L’adverbe de degré ou intensificateur <i>so</i> suivi d’un adjectif ou d’un adverbe	279
4.3	L’interrogative en <i>What</i>	304
4.4	<i>Nobody</i>	334
4.5	Bilan de l’étude des principaux marqueurs d’emphase (expérimentation COR-AC-N2)	344
5	Étude de marqueurs non emphatiques ou à l’emphase réduite chez les anglophones natifs (expérimentations COR-DIV-N3).....	345
5.1	La phrase “neutre” (ou au degré d’emphase minimal)	346
5.2	Décrochage mélodique sur <i>all</i>	355
5.3	Désaccentuation pronominale en fin d’énoncé	362
5.4	Étude temporelle du rythme par durée syllabique	364
5.5	Placement et degré accentuels lexicaux	366
5.6	Synthèse de l’étude des marqueurs à l’emphase réduite (expérimentations COR-DIV-N3)	371

6	Conclusion : analyse prosodique des natifs anglophones sous l’angle de l’emphase	
	373	
6.1	Validation de l’Hypothèse 2	374
6.2	Sélection des marqueurs déterminés par l’emphase chez les natifs.....	374
6.3	Conclusion sur la nature et le potentiel du phénomène de l’emphase	375
CHAPITRE IV – Expérimentation de création de profils et de critères d’évaluation		
	prosodiques d’apprenants.....	379
1	Résumé du Chapitre IV	379
2	Objectifs, enjeux et profilage de l’analyse prosodique des apprenants.....	380
3	Création d’une méthodologie d’organisation et de hiérarchisation temporelle des	
	apprenants (expérimentation de sélection des fichiers d’apprenants étudiés STAT-NN1)	382
3.1	Phase 1 : sélection en fonction des influences linguistiques.....	383
3.2	Phase 2 : classement par la durée de lecture	384
3.3	Phase 3 : sélection selon des repères de représentativité temporelle de l’effectif	
	385	
3.4	Phase 4 : sélection en fonction de critères qualitatifs	386
3.5	Phase 5 : sélection de locuteurs à retenir	386
3.6	Phase 6 : réduction du nombre de locuteurs selon le critère de faisabilité	390
3.7	Résultat obtenu en fin des six phases du protocole de sélection des fichiers	
	d’apprenants	391
4	Méthodologie du profilage (expérimentations COR-AC-NN, PER-LEX-NN, PER-	
	THEM-NN)	392
4.1	Base méthodologique de création des variables d’origine acoustique	
	(expérimentation COR-AC-NN) ou perceptive (expérimentations PER-THEM-NN et	
	PER-LEX-NN).....	392
4.2	Profilage du respect de la non proéminence d’un pronom en fin d’énoncé par test	
	perceptif (expérimentation PER-THEM-NN).....	394
4.3	Profilage du respect du patron accentuel lexical par test perceptif (expérimentation	
	PER-LEX-NN).....	394

5	Analyse par profilage de groupes par archétypes de conformité selon la vitesse de lecture (expérimentation STAT-NN2)	396
5.1	Visée et enjeux des profils par archétypes de conformité.....	396
5.2	Profil des apprenants dans le domaine temporel : vitesse et débit d'élocution ...	397
5.3	Profil des apprenants selon la variable du degré d'ampleur de registre.....	399
5.4	Profil des apprenants dans leur qualité mélodique	404
5.5	Profil des apprenants dans leur maîtrise du rythme	408
6	Analyse par profilage de groupes par archétypes stéréotypiques (expérimentation STAT-NN3)	415
6.1	Objectifs et enjeux	415
6.2	Analyse des données par plan factoriel.....	415
6.3	Analyse des données par dendrogramme avec 28 variables	421
6.4	Analyse des données par taux de réussite et d'erreur dans chacune des 28 variables 422	
6.5	Analyse des données par cercles de corrélations	429
6.6	Analyse des données par profils stéréotypiques	439
7	Analyse évaluative des données par profilage modalisateur de critères distinctifs (<i>critical features</i>) : expérimentation CRIT-LEC-NN	445
7.1	Objectifs et enjeux	445
7.2	Profils de critères d'évaluation temporels (par durée et débit)	446
7.3	Profils de critères d'évaluation selon le registre	449
7.4	Profilage de critères d'évaluation mélodiques.....	452
7.5	Profilage de critères d'évaluation rythmiques	474
8	Synthèse du profilage des traits distinctifs et prototypes de grilles critériées pour évaluer la prosodie d'apprenants francophones	482
8.1	Traits prosodiques distinctifs principaux pressentis à l'issue de cette recherche	483
8.2	Prototypes de grilles évaluatives de critères discriminants temporels.....	483
8.3	Prototypes de grilles évaluatives de critères discriminants de registre.....	484

8.4	Prototypes de grilles évaluatives de critères discriminants mélodiques	485
8.5	Prototypes de grilles évaluatives de critères discriminants rythmiques.....	487
9	Conclusion du Chapitre IV	488
CHAPITRE V — Vérification de la fiabilité du système d'évaluation prototypique semi-automatique par variables.....		
		492
1	Résumé du Chapitre V.....	492
2	Objectif, enjeu et profilage de la validation d'un classement semi-automatique par adéquation avec un classement perceptif classique.....	493
3	Étude perceptive (expérimentations PER-LEC-NN et STAT-NN4).....	495
3.1	Profilage perceptif des réalisations non natives à visée comparative (expérimentation PER-LEC-NN).....	495
3.2	Profilage des réalisations non natives par variables en symétrie avec le profilage perceptif.....	507
3.3	Analyse des résultats obtenus selon le profilage perceptif (expérimentation STAT-NN4)	508
3.4	Synthèse des classements obtenus par profilage perceptif	523
4	Étude des résultats obtenus selon le profilage par variables (expérimentation STAT-NN5)	525
5	Analyse comparative des résultats obtenus selon les deux méthodes : par variables et par perception (expérimentation STAT-NN6)	528
6	Synthèse.....	530
Conclusion.....		532
1	Rappel des objectifs.....	532
2	Hypothèses, Questions de Recherche et structure de la thèse	532
3	Résultats obtenus	535
4	Considérations sur ce travail de recherche et perspectives futures	536
Bibliographie.....		542
1	Classement par auteur.....	542
2	Sitographie et logiciels	551

Annexes.....	555
1 Préambule aux annexes	555
2 Texte de lecture : <i>The Selfish Giant</i>	556
3 Segmentation des polysyllabes.....	558
4 Rappel des 28 variables créées	560
5 Rappel des codes des expérimentations.....	561
6 Correspondance des codes des locuteurs.....	563
7 Dénomination des expérimentations	564
8 Expérimentation perceptive PER-LEC-N (Chapitre II)	565
9 Expérimentation EXP-LEC-N (Chapitre II).....	565
10 Expérimentation perceptive PER-EMPH-N (Chapitre III)	567
11 Expérimentation perceptive PER-THEM-NN (Chapitre III)	567
12 Explications du fichier « Mesures_apprenants10.xlsx » (Chapitres IV et V)	567
13 Expérimentation perceptive PER-LEC-NN (Chapitre V)	577
13.1 PER-LEC-NN : Sélection des parties de texte pour évaluation par perception.....	577
13.2 Questionnaire en ligne de l'expérimentation PER-LEC-NN	579
13.3 Fichier <i>Excel</i> d'analyse des données de l'expérimentation PER-LEC-NN	579
14 Scripts R (Chapitre V)	579
15 Annexe des travaux auxquels la thèse a donné lieu.....	579
15.1 Communications.....	579
15.1.1 Colloques de phonétique et phonologie anglaises.....	579
15.1.2 Colloques de linguistique	580
15.1.3 Colloques de didactique et acquisition des langues	580
15.2 Publications d'articles	581
15.2.1 Publication en linguistique, phonétique et phonologie	581
15.2.2 Publications en didactique et acquisition des langues.....	581
Index.....	582

Table des illustrations¹

1 Figures

Figure 1 – Schéma des correspondances entre les composantes formelles et matérielles de la prosodie (Di Cristo, 2013 : 83).....	46
Figure 2 – Schéma de la communication verbale vive voix. — L’encodage linguistique est suivi d’un deuxième acte d’encodage à partir d’un « code naturel », pré-linguistique (Fónagy ([1983] 1991 : 14)	49
Figure 3 – Ressources mises en jeu dans la pratique de la communication orale inter-individuelle (Di Cristo 2013 : 36).	51
Figure 4 – Place de la prosodie (du compilateur prosodique), au regard des systèmes de production et de compréhension de la parole (Di Cristo 2013 : 44)	54
Figure 5 – Raisonnement séquentiel dans une démarche validante selon <i>ALTE</i> , adapté à partir des travaux de Kane, Crooks et Cohen 1999, et Bachman 2005 (2011 : 15).....	56
Figure 6 – Quelques sources d’erreur dans un score évaluatif (<i>ALTE</i> , 2011 : 16)	57
Figure 7 – Cycle normal de création d’un test (<i>ALTE</i> , 2011 : 16)	62
Figure 8 – Processus de développement d’un test (<i>ALTE</i> , 2011 : 20)	63
Figure 9 – Unités prosodiques selon les chercheurs de l’anglais et du français (Di Cristo, 2004 : 161)	67
Figure 10 – Contour typique de F0 à l’échelle de l’énoncé, dans diverses langues (Vaissière, [1983] 2016 : 106).....	75
Figure 11 – Illustration de la normalisation de l’effet de déclinaison (P1 et P2 désignent deux proéminences consécutives) (Di Cristo, 2013 : 94)	76
Figure 12 – Illustration de quelques contraintes auditivo-perceptives concernant F0 (Di Cristo, 2013 : 95)	77
Figure 13 – Représentation schématique du continuum d’aisance (<i>fluency</i>) dans les six niveaux du <i>CECRL</i>	106
Figure 14 – Modèle <i>CRISP-DM</i> 1.0 : Modalisation des phases du processus (Chapman <i>et al.</i> , ([1999] 2000).....	124

¹ Toutes les mesures de F0 prises en Hertz ont été converties en demi-tons par la formule $\log_2(\text{Hz}/100)*12$ (demi-tons par rapport à 100 Hz) ainsi que le précisent les fichiers *Excel* mis en annexes.

Figure 15 – Décomposition en quatre niveaux selon la méthodologie <i>CRISP-DM</i> (Chapman <i>et al.</i> , ([1999] 2000 : 6)	126
Figure 16 – Domaines de classification de locuteur judiciaire (d’après Jessen, 2010 : 381). 132	
Figure 17 – Tâches spécifiques au profilage de locuteurs et à la comparaison de locuteurs (Jessen, 2010 : 386).....	134
Figure 18 – Plan factoriel en analyse par composantes principales de l'analyse discriminante des 1260 phrases produites par les 63 locuteurs (Tortel 2009 : 181).....	166
Figure 19 – Visualisation des points de dispersion (gauche) et des erreurs-types (droite) selon le plan (%V, ΔC) (Tortel, 2009 : 185)	167
Figure 20 – Répartition des votes « meilleur lecteur » des étudiants selon le niveau de leur groupe classe	190
Figure 21 – Répartition des votes « moins bon lecteur » des étudiants selon le niveau de leur groupe classe	190
Figure 22 – Profilage acoustique des lecteurs anglophones natifs en 4 niveaux selon la méthodologie <i>CRISP-DM</i> (Chapman <i>et al.</i> , [1999] 2000 : 6).....	196
Figure 23 – Continuum qualitatif selon le caractère « naturel » ou « surfait » de la lecture des « bons » lecteurs natifs	203
Figure 24 – Continuum qualitatif selon le degré de vitesse de lecture des « mauvais » lecteurs natifs	204
Figure 25 – Plan factoriel des six lecteurs et des variables qualitatives censées les distinguer (15 auditeurs dont 9 natifs et 6 non natifs).....	207
Figure 26 – Cercle des corrélations représentant les 14 variables continues dans leur dépendance et leur pertinence	210
Figure 27 – Dendrogramme de hiérarchisation des 6 lecteurs natifs selon leurs variables quantitatives avec la fonction <i>centroid2</i>	211
Figure 28 – Dendrogramme de hiérarchisation des 6 lecteurs natifs selon leurs variables quantitatives avec la fonction <i>Complete</i>	212
Figure 29 – Schématisation de la recherche d’une norme prosodique native selon le profilage <i>CRISP-DM</i> (à partir du <i>CRISP-DM Process Guide and User Manual</i>)	217
Figure 30 – Schématisation du rôle de l’emphase dans l’expérimentation prosodique PER-EMPH-N.....	233
Figure 31 – Profilage par la saillance dans les expérimentations sur anglophones natifs	236
Figure 32 – Situation des trois durées de lecture sur l’axe temporel	250
Figure 33 – Extrait du site expérimental de détection de l’emphase	251

Figure 34 – Extrait de repérage d’emphase dans un script annoté par un auditeur anglophone	252
Figure 35 – Extrait de repérage d’emphase dans un script annoté par un auditeur francophone	252
Figure 36 – Extrait de feuille de tableur et vue rapprochée dans le repérage de l’emphase ..	253
Figure 37 – Représentation graphique de la corrélation entre le pourcentage de pointage d’emphase et le degré moyen d’emphase de chaque mot	254
Figure 38 – Visualisation du degré moyen d’emphase (*10) comparé au nombre de mots de chaque phrase	256
Figure 39 – Rapport du degré d’emphase par nombre de mots dans chacune des 26 phrases du texte	258
Figure 40 – Degré moyen d’emphase de chaque mot de la Phrase 9	259
Figure 41 – Degrés d’emphase cumulés dans la diction des trois anglophones pour chaque mot de la Phrase 9	259
Figure 42 – Degré moyen d’emphase de chaque mot de la Phrase 6	260
Figure 43 – Degrés d’emphase cumulés dans la diction des trois anglophones pour chaque mot de la Phrase 6	261
Figure 44 – Exclamative en <i>How</i> : Durée moyenne des syllabes et de la pause éventuelle ..	275
Figure 45 – Moyenne de F0 par syllabe de l’exclamative en <i>How</i> (sur 31 occurrences)	275
Figure 46 — Répartition de la hauteur mélodique maximale parmi les 38 lecteurs par rapport à leur âge	276
Figure 47 — Exclamative en <i>How</i> : Diagramme représentant sous forme de courbes les contours mélodiques du « profil jeune » et du « profil mature »	277
Figure 48 – F0 des profils de « bons » lecteurs (L1 et L5)	278
Figure 49 – Entrée lexicale de <i>so</i> (<i>Longman Pronunciation Dictionary</i>)	288
Figure 50 – Entrée lexicale de <i>so</i> (<i>Cambridge English Pronouncing Dictionary - 18th edition</i>)	288
Figure 51 – Valeur moyenne de F0 dans le tronçon englobant S4 (syntagme intonatif de cinq syllabes, 34 lecteurs)	292
Figure 52 – S4 : Moyenne de F0 dans les contours ayant /s@U/ et /swi:t/ comme valeurs maximales (34 lecteurs)	293
Figure 53 – Textgrid de L1 dans <i>sang so sweetly</i> (P4, v1)	294
Figure 54 – Textgrid de L5 dans <i>sang so sweetly</i> (P4)	295

Figure 55 – Schématisation tonale des syllabes /s@U/ et /swi:t/ de P4 par L1 et L5 (mesures sur voyelles)	295
Figure 56 – F0 moyenne dans le syntagme S22 (8 syllabes, 34 lecteurs).....	296
Figure 57 – Moyenne de F0 dans les contours ayant /s@U/ et /swi:t/ comme valeurs maximales dans P22 (34 lecteurs).....	297
Figure 58 – Textgrid de L1 dans <i>sounded so sweet</i> (P22, v1)	298
Figure 59 – Textgrid de L5 dans <i>sounded so sweet</i> (P22)	299
Figure 60 – Schématisation tonale des syllabes /s@U/ et /swi:t/ de L1 et L5 dans P22 (mesures sur voyelles).....	299
Figure 61 – Contour des moyennes de F0 dans le segment S19 (9 dernières syllabes d’un long syntagme intonatif, 30 lecteurs)	301
Figure 62 – Segment S19 : F0 moyenne des réalisations avec proéminences mélodiques sur /s@U/ et /leIt/ (30 lecteurs).....	302
Figure 63 – Textgrid de L1 dans <i>why the spring is so late in coming</i> (P22, v1).....	302
Figure 64 – Textgrid de L5 dans <i>why the spring is so late in coming</i> (P22).....	303
Figure 65 – Schématisation tonale des syllabes /s@U/ et /leIt/ de P19 par L1 et L5 (mesures sur voyelles)	303
Figure 66 – Question en <i>What</i> : Durée moyenne des syllabes (37 lecteurs).....	314
Figure 67 – Question en <i>What</i> : Représentation par boîtes à moustaches de la F0 moyenne par syllabe (30 lecteurs)	315
Figure 68 – Question en <i>What</i> : Représentation en lignes de la F0 moyenne par syllabe (30 lecteurs)	316
Figure 69 – Question en <i>What</i> : Répartition et pourcentage des proéminences dans le groupe des jeunes lecteurs.....	317
Figure 70 – Question en <i>What</i> : Assimilation des trois premières syllabes et proéminence sur /iN/ d’une jeune lectrice (SG-UP-Ro-f-26).....	318
Figure 71 – Question en <i>What</i> : Répartition et pourcentage des proéminences dans le groupe des lecteurs matures	319
Figure 72 – Question en <i>What</i> : Répartition et pourcentage des proéminences dans le groupe des lecteurs professionnels	320
Figure 73 – Question en <i>What</i> : Degré moyen de proéminence par groupes de lecteurs.....	320
Figure 74 – Question en <i>What</i> : Focus large d’un lecteur mature (SG-CA-J5-m-62)	321
Figure 75 – Question en <i>What</i> : Focus étroit sur <i>you</i> d’une lectrice non professionnelle de 58 ans (SG-UP-A2-f-58).....	322

Figure 76 – Question en <i>What</i> : F0 basse et forte intensité dans la réalisation « sans focus » d’une lectrice professionnelle (SG-PRO-JG-f)	323
Figure 77 – Question en <i>What</i> : F0 basse et forte intensité dans la réalisation « sans focus » avec pauses d’un lecteur mature (SG-PD-W-m-50)	323
Figure 78 – Question en <i>What</i> : Diversification de l’annotation malgré des contours intonatifs similaires	327
Figure 79 – Question en <i>What</i> : Pourcentage d’occurrence des divers tons par catégorie de lecteurs	328
Figure 80 – Question en <i>What</i> : Pourcentage d’occurrence du ton « descendant-montant » sur une même syllabe	329
Figure 81 – Question en <i>What</i> : Pourcentage d’occurrence du ton montant sur une même syllabe.....	329
Figure 82 – Question en <i>What</i> : Pourcentage d’occurrence du ton descendant sur une même syllabe.....	330
Figure 83 – Question en <i>What</i> : Pourcentage d’occurrence du ton « montant-descendant » sur une même syllabe	331
Figure 84 – Question en <i>What</i> : Pourcentage d’occurrence des tons « montant », « descendant » et « montant-descendant » sur l’ensemble des syllabes	332
Figure 85 – Exemple de perte de proéminence de <i>nobody</i> (Bolinger, 1989 : 179).....	338
Figure 86 – Exemples de perte de proéminence de <i>nobody</i> (Bolinger, 1989 : 180)	338
Figure 87 – S11_Nobody : ton hypercomplexe $\vee\setminus$ (lecteur SG-Pro-CK-m)	340
Figure 88 – S11_Nobody : ton hypercomplexe $\wedge\setminus$ (lectrice SG-UP-A2-f-58).....	341
Figure 89 – S11_Nobody : ton hypercomplexe $\vee\setminus$ (lecteur SG-UP-M1-m-65)	341
Figure 90 – S11_Nobody : ton complexe \wedge (lecteur L1 ; SG-UP-A1-m-56).....	342
Figure 91 – Relation entre la complexité du patron tonal de la syllabe /n@U/ en fonction du registre du mot <i>nobody</i> (42 natifs)	342
Figure 92 – Complexité du patron tonal de la syllabe /n@U/ en fonction de la durée du texte (42 natifs)	343
Figure 93 – Contour intonatif descendant de base en français (Huart 2010 : 154).....	347
Figure 94 – Contour intonatif descendant de base en anglais (Huart 2010 : 155).....	347
Figure 95 – Contour intonatif descendant de base en anglais avec remontée finale (Huart 2010 : 155)	348
Figure 96 – Structuration syntagmatique et prosodique de la Phrase 6	350
Figure 97 – Structuration syntaxique et prosodique de la Phrase 9	351

Figure 98 – Spectrogramme de la Phrase 6 par le Lecteur L1 (v1) (SG-UP-A1-m-56)	352
Figure 99 – Spectrogramme de la Phrase 6 par le Lecteur L5 (SG-PD-He1-f-23).....	352
Figure 100 – Spectrogramme de la Phrase 9 par le Lecteur 1 (v1) (SG-UP-A1-m-56).....	353
Figure 101 – Spectrogramme de la Phrase 9 par le Lecteur 5 (SG-PD-He1-f-23)	353
Figure 102 – <i>All</i> + relative en <i>that</i> : contour mélodique normalisé (F0 max).....	358
Figure 103 – <i>All</i> + relative en <i>that</i> : contour mélodique de L1 (v1) (SG-UP-A1-m-56)	359
Figure 104 – <i>All</i> + relative en <i>that</i> : contour mélodique de L5 (SG-PD-He1-f-23).....	359
Figure 105 – <i>All</i> + préposition de lieu <i>round</i> : contour mélodique normalisé (F0 max)	360
Figure 106 – <i>All</i> + préposition de lieu <i>round</i> : contour mélodique de L1 (v1) (SG-UP-A1-m-56).....	360
Figure 107 – <i>All</i> + préposition de lieu <i>over</i> : contour mélodique normalisé (F0 max).....	361
Figure 108 – <i>All</i> + préposition de lieu <i>over</i> : contour mélodique de L1 (v1) (SG-UP-A1-m-56)	361
Figure 109 – Schéma de proéminences attendues, basé sur le degré d’emphase de chaque mot	362
Figure 110 – Degrés moyens d’emphase des mots de la Phrase 4	363
Figure 111 – Comparaison de la durée syllabique dans <i>to lis(ten)</i> et <i>sat at</i>	365
Figure 112 – Cercle des correspondances des variables du Chapitre II avec les variables rythmiques syllabiques « lis », « sat » et « satlis ».....	365
Figure 113 – Dendrogramme des variables du Chapitre II avec la variable rythmique syllabique « sat »	366
Figure 114 – Assignation attendues des proéminences des natifs et des non natifs en fin de Phrase 22	370
Figure 115 – Répartition du nombre de fichiers S001R selon la durée de lecture du texte ...	385
Figure 116 – Durée de lecture du texte par catégorie d’apprenants.....	397
Figure 117 – Durée moyenne du total des tronçons A, B et C pauses incluses et pauses exclues par catégorie d’apprenants	398
Figure 118 – Nombre moyen de syllabes par seconde par catégorie d’apprenants dans l’ensemble du texte en prenant en compte les 541 syllabes prononcées en moyenne par L1 et L5 (lecteurs repères).....	398
Figure 119 – Nombre moyen de syllabes par seconde et par catégorie d’apprenants dans l’ensemble des trois tronçons A, B et C	399
Figure 120 – Comparaison de l’ampleur du registre dans l’exclamative avec adjonction du discours citant (Phrase 5) par catégorie d’apprenants selon la vitesse de lecture	400

Figure 121 – Comparaison de l’ampleur du registre sur <i>so</i> + adjectif/adverbe dans les Phrases 4 et 22 par catégorie d’apprenants.....	401
Figure 122 – Registre du syntagme intonatif S9 de la Phrase 9 par catégorie d’apprenants .	402
Figure 123 – Ampleur du registre de la syllabe emphatique /n@U/ par catégorie d’apprenants	402
Figure 124 – Registre maximal des phrases étudiées et des moyennes de ces registres par catégorie d’apprenants.....	403
Figure 125 - Contour intonatif normalisé moyen par catégorie d’apprenants dans l’exclamative avec discours citant.....	404
Figure 126 – Contour intonatif normalisé moyen par catégorie d’apprenants dans les syllabes pertinentes d’un syntagme intonatif dépourvu d’emphase.....	405
Figure 127 – Contour intonatif normalisé moyen du ton bidirectionnel descendant-montant dans la seconde syllabe du mot <i>arrived</i> par catégorie d’apprenants	406
Figure 128 – Contour tonal normalisé moyen de la syllabe /n@U/ de <i>nobody</i> par catégorie d’apprenants (F0 en demi-tons).....	407
Figure 129 – Variable tonale syllabique \wedge par catégories (barème qualitatif)	407
Figure 130 – Contour intonatif par catégorie du tronçon comprenant <i>so sweetly</i> de la Phrase 4 (F0max en demi-tons)	409
Figure 131 – Contour intonatif par catégorie du tronçon comprenant <i>so sweet</i> de la Phrase 22 (F0max en demi-tons)	409
Figure 132 – Contour intonatif par catégorie du tronçon comprenant <i>all around</i> (Phrase 12) (F0max en demi-tons)	410
Figure 133 – Contour intonatif par catégorie du tronçon comprenant <i>all over</i> (Phrase 16) (F0max en demi-tons)	410
Figure 134 – Placement accentuel lexical correct par catégorie d’apprenants (en pourcentage)	411
Figure 135 – Pourcentage moyen par catégories d’apprenants d’un repérage de proéminence sur <i>them</i> (en pourcentage, 0 = qualité maximale de la réalisation)	412
Figure 136 – Saillance par la différence de durée de deux syllabes non emphatiques contiguës (en pourcentage).....	412
Figure 137 – Durée syllabique moyenne par catégorie des syllabes /s{t/ et /@t/ (en millisecondes).....	413
Figure 138 – Plan factoriel des quinze apprenants avec 28 variables (ACP)	418
Figure 139 – Plan factoriel des quinze apprenants avec les 28 variables et leurs données....	419

Figure 140 – Plan factoriel des quinze apprenants à partir de 25 variables	420
Figure 141 – Dendrogramme de Ward pour les 28 variables des quinze apprenants	421
Figure 142 – Hiérarchisation de la qualité prosodique des six profils d'apprenants selon le dendrogramme de McQuitty	422
Figure 143 – Dendrogramme de McQuitty pour les 28 variables des quinze apprenants	422
Figure 144 – Répartition de quatre graduations qualitatives (20% chacune) de la variable RYT_lex à partir des résultats numériques en pourcentage de réussite des 15 apprenants ...	423
Figure 145 – Extraits du tableau évaluatif général des apprenants par variable : discrimination qualitative visuelle des étiquettes appréciatives des variables DUR_T et RYT_lex (ordre des colonnes pour chaque tableau : liste des apprenants, leur catégorisation d'après leur vitesse de lecture du texte entier, résultat quantitatif en secondes et en pourcentages pour DUR_T et RYT_lex respectivement et catégorisation évaluative à partir des résultats quantitatifs).....	423
Figure 146 – Classement des six profils d'apprenants selon 28 variables prosodiques en fonction de la proportion du taux d'erreur (bilan négatif de malus en rouge)	424
Figure 147 – Classement des six profils d'apprenants selon 25 variables prosodiques en fonction de la proportion du taux d'erreur (bilan négatif de malus en rouge)	425
Figure 148 – Cercle de corrélations des 28 variables prosodiques des apprenants sur le premier plan (dimensions 1 et 2).....	429
Figure 149 – Cercle de corrélations des 25 variables prosodiques des apprenants sur le premier plan (dimensions 1 et 2).....	430
Figure 150 – Cercle de corrélations des 6 variables temporelles des apprenants	431
Figure 151 – Cercle de corrélations des 6 variables de registre des apprenants	433
Figure 152 – Cercle de corrélations des 6 variables mélodiques des apprenants	434
Figure 153 – Cercle de corrélation des 6 variables mélodiques des apprenants avec RM (RYTMEL_so4)	436
Figure 154 – Cercle de corrélations des 7 variables rythmiques des apprenants.....	437
Figure 155 – Cercle de corrélations des 7 variables rythmiques des apprenants avec RM (RYTMEL_so4)	438
Figure 156 – Copie d'écran du spectrogramme et TextGrid Praat du « pianiste » DID0078-S001R pour l'emphase de <i>nobody</i> (Phrase 11)	441
Figure 157 – Copie d'écran du spectrogramme et TextGrid Praat du « pianiste » DID0078-S001R pour l'emphase de l'exclamative (Phrase 5)	442
Figure 158 – Copie d'écran de spectrogramme et TextGrid Praat du « pianiste » DID0062-S001R (S7) pour l'emphase sur l'exclamative (Phrase 5)	444

Figure 159 – Copie d'écran du spectrogramme et TextGrid <i>Praat</i> du « pianiste » DID0062-S001R (S7) pour l'emphase sur <i>nobody</i> (Phrase 11)	444
Figure 160 – Vitesse de lecture du texte par les catégories de lecteurs apprenants francophones, lecteurs anglophones jeunes, matures et professionnels (en secondes).....	446
Figure 161 – Profil de registre pour chaque phrase ou entité étudiée, toutes catégories d'apprenants confondues (en demi-tons base 100)	450
Figure 162 – Comparaison des moyennes de registres par entité étudiée et par catégorie d'apprenants selon la rapidité de lecture (en demi-tons base 100)	451
Figure 163 – Contour intonatif de la Phrase 5 (exclamative et discours citant) du groupe D1 (F0max en demi-tons)	454
Figure 164 – Contour intonatif de la Phrase 5 (exclamative et discours citant) du groupe Q1 (F0max en demi-tons)	454
Figure 165 – Contour intonatif de la Phrase 5 (exclamative et discours citant) du groupe de la Médiane (F0max en demi-tons)	454
Figure 166 – Contour intonatif de la Phrase 5 (exclamative et discours citant) du groupe Q3 (F0max en demi-tons)	455
Figure 167 – Contour intonatif de la Phrase 5 (exclamative et discours citant) du groupe D9 (F0max en demi-tons)	455
Figure 168 – Qualité mélodique de l'exclamative et de la Phrase 5 calculée par formules critériées (barème qualitatif)	457
Figure 169 – Contour intonatif du syntagme intonatif S9 du groupe D1 (F0mean en demi-tons).....	459
Figure 170 – Contour intonatif du syntagme intonatif S9 du groupe Q1 (F0mean en demi-tons).....	459
Figure 171 – Contour intonatif du syntagme intonatif S9 du groupe de la Médiane (F0mean en demi-tons).....	459
Figure 172 – Contour intonatif du syntagme intonatif S9 du groupe Q3 (F0mean en demi-tons).....	460
Figure 173 – Contour intonatif du syntagme intonatif S9 du groupe D9 (F0mean en demi-tons).....	460
Figure 174 – Conformité du contour descendant général sur le syntagme intonatif neutre S9 par barème qualitatif (histogramme détaillé)	462
Figure 175 – Conformité du contour descendant général sur le syntagme intonatif neutre S9 par barème qualitatif (histogramme résumé)	462

Figure 176 – Schématisation normalisée du ton descendant-montant (∨) sur une syllabe devant frontière mineure (<i>arrived</i>) dans le groupe D1 (F0max en demi-tons)	464
Figure 177 – Schématisation normalisée du ton descendant-montant (∨) sur une syllabe devant frontière mineure (<i>arrived</i>) dans le groupe Q1 (F0max en demi-tons)	464
Figure 178 – Schématisation normalisée du ton descendant-montant (∨) sur une syllabe devant frontière mineure (<i>arrived</i>) dans le groupe Médiane (F0max en demi-tons)	465
Figure 179 – Schématisation normalisée du ton descendant-montant (∨) sur une syllabe devant frontière mineure (<i>arrived</i>) dans le groupe Q3 (F0max en demi-tons)	465
Figure 180 – Schématisation normalisée du ton descendant-montant (∨) sur une syllabe devant frontière mineure (<i>arrived</i>) dans le groupe D9 (F0max en demi-tons)	465
Figure 181 – Schéma détaillé de la réalisation du ton bi-directionnel ∨ dans <i>arrived</i> selon un barème qualitatif.....	467
Figure 182 – Schématisation normalisée du mot <i>nobody</i> dans le groupe D1 (F0 en demi-tons)	469
Figure 183 – Schématisation normalisée du mot <i>nobody</i> dans le groupe Q1 (F0 en demi-tons)	469
Figure 184 – Schématisation normalisée du mot <i>nobody</i> dans le groupe Médiane (F0 en demi-tons).....	470
Figure 185 – Schématisation normalisée du mot <i>nobody</i> dans le groupe Q3 (F0 en demi-tons)	470
Figure 186 – Schématisation normalisée du mot <i>nobody</i> dans le groupe D9 (F0 en demi-tons)	471
Figure 187 – Complexité de la variation tonale sur la syllabe accentuée de <i>nobody</i> par barème qualitatif	472
Figure 188 – Contour tonal normalisé moyen de la syllabe /n@U/ de <i>nobody</i> par schéma tonal (F0 en demi-tons)	473
Figure 189 – Contour mélodique de la paire <i>so sweetly</i> de la Phrase 4 dans le groupe D1 (F0 en demi-tons).....	475
Figure 190 – Contour mélodique de la paire <i>so sweetly</i> de la Phrase 4 dans le groupe Q1 (F0 en demi-tons).....	475
Figure 191 – Contour mélodique de la paire <i>so sweetly</i> de la Phrase 4 dans le groupe Médiane (F0 en demi-tons)	476
Figure 192 – Contour mélodique de la paire <i>so sweetly</i> de la Phrase 4 dans le groupe Q3 (F0 en demi-tons).....	476

Figure 193 – Contour mélodique de la paire <i>so sweetly</i> de la Phrase 4 dans le groupe D9 (F0 en demi-tons).....	476
Figure 194 – Détail des données constitutives des variables RYT_so4 et RYTMEL_so4 par catégorie d'apprenants selon la vitesse de lecture (barème qualitatif).....	478
Figure 195 – Placement accentuel lexical conforme sur 6 polysyllabes par apprenant (accentuation primaire et secondaire)	480
Figure 196 –Réalisations correctes de patrons accentuels sur polysyllabes par mot (en pourcentage).....	481
Figure 197 – Modélisation de la dernière phase du profilage des traits distinctifs évaluatifs	493
Figure 198 – Échelle d'Osgood utilisée dans le questionnaire PER-LEC-NN	503
Figure 199 – Résultats de l'évaluation des quinze fichiers d'apprenants (abscisse) par perception des 68 évaluateurs (ordonnées) selon le code : 1 = <i>much worse</i> , 2 = <i>worse</i> , 3 = <i>similar to File 0</i> , 4 = <i>better</i> , 5 = <i>much better</i>	509
Figure 200 – Plan factoriel de l'analyse des correspondances multiples de l'évaluation perceptive des quinze apprenants par 68 évaluateurs spécialistes	510
Figure 201 – Arbre d'inférence conditionnelle catégorisant et ordonnant les prestations des apprenants selon l'ensemble des 68 évaluateurs spécialistes par perception auditive.....	511
Figure 202 – Classement perceptif catégoriel des fichiers d'apprenants par arbre d'inférence selon les 68 experts	511
Figure 203 – Classement perceptif des fichiers d'apprenants par moyenne des résultats attribués par les 68 experts	511
Figure 204 – Résultats de l'évaluation des quinze apprenants (abscisse) par perception des 40 enseignants francophones (ordonnées) selon le code : 1 = <i>much worse</i> , 2 = <i>worse</i> , 3 = <i>similar to File 0</i> , 4 = <i>better</i> , 5 = <i>much better</i>	512
Figure 205 – Plan factoriel de l'analyse des correspondances multiples de l'évaluation perceptive des quinze apprenants par 40 experts francophones.....	513
Figure 206 – Arbre d'inférence conditionnelle catégorisant et ordonnant les prestations des quinze apprenants par perception auditive selon les 40 enseignants francophones.....	513
Figure 207 – Classement perceptif catégoriel des fichiers d'apprenants par arbre d'inférence selon les 40 enseignants francophones	514
Figure 208 – Classement perceptif des fichiers d'apprenants par moyenne des résultats attribués par les 40 experts francophones	514
Figure 209 – Détail de l'évaluation perceptive des 14 apprenants par chacun des 40 experts francophones	514

Figure 210 – Résultats de l'évaluation des 14 apprenants (abscisse) par perception des 25 enseignants francophones (ordonnées) selon le code : 1 = <i>much worse</i> , 2 = <i>worse</i> , 3 = <i>similar to File 0</i> , 4 = <i>better</i> , 5 = <i>much better</i>	515
Figure 211 – Arbre d'inférence conditionnelle catégorisant et ordonnant les prestations des 15 apprenants par perception auditive selon les 25 évaluateurs spécialistes anglophones natifs	516
Figure 212 – Classement perceptif catégoriel des fichiers d'apprenants par arbre d'inférence selon les 25 spécialistes anglophones natifs.....	516
Figure 213 – Classement perceptif des fichiers d'apprenants par moyenne des résultats attribués par les 25 experts anglophones.....	516
Figure 214 – Détail de l'évaluation perceptive des 14 apprenants par chacun des 25 experts anglophones.....	517
Figure 215 – Résultats de l'évaluation des 15 apprenants (abscisse) par perception des 18 experts profilés (ordonnées) selon le code : 1 = <i>much worse</i> , 2 = <i>worse</i> , 3 = <i>similar to File 0</i> , 4 = <i>better</i> , 5 = <i>much better</i>	518
Figure 216 – Arbre d'inférence conditionnelle catégorisant et ordonnant les prestations des 15 apprenants par perception auditive selon les 18 experts profilés	519
Figure 217 – Classement perceptif catégoriel des fichiers d'apprenants par arbre d'inférence selon les 18 experts profilés	519
Figure 218 – Classement perceptif des fichiers d'apprenants par moyenne des résultats attribués par les 18 experts profilés.....	519
Figure 219 – Résultats de l'évaluation des 15 apprenants (abscisse) par perception des 68 experts (ordonnées) selon le code : 1 = <i>worse</i> , 2 = <i>similar to File 0</i> , 3 = <i>better</i>	520
Figure 220 – Arbre d'inférence conditionnelle de l'évaluation perceptive en trois catégories des 15 apprenants par les 68 experts	521
Figure 221 – Classement perceptif des fichiers d'apprenants en trois catégories par arbre d'inférence selon les 68 experts profilés	521
Figure 222 – Classement perceptif des fichiers d'apprenants par moyenne des résultats attribués par l'ensemble des 68 experts.....	521
Figure 223 – Détail de l'évaluation perceptive des apprenants en trois catégories par les 68 experts	522
Figure 224 – Résultats de l'évaluation des 15 apprenants (abscisse) par perception des 18 experts profilés (ordonnées) selon le code : 1 = <i>worse</i> , 2 = <i>similar to File 0</i> , 3 = <i>better</i>	522
Figure 225 – Arbre d'inférence conditionnelle de l'évaluation perceptive en trois catégories des 15 apprenants par les 18 experts profilés	523

Figure 226 – Classement perceptif des fichiers d’apprenants par arbre d’inférence en trois catégories selon les 18 experts profilés	523
Figure 227 – Classement perceptif des fichiers d’apprenants par moyenne des résultats attribués par les 18 experts profilés.....	523
Figure 228 – Bilan général et par domaine prosodique des 15 apprenants selon le système de bonus/malus sur 18 variables	528
Figure 229 – Bilan général et par domaine prosodique des 15 apprenants selon le système de bonus/malus sur 25 variables	528
Figure 230 – Modélisation de la recherche de critères distinctifs évaluatifs chez les non natifs anglicistes francophones selon le profilage <i>CRISP-DM</i> (à partir du <i>CRISP-DM Process Guide and User Manual</i>)	534
Figure 231 – Organisation des expérimentations sur les natifs et les non natifs pour déterminer les traits discriminants en prosodie de l’interlangue	535
Figure 232 – Organisation des expérimentations sur les natifs et les non natifs pour déterminer les traits discriminants en prosodie de l’interlangue (rappel de la Figure 231, page 535).....	561

2 Tableaux

Tableau 1– Valeurs moyennes et écarts types (entre parenthèses) de cinq variables de l’aisance en lecture chez les natifs et non natifs anglophones (***) = significatif avec $p < 0,001$; ** = significatif avec $p < 0,01$, * = significatif avec $p < 0,005$) d’après Gut (2009 : 99).....	73
Tableau 2 – Corrélation des classements hiérarchiques de taux d’articulation chez chacun des locuteurs non natifs dans chaque style de parole (***) = significatif avec $p < 0,001$, ** = significatif avec $p < 0,01$) de Gut (2009 : 103)	73
Tableau 3 – Longueur moyenne des énoncés et pourcentage des énoncés interrompus en anglais non natif selon le style de parole (***)=significatif avec $p < 0,001$; *=significatif avec $p < 0,05$) (d’après Gut 2009 : 229).....	80
Tableau 4 – Fréquence de tons nucléaires (en %) en anglais non natif selon le style de parole (d’après Gut 2009 : 233)	82
Tableau 5 – Comparaison de la fréquence de tons nucléaires (en %) en anglais non natif et natif (d’après Gut 2009 : 233).....	82

Tableau 6 – Comparaison de l’envergure du mouvement mélodique nucléaire chez les natifs et non natifs dans la moyenne des chutes et des montées (d’après Gut 2009 : 237).....	83
Tableau 7 – Contours intonatifs des deux questions polaires dans les différents groupes (Herment <i>et al.</i> 2014)	84
Tableau 8 – Contours intonatifs des deux questions partielles dans les différents groupes (Herment <i>et al.</i> 2014)	85
Tableau 9 – Moyenne du registre maximal des natifs et non natifs selon le style de parole (***=significatif avec $p=0.001$; **=significatif avec $p<0.01$) (d’après Gut 2009 : 238, 240)87	
Tableau 10 – Moyenne du registre maximal des natifs et non natifs des énoncés avec discours direct, rhème et thème en lecture (***=significatif avec $p=0.001$; **=significatif avec $p<0.01$) (d’après Gut 2009 : 240).....	87
Tableau 11 – Hauteur mélodique (en %) en début d’énoncé (H=high, M=mid, L=low) des natifs et non natifs dans les phrases contenant rhème, détails additionnels et thème en lecture (Gut 2009 : 241)	87
Tableau 12 – Positionnement du noyau (en %) sur le verbe dans les structures verbe+pronom, et sur le nom dans les structures nom+verbe chez les locuteurs natifs et non natifs ; désaccentuation (en %) des mots donnés ou accessibles (Gut 2009 : 244).....	89
Tableau 13 – Distribution (en %) des mouvements mélodiques dynamiques simples et complexes en position pré-noyau en lecture suivant les locuteurs natifs et non natifs (Gut 2009 : 244)	89
Tableau 14 – Données rythmiques et durée moyenne des syllabes accentuées et inaccentuées en anglais britannique d’après Gut (2009 : 170)	97
Tableau 15 – Données rythmiques et leur écart type chez les natifs anglophones et non natifs de Gut (2009 : 179) (** = significatif avec $p<0.01$, *** = significatif avec $p< 0.001$)	98
Tableau 16 – Récapitulatif des différentes métriques présentées en fonction de la tâche de discrimination visée (Tortel 2009 : 201)	99
Tableau 17 – Récapitulatif des meilleures prédictions pour chacune des combinaisons factorielles (Tortel 2009 : 202).....	99
Tableau 18 – Tableau récapitulatif des marqueurs prosodiques de l’aisance selon le <i>CECRL</i> et <i>CEFR</i> (d’après <i>CECRL</i> , 2000 : 28-29, <i>CEFR</i> , 2001 : 28-29).....	106
Tableau 19 – Schématisation des opérations de profilage d’après la définition de Hildebrandt et Backhouse (2005 : 106).....	121
Tableau 20 – Caractéristiques de détermination de locuteurs (Jessen, 2010 : 387).....	134

Tableau 21 – Fréquence moyenne (en %) des tons dans l’anglais conversationnel d’après Crystal ([1982] 1992 : 117).....	143
Tableau 22 – Reprise des variables de Gut (2009) dans mon profilage de natifs en lecture .	154
Tableau 23 – Reprise des variables de Gut (2009) dans mon profilage de natifs en lecture avec transformation	156
Tableau 24 – Métriques des 9 paramètres rythmiques de Tortel (2009 : 178).....	165
Tableau 25 – Ontologie des syllabes annotées dans l’étude Ballier <i>et al.</i> (2016).....	177
Tableau 26 – Récapitulation de la sélection des auditeurs pour sélectionner le « meilleur » et le « moins bon » lecteur	191
Tableau 27 – Moyennes du nombre global de proéminences et des niveaux de discrimination dans la détection des proéminences	198
Tableau 28 – Matrice récapitulative des données temporelles pour plan factoriel	200
Tableau 29 – Matrice récapitulative du registre de lecture pour plan factoriel.....	201
Tableau 30 – Matrice récapitulative des données rythmiques pour plan factoriel (15 auditeurs)	202
Tableau 31 – Matrice récapitulative des données rythmiques pour plan factoriel (9 auditeurs natifs).....	202
Tableau 32 – Matrice récapitulative des données rythmiques pour plan factoriel (7 auditeurs natifs).....	202
Tableau 33 – Variables permettant de déterminer statistiquement les profils prosodiques de natifs anglophones.....	206
Tableau 34 – Plan factoriel situant les 6 lecteurs (15 auditeurs dont 9 natifs et 6 non natifs)	208
Tableau 35 – Dendrogramme de hiérarchisation des 6 lecteurs natifs selon leurs variables quantitatives avec la fonction <i>Ward</i>	212
Tableau 36 – Décomposition du processus de profilage des lecteurs natifs en 4 phases.....	218
Tableau 37 – Protocole d’expérimentation d’Herment (2001 : 146) sous forme de tableau .	244
Tableau 38 – Classement par les anglophones natifs des quinze premiers mots les plus emphatiques du texte <i>The Selfish Giant</i>	255
Tableau 39 – Caractéristiques des deux phrases « neutres » courtes	261
Tableau 40 – Tableau récapitulatif des dix structures tonales appliquées aux exclamatives en <i>how</i> et l’explication sémantique correspondante (d’après O’Connor et Arnold, 1961 : 41-45 ; 106-252)	271
Tableau 41 – Ampleur du registre des lecteurs L1 et L5 par différence entre F0max et F0min	279

Tableau 42 – Fréquence des amplificateurs les plus courants en conversation britannique et américaine et en discours académique (Biber <i>et al.</i> , 1999 : 565).....	281
Tableau 43 – Fréquence des constructions comparatives selon les registres (Biber <i>et al.</i> , 2005 : 528)	283
Tableau 44 – S4 : Répartition des catégories de lecteurs suivant leur choix mélodique sur <i>so sweetly</i>	294
Tableau 45 – Répartition des catégories de lecteurs suivant leur choix mélodique pour S22 (34 lecteurs)	298
Tableau 46 – Répartition des types de questions selon la nature des corpus (Biber <i>et al.</i> , 1999 : 212).....	306
Tableau 47 – Récapitulatif des 10 structures tonales appliquées aux questions en <i>wh-</i> et l'explication sémantique correspondante (d'après O'Connor et Arnold, 1961 : 49-89).....	308
Tableau 48 – Récapitulatif des dix structures tonales appliquées aux questions en <i>wh-</i> avec leur description mélodique (d'après O'Connor et Arnold, 1961 : 41-45).....	310
Tableau 49 – Symboles de transcription des mouvements tonaux syllabiques.....	326
Tableau 50 – Question en <i>What</i> : Pertinence de la question en <i>What</i> de P10 comme modèle pour évaluation.....	333
Tableau 51 – Distribution des pronoms indéfinis (Biber <i>et al.</i> , 1999 : 352).....	335
Tableau 52 – Sélection des variables à la suite de l'expérimentation COR-AC-N2.....	345
Tableau 53 – Variation tonale des 42 natifs sur les Phrases 6 et 9	354
Tableau 54 – Phrase candidate à l'étude de la désaccentuation d'un pronom en fin d'énoncé	362
Tableau 55 – Vérification que Mean F0 sur voyelles dans la Phrase 4 est la plus basse sur <i>them</i>	363
Tableau 56 – Sélection des syllabes pour étude temporelle du rythme	364
Tableau 57 – Récapitulatif des informations trouvées sur le site <i>Word and Phrase . Info</i> pour les six polysyllabes étudiés	367
Tableau 58 – Schémas accentuels canoniques des polysyllabes étudiés	368
Tableau 59 – Tronçons de phrases sélectionnés pour étude acoustique du rythme par prééminences ou absence de prééminence.....	372
Tableau 60 - Mots sélectionnés pour étude rythmique	372
Tableau 61 – Contours mélodiques vérifiés sur la Phrase 9 (déclarative non emphatique)...	373
Tableau 62 – Sélection des variables déterminées par les expérimentations COR-AC-N2 et COR-DIV-N3 (dont PER-ACLEX-N)	375

Tableau 63 – Ma position sur la nature catégorielle ou continue du phénomène de l’emphase	377
Tableau 64 – Tableau récapitulatif des apprenants écartés de cette étude	384
Tableau 65 – Positionnement des fichiers S001R sélectionnés pour la catégorie D1.....	387
Tableau 66 – Positionnement des fichiers S001R sélectionnés pour la catégorie Q1.....	387
Tableau 67 – Positionnement des fichiers S001R sélectionnés pour la catégorie Médiane...	388
Tableau 68 – Positionnement des fichiers S001R sélectionnés pour la catégorie Q3.....	388
Tableau 69 – Positionnement des fichiers S001R sélectionnés pour la catégorie D9.....	389
Tableau 70 – Tableau de positionnement dans le corpus des fichiers S001R destinés à une étude longitudinale	390
Tableau 71 – Tableau de positionnement dans le corpus des quinze fichiers S001R sélectionnés pour l’étude de cette thèse	391
Tableau 72 – Sélection des variables pour non natifs consécutive aux Chapitres II et III.....	393
Tableau 73 – Variables utilisées pour l’établissement des profils des apprenants, regroupées par domaine prosodique	417
Tableau 74 – Moyenne des 25 variables distinctives par domaine dans les six profils d’apprenants (en pourcentage, 0 = non marqué)	426
Tableau 75 – Comparaison des classements selon le système de bonus-malus prosodique et celui de vitesse de lecture, avec catégories de vitesse correspondantes.....	428
Tableau 76 – Résultats de l’étude du profil « pianiste » sur 20 apprenants de vitesse de lecture rapide et moyenne (expérimentation REP-PIAN-NN).....	443
Tableau 77 – Grille évaluative de la durée de lecture du texte <i>The Selfish Giant</i> (en secondes) correspondant à la variable temporelle DUR_T (D1)	447
Tableau 78 – Grille évaluative de la durée des tronçons du texte <i>The Selfish Giant</i> avec pauses (en secondes) correspondant à la variable temporelle DUR_ABCp (D2)	448
Tableau 79 – Grille évaluative du débit d’élocution des tronçons du texte <i>The Selfish Giant</i> avec pauses (nombre de syllabes par seconde) correspondant à la variable temporelle DEB_ABCp (D3)	448
Tableau 80 – Grille évaluative critériée du degré d’ampleur du registre dans l’exclamative en <i>How</i> suivie de son énoncé citant (F0 en demi-tons sur base 100).....	451
Tableau 81 – Grille évaluative critériée du degré d’ampleur du registre dans l’exclamative en <i>How</i> sans énoncé citant (F0 en demi-tons sur base 100).....	452
Tableau 82 – Grille évaluative critériée de la qualité mélodique de l’exclamative en <i>How</i> sans son énoncé citant	458

Tableau 83 – Grille évaluative critériée de la qualité mélodique de l'exclamative avec son énoncé citant.....	458
Tableau 84 – Grille évaluative critériée de la qualité mélodique du syntagme intonatif neutre S9.....	463
Tableau 85 – Grille évaluative critériée de la qualité mélodique du ton descendant-montant (∨) sur une syllabe devant frontière mineure	468
Tableau 86 – Grille évaluative critériée de la complexité tonale sur syllabe accentuée de <i>nobody</i>	474
Tableau 87 – Grille évaluative avec critères qualitatifs de la variable RYT_so4 (paire rythmique « <i>so</i> + adverbe » de la Phrase 4)	479
Tableau 88 – Grille évaluative avec critères qualitatifs de la variable RYTMEL_so4 établie à partir de la variable RYT_so4 et d'une éventuelle remontée mélodique sur dernière syllabe inaccentuée	479
Tableau 89 – Grille évaluative de respect du patron accentuel sur six polysyllabes en contexte (pourcentage de réussite).....	482
Tableau 90 – Traits distinctifs principaux pressentis, sous forme des variables utilisées pour l'établissement des profils des apprenants, par domaine prosodique.....	483
Tableau 91 – Grille évaluative de la durée de lecture du texte <i>The Selfish Giant</i> (en secondes) correspondant à la variable temporelle DUR_T (D1)	483
Tableau 92 – Grille évaluative de la durée des tronçons du texte <i>The Selfish Giant</i> avec pauses (en secondes) correspondant à la variable temporelle DUR_ABCp (D2)	484
Tableau 93 – Grille évaluative du débit d'élocution des tronçons du texte <i>The Selfish Giant</i> avec pauses (nombre de syllabes par seconde) correspondant à la variable temporelle DEB_ABCp (D3)	484
Tableau 94 – Grille évaluative critériée du degré d'ampleur du registre dans l'exclamative en <i>How</i> suivie de son énoncé citant (F0 en demi-tons sur base 100).....	484
Tableau 95 – Grille évaluative critériée du degré d'ampleur du registre dans l'exclamative en <i>How</i> sans énoncé citant (F0 en demi-tons sur base 100).....	485
Tableau 96 – Grille évaluative des critères qualitatifs de la qualité mélodique de l'exclamative en <i>How</i> sans son énoncé citant	485
Tableau 97 – Grille évaluative des critères qualitatifs de la qualité mélodique de l'exclamative avec son énoncé citant.....	485
Tableau 98 – Grille évaluative des critères qualitatifs de la qualité mélodique du syntagme intonatif neutre S9	486

Tableau 99 – Grille évaluative des critères qualitatifs de la qualité mélodique du ton √ sur une syllabe devant frontière mineure	486
Tableau 100 – Grille évaluative des critères qualitatifs de la complexité tonale sur syllabe accentuée de <i>nobody</i>	486
Tableau 101 – Grille évaluative des critères qualitatifs de la variable RYT_so4 (paire rythmique « <i>so</i> + adverbe » de la Phrase 4.....	487
Tableau 102 – Grille évaluative des critères qualitatifs de la variable RYTMEL_so4 établie à partir de la variable RYT_so4 et d’une éventuelle remontée mélodique sur dernière syllabe inaccentuée	487
Tableau 103 – Grille évaluative de respect du patron accentuel sur polysyllabes en contexte (pourcentage de réussite).....	487
Tableau 104 – Vérification de la similarité des critères distinctifs évalués par 24 variables et par test perceptif PER-LEC-NN (Phase 1).....	496
Tableau 105 – Résultats des trois bilans de réussite, échec et bilan global par analyse semi-automatique à partir des traits discriminants	498
Tableau 106 – Vérification de la similarité des critères distinctifs évalués par 18 variables et par test perceptif PER-LEC-NN (Phase 2).....	499
Tableau 107 – Hiérarchisation de la qualité prosodique obtenue par application du profilage par 25 et par 18 variables	500
Tableau 108 – Récapitulatif du classement évaluatif catégoriel des 15 apprenants par les quatre groupes d’experts étudiés avec regroupement coloré des zones de stabilité.....	524
Tableau 109 – Hiérarchisation évaluative des 15 apprenants selon le système de bonus-malus global et par domaine prosodique selon 18 et 25 variables et les deux systèmes de codage des fichiers utilisés.....	527
Tableau 110 – Moyenne générale de correspondance entre les classements perceptif et semi-automatique en fonction du nombre de catégories retenues et du profilage des experts évaluateurs (en pourcentage).....	529
Tableau 111 – Classements individuels des fichiers non natifs sur axe évaluatif continu.....	530
Tableau 112 – Variables utilisées pour l’établissement des profils des apprenants, regroupées par domaine prosodique (rappel du Tableau 73, page 417)	560
Tableau 113 – Liste des expérimentations de la thèse par chapitre, REP-PIAN-NN exceptée	562
Tableau 114 – Caractéristiques principales des locuteurs stéréotypiques.....	563

Tableau 115 – Correspondance des codes des locuteurs natifs participant à plusieurs expérimentations	563
Tableau 116 – Correspondance des codes des locuteurs non natifs participant à plusieurs expérimentations	564

Introduction

1 Résumé de l'Introduction

L'introduction à ce travail de recherche sur les critères d'évaluation prosodiques a pour fonction de le présenter et de le justifier. Son dessein est de montrer sa spécificité et ses enjeux.

Pour atteindre ce but, elle aborde les deux notions fondamentales que sont la prosodie et l'évaluation par leurs concepts définitoires, leur fonctionnement et leur organisation respectives. Elle souligne l'enjeu essentiel et paradoxal de cette étude : une recherche de compatibilité entre la variabilité extrême de la prosodie anglaise et la rigueur au fondement de la nature évaluative.

Cette introduction répond à la difficulté de la tâche en montrant la nécessité d'une prise en compte méthodologique destinée à éclairer les passages obligés et les écueils à éviter. Elle annonce une présentation de l'état de l'art tout comme de ses applications dans les domaines de la prosodie anglaise et de son évaluation chez les apprenants, laquelle devrait pouvoir structurer cette thèse dans la poursuite de ses objectifs.

2 Présentation générale et justification de ce travail de recherche

Ce travail de recherche est motivé par l'importance croissante des besoins économiques internationaux en matière de circulation des personnes afin de combler les offres de recrutement. Les demandeurs d'emploi à l'étranger ou de mobilité internationale doivent pouvoir justifier de leur niveau en langue étrangère, voire seconde, afin de faciliter la tâche des recruteurs. Aussi, un consensus institutionnel international vise à promouvoir un étalonnage évaluatif commun à moindre coût pouvant favoriser cette mobilité (Huver 2014). L'association *ALTE* (*Association of Language Testers in Europe*²), fondée à l'origine par l'*Université de Cambridge* et l'*Université de Salamanque*, ainsi que le *Cadre Européen*

² [<http://www.alte.org/>] dernièrement consulté en 2017.

Commun de Référence pour les Langues – Apprendre, Enseigner, Évaluer (2001) établi par le *Conseil de l'Europe*, font figures de proue en la matière. Étant donné qu'une évaluation automatisée correspondrait parfaitement aux prévisions de la demande croissante, des critères d'évaluation doivent être trouvés pour chaque langue et principalement pour la langue anglaise, qui, bien que détrônée par le chinois mandarin en nombre de locuteurs, qu'ils soient uniquement natifs, ou bien natifs et non natifs, n'en reste pas moins la langue la plus hégémonique selon Chan (2016)³. Son étude d'indexation de l'importance des langues avec le *Power Language Index* comprenant 20 indicateurs regroupés en cinq catégories (*opportunities*) place l'anglais au premier rang :

English is by far the most powerful language. It is the dominant language of three G7 nations (USA, UK and Canada), and British legacy has given it a global footprint. It is the world's *lingua franca*.

Dans cette ligne d'action, et afin d'œuvrer à la construction de bases de données effectives et opérationnelles en enseignement et évaluation de la langue anglaise, le projet *English Profile* (2011 : 2)⁴ ambitionne :

- D'établir une description référentielle de la langue anglaise par niveau correspondant aux prescriptions émises par le *Cadre européen commun de référence pour les Langues – Apprendre, Enseigner, Évaluer* (CECRL),
- De fournir un ensemble fondamental d'outils de référence destiné aux professionnels de l'enseignement de la langue.

D'énormes avancées ont vu le jour en fournissant des critères pour distinguer les niveaux dans les domaines du vocabulaire et de la grammaire, alors que la phonologie reste le parent pauvre. Le domaine prosodique est le moins avancé, bien que des études sur le rythme et la tessiture permettent des repérages en catégorisations de niveaux globales. Il reste donc à se pencher sur l'ensemble des sous-domaines prosodiques et notamment sur la mélodie pour tenter de contribuer à l'effort international général.

³ Chan, K. L., « Power Language Index » (25 mai 2016). *INSEAD*.

[https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&cad=rja&uact=8&ved=0ahUKEwjy2Yzd_8nVAhUJB8AKHSYvCQcQcOFggyMAI&url=http%3A%2F%2Fwww.kailchan.ca%2Fwp-content%2Fuploads%2F2016%2F09%2FKC_PLIpresentationatINSEAD_25May2016_v2.pdf&usq=AFQjCNFttSlGmEli32oyvePcX6bsnZpWQJ], consulté en 2017.

⁴ *English Profile Programme (EPP)*, initié par le groupe *Cambridge ESOL* des certifications de Cambridge en collaboration avec les éditions *Cambridge University Press* et d'autres participants en 2005.

Cette thèse a pour titre : « Élaboration de critères d'évaluation prosodique pour anglicistes francophones à partir de la notion de « profil » et du phénomène d'emphase ». L'objet de recherche est donc la prosodie anglaise. Non prise en compte pour elle-même dans ses particularités, elle représente le terrain, la base, le fondement sur lesquels la parole non native, spécifiquement francophone, est étudiée. Pourtant, l'objectif se centre non pas sur l'apprentissage de la langue, mais sur l'évaluation de son degré de maîtrise et ce qui le justifie.

Ce sont essentiellement ces deux notions d'« évaluation » et de « maîtrise prosodique en anglais » qui bornent le parcours de cette recherche. Les autres termes n'en sont que des corollaires que je justifierai par la suite. Éclairons tout d'abord ces deux aspects fondamentaux.

3 Spécificité de la prosodie

On a déjà qualifié la prosodie de « parent pauvre » du domaine phonéto-phonologique. Afin de présenter ses nature et particularités, cette section apporte les concepts fondamentaux la définissant, puis aborde le fonctionnement prosodique et précise en quoi son organisation relève d'une conjonction spécifique de corrélats.

3.1 Concepts prosodiques définitoires

De nombreuses définitions de la prosodie de la parole émaillent les dictionnaires linguistiques et la littérature spécialisée. Vaissière (2006 : 100-101) en donne une vue détaillée visant l'exhaustivité, qui a toute sa place ici :

La prosodie est une notion difficile à définir. Du point de vue *acoustique*, elle correspond aux variations de la fréquence fondamentale, de la qualité de voix (variations dans le mode de vibration des plis vocaux), aux variations de la durée et de l'intensité physique des sons, ainsi qu'aux variations allophoniques [...]. Les modifications de ces paramètres impliquent parfois la participation de l'ensemble des *organes* dits « de la parole », aux niveaux glottique, sous-glottique et supraglottique : rythme de parole, mode de vibrations des plis vocaux, intensité du geste d'expulsion de l'air des poumons, vitesse, force et précision des gestes de la langue et des lèvres. En fonction de la déviation de ces paramètres par rapport aux valeurs attendues pour le même énoncé qui aurait été énoncé de façon neutre, le « ton » de la voix sera perçu comme triste et mélancolique, ou encore joyeux et enjoué. D'un autre point de vue, la prosodie peut être définie par ses fonctions : fonction lexicale, démarcative, pragmatique, comportementale, émotionnelle,

identificatrice, stylistique ; du point de vue *linguistique*, la prosodie est souvent décrite comme la somme des phénomènes d'accentuation lexicale et d'intonation, ainsi que des facteurs de performance, dont le rythme. Toutes ces définitions sont acceptables, mais il faut veiller à ne pas mélanger les différents points de vue dans une description des faits.

Ainsi, cette présentation concerne divers niveaux d'analyse ou d'interprétation et peut être résumée comme le fait Di Cristo (2013 : 1), en écrivant que la prosodie et les éléments qui la déterminent peuvent être définis de plusieurs manières : par leur nature formelle, par leur matérialité physique ou auditive, par leur fonctionnalité ou même par leur contribution à la signification du message. Reprenant les diverses définitions du terme à partir de nombreux ouvrages de référence, Di Cristo (2013 : 20-21) récapitule que, de par ses racines grecques, signifiant « chant mélodique » ou « métrique », le terme « prosodie » est souvent confondu avec celui d'« intonation » dont il peut être synonyme. Sa particularité est de se référer à des phénomènes qui, échappant à la double articulation du langage et ne recoupant pas les domaines traitant des segments, phones et phonèmes, on leur attribue un caractère « suprasegmental » puisque leur unités linguistiques (« prosodèmes ») englobent et dépassent celle des éléments phonétiques et phonémiques.

Si Wells (2006 : 2) assimile les sens des termes *intonation*, *prosody* et *suprasegmentals*, Crystal (2008 : 393) stipule que le domaine prosodique est restreint par rapport à celui du domaine suprasegmental, lequel englobe celui du paralinguistique.

prosody (*n.*) A term used in suprasegmental phonetics and phonology to refer collectively to variations in pitch, loudness, tempo and rhythm. Sometimes it is used loosely as a synonym for 'suprasegmental', but in a narrower sense it refers only to the above variables, the remaining suprasegmental features being labelled paralinguistic.

Ainsi, la prosodie se limite à l'étude des quatre éléments qui la composent : la hauteur mélodique, l'intensité, le tempo et le rythme. Néanmoins, Crystal (2008 : 349) note dans sa définition du paralinguistique (*paralanguage*), que ce terme correspond à des variations de la voix d'une fréquence inférieure à celle des corrélats prosodiques, surtout en ce qui concerne l'intonation et l'accentuation. Le domaine paralinguistique semble se définir par un « ton de voix » indiquant un comportement spécifique, un rôle social, ou quelque signification se rapportant à la manière de s'exprimer.

paralanguage (*n.*) A term used in suprasegmental phonology to refer to variations in tone of voice which seem to be less systematic than prosodic features (especially intonation and stress). Examples of **paralinguistic features** would include the controlled use of breathy or creaky voice, spasmodic features (such as giggling while speaking), and the use of secondary articulation (such as lip-rounding or nasalization) to produce a tone of voice signalling attitude, social role, or some other language-specific meaning. Some analysts broaden the definition of paralanguage to include kinesic features; some exclude paralinguistic features from linguistic analysis.

Cette définition de la paralangue ne me semble pas convaincante pour clairement délimiter les domaines prosodiques et paralinguistiques dans la mesure où certains des quatre éléments prosodiques sont susceptibles d'avoir une incidence sur l'inflexion de la voix (*tone of voice*) caractérisant le locuteur et non plus le message linguistique. Cet aspect méritera une certaine vigilance dans l'étude en cours.

Tenant compte des diverses définitions, Di Cristo propose un essai de définition synthétique (2013 : 21) :

La prosodie est une branche de la linguistique consacrée à l'analyse des propriétés formelles (phonologie prosodique), de la matérialité (phonétique prosodique) et de la fonctionnalité des éléments non verbaux de l'expression orale, non coextensifs aux phonèmes, tels que l'accent, les tons, l'intonation, la quantité, le tempo et les pauses, que l'on qualifie collectivement de prosodèmes. La matérialité physique des prosodèmes se manifeste par les variations de la fréquence fondamentale (FO), de la durée et de l'intensité (paramètres prosodiques acoustiques), ces variations étant perçues par l'auditeur comme des changements respectifs de hauteur, de longueur et de volume sonore (paramètres prosodiques auditifs). Les informations prosodiques ont la particularité d'être polysémiques et de véhiculer à la fois des informations d'ordre linguistique, paralinguistique et extralinguistique, qui se révèlent déterminantes pour l'interprétation des énoncés du discours et de la conversation.

On peut remarquer que Di Cristo élargit ici le champ des domaines sous incidence prosodique, mais réduit les corrélats acoustiques des prosodèmes à trois : la fréquence fondamentale (FO), la durée et de l'intensité (paramètres prosodiques acoustiques), chacun permettant de mesurer la hauteur mélodique, la longueur et le volume sonore, dont la raison d'exister est d'être entendus, c'est-à-dire perçus auditivement.

Par ailleurs, Laver et Hutcheson (1972 : 12) considèrent d'autres distinctions, des canaux qui sont le « vocal » et le « non vocal » (*non-vocal*) ainsi que le « verbal » et le « non verbal » (*non-verbal*). Wacewicz et Żywiczyński (2012 : 118) résument leur classification des signaux communicatifs en quatre catégories :

- communication vocale et verbale, c'est-à-dire les énoncés construits avec des unités linguistiques,
- signaux vocaux et non verbaux, qui correspondent à l'intonation, l'accentuation, le rythme ou la qualité vocale, lesquels excluent les unités linguistiques,
- signaux verbaux non vocaux: expression d'unités linguistiques excluant le domaine de la parole, mais englobant d'autres canaux comme l'écriture ou le langage des signes,
- signaux ni verbaux ni vocaux : signes ne relevant ni d'unités linguistiques ni d'une production vocale, mais pouvant inclure contact oculaire, postures, gestes manuels ou corporels, à condition que ceux-ci ne soient pas des signaux (dotés d'une signification établie par convention).

Ce classement recadre la place de la prosodie parmi les signes et signaux émaillant un acte de communication en présentiel. Si l'on se fie à ce classement, le sujet de cette thèse cible donc les activités vocales mais écarte celles qui sont non vocales étant donné que ces dernières ne sont aucunement liées au domaine de la parole (*speech production*)⁵. Laver et Hutcheson (1972 : 13) ajoutent une nouvelle catégorie à leur classement : l'extralinguistique, dont les signes se caractérisent par leur incapacité à varier selon la situation, et incluent par exemple la qualité vocale du locuteur. Cette catégorie sera elle aussi à écarter des analyses évaluatives car elle est soumise aux caractéristiques intrinsèques des individus, lesquelles ne sont pas transférables et ne peuvent dans ce cas être prises en compte. Dans son étude sur « la relation voix-parole », Lothe (1982 : 233) rappelle que la distinction de Laver et Hutcheson entre « linguistique », « extralinguistique » et « paralinguistique » provenait au départ des travaux de Trager (1958)⁶, Crystal et Quirk (1964)⁷ et Abercrombie (1967-1968) et que l'ensemble de

⁵ Ce point trouvera son importance lors de la présentation du corpus de travail : parmi les signaux vocaux non verbaux, on pourra inclure les pauses pleines, exprimant l'état physique ou émotionnel du locuteur (*corporeals* selon J. J. Ohala (1994 : 2) précisent Wacewicz et Żywiczyński). Le langage corporel (*body language*, Pease 1981), faisant lui aussi partie des signes non vocaux, sera écarté lui aussi étant donné que le corpus de travail n'est pas multimodal.

⁶ Trager, G. L. (1958). "Paralanguage: A First Approximation". *Studies in Linguistics*, vol. 13 (1-12).

⁷ Crystal, D., Quirk, R. (1964). *Systems of Prosodic and Paralinguistic Features in English*. De Gruyter Mouton.

ces chercheurs soutenaient que ces phénomènes n'étaient ni personnels, ni idiosyncratiques mais de nature culturelle et variaient donc selon les groupes (travaux d'Abercrombie 1967-1968)⁸.

La préférence de Lothe (1982 : 235) va à une distinction prosodique en niveaux d'analyse et considère que l'on observe une plus grande généralisation et une plus grande abstraction au fur et à mesure que l'on étudie ce qui caractérise la parole et/ou la voix d'un locuteur, de plusieurs locuteurs, ou ce qui caractérise les patrons linguistiques de la parole quel que soit le locuteur. Elle essaie de réconcilier le point de vue du locuteur et celui de l'auditeur et s'appuie sur l'analyse auditive pour réussir cette alliance.

En effet, il est difficile d'obtenir un consensus sur les liens unissant matérialité prosodique et abstraction. Sur le plan phonétique et jusqu'à nos jours, l'École britannique de l'intonation utilise la catégorisation de Halliday (1963) en *tonality*, *tonicity* et *tone* (3Ts) signifiant respectivement le découpage en unités intonatives (*intonation phrases*, *IPs*), l'emplacement des proéminences dans la chaîne parlée et la mélodie (Wells 2013). Ce n'est pourtant pas le cas de Di Cristo (2013 : 83), qui propose le schéma ci-dessous pour retrouver l'organisation des correspondances entre les faits concrets et les entités abstraites.

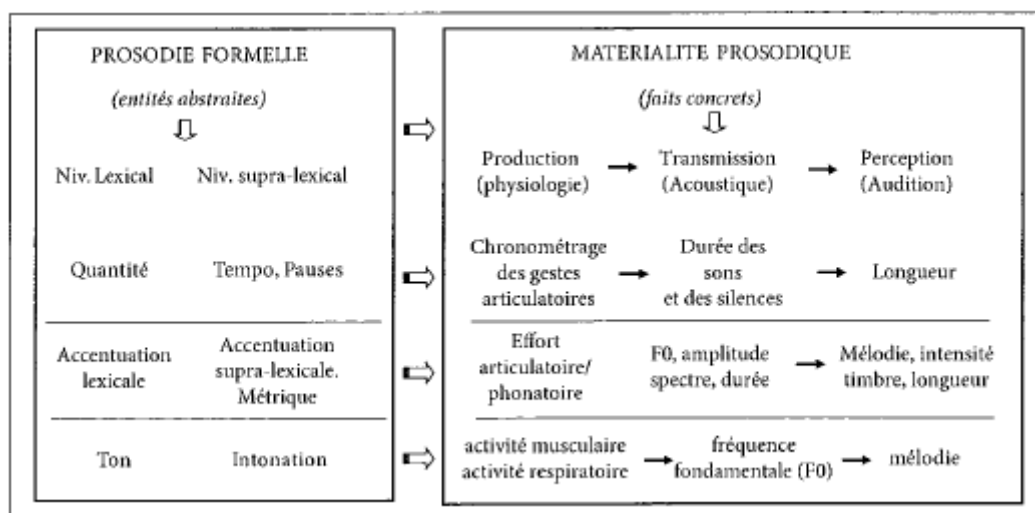


Figure 1 – Schéma des correspondances entre les composantes formelles et matérielles de la prosodie (Di Cristo, 2013 : 83)

Ce schéma présente deux tableaux dont celui de gauche synthétise la manière dont les entités abstraites sont traitées en prosodie formelle. Aux deux niveaux que sont le domaine lexical et le domaine supra-lexical, on observe une adaptation des trois T ciblées sur le rythme par

⁸ Abercrombie, D. (1967). *Elements of General Phonetics*. Edinburgh University Press.

Abercrombie, D. (1968). "Paralanguage". *British Journal of Disorders of Communication*, vol. 3, p. 55-59; in: Laver-Hudcheson (1972: 64-70).

l'accentuation lexicale ou les proéminences supra-lexicales, et sur la mélodie par les tons ou l'intonation, à l'exception de *tonality*, qui est remplacée par la notion de « quantité » au niveau lexical, et « tempo, pauses » au niveau supra-segmental. Le second tableau, intitulé « Matérialité prosodique », est subdivisé en trois colonnes représentant les étapes chronologiques que sont la production physiologique, la transmission (phénomène physique dont on peut rendre compte en acoustique) et la réception par perception. L'humain est à la fois source et aboutissement, mais faute de bénéficier d'un encadrement spécifique à la production et à la réception, le chercheur doit se contenter des mesures acoustiques qu'il pourra effectuer à partir d'enregistrements. Vaissière (2006 : 43) affirme que certaines données ne peuvent être prises qu'en milieu hospitalier ou en laboratoire de phonétique à cause du matériel lourd généralement utilisé par les cliniciens pour mesures électromyographiques de l'activité électrique, ou prises de données cinéradiographiques (par flexoscope, variété d'endoscope). Elle note l'utilisation de plus en plus fréquente de « *L'imagerie par résonance magnétique (IRM), la transillumination du larynx, l'imagerie cérébrale fonctionnelle, l'électroencéphalographie (EEG) et la magnétoencéphalographie (MEG)* ». Ces traitements et outils permettent alors d'étudier précisément gestes et activité corporelle pendant la phonation. La phase de transmission ne peut être détaillée finement que grâce à un enregistrement effectué pendant la phase de production pour étude acoustique de ce que l'on peut mesurer : durée, fréquence fondamentale, amplitude et spectre. La phase de réception par audition ne peut être sondée que par le truchement de questionnaires ou leurs équivalents pour tenter d'évaluer ce que l'humain perçoit comme longueur, mélodie, intensité ou timbre de voix. Ce schéma représente la vue la plus synthétique et complète des divers aspects de la parole telle qu'elle peut s'inscrire dans cette recherche.

De même que la parole est destinée à être perçue (ce qui présuppose un auditeur ou interlocuteur), et non prioritairement à être mesurée, cette même parole est appelée à avoir un impact sur la réalité du monde dans laquelle elle s'inscrit. Lacheret (2007) affirme :

Concernant plus spécifiquement la prosodie, l'objectif est de comprendre comment les phénomènes intonatifs contribuent à l'inscription des sujets dans les situations dans lesquelles ils sont engagés (environnement matériel, social et cognitif). [...] La prosodie et l'action, la prosodie en action, l'angle d'attaque choisi insiste sur l'aspect *situé* de la cognition.

Elle préconise comme angle d'attaque la vision incarnée de la cognition, que l'on trouve dans la théorie de l'énaction chez Varela *et al.* (1993) et chez Bottineau (2007). Bottineau (2008)

insiste sur cette faculté révolutionnaire qu'a la parole de modifier non seulement le monde mais aussi les acteurs que sont le locuteur et l'auditeur-interlocuteur : « [...] language is to be reconsidered in terms of sensorimotor interactions with an environment in which both the individual and the environment are modified [...] ». Il devient très clair que l'on ne peut étudier la prosodie sans prendre en compte ses multiples facettes, ni le monde dans lequel elle évolue : la matérialité prosodique doit être préservée dans sa diversité, l'objet d'étude provenant de la matérialisation de l'activité humaine.

En conséquence, ce travail de recherche prend appui sur les considérations exposées précédemment en favorisant l'organisation prosodique de Di Cristo. Il prend aussi en compte la théorie de l'énaction car l'individu est un tout qu'il faut relier à son environnement, ses expériences passées, ses modes de pensée et éventuellement son profil d'apprentissage pour comprendre les phénomènes en jeu. Il semble peu approprié à la difficulté de la présente tâche de penser que la seule étude automatique ou manuelle de corpus de plus en plus vastes puisse compenser un manque de discernement que seul le choix raisonné d'outils et de conceptions appropriés pourra éclairer.

Pour clarifier les notions utilisées dans ce travail, voici comment je désignerai les domaines étudiés. Je garderai au mot « intonation » le sens de « mélodie » et utiliserai ce dernier terme afin d'éviter toute confusion. La « prosodie » regroupera les domaines mélodiques, rythmiques (accentuation et métrique) et de temporalité (tempo et pauses). Je rejoins ainsi l'organisation prosodique présentée par Di Cristo en amont.

3.2 Fonctionnement prosodique

Suite à la présentation de la caractérisation de la prosodie, procédons à l'étude de son fonctionnement. La citation de Vaissière liste plus haut les fonctions de la prosodie dans la parole : lexicale, démarcative, pragmatique, comportementale, émotionnelle, identificatrice, stylistique. Elle se situe alors essentiellement dans une communication orale inter-individuelle. Comment le fait-elle ? Par quels moyens y parvient-elle ? Comment s'inscrit-elle au sein des systèmes de production et de compréhension de la parole ?

Voyons tout d'abord comment la parole est utilisée en situation. Selon Fónagy ([1983] 1991 : 36), la communication verbale à vive voix subit deux actes successifs d'encodage. Le premier est linguistique et transforme la teneur du message en une séquence de phonèmes alors que le second vient se superposer au premier. Le premier encodage est celui du code

linguistique et le second est un codage paralinguistique. Le graphique ci-dessous schématise le processus.

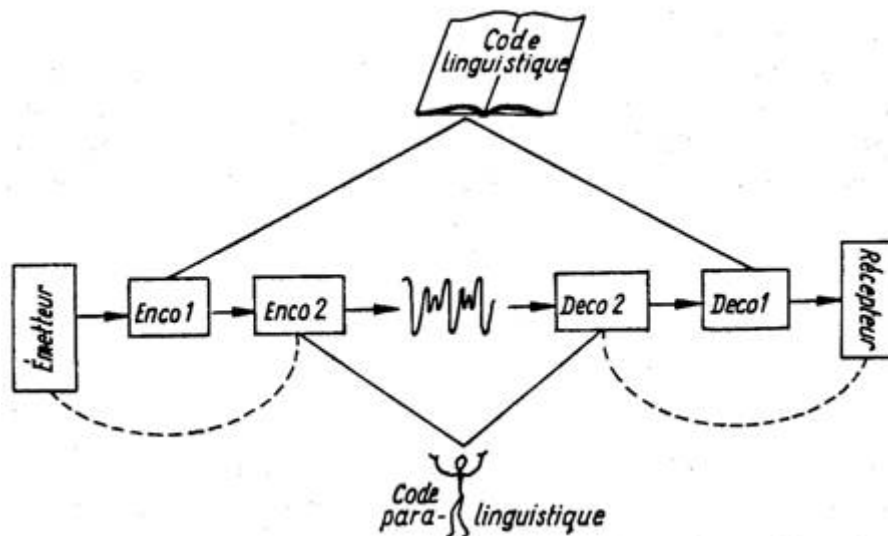


Figure 2 – Schéma de la communication verbale vive voix. — L’encodage linguistique est suivi d’un deuxième acte d’encodage à partir d’un « code naturel », pré-linguistique (Fónagy ([1983] 1991 : 14)

Fónagy ([1983] 1991 : 15) explique son schéma en faisant remarquer que la mention des codes se trouve soit au-dessus de la ligne, soit en dessous, afin de suggérer que ces deux opérations se situent à deux niveaux mentaux différents. Il poursuit : « Le message échappe à la conscience : on ne perçoit que son effet. ». Ce second encodage ou « code naturel » est subconscient et s’appuie sur les données suprasegmentales. Le « double codage » obéit au principe de la condensation inhérente à la parole comme instrument de communication orale. Cette théorie du « double codage » est mise en évidence par la fragilité psychique de certains individus (enfants, malades psychotiques, etc.), qui occultent le premier niveau pour ne prendre en considération que le deuxième, plus primitif et donc davantage à leur portée. Ce deuxième niveau déclenche le passage par une troisième dimension d’ordre psychanalytique. Le niveau inconscient commande une reconstitution d’opérations mentales similaires préconscientes. Ainsi, Fónagy tente de relier la théorie freudienne des pulsions à la théorie de la forme phonique⁹.

Fónagy note la nécessité pour le code paralinguistique d’être indépendant du premier codage : « le principe de condensation inhérente à la communication à vive voix suppose la distorsion d’un message linguistique primaire » ([1983] 1991 : 17). Il insiste sur le caractère

⁹ [<http://www.universalis.fr/encyclopedie/ivan-fonagy/>], consulté en 2017.

non anarchique de cette opération et dont les règles sont « strictes et relativement simples ». Les trois règles fédératrices figurent ci-dessous ([1983] 1991 : 17-18) :

« 1. La reproduction volontaire des *symptômes* vocaux d'une émotion signale la présence de cette émotion. »

« 2. Les organes de la parole peuvent *représenter*, symboliser d'autres objets animés ou inanimés qui leur sont associés par la ressemblance ou une analogie fonctionnelle. »

« 3. Le principe de l'isomorphisme de l'expression et du contenu exige qu'aux différents degrés d'intensité sémantique correspondent différents degrés d'intensité sur le plan de l'expression sonore ainsi

3.1. à l'intensité de l'émotion correspond l'intensité de l'activité musculaire [...] »

Fónagy voit l'encodage secondaire comme « une déformation sémiotique (significative) du message primaire » à l'instar d'un modulateur non incorporé à la grammaire, mais par lequel tous les phones passeraient obligatoirement. Ainsi, tous les sons concrets sont expressifs, et « La non distorsion n'est pas nécessairement moins suggestive que la distorsion, elle exprime ou peut exprimer autre chose (une attitude froide, volontairement neutre, etc. conformément au texte et à la situation). Il considère que les éléments du code paralinguistique (comparés à ceux du code linguistique) sont ([1983] 1991 : 19) :

- Des signes entiers, non articulés,
- Des signes toujours motivés, non arbitraires,
- Des signaux de caractère continu, non discret, car ils se conforment et sont proportionnés au degré d'intensité de l'émotion.

Il ressort de tous ces éléments que l'usage prosodique n'est jamais neutre. La difficulté est alors de ne pas se fourvoyer en utilisant un signe à la place d'un autre ; cette méprise ne serait pas forcément détectée et le message serait dévié de son objectif.

Nous avons jusque là essentiellement abordé le domaine vocal. Dans l'acte de parole, le locuteur se trouve dans une situation qui englobe plusieurs autres ressources, qui peuvent être utilisées pour appuyer son message. Di Cristo (2013 : 36) les représente toutes sous forme de schéma.

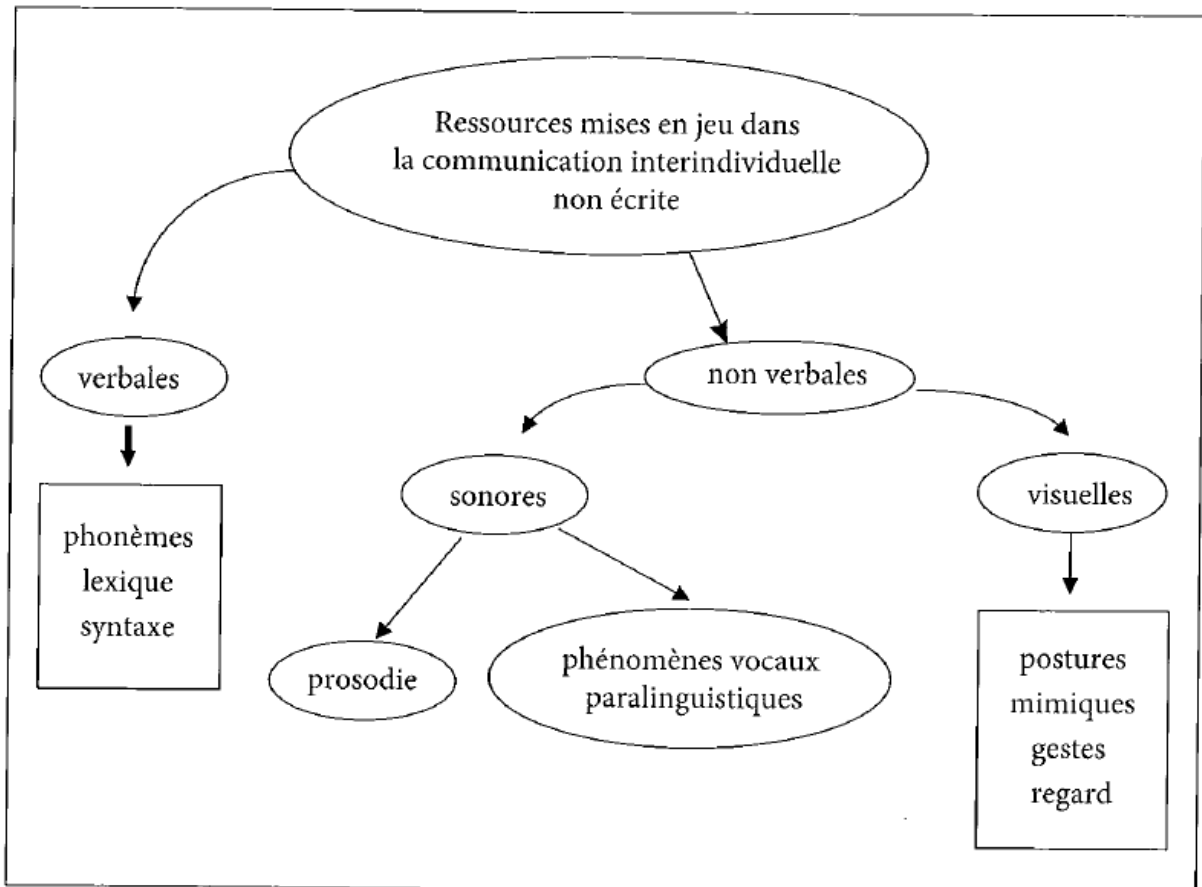


Figure 3 – Ressources mises en jeu dans la pratique de la communication orale inter-individuelle (Di Cristo 2013 : 36).

Ce graphique nous fait part de la dimension non verbale non vocale (extra-linguistique) signalée plus haut. L'émission et le décryptage des comportements manuels ou corporels, attitudes faciales ou oculaires peut, tout comme le second codage (paralinguistique verbal) appuyer, moduler ou contredire le contenu du message verbal. Il s'agit ici d'un troisième canal. Contrairement aux deux autres canaux sonores, celui-ci est visuel. En conséquence, il ne peut être d'utilité que si l'enregistrement est audio-visuel et l'étude multimodale. On peut constater à présent la complexité des phénomènes en jeu dans l'étude de réalisations prosodiques. Il s'agirait de décomposer les niveaux de sens, qui se superposent et agissent en synergie.

3.3 L'organisation prosodique : une conjonction spécifique de corrélats

La présentation ci-dessus scindée en domaines temporel, intonatif et rythmique est exposée de manière abstraite. En fait, on verra par la suite que les divers corrélats acoustiques fonctionnent en synergie pour créer une relation d'échange dans la prosodie, tant au niveau des prises de mesure que dans les effets obtenus. On pourrait même avancer que l'organisation prosodique relève du principe d'autopoïèse, défini par les biologistes chiliens

Maturana et Valera dans leur description abstraite du métabolisme cellulaire, qu'ils qualifient de « machines vivantes » :

An autopoietic machine is a machine organized (defined as a unity) as a network of processes of production (transformation and destruction) of components that produces the components which: (i) through their interactions and transformations continuously regenerate and realize the network of processes (relations) that produced them; and (ii) constitute it (the machine) as a concrete unity in the space in which they (the components) exist by specifying the topological domain of its realization as such a network. (Maturana, Valera, 1980 : 78-79)

Cette propriété qu'a l'organisme vivant à se produire en se régénérant et à interagir en permanence avec un environnement en mutation perpétuelle auquel il doit s'adapter, est similaire, me semble-t-il au système régissant l'organisation prosodique. C'est ce que l'on peut constater dans l'agencement des corrélats acoustiques qui la composent : leur interaction incessante, à divers niveaux, s'organise de telle manière que leur relation systémique est préservée afin de véhiculer le sens attendu. Le chaos de la variation acoustique fait sens perceptivement lorsque les « petits bouts » composant ce « bric-à-brac » sont observés. J'oserais même une comparaison avec la technique impressionniste, qui révèle une forme significative lorsque l'observateur prend de la distance par rapport au tableau exposé, et la soustrait aux yeux inquisiteurs, trop rapprochés pour obtenir une vision d'ensemble. Dans ces systèmes mouvants, la cohérence de l'ensemble l'emporte sur les détails, jugés incidents s'ils sont considérés individuellement. La dynamique du vivant prend le pas sur la stabilité d'une structure rigide, qui ne survivrait pas et deviendrait rapidement obsolète et donc éphémère, alors qu'elle doit être maintenue pour atteindre ses objectifs.

Aussi, les mesures des divers corrélats doivent être interprétées en conjonction. C'est ce que constatait déjà en 1954 Stannard Allen, en soulignant le caractère fluide et mouvant, « vivant » en quelque sorte, de la parole :

Stress, rhythm and intonation should really be considered as a whole, for they are very closely connected elements of a single aspect of the language that we might call Speech Flow. Speech is essentially movement. (Stannard Allen, 1954 : xiii)

Cette complexité ne doit pas occulter celle dans laquelle la prosodie s'insère. Le schéma ci-dessous propose une esquisse de représentation de la communication orale interindividuelle, dans lequel le « compilateur prosodique » opère (Di Cristo : 2013 : 44).

Centré sur le verbal, il détaille les différentes phases auxquelles le locuteur et l'allocutaire sont confrontés lors de l'émission et la réception du message verbal. On y constate la place centrale de la prosodie dans les phases de conceptualisation, formulation, planification phonétique et articulation du locuteur, et les différents niveaux de traitement mis en place par l'auditeur à qui le message est destiné. C'est la maîtrise de cette complexité qui doit être évaluée chez l'apprenant. Je reproduis page suivante la représentation de la complexité que propose Di Cristo (2013 : 44), avec la quelle je suis particulièrement d'accord.

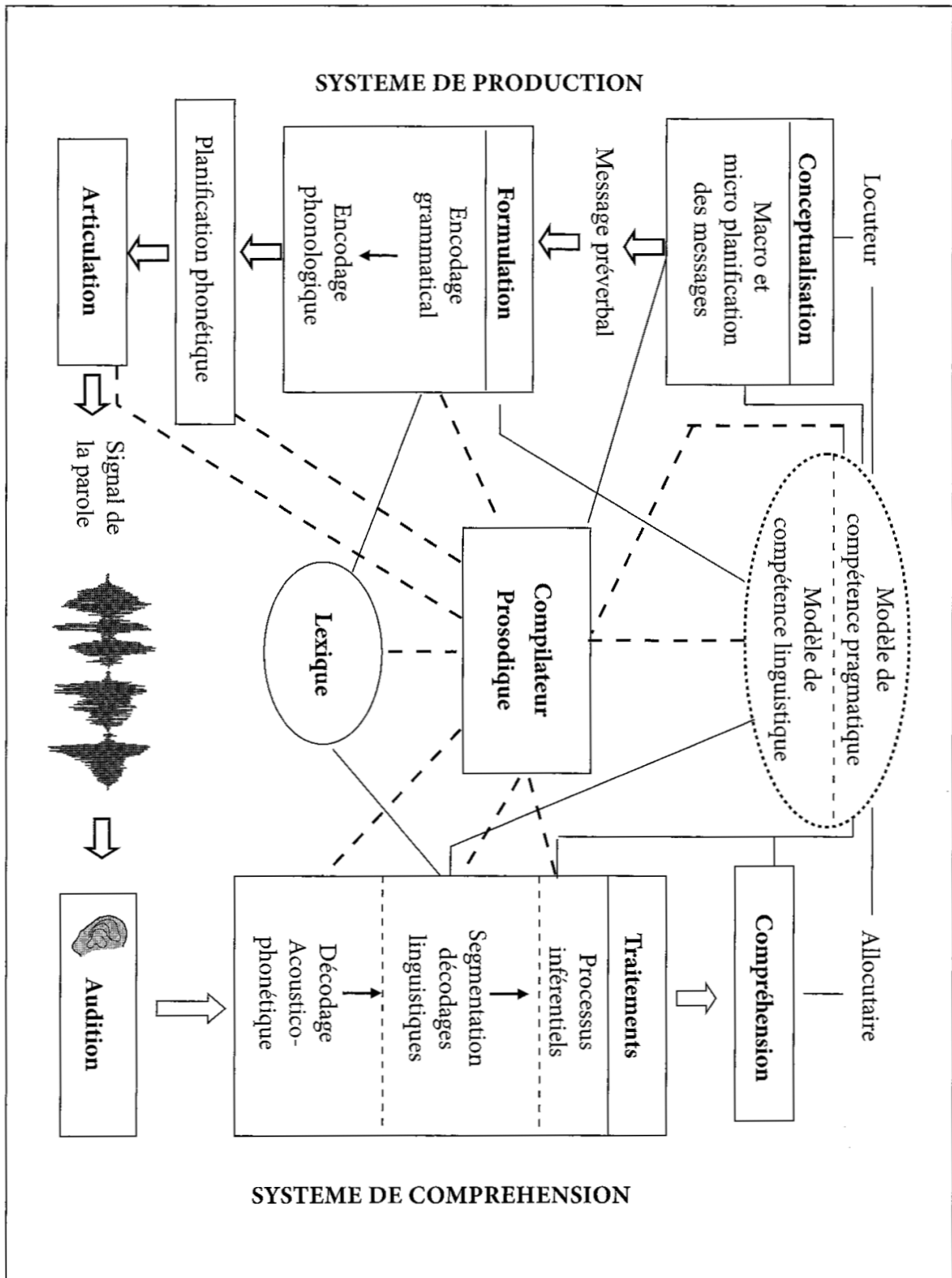


Figure 4 – Place de la prosodie (du compilateur prosodique), au regard des systèmes de production et de compréhension de la parole (Di Cristo 2013 : 44)

4 Enjeux de l'évaluation prosodique

Pour cerner les tenants et les aboutissants du phénomène évaluatif, une première sous-section se focalisera sur ses concepts définitoires, la seconde sur ses fonctions et la troisième sur les phases de développement d'un test évaluatif.

4.1 Concepts définitoires fondamentaux en évaluation

Dans son étude sur l'évaluation, Tardieu (2013 : 248) rappelle les définitions qu'en donne De Ketele : « [...] évaluer c'est confronter un ensemble d'informations à un ensemble de critères (référentiel) en vue de prendre une décision (De Ketele, 1980) ». Tardieu remarque alors que cette évaluation « implique une prise de décision qui échappe au sujet évalué ». Elle note aussi que selon Gardner (1996 : 114), « Évaluer, c'est recueillir des données sur les compétences et les potentiels des individus dans le double objectif de leur en faire retour utilement et de procurer des informations indispensables à la communauté environnante ». Cette seconde définition établit deux destinataires des informations obtenues via l'évaluation.

Qu'elle prenne la forme de tests, de certifications, d'examens ou de contrôles, l'évaluation repose sur trois concepts fondamentaux qui sont la validité, la fidélité (aussi dénommée fiabilité) et la faisabilité. Horner (2010 : 17) ajoute deux autres concepts à cette liste pour évaluer la performance : l'objectivité et l'équité. Reprenons les définitions attribuées à ces notions.

La validité (validity) : « Une activité d'évaluation est valide lorsqu'elle mesure exclusivement et exactement ce qu'elle est censée mesurer. » (Horner, 2010 : 17). Ce qui est effectivement évalué est le *construct*¹⁰. Ce modèle de la compétence évaluée doit être défini clairement et précisément. Il vise les connaissances nécessaires et leur mise en œuvre dans un contexte donné à travers un fonctionnement cognitif (Horner, 2010 : 17). L'*Association des évaluateurs de langue en Europe (ALTE, 2011 : 14)* précise que le degré de validité dépend en outre de la manière dont le test est utilisé : si l'accent est mis sur les compétences en langue ou bien sur les comportements sociaux dans le cas des évaluations liées au *CECRL*. Elle préconise les étapes suivantes en se basant sur les travaux de Kane, Crooks et Cohen (1999) et de Bachman (2005) pour vérifier la validité du processus de la construction du test jusqu'à la vérification de l'aptitude du candidat dans sa vie personnelle ou professionnelle.

¹⁰ Selon *ALTE* (2011 : 82) : Le *construct* se définit : *A hypothetical or mental TRAIT which cannot necessarily be directly observed or measured, for example, in language testing, listening ability.*

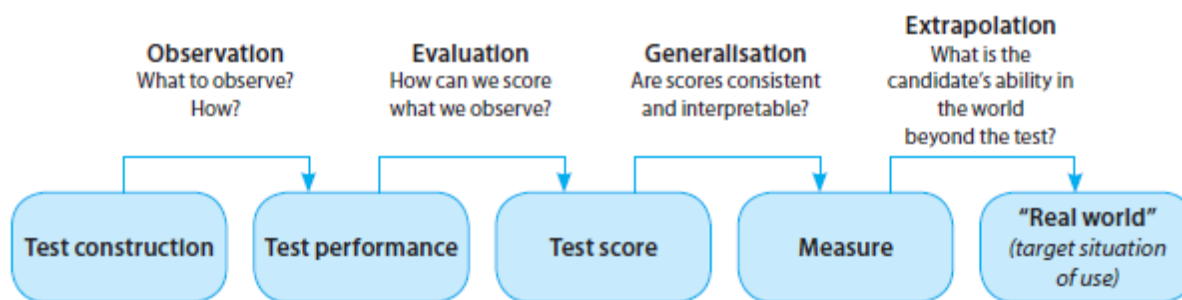


Figure 5 – Raisonnement séquentiel dans une démarche validante selon ALTE, adapté à partir des travaux de Kane, Crooks et Cohen 1999, et Bachman 2005 (2011 : 15)

Le succès des diverses étapes conditionne la validité de l'ensemble (ALTE 2011 : 15). La première phase se concentre sur l'interprétation d'un échantillon de réalisation basé sur un exemple représentatif de ce que l'apprenant a réalisé dans la compétence. L'élaboration du barème qui s'ensuit concerne la phase 2. L'étape 3 considère la généralisation des résultats numériques obtenus à d'autres situations et aborde le domaine de la fiabilité (voir ci-dessous). La phase quatre consiste en l'extrapolation de l'évaluation de la performance au monde extérieur au test en le transposant sous la forme de compétences traduites par *I can ...*. Dans la phase 5, le candidat peut alors être profilé en conséquence.

La fidélité (*reliability*), aussi appelée fiabilité, est « l'assurance que les résultats obtenus restent toujours stables et constants quelles que soient les conditions de passation et de correction. Une évaluation est fidèle si l'on peut se fier aux résultats obtenus, soit de classement (les sujets se retrouvent toujours classés dans le même ordre) soit de niveau (les sujets sont toujours placés au même niveau de compétence : par exemple, A2 ou B1 sur l'échelle du *Cadre européen commun de référence en langues*, ou *CECRL*). C'est cette fiabilité qui nous permet d'avoir confiance en l'exactitude des résultats de l'évaluation. » (Horner, 2010 : 18). Pour obtenir la fiabilité de ses résultats et atteindre une norme requise, il est donc nécessaire de standardiser tous les procédures de l'évaluation, notamment l'objectivité des corolaires et la clarté des grilles évaluatives. Pour les concepteurs de tests, l'objectif recherché est d'obtenir le plus haut degré de validité et de fiabilité. Il est ordinairement souhaitable que deux tests évaluant le même *construct* soient corrélés par « validité convergente ». ALTE (2011 : 16) assimile *reliability*, *consistency* et *dependability*, et précise que la fiabilité d'un test n'implique ni une qualité satisfaisante du test ni la validité de l'interprétation donnée à ses résultats. La variabilité des divers résultats obtenus, ou *error*, provient des facteurs représentés dans la figure ci-dessous, *other* signifiant que ces facteurs sont hors contrôle. Quelle que soit son origine, le degré de ce taux d'erreur présent dans tout

test évaluatif doit être réduit au maximum pour que les résultats obtenus soient attribuables au degré de compétence du candidat et non à des facteurs différents. Les efforts dans ce sens porteront essentiellement sur le degré d'erreur provenant des évaluateurs et des diverses versions du test.

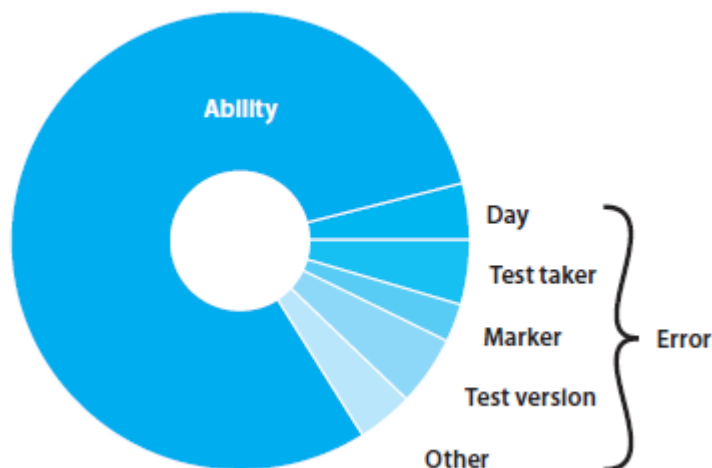


Figure 6 – Quelque sources d'erreur dans un score évaluatif (ALTE, 2011 : 16)

L'objectivité est définie comme « la qualité d'une évaluation qui fait que les résultats obtenus par un sujet donné seront constants quel que soit l'examineur. » (Horner, 2010 : 21). Ce critère de fidélité est essentiel pour l'évaluation de la performance. On a vu précédemment que ALTE (2011 : 16) incluait le critère d'objectivité dans la notion de fiabilité du test.

L'équité (*fairness*) permet d'éviter toute contestation en n'établissant aucune différence de traitement entre les candidats au test (Horner, 2010 : 21). Le partenariat croissant entre les institutions et les créateurs de certifications encourage le respect de « codes de bonne conduite » (*Codes of Practice, Codes of Fairness*) (ALTE, 211 : 17).

Une notion reste fondamentale : la faisabilité (*practicality* ou *feasability*). La procédure d'évaluation doit être « pratique à organiser, administrer et corriger » (Horner, 2010 : 22).

À part ces notions générales sur les critères qui définissent la nature et la qualité de toute évaluation, d'autres, plus spécifiques, sont néanmoins centrales pour cerner le sujet de cette thèse.

Un descripteur (*descriptor*) : ALTE (2011 : 82) la définit comme :

A brief description accompanying a band on a rating scale, which summarises the degree of proficiency or type of performance expected for a test taker to achieve that particular score.

Le *CECRL* (2001) utilise abondamment cette notion, par exemple sous la forme de *can do*, et Horner (2010 : 80-81) clarifie la distinction entre « descripteur de compétence » et « critère de mise en œuvre d'un descripteur » :

Un descripteur décrit ou définit un niveau de compétence (A1, A2, B1, B2...) ; un critère est la mise en œuvre d'un descripteur dans une situation d'évaluation. Ainsi, il se peut que l'on soit obligé de modifier le descripteur afin de le rendre opérationnel.

Ainsi, trouver les descripteurs pertinents, révélateurs, est une première étape avant de pouvoir les rendre opérationnels sur le plan évaluatif sous la forme de critères distinctifs. Dans leur *Dictionnaire pratique du CECR*, Robert et Rosen (2010 : 80-81) citent les préconisations de Tagliante (2005 : 47) en matière de descripteurs : « un bon descripteur doit :

- être formulé de façon positive [...] ;
- être précis et décrire explicitement la capacité attendue [...] ;
- être clair, c'est-à-dire rédigé sans métalangue linguistique, de façon à être compris par tous, enseignants et apprenants ;
- être bref, c'est-à-dire ne décrire que l'essentiel ;
- être indépendant, ne pas dépendre d'un ou plusieurs autres descripteurs ».

Horner mentionne la difficulté à rédiger des descripteurs de niveau ou de compétence (2010 : 92) :

Il faut, à la fois, identifier la ou les composantes à évaluer, décrire une performance possible au niveau requis, permettre une notation à l'intérieur de ce niveau afin de différencier les élèves, et être clair, précis et succinct, tout en créant des descripteurs pratiques lesquels permettent de différencier non seulement entre élèves et niveaux mais aussi entre compétences.

En effet, la tâche de dégager des descripteurs étalonnés est ardue, et de distinguer des niveaux plus fins ou étroits l'est encore davantage. C'est pourquoi Horner (2010 : 97) recommande d'adapter des descripteurs existants soit en combinant plusieurs, soit en fractionnant d'autres.

Un trait distinctif ou discriminant (*criteria feature*) : c'est un critère évaluatif, qu'on obtient par conversion d'un descripteur opérationnel. On trouve aussi dans la littérature l'expression « trait critérié » ou « trait critérisé ». Cette notion se base sur le pouvoir qu'a un élément de discriminer des niveaux chez les candidats (ALTE, 2011 : 82). C'est à partir de ce critère distinctif que peut être créée une grille de notation (Horner, 2010 : 118), laquelle indique des franchissements de seuil d'un niveau à un autre. C'est Hawkins et Buttery (2010) qui ont proposé le « concept de trait critérié » comme indicateur de maîtrise d'une langue étrangère ou seconde. Selon eux, certaines propriétés linguistiques caractérisent et identifient chaque niveau de compétence d'une L2 et c'est ce qui permet aux examinateurs d'évaluer instinctivement leurs candidats avec précision. Ce concept de trait critérié repose sur un défi fondamental : identifier des propriétés distinctives. Il se décline en quatre types selon ses propriétés positives ou négatives. Les propriétés positives d'un trait distinctif dépendent de deux aspects : d'une part, le fait que ses propriétés linguistiques soient conformes à la grammaire et qu'il soit utilisé par les natifs de la langue, et d'autre part, la fréquence avec laquelle le non natif réalise ce trait dans sa production. Un trait distinctif négatif n'est pas généré par la grammaire de la langue (les locuteurs natifs le jugent mal formé) et la fréquence de ces « erreurs » doit être réduites. Ainsi, le travail de l'apprenant est d'accumuler les traits positifs pour se rapprocher des réalisations du natif, alors qu'il lui faudra au contraire réduire les traits négatifs puisqu'ils l'éloignent des réalisations natives.

L'effet de retour (*washback effect* ou *backwash effect* voire *instructional value*) : c'est l'impact ou l'influence des pratiques évaluatives (tests, examens ou autres évaluations) sur tous les individus impliqués dans le processus d'enseignement ou d'apprentissage. Qu'elles soient positives ou négatives, ces conséquences, en rapport avec la valeur sociale du phénomène évaluatif, sont largement reconnues (Bailey, 1999). Elle précise que pour certains, il constitue un critère à part entière de la valeur d'une évaluation.

Ces quelques concepts définitoires établissent des bases pour la présentation de la suite.

4.2 Objectifs fonctionnels de l'évaluation

L'objectif assigné à une évaluation pourra être divers ou pluriel. Sont présentés ici les principaux cas.

Évaluation diagnostique : elle « cherche à identifier les forces et les faiblesses d'un ou de plusieurs élèves avec pour objectif de bâtir un parcours pédagogique qui lui/leur permettra

de progresser » (Horner, 2010 : 23). Il s'agit d'une évaluation formative de dépistage destinée à former des groupes de compétences.

Test de positionnement (*placement test*) : son objectif est de classer les apprenants pour constituer des groupes de niveau (Horner, 2010 : 24).

Évaluation formative (*formative assessment*) : elle suit « un processus continu qui permet de recueillir des informations sur les points forts et les points faibles » des apprenants (CECRL ([2001] version restructurée : 141). Elle vise une amélioration de l'apprentissage (Horner, 2010 : 24-25).

Auto-évaluation : c'est « le jugement que l'on porte sur sa propre compétence ». Elle agit comme « un facteur de motivation et de prise de conscience » (Horner, 2010 : 26).

Inter-évaluation : elle permet d'évaluer autrui, par exemple les autres individus d'un groupe (Horner, 2010 : 34).

Évaluation sommative (*summative assessment*) : elle est destinée à contrôler les acquis à la suite d'une formation ou d'un cursus et attribue une note ou un rang (Horner, 2010 : 35). Il s'agit d'une évaluation de compétence.

Évaluation du savoir ou du niveau (*achievement assessment*) : *is the assessment of the achievement of specific objectives – assessment of what has been taught. [...] Achievement assessment is oriented to the course. It represents an internal perspective. CEFR ([2001] restructured version: 183).*

Évaluation de la performance ou capacité (*proficiency assessment*) : *is assessment of what someone can do/knows in relation to the application of the subject in the real world. It represents an external perspective. (CEFR, [2001] restructured version: 183).* C'est la mise en œuvre de la compétence, c'est-à-dire l'évaluation de ce que l'on peut ou sait faire « en rapport avec son application au monde réel » (Horner, 2010 : 36). Il s'agit de mesurer avec la plus grande précision le niveau atteint par un candidat dans l'absolu. Ce genre de test doit être valide et sa fiabilité maximale, tout en préservant la grande finesse discriminatoire d'un pouvoir classant (Horner, 2010 : 37).

Évaluation normative (*proficiency assessment*) : elle crée une hiérarchie parmi les candidats en les ordonnant les uns par rapport aux autres¹¹. Elle nécessite une notation chiffrée.

Évaluation critériée (*criterion-referencing*) : c'est une évaluation analytique consciente guidée par une démarche prescrite s'appuyant sur l'utilisation constante de critères spécifiques (*criterial features*, abordés en amont). En cela, elle s'oppose à une évaluation holistique ou globale. Elle est destinée à réduire radicalement la variation des jugements. Horner précise que « les examinateurs éprouvent des difficultés à mettre en œuvre en même temps un nombre important de catégories : il semblerait, par exemple que manipuler quatre ou cinq catégories présente déjà des difficultés, et qu'à partir de sept catégories, un examinateur n'est plus fiable » (Horner, 2010 : 38). Ce type d'évaluation s'oppose aussi à l'évaluation normative, qui sert à classer les apprenants les uns par rapport aux autres. Le *CEFR* ([2001] restructured version: 184) ajoute :

Criterion-referencing implies the mapping of the continuum of proficiency (vertical) and range of relevant domains (horizontal) so that individual results on a test can be situated in relation to the total criterion space. This involves (a) the definition of the relevant domain(s) covered by the particular test/module, and (b) the identification of 'cut-off points': the score(s) on the test deemed necessary to meet the proficiency standard set. The scales of illustrative descriptors are made up of criterion statements for categories in the descriptive scheme.

L'évaluation critériée permet ainsi une évaluation à la fois plus objective et plus fine. Elle est centrale pour une meilleure compréhension des phénomènes d'acquisition par la pertinence des domaines couverts par le test et l'identification des « points de césure » ou « seuils » (*CECRL*, [2001] version restructurée : 140), qui délimitent les catégories de niveaux. Le supplément au *CECRL* sur *l'Évaluation de compétences en langues et la conception de tests* définit la notion de validité critérielle s'appliquant aux tests : elle correspond à la possibilité de « démontrer le rapport entre les notes obtenues et un critère externe qui est censé mesurer la même capacité » (Milanovic *et al.* 2002 : 52).

Évaluation directe et évaluation indirecte : l'évaluation directe privilégie la validité en évaluant ce que le candidat est en train de faire, ou ce qu'il a fait par le truchement d'un enregistrement. L'évaluation indirecte cible avant tout l'évaluation d'un niveau de

¹¹ *Norm-referencing* is the placement of learners in rank order, their assessment and ranking in relation to their peers. » *CEFR* ([2001] restructured version: 184).

compétence et peut recourir à des QCM pour évaluer indirectement une production par exemple. Elle fonctionne par corrélation avec les résultats d'une évaluation directe. La validité de ses résultats en sera amoindrie, mais suivant les moyens technologiques utilisés, elle pourra être plus rapide, moins coûteuse et d'une fiabilité très élevée (Horner, 2010 : 40-41). Contrairement à l'évaluation directe qui incite à mettre en place des stratégies d'apprentissage propres à obtenir un bon score, elle ne favorise pas l'effet de retour (*washback effect*).

Les types évaluatifs peuvent être nombreux (par exemple, *CEFR* ([2001] restructured version: 183), et son impact encore davantage (Huver, 2014). Ils se concrétisent dans la construction d'un test.

4.3 Phases de développement d'un test évaluatif

Les étapes de développement et mise en place d'un test d'évaluation suivent celle d'un profilage où chaque étape conditionne celles qui suivent. Ce cycle est représenté par la figure ci-dessous.

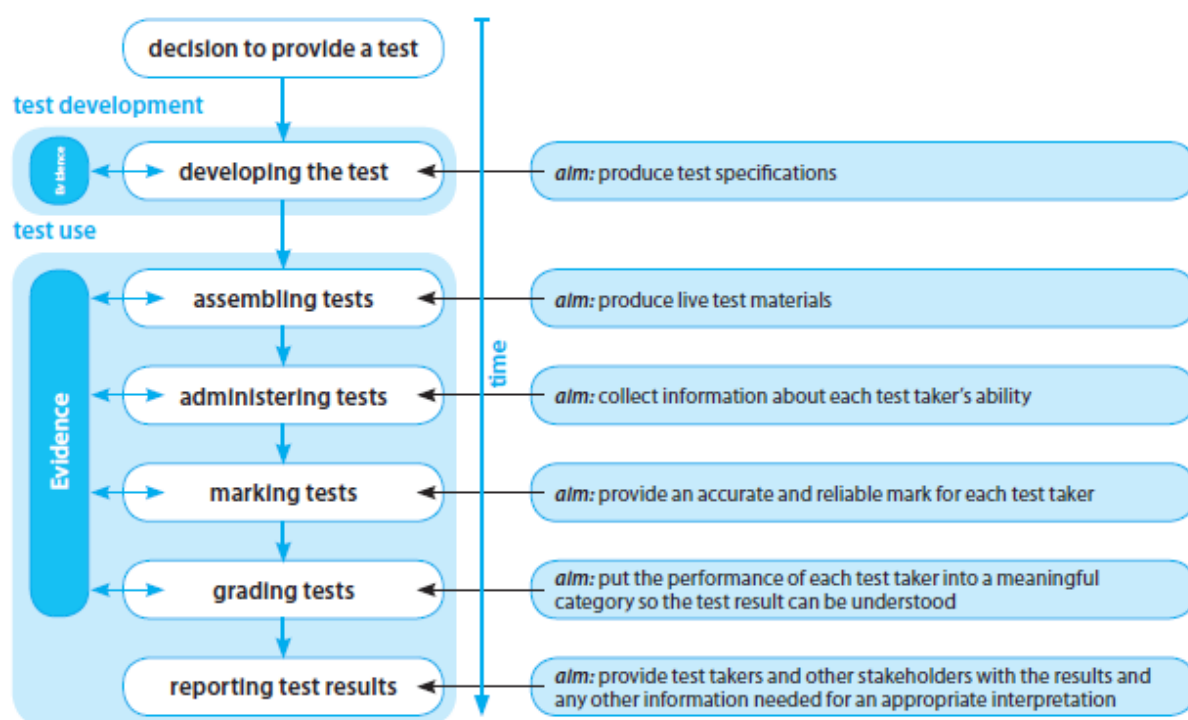


Figure 7 – Cycle normal de création d'un test (ALTE, 2011 : 16)

Quelle que soit la provenance de la décision, les phases d'un projet de création de test seront toujours les mêmes : d'abord en spécifiant sa nature et ses caractéristiques, puis en procédant aux phases liées à son utilisation. Les sous-tâches (micro-tâches) liées à chaque phase figurent

en bleu à droite de la figure et en face de leur phase respective. Ici aussi, un cycle chronologique est symbolisé par les flèches verticales reliant chaque phase à la suivante. À chaque administration du test correspondront des résultats pouvant être comparés.

Le développement du test fait suite à la décision de sa création. Sa finalité est de fournir les *specifications*, c'est-à-dire de spécifier les caractéristiques de l'examen telles que ce qui est testé et la procédure utilisée (ALTE, 2011 : 85). La figure ci-dessous en décompose sommairement les sous-phases. Milanovic *et al.* (2002 : 4) présente un schéma détaillant davantage et mettant en saillance le caractère cyclique du processus d'élaboration des tests. Milanovic *et al.* (2002 : 4) stipulent que l'objectif sous-jacent d'un test, à considérer de pair avec ses spécifications, est sa validité, sa fidélité, son impact et sa faisabilité.



Figure 8 – Processus de développement d'un test (ALTE, 2011 : 20)

La phase *planning* concerne la collecte d'informations préparatoires d'après les commanditaires, institutions, codes, experts, parties prenantes, etc. La phase *design* cible la détermination de la nature du test et de ses principales caractéristiques, de la structure générale à ses contenus. L'étape de *try-out* est destinée à faire un premier essai sur « cobayes » (sujets d'expérimentation), de récolter l'avis de collègues, de solliciter l'avis critique des parties prenantes. La dernière phase du développement consistera à fournir les éléments d'information sur le test correspondant suivant le public concerné (version simplifiée ou complète, détails techniques, etc.), ainsi qu'un échantillon du test.

Les étapes suivantes concernent davantage la constitution du test lui-même, ce qui éloigne de l'objectif de la présente recherche.

5 Problématique : Cadrer la variabilité et repérer des tendances

Les sections précédentes ont surtout montré l'extrême variabilité de la prosodie et le caractère rigide de l'évaluation. Dans les évaluations de type traditionnel et classique, les enseignants repèrent globalement une impression de maîtrise prosodique, à laquelle ils

attribuent une note ou un score. Les grilles évaluatives dépendront de la nature de l'enseignement dispensé, du cursus suivi, ou des critères à prendre en compte par l'évaluateur parmi tous ceux provenant de descripteurs importants. On a vu qu'il était difficile voire impossible pour un enseignant d'évaluer objectivement en se focalisant sur un grand nombre de critères à la fois. Par ailleurs, une trop grande variabilité est incompatible avec une évaluation automatique, qui ne peut être que normative. On comprend alors mieux les réticences que suscite une telle gageure.

Pour répondre à la demande de critères évaluatifs fiables de l'interlangue prosodique susceptibles d'aboutir à une possibilité d'évaluation automatique des réalisations des apprenants anglicistes, il est nécessaire de cadrer la variabilité et repérer des tendances chez les locuteurs natifs et non-natifs. La question est de trouver comment procéder.

La réponse à cette question implique de procéder à un état de la recherche préalable dans le domaine de la prosodie anglaise et de ce qui la caractérise chez les apprenants francophones dans une optique évaluative. C'est ce que présente le Chapitre I. À partir de cette étude, une méthodologie d'élaboration de critères prosodiques devrait pouvoir être dégagée afin de correspondre aux besoins de cette recherche. C'est elle qui structurera cette thèse dans la poursuite de ses objectifs, tant au niveau de ses choix que de l'utilisation de ses résultats à chaque étape de son parcours.

CHAPITRE I – Prosodie anglaise et évaluation prosodique d'apprenants

1 Résumé et introduction du Chapitre I

Ce premier chapitre a pour fonction de présenter l'état de l'art de la recherche en prosodie anglaise en liaison avec celle de l'évaluation prosodique des apprenants francophones. Son objectif fondamental est de dresser un bilan de la recherche effectuée dans ces domaines afin de dégager une méthodologie adéquate pour élaborer les critères prosodiques visés chez les natifs francophones.

Pour cela, ce Chapitre I introduit les grands principes traditionnels d'organisation prosodique par domaines (temporel, mélodique et rythmique) pour pouvoir situer la prosodie des anglicistes francophones et ensuite mettre en perspective les contraintes du cadre évaluatif. Les conditions expérimentales dans lesquelles se situe cette recherche sont alors énoncées.

À l'aune de ces avancées, possibilités et contraintes, il est alors possible de pleinement apprécier la complexité de la tâche à entreprendre, de dégager des pistes de travail propres à répondre à ces difficultés, et de baliser une progression vers l'objectif assigné sous forme d'hypothèses ou de questions de recherche, lesquelles structurent cette étude en chapitres. Chacun d'entre eux débouche sur des expérimentations dont les résultats conditionnent la progression vers la phase suivante de cette recherche.

L'Introduction a spécifié quelques unes des difficultés les plus marquantes auxquelles l'analyse de la prosodie et son évaluation seront confrontées dans cette étude dans une recherche de régularités pouvant déboucher sur la création de traits distinctifs évaluatifs. Afin de commencer cette étude, ce premier chapitre se focalise tout d'abord sur les critères d'organisation prosodique fondamentaux que sont les dimensions temporelle, mélodique et rythmique, et introduit en son sein la prosodie des apprenants anglicistes francophones. Les principes fondamentaux régissant l'organisation de l'évaluation prosodique des apprenants seront alors énoncés. Dans un troisième temps seront présentés les conditions d'étude et le corpus de travail. L'implication de ces avancées et cadrage des conditions de recherche

orientent cette étude vers les moyens et méthodes les plus appropriés à mettre en œuvre pour remédier à la difficulté de modaliser l'interlangue prosodique. Ce chapitre conduit à une présentation de la démarche suivie dans cette thèse et à son organisation générale en chapitres, chacun aboutissant à une phase expérimentale.

2 Organisation prosodique

Nous avons vu dans l'Introduction que la définition de la prosodie formelle selon Di Cristo (2013 : 83) ciblait les entités abstraites que sont le tempo, l'accentuation et la mélodie. Chaque domaine organise son fonctionnement selon certains critères, qui sont présentés ci-dessous.

2.1 L'organisation temporelle de la prosodie

Selon Di Cristo (2013 : 13), la dimension temporelle de la prosodie regroupe trois classes de phénomènes articulés autour de la durée des unités linguistiques, des pauses et du tempo.

En ce qui concerne la prosodie des apprenants, Gut mentionne les conceptions psycholinguistiques et didactiques pour aborder la notion d'aisance (*fluency*) en production orale (2009 : 79). Ainsi, les études sur les corrélats temporels de l'aisance en production orale se fondent sur l'hypothèse que l'aisance constitue un élément isolable de la maîtrise langagière indiquant la nature et l'efficacité des processus en jeu dans la production orale et pouvant être quantifié en mesurant la vitesse d'élocution, la durée des tours de parole, et les marqueurs de disfluence tels que les répétitions, reprises et pauses, entre autres mesures. Selon une conception récente, Gut (2009 : 80-81) mentionne une répartition en trois groupes que sont les caractéristiques temporelles, les marqueurs de disfluence et les traits d'interactivité. Les deux derniers groupes concernent essentiellement les critères d'étude de la conversation. Ainsi, les mesures temporelles destinées à apprécier quantitativement l'aisance sont les pauses, le débit de parole et la proportion de pauses dans le discours.

2.1.1 La durée des unités linguistiques

Alors que l'étude du phonème relève du domaine segmental, le domaine prosodique peut prendre en compte une grande diversité d'unités lorsqu'il s'agit de mesurer la variation. Di Cristo (2005 : 152) compile dans une figure synthétique la variation terminologique des diverses unités privilégiées par des chercheurs en fonction de la nature prosodique.

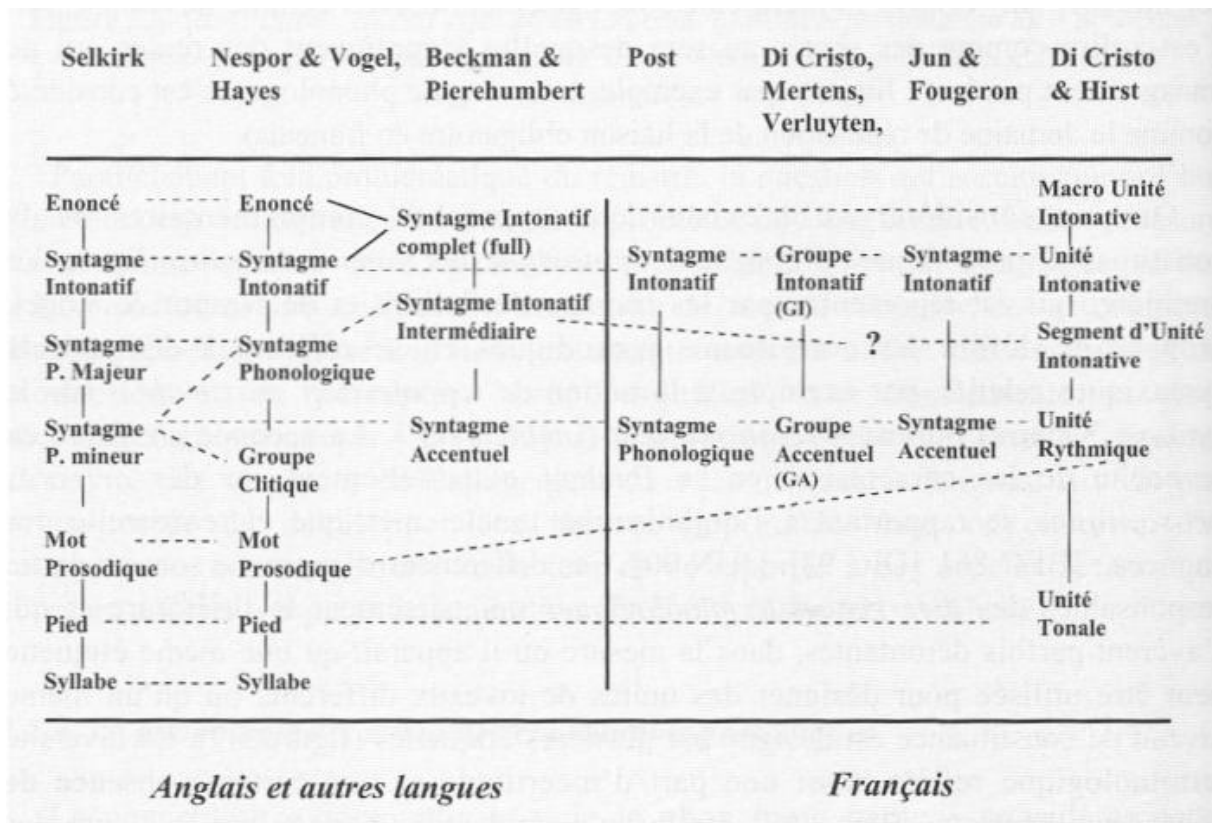


Figure 9 – Unités prosodiques selon les chercheurs de l'anglais et du français (Di Cristo, 2004 : 161)

Dans cette présentation récapitulative, on observe une hiérarchisation des unités prosodiques bien plus fine et élaborée en anglais qu'en français. Alors que l'unité la plus large est répertoriée sous les termes d'« énoncé » (Selkirk, Nespors et Vogel, Hayes), de « syntagme intonatif complet » en anglais (Beckman et Pierrehumbert), et de « macro unité intonative » en français (Di Cristo et Hirst), l'unité la plus réduite est la syllabe, mais uniquement chez Selkirk, Nespors et Vogel, et Hayes. L'analyse prosodique de la syllabe n'obtient pas les faveurs des prosodistes du français. Par contre, parmi les petites entités, le pied est analysé par les prosodistes de la syllabe anglaise ainsi que par Di Cristo et Hirst pour le français. Le syntagme intonatif, dont l'appellation peut varier suivant les chercheurs prosodiques, est l'unité reconnue par tous.

Gut récapitule les résultats de l'étude de Grosjean et Deschamps (1975) sur les corrélats temporels de l'aisance en conversation libre des natifs anglophones (2009 : 83). La longueur moyenne d'une entité (*mean length of run*) est de 12,61 syllabes.

Dans son étude, Gut considère les mesures quantitatives des unités suivantes : l'intégralité du discours, la syllabe, la pause silencieuse, la pause remplie, tronçon de parole sans pause, le mot (2009 : 93-94). Elle combine leur nombre ou durée pour obtenir des

résultats. Ainsi, dans sa comparaison entre natifs et non natifs en lecture, Gut signale (2009 : 98) que les non natifs auront tendance à insérer davantage de mots dans leur lecture, sans que cette augmentation soit significative, mais cela indiquerait une propension à prendre de faux départs et à se reprendre.

L'étude menée ici relève essentiellement du français et de l'anglais. Il me semble alors nécessaire d'utiliser des unités de mesure communes à ces deux langues à titre comparatif. Bien que des recherches très intéressantes aient apporté une vue globalisante des phénomènes prosodiques au niveau de l'énoncé ou paragraphe (par exemple, en anglais, Wichmann 2000, Wells 2006, et Di Cristo 2013 en prosodie générale), il me semble préférable de surtout cibler des unités linguistiques de petite et moyenne durée afin de cerner et décrire finement les caractéristiques prosodiques à ces deux niveaux. En conséquence, le choix du syntagme intonatif devrait être privilégié pour l'unité de taille moyenne. Par contre, étant donné le peu de consensus sur les unités de petite taille, cette étude devrait se référer à la fois à la syllabe et au pied pour ancrer ses investigations.

2.1.2 Les pauses

La pause possède un statut assez particulier. Di Cristo (2013 : 14) la définit comme « la manifestation physique (pause objective) ou perceptive (pause subjective) d'une interruption ponctuelle du flux régulier de la parole ». Il précise que sa nature et sa fonctionnalité sont complexes. On peut les catégoriser suivant leur perception : la « pause silencieuse » marquera une absence de parole (un silence), alors que la « pause remplie » sera marquée par « l'allongement exceptionnel d'une syllabe ou par la production d'interjections telles que « euh » ou « hum » » (2013 : 14-15). Cette seconde catégorie de pauses englobe les hésitations, les « faux-départs » et les répétitions. Les pauses silencieuses, quant à elles, sont dites « respiratoires » étant donné qu'elles coïncident souvent avec la nécessité qu'a le locuteur de prendre son souffle pour parler ensuite. Pour cela, elles surviennent généralement entre les unités grammaticales de manière à ne pas désorganiser le message sémantique. Ainsi que le résume Di Cristo (2013 : 15), « Les pauses respiratoires sont donc la conséquence d'un compromis entre l'exécution de contraintes de bas niveau (ou physiologiques) et de contraintes linguistiques de haut niveau (ou cognitives). » Cette réflexion de Di Cristo semble s'appliquer davantage à certains styles de parole, telle la parole spontanée, plutôt qu'à la lecture. Pourtant, l'existence de pauses et leur nature revêt une importance considérable, qui doit être prise en considération.

Gut rapporte que les chercheurs catégorisent les pauses selon leur durée Gut (2009 : 80-81). Riggenbach (1991) considère ainsi les micro-pauses (< 0,2 secondes), les hésitations (entre 0,3 et 0,4 seconde) et les pauses véritables (entre 0,5 et 3 secondes). Elle ajoute que la durée minimale d'une pause est de 0,1 seconde selon Trofimovich et Baker (2006), mais peut se rapprocher de trois secondes suivant d'autres chercheurs. Gut précise que Campione et Véronis (2002) catégorisent les pauses silencieuses (*unfilled pauses*) en trois sortes : courtes (< 0,2 secondes), moyennes (de 0,2 à 1 seconde) et longues (> 1 seconde), la dernière catégorie ne concernant que la parole spontanée. Les pauses remplies (*filled pauses*), quant à elles, seraient constituées de termes non lexicaux marquant une hésitation ou d'une élongation de certains sons (*drawls*). Ainsi les mesures des corrélats de l'aisance en termes de pauses se basent sur le nombre et la durée de ces pauses, le pourcentage entre la durée totale des pauses et la durée de parole effective (*pause ratio*), le nombre de pauses vides ou remplies par minute ou toutes les 100 syllabes.

Gut présente la proportion des pauses dans le discours sous la dénomination de *phonation/time ratio*, reprenant Towell (2002), Cucchiari *et al.* (2000, 2002) (2009 : 81). On l'obtient en divisant la durée totale de parole par la durée totale de l'enregistrement. Riggenbach (1991), Freed (1995), Freed *et al.* (2004), Segalowitz et Freed (2004), O'Brien *et al.* (2007) se réfèrent au nombre de mots dans le discours pour mesurer la durée de parole, alors que O'Brien *et al.* (2007) ne prennent en considération que le tronçon sans pause le plus long. Le nombre de syllabes par seconde (*mean speech rate*) des natifs anglophones en lecture (Gut 2009 : 83) est de 4,5 à 5,9 selon les travaux de Dauer (1983) et de 4,7 chez Fant *et al.* (1991). En conversation libre, Deese (1980) l'évalue entre 5,48 et 11,25 alors que Grosjean et Deschamps (1975) l'estiment à 5,17. Gut rapporte que Deese (1980) retient une moyenne constante de 5,5 syllabes par seconde panachée d'occasionnelles pointes en conversation libre (2009 : 84). Ces passages accélérés correspondent à des propositions dépourvues d'emphase à l'intensité réduite et à l'intonation plate, ou bien à des ponts reliant des sections dotées d'une intonation normale. Selon Fant *et al.* (1991), la proportion de pauses dans le discours est de 31,8% lorsqu'on raconte une histoire (*retelling*). On remarque qu'elle est plus basse dans les entretiens radiophoniques d'anglais spontané (16,85% selon Grosjean et Deschamps (1975)). La durée moyenne des pauses varie selon le style de parole (*speaking style*). On doit noter les différences entre production en lecture (ou *retelling*) et en conversation libre. Le nombre de syllabes par seconde est manifestement plus important en conversation libre qu'en lecture. Elle note remarque en outre que la proportion de pauses silencieuses ne varie pas

suivant le style de parole alors que les pauses remplies sont quasiment inexistantes en lecture. Après avoir répertorié les principales études sur les corrélats temporels de l'interlangue de 1980 à 2007, Gut remarque (2009 : 92) que seuls la vitesse d'élocution, le nombre et la durée des pauses silencieuses, la durée moyenne d'un tronçon compact et la quantité de parole exprimée sont des indicateurs d'aisance validés dans plus d'une recherche.

Dans son expérimentation sur les 161 enregistrements des non-natifs de l'anglais (expérimentation décrite dans le Chapitre II), Gut constate qu'en moyenne, les pauses silencieuses constituent 9,75% du corpus et les pauses pleines 15,1% (2009 : 95). En lecture, les non natifs prononcent quelques pauses remplies (2009 : 98).

2.1.3 Le tempo

Di Cristo (2013 : 16) utilise le terme de « tempo » pour signifier ce que d'aucuns nomment « débit », « vitesse d'élocution » ou même « vitesse d'articulation ». Ces notions ont pour finalité d'« évaluer subjectivement, ou quantifier, la vitesse de déroulement de la parole (estimée concrètement en nombre de syllabes par unité de temps, celle-ci étant habituellement la seconde) ». Il souligne que le tempo est souvent assimilé à un trait idiolectal caractérisant un sociolecte, une langue, un dialecte, ou un parler régional. Elles assurent alors une « fonction identificatrice plurielle », mais peuvent aussi être sources d'informations de nature linguistique ou paralinguistique. Les variations du débit de parole sont en outre liées à certains styles de discours, dont le commentaire sportif est un exemple révélateur, ou à des critères plus individuels tels que l'expression des attitudes et des émotions. La parole s'inscrivant dans la dimension temporelle, Hirst prévoit une influence de la vitesse d'élocution sur les autres domaines prosodiques dans la parole d'un locuteur.

De Looze (2010 : 61) restreint l'hypothèse de Di Cristo sur l'importance de cette incidence. À la suite des premières études menées sur l'anglais et le français résumées dans Goldman-Eisler (1968), Lane et Grosjean (1973), Grosjean et Deschamps (1975) et Grosjean (1977), elle affirme que « la variation de la vitesse de parole est le résultat de la variation de la durée et du nombre de pauses plutôt que de la vitesse d'articulation ». Elle cite ensuite les résultats rapportés par Goldman-Eisler (1956) à partir d'un corpus de parole spontanée (entretiens en anglais) : « moyenne de vitesse d'articulation de 4.4 syllabes par seconde [...] pour les locuteurs les plus lents et de 5.9 syll/sec pour les locuteurs les plus rapides, alors que cette variation inter-sujets est cinq fois supérieure pour la durée des pauses ». En effet, De Looze avance que ces travaux (Goldman-Eisler, 1958; Goldman-Eisler, 1961, 1968) prouvent

que le phénomène de durée et de nombre de pauses d'une part, et celui de la vitesse d'articulation d'autre part sont complètement indépendants l'un de l'autre. Elle rapporte la conclusion de Goldman-Eisler (1961) : la vitesse d'articulation est une « constante personnelle d'invariance remarquable » (*a personality constant of remarkable invariance*), en conformité avec d'autres expérimentations postérieures.

Léon (2011 : 146) note que le débit révèle des constantes modelant l'image de la langue. À la suite d'une étude sur le discours spontané académique en français, il remarque que le débit des femmes (5,9 syllabes par seconde) est très légèrement supérieur à celui des hommes (5,7 syllabes par seconde), mais pas de manière significative. En ce qui concerne l'anglais, Rogers (2000 : 286) rapporte le débit de parole moyen de cinq syllabes par seconde selon Laver (1994). Par ailleurs, Léon (2011 : 139) conçoit une variation temporelle de nature phonostylistique : à la « variation émotive incontrôlée généralisée » s'oppose le « style caractéristique d'un tempérament », lequel, à partir d'un « indice », devient parfois un « signal » phonostylistique, tout en pouvant être plus ou moins conscient. Ainsi les phonostyles peuvent provenir de la perception phonostylistique d'un dialecte par exemple, d'une émotion, d'une attitude, ou bien d'un type de discours ((2011 : 147). Cependant, sur le plan de la perception, Léon dresse le bilan de l'étude de Fant, Kruckenberg et Nord et mentionne leur observation : la perception du tempo est concomitante du temps de pause.

Pour analyser le discours des apprenants, le débit de parole a été calculé de manières diverses détaillées par Gut (2009 : 81) :

- Nombre de mots par minute, soit brut, soit après qu'aient été écartées reprises, répétitions, apartés et autres ajouts au discours de base (Lennon 1990),
- Nombre de mots par minute (O'Brien *et al.* 2007),
- Nombre de syllabes par minute (Towell 2002, Trofimovich et Baker 2006),
- Nombre de phones par unité de temps (Cucchiarini *et al.* 2000, 2002),
- Durée moyenne d'un tronçon dépourvu de pauses (Deschamps 1980, Lennon 1990, Freed 1995, Towell 2002, Cucchiarini *et al.* 2000, 2002, Freed, Segalowitz et Dewey 2004, Segalowitz et Freed 2004, O'Brien *et al.* 2007),
- Fréquence d'occurrence et longueur des pauses à l'intérieur d'une *t-unit*, c'est-à-dire une proposition principale suivie de ses propositions subordonnées (Lennon 1990),
- Pourcentage des *t-units* suivis d'une pause (Lennon 1990),

- Pourcentage et durée moyenne des pauses aux frontières des *t-units* (Lennon 1990).

On remarque dans cette liste, l'importance non négligeable que jouent les pauses dans les mesures du débit.

Au regard de ces travaux, je peux déduire que la dimension temporelle de la parole pourrait être interprétée, voire métaphoriquement figurée, par le schéma ci-dessous à trois termes :



Si l'on pouvait comparer ce schéma à une équation¹², on verrait que l'on peut déduire un terme à partir des deux autres, l'un d'eux devenant l'inconnue. Il sera donc inutile d'établir des prises de mesures sur les trois éléments, puisque deux suffiront. Par contre, afin de comprendre et d'identifier la source de la variation temporelle par ses multiples facettes, la nature phonostylistique de l'acte de parole devra aussi être déterminée.

En ce qui concerne les résultats de son expérimentation sur les non natifs germanophones de l'anglais, Gut note que la vitesse d'élocution moyenne est de 2,63 syllabes par seconde avec pauses et 4,19 syllabes par seconde sans tenir compte des pauses. La longueur moyenne d'un tronçon sans pause est de 6,22 syllabes (2009 : 95). La vitesse d'élocution est plus rapide en lecture, de même que la longueur d'un tronçon dépourvu de pauses (2009 : 97). En comparant les natifs et non natifs selon diverses variables en lecture (tableau ci-dessous), Gut note que la vitesse d'élocution (*speechrate*) des non natifs et leur taux d'articulation (*artrate*) sont plus faibles, leurs tronçons dépourvus de pauses (*mlr*) sont plus courts (2009 : 98). Gut précise en outre que le degré d'homogénéité parmi les non natifs est plus important que celui des natifs en ce qui concerne la vitesse d'élocution et la longueur moyenne des tronçons sans pauses (2009 : 99).

¹² Cette représentation reste métaphorique : on ne saurait additionner des pauses (en millisecondes) et des vitesses d'articulation (par exemple en syllabes par seconde).

	silratio	fillratio	speech rate	art rate	mlr	words
non-native English reading passage style	9.81% (6.27)	1.35% (1.67)	3.3 (0.82)	4.4 (0.76)	7.3 (1.78)	273.3
native English reading passage style	9.03% (11.05)	0% --	4.1 (0.76)	5.6 (0.64)	9.44 (2.19)	264
	<i>n.s.</i>	<i>n.s.</i>	*	**	**	<i>n.s.</i>

Tableau 1– Valeurs moyennes et écarts types (entre parenthèses) de cinq variables de l’aisance en lecture chez les natifs et non natifs anglophones (** = significatif avec $p < 0,001$; * = significatif avec $p < 0,01$, * = significatif avec $p < 0,005$) d’après Gut (2009 : 99)

Gut précise (2009 : 102) que le seul point commun entre la production native et non native en termes d’intercorrélations des variables temporelles de l’aisance est la corrélation entre la vitesse d’élocution et le taux d’articulation, ainsi qu’avec la longueur des tronçons dépourvus de pauses. Soucieuse de discriminer qualitativement les différences individuelles en matière d’aisance selon les types de tâches, Gut classe hiérarchiquement les locuteurs non natifs suivant leur taux d’articulation dans les trois styles de parole induits par les tâches de lecture, la reformulation d’une histoire et la conversation libre (2009 : 102). Cette analyse débouche sur la constatation que tous les locuteurs tendent à observer le même rang dans les trois tâches assignées, ainsi que le montre son tableau ci-dessous.

	reading passage	retelling	free
reading passage	--		
Retelling	.75***	--	
Free	.42**	.41**	--

Tableau 2 – Corrélation des classements hiérarchiques de taux d’articulation chez chacun des locuteurs non natifs dans chaque style de parole (** = significatif avec $p < 0,001$, * = significatif avec $p < 0,01$) de Gut (2009 : 103)

La variable du taux d’articulation est donc relativement indépendante des styles de parole.

Les conclusions que Gut tire de son étude sur l’aisance est que le discours des non natifs se caractérise par une grande proportion de pauses, pleines ou silencieuses. La proportion de pauses, la vitesse d’élocution et la longueur moyenne des tronçons non ponctués permettent systématiquement de différencier le discours natif du discours non natif.

Horgues (2010) avait constaté que le débit des apprenants francophones était plus lent que celui des natifs anglophones. Ferragne (2013 : 156) constate la supériorité de durée globale de lecture des apprenants par rapport à celle des natifs pour chacun des quatre textes lus (différence significative vérifiée par test de *Tukey-Kramer*). Il en conclut la robustesse de la variable de vitesse de lecture pour discriminer les apprenants des natifs.

2.2 L'organisation mélodique

Cette sous-section s'organise autour de trois axes : 1) les rapports entre les termes « mélodie » et « intonation », 2) l'adaptation perceptive de la hauteur mélodique aux contraintes physiologiques et acoustiques, 3) les approches de la mélodie anglaise, puis 4) les résultats de recherches sur la mélodie non native.

2.2.1 Mélodie et intonation

Dans son *Dictionary of Linguistics and Phonetics*, Crystal nomme la mélodie « intonation » et la définit ainsi :

intonation (*n.*) A term used in the study of SUPRASEGMENTAL PHONOLOGY, referring to the distinctive use of patterns of PITCH, or melody. The study of intonation is sometimes called **intonology**. [...] (Crystal, 2008: 252)

Il ajoute que, parmi les finalités de l'utilisation de l'intonation, la détermination de la structure grammaticale présente une similarité avec la ponctuation dans l'écriture, bien que s'opérant de manière beaucoup plus contrastée. Outre cette première fonction, la mélodie est aussi utilisée pour traduire un état mental personnel (*personal attitude*) tel que le sarcasme, la perplexité, la colère, etc. Les corrélats acoustiques correspondants sont alors la différence de hauteur mélodique (marquant un contraste), ainsi que d'autres traits prosodiques ou paralinguistiques. D'autres rôles sont dévolus à la mélodie, dont celui de marquer une appartenance à un groupe social (Crystal, 2008: 252). Comme on l'a vu dans l'Introduction, d'autres prosodistes différencient les notions de « mélodie » et d'« intonation », attribuant à cette dernière le sens de « prosodie ». Aussi, afin d'éviter toute ambiguïté, je précise que je rejoins la définition de Crystal et utiliserai essentiellement le terme de « mélodie » dans cette étude pour éviter toute ambiguïté.

2.2.2 Adaptation perceptive de la hauteur mélodique aux contraintes physiologiques et acoustiques

L'organisation mélodique de la parole est basée sur la hauteur de la voix, laquelle dépend essentiellement de la longueur des cordes vocales du locuteur. Ainsi, lors de la tension et la vibration des cordes vocales¹³, les voix d'hommes se situent autour de 120 Hertz en moyenne, celles de femmes vers 240 Hz, celles d'enfants vers 350 Hz et celle de nouveau-nés peuvent atteindre 3 000 Hz (Léon, 2011 : 77). Ces mesures correspondent à la fréquence

¹³ Léon ajoute à ces facteurs la pression sous-glottique (2011 : 77).

fondamentale (F0) de la voix, c'est-à-dire aux mesures de l'harmonique de premier rang. Ces mesures sont le plus souvent effectuées en Hertz (Hz) ou en demi-tons (DT). Cette moyenne correspond au « fondamental usuel », équivalant souvent à la note musicale du rire, des pauses pleines, et des syllabes inaccentuées dans la parole spontanée (Léon, 2011 : 78). Il précise que c'est à partir du « fondamental usuel » de chaque locuteur qu'un auditeur pourra distinguer les mouvements mélodiques entendus et leur ampleur.

L'organisation mélodique d'un locuteur s'inscrit dans un « registre », lui-même inséré dans une « tessiture ». « La tessiture de la voix est l'étendue des notes où l'on parle ou chante sans difficulté. » (Léon, 2011 : 78). En d'autres mots, il s'agit là de l'étendue des possibles, alors que « Le registre est la zone où l'on parle ou chante naturellement. » (*idem*). Le « registre » est donc l'actualisation du possible dans une occurrence déterminée. Soumis aux contraintes biologiques, il est tributaire de la longueur du groupe de souffle. Vaissière (2006 : 108) schématise la diminution de l'ampleur du registre de parole en fonction de la durée du groupe de souffle.

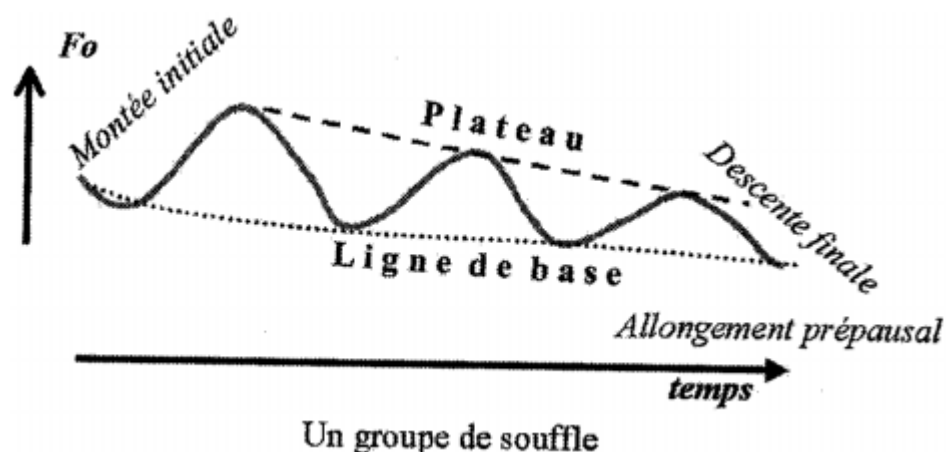


Figure 10 – Contour typique de F0 à l'échelle de l'énoncé, dans diverses langues (Vaissière, [1983] 2016 : 106)

On voit qu'entre deux pauses respiratoires, indépendamment des divers contours mélodiques, la fréquence fondamentale de la parole adopte une tendance à la baisse tandis que les poumons se vident (cause physiologique). Ainsi, le début de l'énoncé sera marqué par une F0 et une intensité maximales, en opposition à la fin de l'énoncé où F0 et intensité seront minimales.

Di Cristo note un ajustement de la perception face à ces contraintes physiologiques (2013 : 94). Il s'agit là d'une normalisation perceptive de l'effet de déclinaison observé

acoustiquement. La figure suivante distingue le schéma acoustique *a* de la normalisation opérée perceptivement représentée en *b*.

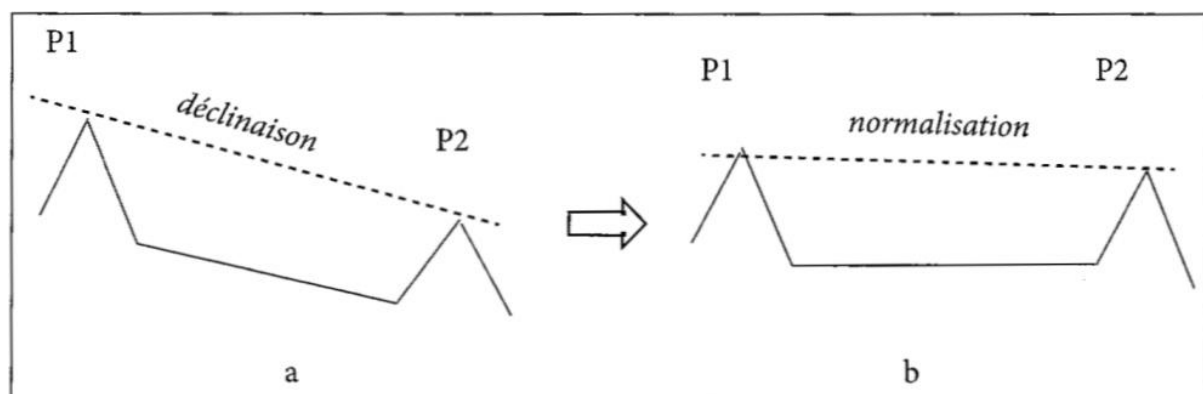


Figure 11 – Illustration de la normalisation de l'effet de déclinaison (P1 et P2 désignent deux proéminences consécutives) (Di Cristo, 2013 : 94)

Ce que Di Cristo nomme « les ruptures de correspondances entre la matérialité acoustique et la substance auditivo-perceptive de la prosodie, dans le cas du paramètre F0 » est aussi évoqué par d'autres ajustements perceptifs par rapport aux données acoustiques. Ainsi la « perception différentielle de F0 » nécessite une différence de F0 suffisante (supérieure au « seuil différentiel de fréquence ») pour que l'oreille humaine soit capable de percevoir une différence de hauteur mélodique entre deux syllabes contiguës (2013 : 94), comme le schématisent les graphiques *a (1)* et *a (2)* ci-dessous (Di Cristo, 2013 : 95). D'autre part, lorsque la courbe mélodique n'est pas statique (grande majorité des cas), la perception de la variation mélodique est tributaire de la durée du mouvement ainsi que de son ampleur : un mouvement mélodique de faible durée (inférieure à 200 millisecondes) sera perçu comme statique aux deux tiers de son mouvement de F0 (voir graphique *b (1)* ci-dessous). De même, lorsque la variation est audible, la hauteur mélodique perçue sera réduite par rapport à celle détectée acoustiquement (voir graphique *b (2)* (2013 : 95 et 100)).

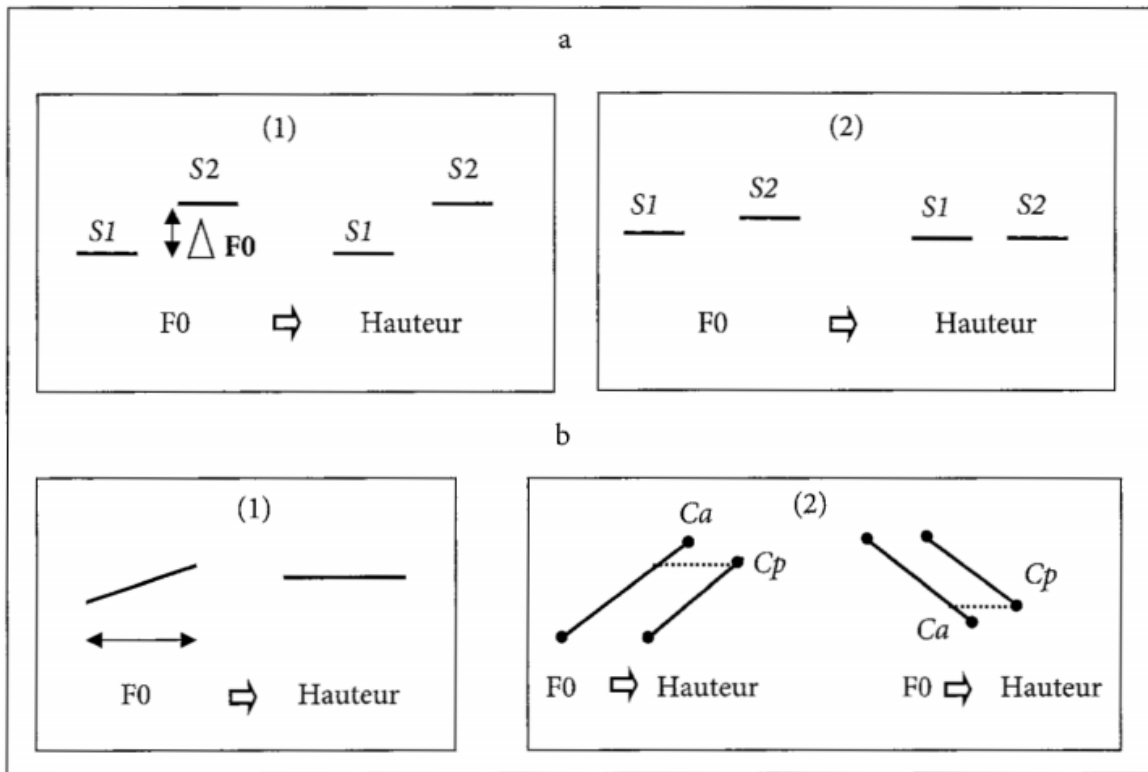


Figure 12 – Illustration de quelques contraintes auditivo-perceptives concernant F0 (Di Cristo, 2013 : 95)

Par ailleurs, les microvariations de la fréquence fondamentale sur certaines consonnes, bien qu'elles soient audibles, sont interprétées comme négligeables par le système perceptif de l'auditeur par rapport à la F0 des voyelles (Di Cristo, 2013 : 91). Ce dernier précise que les consonnes non voisées devant voyelle occasionnent un bombement de la F0, alors que les consonnes voisées provoquent une dépression de la F0 d'importance variable : les consonnes sonorantes telles que [l, m, n et r] sont bien moins affectées que les consonnes obstruantes voisées (occlusives et constrictives).

Ainsi, Di Cristo résume ces adaptations en expliquant qu'il s'agit de « contraintes de bas niveau imposées par l'organe auditif et à des effets de normalisation de niveau plus élevé, qui sont attribuables au système de perception. » (2013 : 91). Afin de déterminer le profil suprasegmental des unités étudiées, la prise de mesures devra donc tenir compte des incidences liées à ces phénomènes en les neutralisant ou en en réduisant la portée.

Les faits rapportés ci-dessus montrent la « nature pluriparamétrique des signaux prosodiques » (Di Cristo, 2013 : 99-100), à la fois physiologique, acoustique et perceptive. Ainsi, chacun de ces paramètres doit être analysé à la lumière des autres pour comprendre et estimer à leur juste valeur les ruptures de correspondance entre elles, et surtout entre les

mesures acoustiques et leur réception auditivo-perceptive. C'est ainsi que pourront être interprétés linguistiquement les différents paramètres prosodiques qui établissent les faits.

2.2.3 Approches de la mélodie anglaise

Définie avec concision comme l'utilisation linguistique de la hauteur mélodique et de ses variations d'une manière systématique et spécifique à chaque langue¹⁴, Gut la précise (2009 : 195) :

The broad definition of intonation includes paralinguistic features such as tempo, voice quality and loudness, which can be employed in order to express the emotional state of a speaker. Less broad definitions of intonation include only linguistic, i.e. meaning distinguishing, phenomena produced with the prosodic features tone, stress and quantity and their physical correlates fundamental frequency, intensity, duration and spectral characteristics.

The narrowest definitions of intonation are restricted to postlexical phonological phenomena, thus excluding word stress, tone and quantity (Ladd 1996, Hirst and Di Cristo 1998). In this conception, which is the one adopted in this chapter, intonation refers to the use of suprasegmental features to convey postlexical meanings in a linguistically structured way. Morphosyntactic and semantic structures as well as the prosodic features stress, accent and tone, which form part of the lexicon, are thus not included in this definition.

Ces définitions posent la question du statut phonologique de l'intonation, de son caractère catégoriel aux unités discrètes.

Gut récapitule les deux modélisations rivales de l'intonation que sont l'École britannique de l'intonation et l'approche métrique autosegmentale. L'École britannique (Palmer 1922, Kingdon 1958, O'Connor et Arnold 1961, Halliday 1967) décrit les contours mélodiques semblablement à une partition de musique, signalant les propriétés accentuelles, de hauteur et de mouvement mélodique de chaque syllabe. Le système se base sur le contour mélodique (*tone*), *a specific pitch movement on a stressed syllable* (2009 : 196). Sa structure est hiérarchisée en unités tonales (*tone units, TU*). Trim (1959) distingue les *major tone units* (*MU*, signalés par ||), correspondant souvent à une phrase, des *minor tone units* (*TU*, représentés par |). Quatre degrés accentuels sont associés à cette approche : absence d'accent

¹⁴ Gut (2009 : 195) : *The term intonation refers to the linguistic use of pitch and pitch movement in a systematic, language-specific way.*

lexical (*unstressed syllables*), syllabes portant l'accent lexical sans correspondre à une variation mélodique (*stressed syllables*), syllabes marquant une prééminence mélodique (*accented syllables*) et syllabes portant le mouvement mélodique principal d'une unité intonative (*nucleus, nuclei*). L'unité intonative de base comprend obligatoirement un noyau (*nucleus*) auquel peuvent s'adjoindre une tête (*head*) ou même une pré-tête (*prehead*) ; les syllabes suivant le noyau forment la queue (*tail*). Gut souligne (2009 : 198) que l'approche britannique se distingue par la conception des cibles tonales, les contours mélodiques (*dynamic pitch contours*) en représentant le fondement.

L'approche métrique autosegmentale (*autosegmental-metrical, AM*) (Pierrehumbert 1980, Liberman and Pierrehumbert 1984, Beckman and Pierrehumbert 1986, Pierrehumbert and Beckman 1988) se distingue par une division de l'intonation en cibles tonales discrètes et indépendantes (2009 : 198). Un énoncé intonatif anglais est caractérisé par les accents de hauteur, ou accents tonals, chromatiques ou mélodiques, un relief tonal donné mélodiquement à une syllabe (*pitch accents, PA*), les accents de phrase se situent à la fin des segments de phrases (*phrase accents*), et les tons de frontières sont les unités mélodiques marquant la fin des énoncés intonatifs (*boundary tones*). Deux niveaux tonals (H pour *high*) et L (pour *low*) marquent les hauteurs mélodiques hautes ou basses les unes par rapport aux autres. Le système permet une structuration hiérarchique et est compatible avec la transcription *ToBI* (*Tone and Break Indices*), adaptable à de nombreuses langues (pour l'anglais américain : Silverman *et al.* 1992, Beckman et Ayers 1994). Le système *ToBI* est notamment compatible avec l'utilisation de logiciels de transcription automatique, même si la transcription est effectuée manuellement sur une ligne d'annotation qui lui est réservée. Gut rapporte que Roach (1994) mentionne une tentative de conversion automatique de l'annotation britannique en *ToBI* (2009 : 201).

Ces définitions spécifiques sur les divers positionnements et le traitement de la matière intonative éclaireront les sections suivantes ainsi que les choix et descriptions effectués au cours de l'étude de cette thèse.

2.2.4 Résultats de recherches sur la mélodie non native

Cette présentation de résultats de travaux de recherche sur la mélodie non native est essentiellement agencée en quatre sous-sections thématiques et synthétiques, qui abordent 1) le découpage en unités intonatives, 2) la qualité tonale et des contours mélodiques, 3) la gamme de fréquence fondamentale (registre), et 4) le positionnement de l'accent mélodique.

2.2.4.1 *Découpage en unités intonatives*

Quelle que soit l'approche envisagée, il est essentiel de pouvoir délimiter l'unité intonative (*intonation phrase*), qui ne correspond pas toujours à une unité syntactique. Gut réaffirme (2009 : 209) la contrainte de la distribution des unités intonatives en lecture de l'anglais. Elle rapporte le consensus des lecteurs sur la localisation des unités tonales majeures (*MU*) dans 100% des cas, et sur celle des unités tonales mineures (*TU*) dans 80% des cas. Les frontières des unités tonales majeures correspondent toujours à celles des propositions syntaxiques, ce qui ne se produit que parfois pour les frontières des unités tonales mineures, plus sujettes à la variabilité personnelle des locuteurs. Yang (2004) corrobore la recherche de Gut (2005) : dans les narrations 73% de toutes les frontières syntaxiques sont suivies d'une pause.

En ce qui concerne le découpage des unités intonatives, Gut résume la tendance générale des non natifs d'accroître le nombre des unités et de diminuer le nombre des mots qu'elles contiennent, ce qui en réduit la durée. Ainsi, les frontières ne correspondent plus à celles des locuteurs natifs. Un consensus existe pour en attribuer la cause à une moindre capacité à procéduraliser la lecture en L2.

Suite à ses travaux, Gut affirme que la longueur moyenne de tous les énoncés anglais non natifs est de 3,96 mots. Le pourcentage des énoncés interrompus sur l'ensemble des énoncés est de 15,81%. À partir du tableau de Gut 8.2 (2009 : 229), j'ai réduit les résultats non natifs en fonction du style de parole à l'anglais (voir tableau ci-dessous).

	mean length of phrases (in words)	percentage interrupted phrases
reading passage	5.13	7.19
retellings	3.3 ***	21.25 ***
free speech	3.36 ***	20.26 ***

Tableau 3 – Longueur moyenne des énoncés et pourcentage des énoncés interrompus en anglais non natif selon le style de parole (***=significatif avec $p<0.001$; *=significatif avec $p<0.05$) (d'après Gut 2009 : 229)

On remarque de grandes disparités suivant les styles de parole. La longueur moyenne des énoncés (on peut penser qu'il s'agit des *IPs*) en lecture est de deux mots supérieure à celle des reformulations et des conversations. De même, la proportion des énoncés interrompus est très nettement supérieure dans ces deux derniers cas.

La comparaison entre natifs et non natifs montre que le nombre d'énoncés moyen en lecture des non natifs (58,58) est bien supérieur à celui des natifs (37). Son corollaire est que

ces énoncés sont plus courts par le nombre de mots qui les composent (natifs : 7,62 ; non natifs : 5,13). Les passages de lecture ne contiennent pas d'énoncés interrompus chez les natifs, alors qu'en reformulation leur proportion est plus faible (11,88%) que chez les non natifs (21,25%).

2.2.4.2 Qualité tonale et des contours mélodiques

Positionnement de Vaissière (2006)

Vaissière note que « Le français est souvent décrit comme une langue « montante » » (2006 : 111). D'après cette description, ce qui caractérise l'interlangue des anglicistes francophones devrait être la continuation marquée en fin de groupe de souffle.

Positionnement et résultats de Gut (2009)

En ce qui concerne les tons chez les locuteurs non natifs, Gut (2009 : 217) mentionne la difficulté qu'ont les locuteurs d'anglais nigériens à produire des tons complexes bi-directionnels (*fall-rise* et surtout *rise-fall*) par rapport aux locuteurs d'anglais britannique (Udofot 1997, 2003, Gut 2004, Eka 1985). Elle rapporte aussi l'expérimentation de Grover *et al.* (1987) où la phrase *I saw Anne but not Eric* était lue : *Anne* était lue avec une chute par les natifs anglophones et avec une montée par les natifs francophones. De nombreuses études interphonologiques tendent à prouver que l'intonation en L2 subit l'influence de la L1, à plus forte raison si les tons de la L2 n'existent pas dans la L1. Selon Grosser (1997), les natifs britanniques produisent 20% de tons descendants-montants ; il présente une hiérarchie acquisitionnelle tonale : les tons simples sont acquis avant les tons complexes, et les chutes sont produites avant les montées. En outre, les non natifs ne produisent pas toujours les tons aux endroits adéquats.

Le tableau suivant détaille la fréquence relative de chaque type de ton nucléaire dans les trois styles de parole que sont la lecture, la reformulation d'histoire et la conversation libre en anglais chez les non natifs. On peut voir que seul le ton *rise-fall-rise* est quasiment absent, tous les autres tons se produisent dans tous les styles de parole. Toutefois, c'est en conversation libre que se produisent davantage les tons plats (c'est-à-dire dans environ 80% de l'ensemble des cas), alors qu'en lecture et en reformulation leur réalisation correspond à 63% des cas). Par contre, les chutes sont bien plus fréquentes en lecture qu'en reformulation et en conversation. Quel que soit le style de parole, le *rise-fall* se rencontre deux fois moins souvent que le *fall-rise*.

	fall	rise	high	mid	low	rise-fall	fall-rise	rise-fall-rise	<i>n</i>
reading passage	23.76	14.68	14	26.74	19.31	0.44	1.07	--	2723
retelling	11.03	21.14	26.22	31.54	8.7	0.4	0.97	--	1541
free speech	10.24	8.59	17.29	39.54	23.02	0.34	0.88	0.1	908

Tableau 4 – Fréquence de tons nucléaires (en %) en anglais non natif selon le style de parole (d'après Gut 2009 : 233)

En fin d'énoncé, l'analyse des tons de frontières indique que c'est en lecture que l'on trouve les tons les plus bas (43,55%) alors que les tons hauts se produisent à 29,75% et les tons moyens à 6,7%¹⁵.

L'analyse acoustique révèle l'envergure du mouvement mélodique dans les chutes et les montées selon le style de parole : les chutes sont légèrement plus fortes (4,38 demi-tons) dans la lecture que dans la reformulation (3,47 DT) et la conversation (3,09 DT). Le phénomène est quasiment inversé pour le ton montant avec 3,95 DT en lecture, 4,28 DT en reformulation et 4,15 DT en conversation.

Le tableau suivant de comparaison des tons entre natifs et non natifs est une reproduction partielle de celui de Gut (2009 : 235). Elle remarque que les natifs produisent moins de tons plats que les non natifs. Les chutes sont deux fois plus utilisées par les natifs que les non natifs. Il en est de même pour le ton *fall-rise*, bien que le ton montant soit plus fréquent chez les non natifs. Gut précise aussi que les locuteurs anglophones natifs ne produisent pas de tons *rise-fall-rise* dans les noyaux.

	fall	rise	level	rise-fall	fall-rise	rise-fall-rise	<i>n</i>
non-native English	17.64	15.54	65.4	0.4	1	0.02	5172
native English	36.8	12.8	47.9	0.4	2.1	--	234

Tableau 5 – Comparaison de la fréquence de tons nucléaires (en %) en anglais non natif et natif (d'après Gut 2009 : 233)

Les tons de frontières bas sont plus fréquents chez les natifs. Contrairement aux non natifs, les locuteurs natifs marquent une forte préférence pour le ton montant dans les syllabes précédant le noyau et le ton descendant en position nucléaire.

¹⁵ La forte proportion des tons hauts présentée pour la lecture dans le tableau de la page 234 de Gut (2009) est confirmée par le commentaire qui en est fait page 233. Il semble néanmoins qu'il y ait une erreur dans les pourcentages car le total des occurrences correspond à 80% seulement et non 100%, contrairement aux données correspondant à la reformulation et à la conversation.

Les non natifs réalisent des chutes de moindre envergure que les natifs : les natifs parviennent à un mouvement moyen de 7,81 demi-tons, ce qui représente plus du double de l'envergure de la chute moyenne des non natifs. Les montées nucléaires chez les natifs anglophones sont de moindre envergure que les chutes, mais sont de taille inférieure à celles des non natifs (voir tableau ci-dessous).

	fall	rise
non-native English	3.64	4.129
native English	7.81	3.8
	**	

Tableau 6 – Comparaison de l'envergure du mouvement mélodique nucléaire chez les natifs et non natifs dans la moyenne des chutes et des montées (d'après Gut 2009 : 237)

Positionnement et résultats de Horgues (2010)

Dans ses expérimentations et rapports de résultats, Horgues (2010) ne dissocie pas les phénomènes de registre de ceux de mélodie. Aussi, c'est dans cette section que je rapporterai conjointement les résultats de ces deux domaines.

Horgues réserve une large part de ses expérimentations à la perception et son test **TPer1** montre que les seuls indices prosodiques (rythmiques et mélodiques) qui résistent au filtrage et à la resynthèse sont suffisants pour que les auditeurs détectent les locuteurs francophones en anglais L2. Elle observe une montée mélodique devant frontière. Dans le cadre de la focalisation intonative, Horgues remarque que les francophones éprouvent une difficulté à réaliser une chute mélodique de forte amplitude sur un nombre restreint de syllabes et un noyau trop tardif dans les cas de désaccentuation de mots grammaticaux. La désaccentuation prosodique post nucléaire est moins marquée que chez les natifs : chute mélodique moins ample, présence de fluctuations mélodiques dans la queue de l'IP, irrégularité du décroscendo d'intensité. Dans l'intonation des structures interrogatives, les questions partielles sont associées à un contour descendant dans les deux groupes de lecteurs alors que les questions totales sont systématiquement caractérisées par une montée chez les francophones alors qu'elles sont « panachées » dans le groupe contrôle. Le contour bidirectionnel *fall-rise* pose problème aux francophones dans la mesure où ils éprouvent des difficultés à le reproduire en répétition sur une seule syllabe (son ampleur est très restreinte si elle ne disparaît pas complètement).

Dans l'étude des réalisations acoustiques de son test TPer2, Horgues repère deux influences de la L1 surtout dans le contour mélodique montant sur syllabe devant frontière et le décroscendo d'intensité très restreint après noyau. La modification des stimuli par

resynthèse dans TPer3 confirme le rôle mélodique et accentuel de la prosodie non native dans l'impression d'accent français. Cela corrobore les observations de Vaissière (2006) énoncées plus haut pour la mélodie et plus bas pour le rythme.

Le test TPer4 a repéré l'accent français dans les stimuli en conversation spontanée filtrée. Horgues remarque que les auditeurs anglophones ne saisissent pas mieux le degré d'accent français qu'en lecture. Horgues observe que le ton *fall-rise* n'est utilisé que par les apprenants les plus avancés, et remplacé le cas échéant par un ton simple.

Positionnement et résultats de Herment *et al.* (2014)

L'étude Herment *et al.* (2014) compare des réalisations prosodiques de questions polaires et en *wh-* en anglais étudiées chez des natifs anglophones et des apprenants anglicistes francophones à partir du corpus de lecture *AixOx*. Elle trouve sa justification dans le souhait de prouver l'utilité pédagogique du corpus et de la visualisation des contours et de leur annotation en enseignement des langues étrangères française et anglaise.

Dans le cas des questions totales (*yes/no*), la structure française sans inversion oblige le locuteur à adopter un schéma intonatif montant pour différencier la question de l'assertion. Ainsi, le tableau suivant récapitule les résultats obtenus en termes de pourcentages de réalisation selon le patron intonatif (F pour *fall* et R pour *rise*, FR pour *fall-rise*) dans les divers groupes de locuteurs.

Yes-no question	natives	independent	proficient
Can you tell me...	F 60%	R 90%	R 60%; FR 20%; F 20%
Vous pourriez me dire ...	R 100%	R 70%	R 70%

Tableau 7 – Contours intonatifs des deux questions polaires dans les différents groupes (Herment *et al.* 2014)

On constate que 100% des natifs francophones adoptent un schéma montant et que l'on ne peut différencier les résultats des apprenants anglophones indépendants de ceux des expérimentés. Contrairement aux descriptions trouvées dans la littérature, les anglophones natifs se départagent puisque 60% d'entre eux adoptent un patron descendant. Quant aux apprenants francophones, ils marquent une nette préférence pour le contour montant.

L'analyse de la courbe de F0 sur spectrogramme avec alignement montre que la tête de l'unité intonative chez les apprenants francophones se trouve sur *tell* et non sur *can*, contrairement à la littérature qui stipule que la tête doit se trouver sur le premier auxiliaire. Le

schéma intonatif des natifs francophones est très différent de celui des natifs anglophones (contour global descendant). Les apprenants francophones ne différencient pas correctement les mots outils (*tool words*) des mots lexicaux (*lexical words*), ce qui les induit à prendre des repères accentuels inadaptés pour établir leur contour intonatif. Le contour montant avec accentuation sur le dernier mot ne semble pas naturel et est typique d'une intonation française bien que le contour global ne soit pas faux car c'est au niveau du rythme que le choix des mots à accentuer est effectué.

En ce qui concerne les questions partielles (*wh-*), le contour descendant est généralement adopté par les natifs anglophones, alors qu'en français, les deux schémas montants et descendants sont acceptables, même si le patron descendant l'emporte à 70% (voir tableau ci-dessous).

<i>Wh</i> -question	natives	independent	proficient
What can I have...	F 100%	F 50%	F 60%
Qu'est-ce que je vais ...	F 70%	F 80%	F 100%

Tableau 8 – Contours intonatifs des deux questions partielles dans les différents groupes (Herment *et al.* 2014)

La littérature atteste un patron montant pour la question en *wh-*, contrairement aux réalisations résultant de l'expérimentation. Les auteurs soulignent néanmoins un contexte particulier, car il s'agit d'une question rhétorique à laquelle aucune réponse n'est attendue. L'hypothèse est avancée selon laquelle le contour intonatif serait influencé par la nature syntaxique de la question mais aussi par l'aspect pragmatique de la situation. Du côté des apprenants anglicistes, on ne remarque quasiment aucune différence entre les deux groupes mis à part une légère préférence du groupe le plus avancé pour le patron descendant. Les apprenants francistes font le choix du contour descendant dans leurs réalisations.

L'analyse sur spectrogramme révèle d'autres écarts par rapport à la littérature. Les dix natifs anglophones ont placé la tête du syntagme intonatif sur l'auxiliaire de modalité *can* alors qu'il était attendu sur le pronom interrogatif *What*. Dans les réalisations des anglicistes, on remarque les contours montants observés dans les phrases françaises avec ici une montée sur la dernière syllabe de *dinner* et de *tonight*. Étant donné que la moitié des apprenants anglicistes adoptent un contour montant dans la question partielle et encore davantage dans la question totale, Herment *et al.* se posent la question de savoir si les apprenants francophones sont effectivement capables de différencier l'intonation propre à ces deux sortes de questions.

2.2.4.3 Gamme de fréquence fondamentale (registre)

Positionnement et résultats de Gut (2009)

Gut conçoit la gamme de fréquence ou tessiture (*pitch range*) comme faisant partie de l'intonation et la définit (2009 : 213) comme *the difference between the top and bottom limits of the pitch movement across an utterance*, ce qui en fait une notion relative car elle variera selon chaque locuteur. Elle rapporte deux notions de Ladd (1996) et Patterson (2000), qui marquent une différence entre le niveau moyen de fréquence mélodique d'un énoncé (*overall pitch level*) et la gamme ou étendue de fréquence (*pitch span*) couverte par un locuteur dans un énoncé. Deux manières d'en prendre mesure sont proposées par Patterson (2000) : la large gamme de fréquences (*wide pitch range*) est calculée en soustrayant la moyenne des fréquences minimales de l'énoncé de la moyenne des fréquences maximales de l'énoncé ; la gamme de fréquences étroite (*small pitch range*) correspond à la différence de fréquence entre les sommets non initiaux et les vallées non terminales. Gut mentionne en outre la fonction linguistique de l'étendue de fréquence qui est d'organiser le discours en unités selon une structure hiérarchique. De nombreuses recherches (Lehiste 1975, Brazil 1975, Brown *et al.* 1980, Wichmann 2000) montrent que la hauteur mélodique maximale se trouve souvent en début d'énoncé et que le niveau mélodique minimal marque la fin d'un énoncé, surtout en lecture (Clark 1999). Yule (1980) introduit le terme de *paratone*, équivalent du paragraphe dans le discours parlé. Un autre type de paratone étudié par Wennerstrom (2001) rend compte de la gamme de fréquence très réduite que l'on trouve dans les énoncés parenthétiques tels que les explétives, incises et autre remarque métalinguistique. Peu étudiée jusqu'alors chez les non natifs, la gamme de fréquence fondamentale peut différer des normes natives, parfois en relation avec la L1. La manière de marquer la structure discursive par la hauteur mélodique diffère parfois de celle des natifs.

Gut remarque (2009 : 239) que c'est en lecture que le registre est maximal pour les natifs et non natifs, bien que celui des non natifs lui soit nettement inférieur. Elle précise que, parmi ses 46 non natifs, 18 réalisaient un registre comparable à celui des natifs en reformulation, et 12 en lecture. Parmi ces derniers, cinq avaient un registre supérieur à celui des natifs. Le tableau suivant récapitule les données fournies par Gut.

	max.pitch range reading	max.pitch range retelling	max.pitch range free speech
non-native English	12	10.3	10.75

native English	17	12.7
	**	***

Tableau 9 – Moyenne du registre maximal des natifs et non natifs selon le style de parole (***=significatif avec $p=0.001$; **=significatif avec $p<0.01$) (d'après Gut 2009 : 238, 240)

Les mêmes stratégies prévalent suivant les fonctions discursives chez les natifs et non natifs. L'information connue possède un registre plus étroit que l'information nouvelle ou celle contenue dans le discours direct (voir tableau suivant).

	direct speech	new topic	background information
non-native English	6.87/7.17 ¹⁶	5.79	4.91
native English	12.9	11.78	4.76
	*	*	n.s.

Tableau 10 – Moyenne du registre maximal des natifs et non natifs des énoncés avec discours direct, rhème et thème en lecture (***=significatif avec $p=0.001$; **=significatif avec $p<0.01$) (d'après Gut 2009 : 240)

Gut remarque que le registre est similaire chez les deux types de locuteurs lorsqu'il s'agit d'informations déjà posées.

Le tableau suivant détaille la répartition de hauteur mélodique maximale (selon les catégories haute, moyenne ou basse) sur le sommet en début d'énoncé des natifs et non natifs suivant leur classement en information nouvelle, posée et détails additionnels.

	new topic	elaboration	background information
non-native English	<i>H</i> : 74.2% <i>M</i> : 25.2% <i>L</i> : 0.6%	<i>H</i> : 57.3% <i>M</i> : 38.5% <i>L</i> : 4.2%	<i>H</i> : 3.7% <i>M</i> : 53.7% <i>L</i> : 42.6%
native English	<i>H</i> : 94.4% <i>M</i> : 5.6%	<i>M</i> : 100%	<i>L</i> : 100%

Tableau 11 – Hauteur mélodique (en %) en début d'énoncé (H=high, M=mid, L=low) des natifs et non natifs dans les phrases contenant rhème, détails additionnels et thème en lecture (Gut 2009 : 241)

Gut relève des différences stratégiques : les locuteurs natifs choisissent une fréquence essentiellement haute (dans 94,4% des cas) pour aborder des sujets nouveaux, une fréquence moyenne (100%) pour ajouter des détails et une fréquence basse (100%) pour mentionner une information déjà connue. Davantage de variabilité caractérise les choix des non natifs. Les nouveaux sujets sont principalement abordés avec une fréquence haute (dans 74,2% des cas), mais qui peut aussi être moyenne (25,2%) ; les détails additionnels le seront avec une

¹⁶ Les tableaux des pages 238 et 240 fournissent des données différentes. Aussi, je les ai notées conjointement.

fréquence haute (57,3%) ou moyenne (38,5%) ; les informations posées sont apportées avec une fréquence moyenne (dans 53,7% des cas) ou basse (42,6%).

Positionnement et résultats de Horgues (2010)

En ce qui concerne la mélodie des apprenants francophones de l'anglais, Horgues (2010) considère que mélodie et registre doivent être étudiés de pair. C'est la raison pour laquelle ses résultats figurent dans la rubrique sur la mélodie en amont. Je rappellerai seulement ici qu'elle constate une étendue de registre moindre.

2.2.4.4 Positionnement de l'accent mélodique

Gut note (2009 : 210) que la position du noyau correspond à un mouvement de la hauteur mélodique distinctif et que, dans un énoncé anglais non marqué, il se positionne sur le dernier mot non grammatical de l'énoncé. C'est ce qui se produit dans les phrases simples de base. Elle ajoute qu'en anglais, la localisation du noyau est intimement liée au concept de focus. Elle le définit comme *that part of a sentence which is prominent in terms of content and which forms a contrast to the background information of the sentence* et mentionne les recherches de Féry 1988 et Ladd 1980, qui distinguent le focus large (*broad focus*) portant sur plusieurs constituants et même la phrase entière, du focus étroit (*narrow focus*), qui s'applique à des items lexicaux particuliers. Le focus est essentiellement marqué par une accentuation combinée à des mouvements de hauteur mélodique. Il dépend de la structure informationnelle (*information structure*) introduite par Halliday (1967) selon laquelle les traits discursifs d'information nouvelle (*new*) et d'information ancienne (*given*) créent un différent niveau structurel.

La position du noyau a été vérifiée sur des énoncés de lecture contenant un verbe suivi d'un pronom d'une part, et d'une construction avec nom suivie d'un verbe d'autre part. La désaccentuation de l'information donnée ou accessible est aussi vérifiée. Les résultats en sont consignés dans le tableau suivant. Les locuteurs natifs positionnent le noyau sur le verbe dans la structure « verbe+pronom » dans 100% des cas, alors que les non natifs ne le font qu'à 83,6%. Les choix des deux types de locuteurs sont par contre équivalents dans le positionnement du noyau sur le verbe dans la structure « nom+verbe », et la désaccentuation de l'information donnée ou accessible est vérifiée dans 100% des occurrences chez les natifs, mais seulement à 57,7% chez les non natifs.

verb+pronoun	noun+verb	deaccenting
---------------------	------------------	--------------------

non-native English	83.6	45.3	57.7
native English	100	46.6	100

Tableau 12 – Positionnement du noyau (en %) sur le verbe dans les structures verbe+pronom, et sur le nom dans les structures nom+verbe chez les locuteurs natifs et non natifs ; désaccentuation (en %) des mots donnés ou accessibles (Gut 2009 : 244)

Le mouvement mélodique en position pré-nucléaire a donné les résultats qui figurent dans le tableau suivant. Gut a catégorisé les 226 mots d'un texte (81 mots lexicaux et 145 mots outils) par rapport aux tons dynamiques simples et complexes (*fall*, *rise*, *rise-fall* et *fall-rise*).

	fall	rise	rise-fall	fall-rise	n
content words non-native English	9.3	3.4	0.1	0.02	4617
function words non-native English	1	1.1	--	--	8265
content words native English	4.3	6.7	0.3	--	324
function words native English	1.2	2.4	--	--	580

Tableau 13 – Distribution (en %) des mouvements mélodiques dynamiques simples et complexes en position pré-noyau en lecture suivant les locuteurs natifs et non natifs (Gut 2009 : 244)

Les deux types de locuteurs sélectionnent les mots lexicaux pour leur faire porter des tons mélodiques. Gut remarque qu'alors que les non natifs préfèrent les chutes, les natifs favorisent les montées pour les deux catégories de mots.

La proportion d'accents mélodiques réalisés par les non natifs est surdimensionnée par rapport à la parole native. Pourtant, elle se caractérise par une accentuation des mots à la fois excessive en nombre et insuffisante en qualité.

Dans son étude, Horgues (2010) oriente ses constatations dans le même sens. L'analyse prosodique de quatre conversations d'apprenants a confirmé ses conclusions émises à partir de la parole lue : accentuation des formes grammaticales finales (surtout les pronoms), seuls les apprenants les plus avancés désaccentuent l'information ancienne.

2.3 L'organisation rythmique

Cette sous-section se centrera sur l'organisation rythmique en interrogeant sa nature, puis en abordant deux procédés organisationnels possibles : celui déterminé par les phénomènes de proéminence et d'accentuation, puis celui d'alternance de durées (métriques).

2.3.1 Quiddité du rythme

Ainsi que l'écrit Grammont (1937 : 85), « Le rythme, on le sait, est constitué par le retour des temps marqués à intervalles théoriquement égaux. ». Toute la question est de savoir

quels éléments marquent ce rythme. Selon un consensus, le rythme s'inscrit dans le temps ; mais j'observe que la notion de rythme a deux sens. Le premier équivaut à un rythme naturel. Rythme cardiaque (battement du cœur), rythme des vagues, de la houle (battement, balancement des vagues ou de la houle) n'en sont que quelques exemples, qui ont en commun de correspondre à un bruit causé répétitivement par une action à intervalles perçus comme réguliers. Le mouvement qui l'a engendré est celui d'un aller-retour mécanique, lequel crée la perception du rythme par le bruit répétitif qu'elle cause. Ces mouvements « ne communiquent pas entre eux », ils sont « juxtaposés », ils sont « mis bout à bout »¹⁷. C'est une sorte d'alternance de temps forts et de temps faibles dénués de sens.

À l'inverse, le second rythme est produit artificiellement. Il s'inscrit dans le temps afin d'établir des repères réguliers déclencheurs d'action. La danse en est un excellent exemple : la perception de temps forts et faibles se succède et guide un mouvement continu et cohérent. Le rythme n'est donc plus signe, mais signal déclencheur des pas de danse. Temps forts et temps faibles n'ont de cohérence que les uns par rapport aux autres, « ils n'existent que comme des différences dans le continu du mouvement » et « chaque moment contient le précédent et le suivant »¹⁸. L'alternance et la rapidité des signaux oriente et guide le mouvement de « rencontre, enlacement et séparation », lequel peut alors prendre forme et sens. La beauté et l'émotion créées (ou leur absence) sont parties intégrantes du message véhiculé par cette forme fugitive et éphémère. De même que pour la parole, il ne reste rien de la danse lorsqu'elle prend fin, sinon l'impact produit sur nos sens et notre entendement. Aussi, on peut affirmer que le rythme est l'essence même de la parole car il en soutient l'agencement, quels que soient les procédés utilisés pour remplir sa mission.

D'emblée, Gut associe le rythme en L2 à la réduction vocalique. Elle le présente comme *one of the major organising principles of speech* (2009: 159) et rappelle la nature structurante du rythme, prégnante dans toutes les conceptions qui lui sont associées. Elle récapitule les deux manières fondamentales de concevoir le rythme qui sont la hiérarchie prosodique et la dichotomie des langues en classes rythmiques. La centration sur la localisation du rythme dans la Hiérarchie Prosodique de Selkirk 1986, Beckman et Pierrehumbert 1986, Nespor et Vogel 1986, etc., dont les unités de base sont les syllabes (*Syllable*), sont regroupées en unités supérieures : pieds (*Foot*), puis mots phonologiques

¹⁷ Ancet, Jacques, Ancet, Lucienne. *La Caminata*. Du rythme : de l'importance du rythme dans la danse d'improvisation qu'est le tango [<https://sites.google.com/site/annecytango/chroniques/du-rythme>], consulté en janvier 2017.

¹⁸ Idem, [<https://sites.google.com/site/annecytango/chroniques/du-rythme>], consulté en janvier 2017.

(*Phonological Word*), groupes clitiques (*Clitic Group*), phrases intermédiaires (*Intermediate Phrase*) ou groupes phonologiques (*Phonological Phrase*), syntagmes intonatifs (*Intonation Phrase*) et syntagmes phonologiques (*Phonological Utterance*). Destinés à être des repères de production de la parole, ils constituent en outre des domaines de restriction dans l'application de diverses règles phonologiques (Selkirk 1986).

Gut rappelle (2009 : 160) la division rythmique en classes de langues de Pike (1945) puis d'Abercrombie (1967) selon laquelle chaque langue serait régie par la prédominance soit de la syllabe (les syllabes seraient isochrones, par exemple le français), soit du battement rythmique (les pieds, intervalles entre deux battements rythmiques, seraient isochrones eux aussi, comme en anglais). Les nombreuses études acoustiques qui ont suivi n'ont pu corroborer l'ajustement temporel de la durée des syllabes dans l'un ou l'autre système sur lequel se fonde le principe de l'isochronie. Aussi, il a été relégué au rang de phénomène perceptif ou même tendanciel.

Tortel (2010 : 30) rappelle elle aussi qu'il existe deux sortes de rythme : la rythmicité accentuelle et la rythmicité syllabique, différence introduite par Pike (1945) suivi de Bolinger (1981). Elle précise que Di Cristo (2002) décrit l'architecture prosodique comme « un supra-système phonique constitué de trois sous-systèmes », lesquels sont métrique, tonal et temporel. Mètre et rythme doivent alors être dissociés. Di Cristo (2013 : 12) spécifie l'utilisation qu'il fait de chacun d'entre eux :

- (i) le terme de métrique pour désigner l'étude formelle de la structure sous-jacente (ou abstraite) du phénomène rythmique (et des représentations abstraites qui s'y rattachent)
- (ii) celle de rythme, pour décrire la manifestation concrète de ce phénomène, au niveau des structures de surface qui actualisent la prononciation des énoncés

L'étude de la structure rythmique sous-jacente (structure métrique, reflet des compétences linguistique et rythmique) ne peut qu'éclairer les observations faites à partir de la substance concrète du phénomène rythmique dans la quête de régularités prosodiques. Aussi, des deux faces du même phénomène (Di Cristo, 2013 : 121), c'est le rythme qui sera l'objet de recherche. Pourtant, pour en revenir à la question de départ, aucune des deux pistes que constituent le phénomène accentuel et l'organisation temporelle des syllabes ne devrait être négligée. Les procédés rythmiques mis en œuvre dans cette « rythmicité binaire » (Tortel, 2010 : 29) peuvent être très divers, mais, qu'ils correspondent à des prééminences basées sur

l'intensité, la hauteur mélodique, les mouvements de fréquence fondamentale, ou à des alternances de durées, ou de parole et de silences, tous convergent vers la même finalité : créer un maillage de temps forts et de temps faibles pour soutenir techniquement l'organisation du message en le structurant.

L'accentuation correspond à une fonction démarquative, alors qu'à d'autres endroits, elle aura une fonction expressive avec un « accent d'insistance [qui] permet une mise en relief d'une unité généralement plus petite que le syntagme » (Léon, 2011 : 157). Cet accent d'insistance aura l'une des trois fonctions suivantes : *oppositive*, *emphatique* ou *différenciative* (2011 : 158).

Le rythme martèle un sens basique et vital dans la mesure où il constitue la trame même sans laquelle le sens ne pourrait matériellement s'inscrire dans la parole. On se souvient tous de la voix monocorde, dépourvue d'un quelconque rythme, des premiers essais sonores synthétiques. Seule l'introduction de pauses permettait le nécessaire découpage en unités de sens. Que l'on rende compte du rythme par l'accentuation ou bien par des mesures de durée, il est essentiel qu'il soit représenté de la manière la plus complète pour être la plus représentative, sans exclusive. Sa nature consiste en une alternance forte-faible. L'accentuation pourra donc être étudiée au niveau phrastique de deux manières : en comptant le nombre de proéminences dans une unité prosodique pour le comparer à celui d'un modèle natif ou non-natif, ou en repérant la localisation accentuelle attendue et en vérifiant sa présence au même endroit dans la parole d'apprenant, par exemple. La qualité rythmique dans l'accentuation lexicale sera vérifiée de la seconde manière. Les comparaisons de durées sont aussi à privilégier pour ne pas réduire le kaléidoscope des possibles rythmiques.

Afin de repérer les faits pertinents sans être assujettie à l'optique d'un système particulier, cette étude, se voulant tout d'abord descriptive, privilégiera les repères les plus « neutres » représentés par les corrélats acoustiques.

2.3.2 Rythme déterminé par phénomènes de proéminence et d'accentuation

Rythme et proéminences sont au cœur de l'analyse prosodique, ainsi que le rappellent Avanzi, Lacheret et Victorri (2008)¹⁹ :

¹⁹ Avanzi, M., Lacheret-Dujour, A., Victorri, B. (2008). « ANALOR : Un outil d'aide pour la modélisation de l'interface prosodie - grammaire », *Cahiers du CERLICO*, 21, 27-46.

Nous entendons par proéminence mise en valeur perceptive d'une syllabe par rapport à son environnement, et qui se manifeste par la perception d'un « accent » [Lacheret & Beaugendre 1999 : *glossaire*]. La détection des proéminences constitue la pierre angulaire de toute étude prosodique, quelle qu'elle soit [Avanzi *et al.* 2007].

Il s'agit là d'une rythmicité basée sur un procédé accentuel. Cela correspond aussi à la posture de Léon : « *Un rythme est instauré quand il y a perception d'une proéminence accentuelle répétée* » (2011 : 162). Di Cristo rappelle que, bien que les corrélats acoustiques étudiés pour en rendre compte soient la F0, la durée et l'intensité, l'explicitation du phénomène rythmique accentuel réside dans la « construction subjective de la perception » (2013 : 120). Il explique que « l'accent est vu principalement comme une entité relationnelle, le produit d'un réseau de relations hiérarchiques fondé sur la syllabe » (2013 : 123). En effet, selon lui, et appuyant les vues du linguiste américain Hockett, la hiérarchie accentuelle fonde un « mode d'organisation phonologique ». Les divers niveaux relatifs composant la structure hiérarchique accentuelle de la chaîne sonore sont souvent représentés, dans l'approche de la phonologie métrique et sur le plan formel, par des « arborescences à branchements binaires » se ramifiant jusqu'à la syllabe « dans lesquelles tous les nœuds (à l'exception de celui de la racine) sont étiquetés au moyen des symboles *s* (*strong*) ou *w* (*weak*) ». Se référant à l'accent démarcatif, Fónagy relie forme et sens, et ainsi justifie sémantiquement le caractère fondamental du rythme dans la parole :

L'accent articule et organise la parole. Il divise la chaîne parlée continue en séquences, en « groupes rythmiques ». L'accent établit d'autre part une certaine *hiérarchie* sémantique dans la phrase, en prêtant plus ou moins d'intensité aux mots conformément à leur poids sémantique et à l'importance actuelle des mots dans le message concret. (Fónagy, [1983] 1991 : 107)

Vaissière différencie les prosodies typiques du français et de l'anglais (2006 : 111-112) :

Les voyelles en position finale des groupes de sens dominant perceptivement en français. C'est le retour des syllabes allongées, avec des intonations montantes qui définissent essentiellement le rythme en français. En anglais, ce qui frappe un Français, c'est la récurrence énergique et quasi régulière de syllabes fortement accentuées, avec une forte attaque consonantique, qui alternent avec des syllabes réduites. Ce type d'accentuation évoque pour l'oreille française l'*accent d'insistance*, d'où l'impression d'une insistance permanente que peut donner la

langue anglaise à une oreille non entraînée. L'unité principale de rythme en anglais est le « stress group ».

D'après cette description, l'interlangue des anglicistes francophones se caractérise par la perception d'un poids plus égal des syllabes les unes par rapport aux autres.

Sur le plan acoustique, il est complexe de déterminer l'incidence de chacun des trois corrélats dans le phénomène d'accentuation, qu'il s'agisse de l'accent lexical ou de phrase. Ainsi que le souligne Vaissière, « L'accentuation lexicale est une notion abstraite. Elle est une caractéristique intrinsèque des mots et des morphèmes, stockée dans le lexique mental » (2006 : 101). Au niveau du syntagme intonatif, l'accent de phrase relève avant tout d'un phénomène perceptif, pouvant être expliqué par une convergence des corrélats acoustiques chez les anglophones natifs. Rien ne peut nous garantir que les non natifs observeront les mêmes critères de production que les natifs. Alors que c'est la fréquence fondamentale (Herment, 2001), ou dans certains cas l'intensité (Herment)²⁰, qui revêt une importance essentielle pour marquer l'accentuation en anglais, c'est le phénomène de durée qui est fondamental en français, et parfois l'intensité, comme en français québécois (Léon, 2011 : 170). Aussi, il sera nécessaire de différencier la méthodologie de repérage accentuel de celle d'explication du phénomène. La première se doit d'être perceptive, alors que la seconde peut être acoustique.

On a déjà noté que les expérimentations perceptives menées par Horgues (2010) montrent que les indices rythmiques et prosodiques résistant au filtrage et à la resynthèse étaient suffisant pour signaler aux anglophones natifs l'accent étranger des locuteurs francophones parlant anglais. Cela octroie au rythme une importance particulière. Dans l'analyse acoustique des pieds accentuels, elle constate que la variabilité des pieds accentuels est plus importante chez les apprenants francophones que chez les natifs anglophones. Elle remarque en outre que l'accentuation lexicale est difficile à déterminer perceptivement sans contexte chez les apprenants les moins avancés. Son test **TPer2** souligne la difficulté des auditeurs natifs pour situer l'accentuation lexicale des stimuli sans contexte ; ils notent 2,5 fois plus de déplacements accentuels chez l'apprenant que chez le natif. Horgues signale la réalisation incorrecte du patron accentuel dans les contextes de frontière droite de syntagmes intonatifs (*IP*) non finaux et de contour montant dans les questions.

²⁰ Herment, S. (article non publié). "Is intensity a relevant criterion in the perception of spontaneous speech? The case of emphasis in English".

2.3.3 Rythme par phénomènes d’alternance de durées

Léon (2011 : 169) rapporte l’observation par Konopczynski de l’établissement d’un protolangage chez le jeune enfant apprenant à parler en situation d’interaction. Le début de structuration de la parole se traduit par une organisation syllabique tendant à leur octroyer une durée similaire. L’acquisition des phénomènes rythmiques et intonatifs lui permettra de se conformer graduellement aux règles en vigueur dans sa langue maternelle avec constitution des groupes rythmiques normés. Elle avance l’explication selon laquelle il s’agirait d’une étape normale dans une phase d’apprentissage, qu’il s’agisse de parole ou de gestuelle :

Avec l’acquisition du système de sa langue maternelle, l’enfant est soumis à de nouvelles contraintes rythmiques. Il passe d’une rythmicité biologiquement motivée à une structuration temporelle et accentuelle, stylisée selon des normes conventionnelles. (Léon, 2011 : 169)

La diction rythmique de l’enfant pourra alors se soumettre à la thèse de Pike (1946) selon laquelle les langues se subdivisent en deux catégories distinctes : langues isosyllabiques (à rythmicité syllabique, *syllable timing*), et langues à isochronie accentuelle (*stress timing*). Le français appartient à la première catégorie et l’anglais à la seconde. Le « principe d’isochronie » se base sur le sentiment que les durées des syllabes inaccentuées entre deux prééminences se réduisent ou s’allongent pour que leur total s’équivalle. À la suite de nombreux débats sur le sujet, il semblerait que le fondement de la thèse de Pike se vérifie non pas acoustiquement, car il s’agit d’une tendance, mais surtout perceptivement (Vogel, 2016)²¹.

Gut (2009) note elle aussi que des chercheurs se dispensent carrément de la notion de rythme accentuel pour faire valoir les propriétés structurelles des langues (Dauer 1983, Barry 2007) reposant fondamentalement sur la structure syllabique. Elle ajoute que la réduction vocalique de certaines syllabes intervient essentiellement dans les langues à battement rythmique. De plus, l’accent lexical des langues dites syllabiques, lorsqu’il existe, est marqué par des variations mélodiques alors que l’accent lexical des langues à battement rythmique est fortement imprimé par une combinaison acoustique alliant la durée, la hauteur mélodique, l’intensité et les changements de qualité de la voix. Gut aborde les approches actuelles selon lequel *rhythm is a multidimensional concept that includes various phonological properties of languages* (2009 : 162). L’approche qui prévaut sur la dichotomie en classes de langues est

²¹ Irène Vogel : Chaire internationale du Laboratoire d’excellence « Fondements Empiriques de la Linguistique » *Labex EFL* à Paris (juin et juillet 2016)

celle de leur ancrage le long d'un continuum qui marquerait leur degré rythmique les unes par rapport aux autres. La réduction vocalique est à présent la source des mesures rythmiques, que ce soit au niveau de la syllabe qui l'inclut ou du segment lui-même. Les mesures peuvent soit être globales, soit se focaliser sur des comparaisons de durée très ciblées. Selon les mesures globales des traits phonétiques de Ramus *et al.* (1999), la chaîne sonore est divisée en segments qui sont regroupés selon leur nature vocalique ou consonantique. Les données qui en découlent sont la proportion des intervalles vocaliques d'un énoncé (%V), l'écart type de ces intervalles (ΔV) et l'écart type des intervalles consonantiques (ΔC). D'autres mesures du rythme se basent sur des comparaisons de durées sur des éléments très ciblés. Gut mentionne la différence de durée entre deux voyelles successives du *Pairwise Variability Index (PVI)* de Low et Grabe (1995), ainsi que son corollaire *PVI* pour les consonnes (Grabe et Low 2002). Le *Rhythm Ratio (RR)* de Gibbon et Gut (2001), dans chaque paire d'unités adjacentes (soit de voyelles, soit de syllabes), la plus courte est divisée par la plus longue et la moyenne des résultats est multipliée par 100. Le *Variability Index (VI)* de Deterding (2001) mesurent les différences de durée moyenne entre des syllabes qui se suivent à l'exclusion de celles en fin de mots. Gut (2009 : 165-167) remarque une très grande variabilité suivant les diverses variétés d'une même langue : les mesures de l'anglais de Singapour, de Hong Kong, de Taiwan, du Nigéria classent ces variétés parmi les langues traditionnellement rythmiquement classées sur la syllabe et non sur le battement syllabique comme l'anglais britannique. Toutes ces variétés se positionnent individuellement sur le continuum rythmique. Les résultats empiriques des recherches antérieures répertoriées par Gut pour l'anglais figurent dans le tableau ci-dessous. Parmi les résultats relevés, on peut vérifier que les syllabes accentuées sont une fois et demie plus longues que les syllabes inaccentuées. La variation selon la vitesse d'élocution semble devoir être prise en compte dans l'appréciation des résultats : un accroissement de la vitesse d'élocution tend à restreindre les différences de durée entre les syllabes adjacentes en anglais de Singapour et en anglais britannique (Deterding 2001). D'après les études mentionnées par Gut (Dellwo et Wagner 2003, White et Mattys 2007), %V semble fournir des données plus stables indépendamment de la vitesse d'élocution. Le style de parole peut aussi avoir une incidence dans les principes organisationnels du rythme : Gut rapporte que, selon Guaitella 1999, la lecture permet une analyse acoustique quantitative, alors que l'analyse de la conversation est plus propice à un traitement perceptif.

British English	
%V	41.1% (Grabe and Low 2002)

	40.1% (Ramus et al. 1999)
delta C	51.7 (Grabe and Low 2002) 53.5 (Ramus et al. 1999)
PVI	vocalic: 57.2 (Grabe and Low 2002) consonantal: 64.1 (Grabe and Low 2002)
mean length of stressed syllable	162-239 ms (Williams and Hiller 1994) 300 ms (Fant et al. 1991) 294 ms (Campbell 1989) 195-221 ms (Hoequist 1983)
mean length of unstressed syllable	164-183 ms (Williams and Hiller 1994) 140ms (Fant et al. 1991) 138 ms (Campbell 1989) 138-158 (Hoequist 1983)
mean ratio stressed/unstressed syllable	1.45:1 (Hoequist 1983) 1.5:1 (Dauer 1983) 1.7:1 (Delattre 1966)

Tableau 14 – Données rythmiques et durée moyenne des syllabes accentuées et inaccentuées en anglais britannique d’après Gut (2009 : 170)

Gut liste aussi les études faites sur le rythme en anglais des non natifs (2009 : 171-172). On peut retenir celle d’Adams (1979) selon laquelle l’accent étranger en anglais est notamment perçu par insuffisance de durée entre les syllabes accentuées et inaccentuées, absence des mécanismes de liaison natifs, pauses erratiques et déplacement accentuel. Wenk (1985) remarque qu’en lecture de l’anglais après écoute, sur 43 non natifs francophones, la réduction vocalique était effective dans 40% des cas de syllabes non accentuées avant accent tonique et 81% des syllabes inaccentuées après accent tonique. Gut remarque (2009 : 174) l’utilisation d’un groupe de contrôle de natifs dans le profilage évaluatif et explique les différences rythmiques par la différence structurelle des L1 et L2.

Les données rythmiques qu’obtient Gut à la suite de sa recherche figurent dans son tableau 7.4, que j’ai reproduit ci-dessous. On peut voir que les syllabes (*mean length sfv*) des non natifs sont toujours de durée bien supérieure à celle des natifs. Leurs syllabes aux voyelles réduites (*mean length srv*) ou effacées (*mean length sdv*) sont moins nombreuses et la différence de durée moins importante entre les syllabes adjacentes avec voyelles pleines (*ratio full/full*) qu’avec voyelles réduites (*ratio full/red*).

mean length sfv	mean length srv	mean length sdv	perc. red/del	ratio full/red	ratio full/full	n
------------------------	------------------------	------------------------	----------------------	-----------------------	------------------------	----------

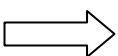
non-native English	236.1 (44.38)	155.07 (41.02)	157.07 (63.43)	24.01% (6.9)	1.98:1 (0.4)	1.22:1 (0.14)	41670
native English	210.75 (19.9)	101.8 (13.4)	85 (39.04)	30.65% (5.74)	2.45:1 (0.33)	1.18:1 (0.07)	2492
	***	***	***	**	**	n.s.	

Tableau 15 – Données rythmiques et leur écart type chez les natifs anglophones et non natifs de Gut (2009 : 179) (** = significatif avec $p < 0.01$, *** = significatif avec $p < 0.001$)

Gut note en outre que seules les syllabes adjacentes avec voyelles pleines ont des durées correspondantes chez les natifs et les non natifs. La variabilité des mesures est plus importante chez les non natifs. Les différences de style de parole montrent que les natifs anglophones ont tendance à observer les mêmes proportions de durée des paires de syllabes entre celles avec voyelles pleines suivies de celles avec voyelles réduites ou effacées, bien qu'elles soient légèrement plus longues. La comparaison concerne la lecture et la reformulation, aucune donnée n'étant disponible en conversation. L'interdépendance entre le rythme de parole (*speech rhythm*) et le taux d'articulation (*speech rate*) ne fournit aucune corrélation pour les locuteurs non natifs de l'anglais.

Dans l'étude des non natifs spécifiquement francophones de l'anglais, Horgues (2010) compare ses résultats acoustiques avec ceux des natifs et observe une différence d'intensité moindre entre les voyelles accentuées et inaccentuées. Elle ajoute que le rapport entre syllabes accentuées et inaccentuées est moindre chez les apprenants. Son bilan de la métrique Δ_{intV} , Cv_{intV} , PVI vocalique (variabilité temporelle des intervalles vocaliques, coefficient de variation des intervalles vocaliques, variabilité temporelle des intervalles vocaliques) donne un résultat inférieur aux francophones par rapport aux anglophones natifs.

Ciblant les mêmes groupes, Tortel (2009) a poursuivi l'étude des différentes métriques proposées par la littérature afin d'évaluer la maîtrise prosodique chez les apprenants francophones anglicistes. En déterminant les profils rythmiques d'apprenants francophone, elle parvient à prédire à 69,50% de fiabilité leur niveau d'anglais en trois catégories. Tortel souligne que certaines variables, associées ou non, permettent de davantage cibler la distinction entre les deux groupes d'apprenants alors que d'autres différencient les non natifs des natifs, ou d'autres encore les trois catégories ensemble (2009 : 201). Son tableau, reproduit ci-dessous, en rend compte :

Niveaux de distinction		apprenant vs natif	FR1 vs FR2	Gradation (FR1-FR2-GB)
------------------------	---	--------------------	------------	---------------------------

1 PARAMETRE	% V	ΔC	ΔC
	cvC	nPVI_C	rPVI_C
		nPVI_V	cvV
COMBINAISON	% V	%V, ΔC	rPVI_V, rPVI_C
	cvV	nPVI_C, nPVI_V	cvV, cvC
			%V, cvC

Tableau 16 – Récapitulatif des différentes métriques présentées en fonction de la tâche de discrimination visée (Tortel 2009 : 201)

En fonction du nombre de variables mises en œuvre, Tortel en présente les scores obtenus :

VARIABLE	SCORE OBTENU
cvV	45,79%
ΔC , cvC	62,30%
%V, ΔV , cvV	63,09%
%V, ΔV , cvV, nPVI_V	66,74%

Tableau 17 – Récapitulatif des meilleures prédictions pour chacune des combinaisons factorielles (Tortel 2009 : 202)

Les résultats les plus fiables sont ceux des combinaisons à quatre facteurs. La fiabilité la meilleure est obtenue à 66,74% par la combinaison %V, ΔV , cvV, nPVI_V grâce aux paramètres rythmiques des intervalles vocaliques. Malgré une précision accrue de quatre paramètres coassociés, la fiabilité à 62,30% de la combinaison à deux variables ΔC , cvC semble à Tortel plus satisfaisante

Léon note que « D'une manière générale, une syllabe accentuée est en moyenne deux fois plus longue qu'une syllabe inaccentuée, en français standard. » (2011 : 154). Bien sûr, le nombre de phones composant la syllabe intervient dans la variation de durée, mais celle-ci est statistiquement vérifiée. Il précise que « Dans le français standard, on dit que l'accentuation est *oxytonique*, c'est-à-dire qu'elle tombe sur la *dernière syllabe prononcée* du groupe *sémantique* » (2011 : 155), surtout en lecture non expressive, au caractère « neutre », alors qu'elle se produit parfois sur la première syllabe d'un groupe phonique en conversation spontanée. En anglais, la tendance est plutôt inversée : le schéma de l'accentuation lexicale

(« ce squelette accentuel [dont] partent les règles phonologiques »²²) n'est pas fixe car l'anglais est une langue à accent lexical libre, et l'accent de phrase n'est pas systématiquement porté par la dernière syllabe non plus.

Dans l'étude de Ferragne (2013), la segmentation automatique des voyelles fournit des intervalles de durée plus petits et semblables entre eux, indépendamment de son aptitude à détecter le nombre de voyelles de manière fiable. Il en déduit que la métrique PVI des voyelles (*Pairwise Variability Index*) mesurant la différence de durée moyenne entre toutes les paires d'intervalles vocaliques adjacents sera relativement réduite et donc de valeur discriminatoire peu fiable, ce qui l'oriente vers d'autres paramètres extractibles automatiquement plus fiables.

Ferragne (2013 : 160-161) différencie les intervalles voisés des non voisés et trouve que le pourcentage de durée de voisement est le plus réduit chez les hommes anglophones (moyenne de 26,00) et le plus ample chez les femmes francophones (moyenne de 48,12). L'analyse du pourcentage de durée d'intervalle non voisé ne fournit aucun résultat significatif. Aussi, il en conclut que seul le pourcentage de durée de voisement fournit un paramètre pouvant discriminer les anglophones natifs des apprenants anglicistes francophones, bien que cette pertinence ne concerne que les hommes.

Dans les résultats de la métrique *intensity PVI_i* ou *PVI_i*, seul le facteur de groupe est significatif. *PVI_i* est plus bas chez FR1 (moyenne de 1,48) que GB (moyenne de 1,72). Cette métrique est donc conforme aux conditions requises pour distinguer les apprenants des natifs. C'est l'intensité marquant l'alternance entre voyelle accentuée et voyelle non accentuée contigües qui devient indicateur rythmique.

Bien que les segmentations manuelles et automatiques fournissent des résultats différents, Ferragne arrive à opérer un classement automatique en groupes FR1 et GB (c'est-à-dire deux groupes) avec 8,65% de taux d'erreur. Sa hiérarchisation des facteurs de classification les plus performants est la suivante : 1. Durée vocalique totale du passage ; 2. Pourcentage de durée de voisement ; 3. Nombre de voyelles ; 4. Écart type de la durée vocalique ; 5. *PVI_i*. Il note que l'écart interquartile de registre n'est que marginalement significatif ($p=.057$) ; la durée vocalique moyenne ($p=.952$) ainsi que la variation de coefficient de durée d'intervalles non voisés ($p=.479$) ne sont pas significatifs.

²² DUCHET, J.-L., FRYD, M. (1998). *Manuel d'anglais oral pour les concours*. Paris : CNED. Page 131.

Ballier, Martin et Amand (2016) ambitionnent de trouver une méthode d'évaluation de la qualité de l'interlangue prosodique des francophones anglicistes plus respectueuse de la réalité de la nature linguistique, qu'ils jugent malmenée par le recours unique aux métriques. Pour cela, ils analysent la segmentation syllabique des francophones du corpus d'apprenants anglicistes ENGLISH de Tortel & Hirst (2010).

Les résultats obtenus par la durée moyenne des types de syllabes ne permettent pas de démontrer la supériorité de niveau prosodique des apprenants de niveau avancé (FR2) par rapport au groupe intermédiaire (FR1). Cela corrobore les résultats obtenus par Tortel (2009). La différence de durée des syllabes réduites des différents groupes n'est pas marquante non plus, même pour les natifs anglophones.

La comparaison des médianes des différents groupes n'est pas significative non plus entre les natifs et les apprenants avancés. Par contre, le groupe intermédiaire est plus lent de 23,7 ms que le groupe des natifs, et le groupe avancé est plus rapide que le groupe intermédiaire. Le seuil de la valeur observée (p) choisi pour valider la signification est 0,01, ce qui démontre une évidence forte. La comparaison met en outre en évidence un effet de la différenciation par genre car les hommes sont significativement plus rapides que les femmes en moyenne.

Les tests de comparaison par type de syllabe chez les natifs montrent une différence significative entre les syllabes inaccentuées et les syllabes accentuées ou finales. Par contre, la durée des syllabes sous focus ne semble pas pertinente *a posteriori* pour l'étude car elle ne se démarque pas suffisamment des syllabes finales. La structure syllabique des syllabes accentuées semble pertinente, mais pas la position de la syllabe réduite, qu'elle soit prétonique, intertonique ou syllabe ouverte.

2.4 Synthèse de l'organisation prosodique

Cette section a considéré les aspects les plus fondamentaux de la prosodie en général : aspects temporels, mélodiques et rythmiques. Ainsi qu'il a été mentionné, pour détecter l'accent français Horgues (2010) s'est efforcée d'étudier le phénomène prosodique sous de multiples aspects. Son jugement sur les particularités prosodiques énoncées est donc de grand intérêt. Son test **TPer5** montre que la sévérité de l'évaluation de l'accent français par les auditeurs natifs est proportionnelle au manque d'aisance globale de la production. Horgues note en outre que si la corrélation du degré d'accent français et de fluidité en lecture de texte est bonne, elle l'est moins en conversation spontanée et encore moins en lecture de phrases

déconnectées. Elle conclut en signalant le risque d'inintelligibilité lorsque les facteurs prosodiques tels que l'accentuation lexicale, le rythme, les contours intonatifs procédaient de choix et de réalisations erronés.

Ferragne (2013) conclut que les variables suprasegmentales de durée vocalique, registre et intensité permettent de classer automatiquement et efficacement les locuteurs en groupes de natifs britanniques et de non natifs francophones. Il précise que la durée de lecture est un critère fiable d'évaluation de l'aisance (*fluency*) et que la durée vocalique calculée par segmentation manuelle est plus propice à la variabilité et donne de meilleurs résultats de distinction rythmique.

Ce programme de recherche n'étant pas destiné à l'étude de la prosodie native, française ou anglaise, pour elle-même mais pour l'évaluation de l'interlangue des apprenants, tournons-nous à présent vers cette phase de l'étude.

3 Principes évaluatifs

Ayant préalablement caractérisé la prosodie anglaise des non natifs, on va pouvoir à présent aborder le concept d'interlangue. On poursuivra cette présentation en énonçant les descripteurs prosodiques généraux du *Cadre européen commun de référence pour les langues*, et la notion de traits discriminants (*crierial features*) spécifiques à une langue précise.

3.1 Concept d'interlangue

Le terme d'interlangue s'insère dans un processus d'apprentissage ou d'acquisition d'une langue étrangère ou seconde. Le *Glossaire de didactique de l'anglais* (1997 : 144) le définit ainsi :

Langue intermédiaire que l'apprenant constitue à partir de tous les matériaux à sa disposition — verbaux ou non verbaux — issus de sa langue de départ ou de la langue-cible. Il les intègre dans un système provisoire d'hypothèses sujettes à révision, système qui lui est personnel et manifeste l'état ponctuel de sa connaissance du monde. (Quivy et Tardieu, 1997 : 144)

Observant que la très grande majorité des apprenants ne réalisait pas des énoncés en « langue cible » (*target language, TL*) de la même qualité qu'un natif et rencontrait des difficultés à se rapprocher du modèle natif, Selinker a créé le concept d'« interlangue » (*interlanguage*) pour désigner le système linguistique créé par l'apprenant pour s'exprimer

dans cette nouvelle langue en essayant de se rapprocher de la norme langagière (Selinker, 1972 : 214) :

Since we can observe that these two sets of utterances are not identical, then in the making of constructs relevant to a theory of second-language learning, one would be completely justified in hypothesizing, perhaps even *compelled* to hypothesize, the existence of a separate linguistic system based on the observable output which results from a learner's attempted production of a TL norm. This linguistic system we will call 'interlanguage' (IL).

Il ajoute que les données observables destinées à étudier cette interlangue proviennent de trois sources différentes qui sont (1972 : 214) :

- (1) les énoncés que l'apprenant prononce dans sa langue maternelle (*native language, NL*),
- (2) les énoncés d'interlangue produits par l'apprenant,
- (3) les énoncés produits par les natifs de la langue objet d'apprentissage (langue cible) dans leur langue.

Dans l'apprentissage d'une langue autre que la langue maternelle, les processus psycholinguistiques à l'œuvre dans la structuration de l'apprentissage sont bloqués par le phénomène qu'il nomme la « fossilisation » :

Fossilizable linguistic phenomena are linguistic items, rules and subsystems which speakers of a particular NL will tend to keep in their IL relative to a particular TL, no matter what the age of the learner or amount of explanation and instruction he receives in the TL. (Selinker, 1972: 215)

La difficulté inhérente à l'apprentissage sera de surmonter ce phénomène de fossilisation, induit par l'influence de la langue maternelle, pour progresser en minimisant les écueils fondamentaux, conscients ou inconscients, qu'il nomme : *language transfer, transfer-of-training, strategies of second-language learning, strategies of second-language communication, overgeneralization of TL linguistic material* (Selinker, 1972: 215). Une combinaison de plusieurs d'entre eux est possible. Selinker souligne le fait que certaines structures linguistiques ne disparaissent jamais vraiment, surtout en cas de stress, de déstabilisation ou de relâchement, et qu'elles constituent un matériau de grande valeur pour le chercheur acquisitionniste (Selinker, 1972: 221). Selinker et Lamendella (1981 : 203) dissocient *overt microbehaviour* (que le laïque, comme le spécialiste, peut observer) de *covert*

microbehaviour (dont la connaissance n'est accessible que par expérimentation neurophysiologique). Selon eux, la conjonction de ces deux approches devrait aider à modéliser transferts et phases plateaux de l'interlangue en révélant le processus dynamique en œuvre dans l'acquisition. Ils considèrent que la réponse aux questions posées ne peut provenir, non pas de l'accumulation d'un grand nombre de données, mais que de leur pertinence, de leur organisation judicieuse et surtout de l'interprétation donnée par rapport au positionnement théorique (1981 : 218).

Corder a analysé le phénomène d'erreur dans le discours des apprenants et considère que l'interlangue n'est pas simplement à la croisée de la langue maternelle et de la langue cible, mais qu'elle est une langue à part entière, avec ses propres règles :

“Thus we come to the conclusion that the concept of ungrammaticality or deviance is not applicable to the learner. Everything he utters is by definition a grammatical utterance in his dialect.” (Corder 1981 : 32)

“It is proposed that a description of the learner's 'état de dialecte' can be better achieved by a recognition of what he speaks is not an inadequate or incorrect form of the target language but a peculiar transitional idiolect, which should be approached in the same way as the language of an infant or some unknown language. (Corder 1981 : 34)

L'interlangue est donc considérée comme une nouvelle variété dialectale en perpétuelle mutation car ses composants se situent le long d'un continuum qui reste à être étudié et analysé.

3.2 Descripteurs prosodiques du CECRL

Les critères permettant d'accéder à une maîtrise de la compétence phonologique sont mentionnés aux pages 91 et 92 du CECRL. Ils sont minimalistes et semblent présentés comme déconnectés de l'utilisation de la langue en situation de communication.

Un locuteur de niveau A1 sera « compréhensible avec quelque effort », la prononciation d'un locuteur de niveau A2 sera « en général suffisamment claire pour être comprise malgré un net accent étranger » pour devenir « clairement intelligible » à un niveau B1 « même si un accent étranger est quelquefois perceptible et si des erreurs de prononciation proviennent occasionnellement ». Un apprenant de niveau B2 aura « une prononciation et une intonation claires et naturelles » alors qu'aux niveaux C1 et C2, il sera capable de « varier l'intonation et placer l'accent phrastique correctement afin d'exprimer de fines nuances de

sens ». Un encart précise néanmoins qu'il appartient aux utilisateurs du *CECRL* (enseignants, institution, évaluateurs) d'envisager et d'expliciter « les aptitudes phonologiques nouvelles exigées de l'apprenant » ainsi que « l'importance relative des sons et de la prosodie ». « L'exactitude phonétique et l'aisance » sont aussi laissées à la discrétion de l'utilisateur.

Les préconisations du *CECRL* devant être applicables à de nombreuses langues, il est compréhensible qu'elles ne soient pas plus précises en la matière. Néanmoins, il est déplorable que la compétence phonologique ne soit pas davantage incluse dans d'autres compétences, telles que la compétence lexicale. En effet, forme et sens sont indissociables dans un apprentissage performant, notamment en anglais où l'accent lexical est déterminant pour la compréhension.

C'est dans les descripteurs des aspects qualitatifs de l'utilisation de la langue parlée que l'on trouve des critères phonologiques dont la réalisation permet de franchir les divers seuils des niveaux de compétence (2001 : 28, Tableau 3). Le tableau présentant les aspects qualitatifs de l'utilisation de la langue parlée (*CECRL*, 2000 : 28-29, *CEFRL*, 2001 : 28-29) ne mentionne que deux rubriques en lien avec la prosodie selon les divers niveaux communs : il s'agit de l'aisance (*fluency*) et de l'interaction (*interaction*). Cette dernière se focalise essentiellement sur les tours de parole entre interlocuteurs, bien que soient mentionnés des indices intonatifs (*picking up and using non-verbal and intonational cues apparently effortlessly*), mais seulement au niveau C2. En conséquence, on doit se tourner vers les descripteurs d'aisance dont les six niveaux rendent compte. Le tableau ci-dessous les récapitule pour la conversation spontanée (j'ai noté en caractères gras les termes essentiels) :

Niveau	Synthèse prosodique	Citations
C2	Capacité à « s'exprimer longuement, spontanément dans un discours naturel » sans que des auditeurs natifs puissent vraiment détecter d'éventuelles difficultés. Caractère fluide et naturel de la conversation, similaire à celui d'un natif.	Can express him/herself spontaneously at length with a natural colloquial flow , avoiding or backtracking around any difficulty so smoothly that the interlocutor is hardly aware of it.
C1	Capacité à s'exprimer avec « aisance et spontanéité presque sans effort » en gardant « le flot naturel et fluide du discours » dans la plupart des circonstances. Caractère fluide, spontané et naturel du discours.	Can express him/herself fluently and spontaneously conceptually difficult subject can hinder a natural, smooth flow of language.
B2	Capacité à « parler relativement longtemps avec un débit assez régulier », quelques hésitations mais peu de longues pauses. Caractère fluide et régulier du discours rarement interrompu par de longues pauses.	Can produce stretches of language with a fairly even tempo ; although he/she can be hesitant as he/she searches for patterns and expressions. There are few noticeably long pauses.
B1	Capacité à se faire comprendre malgré des pauses évidentes, surtout en prise de parole longue. Caractère moins fluide du discours entrecoupé de pauses évidentes survenant en prise de parole longue.	Can keep going comprehensibly , even though pausing for grammatical and lexical planning and repair is very evident, especially in longer stretches of free production.
A2	Capacité à se faire comprendre dans de brèves interventions, malgré de nombreuses pauses. Caractère haché du discours ne remettant pas en question la compréhension générale du message.	Can make him/herself understood in very short utterances, even though pauses, false starts and reformulation are very evident.
A1	Capacité à produire « des énoncés très courts, isolés et généralement stéréotypés » marqués par de nombreuses pauses. Difficulté de la communication. Caractère très haché du discours constitué d'énoncés très courts, rendant la communication difficile.	Can manage very short, isolated, mainly pre-packaged utterances, with much pausing to search for expressions, to articulate less familiar words, and to repair communication.

Tableau 18 – Tableau récapitulatif des marqueurs prosodiques de l'aisance selon le *CECRL* et *CEFR* (d'après *CECRL*, 2000 : 28-29, *CEFR*, 2001 : 28-29)

Chaque niveau de ce tableau fournit un continuum d'éléments pouvant se traduire par des mesures acoustiques : la durée des pauses et des unités inter-pausales.

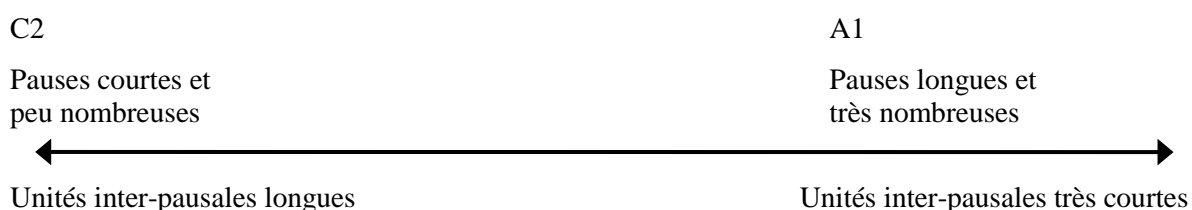


Figure 13 – Représentation schématisée du continuum d'aisance (*fluency*) dans les six niveaux du *CECRL*

Ensemble, ces mesures constituent des paramètres de la vitesse générale d'élocution et fournissent une première approche de ce que peuvent être des critères d'évaluation acoustiques dans le domaine prosodique. Chaque niveau correspond à une maîtrise accrue du rythme. Au niveau A1, les pauses seront longues et nombreuses, hachant et allongeant le discours, puis diminueront au fur et à mesure des progrès réalisés en expression orale, jusqu'à ce que ce discours, en niveau C2, devienne fluide et spontané, permettant d'interagir avec facilité tout en prenant son tour de parole naturellement. Les contours prosodiques seront réduits lorsqu'ils seront réalisés avec quelques mots (niveau A2, colonne « Étendue ») alors qu'aux niveaux B2 et C1, il est implicite que la longueur des phrases et la cohérence de leur structure (colonne « Cohérence ») se traduiront par la production de plus longs segments prosodiques. Ce n'est qu'à partir du niveau C2 que sont abordés les aspects qualitatifs de la prosodie : « [...] permettant de transmettre avec précision des nuances fines de sens afin d'insister, de discriminer ou de lever l'ambiguïté » et « [...] en utilisant les indices non verbaux et intonatifs sans effort apparent ».

Pourtant, nous allons voir que seul le rythme a été retenu comme révélateur du niveau d'un apprenant.

3.3 Criterial features (critères d'évaluation) adaptés à la prosodie anglaise

McCarthy (2010), se focalisant sur la composante phonologique, cible la composante temporelle (en particulier la fréquence des pauses) pour évaluer de façon fiable monologues et interactions selon les niveaux de maîtrise prescrits par le *CECRL*. Il se base sur de nombreuses études qui montrent que les pauses sont le résultat de phénomènes complexes très coûteux sur le plan cognitif lors de l'agencement des idées ou de la recherche de lexique. Il classe les recherches sur l'aisance (*fluency*) en quatre catégories :

1. La vitesse d'élocution et la fluidité de parole (*rate and smoothness of delivery*), lesquelles peuvent être objectivées par le calcul du nombre de mots par minute ou unité de parole ainsi que la localisation, la distribution et la durée des pauses, la régularité sans accroc du tempo et ses corolaires ;
2. Le caractère automatique (*automaticity*), c'est-à-dire un accès lexical immédiat et sans effort de l'apprenant aux mots, expressions et propositions ;
3. L'aisance perçue par les professionnels que sont les enseignants, les concepteurs de programmes et les évaluateurs ;

4. Les personnes n'appartenant pas au groupe 3, que représentent tout un chacun, les pairs, les employeurs, etc.

Selon lui, d'autres critères entrent en jeu dans les interactions, notamment celui de *confluence*, surtout pour les niveaux les plus élevés ([...] *speed of delivery alone is clearly a blunt instrument for assessing fluency and cannot tell the whole story. [...] Pauses may not necessarily be a sign of communicative failure but may indicate complex planning and boosted cognitive effort.*). Il rapporte que l'interaction inter-locuteurs accélère la vitesse de parole selon Wolf (2008), et que Bavelas *et al.* (2000) remarquent la corrélation entre la capacité du locuteur à raconter et l'activité de l'interlocuteur (*backchanneling*). McCarthy souligne le caractère interactif de l'aisance étant donné qu'il s'agit là d'une responsabilité partagée (*fluency as a co-created achievement*) pour assurer la réussite d'une interaction efficace (*confluence*) dans le monde réel, ainsi que la nécessité de s'appuyer sur l'étude de corpus afin d'établir des échelles de niveaux, dimension absente du *CECRL*.

Gut (2009 : 79) mentionne elle aussi que la vitesse d'élocution a longtemps été synonyme d'automatisation et de procéduralisation des savoir faire. Barker *et al.* (2011) se basent sur la nature de la langue maternelle de l'apprenant pour vérifier les incidences du rythme. Ils s'appuient sur la dichotomie entre langues accentuelles et langues syllabiques et observent que leur expérimentation sur l'espagnol, langue syllabique, est concluante.

Cette conception minimaliste de la prosodie occulte la composante intonative qui joue un rôle fondamental dans l'impact que peut avoir un discours sur l'interlocuteur. Des études comparatives de corpus oraux d'anglophones et d'apprenants anglicistes francophones montrent que le manque d'adéquation entre les choix lexicaux ou grammaticaux effectués par le locuteur et l'intonation utilisée est de nature à nuire à l'interlocution. Aussi, il semble qu'il faille aller plus loin dans la recherche de traits critériés pour l'évaluation de l'oral, si possible en relation avec des objectifs communicationnels. Le programme de recherche interdisciplinaire *English Profile*²³ est destiné à appliquer les prescriptions de *CECRL* à la langue anglaise. Il s'en différencie par son approche verticale (par domaines d'investigation) alors que le *CECRL* privilégie une description horizontale (par niveaux de référence). Les banques de données lexicales et grammaticales ne cessent de croître, cependant les particularités phonologiques et surtout intonatives restent à développer. Pourtant, la vitesse

²³ Projet *English Profile* : [<http://www.englishprofile.org/>], dernièrement consulté en 2017.

générale d'élocution représentée par les mesures de durée semble être un indicateur digne d'intérêt pour ordonner les fichiers sons d'un texte lu.

Jusqu'à présent, ce chapitre s'est focalisé sur la teneur des objectifs à atteindre. Orientons-nous maintenant vers les conditions expérimentales attribuées à cette recherche.

4 Conditions expérimentales de cette recherche

Cette thèse prend aussi son origine dans le projet de recherche *Longdale*²⁴, fruit d'une coopération entre l'*Université Catholique de Louvain* en Belgique et huit autres partenaires internationaux dont l'*Université Paris Diderot - Paris 7*.

4.1 Corpus *Longdale-Charliphonia*

Dans le cadre des projets *Diderot-Longdale*²⁵ et *Longdale-Charliphonia*²⁶ en découlant pour l'Équipe d'accueil du laboratoire *CLILLAC-ARP* à l'*Université Paris 7*, 205 étudiants volontaires non natifs de la langue anglaise ont été enregistrés en lecture puis en conversation sur une période d'au moins trois ans (à partir d'octobre 2009), constituant un corpus longitudinal. Le premier texte enregistré consiste en un collage²⁷ à partir d'extraits du conte d'Oscar Wilde *The Selfish Giant* (426 mots²⁸). La lecture de ce texte (S001R) a immédiatement été suivie de l'enregistrement d'une conversation guidée avec lecteur anglophone privilégiant un thème de discussion parmi trois proposés (S001). En phase finale, c'est-à-dire en fin de la troisième année de licence pour les étudiants, le même texte a été relu et réenregistré, ainsi que de nouveaux fichiers de conversation guidée. Les étapes intermédiaires sont ponctuées d'enregistrements de lecture et de conversation guidée.

Des questionnaires de métadonnées en ligne ont systématiquement été complétés en anglais par les étudiants juste avant les différentes phases d'enregistrement par des lecteurs anglophones natifs qui conduisaient l'entrevue en anglais. Les questions posées dans ce questionnaire de départ regroupaient trois catégories :

²⁴ *Longitudinal Database of Learner English* [<http://www.clillac-arp.univ-paris-diderot.fr/projets/longdale>], dernièrement consulté en 2017.

²⁵ [<http://www.clillac-arp.univ-paris-diderot.fr/projets/longdale>]

²⁶ [<http://www.clillac-arp.univ-paris-diderot.fr/projets/charliphonia>]

²⁷ Voir en annexe.

²⁸ La définition d'un mot sera ici d'ordre graphique : un mot sera délimité par des espaces ou signes de ponctuation.

- Questions en lien avec l'établissement du questionnaire (identification, vérification, date de l'enregistrement, vérification de la finalisation du questionnaire),
- Questions centrées sur l'apprenant concernant sa date de naissance, son sexe, ses caractéristiques linguistiques, ses conditions d'apprentissage linguistique, les autres langues apprises ou pratiquées, les études poursuivies, la qualification recherchée, le lieu et conditions d'étude, le niveau atteint en termes de réussite à des certifications, le temps passé dans des pays anglophones,
- Questions concernant les tâches à accomplir pendant l'entrevue : le sujet de conversation choisi parmi trois possibilités et le titre du texte à lire.

Les apprenants pouvaient en outre renseigner un champ supplémentaire pour mentionner toute autre information susceptible d'être utile. Le corpus *Longdale-Charliphonia* était destiné à une étude large, comprenant de nombreux domaines de recherche. Contrairement aux questionnaires complétés lors de l'élaboration du corpus *LeaP (Learning Prosody in a Foreign Language)*, spécifiquement créé dans une optique de recherche prosodique, aucune métadonnée n'a été demandée en lien avec la prononciation, ni la prosodie, ou même l'intérêt, la compétence ou l'expérience des participants dans le domaine musical ou théâtral. La question de leur stratégies d'apprentissage n'a pas été non plus abordée.

Ces métadonnées manquantes, utiles à la compréhension du fait linguistique et à l'élaboration du profil de l'apprenant, peuvent néanmoins être retrouvées en analysant les fichiers de conversation guidée. Elles sont présentées oralement, de manière indirecte, grâce à des questions informelles générales posées au cours de l'échange conversationnel afin d'alimenter la prise de parole. L'apprenant avait tout loisir de formuler une réponse comprenant détails et exemples s'il le souhaitait et le lexique utilisé pour s'exprimer sur ces sujets étant largement à sa portée. Au cours de cet échange informel avec un lecteur de langue anglaise, on peut penser qu'il a pu se sentir davantage libre de s'exprimer selon son souhait que s'il lui avait fallu compléter un formulaire écrit avant l'interview. Les conditions de sollicitation étant inscrites dans un échange informel après avoir partagé plusieurs minutes à parler du premier sujet, il est probable que les étudiants ont répondu à ces questions avec franchise et confiance, en précisant les détails qui leur semblaient pertinents, sans prendre le recul dont on fait preuve en complétant un questionnaire et qui peut biaiser les réponses.

Ces métadonnées, obtenues via questionnaire ou question informelle pendant l'échange conversationnel, ont été précieuses lors des choix à opérer au cours de notre étude, en tout premier lieu pour la sélection des apprenants à effectuer.

4.2 Parole lue

Les enregistrements du *Selfish Giant* de la première phase ont constitué la matière brute de départ sur laquelle je devais m'appuyer pour étayer ma recherche sur l'évaluation de la prosodie des apprenants. L'avantage indéniable était le nombre important des enregistrements, réalisés dans les mêmes conditions avec une population homogène d'apprenants volontaires et des métadonnées recueillies sur tableur. La lecture permet de comparer très précisément un grand nombre de réalisations quasiment identiques et, en cela, de minimiser la variabilité du support pour étudier en détail la variation de la parole émise dans des conditions identiques.

Le point négatif était la nature de l'enregistrement. On a vu dans les résultats des expérimentations présentés en amont dans ce chapitre que lecture et conversation spontanée n'obéissaient pas aux mêmes règles et n'obtenaient pas les mêmes résultats. De plus, il s'agit d'un conte, c'est-à-dire un type de texte spécial, narratif, doté d'un rythme particulier et ne représentant pas la langue parlée. Les contes sont ordinairement lus à voix haute aux enfants et la lecture de ce support doit s'adapter naturellement et instinctivement à ce genre de public. Deux autres écueils supplémentaires complètent le tableau. L'un d'eux est la langue archaïque faisant buter apprenants mais aussi anglophones natifs sur des structures et expressions inattendues. L'autre étant que les quelques rares phrases au discours direct sont cadrées par un discours citant donnant des directives très précises sur la manière dont le lecteur doit maquiller sa voix. L'ensemble de l'enregistrement résulte finalement en une lecture contrefaite, très peu propice à trouver des traits distinctifs dans la parole générique des apprenants. Par ailleurs, plusieurs fichiers du début de l'expérimentation sont bruités.

5 Démarche et organisation générale de la thèse

Le Chapitre I a jusque là exposé le panorama théorique et expérimental dans lequel se situent les études sur le sujet. Il a aussi présenté le cadre de travail constituant la présente recherche. Prenant appui sur ces éléments, la question de recherche fondamentale peut se traduire en d'autres questions de recherche subalternes et en hypothèses propres à atteindre l'objectif.

5.1 Choix méthodologiques pour remédier à la difficulté de modéliser l'interlangue prosodique

Selon les préconisations méthodologiques de Martinez et Berkhout (2009), une bonne question de recherche ou hypothèse doit être faisable, « susciter l'intérêt du ou des investigateurs », être « originale dans le sens d'innovante », éthique et surtout pertinente. On a déjà vu le degré de pertinence de cette recherche sur la modélisation de l'interlangue prosodique par l'intérêt pédagogique, économique et social qui s'y rattache. Cela est de nature à susciter l'intérêt de l'investigateur mais aussi de collaborateurs potentiels sollicités pour mener à bien ce projet de recherche. Si l'on considère le phénomène évaluatif comme égalitaire, favorisant la valeur des individus, cette recherche peut être éthique ; la méthodologie mise en œuvre pour atteindre l'objectif devra l'être également. Par contre, la difficulté prévisible est celle de la faisabilité : les analyses précédentes ont montré l'extrême variabilité des données, peu propices à une modélisation de la prosodie des apprenants. C'est d'ailleurs la raison pour laquelle ce type de recherche est peu avancé par rapport aux autres. Ce point est fondamental : ce travail ne peut aboutir qu'en cadrant la variabilité qui lui est inhérente. La question de recherche qui s'impose porte donc sur le choix de la méthodologie à utiliser pour atteindre le but. Étant donné les difficultés prévisibles, il semble particulièrement indiqué de ne pas se cantonner aux méthodologies conventionnelles dans l'exploration de la matière prosodique, mais de s'appuyer fermement sur l'originalité des recherches antérieures afin de bâtir une méthodologie encore plus originale et innovante, de nature à se dégager des ornières sclérosantes et ainsi atteindre un degré élevé dans la compréhension du phénomène prosodique. Originalité et innovation ne sont pas des conditions supplémentaires, elles constituent l'essence de la faisabilité de ce travail de recherche.

En conséquence, je suivrai l'exemple de Horgues (2010) en ne sélectionnant aucun cadre de recherche d'emblée, qui pourrait restreindre une vue d'ensemble des phénomènes ou occulter certaines approches fécondes. Ainsi, je pense favoriser les possibilités les plus nombreuses pour répondre adéquatement aux difficultés à résoudre et élargir les possibilités de choix à opérer en fonction des nécessités et possibilités du terrain. Dans une optique descriptive, j'utiliserai les termes de l'École britannique de l'intonation.

Cadrer la variabilité signifie trouver des repères objectifs, alors que la prosodie est par nature l'expression de la subjectivité. On a vu que la langue cible, l'anglais, était le repère de base modélisateur choisi dans les études. Cela ne résout aucunement le problème car la langue

native est elle-même dotée d'une grande variabilité. Il est donc nécessaire de préalablement comprendre la nature de la variabilité chez les natifs anglophones. Quatre axes d'investigation de la parole s'offrent à nous pour scinder la difficulté et ainsi la réduire : d'une part au niveau des locuteurs les axes d'analyse intra-locuteur et interlocuteurs, d'autre part au seul niveau de leurs réalisations les axes syntagmatique et paradigmatique. Pour sonder la complexité prosodique selon l'axe syntagmatique, j'appliquerai les notions de « profil » et « profilage » et pour faire de même sur l'axe paradigmatique j'utiliserai la détection du degré d'emphase. Ces concepts sont développés en aval, je ne m'y attarderai donc pas ici. Je me contenterai d'expliquer leur choix non pas par leur nature, mais par leur fonction en tant qu'outils de catégorisation et de hiérarchisation. Chacun des quatre est destiné à scinder le matériau prosodique en le catégorisant, en le hiérarchisant puis en opérant des sélections (locuteurs ou réalisations prosodiques conservés, regroupés, écartés de l'analyse). Dans la résolution de problèmes, peu importe la logique du chercheur, s'il planifie sa méthodologie d'après la phase initiale de description de la situation ou suivant l'objectif final à atteindre, l'essentiel est d'organiser une méthodologie en mesure d'atteindre l'objectif (Barsalou 1999)²⁹. Le but recherché est de favoriser l'opérationnalisation des choix en faisant éclater le matériau prosodique pour le réagencer selon des critères différents de ceux traditionnellement utilisés, et ceci avec une différence fondamentale : il ne s'agit pas de l'atomiser pour observer des parties encore plus réduites réagencées sous forme de métriques, dissociées de leur nature profonde : l'humain, le vivant. Bien au contraire, l'objectif méthodologique est de replacer le phénomène prosodique dans son environnement afin de l'observer vivre en symbiose avec la subjectivité humaine et favoriser autant que faire se peut l'ancrage interlocutif (Kerbrat-Orecchioni [1999] 2009, Douay 2000) de la langue pour repérer les tendances. La quête de repères objectifs et tangibles est limitée si les choix sont autoréférencés. Ainsi devrait être favorisée l'aptitude de la prosodie à refléter l'auto-organisation du cerveau en fonction du vécu des situations. À partir des tendances dégagées, le lien forme sonore-sens devrait émerger et établir une hiérarchie organisationnelle nouvelle au sein même des sous-domaines temporel, mélodique, de registre et accentuel. Bien qu'elles fonctionnent en synergie, il est intéressant de garder ces catégorisations afin d'analyser plus finement leur structure organisationnelle en interaction. Cela devrait permettre de privilégier la précision ainsi que de

²⁹ *From the perspective of perceptual symbol systems, problem solving is the process of constructing a perceptual simulation that leads from an initial state to a goal state. Problem solvers can work forward from the initial state or backward from the goal state, but in either case they attempt to simulate a plan that achieves the goal. In novice problem solving, the difficulty is finding satisfactory components of the simulation and ordering them properly.* Barsalou (1999 : 605)

favoriser le croisement des données pour vérifier la convergence des résultats sous divers angles d'étude.

Je résumerai très sommairement ces propos en écrivant que pour mettre en application ces idées sur la manière d'organiser cette recherche, il est donc nécessaire de créer un corpus parallèle et complémentaire de locuteurs natifs ayant lu le même texte dans les mêmes conditions. Ce nouveau corpus sert de base à l'analyse par profils pour obtenir une sélection des locuteurs natifs à analyser acoustiquement (Chapitre II), puis il est utilisé pour l'analyse par l'emphase afin d'obtenir une sélection des zones du texte à privilégier pour l'étude acoustique (Chapitre III). En fonction de ces repères natifs, le Chapitre IV vérifie l'hypothèse selon laquelle la vitesse d'élocution serait révélatrice du niveau prosodique des apprenants et propose des critères de catégorisation et hiérarchisation de l'interlangue prosodique par variables pour une utilisation dans une évaluation semi-automatique. Le Chapitre V est destiné à vérifier le degré de fiabilité des résultats issus de l'évaluation prosodique semi-automatique par comparaison avec ceux attribués par méthode traditionnelle.

5.2 Questions de recherche, hypothèses et structure de la thèse

En conséquence, le travail de recherche de cette thèse s'articule tout d'abord autour de trois hypothèses, chacune soumise à validation en fin de chapitre :

H1 — Étant donné la forte variabilité des réalisations prosodiques natives et d'apprenants, des outils spécifiques doivent être utilisés afin de l'analyser. La création de « profils prosodiques » devrait permettre de mieux organiser et hiérarchiser les réalisations inter-locuteurs en fonction de regroupements des variables. (expérimentation soumise à validation en fin de Chapitre II)

H2 — Le phénomène d'emphase favorise la variabilité des réalisations prosodiques. Aussi, son repérage permet de limiter l'étude à des entités concentrant la variation maximale. Le phénomène d'emphase devrait alors permettre de mieux comprendre, organiser et hiérarchiser la variabilité des phénomènes prosodiques intra-locuteurs en graduant la variation pour l'analyser plus précisément ou au contraire pour l'exclure des analyses. (expérimentation soumise à validation en fin de Chapitre III)

H3 — Le paramètre de rapidité d'élocution étant un indicateur déterminant de niveau des apprenants, la catégorisation de ceux-ci selon ce paramètre devrait permettre d'organiser la variation inter-locuteurs des résultats des autres variables prosodiques afin de trouver des critères d'évaluation fiables. (expérimentation soumise à validation en début de Chapitre IV)

L'investigation se poursuit et prend la forme de deux questions de recherche étant donné que dans QR1 la tâche assignée est de parvenir à créer des traits distinctifs évaluatifs via profilage par variables semi-automatiques, et que QR2 se réfère à l'évaluation du travail de thèse en calculant le degré de fiabilité de l'évaluation par variables.

QR1 — Étant donné la forte variabilité des réalisations prosodiques natives et d'apprenants, des outils spécifiques doivent être utilisés ou créés afin de l'analyser. La validation des Hypothèses H1, H2 et H3 devrait permettre de créer des profils prosodiques de non natifs permettant d'orienter la création de traits discriminants évaluatifs. (expérimentation effectuée dans le Chapitre IV)

QR2 — À quel degré de fiabilité les variables à la base des profils prosodiques d'apprenants et de leurs grilles évaluatives sont-elles validées par comparaison de leurs résultats à ceux d'un test perceptif effectué par des experts ? (expérimentation effectuée dans le Chapitre V)

La validation des trois hypothèses présentées ci-dessus charpente ce travail de recherche en cadrant les principes de sa mise en œuvre et conduisant à sa validation ou son manque d'adéquation à la problématique. Les deux questions de recherche suivantes aboutissent à la véritable finalité de cette thèse, qui est la création effective de profils et de critères d'évaluation prosodiques homologués. Cette thèse est ainsi structurée aux cinq chapitres conduisant à la conclusion récapitulative de la cohérence de cet agencement, lesquels figurent ci-dessous :

CHAPITRE I - Prosodie anglaise et évaluation prosodique d'apprenants (état de l'art)

CHAPITRE II - Profils et profilage : mode d'organisation hiérarchique de la prosodie des anglophones natifs (vérification de l'Hypothèse 1)

CHAPITRE III - Emphase : mode d'organisation paradigmatique des réalisations prosodiques des anglophones natifs (vérification de l'Hypothèse 2)

CHAPITRE IV – Expérimentation de création de profils et de critères d'évaluation prosodiques d'apprenants (vérification de l'Hypothèse 3, Question de recherche 1)

CHAPITRE V - Vérification de la fiabilité du système d'évaluation prototypique semi-automatique par variables (Question de recherche 2)

CHAPITRE II – Profils et profilage : mode d’organisation hiérarchique de la prosodie des anglophones natifs

1 Résumé du Chapitre II et introduction

Ce deuxième chapitre a pour fonction d’examiner la prosodie des anglophones natifs selon le mode d’organisation en profils. Son objectif fondamental est de repérer les variables à prendre en compte sur le plan syntagmatique (agencement des corrélats le long de la chaîne sonore) afin de contribuer à construire un modèle prosodique évaluatif natif.

Pour cela, ce chapitre introduit les notions de « profil » et de « profilage ». Il se centre sur leur application linguistique ou prosodique dans l’étude d’individus puis de groupes. Le bilan tiré permet d’établir un profilage de la norme native en lecture par tests croisés (perceptif PER-LEC-N, introspectif EXP-LEC-N, acoustique COR-AC-N1, puis statistique STAT-N1).

Ce chapitre aboutit à deux sortes d’avancées dans la création d’outils pour analyser l’interlangue.

Dans le domaine des données, j’obtiens :

- Des profils d’anglophones natifs pour établir des modèles d’évaluation prosodiques,
- Des critères acoustiques pour modéliser la norme évaluative,

lesquels valident l’Hypothèse 1.

Dans le domaine méthodologique, l’état des lieux du potentiel et des limites des méthodes de profilage dans la littérature rend possible la détermination de profilage pour la suite de ma recherche.

Si la première étape d’une étude de l’interlangue sur corpus est de constituer ces corpus de travail, la deuxième est de trouver un point d’attaque pour en extraire les données les plus pertinentes, fidèles et représentatives de la description de l’interlangue. Dans cet esprit, Granger et Rayson (1998 : 131) prônent la notion de « profilage » :

[...] we have shown that automatic profiling can help researchers form a quick picture of the interlanguage of a given learner population and that it opens up interesting avenues for future research.

Ce chapitre est consacré à l'étude des profils et du profilage afin de poursuivre les objectifs suivants :

- établir des profils des anglophones natifs afin de sélectionner celui ou ceux qui peuvent être pris comme modèles d'évaluation prosodiques,
- à partir de ces modèles, obtenir les critères acoustiques correspondant à une bonne lecture native, en évitant tout excès,
- valider par diverses méthodes la fiabilité des profils obtenus,
- étudier diverses méthodes de profilage afin de sélectionner celles qui peuvent être reprises dans la présente étude ou d'envisager d'autres méthodes plus adéquates,
- vérifier le potentiel de la méthodologie par profils et ses limites afin de valider ou non l'Hypothèse 1, rappelée ci-dessous :

H1 — Étant donné la forte variabilité des réalisations prosodiques natives et d'apprenants, des outils spécifiques doivent être utilisés afin de l'analyser. La création de « profils prosodiques » devrait permettre de mieux organiser et hiérarchiser les réalisations inter-locuteurs en fonction de regroupements des variables.

Le processus de profilage et son corollaire, le profil, sont présentés comme des outils qui facilitent l'analyse de nombreux individus et semblent promettre des avancées nouvelles. Le champ couvert par les notions de « profil » et de « profilage » est extrêmement vaste. Aussi, afin d'en cerner l'essence, je vais tout d'abord définir la notion générale de « profil », puis orienter mes investigations vers son application aux domaines se rapprochant du sujet de cette étude en me focalisant sur les profils personnalisés puis sur les profils de groupes. Je m'interrogerai sur les opérations de profilage de ces divers profils pris comme exemples et présenterai alors le choix d'utiliser la notion de profil dans ma méthodologie pour la constitution d'un profil prosodique de natif.

2 Profil et profilage : définitions propres au domaine

Cette définition de la nature des termes « profil » et « profilage » est déclinée par analyse lexicographique avant que la notion soit élargie au sens qui converge avec cette étude.

2.1 Analyse lexicographique du mot « profil »

Afin de vérifier les divers sens que peuvent avoir les mots « profil » en français et « profile » en anglais, dix dictionnaires en ligne (trois de langue française et sept de langue anglaise) ont été consultés les 8 et 9 novembre 2014 :

- *Cambridge American English Dictionary* (CAED)
- *Cambridge British English Dictionary* (CBED)
- *Dictionnaire de français Larousse* (DFL)
- *Le grand dictionnaire terminologique* (GDT) de l'Office québécois de la langue française
- *Le Trésor de la Langue Française Informatisé* (TLFI)
- *Longman Dictionary of Contemporary English* (LDOCE)
- *Macmillan Dictionary* (MMD)
- *Merriam-Webster's Online Dictionary* (MWD)
- *Oxford English Dictionary* (OED)

ainsi que *Le Petit Robert* (PR, version électronique du *Nouveau Petit Robert*, version VUEF, 2.1, 2001).

Les mots « profil » ou *profile* ont respectivement été introduits dans les langues française et anglaise vers le milieu du XVII^{ème} siècle à partir de l'italien (PR, DFL, MWD). Le sens premier porté par le mot français est « Contours, traits d'un visage vu par un de ses côtés » (TLF). À partir de ce sens concret, un élargissement de la notion aux éléments inanimés sera : « aspect de quelque chose, vu latéralement, et dont le contour se détache » (DFL). Un glissement de sens donnera un « ensemble des traits caractéristiques d'une chose, d'une situation, d'une catégorie de personnes. » (TLF). À partir de cela, le sémantisme du vocable sera lié à sa finalité : on trouve un « ensemble des traits psychologiques marquants qui définissent le type de personne adéquat pour exercer une fonction, remplir une tâche » (TLF). Un nouvel élargissement du sens aux éléments inanimés deviendra un « ensemble de caractéristiques qui définissent fondamentalement un type de chose ; configuration de quelque chose à un moment donné » (DFL).

Le mot anglais recoupe ces mêmes notions, soulignant la vue de côté (*An outline of something, especially a person's face, as seen from one side* (OED)), la description par les caractéristiques saillantes d'une personne (*a description of someone containing all the most important or interesting facts about the person* (CAED)), définition pouvant être élargie à un groupe de personnes (*a description of a person, group, or organization that contains all the details that someone needs* (MMD)), ou même à des inanimés (*a brief written description that provides information about someone or something* (MWD)). Par contre, l'ordre de présentation est parfois différent. Certains dictionnaires privilégient la notion de *degree or level of public exposure* (MWD), *the amount of public attention that someone or something receives* (CAED), qui ne reçoit que peu d'écho en français, sauf à la rigueur pour évoquer lointainement l'« hypothèse maximale, minimale d'un programme d'action » (TLF).

À partir de ces définitions de base, dont le point commun est de représenter essentiellement une personne sommairement en mettant en lumière ses caractéristiques les plus distinctives au détriment de celles qui sont plus communes, j'avance que la spécificité d'un profil est l'immédiateté de sa représentation, tout en restant fidèle à la réalité. Par dérivation de l'entrée concernant la description d'un type par les caractéristiques les plus représentatives des personnes ou des éléments concernés, de nombreuses applications de profils ont vu le jour dans les domaines tels que la psychologie, la médecine, les assurances, l'éducation, le travail, etc. (GDT). C'est prioritairement vers ces profils de personnes ou d'objets, que notre étude va s'orienter.

2.2 Étude générale de la notion de profil et profilage

En 2005 est parue une étude éditée par Hildebrandt et Backhouse pour le Consortium *Future of Identity in the Information Society (FIDIS)*. Vingt-quatre universités, instituts ou centres d'étude ont collaboré afin d'arriver à décrire, inventorier et analyser la mise en place de profils. Cette pratique du profilage est présentée comme un phénomène technologique puissant, crucial, mais, dans certains de ses aspects, inquiétant dans la mesure où il permet d'exploiter des données innombrables concernant le comportement d'individus ou de groupes, que les objectifs poursuivis soient infâmes ou salutaires (2005 : 1). De plus, selon cette étude, la sélection des éléments d'information pertinents semble plus importante que la récupération du plus grand nombre de données (2005 : 9). Nous prendrons les résultats de cette étude comme point de départ pour aborder la notion de « profil » et de « profilage ».

2.2.1 Profil et notions de base associées

2.2.1.1 Définition générale d'un profil

En opposition à une somme de données brutes, la définition d'un profil est un ensemble de données mises en correspondance. Alors que les données ne sont que des renseignements (des informations), les profils (lorsqu'ils sont interprétés) sont des constructions du savoir³⁰ (*knowledge constructs*) (2005 : 9). Ils représentent un sujet, qu'il soit une personne, une chose, une organisation ou une autre entité. Le savoir en question est constitué de *patterns of correlated data* (2005 : 13) s'appuyant le plus souvent sur des données glanées au fil du temps. Ce profil peut être défini comme « un ensemble de données mises en correspondance qui identifie et représente un sujet de données ».

2.2.1.2 Profil personnalisé et profil de groupe

Si le sujet de données se réfère à un seul individu, on utilisera la dénomination de *personalised profile*. Un profil personnalisé (*personalised*) se centrera sur les attributs d'un individu afin de l'identifier et le représenter d'après ses caractéristiques et son comportement dans le passé.

Par contre, s'il s'agit d'un groupe, d'une catégorie ou d'un ensemble, on parlera de « group profile ».

profile: set of correlated data that identifies and represents a data subject. If the data subject is a group/a category/or a cluster we speak of a group profiles, when the data subject is a single person we speak of a personalised profile (2005 : 106)

Le profil de groupe est aussi défini comme suit : *a group profile is a set of correlated data that identifies a group, and/or when applied identifies a person as a member of a group* (2005 : 16). Le profil de groupe sera une catégorisation se référant à une personne abstraite, une sorte de prototype, et non pas à des individus en particulier (2005 : 17). La mise en conformité par catégorisation s'effectuera par comparaison avec le prototype (2005 : 53).

2.2.1.3 Profil distributif et profil non distributif

Un profil distributif pose comme condition que chaque élément du groupe possède tous les attributs définissant le profil. Au contraire, dans les cas de profils non distributifs, les traits caractéristiques du groupe ne s'appliquent pas obligatoirement à tous les éléments de la catégorie.

À partir de ces notions, voyons comment s'organise l'opération de profilage.

³⁰ Ma traduction.

2.2.2 Opération de profilage

2.2.2.1 Définition

L'opération de profilage réussie, dans son sens premier, résulte en la création de profils pour identifier et classer un sujet de données. Dans un second temps, l'opération de profilage signifiera l'application des profils à des individus afin de les catégoriser. Elle nécessite l'adjonction de plusieurs notions inhérentes au concept de profilage : celles du sujet de données (personne concernée), du responsable du traitement et de l'utilisateur final (sujet visé), lesquelles sont définies ainsi :

data subject: the subject (human or non-human, individual or group) the data refer to (2005 : 105)

data controller: the subject (person or organisation) that determines the purposes and means of the processing of data (2005 : 105)

(end) user: the profiled data subject using a certain device (web, customer loyalty card) that facilitates the recording of data to be stored and processed for the data controller (2005 : 106)

profiling: the process of constructing profiles (correlated data), that identify and represent a data subject (either a person or a group/category/cluster), and/or the application of profiles (correlated data) to identify and represent a person as a specific person or as member of a specific group/category/cluster, aiming at the assessment of risks and/or opportunities for the data user (inferred from risks and opportunities concerning the data subject) (2005 : 106)

2.2.2.2 Processus et finalité du profilage

La finalité ultime des opérations de profilage est donc la sélection d'individus ou leur exclusion. Elle comporte quatre étapes, que je schématise dans le graphique ci-dessous.

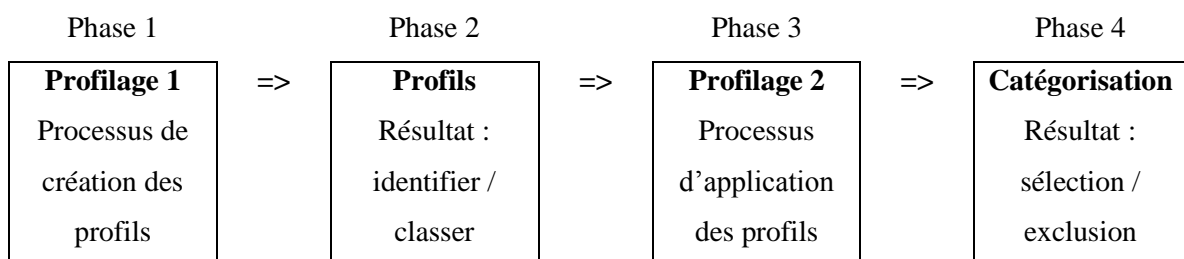


Tableau 19 – Schématisation des opérations de profilage d'après la définition de Hildebrandt et Backhouse (2005 : 106)

La finalité du processus dans sa globalité est similaire à notre objectif pour trouver des critères d'évaluation fiables en prosodie anglaise. Ce processus englobe des techniques, technologies et savoir-faire spécifiques, qu'il convient d'examiner avec attention.

2.2.2.3 *Méthodologie du profilage*

2.2.2.3.1 Présentation générale et avertissement

Les processus de création ou d'application de profils, qu'ils concernent un individu ou un groupe, mettent en jeu des techniques (méthodes), des technologies (association d'instruments tangibles et de techniques éprouvées, de matériel informatique et logiciels) ainsi qu'un savoir faire permettant de s'adapter à des contextes et objectifs spécifiques (2005 : 12). Hildebrandt et Backhouse (2005 : 13) insistent sur cette compétence personnelle et artisanale, cette spécialisation, adossée à une formation professionnelle adéquate, ne pouvant s'acquérir qu'au fil du temps et des expériences de terrain car elle nécessite des connaissances explicites et un savoir faire implicite. En aucun cas il ne s'agirait simplement que d'appliquer mécaniquement règles et méthodologies toutes prêtes.

2.2.2.3.2 Étapes du processus de profilage

Qu'il s'agisse de profils de groupe ou personnalisés, on trouve néanmoins des phases automatisées dans les étapes suivantes (2005 : 19) :

- recording data (taking note of them in a computable manner)
- storing data (in a way that makes them accessible, aggregated in a certain way)
- tracking data (recording and storing over a period of time, linking data to the same data subject)
- identifying patterns and trends in the data (by running algorithms through the data base) and
- monitoring data (checking whether new data fit the pattern or produce outliers).

Les données brutes (*raw data*) à prendre en compte doivent donc toutes avoir subi une conversion pour devenir des données utilisables informatiquement.

2.2.2.3.3 Profilage pour « profil personnalisé »

Un profil personnalisé est destiné à identifier et représenter un sujet de données individuel, c'est-à-dire une seule personne (Hildebrandt et Backhouse, 2005 : 32). Non axé sur des approches probabilistes, il s'appuiera au contraire sur des approches cognitives et fondées sur la connaissance et l'individualité de la personne pour répertorier les caractéristiques personnelles d'un individu afin de pouvoir répondre à des besoins

spécifiques. Les sources d'information fournissant les données pourront être diverses et variées.

2.2.2.3.4 Profilage de groupes

Lors de l'analyse d'un ensemble de données mises en corrélation qui identifient et représentent un regroupement, une catégorie ou un faisceau, une méthodologie novatrice se distingue de la recherche traditionnelle en sciences sociales : c'est le profilage de groupes.

KDD : extraction du savoir à partir de bases de données

Dans notre société numérique, un processus émergent est de plus en plus utilisé : le KDD, dont la définition est donnée :

KDD: knowledge discovery in data bases, comprising the whole process of recording, storing, tracking of data, including discovery of patterns by means of data mining techniques and interpretation of the correlated data (2005 : 105)

En fait, le but ultime n'est pas l'extraction de données mais l'extraction du sens, auparavant indétectable, qui pourra prendre une forme structurée à partir de masses de données informes et sans signification.

L'extraction de connaissances à partir de données (KDD) est surtout utilisée dans la création de profils de groupes. Elle se déroule en six étapes (2005 : 10) :

- (1) collection and recording of data;
- (2) aggregating or preparing of the data in databases;
- (3) data mining or running algorithms through the database;
- (4) interpretation of the emerging patterns;
- (5) decision taking on the basis of the resulting profiles, and
- (6) proliferation of these profiles in the relevant social contexts.

Modèle de profilage CRISP-DM

Parmi les modèles de profilage proposés, celui publié en 2000 par le consortium *Cross-Industry Standard Process for Data Mining (CRISP-DM)* et en partie financé par la *Commission européenne*, emporte de nombreux suffrages. Orienté au départ vers la sphère économique, le processus qu'il propose comme modèle se veut applicable à n'importe quel domaine utilisant une extraction de données. La figure ci-dessous représente les six phases du processus.

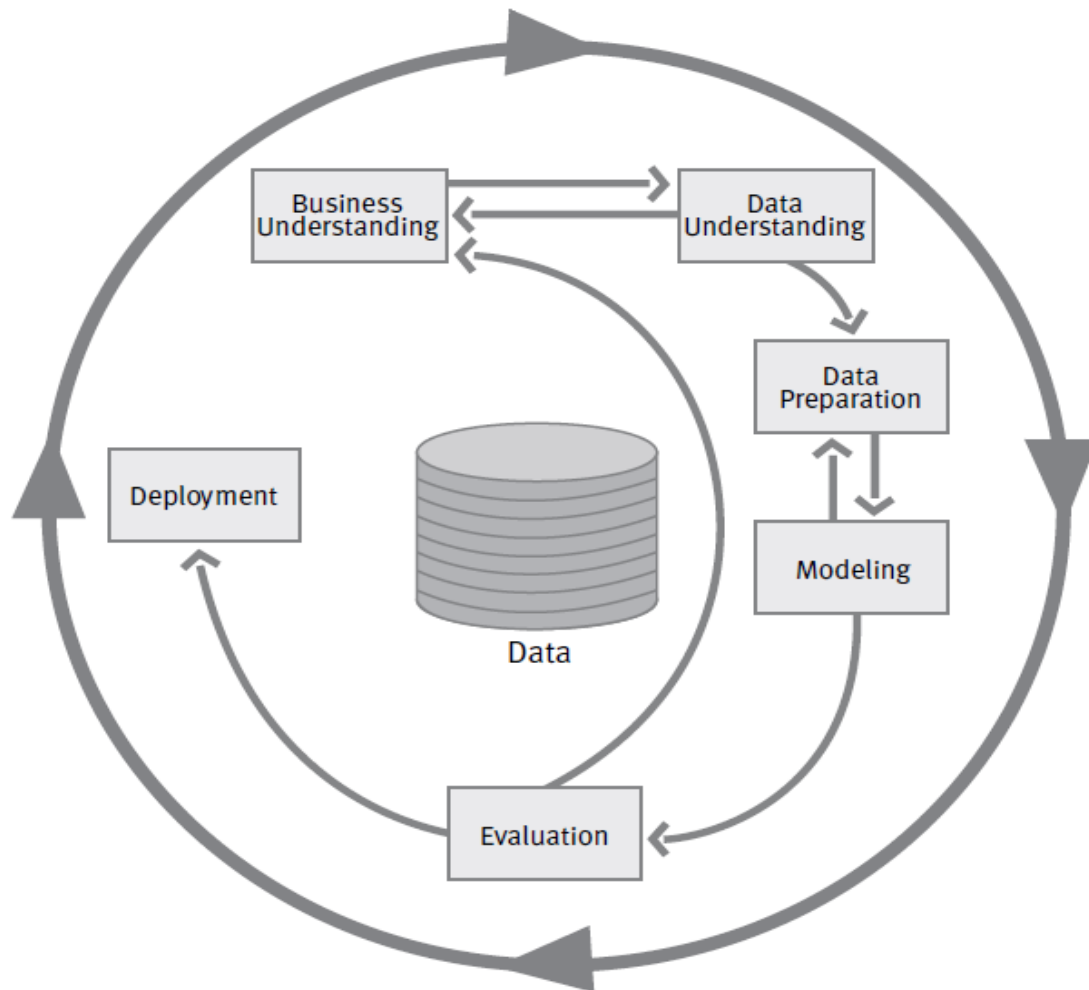


Figure 14 – Modèle *CRISP-DM 1.0* : Modalisation des phases du processus (Chapman *et al.*, ([1999] 2000))

Ce diagramme montre clairement les relations entre les différentes phases de ce modèle, ainsi que la nature réursive du processus, alimenté par ses propres résultats. Ces phases sont résumées ci-dessous à partir du document original : *CRISP-DM Step-by-step data mining guide* ([1999] 2000 : 10-11).

- ***Business understanding*** : Cette première phase est destinée à identifier les objectifs du projet et ses implications en termes de perspective commerciale, puis, à partir de cela, à définir le problème de l'exploitation des données et d'en dresser les grandes lignes d'action.
- ***Data understanding*** : Il s'agit ici de recueillir les données, de se familiariser à leur contact afin de repérer d'éventuelles déficiences, d'orienter l'approche et de repérer des sous-catégories sur lesquelles seront appliquées des hypothèses permettant de découvrir des tendances imperceptibles.

- **Data preparation** : Cette phase de préparation des données englobe toutes les activités destinées à l'élaboration de la base de données finale à partir des données brutes, afin de la saisir dans l'outil de modélisation. Les tâches de préparation des données sont susceptibles d'être réitérées de nombreuses fois et sans ordre prescrit. Elles comprennent la sélection des tableaux, des jeux de données, des valeurs caractéristiques, ainsi que la transformation et le nettoyage des données pour les rendre utilisables.
- **Modeling** : Lors de cette phase, on sélectionne puis applique diverses techniques de modélisation ; leurs paramètres sont calibrés de manière optimale. On trouve d'ordinaire plusieurs techniques susceptibles de résoudre un même type de problème pour la classification, à ce stade, de données. Certaines techniques nécessitant une configuration spécifique des données, la phase de préparation des données devra souvent être réétudiée.
- **Evaluation** : À ce stade du projet, on a construit un ou des modèles qui semblent pouvoir s'inscrire adéquatement dans une perspective d'analyse de données. Avant de procéder à la finalisation du modèle, ce dernier doit subir une évaluation minutieuse et une vérification des étapes de sa création afin de s'assurer de l'adéquation du modèle aux objectifs assignés. Un objectif fondamental est de vérifier que tous les problèmes déterminants aient trouvé une solution adaptée. La fin de cette phase doit être marquée par une prise de décision sur l'utilisation qui sera faite des résultats de l'extraction.
- **Deployment** : La création d'un modèle signifie rarement l'achèvement du projet. Même si l'objectif poursuivi par la création du modèle est de mieux comprendre les données, l'information obtenue devra être organisée et présentée de manière à ce que le client puisse l'utiliser. Cela implique souvent l'application de modèles faits « d'après nature » au sein des processus de prise de décision d'une organisation (par exemple, la personnalisation de pages sur la Toile en temps réel, ou l'évaluation en continu des bases de données de commercialisation). Selon les besoins, la phase d'exploitation sera aussi simple que de rédiger un rapport, ou aussi complexe que de mettre en place un processus d'exploration des données répétitif à l'échelle de l'entreprise. Dans de nombreux cas, ce sera le client, et non l'analyste de données, qui procédera aux étapes de généralisation. Cependant, même si l'analyste se montre en

mesure de s'acquitter de cette tâche d'exploitation, il est important que le client comprenne auparavant quelles opérations sont impliquées par l'utilisation des modèles créés.

Schématiquement, ces différentes phases peuvent être représentées hiérarchiquement, ainsi que le montre le graphique ci-dessous. Un ensemble de tâches se décompose en quatre niveaux d'abstraction, du plus général au plus spécifique : grandes phases, tâches génériques, tâches spécialisées, procédures spécifiques. Une mise en correspondance articule les niveaux 1 et 2 du modèle CRISP avec les niveaux 3 et 4 ciblant les contraintes techniques d'accès aux données. Alors que le niveau *generic tasks* est général afin de couvrir toutes les situations possibles d'extractions de données et de rester ouvert à des techniques modélisatrices nouvelles, le troisième niveau *specialized tasks* bascule vers la réalisation concrète des tâches génériques dans des situations spécifiques : il décrit comment les tâches de deuxième niveau doivent être exécutées selon la situation. Le quatrième niveau *process instances* décrit précisément ce qui se passe sur le terrain.

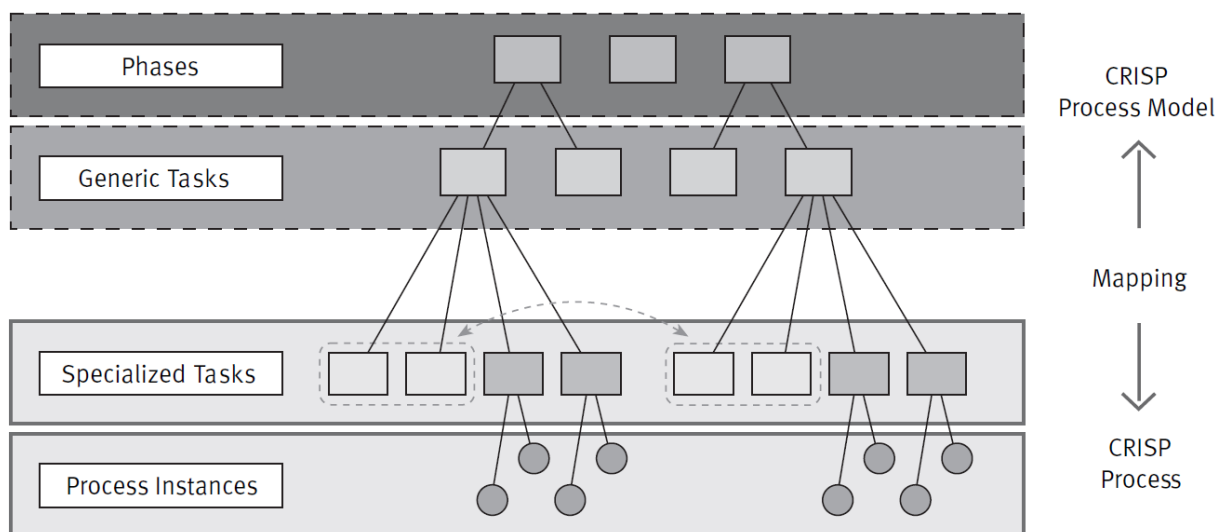


Figure 15 – Décomposition en quatre niveaux selon la méthodologie *CRISP-DM* (Chapman *et al.*, ([1999] 2000 : 6)

L'aboutissement de l'application du modèle de référence *CRISP-DM* doit pouvoir se traduire selon la formule suivante (Chapman *et al.*, ([1999] 2000 : 51) :

$$\text{RESULTS} = \text{MODELS} + \text{FINDINGS}$$

Profilage par analyse ascendante ou analyse descendante

La fouille des données (*data mining*) est un champ pluridisciplinaire fortement enraciné dans l'analyse quantitative, en quête de nouvelles configurations et corrélations au

sein de bases de données pré-existantes³¹. Elle utilise deux approches : l'analyse ascendante et l'analyse descendante (Hildebrandt et Backhouse, 2005 : 26-27).

L'utilisation de l'analyse descendante se justifie pour tester des hypothèses, pour appliquer des schémas sur une base de données, ou pour identifier des comportements dérogeant à la norme. On l'appelle *On-Line Analytical Processing (OLAP)*.

L'analyse ascendante, au contraire, plutôt que de vérifier des configurations hypothétiques, recherche ces schémas par procédure automatique de manipulation des données. Ce procédé essentiellement inductif est destiné à générer des hypothèses susceptibles d'expliquer les comportements observés ou de prédire les comportements futurs. Ce procédé ascendant peut être orienté et, dans ce cas, traduit l'existence d'une hypothèse, qu'elle soit intuitive ou plus réfléchie.

Profilage dans les profils distributifs ou non-distributifs

Les techniques de profilage de groupe par extraction de données sont appelées à revêtir une importance sans limite pour s'appliquer à des groupes de personnes et non plus à des individus isolés, prenant une extension généralisatrice sans précédent. Pour cette raison, la distinction entre profils distributifs et profils non-distributifs, dont Hildebrandt et Backhouse (2005 : 30) donnent les définitions suivantes, sera amenée à jouer un rôle prépondérant dans la nature même des profils créés, la mise en relation des divers profils, ainsi que les opérations de profilage qui en découlent :

Distributive profiles assign certain properties to a data or information subject, consisting of a group of persons however defined, in such a way that these properties are actually and unconditionally manifested by all members of that group. Distributive generalisations and profiles are phrased in the form of down-to-earth, matter-of-fact statements. *Non-distributive* profiles are framed in terms of probabilities, averages and medians, significant deviancies from other groups, etc. They are based on comparisons of members of the group with each other and/or on comparisons of one particular group with other groups. *Non-distributive* profiles are, therefore, significantly different from distributive profiles.

Dans un profil distributif, tous les éléments du groupe sont dotés des mêmes caractéristiques définitoires. Le seul fait de ne pas correspondre à un seul des critères exclut l'élément du

³¹ *Data mining can be defined as data processing, using sophisticated data search capabilities and statistical algorithms to discover patterns and correlations in large pre-existing databases; a way to discover new meaning in data. Data mining is a multidisciplinary field with strong quantitative roots.* (Hildebrandt et Backhouse (2005 : 26).

groupe. Aussi, le caractère des conditions définitoires est très clair et tranché dans la manière dont il est présenté pour ne laisser aucune place à l'approximation. Inversement, un profil non distributif représente une tendance à laquelle sont sujets un certain nombre d'éléments du groupe. L'appartenance à un groupe non distributif signifie qu'il existe une certaine probabilité pour qu'un certain nombre d'éléments de ce groupe possède telle ou telle caractéristique. Un profil non-distributif fonctionne selon le principe de l'inclusion alors qu'un profil distributif fonctionne selon le principe de l'exclusion.

2.3 Bilan et discussion

On a vu que le processus de profilage se décompose en étapes et que les données recueillies doivent être catégorisées pour faire apparaître des corrélations intéressantes, trouver les algorithmes qui localiseront les configurations judicieuses, comprendre sens et tendance de ces configurations naissantes et, en connaissance de cause, agir de manière optimale tout en ayant à l'esprit les répercussions à prévoir (Hildebrandt et Backhouse, 2005 : 10).

La tendance exacerbée à l'utilisation exponentielle de données numériques ne doit pas faire oublier une méthodologie « faite main » qui a permis de tout temps à des organisateurs de repérer et d'identifier les candidats à leur sélection. Les études mentionnées soulignent le nécessaire travail artisanal (*handiwork*), toujours nécessaire de nos jours, malgré le progrès des technologies numériques et ce, pour deux raisons :

- Le profilage peut, dans certains cas, être effectué sur la base de peu de données quantifiables,
- Quelle que soit la quantité de données, des étapes artisanales sont nécessaires dans l'élaboration des algorithmes et l'évaluation des résultats.

Ainsi que le décrivent Hildebrandt et Backhouse (2005 : 10), ce travail relève de la combinaison d'une expérience des phénomènes en jeu pertinente, d'un scepticisme de bon aloi, et d'une compétence intuitive éprouvée.

Le terme de « profil » est de nos jours très courant et est appelé à le devenir encore davantage. Ainsi, l'appartenance à un profil est un signe de reconnaissance, une détermination, une spécification grâce à des signes communs ou distinctifs. Les mentions précédentes balisent la suite de cette étude sur la nature et le fonctionnement des deux sortes de profils que nous utiliserons : profils personnalisés et profils de groupe. Tout profil est basé

sur l'existence de données. C'est à partir de leur interprétation que pourront être construits ces profils. Ils pourront alors remplir leur rôle qui est d'inclure ou d'exclure des locuteurs en fonction des objectifs à atteindre, lequel est de cerner les formes les plus représentatives sans exclure celles qui, malgré leur marginalité, sont néanmoins acceptables pour modéliser des critères évaluatifs.

Je propose celle définition des profils prosodiques, que j'ai eu l'occasion de défendre ailleurs³²:

Nous définissons un profil comme un faisceau cohérent de réalisations de traits pertinents susceptibles de caractériser et catégoriser certains locuteurs natifs ou non natifs. Un profil prosodique se définira par divers corrélats acoustiques prépondérants et récurrents dans les réalisations, catégorisant les locuteurs dans une typologie en vue de les organiser.

D'aucuns les appelleraient « phonostyles » au sens large du terme, avoisinant parfois l'idiosyncrasie.

3 Profils personnalisés : centrage sur l'individu en vue d'une identification ou d'une remédiation

Les profils personnalisés et leur profilage corollaire sont utilisés dans de nombreux domaines proches de notre étude : les secteurs médico-légal, linguistique et plus particulièrement prosodique. Nous allons ainsi en examiner quelques particularités parce qu'ils sont à la source de la notion de profil.

3.1 Profilage médico-légal (*offender profiling* et *forensic profiling*)

Le profilage médico-légal se rattache à cette étude parce qu'il fournit un éclairage particulier mais pertinent à une application de la phonétique peu valorisée dans le monde de la recherche phonétique. Après avoir présenté sa nature et son fonctionnement général, je ciblerai l'utilisation qui est faite des profils phonétiques.

3.1.1 Nature et fonctionnement du profilage médico-légal

³² Cauvin (2014). « Plaidoyer pour une politique linguistique LANSAD intégrant la phonologie anglaise ». Dans René Daval, Émilie Hilgert, Thomas Nicklas, Daniel Thomières (dirs.), *Mélanges en hommage à Pierre Frath*. Reims : Éditions et Presses universitaires de Reims.

Les cas de profilage dans le domaine médico-légal sont innombrables, un des plus connus étant sans doute le cas de Jack l'Éventreur, le tueur en série qui sévit en 1888 à Londres. Il s'agit là d'un profil personnalisé, où toutes les sources possibles d'information ont été privilégiées afin d'opérer un recoupement de données (*data matching*).

De nos jours, le terme de « profil » est surtout utilisé en psychologie et en criminologie dans le domaine de l'analyse comportementale. L'analyste comportemental, aussi appelé psycho(logue)-criminologue, est chargé de réaliser l'ébauche du type de portrait psychologique d'une personne recherchée, de décrire les caractéristiques de son comportement et de ses motivations, afin de l'aider par des conseils, d'orienter son comportement contrairement à ses intérêts propres, ou de l'empêcher de nuire. Dans tous les cas, se vérifie ce que rappelle Jean-Daubias (2011 : 32) : « le terme de profil fait référence aux informations concernant un individu donné dans un contexte donné ».

Quant au processus mis en œuvre, l'étude conduite par Geradts et Sommer (2008 : 13), s'appuyant sur celle de Hildebrandt et Backhouse (2005), définit la collocation *forensic profiling* comme *the exploitation of traces in order to draw profiles that must be relevant to the context of supporting various security tasks, mostly in the criminal justice system*.

Cette définition inscrit ce type de profilage dans une démarche ascendante étant donné que l'analyste aura comme point de départ des données plus ou moins éparses. Sa mission sera de les organiser pour leur donner un sens. Afin d'englober les pratiques anciennes (*data matching*) et modernes (*data mining*), Bygrave (2002), cité dans Geradts et Sommer (2008 : 39), précise l'inférence de la caractérisation d'un ou plusieurs individus, qui seront alors traités à la lumière de cette caractérisation. L'exploitation des indices relevés permet de dresser un profil personnalisé, lequel est ensuite mis en correspondance avec des profils de groupe ou d'immenses bases de données. On trouve néanmoins la définition d'*offender profiling* et ses équivalents *criminal profiling*, *criminal personality profiling*, *criminological profiling*, *behavioral profiling* ou *criminal investigative analysis* pour désigner des pratiques comportementales et d'investigation destinées à aider des enquêteurs à prédire et décrire avec précision les traits caractéristiques de criminels non identifiés en dressant leur profil³³. La notion de « profilage criminel » se rapportera davantage à la mise en correspondance de données. Dans la sphère criminelle, le décryptage prosodique des suspects lors des

³³ [http://en.wikipedia.org/wiki/Offender_profiling], consulté en 2016.

interrogatoires avec la police peut fournir des éléments à charge contre eux ou non (Fadden, 2006).

3.1.2 Profils phonétiques médico-légaux (*forensic phonetics*)

Selon Jessen (2010 : 378-9), la phonétique médico-légale porte les noms de *Forensic Phonetics and Acoustics* ou *Forensic Speech and Audio Analysis*, ou même *Forensic Speech Science*. L'identification judiciaire d'un locuteur se décompose en plusieurs sous-tâches. Cependant, un éventuel profilage du locuteur aura lieu sous deux conditions : (1) un enregistrement audio du criminel ou contrevenant est disponible, (2) il n'existe aucun suspect. On l'appellera *speaker profile*, *voice analysis* ou *voice profiling*. L'application de cette méthodologie pour trouver un suspect s'orientera vers la détermination de l'âge, du sexe, de la provenance géographique, de la condition sociale et d'un éventuel accent étranger pour dresser un profil.

3.1.2.1 Méthodologie de profilage

Jessen (2010 : 381), distingue deux aspects du domaine de profilage judiciaire :

- la tâche de classification d'un locuteur (*speaker classification*) :

Speaker classification (see Müller 2007 for the full range of that term) can be understood as the task of inferring from speech evidence the “class” or “category” to which a speaker belongs. As mentioned above, this includes region, age, sex, and social background.

- la tâche d'identifier les traits caractéristiques les plus saillants d'un locuteur (*speaker-specific characteristics*) :

But a speech sample might also contain other features that are striking and noteworthy speaker characteristics in the perception and categorisation of a layperson, although these features might have nothing to do with speaker classification. For example, a voice might be very high- or low-pitched, the speaker might speak in a very fast or slow manner or in a very careful or sloppy manner; there might be creaky, breathy, harsh or nasal voice quality and so forth.

Cette dichotomie permet de dissocier les destinataires du profil. Les policiers, ainsi qu'un large public, après avoir écouté la voix, seront en mesure de pouvoir proposer des candidats au statut de suspect, l'audition de la voix étant plus « parlante » qu'une classification abstraite pour des non spécialistes. Une autre tâche pertinente dans une optique de profilage est le repérage d'indices permettant de détecter les états transitoires d'un locuteur, indiquant un

3. Profils personnalisés : centrage sur l'individu en vue d'une identification ou d'une remédiation

CHAPITRE II – Profils et profilage : mode d’organisation hiérarchique de la prosodie des anglophones natifs

stress, une émotion ou même la consommation de drogue ou d’alcool. Ainsi, les traits répertoriés dans la classification des locuteurs se retrouvent chez un groupe de locuteurs, alors qu’une voix aiguë ou un tempo rapide relèvent de caractéristiques individuelles. La vitesse d’élocution et la hauteur mélodique ont toute leur importance dans la discrimination de locuteurs.

La figure suivante montre la classification des domaines pris en compte dans des affaires ou susceptibles de l’être plus encore à l’avenir (Jessen, 2010 : 381).

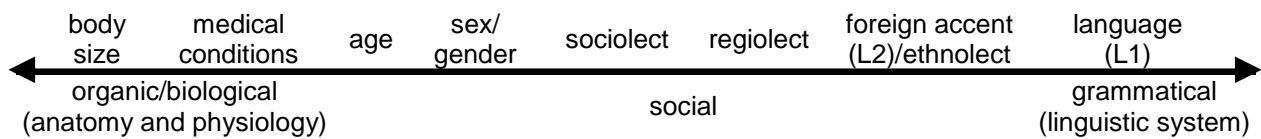


Figure 16 – Domaines de classification de locuteur judiciaire (d’après Jessen, 2010 : 381)

Il rapporte des corrélations positives entre la taille, et dans une moindre mesure le poids des locuteurs et la longueur de leur larynx. Chez les locuteurs de taille extrême, des études ont constaté que les formants de voyelles très bas correspondaient le plus souvent à des personnes de grande taille alors que les formants très hauts étaient ceux de locuteurs de petite taille (Jessen, 2010 : 382) :

Although the correlation is very weak in the mid-range of formant frequencies and body height measures, there are clear patterns at the extremes. Speakers with formants in the high range above 2500 Hz have small or medium body size, whereas speakers with formants in the low range below 2200 Hz have above-average body size. In addition, the results for the average second formant (not shown) are similar to those of the third formant.

Pour ce qui est de l’âge, Jessen (2010 : 382) rapporte que l’information disponible concerne essentiellement l’influence de l’âge sur la fréquence fondamentale, mais que la vitesse d’élocution et l’intensité peuvent aussi jouer un rôle non négligeable dans la détection de l’âge :

For male adults there is a decrease in f_0 up to about the age of 40, then from about 50 an increasing pattern that continues into old age. For female adults there is a gradual decrease from young adulthood on, sometimes accelerated during menopause; there might be a slight increase again at old or very old age (Baken and Orlikoff 2000: 173–76; Schötz 2006: 83). In a recent large-scale study, Schötz confirmed this pattern and also found that speech rate is a good age correlate,

3. Profils personnalisés : centrage sur l’individu en vue d’une identification ou d’une remédiation

because it decreases gradually for both males and females (Schötz 2006: 110). She found another salient correlate – increasing intensity range – which requires further research.

Dans le domaine médico-légal, la classification des locuteurs selon leur âge relève de la perception auditive avant tout. Les erreurs d'appréciation excèdent rarement un écart de six ans et les experts en la matière ont un jugement légèrement meilleur que les amateurs. Somme toute, les corrélats acoustiques dépendent de paramètres biologiques, sociaux (genre et statut social), ainsi que du système linguistique lorsqu'il traduit une modification de la prononciation (Jessen, 2010 : 383).

Spécialistes ou novices devinent facilement le sexe du locuteur, c'est-à-dire sa détermination biologique, la moyenne de la fréquence fondamentale d'une voix de femme étant bien plus élevée que celle d'un homme, du moins avant que la vieillesse ne s'installe. Jessen (2010 : 383-384) justifie leur position par la différence de la longueur des plis vocaux (Vaissière 2006 : 48-55), qui sont 60% plus longues chez les hommes que chez les femmes d'après Titze (1994: 173). Selon lui, la distribution des valeurs de F0 masculines et féminines dans un échantillon représentatif de la population est plus significative encore (2010 : 383-384) :

The range of mean f0 from the lowest-pitched to the highest-pitched voices is from about 80 to 170 Hz in men (Künzel 1989; Jessen *et al.* 2005) and from about 165 to 260 Hz in women (Künzel 1989; Simpson and Ericsson 2007). This information about f0 distribution can prove useful in situations where perceptual sex classification is difficult because the pitch of the unknown voice is too low for a typical female and too high for the typical male voice.

Une restriction pèserait sur un discours émotionnel, lequel montre des intersections entre les deux distributions, car les expérimentations citées concernent la parole de laboratoire.

Le chercheur mentionne ensuite les cas où la hauteur mélodique ne peut être utilisée afin de repérer le sexe du locuteur, lorsque la voix est déguisée en utilisant une voix craquée par exemple, ou lorsque le discours est dévoisé, comme par un chuchotement. Dans ces cas, d'autres stratégies permettent de décrypter les caractéristiques du locuteur, notamment en déduisant la longueur du larynx par des mesures des fréquences formantiques F1 et F3, à moins qu'il ne s'agisse d'enregistrements téléphoniques, dans lesquels F3 n'est souvent pas accessible.

Le genre joue un rôle non négligeable dans la détermination biologique car les facteurs socio-culturels peuvent augmenter ou au contraire minimiser les caractéristiques sexuelles de la parole. Jessen mentionne ainsi l’orientation de l’influence sexuelle, qui conduira certains locuteurs à adopter les caractéristiques propres à la voix d’un sexe déterminé auquel ils s’identifient, même s’il n’est pas leur sexe biologique. Jessen rapporte aussi (2010 : 384-385) que *both male and female speakers have a high average f0 in loud speech and can be difficult to distinguish* (2010 : 384). L’arrondissement des lèvres peut aussi déguiser une voix en faisant baisser les résonnances : une voix de femme peut ainsi être prise pour une voix d’homme.

Ces quelques différences sont universelles. En fonction des diverses cultures, des particularités se font jour dans telle ou telle langue. Des études ciblées et poussées ont été faites sur la classification juridique des locuteurs

3.1.2.2 Comparaison de locuteurs

La comparaison de locuteurs s’appuie sur les corrélats acoustiques de voix réelles (*tokens*) tandis que les opérations de profilage concernent des archétypes (*types*). Les tâches y afférant en tiennent compte, ainsi que le montre le graphique de Jessen (2010 : 386).

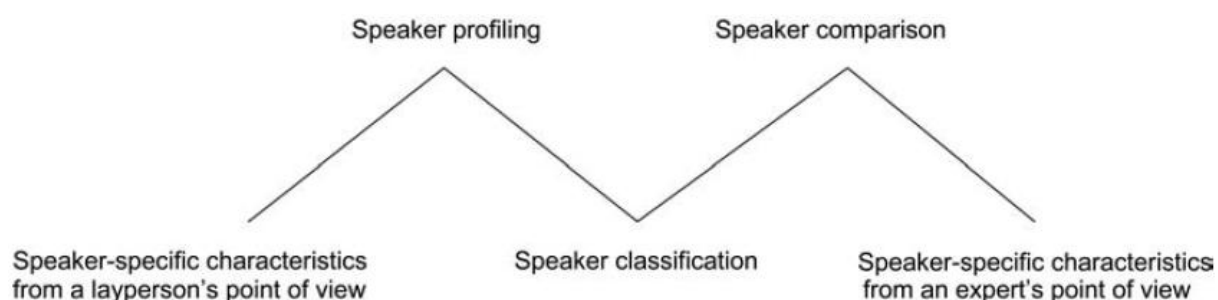


Figure 17 – Tâches spécifiques au profilage de locuteurs et à la comparaison de locuteurs (Jessen, 2010 : 386).

Les auditeurs non spécialistes procéderont à une classification dont les résultats en termes de sélection ou exclusion par catégorisation sera vérifiée par comparaison avec les données acoustiques que seul un phonéticien peut analyser et apprécier selon leur nature. Le graphique ci-dessous classe ces corrélats en trois grandes catégories à cet effet.

<i>Organic</i>	<i>Idiolectal</i>	<i>Habitual</i>
1. average fundamental frequency (f_0)	4. individual aspects of sociolect, regiolect and foreign accent	7. articulation rate
2. (long-term) formant frequencies	5. linguistic-phonetic details	8. f_0 -variability
3. voice quality	6. forensic-linguistic features	9. dysfluent behaviour

Tableau 20 – Caractéristiques de détermination de locuteurs (Jessen, 2010 : 387)

3. Profils personnalisés : centrage sur l’individu en vue d’une identification ou d’une remédiation

L'augmentation de l'intensité est corrélée à celle de la F0 (Jessen, 2010 : 387). Il arrive fréquemment que l'intensité joue un rôle prépondérant dans les analyses médico-légales de la voix. On parlera alors de variation intra-locuteur (*intra-speaker variation*) (2010 : 388). Il semblerait que les moyennes de F0 aient aussi tendance à varier selon les communautés linguistiques.

Jessen (2010 : 391) avance que la qualité de la voix peut rarement être mesurée selon les critères de Laver (1980), même pris au sens large (*including both laryngeal characteristics such as creaky, breathy, harsh, and pressed voice and supralaryngeal settings such as (open and closed) nasality, lip rounding and jaw lowering*) étant donné que la plupart des enregistrements proviennent de conversations téléphoniques. La perception auditive est donc primordiale.

L'idiolecte peut aussi permettre d'identifier un individu. Jessen (2010 : 391-392) l'identifie de trois manières : 1) un locuteur combine les caractéristiques de plusieurs variétés régionales, sociales et multilingues, 2) caractéristiques phono-linguistiques non uniquement attribuées à des dialectes ou des variétés de langues, 3) davantage linguistique par ses aspects idiosyncratiques de la syntaxe et du lexique (Coulthard 2004).

3.1.3 Bilan et discussion

Cette section a souligné l'importance que pouvaient revêtir les corrélats acoustiques dans le repérage de traits permanents ou d'états transitoires, de la discrimination en caractéristiques individuelles ou se rapportant à un groupe de locuteurs. Que les corrélats acoustiques se rapportent à des paramètres biologiques, sociaux, ou au système linguistique, ce panorama a montré le potentiel de l'analyse acoustique pour dresser des profils personnalisés, mais aussi ses limites lorsque le signal sonore était de qualité réduite, ce qui est souvent le cas dans les enquêtes médico-légales. Cette étude a aussi montré la nécessité de recourir aux tests perceptifs, non seulement pour pallier ces manques, mais aussi afin de procéder à une première catégorisation des données, lesquelles sont analysées plus scientifiquement grâce aux techniciens spécialistes. Il a mis en valeur les tests perceptifs dans le profilage personnalisé et leur complémentarité avec l'analyse acoustique.

Les critères utilisés en profilage personnalisé dans le domaine médico-légal nous montrent en outre qu'il faut instaurer une hiérarchie entre les traits permanents et transitoires

car des critères d'évaluation doivent se baser sur les données les plus fiables, que sont les traits permanents. De même, les caractéristiques individuelles, biologiques et sociales sont à exclure au profit de celles qui permettent un classement des locuteurs en catégories relevant du niveau de maîtrise langagière.

En conclusion, un profil médico-légal se concentre sur l'individualité de la personne, les traits linguistiques et extralinguistiques qui la différencient du groupe, d'autres individus. L'essence de cette notion est reprise par Crystal dans son étude sur les profils linguistiques.

3.2 Approche pionnière des profils linguistiques (Crystal 1982)

Dans son ouvrage *Profiling Linguistic Disability*, Crystal ([1982] 1992) se penche méthodiquement sur la notion de profil linguistique. Nous reprenons ici ses arguments pour définir sa notion de profil linguistique, préciser sa finalité et les éléments dont il est constitué, rendre compte du traitement appliqué aux données, ainsi que du processus d'élaboration qui conduit à sa création.

3.2.1 Définition

Crystal ([1982] 1992 : 1) définit ainsi un profil linguistique :

A linguistic profile is a principled description of just those features of a person's (or group's) use of language which will enable him to be identified for a specific purpose. Profiles could be constructed for any area of linguistic inquiry, such as the study of literary style, the investigation of disputed authorship, or the analysis of achievement in foreign-language or mother-tongue learning.

Il associe donc la notion de profil linguistique aux choix linguistiques caractérisant une personne dans sa manière de s'exprimer. Ces choix, qu'ils soient conscients ou inconscients, permettront un repérage des traits les plus saillants et une classification destinée à diagnostiquer les écarts par rapport à une norme. Sa définition précise que les individus étudiés sont des personnes et elle n'exclut pas un groupe de personnes, formé alors par les individus étudiés.

3.2.2 Finalité

Crystal ([1982] 1992 : 1) précise la raison d'être d'un profil linguistique : *The primary purpose of profile-construction is to enable an accurate assessment of P's disability to be made, sufficient to provide a basis for remedial intervention.* Qu'il soit patient (P pour *patient*) ou élève (P pour *pupil*), Crystal explique que le but à atteindre est de diagnostiquer la

nature de la pathologie ou du trouble que des examens futurs confirmeront ou infirmeront. Pour cela, deux objectifs primordiaux doivent être atteints ([1982] 1992 : 2) :

(a) to identify the linguistic level P has achieved, in relation to the level he should be achieving;

(b) to suggest a remedial path, which will take him from where he is, to where he ought to be.

Un profil linguistique se situera donc à la croisée des domaines pratiques et théoriques. Pour remédier aux troubles dont souffre un individu spécifique, il sera nécessaire de répondre aux exigences de la recherche théorique ([1982] 1992 : 2) : *Linguistic profiles provide one method of bridging the gap between the demands of theory and the exigencies of practice.*

3.2.3 Spécificité des profils

Crystal met en garde contre l'assimilation d'un profil à un test. Ils les oppose afin de mieux faire distinguer leur caractéristiques respectives ([1982] 1992 : 2-3) :

- Un test est destiné à évaluer le terrain par des questions ou tâches standardisées. Bien que facile à administrer, il peut ne pas identifier des points importants, alors qu'un profil se doit d'être avant tout global dans son approche des domaines linguistiques étudiés.
- En outre, un test n'a pas pour objectif la remédiation, alors qu'un profil se doit de montrer la voie vers une amélioration possible étayée systématiquement et justifiée théoriquement.
- Un test fournit un résultat souvent chiffré qui se veut objectif ; un profil est de nature plus subjective.

Le rôle spécifique d'un profil est de produire un diagnostic systématique et ponctuel qui permettra à l'enseignant ou au thérapeute d'agir en connaissance de cause, mais ne se substituera pas à lui : *Linguistic profiles are only concerned with the 'what' and the 'why' of language remediation, but as such they give no guidance as to the 'how'.* ([1982] 1992 : 3).

En outre, un profil ne décrit pas la langue et n'a pas de validité universelle. Son cadre descriptif est homogène. Les contraintes dues à sa création ne correspondent pas aux procédures habituellement suivies en linguistique malgré quelques similitudes avec la linguistique appliquée. Un profil n'est que l'approximation d'une description linguistique, sa

seule justification résidant dans son objectif clinique, et il ne prend sens qu'après qu'une description linguistique ait été faite.

3.2.4 Éléments constitutifs des profils

Crystal ([1982] 1992 : 4) insiste sur les trois dimensions nécessaires à la construction d'un profil et regrette leur absence dans les études menées alors (T signifiant l'enseignant ou thérapeute) :

Ideally, any profile should contain three dimensions: it should provide a comprehensive description of P's data; it should provide a principled grading of the data; and it should show the influences operating on P, as he interacts with T, the clinical setting, the clinical materials and so on.

Selon lui, un profil devrait analyser exhaustivement les données signalant le trouble linguistique dans la limite du corpus observable. Il est cependant délicat de sélectionner par avance les aspects qui se révéleront *a posteriori* plus riches en signification ou d'éliminer ceux qui ne présenteront que peu d'intérêt. Crystal, conscient de l'aspect chronophage de la procédure d'investigation, propose quelques pistes pragmatiques.

D'un point de vue descriptif, Crystal rend compte d'un profil par un tableau où figure une série de catégories organisées plaçant en évidence les contrastes structuraux attendus dans la langue, qu'ils soient sons, structures grammaticales, éléments lexicaux, etc. La difficulté réside dans le choix et le classement des rubriques afin d'obtenir des résultats d'analyse utiles. Le processus de sélection repose sur un objectif clairement défini et de nombreuses tentatives pour s'en rapprocher. Un nombre de variables trop élevé ou trop restreint est de nature à rendre les profils obtenus embrouillés, ambigus ou déroutants, sans compter un suivi longitudinal rendu peu probant. Selon lui, toutes les catégories devraient avoir leur pertinence et figurer dans le tableau au départ, les plus révélatrices bénéficiant d'une place de choix.

Un profil étant par définition une approximation, il peut être judicieux d'opérer une investigation poussée sur une partie du tableau, c'est ce que Crystal ([1982] 1992 : 6) nomme « micro-profil » :

A micro-profile is a closer look at an area of the chart felt to be of particular clinical significance for an individual P, using the same general procedures as were used to construct the chart as a whole.

La localisation ainsi que les mesures effectuées doivent être soigneusement étudiées préalablement afin d'éviter une extrapolation des résultats ne correspondant pas à l'ensemble du fichier et biaisant ainsi l'analyse.

3.2.5 Traitement des données

Les données obtenues à partir d'un échantillon doivent être calibrées en prenant comme repère la norme des adultes ([1982] 1992 : 7). Une graduation rend les évaluations possibles et ouvre la voie vers une remédiation. Crystal précise néanmoins que ce sont les données qualitatives et non quantitatives qui sont de loin les plus utiles, mais que ces dernières peuvent être quantifiées avec profit, la plus utile étant l'ordre d'émergence des catégories : *The most illuminating approach, in terms of providing detailed assessments and remedial goals, is acquisitional in character*. Crystal aimerait donner à tous les profils une dimension acquisitionnelle. Même s'il reconnaît que les premières étapes du développement langagier (dont certains aspects de l'intonation) sont propices à fournir une perspective normative utile sur le plan clinique, ce n'est pas le cas des profils ([1982] 1992 : 7) :

By contrast, profiles are elementary descriptive devices, based on a synthesis of the empirical findings of the research literature. Where possible, categories are grouped into discrete stages of development, and associated with chronological age—but this is not always feasible.

Inscrire un profil dans une perspective acquisitionnelle permet d'évaluer les données par rapport à une norme chronologique, ce qui met en lumière tout écart par rapport à la norme des catégories linguistiques en ce qui concerne leur nombre, leur portée et leur ordre d'apparition.

Crystal met en garde contre une insuffisance des données sociolinguistiques répertoriées en ce qui concerne la situation d'interaction et les conditions de communication. Il préconise une prise en compte effective des facteurs extralinguistiques et des schémas d'interaction lors de l'élaboration des profils.

De manière similaire, des notes précises doivent rendre compte de la situation d'échantillonnage et de la stratégie d'échantillonnage mise en œuvre. L'objectif à atteindre étant l'émergence d'une tendance à partir des données, il recommande de spécifier les caractéristiques de l'échantillon dans le tableau du profil et de signaler le degré de représentativité de l'échantillon ([1982] 1992 : 9).

Crystal conseille une grande circonspection quant à la dichotomie compétence-performance. Les profils ne font que résumer la performance d'un élève ou patient en indiquant son degré de maîtrise d'une catégorie linguistique. Ils ne donnent pas d'indication sur sa compétence, qu'elle soit en compréhension ou en production.

3.2.6 Processus d'élaboration d'un profil

Crystal ([1982] 1992 : 9) préconise six étapes dans l'élaboration d'un profil :

- (i) a sample of P data is obtained;
- (ii) the sample is transcribed;
- (iii) the transcription is analysed;
- (iv) the analysis is profiled on a summary chart;
- (v) the pattern on the profile chart is assessed;
- (vi) the profile pattern is given an interpretation in remedial terms.

Selon lui, la transcription est une étape délicate : elle nécessite qu'un résumé de la production de l'élève ou patient figure dans le tableau sous forme catégorielle accompagné de statistiques afin de faire ressortir des tendances imperceptibles jusqu'alors. Cette transcription se doit d'être fiable, même si elle représente la phase la plus chronophage du travail de l'analyste dans l'élaboration du profil ([1982] 1992 : 10) :

The data transcription, whether phonological, grammatical or semantic, is likely to be the most time-consuming part of the whole profile exercise, but it must not be skimmed—the validity of a profile depends entirely upon its accuracy. A small amount of data well transcribed is far more valuable than a casual transcription of a large corpus.

3.2.7 Bilan et discussion

En conclusion, Crystal souligne ([1982] 1992 : 12) l'avancée résultant de la méthodologie propre à l'élaboration de profils et sa finalité (T représente l'enseignant-*teacher* ou le thérapeute-*therapist*) :

[...] the essential strength of a profile procedure is the way in which it makes T look at P's language in a systematic and detailed way, and organizes the data in such a way that clinically relevant patterns emerge.

Il met ainsi en relief les trois points fondamentaux sur lesquels s'articule la notion même de profil linguistique :

- L'étude systématique et détaillée des données sélectionnées,
- La création de catégories propres à classer les données obtenues,
- La création de tendances modélisables à partir de ce classement.

Il minimise néanmoins cette avancée du point de vue linguistique, pourtant essentielle sur le plan clinique, et laisse entrevoir le pendant du « micro-profil » :

From the viewpoint of linguistics, these aims are not very ambitious; but from the clinical viewpoint, they constitute a major advance. It must be recalled that few published descriptions of individual Ps exist—and I know of no description of any P which provides a 'macroprofile', in which all the factors addressed in this book are incorporated, and the interactions between them probed.

Le « macro-profil » serait donc la conjonction des divers profils d'une personne afin de cerner tous les aspects humains grâce à une mise en relation exhaustive des données. Bien qu'approximation à la croisée des domaines pratiques et théoriques, le « macro-profil » est justifié par sa finalité, c'est-à-dire les possibilités cliniques et de remédiation qu'il autorise après description linguistique. Un profil est destiné à faire émerger des tendances, Crystal insiste donc sur la précision et la fiabilité des transcriptions, préférant la qualité à la quantité des données. De plus, un profil cible une personne : l'analyste doit pouvoir comprendre la situation d'interaction et les conditions de communication. Il précise en outre qu'un profil tel qu'il le conçoit n'est pas propice à fournir une perspective normative utile sur le plan clinique.

Crystal affine son étude en distinguant trois sortes de profils linguistiques : un profil phonologique segmental, un profil phonologique prosodique et un profil d'ordre sémantique. La partie suivante va ainsi se focaliser sur la notion de profil prosodique.

3.3 Approche pionnière des profils prosodiques (Crystal 1982)

La présentation suivante permettra de définir un profil prosodique selon Crystal, de décrire les éléments qui le constituent et le traitement subi par les données afin de comprendre le processus qui permet d'élaborer un profil.

3.3.1 Définition

Crystal ([1982] 1992 : 114) définit un profil prosodique comme un outil concret. Son approche descriptive est d'en donner un exemple précis, tangible, source de commentaires généralisateurs :

The nonsegmental phonological profile known as PROP (*'Prosody Profile'*) is a single-page chart, on which can be placed information about the main prosodic patterns encountered in a sample of clinical data. The term 'prosody' here refers to the *linguistic* use of pitch, loudness, speed of speech, pause, and rhythm—in other words, to the way in which these variables can alter the meaning of what we say.

Il écarte d'emblée les traits extralinguistiques (privilégiés dans une analyse médico-légale) pour se concentrer sur une approche linguistique, centrée sur le sens du message véhiculé.

3.3.2 Éléments constitutifs des profils

Crystal ([1982] 1992 : 10) insiste sur l'importance de la transcription prosodique, qui ouvre la voie à la compréhension de l'organisation grammaticale dans le discours de l'élève ou du patient :

The way P organizes his speech prosodically is the main clue to the organisation of his grammar, and a full account of grammatical disability cannot be obtained without this dimension.

Il liste les cinq principaux traits prosodiques d'une transcription : la délimitation des unités tonales, la direction du ton nucléaire, l'emplacement du ton nucléaire, les autres syllabes proéminentes, la durée des pauses. Selon lui, une transcription orthographique est suffisante pour procéder à l'analyse d'un trouble prosodique, grammatical ou sémantique, mais se montre inadéquate pour construire un profil phonologique segmental.

D'après lui, c'est la mélodie qui est responsable de la majorité des difficultés dans le milieu clinique. Mais c'est l'ensemble des traits prosodiques qui doivent être étudiés, selon le *principle of profile comprehensiveness* ([1982] 1992 : 4-5, 114).

3.3.3 Traitement des données

Le traitement des données s'organise autour de leur classement en catégories et de la constitution d'un système de transcription des données. Le tableau récapitulatif des données d'un profil organise la prosodie selon trois catégories :

- L'organisation de la parole en discours suivi (*connected speech*) en fonction du découpage en unités tonales,
- L'utilisation dans les unités tonales de tons spécifiques,
- Le phénomène de tonicité à l'intérieur des unités tonales.

3. Profils personnalisés : centrage sur l'individu en vue d'une identification ou d'une remédiation

L'unité tonale, sur laquelle repose toute cette organisation est définie par Crystal ([1982] 1992 : 114, 116) ainsi :

A tone unit is a finite set of pitch movements, grouped into a distinctive contour and uttered with a distinctive rhythm. Tone units are often bounded by pauses, and these boundaries generally coincide with points of grammatical junction. [...] In very fast speech, [...] tone units run together and lose their clear relationship to grammar. And in highly emotional speech, each word may be given a separate tone unit [...].

Chaque unité tonale est dotée d'un contour mélodique composé d'un nombre limité de tons définis chacun par un mouvement ou hauteur mélodique spécifique. En son sein se détache la syllabe nucléaire. Sa localisation dans l'unité tonale (*tonicity*) modifie le sens véhiculé en offrant un nouveau contraste ou possibilité d'emphase ([1982] 1992 : 116). Les tons portés par les syllabes nucléaires, dites toniques, se définissent par leur direction mélodique et leur registre. Crystal considère trois catégories fondamentales : les tons simples (unidirectionnels), les tons complexes (avec changement de direction mélodique à l'intérieur de la syllabe) et les tons composés (mouvement des tons complexes répartis sur plusieurs syllabes). Le tableau suivant présente le pourcentage moyen de fréquence de ces tons dans l'anglais conversationnel d'adultes rapporté par Crystal ([1982] 1992 : 117)³⁴.

<i>Dénomination tonale</i>	<i>Fréquence d'occurrence moyenne en pourcentage</i>	<i>Catégorie tonale</i>
Ton plat	4,9%	Tons simples (unidirectionnels)
Ton descendant	51,2%	
Ton montant	20,8%	
Ton descendant-montant	8,5%	Tons complexes (changement de direction mélodique à l'intérieur d'une syllabe)
Ton montant-descendant	5,2%	
Ton descendant puis montant	7,7%	Tons composés (changement de direction mélodique sur plusieurs syllabes)
Ton montant puis descendant	1,7%	

Tableau 21 – Fréquence moyenne (en %) des tons dans l'anglais conversationnel d'après Crystal ([1982] 1992 : 117)

Crystal ([1982] 1992 : 117) utilise un système de transcription détaillé permettant d'annoter le script des enregistrements par des signes diacritiques divers insérés sur la ligne d'écriture (*caret transcription*).

³⁴ Crystal ne donne pas de précisions quant à la nature des données sur lesquelles il se fonde.

3.3.4 Processus d'élaboration d'un profil

Voici une description détaillée du traitement des données destinées à être classées en catégories et transcrites : les unités tonales, les variations tonales et le rythme.

3.3.4.1 Nombre et qualité d'unités tonales analysées

Crystal ([1982] 1992 : 124) note qu'un profil doit être basé sur l'analyse d'une centaine d'unités tonales validées (*unproblematic*), ainsi que définies plus bas, bien que, pour certains locuteurs offrant peu de variation prosodique, une petite quarantaine puisse suffire. Le nombre moyen de mots par unités tonales revêt un certain intérêt comparatif.

D'emblée, Crystal ([1982] 1992 : 118) précise le rôle fondamental des unités tonales dans la structuration grammaticale du discours :

The most important function of tone units is grammatical—the integration and delimitation of grammatical patterns. The primary focus of attention, therefore, is on the way in which tone units relate to grammatical structure.

Il recommande de n'analyser, dans une optique grammaticale, que des unités non tronquées (*complete tone units*). Mais d'autre part, étant donné qu'il peut être très difficile d'identifier des unités tonales complètes, surtout dans la parole rapide, il recommande de prévoir une catégorie d'unités tonales indéterminées (*indeterminate*) pour les classer à part. À l'inverse, l'analyste pourra en outre trouver des schémas prosodiques stéréotypés (*stereotyped*) au rythme et contour mélodique standardisés tels que ceux correspondant aux proverbes, comptines (*nursery rhymes*) et *limericks*. Lorsqu'il s'agit de reprises ou de répétitions de l'interlocuteur, l'analyste ne peut décider si le schéma prosodique est révélateur de la maîtrise langagière ou d'une faculté d'imitation.

Une fois ces cas identifiés, Crystal recommande d'analyser les unités tonales selon les structures grammaticales générales qu'elles contiennent, qu'il s'agisse de propositions, d'expressions ou de mots isolés, à condition de toujours réserver la priorité du découpage prosodique à l'unité tonale et non à la structure grammaticale. Il recommande de ne pas prendre en considération la valeur sémantique ou pragmatique (l'acte de parole potentiellement associé) dans le lien entre grammaire et unités tonales, et justifie ce manque de classification par le manque de précision que cela induirait sur l'analyse :

No reference is made to the semantic or social function of the tone unit under the above headings—for example, whether the 'meaning' of the tone unit is to question, persuade, command, excite, and so on. The only reason for the lack of a

classification here is the difficulty of demarcating and formally defining these notions. [...] On a profile chart, it is not possible to anticipate research, nor is it possible to list all possible contextual influences. On the other hand, the chart does need to allow for the importance of these attitudes, in the evaluation of speech. ([1982] 1992 : 123)

Conscient de l'importance de cette restriction, Crystal prévoit une case dans son tableau intitulée « Fonctions » (*Functions*) destinée à mentionner tout contexte social ou attitudinal conduisant le patient ou apprenant à diverger de la norme, ainsi que le nombre de fois où ce phénomène se produirait. Ce dernier aspect souligne les limites du champ des profils prosodiques tels que le conçoit Crystal dans son ouvrage.

3.3.4.2 Variations tonales

Dans un tableau rendant compte d'un profil prosodique, la variation tonale des syllabes nucléaires occupe la seconde position, toute autre distinction tonale trouvant sa place au bas du tableau. Crystal précise ([1982] 1992 : 125) : *Tone here refers to the nuclear tones in tone units, i.e. the maximally prominent tone.*

La première difficulté est de transcrire fidèlement les variations tonales. À cet effet, Crystal préconise une méthode de transcription interlinéaire, qu'il justifie par le caractère imprévisible des réalisations pathologiques ou déviantes :

An interlinear transcriptional method is used, because there is no way of knowing in advance what phonological use is being made of the tones. Only a detailed phonetic record provides the evidence as to whether there is any system present, and this is what an interlinear transcription can provide. ([1982] 1992 : 125)

La transcription interlinéaire permet d'identifier d'emblée visuellement un écart par rapport à la norme constituée par un échantillon correspondant à l'intonation d'un adulte ou d'un enfant à divers stades d'usage de ce dialecte. Les tons spécifiques sont notés par des lettres, qui sont ensuite comptées pour alimenter une analyse statistique, et vérifier les progrès effectués par rapport à ce stade d'élocution. Dans son repérage des tons complexes, Crystal ([1982] 1992 : 129) préfère garder une transcription non affinée, regroupant les mouvements les plus perceptibles, sans sous-catégorisation :

For example, the normal phonetic form of the fall-rise in English is \bar{v} . A 'widened' version of this tone could therefore be any of the following:



3. Profils personnalisés : centrage sur l'individu en vue d'une identification ou d'une remédiation

Les éléments de transcription présentés comme exemples montrent une grande proximité phonétique. Une réduction du nombre de ces éléments est davantage signe d’une proximité phonologique. Par ailleurs, Crystal réduit le nombre de symboles tonals complexes aux bi-directionnels. Il assimile les tons complexes tri-directionnels aux tons bi-directionnels dont ils ne seraient que des variantes :

[...] Other kinds of complex tone (such as rise-fall-rise) are considered to be variants of the two basic types recognized on the chart: a rise-fall-rise is seen as a variant of fall-rise; a fall-rise-fall as a variant of rise-fall; and so on.

La transcription que présente Crystal est manuelle dans sa conception et artisanale dans sa précision pour préserver la fidélité maximale à la réalisation acoustique, bien que les tons tri-directionnels soient assimilés à des tons bi-directionnels, restant par là fidèle aux principes de l’École britannique de l’intonation.

3.3.4.3 Le rythme

Tonicity se définit ici par *the placement within a tone unit of its nuclear tone* (Crystal, [1982] 1992 : 129). Trois types d’accents toniques doivent être repérés d’emblée car ils posent des difficultés à éliminer. Il s’agit des cas d’*indeterminate tonicity* (difficulté de repérer l’accent tonique dans une unité tonale complète), de *stereotyped tonicity* (la localisation de l’accent nucléaire est fixe, comme dans les limericks, les tables de multiplication, les tournures idiomatiques). Crystal ([1982] 1992 : 131) souligne un écho dans l’occurrence de patrons standardisés, révélateur de la cohérence d’un profil :

A stereotyped tonicity generally coincides with stereotyping elsewhere—often the tone-unit structure as a whole is stereotyped, as is the choice of nuclear tone, and the underlying grammatical pattern.

Le troisième cas à repérer est celui de l’imitation du stimulus de l’interlocuteur ou modèle.

Une fois ces cas éliminés, le tableau matriciel indique si l’accent nucléaire des unités tonales se trouve en position finale ou non finale, et ensuite, s’il concerne un item lexical ou grammatical. L’analyste répertorie le nombre de positions correctes ou erronées.

3.3.4.4 Autres traits

La dernière section du tableau, *Other*, permet une rapide mention de quelques traits prosodiques classés en quatre catégories :

- *Tone unit pitch* : utilisation de patrons intonatifs différents

3. Profils personnalisés : centrage sur l’individu en vue d’une identification ou d’une remédiation

- *Tone unit other* : signalement de variation dans l'intensité, le tempo, l'accentuation ou les pauses
- *Prosodic features (TU+)* : signalement de variation dépassant l'unité tonale
- *Paralinguistic features* : signalement de rires, chuchotements, voix de fausset ou soufflée, changement de voix, etc.

Ainsi se sont succédées les trois phases de transcription, de profilage puis d'interprétation des résultats.

3.3.5 Bilan

L'ambition de Crystal est de présenter un outil linguistique sur le plan technique dans une optique de remédiation pour une personne. Son objectif en 1982 reste modeste, mais optimiste : *However, profiling brings to light a wide range of novel possibilities for diagnoses in purely linguistic terms.* ([1982] 1992 : 214). Néanmoins, conscient des limites de cet outil, il entrevoit des possibilités qui ouvrent la voie à des analyses à grande échelle avec l'avènement des grands corpus. Les instruments « bruts » et « non sophistiqués » que représentent ces profils forment un premier pas décisif vers l'établissement d'une base de travail en vue de la création de traits distinctifs prosodiques. C'est en se greffant sur un cadre théorique linguistique et technique ancré préalablement sur une pratique de terrain, que la notion de profil introduite par Crystal peut se révéler la plus fructueuse :

Profiles are not introductions to linguistics, and should not be taught as such. Rather, they begin to make sense only *after* a framework of linguistic theory and technique has been assimilated. To use a profile efficiently, one must be in a position to see its strengths and its limitations, the criteria for which it is the business of linguistics to provide.

3.4 Synthèse

La présentation des trois profils qui ont été présentés dans les domaines médico-légal, linguistique et prosodique a mis en relief les points communs de ces notions dans un cadre cette fois technique parce qu'associé à des situations concrètes. La méthode de caractérisation se base sur l'observation fine, la description fidèle et l'analyse des données provenant d'un individu réel, connu ou pressenti. Il s'agit de décrire fidèlement ce qui le caractérise, de relever la saillance de certains traits pertinents en extrayant les particularités de ses réalisations et en les comparant à un échantillon de référence (adulte normal, etc.). Pour ce faire, sont requises la minutie de l'étude systématique et détaillée, l'importance de

3. Profils personnalisés : centrage sur l'individu en vue d'une identification ou d'une remédiation

l’exactitude des données, des choix catégoriels, de la nécessité de faire ressortir des tendances modélisables, de l’objectif qualitatif à atteindre. Pour obtenir avec précision la description des traits les plus saillants, une analyse fine et extensive doit être pratiquée. Crystal mentionne pourtant la possibilité de concentrer l’étude fine sur certaines zones spécifiques (échantillon), pour obtenir un « micro-profil », étude ciblée non chronophage.

L’étude de profils personnalisés ne saurait être basée sur des approches probabilistes mais au contraire sur des approches cognitives et fondées sur la connaissance et l’individualité de la personne grâce à des sources d’information dont les données pourront être diverses et variées. En fait, un profil ne se limite pas à une simple description mais est un outil qui s’inscrit dans un processus avec un objectif à atteindre. Un profil n’est pas une fin en soi mais un moyen. Les trois catégories de profils présentées utilisent une approche ascendante : l’analyste part des données brutes pour rechercher d’éventuels schémas qui se dessineraient.

Après avoir effectué une description de profils personnalisés, voyons à présent comment s’organisent les profils de groupe dans des corpus oraux.

4 Profils de groupes : catégorisation natifs-non natifs

Afin de rester thématiquement proches du sujet de notre étude, les exemples de profils de groupes seront extraits de recherches à partir de corpus d’apprenants oraux. Quelques généralités sur ces études introduisent les travaux de profilage de groupes de Gut (2009), Tortel (2009), Horgues (2010), Ferragne (2013) et Herment *et al.* (2014). Puis j’aborderai la méthodologie utilisée dans la recherche de Laufer et Nation (1995) qui, bien que centrée sur le profilage de la richesse lexicale dans la production écrite, traite la question de la hiérarchisation qualitative des individus et des groupes d’apprenants en L2.

4.1 Introduction : corpus d’apprenants oraux

Ainsi que le rappellent Ballier et Martin (2015), les corpus oraux sont toujours rares par rapport aux corpus écrits (27% des entrées par rapport aux corpus écrits selon la base de données des corpus d’apprenants répertoriés³⁵ en 2014 avec 38 corpus oraux ; 48 corpus oraux en 2015). Le coût élevé du traitement d’alignement et d’annotation en est la cause. Ils distinguent trois sortes de corpus oraux :

³⁵ Université Catholique de Louvain [<http://www.uclouvain.be/en-cecl-lcworld.html>] (consulté le 30 octobre 2015)

- transcriptions de corpus oraux (*mute spoken corpora*),
- corpus sonore avec alignement temporel (*speaking corpora*) pour étude sur l'axe temporel et disfluences,
- corpus avec visualisation du signal acoustique (*phonetic corpora*).

Seuls les corpus avec visualisation du signal acoustique permettent une étude phonétique approfondie quels que soient les corrélats. L'alignement du signal avec lignes d'annotation sert de base aux mesures et à la transcription au niveau des phones, de la syllabe ou d'unités supérieures. Les domaines étudiés les plus courants en prosodie interphonologique sont le rythme, les proéminences, le registre de parole. Parmi les déviations par rapport à la norme native, figurent un déplacement de l'accentuation lexicale, de l'accent de phrase, des réalisations non attendues, résultant souvent d'un transfert à partir de la L1. Les traits discriminants (*critical features*)³⁶ prosodiques pressentis concernent essentiellement le déplacement de l'accent lexical, la durée syllabique en général, la durée de la syllabe accentuée, les transferts intonatifs, le découpage intonatif en rupture avec la syntaxe, le déplacement d'accent de phrase et la substitution de contour intonatif. La plupart de ces études se fondent sur un découpage syllabique, lequel donne accès à une prise de mesures automatique et à un traitement statistique des données (*metrics*).

Que le mot « profil » soit clairement exprimé ou pas, cette notion est corollaire de l'utilisation d'un corpus. Les exemples suivants indiquent quels choix ont été opérés par inclusions ou exclusions parmi leurs données, tant au niveau de la sélection des locuteurs que dans le traitement des formes étudiées.

4.2 Profilage prosodique multidimensionnel : une approche pionnière de l'interlangue (Gut 2009)

Les résultats principaux de recherche de Gut (2009) ont déjà été présentés au Chapitre I. Aussi, cette section se focalise sur la méthodologie qui lui a permis de les obtenir en aisance, rythme et intonation, avant d'en faire le bilan pour ma recherche.

4.2.1 Obtention et traitement général des données

La recherche de Gut a été effectuée sur le corpus *LeaP*, corpus oral intégralement aligné (Gut 2009). Il se subdivise en un sous-corpus de langue anglaise et un autre de langue allemande. Le sous-corpus de langue anglaise comprend 46 locuteurs non natifs de divers

³⁶ Présentés au Chapitre I.

niveaux d’acquisition (dont certains quasiment bilingues) et quatre locuteurs natifs de référence (2009 : 65). Quatre styles de parole (*speaking styles*) y figurent : conversation, lecture à voix haute, restitution orale de l’histoire précédemment lue et lecture de mots épars (2009 : 64). Les formes étudiées sont clairement présentées et ce classement est propice à l’application d’un traitement statistique quantitatif.

Le système d’annotation utilise le logiciel *Praat* avec six lignes d’annotation manuelle (mots prosodiques et pauses, mots, syllabes, segments, accents et frontières (*tones*), hauteur mélodique) et deux lignes d’annotation automatique (*part-of-speech*, *lemmes*). La ligne d’annotation de l’intonation non-native a été complétée par l’adaptation du système ToBI à l’allemand GToBI. Cette stratégie d’annoter manuellement et extensivement l’ensemble des enregistrements implique au préalable une formation lourde de trois mois pour chacun des six annotateurs (sauf 1 ayant intégré le projet plus tard) ; un test statistique vient compléter le dispositif afin de vérifier la cohérence des annotations d’un même annotateur (écart intra-annotateur) et l’homogénéité des annotations des différents annotateurs entre eux (écart inter-annotateur). Les mesures *Mean kappas* de l’écart inter-annotateur indiquent un taux d’équivalence moyen dans l’annotation des syllabes et des tons mélodiques alors que ceux des mots, segments et hauteur mélodique atteignent un taux d’accord presque parfait. Avec une période de deux ans entre la première et la seconde annotation des mêmes annotateurs, les taux d’équivalence intra-annotateur sont similaires à ceux de l’écart inter-annotateur. Gut conclut (2009 : 72) en mentionnant la proportionnalité entre la fiabilité de l’annotation manuelle et la complexité de la tâche demandée. Les données alignées ont ensuite été converties au format *TASX* (*Time Aligned Signal data eXchange format* ; *Milde and Gut 2002a, 2002b, Pitsch et al. 2003*), format pour les données en XML. Les métadonnées y sont intégrées via un format IMDI/ISLE.

Une étude statistique a été menée grâce aux outils d’analyse intégrés dans *TASX* ou, après conversion dans le format de fichier *NITE* (*Natural Interactivity Tools Engineering*), la partie *NXT Search* du logiciel *NITE XML Toolkit* (*NXT*)³⁷. En liaison avec la présente étude, trois domaines sont regardés avec précision : l’aisance de la parole (*fluency*), le rythme et l’intonation.

³⁷ Carletta *et al.* (2006) précisent qu’il s’agit d’un logiciel de traitement de corpus langagiers capable de gérer des ensembles de données multimodales pourvues d’annotations complexes sur divers niveaux. Les sous-programmes des bibliothèques de *NITE XML Toolkit* (*NXT*) permettent de mettre en évidence les relations structurelles et temporelles qu’entretiennent les données. Les domaines étudiés par Gut chez les non natifs englobent aussi bien les phénomènes segmentaux que suprasegmentaux dans leurs interactions.

4.2.2 L'aisance (fluency)

Gut donne de l'aisance (*fluency*) diverses définitions selon le domaine de recherche (2009 : 77). Le sens le plus commun est celui de la maîtrise générale d'une langue étrangère lorsqu'elle est parlée. Elle se réfère ainsi à une production de discours oral en termes de fluidité, de continuité, d'automatisme et d'uniformité (*smoothness*). Elle précise (2009 : 78) que dans le domaine de l'enseignement et l'évaluation des langues, l'aisance n'est qu'un trait distinctif (*differentiable feature*) parmi d'autres, tels que la précision, la prononciation, le caractère idiomatique et la pertinence. Elle souligne que c'est en didactique des langues essentiellement que la ligne de démarcation entre l'aisance et la précision est la plus claire. Elle mentionne Brumfit (2000) selon lequel l'aisance langagière peut être présente dans une production orale par son caractère fluide et continu sans que soient respectées les règles et structures grammaticales ni pragmatiques. Elle ajoute qu'en sociolinguistique, la définition de l'aisance englobe davantage de capacités et savoir faire tels que l'emploi et la production corrects d'expressions gestuelles et faciales (selon Bavelas 2000) et même la maîtrise des conventions conversationnelles comme la prise de tours de parole (Fiksdal 2000).

Dans son étude, Gut (2009 : 78) se base essentiellement sur la définition en usage en psycholinguistique. Elle rapporte que Schmidt (1991) la décrit comme *the qualitative outcome of the cognitive production processes involved in the planning and delivery of utterances*. Basés sur la modélisation du discours monolingue de Levelt (1989) et de Bot (1992) dans son application aux locuteurs de deux langues, la production langagière repose sur deux sortes de connaissances que sont les connaissances déclaratives et les connaissances procédurales. Les connaissances déclaratives sont factuelles alors que les connaissances procédurales sont des savoir faire. Chez les locuteurs natifs d'une langue, les deux sortes de connaissances jouent de pair, en parallèle et automatiquement. L'automatisme de la production permet de s'exprimer sans effort ni contrôle conscient, ce qui lui permet de concentrer son attention sur le sens de ses propos ou la situation interlocutive. Ce n'est pas le cas de l'apprenant, lequel doit automatiser ses connaissances déclaratives afin de procéduraliser leur utilisation. Gut mentionne Towel *et al.* (1996) selon lesquels la tâche de l'apprenant est de développer ses savoirs procéduraux à partir de ses connaissances déclaratives. Les disfluences constatées chez les non natifs révèlent alors le degré du manque d'automatisation des processus de bas niveau en jeu dans la production orale, l'aisance absolue étant atteinte lorsque les capacités procédurales en production ne requièrent ni attention ni effort.

De la définition de l'aisance en production orale selon les études psycholinguistiques, Gut (2009 : 79) rapporte qu'ont été tirées des mesures temporelles pour mesurer la vitesse et la qualité des processus de production orale. Plusieurs études montrent que l'aisance absolue ne se réalisait pas toujours chez les natifs, que l'on prend pourtant comme norme. Très étudiées, les pauses jalonnant l'élocution ont une influence sur l'aisance perçue par l'auditeur.

4.2.2.1 Objectifs et méthodologie de Gut

Gut présente (2009 : 93) son objectif central :

The aim of this study is to provide a corpus-based large-scale analysis of non-native fluency that includes a great number of speakers of different levels of competence producing a large sample of speech in different speaking styles.

Cette recherche extensive de l'aisance ne saurait se couper de celle de la prononciation, de l'intonation, ni de l'accentuation (autres compétences phonologiques), ou celles de métadonnées sur les apprenants. Elle peut permettre de savoir :

- quelles méthodes sont les plus appropriées pour analyser l'aisance,
- si l'aisance peut être corrélée avec d'autres variables phonologiques ou si c'est une compétence qui leur est indépendante,
- si l'aisance est tributaire d'un des quelconques facteurs qui influent sur l'apprentissage général d'une langue.

Elle peut aussi contribuer à:

- fournir des valeurs de référence aux évaluateurs de langue,
- de fournir une réponse quant à quel style de parole devrait être évalué.

En conséquence, Gut a sélectionné (2009 : 93-94) les mesures quantitatives de l'aisance suivantes pour être appliquées à l'intégralité du corpus *LeaP* :

- **len**: total length of speech (durée totale du discours y compris de tout ce qui est prononcé, pauses pleines ou silencieuses et événements paralinguistiques tels qu'un éclat de rire)
- **art**: total length of articulation (durée totale de l'ensemble des phrases)
- **silratio**: silent pause ratio (proportion en pourcentage de la durée totale des pauses silencieuses par rapport à la durée totale d'élocution)

- **fillratio**: filled pause ratio (proportion en pourcentage de la durée totale des pauses remplies par rapport à la durée totale d'élocution)
- **speechrate**: speech rate (nombre total de syllabes divisé par la durée totale du fichier)
- **artrate**: articulation rate (nombre total de syllabes divisé par la durée totale du discours parlé)
- **mlr**: mean length of run (longueur moyenne d'un tronçon sans pause mesuré en syllabes)
- **words**: total number of words (nombre total de mots dans la seule reformulation)

Gut précise qu'outre ces analyses quantitatives, des analyses qualitatives ont été ajoutées à l'étude destinées à comparer l'aisance des apprenants avant et après un séjour à l'étranger, avant et après un stage de prononciation, ou même comparer l'aisance en L2 et en L3, parmi une collecte de données métalinguistiques. Elles sont destinées à expliquer la variation pouvant être repérée dans les données (2009 : 56).

4.2.2.2 Bilan de l'étude de la fluidité dans Gut (2009) et implications pour mon travail

Par son récapitulatif des données antérieures et les siennes, Gut fournit une base comparative du plus grand intérêt que ce soit dans l'étude des natifs anglophones ou des non natifs.

4.2.2.2.1 Méthodologie

Son profilage s'est basé sur un grand corpus (*LeaP*) selon une approche linguistique d'analyse de corpus. Ses objectifs ont été d'explorer les propriétés générales de l'aisance des non natifs et de trouver les contraintes auxquelles étaient confrontés les non natifs dans leur acquisition, et leur degré de réussite au cours de leur apprentissage. Son troisième objectif a été de tester la validité de certains paramètres temporels de l'aisance afin d'en extraire des pistes destinées à l'évaluation de l'aisance. En ce qui concerne l'étude de l'anglais, cette finalité a nécessité la création de corpus de natifs et de non natifs anglophones, ainsi que d'analyses quantitatives et qualitatives, lesquelles s'inscrivent dans une continuité des recherches sur le sujet, testées alors à plus grande échelle.

Les résultats des métriques sont parfois difficiles à interpréter qualitativement car la diversité des styles de parole et celle du niveau des apprenants n'est que peu exploitée. On constate un regroupement des locuteurs non natifs (sans exploiter vraiment les différences de

CHAPITRE II – Profils et profilage : mode d’organisation hiérarchique de la prosodie des anglophones natifs

niveau, la dichotomie fondamentale étant : natif/non natif) et des styles de parole, la priorité étant laissée aux formes étudiées. Les profils résultants mettent davantage en valeur les formes étudiées que les locuteurs qui les ont produites ou leur provenance.

Quant aux pauses silencieuses, la bi-modalité de durée des pauses renseigne sur la fourchette de durée à prendre en compte pour étudier les pauses en lecture : 0,1 seconde semble être la durée minimale à sélectionner.

4.2.2.2.2 Sélection des variables et implication pour mon étude

Son profilage implique un investissement très important en termes de moyens techniques et humains mis à disposition. Il ne peut en être question dans cette étude. En conséquence, j’ai repris et adapté les variables qui semblaient intéressantes suite aux résultats exposés précédemment pour les rendre gérables dans cette étude, qui est centrée sur la lecture à haute voix et qui ne nécessite pas une entreprise comparable. Cependant, malgré ces restrictions, je reste convaincue de l’importance du lien que peut avoir l’évaluation prosodique en lecture et en conversation (même si cela n’est pas le sujet de cette étude). Aussi, dans le choix de mes variables, je privilégierai celles qui seront transposables à d’autres styles de parole, dont la conversation, dans un souci de généralisation et d’uniformisation des pratiques, ce qui correspond à une notion d’élégance dans le profilage.

Bien que ce profilage de Gut soit adapté à une multiplicité de styles de parole, je sélectionnerai les variables suivantes pour mon étude sur les profils de natifs en lecture :

Catégories des variables	Variables de Gut	Mes variables	Caractéristiques de mes variables
Durée de lecture	<i>len</i>	VIT.T	Durée totale du texte
Débit d’élocution	<i>speechrate</i>	DEB.T	Nombre de syllabes par seconde dans le texte (pauses incluses)

Tableau 22 – Reprise des variables de Gut (2009) dans mon profilage de natifs en lecture

Le tableau ci-dessus reprend les deux variables de base que sont la vitesse de lecture du texte (VIT.T) et le débit d’élocution du texte (DEB.T).

N’étant pas dotée des moyens dont a bénéficié la recherche de Gut (2009), j’ai sélectionné un segment de parole non ponctué où l’on peut raisonnablement penser que les locuteurs s’abstiendraient de marquer une pause pendant la lecture de ce tronçon, que je

nomme A. Ainsi, on peut espérer que cette mesure pourrait refléter l'ensemble du discours à bien moindre frais que la variable de Gut *art*. Toute règle comportant des exceptions, j'ai prévu de doubler cette variable par sa consœur VIT.Ap, qui inclut le temps de pauses éventuelles alors que VIT.A les exclut. Cette distinction permet en outre de prendre en compte le débit d'élocution par un travail réduit et très ciblé : DEB.A est destiné à rendre compte du débit d'élocution moyen par syllabe dans un segment comportant uniquement de la parole, alors que DEB.Ap fera de même avec modulation générée par des pauses éventuelles. DEB.A devrait correspondre à la variable de Gut *artrate* sur partie ciblée du texte alors que son corollaire DEB.Ap serait approximativement le reflet de *spechrates*, dont la prise de mesures est particulièrement chronophage sur un texte long. Le choix de sélectionner un tronçon pré-déterminé permet, outre un gain de temps appréciable, de facilement adapter cette technique de mesure à la parole spontanée, même si la précision est bien moindre. Par ailleurs, ce ciblage sur une zone, que l'on peut rapprocher d'un micro-profil selon la définition de Crystal (voir plus haut), permet de s'affranchir de la lourdeur de l'annotation en équipe et, en conséquence, de la formation préalable des annotateurs ainsi que des calculs d'écart inter-annotateurs par la suite.

Catégorie des variables	Variables de Gut apparentées	Mes variables	Caractéristiques de mes variables
Durée de lecture et des pauses	<i>art</i>	VIT.A	Durée d'un segment déterminé par l'improbabilité de ses pauses (sans pauses)
	<i>art/len</i>	VIT.Ap	Durée d'un segment déterminé par l'improbabilité de ses pauses (avec pauses éventuelles)
	<i>silratio</i>	PAUvir	Durée moyenne d'une pause à l'emplacement de virgule
Débit d'élocution	<i>artrate</i>	DEB.A	Nombre de syllabes par seconde dans le tronçon A (sans pauses)
	<i>spechrates</i>	DEB.Ap	Nombre de syllabes par seconde dans le tronçon A

			(pauses incluses)
--	--	--	-------------------

Tableau 23 – Reprise des variables de Gut (2009) dans mon profilage de natifs en lecture avec transformation

La prise en considération des pauses dans le segment A est doublée d’une autre variable : PAUvir. Liée à la lecture de l’écrit, il s’agit de la durée moyenne d’une pause marquée à l’endroit où le texte est ponctué d’une virgule. Cette variable se rapproche d’une certaine mesure de la variable *silratio* bien qu’elle cible avec précision l’endroit où cette pause se produit, en excluant tout autre signe de ponctuation. Les pauses remplies n’ayant pas de pertinence dans ce profilage de natifs (corrélée au style de parole et absente de la lecture à haute voix des natifs), je n’en ai pas tenu compte de cette variable.

La variable de Gut *mlr* n’apparaît pas dans le tableau car je l’ai transposée en constante : il s’agit du tronçon A, dont la variante avec pauses éventuelles est Ap. Elles servent de points d’appui aux variables de vitesse (VIT.A et VIT.Ap) et de débit (DEB.A et DEB.Ap)³⁸.

Ainsi seront plus ou moins approximativement préservées les variables que Gut a reconnues comme essentielles dans sa recherche présentée en amont. Leur nature acoustique et donc numérique leur permet de s’inscrire dans un continuum. Leur corollaire qualitatif devra catégoriser ces données en opérant des choix arbitraires.

4.2.3 Le rythme

Gut remarque que, si le modèle de ces mesures implique une alternance binaire de type long-bref ou fort-faible, les résultats dépendent surtout de la phrase dans laquelle les mesures ont été faites. Elle rappelle sa recherche à partir de l’hypothèse de Bolinger (1981) selon laquelle le rythme de la langue anglaise est conditionné par l’alternance de syllabes avec voyelles pleines et d’autres aux voyelles réduites. Bolinger met en garde contre la succession de deux syllabes aux voyelles pleines car la première est bien plus longue que si elle était suivie par une syllabe à voyelle réduite. Les proportions de moyennes à mesurer sont donc celles de paires de syllabes à voyelles pleines adjacentes d’une part, et celles de voyelles pleines suivies de syllabes aux voyelles réduites d’autre part. Par précaution, Gut exclut (2003a) du profilage les syllabes précédant une pause. Toutes ces mesures d’approche

³⁸ La méthodologie de création de ces variables est exposée dans la sous-section 6.3.2 « Méthodologie du profilage ».

phonétique réduisent néanmoins le rythme à un niveau purement segmental selon le reproche qui lui est fait par ailleurs.

Ce panorama des diverses approches du rythme s'appliquent à l'anglais avec les résultats suivants.

4.2.3.1 *Objectifs et méthodologie de Gut*

Afin de pouvoir décrire les corrélats acoustiques qui différencient le rythme natif de celui des non natifs selon le style de parole et selon les niveaux d'apprentissage, Gut liste (2009 : 175-176) les mesures qu'elle utilise :

- **mean length sfv** : durée moyenne de toutes les syllabes comprenant une voyelle pleine
- **mean length srv** : durée moyenne de toutes les syllabes comprenant les voyelles réduites (en anglais : /ə/ et /ɪ/)
- **mean length sdv** : durée moyenne de toutes les syllabes comprenant une voyelle effacée
- **percentage red/del** : nombre de toutes les syllabes comprenant une voyelle soit réduite soit effacée divisé par le nombre total de syllabes et multiplié par 100
- **ratio full/red** : proportion de la durée moyenne de toutes les paires de syllabes dans lesquelles une syllabe comprenant une voyelle pleine est suivie d'une syllabe avec voyelle soit réduite, soit effacée
- **ratio full/full** : proportion de la durée moyenne de toutes les paires de syllabes dans lesquelles une syllabe comprenant une voyelle pleine est suivie d'une autre syllabe avec voyelle pleine

Une analyse de corpus qualitative a été menée afin d'approfondir et d'expliquer les résultats quantitatifs.

4.2.3.2 *Bilan du rythme*

Les métriques que Gut a utilisées dans son étude sur le rythme des non natifs sont %V, ΔC , le taux syllabique, et la qualité et longueur des voyelles. Gut émet des réserves sur l'utilité de la variable ΔC de Ramus *et al.* (1999) car, selon elle, ΔC permet de distinguer la parole des natifs de celle des non natifs sans pouvoir localiser cette dernière sur l'axe du continuum rythmique des langues.

Les données ont montré que les syllabes de toute nature des non natifs étaient plus longues que celle des natifs et que son corollaire était que la parole non native était plus lente que celle des natifs. Les non natifs ne parviennent pas à réduire ou supprimer leurs voyelles : cela se traduit pas une différence insuffisance de durée entre les syllabes avec voyelles pleines et celles aux voyelles réduites ou effacées.

En conclusion de l'étude de Gut sur le rythme, je retiens le caractère chronophage de l'analyse du rythme par approche inter-consonantique ou battement rythmique. Sa préférence va vers une combinaison de la prédominance syllabique avec celle du battement rythmique lorsqu'elle retient la mesure de durée des syllabes en tenant compte de la qualité vocalique (pleine-réduite ou effacée), qui implique un degré accentuel (fort-faible). Le profil de groupe inclut toutes les syllabes du corpus correspondant à la définition. Cependant, la notion d'exclusion est présente : suivant les instructions du *Variability Index* de Deterding (2001) écartant les syllabes finales, Gut précise (2003) l'exclusion de sa propre étude des syllabes précédant une pause. En cela, les règles et exceptions des métriques ne font sens que par leurs résultats et non par la cohérence ni la cohésion de leurs principes fédérateurs sous-jacents, s'ils existent. D'ailleurs, Gut remarque l'influence que peut exercer la phrase dans laquelle est insérée la paire syllabique mesurée. Aussi, très traditionnellement, ma préférence pour organiser les profils de natifs, ne se détournera pas d'une analyse perceptive, tenant compte du nombre de proéminences et de la vitesse d'élocution. Pour l'analyse des non natifs, qui nécessite des mesures acoustiques fiables, je retiens néanmoins les résultats d'Adams (1979) caractérisant le rythme des non natifs : insuffisance de durée entre les syllabes accentuées et inaccentuées, pauses erratiques et déplacement accentuel. De même, je favoriserai les métriques simples du type pourcentage et proportion de données facilement et rapidement mesurables. Élément important du profilage des données mais aussi des variables à fin évaluative, j'utiliserai aussi un corpus contrôle de natifs.

Troisième volet essentiel en prosodie reposant sur le rythme, examinons à présent l'intonation.

4.2.4 L'intonation

Cette sous-section se focalise sur les objectifs poursuivis par Gut et la méthode d'analyse de corpus, puis mon approche au regard de sa méthodologie pour cette thèse.

4.2.4.1 Objectifs poursuivis par Gut

Regrettant le peu de recherches consacrées à l'étude de l'intonation non native, Gut vise à fournir un panorama systématique de l'intonation non native en faisant l'inventaire des tons et de leur signification dans l'interlangue, en décrivant ses contours mélodiques, son accentuation et l'utilisation de son étendue de fréquence, et en étudiant ces éléments interphonologiques en anglais et en allemand. Il faut y ajouter leurs caractéristiques suivant les styles de parole et une analyse qualitative complémentaire sur les contextes d'apprentissage sur l'acquisition de l'intonation en L2 ainsi que la relation entre L1 et L2 dans ce domaine. Les mesures prises pour obtenir les données quantitatives sont répertoriées ci-dessous.

4.2.4.2 *Méthodologie utilisée par Gut*

4.2.4.2.1 Découpage en unités intonatives

Dans les 359 enregistrements, le profilage a consisté à distinguer les unités intonatives (*intonation phrases, IPs*) correspondant à un schéma opératoire, des énoncés interrompus (*interrupted phrases*). Les *IPs* retenus correspondent à un profil de groupe distributif, c'est-à-dire que tous leurs éléments doivent se conformer à la définition. Ils se définissent par l'existence d'un noyau et au moins l'un des critères suivants : allongement final, *IP* suivi de pause, anacrusse et augmentation de la fréquence fondamentale dans l'*IP* suivant. Un critère non définitoire est que l'*IP* correspond souvent à une phrase ou proposition syntaxique. Les « énoncés interrompus » sont définis comme ne possédant pas les caractéristiques des *IPs*.

4.2.4.2.2 Inventaire tonal

À partir des 359 enregistrements, le script *Prosogram* (Mertens 2004) compatible avec le logiciel *Praat* a élaboré des prosogrammes permettant de visualiser une schématisation des contours mélodiques à l'exclusion de dix enregistrements dont la qualité sonore était insuffisante pour calculer automatiquement la fréquence fondamentale. L'analyse a porté sur les tons nucléaires et pré-nucléaires des seuls *IPs* avec noyau. Le classement tonal a été opéré en fonction des catégories : *fall, rise, low level, mid level, high level, rise-fall, fall-rise, rise-fall-rise*. Les mesures des tons dynamiques ont été effectuées en demi-tons (différence entre F0 maximale et F0 minimale) en incluant les mots précédant le noyau. La hauteur mélodique pré-nucléaire a été étudiée pour les classes de mots lexicaux (*content words*) et de mots outils (*function words*).

4.2.4.2.3 Gamme de fréquence fondamentale

L'étendue maximale de fréquence de chacun des 359 enregistrements a été mesurée en demi-tons. Une classification en trois catégories qui sont le discours direct, le rhème (*new topics*) et le thème (*background*) doit permettre d'analyser les éventuelles corrélations entre la gamme de fréquence et les différents types d'énoncés discursifs. Pour vérifier les thèses de Wichmann (2000), une mesure de la hauteur initiale de chacun de ces énoncés est destinée à être classée dans une des trois catégories de la gamme de chaque enregistrement *high*, *mid* et *low*.

4.2.4.2.4 Placement de l'accent mélodique

Huit phrases ont été extraites des passages de lecture afin de vérifier si les réalisations non natives correspondaient à celles des modèles natifs. Dans trois d'entre elles, la structure verbe + pronom permet de vérifier l'accentuation sur le verbe. Dans trois autres phrases avec nom + verbe, la localisation du noyau est libre. Deux autres phrases sont destinées à vérifier la désaccentuation du thème. Les analyses qualitatives vérifient l'influence pour la même L1 du contexte d'apprentissage sur l'acquisition de la L2.

4.2.4.2.5 Analyses qualitatives

L'influence de la L1 sur l'intonation de la L2 a été vérifiée par ANOVA avec seuil de signification inférieur à 0,05 pour différencier les données acoustiques selon trois groupes de L1. Dans la seconde expérimentation d'une modification des critères acoustiques spécifiques aux difficultés intonatives des apprenants après un cours intensif ou un séjour à l'étranger, Gut ne mentionne pas les tests utilisés à l'exception des seuils de signification des résultats qui sont inférieurs à 0,05 ou 0,01.

4.2.4.3 **Bilan de l'intonation**

L'introduction méthodologique de Gut présentant les notions et recherches antérieures est du plus grand intérêt car elle permet d'asseoir la présente étude. Ayant précédemment présenté le travail de Gut, je ne m'attarderai pas à le résumer davantage, mais m'orienterai uniquement vers une analyse du profilage utilisé et de son influence sur les choix que j'opèrerai dans ma propre recherche.

4.2.4.3.1 Analyse du profilage intonatif de Gut

Devant analyser le grand corpus *LeaP*, Gut opte pour la définition de l'intonation étroite : *intonation refers to the use of suprasegmental features to convey postlexical meanings in a linguistically structured way*. Elle exclut les traits paralinguistiques et la

focalisation, les structures morphosyntaxiques et sémantiques, les traits suprasegmentaux lexicaux afin de mieux cerner les unités phonologiques et discrètes qui permettent d'analyser l'interlangue en termes linguistiques et catégoriels. Elle adopte l'approche métrique autosegmentale ainsi que le système de transcription *ToBI* pour décrire et catégoriser les réalisations prosodiques.

Bien que très proche par ses objectifs, ma recherche est bien plus modeste, sans aucune mesure avec celle d'un grand corpus, ni avec les moyens humains mis en œuvre dans l'étude de Gut. Aussi, les écueils à éviter sont différents : n'étudiant qu'un (relativement) petit corpus, dont l'objectif est évaluatif, je dois être aussi précise que possible dans mes mesures et les résultats des analyses. C'est pour cela que je n'opterai pas pour l'approche métrique autosegmentale, ni l'annotation *ToBI*. Ces approches consistent à d'abord catégoriser pour décrire dans un second temps, favorisant une possible perte d'information sur les données. Je garderai de l'intonation les critères suprasegmentaux qui me permettront d'atteindre les objectifs catégoriels évaluatifs en favorisant une compréhension fine des phénomènes en jeu afin de pouvoir écarter à bon escient les éléments inutiles ou perturbateurs par la suite. Il s'agit là d'une question méthodologique d'ordre chronologique. En outre, le caractère de l'annotation britannique est davantage visuel, faisant ainsi appel à l'intuition du chercheur. Mon objectif étant de décrire pour comprendre, puis de catégoriser pour hiérarchiser et ensuite d'évaluer, il me semble que l'approche de l'École britannique de l'intonation devrait davantage permettre de saisir la cohérence entre les unités intonatives que l'approche métrique autosegmentale, qui *divides intonation into a sequence of discrete and independent tone targets* (Gut, 2009 : 198).

Gut regrette le peu de recherches consacrées à l'étude de l'intonation non native. En effet, son analyse systématique d'un si grand corpus multilingue (359 enregistrements) est exceptionnelle. Son profilage des unités intonatives a d'abord consisté à les découper. Pour cela, une première étape a constitué deux groupes distributifs : unités intonatives devant par la suite être analysées finement et énoncés interrompus de moindre intérêt pouvant alors être écartés de l'analyse tonale. Un nombre très élevé de catégories tonales (8 en tout) a permis de classer les tons relevés automatiquement par prosogrammes (à l'exception des enregistrements inutilisables) et de comparer leur mesure en demi-tons sur le noyau et les éléments le précédant (catégorisés en mots lexicaux et mots outils). Gut est en outre parvenue à faire correspondre trois niveaux de la gamme de fréquence fondamentale avec trois types d'énoncés discursifs. La stratégie méthodologique a consisté à créer des catégories générales

CHAPITRE II – Profils et profilage : mode d’organisation hiérarchique de la prosodie des anglophones natifs

dans lesquelles sont alors consignées les mesures acoustiques afin de constituer des groupes et de comparer leur effectif ensuite. Le profilage consiste à corréliser tendances et catégories ; les différences qui apparaissent dans le classement peuvent alors être analysées et organisées pour faire sens. On part des mesures (données continues) pour les intégrer dans un classement qualitatif (catégoriel) tendant à l’exhaustivité des possibles. Les classes acoustiques sont distributives et les catégories de locuteurs sont non distributives. L’analyse systématique doit regrouper le plus grand nombre de données pour les organiser et rendre compte du classement et des croisements obtenus.

Le profilage du placement de l’accent mélodique correspond à une méthodologie différente. Gut a créé des classes catégorielles *ad hoc* et a vérifié ensuite en lecture si le phénomène d’accentuation leur correspondait en validant ou non la catégorie. Ce profilage permet d’obtenir des données quantitatives à partir de choix qualitatifs. Les analyses qualitatives sont effectuées à partir de données métalinguistiques et non acoustiques.

En fonction de la méthodologie de profilage utilisée par Gut, voyons ce qu’il est intéressant de garder dans mon étude.

4.2.4.3.2 Éléments favorisant la sélection des variables intonatives

Découpage en unités intonatives

Tout comme Crystal, Gut établit deux catégories : les unités pouvant être analysées avec profit et celles qui ne sont pas conformes à la standardisation requise. Gut montre que c’est en lecture que l’on trouve le plus grand pourcentage d’unités intonatives analysables. Me basant sur un texte écrit, je retiendrai les unités intonatives correspondant sensiblement au texte de Wilde dans la mesure où des pauses incongrues ne perturbent pas la cohérence de l’intonation.

Inventaire tonal

L’inventaire le plus complet possible me semble préférable. L’objectif comparatif rend nécessaire une unité de mesure commune telle que les demi-tons.

Gamme de fréquence fondamentale

Je reprendrai les mesures de Gut selon lesquelles, dans l’enregistrement effectué, la différence entre la hauteur mélodique la plus haute et la plus basse définissent le calcul du registre. Ce style de mesure manuelle est davantage adapté à un petit corpus sur un texte court

si l'on veut pouvoir vérifier qu'il n'y a pas de sauts d'octaves. Ici aussi la conversion des Hertz en demi-tons permet des calculs des mesures de fréquence fondamentale.

Positionnement de l'accent mélodique

Le ciblage que fait Gut de structures contraintes concentrant les déviations des non natifs par rapport à la norme native correspond à un profilage efficace des non natifs.

Analyses qualitatives

Les questionnaires permettent d'obtenir des données métalinguistiques qui peuvent aider à expliquer les réalisations des locuteurs, mais aussi, comme on l'a vu dans l'expérimentation du Chapitre I, à exclure du corpus ceux dont les mesures acoustiques pourraient biaiser les résultats.

4.2.5 Bilan et discussion de l'étude de Gut (2009)

Le compte rendu détaillé que j'ai fait de la recherche de Gut dans les trois principaux domaines intonatifs est extrêmement riche sur le plan des données natives et non natives, mais aussi sur le profilage mis en œuvre pour les obtenir. Ces données, provenant de la recherche de Gut ou antérieures à celle-ci, sont un point d'appui comparatif fondamental à mes propres données. Ces profils et profilages ont mérité d'être étudiés en détail. Ils suivent une approche ascendante : Gut est partie des données pour chercher une vision d'ensemble, ensuite elle a exclu les éléments qui ne sauraient être intégrés dans des catégories pré-établies ou susceptibles de fausser l'analyse (dans le découpage des unités intonatives, les tons et le registre). Elle a aussi utilisé une approche descendante dans la localisation de l'accent mélodique (les structures syntaxiques à vérifier sont prédéterminées) et le choix de codage *ToBI* car, dans la mesure où l'on annotait manuellement en fonction de codages phonologiques, on imprègne les données de ces visions.

Par ailleurs, les analyses statistiques utilisées par Gut sont principalement les corrélations et ANOVA avec seuils de signification inférieurs à 0,05 ou 0,01.

4.3 Profilage du rythme par comparaison de métriques : l'importance de la parcimonie du modèle (Tortel 2009)

4.3.1 Objectifs

La recherche d'Anne Tortel porte sur le rythme des apprenants francophones anglicistes. L'absence de critères d'évaluation prosodique la conduit à considérer les facteurs

rythmiques en L2 anglaise en analysant l’influence du rythme du français sur la production anglaise, puis en proposant des critères évaluatifs du rythme après avoir testé différentes métriques destinées à évaluer le rythme des apprenants anglicistes. Ses hypothèses se basent sur la dichotomie des langues selon la catégorisation en langue isosyllabique ou langue isoaccentuelle.

Elle met en avant un modèle qui « permet de construire un feedback précis, sur des erreurs ciblées, et de comprendre où se situent les difficultés ». Je détecte là une question portant sur la finalité de l’évaluation, à savoir si sa nature est formative (test de positionnement en vue de détecter les faiblesses pour remédiation) ou sommative (hiérarchisation des performances des apprenants pour les catégoriser).

4.3.2 Méthodologie d’investigation

À partir du corpus multilingue *EUROM 1*, lui-même provenant d’un projet préliminaire *SAM (Speech Assessment Methods)*, Tortel a constitué le corpus *ENGLISH* dont la partie lecture a servi à son expérimentation sur l’évaluation du rythme des apprenants anglicistes francophones. Il s’agit d’une tâche de lecture de quatre passages de cinq phrases chacun, lus par 63 locuteurs (2009 : 159). Chaque texte a été soigneusement sélectionné pour éviter les difficultés de vocabulaire, de prononciation, les phrases trop courtes, ou trop longues, ou en calibrant les polysyllabes (2009 : 158). Une fois ces critères préalables définis, les paramètres sont restés constants dans l’analyse. Par contre, les variables ont concerné :

- Les locuteurs : ils répartis en trois groupes de niveaux (2009 : 162) : 23 locuteurs natifs anglophones (GB), 20 apprenants étudiants de fin de 2^{ème}/début 3^{ème} année d’études (FR2) et 20 apprenants ayant niveau Bac anglais L2 (FR1).
- Le découpage des unités (2009 : 167) : segmentation en unités intonatives, en mots, en syllabes, en unités rythmiques, et en segments (CVC) codés C (pour consonnes) et V (pour voyelles).
- Les métriques (2009 : 174-175) : utilisation de neuf paramètres rythmiques à fin de comparaison entre les locuteurs anglophones natifs et les apprenants. Elles sont définies dans le tableau ci-dessous :

% V	la somme de la durée des intervalles vocaliques divisée par la durée totale des intervalles vocaliques et consonantiques, multipliée par 100
-----	--

ΔC	écart type des durées des intervalles consonantiques
ΔV	écart type des durées des intervalles vocaliques
rPVI_C	moyenne des différences de durées entre les intervalles consonantiques successifs (brut)
nPVI_C	moyenne des différences de durées entre les intervalles consonantiques successifs divisée par la moyenne de ces mêmes intervalles et multipliée par 100 (normalisé)
rPVI_V	moyenne des différences de durées entre les intervalles vocaliques successifs (brut)
nPVI_V	moyenne des différences de durées entre les intervalles vocaliques successifs divisée par la moyenne de ces mêmes intervalles et multipliée par 100 (normalisé)
cvV	(coefficient de variation=) écart type des durées des intervalles vocaliques divisé par la moyenne des durées des intervalles vocaliques
cvC	(coefficient de variation=) écart type des durées des intervalles consonantiques divisé par la moyenne des durées des intervalles consonantiques

Tableau 24 – Métriques des 9 paramètres rythmiques de Tortel (2009 : 178)

La première phase de l'étude de Tortel a été de procéder à une analyse en composantes principales (ACP) pour visualiser son hypothèse discriminatoire entre les trois groupes natifs (GB), apprenants avancés (FR2) et apprenants intermédiaires (FR1) en fonction de leur rythme respectif suivant la dichotomie des langues en langues isosyllabiques et isoaccentuelles (voir graphique ci-dessous).

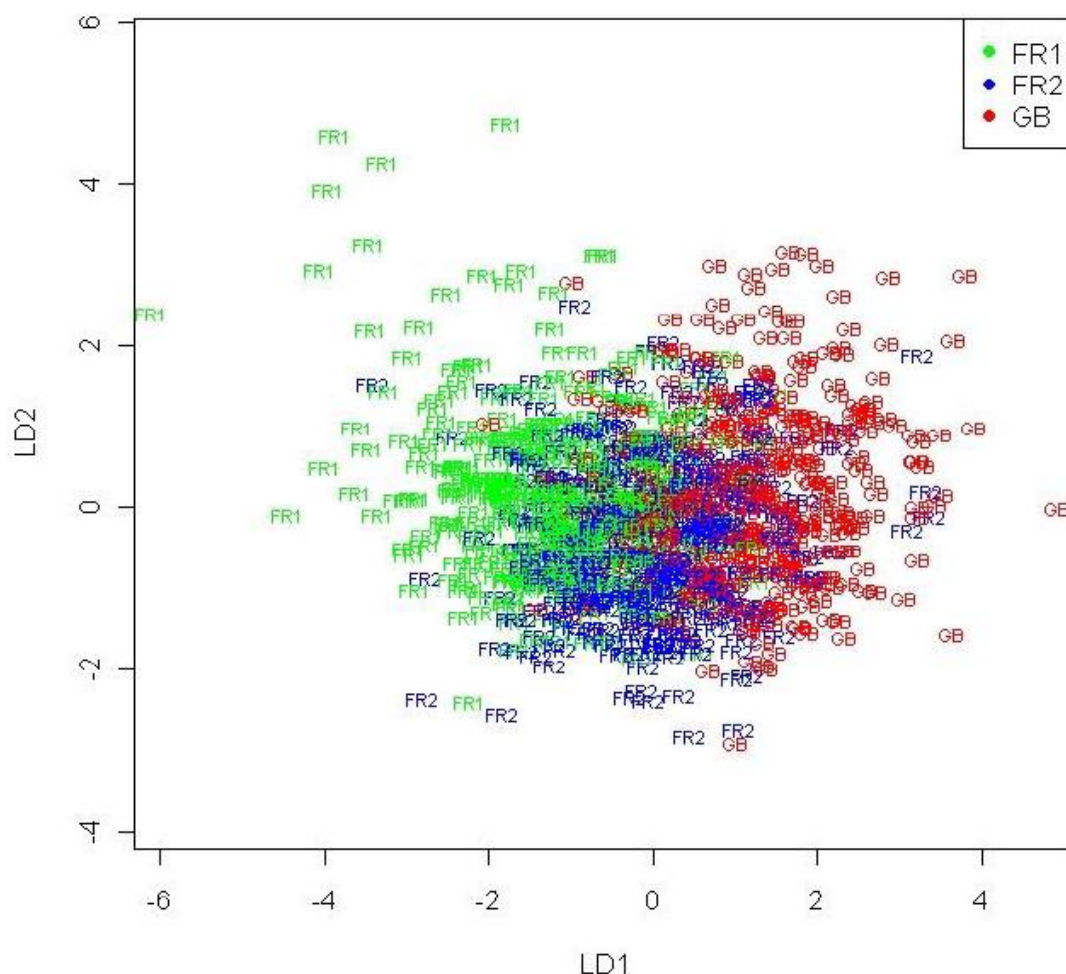


Figure 18 – Plan factoriel en analyse par composantes principales de l'analyse discriminante des 1260 phrases produites par les 63 locuteurs (Tortel 2009 : 181)

Elle a chiffré ses prédictions en regroupant les neuf paramètres rythmiques et utilise des vérifications par corrélation. Ensuite, elle a révélé les différences entre les productions natives et celles des non natifs au moyen d'analyses discriminantes (homoscédasticité ou paramétrique linéaire) des erreurs types (*standard error*) des trois groupes selon les neuf paramètres rythmiques deux par deux pour vérifier s'il est possible de déterminer avec précision les locuteurs selon les trois groupes d'appartenance. Tortel teste les métriques rencontrées dans la littérature. Voici ci-dessous un exemple des représentations graphiques provenant de l'étude de Ramus *et al.* (1999) avec les métriques %V- Δ C. La figure de gauche situe tous les locuteurs alors que celle de droite ne représente que leur moyenne par vue rapprochée de la partie centrale de la figure de droite.

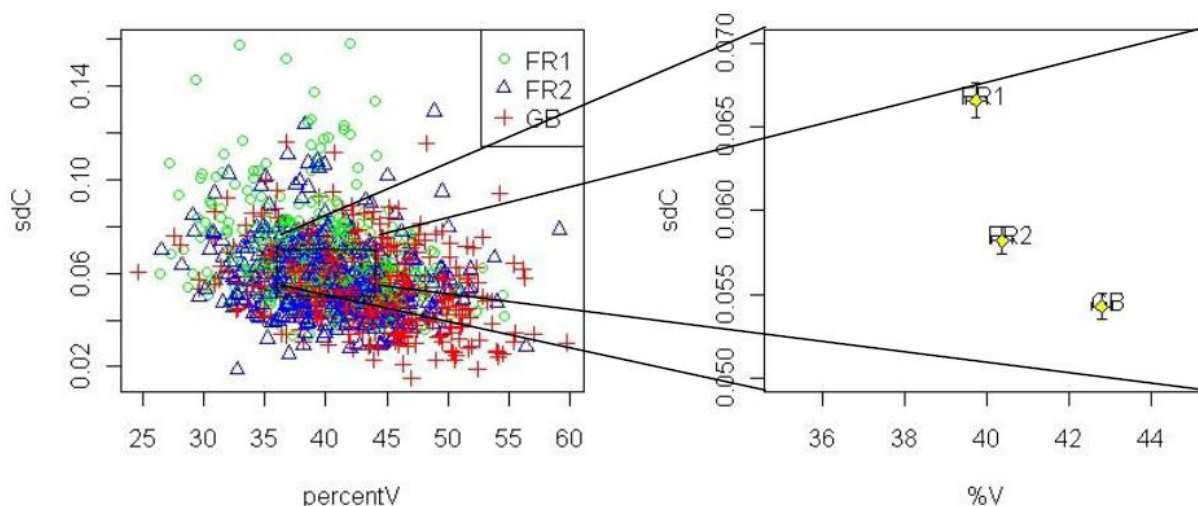


Figure 19 – Visualisation des points de dispersion (gauche) et des erreurs-types (droite) selon le plan (%V, ΔC) (Tortel, 2009 : 185)

Le paramètre %V permet de distinguer les natifs des apprenants, mais FR1 et FR2 sont trop rapprochés sur l'axe %V pour les distinguer l'un de l'autre, ce que corrobore la corrélation non significative entre FR2 et %V ($p_{val}=0,0846$). Par contre, le paramètre ΔC distingue adéquatement les deux niveaux d'acquisition de FR1 et FR2 : on peut clairement visualiser leur éloignement de l'axe de l'écart type des consonnes (sdC). En conséquence, la représentation graphique de la combinaison %V- ΔC permet de distinguer les trois groupes de locuteurs.

Après visualisation des métriques, elle les teste statistiquement pour les confirmer et évaluer leur pertinence respective dans une optique évaluative. Parmi plus de 5 000 combinaisons possibles générées par les neuf paramètres, 255 ont été calculées par analyse complète SVM (machine à vecteurs de support) grâce à un script R de Hirst pour étudier ensemble toutes les combinaisons possibles sur la base de un à quatre paramètres.

4.3.3 Bilan et discussion

La recherche d'A. Tortel (2009) s'est basée sur des paramètres rythmiques pour discriminer des locuteurs natifs et non natifs. Bien que marquant une grande avancée dans le domaine évaluatif, ses métriques ne sont pas intuitives tout en mobilisant des moyens considérables, tout en ignorant d'autres aspects du rythme, tels que les proéminences qui scandent le rythme.

Sa méthodologie consistant à visualiser les données par des analyses en composantes multiples (ACM) et plans factoriels afin de les vérifier ensuite statistiquement me semble d'un grand intérêt car elle fait appel au jugement subjectif de l'analyste dans son

cheminement pour élaborer des hypothèses de recherche. La subjectivité du chercheur est souvent occultée au détriment de l'utilisation raisonnée que l'on peut en faire dans un protocole méthodologique de recherche.

4.4 Profilage prosodique des anglicistes francophones : réalisations et accent français (Horgues 2010)

Considérons les objectifs de la recherche de Horgues et sa méthodologie pour en dresser un bilan.

4.4.1 Objectifs

Les objectifs de Horgues (2010) sont les suivants :

- Prouver l'existence d'une prosodie distinctive des apprenants anglicistes francophones,
- Décrire les corrélats acoustiques différenciant les natifs et non natifs,
- Vérifier l'influence de la langue maternelle sur les réalisations en L2,
- Trouver les conditions optimales de déviance des réalisations L2 par rapport à la norme en L1.

4.4.2 Méthodologie d'investigation

Dans son étude sur les réalisations prosodiques des anglicistes francophones et de leur perception par des auditeurs anglophones, Horgues a mis en place deux sortes d'expérimentations de trois grands volets chacune : test de production (TPro), test de perception (TPer) et analyse acoustique.

Horgues procède en deux temps : une première série d'investigations cadre un domaine d'application général de sa recherche, puis la seconde série cible des contextes contraignants pour étudier les phénomènes de manière plus approfondie.

La première série expérimentale débute avec **TPro1** pour constituer le **CORPUS1** (répétition, lecture de phrases isolées et d'un texte suivi, conversation avec un locuteur anglophone, par 12 apprenants anglicistes francophones de niveau intermédiaire-avancé). Ensuite, en se basant sur ce corpus, le test de perception **TPer1** fait écouter à 18 auditeurs anglophones (non enseignants et non linguistes) des stimuli provenant de trois classes de locuteurs (anglophones, francophones ou d'une autre L1) présentés sous quatre formes (originaux, filtrés, resynthétisés et monotonisés). Une analyse acoustique compare ensuite les corrélats

qui différencient les réalisations des natifs anglophones de celles des natifs francophones en anglais dans leur rythme et leur mélodie.

La deuxième série de tests expérimentaux s'est focalisée sur des mots clefs afin de comparer leur variation dans la lecture d'un anglophone natif et d'un francophone dans des contextes contraints. Le test **TPro2** a constitué le CORPUS2 pour favoriser l'occurrence de mots dans des contextes contraints et un dialogue interprétés par un anglophone et un francophone. Dans le test **TPer2**, les auditeurs anglophones doivent déterminer la syllabe accentuée de mots dépourvus de leur contexte. **TPER3** utilise la resynthèse pour confirmer le rôle des déviations temporelles et mélodiques dans la perception de l'accent français.

Horgues rappelle la différence entre les styles de parole en conversation spontanée et en lecture : leur prosodie diffère par la différence de leurs mécanismes cognitifs et de leurs enjeux communicatifs. Elle introduit alors le test **TPer4**, symétrique du test TPer1, mais avec conversation spontanée filtrée. Il fait intervenir les conversations de quatre apprenants. Le test **TPer5** explore une éventuelle correspondance du degré d'accent français avec celui d'aisance (*fluency*) .

À chaque étape de ses expérimentations, Horgues a évalué statistiquement les résultats de sa recherche.

4.4.3 Bilan et discussion

L'étude de Horgues tente de couvrir l'ensemble des domaines liés à la perception de l'accent français grâce à une classification très fine des phénomènes dont elle rend compte à chaque étape sous diverses formes dont statistique.

Le cycle production, perception, analyse acoustique permet de concevoir les répercussions de chaque mode phonétique d'abord dans leur globalité, puis d'en organiser une étude plus ciblée. Comme Gut (2009), elle discrimine systématiquement les styles de parole et les divers aspects prosodiques étudiés, sans toutefois bien discerner ni les locuteurs natifs entre eux, ni les apprenants francophones : la variation obtenue n'est pas systématiquement exploitée, bien qu'elle montre par ses commentaires qu'elle y est très attentive. Néanmoins, les résultats obtenus sont de la plus grande valeur en référence à ma propre recherche.

4.5 Profilage suprasegmental par analyse automatique (Ferragne 2013)

Comme précédemment pour les autres recherches, voyons les objectifs assignés et la méthodologie utilisée par Ferragne avant d'en faire le bilan.

4.5.1 Objectifs

Ferragne (2013) souligne le caractère crucial du rythme dans la détection d’un accent étranger et propose de rendre compte des productions d’apprenants anglicistes francophones par analyse automatique du rythme et d’autres traits suprasegmentaux comme l’étendue de registre ou l’intensité.

Pour cela, il se fixe deux objectifs (2013 : 154). Le premier est d’établir une comparaison entre, d’une part, l’annotation manuelle d’intervalles vocaliques et consonantiques, et d’autre part, une segmentation entièrement automatisée. Le deuxième objectif est de classer automatiquement les résultats obtenus pour catégoriser apprenants francophones et natifs anglophones.

4.5.2 Méthodologie d’investigation

Afin de calculer automatiquement les paramètres rythmiques et de registre pouvant contraster les réalisations suprasegmentales natives et celles d’apprenants, Ferragne utilise une partie du corpus *ENGLISH* (Tortel 2009, voir plus haut, ou [2008] 2013), c’est-à-dire les fichiers de lecture du groupe de 23 locuteurs natifs d’une part (GB) et des 20 apprenants ayant un niveau d’anglais L2 du Baccalauréat d’autre part (FR1). L’annotation manuelle est fournie avec le corpus *ENGLISH*. Le profilage consiste à extraire une moyenne des données (inclusion) de chaque paramètre afin de la contraster avec d’autres moyennes (exclusion) de même type. L’individu n’a de signification que dans sa relation au groupe auquel il appartient ou auquel il s’oppose.

Ferragne établit une comparaison de la durée de lecture dans les groupes GB et FR1 pour chacun des textes de lecture. Il limite l’étendue de la plage de variations mélodiques à l’écart interquartile de F0, qu’il calcule automatiquement avec *Praat*³⁹ en balisant la F0 à 75 et 400 Hertz pour chaque locuteur et chaque passage de lecture, qu’il convertit en demi-tons. Il justifie son choix d’utiliser l’écart interquartile des données recueillies plutôt que l’écart type par la difficulté d’obtenir des mesures fiables de F0.

L’analyse de l’étendue de registre montre que son écart interquartile dépend à la fois de la différence natif-non natif, du genre du locuteur, ainsi que du style de texte. Ferragne reste confiant dans la robustesse de l’écart interquartile pour mesurer le registre bien que l’un

³⁹ [<http://www.fon.hum.uva.nl/praat/>], dernièrement consulté en 2017.

de ses résultats, 32 demi-tons, équivalent à trois octaves, lui semble excessif dans le recueil de mesures d'un des textes.

Pour comparer les intervalles vocaliques dans les groupes GB et FR1, il corrèle les résultats de l'algorithme de Pellegrino & André-Obrecht (2000) de repérage des voyelles à ceux de l'analyse manuelle de détection des voyelles et montre que l'algorithme tend à détecter plus de voyelles que l'annotation manuelle. Examinant les frontières des intervalles, Ferragne (2013 : 158) vérifie statistiquement (test de Wilcoxon) que la durée moyenne de chaque méthode est significative et utilise le test de Fisher pour appuyer sa conclusion sur la différence d'écart interquartile.

Ferragne utilise la variable PVI (*Pairwise Variability Index*) pour mesurer l'intensité (au lieu de la durée) et appelle *intensity PVI*s ou PVI_i cette différence moyenne entre toutes les paires de voyelles successives dans un énoncé.

Pour atteindre son deuxième objectif de classification automatique, Ferragne applique des régressions logistiques par variables binaires à chacun des textes et utilise la validation croisée pour estimer le taux d'erreur moyen.

La méthodologie de Ferragne se base sur une démonstration argumentée étayée à tous les niveaux de vérifications statistiques. Elle se caractérise par le souci d'effectuer une recherche cohérente et utile, qui fait sens sans se limiter au seul domaine phonétique.

4.5.3 Bilan et discussion

Lors de l'extraction des mesures de F0, Ferragne ne prend en compte que l'écart interquartile pour éviter les erreurs de détection automatique : il filtre ainsi une partie des données possibles en ciblant son objectif. Par ailleurs, toujours pour étudier la F0, il subdivise le corpus de lecteurs selon le sexe. De même, dans les mesures automatiques des voyelles, il signale qu'il y a plus de voyelles repérées automatiquement que manuellement (2013 : 159). En conséquence, il décide d'exclure les mesures des voyelles de son étude (2009 : 165). Le profilage opéré est organisé en amont des mesures par la création de certaines catégories par subdivision ou au contraire par une exclusion radicale. Il s'agit là d'un profilage de groupes à tous les niveaux de l'expérimentation, dont le résultat final est un profil suprasegmental des apprenants anglicistes et celui des locuteurs natifs selon les paramètres utilisés.

Ferragne aborde un profilage d’inclusion, exclusion et classement entièrement automatisé. Les variables qu’il utilise diffèrent de celles obtenues par segmentation manuelle, mais elles fonctionnent, au moins pour les deux groupes de locuteurs FR1 et GB. Selon lui (2013 : 155), un avantage de la segmentation automatique est que les décisions arbitraires en fonction de la langue étudiée font place à une harmonisation des pratiques conduisant à une élégance du système méthodologique (2013 : 155).

Ferragne dresse un profil de la prosodie des apprenants anglicistes francophones à partir des variables de durée de lecture (aisance), d’intensité et de durée vocalique (rythme), et de la plage de registre des variations mélodiques (mélodie). Le bilan que je peux tirer de sa recherche est que la durée de lecture et l’étendue de registre sont facilement transposables à mon étude. Je garderai cependant l’annotation manuelle de F0, pour obtenir précision et fiabilité maximales des mesures, certains des fichiers de mon corpus n’étant pas de qualité maximale. Je ne retiendrai pas le facteur d’intensité car mes corpus n’ont pas été élaborés en fonction de cette variable : les locuteurs étaient libres de se mouvoir pendant l’enregistrement et je ne peux garantir la fiabilité de la correspondance entre les mesures et les réalisations originelles. Le critère de durée vocalique, bien que prometteur, me semble difficile à mettre en œuvre sur l’ensemble d’un fichier par son caractère chronophage si on y inclut les vérifications nécessaires à la fiabilité des données. Des aménagements devraient pouvoir le modifier pour cibler davantage les éléments mesurés selon un profilage par inclusion et exclusion.

4.6 Profilage intonatif par étude multidimensionnelle quantitative et qualitative (Herment *et al.* 2014)

Voyons les objectifs fixés et la méthode d’investigation utilisée, puis dressons-en un bilan.

4.6.1 Objectifs

On a vu dans le chapitre précédent que l’étude de Herment *et al.* (2014) concerne une comparaison des réalisations prosodiques de questions polaires et en *wh-* en anglais étudiées chez des natifs anglophones et des apprenants anglicistes francophones à partir du corpus de lecture *AixOx*. Ils la présentent comme exemple d’application des outils d’annotation automatique que sont le logiciel *SPPAS* (Bigi 2012) et l’algorithme *MoMel* et le système de transcription *INTSINT* (Hirst 2007). Chez les natifs anglophones, la littérature assigne une intonation montante aux questions ouvertes (*wh-*) ainsi qu’aux questions polaires ou totales

(*yes/no*) sauf dans les cas d'un ton descendant, où la question semblera « plus appuyée, méthodique, très assurée et brusque ». Herment *et al.* défendent au contraire la thèse selon laquelle le schéma de base des questions ouvertes est la chute intonative sauf dans le cas où le locuteur souhaite adoucir l'exigence portée par sa question, ou qu'il opère une répétition.

En règle générale, le français est caractérisé par une montée intonative dans les questions totales alors que les questions partielles se distinguent par une chute, mais les deux types de contours se rencontrent dans les deux structures interrogatives. Herment *et al.* étayaient leur affirmation en mentionnant Beyssade *et al.* (2007), qui affirment que le choix du contour se fait selon le contexte et non la structure syntaxique des questions.

Les attentes, en matière d'expérimentation contrôlée comme l'est la lecture du corpus *AixOx*, sont que les francophones réalisent les questions totales avec un contour montant et les questions partielles avec l'un ou l'autre des contours distribués suivant le contexte.

4.6.2 Méthodologie d'investigation

L'étude en cours est basée sur quatre questions (deux en anglais et leur traduction en français) parmi les 22 que comprend le corpus de lecture de passages dans chacune des deux langues :

- Can you tell me what's on television tonight?
- What can I have for dinner tonight?
- Vous pourriez me dire ce qu'il y a à la télévision ce soir ?
- Qu'est-ce que je vais manger ce soir ?

Les lecteurs du corpus sont au nombre de 60, entre 20 et 35 ans, répartis comme ci-dessous :

- dix natifs francophones (FRFR),
- dix natifs anglophones (ENEN),
- dix anglicistes francophones de niveau indépendant (FRENB : B1-B2),
- dix anglicistes francophones de niveau expérimenté (FRENC : C1-C2),
- dix francistes anglophones de niveau indépendant (ENFRB : B1-B2),
- dix francistes anglophones de niveau expérimenté (ENFRC : C1-C2).

Dans chacun des groupes, on observe la même répartition des genres : cinq hommes et cinq femmes. Des questionnaires viennent compléter les données recueillies pour chaque locuteur.

Les réalisations des natifs sont analysées puis comparées à celle des apprenants. Les 120 occurrences sont étiquetées par un alignement multi-niveaux avec annotation en mots et phonèmes grâce au logiciel *SPPAS*, et avec tons mélodiques et valeurs de F0 selon le système *INSINT*. Une chute intonative est attendue sur la question « Qu'est-ce que je vais manger ce soir ? », justifiée par sa fonction rhétorique.

Sur un plan pédagogique, la visualisation sur spectrogramme permet aux apprenants de constater les erreurs commises et de leur en donner l'explication. L'outil utilisé dans l'analyse du corpus *AixOx* comprend à la fois la comparaison des productions de l'apprenant avec celle d'un natif grâce à la courbe de F0 afin de rendre plus performant le travail d'apprentissage prosodique.

4.6.3 Bilan et discussion

Dans le domaine pédagogique, rien ne prouve que les progrès des apprenants soient dus à l'utilisation de cette nouvelle approche. D'ailleurs l'article d'Herment *et al.* précise que la recherche est en cours, mais je note que des avancées sont déjà sensibles tant au niveau de la méthodologie de profilage, que des résultats obtenus présentés ci-dessus.

En matière de profilage, on constate qu'une remise en cause des règles édictées dans la littérature résulte de l'utilisation d'une nouvelle approche technique et méthodologique. Dans cette recherche, le corpus ne se limite pas à deux groupes (natifs de la langue cible et non natifs), il inclut aussi un groupe de langue maternelle française et un groupe d'apprenants anglophones francistes. À cela s'ajoute le point de vue exprimé dans la littérature et une différenciation des niveaux des apprenants. La vision obtenue englobe tous les points de vue, qui peuvent alors être comparés avec grand profit.

L'analyse acoustique se distingue aussi par cette volonté de regrouper et croiser les approches. Intonation et rythme sont traités conjointement, ce qui leur permet d'éclairer leurs résultats mutuels. En outre, Herment *et al.* n'utilisent pas de test perceptif pour prendre des repères acoustiques : la courbe de F0 est souvent assez nette pour mettre en évidence les ancrages du rythme et ainsi situer les proéminences sur les syllabes. Cette approche méthodologique permet une économie de moyens tout en préservant la rigueur scientifique.

Les dix réalisations natives dans chaque langue servent de repères, non pas selon une dichotomie en juste ou faux, mais en fréquence d'utilisation des contours lorsque des alternatives coexistent. Il en est de même pour les apprenants, ce qui permet de comparer les

degrés d'acquisition au cours de l'apprentissage et d'en extraire des tendances. Ainsi, la description des contours mélodiques est une hiérarchisation qualitative implicite. Herment *et al.* parviennent donc à effectuer une analyse multidimensionnelle à la fois quantitative et qualitative dans l'étude mélodique et rythmique de L1, L2 et interlangue, du français et de l'anglais.

Ce genre de profilage quantitatif et qualitatif est très adapté à un domaine aussi fluide et empreint de variabilité que l'intonation, surtout lorsqu'il s'agit de cadrer l'évolution de l'acquisition au cours de l'apprentissage en L2. L'application d'une telle méthodologie n'est envisageable que sur un corpus restreint : ici, les phrases interrogatives. Une sélection analogue permettrait de conserver les aspects essentiels de ce profilage.

4.7 Profilage du rythme par durée syllabique (Ballier, Martin, Amand 2016)

4.7.1 Objectifs

Ballier *et al.* (2016) analysent la segmentation syllabique des francophones du corpus d'apprenants anglicistes ANGLISH de Tortel & Hirst (2010). L'objectif final est de trouver une méthode d'évaluation de la qualité de l'interlangue prosodique des francophones anglicistes plus adéquate que les précédentes. Pour l'atteindre, il leur est nécessaire de trouver la meilleure manière de rendre compte du rythme et de « sortir de l'impasse des métriques », lesquelles ont fait beaucoup d'émules depuis Ramus *et al.* (1999). Ils soulignent que cela permettrait de « restituer les constituants de la hiérarchie prosodique et de montrer la pertinence de la syllabe accentuée et du pied pour l'analyse de l'anglais, tout en abordant la question de l'isosyllabité et de la structuration accentuelle ».

Même si un modèle de la langue cible est jugé nécessaire pour amorcer cette étude, Ballier *et al.* posent très clairement la question de la variabilité dans les réalisations natives et de la norme de référence pour la prononciation des non natifs.

Les deux problématiques sont donc alors : 1) Quelle est l'homogénéité des réalisations des anglophones (types de syllabes, proportion des durées) ? 2) Quelles sont les prévisions permises par l'alignement pour les réalisations des francophones (allongement et présomption de voyelles épenthétiques, proportion des durées selon les types de syllabes) ?

L’homogénéité des réalisations natives est recherchée dans une classification très affinée de la structure et position syllabique. À partir de ces 14 cas envisagés, des prédictions sont effectuées sur les interférences et transferts possibles de la langue maternelle dans les réalisations des locuteurs non natifs.

4.7.2 Méthodologie de classification

Pour proposer un différentiel natif/non natif, Ballier *et al.* cherchent à déterminer des zones de stabilité parmi les réalisations des locuteurs natifs dans la parole lue, lesquelles seraient retenues pour établir des patrons de comparaison, des « modèles » robustes peu sensibles à la variabilité. Ballier *et al.* utilisent une méthode d’alignement forcé sur des séquences identiques de texte lu afin d’en comparer la prosodie en minimisant la variabilité. Leur analyse de l’interlangue basée sur la durée syllabique prend en compte leurs propriétés distributionnelles, accentuelles et structurelles afin d’étudier la variabilité des réalisations syllabiques.

Pour cela, ils établissent un patron détaillant les réalisations des syllabes anglophones. Ensuite, après avoir découpé les réalisations des anglophones en syllabes, ils comparent les réalisations des apprenants au modèle obtenu. La pertinence des séquences est repérée grâce à la configuration du logiciel *WinPitch LTL*, dont la paramétrisation permet de repérer par colorisation des segments temporels de la F0, l’intensité ou le spectrogramme. Seule la mesure de différentiel temporel inter-syllabique est prise en compte pour comparer les réalisations des anglophones à celles de l’interlangue. Un distinguo est fait entre syllabes accentuées et syllabes inaccentuées, ce qui permet de réintroduire l’unité prosodique de pied.

L’étude sur corpus consiste en une comparaison de quatre passages lus du corpus *ENGLISH* (Tortel & Hirst 2010) : 20 anglophones natifs, 20 francophones de niveau intermédiaire (FR1) et 20 francophones en troisième année de Licence d’anglais (FR2).

La méthodologie d’analyse procède en deux temps. Deux experts ont annoté les 65 mots qui ont été sélectionnés pour établir 99 « valeurs cibles » en privilégiant le rattachement des consonnes en coda en cas de syllabe accentuée (Wells 1990). En fonction de la configuration des syllabes (CV, CVC, CVCC) et des transferts attendus chez les francophones (« réalisation d’une voyelle épenthétique en lieu de la voyelle réduite ou de la réalisation syllabique de la consonne [n] ou [l]) ainsi que leur statut accentuel dans la hiérarchie accentuelle », 14 types de syllabes ont été analysés (voir tableau ci-dessous).

type de syllabe	caractérisation	interférences et transferts possibles
Focus	syllabe obligatoirement noyau de l'unité intonative (ex: <i>do</i> d'emphase)	réalisation de contour prosodique différente
Final	syllabe en fin d'unité intonative	risque de surallongement par transfert du français (Delais, Herment <i>et al.</i>)
MaxCoda	syllabe accentuée fermée par une ou plusieurs consonnes VC(C)(C) ex : /'brɪt n/	possibilité de voyelle épenthétique, de rattachement de la coda à la syllabe suivante (['bri 'toen])
RedFinCC	syllabe réduite en VC(C)	possibilité de voyelle épenthétique
Syllabiq	consonne syllabique, sommet de syllabe (n, l ou r) ex : /'brɪt n/	possibilité de voyelle épenthétique [bri toen]
Interton	schwa en position médiale	risque d'allongement
RedFin	syllabe réduite ouverte (en milieu de mot portant un accent)	allongement (risque de transfert d'accent démarcatif final du groupe accentuel)
RedCoda	syllabe réduite en coda CVC	
happyY	syllabe inaccentuée ouverte (en finale de mot, happy /'hæp i/)	risque d'allongement
preCV	syllabe pré-tonique ouverte (<i>a</i> dans <i>along</i>)	risque d'allongement vocalique
preCC	syllabe pré-tonique fermée (<i>in</i> dans <i>inspector</i>)	risque d'allongement, moindre que pour preCV.
monCV	monosyllabe en syllabe ouverte	risque de réalisation isosyllabique, allongements des voyelles inaccentuées
monCC	monosyllabe en syllabe fermée	risque de réalisation isosyllabique, allongements des voyelles inaccentuées
Ff	forme faible : marqueur grammatical réalisé en voyelle réduite sauf en cas de focus (<i>of, the, a</i>)	risque d'isosyllabité par allongement et réalisation de timbre vocalique autre que schwa (voyelle pleine).

Tableau 25 – Ontologie des syllabes annotées dans l'étude Ballier *et al.* (2016)

La hiérarchie accentuelle adoptée comprend quatre degrés (syllabe sous accent principal, syllabe sous accent secondaire, syllabe inaccentuée non-réduite et syllabe inaccentuée réduite), lesquels ne prennent pas en compte un accent secondaire au sens de Guierre (1979), mais la qualité du timbre de la voyelle. Une hiérarchie de critères a été adoptée pour analyser les syllabes relevant de plusieurs cas différents : un codage unique tient compte en premier lieu de la propriété *focus*, puis de celle de la position dans l'énoncé, puis du nombre de ses syllabes et en dernier de la structure de cette syllabe.

Le logiciel *WinPitch LTL* (*WinPitch 2016*)⁴⁰ a permis de comparer spectres et intervalles temporels par alignement forcé. Contrairement à *SPPAS*⁴¹ de Bigi 2012, il

⁴⁰ Martin Ph. [<http://www.winpitch.com/>], consulté en 2017.

s'affranchit de modèles de reconnaissance en vigueur pour permettre une segmentation en phones et syllabes selon la division de la syllabe chez Wells. Il n'est pas mentionné quel profilage a permis d'identifier le degré de prototypicalité des fichiers natifs qui ont servi de patrons. Afin de comparer les fichiers deux par deux, Ballier *et al.* ont développé un algorithme qui permet un alignement sans tenir compte des pauses.

Dans l'étude statistique, pour analyser la variance de certains types de syllabes, ont été utilisés les tests statistiques d'Anderson-Darling (non probant), puis celui de Kruskal-Wallis pour comparer les médianes de divers groupes (alternative non-paramétrique à l'ANOVA d'ordre Un), lequel est basé sur des rangs. Ensuite, des tests de comparaison multiple deux à deux ont permis de rendre compte des durées moyennes des syllabes les plus caractéristiques dans les réalisations des anglophones natifs afin d'opposer les types syllabiques par leur durée. La classification hiérarchique ascendante souligne les différences ou l'indistinction entre les types de syllabes.

Selon Ballier *et al.* cette approche fonctionnelle est plus pertinente que celle d'A. Tortel (2009) car elle repose sur les entités linguistiques attestées que sont les syllabes, ce qui n'est pas le cas des métriques intervocaliques et interconsonnantiques utilisées par les héritiers de Ramus *et al.* (1999) afin de discriminer le niveau de maîtrise des locuteurs.

4.7.3 Bilan et discussion

Cette étude sur la méthode d'évaluation du rythme des réalisations des apprenants anglicistes ne peut avoir que de nombreux échos dans ma propre recherche.

Le problème du choix du modèle, des patrons de comparaison, des repères acoustiques considérés comme la norme de référence, est soulevé, mais la question de la variabilité ne reçoit pas toujours de réponse claire. Si la diversité des styles de lecture est testée sur quatre passages lus, et la variabilité parmi les natifs anglophones sur 20 lecteurs, la seule indication donnée est leur temps de résidence en France réduit. Les annexes d'A. Tortel (2009 : 21-22) en précisent le sexe parmi les métadonnées. Après croisement des données, on peut supposer que sur les 13 locutrices natives d'A. Tortel (2009), les trois ayant le plus longtemps séjourné en France ont été écartées. L'observation du tableau montre que ce sont quasiment les moins jeunes (45, 44 et 28 ans), dont les durées de séjour en France sont respectivement de 17, 10, et 4 ans. Dans le cas des dix hommes, les durées de séjour en France peuvent

⁴¹ Bigi, B. [<http://www.sppas.org/>], consulté en 2017.

atteindre 10, 29 voire 39 ans et leurs âges respectifs sont 58, 55 et 61 ans. Il est probable que tous les dix aient été gardés pour constituer le corpus de dix sujets masculins. La moyenne d'âge du sous-corpus féminin (devrait avoisiner 22,8 ans) sera donc bien plus basse que celle du sous-corpus masculin (35,4 ans). Les sujets des deux sexes ont un niveau minimum d'études « bac + 2 ». Dans ce cas, la distribution de la variable « âge » serait limitée et pourrait représenter un biais pour l'analyse. Plus tard, afin de tester l'alignement des segments annotés, l'analyse est conduite à partir « d'un fichier natif identifié comme particulièrement prototypique », puis d'un deuxième dont on peut penser qu'il obéit aux mêmes particularités. Le manque de précision à ce sujet conduit à penser que ce sont les syllabes de ces 20 lecteurs qui ont subi un profilage méthodique, plus que les lecteurs eux-mêmes, sauf si l'on considère la probable exclusion des lecteurs natifs dont le temps de résidence aurait fortement excédé un an. Les tableaux dans les annexes de Tortel précisent en outre le lieu de résidence des locuteurs jusqu'à l'âge de 18 ans, qui détermine le géolecte⁴² le plus marquant) : on constate que les locuteurs sont tous d'origine britannique, ce qui réduit en conséquence la variabilité géolinguistique. Les sociolectes (Bavoux 1997, 265⁴³) sont aussi susceptibles d'influer sur les réalisations syllabiques. Les caractéristiques des locuteurs peuvent ainsi orienter les propriétés des données analysées et biaiser les résultats obtenus car la variabilité des données acoustiques obtenues peut résulter de la variabilité sociolinguistique des locuteurs.

Le prérequis pour minimiser la variabilité est qu'il faut au préalable « déterminer des zones de stabilité des réalisations des natifs ». L'outil de profilage utilisé dans la recherche d'homogénéité des syllabes est une méthode d'alignement forcé des textes lus, et le résultat obtenu est une classification en 14 types. Une première étape est d'exclure les syllabes non porteuses de voyelles (consonnes syllabiques), pour obtenir une homogénéité des résultats. L'objectif étant de mesurer la durée syllabique (différentiel inter-syllabique), le profilage consiste en un travail d'annotation pour appliquer une série de critères afin de classer les syllabes en fonction leur accentuation, de leur position et de leur structure. Certaines syllabes correspondant à plusieurs critères, une hiérarchie des critères a été élaborée, ainsi qu'une hiérarchie accentuelle de quatre degrés. La typologie des 14 syllabes étudiées repose sur deux critères : leur caractérisation (étude des natifs) et les possibles effets sur l'interlangue (prévisions sur non natifs). Ces interférences prévisibles concernent essentiellement un

⁴² « En linguistique, dialecte propre à une zone géographique ». Définition de l'*Encyclopédie Universalis* [<http://www.universalis.fr/dictionnaire/geolecte/>], consultation en 2016.

⁴³ Définition de BAVOUX, Claudine dans MOREAU, Marie-Louise (éd.) (1997), *SOCIOLINGUISTIQUE : Les concepts de base*. Sprimont : Mardaga.

CHAPITRE II – Profils et profilage : mode d’organisation hiérarchique de la prosodie des anglophones natifs

contour prosodique différent, un allongement de la syllabe, une voyelle épenthétique ; les deux dernières réalisations résultant du phénomène d’isosyllabicit . Les réalisations prosodiques  tant par essence d’une grande variabilit , je suis d’avis qu’il est essentiel de la ma triser autant que faire se peut en faisant reposer les analyses sur des zones de faible variabilit  o  les réalisations de natifs seront stables.   partir de l’ tude de ces formes, les mesures acoustiques devraient permettre de cat goriser puis de classer les r sultats, ce qui serait un guide, un mod le pour les r alisations des apprenants qui seraient  talonn es. L’op ration de profilage consistera donc    carter les zones instables, puis, en fonction des r sultats acoustiques obtenus,    tablir des cat gories permettant de classer les r sultats (travail d’annotation) et ensuite d’essayer de les interpr ter en leur donnant une valeur hi rarchique les unes par rapport aux autres. Comme l’ont fait Ballier *et al.*, les profils de marqueurs r sultant du profilage des r alisations natives serviront de bases de comparaison aux productions non natives, lesquelles seront aussi balis es par les interf rences attendues par rapport   la langue cible.

Le d coupage syllabique utilis  dans cette  tude est bas  sur une conception non consensuelle.  tant donn  la comparaison n cessaire des deux langues que sont le fran ais et l’anglais, l’application d’un principe analogue me semble pr f rable. Je choisirai pour les deux langues un d coupage qui suit le m me principe de l’attaque maximale (MOP) et opterai pour celui du *Cambridge English Pronouncing Dictionary – 18th Edition (CEPD18) on CD-ROM* (Roach *et al.* 2011).

Une premi re  tude s lectionne des « valeurs cibles » peu sensibles   la variabilit , une deuxi me  tude  tablit les diff rences entre r alisations natives et non natives. L’outil utilis  pour comparer les r alisations non natives   celles des natifs est le logiciel *WinPitch LTL*. Son fonctionnement  tablit un distinguo sur deux canaux : l’analyse acoustique est pr sent e sous forme num rique ; elle est aussi doubl e d’un signal visuel par repr sentation graphique color e. L’utilisateur, qu’il soit chercheur ou apprenant, est doublement sensibilis  et par sa perception subjective, et par les r sultats sous forme num rique. La prise en compte de la perception subjective de l’analyste au m me titre que la n cessit  de fournir des r sultats objectifs et tangibles doit  tre soulign e car elle est de nature   aider le chercheur   formuler des hypoth ses et contribuer   y r pondre.

Leur profilage qu’ils d finissent par « approche paradigmatique » («  v nements prosodiques produits sur un intervalle donn  ») et « syntagmatique » (comparaison des

différences de durée sur des empan comparables) permet d'élargir le champ de recherche sans rester limité aux questions de temporalité. Bien que les pauses puissent revêtir une grande importance en prosodie, il est important qu'elles n'occultent pas d'autres phénomènes du domaine étudié. Ainsi, le fait de pouvoir les écarter dans la prise de certaines mesures ouvre des perspectives intéressantes.

Quant au profilage sur le plan statistique, le calcul de l'écart inter-annotateur valide l'homogénéité des annotations dans la mesure où plusieurs annotateurs interviennent. Par ailleurs, divers tests statistiques (Kruskal-Wallis, test de comparaison multiple) ou graphiques statistiques (classification hiérarchique ascendante) permettent d'obtenir des résultats éclairants susceptibles d'être utilisés pour valider ceux obtenus dans ma propre étude.

Les résultats eux-mêmes sont quelque peu décevants, même si le seuil de valeur observée est très restrictif. Une hiérarchie des durées peut néanmoins être établie. La différenciation par genre est peu surprenante si l'on considère que les femmes anglophones parlent souvent sur un registre très ample, ce qui prend plus de temps que de garder un registre étroit. Cette distinction est à prendre en compte. La différence de durée entre les syllabes inaccentuées et les syllabes sous l'accent ou finales est d'un grand intérêt, de même que le manque de pertinence des syllabes sous focus pour ce type d'analyse ou celui de la position de la syllabe réduite.

Le profilage mis en œuvre dans cette recherche a aussi le mérite de reposer sur des entités linguistiques attestées, ce qui accroît son élégance.

4.8 Profilage évaluatif de la qualité lexicale (Laufer et Nation 1995)

Bien que destinée à une mesure linguistique de l'écrit, cette recherche basée sur la mesure de la fréquence lexicale revêt un intérêt par sa méthodologie plus que pour ses résultats, qui sont assez éloignés de la présente étude. Elle est étroitement reliée à un profilage destiné à une évaluation de la richesse lexicale en Langue 2 par profilage de groupes et en cela se rapproche de mon objectif d'évaluer la qualité prosodique des apprenants anglicistes francophones.

4.8.1 Présentation de la méthode de profilage par fréquence lexicale

Dans leur article *Vocabulary Size and Use: Lexical Richness in L2 Written Production*, Laufer et Nation (1995) présentent ce qu'ils nomment le *Lexical Frequency*

Profile (LFP), qui est en fait une nouvelle manière de mesurer la richesse lexicale en production écrite des apprenants :

The LFP shows the percentage of words a learner uses at different vocabulary frequency levels in her writing—or, put differently, the relative proportion of words from different frequency levels. (1995 : 311)

Ce profilage (1995 : 311-312) se fonde sur une répartition des types de mots selon leur fréquence d'utilisation dans la langue, lesquels sont catégorisés en tranches de 1 000 mots appartenant à un registre basique ou plus recherché. Le pourcentage d'occurrence du lexique correspondant à chaque tranche indique le niveau de maîtrise lexicale de l'apprenant lorsque les mots de son écrit sont mis en relation avec les trois listes de mots du programme informatique, le coffret comprenant le logiciel et les listes de mots s'appelle le *VocabProfile package*. Le lexique de nature universitaire (*UWL*, sigle d'*University Word List*) est lui aussi pris en compte, ainsi que les mots non répertoriés dans le système. Le résultat du test informatique se présente sous la forme de quatre pourcentages correspondant à la répartition des mots trouvés dans la composition de l'apprenant par rapport aux catégories pré-établies (premiers 1 000 mots les plus fréquents, seconds 1 000 mots les plus fréquents, liste de lexique universitaire (*UWL*) et mots non répertoriés. Bien que le logiciel distingue et prenne en considération dans la création des profils les occurrences, les types ainsi que les familles de mots, ce sont ces dernières que Laufer et Nation préfèrent pour calculer la richesse du lexique, choix justifié par un critère pragmatique : la distinction qui fait sens chez l'apprenant :

It is the latter calculation [word families] that we consider more revealing as an indication of lexical richness, because it uses a definition of what should be counted as a word which most closely matches how learners view words. (1995: 312)

D'après Laufer et Nation (1995: 312-313), cette nouvelle méthode pour mesurer la richesse lexicale par calcul de la fréquence d'occurrence est supérieure aux précédentes par son objectivité (*it is independent of the learner's environment*), son autonomie (*largely independent of syntax and text cohesiveness*), sa précision (*a more detailed picture of the different types of words (as defined by frequency levels)*), [...] *it can be used when comparing groups from different educational systems, or groups acquiring language outside the classroom*), sa mesure fine du degré de maîtrise (*The LFP will discriminate between subjects who use frequent and less frequent vocabulary*) et sa fiabilité (*A word used incorrectly is not considered to be part of the learner's lexicon*). Ils précisent en outre que la validité du

profilage par fréquence lexicale (LFP) doit être vérifiée par similarité des résultats obtenus avec divers extraits de production écrite d'un même apprenant, mais aussi avec ceux fournis par d'autres méthodes de mesure. C'est ce qu'ils mettent en application dans une expérimentation.

4.8.2 Expérimentation portant sur la validité et la fiabilité du LFP

L'objectif poursuivi par Laufer et Nation dans cette expérimentation est de démontrer la fiabilité et la validité du Profilage par Fréquence Lexicale (*LFP*) pour mesurer la qualité lexicale en production écrite. Trois groupes d'apprenants de différents niveaux totalisant 65 individus ont leurs productions écrites comparées (300 premiers mots de productions écrites ne devant pas dépasser les 350 mots).

Les deux questions de recherche posées pour vérifier la validité du profil sont :

- a Will there be a significant difference between the LFPs of learners of different language proficiency levels?
- b Will the LFP of the compositions correlate highly with the scores of the same learners on the active version of the Vocabulary Levels Test? (1995: 314)

Pour répondre à la question a, la validité du profilage par fréquence lexicale est établie en comparant la moyenne des pourcentages de résultats et l'écart-type de deux rédactions par apprenant et de chaque groupe au moyen d'une ANOVA (pour comparer les moyennes des trois groupes à chaque niveau de maîtrise) et du test de Fisher (pour en évaluer le degré de signficativité). Les résultats montrent que les étudiants les plus faibles ont davantage tendance à utiliser les 1 000 mots les plus fréquents dans la langue. La deuxième tranche des mots les plus fréquents sera plus utilisée par les étudiants faibles. Le lexique sophistiqué, universitaire et non répertorié, renforce les différences entre chaque groupe. Laufer et Nation rapportent leur analyse à une définition générale du concept de maîtrise langagière et en déduisent la validité de leur profilage selon ce critère :

These differences are in accordance with the concept of language proficiency which assumes that richer vocabulary is characteristic of better language knowledge. If the LFP has tapped these differences, this is evidence for its validity. (1995: 316)

Une seconde analyse est effectuée pour répondre à la question b en vérifiant l'adéquation des résultats du test de référence *Vocabulary Levels Test* avec le *Lexical*

Frequency Profile. Les résultats des tests de corrélation sont jugés éclairants pour la tranche des 1 000 premiers mots les plus fréquents (corrélation négative) mais non concluants pour la seconde tranche des 1 000 mots.

En ce qui concerne la fiabilité du profilage, les questions de recherche sont :

c Will the LFPs in two sets of compositions written by the same learners correlate highly with each other?

d Will the percentages of words at each frequency level correlate highly with each other in the two sets of compositions?

Pour répondre à la question c, Laufer et Nation privilégient une analyse intra-locuteur (*within-subject analysis*) pour comparer les résultats de deux productions d'un même apprenant avec le *t-test* de Student pour données appariées (comparaison de deux résultats émanant des mêmes individus) afin de vérifier la fréquence lexicale, et MANOVA pour rendre compte des proportions entre les niveaux de maîtrise. Les résultats révèlent une stabilité du profil dans les deux groupes les moins avancés alors que le vocabulaire du groupe avancé semble dépendre davantage du sujet traité.

Un autre procédé de vérification de la fiabilité du profilage consiste à écarter des calculs les 1 000 mots de la première tranche étant donné qu'ils ne représentent pratiquement que les mots outils grammaticaux ainsi que les mots les plus basiques de la langue. En conséquence, ils ne peuvent être aucunement indicatifs d'une quelconque qualité lexicale car ils constituent un bagage nécessaire à tout étudiant de niveau avancé. Dans cette nouvelle distribution, les analyses montrent qu'aucune des différences entre les deux compositions n'est significative, ce qui démontre une stabilité du profil indépendamment du sujet traité dans la production écrite, plaçant ce profilage en position de choix de fiabilité pour analyser la qualité lexicale à tous les niveaux de maîtrise et en tout cas le plus fiable pour analyser les écrits des niveaux avancés.

Par la suite, Laufer répond aux critiques de Meara (Laufer 2005 : 584-585) en ajoutant les résultats de corrélations au *t-test* de Student pour données appariées étant donné qu'il réfutait la fiabilité du profil sur cette base. Laufer met néanmoins en garde contre une confusion entre l'étendue du vocabulaire d'un apprenant et l'utilisation qu'il en fait en production car leur évolution n'est pas similaire. De même, elle insiste sur la nécessité d'une approche de terrain qui, seule, permet de valider ou d'invalider des hypothèses :

[...] all pseudo-experiments are inconclusive unless supported by real data, collected from real learners performing real language tasks. (2005: 586).

4.8.3 Bilan et discussion

Ce profilage de groupe repose sur la définition de l'unité servant de mesure : le mot. Ici ce sont des considérations didactiques et pragmatiques qui déterminent les choix de classification linguistique puisque les catégorisations lexicales faisant sens pour les apprenants prennent le pas sur d'autres possibilités utilisées en linguistique. Par ailleurs, Laufer (2005) insiste aussi sur l'importance de l'usage et de l'approche de terrain, laquelle est inconcevable sans corpus, dans toute expérimentation probante en L2.

Cette évaluation de la qualité de la production se fait de manière précise et objective grâce à des mesures quantitatives automatiques. La mesure du degré de maîtrise lexicale est double : elle est établie non seulement par le critère d'occurrence ou de non occurrence d'un mot, mais aussi par la fréquence d'occurrence de ce mot. Ainsi que Laufer l'affirme clairement (2005 : 583), *An LFP is basically a measure of lexical use in writing*. Il ne s'agit pas de mesurer les connaissances des apprenants, mais leurs choix d'utiliser un lexique fréquent ou recherché, le rendu se faisant sous forme de proportion (par pourcentage). Cette méthode d'analyse a fait de nombreux adeptes en acquisition des Langues 2, que ce soit dans ses principes (attestés comme le fait Chitez 2014, ou subissant une influence probable comme chez Hawkins et Buttery 2010) ou encore plus carrément dans la création de logiciels en ligne (Tom Cobb's website⁴⁴ ou projet ADELEX⁴⁵).

Cette étude montre en outre l'importance que peut revêtir le choix des tests statistiques afin d'asseoir sa démonstration. Si ANOVA, MANOVA, t-test et test de Fisher sont acceptables, les corrélations semblent être les tests statistiques les plus convaincants. Écarter de l'analyse la tranche des 1 000 mots les plus fréquents est bienvenue pour simplifier l'analyse, cependant il n'en va pas de même du choix des tests statistiques : écarter un test de corrélation favorise une remise en cause d'une démonstration.

⁴⁴ COBB, Tom, *Compleat Lexical Tutor v.8* : [<http://www.lextutor.ca/>], consulté en 2016.
[<http://www.lextutor.ca/vp/eng/>] consulté en juillet 2016.

⁴⁵ Projet de recherche ADELEX (*Assessing and Developing Lexical Competence through the Internet*) de l'Université de Grenade ; ADA ADELEX ANALYSER :
[<http://www.ugr.es/~inped/ada/ada.php?ada=nttpbaat6jroctoobk8lsd9qv2&lng=english>];
ADELEX ANALYSER Lexical Profile :
[<http://www.ugr.es/~inped/ada/perfil.php?ada=nttpbaat6jroctoobk8lsd9qv2&lng=english>] consulté en 2016.

Ces critères d'analyse de profils de groupes me semblent particulièrement pertinents pour une éventuelle transposition dans l'étude de domaines de grande variabilité et qualitatifs, en particulier celui de la mélodie.

5 Bilan des phénomènes de profils et de profilage

L'opération de profilage est vue comme le processus méthodologique permettant de créer l'objet qu'est un profil spécifique, et le profil d'un individu sa caractérisation. On a vu précédemment que les notions de profil et profilage prennent leur sens dans une situation spécifique qui est à préciser en fonction de l'objectif à atteindre. Aussi, les diverses procédures repérées en amont vont servir à orienter et déterminer la méthodologie destinée d'abord au profilage des locuteurs natifs, puis dans un second temps celui des apprenants. Les exemples de profils et profilages présentés dans ce chapitre ont permis de sonder le rôle d'un profil. À la croisée des domaines pratiques et théoriques, les profils doivent répondre précisément aux besoins concrets d'une situation spécifique à un moment donné avec la méthodologie adaptée mise en œuvre pour les créer, c'est-à-dire le profilage. La description de chacun d'entre eux a consisté à présenter des résultats essentiels pour la présente recherche, mais a aussi fourni un éventail de questionnements et de pratiques méthodologiques pour y répondre.

Ainsi, dans le cadre de cette étude, un profilage de natif ne peut être que différent d'un profilage de non natif étant donné que les objectifs à atteindre pour ces deux catégories de locuteurs sont différents. L'objectif final de cette thèse est d'élaborer des critères fiables d'évaluation de l'interlangue prosodique. Les profils prosodiques non natifs sont destinés à atteindre cet objectif en contribuant à comprendre leur fonctionnement prosodique afin de les catégoriser selon leur degré d'acquisition. Ce dernier se caractérise par la meilleure adéquation possible à une norme, le modèle natif. Or, si dresser le profilage des non natifs signifie mesurer la distance entre leurs réalisations et celles de natifs, quel profil prosodique natif faut-il prendre comme modèle ? On peut difficilement opérer une évaluation précise et fiable en gardant une variabilité rendant un modèle évaluatif impossible à concevoir parce que trop flou dans ses implications acoustiques. Le profilage des non natifs nécessite comme norme un profil natif minimisant la variabilité.

La plupart des études décrites en amont se basent sur une norme (norme des locuteurs sains et adultes, norme des locuteurs natifs, norme énoncée dans la littérature, etc.). Cette

norme établit une base de comparaison servant d'étalon, que ce soit en profil personnalisé ou en profil de groupe, pour faire ressortir la particularité des traits étudiés. Les expérimentations décrites ont pu témoigner de la grande disparité du nombre de locuteurs natifs sélectionnés comme norme. Les réalisations acoustiques de ces locuteurs sont souvent réduites à une moyenne, le profil de groupe obtenu effaçant l'individualité de leur nature. Dans un cadre évaluatif, la notion de norme est cruciale. La question qui se pose alors est simple : comment établir une norme prosodique pour qu'elle soit fiable quantitativement et qualitativement ? En bref, pour les natifs, comment passer d'un profilage descriptif individuel à un profilage normatif hiérarchisable suivant la qualité de lecture ? Pour répondre à cette question (incongrue dans la littérature car elle en est officiellement absente, mais pourtant insidieusement présente lorsqu'on observe les données métalinguistiques de locuteurs natifs instruits recueillies), il me paraît nécessaire de trouver la qualité maximale en lecture à haute voix pour modéliser la norme par rapport à laquelle les réalisations acoustiques de non natifs seront étalonnées.

On a déjà vu dans le premier chapitre comment la durée de lecture avait été préalablement choisie afin de hiérarchiser les réalisations des apprenants. Ce choix repose sur des données quantitatives et il est justifié par une approche probabiliste pour déterminer une norme qualitative. La finalité évaluative se doit d'être qualitative et le processus de mise en œuvre pour parvenir à atteindre l'objectif est quantitatif par souci de précision et d'objectivisation. Quantitatif et qualitatif sont donc intimement liés et doivent mutuellement s'étayer et se corroborer. La création d'une norme native devant modéliser un système évaluatif de la prosodie non native nécessite un profilage complexe que je mets en place au moyen de l'expérimentation qui suit.

Un profilage étant un processus *ad hoc*, il me paraît plus adapté de présenter l'adaptation des notions que j'utiliserai lors de leur mise en place circonstanciée.

6 Profilage de la norme native en lecture

Lorsque l'on cherche à profiler une norme native en lecture, les objectifs fixés à cette expérimentation sont :

- D'obtenir des repères prosodiques acoustiques de la qualité native en lecture,
- D'établir une hiérarchie entre les variables pour sélectionner celles qui sont les plus pertinentes,

- Réduire la variabilité prosodique de la norme,
- Et ainsi valider une norme aux paramètres clairs, fiables et restreints.

Pour y parvenir, j’aurai recours à la création de profils prosodiques d’anglophones natifs en lecture. Ces profils seront à la fois quantitatifs (pour modéliser les mesures acoustiques) et qualitatifs (pour être hiérarchisables). Ils ne peuvent être que des profils d’ensemble, regroupant une multiplicité de variables puisqu’en prosodie, elles fonctionnent en synergie. Outre le travail de constitution d’un corpus de locuteurs natifs de l’anglais dans les normes (pour chacun des 42 locuteurs : sollicitation, acceptation de participation écrite⁴⁶, recueil de métadonnées, enregistrement(s), codage des fichiers), ce profilage complexe suit plusieurs phases :

- Profilage perceptif de lecteurs (expérimentation PER-LEC-N),
- Analyse des commentaires de lecteurs sur leur prestation et leur expérience dans le domaine (expérimentation EXP-LEC-N),
- Profilage acoustique des lecteurs (COR-AC-N1),
- Profilage statistique des lecteurs avec hiérarchisation (expérimentation STAT-N1).

Ces étapes s’étayent mutuellement pour pouvoir confirmer ou infirmer leurs résultats. Ainsi, un bon ou mauvais lecteur sera repéré par test perceptif (donc subjectif), l’analyse acoustique de sa lecture (objective) sera comparée aux autres résultats et l’ensemble des données pourra être évalué statistiquement pour vérifier les données de départ.

6.1 Profilage prosodique qualitatif par test perceptif (expérimentation PER-LEC-N)

6.1.1 Objectif de l’expérimentation

Ce test perceptif est destiné à catégoriser et hiérarchiser divers styles de lecture d’anglophones natifs sur un plan qualitatif. Prenant en compte l’approche intersubjective de l’interlocution dans la lecture, je considère la notion d’asymétrie dans cette relation : le lecteur occupe un rôle de premier plan alors que l’auditeur (virtuel lors de l’enregistrement) n’apparaît que dans la méthodologie du test.

⁴⁶ Consentement écrit à l’exception de plusieurs locuteurs ayant librement donné accès à leurs lectures du conte *The Selfish Giant* sur Internet.

6.1.2 Méthodologie du profilage

Ce test consiste en l'évaluation subjective de la lecture des cinq premières phrases du texte *The Selfish Giant* par douze anglophones natifs évalués par 52 auditeurs : huit anglophones natifs, huit enseignants d'anglais francophones, 36 étudiants anglicistes (douze de niveau B1+/B2, neuf de niveau B2 et quinze de niveau B2/C1⁴⁷). La diversité de niveau de maîtrise des auditeurs est justifiée par l'objectif de pouvoir profiler les jugements subjectifs selon ce critère.

Les critères d'évaluation pour déterminer une typologie prosodique de « bons » et « mauvais » lecteurs ont été recueillis par questionnaires portant sur la clarté de l'information véhiculée et sur la capacité à transmettre un contenu émotionnel sous forme de vote, de questions à choix multiples, de questions à échelles de valeurs et de questions ouvertes. Les auditeurs devaient se positionner par rapport à un questionnaire leur demandant de sélectionner la manière de lire qu'ils jugeaient la meilleure, celle qu'ils aimaient le moins sans se laisser influencer par les paramètres paralinguistiques ou techniques. En cas d'hésitation entre deux lecteurs, les auditeurs avaient la possibilité de sélectionner deux candidats au lieu d'un. Un commentaire de leurs choix leur était aussi demandé.

Pour établir ce questionnaire, je me suis basée sur l'hypothèse qu'une bonne lecture se caractérisait par la facilité avec laquelle l'auditeur pouvait comprendre le message transmis et qu'une interlocution réussie était de nature à toucher émotionnellement le lecteur sur un plan intersubjectif.

6.1.3 Analyse des données

Les résultats des votes des auditeurs figurent sous la dénomination de « meilleur » ou de « plus mauvais » lecteur pour représenter un point de vue subjectif par rapport à la préférence personnelle ou la désapprobation des auditeurs.

En ce qui concerne la meilleure lecture, le Lecteur 1 (L1) est très largement plébiscité devant les autres candidats par les natifs anglophones (six voix parmi les huit votants) et les enseignants d'anglais francophones (six voix parmi les huit votants). Les résultats des 36 apprenants, représentés dans le tableau ci-dessous, montrent la répartition des voix selon le

⁴⁷ Évaluation des enseignants de ces groupes par test de positionnement et équivalence selon les critères du *Cadre Européen Commun de Référence du Conseil de l'Europe*. B1 correspond au niveau « indépendant seuil », B2 au niveau « indépendant avancé » et C1 à celui d'un utilisateur « expérimenté autonome » http://www.coe.int/t/dg4/linguistic/Source/Framework_FR.pdf

CHAPITRE II – Profils et profilage : mode d'organisation hiérarchique de la prosodie des anglophones natifs

niveau de maîtrise langagière. On peut constater que les étudiants les plus avancés (groupes classes B2 et B2/C1) préfèrent L5 (neuf voix) puis L1 (sept voix), alors que les moins avancés (B1+/B2) plébiscitent L10 (huit voix) puis L3 (sept voix).

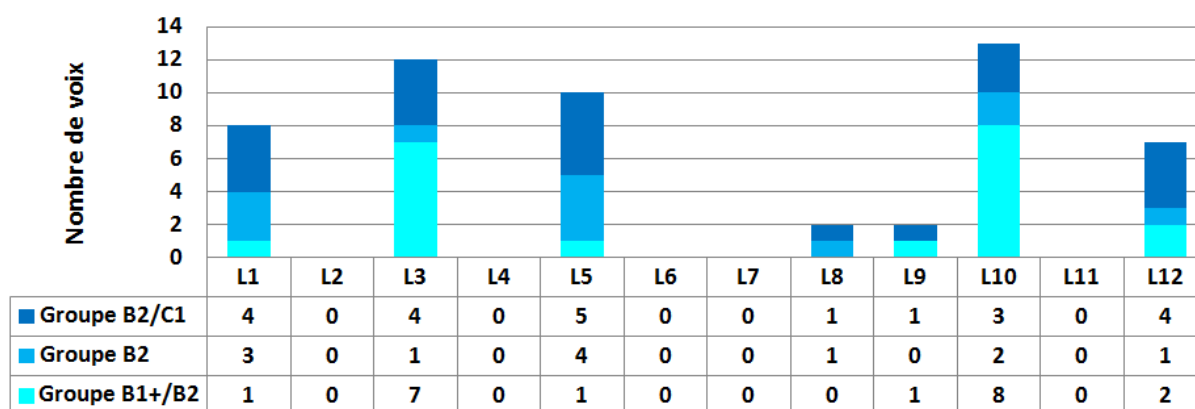


Figure 20 – Répartition des votes « meilleur lecteur » des étudiants selon le niveau de leur groupe classe

À l'autre bout de l'échelle, le lecteur le moins aimé des natifs anglophones est L10 (trois voix) devant L4, L9, L11 et L12 qui en totalisent deux. Les enseignants d'anglais francophones sélectionnent L4 (quatre voix) devant L6 (trois voix). Les apprenants anglicistes ont massivement voté contre L6 (17 voix) quel que soit leur niveau. La répartition détaillée figure ci-dessous.

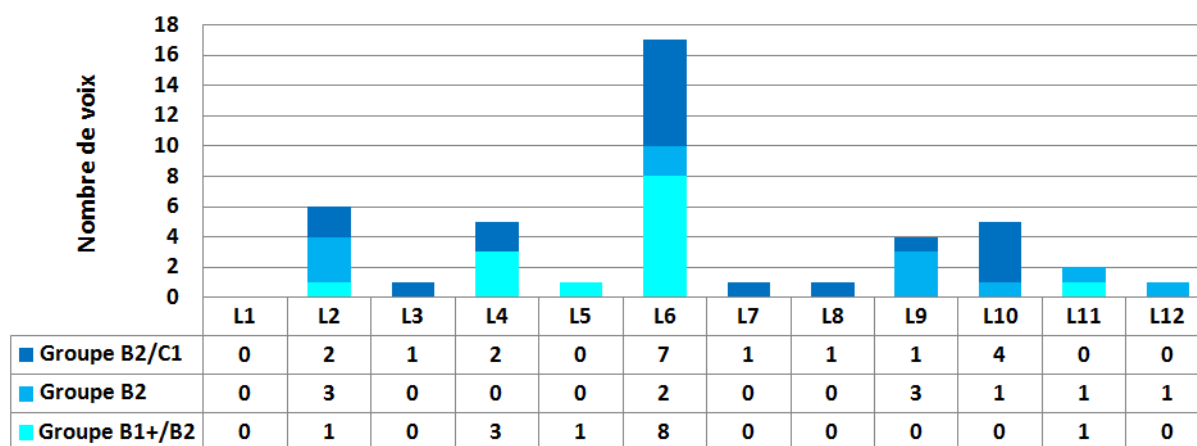


Figure 21 – Répartition des votes « moins bon lecteur » des étudiants selon le niveau de leur groupe classe

Le tableau ci-dessous récapitule les votes des divers groupes. Pour plus de lisibilité j'y ai regroupé les résultats des groupes classes B2 et B2/C1.

	« meilleur » lecteur	« plus mauvais » lecteur
Anglophones	L 1	L 10
Enseignants d'anglais francophones	L 1	L 4

Étudiants anglicistes (groupes B2 et B2/C1)	L 5 (+ L1)	L 6
Étudiants anglicistes (groupe B1+/B2)	L10 (+ L3)	L 6

Tableau 26 – Récapitulation de la sélection des auditeurs pour sélectionner le « meilleur » et le « moins bon » lecteur

6.1.4 Bilan de la phase d'expérimentation PER-LEC-N

Malgré quelques biais dans le profilage (faible nombre d'auditeurs, qualité inégale des enregistrements, enregistrements présentés dans le même ordre d'écoute), il ressort de cette analyse que les préférences correspondent à une tendance claire en fonction du niveau de maîtrise langagière des auditeurs. Les réponses parfois très détaillées des questionnaires le confirment. Le caractère rédhibitoire de la prestation de L6 pour les apprenants est l'extrême rapidité de lecture, alors que la lecture naturelle et envolée de L1 conduit les auditeurs expérimentés à le préférer. Il est à noter que L10 est à la fois le « meilleur lecteur » des apprenants faibles et le « moins bon lecteur » des natifs anglophones. Son style de lecture surfait (*exaggerated, phoney, theatrical, acceptable for children*) semble lasser ou agacer les auditeurs avancés.

Cette expérimentation présente une cohérence qu'il a été intéressant de vérifier dans les réactions que les auditeurs ont consignées dans les questionnaires. Ces impressions subjectives peuvent-elles être corroborées, infirmées ou expliquées par la subjectivité des lecteurs eux-mêmes ? C'est l'objectif de l'analyse suivante.

6.2 Analyse des commentaires de lecteurs sur leur prestation (expérimentation EXP-LEC-N)

6.2.1 Objectif de l'expérimentation

J'ai mentionné les quelques biais de l'analyse perceptive précédente. Un passage obligé d'un profilage méticuleux est la vérification des analyses sur lesquelles sont basées la conception d'un modèle : le profil. Si le lecteur peut se présenter par sa prestation en lecture (enregistrement objectif), il peut aussi le faire en faisant part de sa subjectivité dans cette activité. Je me base en ceci sur la thèse de Bottineau (2008) selon laquelle la parole est le miroir de notre subjectivité marquée par l'empreinte de nos expériences et destinée à les modifier : [...] *language is to be reconsidered in terms of sensorimotor interactions with an environment in which both the individual and the environment are modified*; [...]. Cette expérimentation nouvelle est un profilage de confirmation des analyses perceptives précédentes par un éclairage pertinent sur la raison d'être de ces styles de lecture.

6.2.2 Méthodologie du profilage

Après l’enregistrement du texte *The Selfish Giant*, chaque lecteur a complété un questionnaire fournissant des informations métalinguistiques sur son parcours personnel pour essayer d’établir une correspondance entre la subjectivité des auditeurs et celle des lecteurs. Outre les renseignements de base (âge, genre, variété d’anglais), les questions demandaient principalement à qui il s’imaginait s’adresser pendant la tâche de lecture, quels étaient ses expériences préalables de prise de parole en public et son sentiment sur la question. Les questions fermées se présentaient à choix multiples selon une échelle de valeurs et l’écriture d’un commentaire était systématiquement proposée.

6.2.3 Analyse des données

Je vais présenter chacun des six lecteurs figurant dans le tableau récapitulatif selon les questionnaires complétés. Nous reprenons ici les termes de l’article « Interlocution et variation prosodique dans la lecture à haute voix : de l’influence de la subjectivité des lecteurs sur celle des auditeurs »⁴⁸, qui détaille l’expérimentation⁴⁹.

L1 est un homme de « 50+ » ans, enseignant d’anglais pour spécialistes à l’Université, qui se distingue par une expérience jugée maximale (graduée 4/4) en enseignement, en prise de parole en public et en lecture de contes ou autres histoires enfantines. Il évalue son expérience du chant à 3/4. Il se sent à l’aise (3/4) en parlant à de petits groupes dans sa langue maternelle, mais n’estime sa préférence pour s’adresser à un large public dans sa langue maternelle qu’à 2/4.

Dans la partie commentaire du questionnaire, il répond qu’en lisant le texte, il s’imaginait s’adresser à des enfants. Son commentaire sur la prise de parole en public est : «PREFER PRIVATE, TWO'S COMPANY, ETC. » [sic].

6.2.3.1 Lecteur 3

L3 est un homme de 65 ans, enseignant d’anglais à l’Université, se caractérisant par sa grande expérience (4/4) d’enseignant, d’acteur amateur et de prise de parole en public. Il se sent très à l’aise (4/4) en s’adressant aussi bien à de petits groupes ou à un large public dans sa langue maternelle, qu’à un public quelconque dans une langue étrangère.

⁴⁸ Article édité par C. Douay et D. Rouland (2012) dans *L’interlocution comme paramètre*. Collection Rivages Linguistiques, Presses Universitaires de Rennes, 279-302.

⁴⁹ Voir aussi le fichier « Données EXP-LEC-N.xls » en annexe.

Dans la partie commentaire du questionnaire, il répond qu'en lisant le texte, il s'imaginait s'adresser à des « FIRST YEAR STUDENTS IN LETTRES MODERNES » [sic] et précise sa position sur la prise de parole en public : *no problem if it's essentially an attentive audience, no problem even if it's a (in majority) hostile audience (Student Union AGs for example).*

6.2.3.2 Lecteur 4

L4 est une femme de 31 ans occupant un poste de lectrice à l'Université. Elle juge son expérience d'enseignement importante (4/4) ainsi que celle de prise de parole en public (3/4). Elle n'a quasiment aucune expérience du chant ou d'art dramatique (1/4). Elle écrit se sentir surtout à l'aise en parlant à de petits groupes dans sa langue maternelle (4/4) même si elle évalue sa préférence pour ce qui est de s'adresser à un large public dans sa langue maternelle à 3/4.

Son récepteur imaginaire était de petits enfants, mais elle n'a écrit aucun commentaire dans la rubrique concernant la prise de parole en public.

6.2.3.3 Lecteur 5

L5 est une femme de 23 ans occupant elle aussi un poste de lectrice à l'Université. Elle estime son expérience d'enseignement maximale (4/4), celle de prise de parole en public et celle de lecture de contes ou d'autres lectures à des enfants à 3/4, mais quasiment aucune expérience du chant (1/4). Il semblerait qu'elle ne se sente jamais parfaitement à l'aise en parlant en public étant donné qu'elle juge sa préférence maximale (3/4) pour s'adresser à de petits groupes dans sa langue maternelle. Elle estime sa préférence pour s'adresser à un large public dans sa langue maternelle à 1/4.

En lisant son texte, elle imaginait avoir l'attention d'enfants âgés de six à dix ans. Son commentaire sur la prise de parole en public est le suivant : *I am terrified of public speaking and am not keen about being recorded or even leaving messages on the phone. Therefore, I try to speak in front of audiences as much as possible so as to try and overcome that fear.*

6.2.3.4 Lecteur 6

L6 est une femme de 27 ans, elle aussi lectrice dans une université. Elle considère son expérience maximale (4/4) en ce qui concerne l'enseignement et la prise de parole en public. Elle évalue son expérience dans les autres domaines à 2/4 et ses préférences à 3/4, que ce soit

pour s'adresser aussi bien à de petits groupes ou à un large public dans sa langue maternelle, qu'à un public quelconque dans une langue étrangère.

Elle décrit ainsi son auditoire imaginaire : *researchers for this project (I am not very imaginative)* et donne son sentiment sur la prise de parole en public : *I frequently have to speak in public for my job so I don't really have a problem with it any more.*

6.2.3.5 **Lecteur 10**

L10 est un homme dont je n'ai pas pu obtenir de questionnaire étant donné qu'il s'agit d'un lecteur professionnel qui vend ses lectures sur Internet sous la forme de fichiers sons d'extension MP3 dans un objectif thérapeutique⁵⁰. J'ai néanmoins recueilli sur lui ces informations sur Internet :

Grant holds a B.A. from the University of Michigan with emphasis in German Language, Psychology and Music Performance. He has attended the American Institute For Musical Studies in Graz, Austria and the Flint School Of Performing Arts. He is a Certified Massage Therapist (FSTM 500 Hours) and is excellently trained in Deep-tissue, Swedish, Sports, Shiatsu, Chakra Release and Reflexology.

Grant has been using the powerful effects of guided meditations and positive affirmations in his daily life since 1986.

6.2.3.6 **Résultats de l'analyse**

Les données que je viens d'exposer sur l'expérience et la sensibilité des lecteurs anglophones natifs me permettent à présent de les présenter succinctement par regroupement de profils deux par deux :

- L3 jouit d'une faculté à s'adresser avec assurance à de vastes publics dans une diction appropriée. Il semblerait que L10 le rejoigne dans ce registre à cela près que L3 n'hésite pas à imposer son message à l'auditoire si nécessaire. Ces deux lecteurs, bien que maîtrisant parfaitement leur diction, semblent déclamer le texte, lui ôtant tout naturel, comme dans une prestation théâtrale devant un large public.

⁵⁰ [<http://ia310828.us.archive.org/1/items/TheSelfishGiant-ByOscarWilde-NarratedByGrantRaymondBarrett/01SelfishGiant.mp3>], consulté en 2010. Vente d'enregistrements sur site commercial : [<https://www.amazon.com/Feel-Good-Meditations-Positive-Affirmations/dp/B004YXDIZ8>], consulté en 2016.

- L4 et L6 ne bénéficient ni de formation, ni d'une expérience appropriée comparable. Les questionnaires les révèlent bien moins impliquées dans l'acte d'interlocution.
- L1 possède l'expérience nécessaire à la prise de parole en public, mais sa sensibilité le conduit à préférer le registre intime (où l'acte interlocutoire est de grande portée), ce qui fait sans doute correspondre sa diction aux critères des auditeurs expérimentés. L5 semble dotée d'un profil similaire, accentué par sa jeunesse.

6.2.4 Bilan de la phase d'expérimentation EXP-LEC-N

Après avoir analysé les résultats subjectifs des auditeurs, j'ai pu faire de même avec les données subjectives des lecteurs. La correspondance des deux analyses de deux points de vue diamétralement opposés (émetteurs-récepteurs) relève d'une logique de bon aloi : le tableau récapitulatif des « bons » et « mauvais lecteurs » est expliqué par la relation interlocutive asynchrone à la fois du point de vue des auditeurs que de celui des lecteurs, car ils se complètent et s'étayent mutuellement. À présent, un profilage acoustique des corrélats de ces lectures pourrait justifier objectivement les jugements des auditeurs.

6.3 Profilage des lecteurs par analyse acoustique (expérimentation COR-AC-N1)

6.3.1 Objectif de l'expérimentation

Cette étude s'est jusqu'à présent focalisée sur la subjectivité et le ressenti émotionnel des auditeurs-récepteurs puis des lecteurs-émetteurs lors de l'acte de lecture à voix haute. C'est à présent le profilage des lecteurs par les corrélats acoustiques de leur lecture qui va pouvoir corroborer ou infirmer en toute indépendance les relations de cause à effet pressenties dans les profils précédents.

6.3.2 Méthodologie du profilage

Les étapes du profilage par analyse acoustique sont schématisées dans le graphique suivant selon la hiérarchisation en quatre niveaux d'abstraction du modèle *CRISP-DM*.

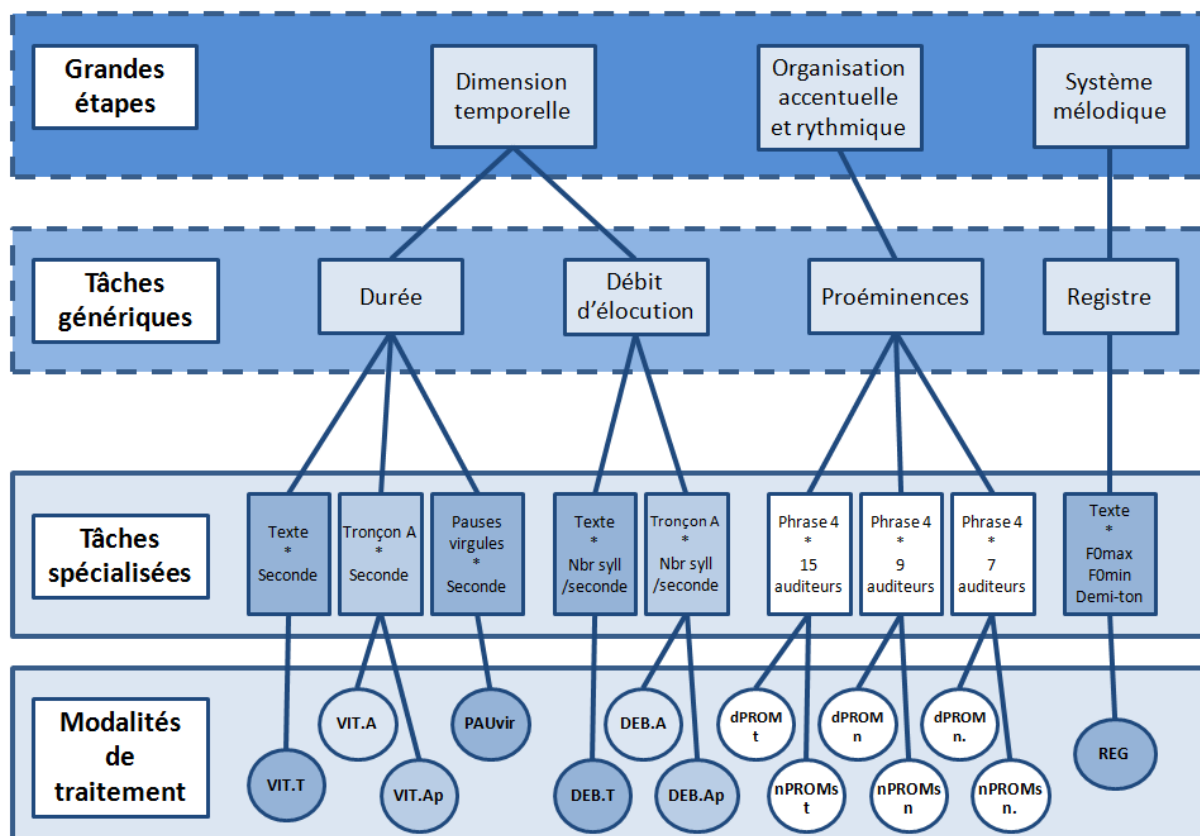


Figure 22 – Profilage acoustique des lecteurs anglophones natifs en 4 niveaux selon la méthodologie CRISP-DM (Chapman *et al.*, [1999] 2000 : 6)

Le profilage est au départ balisé par les trois grands domaines qui sont :

- L'organisation temporelle, subdivisée ensuite en étude de la durée d'une part et du débit d'élocution d'autre part,
- L'organisation accentuelle et rythmique, orientée vers l'analyse des proéminences,
- Le système mélodique, limité au registre.

Les tâches génériques de niveaux 1 (grandes étapes) et 2 (tâches génériques) relèvent d'un choix guidé par les remarques subjectives des auditeurs ainsi que par la relative facilité de prise de mesures dans le texte de cinq phrases. Je vais davantage détailler les niveaux 3 (tâches spécialisées) et 4 (modalités de traitement) ciblant les contraintes techniques d'accès aux données.

La dimension temporelle est mesurée de manière double. Elle l'est de façon absolue par ses mesures d'une durée de parole : c'est le cas des cinq phrases constitutives du texte entendu, du début de la première syllabe à la fin de la dernière (**VIT.T**). J'ai sélectionné le tronçon A pour établir un micro-profil où les pauses ne sont pas censées apparaître. Il s'agit

de la fin de la Phrase 4 : [...] *that the children used to stop their games in order to listen to them*. La version avec éventuelles pauses incluses est **VIT.Ap** et la variable de durée du tronçon à l'exclusion de toute pause d'au moins 0,1 seconde est **VIT.A**. Ces durées sont mesurées en secondes, afin de pouvoir opérer plus facilement des comparaisons avec les études antérieures sur le sujet mentionnées dans la littérature. Un calcul de la durée des pauses est opéré indépendamment : c'est la variable **PAUvir**. Ses données correspondent à la moyenne de la durée des pauses aux cinq endroits du texte ponctués de virgules. On a vu en amont l'importance de la durée ou l'absence de pauses et je formule l'hypothèse que c'est à l'emplacement de virgules que les différences de style de parole sont les plus révélatrices.

Le débit d'élocution permet de rendre compte de la dimension temporelle de manière relative en mesurant ce qu'un intervalle temporel contient. Ces intervalles sont le texte et le tronçon A, leur contenu est le nombre de syllabes prononcées. **DEB.T** est le nombre de syllabes prononcées par seconde dans le texte. **DEB.A** est le nombre de syllabes prononcées par seconde dans le tronçon A en excluant la durée des éventuelles pauses et **DEB.Ap** est le nombre de syllabes prononcées par seconde dans le tronçon A en comptabilisant la durée des éventuelles pauses.

Je prends ici les proéminences comme fondement des mesures rythmiques. La Phrase 4, avec ses 28 syllabes canoniques et son absence de ponctuation, permet de dresser un micro-profil rythmique de grand intérêt pour différencier les profils prosodiques. Deux types de variables sont calculées : la densité des proéminences mesurée par le pourcentage de proéminences en fonction du nombre de syllabes prononcées, et le nombre de proéminences par seconde en fonction de la vitesse de lecture de cette phrase. Étant donné leur nature acoustique complexe, mes variables se basent sur la perception auditive de plusieurs groupes d'auditeurs natifs et non natifs à qui j'ai demandé de sélectionner les syllabes proéminentes de la Phrase 4 par rapport à celles qui ne l'étaient pas, avec deux niveaux de proéminence : présence ou absence. Je nomme les variables aux proéminences détectées par tous les 15 auditeurs, natifs et non natifs, **dpROMt** (pour la densité des proéminences) et **npROMst** (pour le nombre de proéminences par seconde). La moyenne du nombre de proéminences détectées par les 15 auditeurs chez les six lecteurs est de 7,48 et la moyenne des niveaux de discrimination des proéminences par les 15 auditeurs chez les six lecteurs est de 5,06. J'appelle **dpROMn** (pour la densité des proéminences) et **npROMsn** (pour le nombre de proéminences par seconde) les variables dont les proéminences sont repérées par les neuf locuteurs natifs anglophones du groupe (moyenne des proéminences détectées de 6 et

CHAPITRE II – Profils et profilage : mode d’organisation hiérarchique de la prosodie des anglophones natifs

moyenne des niveaux de discrimination de 5,22), en écartant les six enseignants d’anglais francophones, dont un bilingue (moyenne des proéminences détectées de 9,4). Ayant remarqué que la fourchette du nombre moyen de proéminences détectées suivant les lecteurs était estimée par les natifs entre 3 et 8, j’ai alors retiré du groupe les deux auditeurs natifs qui ne discriminaient qu’une différence de trois proéminences dans les lectures de la Phrase 4 afin d’augmenter la moyenne du taux de discrimination, et j’ai donc obtenu un groupe réduit à sept auditeurs natifs, dont les proéminences détectées s’échelonnaient de 4 à 8 (total de cinq niveaux et moyenne de 5,85 ; moyenne du nombre de proéminences de 6) : j’ai nommé les variables correspondantes **dPROMn**. (pour la densité des proéminences) et **nPROMsn**. (pour le nombre de proéminences par seconde). Le tableau suivant récapitule les différentes moyennes obtenues suivant le nombre d’auditeurs dans la mesure des proéminences.

Détection des proéminences dans la Phrase 4	15 auditeurs (natifs et non natifs)	9 auditeurs natifs	7 auditeurs natifs (discrimination fine des proéminences)
Moyenne du nombre de proéminences par lecteur	7,48	6	6
Moyenne des niveaux de discrimination dans la détection des proéminences	5,06	5,22	5,85

Tableau 27 – Moyennes du nombre global de proéminences et des niveaux de discrimination dans la détection des proéminences

Le nombre moyen de proéminences perçues par les 15 auditeurs permet de voir par déduction que les enseignants d’anglais francophones perçoivent beaucoup plus de proéminences que les natifs anglophones. Ainsi, leur exclusion du calcul de la moyenne la fait chuter considérablement (de 7,48 à 6). Le corollaire de cette perception, qui se veut exhaustive, est le faible niveau de discrimination entre les proéminences détectées par les francophones chez divers lecteurs, comme si leurs battements rythmiques étaient similaires. Ainsi, semblablement aux résultats de l’exclusion de ces derniers, la moyenne du niveau de discrimination des anglophones natifs est plus élevée (5,22 au lieu de 5,06). Cette moyenne est même augmentée à 5,85 lorsque l’on écarte les deux auditeurs anglophones qui ne distinguent que trois niveaux de proéminences entre les différents styles de lecture. La réduction graduelle du nombre d’auditeurs suivant leur capacité à discriminer le nombre de proéminences résulte en l’obtention de trois variables dont il serait intéressant de jauger la pertinence.

Le domaine mélodique (que j’appelle intonation) est le parent pauvre de ce profilage étant donné qu’une seule variable lui est apparentée : le registre (**REG**). Un registre large est

signe de mouvements mélodiques de grande ampleur. La qualité mélodique, domaine empreint de la plus grande variabilité, m'a semblé difficilement profilable par des variables à ce stade d'investigation.

En conclusion, les tâches spécialisées et leur modèle de traitement totalisent 14 variables, dont sept pour la dimension temporelle, six pour le rythme, et une apparentée au système mélodique.

6.3.3 Analyse des données acoustiques

Les données à partir des variables décrites ci-dessus ont été consignées dans une matrice accompagnées des niveaux charnières 2 et 3 du modèle de profilage *CRISP-DM* (Chapman *et al.*, ([1999] 2000 : 6) concernant les tâches génériques (durée, débit d'élocution, proéminences et registre) et les tâches spécialisées qui leur correspondent indiquant la localisation de l'intervention et la nature des mesures prises. Les données acoustiques résultant des variables du niveau 4, correspondant aux modalités de traitement, se retrouvent accompagnées d'une catégorisation quantitative et qualitative afin de pouvoir les visualiser comparativement. Une matrice similaire bien qu'aux catégories réduites en a été extraite afin d'adapter la catégorisation qualitative en vue de l'analyse statistique. La matrice étant de dimensions imposantes, je la présente ci-dessous dans sa nouvelle version bien qu'elle corresponde à la section suivante sur l'analyse statistique.

Les mesures des trois variables de vitesse de lecture se corroborent : ce profilage de la vitesse de lecture par micro-profil facilite l'étude de textes longs.

Le critère de durée de lecture correspond aux choix des auditeurs. Le débit d'élocution figurera sous la forme « nombre de syllabes par seconde » sans spécifier sa localisation, par souci de simplicité. La gestion de la durée dans la lecture n'est pas qu'une affaire de durée moyenne, elle porte sur les choix de pauses spécifiques qu'il est important de prendre en considération.

VITESSE GÉNÉRALE DE LECTURE				L1	L3	L4	L5	L6	L10
Durée de lecture	Durée totale du texte (5 phrases) en secondes	Catégories qualitatives	Catégories quantitatives	29,37	35,39	27,67	30,48	25,46	44,21
		Rapide	$20 \leq X < 30$	+	-	+	-	+	-
		Moyenne	$30 \leq X < 40$	-	-	-	+	-	-
		Lente	$40 \leq X < 50$	-	+	-	-	-	+
	Durée du tronçon A pauses exclues (en secondes)	Catégories qualitatives	Catégories quantitatives	2,59	3,37	2,67	2,92	2,42	3,53
		Rapide	$2,3 \leq X < 2,8$	+	-	+	-	+	-
		Moyenne	$2,8 \leq X < 3,3$	-	-	-	+	-	-
		Lente	$3,3 \leq X < 3,8$	-	+	-	-	-	+
	Durée du tronçon A pauses incluses (en secondes)	Catégories qualitatives	Catégories quantitatives	2,59	3,37	2,67	2,92	2,42	3,93
		Rapide	$2,3 \leq X < 2,8$	+	-	+	-	+	-
		Moyenne	$2,8 \leq X < 3,3$	-	-	-	+	-	-
		Lente	$3,3 \leq X < 3,8$	-	+	-	-	-	-
Débit d'élocution	Nombre de syllabes par seconde, pauses incluses (5 phrases)	Catégories qualitatives	Catégories quantitatives	4,18	3,50	4,48	4,03	4,86	2,80
		Très élevé	$4,4 \leq X < 5$	-	-	+	-	+	-
		Élevé	$3,8 \leq X < 4,4$	+	-	-	+	-	-
		Moyen	$3,2 \leq X < 3,8$	-	+	-	-	-	-
		Faible	$2,6 \leq X < 3,2$	-	-	-	-	-	+
	Nombre de syllabes par seconde, pauses exclues (tronçon A)	Catégories qualitatives	Catégories quantitatives	6,54	5,03	6,35	5,80	7,01	4,81
		Élevé	$6,1 \leq X < 7,1$	+	-	+	-	+	-
		Moyen	$5,1 \leq X < 6,1$	-	-	-	+	-	-
		Faible	$4,1 \leq X < 5,1$	-	+	-	-	-	+
	Nombre de syllabes par seconde, pauses incluses (tronçon A)	Catégories qualitatives	Catégories quantitatives	6,54	5,03	6,35	5,80	7,01	4,32
		Élevé	$6,1 \leq X < 7,1$	+	-	+	-	+	-
		Moyen	$5,1 \leq X < 6,1$	-	-	-	+	-	-
Faible		$4,1 \leq X < 5,1$	-	+	-	-	-	+	
Pauses	Durée moyenne des pauses à l'emplacement de virgules (en secondes)	Catégories qualitatives	Catégories quantitatives	0,26	0,30	0,26	0,24	0,17	0,48
		Tr. prononcée	$0,4 \leq X < 0,5$	-	-	-	-	-	+
		Prononcée	$0,3 \leq X < 0,4$	-	+	-	-	-	-
		Perceptible	$0,2 \leq X < 0,3$	+	-	+	+	-	-
		Peu perceptible	$0,1 \leq X < 0,2$	-	-	-	-	+	-
	Durée moyenne d'une pause dans le tronçon A (en secondes)	Catégories qualitatives	Catégories quantitatives	0	0	0	0	0	0,40
		Prononcée	$0,1 \leq X < 0,5$	-	-	-	-	-	+
		Imperceptible	$0 \leq X < 0,1$	+	+	+	+	+	-

Tableau 28 – Matrice récapitulative des données temporelles pour plan factoriel

Les données (tableau ci-dessous) montrent qu'un registre ample est incompatible avec une trop grande vitesse de lecture. Il lui est corrélé dans la mesure où les modifications mélodiques nécessitent le temps de leur réalisation.

		L1	L3	L4	L5	L6	L10
Registre (en demi-tons)	Catégories qualitatives	20,67	18,69	15,24	15,88	14,55	25,50
	Très ample	22 ≤ X < 26	-	-	-	-	+
	Ample	18 ≤ X < 22	+	+	-	-	-
	Normal	14 ≤ X < 18	-	-	+	+	+

Tableau 29 – Matrice récapitulative du registre de lecture pour plan factoriel

Les demi-tons rendent la comparaison des données masculines et féminines comparables. Cette unité de mesure sera donc préservée.

Pour ce qui est de la composante rythmique, les deux modes de calcul à partir des proéminences fournissent deux variables dont les données s'éclairent et s'étayent mutuellement. La rapidité de lecture est logiquement corrélée au nombre de mots et syllabes prononcés, et par voie de conséquence, aux syllabes candidates à la proéminence disponibles. Ainsi, dans les tableaux ci-dessous, le nombre de proéminences devrait être étroitement corrélé à la vitesse de lecture et la densité des proéminences devrait marquer une similitude par rapport à la variable précédente, car la taille des pieds accentuels devrait être proportionnelle au faible nombre de proéminences. Le nombre de syllabes proéminentes semble en effet être inversement proportionnel à la vitesse de lecture, comme le montrent les croix situant les données dans les tableaux ci-dessous, présentées plus ou moins symétriquement, selon le nombre d'auditeurs. Ainsi, les lecteurs caractérisés par une faible vitesse de lecture et un registre ample ont tendance à compenser ces traits par un faible taux de proéminences. Or, L4 fait exception à cette présentation en miroir, quelle que soit la configuration du nombre d'auditeurs. On peut détecter là un manque de cohérence dans la qualité de lecture : sa faible proportion de proéminences, qu'elle soit calculée en fonction de la durée de lecture ou du nombre de syllabes de la phrase, rend la lecture difficile à suivre faute de repères. De plus, le registre étroit de L4 ne peut qu'accentuer le désarroi des auditeurs.

PROÉMINENCES (15 auditeurs N et NN)			L1	L3	L4	L5	L6	L10
Nombre de proéminences par seconde	Catégories qualitatives	Catégories quantitatives	1,28	1,20	1,16	1,29	1,63	1,15
	Important	$1,5 \leq X < 1,75$	-	-	-	-	+	-
	Moyen	$1,25 \leq X < 1,5$	+	-	-	+	-	-
	Faible	$1 \leq X < 1,25$	-	+	+	-	-	+
Densité des proéminences (pourcentage du nombre de syllabes par proéminence)	Catégories qualitatives	Catégories quantitatives	23,57	29,28	22,61	25,47	25,95	33,57
	Très élevé	$32 \leq X < 36$	-	-	-	-	-	+
	Élevé	$28 \leq X < 32$	-	+	-	-	-	-
	Normal	$24 \leq X < 28$	-	-	-	+	+	-
	Faible	$20 \leq X < 24$	+	-	+	-	-	-

Tableau 30 – Matrice récapitulative des données rythmiques pour plan factoriel (15 auditeurs)

PROÉMINENCES (9 auditeurs natifs)			L1	L3	L4	L5	L6	L10
Nombre de proéminences par seconde	Catégories qualitatives	Catégories quantitatives	1,10	1,01	0,86	1,05	1,37	0,99
	Important	$1,2 \leq X < 1,5$	-	-	-	-	+	-
	Moyen	$0,9 \leq X < 1,2$	+	+	-	+	-	+
	Faible	$0,6 \leq X < 0,9$	-	-	+	-	-	-
Densité des proéminences (pourcentage du nombre de syllabes par proéminence)	Catégories qualitatives	Catégories quantitatives	20,23	24,60	16,66	20,63	21,82	28,96
	Très élevé	$27 \leq X < 31$	-	-	-	-	-	+
	Élevé	$23 \leq X < 27$	-	+	-	-	-	-
	Normal	$19 \leq X < 23$	+	-	-	+	+	-
	Faible	$15 \leq X < 19$	-	-	+	-	-	-

Tableau 31 – Matrice récapitulative des données rythmiques pour plan factoriel (9 auditeurs natifs)

PROÉMINENCES (7 auditeurs natifs)			L1	L3	L4	L5	L6	L10
Nombre de proéminences par seconde	Catégories qualitatives	Catégories quantitatives	1,10	1,05	0,76	1,06	1,45	1,08
	Important	$1,2 \leq X < 1,5$	-	-	-	-	+	-
	Moyen	$0,9 \leq X < 1,2$	+	+	-	+	-	+
	Faible	$0,6 \leq X < 0,9$	-	-	+	-	-	-
Densité des proéminences (pourcentage du nombre de syllabes par proéminence)	Catégories qualitatives	Catégories quantitatives	18,87	23,46	14,79	18,87	20,91	29,59
	Très élevé	$27 \leq X < 32$	-	-	-	-	-	+
	Élevé	$22 \leq X < 27$	-	+	-	-	-	-
	Normal	$17 \leq X < 22$	+	-	-	+	+	-
	Faible	$12 \leq X < 17$	-	-	+	-	-	-

Tableau 32 – Matrice récapitulative des données rythmiques pour plan factoriel (7 auditeurs natifs)

Non seulement une diversité des variables renforce-t-elle un diagnostic comparatif des données, mais encore une étude comparative selon les métriques exposées en amont ne permettrait pas de comprendre et vérifier la synergie des phénomènes en jeu, mais tout au plus d'exposer un résultat final.

L’analyse des corrélats acoustiques montre la convergence des diverses variables pour distinguer les profils de locuteurs natifs. Les « mauvais » lecteurs pour les non natifs sont ceux qui lisent très vite (L4 et L6) sans marquer de pauses suffisantes aux endroits stratégiques, alors que les « bons » lecteurs L1 et L5 conjuguent rapidité de lecture, débit élevé, avec des pauses perceptibles aux endroits adéquats et un registre ample, sans alourdir la lecture en martelant des proéminences en trop grand nombre. La rapidité de lecture correspondrait à la vérification du critère de clarté et de facilité de compréhension du message, alors que le critère de qualité relèverait de la transmission de l’émotion par la lecture, dans laquelle une multitude de facteurs interviennent. Le tableau récapitulatif de la sélection des auditeurs pour sélectionner le « meilleur » et le « moins bon » lecteur, bien que très succinct, nous renseigne clairement sur les tendances à l’œuvre dans leurs choix (renseignées par les réponses aux questionnaires). L’axe définissant les « bons » lecteurs, sur lequel on peut situer L1, L5, L3 et L10, pourrait se définir par le caractère naturel, artificiel ou surfait de la lecture, alors que celui des « mauvais » lecteurs, situant L10, L4 et L6, correspondrait à la seule vitesse de lecture (très lente, ou au contraire très rapide), qui soit alourdit la diction en saturant l’auditeur par le trop grand nombre de repères, soit l’empêche de les utiliser s’il y en a).

J’é mets ainsi l’hypothèse que, pour les auditeurs de l’expérimentation, le critère de qualité de lecture discriminant le « meilleur » du « plus mauvais » lecteur peut être représenté le long d’un continuum graduant le style de lecture selon le critère correspondant au caractère « naturel » ou « surfait » :

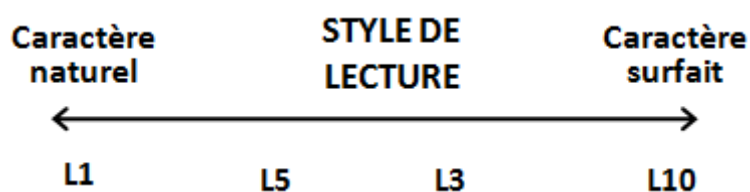


Figure 23 – Continuum qualitatif selon le caractère « naturel » ou « surfait » de la lecture des « bons » lecteurs natifs

J’é mets en outre l’hypothèse que, pour les auditeurs de l’expérimentation, le critère de qualité de lecture discriminant le « meilleur » du « plus mauvais » lecteur peut être représenté le long d’un axe symbolisant le degré de vitesse de lecture :

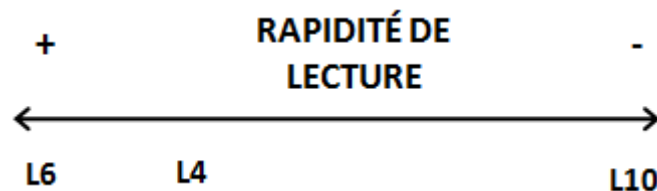


Figure 24 – Continuum qualitatif selon le degré de vitesse de lecture des « mauvais » lecteurs natifs

Ces deux schémas induisent une correspondance entre le caractère surfait de la lecture et sa lenteur exagérée. Cependant, si la rapidité de lecture peut facilement être calculée grâce aux variables de la matrice, il n'en va pas de même du caractère spontané, naturel, appuyé ou sophistiqué du style de lecture. C'est une conjonction de données relevant de plusieurs variables stratégiques qui peut éclairer le sujet en les hiérarchisant.

6.3.4 Bilan de la phase d'expérimentation COR-AC-N1

En fait, la matrice obtenue grâce à l'expérimentation COR-AC-N1 indique les caractéristiques de chaque profil de diction en fonction de la catégorie d'auditeurs à satisfaire. Les résultats montrent que ces derniers marquent une préférence pour une large tessiture avec contours intonatifs variés communiquant sens et émotion. La vitesse de lecture est un facteur essentiel par rapport à la capacité de rétention de l'information par l'auditeur, et donc par rapport à son niveau de maîtrise de la langue. L'usage pertinent des pauses rehausse les fonctions démarcative et psychologique de la prosodie et facilite l'assimilation du contenu informatif. Les prééminences sont perçues comme des repères informatifs. Les bons lecteurs utilisent ces paramètres de façon à agrémenter l'écoute et à faciliter la compréhension et la rétention de l'information. Les divers choix de diction pour chaque paramètre font correspondre ou non les prestations des lecteurs aux attentes (et besoins) de tel ou tel profil d'auditeur. L'analyse acoustique corrobore les résultats des expérimentations subjectives et en cela souligne leur pertinence et leur robustesse.

Cette analyse contrastive des données acoustiques d'un corpus sonore de lecture à voix haute a permis de constituer un socle destiné à une étude statistique pour hiérarchiser les variables.

6.4 Profilage des lecteurs par analyse statistique (expérimentation STAT-N1)

6.4.1 Objectif de l'expérimentation statistique

Afin de relever les critères prosodiques acoustiques susceptibles de constituer des profils de lecteurs à cibler ou à éviter, j'ai jugé nécessaire de vérifier par voie statistique la cohérence de ces jugements subjectifs et qualitatifs sur ces six profils. Pour atteindre cet objectif, j'ai suivi la démarche modélisée ci-dessus par le modèle *CRISP-DM*, que j'ai adapté à la présente étude. Chaque phase de ce profilage est détaillée dans les rubriques suivantes.

6.4.2 Méthodologie du profilage statistique

À partir des résultats subjectifs de l'expérimentation perceptive de la qualité de lecture d'anglophones natifs, plusieurs paramètres ont été utilisés à fin de comparaison pour établir une typologie de « bons » lecteurs anglophones natifs dont des modèles pour apprenants pourraient être extraits par voie statistique. Ils s'apparentent aux grandes composantes d'analyse prosodique à l'exclusion de l'étude mélodique, étant donné la difficulté à rendre compte de la qualité globale de l'intonation d'un locuteur natif par une seule variable. De chacune de ces variables continues a été dérivée une variable qualitative après visualisation sur histogramme et catégorisation manuelle des mesures obtenues. Ce traitement de la multicolinéarité permet d'étudier ces variables sous deux angles différents : d'une part par cercle de corrélations ACP (Analyse en Composantes Principales) et d'autre part par plan factoriel ACM (Analyse en Composantes Multiples). Il est alors possible de s'appuyer conjointement sur leurs résultats pour déterminer les variables les plus représentatives de la qualité prosodique des locuteurs, et pour identifier les profils prosodiques principaux de lecteurs natifs. L'analyse par dendrogramme fournit une représentation graphique complémentaire aux précédentes. Le tableau suivant présente ces variables destinées à déterminer les profils prosodiques de natifs anglophones. Le nom des variables qualitatives se distingue de leur pendant quantitatif par l'adjonction de « .1 ». Pour obtenir un plan factoriel performant, les données des variables ont été une nouvelle fois catégorisées selon des échelles de valeurs quantitatives et qualitatives resserrées. Elles ont été présentées plus haut pour éviter les doublons de versions similaires des tableaux.

CHAPITRE II – Profils et profilage : mode d'organisation hiérarchique de la prosodie des anglophones natifs

Apparentées aux catégories		Variables continues et catégorielles	Signification, mesures et degrés
Dimension temporelle	Durée	VIT.T VIT.T.1	Vitesse de lecture du texte de 5 phrases Mesure en secondes ; degré : rapide/moyenne/lente
		VIT.A VIT.A.1	Vitesse de lecture du tronçon A (pauses exclues) Mesure en secondes ; degré : rapide/moyenne/lente
		VIT.Ap VIT.Ap.1	Vitesse de lecture du tronçon A (pauses incluses) Mesure en secondes ; degré : rapide/moyenne/lente/très lente
		PAUvir PAUvir.1	Durée moyenne d'une pause à l'emplacement de virgule Mesure : en secondes ; degré : peu perceptible/perceptible/prononcée/très prononcée
	Vitesse d'élocution	DEB.T DEB.T.1	Débit d'élocution du texte Mesure : nombre de syllabes par seconde ; degré : très élevé/élevé/moyen/faible
		DEB.A DEB.A.1	Débit d'élocution du tronçon A (sans pauses) Mesure : nombre de syllabes par seconde ; degré : élevé/moyen/faible
		DEB.Ap DEB.Ap.1	Débit d'élocution du tronçon A (pauses incluses) Mesure : nombre de syllabes par seconde ; degré : élevé/moyen/faible
Organisation accentuelle et rythmique	Proéminences (15 auditeurs)	dPROMt dPROMt.1	Densité de proéminences Mesure : pourcentage de proéminences parmi les syllabes ; degré : très élevé/élevé/normal/faible
		nPROMst nPROMst.1	Nombre de proéminences par seconde Mesure : nombre de proéminences parmi les syllabes ; degré : important/moyen/faible
	Proéminences (9 auditeurs)	dPROMn dPROMn.1	Densité de proéminences Mesure : pourcentage de proéminences parmi les syllabes ; degré : très élevé/élevé/normal/faible
		nPROMsn nPROMsn.1	Nombre de proéminences par seconde Mesure : nombre de proéminences parmi les syllabes ; degré : important/moyen/faible
	Proéminences (7 auditeurs)	dPROMn. dPROMn..1	Densité de proéminences Mesure : pourcentage de proéminences parmi les syllabes ; degré : très élevé/élevé/normal/faible
		nPROMsn. nPROMsn..1	Nombre de proéminences par seconde Mesure : nombre de proéminences parmi les syllabes ; degré : important/moyen/faible
	Système mélodique	REG REG.1	Registre F0 max - F0 min (en demi-tons) Mesure : demi-tons ; degré : très ample/ample/normal

Tableau 33 – Variables permettant de déterminer statistiquement les profils prosodiques de natifs anglophones

6.4.3 Analyse statistique des données

La représentation des données qualitatives avec plans factoriels issus d'analyse factorielle par correspondances (AFC) est destinée à la gestion de grands nombres de données (Husson *et al.* 2009 : i) afin de repérer des tendances qui resteraient invisibles à l'entendement

humain. Pourtant, c'est cette approche que j'ai choisie ici. Malgré le très faible nombre d'individus, de nettes tendances se dégagent afin d'élaborer des profils de locuteurs anglophones natifs en lecture. Ces profils personnalisés constituent alors des portraits robots, lesquels, par extrapolation, servent de base aux profils recherchés et font office de profils de groupes en constituant des « archétypes stéréotypiques », notion que j'explique ci-après. Destinés à établir une borne des possibles, ils conservent ainsi leurs particularités extrêmes individuelles, qui auraient été gommées en effectuant une moyenne des mesures englobant des individus moins typés.

Le graphique suivant représente donc une analyse multidimensionnelle destinée à situer les six lecteurs parmi les variables qualitatives étudiées.

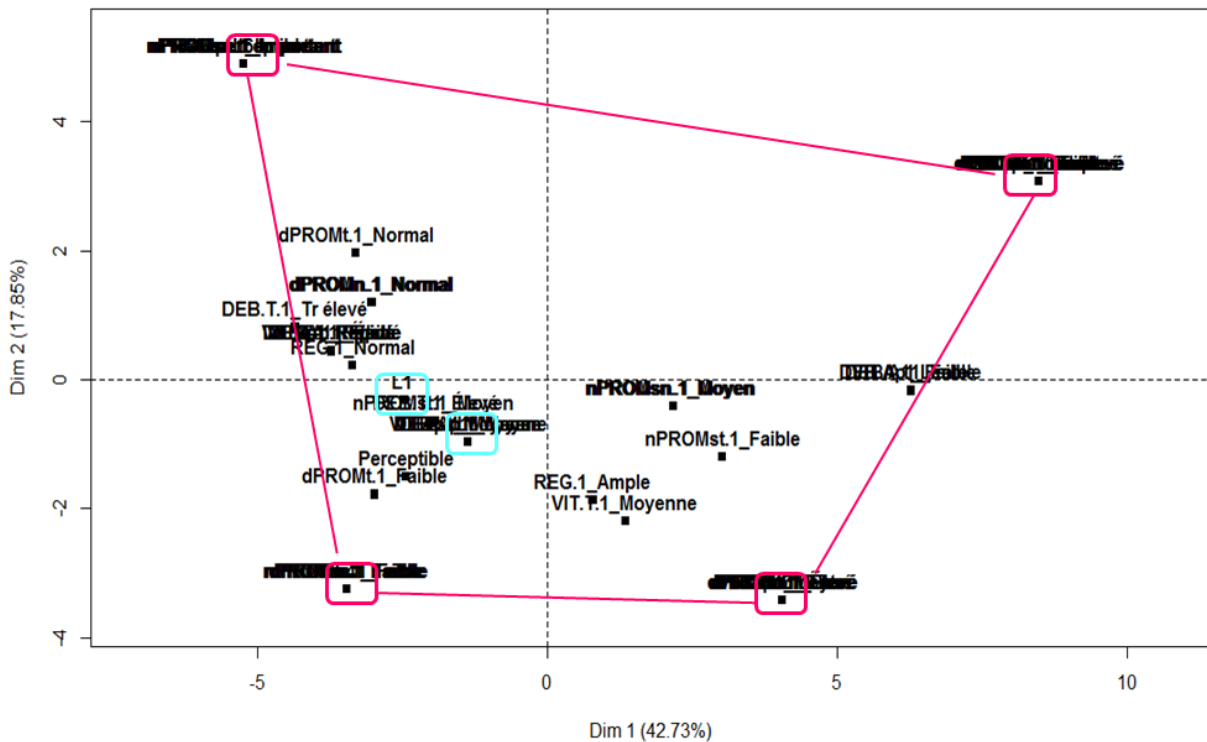


Figure 25 – Plan factoriel des six lecteurs et des variables qualitatives censées les distinguer (15 auditeurs dont 9 natifs et 6 non natifs)

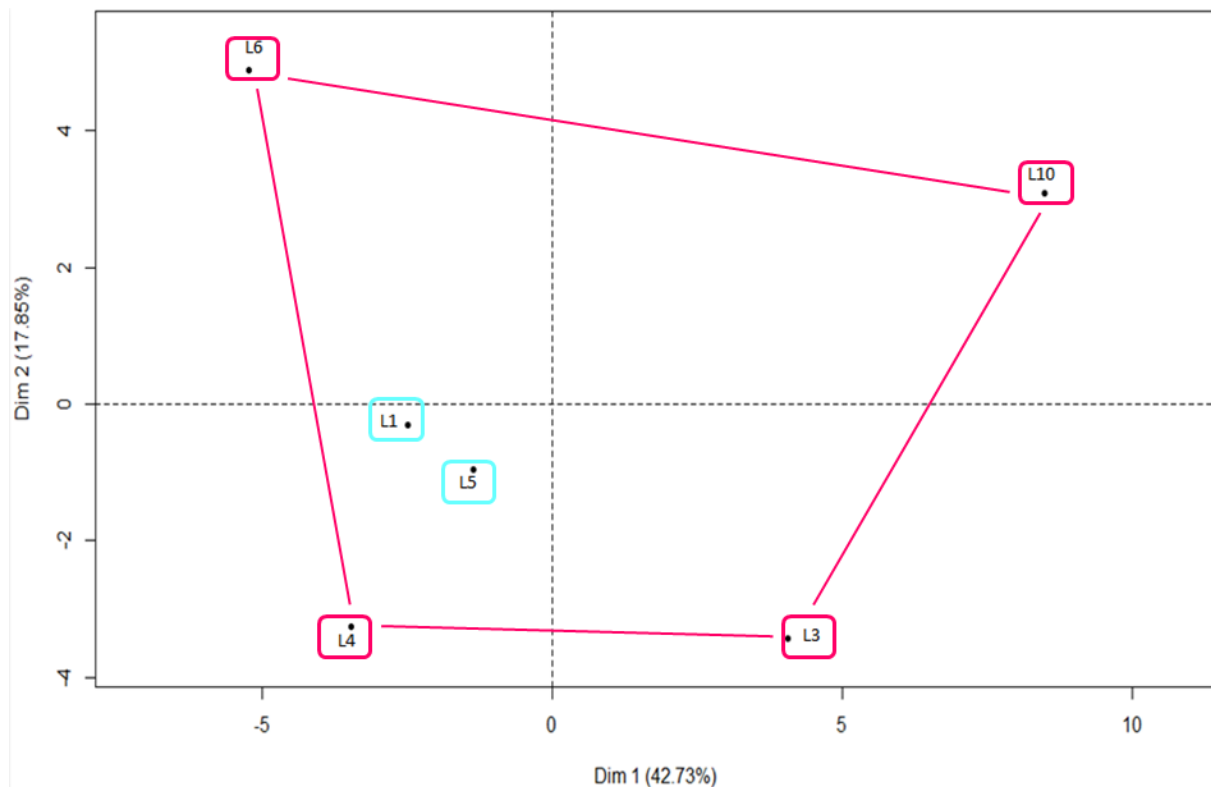


Tableau 34 – Plan factoriel situant les 6 lecteurs (15 auditeurs dont 9 natifs et 6 non natifs)

Les profils caractéristiques se reconnaissent grâce aux mentions qui les entourent et à leur position dans le graphique. Le plan factoriel des lecteurs sans la superposition des termes facilite la lecture du graphique et permet de distinguer l'émergence de quatre profils repérés par un faible nombre de variables qualitatives cerclées de couleur et reliés par des lignes. L3, L4, L6 et L10 sont très typés, représentatifs de mesures extrêmes de certaines variables du corpus, c'est pourquoi ils se situent en périphérie du graphique, loin du barycentre (intersection entre les deux droites perpendiculaires en pointillés) et des autres lecteurs. La dimension 1 est représentée à 42,73 % par les variables de vitesse de lecture. La densité et le nombre de proéminences par seconde constituent une part importante de la dimension 2⁵¹. Les lecteurs L1 et L5 ont une diction prosodique plus équilibrée, éloignée des réalisations des autres profils plus marqués. L5 est le plus proche du barycentre. Sa diction se rapproche de L1, un peu moins centré que L5, mais proche lui aussi de l'intersection des axes en pointillés. Il est similaire aux locuteurs L4 et L6 par sa rapidité de lecture (il est situé à gauche de l'axe vertical). Bien qu'étant sommaire, ce plan factoriel montre que ce sont les lecteurs L1 et L5

⁵¹ Ce type d'analyse permet de calculer l'importance des contributions à la variance et le pourcentage de variance expliquée par les axes factoriels.

qui, par leurs données équilibrées (proches de l'intersection des deux axes), sont susceptibles, parmi les 12 lecteurs sélectionnés au départ de l'expérimentation, de représenter les profils les plus aptes à correspondre à des modèles d'évaluation de la prosodie des non natifs d'après les 14 variables déterminées. On retrouve des lecteurs typés :

- **L3** = l'« acteur de théâtre » : il pose sa voix, qui est lente, avec des pauses prononcées. Il fait grand usage de proéminences.
- **L10** = le « conteur d'histoires » : son registre est extrêmement ample, sa vitesse de lecture très lente et ses pauses prononcées.
- **L6** = La position extrême de L6 est due à sa rapidité de lecture.
- **L1** = le « meilleur » lecteur : son registre est très ample, sa lecture rapide a un débit syllabique très élevé et une diction pauvre en proéminences. Des pauses sont perceptibles. On peut remarquer qu'il se définit à la fois par des caractéristiques appartenant à L6 et L10. C'est cette alchimie qui a su faire de lui le « meilleur lecteur » de l'expérimentation perceptive.
- **L4** = « mauvais » lecteur selon le groupe d'enseignants d'anglais francophones.

Le plan factoriel ne signale pas de caractéristiques prosodiques extrêmes concernant le lecteur L5, lesquelles pourraient le distinguer comme profil particulier. Ses réalisations se rapprochent davantage de celles de L1, L3 et L4 que de celles de L6 et L10. L5 se trouve à l'intersection entre les deux axes. Cette position centrale souligne un équilibre des variables les plus décisives dans l'établissement d'un profil, ce qui représente une moyenne atteignable pour les apprenants afin d'évoluer ensuite vers une typologie plus appropriée suivant la situation d'énonciation. Pour la lecture de ce type de texte, L1 est le modèle visé, car il a été désigné « meilleur lecteur » par de nombreux auditeurs, dont les locuteurs anglophones natifs. On peut rappeler que L5 a été désigné lecteur préféré des apprenants francophones, en particulier ceux de niveau B2/C1. La possibilité de sélectionner L5, non pas comme profil de lecture mais comme modèle à imiter serait en outre justifié par son âge (23 ans) car il correspond à la tranche d'âge de la plupart des étudiants inscrits dans le système universitaire.

Parmi ces 14 variables, certaines peuvent sembler redondantes, alors que d'autres sont davantage représentatives. La question se pose de connaître la pertinence de chacune d'entre elles afin de restreindre leur éventail tout en favorisant la pertinence maximale. D'un point de vue pragmatique, d'autres éléments doivent être pris en considération dans une optique

CHAPITRE II – Profils et profilage : mode d’organisation hiérarchique de la prosodie des anglophones natifs

d’évaluation tels que la facilité de collecte de données ou même la portabilité des critères obtenus vers l’analyse du discours spontané. Là non plus, une analyse par cercle de correspondances, destinée à d’importants effectifs de données quantitatives, n’est pas indiquée dans le cadre d’un faible nombre d’individus. Pourtant, son utilisation est emblématique des pistes à suivre afin de déterminer un prototype de modélisation des variables à sélectionner en fonction de l’objectif. Le graphique suivant permet de visualiser les relations entre les variables quantitatives ainsi que leur degré de pertinence dans l’établissement de profils.

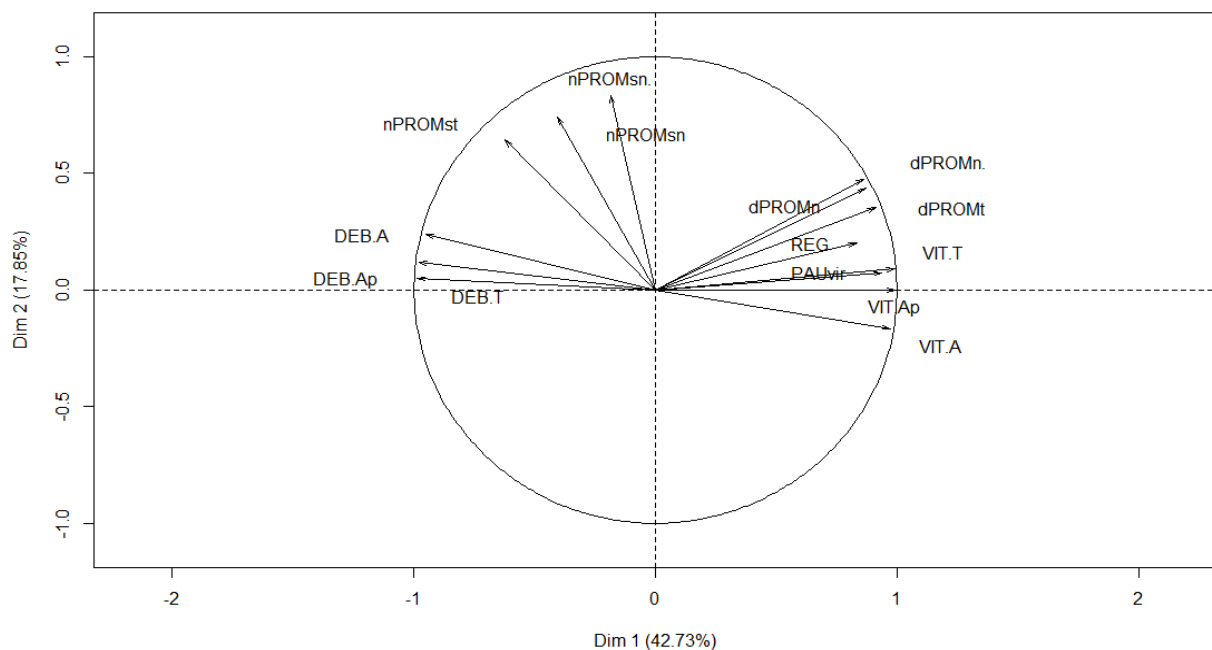


Figure 26 – Cercle des corrélations représentant les 14 variables continues dans leur dépendance et leur pertinence

Dans le graphique ci-dessus, seuls apparaissent les paramètres quantifiés. La longueur des flèches montre le degré de pertinence de la variable qu’elle représente. Le rapprochement de certaines flèches indique leur corrélation car deux individus appartenant au même groupe doivent être très proches en termes de comportement pour garantir une faible variance intra-classe. Ainsi, certaines variables peuvent être écartées par manque de pertinence et par la difficulté de prise de mesures à grande échelle.

Afin de préciser ou d’infirmer les résultats donnés par le plan factoriel et le cercle de corrélation, j’ai opté pour l’utilisation d’une classification ascendante hiérarchique, technique de classification qui regroupe les données quantitatives par similarité. J’ai ainsi testé les huit méthodes de la fonction *hcluster* avec les résultats suivants dans l’environnement de travail *RStudio*. Les six méthodes *single*, *average*, *mcquitty*, *median*, *centroid* et *centroid2* présentent

des dendrogrammes similaires par l'arbre de hiérarchisation des lecteurs et leur position de classement les uns par rapport aux autres. En voici ci-dessous l'exemple de la fonction *centroid2*.

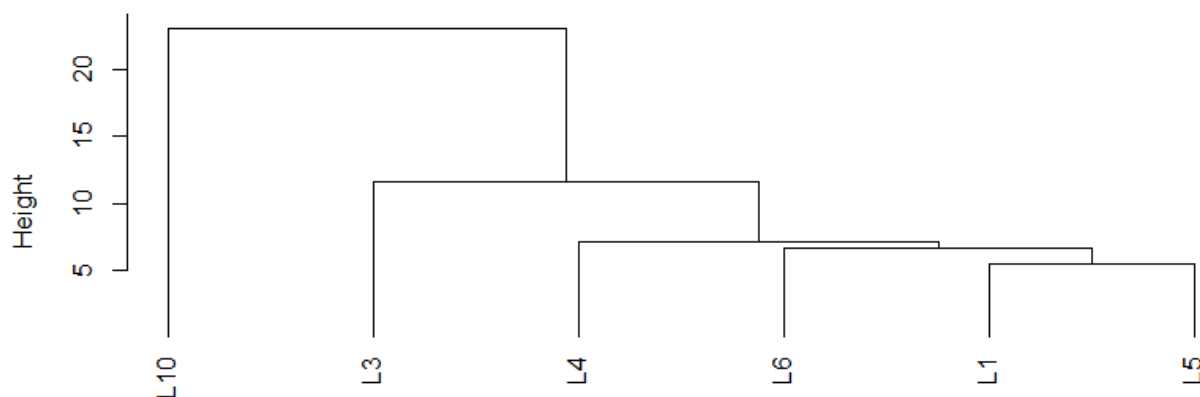


Figure 27 – Dendrogramme de hiérarchisation des 6 lecteurs natifs selon leurs variables quantitatives avec la fonction *centroid2*

On y voit que L1 et L5 sont présentés comme les plus semblables, corroborant les précédentes analyses. Ensuite, quelques similarités sont trouvées avec leur paire et L6. Ce nouvel ensemble est très proche de L4, puis le nouveau groupe marque plus de distance en englobant L3, puis encore davantage en intégrant L10. Le degré de similarité est proportionnel à la hauteur de l'embranchement vertical.

Les regroupements diffèrent pourtant dans les dendrogrammes effectués avec les méthodes *Ward* et *Complete*. Les deux méthodes sont similaires entre elles dans la ramification de l'arbre, mais diffèrent par l'agencement de L4 et L6. On peut remarquer dans le dendrogramme suivant, réalisé avec la fonction de *Complete*, deux groupes distincts : L3 et L10 ensemble en binôme et les autres lecteurs regroupés par ailleurs. La hauteur de l'embranchement vertical montre que, bien qu'ils soient regroupés, ce qui signale des similitudes, L3 et L10 ont moins de points communs entre eux que les autres lecteurs par ailleurs. La position de L4 entre L6 et L1 montre que L4 est considéré être plus semblable à L1 et L5 qu'à L6.

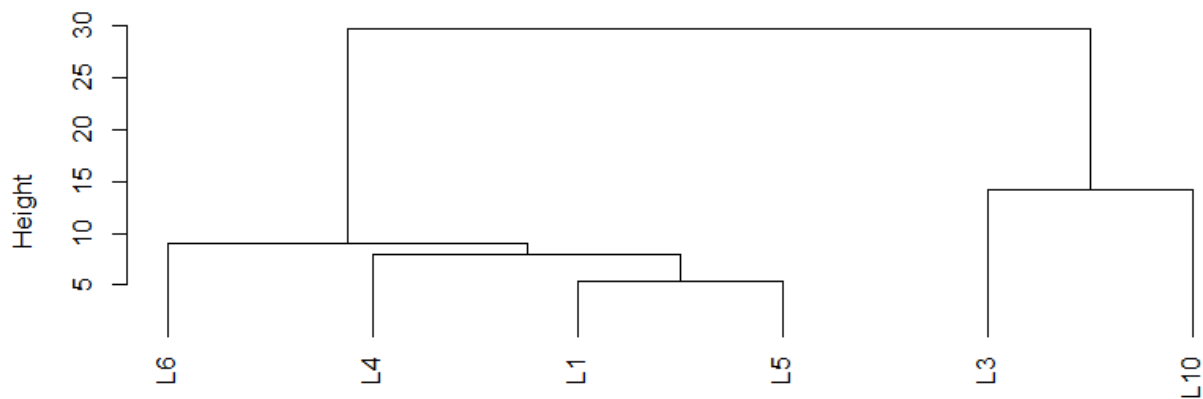


Figure 28 – Dendrogramme de hiérarchisation des 6 lecteurs natifs selon leurs variables quantitatives avec la fonction *Complete*

Il semblerait que la hiérarchisation présentée se base sur la variable de rapidité de lecture étant donné que le classement des six lecteurs les uns par rapport aux autres le vérifie.

Le dendrogramme effectué avec la fonction de *Ward* est semblable au précédent à une différence près : L6 se trouve placé hiérarchiquement entre les niveaux de L1 et de L4, en agencement inversé par rapport au dendrogramme avec la fonction *Complete*.

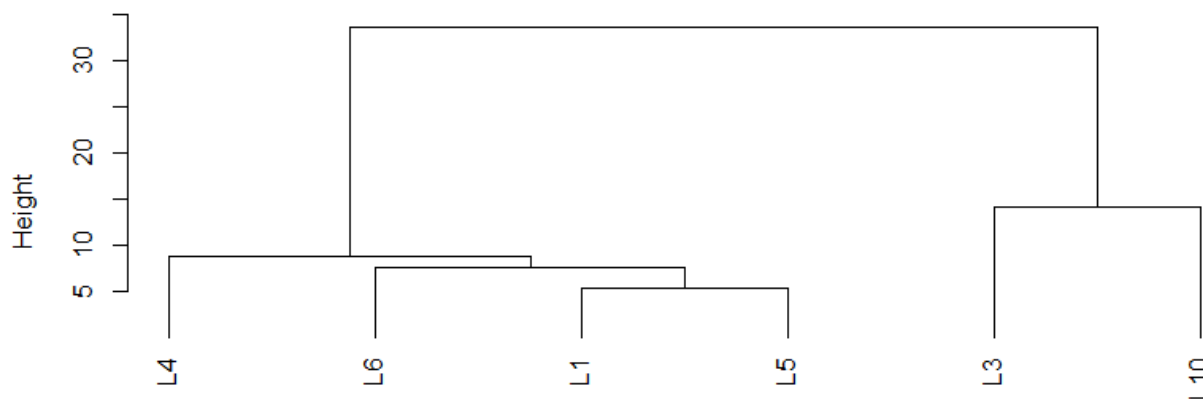


Tableau 35 – Dendrogramme de hiérarchisation des 6 lecteurs natifs selon leurs variables quantitatives avec la fonction *Ward*

Dans ce dendrogramme, L4 est excentré par rapport au groupe formé par L1, L5 et L6. Bien que moins rapide que L6, l'analyse des données acoustiques rythmiques avait déjà signalé le manque de cohérence de sa diction. Aussi, je pense que c'est ce graphique qui explique le mieux que L4 ait été stigmatisé « mauvais » lecteur par les enseignants d'anglais francophones.

6.4.4 Bilan de la phase d'expérimentation statistique STAT-N1

L'objectif final de l'expérimentation statistique STAT-N1 était de corroborer ou d'infirmer la sélection et la hiérarchisation qualitative des lecteurs natifs anglophones obtenue

par expérimentation perceptive et acoustique. L'analyse statistique a confirmé que les lecteurs L1 et L5 pouvaient être pris comme référence dans cadre d'une évaluation des apprenants.

7 Synthèse : création d'outils pour analyser l'interlangue

Ce chapitre sur les profils et son corollaire le profilage a plusieurs objectifs, mentionnés dans sa phase préliminaire, que je rappelle ci-dessous :

- établir des profils des anglophones natifs afin de sélectionner celui ou ceux qui peuvent être pris comme modèles d'évaluation prosodique normatifs,
- à partir de ces modèles, obtenir les critères acoustiques correspondant à une bonne lecture native, en évitant tout excès,
- valider par diverses méthodes la fiabilité des profils obtenus,
- étudier diverses méthodes de profilage afin de sélectionner celles qui peuvent être reprises dans la présente étude ou d'envisager d'autres méthodes plus adéquates,
- vérifier le potentiel de la méthodologie par profils et ses limites pour valider ou non l'Hypothèse 1.

Pour rendre compte des résultats obtenus à chaque étape, je vais aborder chaque question séparément.

7.1 Profils pour établir des modèles d'évaluation prosodique

On a souligné que le domaine prosodique était empreint de la plus grande variabilité. De plus, la lecture à haute voix obéit à des codes interlocutifs : la lecture doit être « réussie » au sens d'Austin (1962), c'est-à-dire que l'interlocution engendrée est « réussie » si la lecture est appréciée des auditeurs ou au contraire « ratée » si elle est jugée neutre ou insignifiante. Aussi, la quête d'un « bon » lecteur natif, qui soit le modèle adéquat pour étalonner les réalisations non natives dans un cadre évaluatif, est essentielle. Pour parvenir à identifier cette norme évaluative, j'ai choisi la méthodologie de repérage et de sélection par profil. Elle a abouti à déterminer une typologie de six lecteurs, à partir de 12 candidats natifs, dont deux constituent une norme adéquate : L1 et L5 sont reconnus les « meilleurs » lecteurs. L1 est un homme d'âge mûr britannique, L5 est une jeune américaine.

7.2 Critères acoustiques pour modèle et norme évaluatifs

À partir de la sélection des six lecteurs les plus typés, les impressions subjectives de tempo, de proéminences, de tessiture, etc. ont été traduites en corrélats acoustiques objectifs de durée (en seconde), de fréquence fondamentale (en demi-tons), et en données numériques du nombre de proéminences, de syllabes, de pauses, etc. Ces résultats ont été consignés dans une matrice de traits prosodiques regroupés en variables.

7.3 Validation de l'Hypothèse 1

En utilisant diverses méthodes d'investigation, les profils obtenus par questionnaires d'auditeurs ont subi plusieurs contre-expertises par profilage subjectif du lecteur lui-même, par profilage acoustique, ainsi que par profilages statistiques divers. Ils ont tous corroboré les premiers résultats de perception auditive, ce qui les a validés.

La lecture à haute voix est un style de parole particulier qui possède ses propres règles et exigences. Les profils ont montré que les variables prosodiques constituaient des styles de lecture et qu'elles étaient solidaires les unes des autres. En conséquence, on ne peut utiliser comme norme des données issues de variables indépendamment des profils : il faut garder la cohérence des profils en trouvant leur principe fédérateur, le dénominateur commun des styles de lecture, pour en extraire une norme.

Chaque profil a sa cohérence. Les commentaires subjectifs des auditeurs ont souligné les caractéristiques servant de base à une typologie des lecteurs. Les commentaires des lecteurs sur leur expérience et parcours ont contribué à expliquer les choix de diction souvent inconscients. La matrice de traits a pu confirmer et parfois infirmer ces commentaires. L'analyse statistique a permis de hiérarchiser les variables et les profils de lecteurs anglophones natifs en observant les similitudes et en écartant les variables non fédératrices. Elle a aussi justifié statistiquement les remarques et choix subjectifs des auditeurs ainsi que l'analyse des corrélats acoustiques. Le profilage par vérifications successives a validé la fiabilité des profils et de leur hiérarchisation car les résultats convergent et s'éclairent mutuellement. Ainsi, l'Hypothèse 1 de ce chapitre (rappelée ci-dessous) a été vérifiée et validée.

H1 — Étant donné la forte variabilité des réalisations prosodiques natives et d'apprenants, des outils spécifiques doivent être utilisés afin de l'analyser. La création de « profils prosodiques » devrait permettre de mieux organiser et

hiérarchiser les réalisations inter-locuteurs en fonction de regroupements des variables.

7.4 Potentiel et limites des méthodes de profilage de la littérature pour ma recherche

En début de chapitre, la présentation des définitions d'un profil et de la méthodologie canonique du profilage a permis d'utiliser les notions correspondantes afin d'asseoir les méthodes présentées sur des bases qui permettaient de comprendre le fonctionnement des profils dans leurs interactions. Comme l'a signalé Crystal, un profil est par définition une approximation. Il ne faudrait cependant pas voir dans cette faiblesse une incompatibilité entre les profils, qui sont des approximations, et l'évaluation, qui doit être rigoureuse et précise. Sa force réside au contraire dans sa faculté d'adaptation à l'objectif assigné. Les recherches présentées en amont ont souvent fait le choix de peu profiler leurs locuteurs. Le nombre de leurs catégories est alors restreint : souvent natifs et non natifs, parfois une diversification du niveau langagier des non natifs. D'autres domaines sont davantage catégorisés. Ainsi la méthodologie de Gut a préféré catégoriser les styles de parole plutôt que le niveau des non natifs dont elle a rendu compte.

Aujourd'hui, le profilage mis en œuvre favorise un nombre maximal de données acoustiques obtenues pour être utilisées en statistique avec les tests appropriés. Les profils personnalisés n'ont plus beaucoup la faveur des chercheurs actuels. Ils sont néanmoins utilisés parcimonieusement pour expliquer ou justifier un point de vue axé sur les résultats d'un profilage de groupe (Herment *et al.* 2014). Le détail du profil personnalisé se dilue dans le profil de groupe, lequel perd en valeur individuelle mais gagne en puissance et rapidité pour cibler une population à privilégier ou exclure. Ainsi le profilage du rythme a évolué de la mesure de l'accent revenant de manière cyclique à la durée de segments pouvant être comparés les uns par rapport aux autres.

On remarque peu d'études de la qualité prosodique à proprement parler : on profile la variation (en termes quantitatifs), sans vraiment savoir la hiérarchiser qualitativement. Laufer et Nation 1995, Horgues 2010, Herment *et al.* 2014 s'y sont attelés : les premiers en hiérarchisant qualitativement l'utilisation du lexique dans les productions écrites non natives, et la suivante en hiérarchisant qualitativement les tons utilisés dans des syllabes contraintes, les derniers en cataloguant finement les étapes de l'évolution de l'acquisition intonative des

apprenants par rapport aux natifs de L1 et L2 dans les questions en lecture suivie. Ces méthodologies à finalité qualitative retiennent toute mon attention.

On remarque en outre qu'une dimension prosodique a été écartée de mon étude, même si le profilage des natifs a trouvé sa cohérence dans les repères trouvés. L'étude de l'interlangue devra être complétée par la celle de la mélodie, qui représente une variabilité rebelle à un classement gradué par rapport à des valeurs, pour pouvoir donner une vue d'ensemble du domaine prosodique. D'après Crystal, c'est la mélodie qui est responsable de la majorité des difficultés dans le milieu clinique. Mais c'est l'ensemble des traits prosodiques qui doivent être étudiés, selon le *principe of profile comprehensiveness* ([1982] 1992 : 4-5, 114). Si la mélodie (courbe mélodique de F0) ne représente pas un enjeu majeur chez les natifs, tous les contours produits étant acceptables, il n'en est pas de même chez les non natifs, chez qui les schémas intonatifs peuvent être en totale contradiction avec ceux de la langue cible. En effet, le registre ne peut à lui seul rendre compte de l'intonation, dans son sens le plus restreint de mélodie. Étant donné son extrême variabilité, réfractaire à son évaluation qualitative selon des critères objectifs, la mélodie a besoin d'un profilage particulier pour prendre cette variabilité en considération, à plus forte raison dans une optique évaluative. En conséquence, un nouveau profilage sera dédié à l'étude de l'intonation dans le prochain chapitre : il utilise l'emphase comme filtre.

Alors que le profilage que j'ai utilisé dans ce chapitre se base sur la nature syntagmatique du discours car les mesures étaient prises le long de l'axe temporel, l'emphase est de nature paradigmatique car elle se définit par son degré sur une syllabe ou zone donnée. En conséquence, l'utilisation du phénomène d'emphase se superpose à la méthodologie de profilage syntagmatique en la complétant sur un plan paradigmatique. Mon hypothèse est qu'une localisation des parties du texte les plus emphatiques devrait fournir les zones d'étude mélodique les plus pertinentes. Ces données nouvelles permettraient de donner aux profils toute leur valeur afin de comprendre et recadrer la portée des variables les unes par rapport aux autres, et ainsi de contribuer à calibrer les critères évaluatifs prosodiques.

7.5 Détermination de ma méthode de profilage

Le profilage que j'ai mis en œuvre n'est qu'une réponse *ad hoc* à une problématique située. Elle se veut avant tout pragmatique pour pouvoir contribuer de la meilleure manière possible à atteindre les objectifs d'évaluation prosodique en lecture des non natifs. Voyons comment les études décrites précédemment contribuent à orienter et façonner le processus

méthodologique de modélisation évaluative dans ses diverses dimensions et phases. Après avoir présenté les principes fondamentaux du profilage et plusieurs de ses applications sur le terrain, je vais à présent pouvoir fournir une approche modélisatrice de mon étude dans la quête de critères évaluatifs.

7.5.1 Schématisation de ma recherche selon le profilage *CRISP-DM*

Même si le profilage mis en œuvre dans cette thèse ne concerne pas des centaines de locuteurs, son organisation peut être modélisée selon les préconisations *CRISP-DM*. En voici ci-dessous une transposition.

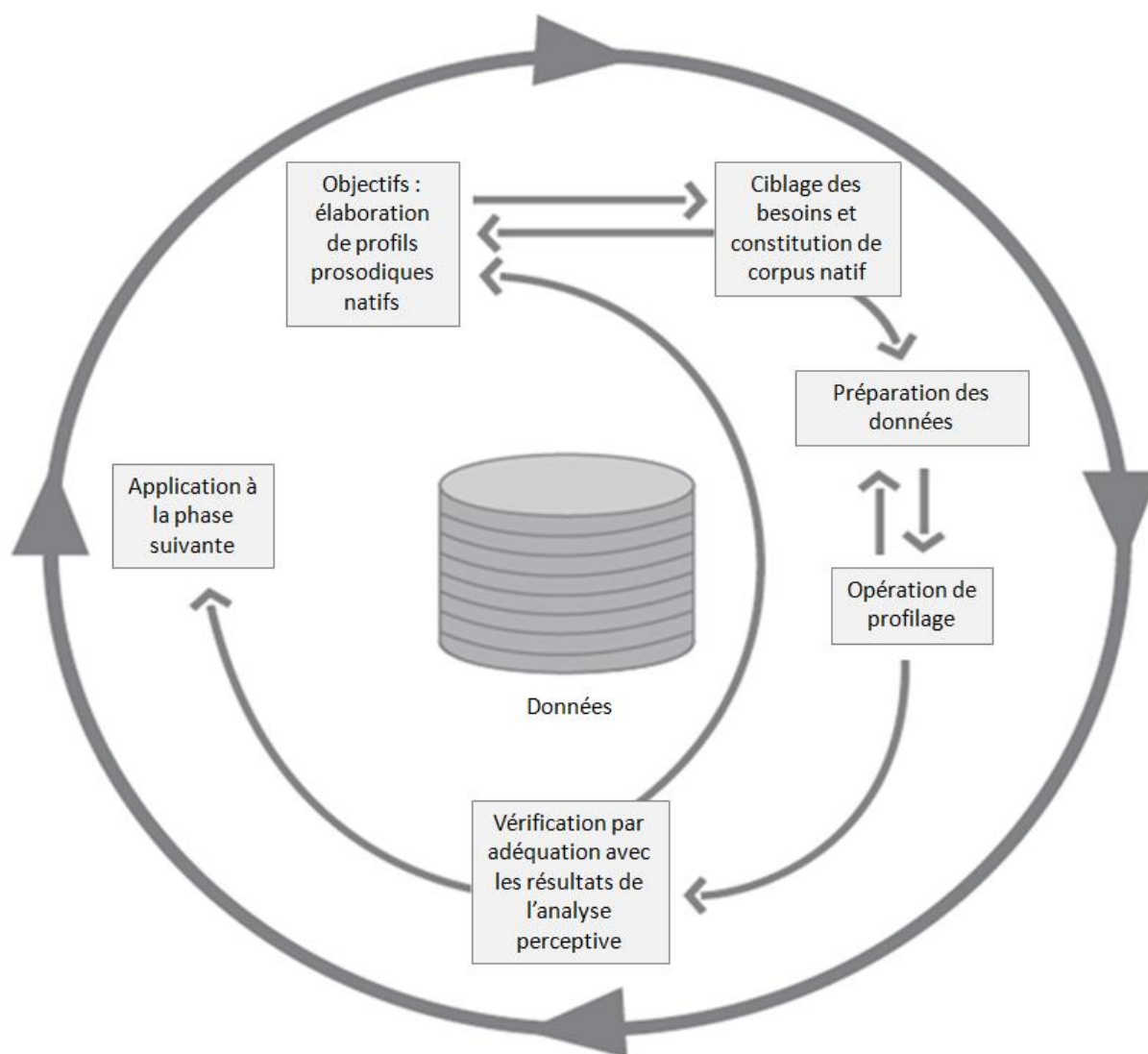


Figure 29 – Schématisation de la recherche d'une norme prosodique native selon le profilage *CRISP-DM* (à partir du *CRISP-DM Process Guide and User Manual*)

Ce schéma montre clairement la phase « opération de profilage » et la nécessaire récursivité de la « vérification par adéquation avec les résultats de l'analyse perceptive ». Cet

CHAPITRE II – Profils et profilage : mode d’organisation hiérarchique de la prosodie des anglophones natifs

agencement des réalisations par rapport au résultat prévu est susceptible de remettre en question sa validité et de conduire à une réadaptation. L’« application à la phase suivante » est l’expérimentation du chapitre suivant sur l’emphase, une fois que les diverses vérifications ont donné satisfaction. Ce schéma devrait pouvoir être utilisé pour modéliser chaque phase expérimentale de ce travail de recherche. Il présente les différentes étapes du processus de manière très générique et peut être décomposé plus précisément en quatre phases, comme ci-dessous.

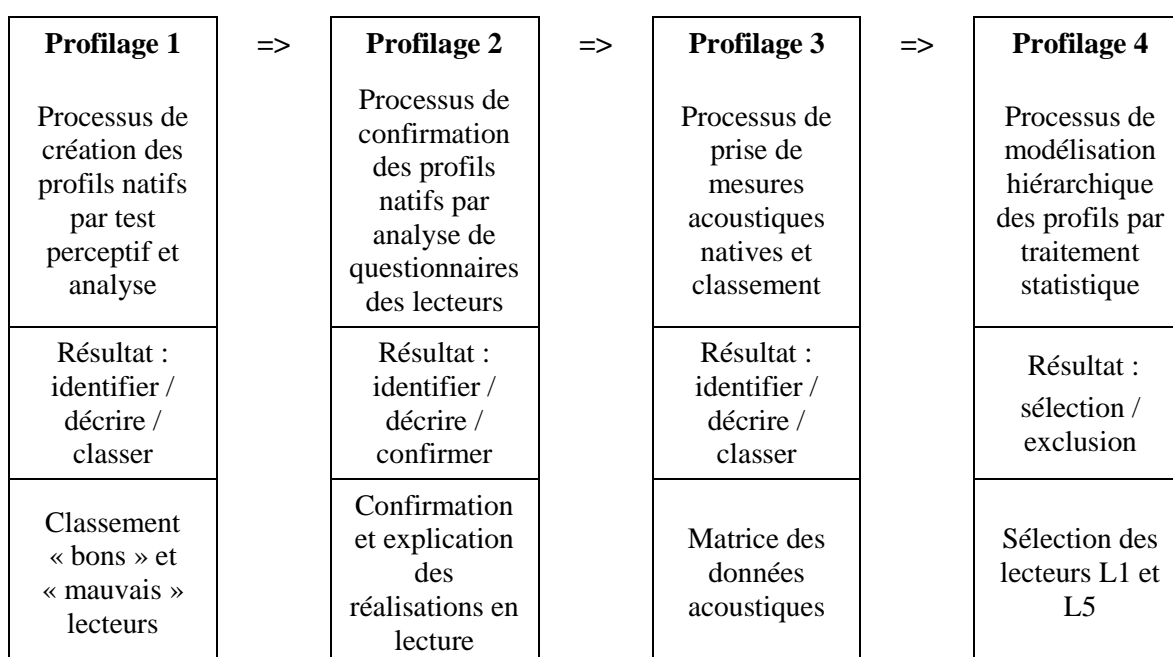


Tableau 36 – Décomposition du processus de profilage des lecteurs natifs en 4 phases

On y voit clairement les quatre étapes de classement, confirmation, analyse acoustique puis statistique qui a permis de créer les profils de lecteurs natifs et de les hiérarchiser dans l’expérimentation de ce chapitre. Le chapitre suivant poursuivra le processus en y ajoutant les données de l’étude de l’emphase. Puis l’analyse des productions non natives complètera le processus de modélisation. On a vu ci-dessus que, selon Chapman *et al.* ([1999] 2000 : 51), l’aboutissement du modèle de référence *CRISP-DM* devait pouvoir se traduire selon leur formule : *RESULTS = MODELS + FINDINGS*. Si l’on conçoit cette thèse comme un processus méthodologique destiné à créer des traits distinctifs de la prosodie non native à finalité évaluative, on devrait sans doute pouvoir la transposer ainsi :



7.5.2 Mes outils de profilage

Pour mener à bien l'expérimentation de ce chapitre, je me suis bien sûr inspirée de nombreuses notions et pratiques de recherche que j'ai présentées en début de chapitre. On a pu les retrouver mentionnées dans la méthodologie de l'expérimentation. En matière de profils, je souhaiterais néanmoins préciser certains concepts utilisés dans mon profilage.

7.5.3 Profils personnalisés : objectifs

On a vu que le profil personnalisé devra faire ressortir les particularités qui distinguent le locuteur étudié des autres. Il s'agit de préciser en quoi il est différent, de cerner les traits les plus déterminants par leur non conformité avec ce que l'on peut supposer être la norme des autres locuteurs, et donc de hiérarchiser ces traits entre eux de manière à pouvoir immédiatement repérer, étiqueter, signaler ces traits en établissant une fiche signalétique du locuteur. Les traits les plus distinctifs prennent alors le pas sur ceux qui sont plus conformes à la norme. Les profils personnalisés sont en fait un passage quasi obligé, une première phase afin d'accéder aux profils de groupes. La question est alors de détecter la saillance de ces traits. En contrepartie, le faisceau de variables marquant l'individualité du locuteur est destiné à être réduit au minimum.

7.5.4 Profils de groupes : objectifs et profilage

Alors que les profils personnalisés se veulent fidèles à la représentation précise d'un individu, les profils de groupes rassemblent certains individus afin de les distinguer d'autres individus, lesquels constituent d'autres groupes. Par son objectif de catégorisation, le profilage consiste à magnifier ou minimiser certaines particularités distinctives jugées peu représentatives au profit d'autres jugées fondamentales. Cette redistribution des caractéristiques définissant les individus permet de faire surgir une répartition susceptible de porter un éclairage différent mettant en valeur de nouveaux critères. Ce profilage est basé sur l'application des notions d'inclusion et d'exclusion. Pour répondre au besoins de catégorisation et de hiérarchisation, je considérerai deux sortes de profilages : celui conduisant à la création d'une catégorie de classement, dont je nommerai le profil résultant « archétype de conformité », et celui déterminant la borne des possibles, qui aboutit au profil « archétype stéréotypique ».

7.5.4.1 Archétype de conformité : catégorie de classement

Comme on l'a vu précédemment, les profils de groupes destinés à établir des catégories de classement obéissent à une autre logique que celle des profils personnalisés. Ils

visent le repérage non plus des points qui les distinguent les uns des autres, mais des points communs à l'intérieur d'un groupe. Le faisceau de variables pourra sceller l'appartenance au groupe dans le cas de profils distributifs (caractéristiques partagées par tous les membres du groupe par leurs propriétés définitoires) ou simplement être indicatif de l'appartenance au groupe dans le cas de profils non-distributifs (tous les individus d'une catégorie ne possèdent pas obligatoirement cette caractéristique : leurs propriétés sont accessoires). La méthodologie de profilage consistera à repérer les traits constituant les variables ainsi que les variables mettant en valeur ces traits. Ce profilage de classement et de catégorisation doit permettre de repérer des traits acoustiques ou perceptifs et vérifier si ils sont bien représentatifs d'une catégorie donnée, qu'elle soit établie *ad hoc* ou découverte par manipulation statistique par la suite⁵². Ce profilage aboutit à un profil que j'appelle « archétype de conformité » parce qu'il établit la conformité d'un individu par rapport aux caractéristiques d'un groupe.

7.5.4.2 Archétype stéréotypique ou de divergence : borne des possibles

Le profil que je nomme « archétype stéréotypique » ou « archétype de divergence » est d'une toute autre nature. On recherche une tendance extrême, une limite à atteindre, ou au contraire à éviter, basée sur des réalisations objectives, qui sont celles de profils personnalisés et des données de leurs variables : on traite donc des données réelles comme si elles provenaient d'un type et étaient virtuelles. Ces tendances extrêmes délimitent un espace de réalisations objectives parce que constatées. Elles permettent de capturer l'essence d'une tendance, sans l'exagérer ou la déformer en la représentant de manière très réductrice, car on se focalise sur ce qui l'éloigne de la norme. Cet archétype permet d'organiser une catégorisation des individus les uns par rapport aux autres en fonction de leurs traits les plus saillants et de repérer les tendances extrêmes d'un individu afin de hiérarchiser leur importance. Pour cela, on écarte les variables peu pertinentes caractérisant un individu au profit de celles qui le sont davantage.

⁵² Husson *et al.* (2009 : 2-3) : « Lorsque les individus sont décrits par un grand nombre de variables, il est nécessaire d'avoir un outil pour explorer l'espace dans lequel elles évoluent. L'étude des individus consiste à appréhender les ressemblances entre individus du point de vue de l'ensemble des variables. En d'autres termes, à dresser une typologie des individus : quels sont les individus les plus proches (resp. les plus éloignés) ? Existe-t'il des groupes d'individus homogènes du point de vue de leurs ressemblances ? »

C'est cette approche typologique que j'ai utilisée dans la hiérarchisation des lecteurs par plan factoriel⁵³. Son avantage est de faire intervenir peu d'individus et d'obtenir des données plus extrêmes que si elles provenaient de moyennes « archétype de conformité ».

Voyons à présent comment l'emphase peut compléter le système de variables pour caractériser l'intonation des locuteurs natifs.

⁵³ Husson *et al.* (2009 : 3) : « Un autre aspect consiste à chercher des dimensions de variabilité communes qui opposent des individus extrêmes avec des individus intermédiaires.

CHAPITRE III – Emphase : mode d’organisation paradigmatique des réalisations prosodiques des anglophones natifs

1 Résumé du Chapitre III

Ce troisième chapitre a pour fonction d’examiner la prosodie des anglophones natifs selon le mode d’organisation en degrés d’emphase. Son objectif fondamental est de repérer les variables à prendre en compte sur le plan paradigmatique afin de contribuer à construire un modèle prosodique évaluatif natif.

Pour cela, ce chapitre introduit la notion d’emphase et de son corollaire la saillance, puis s’attache à élaborer un test perceptif de détection et hiérarchisation de l’emphase chez les anglophones natifs (expérimentation PER-EMPH-N). Ses résultats permettent alors de calibrer chaque mot du texte selon son degré d’emphase pour sélectionner les marqueurs à étudier. Ainsi l’expérimentation COR-AC-N2 se centre sur l’étude acoustique des quatre principaux marqueurs d’emphase (complétée par les expérimentations PER-TON-N, ANN-TON-N1 et ANN-TON-N2 de perception et d’annotation tonale), alors que l’expérimentation COR-DIV-N3 (complétée par l’expérimentation PER-ACLEX-N portant sur l’accentuation lexicale) se focalise au contraire sur cinq catégories de marqueurs à l’emphase réduite.

Ce chapitre aboutit à la validation de l’Hypothèse 2 selon laquelle le phénomène d’emphase amplifie la variabilité des entités. En cela, il permet d’établir une sélection des marqueurs à privilégier ou à exclure de l’étude évaluative chez les non natifs sur l’axe paradigmatique, laquelle affine celle établie dans le chapitre précédent sur l’axe syntagmatique. Ce chapitre montre en outre le potentiel du phénomène d’emphase en tant qu’outil dans une optique de profilage.

Dans le chapitre précédent, nous avons vu que Crystal présentait ([1982] 1992 : 1) la raison d’être de la création d’un profil linguistique, qui est de pouvoir évaluer avec suffisamment de justesse l’état du patient afin de rendre possible une intervention réparatrice. Même s’il n’est qu’une approximation située à la croisée des domaines pratiques et

théoriques, le profil se justifie par son rôle dans un dispositif de remédiation. C'est aussi de cette manière que je concevrai l'emphase : comme clef nécessaire à l'investigation de l'intonation afin que la mélodie puisse prendre sa place parmi les variables constitutives des profils prosodiques. La confrontation à la variabilité des réalisations prosodiques et l'impossibilité de pratiquer une analyse exhaustive de tous les fichiers dans tous leurs aspects implique l'adoption d'une méthodologie particulière destinée à opérer des choix de sélection et de catégorisation. Le phénomène d'emphase appliqué à la lecture des natifs organise la sélection des marqueurs et éléments à étudier dans chaque fichier d'apprenant. Dans certains cas, cette variation fournira un paradigme de réalisations auquel les réalisations des apprenants pourront être confrontées, alors que dans d'autres la variation traduira des choix interlocutifs peu compatibles avec la nécessité de limiter la variabilité à des fins évaluatives. Plus précisément, l'étude sur l'emphase correspond à deux finalités spécifiques :

- Le besoin de distinguer la variation attitudinale de la variation grammaticale : afin de distinguer les entités stables de celles soumises à une variabilité plus ou moins forte,
- Le besoin de sélectionner les entités les plus emphatiques, lesquelles devraient fournir les phénomènes intonatifs les plus intéressants, l'emphase étant attendue pour jouer le rôle d'une loupe grossissante.

La détection du phénomène d'emphase est donc utilisée pour filtrer positivement ou négativement des entités, respectivement à étudier ou écarter de l'étude, déterminant une présélection intra-locuteurs.

Je rappelle mon hypothèse sur le rôle de l'emphase dans cette étude :

H2 — Le phénomène d'emphase favorise la variabilité des réalisations prosodiques. Aussi, son repérage permet de limiter l'étude à des entités concentrant la variation maximale. Le phénomène d'emphase devrait alors permettre de mieux comprendre, organiser et hiérarchiser la variabilité des phénomènes prosodiques intra-locuteurs en graduant la variation pour l'analyser plus précisément ou au contraire pour l'exclure des analyses.

C'est dans cette logique que je vais présenter le phénomène d'emphase en le définissant, puis en présentant ses paramètres perceptifs et ses corrélats acoustiques avant de présenter les expérimentations.

2 Phénomène de l’emphase

Afin d’introduire le phénomène d’emphase, je vais le définir, décrire son fonctionnement perceptif et acoustique, puis expliquer mes choix de corrélats obtenus par perception et de corrélats acoustiques.

2.1 Définitions de l’emphase

2.1.1 Définitions à partir de dictionnaires et encyclopédies

Le premier sens que donne le *Dictionnaire de français Larousse* en ligne⁵⁴ du mot « emphase » est : « Exagération pompeuse dans le ton, le geste, dans les termes employés ; enflure, grandiloquence : Parler avec emphase. ». D’après cette définition négative, l’emphase souligne le caractère théâtral et artificiel du mot. Un deuxième sens en donne une définition plus technique : « Mise en relief d’un des constituants de la phrase par l’intonation ou par l’ordre des mots. » La nature du terme est ici mise en liaison avec les procédés prosodique ou syntaxique qui lui donnent vie. En rhétorique, Calas en donne (2011 : 271) la définition technique suivante, neutre, qui rapprocherait la notion d’emphase de l’écrit puisque la parole n’y est pas mentionnée :

Emphase : terme grammatical et rhétorique. En rhétorique, il désigne tous les procédés qui permettent de renforcer une image, une idée. Elle peut recourir à l’hypothypose, à l’hyperbole ou à la répétition. C’est un terme général qui regroupe un ensemble très vaste de figures.

Pourtant, l’*Encyclopédie, ou Dictionnaire raisonné des sciences, des arts et des métiers* stipule que la rhétorique est l’« art de parler » (Diderot, D’Alembert, 1751-1772 : vol. 14 : 250), insistant sur le caractère oral du terme ; et l’un des collaborateurs de l’*Encyclopédie* qualifie (1751-1772 : vol. 5 : 577) l’emphase dans la même ligne de pensée :

EMPHASE, s. f. (*Belles - Lettres.*) énergie outrée dans l’expression, dans le ton de la voix, dans le geste.

Emphase se prend ordinairement en mauvaise part, & marque un défaut, soit dans les paroles, soit dans l’action de l’orateur. [...] ce n’est sûrement pas un éloge. [...]

On peut y détecter un facteur d’« énergie » relevant de la multimodalité, qui devrait se traduire par l’intensité du signal de parole. L’appréciation très négative qui en est faite dans la diction provient du caractère exagéré de l’expression, lequel suggère alors implicitement

⁵⁴ [<http://www.larousse.fr/dictionnaires/francais/emphase/28913>], consulté en 2016.

comme modèle son contraire : le naturel. Cette définition rejoint la ligne de pensée du premier sens donné par le *Dictionnaire de français Larousse* alors que celle de Calas est plus proche de la deuxième.

Ces quelques approches disparates me conduisent à explorer davantage la notion d'emphase telle qu'on la trouve exprimée dans quelques dictionnaires et encyclopédies de langue française puis anglaise afin de mieux cerner les concepts associés aux termes « emphase » et *emphasis* et en trouver le dénominateur commun. En voici le bilan.

À partir des cinq ouvrages que sont l'*Encyclopédie* de Diderot et d'Alembert (EDRSAM), du *Trésor de la Langue Française Informatisé* (TLFI), du *Petit Robert* (PR), du *Dictionnaire de français Larousse* (DFL) et l'ouvrage sur la *Stylistique* de Calas (SMC), consultés en août 2016, je peux catégoriser les notions qui y sont exposées en quatre groupes. À l'exception de l'ouvrage de Calas, une notion négative correspondant à la perception de l'auditeur y est systématiquement et prioritairement exposée : « énergie outrée », « défaut » (EDRSAM), « force expressive exagérée », « exagération dans la manière » (TLFI), « Exagération pompeuse » (DFL), « Emploi abusif ou déplacé », « ton déclamatoire », « exagération » (PR). L'objectif du locuteur, par contre, n'est exprimé que deux fois : « Mise en relief d'un des constituants de la phrase » (DFL) et « renforcer une image, une idée » de Calas. Les moyens techniques utilisés par le locuteur pour atteindre son objectif sont diversifiés et se déploient des plus larges aux plus spécifiques dans les cinq ouvrages : « tous les procédés » (Calas 2011), « énergie » (EDRSAM et PR), « Accent particulier porté sur un constituant de la phrase » (TLFI), « par l'intonation ou par l'ordre des mots » (DFL). La localisation de l'emphase est mentionnée dans trois ouvrages seulement : alors qu'en grammaire transformationnelle un accent particulier sera « porté sur un constituant de la phrase » (TLFI), le ton de la voix et le geste sont mentionnés dans EDRSAM, TLFI et DFL ; TLFI y ajoute « le comportement » ainsi que « la manifestation d'un sentiment », DFL y inclut « les termes employés ». On peut en conclure que la notion d'« emphase » est de portée négative, voire péjorative, elle se manifeste dans l'expression générale du locuteur, mais surtout le ton de sa voix et sa gestuelle, par le truchement d'une énergie accrue, une intonation particulière, parmi de nombreux procédés stylistiques permettant une mise en relief.

Semblablement à l'étude précédente sur la notion de *profile*, j'ai consulté cinq dictionnaires de langue anglaise pour repérer d'éventuels glissements de sens, similitudes ou divergences par rapport au concept français. Les cinq dictionnaires de langue anglaise

CHAPITRE III – Emphase : mode d'organisation paradigmatique des réalisations prosodiques des anglophones natifs

consultés en août 2016 sont le *Cambridge American English Dictionary* (CAED), le *Cambridge British English Dictionary* (CBED), le *Longman Dictionary of Contemporary English* (LDOCE), le *Macmillan Dictionary* (MMD), le *Merriam-Webster's Online Dictionary* (MWD) et l'*Oxford English Dictionary* (OED). Contrairement à la notion française, le mot *emphasis* ne semble être défini que par quatre notions fondamentales. L'objectif du locuteur est au cœur de toutes les définitions : utilisation quasi systématique des mots *particular* ou *special*, *importance* ou *attention*, avec *impressiveness* (MW), *because it is important or because you want it to be noticed* (CBED en anglais américain), *prominence* (OED). La subjectivité de l'auditeur n'est que peu spécifiée : *force or intensity of expression* (MW), *Vigour or intensity of expression* (OED). Par contre, les moyens techniques mis en œuvre par le locuteur sont mentionnés avec plus ou moins de précision : *particular prominence* (MW), *extra force* (CBED), *stress* (OED), *louder* ou *extra loudness* (MMD, LDOCE), *higher* (LDOCE), et *by printing it in a special way* (LDOCE) lorsqu'il s'agit spécifiquement d'écrit. Sa localisation varie d'une portion de mot au groupe de mots, en passant par le mot lui-même. Si MW ne fait aucunement mention d'un lieu quelconque, les autres dictionnaires précisent la nature orale de l'emphase et LDOCE y ajoute même celle de l'écrit (*or by printing it*). Aucun jugement négatif n'est perceptible. En résumé, le terme *emphasis* est utilisé lorsque le locuteur ou écrivain veut attirer l'attention de son auditoire ou de son lecteur sur une partie du discours pouvant s'étendre d'une portion de mot jusqu'à un énoncé de plusieurs mots, par une force ou vigueur du propos toute spéciale, dont la prééminence se traduit, à l'oral, par une accentuation, une intensité, un mouvement mélodique montant, ou à l'écrit par une typographie différente.

On peut alors constater la différence de sens des deux mots « emphase » et *emphasis*, qui correspondent chacun à deux réalités différentes, le concept français étant subjectif et de connotation défavorable alors que la notion anglaise se présente comme neutre et utilitaire, spécifiant son objectif, les moyens techniques à mettre en œuvre pour l'atteindre ainsi que leur localisation. Leur dénominateur commun est la volonté d'attirer l'attention sur une information spécifique de leur discours par une modification de l'intensité ou la hauteur mélodique à l'oral, ou une modification de la typographie à l'écrit. C'est d'ailleurs l'analyse que fait Roggero (1977 : 259) de la différence entre les deux termes :

Il faut d'abord bien saisir le sens du terme *emphatique* ; on ne peut y parvenir si l'on a présent à l'esprit le mot français «emphase» ; il s'agit en réalité d'un

anglicisme. Le verbe anglais *emphasize* signifie «mettre l'accent sur, souligner» ; le nom *emphasis* correspond donc à l'accent (mis sur tel élément).

Selon Riegel, Pellat, Rioul dans leur *Grammaire méthodique du français* ([1994] 2013 : 667), « L'*emphase* – terme emprunté à la rhétorique – désigne tout procédé d'insistance ou de mise en relief ». Cette nouvelle définition technique permet de cadrer le concept plus finement parce qu'elle me conduit à clairement percevoir les deux faces du phénomène de l'emphase : l'objectif est d'insister sur une partie du discours, et le moyen pour y parvenir est le procédé de mise en relief utilisé. L'application du phénomène implique donc un nécessaire binôme : un émetteur (source de la volonté, de l'assignation de l'objectif) et un destinataire (personne réceptrice du message). Le concept implique en outre deux plans : celui du discours ordinaire, normal, et celui de l'emphase, dont le rôle est de se détacher de cet arrière-plan. On retrouve alors un point commun entre l'emphase et une des définitions du mot « profil »⁵⁵ dans son sens premier (contour, *outline*), sur les deux plans différents constitués d'un second plan (fond à partir duquel se détache le profil) et le premier plan, qui constitue la silhouette qui se détache du fond. Les deux concepts de « profil » et d'« emphase » se rejoignent dans leur mission de mettre en relief des caractéristiques particulières afin qu'elles soient immédiatement perceptibles et accessibles à celui à qui le message est destiné, en dissociant le second plan que sont les traits ou caractéristiques ordinaires, normales, du premier plan, qui marque une différence essentielle en véhiculant une valeur remarquable. Le second plan constitue une norme dont le premier plan s'écarte. On voit à présent clairement que « profil » et « emphase » sont des concepts convergents et représentent deux manières complémentaires d'atteindre un objectif de différenciation.

Je vais pouvoir conclure cette étude définitoire générale en m'orientant vers un approfondissement par la relation que l'emphase entretient avec la notion de « saillance ».

2.1.2 Emphase et saillance

Le terme de « saillance » est défini dans son sens premier par « Qui fait saillie, dépasse, déborde sur la ligne normale » (DFL)⁵⁶ et son sens figuré par « Qui est en évidence, ressort du contexte et s'impose à l'attention » (PR). Même si l'on s'en tient à ces seules définitions, on voit très nettement le caractère prosexigène qui unit les trois notions de

⁵⁵ Références en début de Chapitre II : 2.1 Analyse lexicographique du mot « profil ».

⁵⁶ Consulté en 2016.

« profil », « emphase » et « saillance ». Il justifie l'intérêt porté à la saillance pour l'étude de l'emphase.

2.1.2.1 Définition et propriétés de la saillance linguistique

Inkova rapporte (2012 : 9) que dans la littérature linguistique, le mot « saillance » se réfère à « l'importance, l'émergence, ou encore la récence d'une entité dans un contexte donné, propriété qui lui permet d'être perçue, repérée plus facilement au milieu d'autres entités ». La « saillance » correspondrait en quelque sorte à un « écart » ou une « variation » par rapport à l'ensemble des autres entités (2012 : 19).

Toujours en linguistique, Landragin la définit ainsi (2012) : « La saillance, c'est avant tout l'émergence d'une forme sur un fond, la mise en avant d'une entité par rapport à d'autres entités. Etre saillant, c'est ressortir particulièrement, au point de capter l'attention et de donner une accroche, un point de départ à la compréhension. ». Le caractère saillant d'une entité relève du fait qu'elle est différente des autres par un trait, une propriété qui la singularise, et de ce fait, l'isole. De là provient sa dimension cognitive car la propriété de « capter l'attention », de venir « en premier à l'esprit » (Landragin 2006), valide l'efficacité de la saillance. Que les facteurs de saillance soient prosodiques (par leur prééminence, leur rythme d'élocution fracturé, voire une intonation inattendue), lexicaux ou proviennent d'un autre domaine, c'est la saillance qui donne accès au sens car « l'élément saillant est celui que l'on conceptualise le plus rapidement, c'est le point d'ancrage sémantique. C'est ce qui fait l'importance de la notion dans les études sémantiques et pragmatiques. » (Landragin 2012). En effet, Landragin émet l'hypothèse que le phénomène de saillance est à l'origine d'un « mécanisme cognitif dépassant le langage oral, écrit ou visuel » (Landragin 2011). L'entité mise en saillance par des mécanismes propres à son domaine prosodique, lexical, syntaxique ou sémantique, est de fait favorisée par rapport au reste du message et de son contexte lors du processus de compréhension car « L'interlocuteur ou lecteur ne reçoit donc pas uniformément le contenu d'un message, mais en perçoit certaines parties plus aisément et rapidement que d'autres. ». Cet élément devenant linguistiquement saillant dans la présentation du message de par son importance cognitive particulière, il sera traité prioritairement jusqu'à devenir central dans la réception et la réaction du destinataire. Le phénomène de saillance est inhérent à la communication linguistique dans la construction du sens, tant au niveau de la réception que de la production.

La saillance [...] relève moins d'un phénomène de langue que d'un mécanisme cognitif général. [...] elle peut intervenir dans toute interprétation et production linguistiques, en tant que point de départ pour la construction du sens (interprétation) et pour l'articulation d'un message (génération). (Landragin 2011)

À chaque étape de la génération linguistique correspondent des choix spécifiques qui peuvent être traduits en terme de saillance (Landragin 2006). Étant donné que la saillance est une propriété commune à de nombreux éléments, Landragin en conclut (2006) le caractère relatif du phénomène, dépendant de divers facteurs et des niveaux d'analyse adoptés, excluant la possibilité d'une saillance absolue. Ce caractère relatif aboutit logiquement à une gradation des niveaux de saillance des entités, dont des échelles de saillance pourraient rendre compte, lesquelles aboutiraient à des hiérarchisations.

2.1.2.2 Fonctionnement de la saillance (en vue de modélisation)

Landragin s'interroge (2006) sur les propriétés linguistiques qui constituent les facteurs de saillance afin de pouvoir les regrouper en diverses catégories pour faciliter une modélisation. Il remarque que c'est une combinaison de plusieurs critères qui est à l'origine du degré de saillance et précise que certaines constructions syntaxiques sont « quasiment dédiées à la saillance ». Il constate qu'il est difficile d'opérer une hiérarchisation en terme de saillance et que l'on doit se résoudre à utiliser les propriétés linguistiques classiques pour s'en approcher.

Landragin présente (2015) une notion multicritère et multidimensionnelle propre à englober la multiplicité des cas pouvant intervenir. Cette modélisation inclut cinq dimensions principales en se focalisant sur les effets interprétatifs (saillances globales versus saillances locales) et la nature et l'objectif de la communication inter-interlocuteurs, tout en abordant les styles de parole. Chacune d'entre elles est au minimum binaire, ce qui entrave leur modélisation. Il s'agit de la :

- Saillance visuelle et saillance linguistique
- Saillance physique et saillance cognitive
- Saillance à effet immédiat et saillance à effet continu
- Saillance préalable et saillance nouvelle
- Saillance informative et saillance rhétorique

En ce qui concerne la saillance physique et la saillance cognitive, Landragin (2015) distingue trois types de saillance : la saillance physique (P-saillance), la saillance cognitive (C-saillance) et la saillance située (S-saillance). Les deux premières dimensions dissocient (Landragin 2011) forme et contenu.

La P-saillance ou saillance physique concerne les « aspects de saillance détectables via une trace physique contenue dans le signal en tant que matériau physique » (2015). Il s'agit d'une saillance objective relayée par des « critères physiques », « objectifs » et « explicites » (Landragin 2011), parfois simplement inférables à partir du message transmis. Leurs traces mesurables sont compatibles avec une modélisation computationnelle de la saillance. Parmi ces critères physiques, certains sont inhérents à leur nature. D'autres dépendent d'autres éléments par mise en relief contextuelle. Ainsi, dans un ensemble d'éléments, celui qui se différencie des autres parce qu'il ne sera pas doté des mêmes particularités, sera saillant par rapport aux autres : « être un **singleton** est un critère de saillance, et la saillance ne découle pas d'une propriété particulière, mais du fait d'être un singleton pour une propriété » (Landragin 2011). Cette mise à l'écart pourra être organisée comme à l'oral par une mise en évidence prosodique ou à l'écrit par un procédé typographique.

La C-saillance ou saillance cognitive est résumée par Landragin (2015) comme appartenant au domaine de l'immatériel parce qu'elle concerne des aspects relevant d'inférences basées sur l'interprétation que fait le sujet à partir du message, qui lui est personnelle et qui dépend de facteurs dont « l'attention, l'intention, la mémoire, les affects, et d'une manière générale tout ce qui est implicite mais fait partie du message communiqué et permet de l'expliquer en contexte ». La saillance cognitive intègre en outre « les présuppositions que certains sujets font et d'autres pas, ou encore les sous-entendus et tous ces messages cachés que seuls certains sujet reçoivent ». Cette saillance de caractère subjectif est par nature réfractaire à la formalisation mais aussi par sa complexité. En effet, les critères physiologiques et psychologiques de saillance sont dépendants de nos mécanismes de perception, d'attention et de mémoire (Landragin 2011). En conséquence, ils sont individuels dans leur essence. Néanmoins, Landragin tente d'en catégoriser les critères, parmi eux :

- Critère physiologique de saillance par la proximité spatiale : une entité proche et nette sera saillante, ce qui se traduit en linguistique par la proximité pour l'écrit et la récence pour l'oral ;

- Critère psychologique de saillance liée à l'intention : influence de la tâche en cours sur la perception ;
- Critère sémantico-pragmatique de saillance selon la dichotomie connu *versus* nouveau (voire thème *versus* rhème) : relevant ici de l'intention et non plus des niveaux prosodique, morphologique, lexical ou syntaxique ;
- Critère psychologique de saillance liée à l'attention visuelle ;
- Critère psychologique de saillance liée à la mémoire sensitive et à la mémoire à court terme.

La S-saillance ou saillance située est expliquée dans Landragin (2015) et est en quelque sorte le prolongement de la saillance cognitive. Elle signifie « saillance contextuelle au sens large » et englobe l'énoncé linguistique imprégné de son contexte immédiat et des connaissances partagées par les interlocuteurs (du champ pragmatique). Landragin y ajoute « une mise en perspective de cet énoncé compte tenu des relations sociales, socioprofessionnelles, psychologiques et relationnelles des interlocuteurs ». Il précise que « tout acte de communication peut être doté d'une S-saillance de par ses conséquences sur les relations qu'entretiennent les interlocuteurs » et considère que même un énoncé décrété « neutre » du point de vue linguistique peut être doté d'une forte saillance eu égard à l'impact des aspects relationnels sur le discours.

Landragin (2015) conclut en signalant que lors d'une modélisation computationnelle, la fiabilité des données ne peut relever que de la P-saillance car les données des C-saillance et S-saillance proviennent d'hypothèses et d'inférences déduites des traces matérielles et d'éventuelles prospections métalinguistiques.

Landragin (2012) établit en outre une dichotomie entre d'une part la saillance intrinsèque des entités du discours, et d'autre part la saillance discursive provenant d'une mise en relief selon les choix du locuteur lors de la construction du discours.

Landragin (2011) précise cependant que le degré de saillance est constitué par la conjonction de plusieurs facteurs en faveur d'une même entité. Chaque entité fait partie du système et s'inscrit donc naturellement sur une échelle de valeur graduée selon le degré de saillance. Ce principe permet d'opérer une hiérarchisation des entités en fonction de leur saillance. Pour atteindre cet objectif, des critères de saillance sont requis. Aussi, le choix des termes et des constructions singularisantes peuvent être pris en compte, ainsi que le choix des

mécanismes de mise en relief (comme l'accent tonique) parmi d'autres types. Car « la saillance intervient à tous les stades du processus de construction ».

2.1.2.3 *Rôle de la saillance dans le domaine linguistique*

Landragin (2012) l'affirme : « Si la saillance est invoquée, c'est que les notions linguistiques établies ne suffisent plus à expliquer ces phénomènes communicatifs ». Au-delà des classements traditionnels, un besoin se fait sentir pour prendre en compte de nouvelles entités et en faire surgir le sens. De nouveaux principes de catégorisation sont alors nécessaires pour atteindre l'objectif.

Pour Inkova (2012 : 9), le devenir de la notion de saillance réside dans son rôle en tant qu'outil d'analyse, que l'on ne peut opérationnaliser tant qu'elle n'est pas définie en tenant compte des domaines très divers de son application et des relations qu'elle entretient avec d'autres termes proches comme celui de proéminence (*prominence*). Le « caractère flou » (2012 : 10) de la saillance entrave la découverte de critères objectifs propres, qui permettraient de comparer de manière performante des entités linguistiques selon ce nouveau paramètre. De même, Boisseau et Hamm notent (2015) que le phénomène de saillance est un phénomène expressif qui intervient dans toutes les dimensions du langage, et sont en quête d'une objectivation de la notion de saillance qui justifierait une reconnaissance du caractère opératoire de la notion.

Landragin s'interroge (2012) sur la manière d'intégrer la saillance aux analyses linguistiques ainsi qu'aux descriptions formelles. Son souci s'oriente vers une utilisation des phénomènes de saillance en linguistique de corpus et en linguistique automatique. Landragin entrevoit (2015) des perspectives de ce travail basées sur des expérimentations. La vérification de l'influence d'un facteur se ferait par expérimentation psycholinguistique opposant matérialisation et inhibition du facteur étudié. En outre, la vérification de l'importance de sa dimension se ferait par analyse mettant en rapport saillance, structure informationnelle, chaînes de coréférences et transitions référentielles ; c'est-à-dire par l'exploitation des méthodes de la linguistique de corpus : annotation de l'ensemble des facteurs susceptibles d'intervenir, calculs de fréquences, recherches de corrélations, afin de circonscrire l'intervention de la saillance. Landragin reste réservé sur la mise au point d'un système de calcul automatique de scores de saillance, qui reste un objectif à long terme dont la finalité est « d'améliorer significativement les systèmes de compréhension automatique et de traitement automatique de la langue ». Dès 2011, il prônait les modèles formels et même

computationnels capables de classer les entités de discours à l'aide de scores de saillance. Le principe de hiérarchisation, qu'il définit comme la mise en place d'une relation d'ordre selon un critère de saillance, est une première étape vers leur réalisation.

2.1.2.4 *Implications de la saillance pour cette étude*

La présentation du phénomène de saillance a révélé l'impact que ses multiples facettes pouvaient avoir sur la perception de l'emphase dans la lecture du *Selfish Giant*. La rubrique suivante les détaille. Le phénomène de saillance trouve néanmoins de nombreux échos dans les enjeux fondamentaux des autres expérimentations effectuées jusqu'à présent en relation avec les anglophones natifs. Les parties qui suivent précisent leur apport.

2.1.2.4.1 Mise en lumière des enjeux de la perception par l'emphase (PER-EMPH-N)

L'emphase prosodique se présente comme un des processus de mise en saillance et son résultat perceptif. Dans le processus de profilage, elle se situe à l'interface entre le principe général de saillance à mettre en application et le signal sonore à traiter. La saillance est la notion qui permet de mieux envisager une définition et une méthodologie de repérage de l'emphase ainsi que son rôle par rapport au signal sonore pour en extraire les corrélats acoustiques, ainsi que le schématise la figure suivante.

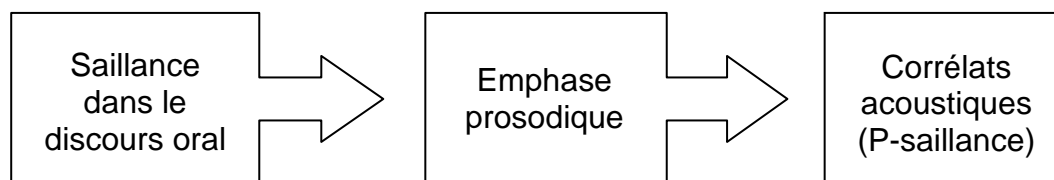


Figure 30 – Schématisation du rôle de l'emphase dans l'expérimentation prosodique PER-EMPH-N

Dans cette expérimentation, le phénomène de saillance est destiné à opérer une dichotomie par exclusion des entités non saillantes par rapport à celles qui le sont. La saillance étant un concept vaste et abstrait, il est nécessaire de lui trouver une interface qui rendra le profilage opérationnel. Le phénomène d'emphase représentant un aspect de la saillance, appliqué à la perception auditive, il peut opérer une dichotomie selon ses propres valeurs. Les entités sélectionnées sont ensuite étudiées acoustiquement. Le repérage de l'emphase utilisera la C-saillance car la saillance cognitive se base sur la perception en fonction de sa compréhension implicite du signal sonore et de sa sensibilité subjective personnelle.

2.1.2.4.2 Mise en lumière des enjeux de la production en lecture (PRO-LEC-N)

Au niveau du texte écrit tel qu’il se présente au lecteur, on remarque notamment les procédés visuels de saillance. C’est le cas du discours direct avec encadrement spatial par guillemets renforcé par le discours citant spécifiant la manière de lire. De même, le texte dans le texte représenté par *TRESPASSERS WILL BE PROSECUTED* est signalé en caractères majuscules doublés d’italiques et est centré sur la page. La centralisation effectuée par la mise en abyme du texte écrit et l’encadrement de la parole constitue des procédés typographiques (Landragin 2012) de P-saillance.

Landragin (2006) écrit: « La production de langage fait elle aussi intervenir la saillance, toute sélection pouvant reposer sur un critère de saillance afin de hiérarchiser les solutions possibles ». En effet, l’activité de lecture du texte elle-même nécessite la mise en œuvre de la C-saillance par la connaissance de la langue écrite qu’elle implique, du sens informationnel à comprendre et à véhiculer. Elle fait aussi intervenir la S-saillance dans sa dimension interlocutive puisque le lecteur, en donnant vie au texte dans une logique de partage d’informations et d’émotions lorsque le texte est correctement lu, effectue des choix interlocutifs, plus ou moins conscients, par profilage hiérarchisant ses priorités en rapport à son objectif. La saillance située établit une relation interlocutive réussie (Austin 1962) ou non. Bien sûr, lire un texte à voix haute est une activité physiologique dont les traces physiques sont les corrélats acoustiques faisant aussi intervenir la P-saillance.

L’activité de lecture à haute voix opérationnalise les trois typologies de saillance et en cela est une activité révélatrice de leur maîtrise.

2.1.2.4.3 Mise en lumière des enjeux de la perception de qualité de lecture (PER-LEC-N)

La tâche d’écoute impliquant une perception du signal sonore dans tous ses aspects, elle fait donc intervenir la P-saillance. Le savoir lire repose sur deux compétences principales : celle de savoir transmettre une information et celle de pouvoir émouvoir l’auditeur. Il met en jeu la C-saillance et la S-saillance. L’expérimentation PER-LEC-N a montré que la détermination de la qualité de lecture était effectuée à partir de la saillance cognitive (être capable de comprendre le message) et la saillance située (par le degré d’émotion transmis dans une relation interlocutive réussie). Le questionnaire s’est efforcé de distinguer ces deux aspects.

L’expérimentation perceptive PER-LEC-N repose elle aussi sur ces trois sortes de saillance.

2.1.2.4.4 Mise en lumière des enjeux de l'exploration de l'expérience personnelle (EXP-LEC-N)

L'expérimentation EXP-LEC-N est introspective, même si elle a fait intervenir la P-saillance dans le passé, cette dernière n'est pas mise en œuvre ici. La réflexion et la mise en perspective de son expérience met en jeu la C-saillance, alors que l'expérience elle-même implique la S-saillance : les résultats de l'expérimentation montrent à quel point certains lecteurs sont éprouvés par la dimension située de s'exprimer en public. Ce « traumatisme » les poursuit parfois jusqu'à influencer de manière importante sur leur manière de s'exprimer en lecture.

2.1.2.4.5 Mise en lumière des enjeux de l'analyse acoustique (COR-AC-N1)

Corollaire des précédentes expérimentations, l'analyse acoustique corrobore par P-saillance les données subjectives de PER-LEC-N et EXP-LEC-N. L'analyse a détaillé ses résultats en amont.

2.1.2.4.6 Mise en lumière des enjeux de l'analyse statistique (STAT-N1)

L'analyse statistique utilise des données objectives et mesurables ou graduables. Elles ont issues des corrélats acoustiques du signal sonore ou d'une objectivation des données perceptuelles, ce qui correspond à la P-saillance. Comme le souligne Landragin (2011), la saillance est « liée à l'isolement ou à l'appartenance à un groupe perceptif ». Dans le traitement statistique, cela est flagrant dans les plans factoriels qui désignent les profils : le Lecteur 4 est ainsi désigné comme un profil à part entière. Il figure comme « singleton », ce qui marque une « rupture » par rapport aux tendances exprimées par les autres dictionnaires. C'est en rapprochant cette saillance (P-saillance) des données acoustiques de COR-AC-N1 (P-saillance) et de la stigmatisation en tant que « mauvais » lecteur de PER-LEC-N (C-saillance et S-saillance) que l'on peut saisir ce profil de lecteur natif par l'incohérence de sa diction. Si L6 et L10 sont stigmatisés « mauvais » lecteurs en raison de leur S-saillance parce que leur diction est inadaptée à une situation interlocutive située, il est probable que L4 le soit aussi pour les mêmes raisons auxquelles s'ajoute la C-saillance car il semble que L4 n'ait pas su ou voulu mettre ses connaissances de lecteur en pratique dans cet exercice.

2.1.2.4.7 Bilan et discussion

On a vu jusqu'ici que la P-saillance de par son caractère objectif n'est pas normative, même si, à partir des mesures sur lesquelles elle se base, il est possible d'en établir une. La C-saillance considère sa propre expérience cognitive : sa subjectivité ôte toute valeur à une

CHAPITRE III – Emphase : mode d’organisation paradigmatique des réalisations prosodiques des anglophones natifs

norme qui ne pourrait avoir de valeur que personnelle. La S-saillance, par contre, implique un jugement de valeur, donc hiérarchisant, et ne peut se situer que par rapport à une norme sociale établie. En conséquence, il semble qu’une évaluation prosodique des apprenants ne puisse être déconnectée des trois catégories de saillance dans une évaluation de type actionnel.

Ce Chapitre III est dédié à l’emphase, qui représente le truchement entre la saillance et l’objectif de sélection des entités les plus riches sur le plan mélodique. Pourtant, en tant que phénomènes, le profilage et la saillance sont omniprésents. Le graphique suivant schématise leur implication dans chacune des expérimentations effectuées ou projetée avec natifs anglophones jusqu’à présent.

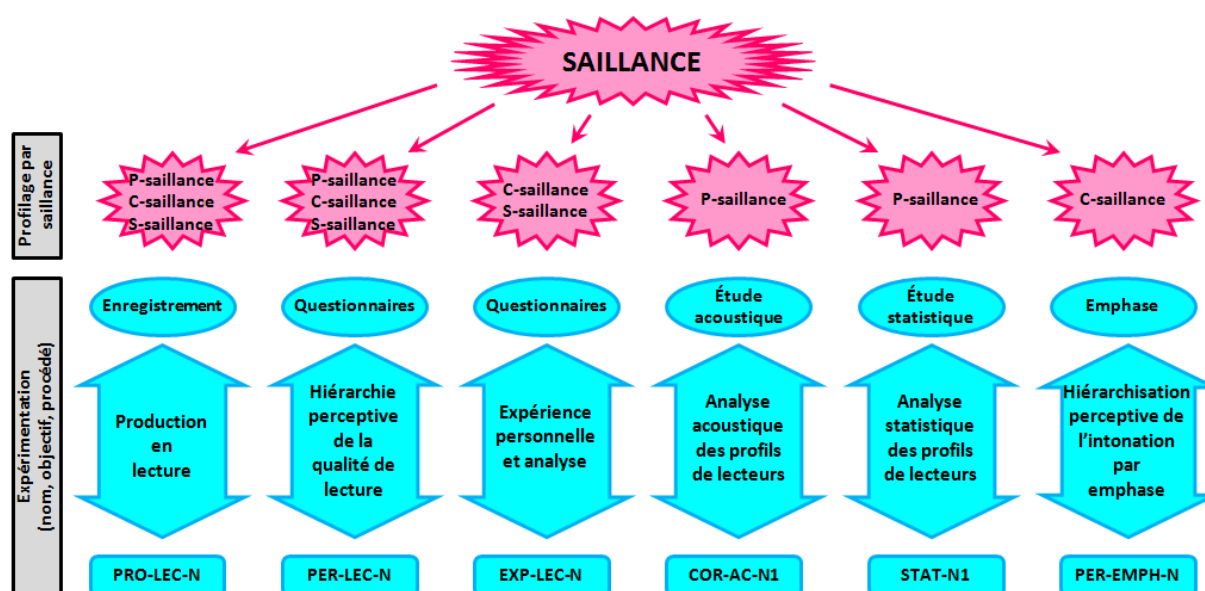


Figure 31 – Profilage par la saillance dans les expérimentations sur anglophones natifs

La saillance est ainsi une notion qui permet de mieux comprendre la congruence des diverses expérimentations effectuées et de mieux envisager les implications d’une méthodologie de repérage de l’emphase. Si les divers types de saillance sont liés parce que complémentaires, seule la P-saillance prend en considération des données objectives et d’une granularité très fine. N’étant qu’un concept générique, c’est le phénomène d’emphase qui va objectiver la notion pour en extraire un profilage adapté à l’obtention de données acoustiques adéquates.

2.1.3 Définition de l'emphase par les phonéticiens

Dans sa thèse, Herment récapitule (2001 : 18-40) en détail l'historique de la notion d'emphase par les linguistes apparentés à la phonétique. En conséquence de quoi, je ne m'attarderai pas sur le sujet. L'essentiel des notions évoquées dans l'exposé sont celles qui ont permis aux linguistes d'analyser la saillance, phénomène dont j'ai rendu compte en amont. Herment constate (2001 : 39-40) au terme de son historique l'absence de consensus sur les termes « topique », « focus », « contraste », « accent », « stress » et « proéminence », qui sont utilisés par les phonéticiens pour évoquer l'emphase. En conséquence, elle opte pour une étude se détachant des traditions d'école et des débats sur le bien-fondé des définitions ou des utilisations du concept. Son optique est de revenir aux sources expérimentales et acoustiques afin d'en extraire une vision objective et scientifique dont elle présente le résultat (2001 : 280). Elle mentionne que le rôle fondamental de la prosodie dans la perception de l'emphase ne doit pas occulter celui du contexte, ni du sémantisme. Le lexique et la syntaxe occupent aussi une place essentielle. Elle signale l'importance des « combinaisons » de facteurs pouvant intervenir dans la perception de l'emphase pour en renforcer l'effet, ou même la substitution d'un facteur par un autre, agissant comme une prise de « relais ». Ces constatations englobent tous les aspects en jeu : linguistiques, sémantiques, pragmatiques, et élargissent le phénomène à d'autres notions nécessaires, comme l'a fait la saillance.

Herment donne aussi une définition plus courte (2001 : 281) :

l'emphase est une mise en relief inattendue qui marque une implication de la part du locuteur plus importante qu'à la normale. Cette mise en relief est obtenue en général grâce à une augmentation de l'énergie dépensée pour prononcer l'occurrence emphatique, et/ou une augmentation de la hauteur et/ou du mouvement mélodique sur cette occurrence.

Cette définition en termes de critères acoustiques et prosodiques uniquement se veut brève et opérationnelle. Elle reprend l'expression « mise en relief » qui fait consensus. C'est cette définition qu'elle réitère (Herment et Hirst 2011) en soulignant l'importance des stratégies et habitudes personnelles dans la perception de ce phénomène complexe.

Je souhaite néanmoins mentionner une référence à Passy (1890)⁵⁷ dans l'ouvrage *Le langage* (Pottier 1973 : 114) et surtout de Peeters (1992 : 111). Mettant en perspective l'économie linguistique de Martinet, Peeters rappelle le « principe d'emphase » de Passy

⁵⁷ PASSY, Paul (1890). *Étude sur les changements phonétiques et leur caractères généraux*. Paris : Didot.

(1890 : §570) selon lequel « le langage tend constamment à mettre en relief ce qui est nécessaire » entrant en rivalité avec le « principe d'économie » en vertu duquel « le langage tend constamment à se débarrasser de ce qui est superflu » (Passy 1890 : §564). La dualité de ces deux tendances qui s'affrontent constamment n'a d'autre but que de parfaire l'acte communicatif au moindre coût afin de satisfaire les besoins. Peeters (1992 : 112) cite Bourciez abordant la double tendance du langage : « [...] *mettre en relief* ce qui est nécessaire [...] repose sur le besoin qu'il a d'être expressif et facilement compris en parlant, puisque le langage est un fait éminemment social ». Bien que le « principe d'économie linguistique » soit appliqué à l'évolution des langues, Martinez lui octroie une dimension panchronique, c'est-à-dire utilisable à la fois en synchronie et en diachronie. Si l'équilibre est parfois rompu entre ces deux tendances en diachronie : magnifier certaines entités et réduire les autres à leur simple expression, il en est de même en synchronie : l'équilibre reste toujours instable par rapport à son objectif interlocutif. C'est cette dualité qui fait de l'emphase un phénomène paradigmatique.

Wells précise (2006 : 11) que parmi les fonctions de l'intonation, une des principales (fonction accentuelle, informationnelle ou de focalisation) est de mettre au premier plan l'information nouvelle par rapport à celle déjà connue, qui reste au second plan. Il donne au verbe *emphasize* l'équivalent *highlight*, pour expliquer l'emploi d'une technique utilisée par le locuteur pour placer certains mots du discours au premier plan pour l'auditeur ; il s'agit là d'une notion linguistique. Il compare par ailleurs (2006 : 212) l'emphase au terme *weightiness*, pour décrire le résultat de la modification intonative consistant à utiliser des têtes intonatives complexes. Il précise (2006 : 215) que le fait d'accentuer la syllabe tonique d'un mot octroie à ce dernier une qualité emphatique, alors que le fait d'utiliser une pré-tête haute (*high pre-head*) dans un syntagme intonatif rend emphatique l'intégralité du syntagme, une seconde possibilité étant d'accentuer plus d'une syllabe dans la tête (*head*). L'emphase peut donc porter sur une entité réduite (mot) ou bien plus vaste (syntagme intonatif) sans intervention sur la nature du noyau. Wells met ainsi en exergue la portée du phénomène emphatique.

2.1.4 Ma définition de l'emphase

On a vu que les termes « emphase » et « saillance » étaient proches l'un de l'autre au point de pouvoir parfois être intervertis. Aussi, je tenterai de définir l'emphase prosodique,

qui est le sujet de mon propos, en m'appuyant sur mon interprétation des recherches précédentes modulées par les résultats des expérimentations que j'ai menées sur le sujet.

L'emphase prosodique se présente comme un des processus spécifiques de mise en saillance et son résultat est prioritairement perceptif. Dans le processus de profilage, elle se situe à l'interface entre le principe générique de saillance à mettre en application et le signal sonore à traiter. Je ferai une distinction de principe entre, d'une part, les termes « accentuation », *stress* et « proéminence », et d'autre part « emphase ». Je considère que la raison d'être des trois premières notions, qui sont d'ordre purement physique, est de communiquer le message informatif effectivement et clairement. L'emphase, quant à elle, se distingue du domaine de l'émotion. Alors que l'émotion se caractérise par l'immédiateté du ressenti par le sujet, et donc par son authenticité, l'emphase est plus « artificielle » car elle implique un décalage par rapport à l'émotion, une prise de recul, un savoir faire pour faire valoir un parti pris interlocutif, une sorte de mise en scène que le locuteur destine à susciter une émotion chez l'auditeur. Ce procédé se rapproche en quelque sorte d'une tentative de « manipulation » de l'auditeur par le locuteur afin de créer une émotion chez lui ; il s'agit d'une tentative de « prise de pouvoir » du locuteur sur l'auditeur, ce qui en fait en outre un exercice de séduction. L'emphase implique alors la maîtrise par le locuteur de sa propre émotion afin de pouvoir agir efficacement sur l'imaginaire et l'émotion de l'auditeur par le truchement de tout moyen adapté. En cela, l'emphase est une émotion moins authentique, sans aller jusqu'à dire qu'elle est contrefaite. Il n'est pas surprenant que cet objectif visé par l'utilisation de l'emphase implique aussi l'utilisation raisonnée et stratégique de tout autre moyen lui permettant d'atteindre son but : un faisceau d'éléments pragmatiques (pour canaliser l'attention et être efficace), sémantiques (par le contraste et le topique), syntaxiques (par les diverses transformations syntaxiques), prosodiques (pour une efficacité d'autant plus puissante qu'elle est primaire par son effet immédiat), et gestuels (en soulignant la saillance prosodique) convergent n'en sera que plus efficace. Pour réussir, la tentative devra correspondre à un degré d'adéquation interlocutive suffisante pour infléchir la volonté de l'auditeur, qu'il s'agisse de séduire ou convaincre. L'emphase devra ainsi surprendre l'auditeur pour le déstabiliser et le faire basculer dans le sens voulu par le locuteur.

Considérer que l'on ne peut transmettre une émotion à quelqu'un si l'on est incapable de la ressentir soi-même, conduit à constater une dualité dans la capacité à influencer sur l'imaginaire du locuteur en suscitant une charge émotionnelle, tout en restant maître de son potentiel à exercer une telle influence sur autrui en calibrant soigneusement les variables pour

qu'en synergie, leur dosage obtienne leur efficacité maximale. L'emphase, semblablement à la saillance, est duelle : elle ne peut exister ni même fonctionner correctement sans sa contrepartie l'absence d'emphase, le calme plat, le ton neutre sans effort ni excès. C'est cette savante alchimie qui peut envoûter l'auditeur. Aussi, les moyens mis en œuvre par chaque locuteur sont éminemment personnels et « situés ».

Reconnaissant que les notions d'« accentuation », de *stress* et de « proéminence » se réalisent souvent acoustiquement de la même manière que l'emphase, je complète ma présentation en précisant la propriété syntagmatique des trois premières notions (leur caractère catégoriel leur permet de poser des repères syntagmatiques pour scander le rythme de la parole) alors que l'emphase se situe à un niveau paradigmatique : il s'agit d'un phénomène systémique continu allant de son absence à des extrêmes. Sur le plan acoustique, j'émet l'hypothèse que les zones les plus emphatiques devraient fournir des « super données », c'est-à-dire des données plus riches et diversifiées, plus significatives et plus révélatrices que celles des autres entités. En conséquence, elles doivent être étudiées prioritairement.

Afin de clarifier le concept d'emphase, j'ai procédé à l'examen des définitions dictionnaires avant de sonder la proximité de la notion de saillance et ses aspects définitoires. J'ai poursuivi mon investigation en mentionnant les recherches liées à la phonétique pour aboutir à l'exposé de mes vues sur la nature du phénomène d'emphase. Gardant à l'esprit que mon objectif est l'organisation d'un profilage grâce à son utilisation, je vais à présent m'orienter vers les choix méthodologiques de tests perceptifs de l'emphase.

2.2 Méthodologie d'investigation perceptive et acoustique de l'emphase

2.2.1 Profilage de la saillance : Landragin (2006 et 2012)

Sensible à l'utilisation de la notion de saillance en linguistique computationnelle, Landragin récapitule (2006) les méthodes de quantification utilisées pour repérer les entités saillantes d'une phrase, d'un énoncé ou d'un discours. Les types d'analyses linguistiques de la saillance reposent sur la détermination d'un ensemble de facteurs de saillance en attribuant un score selon le nombre et l'importance des facteurs convergeant vers le phénomène de saillance. Les données numériques peuvent être comparées objectivement pour déterminer le degré de saillance et hiérarchiser les entités du degré 0 au degré le plus élevé. Landragin énumère cinq méthodes principales de calcul :

- La somme ou la moyenne des facteurs,
- La prise en compte du facteur optimal,
- La moyenne pondérée des facteurs,
- Les méthodes statistiques,
- Les méthodes procédurales.

Elles considèrent toutes des facteurs d'une diversité de domaines catégorisés préalablement, à l'exception de la première. Le calcul de la somme ou la moyenne des facteurs est selon Landragin la méthode la plus simple « car elle consiste à compter les facteurs jouant en faveur de l'entité considérée, en divisant éventuellement ensuite par le nombre de facteurs ». Il regrette néanmoins que tous les facteurs soient placés sur le même plan et ne puissent être distingués selon l'importance de leur influence.

Landragin complète (2012) les méthodes de repérage de l'emphase en prônant les études de corpus dédiées. Selon lui, il existe deux possibilités. La première se base sur la subjectivité des auditeurs auxquels on demande de « repérer les expressions (éventuellement juste les têtes) qui réfèrent à quelque chose de saillant ». Il en est de même pour la détection du degré de saillance par « couples attribut-valeur ». Landragin propose deux traits : un pour le type d'échelle et un pour le degré de saillance. La graduation subjective s'opère selon « l'impression de « *popup* » ressentie, avec par exemple les valeurs « forte », « faible », « nulle » voire « anti » pour les cas d'antisailance ». Landragin en regrette le manque de rigueur mais reconnaît les avantages de son annotation, notamment dans sa mise en œuvre rapide et sa faible implication théorique et sa compatibilité avec l'annotation d'autres facteurs de saillance. La seconde procédure de repérage de l'emphase est rigoureuse et inverse à la première. Il s'agit de préalablement annoter les facteurs de saillance dans chaque domaine (syntaxique, lexical, prosodique, etc.) puis de concevoir un système de calcul automatique pour prendre en compte les annotations effectuées et fournir un tableau des résultats de saillance. Cette procédure objective, qui attribuerait automatiquement des degrés de saillance aux entités, est néanmoins dépendante de sa conception.

Landragin insiste sur l'importance des recherches visant à assigner un degré d'emphase automatiquement car « En complément de la pertinence, la saillance est même un critère de choix idéal dès que plusieurs possibilités se présentent et que le système a besoin d'ordonner ces possibilités ».

2.2.2 Profilage perceptif de l'emphase : Herment (2001)

Pour son travail de recherche, Herment a dû se constituer (2001 : 74) une base de donnée des entités les plus emphatiques à partir d'enregistrements. Elle a fait le choix de styles de parole divers en conversation spontanée naturelle avec les enregistrements d'un débat politique, d'une discussion animée entre deux femmes sur un sujet qui leur tient à cœur (BBC World Service), et d'une conversation informelle et consensuelle entre quatre locuteurs (deux hommes et deux femmes de 24 à 48 ans). Les intervenants sont britanniques, leur accent pouvant être plus ou moins standard, avec, dans le cas de la discussion, une pointe d'accent asiatique. Des extraits de cette base de données dite « secondaire » (d'environ 18500 mots et représentant environ 75 minutes de parole) ont constitué une base de données « primaire » correspondant à une sélection des passages les plus intéressants pour l'étude en cours caractérisés notamment par leur animation (ou leur absence) et leur absence de chevauchements.

Herment a établi sa méthodologie de repérage de l'emphase en s'inspirant des études de Ladd & Morton (1997) et Streefkerk, Pols & ten Bosch (1997) et l'a affinée après deux essais. Elle a fait écouter à trois reprises à des locuteurs anglophones des passages sélectionnés en leur demandant d'annoter le ou les mots ou la syllabe qui lui semblent les plus emphatiques. Les instructions sont présentées par écrit ci-dessous (2001 : 86) :

What you have to do is listen to a TV debate and underline all the passages or words which you feel are EMPHATIC. What I mean when I say emphatic is NOT NEUTRAL, with a special INVOLVEMENT on the part of the speaker. For example, ask yourselves if you would utter the same sentence in the same way in a neutral context. What I want you to tell me is what you think, what you feel is being made prominent in some way.

Par la suite, elle demande d'annoter les scripts avec le degré d'emphase (annexes de sa thèse) :

This time I would like you to distinguish at least two different degrees of emphasis [...] :

- clear emphasis (C under passage or word underlined)
- light emphasis (L under passage or word underlined)

You can add a third one (medium emphasis : M) if you feel the need.

So you first hear a short passage once and then I have selected sentences which you will hear twice and which are written in italics on the script. I would like you

to underline on the text in italics the word(s) which you perceive as being emphatic or prominent.

Un support papier contenant la transcription des passages est remise à chaque auditeur écrite sous forme graphique standard dénuée de ponctuation afin de ne pas interférer avec le repérage auditif. Tout commentaire explicatif était le bienvenu. Ainsi, plusieurs auditeurs ont ressenti le besoin d'ajouter un degré d'emphase *M*, alors que d'autres n'en ont mentionné aucun.

Herment a consigné (2001 : 87) les résultats des sondages sur tableur « en numérotant chaque passage, chaque mot à l'intérieur de chaque passage et chaque syllabe à l'intérieur des mots pour lesquels le découpage en syllabes était nécessaire » afin de procéder à une analyse statistique pour déterminer le degré d'emphase. Les degrés d'emphase ont été convertis numériquement : 1 (pour un degré faible), 2 (pour un degré moyen), ou 3 (pour un degré élevé) en face du mot, et 0 quand aucun degré n'est donné (sondage 1 uniquement). Le pourcentage d'emphase de chaque mot a pu être calculé ainsi que la somme des degrés d'emphase. Un graphique avec « nuage de points » montre que leur corrélation est satisfaisante. Herment remarque que la souplesse de ces sondages a permis d'obtenir des données optimales pour l'étude en cours alors que les mêmes tests sur ordinateur auraient alourdi le processus sans rien apporter de plus qualitativement.

2.2.3 Profilage acoustique de l'emphase : Herment (2001)

Le traitement global du signal acoustique d'Herment a consisté à l'enregistrer ou le numériser, à l'étiqueter, à analyser, manipuler, segmenter et étiqueter la parole avec le logiciel *Praat* (Boersma et Weenik, 1996) grâce à la visualisation d'un spectrogramme et de sa correspondance dans un *TexGrid* avec des lignes d'annotations (*tiers*) textuelles, numériques ou par symboles. Il permet aussi d'en recueillir les résultats automatiquement par programmation de scripts. L'analyse statistique a néanmoins été effectuée avec d'autres logiciels. L'étiquetage en mots a été fait manuellement. Pour chacun des mots, les données sur le pourcentage d'emphase et le degré d'emphase ont été recopiées sur une ligne d'annotation et une transcription phonétique en alphabet *SAMPA* a été ajoutée automatiquement avec la compatibilité d'un dictionnaire rendue possible grâce à la programmation par Hirst de scripts *PERL*. Herment précise (2001 : 113) qu'elle utilise le pourcentage d'emphase dans ses analyses et non le pourcentage de degré d'emphase puisque

les deux variables étant bien corrélées, elles sont redondantes ; de plus, Herment ne possède pas toutes les données en pourcentage du degré d’emphase pour tous les fichiers étudiés.

Après avoir étudié en détail les apports des chercheurs antérieurs à sa recherche sur les corrélats acoustiques de l’emphase, Herment-Dujardin (2001) a focalisé son expérimentation sur les quatre paramètres suivants : la fréquence fondamentale, la durée, l’intensité et les pauses afin de « déterminer les paramètres les mieux corrélés à l’emphase » (2001 : 40). Un rappel de son protocole de prise de mesures figure ci-dessous.

2.2.3.1 Mesures de la fréquence fondamentale

Herment précise (2001 : 146) qu’afin d’alléger la tâche, Hirst a écrit un script pour demander à Praat de calculer pour chaque mot, chaque unité tonale et chaque unité intonative « la valeur minimale de F0 et le temps où elle apparaît, la valeur maximale de F0 et le temps à nouveau, la moyenne de F0, l’écart-type, la pente absolue et la pente absolue sans saut d’octave. »

J’ai résumé le bénéfice attendu par Herment (2001 : 146) des mesures de F0 dans son expérimentation, sous forme de tableau.

Mesures acoustiques	Objectif poursuivi
Les valeurs maximale et minimale de F0	savoir si l’on se situe plutôt dans le grave ou dans l’aigu
La pente	savoir si le contour intonatif est plat ou non (le logiciel PRAAT donne les valeurs absolues de la pente, ce qui ne nous permet pas de savoir s’il s’agit d’une montée ou d’une descente)
L’écart-type	renseigne sur l’ampleur du mouvement mélodique
La pente et l’écart-type	savoir à peu près quel est le type de contour : - si la pente est faible et l’écart type élevé, le contour est ondulateur (type « singsong ») ou tout au moins plusieurs mouvements mélodiques qui s’annulent - si la pente est importante et l’écart-type élevé, cela indique que le mouvement mélodique est important - si la pente est faible et l’écart-type faible aussi, il s’agit d’un schéma plat

Tableau 37 – Protocole d’expérimentation d’Herment (2001 : 146) sous forme de tableau

2.2.3.2 Mesures de la durée

Herment (2001 : 150) fait le choix de se placer au niveau de la syllabe pour prédire la durée afin de minimiser les variations inter-locuteurs. Elle met à profit la théorie de Campbell (1993), et utilise ces données pour comparer les durées prédites par sa théorie (il a extrait des

durées moyennes pour chaque phonème de l'anglais, qu'il a vérifiées dans 86% des cas) aux durées réelles provenant du corpus. Elle justifie ainsi ses positions :

La force de la théorie de Campbell est qu'il se place en fait sur deux niveaux grâce à deux concepts-clé : l'élasticité (*elasticity*) et l'ajustement (*accommodation*). Pour résumer les choses simplement, un ajustement a lieu au niveau de la syllabe grâce à l'élasticité de chaque segment. En d'autres termes, c'est au niveau de la durée de la syllabe que les durées des segments sont ajustées. La théorie de Campbell permet de prédire la durée des phonèmes (cf. Campbell, 1992a) à partir d'un niveau suprasegmental. (Herment, 2001 : 150)

Les paramètres sélectionnés pour mesurer la durée sont donc (2001 : 151) :

- la durée réelle du mot en millisecondes,
- la durée prédite du mot en millisecondes,
- l'allongement (ou raccourcissement si l'on a une polarité négative) en millisecondes,
- l'allongement (ou raccourcissement si l'on a une polarité négative) en pourcentage.

2.2.3.3 Mesure de l'intensité

Regrettant de ne pouvoir procéder à une analyse de balance spectrale, Herment (2001 : 157-158) prend en compte six paramètres pour chaque unité tonale et chaque unité intonative : « l'intensité minimale (dB), le temps de l'intensité minimale (s.), l'intensité maximale (dB), le temps de celle-ci (s.), l'intensité moyenne (dB), l'écart-type (dB) ».

2.2.3.4 Mesure des pauses

Elle constate (2001 :183) que, parmi les quatre paramètres étudiés (nombre de pauses, durée et position), « Seules les pauses ne sont pas aussi révélatrices de l'emphase que les trois autres paramètres ». Les stratégies individuelles des locuteurs ne permettent pas de modéliser les occurrences pausales.

2.2.3.5 Conclusion d'Herment sur les paramètres acoustiques de l'emphase

Herment conclut son étude sur les corrélats acoustiques de l'emphase en mentionnant que « L'emphase semble donc être perçue grâce à une combinaison subtile et surtout variable de plusieurs paramètres. » (2001 : 274) et précise que parmi les quatre paramètres analysés, qui sont la fréquence fondamentale, la durée segmentale, l'intensité globale et les pauses, « la

fréquence fondamentale est le critère qui ressort le plus souvent dans la perception de l'emphase » (2001 : 271).

Elle ajoute néanmoins à la suite de l'analyse statistique, que c'est l'énergie dépensée pour prononcer un mot, paramètre calculé d'après les travaux de Beckman (1986) et qu'elle nomme « amplitude totale », qui « est un paramètre fondamental dans la perception de l'emphase en discours spontané (2001 : 280). L'énergie est en fait une combinaison de la durée et de l'intensité. ». Ce paramètre est donc situé au niveau du mot alors que la portée de la fréquence fondamentale est non limitée et peut donc prendre le relais lorsque le calcul de l'amplitude totale devient inadapté :

La fréquence fondamentale est un autre corrélat de l'emphase, qui peut être associé à l'énergie ou la remplacer pour exprimer l'emphase. C'est le cas lorsque l'on se trouve face à des mots courts, monosyllabiques par exemple, dont l'amplitude totale sera nécessairement faible, la durée d'une syllabe, même allongée, ne pouvant pas atteindre celle de plusieurs syllabes. La fréquence fondamentale prend alors le relais. Elle se manifeste par la hauteur mélodique, que l'on augmente en général sur le mot emphatique, ainsi que par un mouvement mélodique souvent ample sur le mot emphatique. (Herment 2001 : 280)

Les mesures pertinentes de fréquence fondamentale portent sur les différences de hauteur mais aussi sur l'ampleur du mouvement mélodique (la pente ou vitesse).

Elle résume alors les résultats de son analyse sous la forme :

Au terme de ce travail, nous pouvons donc affirmer que l'emphase dans le discours spontané anglais se manifeste au niveau acoustique par une énergie plus importante que sur les mots adjacents, et au niveau prosodique par un mouvement mélodique qui part de plus aigu et est plus ample que sur le reste de l'énoncé. Lorsque l'on applique ces critères sur des mots qui ne se prêtent pas *a priori*, au niveau contextuel ou sémantique, à une mise en relief, l'emphase n'en est que plus importante. C'est à ce niveau que le sémantisme et le contexte jouent un rôle crucial.

Enfin, nous avons vu que l'emphase est en réalité bien souvent exprimée grâce à une combinaison de tous ces paramètres. (2001 : 280-281)

Nous prendrons comme point de départ le protocole d'analyse acoustique qu'a appliqué Herment-Dujardin dans sa thèse.

2.3 Mes choix sur les corrélats perceptifs et acoustiques de l'emphase

Le profilage de détection de l'emphase auquel aboutit Herment (2001) me semble adéquat dans sa phase ultime car il est suffisamment souple et précis pour rendre possible l'obtention de degrés d'emphase. L'entité mesurée est le mot, ce qui manque de précision dans les cas de mots polysyllabiques. Néanmoins, étant donné la charge de travail occasionnée par la longueur du texte et le seul objectif de repérer les zones d'emphase dans sa lecture, j'en reprendrai les grandes lignes.

On a vu ensuite que les paramètres acoustiques entrant en ligne de compte dans la détection acoustique de l'emphase sont nombreux et se relayent. Aussi, l'objectif évaluatif rend inadéquat parce que fastidieuse leur prise en compte intégrale. Je reste aussi réservée sur l'utilisation du paramètre « intensité » car les enregistrements n'ont pas été prévus pour en tenir compte : les locuteurs avaient toute latitude pour se mouvoir en parlant et aucune vidéo ne peut préciser si lors de la prononciation de tel ou tel mot, le locuteur s'est éloigné du micro ou a baissé ou levé la tête.

Par ailleurs, Wichmann ajoute (2004 : 1526) la qualité vocalique (voyelles pleines ou réduites) à l'étude prosodique :

The term 'prosody' includes a number of suprasegmental phonetic features including pitch, loudness, voice quality and tempo; while they can be described separately, they usually operate together in a complex way to create what is loosely described as 'tone of voice'.

Herment mentionne (2001 : 50) que, selon Beckman et Edwards (1994), une voyelle pleine participe à la caractérisation d'une proéminence, et Vaissière ajoute elle aussi le timbre aux paramètres à étudier en prosodie dans certaines langues à tons (Vaissière 2006 : 101)⁵⁸.

En matière de durée, le choix d'Herment a été dicté par la nature de son corpus. Il s'agissait de parole spontanée. Mon corpus premier étant de la parole lue, il est possible de comparer les durées inter-locuteurs (natifs avec non natifs et non natifs entre eux). Je me concentrerai ainsi sur les mesures suivantes :

- la durée réelle des syllabes en millisecondes,

⁵⁸ Bien que la qualité vocalique soit un paramètre perceptif important, le dispositif évaluatif modalisateur supra-segmental des non natifs serait réduit par son alourdissement en prenant en considération un élément segmental sensible non seulement au niveau d'acquisition segmentale des apprenants, mais encore à l'accent des anglophones natifs pris comme modèles. Aussi, je ne tiendrai pas compte de ces considérations.

- la durée modèle des syllabes en millisecondes (à partir de mesures de « bons » lecteurs natifs),
- l'allongement (ou raccourcissement) en millisecondes,
- l'allongement (ou raccourcissement) en pourcentage,
- le rapport de longueur entre les syllabes en pourcentage.

Je reprendrai de l'étude d'Herment la mesure de la F0 et des pauses. En outre, mon étude procèdera similairement à celle d'Herment (2001) : la phase de test perceptif précédant celle d'analyse acoustique, et les résultats perceptifs des zones les plus emphatiques seront étudiées acoustiquement. Alors qu'Herment s'était d'emblée orientée vers la parole spontanée, mes résultats en lecture recouvrant un champ d'investigation différent devrait fournir des données légèrement différentes.

2.4 Synthèse

Dans ce chapitre, j'ai jusqu'à présent analysé le phénomène d'emphase dans ses aspects sémantiques, perceptifs et acoustiques afin de comprendre son fonctionnement et ses implications dans ces divers domaines. Les résultats que j'en retire vont être à présent opérationnalisés dans les expérimentations perceptive et acoustique suivantes. Leurs résultats permettront de compléter la notion de profil par sa composante mélodique.

3 Test perceptif de détection et hiérarchisation de l'emphase chez les anglophones natifs (expérimentation PER-EMPH-N)

Ce test perceptif est un passage obligé pour amorcer le profilage de la parole native par le truchement du phénomène de l'emphase. Sa présentation se déroule en six phases. Le rappel des objectifs introduit le profilage des lecteurs par les auditeurs, puis des auditeurs eux-mêmes. Ensuite est détaillée l'expérimentation résultant dans le classement des données perceptives obtenues. Leur bilan marque l'orientation des expérimentations acoustiques suivantes.

3.1 Rappel des objectifs

L'objectif de cette expérimentation est double :

- catégoriser les mots lus en contexte pour savoir s'il sont emphatiques ou pas afin de prédire leur nature attitudinale ou grammaticale par leur variabilité prosodique,
- catégoriser les mots selon leur degré d'emphase afin de les hiérarchiser pour pouvoir retenir les énoncés les plus emphatiques en vue de leur étude intonative et surtout mélodique.

3.2 Profilage des lecteurs

Parmi les douze lecteurs évalués précédemment, trois ont été sélectionnés pour représenter trois styles de lecture différents selon deux critères : la rapidité de lecture et la qualité de « bon » lecteur.

La rapidité de lecture est un paramètre fondamental dans la mesure où il influe sur les autres paramètres prosodiques. La mélodie pourra être ample avec une lecture lente mais plus difficilement avec une lecture rapide. La durée, le nombre des pauses ou même leur existence devrait varier selon la vitesse de lecture. La présence d'emphase devrait dépendre du paramètre de durée de lecture. Aussi, j'ai sélectionné parmi les douze lecteurs le plus rapide et le plus lent en vérifiant que leur C-saillance n'était pas mise en cause par le jugement des auditeurs natifs. J'ai sélectionné la lecture la plus rapide (de L6, que j'ai renommée Lecteur B, durée de lecture du texte de 110 secondes) et la plus lente (de L10, que j'appelle Lecteur C, pour sa lecture de 180 secondes). L6 n'a été pointée qu'une seule fois « mauvais lecteur » par un auditeur anglophone natif : pour son habileté à susciter l'émotion par sa lecture, elle a obtenu le degré le plus faible sur les quatre degrés proposés avec les commentaires *fast rhythm, intonation not varied enough, clicking noises between phrases* ; par contre, pour la clarté de l'information transmise, elle a obtenu 2/4 avec le commentaire *BBC English*. L'étude acoustique a montré la cohérence de sa diction et sa tessiture relativement ample. Il est très probable que sa vitesse de lecture soit principalement en cause dans ce jugement négatif car c'est le trait que les autres auditeurs ont souligné, ainsi que, parfois, son caractère monotone impropre à transmettre l'émotion. À l'autre bout de l'échelle, j'ai choisi L10, très lent, lecteur professionnel, plébiscité par les apprenants de niveaux B1/B2⁵⁹ comme « meilleur lecteur ». Trois auditeurs natifs l'avaient pointé « mauvais lecteur », mais deux l'avaient trouvé « meilleur lecteur ». L'ambivalence des réactions natives semble avoir été occasionnée par d'autres particularités (comme son accent américain nasillard ou sa tessiture

⁵⁹ Niveaux en langue selon le *Cadre Européen Commun de Référence pour les Langues* du Conseil de l'Europe.

d’ampleur exceptionnelle au point que certains auditeurs n’arrivaient pas à cadrer son genre, ou même sa S-saillance) parce que les évaluations obtenues sont plutôt élogieuses malgré l’étiquette de « mauvais lecteur » : un auditeur n’a pas voulu le noter, mais les deux autres lui ont octroyé 4 et 4 pour la clarté de l’information, et 4 et 3 dans la transmission de l’émotion. J’ai aussi sélectionné L1 (129 secondes), que j’ai renommé Lecteur A, le « meilleur » lecteur, parce qu’il avait été élu « meilleur lecteur » par les anglophones et de nombreux enseignants francophones. Son utilisation de l’emphase devrait être particulièrement judicieuse et équilibrée. La figure suivante montre que la durée de lecture de notre « meilleur » lecteur, par rapport aux deux profils extrêmes que constituent le Lecteur B et le Lecteur C, se situe non pas à mi-parcours entre les deux mais au deux tiers de la longueur de l’axe temporel en se rapprochant du lecteur le plus rapide, ce qui implique qu’une bonne lecture doit être rapide.

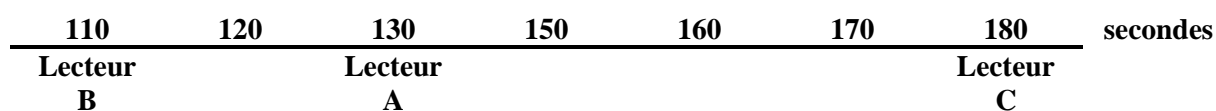


Figure 32 – Situation des trois durées de lecture sur l’axe temporel

Ainsi, l’expérimentation perceptive va pouvoir repérer les zones les plus emphatiques suivant trois styles de lecture différents chez les locuteurs natifs anglophones.

3.3 Profilage des auditeurs

La détection des entités perçues comme emphatiques a été opérée par neuf anglophones natifs, six enseignants d’anglais francophones et 14 étudiants anglicistes francophones de niveau A2 à B2⁶⁰. Ils ont tous été bénévoles pour tenter l’expérience que constituait de test. Un seul auditeur natif l’a abandonné après avoir repéré l’emphase chez le lecteur A. Ce test étant long et rébarbatif, j’ai dû le profiler de manière à ce que toutes les bonnes volontés puissent le faire dans les conditions qui étaient pour les auditeurs les plus favorables et quel que soit leur niveau de maîtrise en anglais.

3.4 Profilage de l’expérimentation

Le profilage de l’expérimentation a suivi les grandes lignes de la méthodologie d’Herment (2001) tout en s’adaptant aux conditions nouvelles (éloignement de certains auditeurs et niveau parfois faible). J’ai créé des pages internet en anglais avec la procédure à suivre décrite avec précision et fourni le lien aux participants.

⁶⁰ Niveaux en langue selon le *Cadre Européen Commun de Référence pour les Langues* du Conseil de l’Europe.

La page d'accueil expliquait les objectifs ainsi que les instructions pour accomplir la tâche. Des liens permettaient de télécharger un fichier texte de réponse (format RTF) différent selon que l'on était étudiant ou non car un très court questionnaire de renseignements sur l'auditeur et le lecteur écouté devait tout d'abord être complété. Les 26 phrases du texte *The Selfish Giant* y étaient inscrites sans ponctuation. Ensuite, en première approche, un lien conduisait l'auditeur vers le Lecteur A, dont il était invité à écouter la lecture en entier éclairée par le script écrit complet. Dans une seconde phase, l'auditeur devait écouter chaque phrase séparément (deux ou trois fois selon les besoins), mettre en surbrillance les mots ou passages emphatiques et écrire le degré d'emphase sous chaque repérage en surbrillance selon les codes fournis au départ en justifiant ses choix par commentaires. La même procédure devait être appliquée aux trois lecteurs. Chaque auditeur m'a envoyé trois fichiers à l'exception d'un auditeur natif. Quelques enseignants et natifs ont préféré utiliser une version papier du script car ils trouvaient que le repérage y était plus facile à inscrire.

Les instructions précisaient la nature de l'aide sollicitée ainsi qu'une définition de l'emphase. J'ai essentiellement repris la définition vague qu'en avait donné Herment (2001) dans ses questionnaires et ai proposé quatre niveaux d'emphase comme l'extrait du site ci-dessous le montre, ce qui totalise cinq niveaux si l'on considère le niveau 0.

What I wish you to do is listen to three texts and highlight all the passages or words which you feel are EMPHASIZED or PROMINENT, that is to say with a special INVOLVEMENT on the part of the speaker (not neutral).

I would like you to distinguish between different degrees of emphasis by writing a code number under the highlighted passages or words. The codes are

- 1 — light emphasis**
- 2 — medium emphasis**
- 3 — heavy emphasis**
- 4 — extreme emphasis**

Figure 33 – Extrait du site expérimental de détection de l'emphase

Voici ci-dessous deux extraits de fiches de réponse envoyées en fichiers joints. On peut y voir qu'il s'agit là de la lecture du Lecteur A (« meilleur » lecteur), qui avait omis deux mots dont *large* dans la Phrase 2.

CHAPITRE III – Emphase : mode d’organisation paradigmatique des réalisations prosodiques des anglophones natifs

Your initials : JNB Age : 62 Gender : M English accent from: Western Canada
Occupation: retired How long have you been living in France ? 0 year(s)

1/6

From **The SELFISH GIANT** by Oscar WILDE
Reader: **A**

1 - Every afternoon, as they were coming from school, the children used to go and **play** in the Giant's garden.

2

Comment: the speaker is showing us the action of the children

2 - It was a **large lovely** garden, with soft green grass.

2

Comment: the speaker omits the word “large” in his reading. “Lovely” paints a picture of the garden

3 - Here and there over the grass stood beautiful flowers like **stars**, and there were twelve **peach-trees** that in the spring-time broke out into delicate blossoms of **pink and pearl**, and in the autumn bore rich fruit.

2

1

1

Figure 34 – Extrait de repérage d’emphase dans un script annoté par un auditeur anglophone

1 - Every afternoon, as they were coming from school, the children used to go and play in the Giant's garden.

2

1

2

2

3

2

2

Comment:

j’ai réservé (4) pour une accentuation abusive

2 - It was a large lovely garden, with soft green grass.

1

3

2

2

2

2

Comment:

je n’ai pas entendu “large”

Figure 35 – Extrait de repérage d’emphase dans un script annoté par un auditeur francophone

On peut voir dans le premier extrait que l’auditeur a mis en surbrillance trois mots d’affilée *pink and pearl*, octroyant de l’emphase à un mot outil. Ainsi, comme l’avait déjà remarqué Herment, ce niveau global de repérage (certains mots entourés de mots emphatiques) doit être pris en compte dans l’analyse.

J’ai ensuite procédé au classement des données, à présent numériques, sur tableur comme l’avait fait Herment (2001) en prenant le mot comme entité d’emphase. Pour faciliter le classement, j’ai considéré qu’un mot se définissait par les deux espaces qui l’entouraient, c’est pourquoi les mots composés graphiquement (liés par un trait d’union) ne comptent que pour un mot et *any one* écrit en deux mots compte pour deux mots. Ce découpage prend en compte la visualisation du script par l’auditeur, lui-même basé sur le texte original d’Oscar Wilde. Les 426 mots ainsi obtenus sont écrits en ordonnée et les initiales des auditeurs en abscisse (la Figure ci-dessous en montre un agrandissement). Avec 88 colonnes contenant

37 488 cellules, chaque ligne correspond à un mot du texte, chaque colonne au repérage par un auditeur de l'emphase d'un enregistrement. Un camaïeu de jaune orangé a permis de percevoir visuellement les zones repérées ainsi que le degré d'emphase attribué. Cette première approche visuelle et subjective avant d'organiser une analyse systématique permet de vérifier, grâce à l'utilisation de formules conditionnelles, les hypothèses qui avaient pu être émises, ainsi que de calculer et hiérarchiser la perception de l'emphase selon le paradigme « natif anglophone », « francophone angliciste de haut niveau de maîtrise » et « francophone angliciste de niveau faible ».

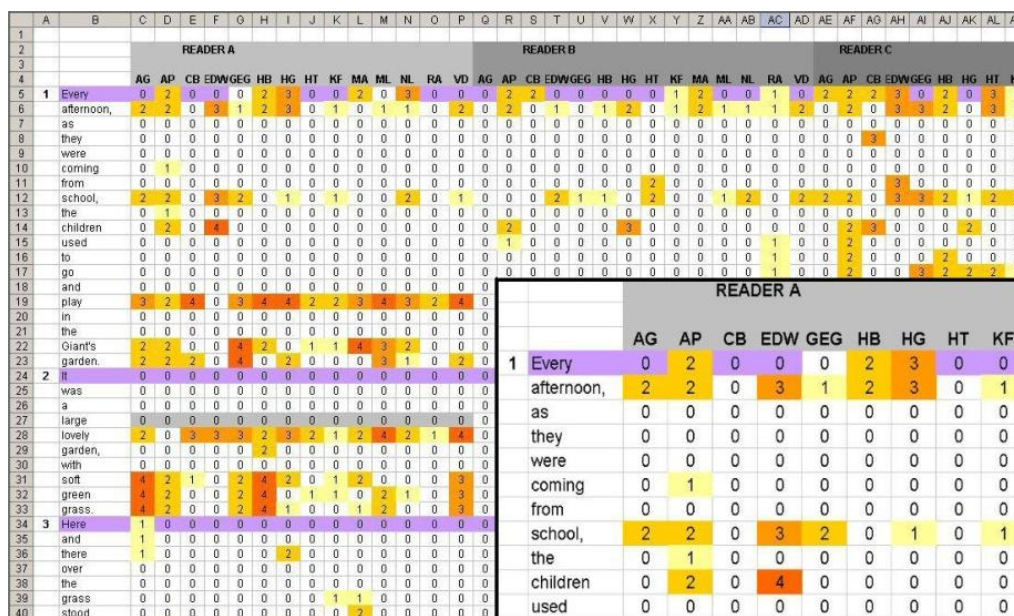


Figure 36 – Extrait de feuille de tableur et vue rapprochée dans le repérage de l'emphase

3.5 Classement et résultat des données perceptives

Après calculs à partir des données, il est possible de vérifier l'hypothèse d'Herment (2001) selon laquelle le pourcentage du pointage d'emphase équivaut à une spécification des degrés d'emphase. Suite à ma propre hypothèse que les entités les plus saillantes, c'est-à-dire les plus emphatiques devraient receler des corrélats intonatifs riches de variation, je vais pouvoir repérer l'emphase maximale de mon corpus afin de sélectionner les entités à étudier.

3.5.1 Vérification de méthodologie d'Herment (2001)

Sur l'ensemble de son corpus, Herment (2001) n'a mesuré que le nombre de fois où était pointée une emphase. Le fait d'avoir pu obtenir pour chacun des auditeurs à la fois le pointage d'emphase et le degré d'emphase permet de vérifier l'hypothèse d'Herment qui avait conclu à leur équivalence puisqu'elle n'avait pas pu obtenir ces données pour l'ensemble de

son corpus étudié. Pour visualiser que le pourcentage du pointage d'emphase équivaut à une spécification des degrés d'emphase, j'utilise une représentation graphique de la corrélation par nuage de points (graphique ci-dessous) et observe une corrélation positive avec peu de points qui s'écartent de la droite.

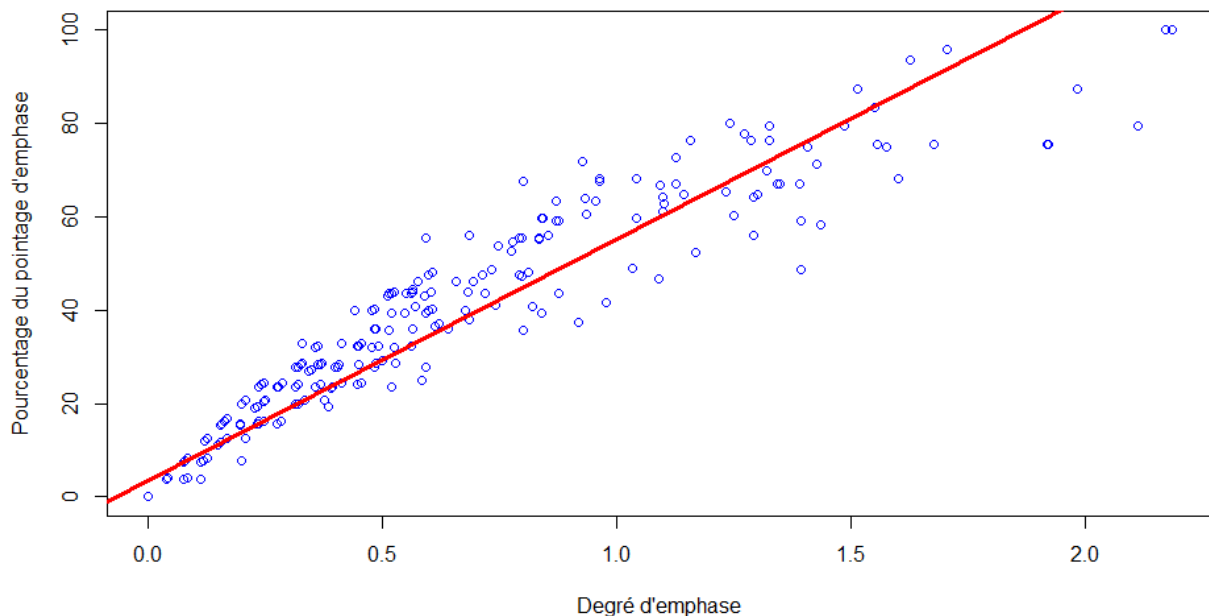


Figure 37 – Représentation graphique de la corrélation entre le pourcentage de pointage d'emphase et le degré moyen d'emphase de chaque mot

La vérification avec la corrélation de Pearson confirme la représentation visuelle. Le coefficient de corrélation est de 0,96 avec un intervalle de confiance entre 0,9552304 et 0,9692016, ce qui indique une très forte corrélation positive entre le pourcentage de pointage d'emphase et le degré d'emphase. R^2 est très important (0,9269), ce qui indique que la variation de pourcentage de pointage d'emphase dépend très étroitement du degré d'emphase. On remarque que la valeur p est particulièrement significative ($< 2e-16$) et l'on peut interpréter qu'à chaque augmentation d'un degré d'emphase, le pourcentage de pointage d'emphase augmente en moyenne de 51,73%. On peut en déduire qu'il existe une très forte corrélation entre le pourcentage de pointage d'emphase et le degré d'emphase, ce qui valide l'hypothèse d'Herment 2001.

3.5.2 Hiérarchisation des mots et zones des plus emphatiques aux moins emphatiques

L'analyse des données permet de déterminer les mots ou zones les plus emphatiques à étudier prioritairement sur le plan acoustique en suivant l'hypothèse que leur degré de variation est maximal dans les zones les plus emphatiques. Elle permet aussi de vérifier l'absence d'emphase sur d'autres mots.

Bien que l'on constate des disparités entre l'emphase perçue par les différents groupes d'auditeurs, ce sont surtout les styles de lecture qui marquent une différence dans le pointage et le degré d'emphase perçus. L'objectif étant d'analyser les réalisations natives, je me focalise ici sur les données fournies par les auditeurs natifs. En calculant la moyenne des résultats pour chaque lecteur, puis en faisant la moyenne pour chaque mot, on obtient le classement suivant. Il doit être précisé que dans le cas des mots non prononcés (*large* et *here*) par le Lecteur A (« meilleur » lecteur), la moyenne a été effectuée à partir des résultats concernant les deux autres lecteurs uniquement, ce qui biaise les résultats. La moyenne des résultats obtenus a permis de classer chaque mot du texte en fonction de sa capacité à focaliser l'emphase dans le contexte du texte *The Selfish Giant*. Arrivent en tête ceux de l'exclamative commençant par *How happy* (11^{ème} et 1^{ère} positions) et l'adverbe intensificateur *so* (2^{ème} et 7^{ème} positions). Ensuite se situent les mots *What* (3^{ème} position), *doing* (5^{ème} position) et *here* (12^{ème} position) de la question (Phrase 10). La quatrième position est occupée par le pronom *nobody*. On remarque que ces mots cités ci-dessus fonctionnent en réseau : lorsque l'un d'entre eux focalise l'emphase, ceux qui lui sont proches syntagmatiquement ont tendance à faire de même, ce qui corrobore l'hypothèse effectuée lors de l'examen visuel des cellules colorées sur tableur.

Rang d'emphase	Mots les plus emphatiques	Degré moyen d'emphase (Lecteurs A+B+C)
1	happy	2,18
2	so (Phrase 22)	2,17
3	"What	2,11
4	nobody	1,98
5	doing	1,92
6	wonderful	1,91
7	so (Phrase 4)	1,70
8	myself."	1,67
8	PROSECUTED	1,67
10	large ⁶¹	1,62
11	"How	1,60
12	here?" ⁶²	1,57
13	stars,	1,55
14	lovely (Phrase 21)	1,55
15	selfish	1,51

Tableau 38 – Classement par les anglophones natifs des quinze premiers mots les plus emphatiques du texte *The Selfish Giant*

⁶¹ Le mot *large* n'est pas prononcé par le Lecteur A. La moyenne n'inclut que les lecteurs B et C.

⁶² Le mot *here* n'est pas prononcé par le Lecteur A. La moyenne n'inclut que les lecteurs B et C.

3.5.3 Sélection d’une phrase « neutre » (non emphatique)

La dichotomie différenciant les phrases se faisant sur la base de l’emphase, il est donc essentiel de poursuivre le raisonnement afin de prendre appui, à titre de comparaison, sur une phrase non emphatique, dite phrase « neutre », qui se situerait à l’autre extrémité du continuum. Mon hypothèse étant que l’emphase concentre la variabilité maximale, une phrase dépourvue d’emphase réduirait au maximum la variabilité prosodique si de surcroît elle était courte. Son nombre de syllabes réduit ferait d’elle une référence stable pour l’analyse de la prosodie des non natifs.

On constate en outre que sur les 426 mots du texte, 148 ne sont dotés d’aucun degré d’emphase, si minime soit-il, alors que tous les autres mots évoluent entre une moyenne de 2,18518 et 0,037037, ce qui les classe selon un continuum et non pas d’emblée en catégories bien distinctes. Les mots non emphatiques sont les « mots outils » dont aucune mise en perspective ne vient modifier la nature. L’étude visuelle sur tableur montre qu’aucune phrase n’est complètement dénuée d’emphase. La saillance est nécessaire, même réduite, pour accrocher l’attention de l’auditeur ou simplement rendre le message compréhensible. La neutralité en termes d’emphase proprement dite n’existe pas.

L’étude du degré emphatique de chaque mot du texte permet de calculer celui de chaque phrase en additionnant le degré de chaque mot et en le divisant par le nombre de mots de la phrase en question. Le graphique suivant en fournit le résultat.

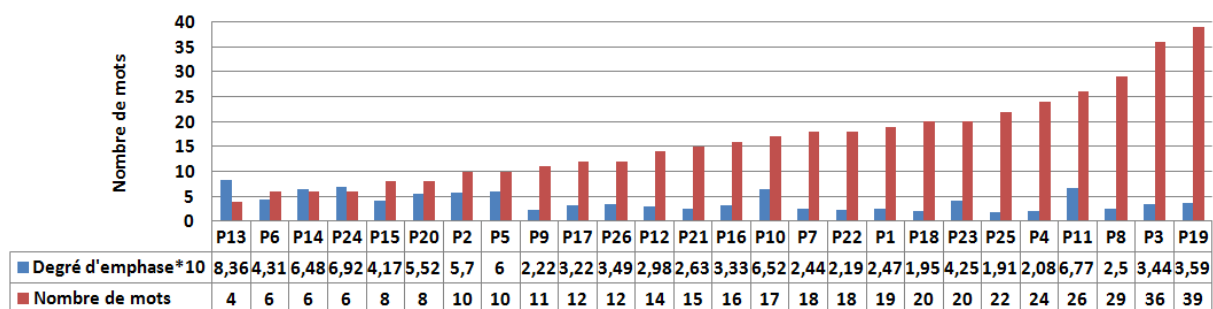


Figure 38 – Visualisation du degré moyen d’emphase (*10) comparé au nombre de mots de chaque phrase

Afin de pouvoir établir une comparaison visuelle, j’ai multiplié ici le degré d’emphase par 10. Le phénomène d’emphase peut donc aussi servir à classer les phrases les unes par rapport aux autres selon leur degré moyen. Si certains mots, le plus souvent de nature grammaticale, sont dotés d’une valeur emphatique nulle, ce n’est pas le cas des phrases, qui regroupent plusieurs mots. On peut constater que la phrase la plus emphatique est aussi la plus courte. Il s’agit de la Phrase 13 (*TRESPASSERS WILL BE PROSECUTED*), dont chacun des quatre mots est

doté d'un haut degré d'emphase, ainsi que le prescrit sa présentation typographique. À l'autre bout de l'échelle, les Phrases 3 et 19 sont les phrases les plus longues du texte (respectivement 36 et 39 mots) et leur degré d'emphase est comparativement très faible. Le degré moyen d'emphase sur l'ensemble des mots du texte est de 0,36. Celui de la Phrase 19 est exactement le même et celui de la Phrase 3 est de 0,34. Plusieurs raisons peuvent être évoquées pour justifier cette tendance :

- Raison stylistique : Le discours direct est un vecteur fondamental des émotions ressenties par les personnages du conte. En conséquence, on trouvera un haut degré d'emphase dans les phrases au discours direct, lequel focalise l'emphase.
- Raison stylistique corollaire : Les phrases longues sont souvent uniquement narratives, alors que le discours direct est davantage véhiculé par des phrases courtes, même si ces dernières sont souvent accompagnées de discours citants.
- Raison phonétique : Le degré d'emphase est plus important avant des frontières. Une phrase perçue sans emphase sera au moins dotée d'une chute finale qui marquera la seule proéminence perceptible. La proportion de frontières dans les phrases courtes est donc, en général, plus importante que dans les phrases longues.

Il faut néanmoins souligner une constante : le phénomène d'emphase est bien présent dans toutes les phrases en proportion et localisations diverses. Si tous les mots d'un énoncé peuvent être dotés d'un degré d'emphase (comme c'est le cas pour la Phrase 13), aucune phrase n'est complètement dénuée d'emphase, aussi minime soit-elle. D'aucuns préféreront alors la rebaptiser « proéminence » ou « accent », pour ne garder au terme d'emphase que ses degrés maximaux. Toujours est-il qu'il sera impossible de trouver une phrase sans emphase aucune, qui permettrait de comparer ses données à celles des phrases porteuses d'emphase.

En outre, plus une phrase sera courte, plus le rapport entre le degré d'emphase et le nombre de mots sera important. Ceci est plus visible dans le graphique suivant, qui établit une comparaison visuelle entre le nombre de mots de chaque phrase et son pourcentage d'emphase par rapport au nombre de ses mots. On peut y voir que la Phrase 9 semble être un bon compromis : alors que le classement des phrases est opéré selon un continuum, du plus petit nombre de mots au plus grand, la proportion d'emphase chute sur la Phrase 9, représentant un décrochage par rapport aux phrases plus courtes, dotées d'une forte valeur emphatique : elles comprennent toutes des focalisateurs d'emphase ou de saillance tels qu'un

CHAPITRE III – Emphase : mode d’organisation paradigmatique des réalisations prosodiques des anglophones natifs

discours direct, une typographie particulière, un adverbe d’intensité, un adjectif à sémantisme fort, des mots à base négative (*nowhere, never, nor*). Seule la Phrase 6 en est exempte.

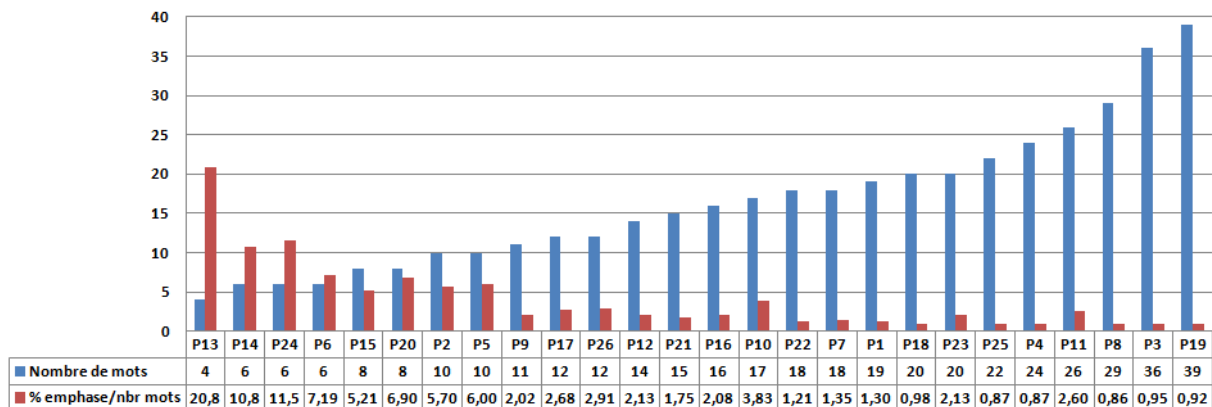


Figure 39 – Rapport du degré d’emphase par nombre de mots dans chacune des 26 phrases du texte

Examinons ces deux phrases de plus près. Les caractéristiques emphatiques spécifiques à chacune des deux phrases sont représentées dans les graphiques suivants et l’on constate que les frontières marquent emphatiquement le mot qui les précède. Les deux phrases sont dénuées de ponctuation.

La Phrase 9 est syntaxiquement complexe⁶³, dépourvue de ponctuation et structurée en trois relations prédicatives, ce qui devrait correspondre à deux syntagmes intonatifs :

When he arrived | he saw the children playing in the garden. ||

La moyenne emphatique de chaque mot est consignée dans le graphique ci-dessous :

⁶³ Khalifa définit (2004 : 43-44) la phrase complexe par deux traits : 1) la présence de plus d’une relation prédicative, 2) la dépendance des propositions les unes par rapport aux autres par imbrication, ou enchâssement, ou emboîtement (*embedding*).

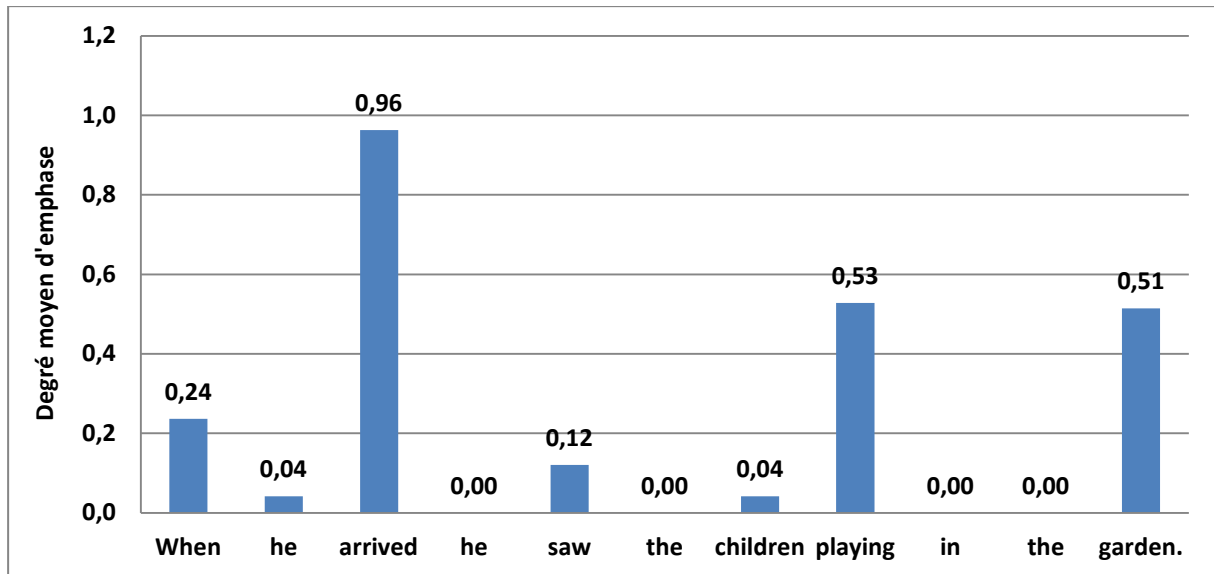


Figure 40 – Degré moyen d'emphase de chaque mot de la Phrase 9

On voit que le degré moyen d'emphase est toujours inférieur à 1. Les emphases sur *arrived* et *garden* correspondent à des tons précédant une frontière, alors que *playing* est un élément post-frontière. Le graphique ci-dessous montre la répartition d'emphase réduite dans la prestation des trois lecteurs par le truchement des degrés d'emphase perçus par les auditeurs natifs. Il permet de constater la prépondérance de l'emphase sur *arrived* et *garden* malgré le degré global : les lecteurs A et B leur accordent plus d'importance, ce qui indique leur stabilité. Le lecteur C octroie à *playing* une importance plus « attitudinale ».

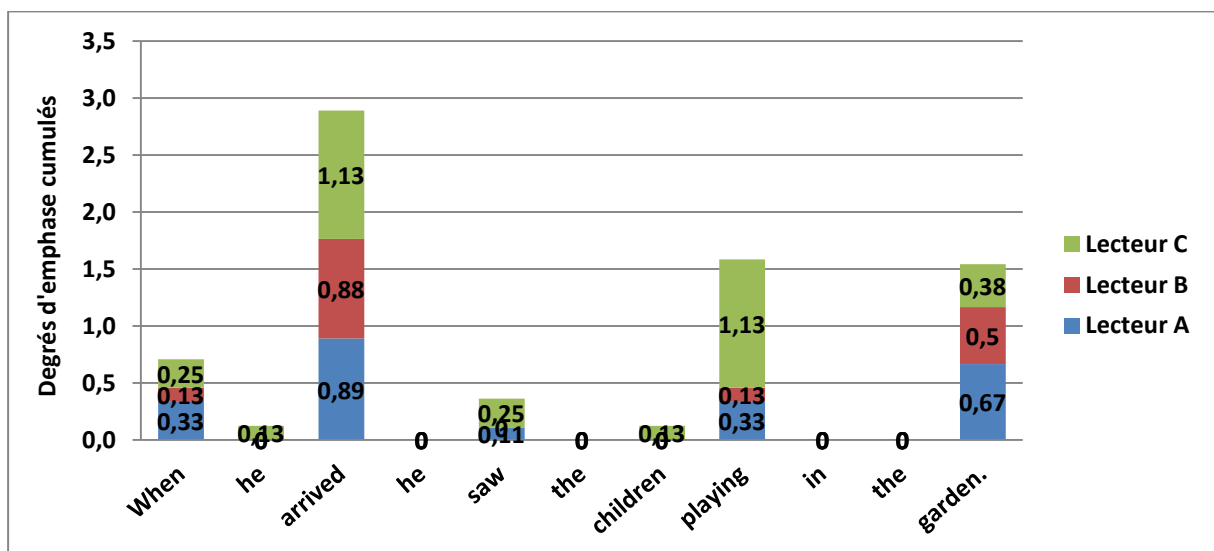


Figure 41 – Degrés d'emphase cumulés dans la diction des trois anglophones pour chaque mot de la Phrase 9

Malgré sa longueur de onze mots, cette phrase semble bien stable et équilibrée pour procéder à une analyse comme repère non attitudinal, surtout en ce qui concerne *arrived* et *garden*.

CHAPITRE III – Emphase : mode d'organisation paradigmatique des réalisations prosodiques des anglophones natifs

La Phrase 6 est plus courte et simple. Elle reprend la structure de la Phrase 9 réduite aux deux premiers syntagmes intonatifs :

One day | the Giant came back. ||

Le mot le plus emphatique de la Phrase 6 est d'un degré d'emphase de 1,10. Une observation plus précise permet de remarquer que quasiment toutes les phrases du texte ont au moins un mot de degré d'emphase analogue ou même supérieur. L'emphase est ici perçue fortement sur le dernier mot de la phrase, donc avant une frontière, et qui se trouve être la particule du *phrasal verb* : *came back*. Par ailleurs, son faible nombre de mots correspond à une phrase de structure syntaxique simple, ce qui est susceptible de réduire encore davantage les dictionnaires possibles. Le graphique suivant en suggère la diversité.

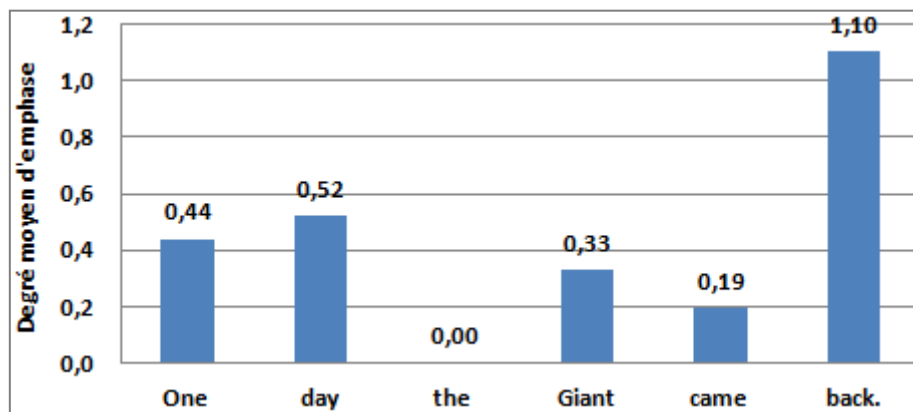


Figure 42 – Degré moyen d'emphase de chaque mot de la Phrase 6

Le détail de l'attribution d'emphase par chacun des trois lecteurs montre une répartition équilibrée.

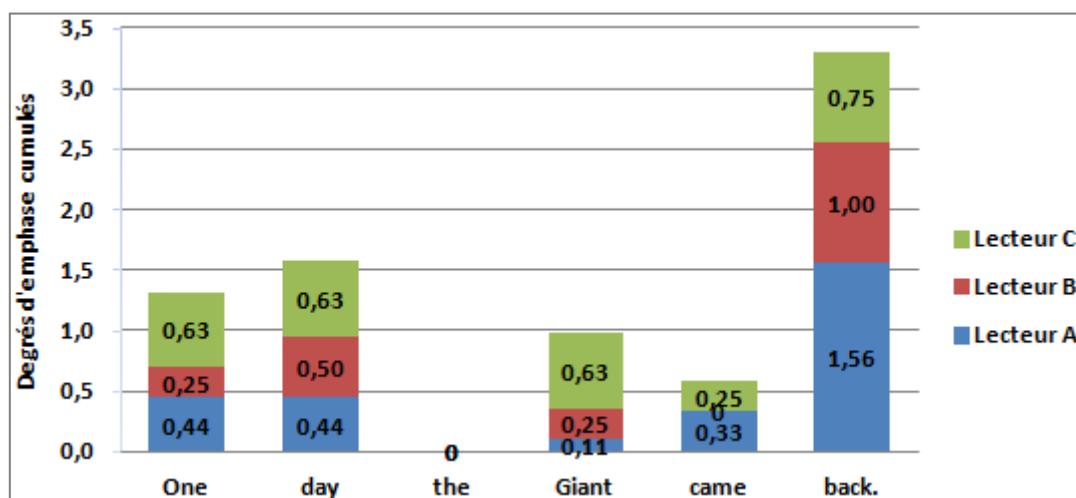


Figure 43 – Degrés d'emphase cumulés dans la diction des trois anglophones pour chaque mot de la Phrase 6

D'après notre analyse précédente, les Phrases 6 et 9 semblent équivalentes pour pouvoir servir de repère normatif dans l'analyse intonative. Voyons à présent le tableau suivant, qui récapitule les points susceptibles de répondre aux critères recherchés.

	Nombre de mots	Moyenne d'emphase	Pourcentage d'emphase par nombre de mots	Degré d'emphase maximal
Phrase 6	6	0,43	7,19	1,10
Phrase 9	11	0,22	2,02	0,96

Tableau 39 – Caractéristiques des deux phrases « neutres » courtes

La Phrase 6 est plus maniable de par son nombre de mots réduit à six alors que la Phrase 9 en totalise 11, son degré d'emphase maximal n'est pas très éloigné de celui de la Phrase 9 et il indique un ton devant frontière, ce qui est normal. Pourtant, les autres critères ne plaident pas en sa faveur car la moyenne d'emphase est le double de celle de la Phrase 9 alors que la valeur attendue devrait en être la moitié, et la proportion de degré d'emphase lui est plus de trois fois supérieure au lieu d'une moitié ou même d'une équivalence. Ainsi, cette disparité d'emphase entre les deux phrases relativement semblables porte à réflexion car il est essentiel que la phrase sélectionnée soit la plus stable possible pour l'analyse à finalité évaluative.

En conclusion, je retiens les deux phrases pour être étudiées acoustiquement. L'analyse ultérieure déterminera laquelle est sélectionnée pour représenter la « phrase neutre » de référence aux données de laquelle les autres éléments seront comparés. Dans la sélection, la méthodologie basée sur la notion d'emphase a joué un rôle fondamental en alertant sur de possibles réalisations attitudinales.

3.6 Bilan de l'étude perceptive de l'emphase (expérimentation PER-EMPH-N)

Cette analyse perceptuelle a permis de corroborer l'hypothèse d'Herment (2001) selon laquelle pourcentage de sélection de locus et degré d'emphase étaient étroitement corrélés.

La hiérarchisation des mots et zones des plus emphatiques aux moins emphatiques a déterminé ceux à étudier acoustiquement en vue de les sélectionner ou de les écarter en vue de l'analyse acoustique des non natifs du prochain chapitre. Tout étant relatif en matière d'intonation, je considérerai étudier les phrases dans lesquelles ces mots sont insérés car il s'agit de discours écrit oralisé. De plus, la lecture à haute voix a tendance à préserver une

correspondance entre la structure syntaxique et la structure phonologique. En conséquence, je retiens comme marqueurs d’emphase à privilégier :

- l’exclamative de la Phase 5 (*How happy we are here!*),
- les trois portions de phrases similaires avec paires « *so* + adjectif ou adverbe » (*trees and sang so sweetly that the children, It sounded so sweet to his ears, why the Spring is so late in coming*),
- l’interrogative de la Phrase 10 (*What are you doing here?*),
- *nobody* de la Phrase 11.

Ces focalisateurs d’emphase ont été détectés en contexte et doivent être étudiés de la même manière pour comprendre leur fonctionnement. Aussi, les entités qui les entourent de part et d’autre sur l’axe syntagmatique, telles que les discours citants, leur donneront une éventuelle base d’appui pour la comparaison de mesures.

Par sa hiérarchisation en degrés d’emphase, cette étude permet aussi de standardiser les réalisations attendues en écartant les forts degrés d’emphase. Ainsi pourront être trouvés des points d’appui stables pour observer les réalisations pointées comme déficientes chez les apprenants et mentionnées en amont (Chapitre II), ainsi que la mélodie de la phrase « neutre ». J’ai pu éclairer mes choix grâce à une hiérarchisation des 26 phrases selon le paramètre d’emphase. Je me propose d’analyser acoustiquement les réalisations des Phrases 6 et 9 afin de sélectionner celle qui correspondra le mieux aux critères de stabilité.

Cette étude corrobore en outre les résultats de l’étude précédente sur les « bons » ou « mauvais » lecteurs. Le Lecteur B, trop rapide, ne marque que peu de *loci*⁶⁴ d’emphase. Le style lent et emphatique du Lecteur C facilite la compréhension en fournissant plus de repères aux apprenants de niveau faible, tandis que le Lecteur A reste très expressif sans jamais sembler pesant : sa maîtrise de l’emphase par rapport au tempo est maximale par rapport aux autres lecteurs étudiés.

L’analyse acoustique implique un travail long et souvent fastidieux. L’application du phénomène de perception d’emphase a permis d’établir un repérage non seulement des mots mais aussi des phrases en vue de leur présélection. L’analyse acoustique va affiner le profilage de sélection par vérification de paramètres objectifs.

⁶⁴ Je reprends cette métaphore du lieu-*locus* de la variation de Ballier *et al.* 2016 , en précisant bien qu’il ne s’agit pas des équations du *locus*.

4 Étude acoustique des marqueurs d'emphase principaux chez les anglophones natifs (expérimentation COR-AC-N2)

L'analyse des résultats de l'expérimentation perceptive par l'application du phénomène d'emphase a permis dans un premier temps de catégoriser et de hiérarchiser les mots et les phrases du texte *The Selfish Giant* en fonction de ce paramètre et grâce à cela de sélectionner des marqueurs d'emphase afin de les analyser acoustiquement. C'est ce que la présente expérimentation des corrélats acoustiques des natifs anglophones (COR-AC-N2) se propose de faire en procédant à l'examen de l'exclamative en *How*, des trois occurrences de l'adverbe de degré ou intensificateur *so* suivi d'un adjectif ou d'un adverbe, de l'interrogative en *What*, ainsi que du pronom indéfini *nobody*.

L'objectif étant de privilégier le lien entre la forme et le fond, entre le discours, le ressenti et la situation énonciative au sens large, bien qu'aucune explication théorique ne soit *a priori* écartée pour éclairer les phénomènes, une approche énonciative, pragmatique pourra venir éclairer le propos le cas échéant. Le cadre théorique sur lequel s'appuient les descriptions est celui de l'approche traditionnelle adoptée par les théoriciens britanniques de l'intonation. Ces présentations, bien loin d'être exhaustives, n'ont pour ambition que de situer les aspects les plus représentatifs du phénomène considéré. Les phénomènes les plus saillants prennent néanmoins appui sur des recherches sur les grammaires et ouvrages de référence ainsi que sur quelques articles se focalisant sur le sujet.

Un exposé préalable dans le cadre d'approches syntaxique, sémantico-discursive puis prosodique serviront d'introduction à une évaluation acoustique de l'intérêt des marqueurs pour l'étude de l'interlangue prosodique qui lui fait suite.

4.1 L'exclamative en *How*

L'expérimentation préalable PER-EMPH-N (test perceptif de détermination d'emphase chez les anglophones natifs) a montré que l'exclamative occupait dans le texte du *Selfish Giant* la place privilégiée parmi les marqueurs d'emphase grâce au mot *happy* (1^{er} rang) puis au mot *How* (XX^{ème} rang) dans la Phrase 5 : "*How happy we are here!*" *they cried to each other.*

4.1.1 Caractérisation linguistique des exclamatives

Le développement ci-dessous reprend les notions syntaxiques générales du marqueur exclamatif anglais ; il le contraste à son équivalent français et le rapproche de structures déclaratives anglaises apparentées. Il souligne en dernier lieu la spécificité de la ponctuation qui l'accompagne.

La phrase exclamative est considérée comme un dérivé de l'interjection selon Montlivault, qui note (1828 : 14) :

C'est une simple voix, un cri inarticulé également arraché à la douleur, à la joie, à la surprise, à l'indignation, à l'admiration, etc. L'*interjection* ou l'*exclamation* n'est donc que la révélation auditive d'un état subit de l'ame [*sic*], une véritable explosion du sentiment qui se manifeste par un cri, tant soit peu modifié par les différents peuples, et que l'Écriture représente ou plutôt note comme elle le peut ; car il appartiendrait plutôt à la musique qu'à la parole ; il serait moins difficile de lui trouver un sens musical que grammatical.

Ainsi, Montlivault rapprochait les termes « exclamation » et « interjection », qu'il regrettait de ne pouvoir classer dans aucune catégorie grammaticale. Il est probable qu'il se soit inspiré des remarques d'Arnauld et Lancelot ([1660] 1803) : 384), qui avaient avant lui qualifié les interjections de « voix plus naturelles qu'artificielles, qui marquent les mouvements de notre ame » [*sic*]. Les dictionnaires de français *Larousse*⁶⁵ en ligne reprennent la substance de la définition de Montlivault et donnent au mot « exclamation » une acception linguistique large englobant interjections et exclamatives : « Type de phrase, parfois réduite à une interjection, exprimant une émotion vive ou un jugement affectif. (Par opposition à l'assertion.) ». Il semble donc opportun de tenir compte des jugements d'Arnauld et Lancelot sur l'expression authentique et naturelle de Montlivault, qui valorise la nécessité d'une congruence affective et linguistique dans le phénomène exclamatif. O'Connor et Arnold insèrent des exemples d'exclamatives dans leurs catégories intitulées *Interjections* ([1961] 1973). Larreya et Rivière avancent eux aussi que la plupart des interjections peuvent être considérées comme des formes particulières d'exclamation (2010 : 274).

Greenbaum distingue quatre types de phrases : l'exclamative est l'une d'elles et son utilisation la plus courante dans le discours est celle d'exprimer des exclamations (1996 : 45). Cruttenden souligne qu'il est bien plus aisé d'identifier l'effet produit que la structure

⁶⁵ [<http://www.larousse.fr/dictionnaires/francais/exclamation/32013#qKVBYY0E55vcSPRY.99>] (consulté en 2015)

exclamative ([1986] 1997 : 166). Aussi, seules les structures canoniques retiendront notre attention.

Les phrases exclamatives, dans leur forme canonique, sont introduites par *what* ou *how*, tout comme les questions, mais le but est seulement d'attirer l'attention de l'interlocuteur. L'ordre canonique est alors modifié, étant donné que l'élément en *wh-*, quelle que soit sa fonction syntaxique, sera antéposé par rapport aux autres constituants afin d'être placé en *position of initial prominence* (Quirk *et al.*, 1985 : 833), sans pour autant occasionner d'inversion du sujet et de l'opérateur, contrairement aux questions en *wh-* (Quirk *et al.*, 1985 : 833).

Quirk *et al.* considèrent la fonction de l'élément en *wh-* (Quirk *et al.*, 1985 : 834) :

The limitation to these functions is not surprising when one realizes that the *wh-* word indicates an extreme position on some scale of value, and therefore can only appear at points in the sentence where an expression of degree is possible. Exactly the same functions are fulfilled (especially in some women's speech) by the emphatic degree items *such* (as a determiner) and *so* (as an intensifier) in statements and questions [...].

Ainsi que le rappellent Downing et Locke ([1992] 2006 : 191), la structure d'une proposition exclamative est donc la suivante : elle débute par un mot en *wh-*, qu'il soit *the determinative what* suivi d'un groupe nominal, ou l'adverbe de degré *how* et d'un adjectif, d'un adverbe ou d'une déclaration (*statement*). La seconde partie de l'exclamative est celle d'une déclarative et l'élément qui suit le mot en *wh-* est un élément de la proposition qui a été antéposé. Larreya et Rivière (2010 : 273) soulignent la différence de construction française, qui rompt la relation de proximité entre l'adverbe exclamatif et l'élément sur lequel il porte :

En français, dans les phrases du type « Comme cet article est intéressant ! » ou « Comme elle chante bien ! », l'exclamatif « comme » est séparé de l'adjectif ou de l'adverbe sur lequel il porte. Dans les équivalents anglais, il faut en revanche placer l'adjectif ou l'adverbe **immédiatement après *how*** : *How interesting this article is!* / *How well she sings!* » [...]

En français, l'exclamative dépourvue du « comme », constitue une déclarative indépendante. Par contre, en anglais, la transformation *wh* (*wh fronting*) fait basculer le lieu de focalisation à l'avant de l'exclamative. À partir de cette comparaison des versions françaises et anglaises, nous constatons donc une inversion symétrique du point focal syntaxique dont les répercussions prosodiques seront examinées dans la partie suivante. Popin attribue (1998 : 33)

4. Étude acoustique des marqueurs d'emphase principaux chez les anglophones natifs (expérimentation COR-AC-N2)

aussi la valeur de morphème exclamatif à « que » dont il en donne un exemple d'emploi : « Qu'il fait beau ! ».

Ce conditionnement syntaxique par la position antéposée du mot en *wh-* fait souligner par Biber *et al.* (1999 : 909) la force exclamatoire que l'on trouve aussi dans des portions déclaratives pourtant dépourvues de *wh-*, notamment celles où il est remplacé par *such*. Ils rappellent que, même si l'objectif de l'utilisation de la structure antéposée nominale semble être de tirer parti de manière maximale des deux points de focalisation de la proposition que sont le début et la fin, certaines structures, telles celle de l'exclamative, concentrent cette focalisation sur le début uniquement (*single focus structure*) (1999 : 902).

Les exclamatives sont le plus souvent ponctuées à l'écrit par un point d'exclamation, ce qui, selon Greenbaum (1996 : 521), a deux finalités :

- celle de faire prononcer la phrase exclamative avec vigueur (*to signal that the sentence is a forceful utterance*),
- celle de marquer la fin de la phrase.

Ce point d'exclamation, en tant que signe de ponctuation, ne peut se trouver que dans un texte écrit, et non pas à l'oral. Il évoque un langage parlé et on le trouve spécifiquement dans des écrits personnels et intimes, des notes informelles et des transcriptions de dialogues (Greenbaum, 1996 : 521). Albrespit souligne que cette marque de ponctuation n'apparaît pas toujours et qu'elle n'est pas obligatoire pour signaler une exclamative (2011 : 107). En cas d'absence, le lecteur se doit d'interpréter l'énoncé et lui donner une valeur exclamative ou pas. Sa présence ne fait donc que renforcer la qualité exclamative de l'énoncé. Popin note que toute phrase énonciative peut être rendue exclamative grâce à la présence d'un point d'exclamation : il est alors intentionnel (1998 : 32).

Les autres marques de ponctuation insérées dans la Phrase 5 sont les guillemets et le point final. L'exclamative de notre corpus constitue une phrase entière au discours direct, ce qui la classe dans le cas n°1 présenté par Popin (1998 : 26) sans toutefois être introduite par deux points, qu'il caractérise ainsi :

- les citations sont annoncées par le deux-points et commencent par une majuscule ;
- elles sont encadrées de guillemets ;

- on place à l'intérieur des guillemets la ponctuation normalement attendue par la (les) phrase(s) entre guillemets qui va se confondre avec la ponctuation de fin de phrase

Les guillemets englobent l'intégralité de l'exclamative, y compris le point d'exclamation, la citation directe faisant partie des cas privilégiés de cette utilisation. Les guillemets utilisés correspondent aux guillemets anglais (Popin, 1998 : 28). Ils marquent le passage de la narration au discours direct et la reprise de la narration : ils sont donc « signes de l'énonciation » (Popin, 1998 : 103-104).

On a ainsi pu voir que, de l'introduction et l'emplacement des signes de ponctuation à l'agencement des catégories grammaticales en son sein, la raison d'être d'une exclamation est de véhiculer une émotion, un émoi, qu'il soit ressenti par le locuteur ou destiné à influencer sur l'auditeur.

4.1.2 Approche sémantico-discursive des exclamatives

Le caractère émotionnel ou attitudinal des exclamatives les classe parmi les modalités appréciatives. La ponctuation renforce par ailleurs une expressivité directe. Il y a plus de 50 ans, une signification générale des exclamatives en *how* était répertoriée. Au jour d'aujourd'hui, on peut se poser la question de la représentativité de cette exclamative en *how* dans la langue courante actuelle.

La non inversion du sujet et du verbe catégorise les exclamatives parmi les exclamatives courantes par opposition aux exclamatives littéraires du genre : « How strange is his appearance! » (Quirk *et al.*, 1985 : 834). En effet, et comme vu précédemment, la structure de l'exclamative courante est semblable à celle d'une déclarative, avec toutefois l'élément suivant le mot en *wh-*, qui est un élément de la proposition antéposé. Ceci fait souligner par Downing et Locke la condition structurelle de la source émotionnelle ou attitudinale véhiculée par l'exclamative ([1992] 2006 : 191) :

For these reasons, exclamative clauses are sometimes considered as an emotive element superimposed on the declarative rather than as a distinct mood.

De plus, une structure de focalisation unique, ce que Biber *et al.* nomment *single focus structure*, rompt avec les schémas environnants, ce qui ne fait qu'amplifier le focus initial. Albrespit reconnaît lui aussi à l'exclamative un « caractère émotionnel » qui classe

l'exclamation parmi les modalités appréciatives, un jugement étant porté sur le prédicat (2011 : 111). Il prend la phrase *Oh how delightful!* comme exemple, et en dit (2011 : 112) :

Après un parcours des degrés de la propriété *be delightful* (avec un adjectif gradable comme *delightful* il s'effectue sur une échelle), il n'est pas possible de trouver un dernier point, une occurrence parfaite de *delight*. Le parcours semble pouvoir avoir lieu à l'infini. Il y a renvoi au haut degré de la notion.

Ce point le conduit à relayer les remarques précédentes de Quirk *et al.* et à approfondir les comparaisons avec d'autres marqueurs d'emphase :

Avec *what* l'opération de parcours n'aboutit pas, comme avec *how*. La différence est qu'avec *how* (*How stupid!*), comme avec *such* (*Such a stupid idea!*) et *so* (*So stupid!*), il y a « auto-référentiation » : la référence se construit à l'intérieur de la propriété.

Albrespit précise que l'utilisation des adjectifs mélioratifs, décrits comme « favorables », indiquant une « bonne valeur », opère un parcours à l'intérieur du domaine, la variation s'orientant positivement vers le centre organisateur. Ceci vaut pour l'adjectif *happy*.

Pour un lecteur francophone aussi, le point d'exclamation est un marqueur privilégié. Seguin (2000 : 95) rappelle les trois grands caractères du point d'exclamation dégagés par Bonnard dans son article *Ponctuation* du *Grand Larousse de la Langue Française* (pages 4455 et 4456) : il est un signe final, intonatif, et modal. Seguin ajoute : « Il concerne donc l'ensemble d'une phrase, le ton de sa profération, et une expressivité, à préciser. » et représente ce signe de « l'expressivité graphique » (Laufer, 1980 : 79) en précisant (2000 : 94-95) son rôle intermédiaire comme ponctuation non pas nécessairement réservée au comédien devant déclamer ses répliques, mais destinée à être lue silencieusement, dans sa tête, c'est-à-dire qu'elle devient une « ponctuation *dramaturgique*, par laquelle la *mise en page* transcrit une mise en scène au sens large du mot », apparentée en quelque sorte au « langage » cinématographique (2000 : 86). Observons la ponctuation utilisée dans la Phrase 5. Elle délimite et renforce la relation interlocutive. Selon Popin (1998 : 26), le rôle fondamental des guillemets est de signaler « toute citation directe, tout emploi en mention, tout emploi avec lequel l'auteur prend ses distances » afin de montrer qu'est opéré un « changement dans l'énonciation ». Ainsi l'exclamative est clairement délimitée aux yeux du lecteur et se détache de l'énoncé citant qui la suit. Ce « signe de l'énonciation » marque un premier degré vers l'émotionnel et l'attitudinal. Ainsi, selon l'analyse de Seguin, le point d'exclamation « tend donc à avoir un contenu plus indiciel qu'informationnel, dont le lecteur

est amené à inventer lui-même la teneur » (2000 : 95). Il devient « cette marque ostensible de l'engagement affectif de celui qui est, à l'écrit, représenté comme sujet parlant » et concourt aux « moyens de l'illusion de la fonction dite « émotive » ». Il devient alors le signe d'une « expressivité directe » en informant des « réactions personnelles immédiates du locuteur ».

Popin classe le point d'exclamation parmi les « signes par évocation », dont la première fonction est d'être « indice de l'oralité et du discours direct », ce qui lui confère la fonction de transformer la narration dans laquelle il est inséré en discours, en révélant un énonciateur (1998 : 100-102). Il présente la seconde fonction comme étant « indice de valeurs modales » et susceptible de charger l'énoncé « de toutes les valeurs affectives que le lecteur veut bien y projeter ». Le point d'exclamation souligne la complexité de la relation interlocutive et des étapes à franchir pour y parvenir :

Il y a nécessairement illusion d'une « interprétation » de l'exclamation. Il est d'ailleurs vraisemblable que dans bien des cas cette interprétation finit par l'emporter sur toute autre fonction. Finalement la part d'information contenue dans l'élément (fonction locutoire) s'avérant faible, c'est plutôt son rôle dans le mécanisme de l'échange verbal qui importe (fonction illocutoire) et peut-être même la signification cachée qui s'y dissimule (fonction perlocutoire).

Le lecteur doit opérer des choix à partir du texte écrit afin de privilégier l'orientation sémantique qu'il souhaite transmettre. C'est ce choix qui déterminera la portée de l'échange interlocutif et privilégiera l'importance donnée à chacune des fonctions (locutoire, illocutoire ou perlocutoire). Il ne peut qu'être très clair pour les anglophones et les francophones.

La structure syntaxique ainsi que sa ponctuation convergent vers l'utilisation du schéma intonatif que le locuteur sélectionnera pour la valeur émotionnelle et attitudinale transmise par la lecture à haute voix. Autrement dit, c'est en fonction de la valeur modale qu'il souhaite transmettre que le lecteur devra sélectionner le schéma prosodique correspondant.

Ainsi O'Connor et Arnold (1961) dégagent dix schémas prosodiques généraux du syntagme intonatif et en répertorient la signification générale pour les *interjections* de toute nature dans lesquelles ils regroupent les exclamatives. Le tableau ci-dessous en rend compte. Les contours 3, 4, 5 et 9 semblent ne pas correspondre aux exclamatives en *how* (aucun exemple donné), c'est la raison pour laquelle ils ne figurent pas dans le tableau suivant. Cette présentation, impressionnante par sa diversité et sa précision, montre à quel point la situation

CHAPITRE III – Emphase : mode d’organisation paradigmatique des réalisations prosodiques des anglophones natifs

d’énonciation indiquée dans la colonne de droite peut faire varier le contour mélodique selon le sens à transmettre. Bien que ce répertoire date de plus de 50 ans et que la prosodie de la langue ait pu avoir beaucoup évolué depuis, ce répertoire méticuleux souligne le lien entre, d’une part, les situations d’énonciation et la valeur attitudinale à prendre en compte, et, d’autre part, leur équivalence prosodique systématisée par l’application de catégories tonales distinctes. La mention « emphatique » ou « non emphatique » indiquée dans chaque structure tonale est destinée à souligner des différences de hauteur mélodique des tons dans les schémas, sans être systématiquement retransmises dans leur signification d’ensemble.

	<i>Contour mélodique</i>	<i>Leur description générale</i>	<i>Explications sémantiques générales</i>
1	THE LOW DROP (UNEMPHATIC)	(Low Pre-head + (High Head +) Low Fall	with no head, calm, unsurprised, reserved, self-possessed; with a high head, very strong.
	THE LOW DROP (EMPHATIC)	High Pre-head +) (High Head +) Low Fall <i>or</i> (Pre-head +) Stepping Head + Low Fall	
2	THE HIGH DROP (UNEMPHATIC)	(Low Pre-head +) (High Head +) High Fall	mildly surprised, not so reserved or self-possessed as with the Low Drop.
	THE HIGH DROP (EMPHATIC)	High Pre-head + (High Head +) High Fall <i>or</i> (Pre-head +) Stepping Head + High Fall	
		(Pre-head +) High Fall(s) + High Fall	
6	THE LONG JUMP (UNEMPHATIC)	(Low Pre-head +) Rising Head + High Fall	protesting, surprised.
	THE LONG JUMP (EMPHATIC)	High Pre-head + Rising Head + High Fall <i>or</i> (Pre-head +) Climbing Head + High Fall	
7	THE HIGH BOUNCE (UNEMPHATIC)	(Low Pre-head +) (High Head +) High Rise	querying all or part of the listener's [...] interjection, but with no critical intention.
	THE HIGH BOUNCE (EMPHATIC)	High Pre-head + (High Head +) High Rise <i>or</i> (Pre-head +) Stepping Head + High Rise	
8	THE JACKKNIFE (UNEMPHATIC)	(Low Pre-head +) (High Head +) Rise-Fall	impressed, sometimes a hint of accusation.
	THE JACKKNIFE (EMPHATIC)	High Pre-head + (High Head +) Rise-Fall <i>or</i> (Pre-head +) Stepping Head + Rise-Fall	
10	THE TERRACE (UNEMPHATIC)	(Low Pre-head +) (High Head +) Mid-Level	(in final word groups) calling out to someone as from a distance.
	THE TERRACE (EMPHATIC)	High Pre-head + (High Head +) Mid-Level <i>or</i> (Pre-head +) Stepping Head + Mid-Level	

Tableau 40 – Tableau récapitulatif des dix structures tonales appliquées aux exclamatives en *how* et l'explication sémantique correspondante (d'après O'Connor et Arnold, 1961 : 41-45 ; 106-252)

Les remarques précédentes conduisent à considérer la représentativité d'une exclamative dans la parole actuelle. Albrespit (2011 : 104) mentionne la remarque de Swan⁶⁶ (1995 : 193) selon laquelle « les exclamatives en *how* sont ressenties comme appartenant à un style un peu formel ou démodé. » Downing et Locke émettent ([1992] 2006 : 191) également des réserves sur la fréquence d'usage de ce genre d'exclamatives dans la langue parlée commune, mais ouvrent une perspective, celle des exclamatives tronquées, très courantes, quel que soit le registre utilisé :

How-exclamative clauses sound somewhat theatrical nowadays, especially when followed by an adverb (*How well he played!*). More commonly heard than clausal

⁶⁶ SWAN, M. (2005). *Practical English Usage*. Oxford, O.U.P.

exclamatives in everyday spoken English are abbreviated noun-headed or adjective-headed forms:

[...] How exciting!

Les diverses facettes sémantiques résultant de l'organisation syntaxique de l'exclamative nous conduisent à présent à étudier leur réalisation prosodique dans le discours.

4.1.3 Approche prosodique des exclamatives

Les considérations syntaxiques et sémantico-discursives présentées en amont nous conduisent à considérer l'intonation de l'exclamative. La chute mélodique représente la forme prosodique canonique des exclamatives malgré de nombreuses variations secondaires.

Wells présente ainsi l'intonation liée aux exclamatives (2006 : 59) :

Exclamations (= expressions of surprise, anger or excitement) virtually always have a **fall**. We call this tone meaning the **exclamatory fall**. It can be seen as a sub-type of the definitive fall.

La chute mélodique représente le ton par défaut, c'est-à-dire celui qui est neutre et non marqué, utilisé dans les déclaratives (Wells, 2006 : 25). Le *definitive fall* est lui-même défini à la même page :

But in general we can say that by using a fall we indicate that what we say is potentially **complete** and that we express it with **confidence**, **definitely** and **unreservedly**. The fall thus also tends to signal **finality**. We call this tone meaning the **definitive fall**.

Cruttenden note que les exclamatives épousent un ton descendant, qui peut parfois devenir un ton descendant-montant (2014 : 296). L'exemple donné d'une exclamative en *how* informe de la présence du noyau sur l'adjectif : « How `stupid he is! ».

Albrespit reprend les assertions de Wells et précise que la chute mélodique, représentant le schéma par défaut, peut ne pas s'appliquer lorsque l'exclamation est teintée « d'ironie, de surprise, de désapprobation » (2011 : 108). Il remarque que c'est une mélodie montante-descendante (*rise-fall*) qui sera choisie pour exprimer une forte implication du locuteur « qui s'extasie, s'émerveille, dit son approbation ou au contraire exprime son empathie ou son désarroi ». Il remarque parfois dans le « caractère émotionnel » de l'exclamative, l'utilisation possible de traits non linguistiques tels que le souffle (*breathy*

voice), tout spécifiquement lorsqu'il s'agit d'échanger informations scandaleuses ou secrets. Il précise en outre que la mélodie peut être descendante-montante (*fall-rise*) dans une exclamation exprimant la surprise, et rappelle qu'Adamczewski et Keen (1973 : 235) avaient noté ce même ton mélodique en cas de contradiction, divergence d'opinions ou regrets.

Dans le tableau récapitulatif des dix structures tonales des exclamatives en *how* (d'après O'Connor et Arnold, 1961 : 41-45 ; 106-252) présenté en amont, l'extrême variation de tons et de hauteur mélodique permet néanmoins d'en repérer les trois éléments distinctifs les plus représentatifs : le noyau, la tête et pré-tête, la queue. En fait, selon Cruttenden ([1986] 1997 : 166), la zone de variation distinguant la prosodie propre à chaque langue semble être celle précédant le nucléus.

La structure syntaxique avec focalisation initiale unique (transformation *wh*) est renforcée par le schéma prosodique de l'exclamative, lui donnant toute sa valeur emphatique. Biber *et al.* notent ainsi (1999 : 904) :

The fronting has an intensifying effect, which is often strengthened by the choice of words (*horrible, bloody, amazing, etc.*), or by emphatic stress when spoken (reflected by exclamation marks). Furthermore, the fronted material is new rather than old information.

4.1.4 Intérêt des exclamatives dans l'étude de l'interlangue prosodique

L'exclamative de notre corpus est conforme à la description de Quirk *et al.* :

How happy we are here ! [C_s S V]

How happy correspond au complément du sujet, *we* est le sujet, *are* la copule, et *here* le circonstanciel de lieu. La tête de l'exclamative est formée de l'adverbe de degré *how* suivi de l'adjectif graduable *happy*. Elle est ponctuée par un point d'exclamation en fin de phrase. L'analyse de ses constituants la classe parmi les exclamatives canoniques sur le plan syntaxique.

En ce qui concerne sa réalisation prosodique, Wells trouve les exclamatives d'un grand intérêt dans l'amélioration de l'interlangue (2006 : 60) :

Exclamations are thus the simplest kind of utterance for the student of EFL. The rule is: if it's an exclamation, say it with a fall. Exclamatory falls are excellent drill material for anyone who wants to practise falling tones.

Cette phrase constitue donc un marqueur susceptible de révéler des réalisations fructueuses pour l’analyse de l’interlangue et de son évaluation prosodique. Elle permettra par ailleurs de vérifier si l’apprenant sait mobiliser sa compétence prosodique afin d’atteindre la performance attendue. Il s’agit ainsi de vérifier le degré d’adéquation entre l’objectif à atteindre et la cohérence des réalisations prosodiques. Pour cela, les mesures d’empan de degré de tessiture sont susceptibles de révéler les réalisations maximales que l’apprenant est capable d’effectuer. Les syllabes sélectionnées pour porter tel ou tel ton, en accord avec la variation de durée entre les syllabes, devraient renseigner sur l’aisance avec laquelle l’apprenant véhicule une dimension attitudinale, révélatrice d’un haut degré de maîtrise prosodique.

Les dimensions émotionnelles et attitudinales évoquées précédemment dressent néanmoins une limite de portée de ces marqueurs, que l’on ne retrouvera que dans certains types de parole, alors qu’il serait préférable de trouver des traits distinctifs en vue d’une évaluation dans une gamme typologique la plus large possible.

Hirst⁶⁷ remarque que ce genre de structure n’est plus utilisé par les jeunes générations et que celles-ci les assimilent à une affirmative du genre (mon exemple) :

How sunny it is today!

It is sunny today!

Le texte étudié ici présente cette phrase exclamative prononcée au style direct. Elle se doit de traduire l’émotion des enfants, la vérité des sentiments, bien davantage que si elle représentait la voix du narrateur. Un cri ne peut se fondre dans la prosodie des phrases adjacentes. Ce sont ces éléments que nous devrions retrouver dans l’analyse acoustique des natifs.

4.1.5 Analyse acoustique de la temporalité

Le graphique ci-dessous montre un certain consensus des réalisations temporelles. Cela est flagrant pour les syllabes /i:/, /wi:/ et /A:/⁶⁸, dont la brève durée laisse penser qu’elles sont inaccentuées. Les variations notoires se produisent sur /haU/ et /hI@/, dont la durée peut être expliquée par les diphtongues qu’elles contiennent, mais surtout sur /h{p/, dont la voyelle est brève, mais qui représente une forte localisation d’emphase. La plus grande

⁶⁷ Conversation en 2016.

⁶⁸ La transcription des phones est opérée avec le jeu de caractères *SAMPA* (Speech Assessment Methods Phonetic Alphabet). [<http://www.phon.ucl.ac.uk/home/sampa/>]

variation est opérée par la pause qui sépare l'exclamative de son énoncé citant. Les lecteurs rapides (souvent les plus jeunes) tendent à ne marquer aucune pause, alors que d'autres utilisent la pause comme effet stylistique afin de mettre l'exclamative en exergue, ce qui explique sa très forte variation.

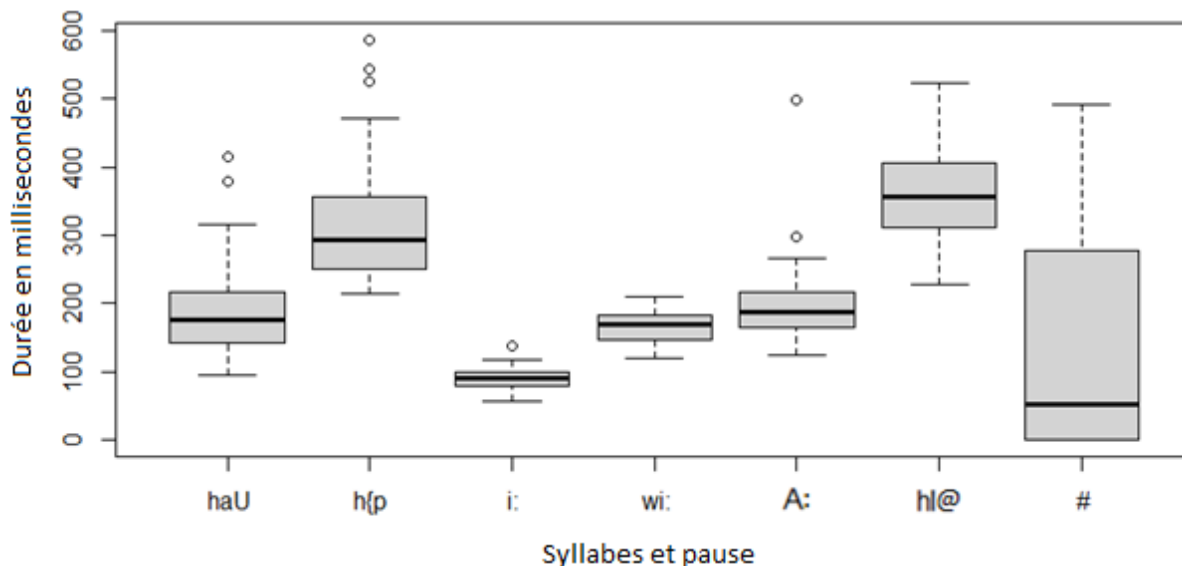


Figure 44 – Exclamative en *How* : Durée moyenne des syllabes et de la pause éventuelle

4.1.6 Analyse acoustique de la mélodie

Le Graphique suivant représente les contours mélodiques de la proposition exclamative. Il a été réalisé à partir des mesures de la moyenne de F0 prise automatiquement par *ProsodyPro* à partir des TextGrids, sur la base de la syllabe.

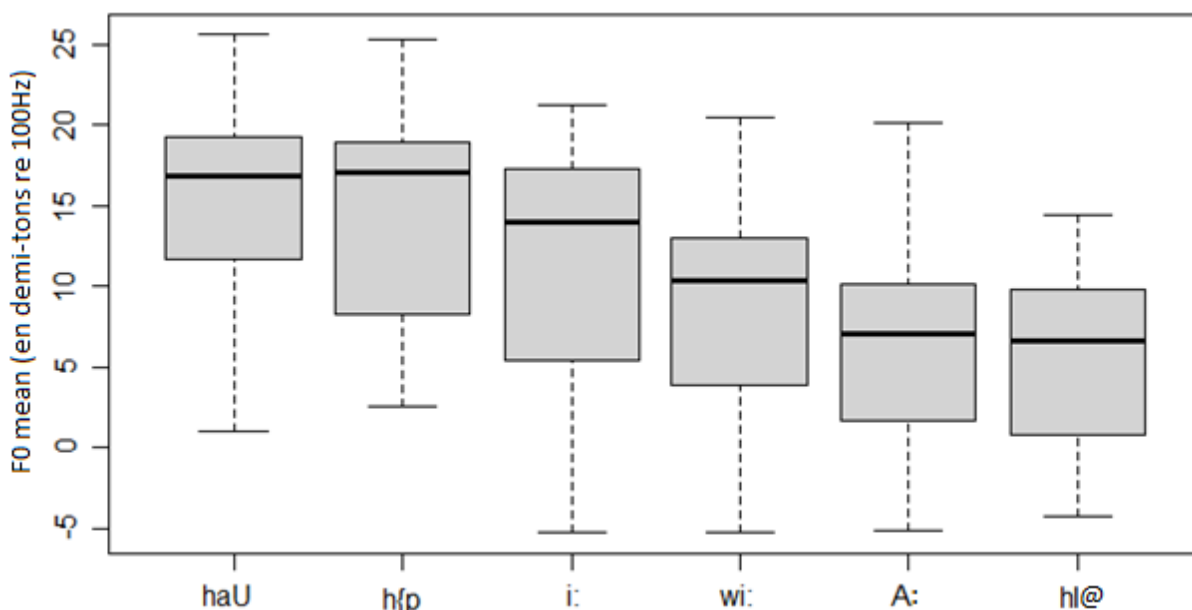


Figure 45 – Moyenne de F0 par syllabe de l'exclamative en *How* (sur 31 occurrences)

Le contour mélodique de F0 corrobore l’assertion de Wells mentionnée en amont sur le modèle de chute tonale que représente l’exclamative anglaise, surtout en ce qui concerne la ligne médiane de chaque boîte à moustaches et la dernière syllabe de la phrase, laquelle révèle peu de variation dans les réalisations. La syllabe /haU/ montre peu de variation dans 50% des cas. Il n’est pas surprenant qu’une proportion très restreinte de courbes mélodiques descendent au-dessous de 0 DT étant donné que cette exclamative est attribuée à de jeunes enfants, dont les voix ne peuvent qu’être aiguës.

Les parties les plus hautes (F0 max) de la mélodie tonale dans *how happy* ont été mesurées manuellement ainsi que les parties les plus basses (F0 min) de la fin de l’exclamative⁶⁹ en faisant usage des TextGrids alignés dans le logiciel *WinPitchW8*⁷⁰. Pour chaque fichier, il a été vérifié que les harmoniques du spectrogramme correspondaient à la courbe de F0. Malgré la qualité médiocre de certains enregistrements, la puissance d’analyse de *WinPitchW8* a permis de mesurer avec précision 38 occurrences. Les données obtenues ont révélé deux catégories en fonction de la hauteur mélodique de /h{p/ par rapport à /haU/. Le graphique représente sur l’axe des abscisses l’âge de tous les lecteurs et sur l’axe des ordonnées la différence de fréquence fondamentale maximale en demi-tons. L’histogramme montre les données aux valeurs positives (noyau d’une F0 max plus élevée que celle de la tête) et aux valeurs négatives (noyau d’une F0 min plus basse que celle de la tête).

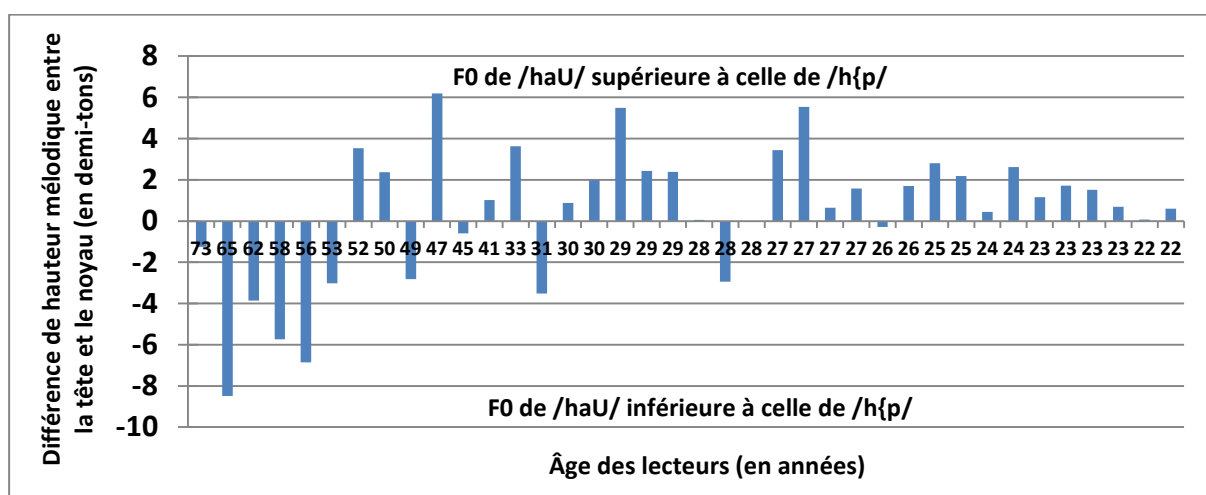


Figure 46 — Répartition de la hauteur mélodique maximale parmi les 38 lecteurs par rapport à leur âge

⁶⁹ 4 lecteurs n’ont pas lu le mot *here*. 4 autres ont lu *are* sur un ton plus bas que *here*. Pour cette raison, seule la fréquence fondamentale la plus basse parmi ces deux mots a été prise en compte.

⁷⁰ *WinPitchW8*, logiciel de Ph. Martin. [<http://www.winpitch.com/>]

Le graphique ci-dessus montre que les jeunes lecteurs ont tendance à prononcer le noyau à une hauteur mélodique supérieure à celle de la tête et que l'inverse se produit pour les lecteurs plus âgés. Destiné à mettre ces catégories en valeur, le graphique en courbes de F0 représente ce qui pourrait être un « profil jeune » et un « profil mature ». Toutes les données positives ont été regroupées pour constituer une moyenne du contour de F0 de 27 lecteurs alors que toutes les données négatives font de même avec une moyenne du contour mélodique de onze lecteurs. Seul un lecteur âgé de 28 ans n'a pas été pris en compte dans l'élaboration du graphique étant donné que les mesures de F0 max étaient identiques.

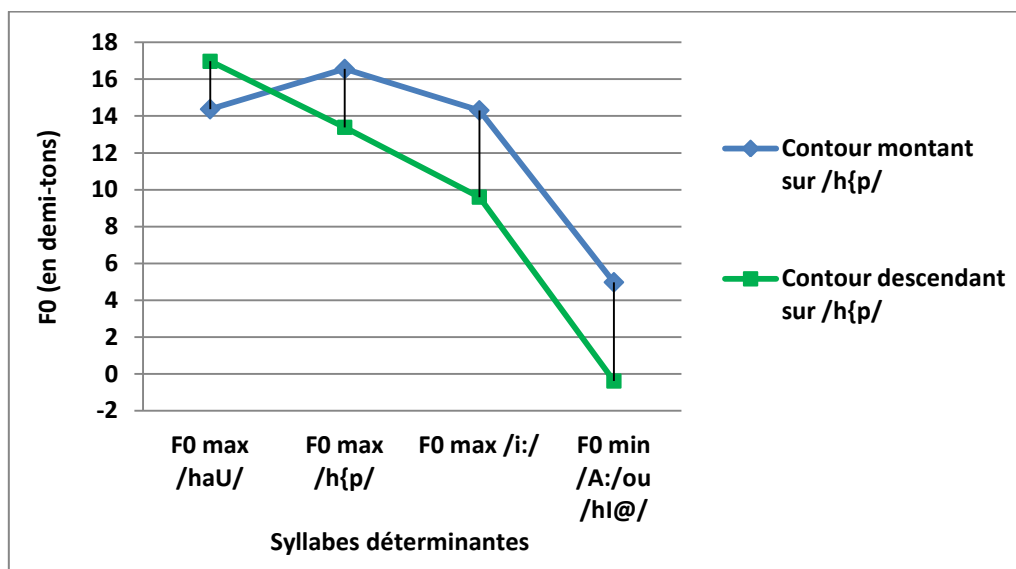


Figure 47 — Exclamative en *How* : Diagramme représentant sous forme de courbes les contours mélodiques du « profil jeune » et du « profil mature »

Ce graphique montre clairement un contour mélodique révélateur d'une typologie indicative de l'âge du lecteur. Le « groupe de plus de 40 ans » comprend quatre exceptions, qui sont des enseignants d'enfants ou de jeunes adultes. On pourrait interpréter cet aspect du « parler jeune » comme un moyen (sans doute inconscient) de prendre une allure plus « branchée », susceptible de mieux les faire accepter, et par là-même respecter, par leurs élèves. À l'autre bout de l'échelle, les trois exceptions du « groupe au-dessous de 40 ans », sont des lecteurs enseignant l'anglais à de jeunes adultes à l'université. L'utilisation d'une intonation aux inflexions plus « matures » a pu leur permettre de mieux asseoir leur statut professionnel face à des étudiants de leur propre âge, ou parfois même plus âgés. Les boîtes à moustaches représentant les deux premières syllabes dans le contour général de F0 ont été impuissantes à établir ou même suggérer ces catégories typologiques.

Une interprétation pertinente est la remarque de Hirst mentionnée en amont. Dans une exclamative, *How* représente la tête précédant le noyau. Dans l’affirmative de référence *We are happy to be here !*, la tête est remplacée par la pré-tête *We*, pour laquelle il est normal de ne pas être d’une F0 élevée ici. La tendance, somme toute subliminale, semble parfaitement logique et en accord avec les observations.

4.1.7 Conclusion

Malgré son faible taux d’occurrences dans le discours oral familial hormis dans sa forme tronquée, l’exclamative reste un moyen fiable de tester la capacité d’un apprenant à appliquer un contour descendant sur un syntagme intonatif de plusieurs syllabes et en cela est un indicateur de registre. Il est néanmoins prudent de concilier le profil mélodique avec l’âge du locuteur.

L’hypothèse de départ H3 concernant l’utilisation de l’emphase est qu’elle favorise la variabilité des réalisations phonétiques. L’étude détaillée des deux mots pointés comme les plus emphatiques (*How happy*) a validé l’Hypothèse 3 en révélant leur variation acoustique à laquelle un facteur d’âge semble être lié via la désuétude de la structure complète de l’exclamative en *How* dans l’usage conversationnel courant.

Nos deux « bons » lecteurs ont lu la Phrase 5 sans prononcer l’adverbe *here*. Les mesures de F0 les concernant sont représentées dans le graphique suivant.

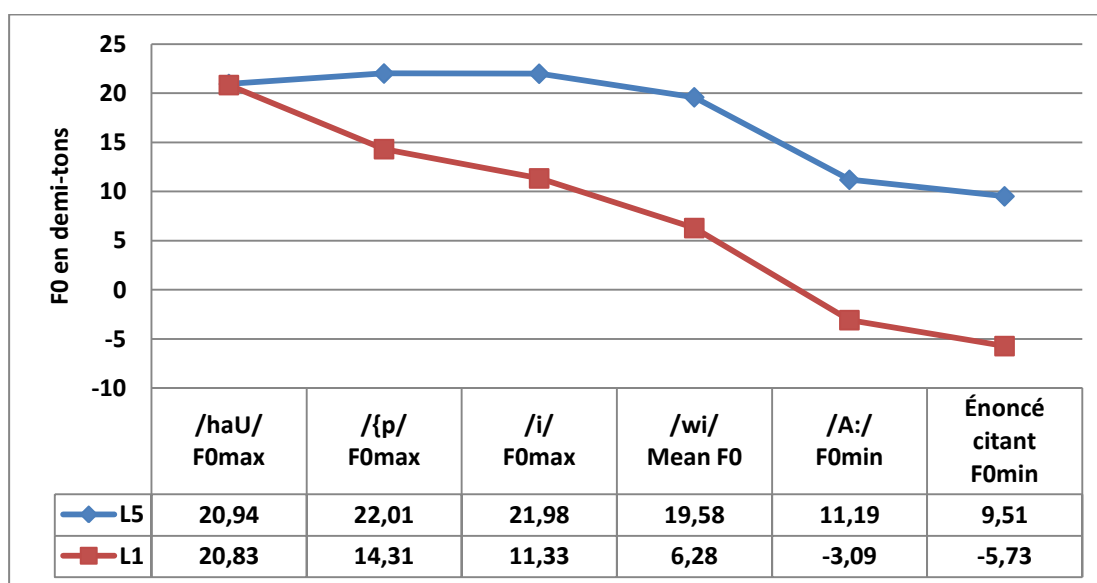


Figure 48 – F0 des profils de « bons » lecteurs (L1 et L5)

On peut voir que ces deux lecteurs sont bien représentatifs du contour intonatif correspondant à leur âge et que l'ampleur du registre est surtout imposante pour notre « meilleur » lecteur L1. Le tableau qui fait suite la précise en montrant que c'est non pas dans la seule exclamative, mais dans le couple « exclamative + énoncé citant », que l'on peut mesurer le plus grand écart entre la fréquence fondamentale maximale et la fréquence fondamentale minimale.

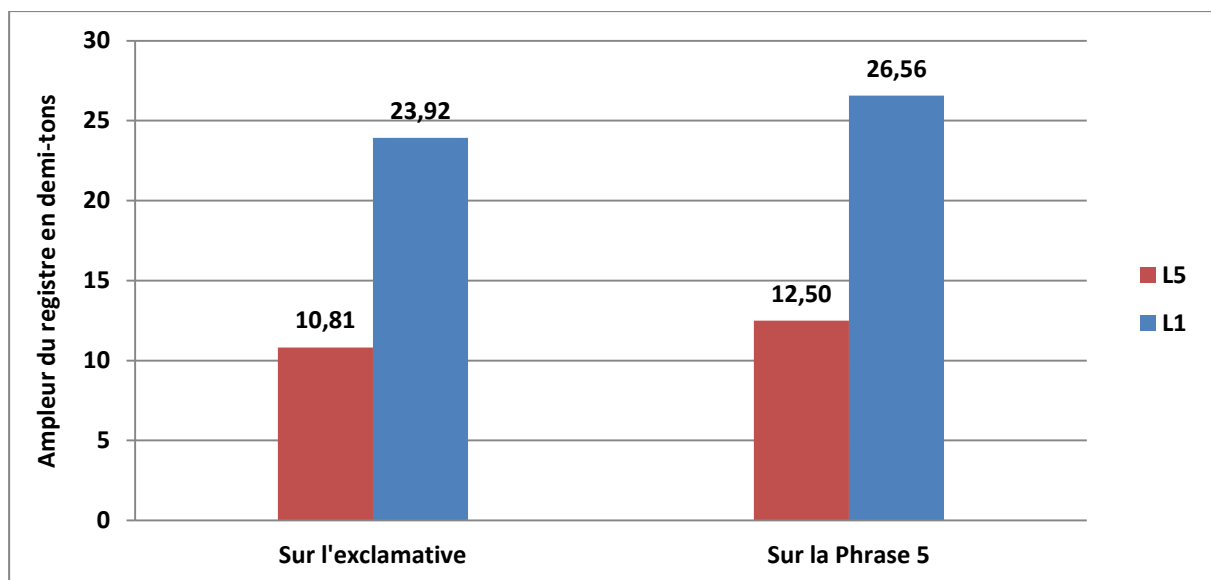


Tableau 41 – Ampleur du registre des lecteurs L1 et L5 par différence entre F0max et F0min

4.2 L'adverbe de degré ou intensificateur *so* suivi d'un adjectif ou d'un adverbe

Même si *so* occupe la deuxième position de proéminence emphatique dans la Phrase 22, puis la 7^{ème} dans la Phrase 4, une troisième occurrence se produit dans la Phrase 19 au 166^{ème} rang *ex æquo* avec deux autres mots. S'agissant de trois portions de phrases similaires syntaxiquement avec paires « *so* + adjectif ou adverbe » (*so sweetly* dans la Phrase 4, *so sweet* dans la Phrase 22, *so late* dans la Phrase 19), il me semble important de les étudier toutes trois afin de comprendre les limites du phénomène emphatique sur cette structure. La présente étude ne porte que sur ces adverbes de degré et d'intensité, suivant la dénomination choisie, étant donné que c'est cette catégorie syntaxique qui est pointée comme emphatique dans notre corpus.

4.2.1 Caractérisation linguistique

Huddleston et Pullum reconnaissent à *so* sa qualité d'adverbe (2002 : 570). Les adverbes représentent une catégorie dont il est difficile de rendre compte. Déjà Montlivault

remettait en question la dénomination d' « adverbe » (1828 : 220), cet « hybride du langage » (1828 : 89). Greenbaum présente d'ailleurs sa perplexité quant à la nature syntaxique des adverbes (1996 : 141) :

Adverbs are a heterogeneous class, varying greatly in their functional and positional ranges. They constitute a series of overlapping subclasses, and some of them belong to more than one subclass.

Quirk *et al.* qualifient la classe des adverbes de *the most nebulous and puzzling of the traditional word classes* et pensent que cet état résulte de la difficulté à faire correspondre cet élément aux critères définissant les autres classes de mots, et qu'il engendrerait la création, par certains grammairiens, de nouvelles classes de mots hors système adverbial (1985 : 438).

Comme tous les adverbes, l'adverbe *so* sert à modifier. Les adjectifs et adverbes graduables (dotés d'un niveau d'intensité) peuvent être modifiés, modulés, grâce à des adverbes de degré. Parmi ses diverses fonctions syntaxiques, l'adverbe *so* peut être adverbe de degré. *Adverbs of degree describe the extent to which a characteristic holds* est la définition qu'en donnent Biber *et al.* (1999 : 554). Ils nomment la catégorie qui en dépend et qui marque le degré supérieur celle des amplificateurs (*amplifiers*) ou intensificateurs (*intensifiers*). Les termes sont donc synonymes. Ils distinguent cependant ceux qui, tout en modifiant des adjectifs graduables, indiquent des degrés sur une échelle (tels *more, very, so, too, extremely, etc.*) alors que d'autres amplificateurs (comme *totally, absolutely* et *quite* dans le sens de *completely*) indiquent l'extrémité supérieure de l'échelle (2005 : 555).

D'après Downing et Locke ([1992] 2006 : 488), le phénomène d'intensification peut être catégorisé en trois degrés : le degré élevé (*high*), le degré moyen (*medium*) et le degré faible (*attenuated*). Ils pensent que cette gradation, même s'ils la présentent sous forme d'échelle, correspond davantage à un continuum qu'à des paliers, étant donné que la gradation est uniquement opérée par des items lexicaux et non par des structures d'ordre syntaxique. L'adverbe de degré *so*, quant à lui, est caractérisé par son fort degré d'intensité tout comme les adverbes *very, really* ou *most*.

Quirk *et al.* ne subdivisent les intensificateurs qu'en deux catégories : les amplificateurs (*amplifiers*) et les atténuateurs (*downtoners*), puis ils raffinent leur hiérarchie (1985 : 445-489). Les amplificateurs, marquant la gradation supérieure, sont répartis en deux classes qui sont les *maximizers* (exemple fourni : *completely*) et les *boosters* (exemple donné : *very much*). Ils envisagent cependant une autre classe d'adverbes qu'ils nomment *emphasizers*

parce qu'ils donnent plus de force à l'adjectif, tout en restant distincts du degré (1985 : 447). Ils s'associent aux adjectifs non graduables, l'effet produit restant similaire à celui des intensificateurs.

L'analyse de corpus de Biber *et al.* révèle l'utilisation plus fréquente des amplificateurs courants en conversation plutôt qu'en prose académique, et précise que, malgré une équitable répartition des amplificateurs en anglais britannique et en anglais américain, *so* est plus fréquent en conversation américaine (1999 : 564). Le tableau ci-dessous montre l'intérêt du marqueur *so* étant donné sa grande fréquence d'occurrence.

Distribution of most common amplifiers (immediately preceding adjectives) in BrE and AmE conversation and in academic prose; occurrences per million words

each • represents 50

	BrE CONV	AmE CONV	ACAD
<i>very</i>	•••••••••• •••••••	•••••••••• ••	•••••••••• ••
<i>so</i>	•••••••••• ••	•••••••••• ••••••••••	•••••
<i>really</i>	••••••••	•••••••••• ••	
<i>too</i>	•••••••	••••••••	••
<i>real</i>	•	••••••••	
<i>completely</i>	•	•	•
<i>absolutely</i>	••	•	
<i>totally</i>	•	••	
<i>damn</i>	•	•	
<i>bloody</i>	••		
<i>extremely</i>			••
<i>highly</i>			••
<i>entirely</i>			•
<i>fully</i>			•

Other amplifiers occurring c. 10 times per million words: *incredibly, perfectly, strongly, terribly.*

Tableau 42 – Fréquence des amplificateurs les plus courants en conversation britannique et américaine et en discours académique (Biber *et al.*, 1999 : 565)

4.2.2 Approche syntaxique de l'adverbe de degré ou d'intensité *so*

En tant que modificateur de degré, l'adverbe *so* est intégré dans le groupe adjectival ou adverbial dont il modifie un élément, celui qui le suit syntaxiquement. Il se place donc en

situation de proximité (Larreya, Rivière, 2010 : 287). Cet élément peut être soit un adjectif, soit un adverbe, ou même un déterminant de quantification (Larreya, Rivière, 2010 : 290). La place canonique de l’adverbe est donc à la tête du groupe adverbial (Biber *et al.*, 2005 : 64). Ce groupe adverbial est en position attribut dans la proposition.

L’adjectif modifié se doit d’être un adjectif graduable. Un adverbe, graduable lui aussi, peut être modifié par un adverbe de degré, mais ce cas est beaucoup moins fréquent, ainsi que le prouvent les analyses de corpus opérées sur le *LSWE Corpus* (Biber *et al.*, 2005 : 546). Huddleston et Pullum soulignent (2002 : 540) que les adjectifs modifiés par *so* ne peuvent fonctionner sous forme d’avant-tête de groupe nominal en tant que complément direct, sauf en tant que prédéterminant devant un article indéfini.

4.2.2.1 Structure « *so* + adjectif ou adverbe »

La structure la plus simple de l’adverbe de degré *so* est celle exemplifiée par Greenbaum (1996 : 51) : *He is so clever. So* se place devant l’adjectif ou l’adverbe qu’il modifie. Quirk *et al.* précisent qu’il peut modifier un adverbe qui, lui-même, peut modifier l’élément qui le suit ([1985] 2005 : 441).

4.2.2.1.1 Structure « *so* » + adjectif + complétive en « *that* »

Ces adjectifs peuvent introduire une proposition ou un groupe de mots pour complément. Les groupes adverbiaux des exemples suivants sont encadrés de crochets (Biber *et al.*, 2005 : 527-529) :

so + adjective + *that*-clause

The murder investigation was [so contrived] **that it created false testimony.**

(NEWS †)

His personality was [so subdued] **that it seemed to fit in with anything he did.**

(FICT)

Biber *et al.* notent que l’adverbe de degré *so* a le rôle d’un intensificateur et porte sur l’adjectif qui le suit, lequel est graduable, formant ainsi un groupe adjectival (entre crochets). Par la notation qu’ils utilisent dans cette structure, c’est ce groupe adjectival, et non simplement l’adverbe *so*, qui déclencherait la proposition subordonnée complétive introduite par *that*, laquelle exprime le résultat d’un tel degré dans la qualité énoncée par l’élément graduable. Le Grézause précise que *that* fonctionne comme une conjonction de subordination (2012 : 38). Ce *that* est en fait facultatif comme le montrent Huddleston et Pullum dans

l'exemple *They were [so small you could hardly see them]* (2002 : 447). La proposition soulignée une fois est introduite sans le *that*, alors que le groupe adjectival *so small* reste nécessaire.

Le tableau ci-dessous présente une distribution des structures comparatives par registre. On peut constater que la structure syntaxique « *so* + adjectif + proposition en *that* » est très courante en fiction, utilisée largement dans la langue journalistique, qu'elle soit britannique ou américaine, assez peu dans la langue académique, et pratiquement pas en conversation britannique, alors qu'elle l'est davantage en conversation américaine.

Distribution of comparative clauses/phrases across registers; occurrences per million words

each ● represents 10

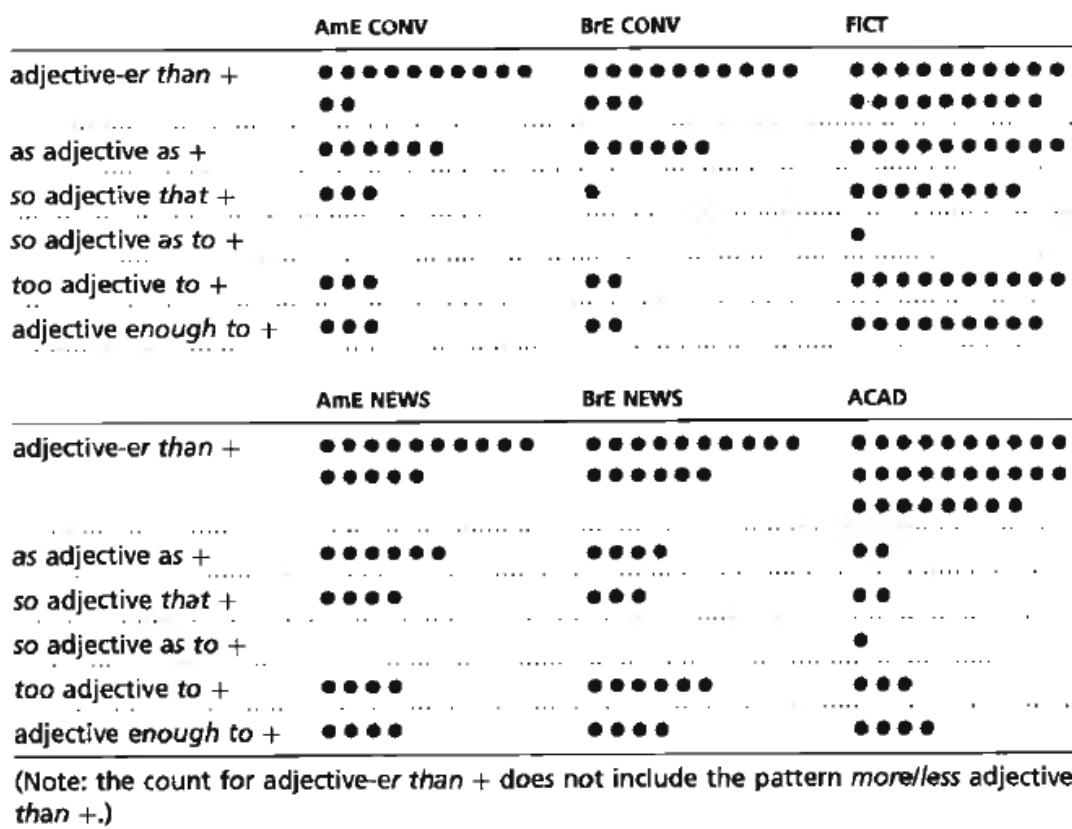


Tableau 43 – Fréquence des constructions comparatives selon les registres (Biber *et al.*, 2005 : 528)

Huddleston et Pullum mentionnent une structure quelque peu hybride, à mi-parcours entre la structure précédente et celle incluant « *so* + adjectif » dans une exclamative avec groupe nominal (2002 : 443). Ils en donnent un exemple :

[so great] a loss that we're likely to go bankrupt [declarative content clause]

De par l'ajout d'un groupe nominal avant ce qui se révélerait dans de nombreux cas une frontière, cette structure est susceptible de modifier les schémas de proéminence de la phrase, surtout si elle est prononcée d'un tempo rapide. Aussi, cette structure, bien que très apparentée à celle de notre corpus, ne sera pas traitée dans cette étude.

4.2.2.2 Structure « *so* » + adverbe + complétive en « *that* »

Les adverbes peuvent avoir des compléments de degré sous forme de propositions ou d'expressions, tout comme les adjectifs. En voici deux exemples fournis par Biber *et al.* (2005 : 550) :

so + adverb + *that*-clause

It happened [*so fast*] that I didn't even realise I had fallen off. (CONV)

He wanted it [*so urgently*] that he fidgeted in his chair. (FICT)

et par Huddleston et Pullum (2002 : 572) :

She spoke [so softly that I couldn't make out what she said].

Contrairement à l'analyse précédente de Biber *et al.*, qui privilégie le groupe adverbial dans son ensemble, Huddleston et Pullum expliquent clairement que ce n'est pas la partie lexicale de tête du groupe adverbial qui déclenche la proposition complétive soulignée une fois, mais l'adverbe de degré *so* lui-même.

Leur recherche sur corpus leur permet d'avancer qu'il est plus fréquent pour un adverbe de modifier un adjectif qu'un autre adverbe (2005 : 546), les groupes adverbiaux *so fast* et *so well* se produisant plus de dix fois par million de mots en anglais américain.

4.2.2.3 Nature syntaxique de « *so* » : adverbe de degré ou intensificateur ?

Boulonnais avance qu'une analyse sémantico-discursive révèle des catégories que ne peut établir une analyse syntaxique, et pour étayer sa thèse, met en lumière les incompatibilités syntaxiques suivantes (2006 : 56-57) :

Les intensifieurs sont souvent confondus avec les adverbes de degré, tels que VERY. Bien qu'ayant également pour cible des adjectifs et adverbes gradués, ils ne répondent pas aux questions en HOW (A : *How tall is he?* B : *Very / Extremely / Rather / *So tall.*). Ils ne peuvent pas non plus être employés seuls. (A : *Is he tall?* B : *Yes, very / extremely / rather / *so.*). [...] De plus, les structures en SO sont difficilement intégrables en position d'épithète ; *a **so reckless** driver est inacceptable, la seule construction possible étant **so reckless** a driver, avec

antéposition du GADJ. La grammaire de *SO* se distingue aussi de celle des adverbes de degré, dont la présence en position interne n'est soumise à aucune contrainte (a **very** tall fellow / a **rather** reckless driver).

Boulonnais prouve ainsi que la syntaxe éloigne *so* de la catégorie des adverbes de degré pour incompatibilité syntaxique. Elle garde, pour le dénommer, le terme d'« intensifieur ». Son éclairage novateur ouvre la voie à une approche fondée sur le sens du discours. J'adopterai la dénomination d'« adverbe d'intensité » en me référant à *so*, tout en tenant compte du fait que la très grande majorité des grammaires se réfèrent à cet adverbe en le nommant « adverbe de degré ». Sur un plan pragmatique, les deux appellations seront donc synonymes dans cette étude, malgré la justesse des observations de Boulonnais.

4.2.3 Approche sémantico-discursive de l'adverbe d'intensité *so*

Biber *et al.* distinguent sept catégories sémantiques principales pour rendre compte de la valeur sémantique des adverbes : adverbes de lieu, de temps, de manière, de degré, additifs ou restrictifs, de modalité (épistémique, d'attitude et de style) et adverbes de liaison ou connecteurs (1999 : 552-560). Ils remarquent que, si les noms et les verbes constituent les classes syntaxiques les mieux représentées dans tous les styles de discours, les adjectifs et adverbes occupent néanmoins une place non négligeable dans tous les registres, bien que les adjectifs soient mieux représentés dans celui de l'écrit, et notamment dans la prose académique, alors que les adverbes sont plus fréquents en conversation et dans les œuvres de fiction (1999 : 504). Ils constatent (1999 : 560) en outre que, selon une répartition sémantique, les adverbes communs de temps, de degré et de modalité sont les plus couramment employés en conversation. Par contre, ce sont les domaines sémantiques du degré et de la liaison qui sont les plus représentés parmi les adverbes courants en discours académique.

Quirk *et al.* définissent l'adverbe de degré comme un dispositif pour identifier, voire évaluer, le degré de la qualité représenté par l'adjectif graduable sur lequel il porte ([1985] 2005 : 441). Quirk *et al.* les nomment « intensificateurs » et affirment qu'ils sont reliés à la catégorie sémantique de degré pour indiquer un point sur une échelle d'intensité abstraite (1985 : 589-90). Parmi eux, les « amplificateurs » se réfèrent à la partie supérieure de l'échelle et se subdivisent en deux catégories : alors que les *maximizers* sont destinés à pointer l'extrémité supérieure de l'échelle, les *boosters* quant à eux dénotent un haut point sur l'échelle, signifiant un haut degré. Ces deux catégories forment des classes ouvertes, surtout celui des *boosters*, le lexique de l'hyperbole devenant rapidement désuet. Ce sont quasiment

les mêmes intensificateurs qui sont utilisés à la fois pour modifier adjectifs et adverbes (Quirk *et al.*, 1985 : 446).

Downing et Locke, quant à eux, considèrent le phénomène d'intensification non pas sous forme de catégories distinctes, mais plutôt comme un continuum dont les éléments pourraient être regroupés dans une quête de visibilité, ce continuum étant constitué par une multitude d'items lexicaux et non pas par une liste fermée de quelques mots ou expressions ([1992] 2006 : 488) :

Intensification is a kind of grading and will be described here in terms of three degrees: 'high', 'medium' and 'attenuated'. They constitute a cline rather than a scale of fixed points, since they are realised exclusively by lexical items rather than by varied structures.

Comme nous l'avons vu ci-dessus, Boulonnais a révélé des incompatibilités syntaxiques entre adverbes de degré et *so* (2006 : 56-57). Elle en dérive son analyse sur la valeur sémantique des intensificateurs dans le discours :

Leur fonctionnement sémantique reposant exclusivement sur la notion de conformité, ils se situent hors du gradient constitué par les adverbes à fonctionnement scalaire du type VERY. [...] Le SO d'intensification signifie la conformité absolue avec un contenu notionnel exprimé par un constituant adjectival ou adverbial. (Boulonnais, 2006 : 56)

Poursuivant sa réflexion, elle approfondit sa thèse en s'appuyant sur une confrontation paradigmatique des éléments de la structure syntaxique, ce qui la conduit à considérer une relation forte entre *so* et des termes se référant à la manière plus qu'au degré :

L'origine de la construction, de même que l'ouverture du paradigme [— + ADJ [ART IND + N]] à des expressions telles que HOW, THIS/THAT et TOO permettent de penser que la classe des adverbes d'intensité est dérivée de la manière et se rapproche par conséquent davantage de cette dernière que de celle du degré. (Boulonnais, 2006 : 57).

Parmi les expressions comparées, on reconnaît le *how* que l'on trouve dans les exclamatives, ce qui présupposerait l'existence d'un réseau de termes emphatiques indépendant des catégories syntaxiques communément utilisées. Ils appartiendraient au domaine de la subjectivité et renverraient directement à la situation d'énonciation. Elle envisage aussi l'analyse de *so* emphatique dans des emplois non standard qui s'éloignent trop de notre corpus pour être abordés dans cette étude.

Si l'on prend en compte l'échelle des modalités de Culioli (Gilbert, 1993 : 93), les structures en *so* relèvent de la modalité de type 3 (modalité appréciative). Le jugement qualitatif du sujet énonciateur S_0 portant sur la validation de la relation prédicative apparaît dans les Phrases 4 et 22 sous la forme d'un « préconstruit » introduit par *that*.

4.2.4 Approche prosodique de l'intensificateur *so*

Downing et Locke notent qu'en anglais parlé, l'expression du degré d'intensité recherché peut être renforcé par l'accentuation et des schémas intonatifs ([1992] 2006 : 488).

Quirk *et al.* font remarquer que le phénomène d'emphase se traduit non seulement par une focalisation informationnelle (*information focusing*), mais peut aussi avoir recours à des procédés permettant de donner de l'emphase purement émotionnelle (*purely emotive emphasis*) ([1985] 2005 : 1414-5). On retrouve ces procédés dans les structures et éléments suivants : exclamatives, *do* persuasif (*persuasive do*) des impératives, interjections, explétifs et intensificateurs, dont les adverbes donnant de l'emphase à la proposition entière (*general clause emphasizees*) tels *actually, really, indeed, etc.*

Quirk *et al.* présentent le *so* caractérisé par sa prosodie *extravagantly emphatic* en conversation familière ([1985] 2005 : 1416). Sur le plan syntaxique, on remarque alors l'absence d'une subordonnée complétive ainsi que la perte du caractère graduable de l'adjectif modifié. Cette force exclamatoire conduit Quirk *et al.* à rapprocher ce *so* non corrélé, des *how* et *what* des exclamatives, ou même des questions et injonctions. Ils en donnent pour exemple l'équivalence suivante :

They were so cross!

How cross they were!

et ajoutent que l'ajout d'une prééminence supplémentaire sur *so* et non seulement sur l'élément modifié accroît encore davantage l'emphase émotionnelle. Ils en donnent cet exemple : *I'm |SÓ PLÉASED|*, et relie ce genre de phénomène à celui des mots à sémantisme émotionnel fort.

Sur un plan plus phonétique, Carr fournit des exemples montrant que le noyau d'un groupe adjectival ou adverbial de degré est normalement porté par la tête du groupe (2013 : 120-121) :

He's so ↘stupid.

He's in'credibly ↘arrogant.

et précise que le noyau peut être porté par l'adverbe de degré pour ajouter de l'emphase :

He's ↗↘so 'stupid !

He's in ↗↘credibly 'arrogant !

La disposition des flèches précise l'orientation de l'intonation sur la syllabe soulignée.

Toujours sur un plan phonétique, Boulonnais remarque que le *so* intensificateur est accentué, mais ne porte pas l'accent de phrase dans la plupart des cas (2006 : 56). Elle note encore que « L'allongement est également caractéristique de l'intensification » et donne pour exemple de cet allongement vocalique la phrase : *haha i love her soo much she sooo awesome and she my big sister... Brittany at gregs!! haha she sooo stupid!! haha that why i love her!* (2006 : 60).

Par contre, dans son étude des formes faibles dans le discours suivi, Cruttenden note que le *so* d'intensification peut être réalisé avec une voyelle faible /sə/ et en donne pour exemple : *ever so (/sə/) many* ([1962] 2014 : 275). Dans le *Longman Pronunciation Dictionary*⁷¹, Wells mentionne cette forme faible :

SO UK  US   səʊ || sou — *There is an occasional weak form sə*

Figure 49 – Entrée lexicale de *so* (*Longman Pronunciation Dictionary*)

Le *Cambridge English Pronouncing Dictionary* est plus explicite sur la question :






SO
UK  US  
normal forms: səʊ,  sou
occasional weak form: sə
so 'long  **so** 'long
Note: Weak-form word. The weak form is used only rarely, and only in casual speech before adjectives and adverbs (e.g. 'Not so bad' /,nɒt.sə'bæd/ /,nɒt-/ , 'Don't go so fast' /,dɒnt.gəʊ.sə'fɑːst/ /,dɒnt.gəʊ.sə'fæst/).

Figure 50 – Entrée lexicale de *so* (*Cambridge English Pronouncing Dictionary - 18th edition*)

Les exemples qu'il fournit montrent que la répartition des proéminences dans le discours est de toute première importance dans ce phénomène de réduction vocalique.

⁷¹ Based on the 18th edition of *Cambridge English Pronouncing Dictionary*. © Cambridge University Press 2011, Version 1.0

Prenant en compte à la fois l'intonation et la durée, l'étude menée par Le Grézause sur le marqueur *so* intensificateur en conversation spontanée montre que, sur son corpus restreint, 58% des marqueurs sont courts, 58% sont diphtongués, et cinq des douze marqueurs, dont la fréquence fondamentale permettait une prise de mesures, avait un contour intonatif plat (100%) (2012 : 154). 40% seulement des marqueurs se démarquent de leur environnement précédent par un ton rehaussé, 20% portent un ton plus bas, et 40% sont de hauteur mélodique semblable. Un quart des marqueurs a une hauteur mélodique supérieure à celle des éléments qui les suivent, 37,5% ont une hauteur mélodique similaire et 37,5% inférieure.

Il ressort de ces études une absence de possibilité modélisatrice de l'intensificateur *so* sur le plan acoustique, bien qu'un surcroît d'emphase émotionnelle soit rendu par une proéminence sur le *so* (Carr et Quirk *et al.*).

4.2.5 Synthèse : Intérêt de l'intensificateur *so* dans l'étude de l'interlangue prosodique

Le corpus de lecture du texte *The Selfish Giant* comprend quatre occurrences du marqueur *so*. L'adverbe de liaison de la Phrase 12 ne sera pas retenu étant donné son incompatibilité avec le pointage de l'emphase. Seuls les marqueurs *so* adverbes de degré, à qualité d'intensificateurs, peuvent être dignes d'intérêt pour cette étude sur l'emphase.

4.2.5.1 Structure « *so* » + adverbe + proposition en « *that* »

Phrase 4 : The birds sat on the trees and sang **so sweetly that** the children used to stop their games in order to listen to them.

La structure syntaxique de cette première occurrence est celle de l'intensificateur *so* suivi d'un adverbe de manière (tête du groupe), formé de l'adjectif *sweet* avec la terminaison *-ly*, introduisant une subordonnée complétive en *that*. Elle survient en début de lecture, alors que quelques phrases ont déjà été lues, participant à l'échauffement et la mise en confiance du lecteur. Elle est insérée au milieu d'une phrase narrative descriptive de 28 syllabes en lecture normale.

Les réalisations attendues sont une proéminence démarquant la paire de son environnement par un ton rehaussé avec une implication supplémentaire du locuteur lorsque l'intensificateur porte une F0 plus élevée.

4.2.5.2 *Structure « so » + groupe adjectival*

Phrase 19 : "I cannot understand why the Spring is **so late in coming**," said the Selfish Giant, as he sat at the window and looked out at his cold white garden; "I hope there will be a change in the weather."

La structure syntaxique de cette occurrence est celle de l'intensificateur *so* suivi d'un groupe adjectival ayant pour tête l'adjectif *late*, suivi d'un groupe prépositionnel. Cette phrase de 48 syllabes normalisées se trouve vers la fin du texte. Elle est composée de deux parties en discours direct séparées par un discours citant de 31 syllabes. C'est dans la première partie de quinze syllabes que se trouve l'intensificateur *so*.

Dans une approche prosodique, il semble nécessaire de noter qu'il s'agit de paroles prononcées par le personnage du Géant Égoïste. La voix devrait donc être bourrue, selon le trait de caractère (*gruff*) introduit dans la Phrase 10, et peut-être en colère car il est contrarié de ne pas voir le printemps se montrer, ou hésitante pour marquer son incompréhension. Un lecteur soucieux d'incarner au mieux la voix du Géant est susceptible d'emprunter un ton grave, des proéminences marquées, ou un allongement de certaines syllabes.

La position contigüe de *is* et de *so* présage des difficultés pour déterminer la frontière entre les deux mots. La marque de voisement attendue pour départager le /z/ du /s/ dans d'autres circonstances devrait être faussée par une assimilation vers la gauche.

De nombreux paramètres peuvent influencer sur la lecture de cette phrase, qui peut révéler l'éventuelle aisance de l'apprenant pour changer sa voix dans l'incarnation du personnage. De nombreux non natifs peuvent la lire sans tenir compte de la qualité attitudinale du message.

4.2.5.3 *Structure « so » + groupe adjectival + proposition en « that »*

Phrase 22 : It sounded **so sweet to his ears that** he thought it must be the King's musicians passing by.

La structure syntaxique de cette occurrence est celle de l'intensificateur *so* suivi d'un groupe adjectival composé de l'adjectif *sweet* précédant un groupe prépositionnel dont il est la tête et introduisant une subordonnée complétive en *that*. Cette phrase narrative descriptive de 22 syllabes normalisées se trouve vers la fin du texte.

De même que pour le groupe adverbial « *so* + adverbe + complétive en *that* », les réalisations attendues sont les mêmes que celles pour l'expression *so sweetly that*.

4.2.6 Marqueurs sélectionnés

Ainsi que l'a montré le tableau de fréquence des amplificateurs les plus courants en conversation et en prose académique (Biber *et al.*, 2005 : 565), l'amplificateur *so* est un marqueur d'occurrence fréquente, et donc susceptible d'être représentatif de la conversation britannique et américaine, ainsi que du discours académique.

Les trois occurrences dans le texte vont permettre de préciser les résultats pressentis, notamment en ce qui concerne les mesures de l'empan de tessiture et la fréquence fondamentale. Le groupe adverbial de la Phrase 19 est susceptible, par sa complexité, d'offrir de nouvelles pistes de réflexion sur de nouveaux paramètres à prendre en considération.

4.2.7 Contextualisation

Dans le texte lu apparaissent trois occurrences de l'adverbe de degré *so* dans trois phrases différentes. Cet intensificateur modifie deux adjectifs (*sweet* et *late*) et un adverbe (*sweetly*) :

P4 - The birds sat on the trees and sang **so sweetly** that the children used to stop their games in order to listen to them.

P19 - "I cannot understand why the Spring is **so late** in coming," said the Selfish Giant [...].

P22 - It sounded **so sweet** to his ears that he thought it must be the King's musicians passing by.

Les phrases P4 et P22 sont interprétées par le narrateur de l'histoire, alors que P19, dans son style direct, incarne la voix du Géant. Des différences de longueur de phrase et de syntagme intonatif peuvent influencer sur les réalisations. Aussi, les segments pertinents pour cette étude (P4 et P22) seront étudiés séparément, et ceux de la Phrase 19 en troisième position, étant donné sa nature interlocutive plus complexe.

4.2.7.1 Résultats et analyse de P4

4.2.7.1.1 Caractéristiques générales

La Phrase 4 est la quatrième phrase du texte et son découpage en syntagmes intonatifs types est le suivant (les nombres entre parenthèses représentent le nombre de syllabes du syntagme intonatif qui les précède) :

The birds sat on the trees (6) and sang so sweetly (5) that the children used to stop their games in order to listen to them.(17)

Le syntagme intonatif étudié sera donc *and sang so sweetly*. À titre de comparaison, les mesures incluront les proéminences adjacentes de part et d’autre du syntagme.

4.2.7.1.2 Mélodie du syntagme intonatif chez 34 lecteurs

Le graphique suivant présente la F0 moyenne de chacune des cinq syllabes du syntagme intonatif S4 pour les 34 lecteurs dont la qualité d’enregistrement était compatible avec l’analyse automatique dans *Praat* pour utilisation du script *ProsodyPro*⁷². On peut y voir *so* doté de la valeur maximale de F0, les autres syllabes portant une F0 décroissante de part et d’autre.

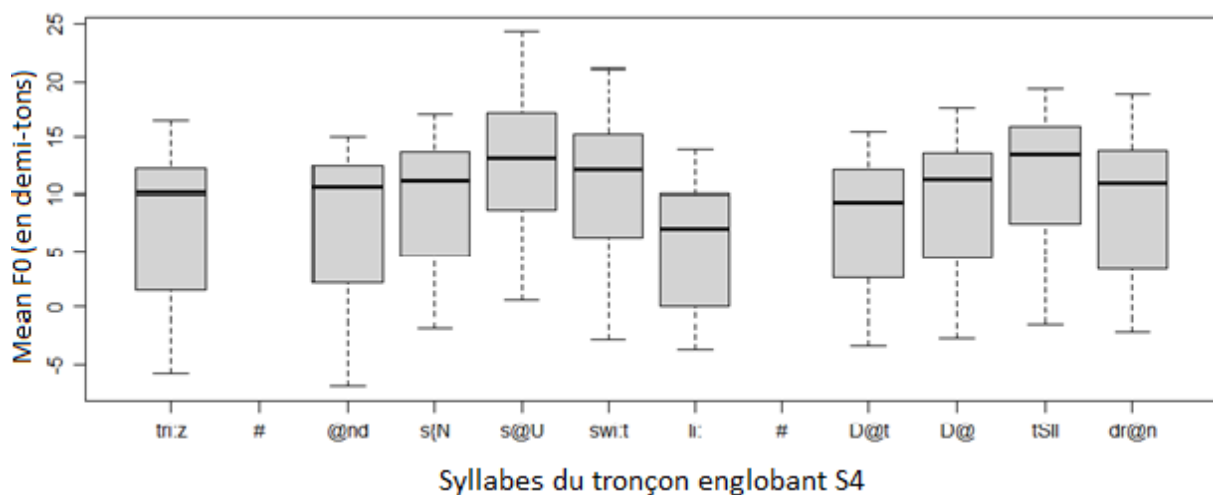


Figure 51 – Valeur moyenne de F0 dans le tronçon englobant S4 (syntagme intonatif de cinq syllabes, 34 lecteurs)

Cette vue uniforme de la totalité des réalisations des 34 lecteurs analysables indique la proéminence de *so* par une F0 supérieure à celle de tous les autres mots, la deuxième position revenant à *sweetly* dont la syllabe accentuée est d’une F0 qui lui est inférieure, mais reste supérieure à celle des autres syllabes du syntagme intonatif. Pour éclairer le phénomène, les 34 fichiers ont été répartis en deux catégories en fonction de la localisation de la valeur maximale de F0 dans le syntagme intonatif : d’une part ceux sélectionnant *so* et de l’autre, ceux privilégiant l’adverbe, c’est-à-dire la syllabe accentuée de ce paroxyton /swi:t/. Cette dichotomie a révélé deux contours mélodiques différents dépeints dans le graphique suivant. La première catégorie de lecteurs (23 lecteurs représentant 67,6% de l’effectif) utilise un empan de tessiture plus large, comme si l’objectif interlocutif était de convaincre, de prendre

⁷² <http://www.homepages.ucl.ac.uk/~uclyyix/ProsodyPro/>, consulté en 2016.

leur auditoire au jeu. La seconde catégorie (onze lecteurs : 32,4%), au contraire, semble plus réservée quant à ses choix interlocutifs : la prééminence mélodique sur le deuxième élément de la paire s'accompagne d'un registre réduit.

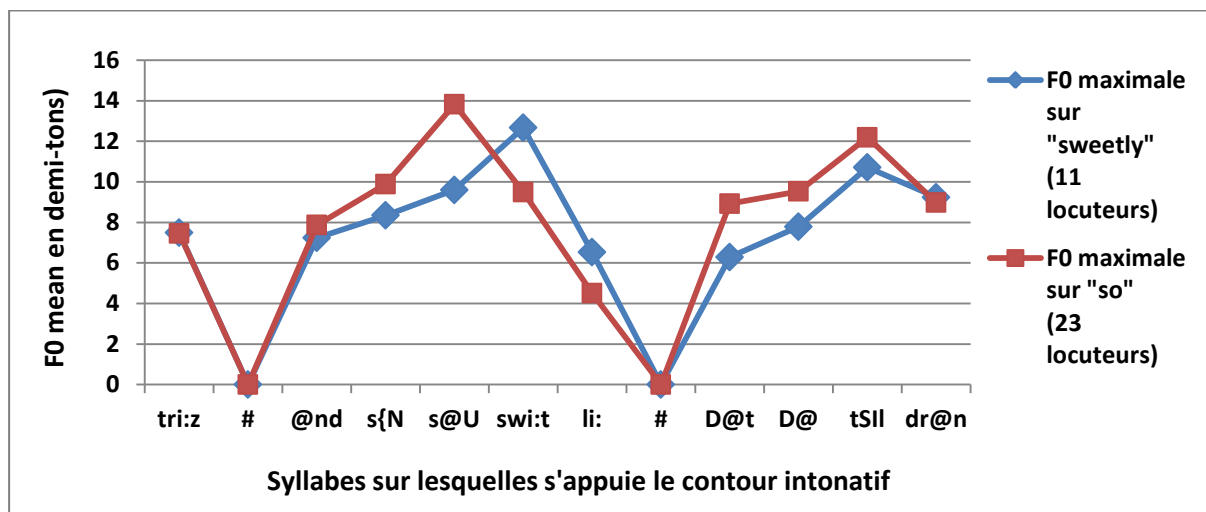


Figure 52 – S4 : Moyenne de F0 dans les contours ayant /s@U/ et /swi:t/ comme valeurs maximales (34 lecteurs)

Une question se pose alors : le choix du contour mélodique dépend-il d'autres traits catégorisables? Quelle serait cette catégorie ? Implique-t-elle un style de parole (Simon *et al.* 2009)⁷³ ? Le tableau suivant précise la répartition des lecteurs suivant leur âge et leur catégorie professionnelle. Le nombre de lecteurs ayant sélectionné *so* pour y placer la moyenne maximale de F0 est quasiment le double de celui ayant choisi *sweetly*. Toutes les catégories sont représentées dans les deux choix. Le pourcentage d'écart entre les choix de chaque catégorie nous renseigne sur leur origine. Les lecteurs matures et professionnels plébiscitent fortement une F0 élevée sur le *so*, la différence d'écart de pourcentage étant respectivement de 50% pour les professionnels et de 63,6% pour les lecteurs matures. Par contre, la différence est relativement bien moins marquée entre les jeunes lecteurs faisant l'un ou l'autre choix, car leur pourcentage d'écart n'est que de 15,8%. Leur grande majorité (57,9%) fait néanmoins le choix de la prééminence mélodique sur *so*. En conséquence de quoi, et à partir de ces seuls éléments, le choix d'une F0 plus élevée sur le *so* devrait être considérée comme étant qualitativement préférable à un choix sur *sweetly*, même si ces deux choix sont acceptables.

⁷³ Simon, A.-C., Auchlin, A., Avanzi, M., Goldman, J.-P. (2009) rappellent la définition de Léon (1993) selon laquelle les styles de parole variés sont « perçus en tant que caractéristiques de circonstances particulières » et dressent un panorama des diverses définitions trouvées dans la littérature.

34 lecteurs	F0 plus élevée sur /s@U/		F0 plus élevée sur /swi:t/		Pourcentage d'écart
	Nombre	%	Nombre	%	%
19 Lecteurs de moins de 40 ans	11	57,9	8	42,1	15,8
11 Lecteurs de plus de 40 ans	9	81,8	2	18,2	63,6
4 Lecteurs professionnels	3	75	1	25	50,0
Total	23	67,6	11	32,4	35,3

Tableau 44 – S4 : Répartition des catégories de lecteurs suivant leur choix mélodique sur *so sweetly*

4.2.7.1.3 Analyse tonale des meilleurs lecteurs

Une investigation plus poussée a été opérée sur les réalisations des lecteurs L1 et L5, que les précédentes expérimentations avaient désignés « bons » lecteurs.

Par l'observation du TextGrid de L1, on remarque qu'il a réalisé la voyelle de /s@U/ avec le ton complexe montant-descendant et celle de /swi:t/ avec un ton simple montant contrastant avec le ton bas mais montant de /li/.

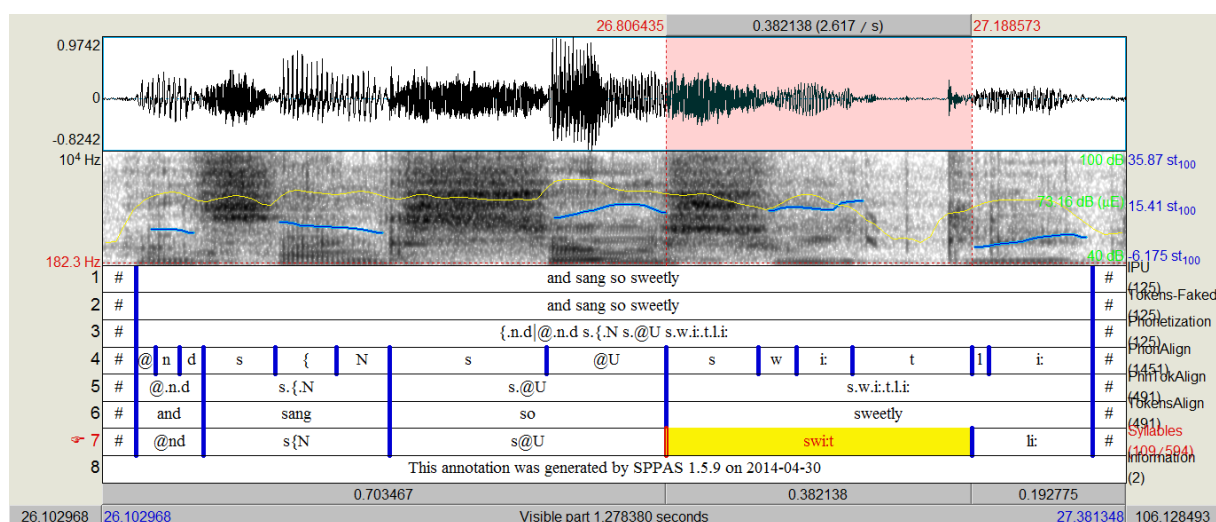


Figure 53 – Textgrid de L1 dans *sang so sweetly* (P4, v1)

L5 a choisi une stratégie plus simple mais inverse : un ton essentiellement montant-descendant sur /s@U/ mais un ton simple descendant sur /swi:t/.

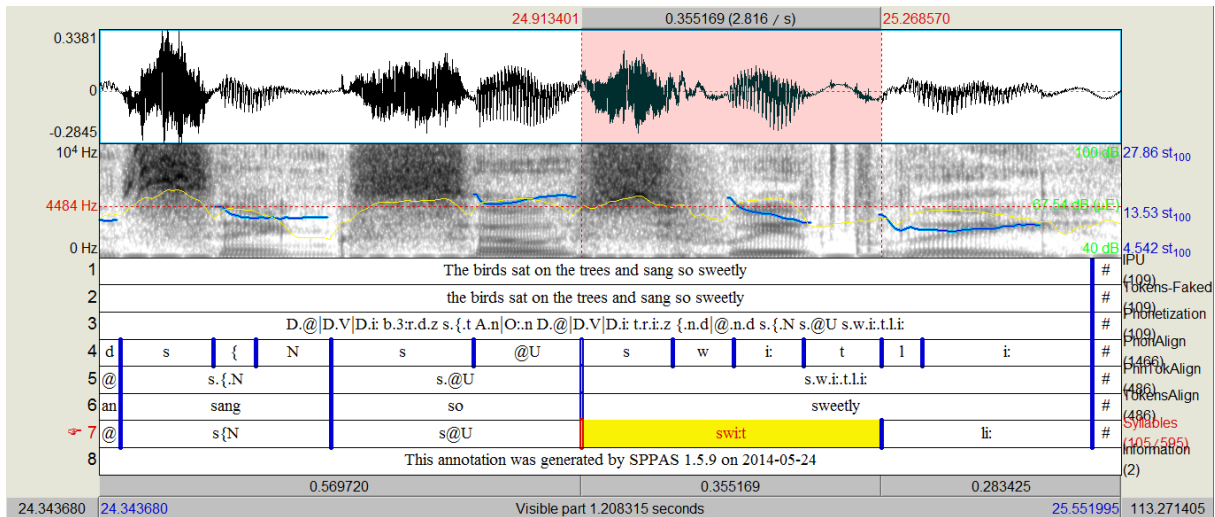


Figure 54 – Textgrid de L5 dans *sang so sweetly* (P4)

Le graphique ci-dessous récapitule et schématise les mesures de F0 aux points stratégiques maximaux ou minimaux suivant la position de début, éventuellement intermédiaire (dans sa partie la plus saillante) ou centrale de la voyelle. Le point commun entre les réalisations de ces deux syllabes est le ton globalement complexe qu’elles adoptent : L1 utilise le ton descendant-montant (\wedge) alors que L5 choisit le ton descendant-montant-descendant (\vee). L’amplitude du ton creusé de L5 est moindre que celle de L1.

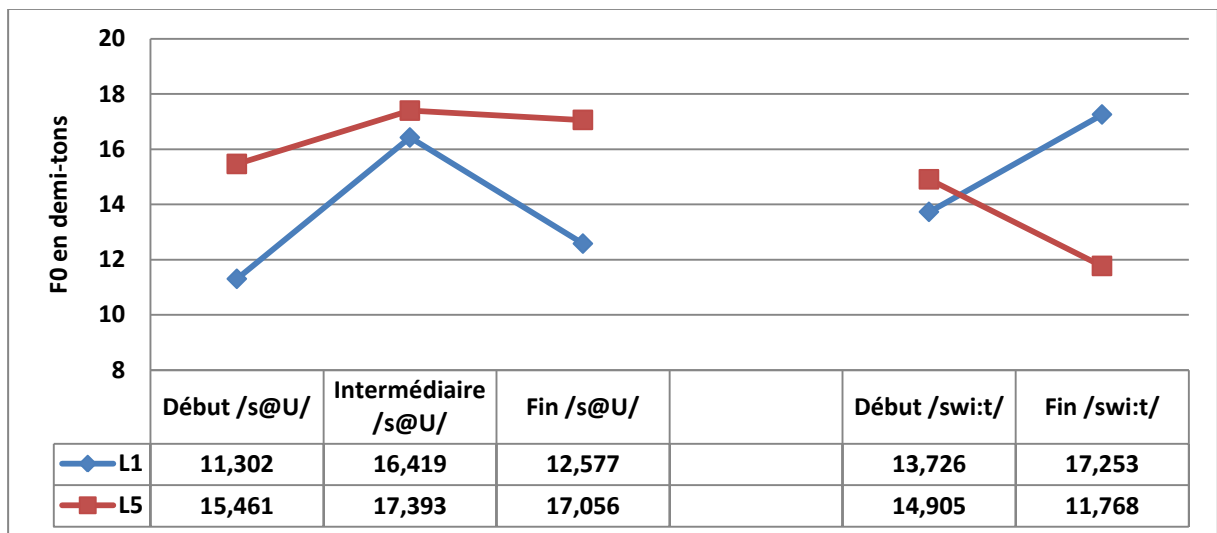


Figure 55 – Schématisation tonale des syllabes /s@U/ et /swi:t/ de P4 par L1 et L5 (mesures sur voyelles)

4.2.7.2 Résultats et analyse de S22

4.2.7.2.1 Caractéristiques générales

La phrase S22 est une des dernières phrases du texte. Son découpage en syntagmes intonatifs types et le nombre de syllabes correspondantes figurent ci-dessous :

4. Étude acoustique des marqueurs d’emphase principaux chez les anglophones natifs (expérimentation COR-AC-N2)

It sounded so sweet to his ears (8) that he thought it must be the King's musicians passing by.(14)

Nombreux sont les lecteurs qui ont marqué une pause après *ears*, mot qui marquera la fin de la partie analysée, dont il sera fait référence sous l’appellation S22.

4.2.7.2.2 Mélodie du syntagme intonatif

Dans le graphique à boîtes à moustaches ci-dessous figure la paire « *so* + adjectif » dans un syntagme intonatif de huit syllabes. On y voit que la syllabe /s@U/ y porte toujours le sommet mélodique le plus élevé en moyenne, car la variation opérée sur cette syllabe est très importante, bien davantage que celle située sur /swi:t/.

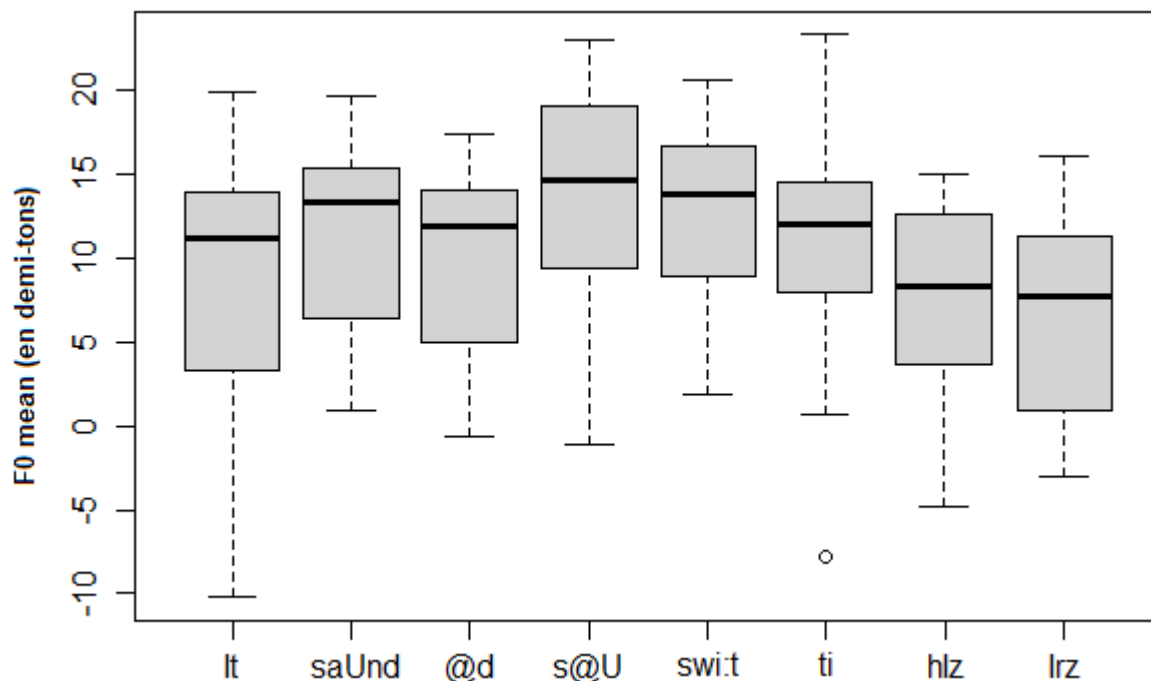


Figure 56 – F0 moyenne dans le syntagme S22 (8 syllabes, 34 lecteurs)

Une fois de plus, le graphique en lignes suivant révèle deux contours au lieu d’un. Les locuteurs ayant utilisé la F0 la plus élevée pour *so* (21 lecteurs représentant 61,8% du corpus du syntagme) ont aussi fait usage d’un large empan de tessiture et tendent à clore leur syntagme par une chute mélodique, alors que les 13 lecteurs (38,2%) ayant doté *sweet* de la F0 la plus élevée ont adopté une voix plus grave avec un empan de tessiture réduit. Le choix d’un contour final ascendant ne concerne que la moitié des lecteurs.

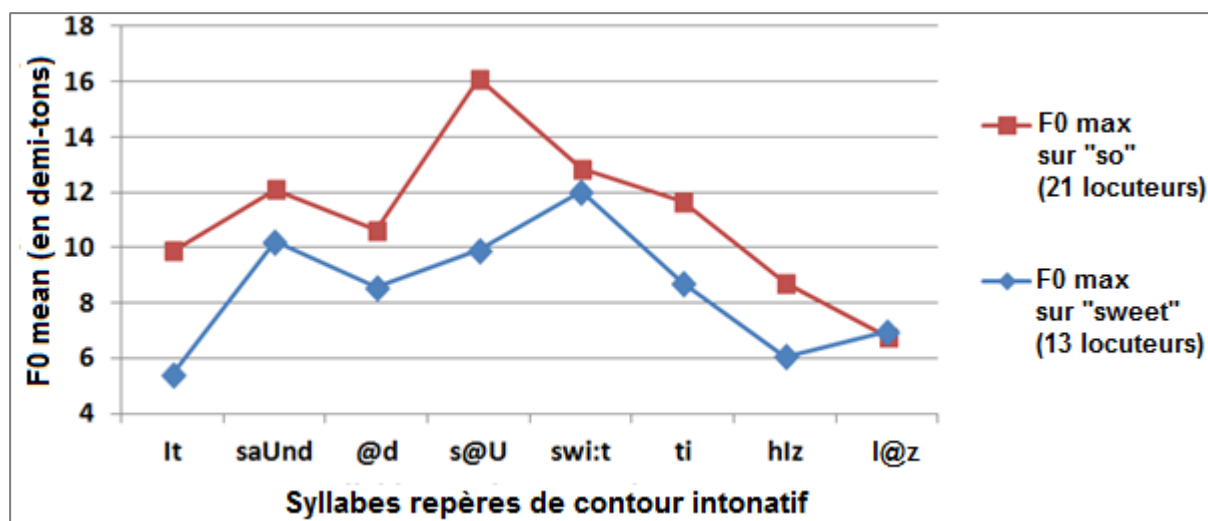


Figure 57 – Moyenne de F0 dans les contours ayant /s@U/ et /swi:t/ comme valeurs maximales dans P22 (34 lecteurs)

Bien que les registres soient respectivement équivalents à ceux mesurés en S4, la différence de hauteur mélodique entre les deux courbes est flagrante. Seule la F0 de *ears* concorde, parce que nombreux sont les lecteurs qui ont à la fois placé une F0 max sur *sweet* et ont utilisé un ton montant en fin de segment pour faire le lien avec la suite du syntagme intonatif pour ceux qui n'avaient pas marqué de pause après *ears*.

Le graphique suivant montre qu'ici aussi, les trois catégories de lecteurs sont représentées dans les deux choix possibles. La constitution des courbes est néanmoins révélatrice d'une préférence marquée des groupes de lecteurs jeunes et surtout matures, leur pourcentage d'écart étant respectivement de 26,3% et 45,5%. Seul celui du groupe de lecteurs professionnels est négatif (-50%). Étant donné que les lecteurs professionnels ne sont que 4, cela ne remet pas en cause les autres résultats, qui donnent une nette préférence au choix d'une F0 élevée sur le *so*, mais interroge sur les raisons de leur choix mélodique.

34 lecteurs	F0 plus élevée sur /s@U/		F0 plus élevée sur /swi:t/		Pourcentage d'écart
	Nombre	%	Nombre	%	%
19 Lecteurs de moins de 40 ans	12	63,2	7	36,8	26,3
11 Lecteurs de plus de 40 ans	8	72,7	3	27,3	45,5
4 Lecteurs professionnels	1	25	3	75	-50
Total	21	61,8	13	38,2	23,5

Tableau 45 – Répartition des catégories de lecteurs suivant leur choix mélodique pour S22 (34 lecteurs)

La sélection de nos deux « bons » lecteurs montre que L1 marque un ton complexe montant-descendant sur la voyelle de *so* d’une amplitude de 6 demi-tons dans sa montée et sa voyelle /i:/ légèrement « bombée » chute de plus de deux demi-tons.

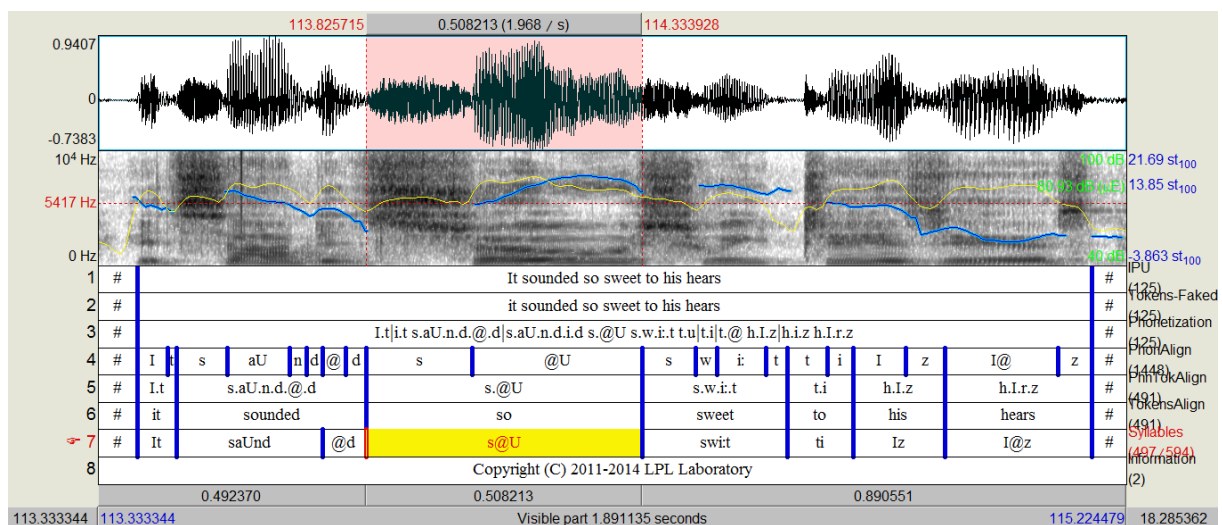


Figure 58 – Textgrid de L1 dans *sounded so sweet* (P22, v1)

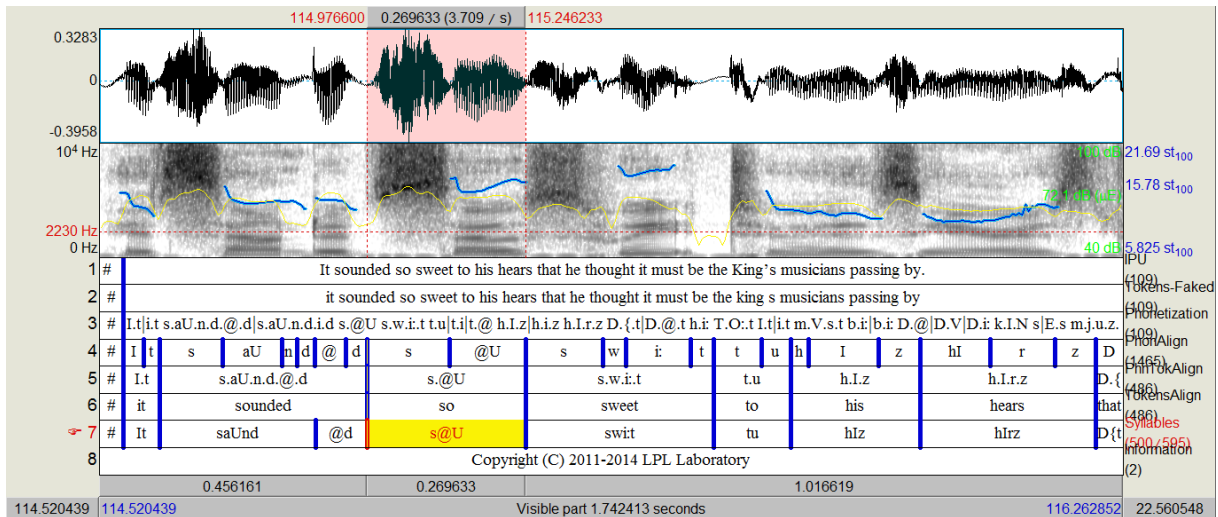


Figure 59 – Textgrid de L5 dans *sounded so sweet* (P22)

L5 module ici aussi les variations de sa F0, mais plus timidement. En raison des micro-mélodies, les extrémités de la ligne de F0 de part et d'autre de la voyelle ne doivent pas être prises en compte pour les mesures mélodiques. La faible amplitude du ton montant-descendant sur *so* et fondamentalement montant sur */i:/*, est plus visible sur la schématisation tonale en demi-tons qui suit.

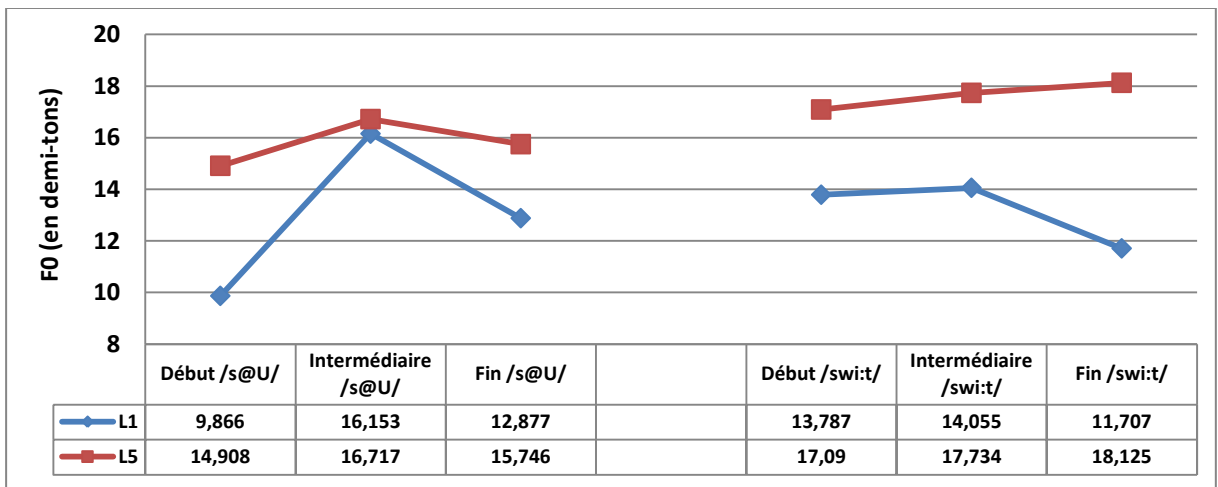


Figure 60 – Schématisation tonale des syllables /s@U/ et /swi:t/ de L1 et L5 dans P22 (mesures sur voyelles)

4.2.7.3 Résultats et analyse de S19

4.2.7.3.1 Caractéristiques générales

La Phrase 19 est longue (48 syllabes), et se trouve au discours direct, entre guillemets, à part un discours citant suivi d'une longue incise :

"I cannot understand why the Spring is so late in coming,"(15) said the Selfish Giant,(6) as he sat at the window (7) and looked out at his cold white garden;(9) "I hope there will be a change in the weather."(11)

La partie étudiée englobera la paire intensificateur + adjectif *so late* ainsi que les prééminences adjacentes, pressenties pour être *why* ou *spring* en contexte gauche, et *coming* en contexte droit, juste avant le discours citant, qui, dans de nombreux cas, établit une frontière de syntagme intonatif. La partie étudiée, même si elle ne cible qu'un syntagme intonatif tronqué, englobe néanmoins neuf syllabes.

Un bon lecteur devra incarner une fois de plus la voix du Géant, qui se parle à lui-même. Aucun état d'esprit n'est précisé dans les indications de lecture, contrairement au discours citant précédent. Le contenu de la phrase permet pourtant de comprendre que le personnage est toujours en colère, et qu'il est étonné et peut-être inquiet de l'absence prolongée du printemps.

4.2.7.3.2 Mélodie du syntagme intonatif

Le graphique ci-dessous illustre le contour mélodique des moyennes de F0 par syllabe sur un long syntagme intonatif dont les neuf dernières syllabes ont été étudiées. Contrairement aux phrases narratives P4 et P22, P19 est au style direct et un bon lecteur se doit d'incarner la voix grave du Géant. Quoi qu'il en soit, toutes les boîtes à moustaches évoquent une grande variation des réalisations, même si la médiane indique que la plupart des lecteurs n'ont pas adopté une F0 basse. Les syllabes marquant une hauteur mélodique sont /kVm/, /sprIN/ et, dans une moindre mesure, /s@U/.

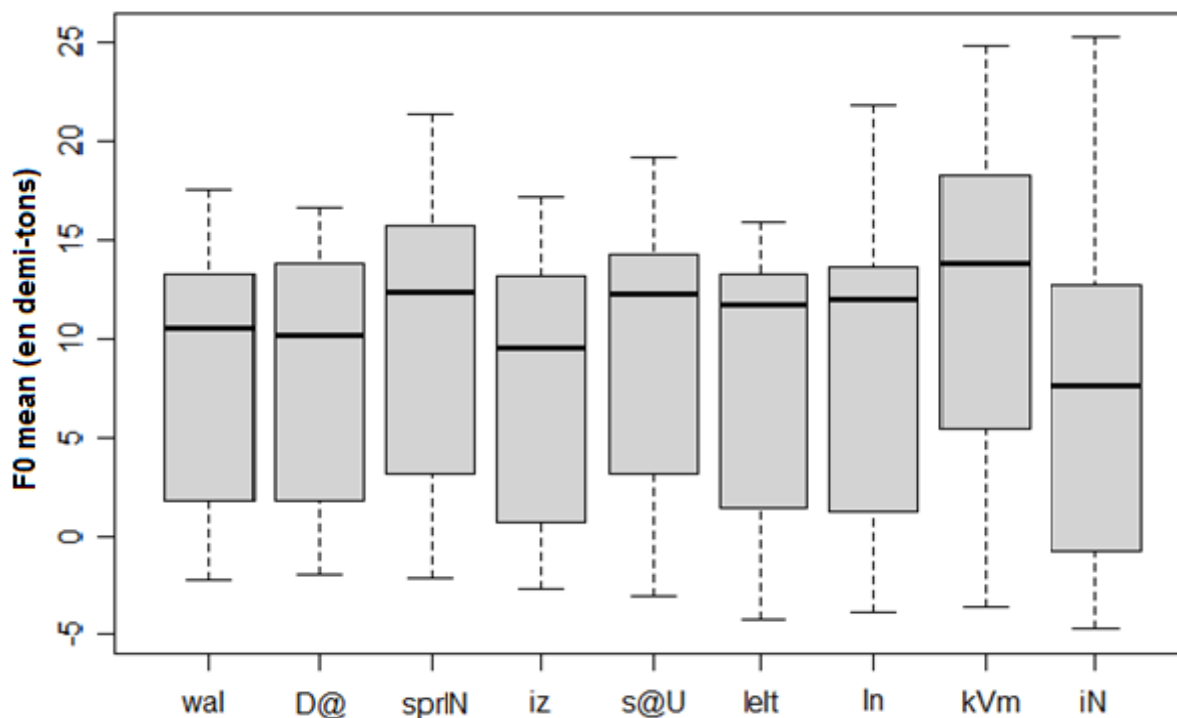


Figure 61 – Contour des moyennes de F0 dans le segment S19 (9 dernières syllabes d'un long syntagme intonatif, 30 lecteurs)

Afin d'analyser ce segment complexe, un graphique en ligne (ci-dessous) représente les quatre catégories de réalisations pouvant aider à comprendre les choix que les trente lecteurs ont été amenés à faire. La ligne A (dix lecteurs, 33,3%) suggère fortement que les proéminences sont portées par les syllabes /sprIN/, /s@U/ et /kVm/ et que la voix utilisée est d'un timbre relativement aigu. Ainsi que l'on pouvait déduire des précédentes analyses empiriques, la ligne C (3 lecteurs, 10%) révèle un timbre plus grave et désigne /sprIN/, /leIt/ et /kVm/ comme syllabes porteuses de proéminences. Les lignes B et D marquent une légère proéminence sur /s@U/, *a fortiori* lorsqu'on la compare à la valeur moyenne de la F0 de /iz/. Elles diffèrent néanmoins par leurs choix stylistiques. La ligne B (onze lecteurs, 36,7%) correspond à un contour mélodique plat et de timbre grave pour incarner la voix du Géant, assorties de très rares proéminences mélodiques puisqu'il aurait fallu élever la F0. La ligne D (6 lecteurs, 20%) correspond au désarroi impatient du Géant en colère : les proéminences marquées par la F0 sont limitées aux syllabes /sprIN/ et /kVm/, laissant /s@U/ et /leIt/ pratiquement hors champ. C'est le contour mélodique marquant les sommets de F0 les plus élevés.

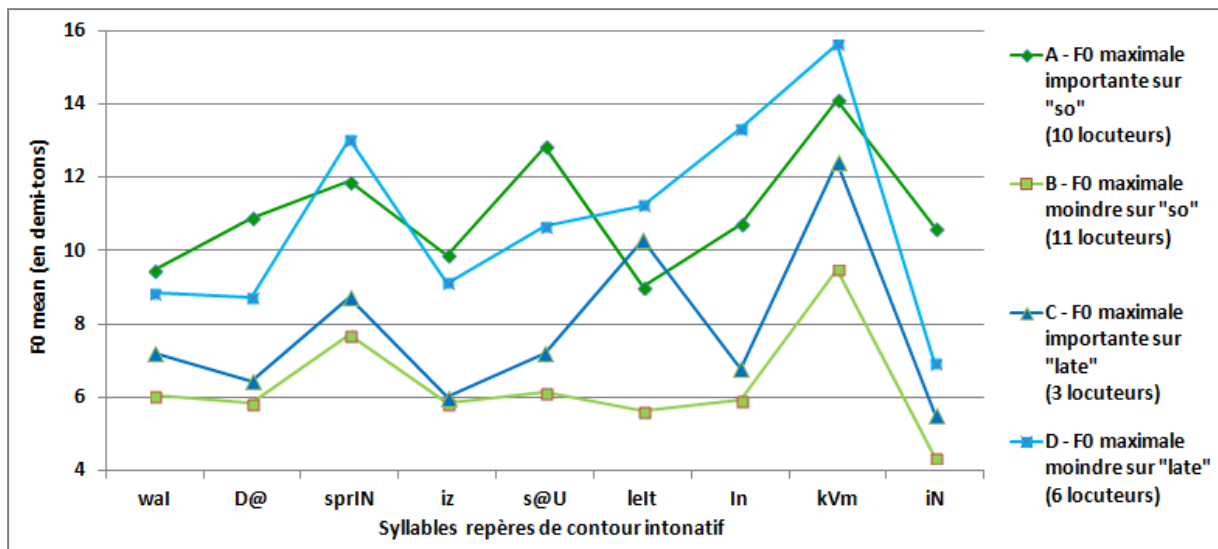


Figure 62 – Segment S19 : F0 moyenne des réalisations avec préominences mélodiques sur /s@U/ et /leIt/ (30 lecteurs)

Pour lire cette phrase, nos deux « bons » lecteurs ont adopté une F0 plus basse que dans les phrases précédentes, surtout L5, dont la F0 n’atteint pas 13 demi-tons. La jeune femme se force au point de rendre sa voix plus grave que celle, masculine, de L1, qui pourtant la baisse aussi. On peut d’ailleurs constater dans les figures suivantes que la courbe de F0 tend à être rectiligne sur l’ensemble des syllabes chez L1 comme chez L5.

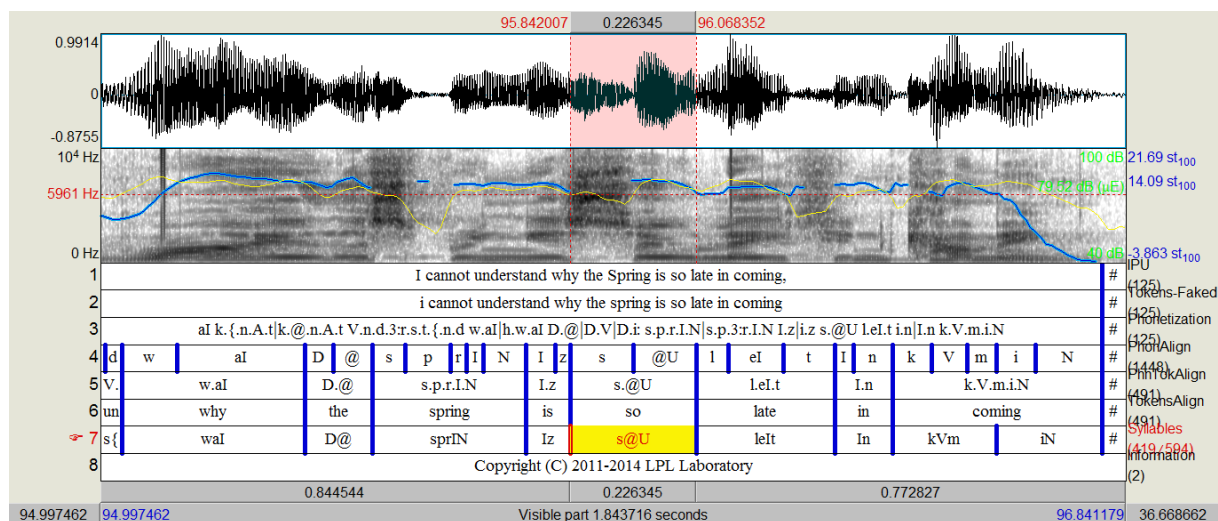


Figure 63 – Textgrid de L1 dans *why the spring is so late in coming* (P22, v1)

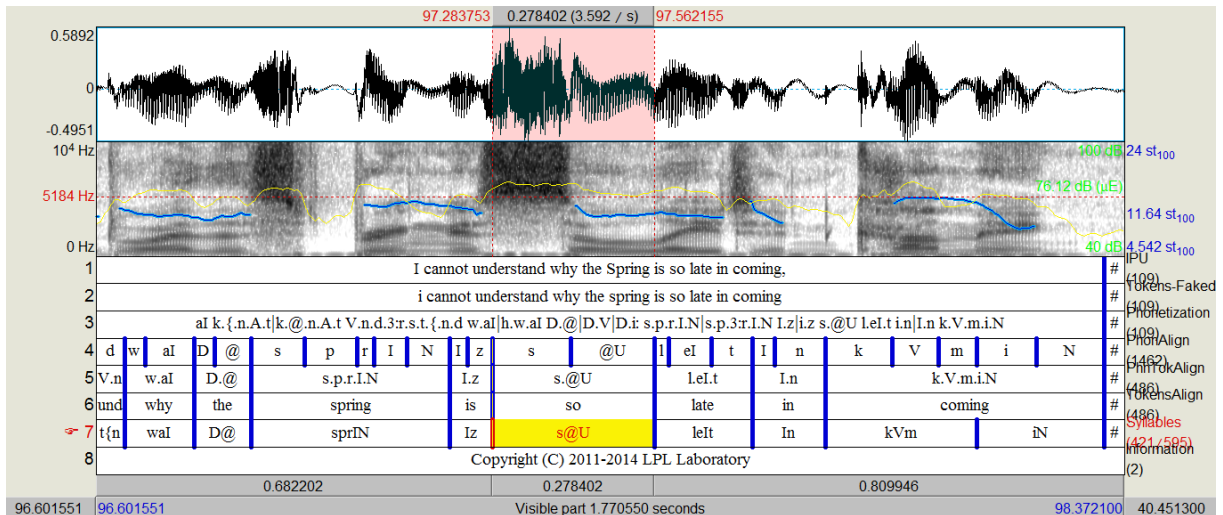


Figure 64 – Textgrid de L5 dans *why the spring is so late in coming* (P22)

Le graphique ci-dessous montre la faible modulation de F0 chez les deux lecteurs. Les tons complexes sur /s@U/ sont peu amples : le ton montant-descendant chez L1 n’atteint pas trois demi-tons, et le ton descendant-montant chez L5, n’atteint pas deux demi-tons. Sur /leIt/, L1 prononce un ton montant-descendant de peu d’ampleur (moins de deux demi-tons) et L5 une chute simple de moins d’un demi-ton.

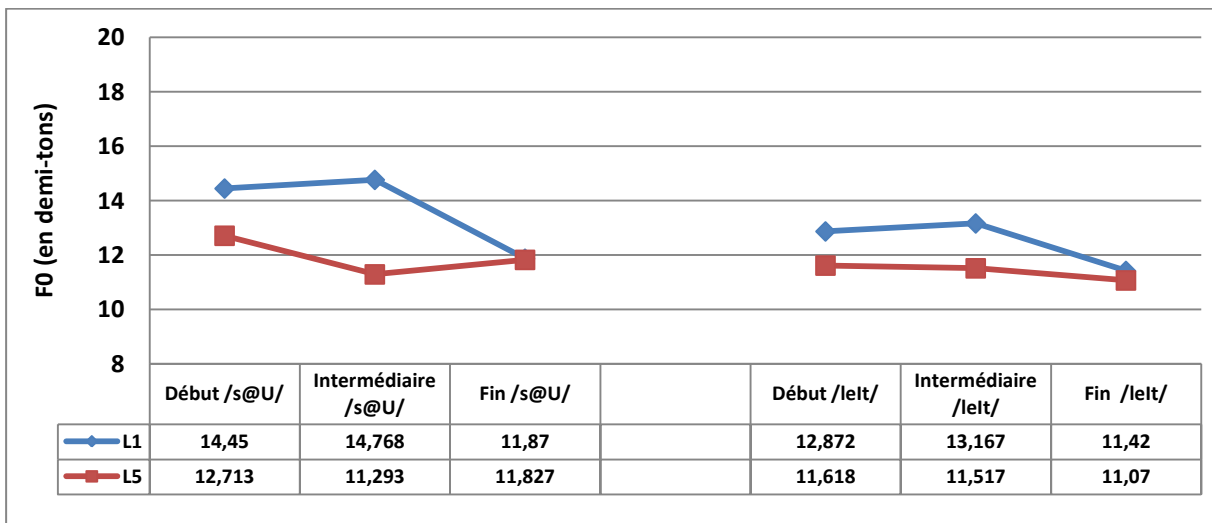


Figure 65 – Schématisation tonale des syllabes /s@U/ et /leIt/ de P19 par L1 et L5 (mesures sur voyelles)

4.2.8 Conclusion

Cette étude sur les trois occurrences de *so* suivi d’un adjectif ou d’un adverbe montre le lien fondamental qui existe entre l’adverbe d’intensité *so* et l’adjectif ou l’adverbe soumis à son action. Dans chacune des trois occurrences, on voit qu’une répartition de la hauteur de F0 est opérée sur la paire « *so* + adjectif ou adverbe » de manière à privilégier l’un des deux termes. Leurs degrés s’équilibrent : le degré d’emphase d’un élément de la paire est d’autant

plus fort que celui adjacent est faible. Il y a équilibre et harmonie des éléments d'emphase entre eux d'une part, et entre ces éléments et ceux qui leur servent d'écrin d'autre part. L'emphase semble alors être un jeu d'échos et de miroirs.

La charge émotionnelle est plus fortement traduite par une hauteur mélodique privilégiant *so* plutôt que le second membre de la paire et est en cohérence avec un registre plus ample que si l'adjectif ou l'adverbe qui le suivent portent la F0 maximale. Cette constatation appuie la thèse de Biber *et al.* selon laquelle c'est l'ensemble des deux termes regroupés en paire qui influe sur la phrase en introduisant sa conséquence. On ne peut ignorer le pouvoir de cette paire sur la phrase entière, qu'elle modifie à son tour. L'introduction de *so*, qui, en synergie avec son binôme, crée un nouvel équilibre et module la qualité interlocutive de la phrase entière, la soumet à de nouvelles contraintes prosodiques. En cela, on constate la forte cohérence de l'ensemble de la prosodie dont quelques réalisations isolées peuvent donner des indications à condition de se donner les outils pour les déchiffrer. Organiser les données en catégories a permis de comprendre et d'interpréter leur signification.

Les catégories présentées ci-dessus n'ont été élaborées qu'à partir de la différence de hauteur mélodique moyenne des paires « *so* + adjectif ou adverbe », lesquelles sont centrales à la compréhension du phénomène d'emphase. Aussi, l'intensificateur *so* en couple avec le terme qui le suit est un marqueur essentiel de qualité prosodique. Je retiendrai les occurrences des Phrases narratives 4 et 22 dans l'étude avec les apprenants. Par contre, bien que très révélatrice sur le fonctionnement de « *so* + adjectif ou adverbe », la Phrase 19 au discours direct est beaucoup trop complexe dans sa nature interlocutive pour être analysée simplement ; aussi, je l'écarterais de l'étude évaluative.

L'étude approfondie de l'intensificateur *so* en contexte a validé l'Hypothèse 2 selon laquelle l'emphase permet d'éclairer le fonctionnement des paires « *so* + adjectif ou adverbe ». C'est aussi l'emphase qui a déterminé ses implications, orienté et justifié la mise à l'écart du *so* de la Phrase 19. En cela, l'emphase devient un principe fondamental dans l'établissement d'un profilage prosodique.

4.3 L'interrogative en *What*

L'interrogative *What are you doing here?* occupe une place privilégiée parmi les marqueurs focalisant l'emphase. La Phrase 10 dans laquelle elle se trouve comprend *What* au 3^{ème} rang des mots les plus emphatiques, *doing* en 5^{ème} position et *here* à la 12^{ème}. De même que précédemment, je vais sonder cette question quant à sa nature linguistique générale par

une approche syntaxique, sémantico-discursive, et prosodique pour évaluer son intérêt dans l'étude de l'interlangue, avant de caractériser son occurrence dans le corpus par sa contextualisation, sa temporalité, sa mélodie, ses proéminences, assimilations ou élisions (expérimentation PER-PROM-N) puis la mélodie des syllabes (expérimentation ANN-TON-N). On remarquera que les deux expérimentations mentionnées sont en marge de l'analyse acoustique par leur protocole mais font partie intégrante de l'étude acoustique des marqueurs d'emphase principaux..

4.3.1 Caractérisation linguistique

Biber *et al.* répertorient (1999 : 204) essentiellement quatre types de questions en anglais. Les trois types de propositions interrogatives indépendantes sont :

- les questions ouvertes, appelées *wh-questions* ou *information questions* en anglais (et questions partielles en français), appellent une réponse informationnelle qui est présentée comme manquante,
- les questions fermées, dites *yes/no questions* en anglais (« polaires » ou « totales » en français), sont destinées à recevoir une réponse en « oui » ou « non » afin de savoir si une proposition est vraie ou non,
- les questions alternatives permettent d'opérer un choix parmi un ensemble fermé d'éléments,
- la quatrième catégorie de questions concerne les question-tags, différente des trois autres par l'usage et la forme.

Biber *et al.* rapportent que, d'après une étude sur corpus⁷⁴ repérant les points d'interrogation, c'est en conversation que l'on trouve le plus de questions (1999 : 211). On en rencontre beaucoup moins en fiction et pratiquement pas dans la langue journalistique et académique. Ils récapitulent dans le tableau ci-dessous (1999 : 212) le fruit de leur étude sur corpus quant à la fréquence et répartition des différents types de questions selon la provenance :

⁷⁴ Il s'agit du *Longman Spoken and Written English Corpus (LSWE Corpus)*, lequel comprend plus de 40 millions de mots écrits et est organisé pour représenter les quatre registres de langue que sont la conversation, la fiction, le discours journalistique et le discours académique (Biber *et al.*, 1999 : 24).

Preference for question type, expressed as a percentage

each ■ represents 5% □ represents less than 2.5%

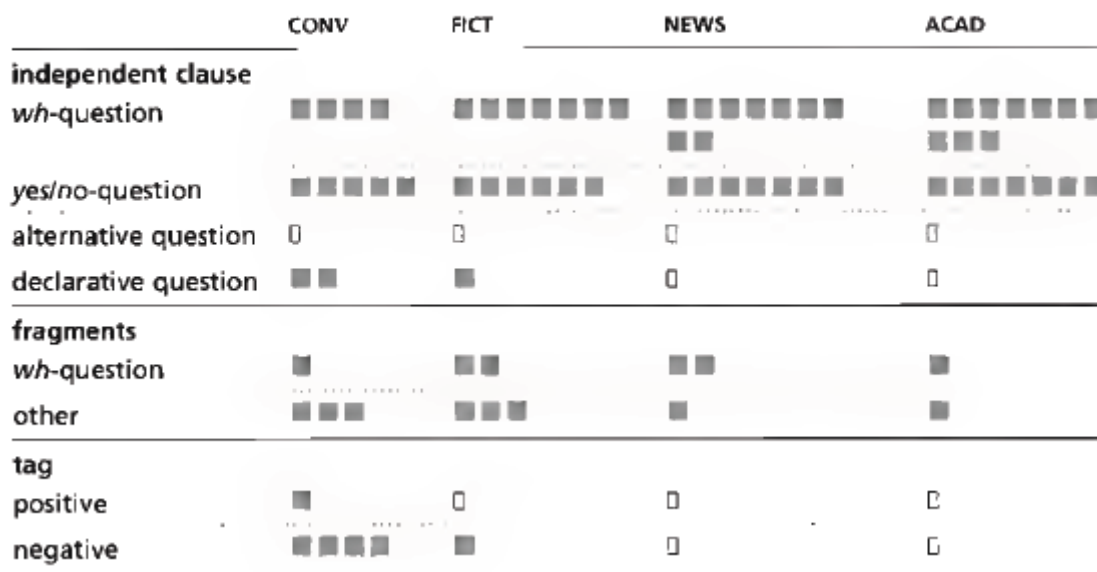


Tableau 46 – Répartition des types de questions selon la nature des corpus (Biber *et al.*, 1999 : 212)

Ces résultats montrent que les questions informationnelles sous forme de proposition indépendante, sont bien représentées dans les quatre catégories de corpus.

4.3.2 Approche syntaxique des questions en *what*

Une question en *wh-* est une question ouverte et est formée avec l’un des mots interrogatifs suivants (Quirk *et al.*, 1985 : 817) : *who / whom / whose, what, which, when, where, how, why*. Greenbaum note que, comme dans l’exclamative, l’élément focalisé est antéposé dans les questions en *wh-* (1996). Biber *et al.* précisent que c’est cet élément qui doit être spécifié à l’interlocuteur, le reste de la proposition reprenant des éléments qui lui sont connus (1999 : 204). L’inversion du sujet et de l’opérateur n’intervient que dans le cas où le mot en *wh-* a la fonction de sujet du verbe de la question.

Quirk *et al.* remarquent qu’il n’existe pas de mot en *wh-* destiné à faire porter la question sur une action, c’est pour cela que *what* est utilisé pour questionner le contenu d’une prédication dans la question (1985 : 818). Biber *et al.* précisent que *what* est un marqueur de proposition interrogative, qui aura la fonction de pronom et non de déterminant (1999 : 87).

4.3.3 Approche sémantico-discursive des questions en *what*

Contrairement aux questions fermées, où seule la polarité est mise en question (Downing et Locke, [1992] 2006 : 183), les questions en *wh-* sont censées fournir une réponse

contenant l'information ciblée par le mot en *wh-* (Greenbaum, 1996 : 47). Quirk *et al.* précisent la nature positive de la question en *wh-*, dans la mesure où une présupposition de réponse est à l'origine de la question (1985 : 819-20-21). Ils envisagent une orientation négative si l'objectif de la question n'est pas de recevoir une information concernant le présupposé du mot en *wh-*. L'élément présupposé (*pushdown element*) est fondamental dans ce type de questions. En outre, Quirk *et al.* notent le style neutre de cette question, étant donné que le mot en *wh-* n'est pas précédé d'une préposition, laquelle catégoriserait la question dans un registre plus formel (1985 : 817).

Quirk *et al.* soulignent que certaines questions, dites questions rhétoriques, possèdent la structure des interrogatives tout en ayant la vigueur d'une forte assertion (1985 : 825) :

The rhetorical question is interrogative in structure, but has the force of a strong assertion. It generally does not expect an answer.

A positive rhetorical *yes-no* question is like a strong negative assertion, while a negative question is like a strong positive one.

Ils précisent que cette particularité s'étend aux questions rhétoriques en *wh-* (1985 : 826) :

The positive question is equivalent to a statement in which the *wh-* element is replaced by a negative element : [...]

What makes YÔU think you can do better? ['Nothing should make you think you can do better.']

En effet, notre interrogative pourrait être paraphrasée par : *You've got NOTHING to do here.* Biber *et al.* confirment que les questions rhétoriques expriment une opinion plus qu'une recherche d'information (1999 : 206) et ajoutent :

By choosing an interrogative form, the speaker appears to let the addressee be the judge, but no overt response is expected. This is therefore a type of question that can just as well occur in monologue as in dialogue.

What, en tant que pronom interrogatif, est non personnel (Huddleston et Pullum, 2002 : 429) et correspond au complément du verbe attendu syntaxiquement.

Il semblerait que la référence en matière de diversité de réalisations, tant prosodiques que sémantiques, soit celle d'O'Connor et Arnold (1961 : 46-90). Le tableau ci-dessous présente les dix grands contours mélodiques s'appliquant aux questions en *wh-* et résume, dans les propres termes de ces phonéticiens, le sémantisme qui en découle.

4. Étude acoustique des marqueurs d'emphase principaux chez les anglophones natifs (expérimentation COR-AC-N2)

CHAPITRE III – Emphase : mode d'organisation paradigmatique des réalisations prosodiques des anglophones natifs

	<i>Contour mélodique</i>	<i>Explications sémantiques</i>
1	THE LOW DROP	With the Low Drop these questions sound <i>searching, serious, intense, urgent</i> , because of the power that the tone group carries. This power may again be used to reinforce both approval and disapproval. Extra power can be conveyed by these questions if the special finite is accented, rather than the interrogative word. [...]
2	THE HIGH DROP	The High Drop is probably the most common way of asking these questions. It avoids the seriousness and urgency of the Low Drop, and such questions sound <i>brisk, businesslike, considerate, not unfriendly</i> . [...] If there is no head and the High Fall nuclear tone occurs on the <i>wh</i> -word, there is no detachment of flatness as with the Low Drop. On the contrary, the questions sound <i>bright and interested</i> .
3	THE TAKE-OFF	When the nucleus is the interrogative word the effect may be either of repeating the listener's question or of asking for information to be repeated. In both cases the questioner's tone is <i>wondering</i> , as though he was mildly puzzled that such a question should have been asked or that he should have been given the information he was given. It is fairly rare to ask any but the above repeated type of <i>wh</i> -question with the Take-Off; any other sounds very calm but very <i>disapproving and resentful</i> .
4	THE LOW BOUNCE	By using the Low Bounce with <i>wh</i> -questions the speaker seeks to establish a bond with the listener, to show interest not only in receiving the information asked for but also in the listener himself. [...] it is a very common way of asking these questions of young children. Among adults too it is often used for an opening question, when the speaker wants to make it absolutely clear that his enquiry is a friendly one, not an attempt to pry or to criticise. [...]
5	THE SWITCHBACK	In echoed questions, whether of the <i>wh</i> - or the yes-no kind, the effect of the Switchback is of <i>astonishment</i> , as if the speaker can hardly believe his ears. The Switchback is also used to make corrections to questions, as to statements.
6	THE LONG JUMP	These give much the same effect as statements; the speaker is asking about something very unexpected to him and perhaps not very pleasing. The <i>protest</i> is still very evident.
7	THE HIGH BOUNCE	When the nuclear tone is on the interrogative word, the High Bounce calls for the <i>repetition of information already given</i> , as does the Take-Off, but the wondering, puzzled flavour of the Take-Off is absent. When the nuclear tone is not on the interrogative word, the speaker is often <i>echoing</i> the listener's question in order to get it clear in his mind before giving an answer; again there is no criticism implied as there is with the Low Bounce. [...] The High Bounce is also used in straightforward <i>wh</i> -questions, that is, not echoes or requests for repetitions; and such questions sound rather like those with the Low Bounce, but very much more <i>tentative and casual</i> , as if to avoid the appearance of prying.
8	THE JACKKNIFE	The Jackknife gives to these questions a note of <i>challenge and antagonism</i> , which is usually equivalent to the word <i>but</i> placed before the question or the word <i>though</i> after it. As with statements, there is often a <i>disclaiming of responsibility</i> for the situation.
9	THE HIGH DIVE	The use of the High Dive with questions of any kind is unusual. When it occurs, the High Fall is normally placed on the <i>wh</i> -word or the special finite, and the effect is of considerable emotion. This emotion may take the form of <i>plaintiveness, despair</i> or the like. Or it may be a matter of <i>gushing warmth</i> . [...]
10	THE TERRACE	The only common use for the Terrace is for <i>non-final</i> word groups; [...] this tone group is readily used to show non-finality with all five sentence types.

Tableau 47 – Récapitulatif des 10 structures tonales appliquées aux questions en *wh*- et l'explication sémantique correspondante (d'après O'Connor et Arnold, 1961 : 49-89)

4.3.4 Approche prosodique des questions en *what*

4.3.4.1 Mélodie

Quirk *et al.* précisent l'intonation descendante de la plupart des questions en *wh-* en stipulant (1985 : 817) :

[c] In a collection of 858 *wh*-questions from the files of the Survey of English Usage, chiefly in surreptitiously recorded spoken unscripted material, 775 had falling intonation.

Quirk *et al.* ajoutent : *Wh-questions generally have a rise-fall tone, less commonly a simple falling tone* (1985 : 826). Wells mentionne aussi que le ton par défaut des questions en *wh-* est une chute mélodique (2006 : 42) :

The default tone for *wh* questions is a **fall**. As with statements, this tone meaning is the **definitive fall** [...].

Il nuance cependant son approche en reconnaissant la possibilité d'une réalisation mélodique pourvue d'une remontée, qu'elle soit simple ou complexe (2006 : 43) :

Nevertheless, a *wh* question can also be said with a non-fall: a rise or, less commonly, a fall-rise. This has the effect of making it *more gentle*, kindly, encouraging, sympathetic or deferential, as opposed to the businesslike fall. We call this tone meaning the **encouraging rise**.

Cruttenden précise quels différents contours intonatifs peuvent adopter les questions en *wh-* selon la précision de la tradition britannique ([1962] 2014 : 295). Le ton par défaut est la chute intonative, qu'elle soit *low fall* ou *high fall*. Un ton montant étant perçu comme moins catégorique, le ton alternatif à cette chute est le *low rise*, et même le *high rise* en anglais américain standard. Le *high rise* pourra être utilisé pour demander à ce que la question soit répétée.

O'Connor et Arnold répertorient les occurrences intonatives des questions en *wh-* et codent leurs réalisations selon la tradition britannique (1961 : 41-45). Le tableau ci-dessous présente les 32 plus grandes catégories intonatives de ces réalisations, dont la signification a été présentée précédemment.

CHAPITRE III – Emphase : mode d'organisation paradigmatique des réalisations prosodiques des anglophones natifs

	<i>Contour mélodique</i>	<i>Description</i>
1	THE LOW DROP (UNEMPHATIC)	(Low Pre-head +) (High Head +) Low Fall
	THE LOW DROP (EMPHATIC)	High Pre-head + (High Head +) Low Fall <i>or</i> (Pre-head +) Stepping Head + Low Fall
2	THE HIGH DROP (UNEMPHATIC)	(Low Pre-head +) (High Head +) High Fall
	THE HIGH DROP (EMPHATIC)	High Pre-head + (High Head +) High Fall <i>or</i> (Pre-head +) Stepping Head + High Fall <i>or</i> (Pre-head +) High Fall(s) + High Fall
3	THE TAKE-OFF (UNEMPHATIC)	(Low Pre-head +) (Low Head +) Low Rise
	THE TAKE-OFF (EMPHATIC)	High Pre-head + Low Head + Low Rise
4	THE LOW BOUNCE (UNEMPHATIC)	(Low Pre-head +) High Head + Low Rise <i>or</i> High Pre-head + Low Rise
	THE LOW BOUNCE (EMPHATIC)	High Pre-head + High Head + Low Rise <i>or</i> (Pre-head +) Stepping Head + Low Rise
5	THE SWITCHBACK (UNEMPHATIC)	(Low Pre-head +) (Falling Head +) Fall-Rise
	THE SWITCHBACK (EMPHATIC)	High Pre-head + (Falling Head +) Fall-Rise <i>or</i> (Pre-head +) Sliding Head + Fall-Rise <i>or</i> (Pre-head +) High Fall(s) + Fall-Rise
6	THE LONG JUMP (UNEMPHATIC)	(Low Pre-head +) Rising Head + High Fall
	THE LONG JUMP (EMPHATIC)	High Pre-head + Rising Head + High Fall <i>or</i> (Pre-head +) Climbing Head + High Fall
7	THE HIGH BOUNCE (UNEMPHATIC)	(Low Pre-head +) (High Head +) High Rise
	THE HIGH BOUNCE (EMPHATIC)	High Pre-head + (High Head +) High Rise <i>or</i> (Pre-head +) Stepping Head + High Rise
8	THE JACKKNIFE (UNEMPHATIC)	(Low Pre-head +) (High Head +) Rise-Fall
	THE JACKKNIFE (EMPHATIC)	High Pre-head + (High Head +) Rise-Fall <i>or</i> (Pre-head +) Stepping Head + Rise-Fall
9	THE HIGH DIVE (UNEMPHATIC)	(Low Pre-head +) (High Head +) High Fall + (Low Accents +) Low Rise
	THE HIGH DIVE (EMPHATIC)	High Pre-head + (High Head +) High Fall + (Low Accents +) Low Rise <i>or</i> (Pre-head +) Stepping Head + High Fall + (Low Accents +) Low Rise
10	THE TERRACE (UNEMPHATIC)	(Low Pre-head +) (High Head +) Mid-Level
	THE TERRACE (EMPHATIC)	High Pre-head + (High Head +) Mid-Level <i>or</i> (Pre-head +) Stepping Head + Mid-Level

Tableau 48 – Récapitulatif des dix structures tonales appliquées aux questions en *wh-* avec leur description mélodique (d'après O'Connor et Arnold, 1961 : 41-45)

4.3.4.2 *Accentuation*

Huart rappelle le choix décisif de l'énonciateur dans la sélection des éléments portant l'accent mélodique, lesquels « renvoient aux termes mis en relation à l'occasion de l'énoncé » (2010 : 147). L'énonciateur spécifie par le mot interrogatif le domaine où l'information fait défaut :

Ces termes sont en général accentués dans les questions parce qu'ils marquent un choix à faire, une case à remplir. En les accentuant, l'énonciateur invite son interlocuteur à spécifier un terme indispensable à la relation [...]

4.3.4.3 *Réduction et assimilation*

Huart souligne que les questions de la langue orale comprennent souvent des réductions des auxiliaires verbaux, même si la langue écrite n'en rend pas compte (2010 : 148-149). De même, un phénomène d'assimilation peut renforcer celui de réduction :

En outre, le rapprochement des formes réduites des auxiliaires avec le pronom sujet YOU inaccentué aboutit souvent à des condensations supplémentaires, car le phonème /j/ au début de ce pronom occasionne fréquemment la palatisation d'un /d/ ou d'un /t/ précédent [...]

What are you talking about?

[ˈwɒt (ə) ju ˈtɔːkɪŋ əbaʊt] → [ˈwɒt jʊ ˈtɔːkɪŋ əbaʊt] → [ˈwɒtʃʊ ˈtɔːkɪŋ əbaʊt]

Cruttenden note que les mots « outils » ont souvent deux formes : l'une forte, et une alternative plus faible dans laquelle opèrent des réductions ([1962] 2014 : 275). Alors bien même que certains mots « outils » ne possèdent pas de formes faibles, ils peuvent néanmoins être l'objet de réductions dans la parole rapide ou en discours familier et se limitent à un nombre restreint d'expressions du type présentement étudié, qu'il donne d'ailleurs en exemple :

What are you doing? /wɒt ə jə ˈduːɪŋ/ or even /wɒtʃə ˈduːɪŋ/

Ainsi réductions et assimilations sont un phénomène naturel et courant dans le parler des anglophones natifs. Ils sont la contrepartie de la focalisation accentuelle.

4.3.5 **Synthèse de l'étude linguistique générale : Intérêt de la question en *what* dans l'étude de l'interlangue prosodique**

La Phrase 10 comprend la seule interrogative du texte, ce qui, d'un point de vue syntaxique est très intéressant pour vérifier sa réalisation acoustique dans l'interlangue. De plus, l'étude sur corpus de Biber *et al.* montre que l'étude de la question en *wh-* indépendante a bien toute sa place dans la présente étude étant donné qu'elle est bien représentée dans les quatre grandes catégories de corpus (1999 : 212).

Cependant, la question en *what* ne peut pas être représentative de toutes les questions étant donné que les questions polaires (*yes-no questions*) se concrétisent par un ton montant par défaut que Wells nomme le *yes-no rise* (2006 : 45). Il arrive néanmoins qu'une question polaire épouse une mélodie tombante, ce qui lui donne une coloration plus insistante, plus professionnelle, plus sérieuse et sans doute plus menaçante (Wells, 2006 : 46).

De plus, il ne s'agit pas d'une vraie question, c'est-à-dire une question appelant une réponse. Le sens de cette question rhétorique est éclairée par le contexte. Elle exprime la colère du Géant, furieux de trouver des enfants dans son jardin. Il s'agit donc davantage d'une réprimande que d'une question, ainsi que l'indique la fin de la phrase puis la phrase suivante. L'effet de la question (et de la voix bourrue avec laquelle elle est prononcée) est de faire fuir les enfants. Le personnage du Géant fait ensuite valoir son titre de propriété en clôturant son jardin d'un haut mur, en y plaçant un écriteau dissuasif et en exprimant au discours direct son refus de partage.

Malgré sa force illocutoire, qui détourne la présupposition de cette question pour en faire un reproche, cette phrase est cependant marquée par un haut degré d'emphase et, ne serait-ce qu'en raison de cet aspect, il semble intéressant d'examiner la manière dont elle est rendue, que ce soit par les natifs ou les non natifs. Il est en outre probable qu'un bon nombre de non natifs ne se sentiraient pas assez confiants au cours de leur lecture pour faire valoir un changement de voix en adoptant une voix très bourrue (*very gruff voice*), peut-être en est-il de même chez quelques natifs. Dans ce cas, cette question redeviendrait une question ordinaire en *what*, qui aurait toute sa place dans cette étude.

Réductions et assimilations mettent davantage en relief les zones d'emphase. Elles sont le signe d'une prise de distance par rapport à la nature isosyllabique de la L1 dans une prise en compte de la qualité rythmique de la langue cible. Il serait donc intéressant de vérifier les réalisations des apprenants dans ce domaine.

Malgré son apparence commune et son fort degré de probabilité dans une conversation courante, cette question, de par sa situation contextuelle et sa force interlocutoire et emphatique, devrait *a priori* fournir des réalisations riches et complexes de grand intérêt, propices à départager les réalisations des apprenants.

4.3.6 Contextualisation

Le texte de la Phrase 10 comprend la question suivie d'un énoncé citant : "*What are you doing here?*" *he cried in a very gruff voice, and the children ran away.*

Il est fondamental de se pencher sur ces instructions de lecture (didascalies) en vérifiant la signification de l'adjectif *gruff*. Dans ses définitions⁷⁵, les termes correspondant aux sonorités sont : *low* (2 occurrences) ou *low in pitch* (une occurrence), *rough* (deux occurrences) et *deep* (une occurrence). *Low* et *low in pitch* se réfèrent à une F0 basse. *Rough*, quant à lui, dans les définitions se référant à la voix, correspond à *hard and loud, harsh and rasping* et *harsh*. *Loud* appartient au domaine de l'intensité. *Harsh* réunit un consensus avec les termes *too strong, strong effect* et *great strength or force* autour de l'intensité lui aussi. *Deep* confirme la F0 basse avec *low, low in pitch* et *low musical pitch or pitch range*. Cette dernière expression signifiant sans doute une réduction de la tessiture. L'expression *full in tone*, elle, renforce la notion d'intensité.

Un bon lecteur (ainsi qu'un bon apprenant) se doit alors d'imiter la voix bourrue du Géant, c'est-à-dire d'adopter une voix à la F0 basse par rapport au registre habituel de sa voix, sans doute de réduire sa tessiture, et d'augmenter l'intensité de sa parole. C'est ainsi qu'il se conformera aux didascalies précisées dans l'énoncé citant et qu'il pourra être crédible en exprimant la colère du Géant propre à faire fuir de peur les enfants qu'il découvre dans son jardin.

4.3.7 Temporalité

Dans le Graphique ci-dessous figurent les boîtes à moustaches de la répartition de la durée moyenne par syllabe des réalisations de 37 lecteurs. Ont été écartés de l'analyse les fichiers dont certaines syllabes étaient assimilées ou qui présentaient des pauses. Les fichiers bruités ou d'une mauvaise qualité d'enregistrement ont néanmoins pu être pris en compte pour cette étude de la temporalité.

⁷⁵ Les définitions de ce terme, ainsi que celles d'adjectifs servant à préciser sa signification, proviennent de trois séries de dictionnaires en ligne (*Cambridge Advanced Learner's Dictionary & Thesaurus, Oxford Dictionaries* et *Merriam-Webster*).

Le graphique ci-dessous révèle un grand consensus de la plupart des lecteurs sur la répartition de durée dans les syllabes, surtout pour la syllabe /du/. Les variations les plus remarquables sont faites sur les syllabes /wVt/, /ju/ et /hI@/. Un grand nombre de lecteurs réalisent différemment la durée de /wVt/ et /hI@/. Par contre, une majorité de lecteurs réalisent /ju/ avec une durée similaire, bien qu’un très faible nombre de locuteurs fassent le choix de doubler, tripler et même quintupler la durée représentée par la médiane. De tels partis pris ne peuvent révéler que des choix de proéminences qui devront être confirmés dans les analyses suivantes sur d’autres corrélats.

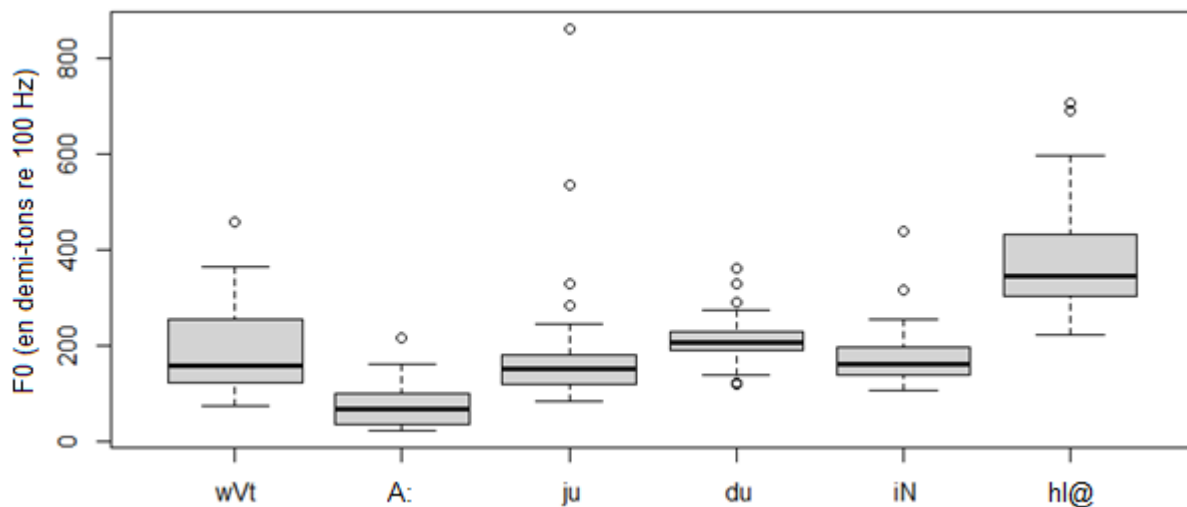


Figure 66 – Question en *What* : Durée moyenne des syllabes (37 lecteurs)

4.3.8 Mélodie du syntagme intonatif

Le graphique ci-dessous présente les boîtes à moustaches de la répartition de la F0 moyenne par syllabe des réalisations de 30 lecteurs.

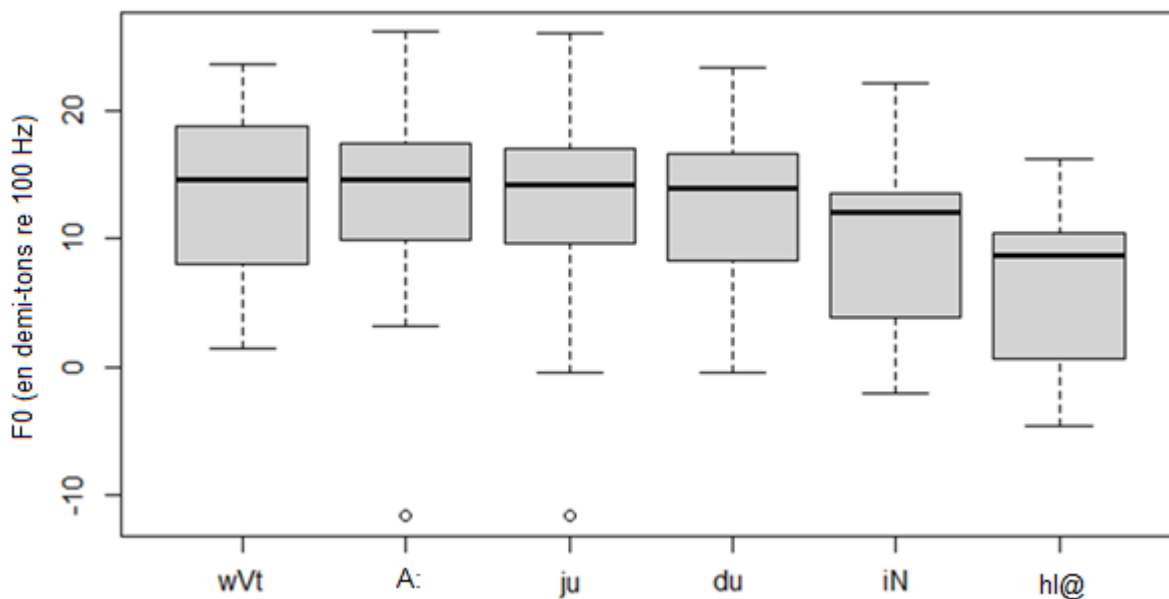


Figure 67 – Question en *What* : Représentation par boîtes à moustaches de la F0 moyenne par syllabe (30 lecteurs)

Il confirme les déclarations de Wells selon lesquelles la mélodie d'ensemble est une chute. De plus, la réaction agressive du Géant rendait la « remontée d'encouragement » carrément inconcevable dans cette situation, même si quelques lecteurs, peu soucieux d'imiter la voix du Géant, l'ont effectivement utilisée. Néanmoins, les quatre premières syllabes déploient une médiane quasiment au même niveau mélodique, comme si la chute ne devait pas se produire. Le graphique en ligne suivant précise le contour normalisé de chacun des lecteurs. Le graphique en lignes ci-après donne une représentation des contours normalisés individuels des 30 lecteurs. On peut y remarquer la grande variation des contours mélodiques utilisés.

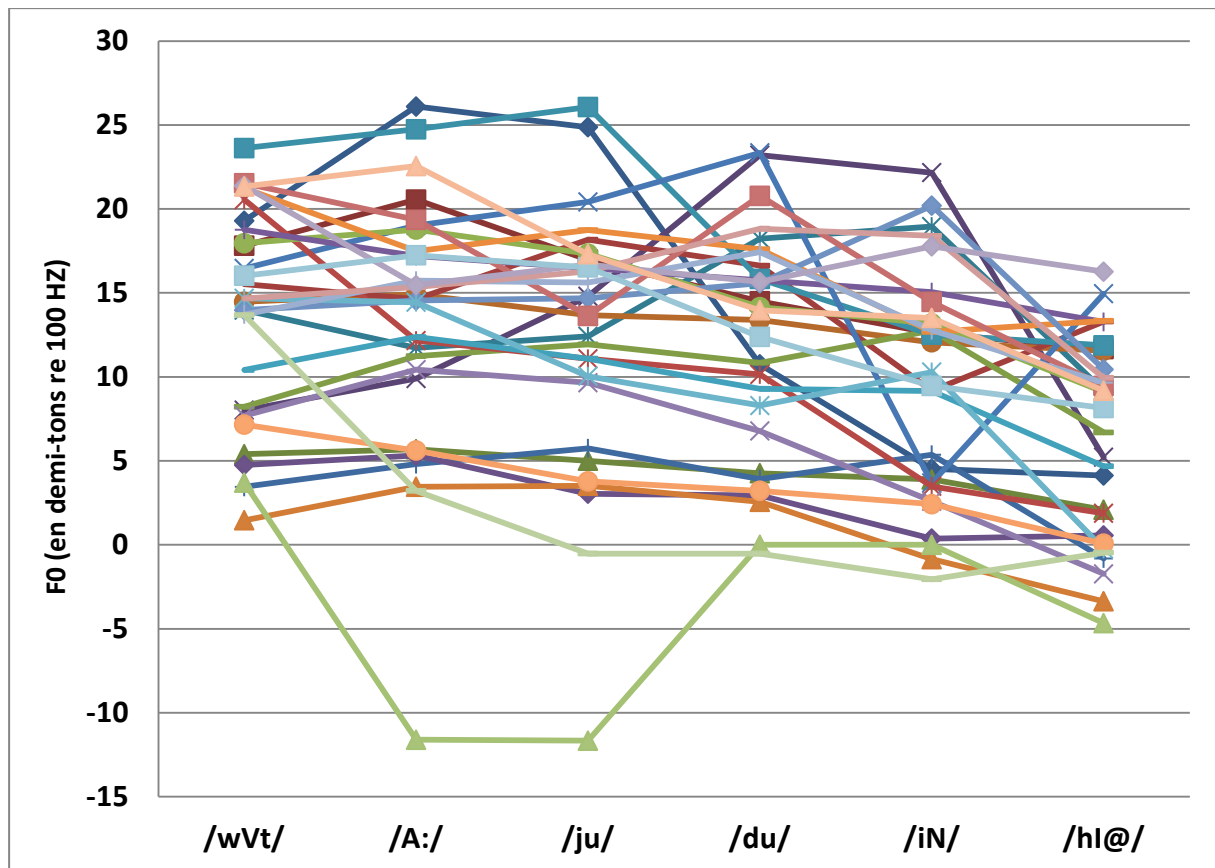


Figure 68 – Question en *What* : Représentation en lignes de la F0 moyenne par syllabe (30 lecteurs)

4.3.9 Proéminences, assimilations et élisions (expérimentation PER-PROM-N)

Selon l'hypothèse émise au départ de cette étude, j'ai précédemment analysé en détail les réalisations des mots pointés comme les plus emphatiques, lesquels se trouvaient juxtaposés. Dans cette phrase, les trois mots les plus emphatiques sont répartis dans la question. Afin d'examiner la variation non révélée par le graphique en boîtes à moustaches mais par celui en lignes, les proéminences de cette question lue par les 42 locuteurs du corpus ont été étudiées plus précisément, non plus par mots mais par syllabes.

4.3.9.1 Choix méthodologiques

Il y a essentiellement deux manières de repérer les proéminences dans la parole :

- La voie traditionnelle, c'est-à-dire par perception auditive.
- La voie acoustique et instrumentale.

Chacune présente avantages et inconvénients, liés à la nature complexe du phénomène, lequel se traduit par plusieurs corrélats acoustiques dont l'oreille humaine sélectionne certains au détriment d'autres en fonction de paramètres complexes eux aussi.

L'objectif poursuivi n'est pas l'analyse des proéminences en elle-même, mais d'obtenir des repère fiables en vue de la compréhension d'autres phénomènes. C'est pourquoi les proéminences de cette question lue par les 42 locuteurs du corpus ont été détectées par perception auditive par deux spécialistes, enseignantes d'anglais non natives. Les distorsions perceptuelles des non natifs par rapport aux natifs revêtent peu d'importance ici étant donné que l'objectif est de comparer les groupes de lecteurs entre eux par un repérage spécifique. C'est la raison pour laquelle les données résultant des tests de perception d'emphase n'ont pas été réutilisés ici. Ni la finalité, ni la méthodologie de repérage ne sont les mêmes. Les travaux de Lacheret sur l'homogénéisation de l'annotation des proéminences avec « p » et « P » (2011 : 24) ont inspiré la première approche. Une deuxième phase a conduit à remplacer les « p » par « 1 » pour indiquer les proéminences faibles, les « P » par « 2 » signifiant les proéminences ordinaires, et ce que j'avais noté « PP », marquant les proéminences perçues comme emphatiques, par « 3 ». L'absence d'annotation signifie, comme chez Lacheret, une absence de proéminence.

4.3.9.2 Analyse des proéminences des jeunes lecteurs

Les graphiques suivants détaillent la répartition des degrés de proéminence repérés chez les jeunes lecteurs.

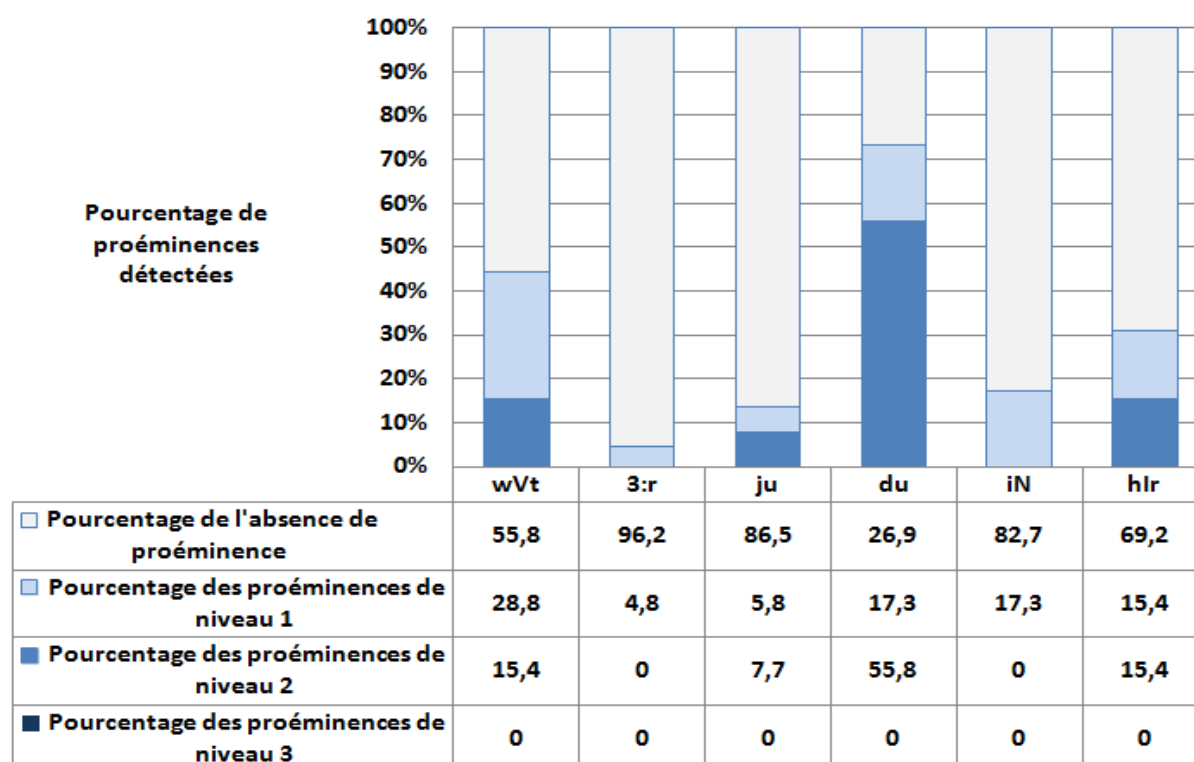


Figure 69 – Question en *What* : Répartition et pourcentage des proéminences dans le groupe des jeunes lecteurs

Ce graphique ne comporte vraiment que trois niveaux de proéminence : « 0 », « 1 » et « 2 ». La mention d’un niveau « 3 » n’est que virtuelle et se justifie par la pertinence de ce niveau pour rendre compte des réalisations des anglophones plus âgés et professionnels. La syllabe /du/ perçoit l’écrasante majorité des choix des jeunes lecteurs, qui la marquent clairement par une proéminence de niveau 2. Cette préférence est souvent accompagnée d’une autre proéminence (de niveau 1 ou 2) sur /wVt/, ou parfois sur le /hI@/ ou même le /iN/. Par contre, elle semble incompatible avec une proéminence sur /ju/ : sur trois cas répertoriés, un seul présente cette compatibilité (niveau 2 sur /du/ et niveau 1 sur /ju/), mais toutes les autres syllabes de la question sont perçues comme proéminentes pour une annotatrice et toutes sauf /iN/ pour l’autre. Le /iN/ n’est jamais perçu comme proéminent sans le /du/ : c’est la raison pour laquelle il relève d’une proéminence de niveau 1 uniquement. Cette proéminence prend la forme d’un *rise-fall* avec F0 élevée. Un exemple en est présenté dans la figure suivante.

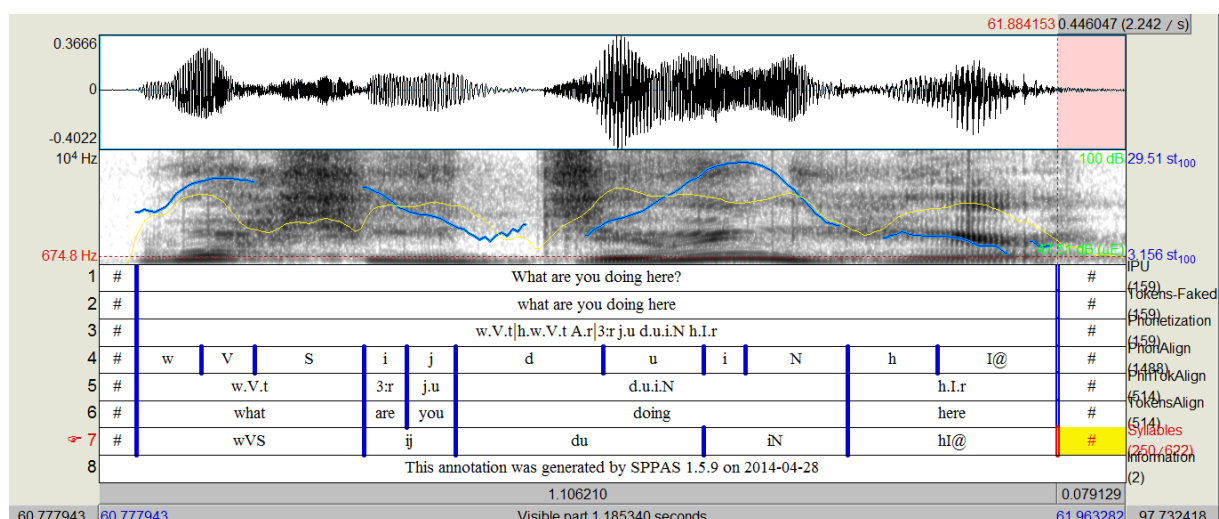


Figure 70 – Question en *What* : Assimilation des trois premières syllabes et proéminence sur /iN/ d’une jeune lectrice (SG-UP-Ro-f-26)

Constatons aussi que chaque fois qu’un lecteur natif a doté le mot *are* d’une proéminence, il en a été de même avec toutes les autres syllabes de la question. L’assimilation ou la disparition de *are*, simple mot outil, justifie la non acceptabilité d’une proéminence sur ce mot seul dans le discours normal des apprenants.

Les jeunes lecteurs n’ont pas systématiquement adopté un changement de voix pour imiter celle du Géant. Nombre d’entre eux ont gardé leur style d’élocution habituel et réduit la durée de la syllabe /A:/. Trois lecteurs l’ont même élidée, l’assimilant aux syllabes adjacentes, comme le graphique précédent en expose un exemple.

4.3.9.3 Analyse des proéminences des lecteurs matures

4. Étude acoustique des marqueurs d’emphase principaux chez les anglophones natifs (expérimentation COR-AC-N2)

Comme indiqué sur le graphique concernant les jeunes lecteurs, les proéminences des lecteurs de plus de 40 ans sont répertoriées sur quatre degrés : « 0 », « 1 », « 2 » et « 3 ». La plupart d’entre eux ont choisi le premier (*What*) et le dernier mot (*here*) de la question, et bien souvent les deux couplés, pour leur faire porter ce poids. Un autre choix, qui peut être indépendant, est de placer l’accent de phrase sur *you* ou le /du/ de *doing* : ce sont les deux seules syllabes qui retiennent une proéminence de niveau 3 dans la question. Les syllabes /A:/ et /iN/ portent un degré d’emphase lorsque le même degré d’emphase ou un degré supérieur est porté par les deux syllabes adjacentes. Il n’y a qu’un seul cas où l’une des deux annotatrices a noté /A:/ proéminente sans que /ju/ le soit aussi.

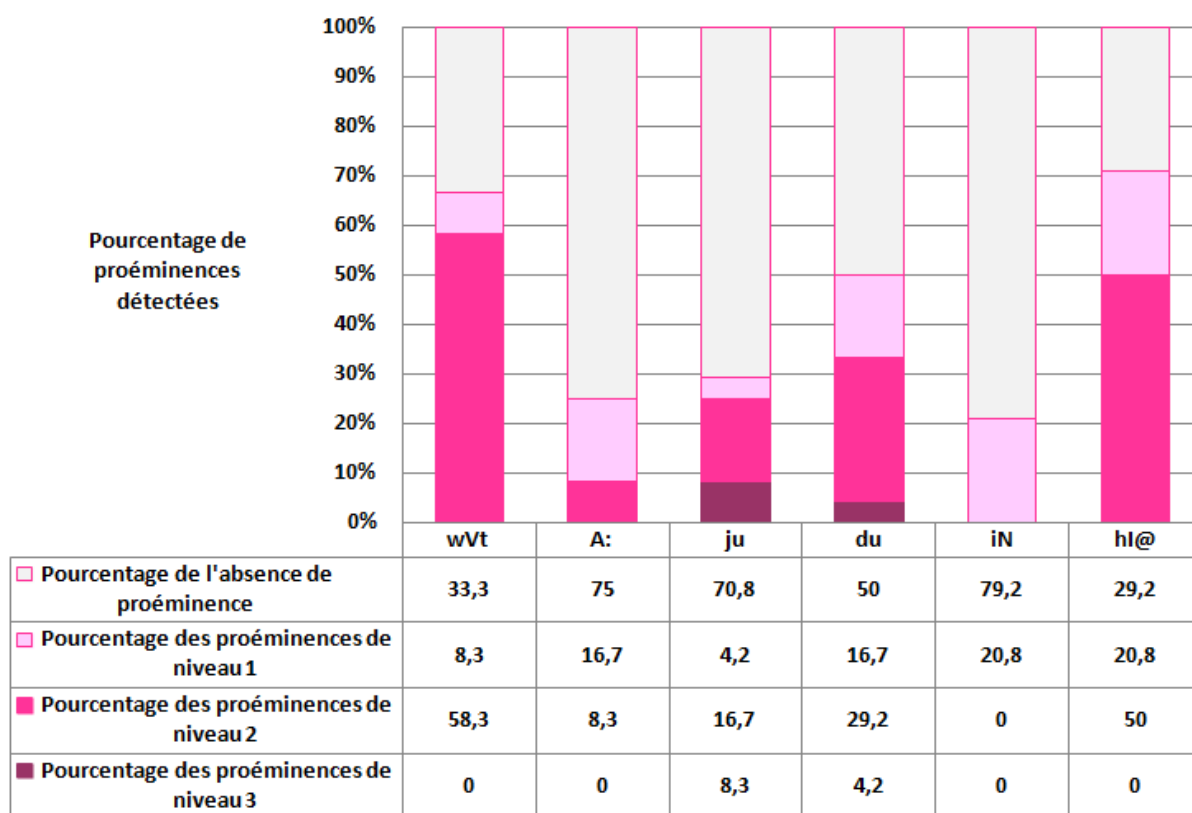


Figure 71 – Question en *What* : Répartition et pourcentage des proéminences dans le groupe des lecteurs matures

4.3.9.4 Analyse des proéminences des lecteurs professionnels

Les proéminences des lecteurs professionnels sont elles aussi répertoriées en quatre degrés : « 0 », « 1 », « 2 » et « 3 ». Deux lecteurs prononcent emphatiquement d’un degré uniforme la totalité des syllabes avec une proéminence particulière sur *you* pour l’un et *what* pour l’autre. Les deux autres lecteurs s’entendent pour marquer /du/ proéminent, et l’un d’eux y ajoute les mots *what* et *here*. La syllabe /du/ obtient donc une proéminence de 100%.

CHAPITRE III – Emphase : mode d'organisation paradigmatique des réalisations prosodiques des anglophones natifs

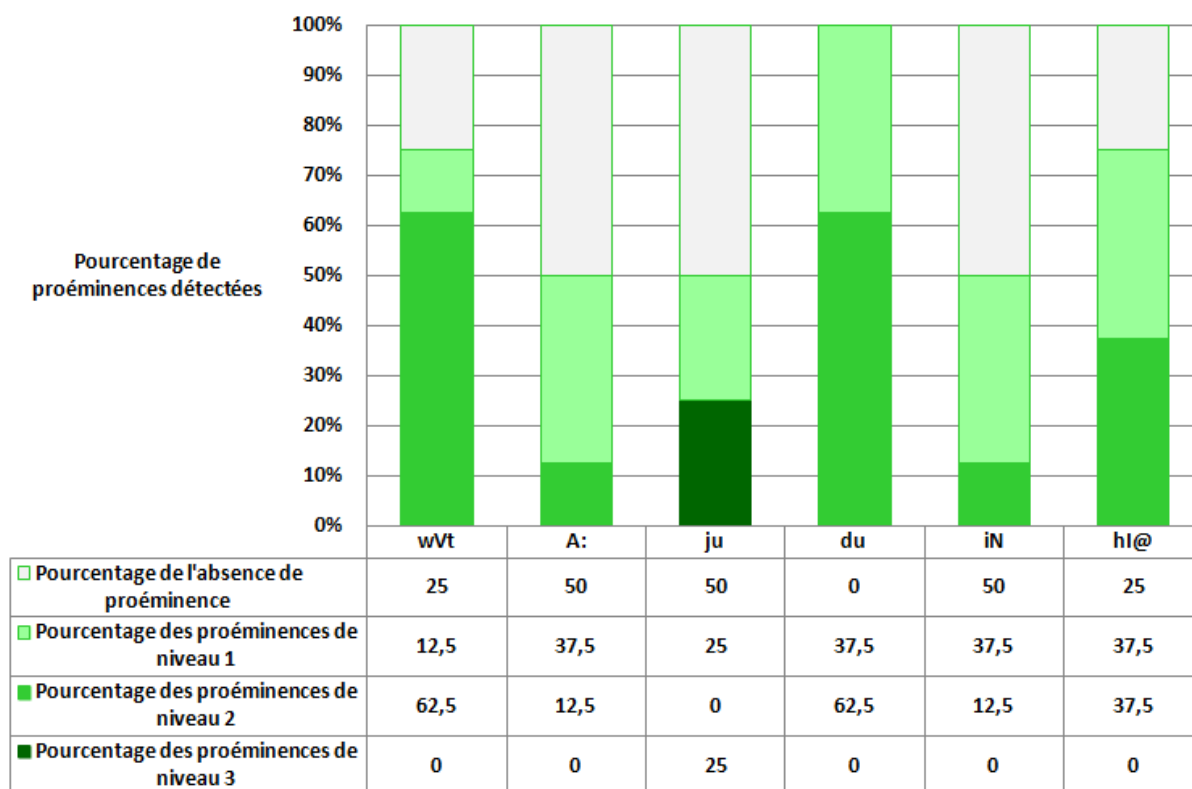


Figure 72 – Question en *What* : Répartition et pourcentage des proéminences dans le groupe des lecteurs professionnels

4.3.9.5 Analyse des proéminences pour l'ensemble des lecteurs

Le graphique suivant résume les résultats globaux de la détection de proéminences. L'axe des « y » ne correspond pas à un pourcentage mais à la moyenne du degré moyen de proéminence calculé pour chacune des deux annotatrices.

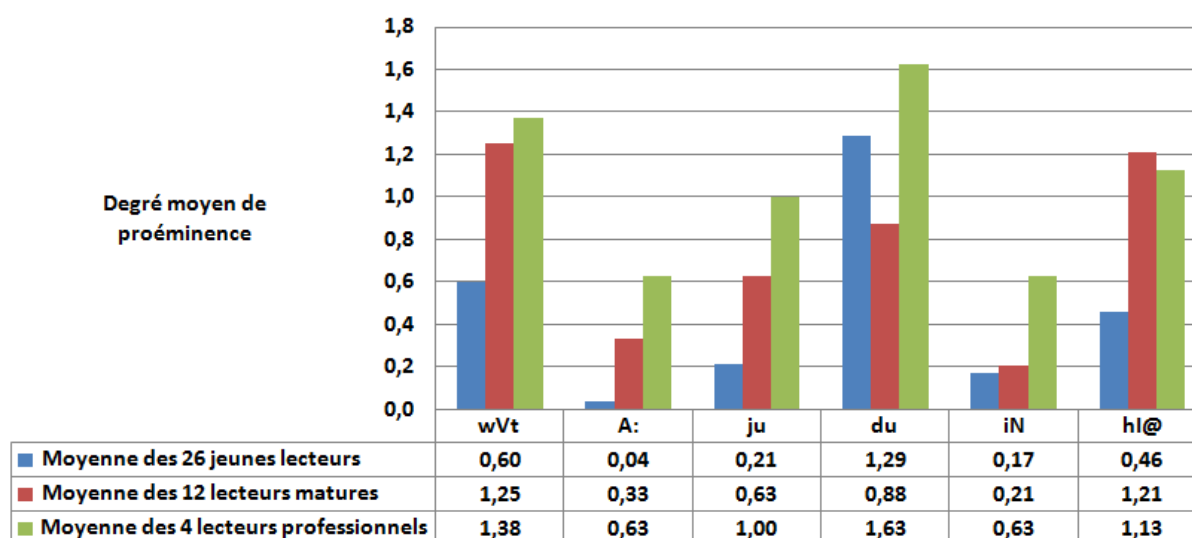


Figure 73 – Question en *What* : Degré moyen de proéminence par groupes de lecteurs

Corroborant les résultats précédents, la syllabe /du/ a été la plus plébiscitée par les lecteurs jeunes et professionnels. Les syllabes /wVt/, /ju/ et /hI@/ ont essentiellement été sélectionnées par les lecteurs de plus de 40 ans et les professionnels. En fait, le décryptage du tableau ne peut se faire sans avoir recours à un classement en catégories prosodiques prenant en compte l'intégralité des corrélats acoustiques.

4.3.9.6 La catégorie d'un focus large

Un focus large signifie que l'accent de phrase est portée par le dernier mot du syntagme intonatif (*here*). Le graphique suivant représente une copie d'écran du TextGrid d'un lecteur non impliqué⁷⁶ dans sa lecture prononçant le focus par défaut sur le dernier mot de la phrase. Le registre est étroit et la courbe de F0 lisse.

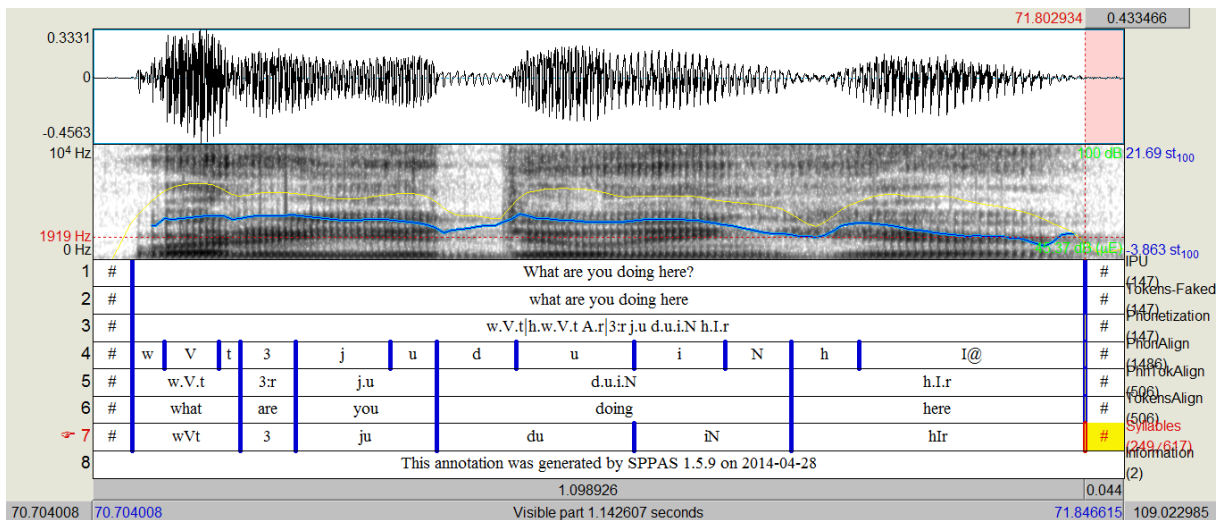


Figure 74 – Question en *What* : Focus large d'un lecteur mature (SG-CA-J5-m-62)

4.3.9.7 La catégorie d'un focus étroit

La plupart des lecteurs ont fait le choix d'un focus étroit. Bien que *doing* recueille la préférence générale, d'autres possibilités sont à envisager :

- *What* en premier signe témoignant de l'accès de rage,
- *here* en contraste à un ailleurs puisque les enfants n'avait pas été autorisés à entrer dans le jardin,
- *you* en opposition avec le légitime propriétaire du jardin.

⁷⁶ « Non impliqué » signifie ici non impliqué dans l'histoire d'un point de vue interlocutif selon les théories de l'interlocution (Douay, 2000 ; Kerbrat-Orecchioni, [1999] 2009).

CHAPITRE III – Emphase : mode d’organisation paradigmatique des réalisations prosodiques des anglophones natifs

Sur le graphique ci-dessous figure le TextGrid d’un focus étroit prenant *you* pour cible. La phrase est tronçonnée par les chutes de F0 sur les syllabes /ju/, /du/ et /hI@/. Ces « tronçons » sont marqués par l’importance que la F0 et l’intensité y prennent.

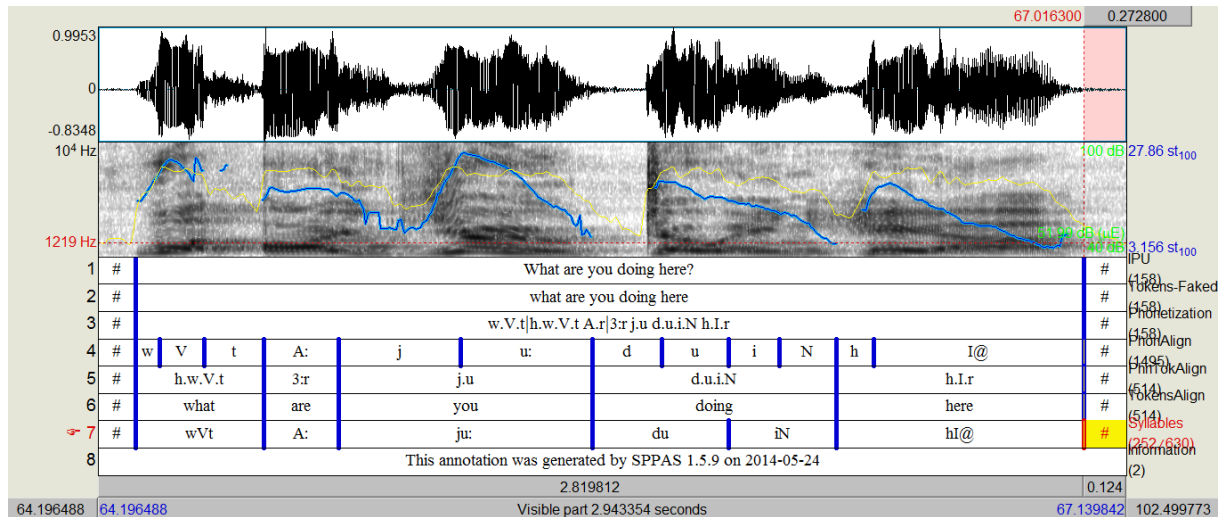


Figure 75 – Question en *What* : Focus étroit sur *you* d’une lectrice non professionnelle de 58 ans (SG-UP-A2-f-58)

4.3.9.8 La catégorie « sans focus »

Cette appellation « sans focus » peut sembler étrange puisque tout syntagme intonatif est notamment caractérisé par son focus, lequel détermine l’emplacement du noyau. Pourtant, elle correspond au choix qu’ont fait d’autres lecteurs en tentant de ne pas identifier de focus afin de déshumaniser la voix du Géant égoïste dans sa réaction contre nature. Sur le plan technique, les corrélats acoustiques tendraient à être :

- F0 basse,
- Intensité accrue sur toutes les syllabes,
- Durée uniforme de toutes les syllabes.

Nous en avons un exemple dans le graphique qui suit.

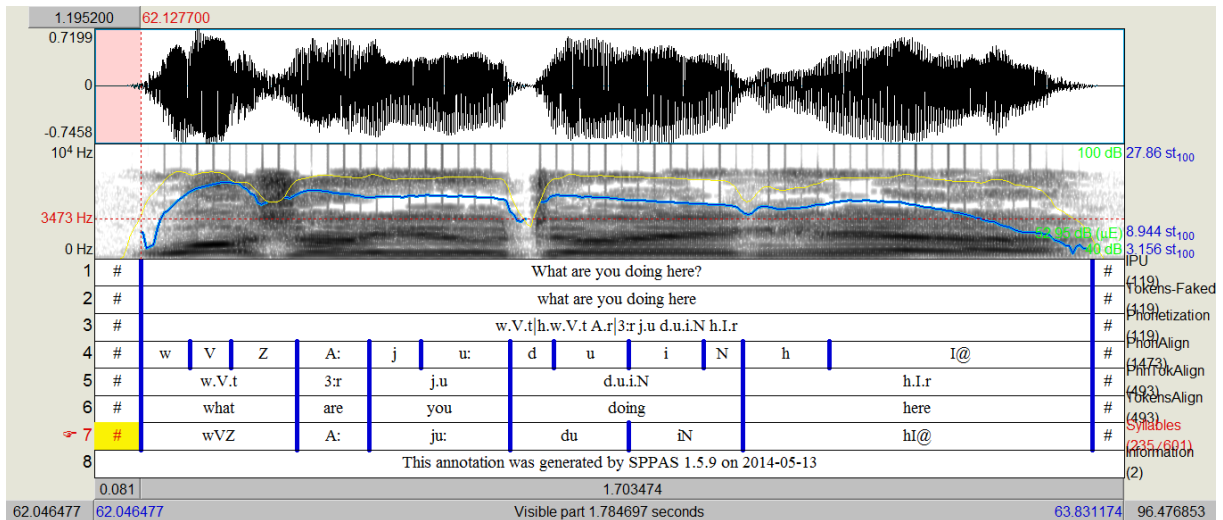


Figure 76 – Question en *What* : F0 basse et forte intensité dans la réalisation « sans focus » d’une lectrice professionnelle (SG-PRO-JG-f)

Des pauses (#) peuvent être insérées, ainsi que le montre le graphique suivant, ce qui met davantage en valeur les syllabes.

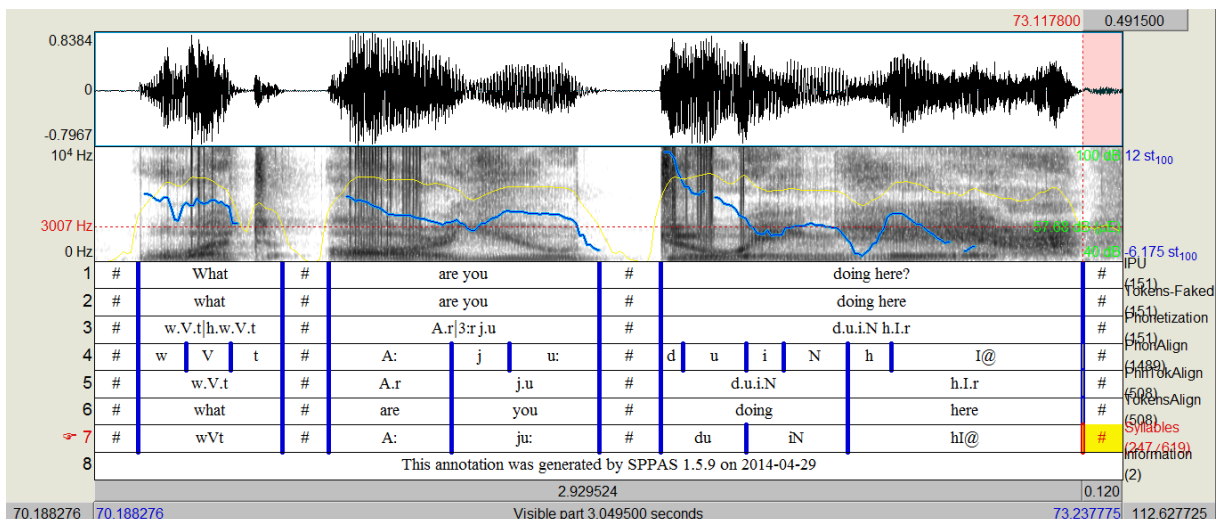


Figure 77 – Question en *What* : F0 basse et forte intensité dans la réalisation « sans focus » avec pauses d’un lecteur mature (SG-PD-W-m-50)

Cette catégorie « sans focus » représente une tendance à laquelle les lecteurs peuvent se conformer ou la combiner avec d’autres critères pour atteindre leur objectif interlocutif. Le groupe des jeunes lecteurs a semblé moins créatif que les deux autres dans leur capacité à incarner la voix du Géant. Le plus souvent, ils n’ont pas opéré de changement de voix et ont gardé leur style de voix coutumière. Nombreux sont ceux qui ont écourté la syllabe /A:/ et trois jeunes lectrices l’ont complètement ignorée, l’assimilant aux syllabes adjacentes. Le Graphique 70 en amont en exemplifie le principe : assimilation des trois premières syllabes par une jeune lectrice (SG-UP-Ro-f-26).

Il est important de remarquer qu’une proéminence sur *are* implique systématiquement celle de toute les autres syllabes de la question. L’assimilation ou l’élision de /A:/, en tant que mot outil, démontre qu’une réalisation d’apprenant en contexte ordinaire n’est pas attendue sur ce mot seul.

Ici encore, le classement en catégories a permis d’éclairer le phénomène de l’emphase dans cette question en *what*, ce qui rend possible une compréhension de ses principes de fonctionnement et de ses implications. On constate la variabilité extrême de cette interrogative dans la variation observée sur son focus, son noyau, son contour mélodique et jusque dans le nombre de ses syllabes.

4.3.10 Tons syllabiques ou mélodie des syllabes (expérimentation ANN-TON-N1)

4.3.10.1 Méthodologie d’annotation

L’observation mélodique a jusqu’à présent porté sur le contour utilisé dans le syntagme intonatif et l’on a pu constater le peu de consensus dans la variation. L’objectif est ici de privilégier l’unité syllabique pour vérifier s’il y en a davantage. Cette unité est très courte, ce qui rend la perception de tons dynamiques extrêmement limitée. Le but ultime recherché est d’identifier les tons susceptibles de révéler une maîtrise mélodique au niveau de la syllabe, afin de les sélectionner comme critères pertinents d’évaluation des apprenants. Comme l’avait mentionné Horgues (rapporté en amont), un ton dynamique est plus difficile à prononcer pour un apprenant francophone qu’un ton statique. La méthodologie consiste à catégoriser les tons simples et dynamiques suffisamment identifiables visuellement pour être considérés comme critères (pour anglophones et apprenants francophones).

Trois questions se posent alors : Qui annote ? L’annotation est-elle phonétique ou phonologique ? Pourquoi ne pas prendre en compte le registre ? L’annotateur est moi-même, c’est-à-dire une seule personne, francophone spécialiste. Cela pourrait sembler problématique aux yeux des puristes. Je tiens à défendre ce choix. Il a été dicté par des contraintes de temps et de faisabilité. C’est la pertinence qui l’a emporté : c’est-à-dire l’adéquation de la méthodologie utilisée à l’objectif à atteindre, à moindre coût. Il s’agit de l’élaboration d’une méthodologie d’annotation fiable et de sa mise en œuvre dans des conditions de faisabilité.

Quel est l’objectif à atteindre ? Il s’agit de schématiser la courbe mélodique afin de la faire correspondre à des tons simples ou complexes et ceci au seul niveau syllabique. L’objectif d’annotation est donc descriptif dans la mesure où la courbe de F0 doit être

fidèlement retranscrite tout en étant schématisée. Ce décalage entre la réalité acoustique continue et la catégorisation phonologique est sujette à la subjectivité de l'annotateur malgré la rigueur souhaitée. Toute schématisation manuelle est par nature subjective et l'utilisation du script *Praat Prosogram* ne rendait compte que des différences perceptibles auditivement, une granularité insuffisante pour cette étude. Ici, l'observation de la courbe de F0 est garante d'objectivité. L'objectif est aussi de nature catégorielle : sélectionner les occurrences susceptibles d'être pertinentes pour l'analyse. Elle se veut phonétique à la base alors qu'elle est indicative de choix et distinctions phonologiques (on s'éloigne du terrain et on se focalise sur l'interprétation, laissant volontairement de côté les corrélats que l'on pense devoir négliger dans l'analyse). Le risque pour l'annotateur (qui façonne sa méthodologie en l'appliquant) est de faire évoluer inconsciemment ses choix en cours de travail d'annotation. La solution trouvée a été de vérifier une nouvelle fois l'ensemble des 42 phrases S10 afin de rectifier et homogénéiser l'annotation. J'ai ainsi constaté que j'avais dû modifier certains signes d'annotation en début et milieu, mais plus en fin de corpus, ce qui m'a conduite à penser que mes critères d'annotation s'étaient stabilisés. Une remarque doit être faite en faveur de l'annotation perceptive : elle permet de prendre en compte tous les fichiers du corpus sans exception, ce qui accroît la représentativité des phénomènes étudiés. Même les fichiers de qualité d'enregistrement médiocre peuvent avoir leur courbe de F0 détectée correctement avec les réglages adéquats, et est utilisable par repérage manuel pouvant discriminer les endroits favorables à une prise de mesures, alors qu'une annotation automatique comprendrait beaucoup trop d'erreurs (sauf si le travail de sélection est opéré manuellement en amont).

Le registre n'a pas été pris en compte dans ce type d'annotation car il dépend grandement de la rapidité de lecture, qui est étudiée par ailleurs. La méthodologie d'annotation arrêtée pour qu'un unique annotateur non natif identifie le ton mélodique révélateur de chaque syllabe est la suivante. Premièrement, une phase d'écoute :

- écoute perceptive de la phrase,
- écoute perceptive la syllabe,
- écoute avec vérification de la courbe de F0.

Deuxièmement, pour vérifier l'hypothèse d'une capacité à réaliser des mouvements mélodiques complexes sur une seule syllabe, la courbe mélodique très lissée de *Praat* est

CHAPITRE III – Emphase : mode d’organisation paradigmatique des réalisations prosodiques des anglophones natifs

décrite visuellement sur tableur *Excel* par un annotateur spécialiste en utilisant les symboles suivants :

Ton statique		Tons dynamiques simples (unidirectionnels)		Tons dynamiques complexes (bidirectionnels)		Tons dynamiques ultra-complexes (tridirectionnels)	
plat	-	montant	/	montant-descendant	∧	descendant-montant-descendant	∩
		descendant	\	descendant-montant	∨	montant-descendant-montant	∪

Tableau 49 – Symboles de transcription des mouvements tonaux syllabiques

Ce système de symboles stylisés et iconiques s’inspire fortement de celui que l’on trouve dans la tradition britannique de l’intonation et des *tonetic stress marks* utilisés dans le *Spoken English Corpus* (Wichmann, 2000 : 10-11). Néanmoins, il ne tient aucunement compte de la position dans le registre et inclut deux symboles supplémentaires, les tons dynamiques ultra-complexes, lesquels sont traditionnellement assimilés à des tons dynamiques complexes. L’indexation que ce système représente se situe à mi-chemin entre l’observation phonétique et la catégorisation phonologique. Les syllabes perçues comme statiques peuvent en réalité être dynamiques par vérification de la courbe de F0. Malgré la pente de F0 constatée, une durée vocalique peu importante, un manque d’intensité ou une réduction vocalique peuvent occulter ce caractère dynamique. Ce ton sera donc annoté comme statique. L’inverse a aussi été constaté : dans le fichier SG-Pro-GRB-m_S10_ (Lecteur C), pour la syllabe /du/, /d/ n’étant pas voisé. De même, un ton dynamique simple peut, après examen de la courbe de F0, se révéler complexe (dans les fichiers de deux lecteurs professionnels) : ils ont été annotés ainsi qu’ils avaient été perçus auditivement. Deux contours mélodiques similaires pourront recueillir des annotations différentes. Dans les réalisations d’un jeune lecteur (SG-PD-So-m-30 dans le graphique ci-dessous) : les trois premières syllabes sont perçues et annotées comme statiques, sans doute à cause de leur brièveté, alors que /du/ et /iN/ sont signalées comme dynamiques. La syllabe /hI@/ est identifiée comme plate.

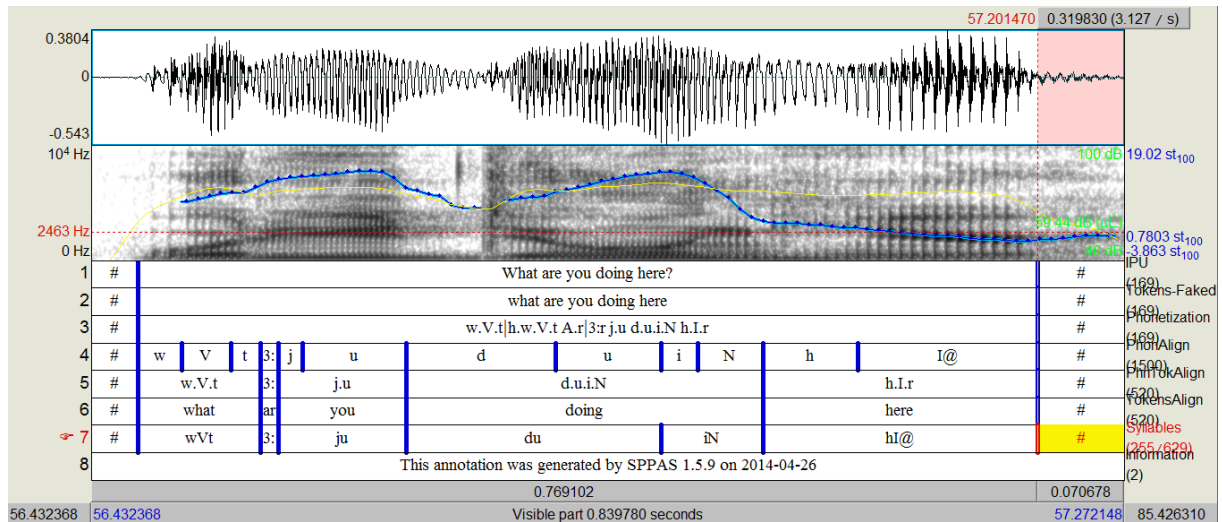


Figure 78 – Question en *What* : Diversification de l'annotation malgré des contours intonatifs similaires

Certains tons dynamiques peuvent être repérables à l'audition de la syllabe, mais pas de la phrase. Par défaut, ces tons sont annotés comme statiques. Il est important de signaler que certains tons complexes, étant prononcés rapidement, ne permettent pas l'identification en temps réel, c'est pourquoi ils ont été annotés comme statiques (exemple du /ju/ dans le fichier SG-Pro-JC-f_S10_ : la courbe de F0 révèle une descente-montée-descente, mais la syllabe ne dure que 0,128 seconde, ce qui rend cet aspect dynamique pratiquement inaudible, sauf à la rigueur pour natif spécialiste expérimenté).

4.3.10.2 Résultats et analyse

4.3.10.2.1 Résultats globaux

Les résultats obtenus figurent dans les graphiques suivants. Le graphique ci-dessous récapitule globalement le pourcentage d'occurrence des tons par catégorie de lecteur.

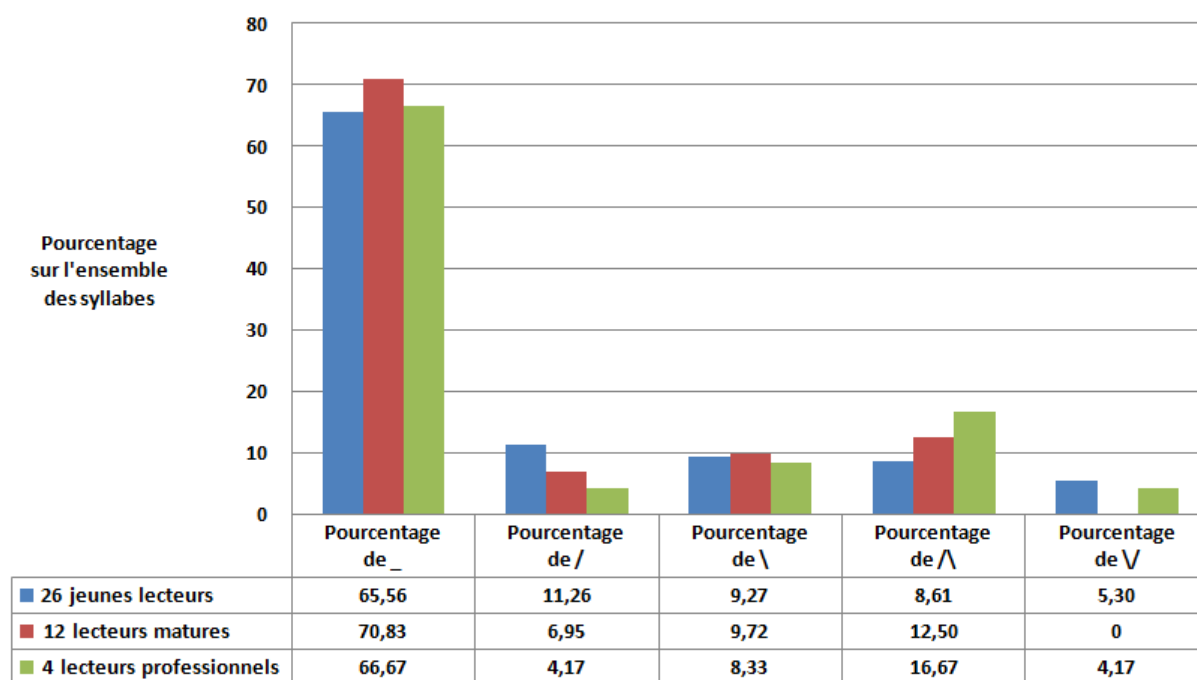


Figure 79 – Question en *What* : Pourcentage d’occurrence des divers tons par catégorie de lecteurs

4.3.10.2.2 Le ton statique

Le ton le plus fréquent, toutes catégories de lecteurs confondues, est le ton par défaut : le ton statique (ou plat). Les pourcentages de fréquence se situent dans une fourchette relativement étroite : entre 65,56% et 70,83% (moyenne de 67,69%), ce qui en fait un ton peu discriminant pour cette étude.

4.3.10.2.3 Le ton descendant-montant

À l’autre bout de l’échelle, le ton *fall-rise* est très peu utilisé (3,15% en moyenne). Les lecteurs matures n’en font aucun usage au niveau de la syllabe. Le graphique ci-dessous précise les lieux d’utilisation des deux groupes restants. Dans 30,77% des cas, les jeunes lecteurs en font usage, mais exclusivement sur le dernier mot *here*. Il s’agit donc d’un *encouraging rise*, sans doute renforcé par le prolongement de la phrase avec un énoncé citant. On peut penser qu’un lecteur peu soucieux de s’impliquer dans les péripéties du conte donnera priorité à l’agencement syntaxique au détriment de la qualité interlocutive. Il est à noter qu’une lectrice professionnelle, représentant donc 25% de son groupe, fait usage de ce ton sur la syllabe /du/. L’observation de la courbe de F0 révèle un ton descendant-montant-descendant à cheval sur les deux syllabes du mot *doing*. Les résultats présentés conduisent à penser que le ton descendant-montant peut être considéré comme témoignage de la jeunesse du lecteur (catégorie sociologique), par contre, il ne peut être discriminant en ce qui concerne

la qualité de lecture des natifs, dans la recherche de *crierial features* pour l'évaluation des apprenants.

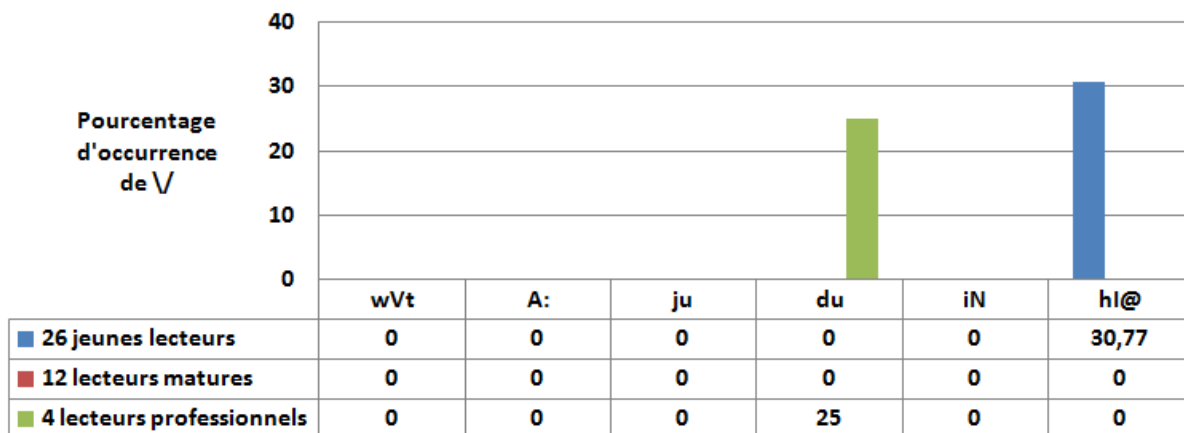


Figure 80 – Question en *What* : Pourcentage d'occurrence du ton « descendant-montant » sur une même syllabe

4.3.10.2.4 Le ton montant

Toutes catégories de lecteurs confondues, le pourcentage des tons montants est de 7,46%. Ainsi que le montre le graphique suivant, on retrouve le ton montant principalement chez les jeunes lecteurs (11,26%) sur les syllabes marquées comme proéminentes (/wVt/ et /du/). Dans une moindre mesure (6,94%), les lecteurs de plus de 40 ans font de même. Un lecteur professionnel (4,17%) en fait usage pour rendre /ju/ proéminent (un examen sur spectrogramme révèle en fait un ton dynamique plus complexe, qui n'a pas été détecté comme tel lors des deux premières écoutes). 7,69% des lecteurs de moins de 40 ans terminent la question par ce ton syllabique pour les raisons évoquées ci-dessus.

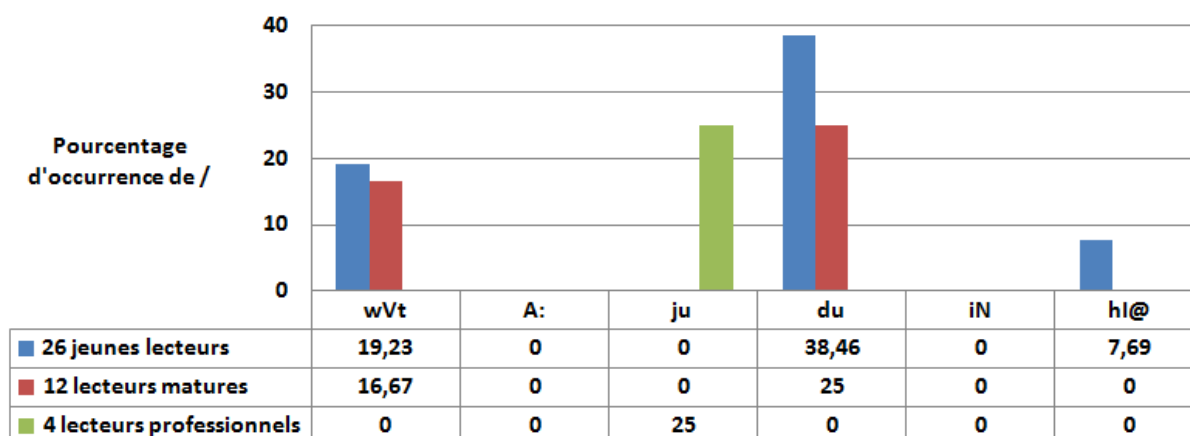


Figure 81 – Question en *What* : Pourcentage d'occurrence du ton montant sur une même syllabe

4.3.10.2.5 Le ton descendant

Toutes catégories de lecteurs confondues, le pourcentage des tons descendants est de 9,11%. Dans la lecture des jeunes lecteurs, ce ton descendant se retrouve sur toutes les syllabes de la question sauf sur /A:/, malgré une moyenne d’utilisation plus faible que pour le ton montant (graphique ci-dessous). Pourtant, les trois catégories de lecteurs se rejoignent dans la moyenne d’utilisation de ce ton (jeunes lecteurs : 9,27%, lecteurs matures : 9,72% et lecteurs professionnels : 8,33%).

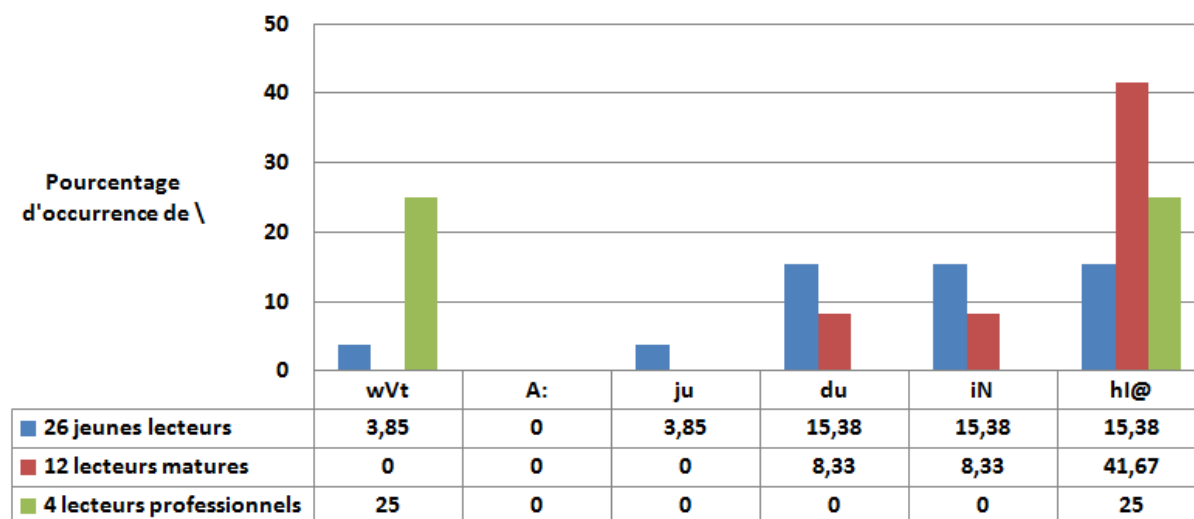


Figure 82 – Question en *What* : Pourcentage d’occurrence du ton descendant sur une même syllabe

4.3.10.2.6 Le ton montant-descendant

Toutes catégories de lecteurs confondues, le pourcentage des montants-descendants représente 12,59%. Il est le moins plébiscité par les jeunes lecteurs (8,61%), alors qu’il représente 12,50% des tons utilisés par les lecteurs plus âgés et 16,67% des lecteurs professionnels. Les jeunes lecteurs en font un usage très limité, sauf dans la caractérisation de *doing* et ce ton montant-descendant est octroyé soit à la syllabe /du/, soit à la syllabe /iN/, mais pas aux deux à la fois sauf dans le cas SG-PD-Ni-f-26_S10_ où ce ton complexe figure sur les quatre dernières syllabes de la question, qui de ce fait devient chantante et presque engageante, même sans *encouraging rise*. Elle semble exprimer davantage la surprise et l’intérêt que le reproche. À ce stade, on ne peut exclure le fait que le ton montant-descendant sur le /iN/ soit indicateur d’un parler « jeune ». Les lecteurs matures et professionnels utilisent le *rise-fall* pour marquer les prééminences.

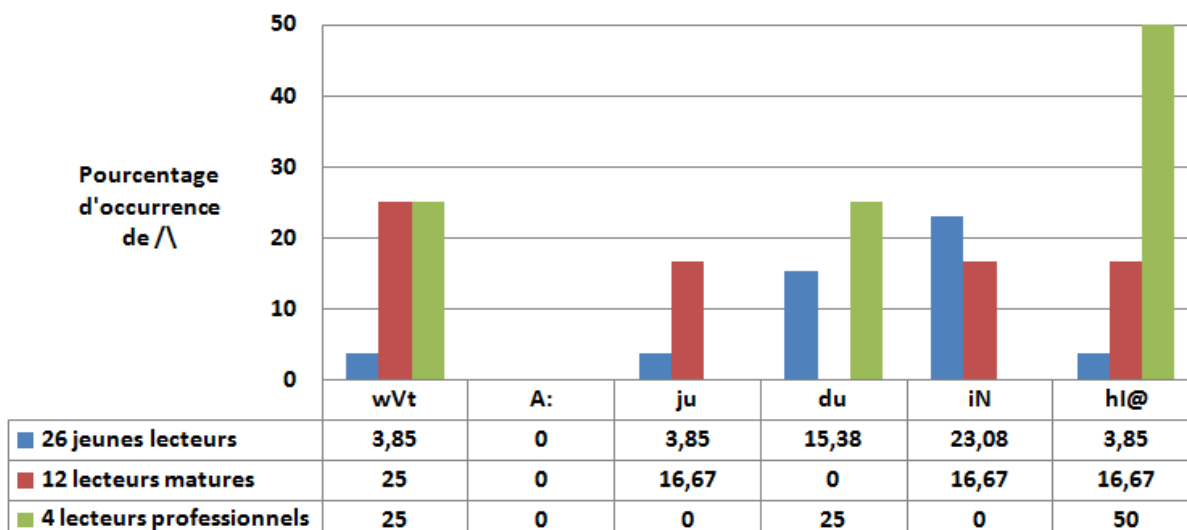


Figure 83 – Question en *What* : Pourcentage d’occurrence du ton « montant-descendant » sur une même syllabe

4.3.10.2.7 Comparaison des trois tons « montant », « descendant » et « montant-descendant »

Jusqu’à présent, chaque ton a été étudié sans tenir compte de son incidence sur les tons voisins. Il semble intéressant d’observer les liens qui peuvent exister entre les trois tons dynamiques « montant », « descendant », « montant-descendant » et les locuteurs qui les utilisent. Reprenons le graphique de départ en examinant plus précisément l’usage de ces tons par les trois catégories de lecteurs. La vue rapprochée présentée dans le graphique ci-dessous a ainsi permis d’établir des lignes de tendance sur le pourcentage d’utilisation des tons sur l’ensemble des syllabes de la question S10. Le premier ton « / » est le plus simple car il part d’une position neutre. Le deuxième « \ » présente une chute par rapport à une montée qui a déjà eu lieu. Le troisième « /\ » montre une montée et une descente dans la même syllabe. L’agencement des trois tons reflète donc une gradation dans la complexité tonale. Les sommets de chaque barre représentant les moyennes de chaque catégorie de lecteurs ont été reliés entre eux par une ligne de même teinte, mettant en évidence la systématicité des choix opérés.

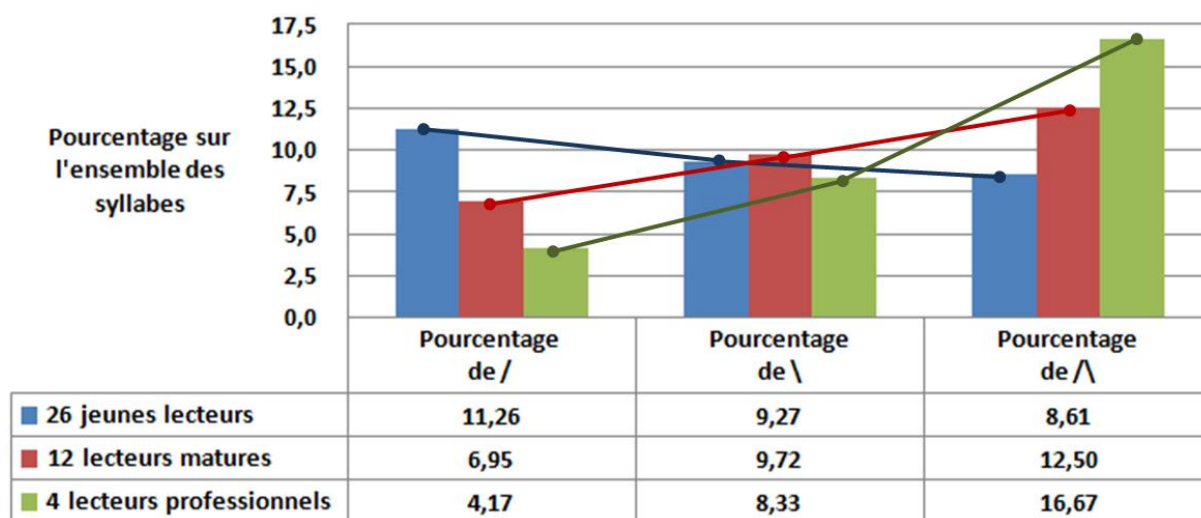


Figure 84 – Question en *What* : Pourcentage d’occurrence des tons « montant », « descendant » et « montant-descendant » sur l’ensemble des syllabes

On remarque que le pourcentage d’utilisation d’un ton descendant est extrêmement proche pour les trois catégories de lecteurs (de 8,33% à 9,72%). Par contre, les différences concernent les autres tons et les tendances. La moyenne générale des jeunes lecteurs montre la primauté du ton montant sur les deux autres. Le ton complexe montant-descendant est le moins utilisé des trois. C’est pourquoi la ligne bleue reliant le sommet des barres bleues est graduellement descendante, signifiant un appauvrissement de la complexité tonale syllabique. Au contraire, comme le montre la ligne rouge ascendante, les lecteurs de plus de 40 ans font peu d’usage du ton montant et préfèrent utiliser le ton complexe « /\ ». Cette stratégie est encore plus évidente chez les lecteurs professionnels dont la ligne de tendance verte est d’une ascension encore plus marquée.

D’après les résultats présentés dans ce graphique, il semblerait que les lecteurs expérimentés aient tendance à utiliser des tons dont la complexité traduirait un enrichissement tonal au niveau syllabique, de nature à « accrocher » l’auditoire en utilisant un second canal de variation mélodique autre que celui du syntagme intonatif, ce niveau se superposant alors à lui. Le ton le plus révélateur d’une maîtrise de l’emphase prosodique au niveau syllabique serait le *rise-fall*.

4.3.11 Conclusion

La question en *What* lue en contexte présente de nombreux aspects complexes dont il faut tenir compte en fonction d’une utilisation évaluative ultérieure. Un bon lecteur opère un changement de voix pour incarner celle du personnage du Géant. D’après l’étude effectuée, deux options principales s’offrent à lui :

- refléter la corpulence massive du Géant en adoptant une F0 basse par rapport à sa voix normale, ce qui le conduit à réduire son registre,
- donner vie et forme à la colère du Géant, c'est-à-dire jouer sur le rythme en utilisant un surcroît d'intensité et des valeurs élevées de F0 pour rendre proéminents les mots sur lesquels porte le reproche.

Ce phénomène est primordial car, suivant l'option choisie, un bon lecteur pourra représenter un mauvais modèle de rendu canonique d'une question en *What* et vice-versa. La matrice suivante récapitule les points saillants à prendre en compte pour déterminer l'adéquation d'une lecture de natif à un bon modèle d'évaluation.

<i>Critères à prendre en compte</i>			<i>Interprétation</i>	<i>Utilisation</i>	
Changement de voix	Large empan de tessiture	Occurrences de Λ	Interprétation suggérée	Bon modèle de lecteur	Bon modèle évaluatif
+	-	+ / -	Cherche à incarner le personnage	+	-
+	+	+	Cherche à incarner le personnage	+	+
-	+	+	Tente d'apporter de la vie à sa lecture, lecteur impliqué	+	+
-	-	+	Lecture normale et naturelle	-	+
-	+	-	Lecture normale et naturelle	+ / -	+
-	-	-	Mélodie monotone, lecteur non impliqué	-	-

Tableau 50 – Question en *What* : Pertinence de la question en *What* de P10 comme modèle pour évaluation

Cette matrice prend en compte trois critères de catégorisation prosodique de cette question. Un changement de voix accompagné d'un registre restreint correspondra à un choix de lecture privilégiant la corpulence massive du Géant au détriment d'une expression de ses sentiments traduits par une variation mélodique. Au contraire, un changement de voix couplé d'un large registre indiquera des variations de F0 représentant des proéminences. Ainsi, un bon lecteur sera caractérisé par au moins l'un des deux points : un changement de voix et un large registre. Un bon modèle de question exclut de fait les lecteurs qui adoptent une F0 basse ou qui montrent un désintéressement en ne correspondant à aucun des trois critères déterminants.

L'hypothèse de départ concernant l'utilisation de l'emphase est qu'elle favorise la variabilité des réalisations phonétiques. Elle est une fois de plus validée par l'étude détaillée de cette interrogative en montrant qu'elle est d'une extrême variabilité due à sa grande complexité. Sa variation porte sur son focus, son noyau, son contour mélodique et jusque sur le nombre de ses syllabes. En conséquence, il ne me paraît pas envisageable que l'interrogative en *What* de la Phrase 10 puisse être sélectionnée comme modèle de réalisation prosodique de question partielle. Par contre, la prosodie syllabique paraît être une piste prometteuse pour discriminer le niveau intonatif d'un apprenant avancé. Cependant, l'absence de locus stable pour la réalisation du ton montant-descendant ne prédispose pas non plus cette interrogative à une utilisation comme modèle évaluatif.

4.4 *Nobody*

Le mot *nobody* se trouve situé au milieu du texte *The Selfish Giant*, dans la Phrase 11, laquelle est presque exclusivement au style direct :

"My own garden is my own garden," said the Giant; "any one can understand that, and I will allow nobody to play in it but myself."

Ce mot se trouve en 4^{ème} position selon son degré d'emphase dans le texte du *Selfish Giant*.

4.4.1 Caractéristiques linguistiques de *nobody*

Nobody est un pronom indéfini. Biber *et al.* introduisent les pronoms en soulignant qu'ils sont avant tout des procédés économiques destinés à remplacer des groupes nominaux en faisant référence au contexte ou à la situation d'énonciation (1999 : 327). En outre, ils peuvent être utilisés lorsque l'élément de référence est inconnu ou très général, ou pour remplir des fonctions très spécifiques de liaison de propositions (*for specific clause-binding functions*). Biber *et al.* en présentent les quatre groupes principaux, tous dérivés d'un quantificateur (1999 : 351) :

the *every* group : *everybody, everyone, everything*

the *some* group: *somebody, someone, something*

the *any* group: *anybody, anyone, anything*

the *no* group: *nobody, no one, none, nothing*

Ainsi que le montre la fréquence d'occurrence des pronoms indéfinis selon les registres (tableau ci-dessous), tous les groupes de pronoms indéfinis sont surtout représentés dans les corpus de conversation et de fiction. Les pronoms se terminant en *-body* sont plus courants en

conversation, alors que ceux en *-one* relèvent davantage de la langue écrite. Les pronoms se terminant en *-body* sont aussi plus fréquents en anglais américain qu'en anglais britannique.

Distribution of indefinite pronouns (frequencies of the quantifiers *every*, *some*, *no*, and *any* are included for comparison); occurrences per million words

each ■ represents 200 □ represents less than 100

	CONV	FICT	NEWS	ACAD
<i>everybody</i>	■	■	□	□
<i>everyone</i>	■	■	■	□
<i>everything</i>	■■	■■	■	□
<i>every</i> (quantifier)	■■	■■	■■	■■
<i>somebody</i>	■■	■	□	□
<i>someone</i>	■	■■	■	■
<i>something</i>	■■■■■■■■	■■■■■■■■	■■■■■■■■	■
<i>some</i> (quantifier)	■■■■■■■■ ■■	■■■■■■■■■■	■■■■■■■■	■■■■■■■■■■ ■■■■
<i>anybody</i>	■	□	□	□
<i>anyone</i>	■	■	■	□
<i>anything</i>	■■■	■■■	■	□
<i>any</i> (quantifier)	■■■■■■■■	■■■■■■■■	■■■■■■	■■■■■■■■
<i>nobody</i>	■	■	□	□
<i>no one</i>	□	■	□	□
<i>none</i> ⁷⁷	□	■	□	□
<i>nothing</i>	■■■	■■■■■■	■	■
<i>no</i> (quantifier)	■■■■■■■■	■■■■■■■■■■ ■■■■■■	■■■■■■■■■■	■■■■■■■■

Tableau 51 – Distribution des pronoms indéfinis (Biber *et al.*, 1999 : 352)

Les quatre grands regroupements sont présentés par Quirk *et al.* ([1985] 2005 : 345), puis Greenbaum (1996 : 192), sous la dénomination d'assertion (*assertive*), de non-assertion (*non-assertive*), de négation (*negative*) et de globalisation (*universal*). Ainsi que le rappellent Quirk *et al.* ([1985] 2005 : 376), les pronoms composés se caractérisent par le regroupement de deux morphèmes qui sont de nature déterminante pour la première partie (*every-*, *some*, *any* ou *no*)⁷⁷ et de nature nominale pour la deuxième (*-one*, *-body*, *-thing*).

⁷⁷ Huddleston et Pullum adoptent (2002 : 539) ce point de vue, qu'ils étayent par l'impossibilité de combiner articles et quantificateurs dans le même groupe nominal. Ils en déduisent qu'*any* ou *no* (entre autres) sont des déterminants.

Nobody est catégorisé comme pronom indéfini négatif (Quirk *et al.*, [1985] 2005 : 376). Une proposition ne pouvant normalement inclure qu'une seule marque de statut négatif, il faut s'attendre à ne trouver que des marques de nature non assertive dans la même proposition, ce qui justifie grammaticalement aux yeux de Quirk *et al.* ([1985] 2005 : 85) le fait que la négation implique la non assertion (*It is this which justifies grammatically the treatment of negation as carrying the implication of nonassertion*). Greenbaum en donne l'exemple suivant : *Nobody ever makes it like she used to make it*, où *nobody* se trouve en position sujet (1996 : 193). Il peut aussi recevoir la fonction d'attribut : *There was nobody at the office*. (Quirk *et al.* ([1985] 2005 : 378)) ou de complément : *I trust nobody but her*. (Huddleston et Pullum (2002 : 461)). Huddleston et Pullum remarquent en outre que de nombreux dialectes de langue anglaise, s'étendant du Cockney des districts est de Londres à l'anglais vernaculaire afro-américain (*African American Vernacular English* ou *AAVE*, entre autre appellations), acceptent des doubles ou triples négations, alors que la langue standard en vigueur interdit la compatibilité des négations absolues avec d'autres négations dans une même proposition (2002 : 846).

Larreyra et Rivière rappellent que, si les pronoms de la série *-body* et *-one* s'accordent au singulier avec le verbe, cet accord se fait au pluriel en anglais courant dans les reprises par pronoms personnels et possessifs, ce qui permet d'éviter le problème du choix de genre ([1991] 2002 : 265).

4.4.2 Approche sémantico-discursive du pronom indéfini « nobody »

Selon Biber *et al.* (1999 : 351), les pronoms indéfinis se réfèrent à des entités que le locuteur ou écrivain ne peut ou ne veut spécifier avec précision. Ils se rapprochent en cela des groupes nominaux incluant des déterminants indéfinis, mis à part la référence faite à une personne ou chose indéterminée. Quirk *et al.* précisent que les composés avec *-one* et *-body* se réfèrent à des personnes, et que, si leur différence fonctionnelle et sémantique reste la même, les pronoms en *-one* sont néanmoins perçus comme plus élégants que ceux en *-body* ([1985] 2005 : 378). Huddleston et Pullum regroupent les quantificateurs en *some*, *any* et *no* dans la catégorie des quantificateurs existentiels, dont la valeur de *any* sera soit la non affirmation, soit le libre choix, et celle de *no* la valeur négative (2002 : 423).

Selon les *Cambridge Dictionaries Online*⁷⁸, *nobody* se réfère à une absence de personnes, de même que *no one*. Leur sémantisme est beaucoup plus marquant et tranché que celui des locutions *not [...] anybody* et *not [...] anyone* :

Nobody, no one, nothing, nowhere are stronger and more definite than *not ... anybody/anyone/anything/anywhere*: [...]

*She told **no one**, not even her mother.* (stronger than *She didn't tell anyone ...*)

Cela explique la raison pour laquelle Huddleston et Pullum (2002 : 823) appellent la classe des composés de *any* des *negatively-oriented polarity-sensitive items* ou *NPIs* que l'on pourrait tenter de traduire par « éléments compatibles avec une polarité négative ». Ils vont (2002 : 812) par ailleurs jusqu'à nommer les formes indépendantes et composées dérivant du déterminant *no* les « négations absolues » (*absolute negators*). *Nobody* en fait partie. Le dictionnaire en ligne *Merriam-Webster*⁷⁹ lui donne comme antonymes *everybody* et *everyone*, ce qui conduirait à penser que *nobody* puisse adopter le processus mental de saisie analytique et synthétique que Lapaire et Rotgé attribuent à *every*, dont le résultat serait ensuite nié dans sa globalité ([1991] 2002 : 189-190).

4.4.3 Approche prosodique du pronom indéfini « nobody »

Larreya et Rivière relèvent qu'en anglais britannique, *anyone* et *nobody* portent seulement un accent principal sur la première syllabe, alors qu'en anglais américain, comme les autres pronoms indéfinis en *-body*, *nobody* porte un accent secondaire sur la première syllabe de *body*, qui a donc une voyelle pleine (2010 : 264).

Huart reprend les propos de Larreya et Rivière ci-dessus tout en précisant que c'est en position sujet que cette accentuation se réalise (2010 : 78-79). Elle signale en outre qu'en fonction de complément, les composés en *some* et *any* « font souvent bloc avec le verbe et perdent ainsi leur accent mélodique » et donne pour exemple :

*Did you **meet** anybody at the **market**?* (chute intonative sur **meet**)

As-tu rencontré quelqu'un au marché? [...]

did you meet anybody = as-tu fait une rencontre ?

Elle justifie cette perte de proéminence par la nécessité pour le verbe de recevoir un complément dont le référent ne relève pourtant pas d'un choix de la part de l'énonciateur. La

⁷⁸ <http://dictionary.cambridge.org/fr/grammaire/grammaire-britannique/no-one-nobody-nothing-nowhere> (consulté en 2015)

⁷⁹ <http://www.merriam-webster.com/dictionary/nobody> (consulté en 2015)

représentation intonative d’exemples donnés par Bolinger (1989 : 180) semble étayer son propos :

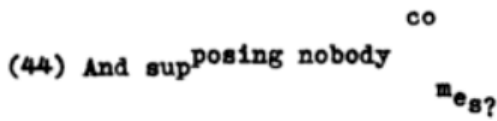


Figure 85 – Exemple de perte de proéminence de *nobody* (Bolinger, 1989 : 179)

L’exemple (44) montre que *nobody* fait partie de la tête du syntagme intonatif étant donné qu’il se trouve sur la même ligne mélodique que *-posing*, lequel débute le pied qui englobe *nobody*. Ce phénomène ne peut se produire qu’en l’absence de proéminence marquée dans *nobody*. Les deux exemples (47) et (48) montrent que la hauteur mélodique de *nobody* se situe à mi-chemin entre celle, plus basse, de la syllabe accentuée du mot qui précède et celle, plus haute, de la syllabe accentuée de celui qui suit.

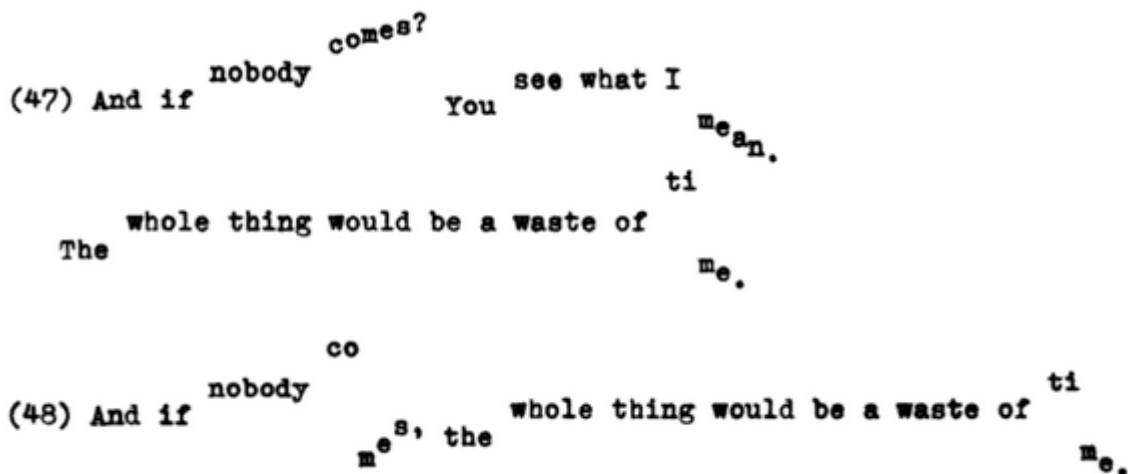


Figure 86 – Exemples de perte de proéminence de *nobody* (Bolinger, 1989 : 180)

Il semblerait néanmoins que la condition (*supposing, if*) introduisant ces phrases prévale en matière de proéminence, ce qui constitue un cas très particulier.

4.4.4 Intérêt de *nobody* dans l’étude de l’interlangue prosodique

L’intérêt d’étudier le pronom indéfini *nobody* est que non seulement ce mot est focalisateur d’emphase, au moins sur sa première syllabe en anglais britannique, par sa forte charge négative, mais encore sa syllabe tonique comprend une consonne nasale et une diphtongue. *A priori*, sa F0 devrait pouvoir être étudiée sur toute la durée de la syllabe car les nasales ne sont pas sujettes aux écarts causés par les micro-mélodies. Il est néanmoins

regrettable que ce mot ne soit pas très couramment employé. Aussi, cette étude sera effectuée, comme les autres, dans un esprit de transférabilité.

4.4.5 Contextualisation de *nobody*

Le pronom indéfini *nobody* se trouve en position complément du verbe *allow*. La longue Phrase 11 dans laquelle il est inséré est au discours direct, entre guillemets, et seule une didascalie minimale *said the Giant* fournit des indications sur la manière de la lire. Pourtant le personnage est déjà présenté physiquement comme étant d'une imposante corpulence : *Giant* est mentionné à trois reprises, doté d'une *very gruff voice* (dans la phrase précédente). Il est aussi dépeint psychologiquement comme *selfish* (dans le titre du conte), d'une intelligence limitée (Phrase 8 : nécessite sept ans pour converser). Le résultat est qu'il fait peur aux enfants qui empiètent sur son territoire au point de s'enfuir. La Phrase 11 doit être comprise dans ce contexte : elle est l'expression de la colère du personnage. Le discours y est emphatique et plusieurs mots ont des degrés d'emphase élevés. *Nobody* est celui qui l'est le plus. Les mots emphatiques les plus proches qui l'entourent sont *that* en 27^{ème} position en amont, et *myself* en 8^{ème} position *ex æquo* en aval ; cinq syllabes séparent *nobody* de ces deux mots.

4.4.6 Analyse de la mélodie tonale à partir d'une annotation manuelle (ANN-TON-N2)

L'exploration de la mélodie tonale du mot *nobody* s'est faite selon le même protocole que celui de l'expérimentation PER-TON-N sur l'interrogative à part qu'il s'agit d'un test de catégorisation et d'annotation tonale manuelle à partir de la visualisation de la courbe de F0. Les 42 lecteurs natifs du corpus anglophone ont pu être pris en compte. Les différences de F0 même minimales (non perceptible auditivement) ont été prises en considération pour cette expérimentation car son objectif est de discriminer la variation tonale en la favorisant.

Comme vu précédemment, la première syllabe du mot *nobody* se révèle particulièrement propice à une expérimentation tonale syllabique de par son accent tonique et sa composition segmentale, spécialement mis à profit par le phénomène d'emphase. Alors que l'expérimentation PER-TON-N ne faisait intervenir que les tons statique (plat), unidirectionnels (montant et descendant) et bidirectionnels (montant-descendant et descendant-montant), on observe ici l'introduction de tons tridirectionnels (montant-descendant-montant et descendant-montant-descendant). Ces tons hypercomplexes ne sont pas reconnus par l'école britannique de l'intonation et sont assimilés à des tons bidirectionnels. Pourtant, les conditions favorables à leur détection incitent à se rapprocher de

la réalité phonétique, même si elle n’est pas validée phonologiquement car l’objectif de cette expérimentation est de vérifier le seuil de complexité tonale que peut atteindre un locuteur. Seuls trois lecteurs natifs sur 42 utilisent un ton tridirectionnel sur la syllabe /n@U/.

Ces tons hypercomplexes ont été détectés chez trois lecteurs dont les copies d’écran suivantes rendent compte. Le point commun entre ces lectures est leur qualité expressive. Le lecteur pour lequel le phénomène est le plus marquant est un lecteur professionnel britannique de renom. La copie d’écran ci-dessous montre clairement le ton descendant-montant-descendant rendu possible par la visualisation de la F0 sur l’intégralité de la syllabe. On remarque en outre la durée exceptionnelle de cette syllabe tonique (0,77 seconde) par rapport aux deux suivantes, indiquant le choix de diction favorisant l’expression de la colère par un registre ample (24,49 demi-tons) et un rythme accentuel très marqué.

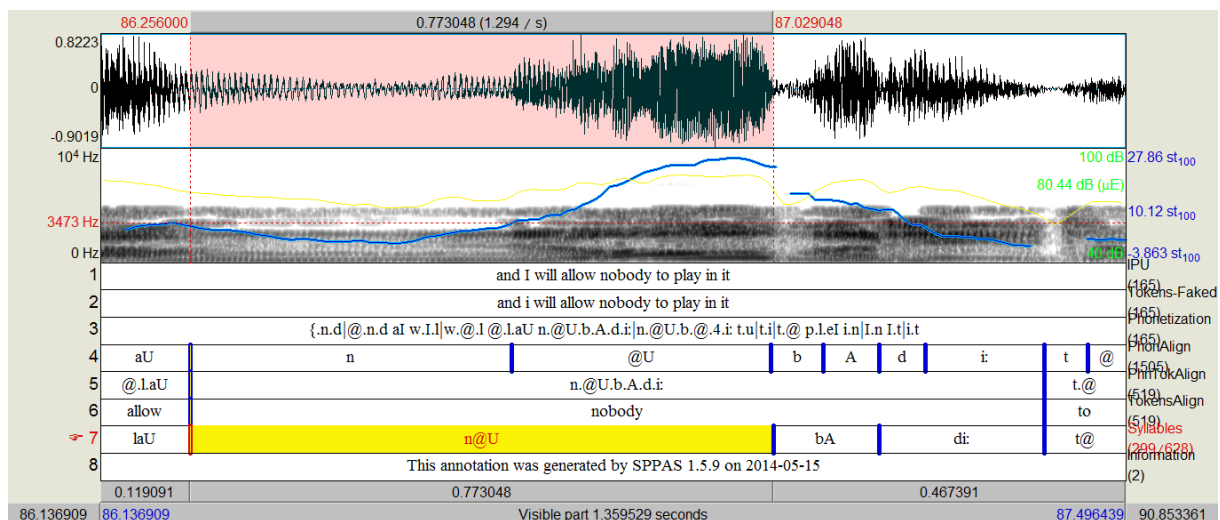


Figure 87 – S11_Nobody : ton hypercomplexe \ (lecteur SG-Pro-CK-m)

La lecture présentée dans la figure suivante est celle d’une anglophone américaine de 58 ans. La durée de la syllabe /n@U/ est moins longue (0,43 seconde) et la remontée finale peu perceptible mais présente. Pour la cohérence de la discrimination, elle a été catégorisée parmi les tons tridimensionnels, lesquels sont très rares.

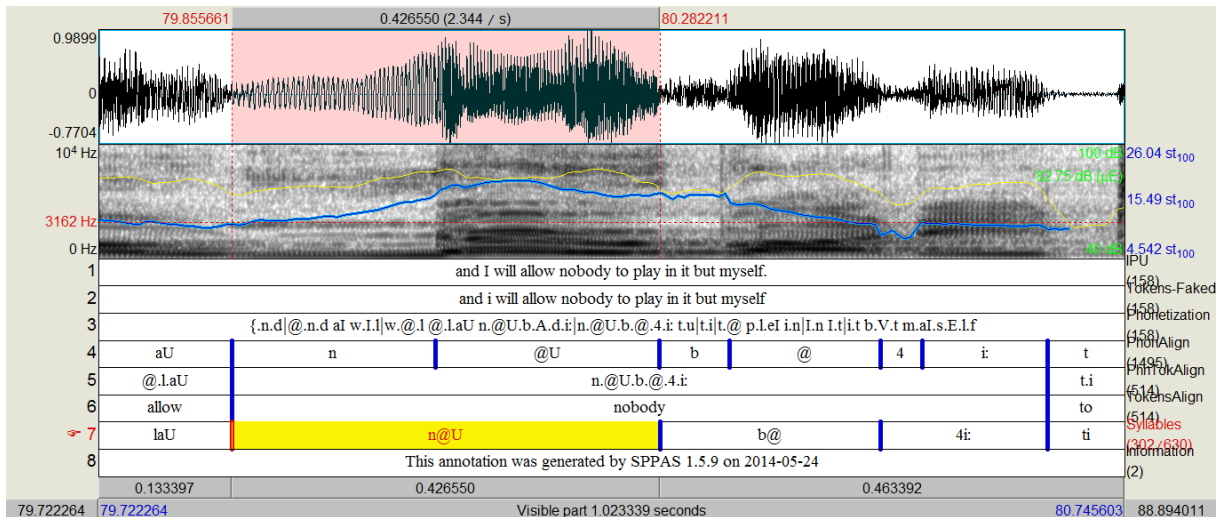


Figure 88 – S11_Nobody : ton hypercomplexe /v/ (lectrice SG-UP-A2-f-58)

La copie d'écran suivante montre une réalisation favorisant l'adéquation de la voix à la corpulence du Géant. Les indicateurs en sont non seulement la F0 très basse sur la totalité de l'énoncé malgré un registre plus ample sur la tonique, mais encore sur le rythme syllabique adopté par la durée similaire des trois syllabes. Il s'agit du lecteur SG-UP-M1-m-65, codé lecteur L3 dans l'expérimentation PER-LEC-N.

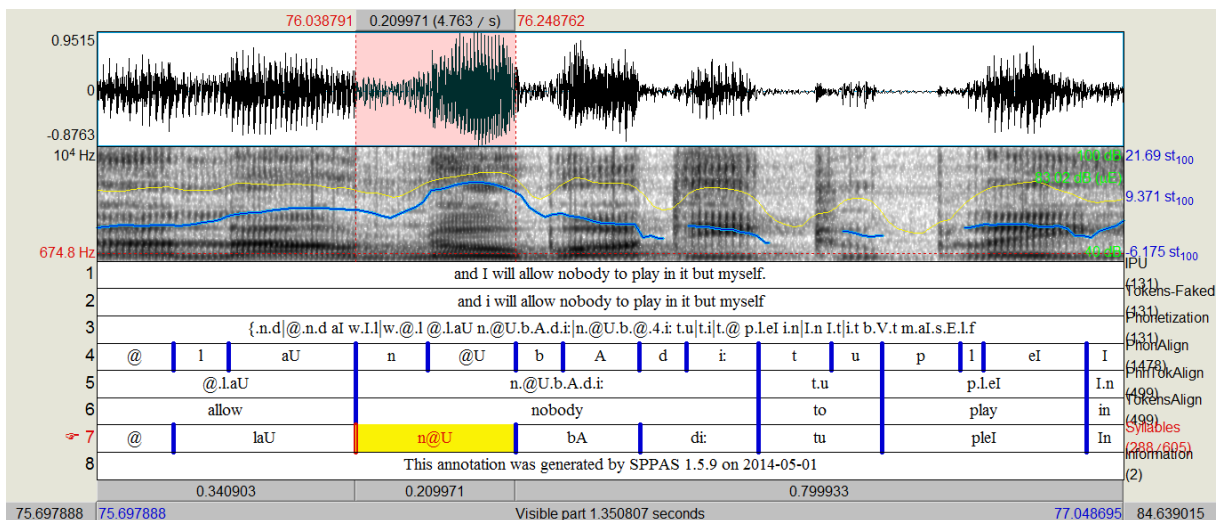


Figure 89 – S11_Nobody : ton hypercomplexe /v/ (lecteur SG-UP-M1-m-65)

Le Lecteur L1 de l'expérimentation PER-LEC-N, « meilleur » lecteur, adopte un ton bidirectionnel comparable aux choix de diction du lecteur professionnel mentionné plus haut sans être autant appuyé (rythme accentuel, durée réduite, ton moins complexe) (voir figure ci-dessous).

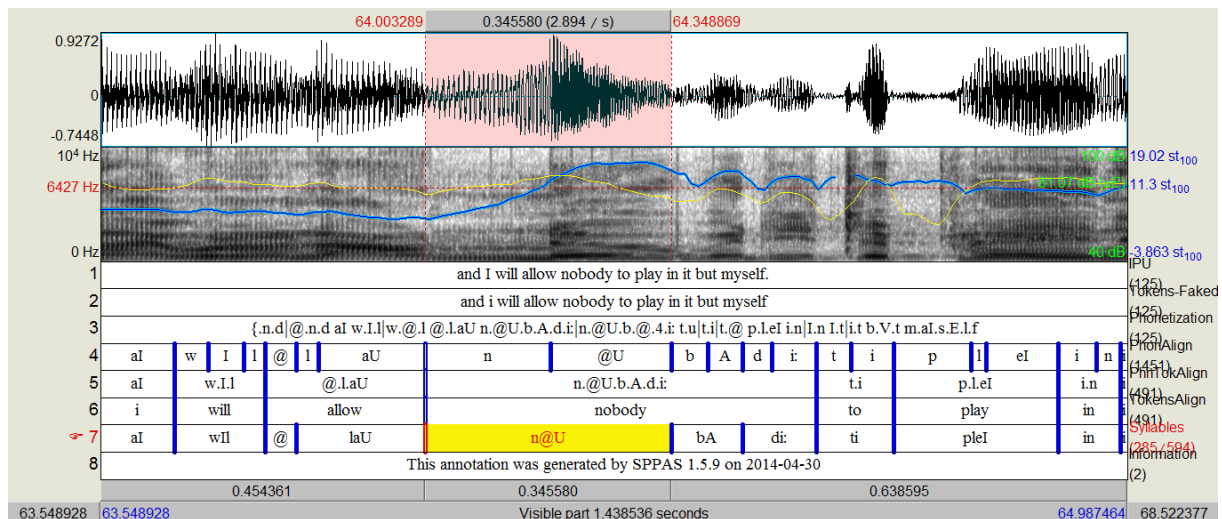


Figure 90 – S11_Nobody : ton complexe \ (lecteur L1 ; SG-UP-A1-m-56)

Les résultats de la relation de la complexité du patron tonal par rapport au registre sont illustrés par les boîtes à moustaches de la figure ci-dessous.

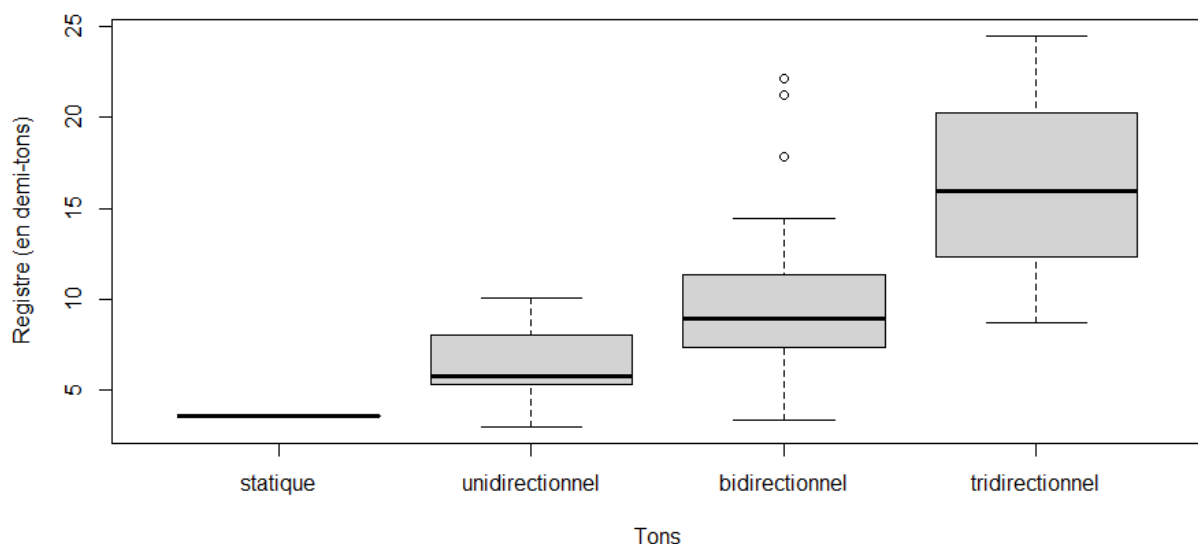


Figure 91 – Relation entre la complexité du patron tonal de la syllabe /n@U/ en fonction du registre du mot *nobody* (42 natifs)

Pour vérifier et affiner ces résultats visuels, j’ai procédé à une analyse de la régression de la complexité tonale de la syllabe /n@U/ par rapport au registre du mot *nobody*. Étant donné le fort effectif de la catégorie tonale bidirectionnelle, je la prends comme modalité de référence pour mesurer les écarts par rapport aux autres valeurs. Je constate que, par rapport au registre bidirectionnel, la catégorie tonale unidirectionnelle est en moyenne moins ample de 3,44 demi-tons (p-valeur = 0,0757 et légèrement supérieure à la valeur critique de 0,05). Par contre, la catégorie tridimensionnelle est en moyenne plus ample de 6,4207 demi-tons, soit une demi-octave (p-valeur = 0,0239). L’analyse de ces résultats tend à montrer que le choix

de la complexité du patron tonal de la syllabe /n@U/ est partiellement corrélé au registre mesuré sur le seul mot *nobody*. Il est d'ailleurs attendu qu'un ton statique n'ait pas besoin d'un registre ample pour être exécuté, alors que les tons complexes nécessitent l'inverse.

Les résultats de la relation de la complexité du patron tonal montrent le faible pouvoir discriminant des choix tonaux unidirectionnels et bidirectionnels par rapport à l'ampleur du registre. Le seuil de discrimination fondamentale se situe entre le choix de ces catégories et celui de tons hypercomplexes (ou tridirectionnels).

On peut aussi constater une relation entre le facteur durée du texte et la complexité tonale (graphiques ci-dessous). On y voit que les cas extrêmes sont bien corrélés (durée brève pour le ton statique et durée longue pour le ton hypercomplexe). Il en est de même pour le ton dynamique simple (durée intermédiaire entre les deux autres tons). Par contre, les tons bidirectionnels démentent une corrélation car la durée de lecture du texte est d'une très grande variabilité, s'échelonnant de très rapide aux durées maximales.

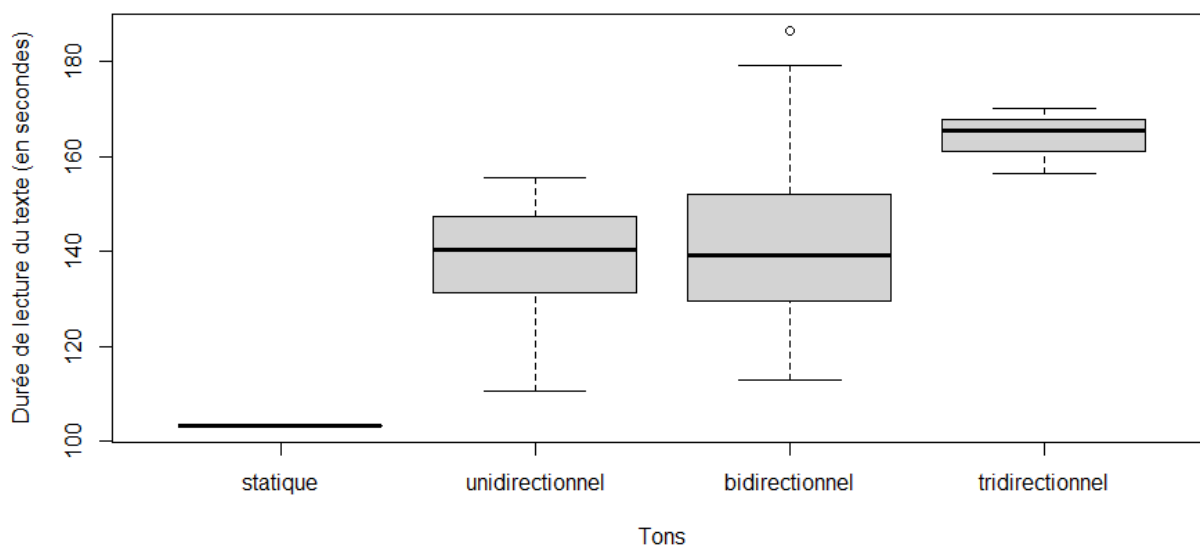


Figure 92 – Complexité du patron tonal de la syllabe /n@U/ en fonction de la durée du texte (42 natifs)

De même que pour la relation entre choix tonaux et registre, examinons l'analyse de la régression de la complexité tonale de la syllabe /n@U/ par rapport à la durée de lecture du texte. Ici aussi la catégorie tonale bi-directionnelle est choisie comme modalité de référence pour mesurer les écarts par rapport aux autres valeurs. On remarque que la catégorie tridirectionnelle est en moyenne plus longue de 20,577 secondes (p -valeur = 0,0763 et légèrement supérieure à la valeur critique de 0,05). Par contre, la catégorie statique est en moyenne plus courte de 40,207 secondes (p -valeur = 0,0407). L'analyse de ces résultats tend

à montrer que le choix de la complexité du patron tonal de la syllabe /n@U/ est partiellement corrélé à la vitesse de lecture du texte.

Les résultats de la relation de la complexité du patron tonal montrent l’écart de durée de lecture faible bien qu’effectif entre les choix complexes (bidirectionnel et tridirectionnel). L’écart de durée entre le choix de tons statiques et bidirectionnels est bien plus discriminant.

4.4.7 Bilan

Bien que cette expérimentation soit très réductrice, elle attire l’attention sur la qualité de lecture via la détection de la complexité tonale syllabique. Il semblerait que le choix de patrons tonaux syllabiques soit doté d’une moindre variabilité que celui sur des syntagmes intonatifs de plusieurs mots, comme l’interrogative étudiée en amont a pu l’exemplifier. Pourtant, les résultats de l’analyse statistique suggèrent que la durée de lecture du texte est davantage discriminante pour les faibles niveaux prosodiques alors que le degré d’ampleur du registre révélerait davantage les meilleurs lecteurs.

On a pu constater la grande variabilité de patrons tonaux pouvant être portés par la seule syllabe /n@U/. Une fois de plus, l’Hypothèse 3 selon laquelle le paramètre d’emphase favorise la variabilité des réalisations est avéré.

4.5 Bilan de l’étude des principaux marqueurs d’emphase (expérimentation COR-AC-N2)

L’expérimentation acoustique sur les quatre principaux marqueurs d’emphase du texte avait comme objectifs de trouver les marqueurs ayant la plus forte variabilité, les étudier acoustiquement et vérifier leur adéquation avec l’objectif d’évaluation de l’interlangue prosodique. Au départ, la sélection de marqueurs parmi tous les mots du texte *The Selfish Giant* a pu être opérée grâce à l’utilisation du phénomène d’emphase, l’hypothèse testée étant que l’emphase favorisait la variabilité des réalisations prosodiques.

À la suite de l’analyse acoustique, on peut affirmer que l’Hypothèse 3 a été validée pour les quatre marqueurs principaux d’emphase.

Les marqueurs retenus pour l’analyse du niveau prosodique des apprenants grâce à cette analyse sont consignés dans le tableau suivant.

	<i>Domaine analysé</i>	<i>Sous-domaines</i>	<i>Variables obtenues et unités de mesure</i>
MÉLODIE	Registre : Ampleur du registre	<ul style="list-style-type: none"> • Sur les marqueurs emphatiques : exclamative, paires <i>so</i> + adjectif/adverbe (Phrases 4 et 22) 	<ul style="list-style-type: none"> • Quantitatives (DT)
	Mélodie dans la phrase : Contour mélodique dans syntagme intonatif	<ul style="list-style-type: none"> • Fall (exclamative) • Rise-fall (<i>nobody</i>) 	<ul style="list-style-type: none"> • Quantitatives (DT) • Qualitatives (réalisation des contours conforme, plutôt conforme, non conforme)
	Mélodie dans la syllabe : Ton syllabique sur nucléus	<ul style="list-style-type: none"> • Rise-fall et tons tridirectionnels (<i>NObody</i>) 	<ul style="list-style-type: none"> • Qualitatives (statique, unidirectionnel, bidirectionnel, tridirectionnel)
RYTHME	Rythme : proéminence de marqueurs spécifiques	<ul style="list-style-type: none"> • Réalisations des paires <i>so</i> + adjectif/adverbe (Phrases 4 et 22) 	<ul style="list-style-type: none"> • Quantitatives (DT)

Tableau 52 – Sélection des variables à la suite de l'expérimentation COR-AC-N2

Poursuivons maintenant l'analyse prosodique via le phénomène d'emphase en vue de trouver les marqueurs les plus pertinents pour construire les variables destinées à déterminer les profils prosodiques d'apprenants et les traits discriminants pour évaluer leur niveau.

5 Étude de marqueurs non emphatiques ou à l'emphase réduite chez les anglophones natifs (expérimentations COR-DIV-N3)

Nous avons étudié jusqu'à présent l'emphase maximale qui permet de favoriser la variabilité. À présent, voyons l'autre face de l'emphase : celle qui permet par son absence ou son faible degré de minimiser la variabilité. C'est ce que l'expérimentation des corrélats acoustiques des natifs anglophones (COR-AC-N2) se propose à présent de faire.

Les recherches exposées dans le Chapitre II sur les profils ont signalé les tendances par lesquelles les apprenants anglicistes francophones étaient susceptibles d'être repérés pour leur faible niveau prosodique. En ce qui concerne le rythme, il s'agit notamment de l'isosyllabité, le déplacement de l'accent lexical, une localisation du martellement rythmique non attendu, la tendance à marquer une proéminence en fin de syntagme intonatif sur les « mots outils », notamment les pronoms. Sur le plan mélodique, on attend le contour

5. Étude de marqueurs non emphatiques ou à l'emphase réduite chez les anglophones natifs (expérimentations COR-DIV-N3)

descendant par défaut de la langue anglaise lorsque le contour par défaut de la langue française est une continuation. Au niveau syllabique, les francophones éprouvent des difficultés à réaliser des tons complexes sur une seule syllabe. Il s’agit alors de sélectionner les marqueurs adéquats pour étudier le phénomène chez les natifs et vérifier leur adéquation dans un objectif évaluatif. Ici encore, l’analyse du phénomène d’emphase va pouvoir aider à profiler la recherche.

5.1 La phrase “neutre” (ou au degré d’emphase minimal)

L’analyse des résultats de l’expérimentation perceptive par l’application du phénomène d’emphase a permis dans un premier temps de catégoriser et de hiérarchiser les mots et les phrases du texte *The Selfish Giant* en fonction de ce paramètre et grâce à cela de sélectionner deux phrases courtes à l’emphase réduite et de les analyser acoustiquement afin de sélectionner celle qui correspondra le mieux aux critères de stabilité.

Afin d’avoir un point d’appui stable et non attitudinal pour observer les réalisations prosodiques d’une phrase simple canonique, j’ai pu éclairer mes choix grâce à une hiérarchisation des 26 phrases selon le paramètre d’emphase. Je me propose à présent de compléter la présentation des Phrases 6 et 9 ébauchée en amont, puis d’analyser acoustiquement leurs réalisations.

5.1.1 Caractérisation linguistique générale

Parmi les quatre types de phrases que Greenbaum (1996 : 45) distingue, les déclaratives (*declarative sentences*) figurent en première position. Leur utilisation la plus courante dans le discours se résume à formuler des déclarations (*statements*). Il précise que les déclaratives représentent la catégorie la plus courante, le type de la phrase de base à partir duquel il est plus aisé de décrire les autres types. Larreya et Rivière examinent le statut de ces « constructions », basées sur la forme et non l’usage, et considèrent que deux niveaux se superposent : le premier catégorisant les quatre types de phrase mentionnés alors que le deuxième niveau consiste à éventuellement ajouter les éléments facultatifs que représentent la négation et l’emphase (2010 : 269).

5.1.2 Approches prosodique et sémantico-discursive de la phrase « neutre »

Wells présente le contour descendant comme le plus courant dans une phrase déclarative, bien que d’autres schémas intonatifs soient aussi possibles (2006 : 25) :

Although simple independent statements can take any tone, they most often have a **fall**. A fall is the **default** (= neutral, unmarked) tone for a statement. We say statements with a fall unless there is a particular reason to use some other tone. [...] by using a fall we indicate that what we say is potentially **complete** and that we express it with **confidence**, **definitely** and **unreservedly**. The fall thus also tends to signal **finality**. We call this tone meaning the **definitive fall**.

Cette position est appuyée par Roach, qui commente de même le contour descendant ([1983] 2009 : 123) :

This is the tone about which least needs to be said, and which is usually regarded as more or less “neutral”. [...] The fall could be said to give an impression of “finality”.

Huart précise les différences intonatives entre le français et l’anglais (2010 : 153) : « Ce qui diffère, dû aux systèmes accentuels des deux langues, est la ligne mélodique générale avant et après la montée ou la descente. » Le français se caractérise par une ligne relativement stable dans chaque groupe, avec des syllabes de durée sensiblement similaire, à l’exception de la dernière car elle est prononcée plus longuement et sur un ton montant jusqu’à la fin de l’énoncé. Les trois dernières syllabes descendent alors par paliers. Les schémas suivants exemplifient les deux cas possibles.

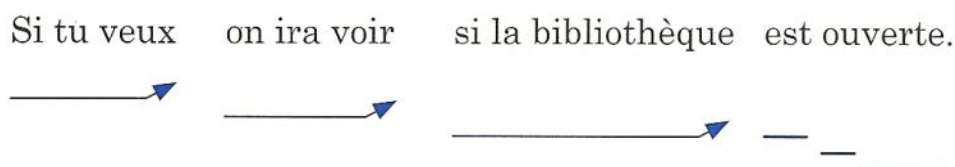


Figure 93 – Contour intonatif descendant de base en français (Huart 2010 : 154)

Huart conçoit le contour intonatif de base de l’anglais comme une sorte de « bosse » (2010 : 154) et en donne l’exemple suivant :

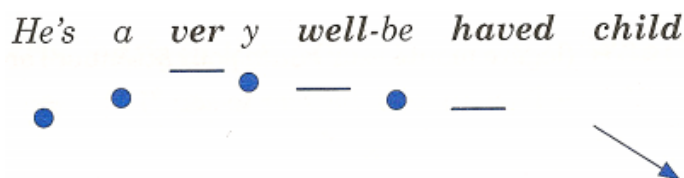


Figure 94 – Contour intonatif descendant de base en anglais (Huart 2010 : 155)

Elle précise qu’une montée, lorsqu’elle doit avoir lieu, « ne commence qu’une fois cette descente initiale effectuée » (2010 : 155), comme dans l’exemple qu’elle donne :

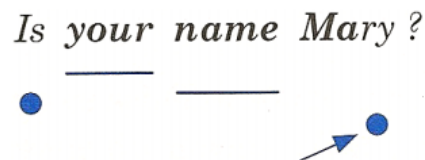


Figure 95 – Contour intonatif descendant de base en anglais avec remontée finale (Huart 2010 : 155)

Il s’agit alors d’un contour descendant-montant, annonçant que tout n’est pas encore dit : *One very typical meaning of a non-fall is **non-finality*** (Wells 2006 : 27). Devant une frontière, le *non-fall* est la plupart du temps un ton descendant-montant auquel Wells attribue le sens de *dependent fall-rise* pour signifier que *there is more material still to come* (2006 : 69). Il exemplifie le phénomène *dependent fall-rise* (ton bidirectionnel sur une syllabe) par l’exemple :

¹After v lunch | we could ¹call on Mary.

qu’il différencie de ce qu’il nomme le *dependent rise* (seul le ton montant est porté par *lunch*) :

¹After / lunch | we could ¹call on Mary.

5.1.3 Intérêt de la phrase « neutre » pour l’étude de l’interlangue prosodique

Suivant la position de Larreya et Rivière mentionnée plus haut (2010 : 269), étant donné que les statuts exclamatifs et interrogatifs ont déjà été abordés, et que le statut impératif transparait dans l’interrogative rhétorique, il est d’autant plus intéressant de vérifier la qualité intonative sur une phrase de statut déclaratif. De plus, les éléments de niveau 2 que sont la négation et l’emphase ont déjà été abordés sous l’angle prosodique. En conséquence, le choix de sélectionner une phrase dépourvue de négation et d’emphase (donc « neutre », et déclarative), est d’un intérêt syntaxique évident pour compléter l’analyse, même si le profilage mis en œuvre n’est pas prioritairement d’ordre syntaxique. On voit que le phénomène d’emphase amène les domaines prosodique et syntaxique à se rencontrer.

La phrase déclarative dénuée d’emphase permet de minimiser la variabilité afin de vérifier si les apprenants sont capables de réaliser un contour descendant, ainsi qu’une remontée adéquate (après ton grave) devant frontière. Le contour descendant serait exécuté sur plusieurs syllabes et le ton descendant-montant sur une seule.

Huart remarque qu'« En anglais, la mélodie est indissociable du rythme déterminé par la distribution des syllabes accentuées » et liste les travers récurrents de l'interlangue prosodique (2010 : 154) :

- une montée systématique en fin d'énoncé, correspondant à la question implicite : « ai-je bien répondu ? est-ce cela que vous attendiez de moi ? »
- des montées et des descentes en zigzag dans un effort pour « mettre l'expression » ou, au contraire
- une ligne mélodique monocorde due à une absence d'implication de la part du locuteur.

Suite à ces remarques, deux points semblent primordiaux : faire ressortir la courbe intonative anglaise de base en forme de « bosse » (Huart, 2010 : 154), et la ligne mélodique plus ou moins rectiligne terminant un groupe de souffle français par une montée (Huart, 2010 : 155). La phrase neutre choisie doit pouvoir mettre ces aspects en évidence.

5.1.4 Description et contextualisation

Les Phrases 6 et 9 sont candidates à l'analyse acoustique de la mélodie. Elles diffèrent par leur longueur (P6 = sept syllabes en six mots et P9 = quinze syllabes en onze mots), leur degré moyen d'emphase (P6 = 4^{ème} rang des phrases les plus emphatiques avec 7,19% ; P9 = 9^{ème} rang avec 2,02% sur les 26 phrases du texte), leur degré de complexité syntaxique (P6 n'a qu'une relation prédicative alors que P9 en a trois).

Dans une analyse syntaxique en constituants immédiats, les deux syntagmes nominaux et le syntagme verbal, constitutifs principaux de la phrase, se rattachent au nœud supérieur, la Phrase 6, comme le montre le schéma suivant.

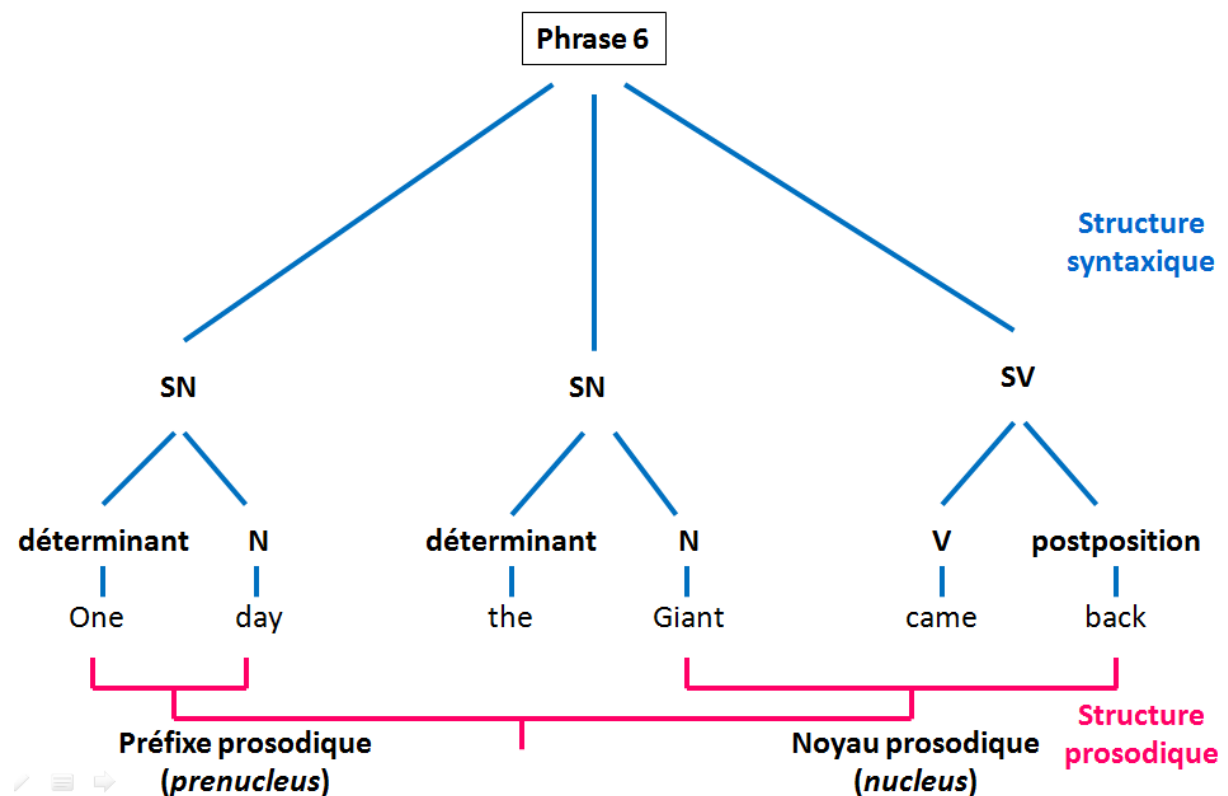


Figure 96 – Structuration syntagmatique et prosodique de la Phrase 6

Dans cette projection effectuée à partir des degrés d’emphase répertoriés pour chaque mot, la structure prosodique est congruente (Martin 2009 : 20, 105-106) à la structure syntaxique comme le montre le schéma. Ainsi, la Phrase 6 comprendrait deux groupes prosodiques, représentés par le préfixe prosodique et le noyau prosodique, et comporterait deux types de frontières :

- Frontière majeure après le noyau, marqué graphiquement à la fin de la phrase par un point,
- Frontière mineure entre le préfixe prosodique et le noyau prosodique (position attendue après *day*).

La Phrase 9 est une phrase complexe par sa structuration syntaxique. Une proposition complément est enchâssée dans la proposition principale avec modification de sa structure de surface (élision de l’auxiliaire *were*). Au niveau prosodique, en se basant sur les prédictions des degrés d’emphase par mot, on peut prévoir deux frontières encadrant deux contours de modalité déclarative. Celui du préfixe prosodique peut se terminer par une remontée pour indiquer la frontière mineure qui le suit. Une frontière majeure est attendue en fin de noyau prosodique. La tête du syntagme intonatif est attendue sur *saw*, premier mot de classe ouverte, en tant que verbe. Le contour comprenant le noyau s’établit sur une suite de onze syllabes,

c'est pourquoi une proéminence sur *playing* est aussi envisageable, selon le débit d'élocution du lecteur (Martin, 2009 : 105).

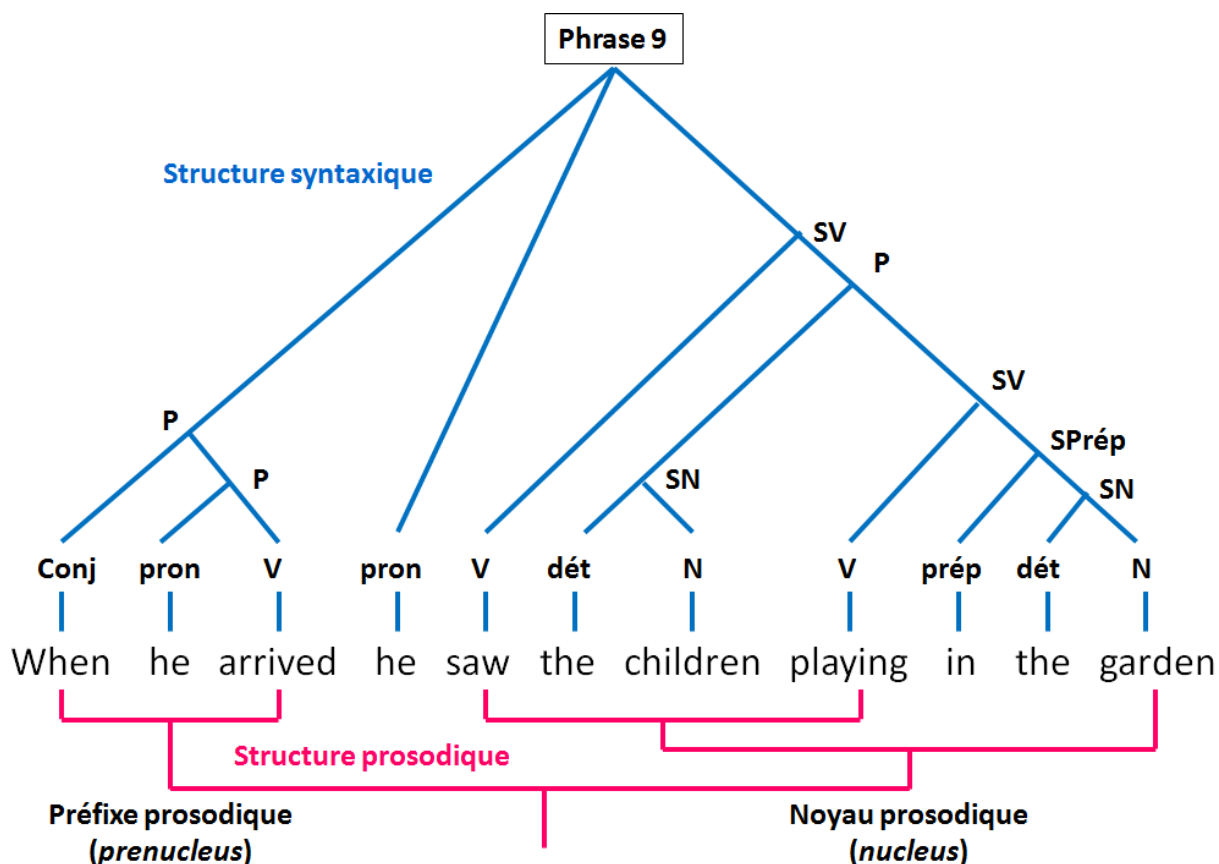


Figure 97 – Structuration syntaxique et prosodique de la Phrase 9

Les deux phrases sont narratives. La Phrase 9 conclut un paragraphe en annonçant la suite diégétique. L'information n'est pas particulièrement nouvelle étant donné que l'on sait déjà que les enfants jouent dans le jardin et que son propriétaire va rentrer chez lui. Par contre, la Phrase 6, située à la charnière de deux paragraphes, réoriente sémantiquement la description édénique vers le chaos. Un bon lecteur se doit de le faire sentir, ne serait-ce que par le rythme. Peut-être est-ce ce qui donne à cette phrase un pourcentage d'emphase de 7,19%, ce qui est relativement élevé par rapport aux autres phrases du texte.

5.1.5 Mélodie dans la phrase « neutre »

Suite aux hypothèses émises à partir des degrés d'emphase et des considérations contextuelles précédentes, les copies d'écran sur les deux bons lecteurs L1 et L5 renseignent sur leur validation. La Phrase 6 lue par L5 montre clairement le mouvement *fall-rise* de la F0 sur la syllabe *day*, et le laisse deviner sur L1 malgré une F0 dévoisée sur /d/. Il s'agit là d'une tendance acoustique, peut-être insuffisamment marquée pour être clairement perceptible et

CHAPITRE III – Emphase : mode d’organisation paradigmatique des réalisations prosodiques des anglophones natifs

être formellement étiquetée en tant que « ton » puisque l’acoustique concerne les phénomènes continus et les tons étant par nature catégoriels.

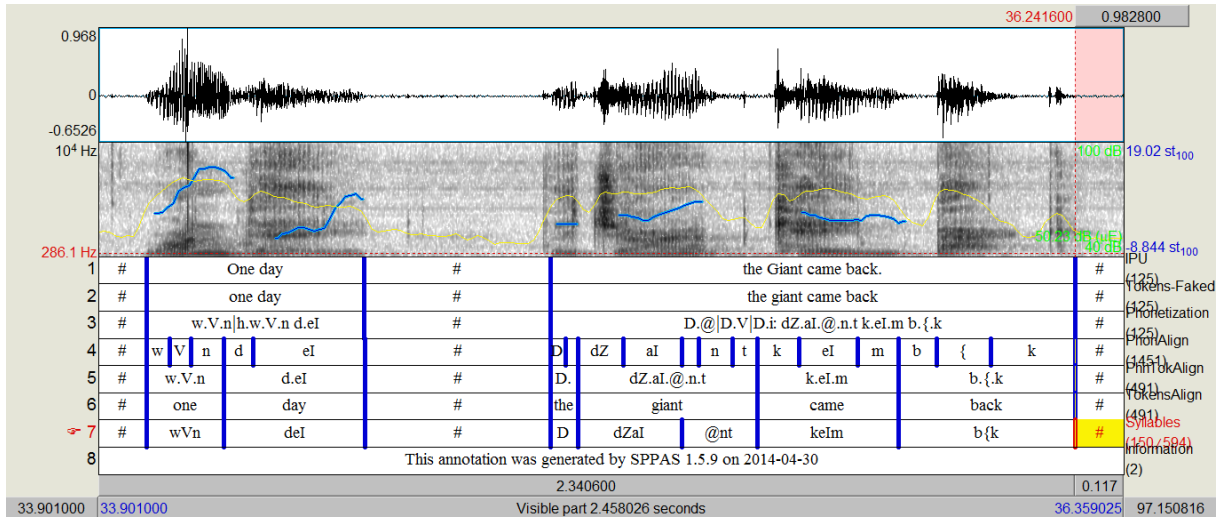


Figure 98 – Spectrogramme de la Phrase 6 par le Lecteur L1 (v1) (SG-UP-A1-m-56)

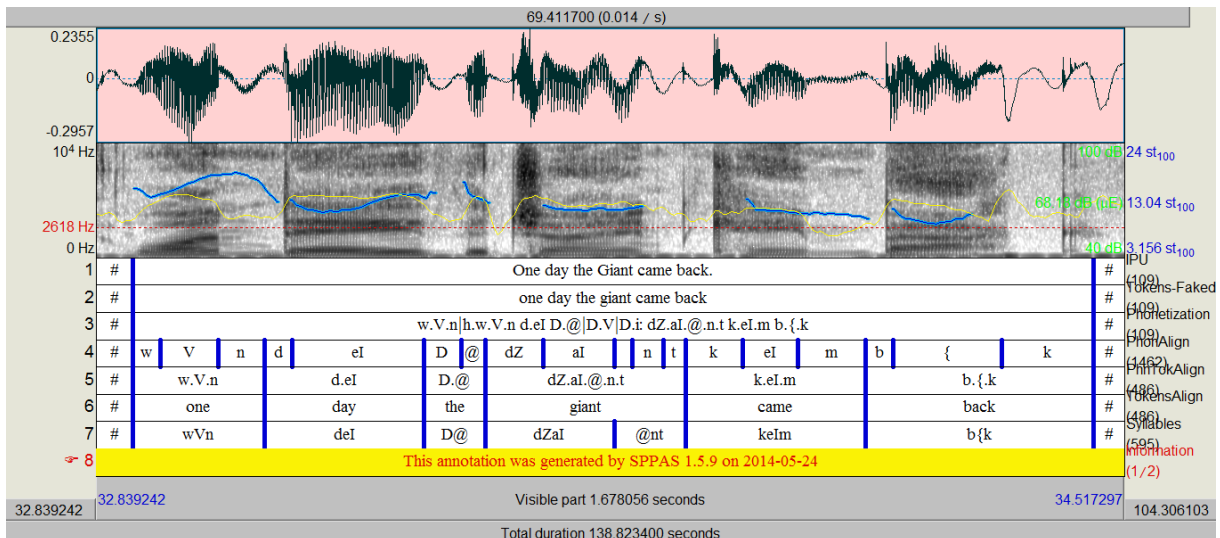


Figure 99 – Spectrogramme de la Phrase 6 par le Lecteur L5 (SG-PD-He1-f-23)

Le contour descendant sur la fin de la phrase est effectif dans les deux cas, bien que l’on ne puisse voir le noyau de L1, l’écoute fait entendre un ton final bien plus grave que celui de *came*. La deuxième syllabe de *Giant* est dotée d’une F0 légèrement plus haute que celle de la première et de ce fait interdit de valider une chute canonique catégorique d’après des critères acoustiques stricts. Je me situe ainsi non plus dans une optique de lien consensuel entre acoustique et perception, mais dans la seule optique acoustique brute et rigoureuse dont la modélisation évaluative de base pourra tenir compte à ce stade.

L'étude de la Phrase 9 montre ainsi que le *fall-rise* est amorcé dans les deux cas, toujours selon mes critères. La chute l'est aussi, mais L1 constitue une série de tons plats se terminant par une chute prononcée sur *garden*, alors que L5 prend *saw* comme tête à partir de laquelle la courbe descend par paliers. On pourrait d'ailleurs considérer que la chute se scinde en deux unités intonatives (frontière après *children*), avant une légère remontée sur *playing*, ce qui ne remet pas fondamentalement en cause son principe.

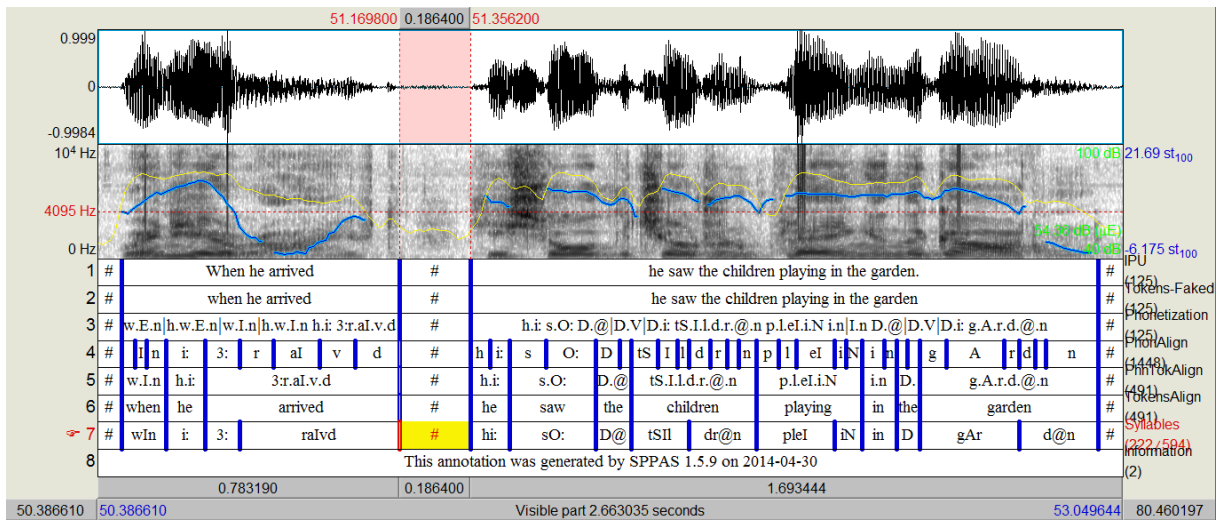


Figure 100 – Spectrogramme de la Phrase 9 par le Lecteur 1 (v1) (SG-UP-A1-m-56)

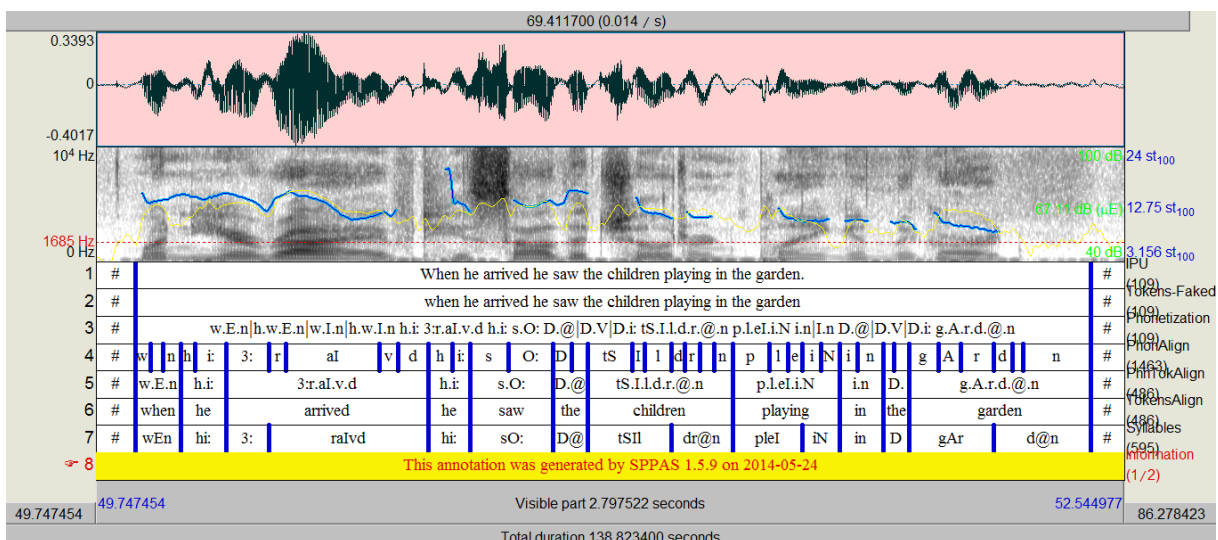


Figure 101 – Spectrogramme de la Phrase 9 par le Lecteur 5 (SG-PD-He1-f-23)

CHAPITRE III – Emphase : mode d’organisation paradigmatique des réalisations prosodiques des anglophones natifs

Pour départager les deux phrases, une annotation manuelle a été faite à partir de la courbe de F0 pour vérifier le nombre de lecteurs qui produisaient un ton descendant-montant sur les syllabes avant frontière (*day*, *-rived*). Les résultats figurent dans le tableau ci-dessous.

42 lecteurs	Phrase 6		Phrase 9	
Tons	V	\	V	\
Jeunes lecteurs	8	14	17	22
Lecteurs mature	7	10	9	11
Lecteurs professionnels	3	3	1	3
Total	18	27	27	36
Pourcentage	42,85	64,28	64,28	85,71

Tableau 53 – Variation tonale des 42 natifs sur les Phrases 6 et 9

On peut voir que les patrons attendus ne sont pas toujours réalisés par les locuteurs. Malgré sa simplicité syntaxique et son faible nombre de syllabes, la Phrase 6 totalise un score de variabilité bien plus important que celui de la Phrase 9. En référence à l’hypothèse testée dans ce chapitre, on peut croire que cette différence inattendue est due à l’importance du degré d’emphase de la Phrase 6, lequel avait prédit une grande variation. En conséquence, c’est la phrase la plus stable chez les natifs anglophones qui sera prise comme référence pour l’étude des apprenants francophones.

5.1.6 Conclusion

Étant donné le souci de minimiser l’emphase et l’absence de négation, la phrase déclarative simple devrait d’autant mieux correspondre au patron par défaut qu’elle est narrative et exempte de discours direct. Les deux Phrases 6 et 9 ont été repérées pour vérifier les patrons tonaux descendants sur syntagme intonatif et descendants-montants sur syllabe devant frontière mineure. Contrairement aux prévisions en matière de longueur et de complexité syntaxique, mais en accord avec les prédictions du degré d’emphase, la Phrase 9 s’est révélée être la plus stable après analyse par annotation manuelle des réalisations des 42 natifs. Aussi, c’est la Phrase 9 qui sera prise comme modèle de référence pour l’étude de la prosodie des apprenants.

5.2 Décrochage mélodique sur *all*

On a vu jusqu'à présent que l'emphase amplifiait la variabilité, surtout mélodique. On a aussi pu constater que l'absence complète d'emphase, lorsqu'elle était possible, ne représentait pas grand intérêt. Par contre, le contraste favorisant un degré d'emphase moyen par rapport à une absence d'emphase, comme l'a révélé l'analyse de la phrase « neutre », permet de repérer des marqueurs précieux pour l'analyse. Aussi, le marqueur *all*, récurrent dans trois phrases, est digne d'attention.

Dans la Phrase 12, *all* occupe le 21^{ème} rang d'emphase avec le degré moyen d'emphase de 1,39 (dans *all round it*), dans la Phrase 8 le 43^{ème} (degré de 1,09 avec *all that he had to say*) et dans la Phrase 16 le 79^{ème} (0,74 dans *all over the country*). Selon la méthodologie adoptée avec l'intensificateur *so*, je vais tenter de sonder la liaison de ce marqueur avec l'emphase pour en détecter ses potentialités et baliser ses limites.

5.2.1 Généralités sur l'opérateur *all*

Lapaire et Rotgé ([1991] 1998 : 195) reconnaissent la difficulté de classer tous les rôles pouvant être endossés par l'opérateur *all* car il peut avoir un fonctionnement adjectival ou déterminatif, pronominal, nominal, adverbial et de préfixe. Les catégories représentées dans le texte étudié étant pronominales et adverbiales, ce sont celles-ci qui seront abordées.

5.2.2 Approche syntaxique liée à *all*

L'adverbe *all* modifie l'adjectif, l'adverbe ou le participe passé qui le suit. C'est l'approche généralement acceptée (Huddleston et Pullum 2002 : 573, Biber *et al.* 1999 : 275) :

She looked *all* upset. (Downing et Locke [1992] 2006 : 490)

They had begun [all enthusiastically]. Huddleston et Pullum (2002 : 573)

Quirk *et al.* les nomment (1985 : 447) *emphasizers* car ils *add to the force (as distinct from the degree) of the adjective* et remarquent que, contrairement aux intensificateurs compatibles uniquement avec des adjectifs de degré, les *emphasizers* le sont avec les adjectifs non graduables. Ils en donnent l'exemple : He looked *all* <confused>.

Le pronom *all* est non personnel (*nonpersonal*) (Quirk *et al.* 1985 : 381), et peut avoir diverses références : situationnelle large ou « universelle », ou explicite. Dans le cas où *all* est suivi d'une subordonnée relative zéro (en \emptyset) ou en *that*, Lapaire et Rotgé proposent de le

paraphraser par l’expression *the only thing that...* et ils en donnent un exemple : *ALL Ø_{REL} you have to do to be strong is stop being weak* ([1991] 2002 : 200). Quirk *et al.* notent cependant qu’il peut être l’équivalent de *everything* (1985 : 382) :

All
Everything } is not lost.

5.2.3 Approche sémantico-discursive de *all*

Lapaire et Rotgé notent le sens de « totalité » et « complétude » véhiculé par *all* : *you’re ALL wrong / You’re COMPLETELY wrong* ([1991] 2002 : 169). Ils mentionnent la conséquence de ce « très haut degré de » en ajoutant que « Ce qui est “total” ou “complet” se présente à l’esprit comme la **limite extrême**, le **maximum**. ». Selon eux, les notions de « **totalité**, **temporalité**, **indifférenciation** et **intensification** sont étroitement liées, chaque contexte se chargeant de les hiérarchiser et de les doser différemment ». Ils nous font percevoir le point commun entre les marqueurs d’emphase repérés jusqu’à présent en expliquant leur fonctionnement sémantique selon ce nouveau point de vue :

Ce phénomène de **glissement de la cardinalité à la modalité**, ou plus précisément du “complet”, “total”, “intégral”, “englobant” à l’“extrêmement” ou au “parfaitement” se conçoit aisément. La “perfection” exploite la double idée “rien ne manque” / “impossible d’avoir ou de faire plus” contenue dans la totalité. L’intensification (“extrêmement”), quant à elle, s’appuie sur l’idée de limite atteinte, elle aussi présente dans la totalité : “aller jusqu’à l’extrémité de”, “pousser au maximum”.

Bien que doté d’un degré d’emphase relativement faible par rapport au *so*, *how*, *what*, *nobody*, etc., on peut percevoir un air de famille entre ces marqueurs. Il n’est pas surprenant que certains aspects de leur fonctionnement, comme la nécessité d’un autre élément pour faire porter leur effet, soient similaires. Dans cette fonction, il est très proche des adverbes intensificateurs comme *so* et *very*, mais alors que ces derniers ne sont compatibles qu’avec des adjectifs graduables, *all* accepte les adjectifs graduables et non graduables.

L’unité à laquelle *all* adverbial vient s’adjoindre peut véhiculer une notion générale de lieu, par exemple *all round*. *All over* signifiera « parTOUT, JUSQUE dans le moindre recoin... » (Lapaire et Rotgé [1991] 2002 : 205). Il sera alors difficile de distinguer entre le « total » et l’« extrême ».

5.2.4 Synthèse : Intérêt de *all* dans l'étude de l'interlangue prosodique

On a vu que le marqueur *all*, ici adverbe et pronom, est doté d'un degré d'emphase variable selon ses occurrences, mais toujours présent à divers degrés sans toutefois focaliser l'emphase au détriment des entités contextuelles. Son degré émotionnel moyen permet de l'utiliser largement dans tout type de discours oral. Aussi, ce mot de classe fermée (« mot outil ») est relativement courant. Une troisième remarque en sa faveur est sa nature segmentale : sa seule syllabe se compose d'une voyelle longue /o:/ suivie de la liquide /l/, lesquelles sont de durée conséquente. Ainsi, grâce au voisement des deux segments, l'analyse de la F0 pourra être effectuée sur l'ensemble de la syllabe. Pour ces raisons, le marqueur *all* est particulièrement intéressant pour l'analyse de l'interlangue.

5.2.5 Analyse contextuelle des occurrences de *all*

L'opérateur *all* apparaît à trois reprises dans le texte dans des phrases déclaratives et narratives.

5.2.5.1 « *All* » + relative en « *that* »

Phrase 8 : After the seven years were over he had said **all that** he had to say, for his conversation was limited, and he determined to return to his own castle.

Dans cette phrase de 29 mots, *all* est un pronom complément du verbe *had said*. Il est suivi de la relative déterminative introduite par *that*. Pouvant être glosé par *everything*, il totalise un degré d'emphase de 1,10. *All* est utilisé ici pour railler la performance du Géant qui a mis sept ans pour venir à bout d'une conversation amicale. La moquerie est renforcée par le paradoxe entre d'une part la durée gigantesque de la conversation et d'autre part l'incapacité du Géant à tenir une conversation riche, selon la teneur de l'incise qui suit. Le Géant est ainsi présenté comme dépourvu de vivacité et d'intelligence⁸⁰.

5.2.5.2 « *All* » + préposition de lieu « *round* »

Phrase 12 : So he built a high wall **all round** it, and put up a notice-board.

En comptant *notice-board* pour un seul mot, cette phrase comprend quatorze mots. Elle relate les agissements du Géant suite à l'intrusion des enfants dans son jardin. Son rôle est d'introduire une qualité emphatique en intensifiant la préposition de lieu *round*. *All round it* implique que le jardin se caractérise par l'absence complète d'issue pour y pénétrer.

⁸⁰ Un niveau de décodage sémantique supplémentaire pourrait être ajouté. Selon N. Ballier, Wilde pourrait faire preuve d'une ironie plus complexe si l'on tient compte du fait qu'il n'est pas donné à tout le monde de tenir une conversation de sept ans.

5.2.5.3 « All » + préposition de lieu « over »

Phrase 16 : ... Then the Spring came, and **all over** the country there were little blossoms and little birds.

Dans cette phrase de seize mots, *all* intensifie la préposition de lieu *over*. Cette phrase marque une rupture entre le calme du jardin et le dynamisme du renouveau printanier, laquelle est reprise dans l’opposition entre *all over the country* traduisant l’universalité du processus et *only* de la phrase suivante marquant l’isolement du Géant.

5.2.6 Approche acoustique de *all*

L’analyse mélodique a été effectuée à partir des réalisations des deux « bons » lecteurs L1 et L5 en mesurant la F0 maximale des voyelles des syllabes adjacentes et de celles les entourant marquées d’une proéminence. La F0 de *all* est faite sur l’intégralité de la syllabe. Le Lecteur L1 ayant effectué deux enregistrements, j’ai utilisé les deux versions pour une analyse interlocuteurs (L1 et L5) et intra-locuteurs (L1v1 et L1v2). C’est la version L1v1 qui avait été utilisée dans l’expérimentation PER-EMPH-N sur la détection de l’emphase.

5.2.6.1 « All » + relative en « that »

Cette phrase place *all* au 43^{ème} rang d’emphase.

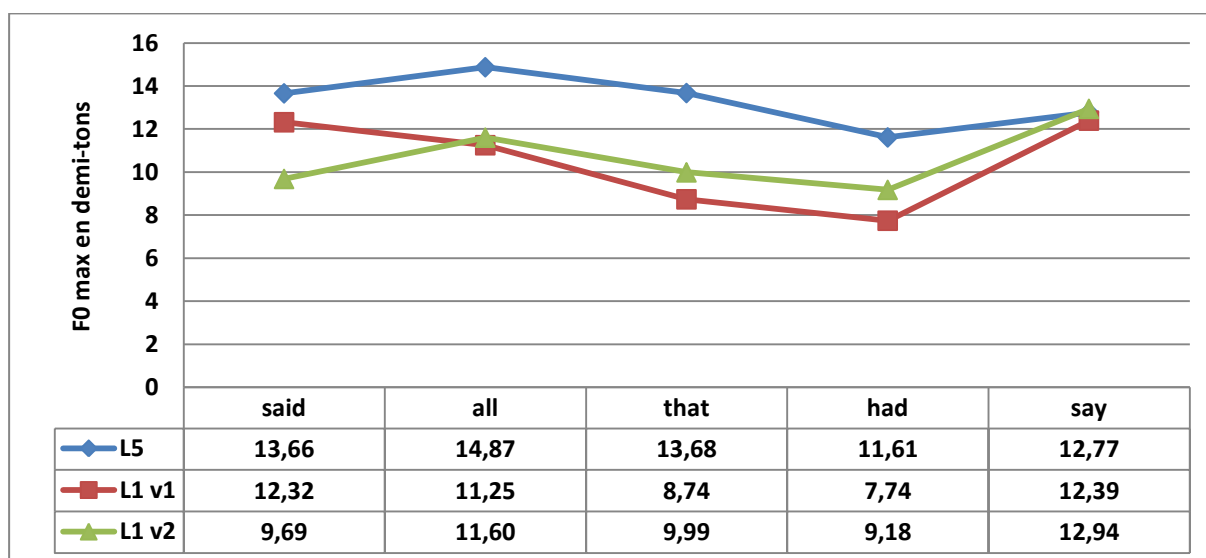


Figure 102 – *All + relative en that* : contour mélodique normalisé (F0 max)

Les trois courbes mélodiques sont semblables à l’exception de L1v1 sur le mot *said*. Alors que L5 et L1v2 utilisent une F0 max plus basse que pour *all*, L1v1 choisit de placer un ton descendant sur *said*, ce qui rehausse la mesure de la syllabe par rapport à *all*, faisant peu de cas de son ton montant-descendant. La différence mélodique semble provenir d’une

accentuation de *said* par L1v1, alors que L5 attend *all* pour marquer une proéminence (voir figures ci-dessous).

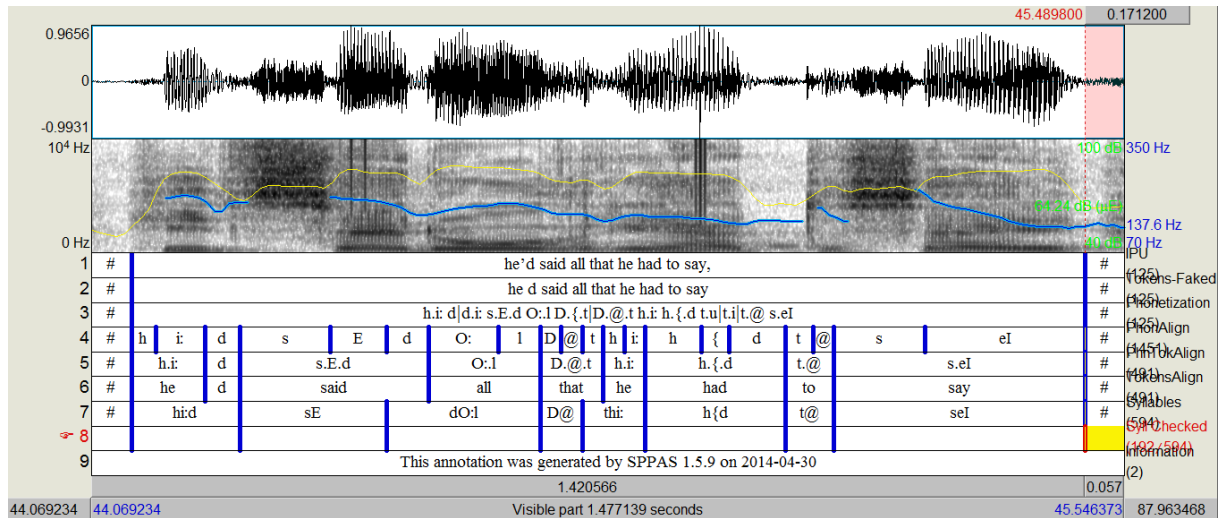


Figure 103 – *All + relative en that* : contour mélodique de L1 (v1) (SG-UP-A1-m-56)

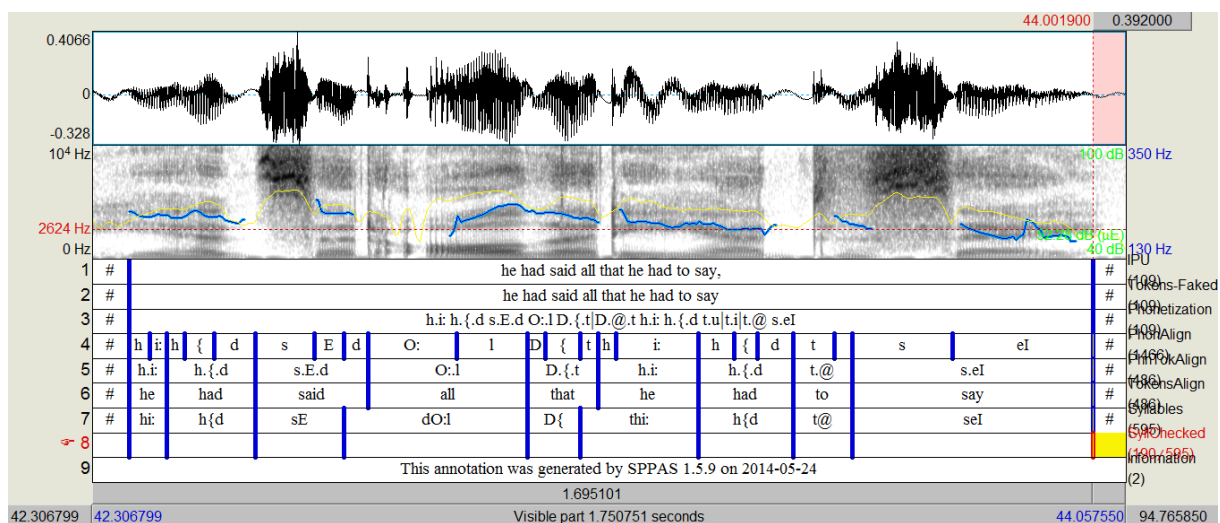


Figure 104 – *All + relative en that* : contour mélodique de L5 (SG-PD-He1-f-23)

5.2.6.2 *All + préposition de lieu « round »*

La nature voisée des segments a permis une prise de mesures sur /O:l/ dans les mots *wall* et *all*, et sur /aUn/ dans *round*, ou *around* (lu ainsi par L5). Le résultat ne fait que renforcer la conclusion de l'analyse en minimisant la différence entre les mesures de F0max.

Il n'est pas étonnant que cette phrase place *all* au 21^{ème} rang d'emphase étant donné que le contour mélodique est le même dans les trois réalisations. L1v1 et L5 produisent une pause après les six syllabes du début de phrase et peuvent ainsi débiter un nouveau syntagme

intonatif pour le terminer à l’incise (voir exemple par copie d’écran ci-dessous). *All* reçoit alors toute la prééminence nécessaire sans rivalité d’éventuelles alternatives accentuelles.

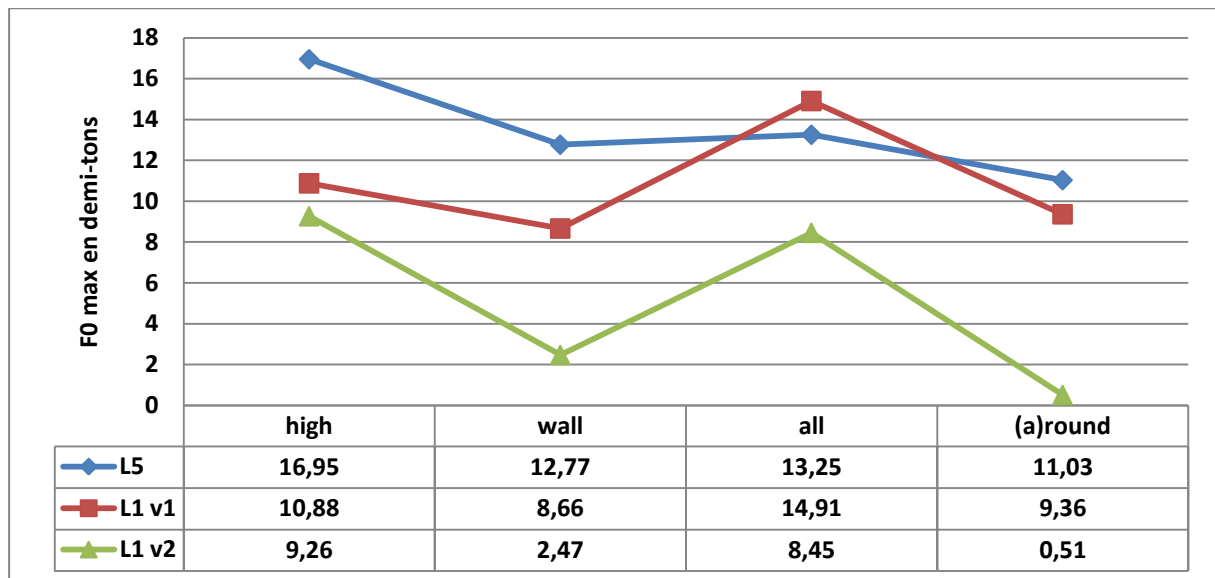


Figure 105 – *All* + préposition de lieu *round* : contour mélodique normalisé (F0 max)

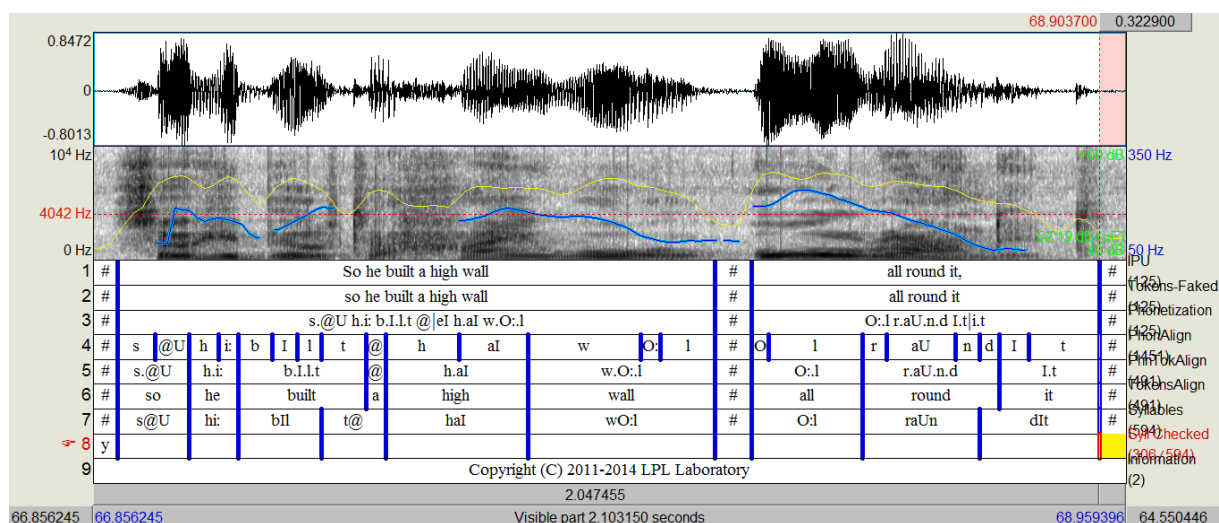


Figure 106 – *All* + préposition de lieu *round* : contour mélodique de L1 (v1) (SG-UP-A1-m-56)

5.2.6.3 « *All* » + préposition de lieu « *over* »

Cette phrase place *all* au 79^{ème} rang d’emphase. L’éloignement de *spring* à trois syllabes de *all* n’influe pas sur son accentuation ni sa mélodie. Pourtant on remarque une variation : alors que L1v2 et L5 placent une F0 max sur la première syllabe de *over*, L1v1 choisit *all*. L’ambivalence du choix rappelle celui constaté sur *so* + adjectif ou adverbe. Dans ce cas, le modèle se ferait à partir de la forme *all* + préposition. Une octave étant composée de

douze demi-tons, le registre de cette phrase pour L1v1 avoisine les deux octaves (différence de F0max entre *and* et *all* = 21,28 DT) alors que les versions de registre avec F0max plus élevée sur *over* sont plus réduites.

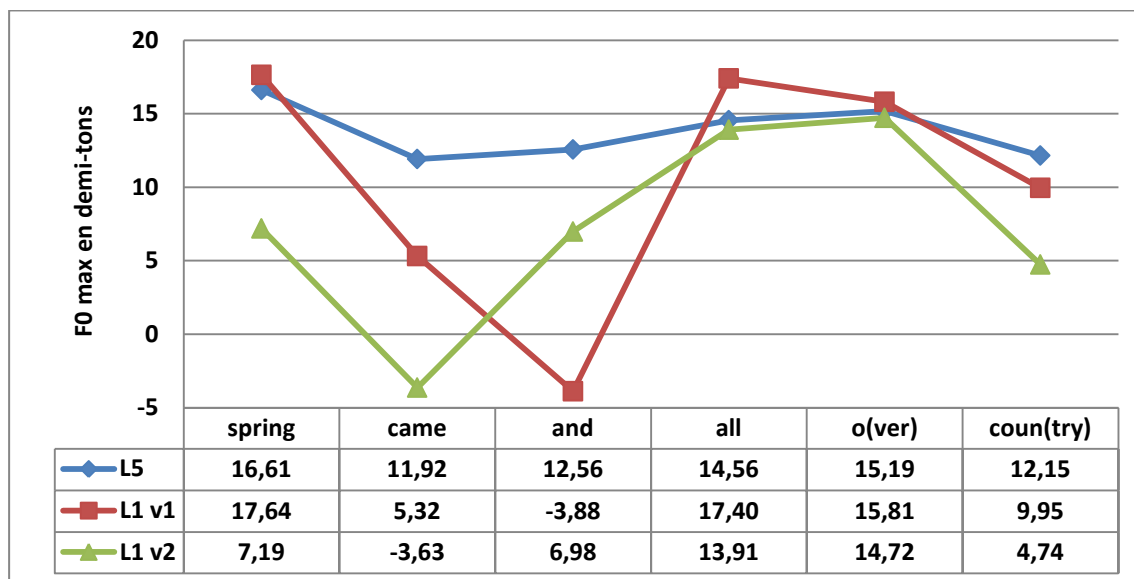


Figure 107 – *All* + préposition de lieu *over* : contour mélodique normalisé (F0 max)

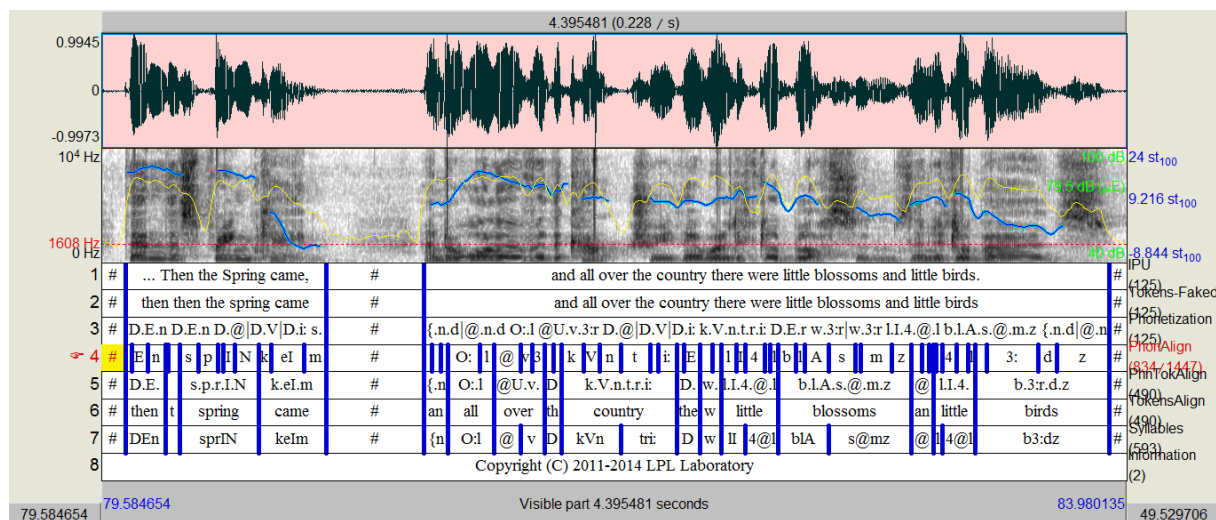


Figure 108 – *All* + préposition de lieu *over* : contour mélodique de L1 (v1) (SG-UP-A1-m-56)

5.2.7 Conclusion

Cette analyse acoustique sur la fréquence fondamentale maximale a permis d'étudier une variation inter et intra-locuteurs du pronom et de l'adverbe *all*. La Phrase 8 (*all that*) recèle une ambivalence rythmique qui rend le marqueur *all* instable dans sa dimension accentuelle et donc mélodique. Pour cette raison, il ne devrait pas être proposé à l'étude de

l’interlangue. La Phrase 12 (*all round*) représente une candidate idéale par son absence de dispersion, comme le montre son haut degré d’emphase. La Phrase 16 (*all over*) rappelle la paire *so* + adjectif ou adverbe pour l’ambivalence de ses résultats. Si ce schéma se vérifie, cette variabilité pourrait induire un trait distinctif à un niveau d’acquisition plus élevé que celui de la Phrase 12.

5.3 Désaccentuation pronominale en fin d’énoncé

Les pronoms appartiennent à la classe fermée des « mots outils ». Procédant par anaphore, ils ne sont pas accentués puisqu’ils représentent une information connue. Seule la Phrase 4 correspond au cas de figure recherché.

<i>Ordre d’apparition dans le texte</i>	<i>Nombre de mots</i>	<i>Nombre de syllabes</i>	<i>Texte de la phrase</i>
Phrase 4	24	28	The birds sat on the trees and sang so sweetly that the children used to stop their games in order to listen to them .

Tableau 54 – Phrase candidate à l’étude de la désaccentuation d’un pronom en fin d’énoncé

Cette longue phrase ne comprend aucun signe de ponctuation : le lecteur peut l’interpréter sans se conformer à un découpage préétabli et a toute latitude pour exprimer ses choix de diction. Le degré moyen d’emphase de la phrase est de 0,21 et le degré maximal moyen par mot portant sur l’adverbe *so* est de 1,70, alors que les autres mots de niveau inférieur ne totalisent pas plus de 0,77, comme c’est le cas de *sweetly*. Cette neutralité globale devrait limiter la variabilité malgré la longueur de la phrase et fournir la stabilité attendue pour ce type d’analyse. Le pronom personnel *them* se trouve en 141^{ème} position d’emphase *ex æquo* avec douze autres mots. Il a pour référent les oiseaux mentionnés en début de phrase.

Dans une telle phrase, le noyau doit être porté par la première syllabe de *listen*, dernier mot de classe lexicale, et les trois dernières syllabes de la phrase doivent constituer une queue basse. Reprenant les caractéristiques de l’annotation de Guierre ([1984] 1992 : 25), le système d’annotation utilisé comprend deux niveaux : proéminence forte et accent nucléaire : /1/, et absence de proéminence : /0/. Le schéma attendu, basé sur le degré d’emphase moyen de chaque mot, est :

[...] to stop their games in order to listen to them.
 0 2 0 1 0 0 0 0 1 0 0 0

Figure 109 – Schéma de proéminences attendues, basé sur le degré d’emphase de chaque mot

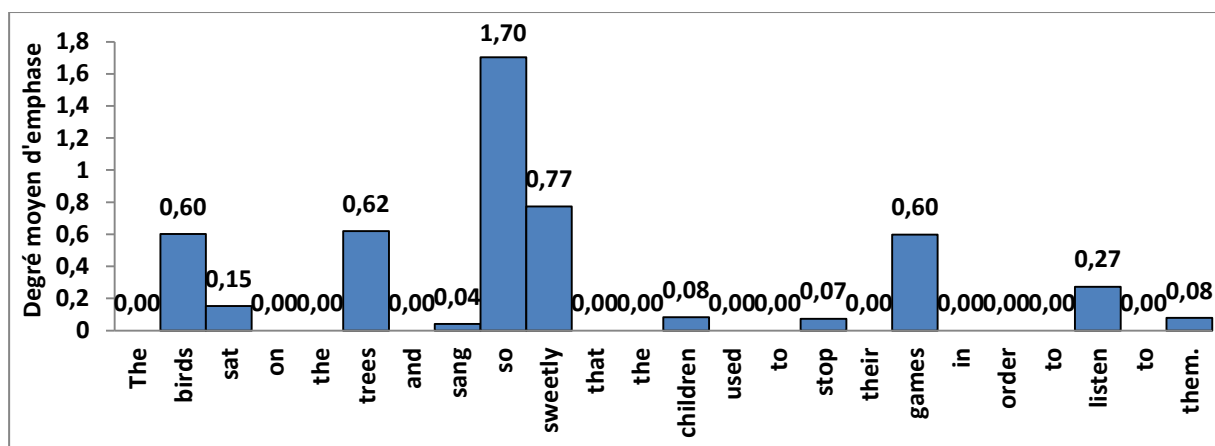


Figure 110 – Degrés moyens d'emphase des mots de la Phrase 4

Dans la figure ci-dessus, le degré moyen de 0,078 sur le mot *them* résulte du degré 1 pointé à deux reprises par un des auditeurs sur les lectures de LA (L1) et LB (L6). Il a annoté une absence totale de proéminence pour la lecture de *them* par LC (L10). Tous les autres auditeurs ont noté un degré 0.

La syllabe /D@m/ est prononcée par L1 et L5 avec la F0 la plus basse de la Phrase 4, ainsi que le montrent leurs mesures dans le tableau de synthèse : la différence de F0 entre la mesure la plus basse de Mean F0 sur voyelle et celle de *them* est toujours négative.

Mesures de F0 de P4 en demi-tons	Mesures de Mean F0 sur voyelles les plus basses		Mean F0 /D@m/	Différence de F0
L1	<i>in</i>	<i>to (them)</i>	<i>them</i>	
	3,607	1,159	-2,107	-3,266
L5	<i>trees</i>	<i>-ly</i>	<i>them</i>	
	11,139	10,749	9,554	-1,146

Tableau 55 – Vérification que Mean F0 sur voyelles dans la Phrase 4 est la plus basse sur *them*

La difficulté attendue est d'arriver à pouvoir maîtriser l'empan de lecture en coordination avec son souffle afin de faire porter la dernière proéminence sur la première syllabe de *listen*. L'erreur prévisible d'un apprenant angliciste francophone est de placer le noyau sur *them*, « mot outil », alors qu'il est en fin de phrase. Ce marqueur semble représenter un candidat idéal pour évaluer l'aptitude à désaccentuer un pronom en fin d'énoncé.

5.4 Étude temporelle du rythme par durée syllabique

En principe, n’importe quelle paire syllabique dont les phones sont bien délimités acoustiquement devrait convenir pour mesurer l’isochronie syllabique à condition de comprendre une syllabe accentuée non emphatique et une autre non accentuée avec voyelle brève. Pourtant, le texte comprend peu de candidats correspondant aux normes optimales. Par exemple, dans la paire /@nd s{N/, la dernière consonne lue par les apprenants a tendance à devenir /Ng/, ce qui rallonge la syllabe en ajoutant un phonème, biaisant les données.

Le rapport de durée inter-syllabique est mesuré sur deux paires de syllabes candidates présentées ci-dessous :

<i>Ordre d’apparition dans le texte</i>	<i>Syllabe forte</i>	<i>Syllabe faible</i>	<i>Texte du tronçon</i>
Phrase 4	/lIs/	/t@/	[...] in order to listen to them.
Phrase 19	/s{t/	/@t/	[...] as he sat at the window [...]

Tableau 56 – Sélection des syllabes pour étude temporelle du rythme

Alors que dans *sat at* la syllabe forte précède la syllabe faible, dans *to listen*, la syllabe forte la suit. La syllabe /t@/ est mesurée à partir de la détente. Un graphique comparatif est établi ci-dessous. On y voit que la hiérarchie de durée des syllabes fortes correspond à celle établie entre lecteurs par la durée du texte entier, cependant, celles mesurées sur *sat* semblent plus fidèles.

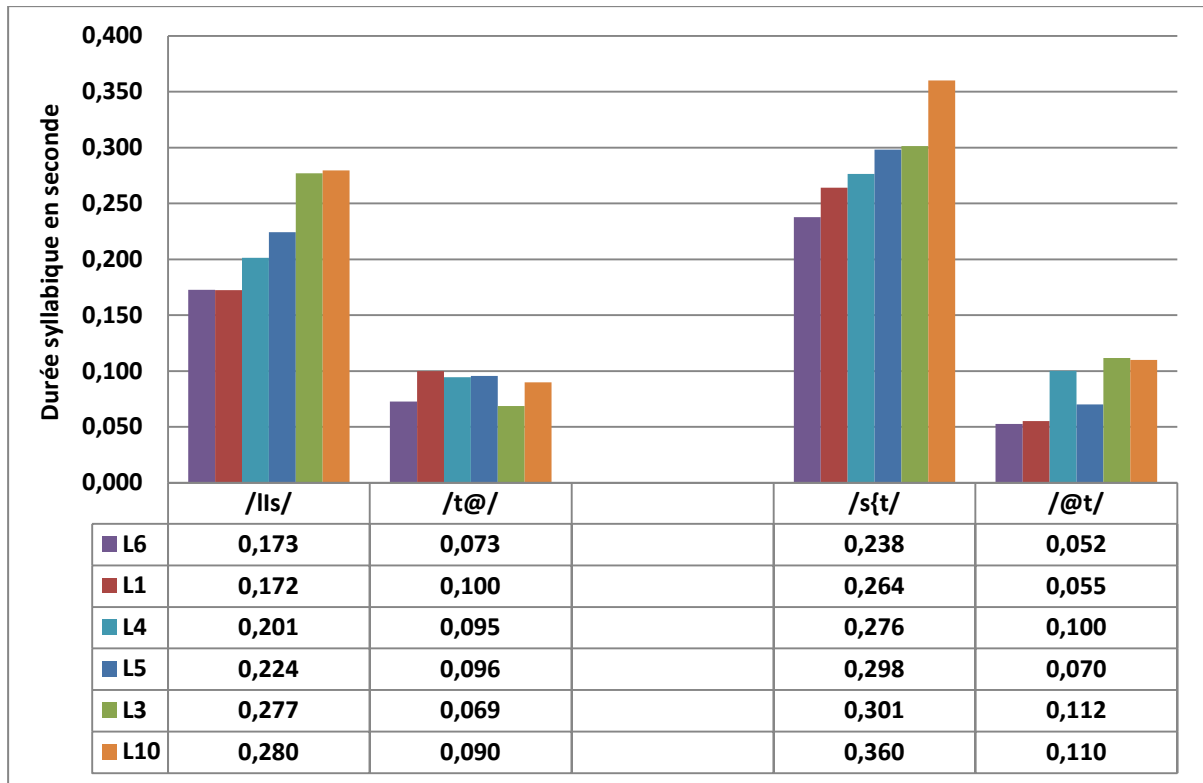


Figure 111 – Comparaison de la durée syllabique dans *to lis(ten)* et *sat at*

La proportion syllabe forte et syllabe faible constitue les variables rythmiques « sil » et « sat », et leur moyenne la variable « satlis », dont le cercle des correspondances avec les variables du Chapitre II figure ci-dessous.

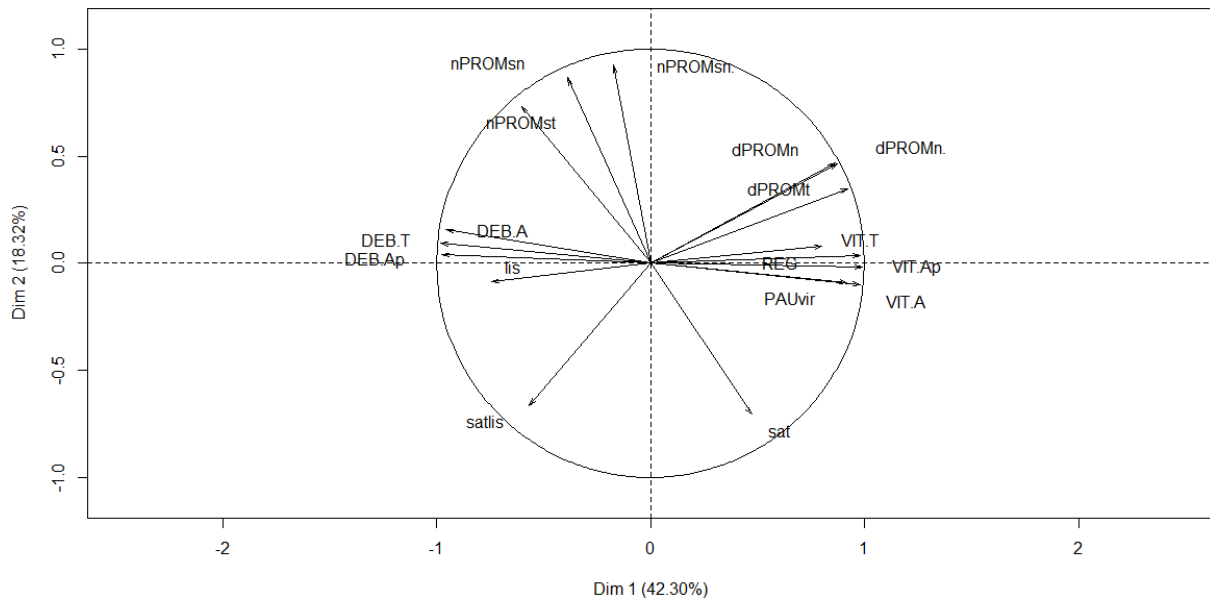


Figure 112 – Cercle des correspondances des variables du Chapitre II avec les variables rythmiques syllabiques « lis », « sat » et « satlis »

CHAPITRE III – Emphase : mode d’organisation paradigmatique des réalisations prosodiques des anglophones natifs

La longueur des flèches « sat » et « satlis » indique leur pertinence par rapport aux variables déjà représentées dans le cercle. Bien qu’elle se rapproche des variables de débit (DEB), la variable « lis » représente un intérêt moindre dans le profilage. Une vérification par dendrogramme montre que seule la variable « sat » conserve les liens établis précédemment. Les autres variables prises ensemble ou isolément modifient la hiérarchie précédemment observée au Chapitre II. Aussi, c’est cette variable qui sera conservée dans l’analyse des non natifs.

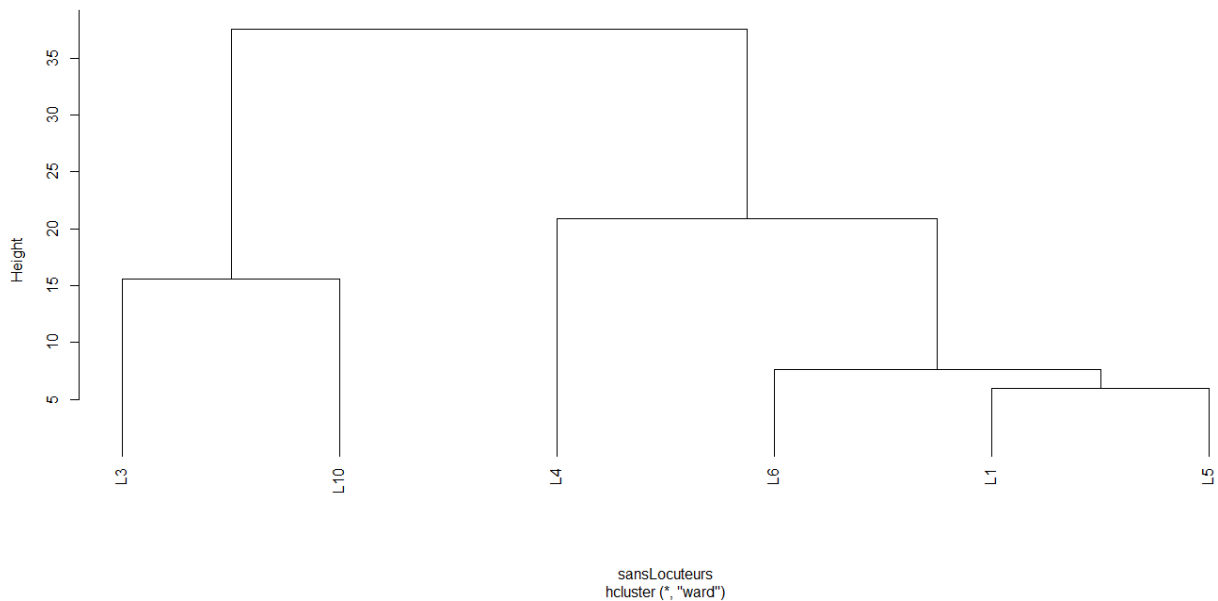


Figure 113 – Dendrogramme des variables du Chapitre II avec la variable rythmique syllabique « sat »

Les trois méthodes présentées jusqu’à présent sont destinées à rendre compte du rythme phrastique de manières diverses. La première a permis de recueillir par perception auditive des données déterminant les profils prosodiques de natifs dans le Chapitre II, mais par sa difficulté de mise en place, ne sera pas utilisée pour profiler la prosodie des non natifs. Elle est remplacée par deux méthodes acoustiques sur marqueurs spécifiques prédéterminés : analyse de la saillance des adverbes intensificateurs *so* et *all* d’une part, et analyse temporelle de syllabe accentuée-inaccentuée d’autre part, chacune devant souligner un aspect du rythme étudié. L’étude lexicale du locus accentuel doit compléter l’analyse.

5.5 Placement et degré accentuels lexicaux

5.5.1 Sélection des polysyllabes pour étude

Afin d’obtenir des informations sur la capacité des apprenants à réaliser un placement accentuel lexical correspondant aux réalisations natives, l’étude se focalise sur une sélection

des polysyllabes d’au moins trois syllabes figurant dans le texte. Ainsi ont été exclus de l’étude les adjectifs *beautiful* et *delicate* de la Phrase 3, et *wonderful* de la Phrase 24, qui représentent le schéma accentuel canonique sur la syllabe antépénultième des mots d’au moins trois syllabes sans suffixe contraignant. Le verbe *understand* (Phrases 11 et 19) a lui aussi été écarté pour son peu d’intérêt : le locus accentuel sur la racine du mot correspond à l’erreur typique que ferait un apprenant faible en accentuant la dernière syllabe. Le pronom indéfini *nobody*, quant à lui, est étudié dans la rubrique des marqueurs d’emphase et représente lui aussi le schéma canonique /100/.

5.5.2 Considérations acquisitionnelles et didactiques

Les six mots d’au moins trois syllabes qui ont été retenus sont *conversation*, *limited*, *determined* (Phrase 8), *trespassers*, *prosecuted* (Phrase 13), et *musicians* (Phrase 22). Bogaards en présentant les processus qui interviennent dans l’apprentissage d’une L2, affirme (1994 : 94) : « L’apprentissage semble donc revenir, en fin de compte, à la mise en place de traces mémorielles et au renforcement des liens qui unissent celles-ci. ». Il fait mention (1994 : 71) des différents types d’informations linguistiques (d’ordre sémantique, morphologique, phonologique, etc.) qui constituent les éléments lexicaux et sur lesquels se basent les relations intrinsèques (ou catégorielles) organisant le lexique mental. Pour être effectif et cohérent, un apprentissage lexical doit donc être opéré en réseaux, sans exclure la composante phonologique. Celce-Murcia *et al.* (1996 : 143) rappelle la priorité pédagogique qu’a un enseignant d’anglais d’inculquer à ses élèves la nécessité de connaître le placement accentuel de chaque mot appris.

5.5.3 Fréquence et domaine d’utilisation des polysyllabes sélectionnés

Le tableau suivant récapitule les données obtenues pour chacun de ces mots selon son rang de fréquence d’après le site *Word and Phrase . Info*⁸¹.

RANK #	PoS	WORD	TOTAL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
773	V	DETERMINE	55918	6530	3608	10413	8308	27059
1165	N	CONVERSATION	36979	10157	9847	6165	5796	5014
1992	J	LIMITED	20108	1935	903	4026	3949	9295
2953	N	MUSICIAN	11609	2537	1345	2121	3046	2560
5141	V	PROSECUTE	5198	2227	260	612	1349	750
26954	N	TRESPASSER	220	10	103	42	41	24

Tableau 57 – Récapitulatif des informations trouvées sur le site *Word and Phrase . Info* pour les six polysyllabes étudiés

⁸¹ [<http://www.wordandphrase.info/frequencyList.asp>] (dernièrement consulté en 2017)

CHAPITRE III – Emphase : mode d’organisation paradigmatique des réalisations prosodiques des anglophones natifs

Ce tableau montre la catégorie grammaticale répertoriée pour le mot analysé, le nombre d’occurrences dans la totalité du corpus ainsi que dans chaque domaine catalogué. On peut s’attendre à ce que le mot *conversation* ait souvent été rencontré à l’oral étant donné sa forte occurrence dans la langue et surtout dans les domaines d’anglais parlé et de fiction. *Musician* et *prosecute* se trouvent dans les corpus oraux mais sont beaucoup moins fréquents dans la langue, ce qui devrait générer davantage d’erreurs. De plus, seul le verbe *prosecute* figure dans la base de donnée des premiers 60 000 mots ; le participe passé correspondant n’apparaît pas. *Determine* se trouve en première position, mais l’on peut observer qu’il appartient au lexique universitaire (de forme souvent écrite et peu connu des étudiants de première année). Il est en outre moins fréquent que le mot *conversation* dans les occurrences orales (anglais parlé et anglais de fiction). Il serait logique de s’attendre à des erreurs sur l’accentuation de ce lemme ainsi que sur *limited*, qui est similaire dans ses occurrences et de fréquence moindre dans la langue. La dernière unité lexicale, *trespasser*, est d’une fréquence d’occurrence très faible par rapport aux autres mots, ce qui devrait générer des difficultés amplifiées par sa morphologie.

5.5.4 Placement et degré accentuels des polysyllabes sélectionnés

Chacun de ces mots représente une source d’erreur possible pour les apprenants francophones. Le tableau suivant, basé sur les informations contenues dans les dictionnaires *Longman Pronunciation Dictionary* (2008) et *Cambridge English Pronouncing Dictionary* (2011), montre l’absence de variantes accentuelles pour les mots sélectionnés, malgré quelques variantes segmentales.

Phrase	Mot	Schéma accentuel LPD	Schéma accentuel CEPD
Phrase 8	CONVERSATION	/2010/	/2010/
Phrase 8	LIMITED	/100/	/100/
Phrase 8	DETERMINED	/010/	/010/
Phrase 13	TRESPASSERS	/100/	/100/
Phrase 13	PROSECUTED	/1000/	/1000/
Phrase 22	MUSICIANS	/010/	/010/

Tableau 58 – Schémas accentuels canoniques des polysyllabes étudiés

Une expérimentation perceptive (PER-ACLEX-N) a permis de déterminer si ces schémas accentuels étaient effectivement produits par les locuteurs L1 et L5 au fil de la lecture. La clarté et la précision de leurs réalisations accentuelles étaient suffisamment nettes pour qu’un spécialiste phonéticien non natif n’ait aucun doute quant à la concordance.

Parmi ces six polysyllabes, l'apprenant est confronté à la gestion des phénomènes de flexion (par ajout de la marque du prétérit, du participe passé ou par éventuelle mise au pluriel) et de dérivation (par ajout d'un affixe lexical). Il doit donc être capable d'appliquer les règles de séparabilité sémantique pour identifier les unités lexicales, puis leur appliquer les règles de position pour placer les accents (Fournier, 2010 : 17).

5.5.5 Hypothèses sur les réalisations non natives

L'apprenant est confronté au repérage et à la gestion des préfixes dans les mots *determined*, *prosecuted* et *trespassers*. Le préfixe d'origine latine *de-* est une base inactive facilement identifiable en anglais contemporain et l'accent tonique est alors porté par la syllabe *-ter-*. Le dictionnaire *Merriam-Webster*⁸² retrace l'origine du verbe *prosecute* à partir du XV^{ème} siècle et signale qu'il provient du mot latin *prosecutus*. La première syllabe *pros-* ne peut être perçue comme un préfixe : dans le mot *prosecute*, l'apprenant devra s'efforcer de distinguer la première syllabe *pros-* du préfixe *pro-* et devra analyser le mot comme une unité sémantiquement inséparable, lui attribuant le schéma accentuel canonique /100/. Il en est de même pour le mot *trespassers* dont l'origine, selon le dictionnaire *Merriam-Webster*⁸³, remonte au XIV^{ème} siècle. Ce mot du Moyen-anglais comprenait le préfixe *tres-*, provenant du préfixe latin *trans-*. Le préfixe se révèle ici aussi être une base inactive dont le schéma accentuel ne peut être que /100/. Il faut mentionner que la Phrase 13 : *TRESPASSERS WILL BE PROSECUTED* est mise en relief typographiquement par l'utilisation de majuscules et par un centrage sur la page de lecture. Certains lecteurs ont ainsi maquillé leur voix pour imiter celle du Géant, voyant l'écriteau comme le prolongement direct de ses paroles menaçantes. Ce changement de voix est susceptible de modifier la prosodie de la phrase et le schéma accentuel des polysyllabes qui la composent.

Les suffixes faibles se trouvent dans les trois verbes *limited*, *prosecuted* et *determined*, ainsi que dans le nom *trespassers*. Dans les verbes, la graphie *-ed* se réalise /d/ après la dernière consonne voisée, sans ajout de syllabe dans le dernier cas, et /id/ après le /t/ non voisé avec ajout de syllabe. Le /z/ correspondant à la graphie du pluriel dans *trespassers* ne nécessite pas l'adjonction de syllabe. Dans ces quatre cas, aucun changement accentuel n'a lieu et le mot est accentué comme s'il ne portait pas de suffixe. Ces suffixés relèvent de la « Loi de la dérivation neutre » (Fournier, 2010 : 20).

⁸² [<http://www.merriam-webster.com/dictionary/prosecute>] consulté en 2015.

⁸³ [<http://www.merriam-webster.com/dictionary/trespasser>] consulté en 2015.

Les deux mots restants confrontent les apprenants à leur gestion des suffixes forts. Dans les deux cas, le suffixe séparable est de forme $-i+V(C_0(e))$. L’accent primaire sera porté par la syllabe précédant immédiatement le suffixe. *Musicians* sera accentué /010/. *Conversation* est plus complexe dans la mesure où « Aucune unité lexicale ne peut commencer par deux syllabes inaccentuées. » et qu’« Il ne peut pas y avoir deux accents successifs au sein d’une même unité lexicale. » (Fournier, 2010 : 12). En conséquence, un accent secondaire devra être ajouté sur la syllabe *con-*, donnant au mot le schéma accentuel /2010/. Si certains apprenants faibles persistent à accentuer ces mots sur la dernière syllabe par transfert de leur langue maternelle, la plupart d’entre eux maîtrise les règles d’accentuation due aux suffixes forts. Néanmoins, ces deux mots ont chacun leurs difficultés propres : les réalisations de *conversation* devront marquer un accent secondaire sur la syllabe *con-*, et le mot *musicians*, quant à lui, devra conserver son accentuation malgré l’organisation du rythme de la phrase dans laquelle il est inséré. En effet, la fin de la Phrase 22 assigne une proéminence d’ordre sémantique sur *King’s* et une autre par nécessité grammaticale sur le deuxième élément du verbe à particule *passing by*.

Phrase 22 : [...] that he thought it must be the King's musicians passing by.

Réalisations de L1 et L5 ⁸⁴ :	2	0	1	0	2	0	2
Assignment attendue d’un apprenant :	1	0	0	1	0	0	1

Figure 114 – Assignment attendues des proéminences des natifs et des non natifs en fin de Phrase 22

Cinq syllabes séparent ces deux proéminences. La tendance naturelle des apprenants est d’accentuer la dernière syllabe du mot *musicians*, renforçant une régularité rythmique qui séparerait les trois proéminences par respectivement deux syllabes. Il s’agirait là pour les apprenants de gérer un problème de rythme, plus qu’un conflit entre *stress* et *accent*.

5.5.6 Bilan de l’étude de placement et degré accentuels lexicaux

Cette rubrique s’est efforcée de présenter et d’expliquer les diverses difficultés que les apprenants devraient surmonter dans leur lecture des polysyllabes. Outre le placement accentuel et son degré, un problème majeur rencontré par les évaluateurs risque cependant d’être le manque de clarté dans les réalisations accentuelles où toutes les syllabes seraient dotées de caractéristiques similaires rendant le repérage difficile.

⁸⁴ Par perception auditive de spécialiste non natif.

5.6 Synthèse de l'étude des marqueurs à l'emphase réduite (expérimentations COR-DIV-N3)

Cette seconde étude dans le cadre du phénomène de l'emphase, cette fois sur les marqueurs de moyen ou faible degré, a porté sur le contour mélodique descendant de la phrase « neutre », le ton bidirectionnel *fall-rise* avant frontière mineure, le décrochage mélodique sur *all*, la désaccentuation pronominale en fin d'énoncé, l'étude temporelle du rythme par durée syllabique et le placement accentuel lexical. On a pu constater que l'absence complète d'emphase n'existait qu'au niveau du mot, mais pas à celui de la phrase, et que si dans certains cas l'absence d'emphase était recherchée pour minimiser la variation prosodique, dans d'autres, elle était nécessaire pour favoriser la variabilité. En effet, c'est grâce à la variabilité que l'on obtient une variation des corrélats graduable, catégorisable et hiérarchisable. Toute hiérarchisation nécessitant au moins un repère stable, l'absence d'emphase devient la référence fondamentale à partir de laquelle les degrés de variation peuvent être calibrés puis hiérarchisés.

Le profilage des corrélats de ces candidats intonatifs à l'évaluation de l'interlangue prosodique a été opéré selon des méthodes diverses : acoustiques, perceptives, répertoriées dans la littérature ou programmées par application de la hiérarchisation du degré d'emphase. Si certains sont retenus pour l'évaluation mélodique, d'autres seront plus adaptés pour décrire le rythme. Ainsi, en fonction des diverses approches pouvant rendre compte du rythme, j'ai établi trois lignes directrices d'investigation présentées ci-dessus ou en amont. Elles sont destinées à se compléter mutuellement car chacune d'entre elles souligne un aspect ignoré par les deux autres. Les divers aspects rythmiques étudiés devraient permettre de rendre compte succinctement de la complexité du phénomène au niveaux phrastique et lexical.

Dans cette sélection de marqueurs rythmiques spécifiques, le tableau suivant récapitule les occurrences de proéminences étudiées sur lieux prédéterminés au niveau supra-syllabique. Les intensificateurs *so* et *all* interviennent au niveau de la phrase. Ils sont très courants dans la langue parlée et, si *so* relève davantage du discours émotif, *all*, moins chargé d'emphase, est très présent dans le discours de registre soutenu.

CHAPITRE III – Emphase : mode d'organisation paradigmatique des réalisations prosodiques des anglophones natifs

<i>Ordre d'apparition dans le texte</i>	<i>Nombre de mots</i>	<i>Nombre de syllabes</i>	<i>Degré d'emphase de la proéminence</i>	<i>Texte du tronçon</i>
Phrase 4	8	10	So : 1,70 Sweetly : 0,77	[...] trees and sang so sweetly that the children [...]
Phrase 22	7	8	So : 2,17 Sweet : 0,39	It sounded so sweet to his ears [...]
Phrase 12	9	9	All : 1,39 Round: 0,51	So he built a high wall all round it, [...]
Phrase 16	5	7	All: 0,74 Over : 0,08	[...] and all over the country [...]
Phrase 4	3	4	Them : 0,07	[...] listen to them .

Tableau 59 – Tronçons de phrases sélectionnés pour étude acoustique du rythme par proéminences ou absence de proéminence

Toujours dans cette sélection de marqueurs rythmiques spécifiques, le tableau suivant récapitule les occurrences de proéminences étudiées sur lieux prédéterminés au niveau du mot.

<i>Ordre d'apparition dans le texte</i>	<i>Mots étudiés</i>	<i>Nombre de syllabes</i>	<i>Degré d'emphase du mot</i>	<i>Nature de l'étude rythmique</i>
Phrase 8	Conversation	4	0,28	Déplacement accentuel dans polysyllabes (étude perceptive)
	Limited	3	0,51	
	Determined	3	0,32	
Phrase 13	Trespassers	3	1,40	
	Prosecuted	4	1,67	
Phrase 22	Musicians	3	0,23	
Phrase 19	Sat	1 + 1	0,12	Étude temporelle du rythme (étude acoustique)
	At		0	

Tableau 60 - Mots sélectionnés pour étude rythmique

On peut y voir que les mots *TRESPASSERS* et *PROSECUTED* sont dotés d'un degré d'emphase très élevé, ce qui présage une grande variabilité. Cependant, la méthodologie d'investigation par perception devrait parvenir à facilement localiser et catégoriser la proéminence globale. À l'autre bout de l'échelle, le degré faible ou nul des mots sélectionnés pour l'étude de la durée syllabique devrait très fortement réduire la variabilité.

Enfin, l'étude mélodique se penchera sur la réalisation du contour descendant par défaut de la langue anglaise au niveau d'une phrase déclarative dépourvue d'emphase, et de celle d'un *fall-rise* sur une seule syllabe (voir tableau suivant). Après l'étude effectuée sur

l'emphase, on ne pourra trouver surprenant de constater que la réalisation de ce ton complexe bidirectionnel nécessite un certain degré d'emphase.

<i>Lieu d'occurrence du contour</i>	<i>Forme du contour</i>	<i>Nombre de syllabes</i>	<i>Texte du tronçon</i>	<i>Degré d'emphase moyen</i>
Contour syllabique	∨	1	[...] (a)rrived [...]	Arrived : 0,96
Contour de syntagme intonatif	\	11	[...] he saw the children playing in the garden.	Tronçon : 0,15

Tableau 61 – Contours mélodiques vérifiés sur la Phrase 9 (déclarative non emphatique)

L'expérimentation COR-DIV-N3 a montré que le côté « en creux » de l'emphase pouvait lui aussi servir à profiler les réalisations prosodiques en les catégorisant pour les sélectionner ou les écarter de l'étude. En cela, elle valide à son tour l'Hypothèse 3. Les marqueurs déterminés par l'absence d'emphase viennent compléter ceux sélectionnés à la suite de l'expérimentation COR-AC-N2. Ensemble, leur étude va pouvoir fournir une vue globale de la mélodie, et du rythme lorsque la mesure de la F0 lui sert de corrélat.

6 Conclusion : analyse prosodique des natifs anglophones sous l'angle de l'emphase

Dans ce chapitre sur l'emphase, une première phase (expérimentation PER-EMPH-N) établit un repérage perceptif des zones d'emphase dans la lecture de trois natifs sélectionnés selon leur qualité de lecture et leur vitesse d'élocution, et une deuxième (expérimentation COR-AC-N2) procède à l'analyse acoustique des marqueurs gradués comme les plus emphatiques, une troisième série d'expérimentations COR-DIV-N3 fait de même sur des marqueurs dépourvus d'emphase ou d'un degré d'emphase réduit. L'expérimentation PER-EMPH-N hiérarchise les mots du texte selon leur degré d'emphase et l'analyse acoustique des syntagmes intonatifs dans lesquels ils sont insérés COR-AC-N2 tend à révéler une variation pouvant être identifiée et catégorisée. Les analyses acoustiques, perceptives ou statistiques COR-DIV-N3 visent une emphase la plus réduite possible afin de minimiser la variabilité des résultats.

L'objectif poursuivi dans ce Chapitre III était double : une validation de l'Hypothèse 3, selon laquelle le phénomène d'emphase exacerbait la variabilité prosodique, opérée par une

sélection pertinente de marqueurs déterminés par l’emphase chez les locuteurs natifs. Cette analyse débouche sur une réflexion sur la nature et le potentiel du phénomène d’emphase.

6.1 Validation de l’Hypothèse 2

Les expérimentations présentées précédemment ont validé l’Hypothèse 2 de cette thèse, que je rappelle ici :

H2 — Le phénomène d’emphase favorise la variabilité des réalisations prosodiques. Aussi, son repérage permet de limiter l’étude à des entités concentrant la variation maximale. Le phénomène d’emphase devrait alors permettre de mieux comprendre, organiser et hiérarchiser la variabilité des phénomènes prosodiques intra-locuteurs en graduant la variation pour l’analyser plus précisément ou au contraire pour l’exclure des analyses.

Poussant plus loin le raisonnement en fonction des résultats obtenus dans les expérimentations COR-DIV-N3, on peut ajouter son corollaire, ce qui modifierait l’hypothèse ainsi :

H2’ — Le phénomène d’emphase favorise la variabilité des réalisations prosodiques. Aussi, son repérage permet de limiter l’étude à des entités concentrant la variation maximale **ou au contraire en la réduisant pour cibler des entités prosodiquement stables**. Le phénomène d’emphase devrait alors permettre de mieux comprendre, organiser et hiérarchiser le degré de variabilité des phénomènes prosodiques intra-locuteurs en graduant la variation pour l’analyser plus précisément ou au contraire pour l’exclure des analyses.

6.2 Sélection des marqueurs déterminés par l’emphase chez les natifs

L’application de l’Hypothèse 2 à l’étude des réalisations prosodiques a permis une sélection réduite des entités analysables afin de compléter les variables destinées à l’analyse de la prosodie des apprenants. Ces marqueurs doivent être ajoutés à ceux résultant de l’expérimentation COR-AC-N1 du Chapitre II sur les profils : ils la complètent. Ces marqueurs déterminés par le phénomène d’emphase sont consignés dans la grille suivante.

	<i>Domaine analysé</i>	<i>Sous-domaines</i>	<i>Variables obtenues et unités de mesure</i>
REGISTRE	Registre : Ampleur du registre	<ul style="list-style-type: none"> Sur les marqueurs emphatiques : exclamative, paires <i>so</i> + adjectif/adverbe (Phrases 4 et 22) Dans phrase déclarative au degré d'emphase faible 	<ul style="list-style-type: none"> Quantitatives (DT)
MÉLODIE	Mélodie dans la phrase : Contour mélodique dans syntagme intonatif ou énoncé plurisyllabique	<ul style="list-style-type: none"> <i>Fall</i> (exclamative, déclarative) <i>Rise-fall</i> (<i>nobody</i>) 	<ul style="list-style-type: none"> Quantitatives (DT) Qualitatives (réalisation des contours conforme, plutôt conforme, non conforme)
	Mélodie dans la syllabe : Ton syllabique sur nucléus	<ul style="list-style-type: none"> <i>Fall-rise</i> (<i>aRRIVED</i>) <i>Rise-fall</i> et tons tridirectionnels (<i>NObody</i>) 	<ul style="list-style-type: none"> Qualitatives (statique, unidirectionnel, bidirectionnel, tridirectionnel)
RYTHME	Rythme : maîtrise du placement accentuel dans les polysyllabes	<ul style="list-style-type: none"> Dans 6 polysyllabes, conformité des schémas accentuels à ceux des natifs par tests perceptifs 	<ul style="list-style-type: none"> Qualitatives (conforme, non conforme)
	Rythme : proéminence de marqueurs spécifiques	<ul style="list-style-type: none"> Paires <i>so</i> + adjectif/adverbe (Phrases 4 et 22) Paires <i>all</i> + préposition (Phrases 12 et 16) Proéminence sur <i>them</i> 	<ul style="list-style-type: none"> Quantitatives (DT) Qualitatives (proéminent, non proéminent)
	Rythme sur syllabes spécifiques (approche temporelle)	Rapport durée inter-syllabique sur la paire de syllabes (<i>sat at</i>)	<ul style="list-style-type: none"> Quantitatives (ms, %)

Tableau 62 – Sélection des variables déterminées par les expérimentations COR-AC-N2 et COR-DIV-N3 (dont PER-ACLEX-N)

6.3 Conclusion sur la nature et le potentiel du phénomène de l'emphase

Herment et Hirst ont fait le lien entre le phénomène d'emphase tel qu'il était conceptualisé par les linguistes phonéticiens et l'extraction des corrélats acoustiques destinés à les repérer (2001, 2003, 2011). Ils en ont donné une définition. J'ai tenté de prolonger leur recherche en utilisant et approfondissant la méthodologie de repérage de l'emphase d'Herment, cette fois sur texte lu, ce qui m'a permis d'établir des comparaisons entre les réalisations en calibrant la variabilité. Au cours de l'étude de ce chapitre, j'ai tenté de dépasser l'étape descriptive de l'emphase pour m'orienter vers celle d'une utilisation comme outil catalyseur de sélection. L'emphase a opéré comme un miroir grossissant, une loupe destinée à agir comme filtre au niveau syntagmatique pour révéler la variation paradigmatique comme un jeu d'échos et de miroirs. Si dans certains cas l'absence d'emphase était recherchée

pour minimiser la variation prosodique, dans d'autres, elle était nécessaire pour la favoriser. En effet, c'est grâce à la variabilité que l'on a obtenu une variation des corrélats graduable, catégorisable et hiérarchisable. Toute hiérarchisation nécessitant au moins un repère stable, l'absence d'emphase devient la référence fondamentale à partir de laquelle les degrés de variation peuvent être calibrés puis hiérarchisés. L'opération de profilage a consisté à établir les modalités de filtrage et leur granularité. Alors que la sélection des syntagmes intonatifs étudiés se fait traditionnellement par le biais de la syntaxe, le phénomène d'emphase a permis de choisir un angle d'attaque différent en hiérarchisant et contenant la variabilité inhérente à la prosodie. Cette étape a été déterminante pour atteindre l'objectif de trouver des traits discriminants sur le plan évaluatif.

J'ai émis l'hypothèse qu'une étude typologique du phénomène d'emphase était de nature à contribuer significativement à rendre compte de la variation entre les données perceptives ou acoustiques hétéroclites et continues. Divers marqueurs d'emphase ont été analysés selon une approche s'appuyant sur l'étude de corpus. Chaque marqueur a débouché sur des patrons spécifiques basés sur les unités de syntagme intonatif ou de syllabe, et a ouvert la voie à une typologie prosodique, ce qui appelle naturellement à s'interroger sur le caractère catégoriel ou continu de l'emphase.

L'emphase n'est pas un phénomène prioritairement acoustique, mais perceptif, subjectif et rhétorique. On peut simplement établir des correspondances à partir des occurrences de parole qualifiées d'« emphatiques ». Si les corrélats acoustiques forment un continuum débutant souvent à partir du seuil 0, la nature catégorielle de l'emphase ne peut être avérée qu'en établissant des seuils induisant une modification typologique des corrélats, qu'ils soient perceptifs ou acoustiques. Sans entrer dans le débat de savoir si l'emphase est un phénomène continu ou catégoriel, lequel n'a pas sa place ici, je peux simplement attirer l'attention sur certains faits. Cette étude oppose nettement les résultats phonétiques et phonologiques d'un haut degré et d'un faible degré d'emphase. Les corrélats acoustiques montrent que les choix opérés lors de l'expression de l'emphase ont tendance à s'exclure et se compenser les uns des autres. Ils permettent alors de catégoriser les réalisations selon une typologie spécifique à l'emphase portant sur une distribution complémentaire des divers corrélats selon les locuteurs afin d'obtenir une réalisation conforme à l'objectif illocutoire (Austin, 1962). Ainsi, sous l'influence attitudinale, l'interrogative de reproche en *What* peut subir un tassement de F0 corrélé à une homogénéisation de l'intensité des syllabes, phénomène s'apparentant aux langues syllabiques. Les exclamatives en *How*, quant à elles,

sont un indicateur du profil d'âge que se donne un locuteur (un « profil jeune » se démarque par son contour intonatif sur exclamative), contribuant à créer un sous-profil de lecteur. Les paires « *so* + adjectif ou adverbe » confirment le degré d'implication interlocutive des lecteurs. Les marqueurs de degré emphatique faible ou nul minimisent la variabilité phonétique mais aussi phonologique. Le noyau d'un mot a tendance à adopter un patron mélodique différent selon que ce mot est doté d'un degré d'emphase élevé (montant-descendant), moyen (descendant) ou faible, voire inexistant (plat). Les résultats obtenus conduisent cependant à constater que l'absence complète de tout degré d'emphase n'est possible que sur de petites entités : le mot et la syllabe, mais pas le syntagme intonatif ou la phrase. Cela prouve sa nature éminemment prosodique, du moins lorsque l'on tient compte d'un faible degré. Elle n'est donc pas un niveau supplémentaire, déconnecté et superposable sur la structure, comme l'est par exemple la négation.

La manière dont je conçois la nature de l'emphase en termes de phénomène catégorisable ou non est schématisée dans la figure ci-dessous.

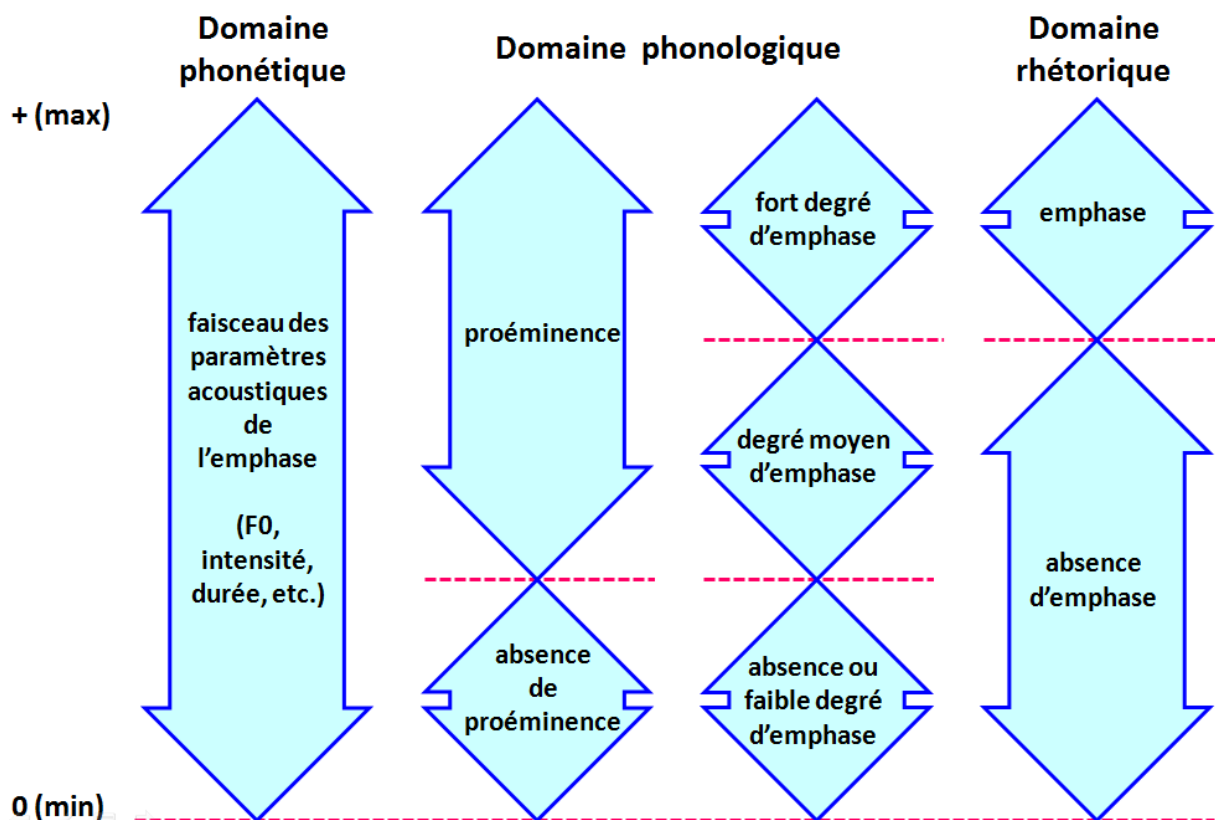


Tableau 63 – Ma position sur la nature catégorielle ou continue du phénomène de l'emphase

La conception sur la catégorisation de l'emphase semble dépendre en fait du domaine étudié. La communauté s'accorde sur le fait que l'emphase soit un phénomène continu dans l'analyse

CHAPITRE III – Emphase : mode d'organisation paradigmatique des réalisations prosodiques des anglophones natifs

de ses paramètres acoustiques, c'est-à-dire en phonétique. Ce n'est pourtant le cas ni en phonologie, ni en rhétorique, bien que les limites départageant les catégories soient laissées à l'appréciation de l'auditeur ou du lecteur. Les travaux d'Avanzi *et al.* (2008) montrent d'ailleurs que l'on peut catégoriser la proéminence : ils l'annotent en la subdivisant en deux degrés (P ou p). Herment (2001) et moi organisons l'emphase en davantage de niveaux encore pour refléter la nature continue et graduelle des corrélats phonétiques.

L'étude de ce chapitre illustre le fait que la classification des données continues de l'emphase en catégories est non seulement possible, mais encore qu'elle constitue un pas obligatoire pour créer et organiser une classification typologique éclairant les réalisations lorsque les catégories traditionnelles demeurent impuissantes à le faire de manière satisfaisante, c'est-à-dire avec élégance.

Cette étude a enfin le mérite d'introduire en quelque sorte le côté subliminal de la prosodie :

- au niveau du syntagme intonatif (variable de l'âge pour les réalisations des locuteurs les plus jeunes de l'exclamative en *How* (devenant une contribution à un « parler jeune »), degré interlocutif dans les paires « intensificateur + adjectif ou adverbe »),
- au niveau de la syllabe (par les tons syllabiques utilisés statiques/dynamiques, simples/complexes).

Destinée à compléter l'analyse prosodique pour servir de base à une modélisation de l'interlangue des anglicistes francophones, le rôle de l'emphase est de contribuer à créer des traits distinctifs dits *criterial features* (Hawkins & Buttery, 2010) prosodiques à visée évaluative. Destinée au départ à éclairer la composante mélodique, elle a contribué significativement à déterminer une variable de registre, et plusieurs variables rythmiques. En cela, l'emphase montre sa puissance en devenant un outil de profilage incontournable. Modéliser et prédire les réalisations des lecteurs natifs a constitué une première étape du travail à visée évaluative. À présent, pourvue de la méthode de profilage et des outils que constituent les profils natifs, la matrice de corrélats des profils natifs et la grille de marqueurs à analyser, il est envisageable de procéder à l'étude des apprenants.

CHAPITRE IV – Expérimentation de création de profils et de critères d'évaluation prosodiques d'apprenants

1 Résumé du Chapitre IV

Ce quatrième chapitre a pour fonction d'analyser la prosodie des anglicistes francophones selon les modes d'organisation profilés dans les chapitres précédents. Son objectif fondamental est de déterminer des critères d'évaluation afin de contribuer à construire un modèle prosodique évaluatif pour anglicistes francophones .

Pour cela et dans un premier temps, ce chapitre met en pratique le critère de rapidité d'élocution présenté au Chapitre I afin d'ordonner les fichiers du corpus d'apprenants et en sélectionner quinze représentatifs (expérimentation STAT-NN1). Ensuite, il rappelle les objectifs, enjeux et résultats des profilages obtenus grâce aux analyses des Chapitres II et III et recadre la méthodologie du profilage pour l'adapter aux quinze non natifs sélectionnés (expérimentations COR-AC-NN, PER-LEX-NN, PER-THEM-NN) pour extraire de nouvelles données. Afin de comprendre le fonctionnement de la diversité prosodique des apprenants, un profilage de groupes STAT-NN2 est alors opéré d'abord par archétypes de conformité selon le critère de rapidité de lecture dans les quatre grands domaines (temporel, de registre, mélodique et rythmique), puis par archétypes stéréotypiques STAT-NN3 par analyse des 28 variables prosodiques par plan factoriel, dendrogramme, taux de réussite et d'erreur, et cercle de corrélation. Il devient alors possible d'envisager un profilage de critères d'évaluation prosodique (*critical features*) CRIT-LEC-NN dans les quatre grands domaines prosodiques et de proposer des prototypes de grilles critériées destinées à évaluer la prosodie d'apprenants anglicistes francophones.

Ce chapitre aboutit à des avancées dans la création d'outils ou l'obtention de leurs résultats pour comprendre et évaluer l'interlangue. Les analyses résultent en :

- Une catégorisation hiérarchisée des apprenants d'après leur vitesse de lecture, laquelle permet de sélectionner quinze fichiers sons représentatifs de l'ensemble du corpus d'après ce critère,
- Une vérification de la pertinence du critère de rapidité de lecture,
- Une détermination des six profils extrêmes d'apprenants parmi les quinze étudiés,
- Un classement quantitatif et qualitatif des réalisations des six apprenants extrêmes,

puis des quinze,

- Une méthodologie pour évaluer toutes les réalisations, y compris les réalisations mélodiques, quantitativement et qualitativement,
- Une sélection de dix critères distinctifs (*critical features*) parmi les 28 variables des quatre domaines prosodiques principaux,
- Une méthodologie pour créer des grilles évaluatives de traits distinctifs,
- Des exemples de grilles évaluatives portant sur les critères distinctifs créés.

2 Objectifs, enjeux et profilage de l'analyse prosodique des apprenants

L'introduction de cette thèse a présenté les objectifs assignés à cette recherche. Le Chapitre I a présenté le critère de rapidité d'élocution comme principe de classement pouvant hiérarchiser des productions non natives. Le Chapitre II a étudié les notions de profils, puis le Chapitre III la notion d'emphase, afin d'en extraire les outils nécessaires en vue d'aborder l'étude de l'interlangue.

Ce chapitre finalise cette étude en mettant en place des expérimentations visant tout d'abord à établir un protocole de sélection d'apprenants basé sur la vitesse de lecture et de sélectionner des fichiers représentatifs du corpus afin de poursuivre l'étude. L'Hypothèse à vérifier préalablement est la suivante :

H3 — Le paramètre de rapidité d'élocution étant un indicateur déterminant de niveau des apprenants, la catégorisation de ceux-ci selon ce paramètre devrait permettre d'organiser la variation inter-locuteurs des résultats des autres variables prosodiques afin de trouver des critères d'évaluation fiables.

Ensuite sont créés des profils prosodiques d'apprenants, lesquels permettent de comprendre le fonctionnement de la prosodie de l'interlangue par les relations que ses diverses composantes entretiennent entre elles (organisation interne), voire avec d'autres caractéristiques de nature non prosodique (organisation externe). Cette hiérarchie nouvelle entre les variables prosodiques apporte des informations fondamentales pour pondérer les données destinées à

créer les critères d'évaluation (*crieterial features*)⁸⁵ et en dresser une grille prosodique. Cela est destiné à valider la Question de Recherche 1, rappelée ci-dessous :

QR1 — Étant donné la forte variabilité des réalisations prosodiques natives et d'apprenants, des outils spécifiques doivent être utilisés ou créés afin de l'analyser. La validation des Hypothèses H1, H2 et H3 devrait permettre de créer des profils prosodiques de non natifs permettant d'orienter la création de traits discriminants évaluatifs.

Pour atteindre ces objectifs, cette expérimentation met en œuvre les outils créés dans les trois premiers chapitres et au début de celui-ci, qui sont :

- Les définitions opérationnelles des trois catégories de profils prosodiques : profils personnalisés, profils de groupes par archétypes de conformité et profils de groupes par archétypes stéréotypiques ou de divergence (Chapitre II),
- La matrice des corrélats acoustiques de profils de lecteurs natifs (Chapitre II),
- La détermination de zones d'emphase graduées chez les natifs (Chapitre III),
- L'étude acoustique des marqueurs d'emphase des natifs (Chapitre III).

Pour atteindre son objectif, après un bref rappel des objectifs visés et de la définition d'un profil (section 4.2), ce Chapitre IV présente le protocole de l'expérimentation STAT-NN1 dont l'objectif est d'opérer un premier classement des fichiers d'apprenants pour pouvoir aborder les profilages suivants avec une réduction de corpus (faisabilité) garantissant sa représentativité (section 4.3).

Ce chapitre introduit ensuite les protocoles des expérimentations COR-AC-NN, PER-LEX-NN, PER-THEM-NN destinées à obtenir des données nouvelles spécifiques à la prosodie non native, puis ses résultats. Il détaille ainsi les considérations prévalant à l'élaboration de la grille d'analyse conduisant à la sélection des variables et des mesures prises en compte, puis il expose la méthodologie de profilage des expérimentations suivantes (section 4.4) et procède à la description de leurs résultats acoustiques ou perceptifs, à l'analyse des données obtenues et à leur traitement en trois étapes, qui sont :

⁸⁵ Hawkins J. A. & Filipović L. (2012 : 11) : "The basic idea is that we try to find properties of learner English that are characteristic and indicative of L2 proficiency at each of the levels and that distinguish higher levels from lower levels."

- La description de diverses variables de l'interlangue selon leur catégorisation en groupes temporels créés sur la base de la vitesse de lecture selon des archétypes de conformité (section 4.5, expérimentation STAT-NN2),
- La détermination de profils prosodiques d'apprenants selon un profilage par archétypes stéréotypiques (section 4.6, expérimentation STAT-NN3),
- La détermination de critères prosodiques distinctifs (*critical features*) dans l'optique de créer un barème évaluatif (section 4.7, CRIT-LEC-NN).

La section 4.8 offre une synthèse du profilage des traits distinctifs obtenus ainsi que des prototypes modalisateurs de grilles critériées pour évaluer la prosodie d'apprenants francophones sur la lecture du texte *The Selfish Giant*.

On a vu précédemment que par son objectif de catégorisation, qu'il soit personnalisé ou de groupe, le profilage consiste à minimiser certaines particularités distinctives jugées peu représentatives au profit d'autres jugées fondamentales, lesquelles sont alors amplifiées. L'objectif poursuivi est d'arriver à définir la caractérisation optimale la plus élégante (au sens mathématique du terme⁸⁶) afin de repérer les caractéristiques définissant un individu. Ce ciblage s'appuie sur des méthodes de catégorisation et de hiérarchisation procédant par inclusion ou exclusion des phénomènes à divers stades du processus de profilage.

3 Création d'une méthodologie d'organisation et de hiérarchisation temporelle des apprenants (expérimentation de sélection des fichiers d'apprenants étudiés STAT-NN1)

On a introduit au Chapitre I les projets *Diderot-Longdale*⁸⁷ et *Longdale-Charliphonia*⁸⁸ dont sont issus les corpus d'apprenants en lecture et conversation. Les métadonnées obtenues via questionnaire ou question informelle pendant l'échange conversationnel ont été précieuses lors des choix à opérer au cours de cette étude, en tout premier lieu pour la sélection des apprenants à effectuer.

⁸⁶ Expression reprise à partir de Durand J. (2005), « Chapitre 3 - Les primitives phonologiques : des traits distinctifs aux éléments », *Phonologie et phonétique : forme et substance*, page 71.

⁸⁷ URL : [<http://www.clillac-arp.univ-paris-diderot.fr/projets/longdale>]

⁸⁸ URL : [<http://www.clillac-arp.univ-paris-diderot.fr/projets/charliphonia>]

Le nombre important des fichiers constituant le corpus *Longdale-Charliphonia* implique une méthodologie de sélection aboutissant à des choix favorisant pertinence et faisabilité. Cette méthodologie se décompose en six phases et sont basées sur les influences linguistiques subies par les locuteurs, la durée de lecture constatée, la représentativité temporelle des fichiers à sélectionner, la valeur qualitative des fichiers et le critère de faisabilité.

3.1 Phase 1 : sélection en fonction des influences linguistiques

Mon sujet concernant l'étude de la prosodie des étudiants francophones enregistrés, j'ai donc dû écarter les locuteurs qui ne correspondaient pas à ce critère. Par « francophones », j'entends les étudiants de langue maternelle française n'ayant pas subi l'influence d'une autre langue que l'anglais (langue qu'ils étudient) et qui serait de nature à modifier leur interlangue lors des phases de production orale enregistrées.

Je me suis donc en premier lieu basée sur le résultat des fiches de renseignements, que les étudiants devaient compléter en ligne avant qu'un anglophone les enregistre. Dans le protocole du projet *Longdale*, trois questions servent à situer l'apprenant dans la genèse de son apprentissage linguistique :

- 1) What is your home country?
- 2) What is your native language?
- 3) What language(s) do you speak at home?

Notre étude ne prendra pas en compte la provenance géographique ou affective des apprenants (1), car son incidence sur le développement linguistique est négligeable. Par contre, j'ai scrupuleusement éliminé les sources linguistiques susceptibles d'avoir une influence prépondérante sur la prosodie. Dans un premier temps, ont été écartés de l'analyse tous les apprenants ayant une langue maternelle autre que le français (2), c'est-à-dire non seulement ceux dont la langue maternelle est différente du français, mais encore ceux mentionnant deux langues maternelles, même si le français est l'une d'elles. La question (3) a révélé que des apprenants considérant la France comme leur pays d'origine et le français comme langue maternelle exclusive, pouvaient néanmoins avoir un accent marqué correspondant à l'une des langues autres que le français mentionnées comme parlées dans le cercle familial. Aussi, par mesure de précaution, tous les apprenants en contact avec une langue autre que le français dans leur environnement familial ont aussi été écartés du présent corpus d'étude.

3. Création d'une méthodologie d'organisation et de hiérarchisation temporelle des apprenants (expérimentation de sélection des fichiers d'apprenants étudiés STAT-NN1)

Nombre total des enregistrements <i>The Selfish Giant</i> (S001R) 93 en 2009, 112 en 2010	205
Apprenants dont le français n'est pas l'unique langue maternelle (question 2)	-37
Apprenants dont le français est l'unique langue maternelle mais n'est pas l'unique langue parlée dans le cercle familial (question 3)	-13
Total des fichiers d'apprenants exploitables dans notre étude sur l'enregistrement S001R	155

Tableau 64 – Tableau récapitulatif des apprenants écartés de cette étude

Par contre, n'ont pas été éliminés les huit apprenants dont la langue anglaise était spécifiée comme langue maternelle (en parallèle ou non avec le français) ou était parlée dans l'environnement familial. Il est probable que cette donnée influe très fortement sur leur interlangue. L'objet de notre étude étant précisément l'élaboration d'une typologie des différents stades de cette langue intermédiaire entre le français et l'anglais d'un bout à l'autre du spectre des possibles, ces fichiers d'apprenants bilingues ont été retenus pour analyse car ils correspondent à un degré d'interlangue particulièrement proche de celui des natifs anglophones monolingues. Ils prendront donc toute leur place dans notre étude. De plus, la place de l'anglais dans l'interlangue tient à mes yeux une place particulière, car je conçois non seulement l'interlangue comme un continuum, mais aussi ce continuum en tant que partie intégrante d'un continuum plus vaste incluant la parole des natifs, des locuteurs les plus maladroits aux plus expérimentés. Ce point sera étudié dans l'analyse comparative des productions des apprenants. Ainsi, les apprenants se présentant comme bilingues ou ayant subi une forte influence de la langue anglaise prennent naturellement leur place dans cet immense continuum.

3.2 Phase 2 : classement par la durée de lecture

Une fois écartés les locuteurs non conformes à l'étude, le corpus se réduisait à 155 apprenants. J'ai dû franchir une autre étape de la sélection car leur nombre était trop important pour les étudier tous en détail. J'ai proposé ailleurs (Cauvin 2012) une analyse des descripteurs du *CECRL* pour caractériser la langue orale selon les systèmes de niveaux de compétence langagière introduits au Chapitre I. Ainsi, l'objectif recherché étant une représentativité des mesures prises, j'ai opté pour la pertinence prosodique de la vitesse d'élocution dans l'évaluation de l'interlangue. J'ai sélectionné les apprenants en fonction de ce critère pour en déterminer des catégories que j'ai étudiées séparément. J'ai ainsi mesuré chacun des 155 fichiers S001R en les élaguant dans le logiciel *WaveLab 6.11* (Steinberg) de manière à ne garder que la partie correspondant au texte lu. Cette mesure est précise à la

seconde près et c'est la seconde inférieure qui a été choisie dans les cas où la mesure se situait entre deux unités de secondes. Selon les préconisations du domaine phonétique, elle ne peut être qu'approximative. En fait, une précision plus poussée n'aurait pas eu de plus grande valeur à cause de la longueur des pauses, et des éventuelles répétitions, reprises, et autres disfluences, très variables d'un locuteur à l'autre.

Selon ces critères de sélection, la lecture la plus courte a totalisé 121 secondes et la plus longue 261 secondes. Le graphique ci-dessous montre la répartition de la durée des 155 lectures sur une étendue de 141 secondes. On remarque que la médiane se trouve autour des 170 secondes.

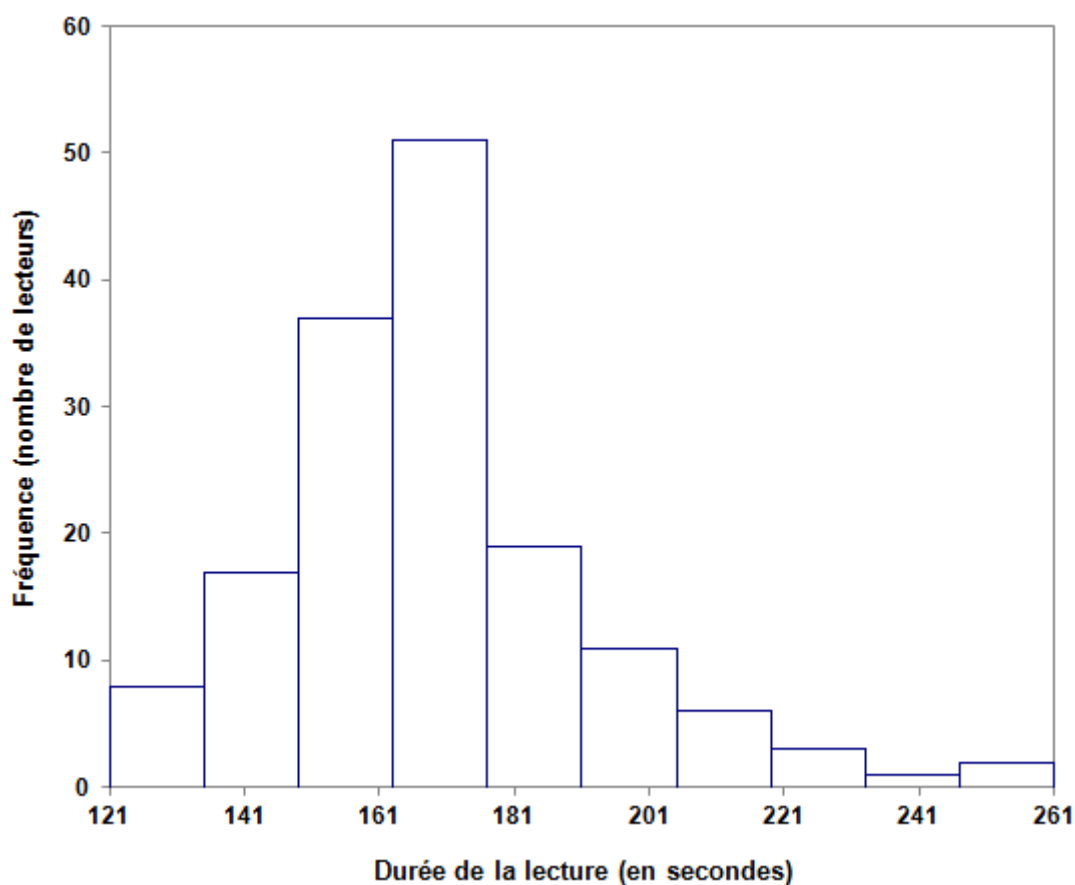


Figure 115 – Répartition du nombre de fichiers S001R selon la durée de lecture du texte

3.3 Phase 3 : sélection selon des repères de représentativité temporelle de l'effectif

Une fois les locuteurs sélectionnés et ordonnés les uns par rapport aux autres selon le critère de durée de lecture, il s'est agi d'organiser des repères destinés à créer des catégories représentatives de l'ensemble des apprenants du corpus dans leur diversité. Ainsi, j'ai choisi de sélectionner les fichiers S001R correspondant aux repères du premier décile, du premier

quartile, de la médiane, du troisième quartile et du neuvième décile pour déterminer les locuteurs à étudier. L'effectif de départ prévu pour chaque catégorie est de huit éléments pour pouvoir capter un mouvement d'ensemble, le confronter à celui des autres, tout en minimisant la variation individuelle et faire surgir la variation révélatrice de la catégorie. Cela constitue 40 apprenants (5 catégories x 8 locuteurs = 40 fichiers S001R).

Une recherche longitudinale requiert un corpus conséquent. Seuls dix étudiants correspondant aux critères linguistiques ont été en mesure de suivre l'intégralité du parcours. Ils ont donc tous été retenus pour cette étude.

3.4 Phase 4 : sélection en fonction de critères qualitatifs

Une fois les repères déterminés, des choix spécifiques ont dû être opérés. J'ai suivi un principe de base enjoignant de sélectionner les fichiers correspondant au repère et ceux répartis de part et d'autre du repère. Néanmoins, certains fichiers ont été éliminés pour favoriser les critères suivants :

- Qualité sonore de l'enregistrement (afin d'éviter les bruits de fond, claquements, voix distantes, etc. qui rendraient l'analyse automatique difficile, trop aléatoire voire erronée). L'écoute entière du fichier a permis d'évaluer la qualité sonore générale.
- Conformité de l'exclamative et de la question en *What* pour étude détaillée (bredouillements évités). Une écoute sélective des fichiers et un visionnement de ces parties dans *Praat* m'ont permis d'opérer les vérifications nécessaires.

Si le cas se présente comme possible, les locuteurs masculins ont été choisis pour tenter de rééquilibrer la parité dans le corpus, qui comprend un très faible nombre de locuteurs masculins.

3.5 Phase 5 : sélection de locuteurs à retenir

Ainsi, les fichiers sélectionnés dans une catégorie par rapport à un repère représentent les caractéristiques suivantes.

Les huit fichiers de lecture ayant le premier décile (D1) pour repère représentent un écart de six secondes sur un total de 141, qui lui-même correspond à six centiles. Les locuteurs masculins (fichiers bleus) et féminins (fichiers roses) sont équitablement répartis (quatre hommes et quatre femmes).

Numéro de fichier	Durée de la lecture (en secondes)	Centiles correspondants	Écart autour de D1
DID0059-S001R	139	7	= 6 secondes = 6 centiles (1 ^{er} décile = 142,4)
DID0117-S001R	140		
DID0191-S001R	140		
DID0210-S001R	142		
DID0078-S001R	143		
DID0018-S001R	144	12	
DID0205-S001R	144	12	
DID0221-S001R	144	12	

Tableau 65 – Positionnement des fichiers S001R sélectionnés pour la catégorie D1

Les huit fichiers de lecture ayant le premier quantile (Q1) pour repère représentent un écart de cinq secondes sur un total de 141, qui lui-même correspond à 6,5 centiles. Un seul locuteur masculin a pu être trouvé (un homme et sept femmes dans la catégorie Q1).

Numéro de fichier	Durée de la lecture (en secondes)	Centiles correspondants	Écart autour de Q1
DID0066-S001R	152	20,5	= 5 secondes = 6,5 centiles (1 ^{er} quartile = 154)
DID0151-S001R	152	20,5	
DID0195-S001R	153		
DID0003-S001R	154	25	
DID0024-S001R	154	25	
DID0108-S001R	156	27	
DID0211-S001R	156	27	
DID0212-S001R	156	27	

Tableau 66 – Positionnement des fichiers S001R sélectionnés pour la catégorie Q1

Les huit fichiers de lecture ayant la médiane pour repère représentent un écart de cinq secondes sur un total de 141, qui lui-même correspond à huit centiles. La catégorie Médiane comprend deux hommes et six femmes.

CHAPITRE IV – Expérimentation de création de profils et de critères d'évaluation prosodiques d'apprenants

Numéro de fichier	Durée de la lecture (en secondes)	Centiles correspondants	Écart autour de la Médiane
DID0189-S001R	166	46	= 5 secondes = 8 centiles (médiane = 168)
DID0062-S001R	167	48	
DID0135-S001R	167	48	
DID0046-S001R	168	50	
DID0121-S001R	168	50	
DID0200-S001R	168	50	
DID0067-S001R	170	53	
DID0112-S001R	170	53	

Tableau 67 – Positionnement des fichiers S001R sélectionnés pour la catégorie Médiane

Les huit fichiers de lecture ayant le troisième quartile (Q3) pour repère représentent un écart de sept secondes sur un total de 141, qui lui-même correspond à 8,27 centiles. La catégorie Q3 regroupe cinq hommes et trois femmes.

Numéro de fichier	Durée de la lecture (en secondes)	Centiles correspondants	Écart autour de Q3
DID0168-S001R	175	70	= 7 secondes = 8,272727 centiles (3 ^{ème} quartile = 178,5)
DID0132-S001R	176		
DID0162-S001R	176		
DID0057-S001R	177		
DID0045-S001R	180		
DID0074-S001R	180		
DID0154-S001R	180		
DID0068-S001R	181	77,272	

Tableau 68 – Positionnement des fichiers S001R sélectionnés pour la catégorie Q3

Les huit fichiers de lecture ayant le neuvième décile (D9) pour repère représentent un écart de sept secondes sur un total de 141, qui lui-même correspond à sept centiles. Un seul locuteur masculin figure dans la catégorie D9 (un homme et sept femmes).

Numéro de fichier	Durée de la lecture (en secondes)	Centiles correspondants	Écart autour de D9
DID0072-S001R	190	85	= 7 secondes = 7 centiles (9 ^{ème} décile = 194,6)
DID0056-S001R	191		
DID0153-S001R	191		
DID0209-S001R	191		
DID0138-S001R	193		
DID0040-S001R	194		
DID0076-S001R	196	91	
DID0192-S001R	196	91	

Tableau 69 – Positionnement des fichiers S001R sélectionnés pour la catégorie D9

Dans chacune des cinq catégories susmentionnées, le rapport entre la différence maximale de durée des lectures de la catégorie et celle de la totalité du corpus est inférieure à 5% ($7/141 \times 100 = 4,96\%$). La moyenne des cinq valeurs est de six secondes. En ce qui concerne les centiles, la moyenne des cinq valeurs est de 7,15 centiles et la valeur maximale est 8,27 centiles.

La représentativité des fichiers comme membre de leur catégorie a été préservée, malgré une qualité d'enregistrement parfois peu conforme aux standards ordinairement utilisés pour ce genre d'étude (enregistrements d'extension mp3 convertis par la suite en wav). Il est de bon ton d'éliminer radicalement tous ces fichiers. Il m'a semblé au contraire nécessaire de privilégier le rôle (et donc la place) des locuteurs en accord avec la méthodologie adoptée, même si cela complexifie l'analyse instrumentale. Les efforts fournis pour traiter des fichiers d'une piètre qualité d'enregistrement seront sans doute récompensés lors de la phase d'analyse, l'objectif ultime étant de faire surgir le sens à partir des données fournies par les apprenants. Quelques tests préliminaires ont permis d'écarter les fichiers non utilisables.

Les locuteurs ayant enregistré la totalité du parcours sont au nombre de dix. Les fichiers S001R pour étude longitudinale se positionnent dans le corpus ainsi que le précise le tableau ci-dessous.

Numéro de fichier	Durée de la lecture (en secondes)	Centiles correspondants	Catégorie d'étude
DID0106-S001R	132	3	Hors catégories
DID0128-S001R	145	14	Hors catégories
DID0024-S001R	154	25	1 ^{er} quartile
DID0108-S001R	156	27	1 ^{er} quartile
DID0062-S001R	167	48	Médiane
DID0135-S001R	167	48	Médiane
DID0168-S001R	175	70	3 ^{ème} quartile
DID0068-S001R	181	77,272	3 ^{ème} quartile
DID0071-S001R	184	79,220	Hors catégories
DID0035-S001R	185	80	Hors catégories

Tableau 70 – Tableau de positionnement dans le corpus des fichiers S001R destinés à une étude longitudinale

L'étude de la lecture du texte *The Selfish Giant* des apprenants peut donc se baser sur un « vivier » comprenant 44 fichiers S001R et dix fichiers S004R1. Cette sélection donne toute latitude pour étudier les phénomènes prosodiques transversalement et longitudinalement.

3.6 Phase 6 : réduction du nombre de locuteurs selon le critère de faisabilité

Parmi ces 54 fichiers, une dernière phase de sélection a été opérée afin de réduire leur nombre à quinze pour raison de faisabilité. Cette nouvelle sélection, bien que présentée ici, n'a pu être effectuée qu'après l'étude des Chapitres II et III, lesquels déterminent les parties de la lecture à privilégier d'après les réalisations natives et la quantité de travail à effectuer sur les fichiers non natifs dans le temps imparti pour cette thèse. Ainsi, seuls trois fichiers S001R ont finalement été retenus dans chacune des cinq grandes catégories temporelles et les fichiers S004R1 sont relégués à une étude post-thèse pour une vérification et un approfondissement des phénomènes. Le choix de ces fichiers tend à favoriser une parité des genres (40% de locuteurs masculins et 60% féminins) ainsi que ceux les plus adaptés à l'étude transversale et celle d'une suite longitudinale. Le tableau suivant en récapitule la sélection.

Numéro de fichier	Durée de la lecture	Catégorie d'étude
-------------------	---------------------	-------------------

	(en secondes)	
DID0210-S001R	142	1 ^{er} décile (= 142,4 secondes)
DID0078-S001R	143	
DID0018-S001R	144	
DID0003-S001R	154	1 ^{er} quartile (= 154 secondes)
DID0024-S001R	154	
DID0108-S001R	156	
DID0062-S001R	167	Médiane (= 168 secondes)
DID0135-S001R	167	
DID0046-S001R	168	
DID0168-S001R	175	3 ^{ème} quartile (= 178,5 secondes)
DID0074-S001R	180	
DID0068-S001R	181	
DID0056-S001R	191	9 ^{ème} décile (= 194,6 secondes)
DID0153-S001R	191	
DID0076-S001R	196	

Tableau 71 – Tableau de positionnement dans le corpus des quinze fichiers S001R sélectionnés pour l'étude de cette thèse

3.7 Résultat obtenu en fin des six phases du protocole de sélection des fichiers d'apprenants

Cette sélection finale prend en compte l'ensemble de multiples critères de représentativité par la qualité linguistique (Phase 1), la durée de lecture (Phase 2) et la représentativité selon ce classement (Phase 3), de diversité par le genre et de qualité acoustique des fichiers sons (Phase 4), de poursuite du travail d'investigation au-delà de la présente thèse dans une optique longitudinale (Phase 5), ainsi que de faisabilité (Phase 6). Ces fichiers sons seront analysés dans la suite de notre étude.

Cette expérimentation STAT-NN1 aboutit à l'élaboration d'une grille méthodologique de classement des apprenants selon la durée de lecture du texte *The Selfish Giant*, et valide l'Hypothèse 3, rappelée ci-dessous :

H3 — Le paramètre de rapidité d'élocution étant un indicateur déterminant de niveau des apprenants, la catégorisation de ceux-ci selon ce paramètre devrait permettre d'organiser la variation inter-locuteurs des résultats des autres variables prosodiques afin de trouver des critères d'évaluation fiables.

Cette avancée concernant l'application du critère temporel est décisive dans une première approche puisqu'elle ouvre la voie à la mise en place d'une vérification de sa pertinence, puis à celle d'autres profilages pour atteindre l'objectif évaluatif. Mais tout

d'abord, il est nécessaire d'exposer les critères méthodologiques mis en œuvre pour y parvenir.

4 Méthodologie du profilage (expérimentations COR-AC-NN, PER-LEX-NN, PER-THEM-NN)

La méthodologie adoptée ici pour l'analyse de la prosodie des apprenants fait suite aux précédentes expérimentations sur les réalisations natives COR-AC-N1 (Chapitre II), COR-AC-N2 et COR-DIV-N3 (Chapitre III) et prend appui sur leurs avancées. En fonction de la nature acoustique ou perceptive, quantitative ou qualitative de chaque variable, une méthodologie particulière a été adoptée pour observer la variation et en rendre compte quantitativement, puis par la suite, qualitativement.

4.1 Base méthodologique de création des variables d'origine acoustique (expérimentation COR-AC-NN) ou perceptive (expérimentations PER-THEM-NN et PER-LEX-NN)

Rappelons le point de départ de ce chapitre : la grille d'analyse des apprenants.

	<i>Domaine analysé</i>	<i>Sous-domaines</i>	<i>Variables obtenues et unités de mesure</i>
TEMPORALITÉ	Durée	<ul style="list-style-type: none"> • Durée du texte brut • Durée des 3 segments non ponctués ABC avec et sans pauses (P4, 16 et 25) 	<ul style="list-style-type: none"> • Quantitatives (seconde)
	Vitesse d'élocution	<ul style="list-style-type: none"> • Vitesse d'élocution dans le texte entier • Vitesse d'élocution dans 3 segments non ponctués ABC avec et sans pauses (P4, 16 et 25) 	<ul style="list-style-type: none"> • Quantitative (nombre de syllabes par seconde)
REGISTRE	Registre : Ampleur du registre	<ul style="list-style-type: none"> • Sur les marqueurs emphatiques : exclamative (P5), paires <i>so</i> + adjectif/adverbe (P4 et P22), syllabe accentuée de <i>nobody</i> (P11) • Dans phrase déclarative au degré d'emphase faible (P9) 	<ul style="list-style-type: none"> • Quantitatives (F0 max - F0 min en demi-tons)
MÉLODIE	Mélodie dans la phrase : Contour mélodique dans syntagme intonatif	<ul style="list-style-type: none"> • <i>Fall</i> : exclamative (P5) et déclarative (P9) 	<ul style="list-style-type: none"> • Qualitatives numériques à partir de mesures quantitatives (DT)
	Mélodie dans la syllabe : Ton syllabique sur nucléus	<ul style="list-style-type: none"> • <i>Fall-rise</i> : <i>aRRIVED</i> (P9) • <i>Rise-fall</i> et tons tridirectionnels : <i>Nobody</i> (P11) 	<ul style="list-style-type: none"> • Qualitatives numériques à partir de mesures quantitatives (DT)
RYTHME	Rythme : maîtrise du placement accentuel dans les polysyllabes	<ul style="list-style-type: none"> • Dans 6 polysyllabes, conformité des schémas accentuels à ceux des natifs par test perceptif 	<ul style="list-style-type: none"> • Qualitative numérique à partir de moyennes (conforme, non conforme)
	Rythme : proéminence de marqueurs spécifiques	<ul style="list-style-type: none"> • Paires <i>so</i> + adjectif/adverbe (P 4 et P22) • Paires <i>all</i> + préposition (P12 et P16) 	<ul style="list-style-type: none"> • Qualitatives numériques à partir de mesures quantitatives (DT)
		<ul style="list-style-type: none"> • Proéminence sur <i>them</i> par test perceptif (P4) 	<ul style="list-style-type: none"> • Qualitative numérique à partir de moyennes (proéminent, non proéminent)
	Rythme sur syllabes spécifiques (approche temporelle)	Rapport durée inter-syllabique sur la paire de syllabes <i>sat at</i> (P19)	<ul style="list-style-type: none"> • Quantitatives (seconde, %)

Tableau 72 – Sélection des variables pour non natifs consécutive aux Chapitres II et III

Cette grille globale montre que les domaines analysés concernent les phénomènes de différences de durée, d'ampleur de registre, de variation mélodique et de choix rythmiques. En incluant la mélodie, elle se veut plus précise que celle qui avait originellement servi à

détecter les profils de natifs et en cela, rejoint les préconisations de Crystal. L'analyse du rythme tient compte de variables accentuelles vérifiées sur *loci* fixes concernant leur proéminence dans la phrase, le mot et la syllabe.

Deux méthodes ont été utilisées pour constituer ces variables, soit par des mesures acoustiques de fréquence fondamentale ou de durée, soit par des tests de perception de proéminence accentuelle. Alors que les mesures acoustiques sont opérées selon les mêmes principes que dans les chapitres précédents, les expérimentations perceptives portant sur le placement accentuel lexical et l'éventuelle proéminence sur pronom en fin de syntagme intonatif ont nécessité une méthodologie adaptée.

4.2 Profilage du respect de la non proéminence d'un pronom en fin d'énoncé par test perceptif (expérimentation PER-THEM-NN)

Ainsi, le repérage de l'accentuation du pronom *them* en fin de syntagme intonatif n'ayant fourni aucune piste intéressante par la seule mesure de la F0, je me suis orientée vers un repérage perceptif car la multiplicité des facteurs intervenant dans le processus aurait été trop chronophage à mettre en œuvre sans même donner l'assurance d'obtenir un résultat adéquat. Pour cela, il a été demandé à cinq auditeurs (un francophone spécialiste et quatre anglophones enseignants d'anglais) de repérer la dernière syllabe proéminente dans la fin de la Phrase 4 chez les quinze apprenants sélectionnés dans le corpus. Les résultats ont révélé de nombreuses disparités. Pourtant, les mêmes résultats ont été observés chez la spécialiste francophone et chez une anglophone. En calculant un pourcentage d'occurrences à partir du pointage du pronom comme *locus* de dernière proéminence, j'ai pu obtenir une variable qualitative numérique qui m'a permis de hiérarchiser les résultats obtenus.

4.3 Profilage du respect du patron accentuel lexical par test perceptif (expérimentation PER-LEX-NN)

Les six polysyllabes lus par les quinze apprenants constituent 90 items à analyser finement en contexte. Pour cela, les auditeurs ont pu les écouter à loisir. Tout d'abord, un premier tri a été effectué par deux enseignantes d'anglais francophones spécialistes. La référence était le schéma accentuel fourni par les dictionnaires de référence *Longman Pronunciation Dictionary* et *Cambridge English Pronouncing Dictionary*, selon la transcription de Guierre. Les enseignantes ont tenté de coder ce qu'elles percevaient avec cette transcription. Ce travail est long, fastidieux et approximatif, mais la formule conditionnelle

=SI(F5<>E5;"à vérifier";SI(E5=D5;"juste";"faux"))

a permis d'effectuer un premier tri dans le tableur *Excel* :

- Schéma « juste » : correspondant exactement au modèle natif et validé par les deux enseignantes,
- Schéma « faux » : aucune des deux enseignantes n'a établi de correspondance avec le modèle natif,
- Schéma « à vérifier » lorsqu'il y a divergence d'appréciation ou qu'il n'y a pas adéquation avec les deux cas précédents.

Les réalisations « à vérifier » sont alors soumises à deux anglophones après leur avoir expliqué le système de codage, et leurs résultats respectifs sont alors codés sur le tableur. Le bilan des auditrices anglophones est traité de la même manière que précédemment avec la formule

=SI(H6<>I6;"à vérifier"; SI(I6=D6;"juste";"faux"))

Le résultat des deux formules est comparé :

=SI(G5=J5;"faux";J5)

Le bilan final est établi par la dernière formule

=SI(G5<>"à vérifier";G5;SI(G5=J5;"faux";J5))

le diagnostic de départ des deux francophones est repris dans les cas où il était « juste » ou « faux ». Dans le cas inverse, s'il est douteux (« à vérifier »), il est considéré comme faux, et dans le cas où les deux anglophones ont perçu le même schéma, c'est leur perception qui est finalement retenue. Les mentions « juste » sont comptabilisées par locuteur pour constituer ici aussi une variable qualitative numérique.

L'ensemble de ces variables d'origine acoustique ou perceptive est ainsi présenté sous forme numérique afin d'être exploité dans les diverses analyses destinées à établir des profils de groupes par archétypes de conformité, puis des profils de groupes par archétypes stéréotypiques, pour aboutir à des prototypes de *criterial features*.

5 Analyse par profilage de groupes par archétypes de conformité selon la vitesse de lecture (expérimentation STAT-NN2)

Cette section présente une première méthodologie de classement des réalisations prosodiques des apprenants et ses résultats. Après avoir déterminé les objectifs et exposé les finalités des profils de conformité chez les apprenants (section 5.1), l'étude procède à un examen des variables dans les quatre grands domaines ciblés que sont la temporalité par la vitesse et le débit d'élocution (section 5.2), le registre (section 5.3), la qualité mélodique (section 5.4) et la maîtrise du rythme (section 5.5).

5.1 Visée et enjeux des profils par archétypes de conformité

La présente étude est destinée à repérer les caractéristiques des groupes d'apprenants en fonction du classement opéré selon leur seule vitesse de lecture du texte *The Selfish Giant*. Le profilage concerne les cinq catégories D1, Q1, Médiane, Q3 et D9, lesquelles sont représentées par trois apprenants chacune. L'objectif est de tenter de dresser une fiche signalétique de chaque groupe, d'en établir la norme, de hiérarchiser ces traits entre eux de manière à pouvoir immédiatement étiqueter l'appartenance d'un locuteur *lambda* à l'un de ces groupes par sa conformité à la norme établie grâce aux points communs observés à l'intérieur de chaque groupe. Cette première phase exploratoire permet d'accéder aux profils de groupes. La question est de détecter la saillance de leurs traits. En contrepartie, le faisceau de variables marquant l'individualité du locuteur est destiné à être réduit au minimum pour constituer la norme censée régir le groupe.

L'enjeu est de savoir si cette catégorisation temporelle détermine les autres catégorisations (rythmique, mélodique ou de registre), ou bien si une hiérarchisation selon ces autres catégorisations devrait être envisagée car elle serait susceptible d'apporter un éclairage plus intéressant pour étudier les phénomènes et les hiérarchiser. Il s'agit là de chercher les limites du classement par seul critère de rapidité de lecture.

Ce profilage par archétype de conformité soumet chacun des cinq groupes aux variables considérées, en utilisant les profils personnalisés pour préciser la nature distributive (par leurs propriétés définitoires) ou non distributive (par leurs propriétés accessoires) de la catégorisation résultant d'un profilage prenant uniquement appui sur la rapidité de lecture.

5.2 Profil des apprenants dans le domaine temporel : vitesse et débit d'élocution

Le domaine temporel est exploré sous deux angles. Tout d'abord la rapidité de lecture est examinée par la seule durée de lecture (section 5.2.1) ; ensuite, elle l'est par le débit d'élocution, c'est-à-dire le nombre de syllabes pouvant être insérées dans cette durée (section 5.2.2).

5.2.1 Rapidité de lecture par la durée de lecture

Cette rubrique se focalise sur la vitesse de lecture du texte d'une part, et de l'autre, la vitesse de lecture dans les trois tronçons non ponctués A, B et C mentionnés précédemment. Bien que son rôle ne soit qu'illustratif, puisque les éléments présentés ici ne correspondent en fait qu'aux données essentiellement produites par les résultats de l'expérimentation STAT-NN1 effectuée en début de chapitre, elle offre néanmoins une symétrie par rapport aux analyses qui suivront et permettra de mieux comparer les écarts par la visualisation des données.

On constate dans le graphique ci-dessous une homogénéité maximale étant donné que la rapidité de lecture constitue le critère de sélection des apprenants étudiés.

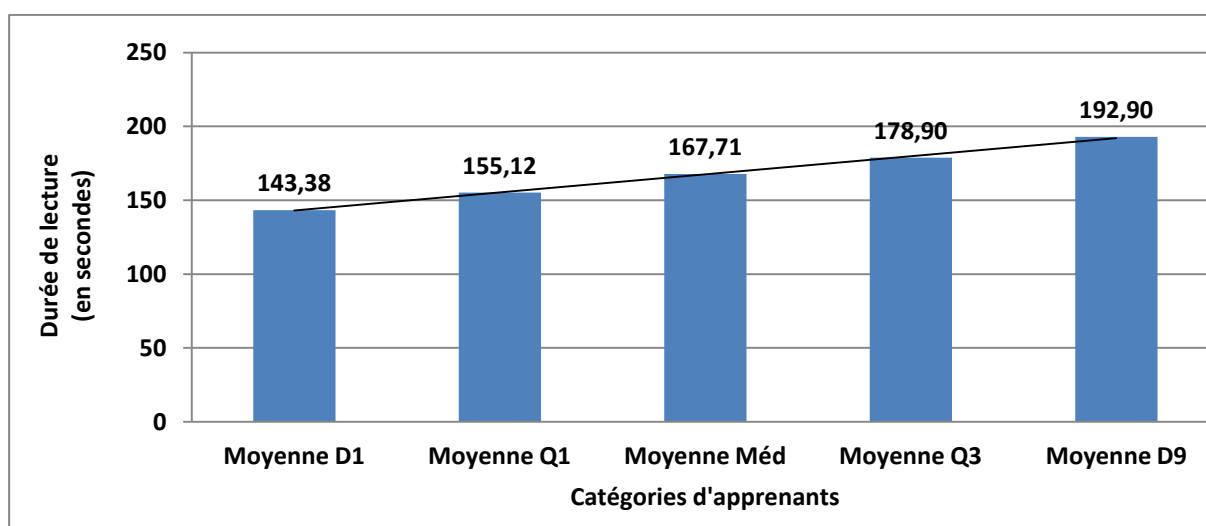


Figure 116 – Durée de lecture du texte par catégorie d'apprenants

Dans le graphique suivant, la hauteur des colonnes coiffée par les lignes droites montre que l'homogénéité n'est pas parfaite entre les diverses catégories temporelles, que les pauses éventuelles aient été incluses ou non. Ces deux droites sont presque parallèles et tendent à s'éloigner lorsque la durée de lecture augmente, ce qui montre que la présence de

pauses dans ces tronçons qui sont censés ne pas en contenir est une caractéristique non négligeable dans la détermination de la variable.

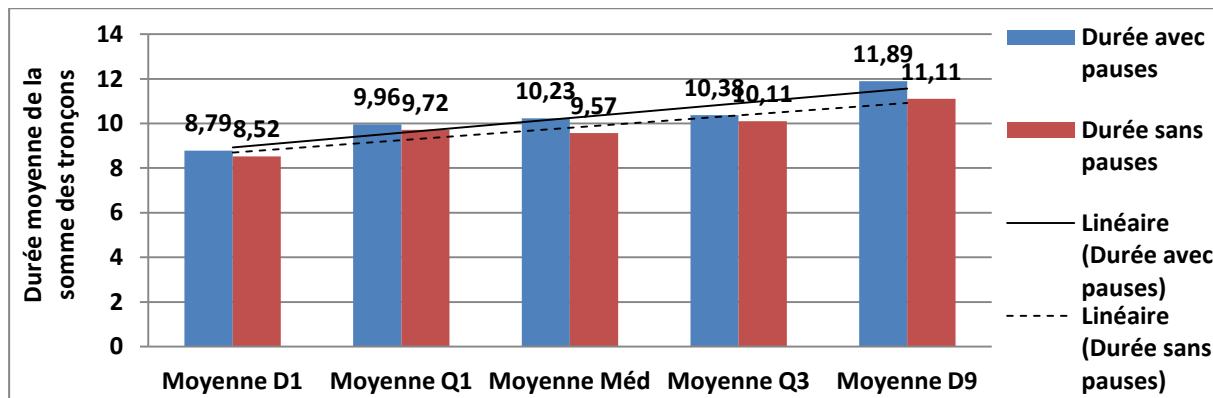


Figure 117 – Durée moyenne du total des tronçons A, B et C pauses incluses et pauses exclues par catégorie d'apprenants

5.2.2 Débit d'élocution par le nombre de syllabes par seconde

Le débit d'élocution correspond ici à la vitesse de lecture du texte par le nombre de syllabes prononcées en moyenne par L1 et L5, c'est-à-dire 541. Il doit être noté qu'il s'agit là d'une constante servant de mesure de base pour faciliter les calculs, étant donné que les disfluences des apprenants (répétitions, reprises, oublis, ajouts) leur ont souvent fait prononcer un nombre de syllabes bien différent et souvent supérieur. Le graphique ci-dessous montre le nombre moyen de syllabes prononcées par catégorie. Bien que moins parfait que le graphique concernant uniquement la vitesse de lecture du texte, ce graphique montre une grande homogénéité et permet d'établir des fourchettes dans la vitesse d'élocution brute en lecture.

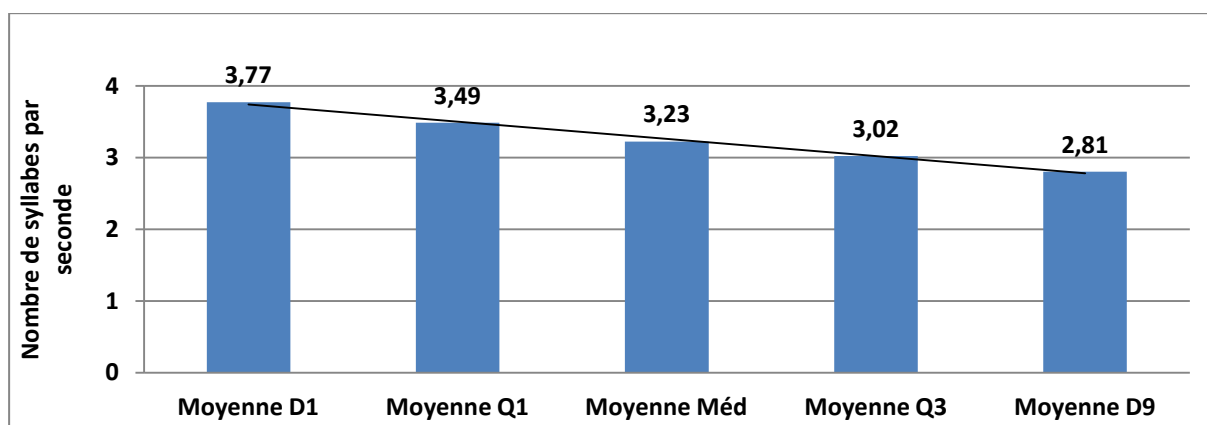


Figure 118 – Nombre moyen de syllabes par seconde par catégorie d'apprenants dans l'ensemble du texte en prenant en compte les 541 syllabes prononcées en moyenne par L1 et L5 (lecteurs repères)

Le graphique qui suit est plus fiable car le nombre de syllabes prononcées a été comptabilisé avec précision. Les résultats montrent les vitesses de pointe des apprenants dans

ces segments. Les deux courbes polynomiales, bien que suivant toujours la direction observée dans la droite de la précédente figure, montrent que, malgré la stabilité de ce type de variables, une multiplicité de données doivent être prises en compte pour évaluer les réalisations liées aux divers aspects de la temporalité.

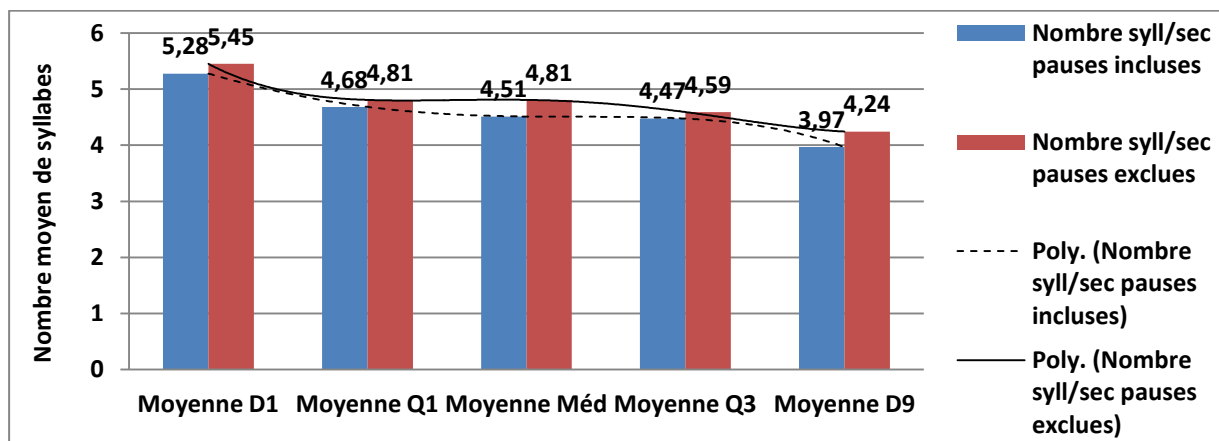


Figure 119 – Nombre moyen de syllables par seconde et par catégorie d'apprenants dans l'ensemble des trois tronçons A, B et C

5.2.3 Bilan

Ces graphiques montrent l'homogénéité générale des résultats. Ils présentent des profils distributifs idéaux pour servir de base à des critères d'évaluation. Une analyse plus poussée pourra déterminer lesquelles de ces six variables sont les plus représentatives dans l'établissement d'un profil. En tout état de cause, il est extrêmement chronophage de compter toutes les syllables d'un texte, alors qu'il est beaucoup plus aisé et rapide de calculer la somme de celles que l'on trouve dans un tronçon non ponctué. De plus, ce modèle de mesure peut être transposé à des réalisations en reformulation ou conversation spontanée, ce qui n'est pas le cas du modèle de mesures prises sur un texte entier, ni de leur résultat.

5.3 Profil des apprenants selon la variable du degré d'ampleur de registre

Cette section examine tout d'abord la méthodologie de mesure utilisée pour le registre (section 5.3.1) avant de s'orienter vers l'étude spécifique de l'exclamative et la Phrase 5 (section 5.3.2), des Phrases 4 et 22 où sont insérés les tronçons « *so* + adjectif ou adverbe » (section 5.3.3), la Phrase 9 pour son syntagme intonatif final au degré d'emphase réduit (section 5.3.4) et la Phrase 11 où figure le mot emphatique *nobody* pour sa syllabe accentuée (section 5.3.5). Un étude comparative des divers registres établis constituera une ouverture sur l'apport de ces variables (section 5.3.6). Par souci de transparence dans la démarche suivie, un

fichier *Excel* récapitulant les mesures et leur conversion figure en Annexe. Les différents traits critériés recueillis dans la thèse en découlent.

5.3.1 Méthodologie de prise de mesures du registre

L'ampleur du registre a été mesurée en tenant compte de F0max et F0min sur les zones sélectionnées comme très emphatiques chez les natifs dans les Phrases 4 et 22 (*so* + adjectif ou adverbe), Phrase 5 (exclamative) et Phrase 11 (*nobody*). À titre de comparaison, le syntagme intonatif correspondant à la proposition principale de la Phrase 9 a aussi été étudié par F0mean pour la mesure la plus haute et F0min pour la mesure la plus basse. Cette phrase étant minimalement emphatique, l'utilisation de F0mean n'apporte pas de grandes modifications par rapport à F0max et permet d'utiliser les mesures effectuées pour l'analyse mélodique présentée en aval. L'ensemble de ces mesures permet d'établir des moyennes ainsi que de déterminer F0min et F0max pour chaque groupe.

5.3.2 Registre de l'emphase dans l'exclamative et la Phrase 5

L'utilisation des demi-tons ne nécessite pas de distinction entre voix d'hommes et voix de femmes. D'ailleurs, la moyenne du troisième quartile est composée uniquement de voix d'hommes et l'on constate dans le graphique ci-dessous, que pour l'intégralité de la Phrase 5, ce sont les locuteurs masculins qui totalisent le registre maximal.

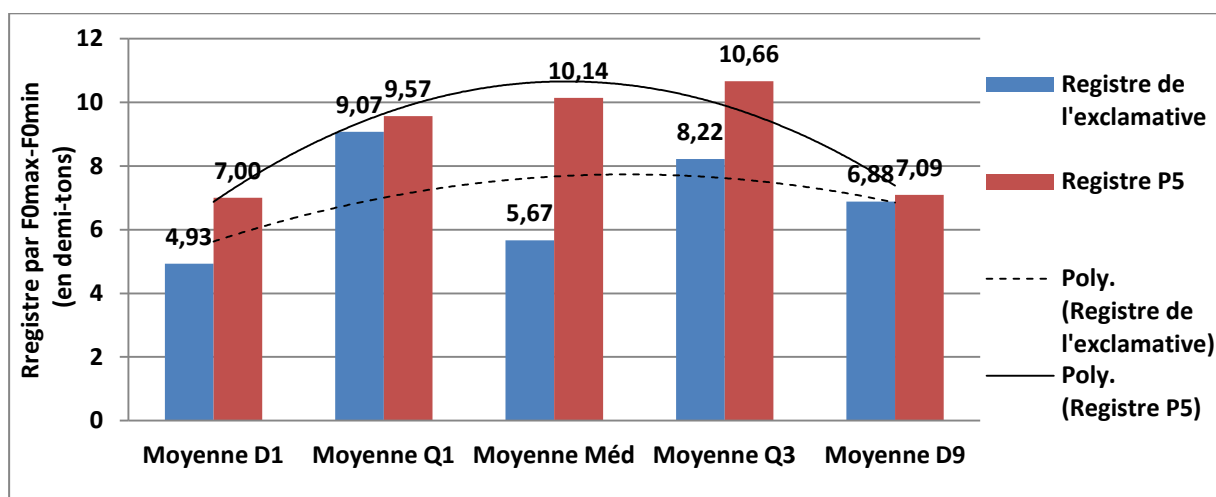


Figure 120 – Comparaison de l'ampleur du registre dans l'exclamative avec adjonction du discours citant (Phrase 5) par catégorie d'apprenants selon la vitesse de lecture

En comparant les diverses moyennes par groupe, on constate que l'ampleur maximale du registre se trouve sur l'association de l'exclamative et de son énoncé citant, c'est-à-dire l'intégralité de la Phrase 5. De plus, c'est elle qui correspond le plus à la catégorisation selon

la vitesse de lecture car l'on voit que les apprenants de la Médiane sont très en deçà de la courbe générale lorsqu'on ne considère que l'exclamative. On peut néanmoins observer que les deux mesures moyennes du premier et du neuvième décile sont les plus basses.

5.3.3 Registre de l'emphase dans la Phrase 4 et le segment de Phrase 22

Le graphique suivant montre que le traitement des deux paires « *so* + adjectif ou adverbe » est très différent. Alors que l'ampleur du registre mesurée dans la Phrase 4 avec F0max sur *so* ou la première syllabe de *sweetly* correspond vaguement à une certaine régularité comme le montre la courbe polynomiale en pointillés, on obtient des résultats quasiment inverses dans la Phrase 22, où l'on constate que l'ampleur est maximale dans le premier et le neuvième décile. Les mesures de la Médiane restent faibles dans les deux cas de figure.

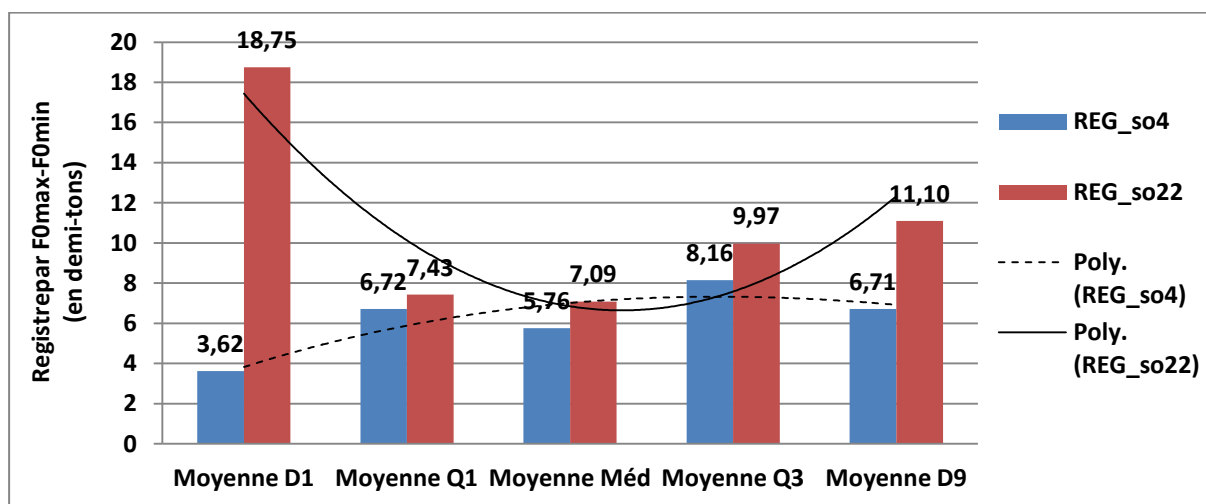


Figure 121 – Comparaison de l'ampleur du registre sur *so* + adjectif/adverbe dans les Phrases 4 et 22 par catégorie d'apprenants

5.3.4 Registre dans le segment à emphase réduite de la Phrase 9

Le graphique représentant l'ampleur du registre dans le syntagme *he saw the children playing in the garden*, rappelle celui de la Phrase 5 et de l'exclamative dans les relations que les catégories entretiennent entre elles. Le registre est néanmoins le plus bas dans D9 alors qu'il est relativement moyen dans D1.

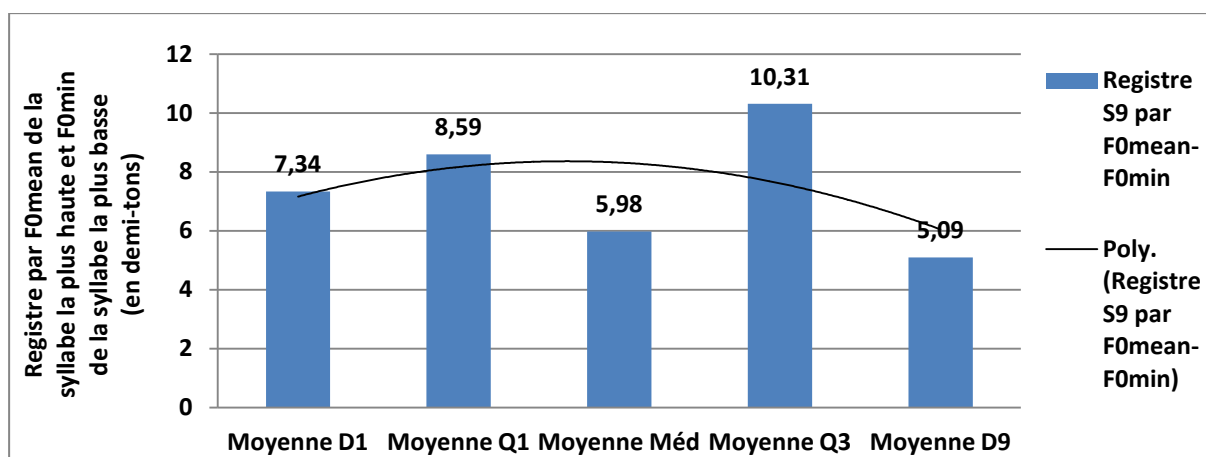


Figure 122 – Registre du syntagme intonatif S9 de la Phrase 9 par catégorie d'apprenants

5.3.5 Registre de l'emphase dans la syllabe /n@U/ de *nobody*

Contrairement aux autres mesures opérées sur phrases entières ou syntagmes intonatifs, ici, l'ampleur du registre porte sur la syllabe accentuée de ce mot, quatrième porteur d'emphase. Ce /n@U/ emphatique s'appuie sur la voyelle diphtonguée ainsi que sur la nasale en attaque. Ainsi, l'ensemble de la syllabe est caractérisé par une F0 exempte de micromélie, ce qui permet d'observer comme à la loupe le développement des réalisations de chaque groupe.

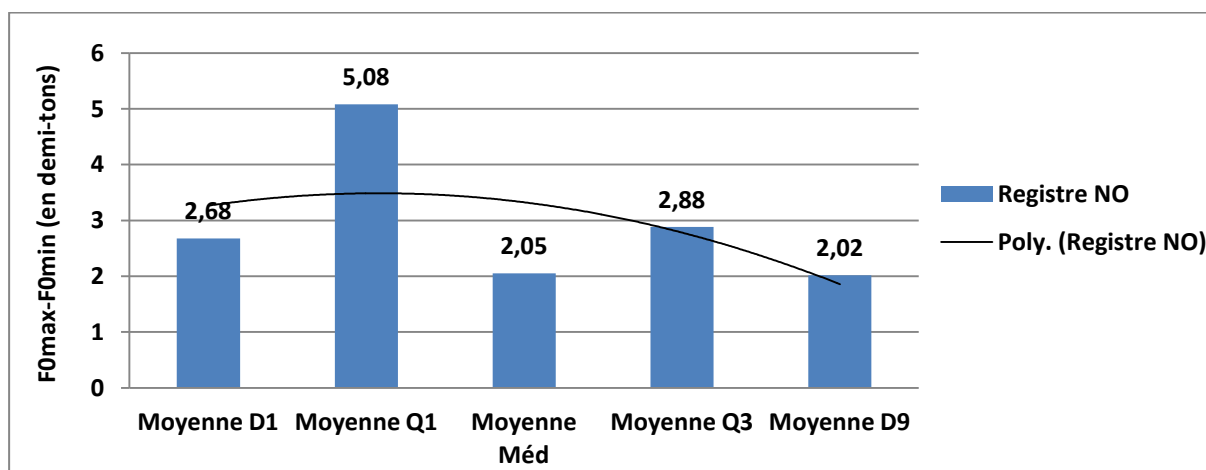


Figure 123 – Ampleur du registre de la syllabe emphatique /n@U/ par catégorie d'apprenants

On peut voir que ce graphique rappelle celui de l'exclamative dans sa tendance générale, même si la moyenne de D1 est plus haute. On constate en outre que les mesures de F0max de /n@U/ se rapprochent de la moyenne des quatre autres zones testées précédemment pour leur registre. Par contre, F0min ne peut atteindre les mesures des autres points emphatiques étant donné qu'il ne porte que sur une seule syllabe alors que les autres mesures de F0min sont prises sur des phrases entières.

5.3.6 Registre global du texte dans les zones étudiées

Le graphique suivant récapitule les mesures précédentes de l'emphase en écartant donc la Phrase 9 des calculs. Alors que la valeur maximale de F0 prend en compte l'exclamative, les paires *so* + *sweet* dans les Phrases 4 et 22, ainsi que la syllabe /n@U/ de *nobody* (Phrase 11), la valeur minimale F0min est calculée en fonction des mêmes éléments auxquels on ajoute la Phrase 5. Ainsi, REG_max détermine le registre maximal qu'un apprenant peut avoir sur l'ensemble de ces éléments alors que REG_moy est la moyenne des quatre F0max d'une part et des cinq F0min d'autre part.

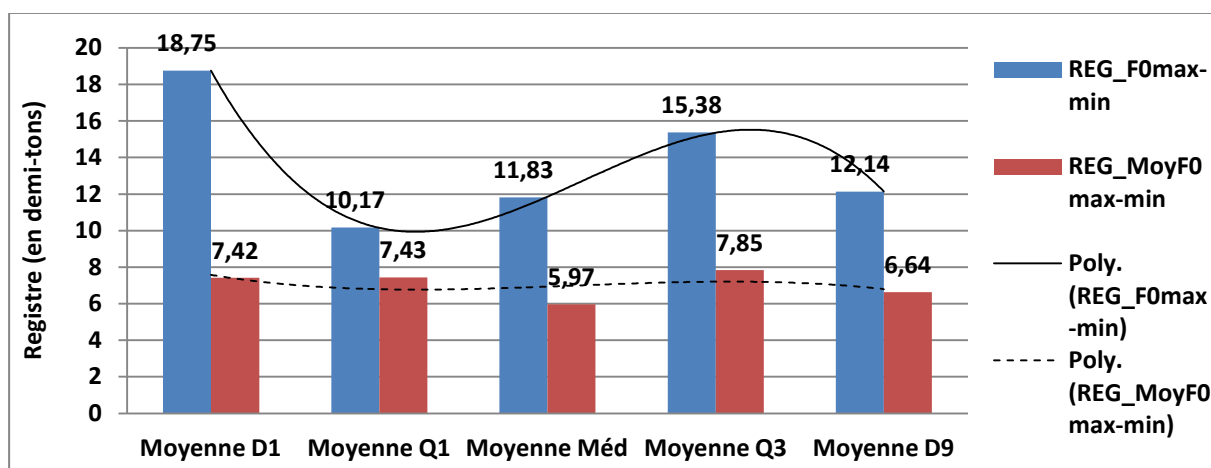


Figure 124 – Registre maximal des phrases étudiées et des moyennes de ces registres par catégorie d'apprenants

Le graphique montre que le registre moyen des F0max et F0min (REG_moy) est très homogène : seules les moyennes de la Médiane et de D9 sont légèrement plus faibles que les autres. Par contre, les limites de registre maximal privilégient D1 et dans une moindre mesure Q3, ce qui n'est en rien proportionnel à la rapidité de lecture.

5.3.7 Bilan

En tout premier lieu, on doit reconnaître la difficulté de pratiquer des mesures de la F0min, la voix craquée faussant parfois le jeu. Un deuxième biais à considérer est la difficulté d'établir des moyennes représentatives avec trois individus maximum par groupe car les différences individuelles sont parfois importantes. Ainsi, il est inattendu que, dans certains cas, l'ensemble des moyennes soit cohérent à l'exception de celle de la Médiane. Davantage de locuteurs aurait pu donner un éclairage constructif sur ce déficit.

Malgré ces biais, on remarque souvent que les profils par rapidité de lecture sont non distributifs, ce qui conduit à rechercher des régularités en utilisant un profilage basé sur d'autres catégories pour trouver des profils distributifs, utilisables dans notre quête de traits

distinctifs dans une optique évaluative. On remarque cependant que les catégories D1 et surtout D9 sont souvent dotées de l'ampleur de registre la plus faible.

5.4 Profil des apprenants dans leur qualité mélodique

Cette étude vérifie si les courbes mélodiques réalisées par les apprenants correspondent aux contours descendants attendus dans la Phrase 5 avec focalisation sur l'exclamative (section 5.4.1) et dans le syntagme intonatif terminal de la Phrase 9 (section 5.4.2). Les tons bidirectionnels sont analysés dans les syllabes accentuées des mots *arrived* (ton descendant-montant) de la Phrase 9 (section 5.4.3) et *nobody* (ton montant-descendant) de la Phrase 11 (section 5.4.4).

5.4.1 Chute mélodique dans syntagme intonatif emphatique (exclamative et Phrase 5)

Afin de mettre les particularités du contour en évidence, les repères les plus sensibles ont été mesurés de la manière la plus révélatrice. Le graphique ci-dessous précise la nature des mesures ainsi que leur localisation.

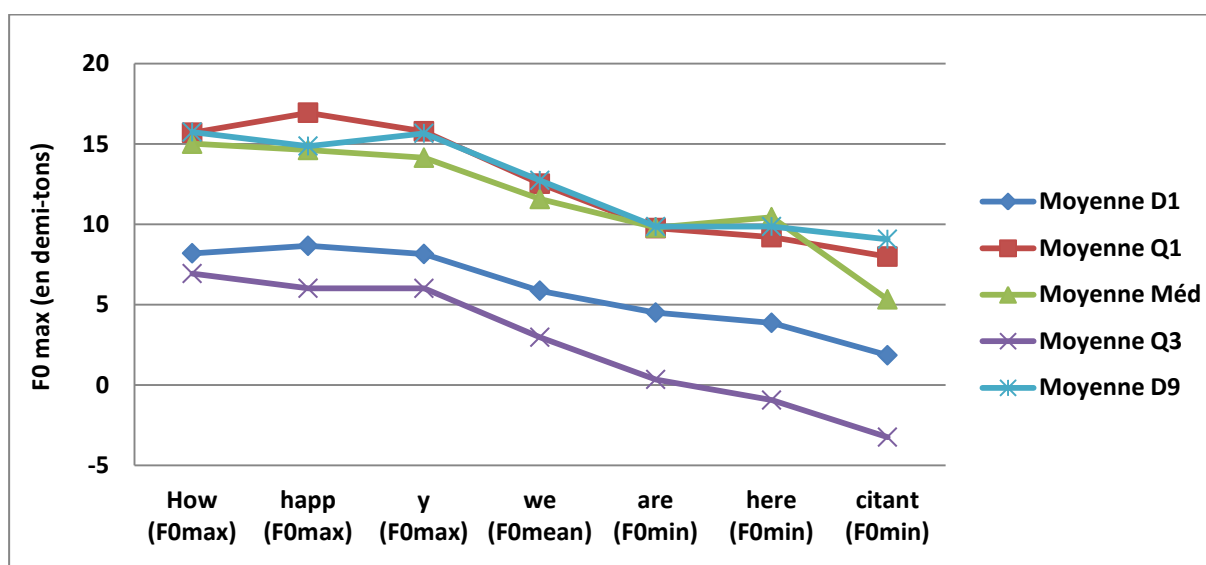


Figure 125 - Contour intonatif normalisé moyen par catégorie d'apprenants dans l'exclamative avec discours citant

On remarque que les contours moyens de Q1 et D9 se superposent presque à l'exception de la première syllabe de *happy*, Q1 étant davantage conforme au modèle natif. Le contour de D1 se rapproche de celui de Q1, bien que plus plat, et le contour de la Médiane rappelle celui de D1. Les meilleures réalisations par rapport à la norme des natifs sont celles de Q1 pour le contour et Q3 pour l'ampleur du registre. On observe en outre que D9 se caractérise par une montée sur la deuxième syllabe de *happy* et que Q3 s'y maintient rectiligne, alors que toutes les autres courbes sont descendantes sur cette syllabe.

5.4.2 Chute mélodique dans syntagme intonatif non emphatique (Phrase 9)

Ici encore, seules les syllabes pertinentes en termes de divergence attendue par rapport au contour canonique ont été mesurées acoustiquement. On remarque dans le graphique ci-dessous, une grande similarité avec les contours de l'exclamative.

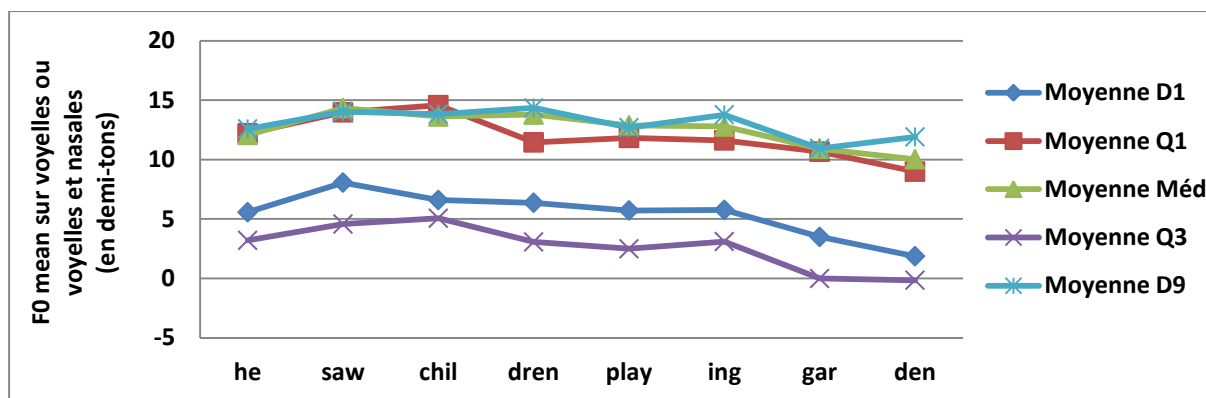


Figure 126 – Contour intonatif normalisé moyen par catégorie d'apprenants dans les syllabes pertinentes d'un syntagme intonatif dépourvu d'emphase

Toutes les courbes ont une avant-tête basse par rapport à la tête *saw*. Les écueils à éviter étaient une remontée en fin de syntagme intonatif comme le fait D9, et une remontée sur la syllabe inaccentuée de *children* et *playing*. D1 et Q1 évitent l'erreur, le contour de la Médiane est relativement plat et indécis, Q3 remonte sur la seconde syllabe de *playing* alors qu'il réalise *children* avec une baisse de F0. Le contour de D9 est le moins conforme dans la mesure où il cumule toutes les erreurs à ne pas commettre : registre très faible, remontée sur *children*, *playing* et *garden*. On peut ici affirmer que la catégorisation par la vitesse de lecture correspond à la qualité mélodique d'un syntagme intonatif descendant non emphatique.

5.4.3 Ton bidirectionnel sur syllabe : descente-montée (*arrived*)

Le mot *arrived* est bisyllabique et d'accentuation iambique. Afin d'éviter l'interférence des micromélodies du dernier phone /d/, je n'en ai pas tenu compte. Le graphique suivant répertorie quatre points de mesure : F0max de la syllabe atone, F0max du début de la séquence /raIv/, F0max en fin de séquence et F0min entre ces deux mesures de F0max. En fait, le phone /v/ n'intervient quasiment pas dans les résultats obtenus, lesquels figurent sous forme de graphique ci-dessous. Observons le ton descendant-montant sur la seule syllabe /raIvd/.

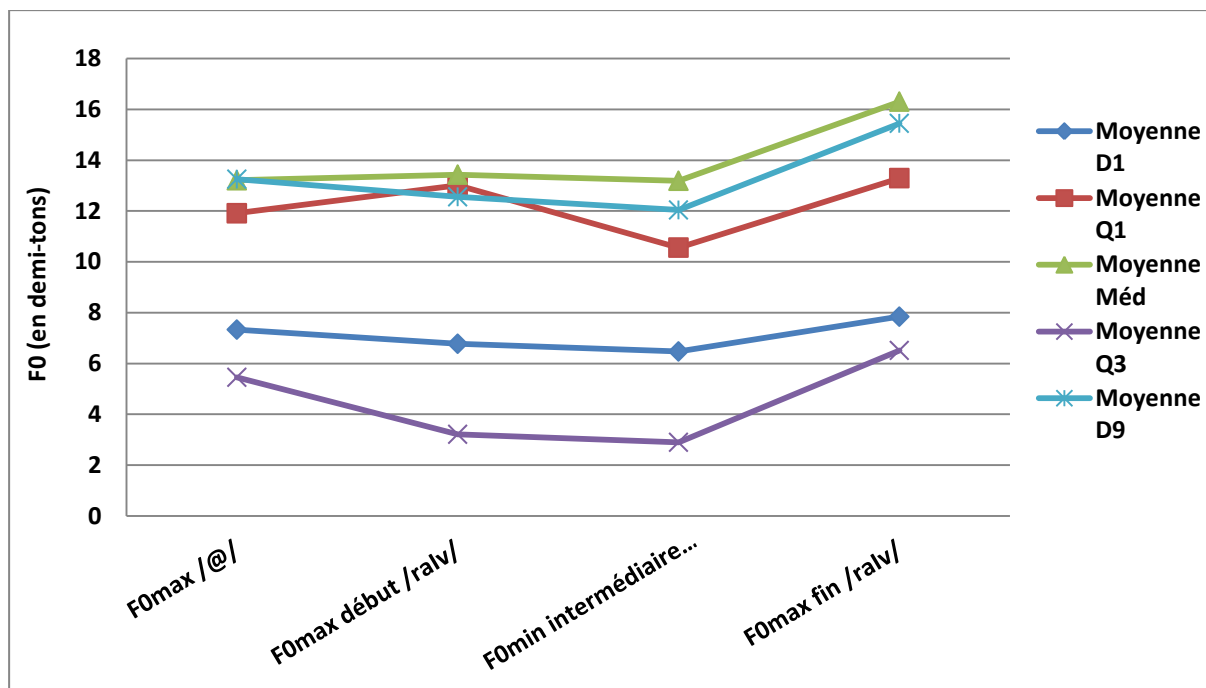


Figure 127 – Contour intonatif normalisé moyen du ton bidirectionnel descendant-montant dans la seconde syllabe du mot *arrived* par catégorie d'apprenants

On remarque que toutes les moyennes remontent en fin de parcours. Par contre, seule Q1 descend significativement sur le début de la syllabe. Les autres moyennes ne descendent qu'imperceptiblement. Q3 montre par la F0 de sa première syllabe que la descente effective s'est produite sur la première syllabe du mot alors que la deuxième porte la remontée, ce qui la place en seconde position qualitative après Q1. D1 révèle le ton descendant-montant le moins bien réussi par son registre trop étroit. Ces résultats montrent que la catégorisation par la rapidité de lecture n'est pas discriminante pour rendre compte de la qualité des réalisations du ton bidirectionnel descendant-montant.

5.4.4 Ton bidirectionnel sur syllabe : montée-descente (*nobody*)

Le ton bidirectionnel montant-descendant est attendu sur la syllabe accentuée du mot emphatique *nobody*. Quatre points de mesure de F0 ont permis de rendre compte des mouvements tonaux à l'œuvre dans cette seule syllabe aux endroits stratégiques (début et fin de syllabe), mais aussi sur la F0 extrême (F0min et F0max). Au besoin, le premier et second tiers ont été choisis arbitrairement pour répondre à la prise de mesures nécessaires à la construction du graphique car une absence de mesures biaisait considérablement la vue d'ensemble. J'ai pu ainsi constater que les réalisations tonales pouvaient être de complexité variable, comme le montre le graphique suivant.

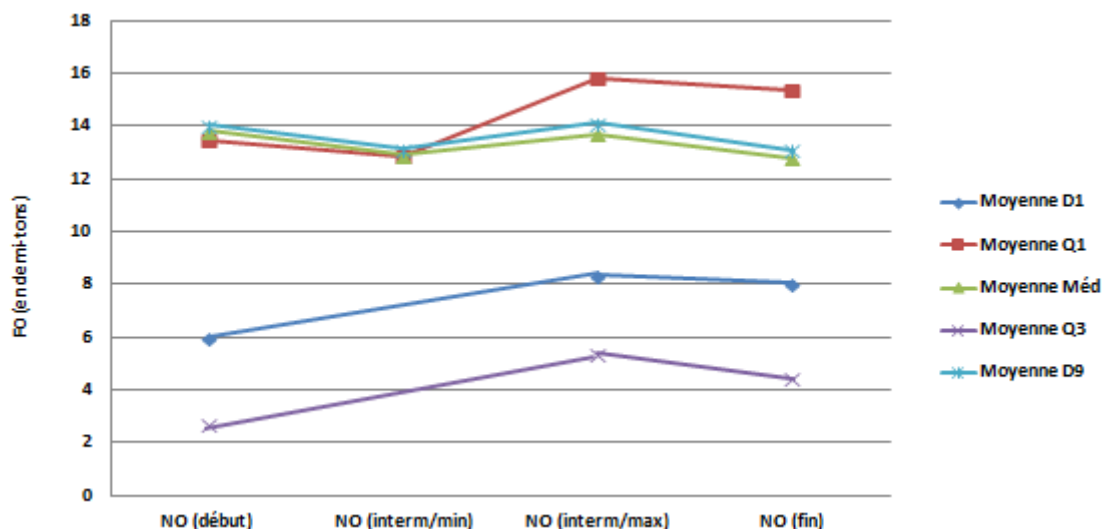


Figure 128 – Contour tonal normalisé moyen de la syllabe /n@U/ de *nobody* par catégorie d'apprenants (F0 en demi-tons)

Dans ce graphique, deux contours sont bidirectionnels et trois tri-directionnels. On constate d'ailleurs que la similitude des patrons tonaux des catégories D9 et Médiane, ainsi que de celles de D1 et Q3, prouvent que la catégorisation selon la vitesse de lecture ne correspond pas à celle qui est nécessaire en matière de mélodie pour mettre en évidence les disparités dans le choix d'utilisation de tel ou tel patron tonal.

En restituant la variation tonale sous forme d'histogramme, on peut davantage discerner l'implication de chaque mouvement mélodique et son ampleur.

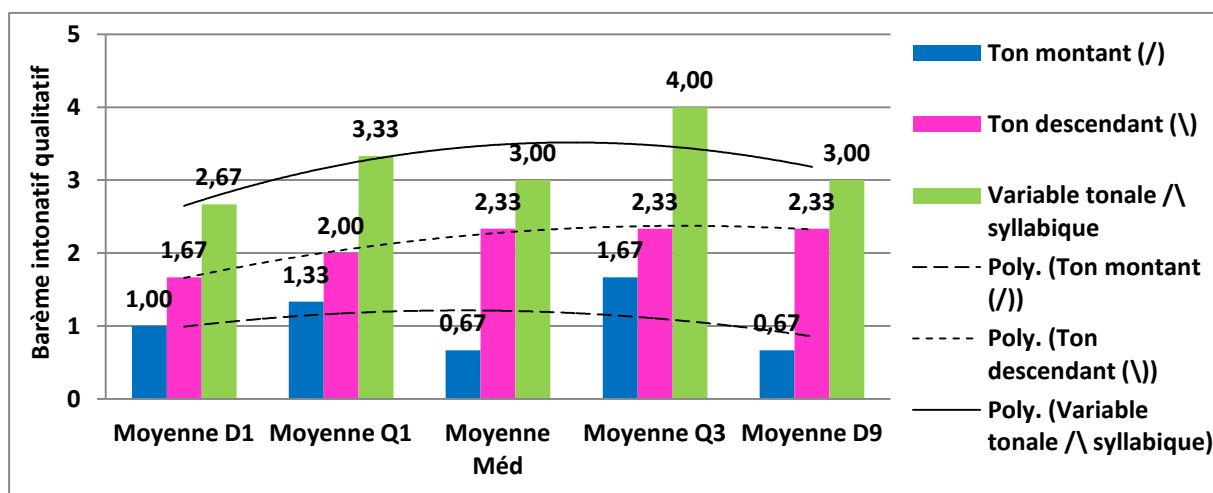


Figure 129 – Variable tonale syllabique Λ par catégories (barème qualitatif)

Le graphique ci-dessus confirme et précise la qualité des réalisations tonales de Q3 puis de Q1. La Médiane et D9 se caractérisent par une trop faible montée alors que D1 présente une répartition tonale équilibrée, mais de registre trop faible. On constate donc ici aussi que la

catégorisation des apprenants par la rapidité de lecture est incompatible avec la qualité mélodique obtenue.

5.4.5 Bilan

Cette étude de la mélodie tonale des apprenants révèle que la catégorisation par la vitesse de lecture ne met pas en évidence la complexité intonative, bien que la qualité de lecture de la Phrase 9 dans son syntagme intonatif à l'emphase minimale corresponde effectivement à la rapidité de lecture. Force est de constater malgré tout que ces profils de groupes ont de grandes chances de ne pas être distributifs. Il faudra vérifier en aval si et à quel point les profils personnalisés qui constituent le profil moyen de leur catégorie peuvent s'en éloigner et, le cas échéant, proposer une catégorisation plus adéquate pour faire ressortir les régularités qualitatives.

5.5 Profil des apprenants dans leur maîtrise du rythme

Cette étude du rythme des apprenants en lecture repose sur l'analyse d'éléments où la saillance se traduit acoustiquement chez les natifs par un décrochement de la F0 ou par une différence de durée. Lors de phénomènes acoustiques plus complexes, la recherche de simplicité et de rapidité recommande de recourir à des tests perceptifs, lesquels fournissent des résultats globaux et en conséquence plus fiables. Ainsi l'analyse du rythme a été effectuée sur la saillance par fréquence fondamentale des paires en « *so* + adjectif ou adverbe » (section 5.5.1) et « *all* + préposition » (section 5.5.2), sur la saillance perceptive des schémas accentuels lexicaux dans les polysyllabes (section 5.5.3) puis du pronom *them* avant frontière majeure (section 5.5.4) et elle se termine par la vérification de la saillance par la différence de durée de deux syllabes contiguës non emphatiques (section 5.5.5).

5.5.1 Saillance mélodique de repères rythmiques : « *so* + adjectif ou adverbe »

Dans chaque syllabe qui a servi de base aux courbes mélodiques, les mesures de F0 ont été prises sur les voyelles à l'exception de celles qui sont suivies de nasales, où la mesure concerne l'ensemble.

Alors que dans la Phrase 4 *so* est doté d'une F0 inférieure à la syllabe précédente, on remarque l'inverse dans la Phrase 22. La baisse sur *so* favorise une remontée sur la syllabe suivante, alors que la Phrase 22 montre que certains apprenants rehaussent la F0 sur *sweet* et que d'autres marquent une baisse mélodique. L'adverbe *sweetly* montre néanmoins la propension qu'ont les apprenants de faible niveau prosodique à prolonger la pente mélodique

ascendante sur la deuxième syllabe. J'é mets l'hypothèse que la différence de traitement mélodique sur *so* résulte de la présence ou de l'absence de proéminence sur la syllabe qui le précède, renforcée par l'importance de la proéminence émise sur l'adverbe *sweetly*. Dans la Phrase 22, *so* est précédé d'une syllabe atone, ce qui lui favorise l'attribution d'une proéminence, alors que la présence de l'élément lexical monosyllabique *sang* attire une proéminence sur lui-même, ainsi que l'adoption du schéma rythmique « fort-faible-fort-faible-fort » sur *trees and sang so sweet-*.

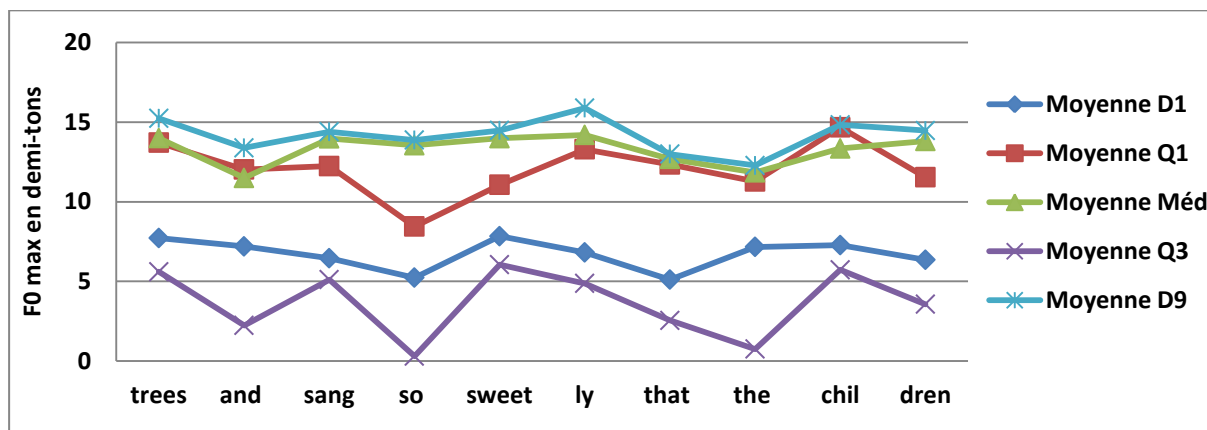


Figure 130 – Contour intonatif par catégorie du tronçon contenant *so sweetly* de la Phrase 4 (F0max en demi-tons)

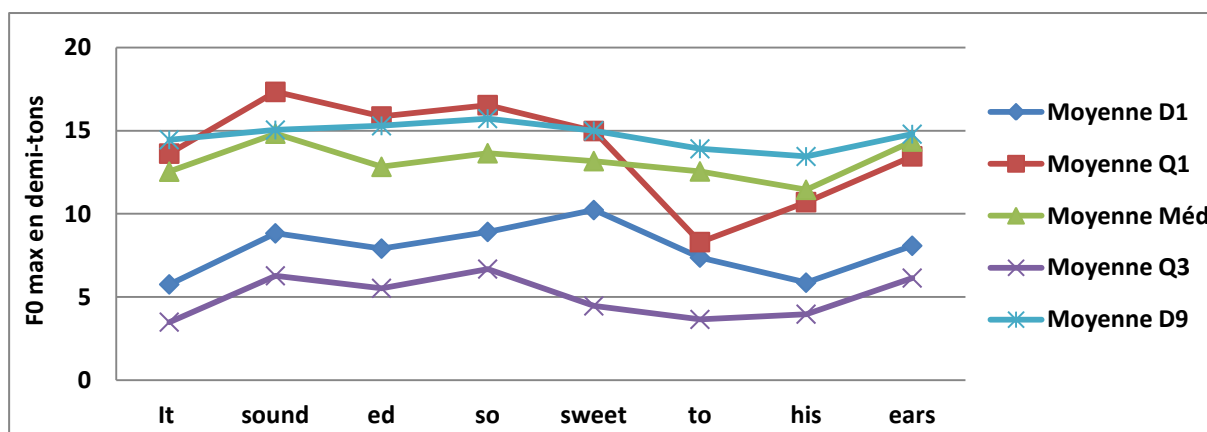


Figure 131 – Contour intonatif par catégorie du tronçon contenant *so sweet* de la Phrase 22 (F0max en demi-tons)

En ce qui concerne les profils par catégories basées sur la vitesse de lecture, le contour optimal de la Phrase 4 est effectué par Q3 et dans une moindre mesure D1, alors que dans la Phrase 22, la meilleure intonation est celle de Q1, suivi de Q3 puis de D1. Dans tous les cas, la distinction par la vitesse de lecture ne fournit pas de hiérarchisation qualitative cohérente.

5.5.2 Saillance mélodique de repères rythmiques : « *all* + préposition »

5. Analyse par profilage de groupes par archétypes de conformité selon la vitesse de lecture (expérimentation STAT-NN2)

CHAPITRE IV – Expérimentation de création de profils et de critères d'évaluation prosodiques d'apprenants

De même que précédemment, les mesures de F0max ont été opérées sur les voyelles ou les voyelles suivies de nasales (-round) à l'exception de *all* dont l'intégralité de la syllabe a été prise en compte (voyelle et liquide).

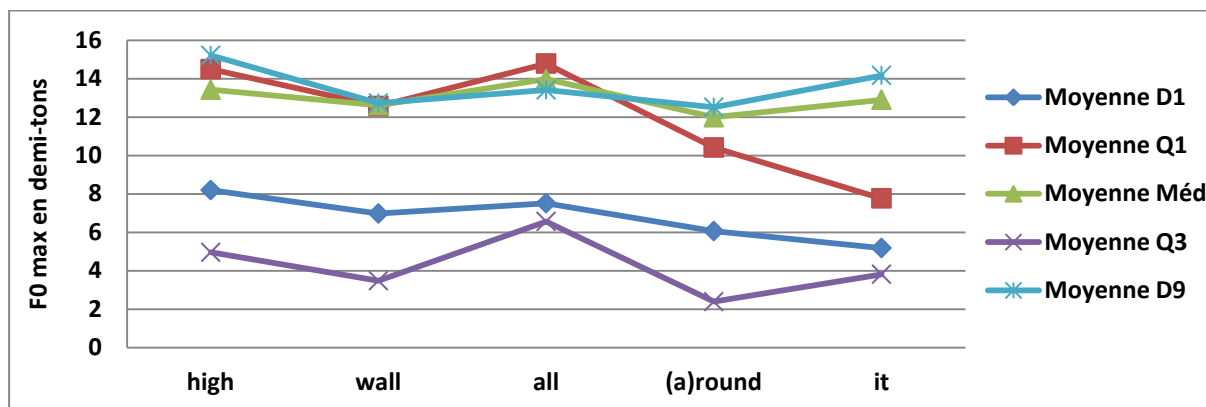


Figure 132 – Contour intonatif par catégorie du tronçon comprenant *all around* (Phrase 12) (F0max en demi-tons)

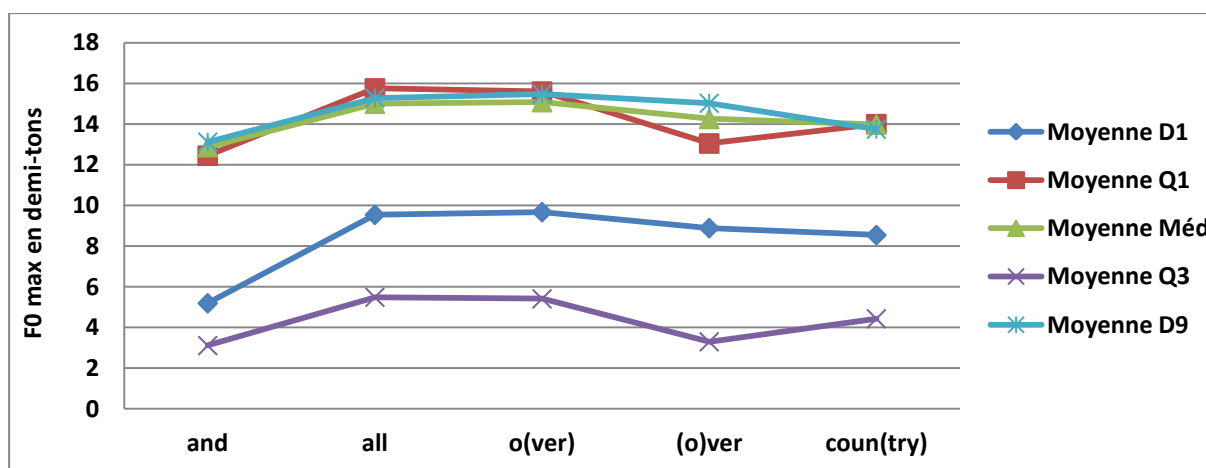


Figure 133 – Contour intonatif par catégorie du tronçon comprenant *all over* (Phrase 16) (F0max en demi-tons)

Dans les graphiques, on remarque le même phénomène que pour les expressions en *so*. Dans la Phrase 16 *all* est précédé de l'élément grammatical *and*, donc non proéminent et de F0 basse, qui résulte en une forte montée sur *all*. Par contre, la Phrase 12 montre *all* précédé de deux mots lexicaux monosyllabiques *high wall*, ce qui limite le potentiel des apprenants à émettre une forte proéminence mélodique sur *all*. La juxtaposition de *high* et *wall* favorise néanmoins le décrochage de F0 puisque la F0 chute sur *wall* après avoir été élevée sur *high*. On remarque en outre que la Phrase 12 favorise une chute mélodique sur la préposition *(a)round* alors que la Phrase 16 tend à maintenir une F0 similaire sur la syllabe accentuée de *over*.

Dans la Phrase 12, la qualité maximale est portée par le groupe Q3. Il en est de même dans la Phrase 16 lorsqu'on y ajoute Q1. Là encore, la catégorisation par la rapidité de lecture ne rend pas compte des différences mélodiques.

5.5.3 Saillance perceptive de patrons accentuels lexicaux dans les polysyllabes

On a vu plus haut la méthodologie utilisée pour quantifier la correspondance des schémas accentuels lexicaux des six polysyllabes étudiés. Le graphique suivant en présente les résultats.

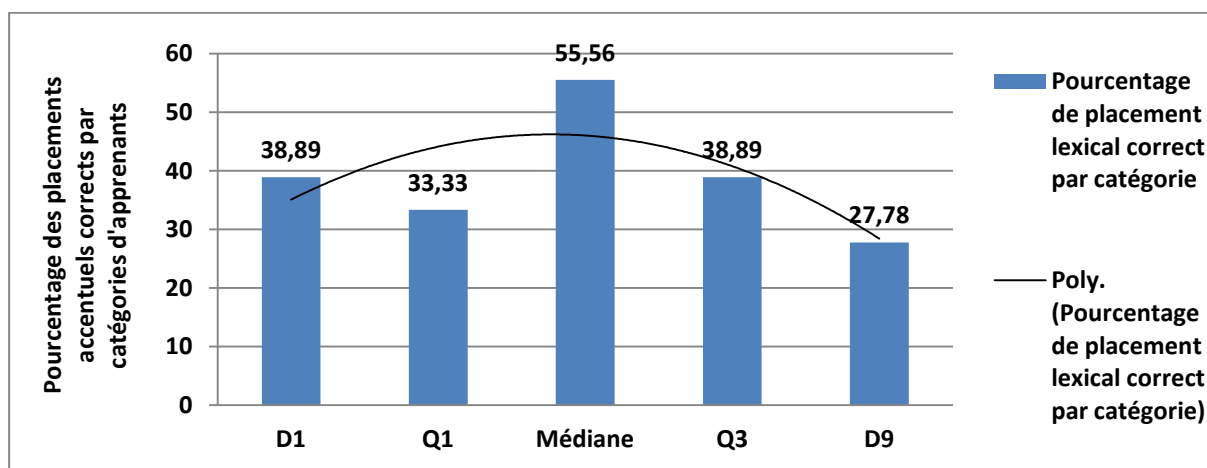


Figure 134 – Placement accentuel lexical correct par catégorie d'apprenants (en pourcentage)

On constate que ces résultats révèlent une relative concordance avec ceux trouvés dans le registre du syntagme intonatif S9, de l'exclamative et surtout de la Phrase 5 (exclamative + énoncé citant) étant donné que les résultats de la Médiane renforcent la régularité de la courbe polynomiale. Cette cohérence par rapport à ces quelques résultats précédents pourrait être relativement conforme à une catégorisation selon la vitesse de lecture dans la mesure où les résultats sont analysés finement.

5.5.4 Saillance mélodique et perceptive d'un pronom avant frontière majeure : *them*

Les résultats des mesures acoustiques de F0 ne concordant pas avec ceux pressentis auditivement, je m'en tiendrai aux résultats du test perceptif dont la méthodologie a été détaillée plus haut.

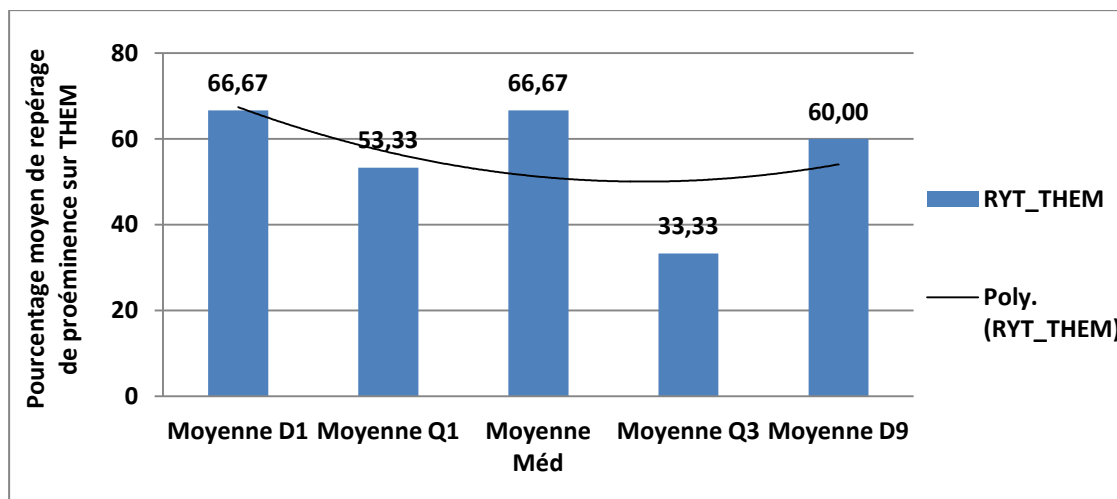


Figure 135 – Pourcentage moyen par catégories d'apprenants d'un repérage de prééminence sur *them* (en pourcentage, 0 = qualité maximale de la réalisation)

Ces résultats d'accentuation du pronom *them* en fin de syntagme intonatif ne semblent pas fournir de cohérence véritable à la catégorisation par la vitesse de lecture car Q1 et Q3 sont très faibles alors que D1, D9 et la Médiane fournissent des résultats élevés.

5.5.5 Saillance par la différence de durée de deux syllabes non emphatiques contiguës

Le graphique suivant montre les résultats en pourcentages de la différence de durée entre la syllabe normalement accentuée de *sat* et celle inaccentuée de *at*. Elle correspond exactement à la durée de /@t/ divisée par celle de /s{t/ multipliée par 100.

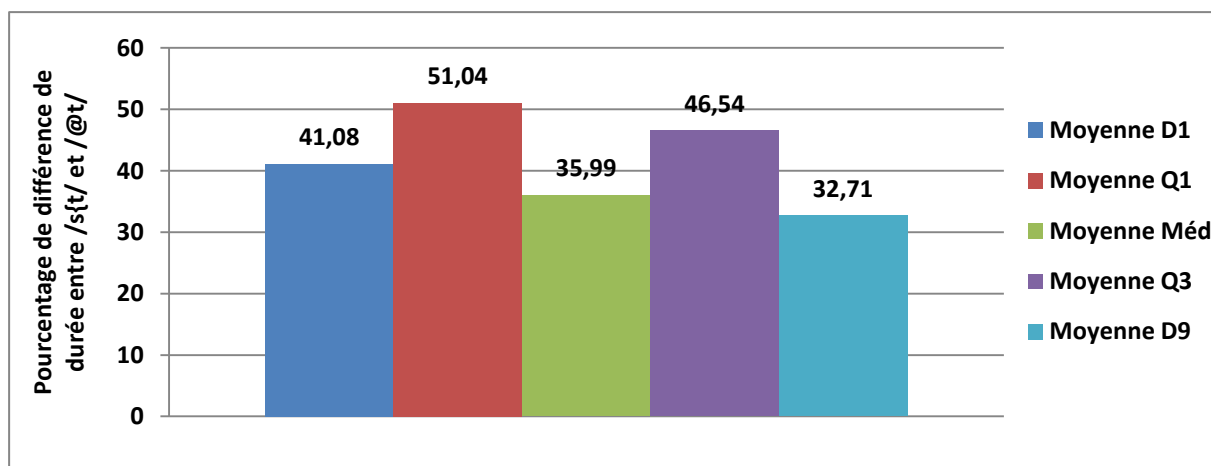


Figure 136 – Saillance par la différence de durée de deux syllabes non emphatiques contiguës (en pourcentage)

L'hypothèse selon laquelle la différence de durée (ici sous forme de pourcentage) équivaudrait à un rythme accentuel plus marqué n'est pas mise en évidence par la catégorisation selon la vitesse de lecture. Au contraire, il semblerait que ce soient les plus lents qui marquent le mieux la différence de durée.

Le graphique ci-dessous précise la durée moyenne des syllabes de chaque groupe. Si pour D1, la Médiane et D9 on constate un allongement de durée de la syllabe /s{t/, cela ne se vérifie pas pour Q1, ni Q3. Par ailleurs, on observe une durée croissante de la syllabe /@t/ chez D1, la Médiane et D9, mais ces distinctions sont insuffisantes pour valider la correspondance de ces mesures avec la rapidité de lecture alors même que l'on avait remarqué une relative concordance chez les natifs.

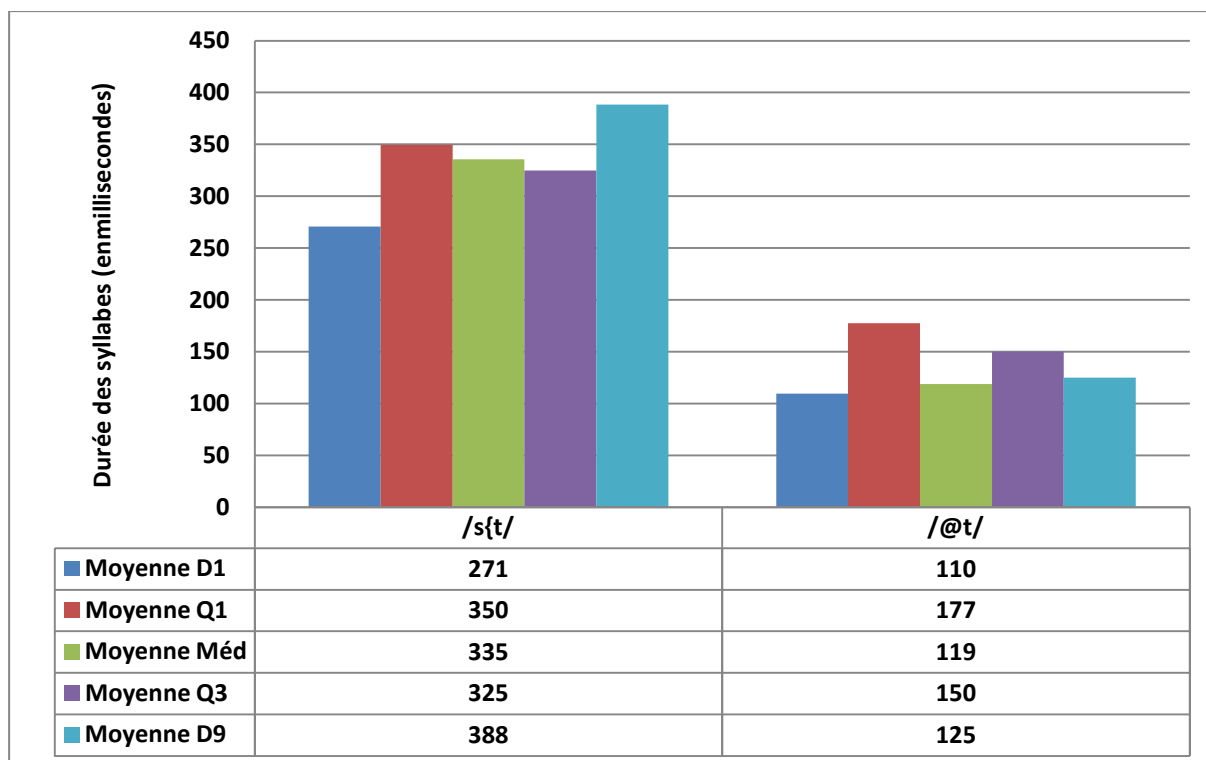


Figure 137 – Durée syllabique moyenne par catégorie des syllabes /s{t/ et /@t/ (en millisecondes)

5.5.6 Bilan

On a vu que la catégorisation des apprenants par la seule vitesse de lecture n'était effectivement fédératrice que pour les mesures liées à la temporalité (durée et nombre de syllabes par seconde). Les variables rythmique, mélodique ou de registre semblent obéir à des principes différents étant donné que les résultats tendant vers ceux observés chez les natifs se retrouvent le plus souvent dans les catégories Q3 et Q1 alors que ceux de la Médiane, D1 et D9 s'en éloignent.

Il faut néanmoins souligner que la catégorie D9 est celle dont les résultats s'éloignent le plus fréquemment de ceux des natifs. Les seuls domaines où D9 obtient des résultats moyens par rapport aux autres catégories sont l'absence de prééminence sur le pronom *them* en fin de syntagme intonatif et la qualité de la prééminence mesurée par la durée syllabique,

où D9 se révèle être le groupe aux résultats les meilleurs. Il s'agit là de quelques variables rythmiques. On peut néanmoins affirmer, à la vue de cette catégorisation par la vitesse de lecture, qu'appartenir à la catégorie la plus lente implique les réalisations mélodiques et de registre les plus éloignées de la norme native.

On peut aussi déduire de cette étude que la catégorie D1 n'obtient pas les meilleurs résultats dans le domaine du registre et de la mélodie ; ils sont parfois proches de ceux de D9 ou inférieurs, ce qui n'est pas très surprenant étant donné que les variations importantes de F0 nécessitent un surplus de temps pour être réalisées : c'est ce qui avait été observé dans l'étude des natifs. On constate néanmoins une aisance dans la production de l'intonation descendante du syntagme intonatif non emphatique S9.

Les résultats de la catégorie de la Médiane semblent à ce stade peu explicables étant donné l'importante différence qui les sépare de ceux des catégories adjacentes. Leur taux de placement lexical correct est le plus élevé et leur non accentuation de *them* aussi. Est-ce la méticulosité de ces détails techniques qui prive les apprenants de la Médiane d'un ressenti du texte favorisant un registre et une intonation plus suggestives ? Est-ce le faible nombre d'individus dans le groupe ? La présente investigation n'est pas assez poussée pour pouvoir répondre à ces questions.

Les différences observées peuvent être imputées au faible nombre d'individus étudiés dans chaque groupe, où un seul élément très typé pourrait faire basculer les moyennes vers ses réalisations, mais il semble plus qu'improbable que la vitesse d'élocution soit un critère suffisant pour évaluer la qualité prosodique. Il faudrait alors trouver la réponse à la question : est-ce qu'une étude par catégorisation des réalisations prosodiques selon la vitesse de lecture est validée par l'adéquation des profils personnalisés des individus qui les composent ? En d'autres termes, le faisceau de variables de chaque individu indique-t-il que les profils de groupes sont distributifs (caractéristiques partagées par tous les membres du groupe) ou non-distributifs (simplement indicatifs de l'appartenance au groupe puisqu'ils ne possèdent pas obligatoirement cette caractéristique) ? Jusqu'à présent, on a pu constater à la vue des résultats que seules les variables temporelles ainsi que l'intonation du syntagme intonatif non emphatique correspondaient à la hiérarchisation par la rapidité de lecture et étaient distributives. L'analyse d'un profilage personnalisé combiné à un profilage de groupe par archétypes stéréotypiques devrait pouvoir contribuer à répondre à cette question.

6 Analyse par profilage de groupes par archétypes stéréotypiques (expérimentation STAT-NN3)

Alors que la section précédente sur les archétypes de conformité a contribué à classer les réalisations des apprenants dans des catégories basées sur la durée de lecture, celle-ci ambitionne au contraire de repérer les tendances divergentes par leurs caractéristiques les plus extrêmes. Après avoir précisé les objectifs et les enjeux de ce profilage de divergence (section 6.1), une analyse par plan factoriel permet de visualiser l'incidence des variables pour établir les profils (section 6.2), suivie d'une confirmation des résultats par dendrogramme (section 6.3). Une analyse par taux de réussite et d'erreur pour chacune des 28 variables explique les profils dégagés (section 6.4), puis une analyse par cercle de corrélations précise la portée de chaque variable dans les quatre grands domaines prosodiques étudiés (section 6.5).

6.1 Objectifs et enjeux

On a vu dans le Chapitre II qu'un archétype stéréotypique, ou de divergence, se basait sur les rapports que les profils personnalisés entretenaient entre eux via les points de convergence ou divergence entre leurs variables. Ils établissent la borne des possibles dans un corpus donné en soulignant une tendance extrême, une limite à atteindre, ou au contraire à éviter, et en cela permettent de hiérarchiser l'importance respective des traits et variables définitoires.

Dans cette optique, ce profilage stéréotypique utilise plans factoriels, dendrogrammes et cercles de corrélation pour hiérarchiser les profils personnalisés de l'étude. Les résultats obtenus sont aussi confrontés à ceux fournis par le tableur *Excel*.

6.2 Analyse des données par plan factoriel

L'avantage du plan factoriel utilisé ainsi est de faire intervenir peu d'individus et d'obtenir des données plus extrêmes que si elles provenaient de moyennes selon l'« archétype de conformité ». Deux profilages sont envisagés ; ils se basent sur les variables constituées à partir des mesures exposées précédemment (voir Tableau Sélection des variables pour non natifs conformément aux Chapitres II et III de la section 4.1, et ajout de variables aux sections 4.2 et 4.3 de ce chapitre), récapitulées dans le tableau suivant (section 6.2.1). Le premier profilage prend en compte l'intégralité des 28 variables (section 6.2.2), alors que le second se concentre sur les 25 variables de base, écartant celles élaborées à partir de moyennes ou de données prélevées des variables de base (section 6.2.3).

6.2.1 Présentation des 28 variables pour analyse prosodique des réalisations d'apprenants

Le tableau ci-dessous détermine les 28 variables utilisées dans le profilage des apprenants. Elles sont regroupées par domaine prosodique. Les codes facilitent une visualisation des données dans les graphiques et figures. À l'exception des variables produites à partir de tests perceptifs (RYT_lex et RYT_them), elles ont toutes été construites à partir de données acoustiques. Les variables portant sur la temporalité et le registre utilisent les mesures acoustiques brutes, alors que celles portant sur la mélodie et le rythme utilisent les corrélats acoustiques de F0 pour les agencer qualitativement et obtenir par formules conditionnelles dans le tableur *Excel* des résultats qualitatifs numériques qu'il est possible de hiérarchiser et d'insérer dans des études statistiques. Afin d'éviter des redites et d'accéder à plus de cohérence dans les explications, ce profilage qualitatif est détaillé en fin de chapitre, dans la rubrique sur les traits d'évaluation distinctifs.

L'intégralité de ces 28 variables se répartit en six variables temporelles, huit variables de registre, six variables mélodiques, sept variables rythmiques et une variable hybride portant à la fois sur un aspect rythmique et un aspect mélodique. Ces variables étant sous forme numérique, il a été possible de caractériser chacune d'entre elles qualitativement en calibrant des catégories pour regrouper les résultats relativement proches et les distinguer de ceux semblant plus lointains (voir détails en Annexe sur fichier « Mesures_apprenants10.xlsx »). Ce sont ces résultats qualitatifs, dérivés des variables numériques (quantitatives ou qualitatives), qui permettent la représentation par plans factoriels.

Domaine	Nom	Code	Caractéristiques
Temporalité	DUR_T	D1	Durée brute du texte <i>The Selfish Giant</i>
	DUR_ABCp	D2	Durée des 3 segments non ponctués ABC avec pauses (P4, 16 et 25)
	DUR_ABC	D3	Durée des 3 segments non ponctués ABC sans pause (P4, 16 et 25)
	DEB_T	D4	Nombre de syllabes par seconde dans le texte <i>The Selfish Giant</i> avec repère moyen de 541 syllabes
	DEB_ABCp	D5	Nombre de syllabes par seconde dans 3 segments non ponctués ABC avec pauses (P4, 16 et 25)
	DEB_ABC	D6	Nombre de syllabes par seconde dans 3 segments non ponctués ABC sans pause (P4, 16 et 25)
Registre	REG_Excl	Re1	Ampleur du registre de l'exclamative (F0max-F0min)
	REG_P5	Re2	Ampleur du registre de P5 (F0max-F0min)
	REG_so4	Re3	Ampleur du registre de P4 (F0max sur <i>so sweet(ly)</i> -F0min P4)
	REG_so22	Re4	Ampleur du registre de P22 (F0max sur <i>so sweet</i> -F0min P22)
	REG_max	Re5	Ampleur maximale des registres Re1-2-3-4-8
	REG_moy	Re6	Moyenne des mesures de registre de l'emphase (Re1-2-3-4-8)
	REG_S9	Re7	Ampleur du registre de S9 (F0mean sur voyelle à hauteur maximale-F0min S9)
	REG_no	Re8	Ampleur du registre de /n@U/ dans <i>nobody</i> P11(F0max-F0min)
Mélodie	MEL_Excl	M1	Qualité mélodique de l'exclamative
	MEL_P5	M2	Qualité mélodique de la Phrase 5
	MEL_S9	M3	Qualité mélodique du syntagme intonatif S9
	MEL_s\	M4	Qualité du ton montant-descendant sur une syllabe
	MEL_s/	M5	Qualité du ton descendant-montant sur une syllabe
	MEL_no	M6	Qualité et complexité tonale réalisée sur la syllabe /n@U/ de <i>nobody</i> (P11)
Rythme	RYT_lex	Ry1	Respect des patrons accentuels canoniques dans les polysyllabes
	RYT_so4	Ry2	Qualité de proéminence par F0 du marqueur <i>so sweet(ly)</i>
	RYT_so22	Ry3	Qualité de proéminence par F0 du marqueur <i>so sweet</i> par F0
	RYT_all12	Ry4	Qualité de proéminence par F0 du marqueur <i>all round</i> par F0
	RYT_all16	Ry5	Qualité de proéminence par F0 du marqueur <i>all over</i> par F0
	RYT_them	Ry6	Absence de proéminence d'un pronom en fin de syntagme intonatif par test perceptif
	RYT_syll	Ry7	Rapport de durée inter-syllabique sur la paire de syllabes adjacentes <i>sat at</i> (P19)
Rythme et mélodie	RYTMEL_so4	RM	Qualité de proéminence par F0 du marqueur <i>so sweet(ly)</i> et de la hauteur mélodique de (<i>sweet</i>) <i>ly</i>

Tableau 73 – Variables utilisées pour l'établissement des profils des apprenants, regroupées par domaine prosodique

Suivant le nombre de variables retenues, il est possible de dresser deux profils prosodiques de la lecture de chaque apprenant et ce que chacun apporte à l'analyse. Voyons tout d'abord les résultats produits avec les 28 variables, puis ceux obtenus avec les 25 variables les plus ciblées.

6.2.2 Profils d'apprenants stéréotypiques avec 28 variables prosodiques

La méthode la plus globale est élaborée à partir de l'intégralité des mesures exposées précédemment : c'est-à-dire un total de 28 variables quantitatives ayant chacune le même poids par rapport aux autres. En insérant ces 28 variables correspondant aux diverses facettes de la prosodie, on obtient le plan factoriel par Analyse en Composantes Principales (ACP) ci-dessous.

Ce type d'analyse concerne des données qualitatives et est ordinairement appliqué à un grand nombre d'individus, surtout lorsqu'il y a de nombreuses modalités dans ces variables (catégories à l'intérieur d'une variable). Il permet cependant de repérer des tendances intéressantes.

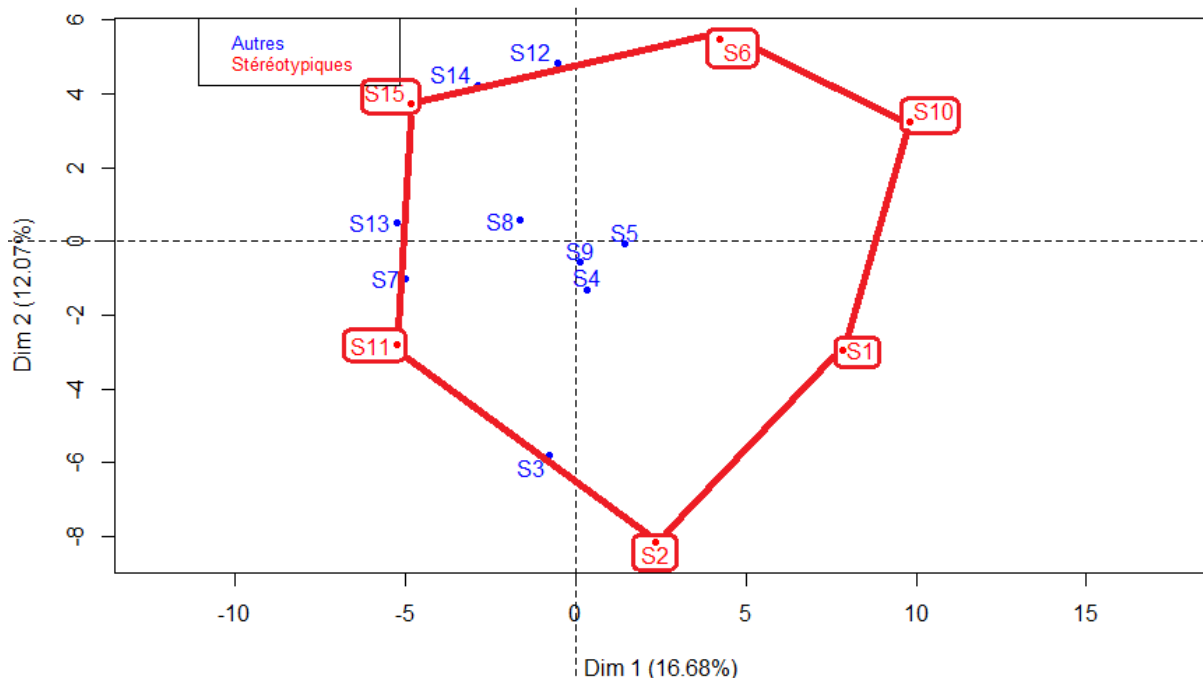


Figure 138 – Plan factoriel des quinze apprenants avec 28 variables (ACP)

Ce plan factoriel montre clairement la position des apprenants les uns par rapport aux autres dans les liaisons que les 28 variables entretiennent entre elles dans les deux principales dimensions. Les deux axes sont déterminés par une multitude de variables comme le montre leur dimension maximale : dimension 1 réduite à 16,68% et dimension 2 à 12,17%. La

faiblesse de ces pourcentages s'explique par la volonté de rechercher par les variables utilisées les plus grandes différences dans les divers domaines prosodiques. Le faible pourcentage de représentativité des deux premières dimensions indique que l'objectif est susceptible d'avoir été atteint. Six apprenants se distinguent par leur position excentrée alors que les autres occupent une place centrale, proche du barycentre signalant l'inertie du nuage de points, marquant l'aspect « moyen » de leurs réalisations, ou bien montrent des tendances moyennes par rapport aux profils « extrêmes » déterminés par les apprenants excentrés. De même, les apprenants 1 et 6 du plan semblent partager des points communs avec les profils adjacents. Cela réduit à quatre le nombre des profils aux particularités très spéciales : le 2, le 10, le 11 et le 15. Le plan factoriel suivant situe tous les apprenants parmi les caractéristiques des données prosodiques les définissant.

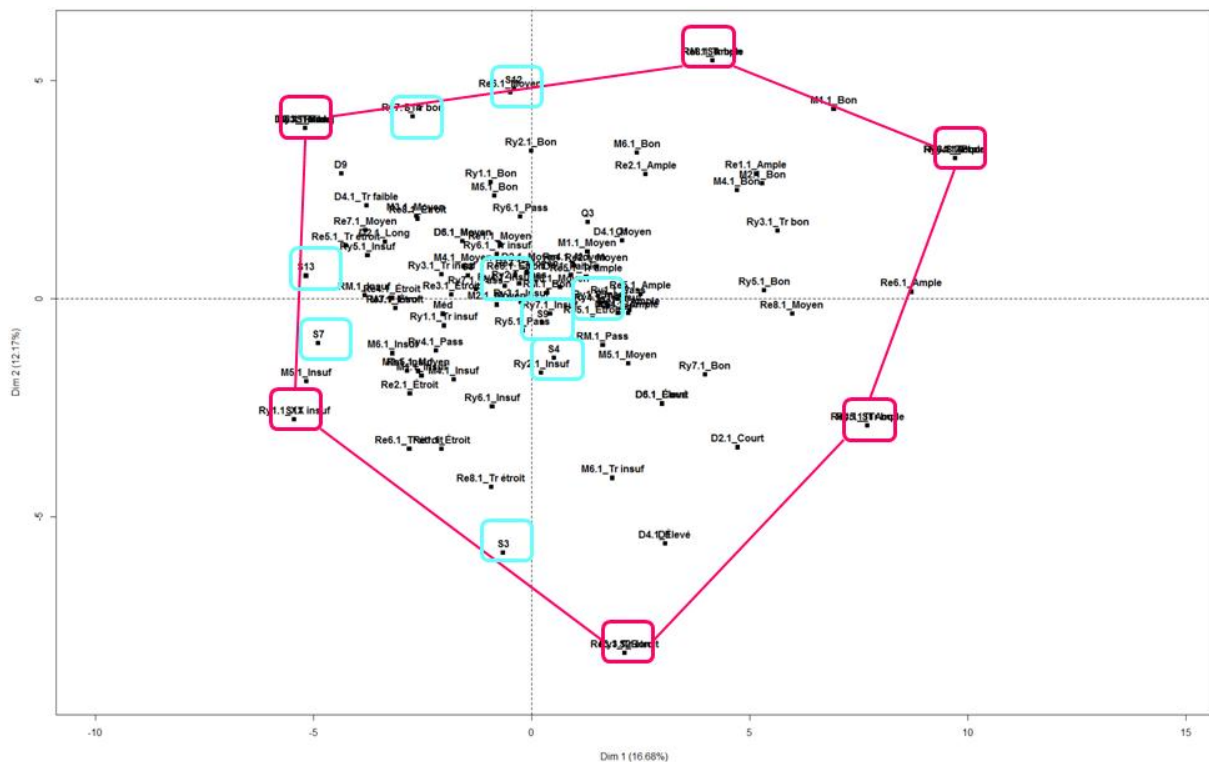


Figure 139 – Plan factoriel des quinze apprenants avec les 28 variables et leurs données

On peut y voir que le quart supérieur droit, où se situent les apprenants 6 et 10, ne contient que des appréciations positives telles qu'« ample », « bon » et « très bon » lorsqu'on s'éloigne du centre. Le quart inférieur gauche, où se situe l'apprenant 11, offre les mentions « insuffisant », « étroit », « très étroit ». À partir de ces deux extrêmes qualitatifs, le quart supérieur gauche (apprenant 15), bien que caractérisé par la mention « très faible », semble plus modulé dans les autres appréciations. Le quart inférieur droit (apprenants 1 et 2) est lui

aussi hybride dans la qualité de ses appréciations : « court », « élevé », « très insuffisant », « bon ».

Ces diverses constatations fournissent un éclairage qu'une réduction des variables pourrait confirmer ou infirmer.

6.2.3 Profils d'apprenants stéréotypiques avec 25 variables prosodiques

Le second profilage réduit le nombre de ses variables à 25 pour gagner en précision en attribuant un poids similaire aux phénomènes étudiés. Les trois variables écartées s'appuient sur des points vérifiés par d'autres variables telles que REG_moy (moyenne des registres de l'emphase), REG_max (extrêmes trouvés dans les registres) et RYTMEL_so4 (données de la variable RYT_so4 auxquelles ont été ajoutées celles de remontée mélodique sur syllabe inaccentuée, déjà partie intégrante de la variable MEL_P9). Le nombre de variables par domaine d'investigation (temporalité, registre, mélodie et rythme) est ainsi plus homogène : 6/6/6/7 au lieu de 6/8/6/7+1. On obtient alors le plan ci-dessous.

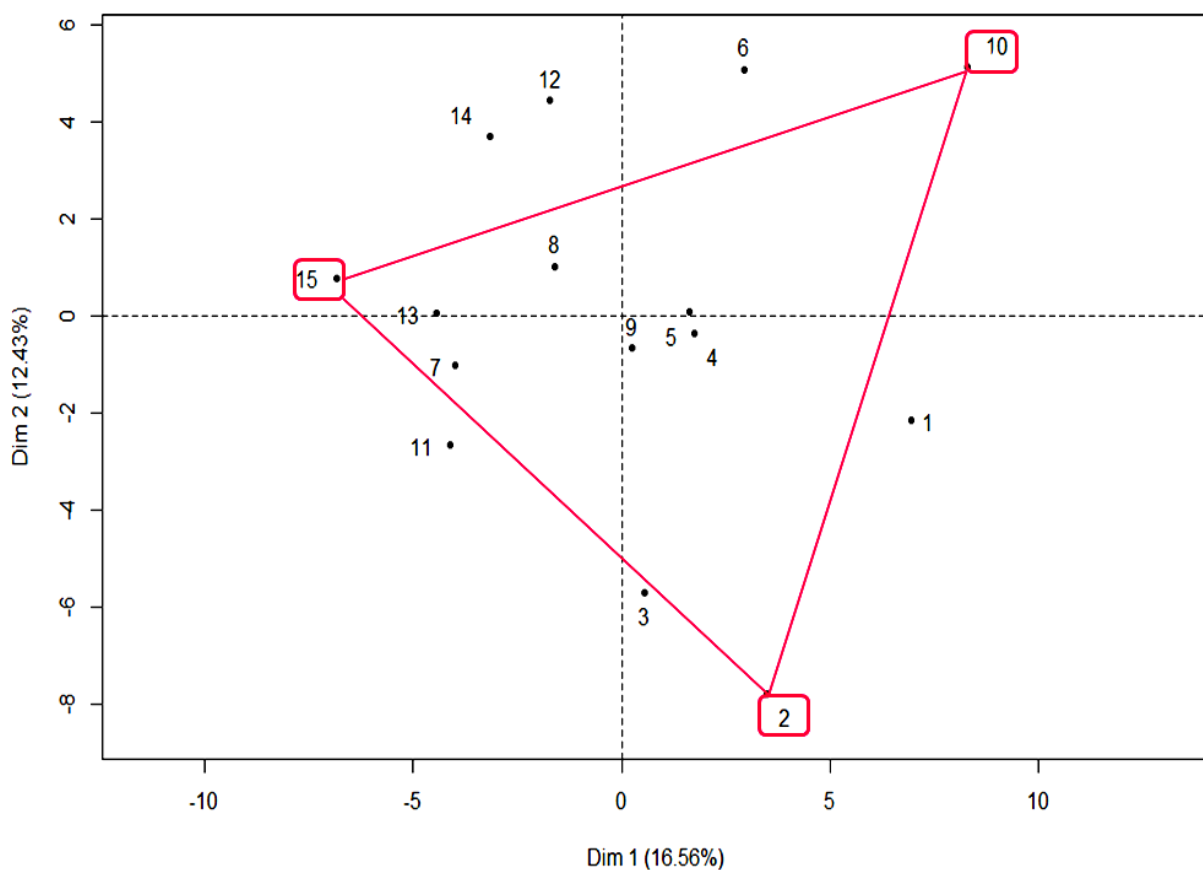


Figure 140 – Plan factoriel des quinze apprenants à partir de 25 variables

On peut y constater que les profils des apprenants 2 et 10 sont confirmés, mais que le choix se porte sur l'apprenant 15 aux dépens du 11.

6.3 Analyse des données par dendrogramme avec 28 variables

Il est opportun de vérifier si les profils obtenus par plan factoriel sont confirmés par dendrogrammes. Comme dans le profilage des natifs, diverses variantes de dendrogrammes ont été testées (Average, Ward, Single, Complete, McQuitty, Median, Centroid et Centroid2). À part celle de Ward, toutes ont placé l'apprenant 10 à l'extrême gauche du dendrogramme. L'intégralité des dendrogrammes a aussi étroitement relié les apprenants 1 et 2. Par contre, aucun ne rejoint étroitement le 11 et le 15. Le dendrogramme de Ward les éloigne même carrément, en les présentant dans deux branches s'éloignant à la base du tronc, bien que proches par leur ramure : le profil S11 est mis à l'écart des autres profils, ce qui contribuerait à montrer sa nature extrême. On observe le même phénomène en n'analysant que 25 variables.

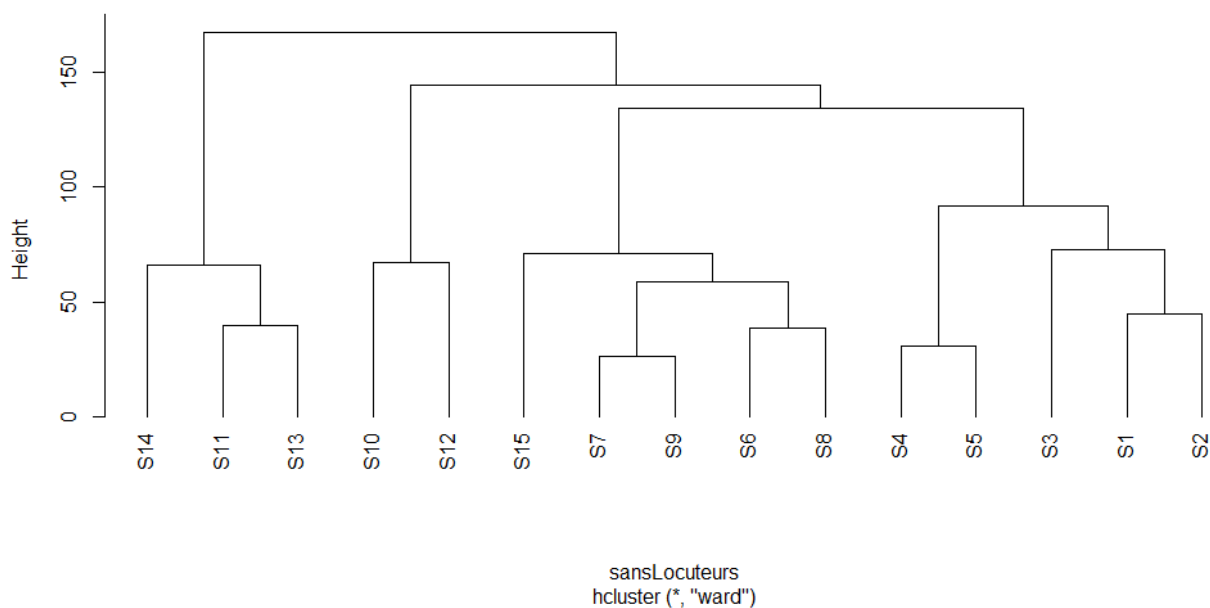


Figure 141 – Dendrogramme de Ward pour les 28 variables des quinze apprenants

Le dendrogramme de McQuitty rejoint celui de Ward en de nombreux points, mais présente ses données avec une particularité intéressante : il semble ordonner les profils des six apprenants selon leur qualité prosodique en fournissant une échelle qualitative allant de l'extrême S10 (meilleur apprenant ?) à l'extrême S11 (le moins avancé ?) en passant par des seuils plus ou moins intermédiaires :

S10 > S1 > S2 > S6 > S15 > S11

Figure 142 – Hiérarchisation de la qualité prosodique des six profils d'apprenants selon le dendrogramme de McQuitty

Cette présentation des réalisations des apprenants par le dendrogramme de McQuitty aurait l'avantage, si confirmée, de représenter linéairement et donc hiérarchiquement les résultats fournis en deux dimensions par le plan factoriel vu précédemment.

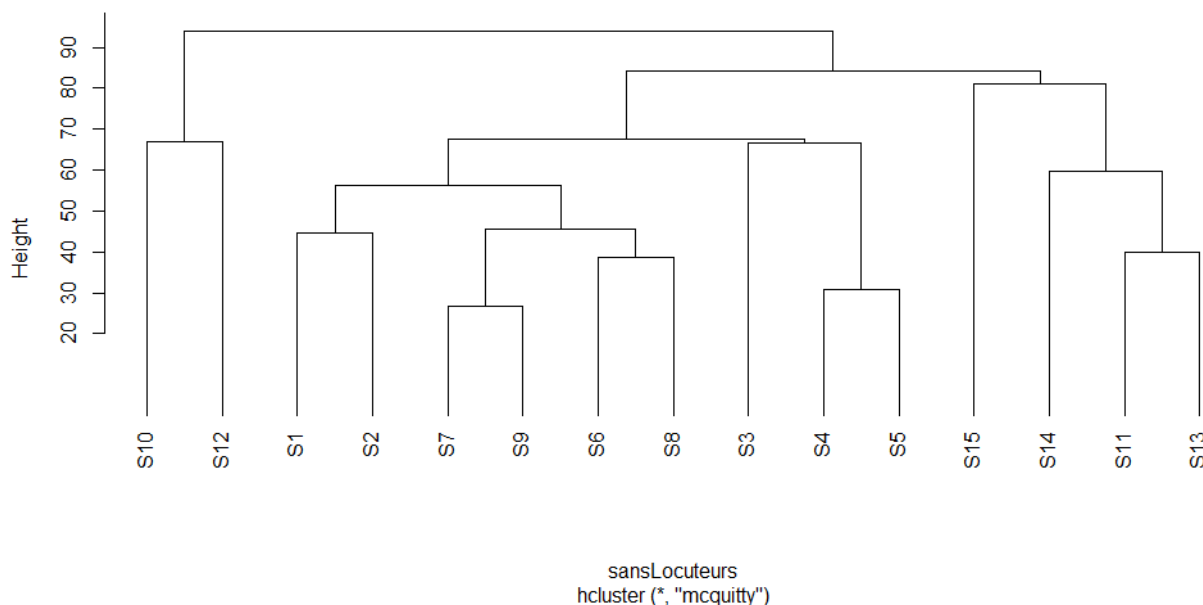


Figure 143 – Dendrogramme de McQuitty pour les 28 variables des quinze apprenants

6.4 Analyse des données par taux de réussite et d'erreur dans chacune des 28 variables

Cette section explique comment j'ai converti les résultats des apprenants, munie de leur classement, en pondération des scores obtenus sur les variables étudiées. J'illustre d'abord ma démarche avec les variables DUR_T et RYT_lex, puis je propose une première analyse avec les moyennes obtenues pour les grandes familles de variables. L'ensemble des procédures de bonification et de pénalisation est détaillé dans la section 7 de ce chapitre.

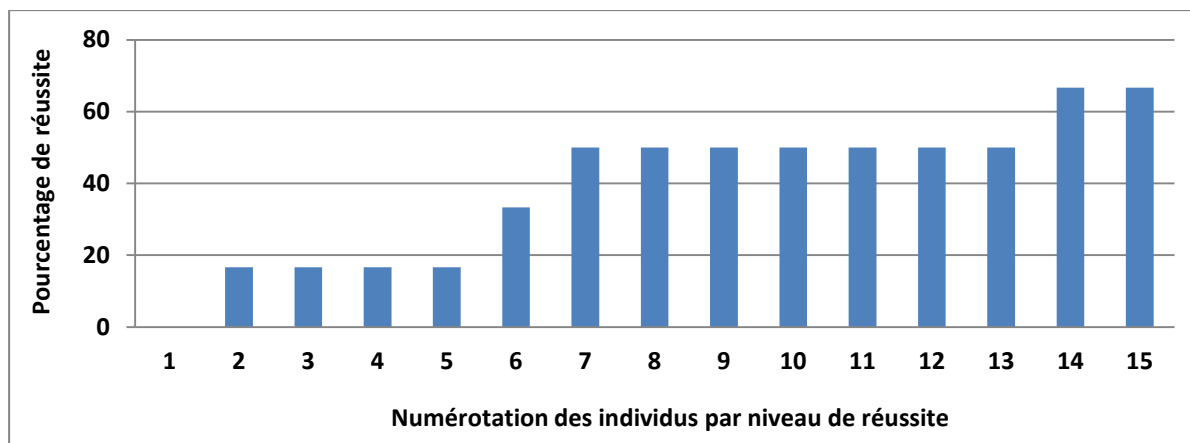


Figure 144 – Répartition de quatre graduations qualitatives (20% chacune) de la variable RYT_lex à partir des résultats numériques en pourcentage de réussite des 15 apprenants

Une étiquette qualitative a été attribuée à chaque catégorie. Les meilleures étiquettes de chaque domaine ont été mises en surbrillance avec la fonction *Excel* grâce à la mise en forme conditionnelle avec, par exemple, repérage du texte « bon » pour surbrillance en vert, et « insuf » pour surbrillance en rouge. Le résultat pour la durée du texte en secondes (DUR_T) et le pourcentage d’accentuation lexicale correctement effectuée (RYT_lex) est le suivant :

		DUR_T	DUR_T
		D1	D1
DID0210-S001R	D1	142,285	D1
DID0078-S001R	D1	143,036	D1
DID0018-S001R	D1	144,825	D1
DID0003-S001R	Q1	154,342	Q1
DID0024-S001R	Q1	154,288	Q1
DID0108-S001R	Q1	156,741	Q1
DID0062-S001R	Médiane	167,654	Méd
DID0135-S001R	Médiane	167,083	Méd
DID0046-S001R	Médiane	168,4	Méd
DID0168-S001R	Q3	175,145	Q3
DID0074-S001R	Q3	179,724	Q3
DID0068-S001R	Q3	181,817	Q3
DID0056-S001R	D9	191,151	D9
DID0153-S001R	D9	190,841	D9
DID0076-S001R	D9	196,717	D9

		RYT_lex	RYT_lex
DID0210-S001R	D1	50	Pass
DID0078-S001R	D1	50	Pass
DID0018-S001R	D1	16,66667	Tr insuf
DID0003-S001R	Q1	16,66667	Tr insuf
DID0024-S001R	Q1	33,33333	Insuf
DID0108-S001R	Q1	50	Pass
DID0062-S001R	Médiane	50	Pass
DID0135-S001R	Médiane	66,66667	Bon
DID0046-S001R	Médiane	50	Pass
DID0168-S001R	Q3	50	Pass
DID0074-S001R	Q3	0	Tr insuf
DID0068-S001R	Q3	66,66667	Bon
DID0056-S001R	D9	16,66667	Tr insuf
DID0153-S001R	D9	16,66667	Tr insuf
DID0076-S001R	D9	50	Pass

Figure 145 – Extraits du tableau évaluatif général des apprenants par variable : discrimination qualitative visuelle des étiquettes appréciatives des variables DUR_T et RYT_lex (ordre des colonnes pour chaque tableau : liste des apprenants, leur catégorisation d’après leur vitesse de lecture du texte entier, résultat quantitatif en secondes et en pourcentages pour DUR_T et RYT_lex respectivement et catégorisation évaluative à partir des résultats quantitatifs)

Les 28 variables ont été soumises à ce traitement de leurs étiquettes appréciatives, ce qui a résulté en une catégorisation en trois groupes qualitatifs : groupe vert (jugé positif dans la variable, équivalent à un bonus dans l’évaluation), groupe rouge (jugé négatif dans la

CHAPITRE IV – Expérimentation de création de profils et de critères d'évaluation prosodiques d'apprenants

variable, associé à un malus), groupe neutre (jugé moyen ou passable, et non spécifiquement marqué dans la variable). Pour chaque variable, les résultats de bonus (en vert) et de malus (en rouge) ont été comptabilisés puis convertis en pourcentage de conformité ou d'erreur. Les graphiques ci-dessous ont été élaborés à partir de ces résultats en deux versions : avec la totalité des 28 variables⁸⁹, puis avec 25 variables, afin d'éviter de donner trop de poids à certains phénomènes examinés sous deux angles différents et alors comptabilisés doublement.

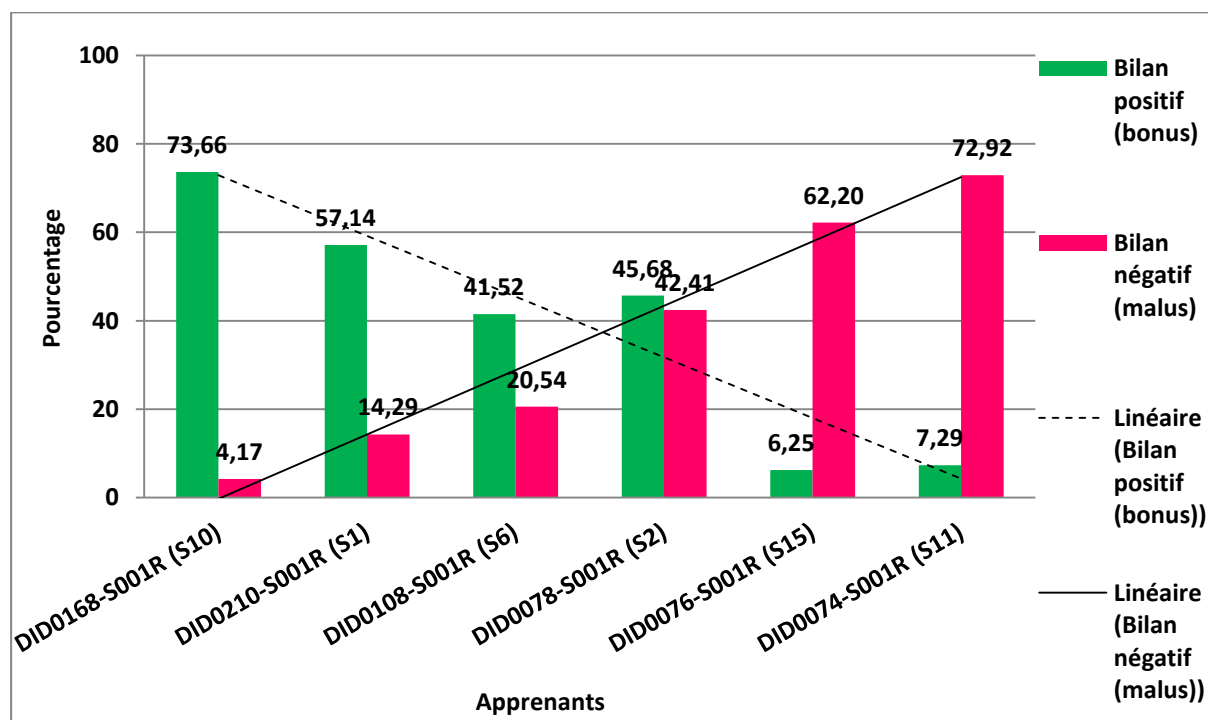


Figure 146 – Classement des six profils d'apprenants selon 28 variables prosodiques en fonction de la proportion du taux d'erreur (bilan négatif de malus en rouge)

⁸⁹ Fichier « Excel Mesures_apprenants10.xlsx » (en annexe).

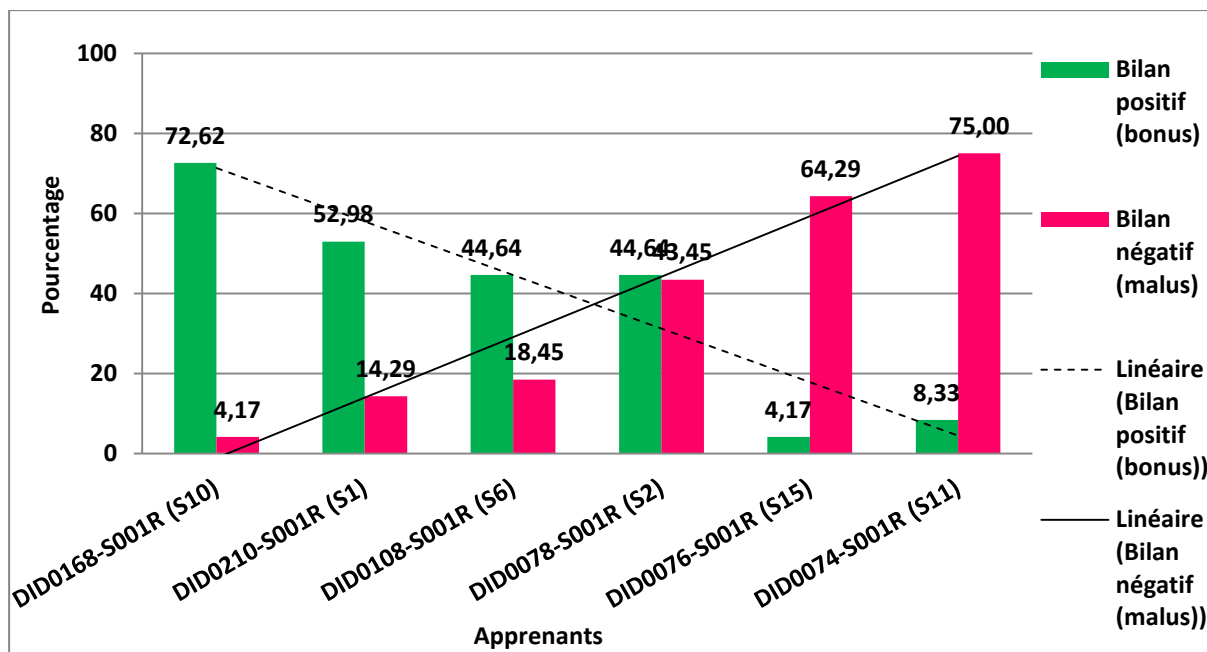


Figure 147 – Classement des six profils d'apprenants selon 25 variables prosodiques en fonction de la proportion du taux d'erreur (bilan négatif de malus en rouge)

Dans les graphiques ci-dessus, les profils d'apprenants ont été classés selon un critère discriminant : par taux d'échec par rapport aux critères relevés et évalués précédemment, le critère de réussite par attribution de bonus étant beaucoup moins révélateur. Ces classements permettent d'emblée de percevoir deux profils : le profil stéréotypique de réussite ou de qualité prosodique maximale (incarné par l'apprenant S10) et le profil stéréotypique d'échec ou de qualité prosodique minimale (représenté par S11). Ils sont tous deux situés aux deux extrêmes du paradigme. Ils sont typiques dans la mesure où ils représentent deux niveaux extrêmes encadrant les autres niveaux intermédiaires entre ceux-ci. Le profil S11 est celui d'un apprenant débutant ou faible. Son taux d'erreurs est maximal et son taux de réussite minimal. En s'améliorant, son taux d'erreurs est censé diminuer et son taux de réussite augmenter. Ces histogrammes confortent les résultats obtenus sur le plan factoriel ACP des quinze apprenants à partir des 25 variables (Figure 143) et le dendrogramme de McQuitty utilisant les 28 variables (Figure 146). Un traitement à partir du taux de malus sur 25 ou 28 variables corrobore les conclusions émises, alors qu'un classement par taux de bonus placerait S2 et S6 au même niveau et intervertirait S11 et S15, que le nombre de variables soit 25 ou 28.

Pour chaque profil, un pourcentage minimal avoisinant les 20-25% correspond à un étiquetage appréciatif non marqué et non discriminant du point de vue évaluatif (en blanc sur le graphique). On remarque pourtant que le profil S2 se démarque des autres profils en

CHAPITRE IV – Expérimentation de création de profils et de critères d'évaluation prosodiques d'apprenants

laissant moins de 12% à cet étiquetage non marqué. On en déduit que ce profil est déterminé à la fois par des taux de réussite (bonus) et d'échec (malus) très importants, laissant peu de place à des résultats moyens ou passables, comme si l'apprenant se situait sur une phase plateau dans son apprentissage. On reconnaît le profil S2 mis en évidence par les plans factoriels en ACP et on commence à percevoir la complexité de ce type de profil grâce à l'analyse par taux de réussite-bonus et d'erreurs-malus.

On observe en outre que ce classement par taux de réussite et d'erreurs opère une hiérarchisation qualitative des réalisations prosodiques en fonction du domaine d'investigation semblable à celle présentée par le dendrogramme de McQuitty. Il y aurait correspondance exacte si les taux de réussite et d'échec étaient inversement proportionnels.

Une investigation de ces pourcentages pour chacun des six profils éclaire leur nature. Le tableau ci-dessous en détaille les résultats. Afin de se rapprocher des réalisations effectives de chaque apprenant pour réaliser le profil, orientons-nous vers une analyse des 25 variables.

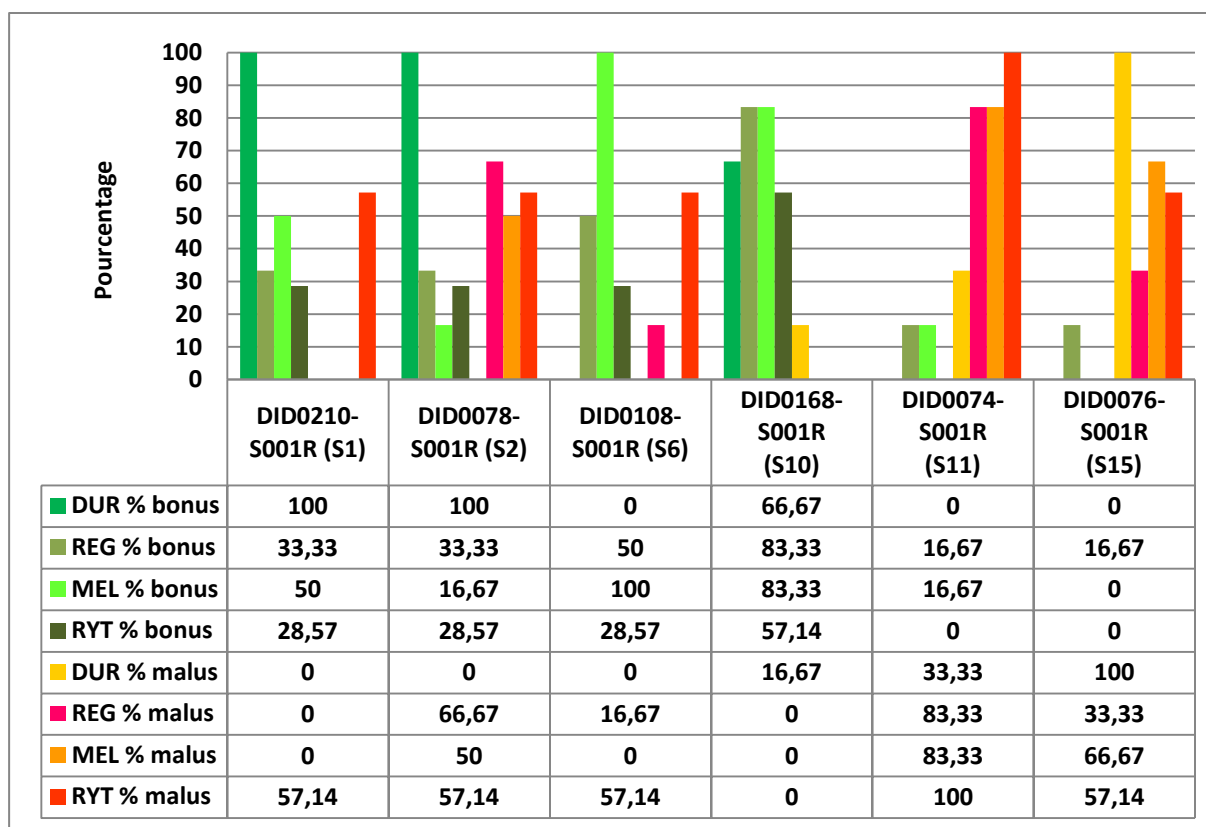


Tableau 74 – Moyenne des 25 variables distinctives par domaine dans les six profils d'apprenants (en pourcentage, 0 = non marqué)

Le graphique qui accompagne le tableau ci-dessus montre que tous les profils, sauf le « meilleur » (S10), atteignent les 100% par une de leurs variables. S1 et S2 se caractérisent

par les 100% de leur vitesse de lecture étant donné qu'ils appartiennent au groupe D1, le plus rapide. S6 se caractérise par un taux de réussite maximal dans sa variable mélodique. S11 produit 100% d'échec rythmique, et S15 une lenteur à 100%.

Bien que n'atteignant pas 100%, S10 n'enregistre que des résultats positifs et élevés dans les quatre domaines (durée : 66,66%, registre : 83,33%, mélodie : 83,33% et rythme : 57,14%) à l'exception de la durée, étant donné que sa faible vitesse d'élocution le classe dans le troisième quartile (Q3), et du rythme. Son succès global le place en position de tête comme « meilleur » étudiant en prosodie.

À l'opposé, le taux d'échec le plus fort est celui de S11. Avec ses 33,33% d'erreurs-malus en durée, 83,33% en registre, 83,33% en mélodie et 100% en rythme, il est ici l'archétype du « mauvais » étudiant en prosodie. Son taux de réussite-bonus en registre et mélodie est particulièrement timide (16,66% et 16,66%) et peu significatif : dans les deux cas, cela résulte d'une réalisation acceptable du syntagme intonatif S9 à l'emphase réduite.

S15 doit son échec maximal à ses 100% de lenteur provenant de son classement dans le groupe D9. Par ailleurs, ses résultats sont nettement moins négatifs que ceux de S11 avec 33,33% d'échec en registre, 66,66% en mélodie et 57,14% en rythme.

S6 réalise le meilleur pourcentage de réussite en mélodie avec 100% et en surpassant S10, alors que son succès en registre et rythme est nettement moindre (50% et 28,57%). En effet, c'est aussi dans ces domaines qu'on le trouve en défaillance : échec de 16,66% et 57,14% respectivement.

S1 et S2 ont de nombreux points communs : réussite temporelle de 100% et échec de 0% puisqu'ils appartiennent tous deux à D1, réussite rythmique de 28,57% et échec rythmique de 57,14%. Ils diffèrent essentiellement par leur registre et leur mélodie : le score de S1 est substantiel avec respectivement 33,33% et 50% de réussite, alors que S2 totalise un taux de réussite moindre (33,33% et 16,66%) et des taux d'échec de 66,66% et 50%. Pourtant, on a précédemment pu observer la place particulière qu'occupe le profil S2 sur les plans factoriels, ainsi que sur les histogrammes. L'analyse à partir des 28 variables, bien qu'elle modifie les données, n'altère en aucune façon la teneur des remarques effectuées.

En conséquence, la synthèse que l'on peut tirer de cette analyse par taux de réussite-bonus et d'erreurs-malus est que :

- Rapidité de lecture n’équivaut pas à qualité mélodique, rythmique ou de registre. Au contraire, elle semble difficilement réconciliable avec un registre conséquent et, par voie de conséquence, une bonne qualité mélodique. Cela signifie que les apprenants les plus rapides ne produisent pas la meilleure qualité prosodique, même si les auditeurs anglophones natifs, hypersensibles au domaine de *fluency* (aisance), favorisent spontanément cette variable au détriment des autres variables prosodiques, qu’ils ne perçoivent alors plus.
- À l’extrême opposé, une grande lenteur de lecture rend l’apprenant capable de respecter un rythme et une mélodie acceptables. Pourtant, on n’observe pas toujours quelque réussite lorsque le degré d’ampleur de registre est élevé. En conséquence, on peut affirmer qu’une vitesse d’élocution très lente indique un faible niveau prosodique, mais pas nécessairement le niveau le plus faible.
- La meilleure qualité prosodique provient des apprenants de Q1 et surtout Q3, ce qui prouve que le critère de lenteur, à condition que celle-ci soit acceptable, est une condition pour obtenir qualité mélodique, rythmique et de registre.
- Les apprenants de la Médiane semblent ne pas particulièrement se distinguer des valeurs centrales, c’est du moins ce que l’on peut observer jusqu’à présent en analysant la prosodie des profils personnalisés sélectionnés.

Après avoir identifié dans la sélection du corpus de départ les quinze individus que j’ai analysés, je peux maintenant ordonner les profils selon une hiérarchisation prosodique qualitative globale, de la meilleure maîtrise prosodique à la plus faible, tous domaines prosodiques confondus :



<p>Classement selon système de bonus/malus et catégorie de vitesse correspondante (25 et 28 variables)</p>	<p>“meilleur” “moins bon”  S10 > S1 > S6 > S2 > S15 > S11 Q3 D1 Q1 D1 D9 Q3</p>
<p>Classement selon rapidité de lecture et catégorie de vitesse correspondante</p>	<p>“plus rapide” “plus lent”  S1 > S2 > S6 > S10 > S11 > S15 D1 D1 Q1 Q3 Q3 D9</p>

Tableau 75 – Comparaison des classements selon le système de bonus-malus prosodique et celui de vitesse de lecture, avec catégories de vitesse correspondantes

Il faut noter qu’une analyse à 25 ou 28 variables conserve le même classement qualitatif. Les profils de la plus grande et la plus faible qualité prosodique dépendant de la même catégorie

de vitesse de lecture, il est fondé d'affirmer que la catégorisation par rapidité de lecture ne fédère pas suffisamment les autres domaines prosodiques pour être discriminante à elle seule. Cela est confirmé par la diversité de l'appartenance des autres individus à d'autres catégories.

Le bilan que l'on peut extraire de ces remarques sur l'analyse par taux de réussite-bonus et d'échec-malus est que le facteur temporel n'étant pas obligatoirement discriminant pour évaluer la qualité prosodique, et remarquant qu'il en est de même en ce qui concerne le registre (facilité à produire un syntagme intonatif non emphatique d'une ampleur très acceptable), ce sont les variables mélodiques et rythmiques qui doivent prioritairement retenir l'attention des évaluateurs pour départager finement et globalement la qualité des réalisations. Cela nous conduit à étudier de plus près la représentativité des variables dans l'opération de profilage.

6.5 Analyse des données par cercles de corrélations

Afin d'évaluer la pertinence de chaque variable dans la détermination des profils, j'ai utilisé la méthode d'analyse par cercles de corrélations. La figure suivante montre l'influence respective des 28 variables numériques.

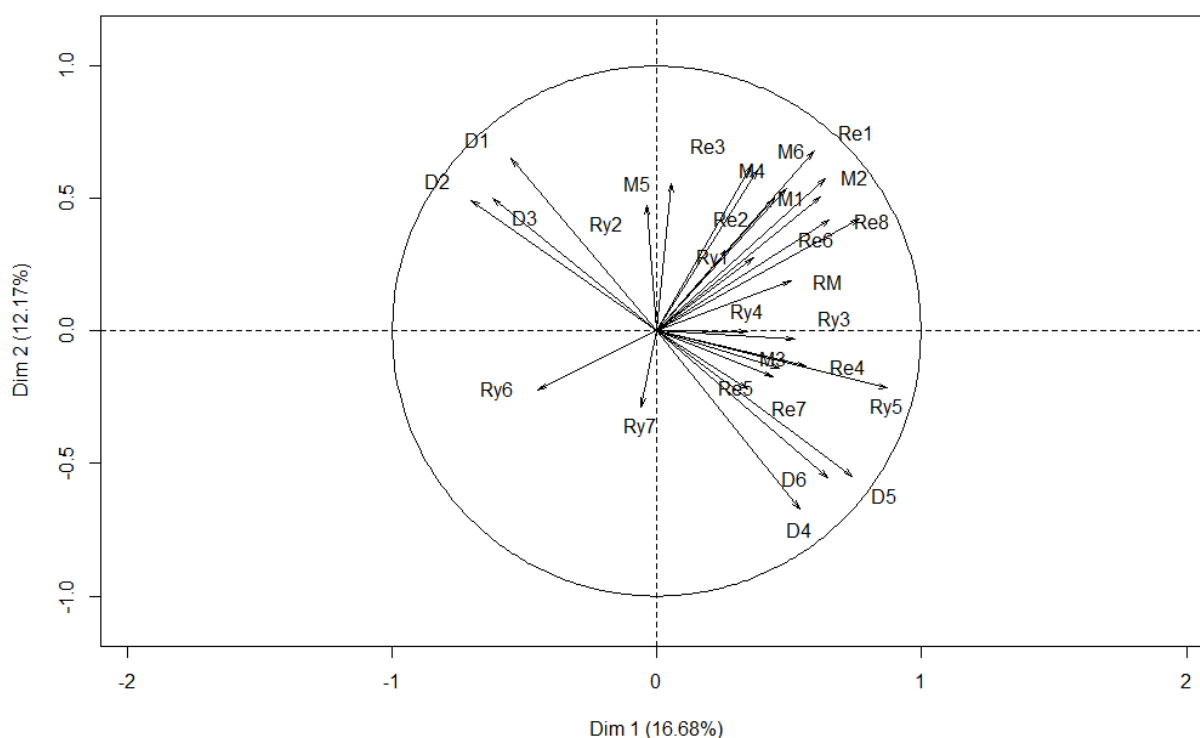


Figure 148 – Cercle de corrélations des 28 variables prosodiques des apprenants sur le premier plan (dimensions 1 et 2)

Le nombre de dimensions sur lesquelles les variables s'inscrivent se traduit par le faible pourcentage des deux premières : Dimension 1 de 16,68% et Dimension 2 de 12,71%. Ces

faibles pourcentages montrent que les variables se répartissent relativement équitablement sur plusieurs dimensions. En d'autres termes, cela équivaut à dire que l'objectif de caractériser les profils par des variables représentatives des diverses particularités prosodiques a de fortes chances d'avoir été atteint. Le cercle de corrélations représentant 25 variables (ci-dessous) ne modifie que très légèrement l'ensemble. L'augmentation des pourcentages des deux premières dimensions ne résulte que du plus faible nombre de variables.

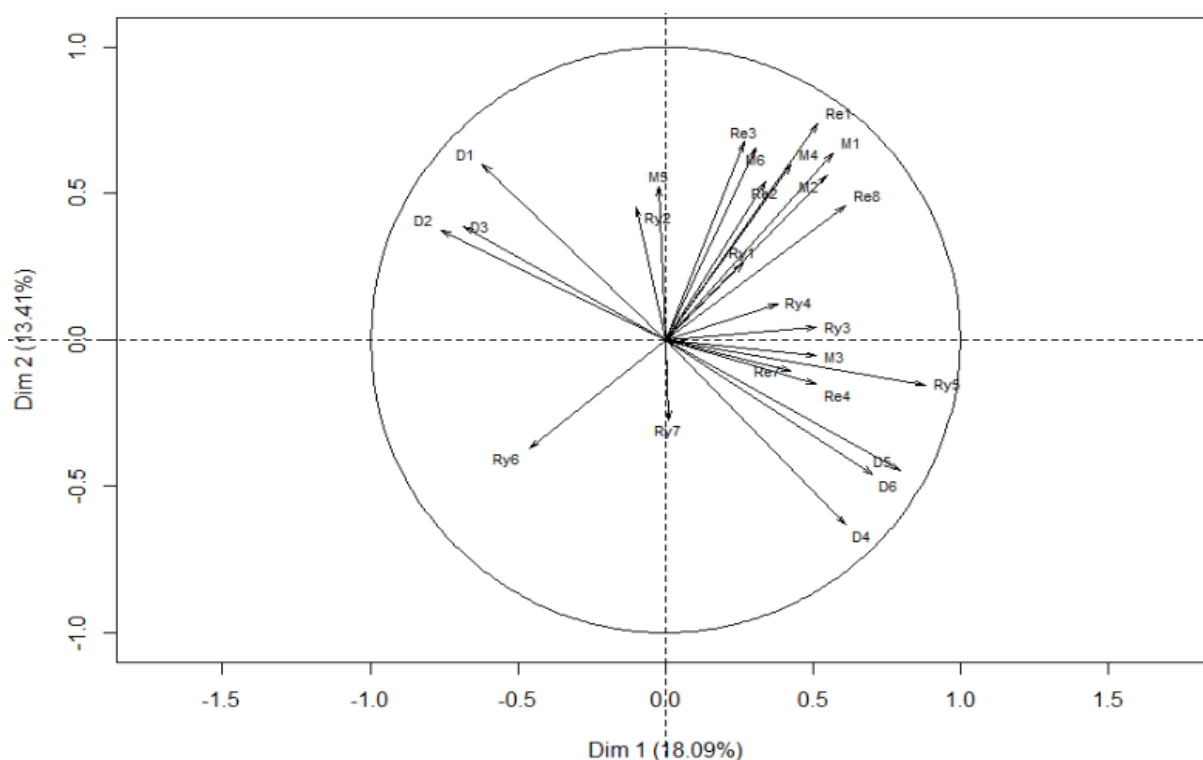


Figure 149 – Cercle de corrélations des 25 variables prosodiques des apprenants sur le premier plan (dimensions 1 et 2)

Le nombre élevé de ces variables ne permet toujours pas de mettre en évidence leur influence dans leur domaine respectif. Aussi, c'est en fonction d'une catégorisation par domaine que l'on peut voir celles qui sont les plus fédératrices et, en conséquence, à privilégier dans l'établissement d'un profilage par rapport à celles dont la pertinence est moindre.

6.5.1 Cercles de corrélations des variables temporelles

Le cercle de corrélations suivant montre l'influence respective des six variables temporelles : D1, D2 et D3 focalisées sur la rapidité de lecture, et D4, D5 et D6 sur le débit de syllabes par seconde. Les 46,45% totalisés par l'addition des pourcentages des deux dimensions, bien que largement supérieur à celui du cercles des 28 variables, est insuffisant pour situer ici six variables.

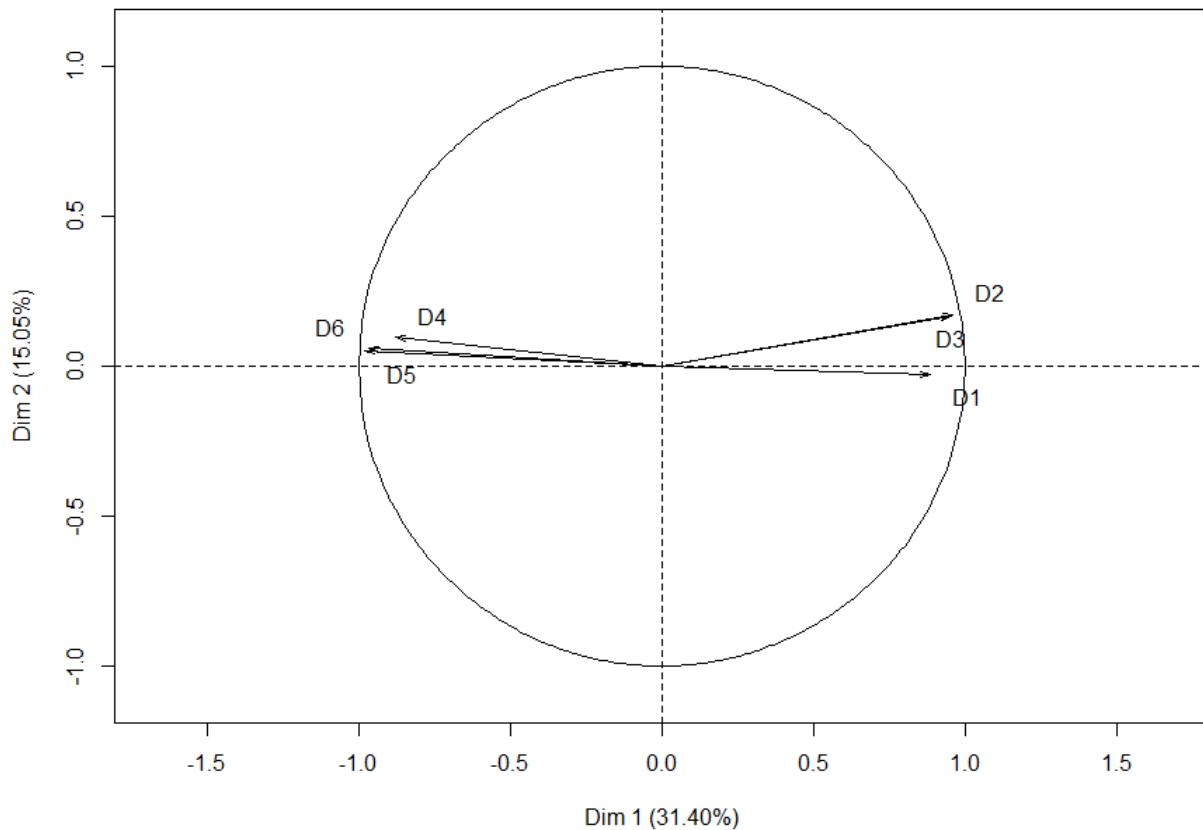


Figure 150 – Cercle de corrélations des 6 variables temporelles des apprenants

On peut voir que les flèches les plus longues sont celles de D2 (DUR_ABCp) et D5 (DEB_ABCp), qui correspondent aux mesures des tronçons non ponctués avec pauses incluses. L'avantage des variables qu'elles représentent est qu'elles restent utilisables pour analyser la prosodie d'une reformulation ou d'une conversation, ce qui est impossible pour une variable centrée exclusivement sur un texte lu. Dans une mesure toute relative, D2 et D5 peuvent donc être considérées comme meilleures représentatives de la vitesse de lecture et du débit de chaque apprenant, ce qui est corroboré par leur correspondance sur deux parties opposées du cercle en indiquant une corrélation relativement négative. Ces flèches étaient aussi les plus longues dans le cercle de corrélations regroupant les 28 variables présenté en amont, ce qui confirme leur représentativité parmi l'ensemble des variables.

On constate en outre la corrélation négative entre le nombre de syllabes prononcées par seconde (D4, D5, D6) et le temps de lecture (D1, D2, D3) puisqu'il s'agit de deux unités de mesure qui s'opposent : un locuteur prononçant un nombre de syllabe élevé en un temps donné aura une durée de lecture plus courte.

En conséquence, il semble plus avisé de choisir la variable DUR_ABCp pour évaluer la vitesse de lecture et DEB_ABCp pour évaluer le nombre de syllabes par seconde en conversation.

6.5.2 Cercle de corrélations des variables de registre

Le cercle de corrélations suivant ne comporte que six variables parce que les variables Re5 et Re6 ont été écartées de l'analyse : elles reprenaient les autres variables par leur moyenne ou leurs extrêmes, ce qui alourdisait l'analyse sans apport positif. La somme des pourcentages sur lesquels fonctionnent les deux dimensions est encore moindre que précédemment, ce qui montre la relativité de leur représentativité.

Dans ce cercle à six variables, la longueur des flèches montre que Re1 (REG_Excl) est la variable la plus fédératrice, suivie de Re3 (REG_so4). Leur proximité montre qu'elles fonctionnent en relative synergie, ce qui est un résultat attendu étant donné que ces deux variables représentent des zones très emphatiques : exclamative au premier rang d'emphase et *so sweet(ly)* au 7^{ème} rang. On retrouve néanmoins Re7 située au troisième rang d'importance sur ce plan. Sa position en angle quasiment orthogonal par rapport à Re1 montre qu'elle n'est aucunement corrélée à Re1, ce qui est normal étant donné leur opposition emphatique/non emphatique puisque S9 a été sélectionné pour son caractère non emphatique. On remarque aussi que Re4 (*so sweet* de la Phrase 22, au 2^{ème} rang d'emphase) n'est en rien corrélée au registre de Re3 (*so sweet(ly)* de la Phrase 4). Cela confirme ce que l'on voyait sur cercles de corrélations avec 28 variables et surtout avec 25 variables (angle quasiment droit entre les deux flèches de Re3 et Re4), et plus en amont, dans l'histogramme du profilage de groupes par archétypes de conformité selon la vitesse de lecture. On peut en déduire qu'ici, Re1 contribue le plus fortement au modèle, suivie de Re3. Re7 semble être une moins bonne indicatrice d'aisance dans la maîtrise du registre, comme le confirment les flèches courtes des cercles de corrélations avec 25 et 28 variables.

Par comparaison, on remarque que le cercle de corrélations représentant les 28 variables favorise Re1 et Re6 quasiment au même niveau, ce qui en fait les deux variables phares de registre. Elles se situent devant Re3 et Re8, ces deux dernières flèches similaires par leur longueur, phénomène davantage visible dans le cercle avec 25 variables. Dans le cercle de corrélations avec 28 variables, Re8 est d'ailleurs relativement proche de Re6 (moyenne des mesures de registre de l'emphase de Re1-2-3-4-8) et bien que moins représentative qu'elle (flèche plus courte), elle est la variable de registre qui lui est la plus

proche, ce qui indique leur forte corrélation. Re4 est non seulement peu représentative par la faible longueur de sa flèche, mais encore par sa position dans le quart bas du cercle avec 25 ou 28 variables, marquant un angle obtus conséquent avec Re1, Re2 et Re3. On observe encore davantage les mêmes caractéristiques pour Re7.

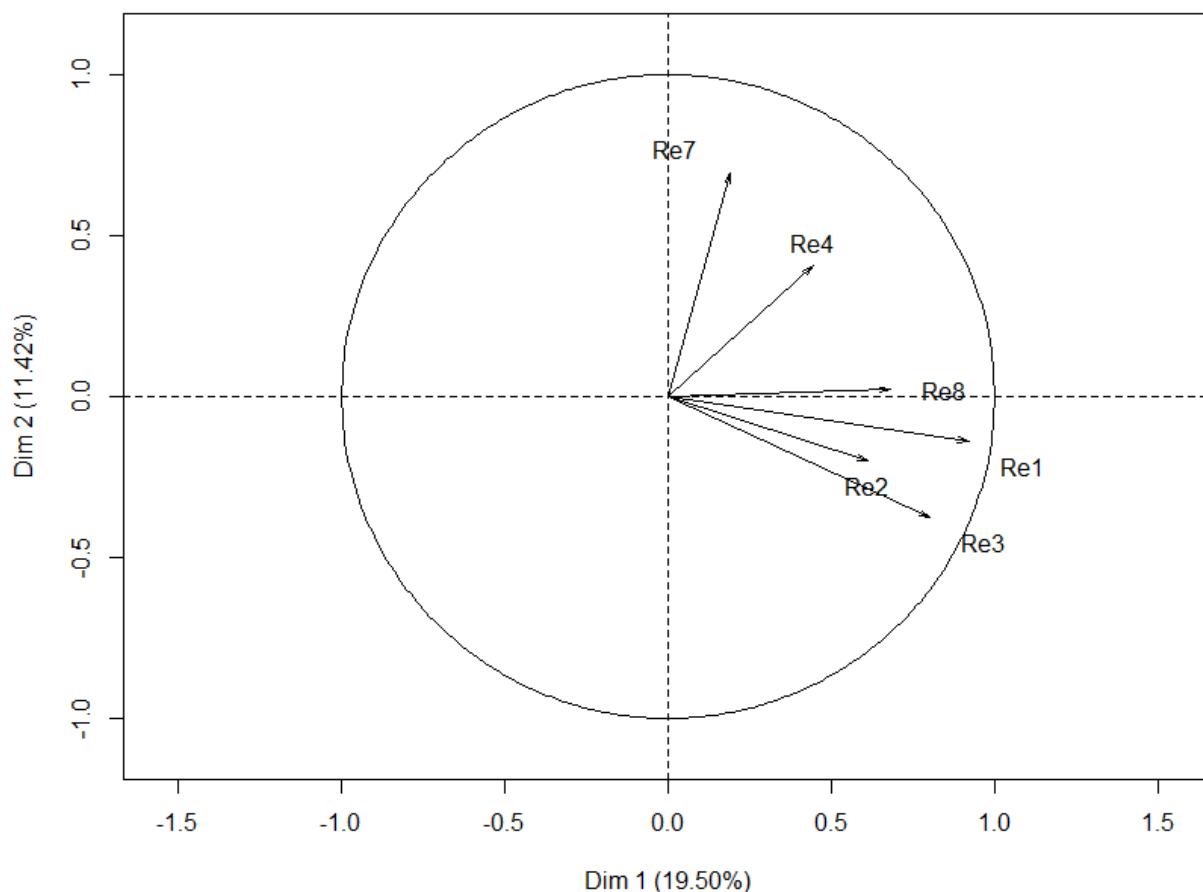


Figure 151 – Cercle de corrélations des 6 variables de registre des apprenants

On déduit de ces analyses que Re1 constitue la variable prioritaire à étudier pour cadrer l'aptitude à réaliser un registre correct, suivie de Re8, qui contribue elle aussi fortement au modèle par sa flèche plus longue que celle de Re2 (Phrase 5 entière) pour un angle similaire dans le cercle avec six variables. L'opposition entre Re3 et Re4 pourrait être attribuée à un schéma rythmique différent dans les deux occurrences de *so sweet-*, ce qui leur confère un caractère évaluatif peu stable.

On a remarqué le caractère peu fréquent et même désuet de l'utilisation d'une exclamative en *how*. La variable Re8 portant sur la syllabe accentuée de *nobody* fournit un repère intéressant par sa facilité de prise de mesures ainsi que par le nombre d'occurrences accru par rapport à Re1, que ce soit en lecture ou en conversation. Les autres variables ne sont

pas retenues car elles contribuent moins fortement au modèle ou sont de caractère aléatoire, sauf très précisément Re3 dans ce texte.

6.5.3 Cercles de corrélation des variables mélodiques

Dans le cercle de corrélations des variables mélodiques, ici aussi, la somme du pourcentage d'influence couvert par les deux premières dimensions (37,40%), est insuffisante pour représenter correctement les six variables mélodiques.

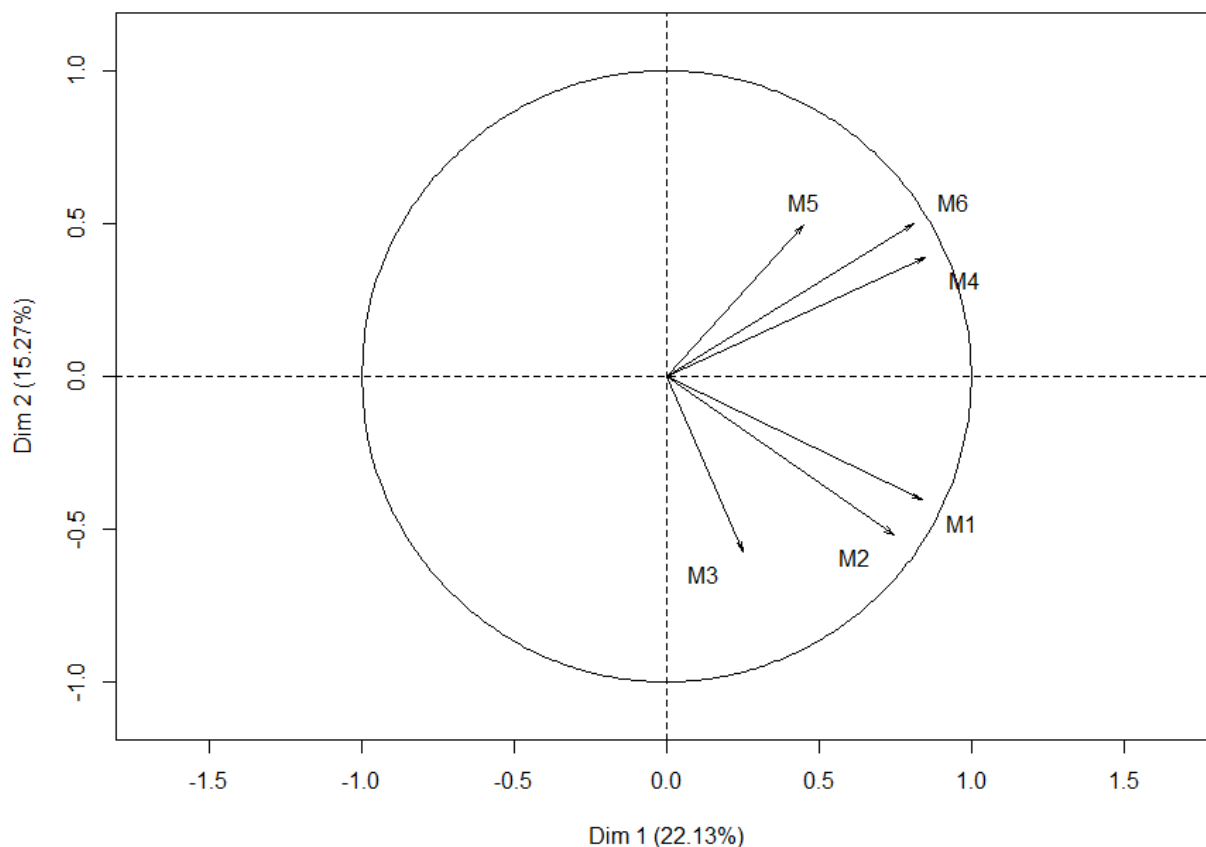


Figure 152 – Cercle de corrélations des 6 variables mélodiques des apprenants

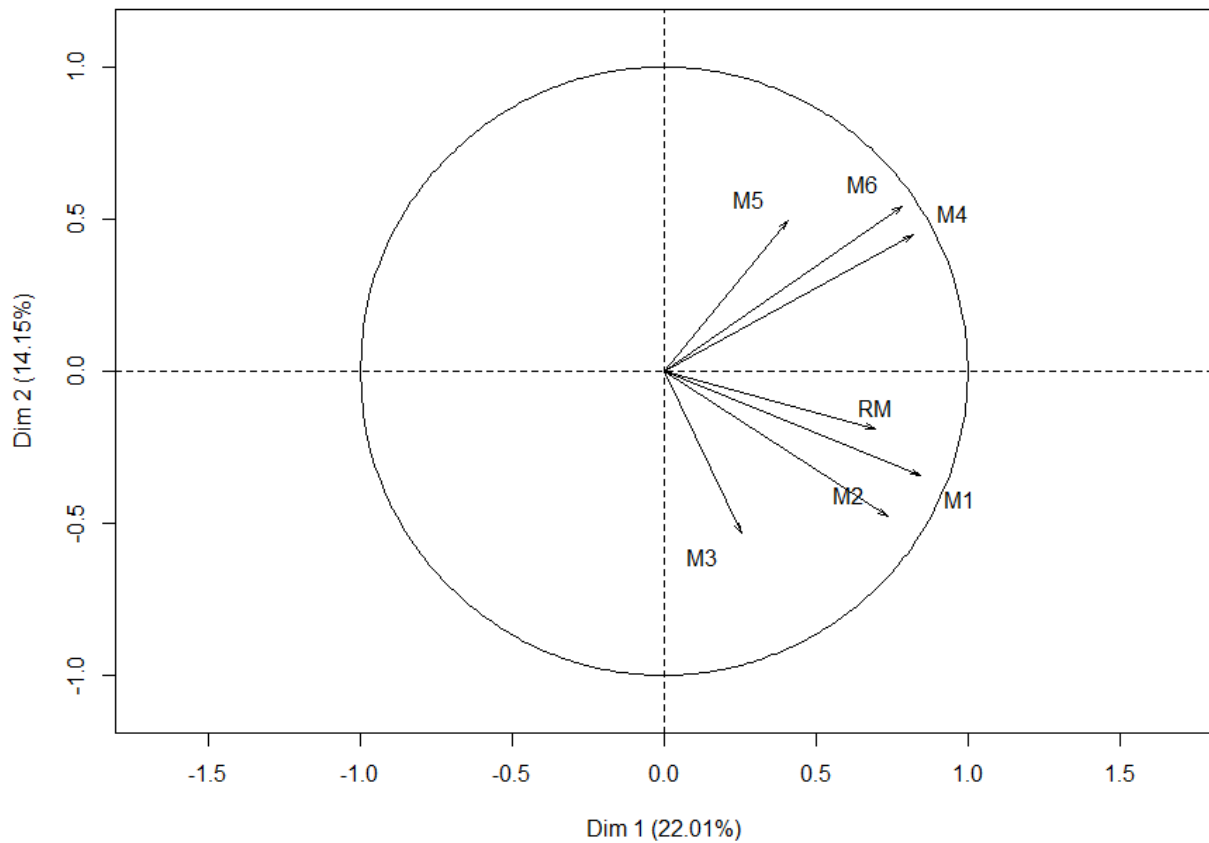
Les six variables ont été conservées parce que chacune apporte une vision différente des phénomènes observés. Dans la partie supérieure du cercle, on remarque, par la longueur des flèches, que la variable la plus fédératrice est M6 (MEL_no), de laquelle M4 (MEL_Λ) se rapproche, suivie de M5 (MEL_∨), montrant le lien qui les unit. Dans le bas du cercle, formant un angle quasiment orthogonal au premier ensemble, on remarque la proximité des flèches M1 (MEL_Excl) et M2 (MEL_P5) au profit de la première par sa longueur. La variable M3 (MEL_S9) rejoint ces dernières en observant quelque distance. La flèche courte et excentrée de M3 (MEL_S9) montre que la mélodie du syntagme intonatif non emphatique n'a que peu de rapports avec les autres mesures mélodiques présentes dans le cercle de

corrélation. Les cercles avec 25 et 28 variables la représentaient très éloignée des autres variables, marquant un angle particulièrement obtus avec M6.

Cette disposition des flèches laisse entrevoir que la véritable aptitude mélodique est représentée par les variables M6 puis M4 et M5. En effet, M1, M2 et M3 se focalisent sur une chute intonative dans une unité intonative de plusieurs mots, phénomène que l'on retrouve en français, alors que M4, M5 et M6 révèlent l'aptitude à moduler la F0 sur une seule et même syllabe, ce qui est typique d'une différence intonative fondamentale entre le français et l'anglais et sans doute la plus difficile à réaliser, parfois même pour les natifs anglophones, comme l'avait souligné Horgues (2010).

Ceci indique que pour minimiser les prises de mesures, on pourrait ne sélectionner qu'une variable dans chacun des deux groupes, chacune destinée à évaluer un niveau de maîtrise différent. On pourrait ainsi se dispenser des variables MEL_∧ (M4) et MEL_∨ (M5) au profit de MEL_no (M6) d'une part, et de MEL_P5 (M2) ou même MEL_S9 (M3) au profit de MEL_Excl (M1) d'autre part. MEL_no et MEL_Excl sont suffisamment représentatives. Si l'on devait ne retenir qu'une variable mélodique parmi les six, il faudrait sélectionner MEL_no (M6), car elle est non seulement un peu plus représentative que MEL_Excl par la longueur de sa flèche, mais encore le mot *nobody* devrait être moins rare dans la langue parlée actuelle qu'une exclamative en *How*. Ici, la complexité mélodique sur une seule syllabe très emphatique peut donc être considérée comme la variable la plus représentative de la qualité mélodique d'un apprenant de niveau avancé et la chute mélodique de MEL_Excl comme celle d'un apprenant débutant.

L'inclusion de la variable RYTMEL_so4 (RM) ne modifie en rien l'ordonnement des variables (voir figure ci-dessous). La petite taille de sa flèche révèle dans les cercles à 6 et 28 variables son manque de pertinence dans l'analyse mélodique par rapport aux autres variables. Ceci est corroboré par une diminution du pourcentage des deux dimensions représentées dans le cercle à six variables, surtout de la dimension 2 : réduction de 15,27% à 14,15%. Cette variable RM, se voulant à la fois mélodique et rythmique, semble être assez peu représentative de la qualité mélodique.



6.5.4 Cercles de corrélations des variables rythmiques

Les deux premières dimensions du cercle de corrélation ci-dessous totalisent un pourcentage de 25,67% pour sept variables. Cette faible représentativité signifie que la notion de rythme a été explorée de manières diverses, utilisant plusieurs dimensions, ce qui était l'objectif à atteindre.

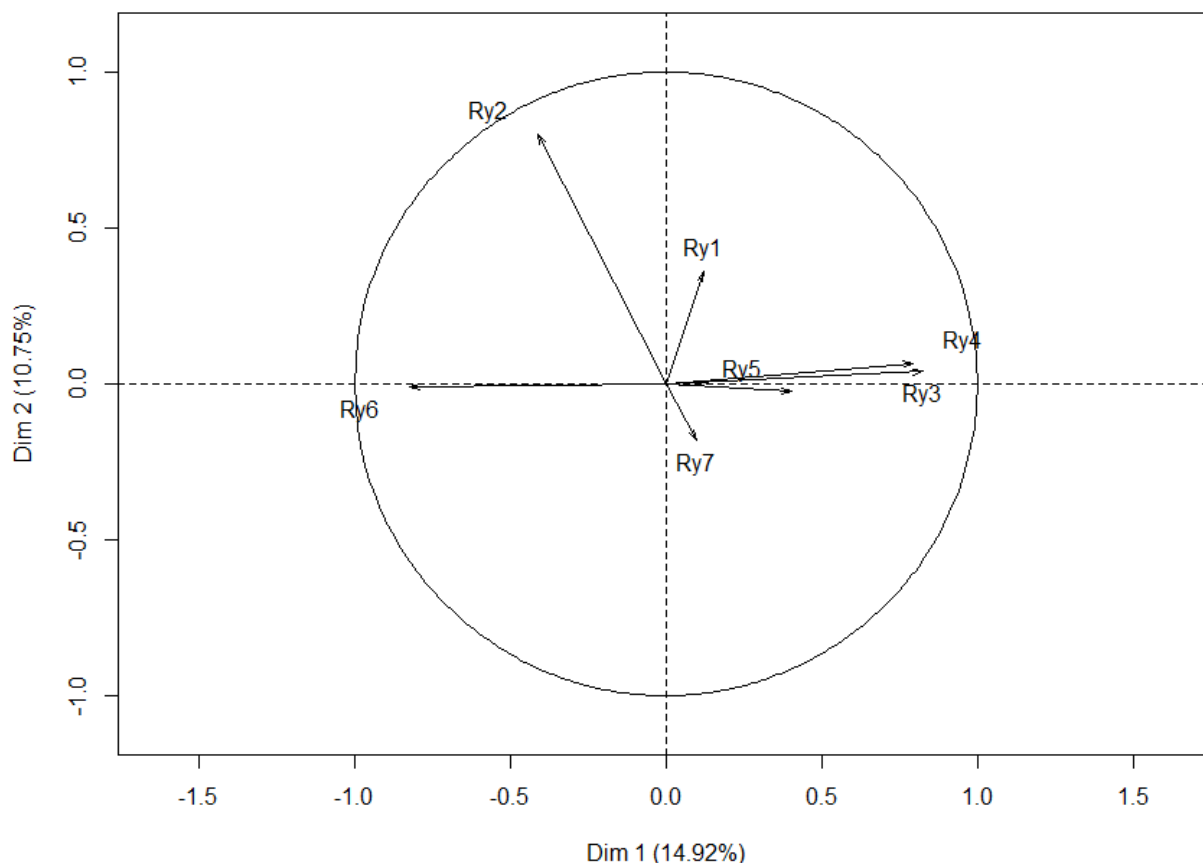


Figure 154 – Cercle de corrélations des 7 variables rythmiques des apprenants

Parmi les sept variables représentées, le cercle ci-dessus montre la prédominance de la variable Ry2 (RYT_so4), suivie dans un tout autre ordre par Ry6 (RYT_them), puis Ry3 (RYT_so22) et Ry4 (RYT_all12), ces deux dernières entretenant une très forte correspondance, laquelle renforce l'hypothèse des liens qu'entretiennent les paires « *so* + adjectif ou adverbe » avec « *all* + préposition ». Contre toute attente, ce cercle montre aussi la forte corrélation négative entre Ry3 et Ry6 (RYT_them), variables qui semblent pourtant éloignées par leur nature et leur méthode de profilage. Alors que la flèche courte de Ry7 (RYT_syll) révèle sa faible pertinence dans ces deux premières dimensions, celle plus longue et plus excentrée de Ry1 (RYT_lex) montre que ses propriétés sont davantage mises en évidence par le cercle de corrélation rythmique tel qu'il est défini dans le plan défini par les dimensions 1 et 2 de l'ACP.

L'ajout de la variable RM (RYTMEL_so4) modifie le cercle de diverses manières. Tout d'abord, les deux dimensions du cercle sont modifiées. Alors que la première subit une réduction de 14,92% à 14,05%, la dimension 2 augmente son pourcentage de 10,75% à 12,62%, l'ensemble totalisant 26,67% en légère augmentation par rapport à 25,67%. Cette avancée est toute relative étant donné que le cercle représente huit variables. On remarque que

RM se situe essentiellement dans la deuxième dimension dont elle est la symétrique de Ry2 (RYT_so4) par rapport à l'axe vertical. RM est donc davantage une variable rythmique qu'une variable mélodique. Mais surtout c'est à présent RM qui devient la variable la plus fédératrice avec la plus longue flèche du cercle. RM a été construite à partir de la paire *so sweetly* : RM pourrait donc être représentative de Ry2 (RYT_so4). On remarque par ailleurs que RM est orthogonale à Ry6 (RYT_them) et Ry4 (RYT_all12). Ry3 (RYT_so22) étant de même longueur et proche de Ry4, Ry6 pourrait représenter Ry3 et Ry4 dont les flèches sont plus courtes. On note ensuite que la variable Ry5 (RYT_all16) se rallonge considérablement et change de place en se situant entre RM et Ry3 (RYT_so22), marquant son lien avec l'emphase de sa famille « adverbe modificateur + adverbe ».

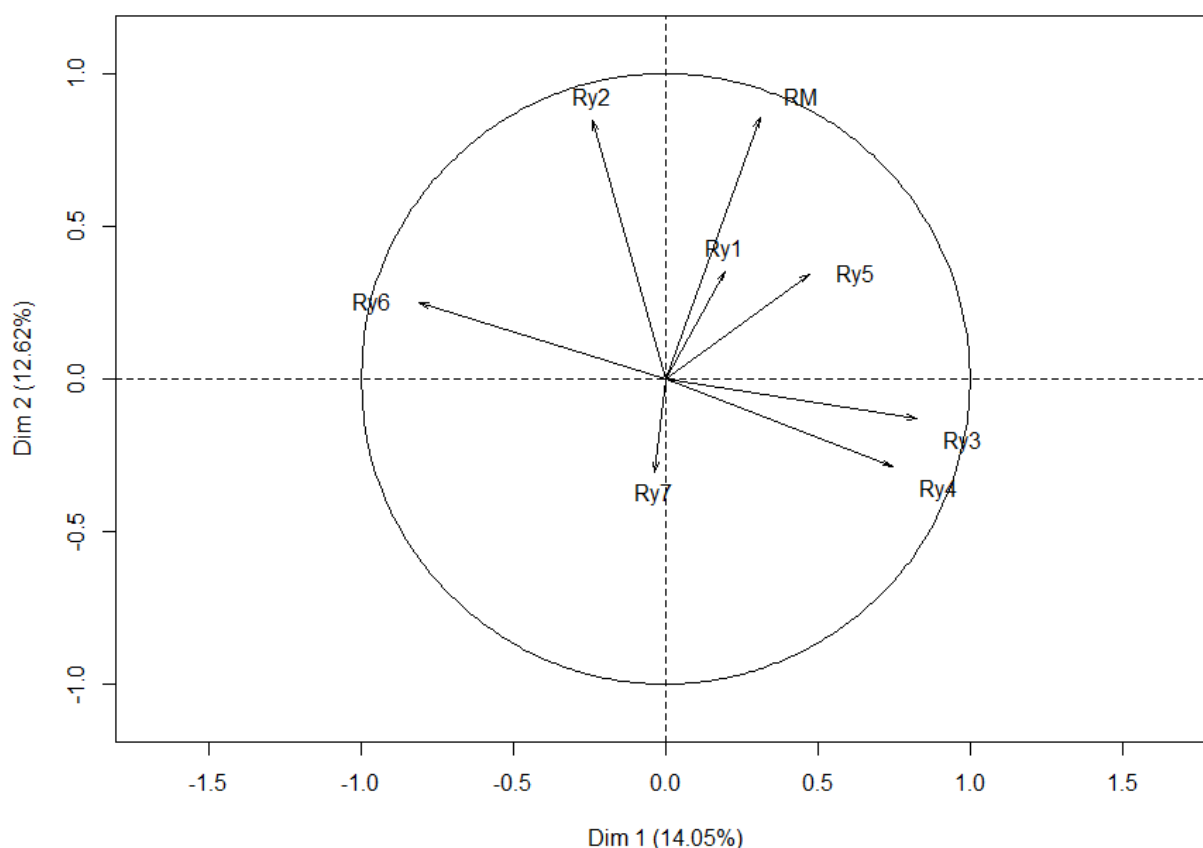


Figure 155 – Cercle de corrélations des 7 variables rythmiques des apprenants avec RM (RYTMEL_so4)

Parmi les variables rythmiques représentées dans le cercle de corrélations à sept variables, devraient être conservées celles qui sont les plus représentatives pour exprimer la particularité du sous-domaine auquel elles appartiennent. Il semble pourtant difficile d'effectuer un choix judicieux.

- Rythme par alternance de proéminences :

- Respect des patrons accentuels canoniques dans les polysyllabes : RYT_lex (Ry1), seule variable de la catégorie
- Absence de proéminence sur pronom en fin de syntgme intonatif : RYT_them (Ry6), seule variable de la catégorie
- Rapport de durée inter-syllabique sur paire de syllabes adjacentes : RYT_syll (Ry7), seule variable de la catégorie
- Qualité de proéminence par F0 sur marqueur : on constate un fort contraste entre la portée des flèches dans les cercles à 25 ou 28 variables (Ry5 extrêmement représentative parmi toutes), et dans ceux à sept ou huit variables (Ry2, Ry3 et Ry4 les mieux représentées)

Si le regroupement des variables par domaine a permis de mettre en évidence leurs caractéristiques communes afin de hiérarchiser leur pertinence dans les opérations de profilage et de contribuer à éclairer les résultats obtenus par plan factoriel, ce n'est pas le cas ici des variables d'ordre rythmique. Afin d'organiser les données pour mettre en lumière leur correspondance sous un autre jour, mettons en œuvre dans la rubrique suivante un profilage stéréotypique.

6.6 Analyse des données par profils stéréotypiques

Cette section a pu utiliser les 56 variables dont 28 quantitatives et 28 qualitatives pour déterminer une typologie d'apprenants en matière prosodique. Pour cela, je me suis appuyée sur les quatre grands domaines prosodiques que sont la vitesse d'élocution, le registre, la qualité mélodique et le rythme. Les outils de profilage façonnés au Chapitre II ont montré leur pertinence dans la découverte de profils de conformité et stéréotypiques.

Ainsi, l'analyse statistique a pu extraire les sept profils stéréotypiques suivants par leurs caractéristiques extrêmes mises en évidence dans l'étude sur tableur :

- Le plus rapide : DID0210-S001R (S1) avec 100% de réussite en vitesse
- Le plus lent : DID0076-S001R (S15) avec 100% d'échec en vitesse
- Le plus mélodieux : DID0108-S001R (S6) avec 100% de réussite en mélodie
- Le moins rythmé : DID0074-S001R (S11) avec 100% d'échec en rythme
- Le meilleur : DID0168-S001R (S10) avec le meilleur bilan positif (73,7%)
- Le plus faible : DID0074-S001R (S11) avec le plus fort bilan négatif (73%)

CHAPITRE IV – Expérimentation de création de profils et de critères d'évaluation prosodiques d'apprenants

- Le « pianiste » : DID0078-S001R (S2) avec un bilan positif de 45,7% et un bilan négatif de 42% (bilan neutre très faible)

On remarque que l'apprenant DID0074-S001R (S11) est à la fois l'apprenant qui incarne le profil du plus fort bilan négatif global et de l'échec maximal en rythme.

La détection de la vitesse d'élocution peut avoir un grand intérêt, même lorsque la prosodie de l'anglais n'est pas primordiale. Il serait pourtant regrettable de penser que cette dimension de la parole fédère une qualité prosodique. On a pu observer à maintes reprises dans ce chapitre que ce n'était pas le cas. Au contraire, une grande rapidité d'élocution peut contribuer à masquer des faiblesses et freiner l'apprenant lors de son processus d'acquisition dans les autres domaines prosodiques, qui nécessitent une certaine durée pour se réaliser.

Ici, en évaluation sommative parmi les quinze profils étudiés, les deux plus pertinents sont ceux dont le bilan global est le plus élevé et plus faible tous domaines confondus :

- Meilleur apprenant en prosodie : DID0168-S001R (S10)
- Plus faible apprenant en prosodie : DID0074-S001R (S11)

Ces profils délimitent l'échelle des réalisations pouvant raisonnablement pouvoir être attendues de la part d'apprenants à un niveau donné, ici : celui des étudiants en première année de licence de spécialistes anglicistes. Une distribution des résultats évaluatifs des candidats est censée s'organiser tout au long de l'axe qualitatif borné par la moyenne de leurs réalisations. Ces profils permettent une hiérarchisation des résultats.

Dans une optique d'évaluation formative, il est pertinent de pouvoir repérer le degré de maîtrise des apprenants par rapport aux quatre domaines prosodiques essentiels. C'est ainsi que les autres profils deviennent à leur tour pertinents. Une fois renseigné, le système d'enseignement peut alors cibler le type de médiation ou de remédiation le plus adapté. Ces profils ne permettent pas de hiérarchisation des résultats, mais un classement en domaines prosodiques.

Il est souvent utile, par souci d'économie et de pertinence, en fait d'élégance de type mathématique, de réduire les variables inutiles au profit des plus influentes. Alors que les autres dénominations sont évidentes pour tous (par exemple, le profil le plus lent correspondant à un apprenant de faible niveau linguistique, aux syntagmes intonatifs courts et

aux pauses longues et nombreuses), il me semble nécessaire de m'appesantir sur la nature et l'intérêt représentés par le profil de S2, dit profil du « pianiste ».

Profil du « pianiste »

Le profil du « pianiste » est une notion abstraite, comme tout profil, et l'apprenant qui l'incarne est (ou a été) l'étudiant répertorié sous le code DID0078-S001R et que j'ai recodé S2 (*student #2*) dans mes graphiques. Comme on l'a vu, le « pianiste » est repérable par ses réalisations acoustiques. Il est relativement rapide, mais peu performant dans son registre, sa mélodie et son rythme. Contre toute attente, il se caractérise à la fois par son fort taux de réussite-bonus et d'échec-malus, laissant peu de place à des résultats moyens. On peut ajouter à sa description une caractéristique extralinguistique : il ne modifie pas sa voix pour incarner correctement les personnages de l'histoire. Je l'ai nommé en me référant à ses corrélats acoustiques de fréquence fondamentale. Les deux copies d'écran ci-dessous exemplifient sa courbe de F0 sur spectrogramme lors de la lecture des paroles les plus emphatiques.

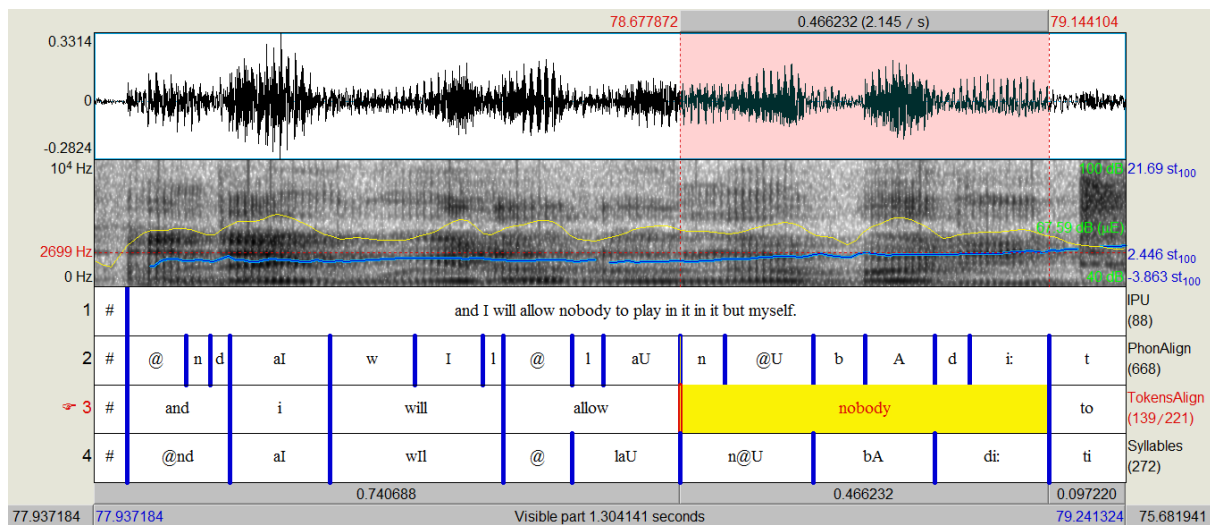


Figure 156 – Copie d'écran du spectrogramme et TextGrid Praat du « pianiste » DID0078-S001R pour l'emphase de *nobody* (Phrase 11)

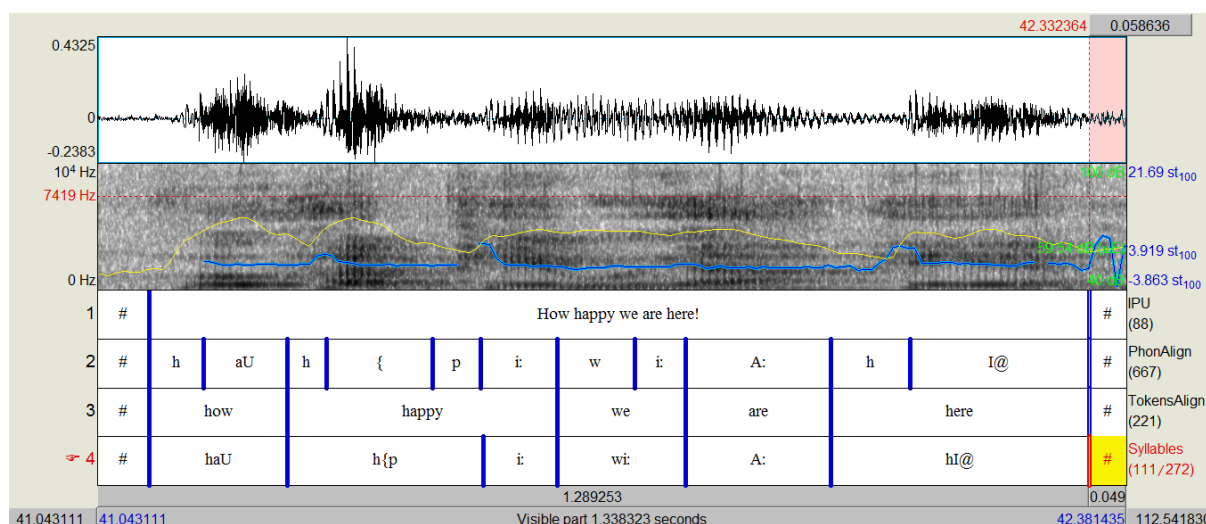


Figure 157 – Copie d'écran du spectrogramme et TextGrid Praat du « pianiste » DID0078-S001R pour l'emphase de l'exclamative (Phrase 5)

On peut constater dans ces deux copies d'écran la fréquence fondamentale horizontale et rectiligne de ses réalisations, lesquelles rappellent la trace acoustique de notes de piano sur spectrogramme, d'où la dénomination de ce profil.

Une étude que j'ai menée sur vingt apprenants du corpus *Longdale-Charliphonia* (dix de lecture rapide et dix de vitesse moyenne) a révélé des éléments intéressants lorsque leurs corrélats acoustiques en lecture ont été confrontés à la teneur de leur discours en conversation guidée. Un locuteur anglophone (statut de lecteur à l'université) finissait la conversation en demandant à l'étudiant pourquoi il avait choisi de faire des études d'anglais. J'ai observé que le profil de « pianiste » correspondait essentiellement aux étudiants qui mentionnaient spontanément leur attirance pour la lecture et parfois l'écriture. Le tableau suivant consigne mes résultats de l'expérimentation de repérage des « pianistes » parmi les apprenants du corpus *Longdale-Charliphonia* (Expérimentation REP-PIAN-NN⁹⁰).

⁹⁰ Expérimentation détaillée dans l'article « Intonational phrasing as a potential indicator for establishing prosodic learner profiles », dans S. Granger, G. Gilquin & F. Meunier (eds) (2013) *Twenty Years of Learner Corpus Research: Looking back, Moving ahead*. Corpora and Language in Use - Proceedings 1, Louvain-la-Neuve : Presses universitaires de Louvain, 75-88.

Caractéristiques mentionnées lors de la conversation	Nombre d'étudiants appartenant à la catégorie	Nombre de « pianistes » avérés dans la catégorie	Pourcentage de « pianistes » dans la catégorie
Amour de la lecture(et de l'écriture)	5	3	60%
Attirance pour la lecture	4	1	25%
Total des étudiants	20	4	20%

Tableau 76 – Résultats de l'étude du profil « pianiste » sur 20 apprenants de vitesse de lecture rapide et moyenne (expérimentation REP-PIAN-NN)

Les résultats ont montré que la tendance à correspondre au profil « pianiste » était proportionnelle au degré d'attirance vers la lecture et souvent l'écriture (adorer lire des romans, écrire des articles, etc.). Ces étudiants baignent dans la langue anglaise écrite au point que l'anglais oral n'est pas une référence pour eux. Par contre, lorsque un étudiant mentionnait qu'il adorait lire mais qu'il ajoutait sa propension à regarder des films en anglais ou discuter en anglais, j'ai constaté que les corrélats acoustiques se démarquaient alors de ceux d'un « pianiste ».

Parmi les quinze apprenants étudiés ici, deux d'entre eux correspondent à ce profil :

- DID0078-S001R (S2) : profil « pianiste » typique. Il fait partie de la catégorie du 1^{er} décile ; il se décrit bon en anglais, affirme adorer Oscar Wilde et avoir écrit en anglais ; il ne modifie pas sa voix pour l'apparenter à celle du Géant ou des enfants.
- DID0062-S001R (S7) : Elle fait partie de la Médiane ; elle lit beaucoup en anglais, surtout les romans d'Agatha Christie ; elle ne modifie pas sa voix pour l'apparenter à celle du Géant ou des enfants. Les copies d'écran ci-dessous confirment sa tendance à utiliser une intonation plate, même en zones d'emphase.

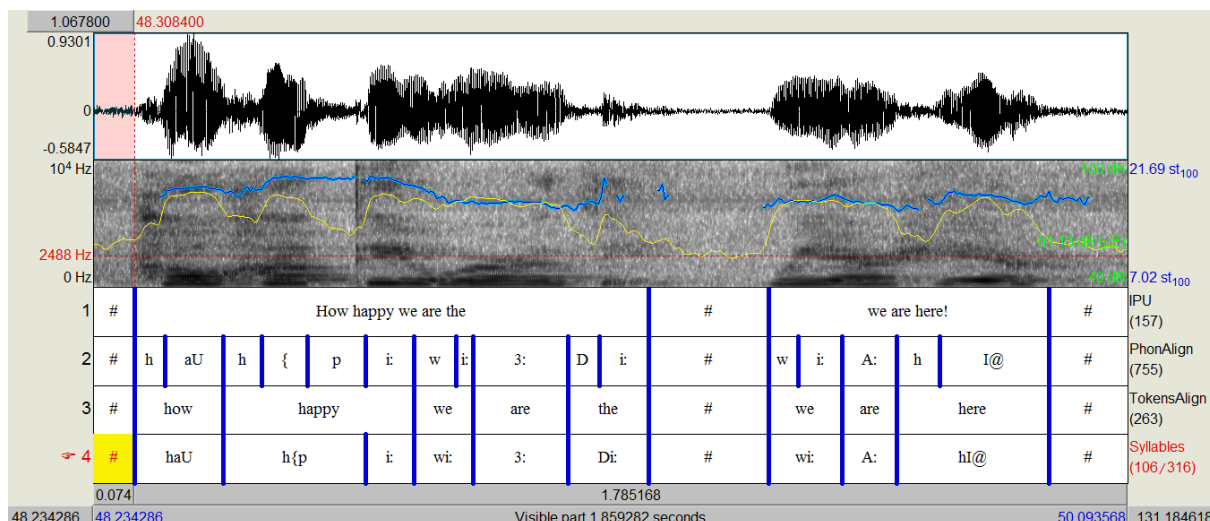


Figure 158 – Copie d'écran de spectrogramme et TextGrid Praat du « pianiste » DID0062-S001R (S7) pour l'emphase sur l'exclamative (Phrase 5)

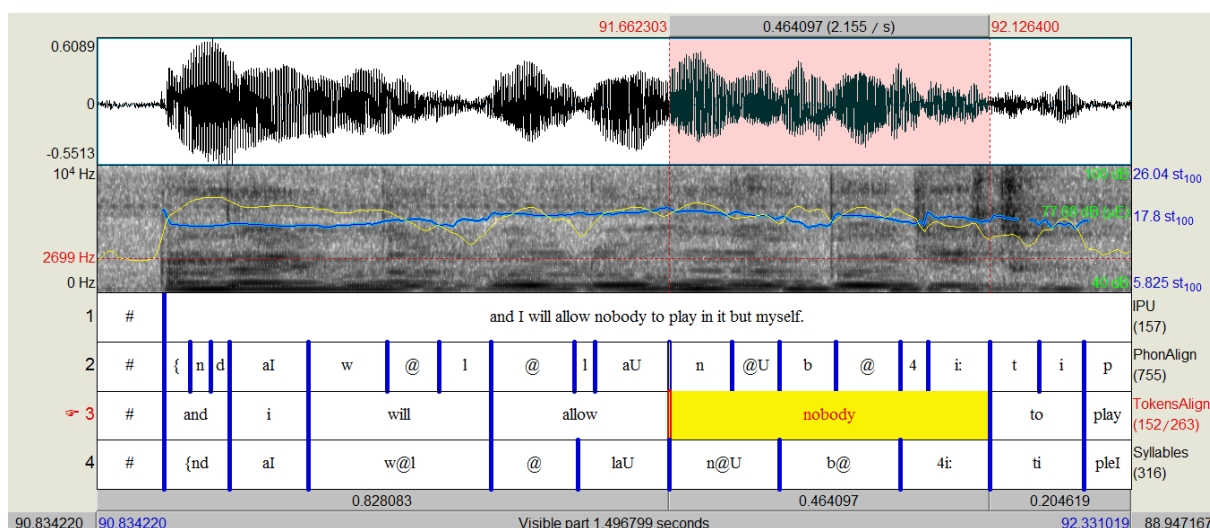


Figure 159 – Copie d'écran du spectrogramme et TextGrid Praat du « pianiste » DID0062-S001R (S7) pour l'emphase sur *nobody* (Phrase 11)

J'émets l'hypothèse que ce profil « pianiste » est le fruit d'un style d'apprentissage⁹¹ basé sur la culture de l'écrit. Ce profil ne peut se concrétiser que chez des apprenants d'un bon niveau d'anglais, qu'ils ont acquis grâce à une culture écrite forte, sans oral conséquent pour rééquilibrer la tendance à exclure les repères oraux de l'apprentissage linguistique. Les traits fédérateurs pour reconnaître le profil ne sont donc pas uniquement de nature linguistique. Ils peuvent néanmoins être repérés par une faible qualité spécifiquement orale

⁹¹ P. Cyr ([1996] 1998 : 83) : « Le style d'apprentissage est habituellement défini comme le mode préféré d'apprentissage de l'individu, sa façon globale d'approcher une tâche d'apprentissage. Il s'agit d'un concept très vaste qui peut englober de nombreux facteurs cognitifs ou socio-affectifs ».

(mélodie, registre, rythme, changement de voix) en opposition avec la vitesse de lecture et le débit par l'aisance que leur donne leur niveau d'anglais général.

Il me semble important de pouvoir diagnostiquer les symptômes caractéristiques de ce profil afin de pouvoir débloquer l'apprenant pour lui permettre de poursuivre son acquisition prosodique.

Après avoir étudié les profils stéréotypiques des apprenants et les apports qu'ils pouvaient fournir en évaluation formative et sommative, mais aussi en apprentissage, abordons à présent les critères d'évaluation en prosodie interlangue.

7 Analyse évaluative des données par profilage modalisateur de critères distinctifs (*criterial features*) : expérimentation CRIT-LEC-NN

Cette fin de chapitre aboutit à la création de *criterial features*. Elle en présente les objectifs et enjeux (section 7.1) avant de procéder au profilage de critères évaluatifs temporels par durée et débit (section 7.2), de critères d'évaluation selon le registre (section 7.3), la mélodie (section 7.4) et le rythme (section 7.5). Chaque section présente un prototype de grille évaluative des traits spécifiques avec suggestion de barème et d'appréciations.

7.1 Objectifs et enjeux

L'étude précédente sur les profils a confirmé les préconisations de Crystal : la caractérisation d'un profil prosodique nécessite une multitude de variables dont chacune est susceptible de rendre compte des traits pouvant rapprocher mais surtout distinguer le locuteur de ses congénères. Ce travail se veut précis et exhaustif, ce qui le rend extrêmement chronophage. Pour cette raison, et par souci de simplification, il est souhaitable de trouver des traits distinctifs à visée évaluative donnant des indications sans doute plus restreintes, mais fiables, sur la performance prosodique générale des apprenants.

Que les réalisations de l'apprenant proviennent d'une lecture de texte, d'une reformulation, d'une conversation spontanée ou plus ou moins guidée, il peut être difficile de trouver les marqueurs susceptibles d'éclairer l'évaluateur sur le niveau prosodique lorsque le support ne s'y prête pas. Aussi, c'est sur les quatre domaines prosodiques que je dresserai un bilan de ce qui peut paraître le plus révélateur.

7.2 Profils de critères d'évaluation temporels (par durée et débit)

L'analyse des réalisations selon un classement par vitesse de lecture a montré les limites d'un tel profilage. Il présente néanmoins des avancées intéressantes. Le graphique ci-dessous montre une comparaison visuelle de la durée du texte lu par l'intégralité des locuteurs étudiés dans le cadre de cette recherche (205 étudiants francophones et 42 natifs anglophones dont 26 natifs de moins de 40 ans, douze lecteurs de plus de 40 ans et quatre professionnels dont on ne connaît pas l'âge). On voit d'emblée que les jeunes lecteurs se démarquent fortement des deux autres catégories par leur rapidité de lecture générale. Il n'existe aucune intersection entre les deuxièmes et troisièmes quartiles de la boîte à moustaches des 26 jeunes lecteurs et de celle des douze lecteurs matures. Par contre, les quatre lecteurs professionnels s'insèrent quasiment dans ces quartiles, ce qui établit entre ces deux catégories une base de comparaison. Il faut néanmoins remarquer que la moustache supérieure des jeunes lecteurs s'approche de la médiane des lecteurs matures et que la moustache inférieure de ces derniers descend au-dessous de la médiane des jeunes lecteurs.

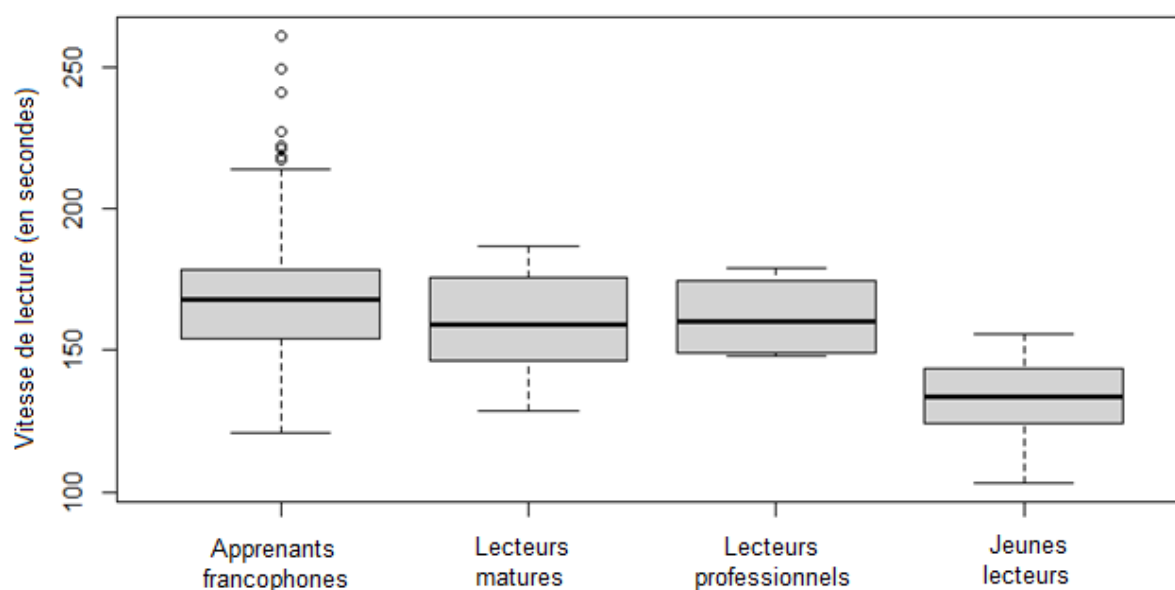


Figure 160 – Vitesse de lecture du texte par les catégories de lecteurs apprenants francophones, lecteurs anglophones jeunes, matures et professionnels (en secondes)

Cette représentation par boîtes à moustaches montre les similitudes temporelles des réalisations des apprenants avec les lecteurs anglophones matures et professionnels, mais peu avec les locuteurs de leur propre tranche d'âge. On remarque pourtant que certains apprenants se rapprochent de la vitesse de lecture des jeunes anglophones, se démarquant alors de celle des autres anglophones. Par contre, le haut de la boîte à moustache des apprenants montre que les plus lents n'entretiennent aucune correspondance avec la diction d'un anglophone. Cet

aspect explique que la fluidité d'élocution soit un critère de niveau fondamental pour un anglophone. Cela pourrait *a priori* avoir deux implications :

- Que les apprenants devraient se rapprocher de la norme constituée par les locuteurs natifs de leur groupe d'âge, dans la mesure où la qualité de lecture est compatible,
- Que, pour être acceptables, les corrélats acoustiques issus des variables autres que la durée devraient correspondre à celles des natifs selon leur vitesse de lecture (lecteurs matures et professionnels).

Pour satisfaire ces directives, le modèle natif optimal utilisé sera L5, la jeune américaine anglophone de 23 ans sélectionnée comme « bonne » lectrice à l'issue de l'expérimentation perceptive PER-LEC-N du Chapitre II sur les profils natifs, d'une rapidité de lecture moins typée que celle du « meilleur » lecteur L1 (56 ans). Pour établir la norme inférieure de limite d'acceptabilité, j'ai sélectionné le locuteur natif le plus lent. Il s'agit d'une femme de 52 ans (SG-UP-S3-f-52), dont une appréciation native avait comparé l'élocution à celle d'une maîtresse d'école.

Le tableau ci-dessous est un prototype de grille d'évaluation de la lecture du texte *The Selfish Giant* se basant, d'une part, sur les moyennes par catégories pour les apprenants, et d'autre part, sur les réalisations de la meilleure lectrice la plus lente (L5) et la lectrice la plus lente chez les 42 lecteurs natifs anglophones. Étant donné la grande rapidité de lecture du lecteur L1, il ne m'a pas semblé utile de tenir compte de ses résultats dans les grilles évaluatives pour apprenants, ceux de L5 étant largement suffisants pour indiquer la borne supérieure.

Locuteurs natifs		Apprenants		Barème	Appréciation
L5	135,3			$x < 140$	Excellent
		D1	143,4	$140 < x < 150$	Très bon
		Q1	155,1	$150 < x < 160$	Bon
		Méd	167,7	$160 < x < 175$	Moyen
SG-UP-S3-f-52	186,5	Q3	178,9	$175 < x < 190$	Passable
		D9	192,9	$x > 190$	Non acceptable

Tableau 77 – Grille évaluative de la durée de lecture du texte *The Selfish Giant* (en secondes) correspondant à la variable temporelle DUR_T (D1)

CHAPITRE IV – Expérimentation de création de profils et de critères d'évaluation prosodiques d'apprenants

Les résultats de ce tableau évaluatif sont uniquement destinés à évaluer les apprenants, étant donné que des lecteurs professionnels et très expressifs seraient qualifiés ici de « moyens ». La durée de la lectrice native la plus lente a été jugée adéquate pour constituer un seuil d'acceptabilité.

Étant donné que l'intérêt de la lecture du texte se réduit essentiellement aux apprenants qui liront ce même texte, l'analyse évaluative de la lecture des tronçons de textes dépourvus de ponctuation et où aucune pause n'est attendue devrait revêtir davantage d'intérêt par sa facilité de prise de mesures. De plus, les résultats obtenus pourraient sans trop de difficulté être transposés à d'autres styles de parole, tels que la reformulation et la conversation. Les prototypes de grilles ci-dessous concernent les mesures avec pauses éventuelles dans l'optique d'une plus grande facilité opérationnelle. La première grille cible la durée de lecture alors que la seconde se centre sur le nombre de syllabes prononcées par seconde. Les apprenants figurant dans les grilles exemplifient la variation des réalisations.

Locuteurs natifs		Apprenants		Barème	Appréciation
L5	7,99			$x \leq 7,8$	Excellent
		DID0210-S001R	8,481	$7,8 \leq x < 8,5$	Très bon
		DID0078-S001R	9,178	$8,5 \leq x < 9,2$	Bon
		DID0024-S001R	9,995	$9,2 \leq x < 10,1$	Moyen
SG-UP-S3-f-52	11,03	DID0062-S001R	10,998	$10,1 \leq x < 11,1$	Passable
		DID0076-S001R	13,720	$x \geq 11,1$	Non acceptable

Tableau 78 – Grille évaluative de la durée des tronçons du texte *The Selfish Giant* avec pauses (en secondes) correspondant à la variable temporelle DUR_ABCp (D2)

Locuteurs natifs		Apprenants		Barème	Appréciation
L5	5,75			$x > 5,5$	Excellent
		D1	5,27	$5,1 \geq x > 5,5$	Très bon
		Q1	4,68	$4,7 \geq x > 5,1$	Bon
		Méd	4,51	$4,4 \geq x > 4,7$	Moyen
SG-UP-S3-f-52	4,17	Q3	4,47	$4,2 \geq x > 4,4$	Passable
		D9	3,96	$x \leq 4,2$	Non acceptable

Tableau 79 – Grille évaluative du débit d'élocution des tronçons du texte *The Selfish Giant* avec pauses (nombre de syllabes par seconde) correspondant à la variable temporelle DEB_ABCp (D3)

Les trois tableaux ci-dessus fournissent des prototypes d'évaluation des réalisations des apprenants anglicistes francophones correspondant aux trois variables temporelles principales DUR_T (D1), DUR_ABCp (D2) et DEB_ABCp (D3).

7.3 Profils de critères d'évaluation selon le registre

Évaluer le degré d'ampleur du registre d'un locuteur représente plusieurs difficultés :

- Prise de mesure : alors que F0max sur une voyelle représente une certaine facilité d'exécution, celle de F0min est plus aléatoire en raison d'un brouillage relevant de diverses raisons dont une voix craquée ou une ligne de F0 irrégulière, voire inexistante,
- Sélection de la phrase : il est préférable de pouvoir établir une moyenne sur diverses phrases, mais l'idéal serait de trouver une phrase type donnant un reflet d'ensemble de la parole du locuteur sans avoir à prendre des mesures sur diverses phrases,
- Barème évaluatif : comme précédemment, proposer un prototype évaluatif.

Pour résoudre le premier problème de la prise de mesures, F0max me semble la prise de mesure idéale pour cibler les valeurs hautes. Par contre, afin de minimiser les erreurs, peut-être serait-il préférable de mesurer les valeurs basses avec F0mean ou F0min stabilisé (F0mean sur la portion de signal jugée la plus révélatrice).

À titre de comparaison, les mesures de F0 des natifs de référence figurent dans les grilles : il s'agit des deux versions de lecture de L1 (« meilleur lecteur », 56 ans) et de celle de L5 (profil similaire bien qu'âgée de 23 ans). Afin de correspondre exactement aux mesures des apprenants, lesquelles ont été effectuées dans *Praat*, elles ont été effectuées sur ce même logiciel, dans les mêmes conditions, à l'exception de celles qui, n'étant pas visibles dans *Praat*, n'ont pu être obtenues qu'en ayant recours à *WinPitchW10*, en variant les approches de détection de la fréquence fondamentale. Parmi les deux versions de lecture de L1, celle qui avait permis de déterminer les profils des natifs au Chapitre II, est la version 2. Alors que les données de L1 et L5 fournissent des points de repère inter-locuteurs parmi les meilleures lectures, la version 1 (L1v1) apporte un éclairage particulièrement pertinent sur la variation intra-locuteur qui peut exister chez un excellent lecteur.

Le graphique suivant permet de visualiser les résultats par phrase en indiquant les moyennes de F0max et F0min par phrase, tronçon de parole ou pour l'ensemble des quinze

CHAPITRE IV – Expérimentation de création de profils et de critères d'évaluation prosodiques d'apprenants

apprenants sélectionnés à partir du corpus. La limite supérieure de registre de S9 a été mesurée par F0mean sur la voyelle de la syllabe de F0 la plus élevée. La hauteur combinée correspond à F0max. On observe le peu d'écart entre les moyennes de F0max, la différence étant bien plus sensible dans celles de F0min. Cette différence montre l'importance d'obtenir une mesure de F0 basse fiable couplée avec une facilité d'exécution dans la prise de mesure.

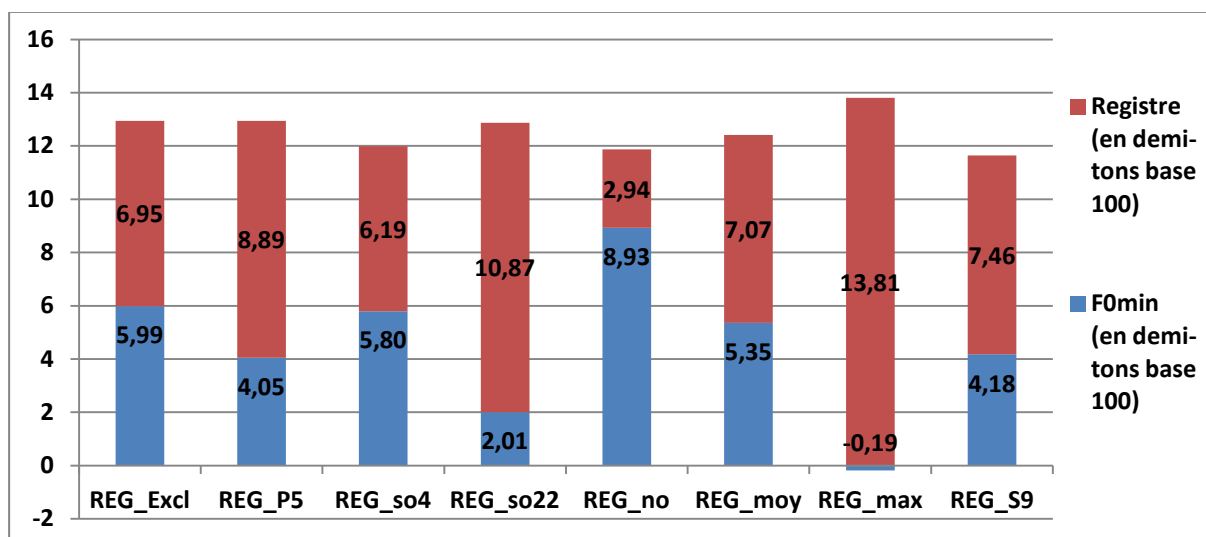


Figure 161 – Profil de registre pour chaque phrase ou entité étudiée, toutes catégories d'apprenants confondues (en demi-tons base 100)

Pour préciser l'analyse, le graphique ci-dessous montre le degré d'ampleur du registre pour chaque lieu de mesure et pour chaque catégorie d'apprenants basée d'après la vitesse de lecture. On voit que chez les apprenants du premier décile, le registre dans le tronçon de Phrase 22 est le plus élevé alors qu'ailleurs le registre est étroit. Dans le premier quartile, le registre de la Phrase 5 se rapproche de celui du tronçon de Phrase 22 ou lui est supérieur. Dans la catégorie Médiane, le registre de la Phrase 5 est de très loin supérieur à celui des autres mesures. Les apprenants de Q3 donnent une certaine équivalence au registre de P5 par rapport à celui des tronçons avec *so*, même si dans la moyenne P5 l'emporte légèrement. Quant aux réalisations du neuvième décile, le registre est étroit, comme l'indique l'homogénéité des mesures de F0max. Par ailleurs, il a été observé lors des prises de mesures que F0min est toujours inférieur à F0mean, notamment dans P4 (*them*) et S9 (*garden*).

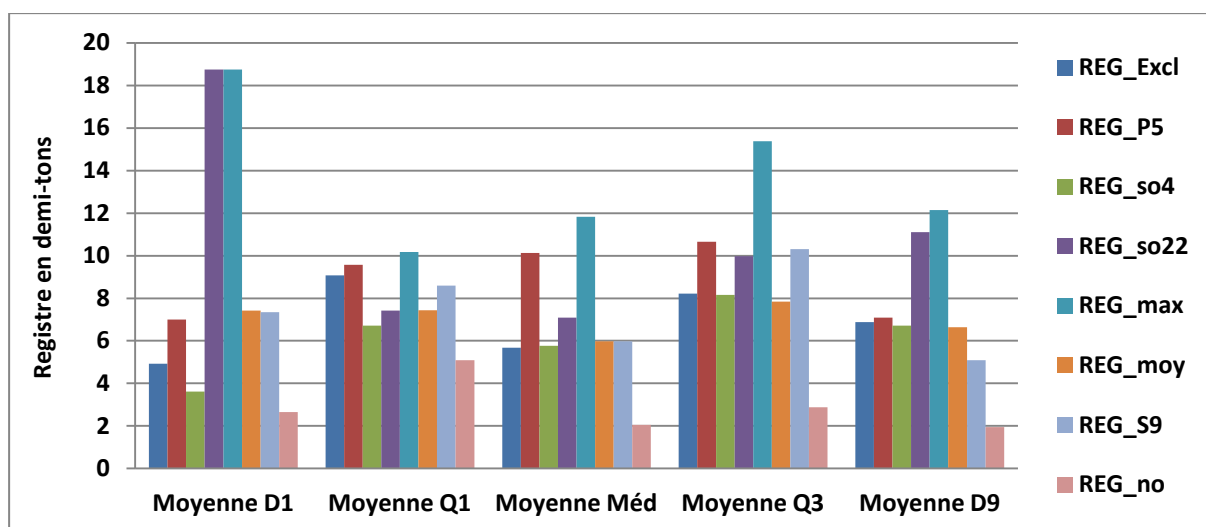


Figure 162 – Comparaison des moyennes de registres par entité étudiée et par catégorie d'apprenants selon la rapidité de lecture (en demi-tons base 100)

On peut déduire de cette analyse que REG_S9 (tronçon au taux réduit d'emphase) permet de jauger les réalisations de l'emphase, qui doivent lui être supérieures. On constate qu'elle n'est pas une variable discriminante de l'emphase qu'un apprenant peut distiller dans ses réalisations, comme l'indique sa stabilité relative par rapport aux autres lieux de mesure. Les catégories D1 et D9 révèlent toutes deux un registre étroit. REG_P5 concentre son ampleur dans les catégories Q1, Médiane et Q3. Par ailleurs, les résultats de registre semblent ne pas correspondre aux catégories de durée lorsque l'on examine les résultats individuels, sauf en ce qui concerne la catégorie D9. D'après ce graphique, REG_P5 semble plus régulière pour l'ensemble des catégories, et plus représentative de l'emphase que les variables REG_so4 et REG_so22. Aussi, on peut constater ici que c'est la Phrase 5 qui est la plus adaptée à l'évaluation du registre sur un temps de parole très réduit en tenant compte de l'exclamative qu'elle comprend ainsi que son énoncé citant. Le graphique suivant en propose un prototype d'évaluation.

Locuteurs natifs		Apprenants		Barème (DT)	Appréciation
L1v1	26,70			16 < x	Excellent
L1v2	17,92				
L5	11,44	DID0068-S001R	15,28	12 < x ≤ 16	Très bon
		DID0108-S001R	10,66	10 < x ≤ 12	Bon
		DID0003-S001R	8,32	8 < x ≤ 10	Moyen
		DID0078-S001R	6,78	6 < x ≤ 8	Passable
		DID0062-S001R	4,02	x < 6	Non acceptable

Tableau 80 – Grille évaluative critériée du degré d'ampleur du registre dans l'exclamative en *How* suivie de son énoncé citant (F0 en demi-tons sur base 100)

CHAPITRE IV – Expérimentation de création de profils et de critères d'évaluation prosodiques d'apprenants

L'analyse statistique par cercles de corrélations STAT-NN3 avait davantage mis en valeur le registre de l'exclamative seule alors que l'analyse du registre par archétypes de conformité STAT-NN2 avait favorisé l'intégralité de la Phrase 5 (exclamative + discours citant). Voici ci-dessous la grille évaluative correspondant à l'exclamative.

Locuteurs natifs		Apprenants		Barème (DT)	Appréciation
L1v1	24,00			13,5 < x	Excellent
L1v2	15,86				
		DID0168-S001R	12,23	11 < x ≤ 13,5	Très bon
L5	9,72	DID0024-S001R	9,47	8,5 < x ≤ 11	Bon
		DID0076-S001R	7,42	6 < x ≤ 8,5	Moyen
		DID0018-S001R	4,62	3,5 < x ≤ 6	Passable
		DID0078-S001R	2,20	x < 3,5	Non acceptable

Tableau 81 – Grille évaluative critériée du degré d'ampleur du registre dans l'exclamative en *How* sans énoncé citant (F0 en demi-tons sur base 100)

Il doit être noté que les repères de la partie supérieure de ces grilles correspondent aux corrélats acoustiques de « bons » lecteurs anglophones natifs, repérés lors de l'étalonnage effectué au Chapitre II (expérimentation PER-LEC-N). Ce n'est pas le cas des autres degrés, surtout du plus faible. Les degrés des divers niveaux regroupés en catégories évaluatives ont été choisis d'après les seuls résultats des quinze apprenants étudiés selon une répartition visuelle des quinze résultats acoustiques sur histogramme. Les seuils évaluatifs et leurs appréciations font figure d'exemples en fonction des données acoustiques disponibles. Ces tableaux ont valeur de prototypes de grilles d'évaluation, leur mérite résidant notamment dans la méthodologie mise en œuvre dans leur construction. Un autre point fort de ces grilles est leur point d'appui sur des corrélats acoustiques vérifiés et représentatifs de la population évaluée d'une part, et des natifs servant de modèles d'autre part. On constate par ailleurs que le meilleur apprenant réalise un registre plus ample que la jeune anglophone L5. Cet enchevêtrement des données acoustiques natives et non natives signale une gradation selon un continuum et réfute l'argument d'une catégorisation antinomique de la parole native par opposition à la parole non native.

7.4 Profilage de critères d'évaluation mélodiques

Alors que précédemment les mesures acoustiques, éventuellement couplées d'un nombre de syllabes, suffisaient à calculer une donnée pour chaque variable, il a fallu procéder à l'application d'une méthodologie différente afin d'aboutir à une seule valeur numérique

rendant compte de la qualité mélodique. Il s'agit de s'éloigner de la catégorisation selon la vitesse de lecture afin de trouver des similarités mélodiques sur lesquelles prendre appui afin de construire des critères évaluatifs opérationnels. Pour cela, chaque contour intonatif est présenté individuellement pour permettre une analyse des différences et similarités avec le modèle natif. Dans une deuxième phase, les résultats de cette analyse sont utilisés pour élaborer des formules de calcul dans le tableur *Excel*. La troisième phase montre le graphique des résultats obtenus grâce à l'application des formules sur les réalisations des apprenants et la quatrième phase est consacrée à la présentation du prototype de grille évaluative correspondant à la variable étudiée. Cette grille se base non seulement sur les réalisations des apprenants mais encore sur celles des natifs anglophones L1 et L5, formant en cela un continuum. Elle comprend ici aussi un exemple de notation et d'appréciation, bien que ces dernières soient destinées à être modifiées suivant le seuil d'acceptabilité correspondant au niveau requis.

Ce profilage a donné comme résultats les variables utilisées pour l'établissement des profils de apprenants en début de ce chapitre. C'est ici qu'est expliquée en détail la méthodologie qui a permis de les créer, basée sur un système de forfait initial sur lequel se greffent des points de bonus et de malus. La présentation de courbes normalisées facilite la comparaison visuelle mélodique par syllabe. Elle est dissociée en cinq groupes (catégorisation par vitesse de lecture) pour encore davantage préciser l'individualité de chaque contour.

7.4.1 Profilage évaluatif de chute mélodique emphatique (exclamative avec discours citant)

Le profilage des deux variables MEL_Excl (M1) et MEL_P5 (M2) est étudié conjointement puisqu'elles se complètent.

7.4.1.1 Phase 1 : Présentation des contours individuels et analyse

Les graphiques suivants permettent de visualiser la courbe intonative normalisée par syllabe des syllabes les plus importantes de la Phrase 5, c'est-à-dire de l'exclamative et de son discours citant. Je rappelle que les mesures ont été effectuées de la même manière que celles pour les natifs anglophones, c'est-à-dire pour mettre en évidence les particularités de la mélodie descendante et son registre maximal. Ainsi, F0max porte sur les trois premières syllabes, F0mean sur *we*, et F0min sur *are, here* et la syllabe la plus basse du discours citant.

CHAPITRE IV – Expérimentation de création de profils et de critères d'évaluation prosodiques d'apprenants

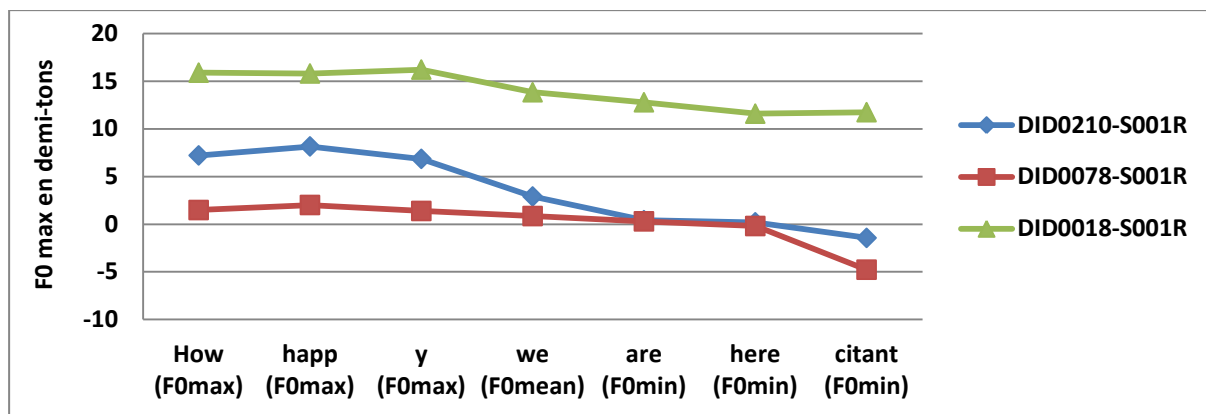


Figure 163 – Contour intonatif de la Phrase 5 (exclamative et discours citant) du groupe D1 (F0max en demi-tons)

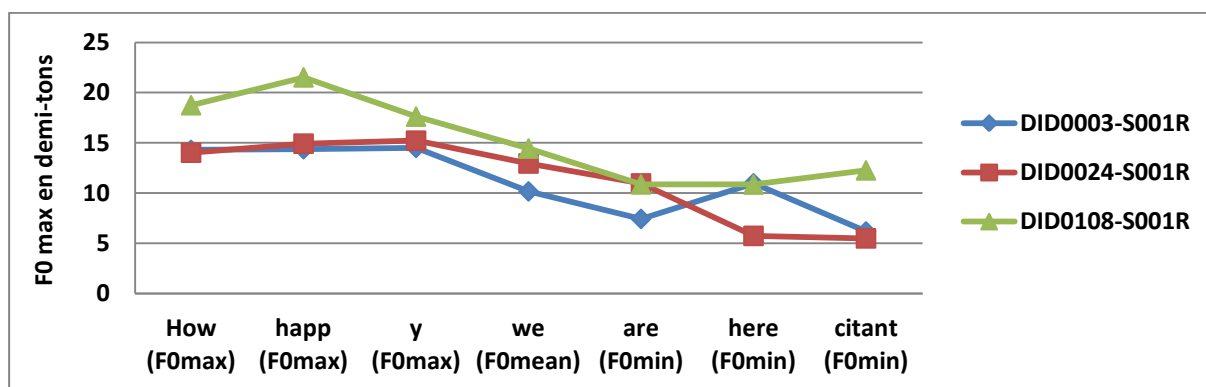


Figure 164 – Contour intonatif de la Phrase 5 (exclamative et discours citant) du groupe Q1 (F0max en demi-tons)

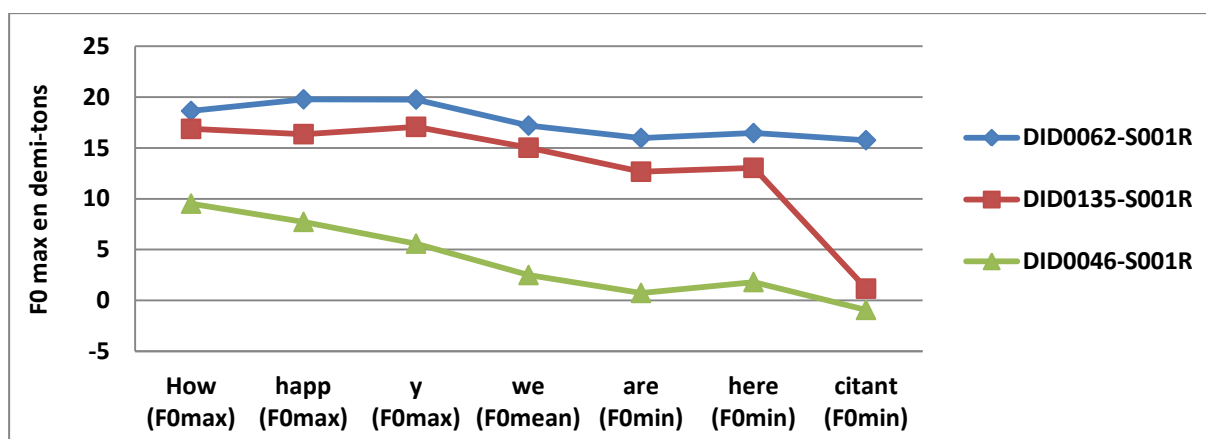


Figure 165 – Contour intonatif de la Phrase 5 (exclamative et discours citant) du groupe de la Médiane (F0max en demi-tons)

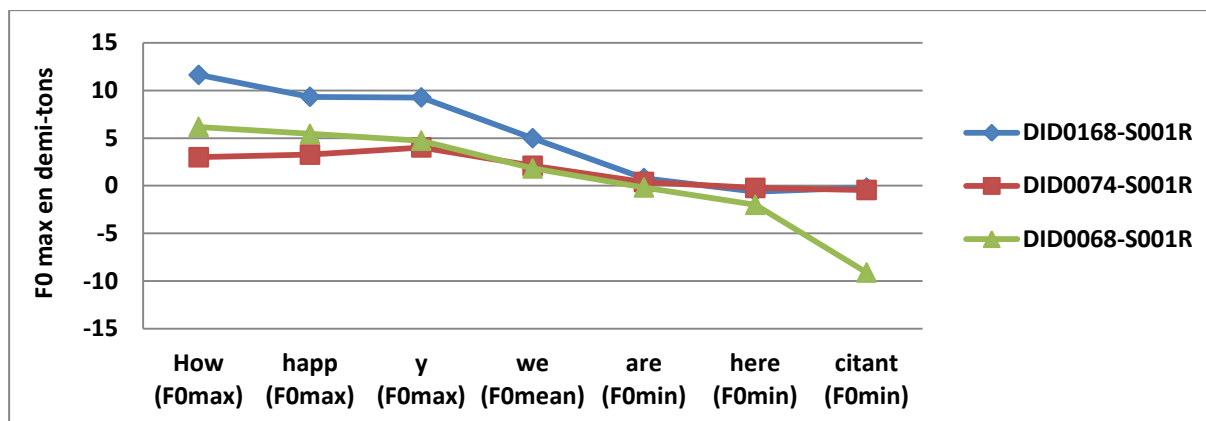


Figure 166 – Contour intonatif de la Phrase 5 (exclamative et discours citant) du groupe Q3 (F0max en demi-tons)

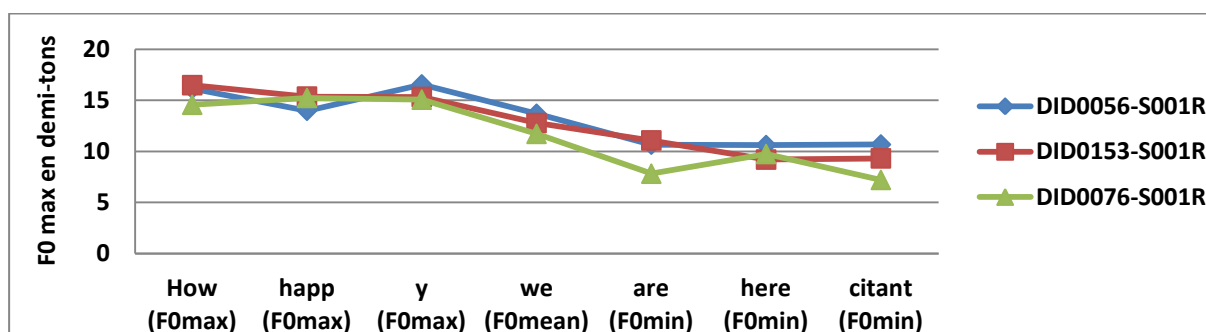


Figure 167 – Contour intonatif de la Phrase 5 (exclamative et discours citant) du groupe D9 (F0max en demi-tons)

Le caractère fondamental de l'exclamative, renforcé par son discours citant, est le degré d'ampleur de son registre car il ne peut y avoir de chute mélodique dans un registre trop étroit. Le dernier mot (*here*) se doit d'être celui à la F0 la plus basse, et non pas celui d'une remontée. La différence de F0 entre le dernier et l'avant-dernier mot de l'exclamative est alors cruciale. Le discours citant doit marquer une F0 encore plus basse.

On constate ensuite que certains apprenants marquent une baisse de F0 sur *happ-* alors que d'autres effectuent une montée. On a vu dans le Chapitre III sur l'emphase que cela relevait d'un choix personnel souvent inconscient. La différence de F0 entre les deux syllabes de *happy* est par contre révélatrice du niveau prosodique des apprenants : s'y exprime la tendance à accentuer la dernière syllabe d'un mot par une continuation, semblablement au français.

7.4.1.2 Phase 2 : Élaboration commentée des formules de calcul sur tableur

7. Analyse évaluative des données par profilage modalisateur de critères distinctifs (criterial features) : expérimentation CRIT-LEC-NN

CHAPITRE IV – Expérimentation de création de profils et de critères d'évaluation prosodiques d'apprenants

À partir des données observées et pertinentes dans une évaluation, des formules dans le tableur *Excel* ont permis de quantifier la qualité des réalisations mélodiques à partir des données acoustiques. Elles ont été construites selon quatre points d'ancrage.

Le calcul de la variable qualitative de l'exclamative repose préalablement sur son registre, établi par différence entre F0max et F0min. La formule posée pour les apprenants est la suivante :

$$=SI(AC15<3,5;1;SI(AC15<5;2;SI(AC15<6,5;3;SI(AC15<8;4;SI(AC15<9,5;5;SI(AC15<11;6;7))))))$$

Afin d'intégrer les natifs, l'échelle a dû être prolongée :

$$=SI(N22<3,5;1;SI(N22<5;2;SI(N22<6,5;3;SI(N22<8;4;SI(N22<9,5;5;SI(N22<11;6;SI(N22<13;7;SI(N22<15;8;9))))))$$

La cellule AC (ou N22 dans la version pour apprenants et natifs) correspond au registre. En fonction de son degré d'ampleur, un forfait de départ est octroyé à la variable. Ensuite, on cible le contour montant sur la deuxième syllabe de *happy* par rapport à la première (/i:/ - /h{p/}) grâce à la formule :

$$=SI(AK15<0;0;-2)$$

Elle signifie que l'on sanctionne de deux points une montée sur la deuxième syllabe⁹², lesquels sont éventuellement retranchés de la note donnée au départ pour le registre. Dans un troisième temps, on opère de manière similaire pour sanctionner un éventuel contour montant en fin d'exclamative (/hI@/ - /A:/ :

$$=SI(AL15>0;-2;0)$$

La variable numérique de la qualité mélodique est obtenue en additionnant les trois résultats obtenus.

⁹² Cette pénalité correspond à la tendance qu'ont les apprenants francophones à accentuer les mots sur la dernière syllabe en y marquant un ton montant en fin de syntagme intonatif. La très grande majorité des natifs procèdent à une descente sur cette syllabe inaccentuée. J'ai pourtant remarqué que certaines locutrices natives très expressives pouvaient marquer une remontée sur cette syllabe. De nouvelles lectures du même style de phrase par l'une d'elles ont confirmé cette remontée sans que ni elle, ni moi, ne puissions l'expliquer. Il s'agissait d'une anglophone britannique d'un accent non marqué. On retrouve ce choix de diction chez L5, ce qui fait artificiellement baisser sa note globale de deux points et modifie le classement par rapport aux apprenants. Cet état de chose souligne la difficulté de cadrer la variation dans toutes ses occurrences et la nécessité de recourir à une variété de critères pour fonder un jugement évaluatif.

Pour obtenir la variable qualitative de la Phase 5, on poursuit les calculs en prenant en compte F0min dans le discours citant. La remontée du discours citant par rapport à F0min de l'exclamative s'effectue par l'application de la formule :

$$=SI(X15>W15;-1;SI(X15<W15;1;0))$$

Ainsi, la variable qualitative de la Phase 5 correspond à l'application de la formule regroupant les quatre résultats, dont les trois premiers ont déjà été calculés en cellule AR :

$$=SOMME(AR16:AS16)$$

La différence de hauteur mélodique entre les deux premières syllabes (*How happ-*) n'a pas été prise en compte parce qu'elle ne correspond pas à une maîtrise mélodique, mais davantage à un choix personnel suivant l'image de lui-même que le locuteur veut donner à son entourage.

7.4.1.3 Phase 3 : Visualisation des résultats obtenus et analyse

À partir de ces données, le graphique suivant a pu être élaboré.

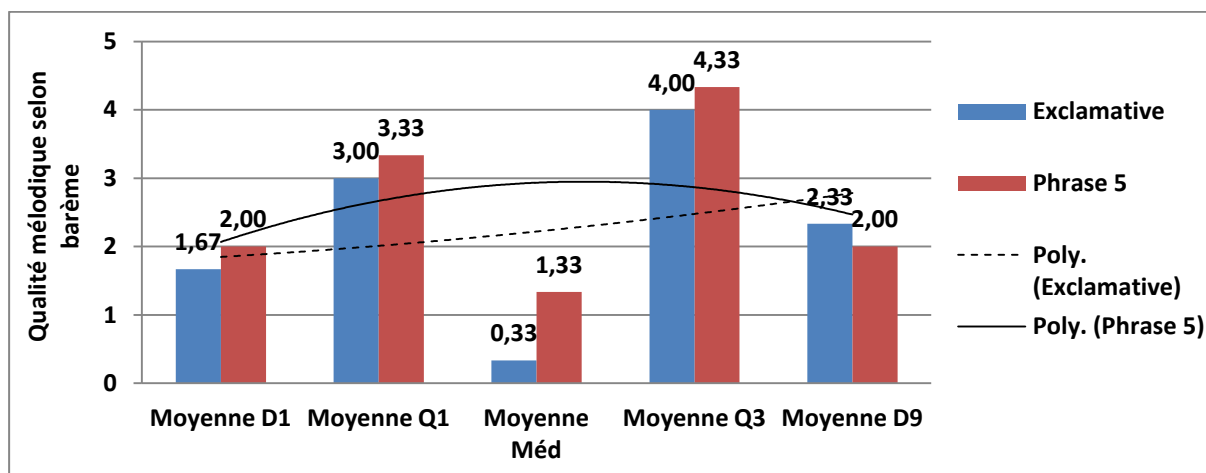


Figure 168 – Qualité mélodique de l'exclamative et de la Phrase 5 calculée par formules critériées (barème qualitatif)

Ce graphique présente des similarités avec celui de la comparaison de l'ampleur du registre dans l'exclamative et avec adjonction du discours citant (Phrase 5) par catégorie d'apprenants. Les résultats de la Médiane sont bien plus faibles, comme on le constate dans plusieurs autres graphiques. On constate néanmoins que la courbe polynomiale est convexe dans sa représentation de la Phrase 5 alors qu'elle est concave dans l'exclamative. Cet aspect conforterait le choix de la Phrase 5 pour évaluer pleinement les réalisations mélodiques et de registre de l'exclamative.

7.4.1.4 Phase 4 : Résultat du profilage par prototype de grille évaluative avec exemple de notation et d'appréciation

Les prototypes de grilles évaluatives sont proposés ci-dessous.

Locuteurs natifs		Apprenants		Barème	Appréciation
L1v1	9			$x \leq 8$	Excellent
		DID0168-S001R	7	$6 \leq x < 8$	Très bon
L5	4	DID0153-S001R	4	$4 \leq x < 6$	Bon
		DID0076-S001R	2	$2 \leq x < 4$	Moyen
		DID0074-S001R	0	$0 \leq x < 2$	Passable
		DID0135-S001R	-2	$x < 0$	Non acceptable

Tableau 82 – Grille évaluative critériée de la qualité mélodique de l'exclamative en *How* sans son énoncé citant

La F0 de la première syllabe de *happy* n'ayant pas pu être mesurée (dévoisement complet par voix soufflée), la notation de L1v2 n'a pu être effectuée, mais l'écoute combinée avec l'examen du spectrogramme invitent à croire que la notation aurait été exactement la même dans les deux versions (L1v1 et L1v2), que ce soit dans la grille mélodique de l'exclamative ou de la Phrase 5.

Locuteurs natifs		Apprenants		Barème	Appréciation
L1v1	10			$8 \leq x$	Excellent
		DID0068-S001R	6	$6 \leq x < 8$	Très bon
L5	5	DID0210-S001R	5	$4 \leq x < 6$	Bon
		DID0076-S001R	3	$2 \leq x < 4$	Moyen
		DID0074-S001R	1	$0 \leq x < 2$	Passable
		DID0135-S001R	-1	$x \leq 0$	Non acceptable

Tableau 83 – Grille évaluative critériée de la qualité mélodique de l'exclamative avec son énoncé citant

7.4.2 Profilage évaluatif de chute mélodique non emphatique (syntagme intonatif S9)

Ce prototype évaluatif est destiné à vérifier la conformité d'un contour descendant général sur syntagme intonatif à l'emphase réduite (S9). Il correspond à la variable MEL_S9 (M3).

7.4.2.1 Phase 1 : Présentation des contours individuels et analyse

Les contours intonatifs qui suivent montrent les réalisations individuelles des quinze apprenants étudiés regroupés selon la catégorisation par vitesse de lecture, ce qui permet de visualiser clairement chacune d'entre elles.

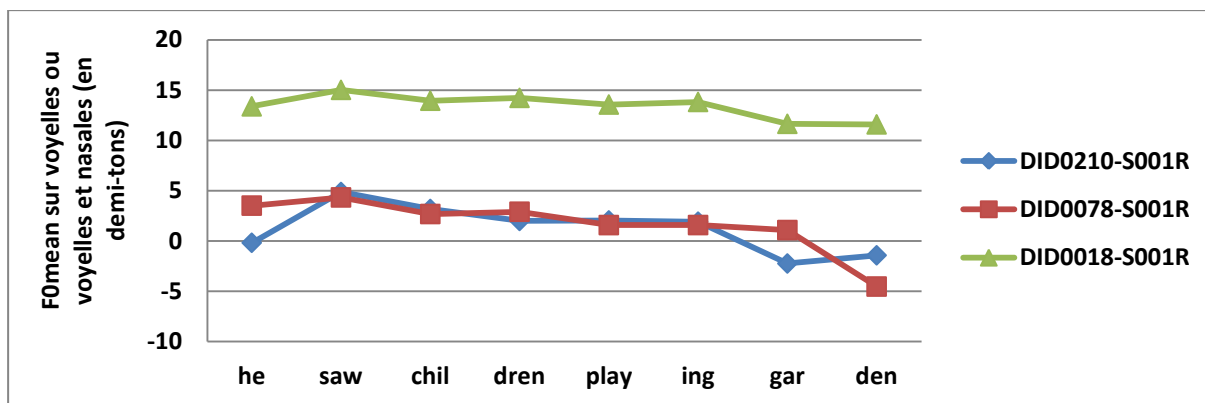


Figure 169 – Contour intonatif du syntagme intonatif S9 du groupe D1 (F0mean en demi-tons)

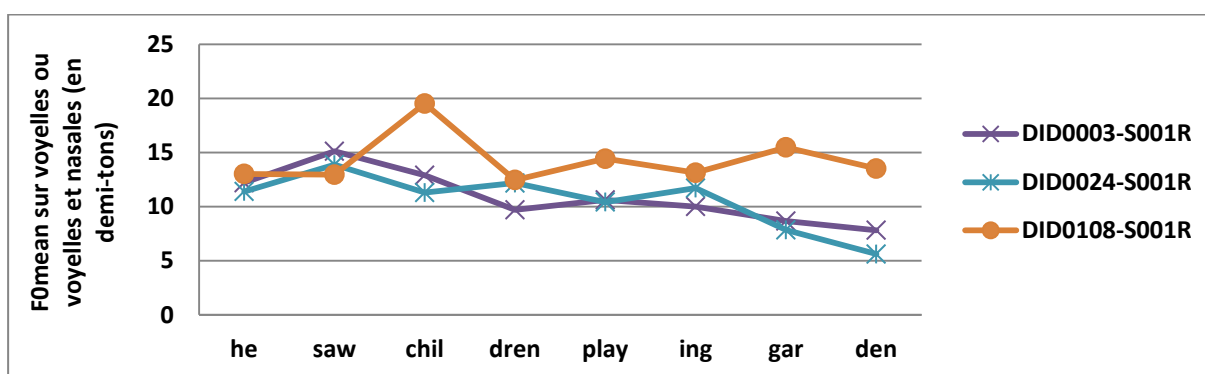


Figure 170 – Contour intonatif du syntagme intonatif S9 du groupe Q1 (F0mean en demi-tons)

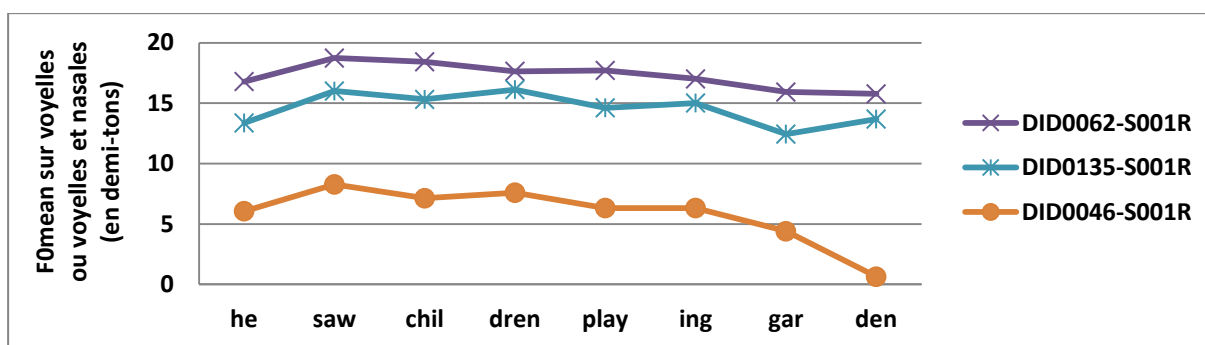


Figure 171 – Contour intonatif du syntagme intonatif S9 du groupe de la Médiane (F0mean en demi-tons)

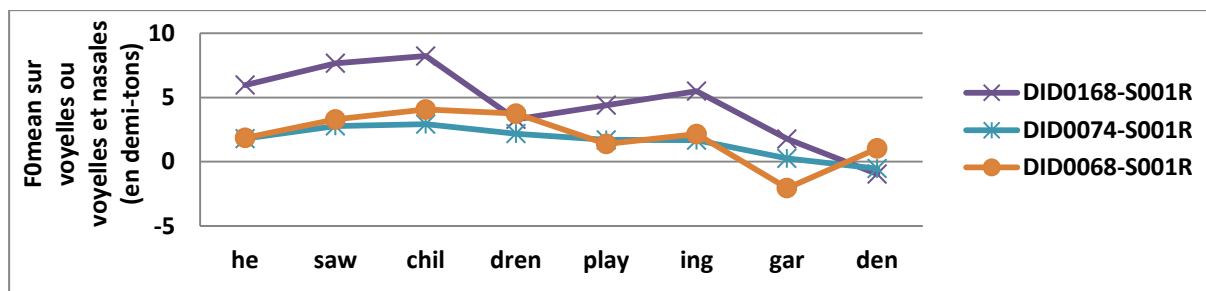


Figure 172 – Contour intonatif du syntagme intonatif S9 du groupe Q3 (F0mean en demi-tons)

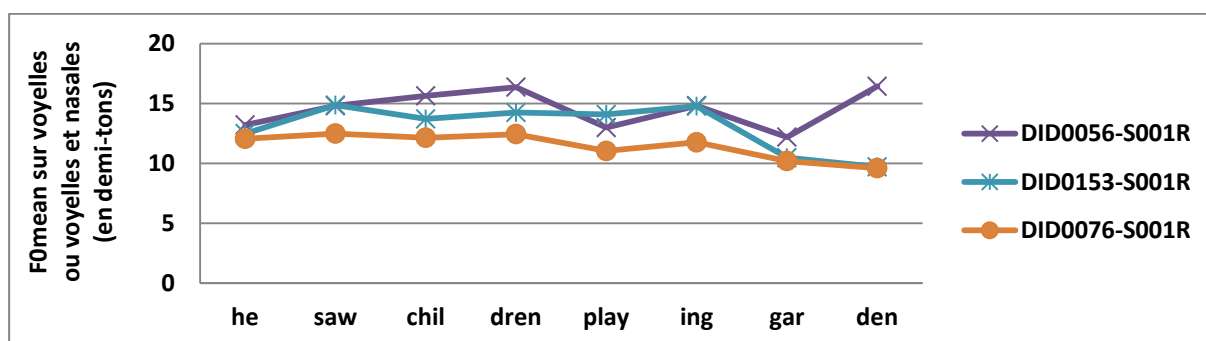


Figure 173 – Contour intonatif du syntagme intonatif S9 du groupe D9 (F0mean en demi-tons)

Ces graphiques se concentrent sur les points sensibles chez les apprenants francophones :

- Avant-tête devant être plus basse que la tête (*he saw*),
- Syllabe inaccentuée ne devant pas être caractérisée par une montée (*children, playing*),
- Mot en fin de syntagme devant être marqué par un ton avec F0 plus basse avec syllabe inaccentuée post-noyau non remontante (*garden*).

7.4.2.2 Phase 2 : *Élaboration commentée des formules de calcul dans le tableur Excel*

Les graphiques précédents ont permis de repérer les quatre écueils à éviter par les apprenants et qui sont évalués par formules de calcul dans *Excel* à partir d'un forfait correspondant au degré d'ampleur du registre. Les différentes étapes du profilage sont expliquées ci-dessous.

Pour mesurer le registre, on a eu recours à F0mean sauf pour la syllabe de F0 la plus basse, qui a été mesurée par F0min afin de ne pas fausser le résultat en cas de remontée en fin de syntagme intonatif. La formule utilisée pour noter le registre est :

$$=SI(AM15<2,5;0;SI(AM15<3,5;1;SI(AM15<4,5;2;SI(AM15<5,5;3;SI(AM15<6,5;4;SI(AM15<7,5;5;SI(AM15<8,5;6;7))))))))$$

AM est la cellule où figure le registre. L'apprenant consigne un point par demi-ton pour constituer un forfait à partir duquel il sera pénalisé pour les erreurs commises ou récompensé suivant les pièges évités, selon une optique de bonus-malus.

Ainsi, dans *he saw*, l'apprenant engrange un ou deux, voire trois points, si la F0 de la tête *saw* monte de deux ou quatre demi-tons ou davantage par rapport à l'avant-tête *he* :

$$=SI(AS15<0,5;0;SI(AS15<2;1;SI(AS15<4;2;3)))$$

Dans les trois cas suivants, on peut observer l'aptitude de l'apprenant à ne pas remonter sa F0 sur la deuxième syllabe d'un mot accentué sur la première syllabe. Cette erreur, typique du français, a déjà été remarquée dans la réalisation de *happy* dans l'exclamative. AT, AU et AW représentent le calcul de la différence de F0 entre la première et la deuxième syllabe. Les formules utilisées sont :

- Children : =SI(AT15>0;-1;0)
- Playing : =SI(AU15>0;-1;0)
- Garden : =SI(AW15<=0;0;-2)

Dans le cas de *garden*, la sanction est double parce qu'il porte aussi le noyau du syntagme. Pour cette nouvelle raison, une remontée doit être évitée. La conformité du contour général descendant sur S9 est calculée par addition des résultats :

$$=SOMME(AX15:BB15)$$

7.4.2.3 Phase 3 : Visualisation des résultats obtenus et analyse

Les données recueillies par notation qualitative à partir du calcul de ces formules sont détaillées dans le graphique ci-dessous, puis schématisées davantage dans celui qui le suit.

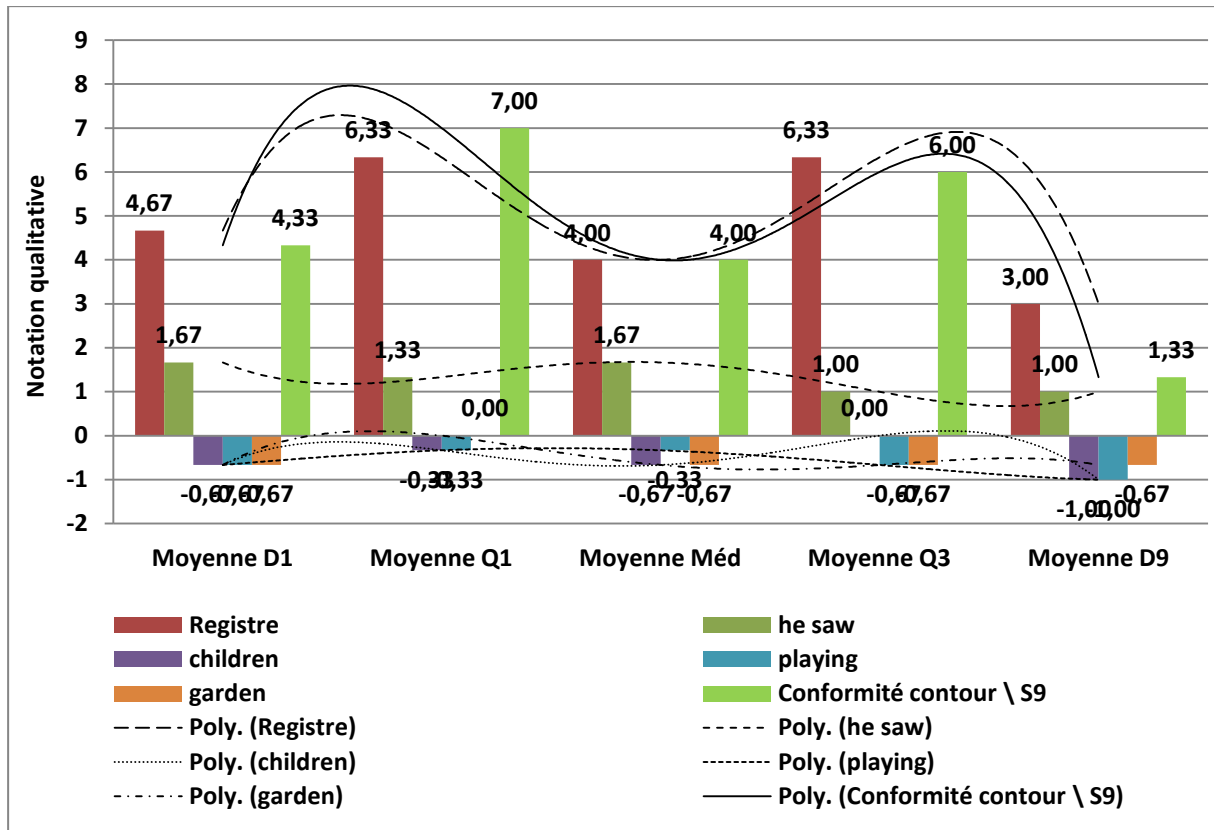


Figure 174 – Conformité du contour descendant général sur le syntagme intonatif neutre S9 par barème qualitatif (histogramme détaillé)

On y voit que les points ajoutés forment des données positives et ceux retranchés, des données négatives. La qualité mélodique est en correspondance avec le registre, à l'exception de la catégorie D9, laquelle se caractérise par une plus forte tendance à remonter sa F0 en fin de mot. Le décrochage de F0 sur la tête *saw* est observé dans toutes les catégories, surtout les plus rapides (D1, Q1 et Médiane). Ces caractéristiques sont résumées dans le graphique qui suit.

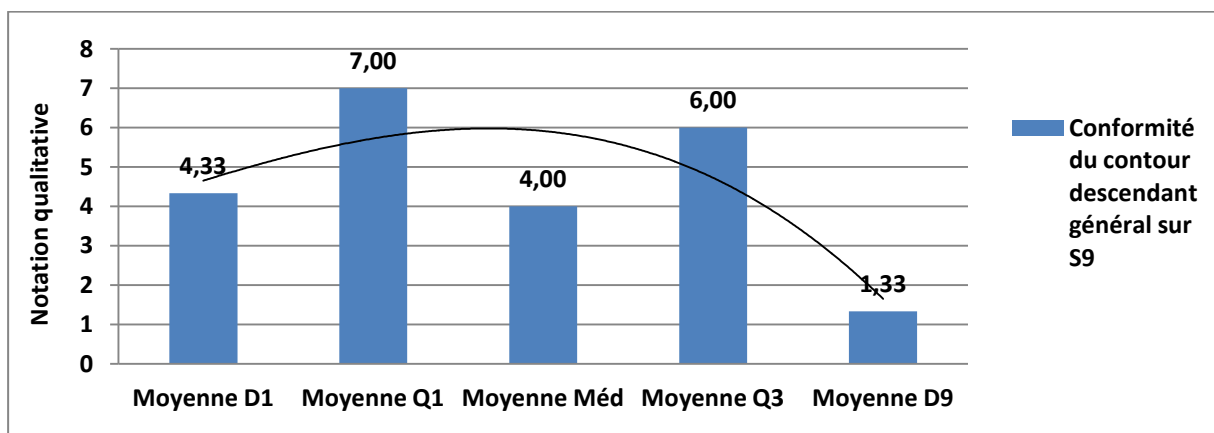


Figure 175 – Conformité du contour descendant général sur le syntagme intonatif neutre S9 par barème qualitatif (histogramme résumé)

Le graphique ci-dessus confirme que les meilleures réalisations mélodiques ne se trouvent ni en D1 ni en D9, mais au contraire en Q1 et Q3. Le groupe le plus faible est nettement D9. Ici encore, on remarque les faibles résultats de la Médiane.

Par ailleurs, le classement par application des formules d'évaluation montre que les apprenants de D9 sont faibles ou moyens et que ceux de la Médiane se répartissent à tous les niveaux.

7.4.2.4 Phase 4 : Résultat du profilage par prototype de grille évaluative avec exemple de barème et d'appréciations

Le prototype de grille évaluative est présenté ci-dessous.

Locuteurs natifs		Apprenants		Barème	Appréciation
				$9 < x$	Excellent
L1v1	9	DID0003-S001R	9	$7 < x \leq 9$	Très bon
L1v2	8				
		DID0168-S001R	7	$5 < x \leq 7$	Bon
		DID0153-S001R	5	$3 < x \leq 5$	Moyen
		DID0068-S001R	3	$1 < x \leq 3$	Passable
		DID0056-S001R	-1	$x \leq 1$	Non acceptable

Tableau 84 – Grille évaluative critériée de la qualité mélodique du syntagme intonatif neutre S9

L'intégralité des mesures n'a pas pu être effectuée sur L5, d'où l'absence de ses résultats sur la grille évaluative. Néanmoins, les données obtenues donnent à penser qu'elle aurait obtenu la note de 3. Par ailleurs, l'absence d'une prestation « excellente » peut s'expliquer par l'absence de phénomène d'emphase. L'hypothèse H3 selon laquelle l'emphase révélerait une variabilité maximale prend tout son sens ici aussi : l'éventail de réalisations possibles et effectives par amplification du registre pourrait être réduit dans ce syntagme d'où l'emphase est quasiment absente. Mais en conséquence, il faudrait reconnaître l'incapacité du syntagme intonatif S9 à discriminer une qualité mélodique maximale, ce qui l'écarte des critères distinctifs révélateurs de niveau dans une évaluation prosodique.

7.4.3 Profilage tonal évaluatif descendant-montant (∨) sur syllabe devant frontière mineure (« arrived »)

La variable MEL_s∨ (M5) est destinée à évaluer la capacité d'un apprenant à réaliser un ton descendant-montant (*fall-rise*) sur une syllabe unique. Cette réalisation est attendue sur la syllabe accentuée de *arrived*, placée devant une frontière mineure.

7.4.3.1 Phase 1 : Présentation des contours individuels et analyse

Les graphiques suivants détaillent les réalisations normalisées des quinze apprenants par catégorisation de vitesse de lecture.

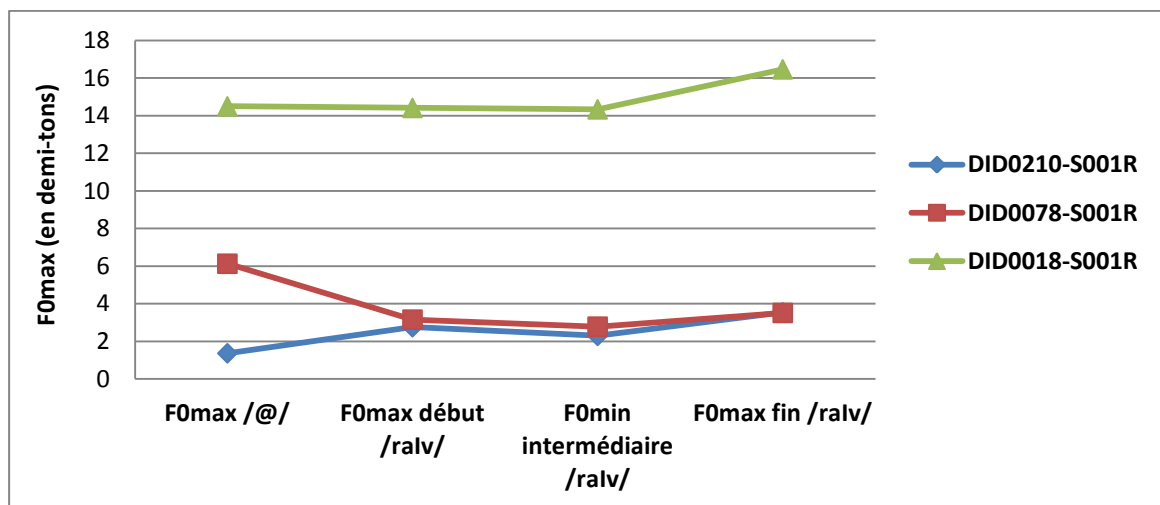


Figure 176 – Schématisation normalisée du ton descendant-montant (∨) sur une syllabe devant frontière mineure (*arrived*) dans le groupe D1 (F0max en demi-tons)

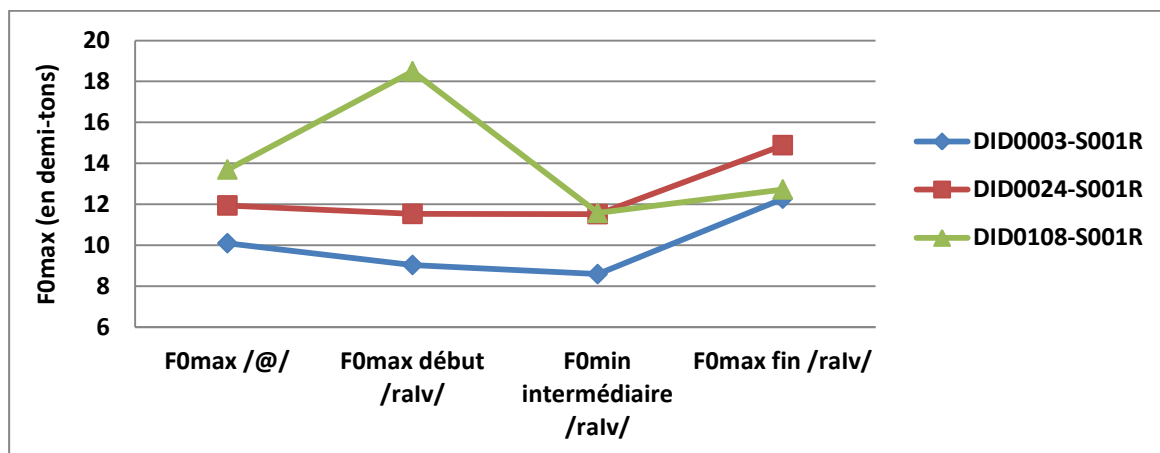


Figure 177 – Schématisation normalisée du ton descendant-montant (∨) sur une syllabe devant frontière mineure (*arrived*) dans le groupe Q1 (F0max en demi-tons)

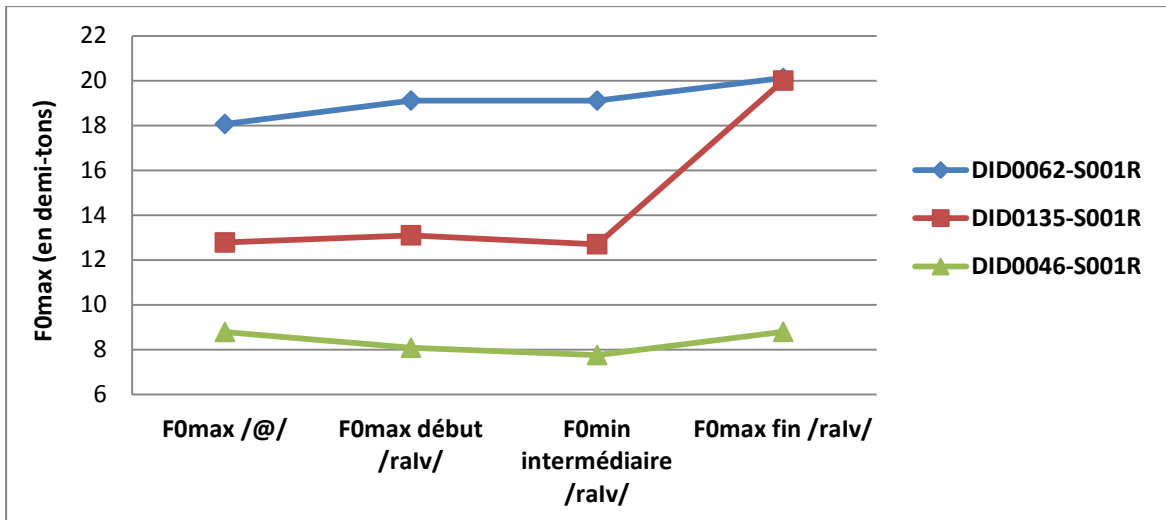


Figure 178 – Schématisation normalisée du ton descendant-montant (∨) sur une syllabe devant frontière mineure (*arrived*) dans le groupe Médiane (F0max en demi-tons)

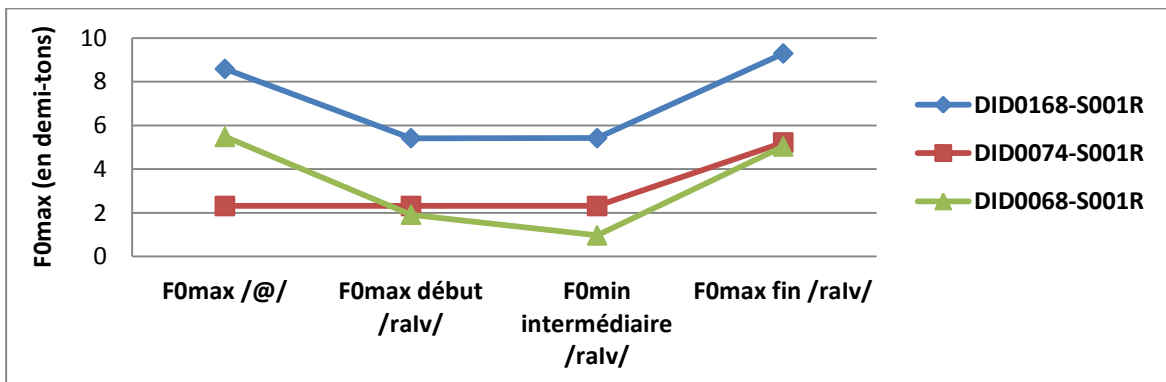


Figure 179 – Schématisation normalisée du ton descendant-montant (∨) sur une syllabe devant frontière mineure (*arrived*) dans le groupe Q3 (F0max en demi-tons)

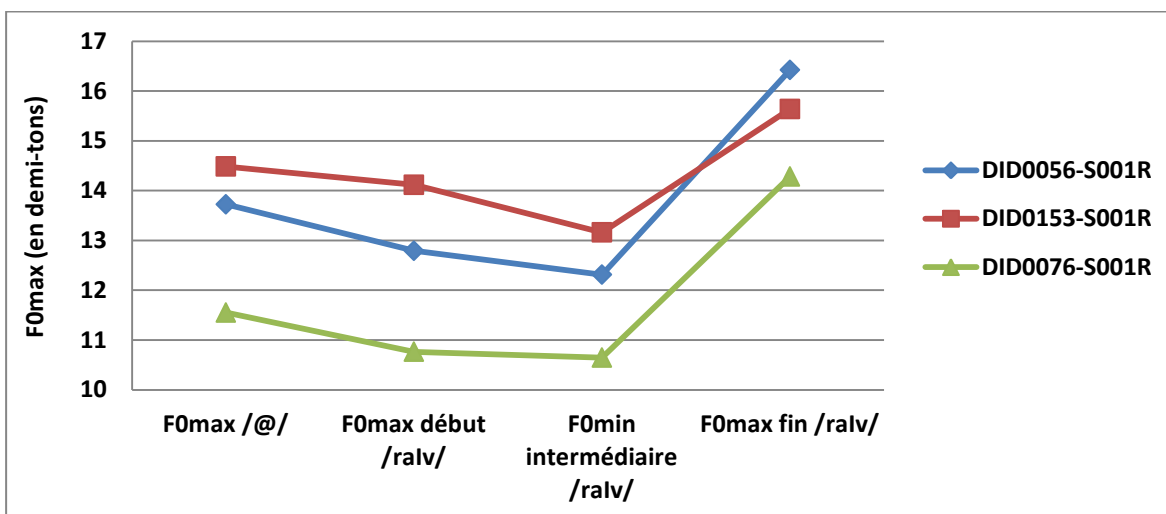


Figure 180 – Schématisation normalisée du ton descendant-montant (∨) sur une syllabe devant frontière mineure (*arrived*) dans le groupe D9 (F0max en demi-tons)

CHAPITRE IV – Expérimentation de création de profils et de critères d'évaluation prosodiques d'apprenants

Ces réalisations confirment une fois de plus la constatation de Horgues (2010) sur la difficulté que rencontrent les apprenants et même les anglophones natifs à réaliser ce ton syllabique dans son intégralité. En effet, il est difficile de bien réaliser un mouvement du ton sans le faire au détriment de l'autre partie : c'est d'ailleurs ce que l'on avait pu constater chez L1, le « meilleur » lecteur. Ici, l'on voit que les deux parties du ton bidirectionnel, montée et descente, ont avantage à être dissociées l'une de l'autre dans une évaluation afin de faciliter l'analyse, puisque les résultats tendent à se compenser à terme. Les critères mis en œuvre ne tiennent compte ni de facteurs de durée, ni de pente de F0, mais uniquement de la hauteur mélodique. Une simplification extrême prévaut pour capter les lignes directrices dans la conception d'un prototype qui, comme les autres, devra être affiné puis complexifié par la suite pour rendre compte des réussites ou des échecs. En effet, les phénomènes prosodiques ne peuvent que bénéficier d'une approche holistique : tempo, registre et mélodie sont étroitement liés. En conséquence, le tempo influe sur le registre, lequel impacte à son tour la mélodie. Pour cette raison, des calculs sur la F0 seule peuvent se dispenser d'autres mesures dans une phase de profilage rudimentaire tout en cernant l'essentiel. Ainsi, une forte montée implique que le locuteur évolue dans un registre ample, ce qu'il ne peut faire sans disposer du temps nécessaire à ce changement de F0. Seuls les locuteurs natifs experts (comme L6 du corpus de natifs, au Chapitre II) peuvent se permettre une telle virtuosité.

Chez les apprenants, on s'attend à ce que la partie tonale descendante de /raIv/ soit sacrifiée au détriment de la montée. En d'autres termes, on peut s'attendre à ce que la montée soit favorisée car la descente est plus facilement réalisable sur la syllabe précédente.

7.4.3.2 Phase 2 : *Élaboration commentée des formules de calcul dans le tableur Excel*

La formule de calcul pour évaluer la descente est :

```
=SI(W15>=0;0;SI(W15>-0,3;1;SI(W15>-0,9;2;SI(W15>-1,5;3;SI(W15>-3;4;SI(W15>-5;5;6))))))
```

Elle considère la différence entre F0max et F0min sur /raIv/. La portion tonale montante de /raIv/ correspond à la formule :

```
=SI(X15<=0;0;SI(X15<1,5;1;SI(X15<3;2;SI(X15<4,5;3;SI(X15<6;4;SI(X15<7,5;5;6))))))
```

La qualité mélodique finale est obtenue en ajoutant les deux résultats précédents :

```
=SOMME(Y15:Z15)
```

7.4.3.3 Phase 3 : Visualisation des résultats obtenus et analyse

Le résultat de ces calculs ainsi que son rendu détaillé figurent dans le schéma *Excel* suivant.

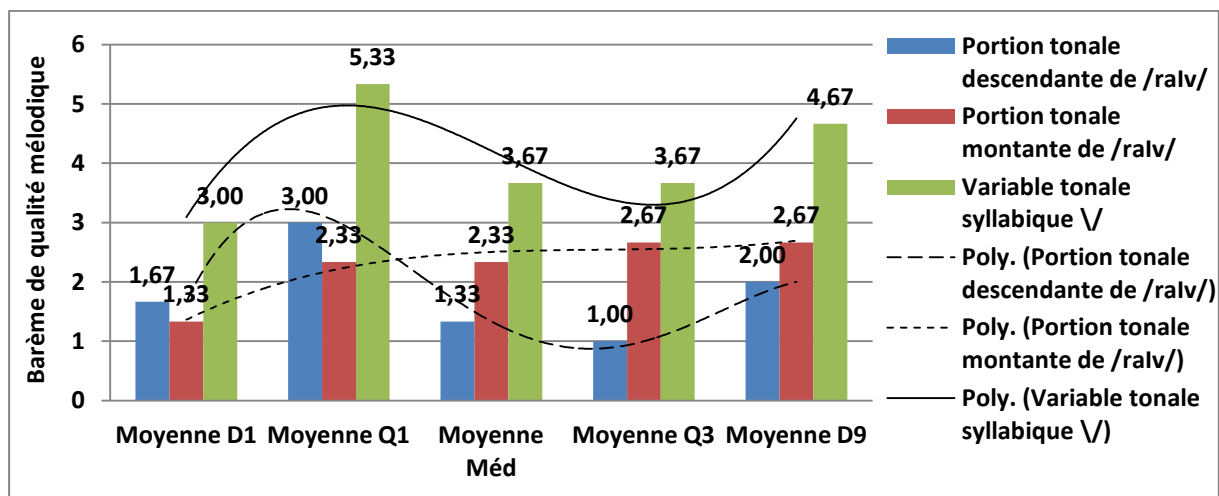


Figure 181 – Schéma détaillé de la réalisation du ton bi-directionnel V dans *arrived* selon un barème qualitatif

On constate que ces résultats sont peu conformes à la catégorisation par vitesse de lecture (surtout pour D9), à l'exception de la portion tonale montante, qui est très stable. On avait d'ailleurs remarqué que les quinze apprenants réalisaient une montée en fin de ton. C'est surtout la portion tonale descendante qui différencie les apprenants : la courbe polynomiale de la portion descendante évoque une similarité avec celle de la variable tonale syllabique dans son intégralité. On observe en outre que D1 et Q1 privilégient la descente alors que les autres catégories s'investissent davantage dans la montée. Les meilleurs résultats sont obtenus par les catégories Q1 et D9, ce qui démontre une fois de plus le caractère quelque peu orthogonal de la rapidité de lecture par rapport à la qualité mélodique.

7.4.3.4 Phase 4 : Résultat du profilage par prototype de grille évaluative avec exemple de barème et d'appréciations

La grille d'évaluation proposée ci-dessous permet un premier classement. On avait observé au Chapitre III les réalisations de L1 dans ses deux versions de lecture. On note que, bien que différentes, ces deux réalisations s'équivalent dans leur notation par leur qualité mélodique. On remarque en outre qu'apprenants et natifs tendent à former un continuum qualitatif remettant sérieusement en cause la dichotomie traditionnelle entre, d'une part, les natifs, et d'autre part, les apprenants.

Locuteurs natifs		Apprenants		Barème	Appréciation
L1v1 L1v2	12			$9 \leq x$	Excellent
L5	7	DID0108-S001R	7	$7 \leq x < 9$	Très bon
		DID0153-S001R	5	$5 \leq x < 7$	Bon
		DID0076-S001R	4	$3 \leq x < 5$	Moyen
		DID0074-S001R	2	$2 \leq x < 3$	Passable
		DID0062-S001R	1	$x < 2$	Non acceptable

Tableau 85 – Grille évaluative critériée de la qualité mélodique du ton descendant-montant (/) sur une syllabe devant frontière mineure

7.4.4 Profilage tonal évaluatif sur syllabe emphatique (dans *nobody*)

La première syllabe du mot emphatique *nobody* se prête parfaitement à deux tâches : celle de vérifier l'aptitude qu'ont les apprenants à réaliser des tons syllabiques montants-descendants (\wedge , *rise-fall*), et aussi, dans un second temps, le degré de complexité qu'ils peuvent atteindre sur une seule et unique syllabe : tons syllabiques unidirectionnels, bidirectionnels, voire tri-directionnels.

7.4.4.1 Phase 1 : Présentation des contours individuels et analyse

Les contours ci-dessous représentent les réalisations des quinze apprenants en fonction de leur catégorie de vitesse de lecture selon l'objectif de départ, qui est la vérification de l'aptitude à réaliser un ton montant-descendant sur une unique syllabe (la syllabe accentuée du mot *nobody*). Les mesures de F0 sont prises en début et fin de syllabe pour /n@U/, en début et fin de voyelle pour /di/, et à l'endroit intermédiaire le plus représentatif pour /n@U/ et /b@/.

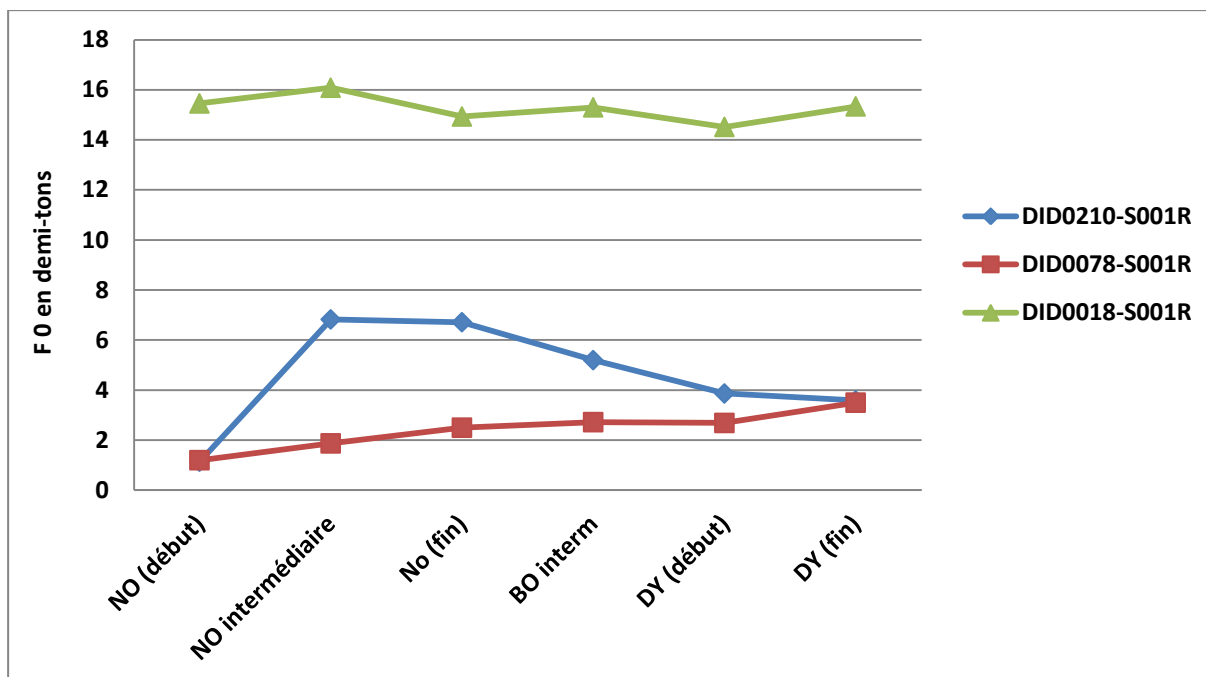


Figure 182 – Schématisation normalisée du mot *nobody* dans le groupe D1 (F0 en demi-tons)

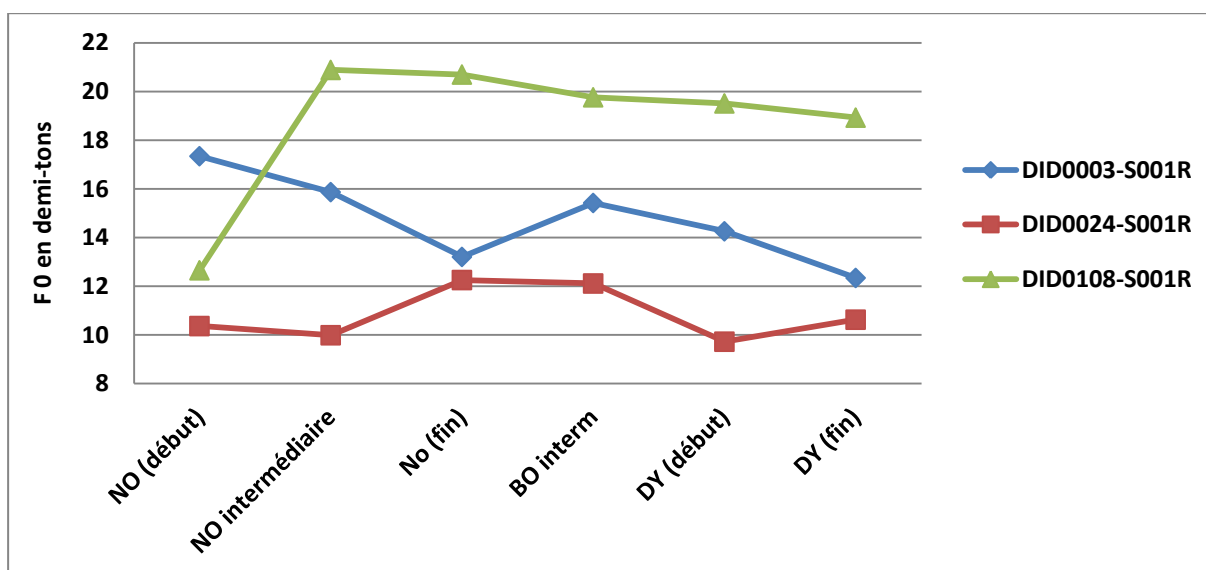


Figure 183 – Schématisation normalisée du mot *nobody* dans le groupe Q1 (F0 en demi-tons)

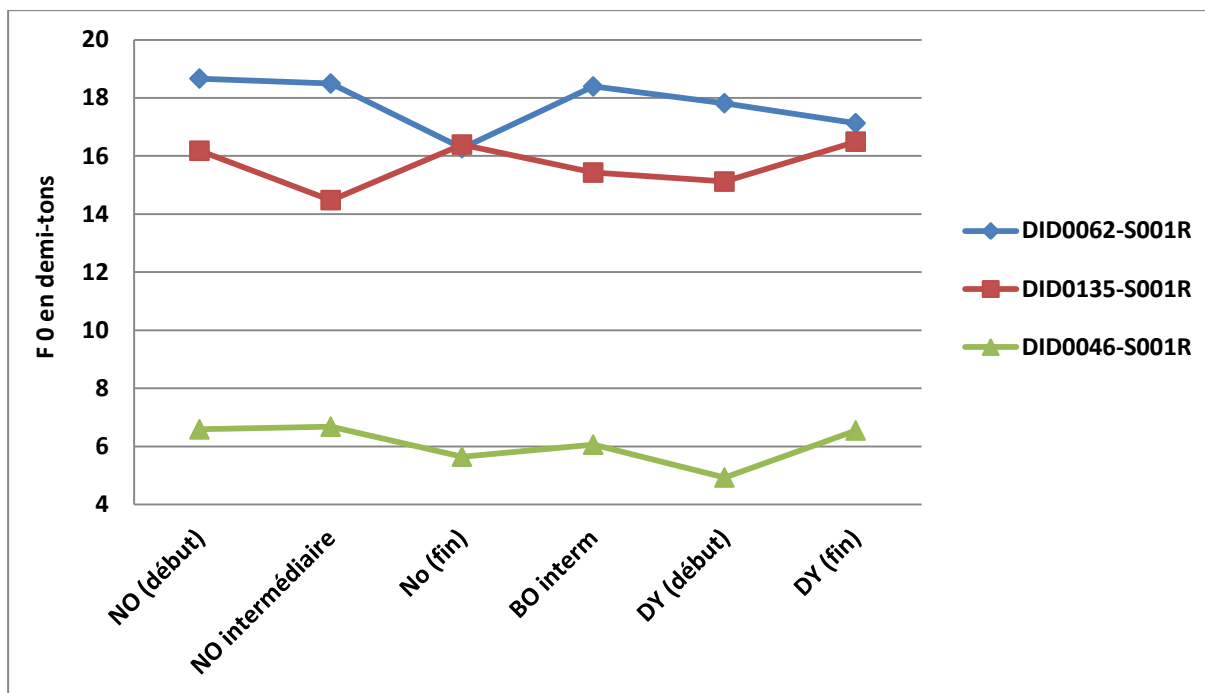


Figure 184 – Schématisation normalisée du mot *nobody* dans le groupe Médiane (F0 en demi-tons)

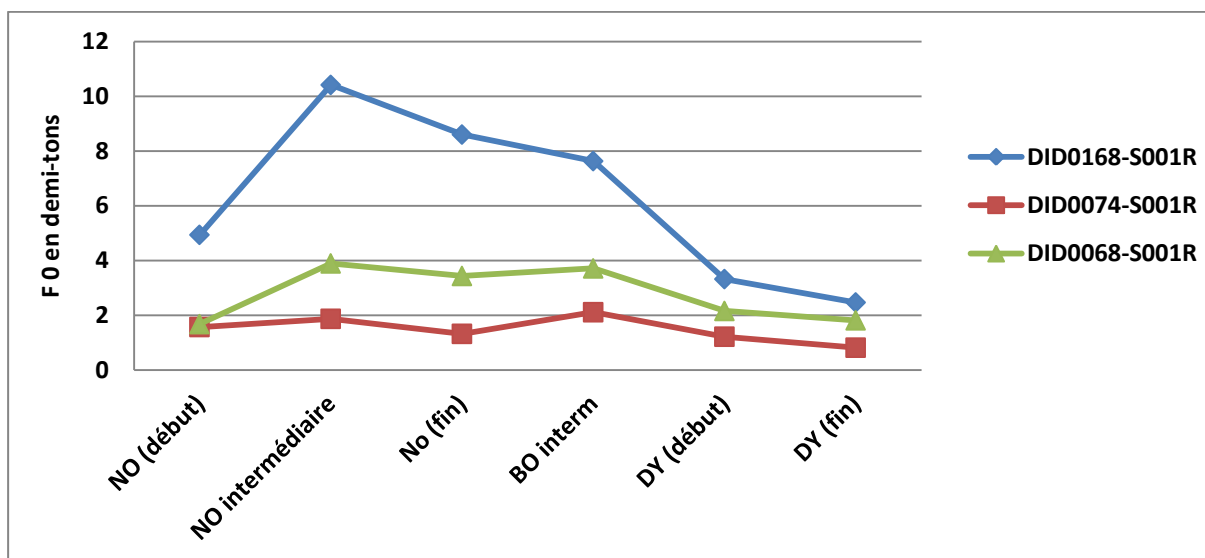


Figure 185 – Schématisation normalisée du mot *nobody* dans le groupe Q3 (F0 en demi-tons)

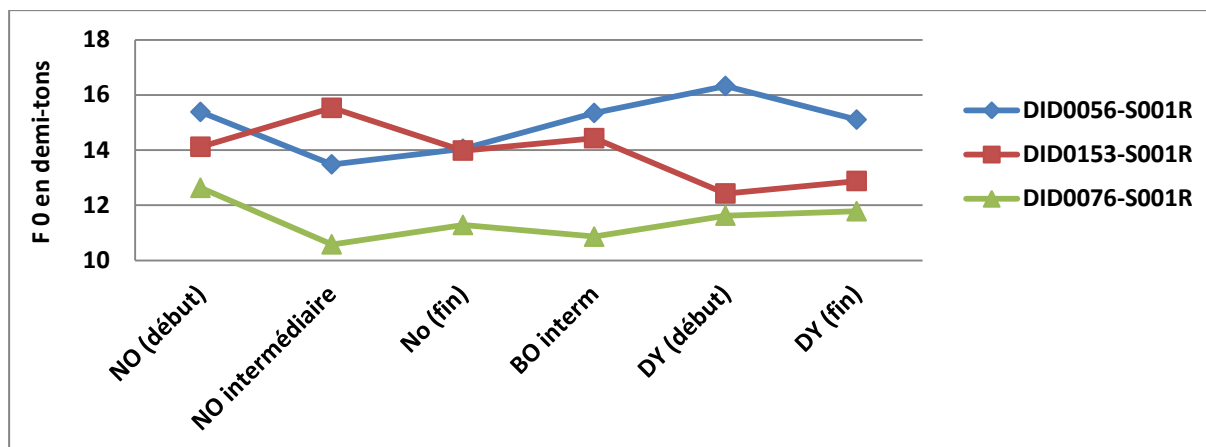


Figure 186 – Schématisation normalisée du mot *nobody* dans le groupe D9 (F0 en demi-tons)

Il ressort de ces graphiques en lignes que certains apprenants réalisent le ton complexe montant-descendant attendu sur la syllabe *no-* accentuée, alors que d'autres choisissent un ton descendant-montant ou un ton simple montant, ou descendant, ou quasiment plat. Sans doute la prise de mesures en trois points est-elle insuffisante pour rendre compte de la variation réalisée avec un degré de précision acceptable.

7.4.4.2 Phase 2 : *Élaboration commentée des formules de calcul sur tableur*

On a vu en amont que les locuteurs natifs pouvaient atteindre un certain degré de complexité sur une seule et unique syllabe : tons syllabiques unidirectionnels, bidirectionnels, voire tridirectionnels. Certains apprenants aussi. Pour le mettre en évidence, le profilage mis en œuvre se limite à des prises de mesure de la F0max : 1. En tout début de syllabe, 2. Intermédiaire ou F0min, 3. Intermédiaire ou F0max, 4. En fin de syllabe. On obtient ainsi le patron mélodique de la syllabe *no-* en tenant compte de la diversité envisagée.

Les formules utilisées sont au nombre de trois : successivement descente, montée et descente éventuelles observées par soustraction des valeurs de F0, puis de l'application de formules de notation qualitative :

- Ton 1 descendant (\) :

$$=SI(AW15="x";0;SI(AW15<-1,5;2;SI(AW15<0;1;0)))$$

- Ton 2 montant (/) :

$$=SI(BC16="x";0;SI(BC16>7;4;SI(BC16>4;3;SI(BC16>1;2;SI(BC16>1;1;0))))$$

- Ton 3 descendant (\) :

$$=SI(BD15<-3;4;SI(BD15<-1,5;3;SI(BD15<0;2;SI(BD15<1,5;1;0))))$$

La variable tonale syllabique \wedge (rise-fall /n@U/), résultant en la variable MEL_s \wedge (M4), est obtenue par l'addition des résultats concernant les deux derniers mouvements :

$$=BG15+BH15$$

La variable qualitative de la mélodie syllabique (MEL_no ou M6) est la somme des trois mouvements éventuels :

$$=SOMME(BF15:BH15)$$

7.4.4.3 Phase 3 : Visualisation des résultats obtenus et analyse

Les résultats de ce barème qualitatif sont représentés dans le graphique ci-dessous.

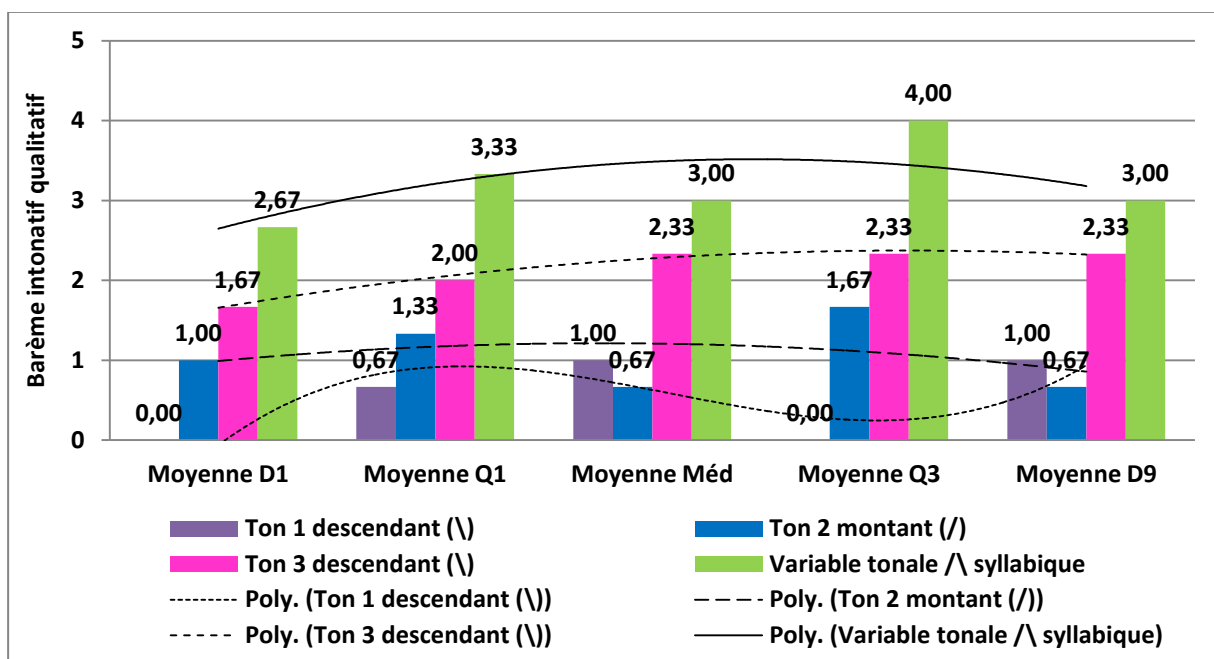


Figure 187 – Complexité de la variation tonale sur la syllabe accentuée de *nobody* par barème qualitatif

On y remarque que le premier ton (Ton 1 descendant) est absent du groupe D1 et de Q3. Alors que les apprenants de D1 s'en tiennent à la variation mélodique et tonale minimale, ceux de Q3 font le choix de se concentrer sur l'amplitude des mouvements montants et descendants, ce qui leur permet d'atteindre la moyenne maximale. Q1, Médiane et D9, investis dans la réalisation des trois mouvements, obtiennent des résultats similaires (Médiane et D9) ou supérieur (Q1) lorsqu'on additionne la note des deux premiers mouvements (tons 1 et 2). Sur les trois mouvements, seul le ton 3 descendant semble davantage déterminé par la vitesse de lecture. Cette catégorisation étant peu satisfaisante pour le phénomène considéré, la

figure suivante montre à présent une catégorisation du contour tonal normalisé sous un autre angle : par patron tonal.

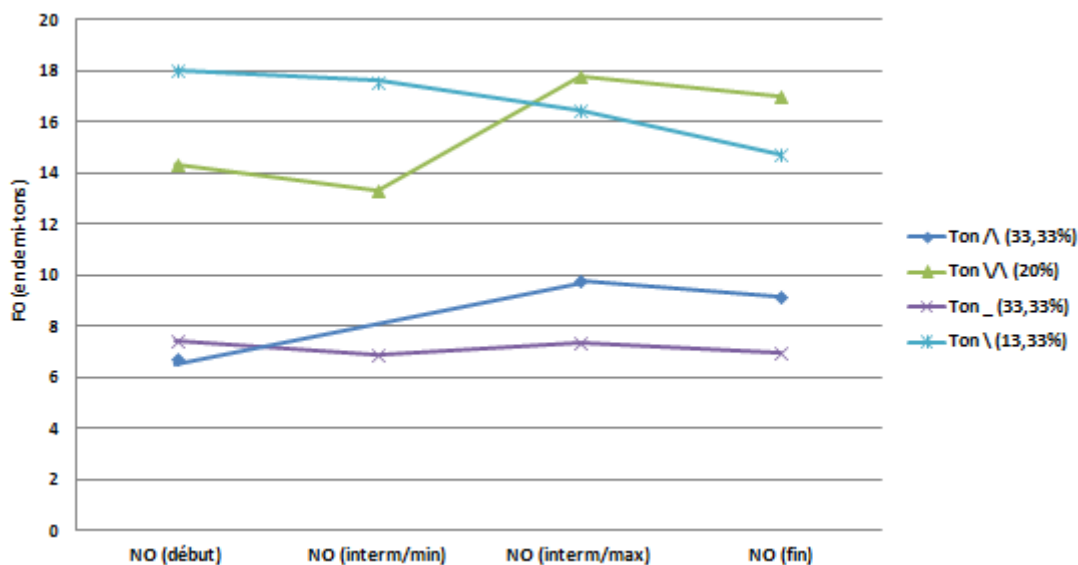


Figure 188 – Contour tonal normalisé moyen de la syllabe /n@U/ de *nobody* par schéma tonal (F0 en demi-tons)

Sa visualisation permet de distinguer clairement les quatre patrons tonaux fondamentaux opérationnels sur la syllabe /n@U/. Leur complexité respective se reflète dans le pourcentage de fréquence dans leur réalisation : le ton plat (le plus basique) est réalisé dans 1/3 des cas (33,33%), alors que le ton tridirectionnel (le plus complexe) ne l'est que dans 20% des cas. J'ai fait l'hypothèse dans mon profilage par formules de calcul que le choix de complexifier le schéma tonal d'une syllabe emphatique exige une maîtrise mélodique dont la qualité peut être quantifiée sur deux axes : nombre de tons mis en œuvre dans la réalisation syllabique et amplitude de ces mouvements par mesure de F0.

7.4.4.4 Phase 4 : Résultat du profilage par prototype de grille évaluative avec exemple de barème et d'appréciations

Le tableau suivant présente un prototype évaluatif qualitatif de cette complexité tonale syllabique, laquelle ne peut être perçue que sur syllabe accentuée d'un mot emphatique regroupant les conditions précisées dans le Chapitre III sur l'emphase.

Locuteurs natifs		Apprenants		Barème	Appréciation
				$8 < x$	Excellent
L1v1	7	DID0108-S001R	7	$6 < x \leq 8$	Très bon
L1v2	6	DID0210-S001R	5	$4 < x \leq 6$	Bon
L5	4	DID0068-S001R	4	$2 < x \leq 4$	Moyen
		DID0076-S001R	2	$1 < x \leq 2$	Passable
		DID0078-S001R	1	$x \leq 1$	Non acceptable

Tableau 86 – Grille évaluative critériée de la complexité tonale sur syllabe accentuée de *nobody*

Ici, les locuteurs L1 et L5 se contentent d'un ton montant-descendant sur la syllabe /n@U/. Dans le cas de L1v2, ce ton *rise-fall* s'apparente à une simple montée, d'où sa note plus basse. La grille montre ici des équivalences qualitatives entre les réalisations des natifs de référence et celles des apprenants. Des natifs à la diction plus « posée » ont doté cette syllabe d'une variation plus importante sur deux plans mélodiques : le choix d'un patron tridirectionnel et un registre plus ample. Bien évidemment, la durée consacrée à la prononciation de cette syllabe était bien plus conséquente : le choix du tempo de lecture est déterminant car il conditionne les contraintes prosodiques qui en découlent.

7.5 Profilage de critères d'évaluation rythmiques

On a vu précédemment que le rythme résultait d'une combinaison de corrélats multiples qu'il peut être extrêmement difficile de dissocier. Le principe organisateur de l'évaluation se doit d'être simple, rigoureux et performant. L'étude des variables présentées ici repose sur une parcimonie de critères, acoustiques ou perceptifs, mais tous ciblent le phénomène de saillance rythmique, pour donner du sens au message en formatant la chaîne de parole. Il s'agit bien là de saillance, c'est-à-dire de celle qui permet d'« accrocher » l'attention du locuteur tout au long du message, de canaliser cette attention pour rendre ce message clair et compréhensible en jalonnant la parole de repères destinés à faciliter la communication.

Aussi, quelle que soit leur nature, les rubriques suivantes détaillent leur fonctionnement ainsi que leurs effets. D'abord est analysée la variable construite à partir de variations de fréquence fondamentale sur la paire « *so* + adverbe » (*so sweetly*), puis dans un second temps la variable RYT_lex (Ry1) déterminée par perception du respect des schémas accentuels sur polysyllabes. Le type de profilage mis en œuvre pour *so sweetly* est à la base des deux variables RYT_so4 (Ry2) et RYTMEL_so4 (RM). Les autres variables rythmiques à mesures acoustiques de F0 sont profilées sur le même modèle.

7.5.1 Profilage mélodique pour évaluation rythmique de la paire « so + adverbe » (Phrase 4)

Les phases de profilage de la paire *so sweetly* sont agencées de la même manière que précédemment pour les traits distinctifs mélodiques. Ainsi, sont d'abord présentés les graphiques en ligne représentant la F0max de chaque syllabe du tronçon dont elle fait partie, toujours selon le classement par catégories de vitesse de lecture. Leur analyse conditionne le profilage des formules *Excel* utilisées pour la construction des données. Le résultat de ces données est visualisé puis analysé. En dernier lieu, ces résultats et analyses de ce profilage aboutissent à un prototype de grille d'évaluation.

7.5.1.1 Phase 1 : Présentation des contours individuels et analyse

L'étude de *so sweetly* est faite à partir de tronçons de parole délimités par deux mots susceptibles d'être marqueurs rythmiques dans la phrase et, en conséquence, d'avoir une incidence sur la prosodie de la paire étudiée.

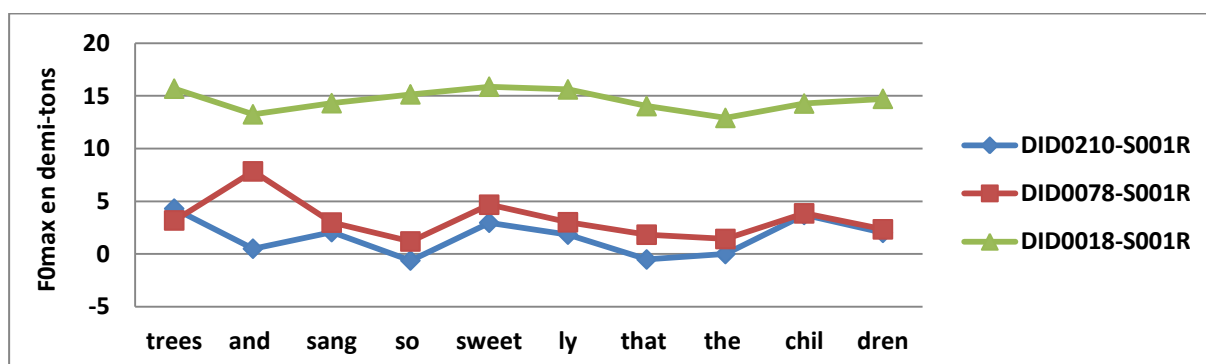


Figure 189 – Contour mélodique de la paire *so sweetly* de la Phrase 4 dans le groupe D1 (F0 en demi-tons)

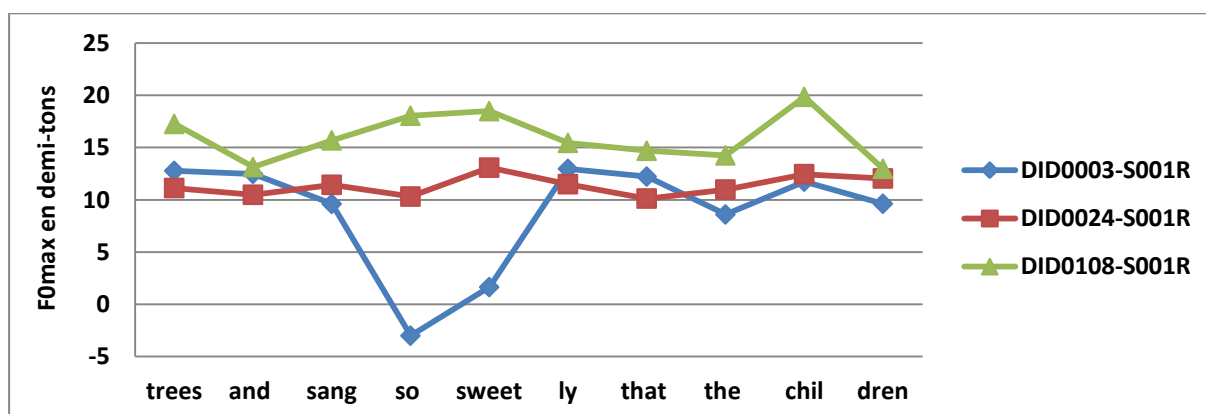


Figure 190 – Contour mélodique de la paire *so sweetly* de la Phrase 4 dans le groupe Q1 (F0 en demi-tons)

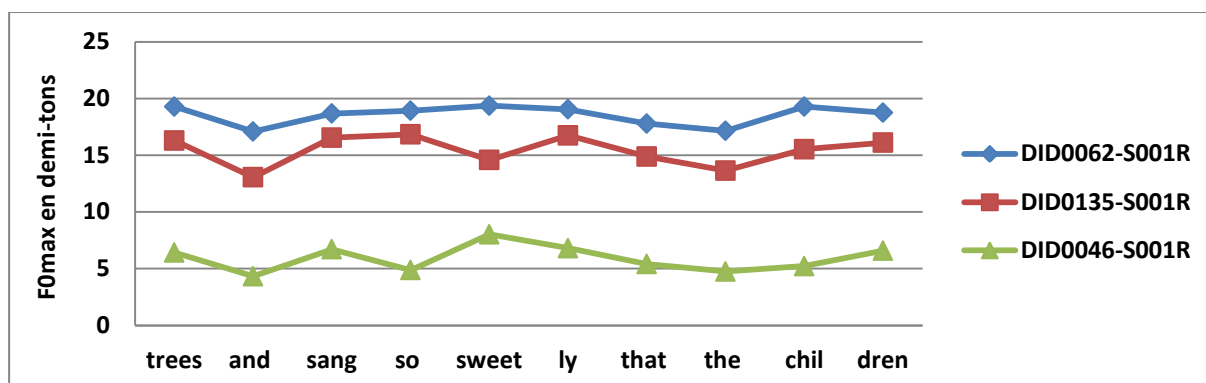


Figure 191 – Contour mélodique de la paire *so sweetly* de la Phrase 4 dans le groupe Médiane (F0 en demi-tons)

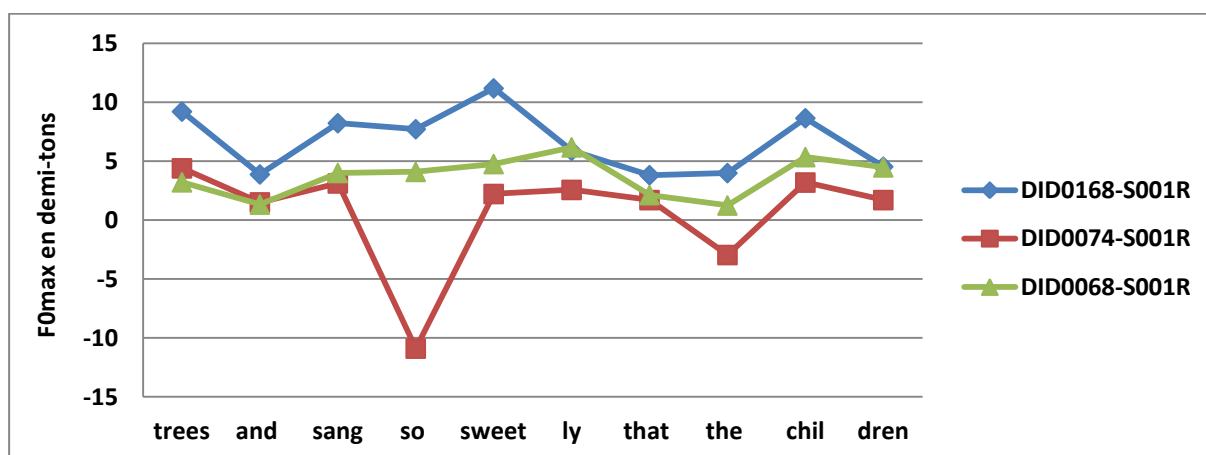


Figure 192 – Contour mélodique de la paire *so sweetly* de la Phrase 4 dans le groupe Q3 (F0 en demi-tons)

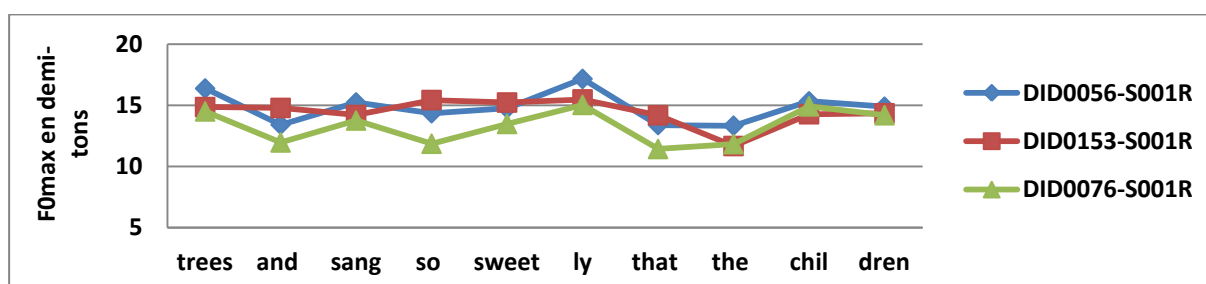


Figure 193 – Contour mélodique de la paire *so sweetly* de la Phrase 4 dans le groupe D9 (F0 en demi-tons)

À partir de la schématisation de la courbe intonative, on observe que l'adverbe *so* reçoit très souvent une F0 plus basse que l'élément qui le précède, ici *sang* (mot de F0 plus haute parce qu'il est accentué). Le rythme obtenu devient alors forte-faible-forte avec *sang so sweet-* ou forte-faible-faible-forte avec *sang so sweetly*, où une continuation s'établit sur la dernière

syllabe de *sweetly*. Les syllabes repères sont donc *sang so sweetly* et la différence de F0 entre chacune de ces quatre syllabes révèle un choix pouvant être traduit en critères évaluatifs.

7.5.1.2 Phase 2 : *Élaboration commentée des formules de calcul dans un tableur*

Afin de rendre le calcul opérationnel, un forfait de 5 points a été attribué dès le départ pour éviter les résultats négatifs. Ensuite s'établissent les différences de F0max entre les quatre syllabes mentionnées précédemment.

Les formules caractérisant la variable RYT_so4 sont présentées ci-dessous.

- sang-so :

=SI(AL15>5;5;SI(AL15>4;4;SI(AL15>3;3;SI(AL15>2;2;SI(AL15>1;1;SI(AL15>0;0;SI(AL15>-1;-1;-2))))))

- so-sweet :

=SI(AM15<-1;1;0)

- sweet-ly :

=SI(AN15>2;-3;SI(AN15>1;-2;SI(AN15>0;-1;SI(AN15>-1;0;SI(AN15>-2;1;2))))

Les données de la variable RYT_so4 correspondent à la somme des résultats (forfait + *sang-so* + *so-sweet*), alors que celles de RYTMEL_so4 impliquent l'adjonction de la différence de F0 entre les deux syllabes du mot *sweetly* (forfait + *sang-so* + *so-sweet* + *sweet-ly*). Cela fait de RYTMEL_so4 une variable hybride, à la fois mélodique et rythmique.

7.5.1.3 Phase 3 : *Visualisation des résultats obtenus et analyse*

La moyenne des données de ces variables obtenues grâce aux formules de calcul expliquées ci-dessus est représentée dans le graphique suivant.

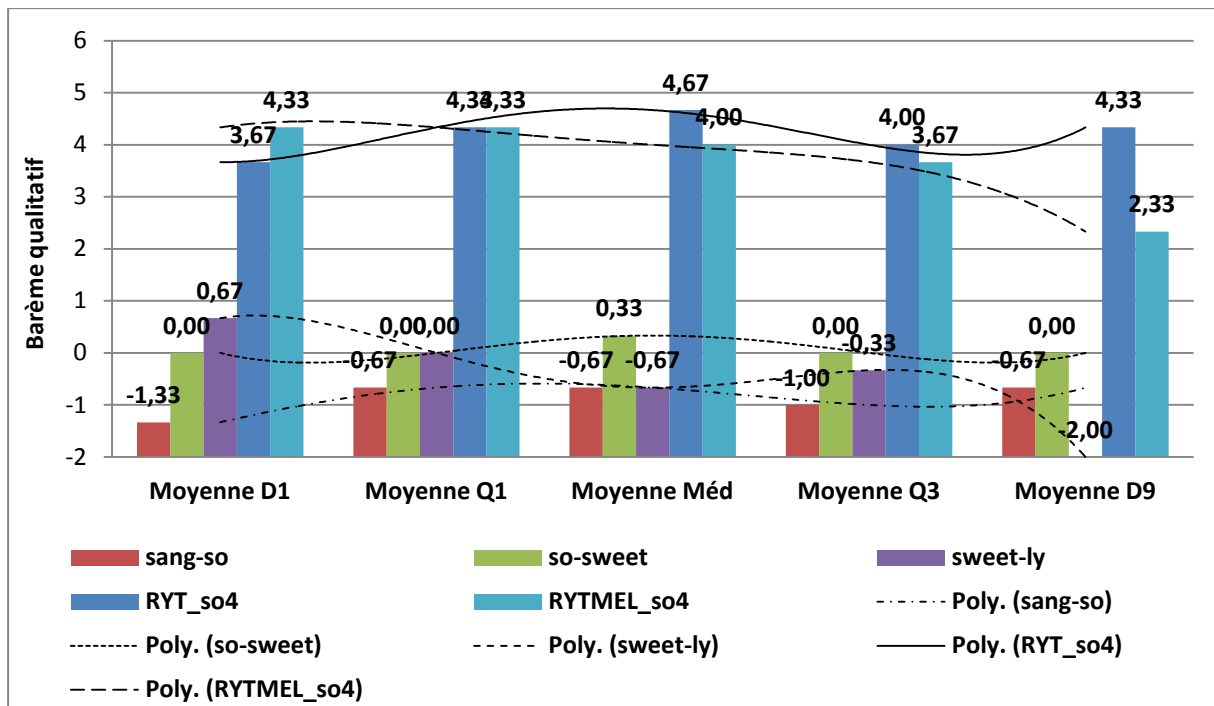


Figure 194 – Détail des données constitutives des variables RYT_so4 et RYTMEL_so4 par catégorie d'apprenants selon la vitesse de lecture (barème qualitatif)

L'histogramme des résultats montre que le critère *sang-so* est réalisé contrairement à la norme par les cinq groupes, et celui de *so-sweet* avoisine 0. Les apprenants de la Médiane obtiennent les meilleurs résultats et ceux de D1 les plus faibles. Le critère mélodique *sweet-ly* est pratiquement réalisé selon le classement en fonction de la vitesse de lecture, ce qui fait grossièrement correspondre la variable RYTMEL_so4 aux variables de durée, lui donnant une certaine régularité.

7.5.1.4 Phase 4 : Résultat du profilage par prototype de grille évaluative avec exemples de barèmes et d'appréciations

Les résultats de l'application des formules des deux variables sur les réalisations des apprenants ainsi que sur ceux des natifs de référence ont permis de réaliser les prototypes de grilles évaluatives aux critères qualitatifs ci-dessous.

Locuteurs natifs		Apprenants		Barème	Appréciation
L1v2	11			$x \geq 8$	Excellent
L1v1	10				
L5	9				
		DID0108-S001R	7	7	Très bon
		DID0135-S001R	6	6	Bon
		DID0062-S001R	5	5	Moyen
		DID0168-S001R	4	4	Passable
		DID0210-S001R	3	$x \leq 3$	Non acceptable

Tableau 87 – Grille évaluative avec critères qualitatifs de la variable RYT_so4 (paire rythmique « so + adverbe » de la Phrase 4)

Locuteurs natifs		Apprenants		Barème	Appréciation
L1v2	13			$9 < x$	Excellent
L1v1	12				
L5	11				
		DID0108-S001R	9	$7 < x \leq 9$	Très bon
		DID0168-S001R	6	$5 < x \leq 7$	Bon
		DID0153-S001R	5	$4 < x \leq 5$	Moyen
		DID0078-S001R	4	$3 < x \leq 4$	Passable
		DID0068-S001R	3	$x \leq 3$	Non acceptable

Tableau 88 – Grille évaluative avec critères qualitatifs de la variable RYTMEL_so4 établie à partir de la variable RYT_so4 et d'une éventuelle remontée mélodique sur dernière syllabe inaccentuée

7.5.2 Profilage de patrons accentuels lexicaux dans polysyllabes

Comme précisé en amont, chaque mot correspond à une difficulté accentuelle spécifique. Ici encore, j'applique l'hypothèse d'Herment, vérifiée dans le Chapitre III sur l'emphase, selon laquelle le nombre de pointages équivaut à l'application d'un degré qualitatif. Contrairement à l'étude de Horgues (2010), la raison du choix du patron accentuel n'entre pas en ligne de compte. Ainsi, après calcul des réponses justes de chaque apprenant et leur transformation en pourcentage, le graphique suivant a pu être construit.

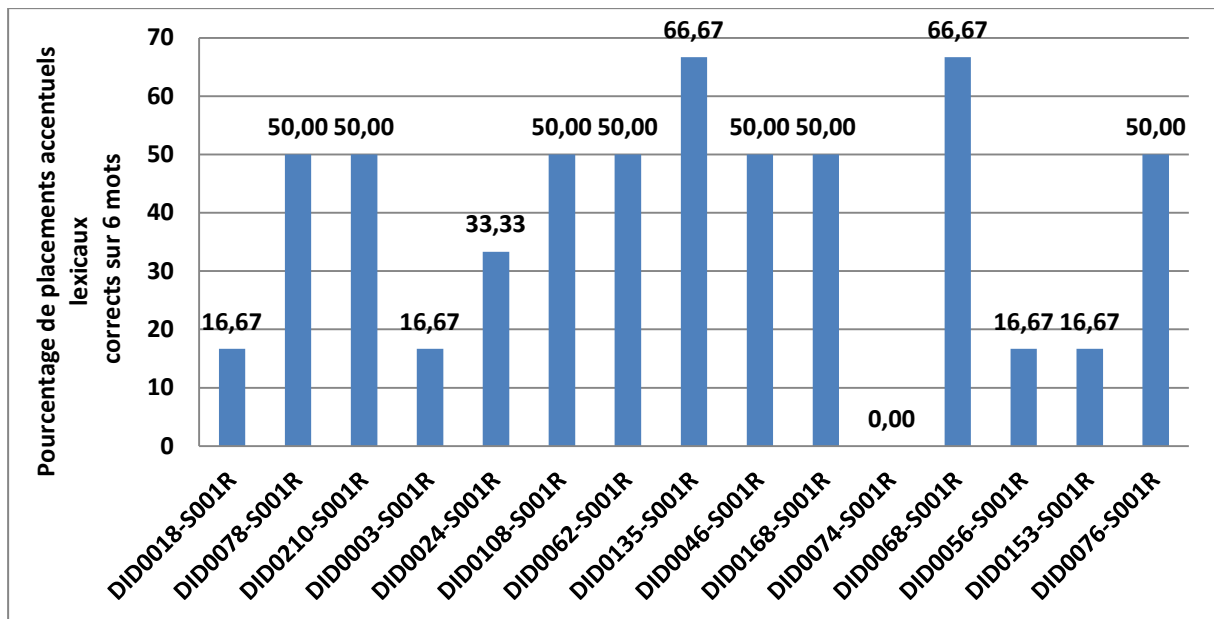


Figure 195 – Placement accentuel lexical conforme sur 6 polysyllabes par apprenant (accentuation primaire et secondaire)

En ce qui concerne la prestation des apprenants, il faut remarquer que l'apprenant DID0076 n'a pas lu la Phrase 13 avec la mention *TRESPASSERS WILL BE PROSECUTED* ; son score est donc amputé de deux points sur les six pouvant être totalisés. L'apprenant DID0074 n'a réalisé aucun patron accentuel correctement. En fait, les seuls aspects positifs de sa prestation prosodique sont centrés sur la réalisation du syntagme intonatif S9 en registre et en intonation. Cet aspect contribue à montrer que, bien qu'en marge des phénomènes prosodiques traditionnellement considérés, le respect des patrons accentuels souligne le degré de maîtrise technique de la langue orale. Ce terrain est propice à la création de traits distinctifs en évaluation dans une dimension prosodique spécifique. On remarque en outre que les meilleurs scores ont été obtenus par des apprenants des groupes de la Médiane et de Q3, et seul le groupe de la Médiane n'obtient aucun résultat inférieur à 50% de réussite. Ceci démontre une fois de plus les limites que peut avoir une catégorisation selon la rapidité d'élocution.

Pour ce qui est des polysyllabes sélectionnés dans cette étude, les réalisations accentuelles adéquates renseignent sur la difficulté de lecture de chaque schéma accentuel. Il est à noter que les mots *trespassers* et *prosecuted*, ayant été omis par l'apprenant DID0076, reçoivent un pourcentage plus faible. Cependant, il faut souligner le fait que cet apprenant a correctement réalisé les schémas accentuels de *conversation*, *determined* et *musicians*, ces deux derniers mots représentant une grande difficulté de réalisation correcte, comme le montre le bilan de réussite par mot ci-dessous.

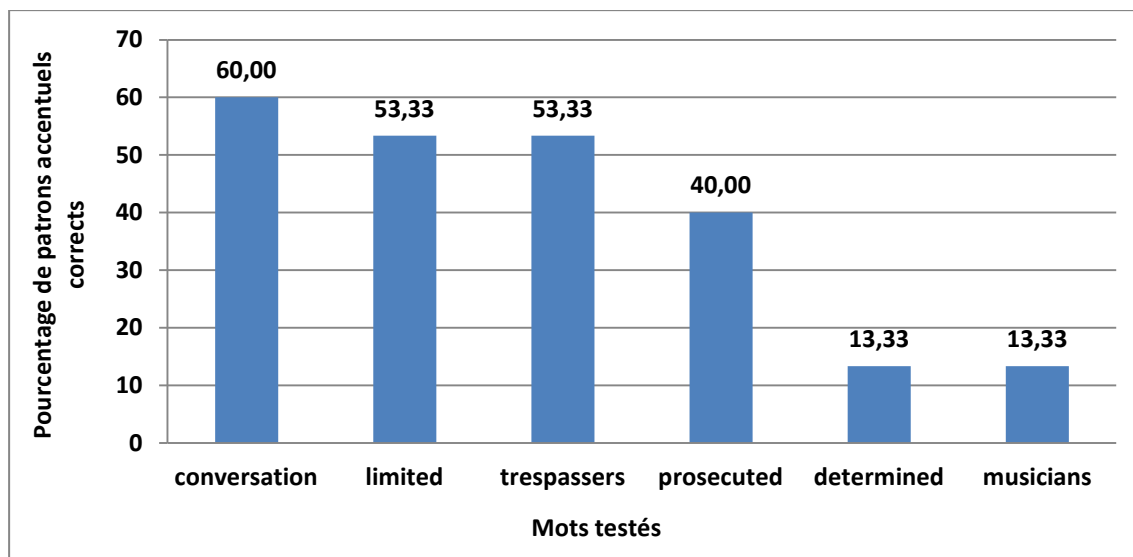


Figure 196 – Réalisations correctes de patrons accentuels sur polysyllabes par mot (en pourcentage)

Les difficultés observées dans l'optique d'une réalisation conforme sont nombreuses. Le *locus* de l'accent secondaire est assez bien maîtrisé dans l'ensemble dans *conversation*. Le patron accentuel /100/ est réalisé par 53,3% des apprenants sur les mots de trois syllabes (*limited*, *trespassers*). Par contre, l'ajout d'un suffixe faible ainsi que d'une diphtongue non accentuée fait chuter les résultats à 40% de réussite (*prosecuted*). Les mots ayant comptabilisé le moins de réussite (13,3%) sont *determined* et *musicians*. *Determined* comprend quatre syllabes, dont un préfixe et un suffixe régulier du prétérit. Nombreux ont été les apprenants troublés au point de réaliser le « i » par /aI/. Le faible pourcentage de réussite sur le mot *musicians* confirme les prévisions émises au Chapitre III : la difficulté majeure étant le patron accentuel à conserver malgré le rythme de la phrase dans les syllabes environnantes. Étant donné la relative réussite accentuelle de *conversation* malgré ses quatre syllabes, seul le contexte rythmique peut expliquer la difficulté du placement accentuel canonique sur *musicians*. Le fort pourcentage d'erreur sur ce mot montre qu'un meilleur niveau de maîtrise du placement accentuel passe, à un niveau plus élevé, par celle de l'influence du rythme dans le flot de la parole.

Le prototype de grille évaluative proposé reprend les résultats observés. Aucun des quinze apprenants n'a réalisé un score de 83,33% de réussite bien que l'étalonnage des possibles prévoit un tel résultat. Aussi, la grille prototype d'évaluation se doit d'y faire référence.

Locuteurs natifs		Apprenants		Barème (%)	Appréciation
L1	100		100	$83,33 < x \leq 100$	Excellent
L5	100				
			83,33	$66,66 < x \leq 83,33$	Très bon
		DID0068-S001R	66,66	$50 < x \leq 66,66$	Bon
		DID0062-S001R	50	$33,33 < x \leq 50$	Moyen
		DID0024-S001R	33,33	$16,66 < x \leq 33,33$	Passable
		DID0056-S001R	16,66	$x \leq 16,66$	Non acceptable
		DID0074-S001R	0		

Tableau 89 – Grille évaluative de respect du patron accentuel sur six polysyllabes en contexte (pourcentage de réussite)

8 Synthèse du profilage des traits distinctifs et prototypes de grilles critériées pour évaluer la prosodie d'apprenants francophones

Dans cette recherche pour créer des traits distinctifs évaluatifs, le profilage a été opéré en quatre phases distinctes, marquant une progression nécessaire dans le processus de création. Tout d'abord, la phase 1 a consisté à observer et analyser les réalisations des apprenants, ensuite, la phase 2 a concrétisé cette analyse en élaborant des formules de calcul. La phase 3 a détaillé et analysé les résultats obtenus par l'application de ces formules aux réalisations des quinze apprenants sélectionnés. La dernière phase, quant à elle, a présenté le résultat du profilage, c'est-à-dire l'outil d'évaluation qu'est la grille d'évaluation critériée opérationnelle. Les résultats obtenus semblent satisfaisants et cohérents lorsqu'ils sont appliqués sur les réalisations des quinze apprenants et de celles des natifs de référence. Il ne s'agit là que de prototypes destinés à être affinés à l'usage. Alors que certaines grilles n'ont de valeur qu'en lecture, d'autres sont applicables à de la conversation, sans doute après un réétalonnage des mesures pour qu'elles correspondent exactement aux conditions d'évaluation.

En synthèse de cette recherche sur les traits discriminants, les voici à présent listés ci-dessous, puis présentés pour certains sous forme de grilles évaluatives.

8.1 Traits prosodiques distinctifs principaux pressentis à l'issue de cette recherche

Pour lister les principaux traits discriminants pressentis à l'issue de la recherche de ce Chapitre IV, je reprends ici le tableau de variables présenté au début de ce chapitre, en ne conservant que celles qui se sont révélées jusque là les plus marquantes. Elles sont ici aussi catégorisées par domaine prosodique.

Domaine	Nom	Code	Caractéristiques
Temporalité	DUR_ABCp	D2	Durée des 3 segments non ponctués ABC avec pauses (P4, 16 et 25)
	DEB_ABCp	D5	Nombre de syllabes par seconde dans 3 segments non ponctués ABC avec pauses (P4, 16 et 25)
Registre	REG_Excl	Re1	Ampleur du registre de l'exclamative (F0max-F0min)
	REG_so4	Re3	Ampleur du registre de P4 (F0max sur <i>so sweet</i> -F0min P4)
Mélodie	MEL_Excl	M1	Qualité mélodique de l'exclamative
	MEL_no	M6	Qualité et complexité tonale réalisée sur la syllabe /n@U/ de <i>nobody</i> (P11)
	RYT_so4	Ry2	Qualité de proéminence par F0 du marqueur <i>so sweet</i> (ly)
	RYT_so22	Ry3	Qualité de proéminence par F0 du marqueur <i>so sweet</i> par F0
	RYT_all12	Ry4	Qualité de proéminence par F0 du marqueur <i>all round</i> par F0
	RYT_them	Ry6	Absence de proéminence d'un pronom en fin de syntagme intonatif par test perceptif

Tableau 90 – Traits distinctifs principaux pressentis, sous forme des variables utilisées pour l'établissement des profils des apprenants, par domaine prosodique

8.2 Prototypes de grilles évaluatives de critères discriminants temporels

Locuteurs natifs		Apprenants		Barème	Appréciation
L5	135,3			$x < 140$	Excellent
		D1	143,4	$140 < x < 150$	Très bon
		Q1	155,1	$150 < x < 160$	Bon
		Méd	167,7	$160 < x < 175$	Moyen
SG-UP-S3-f-52	186,5	Q3	178,9	$175 < x < 190$	Passable
		D9	192,9	$x > 190$	Non acceptable

Tableau 91 – Grille évaluative de la durée de lecture du texte *The Selfish Giant* (en secondes) correspondant à la variable temporelle DUR_T (D1)

CHAPITRE IV – Expérimentation de création de profils et de critères d'évaluation prosodiques d'apprenants

Locuteurs natifs		Apprenants		Barème	Appréciation
L5	7,99			$x \leq 7,8$	Excellent
		DID0210-S001R	8,481	$7,8 \leq x < 8,5$	Très bon
		DID0078-S001R	9,178	$8,5 \leq x < 9,2$	Bon
		DID0024-S001R	9,995	$9,2 \leq x < 10,1$	Moyen
SG-UP-S3-f-52	11,03	DID0062-S001R	10,998	$10,1 \leq x < 11,1$	Passable
		DID0076-S001R	13,720	$x \geq 11,1$	Non acceptable

Tableau 92 – Grille évaluative de la durée des tronçons du texte *The Selfish Giant* avec pauses (en secondes) correspondant à la variable temporelle DUR_ABCp (D2)

Locuteurs natifs		Apprenants		Barème	Appréciation
L5	5,75			$x > 5,5$	Excellent
		D1	5,27	$5,1 \geq x > 5,5$	Très bon
		Q1	4,68	$4,7 \geq x > 5,1$	Bon
		Méd	4,51	$4,4 \geq x > 4,7$	Moyen
SG-UP-S3-f-52	4,17	Q3	4,47	$4,2 \geq x > 4,4$	Passable
		D9	3,96	$x \leq 4,2$	Non acceptable

Tableau 93 – Grille évaluative du débit d'élocution des tronçons du texte *The Selfish Giant* avec pauses (nombre de syllabes par seconde) correspondant à la variable temporelle DEB_ABCp (D3)

8.3 Prototypes de grilles évaluatives de critères discriminants de registre

Locuteurs natifs		Apprenants		Barème (DT)	Appréciation
L1v1	26,70			$16 < x$	Excellent
L1v2	17,92				
L5	11,44	DID0068-S001R	15,28	$12 < x \leq 16$	Très bon
		DID0108-S001R	10,66	$10 < x \leq 12$	Bon
		DID0003-S001R	8,32	$8 < x \leq 10$	Moyen
		DID0078-S001R	6,78	$6 < x \leq 8$	Passable
		DID0062-S001R	4,02	$x < 6$	Non acceptable

Tableau 94 – Grille évaluative critériée du degré d'ampleur du registre dans l'exclamative en *How* suivie de son énoncé citant (F0 en demi-tons sur base 100)

Locuteurs natifs		Apprenants		Barème (DT)	Appréciation
L1v1	24,00			$13,5 < x$	Excellent
L1v2	15,86				
		DID0168-S001R	12,23	$11 < x \leq 13,5$	Très bon
L5	9,72	DID0024-S001R	9,47	$8,5 < x \leq 11$	Bon
		DID0076-S001R	7,42	$6 < x \leq 8,5$	Moyen
		DID0018-S001R	4,62	$3,5 < x \leq 6$	Passable
		DID0078-S001R	2,20	$x < 3,5$	Non acceptable

Tableau 95 – Grille évaluative critériée du degré d'ampleur du registre dans l'exclamative en *How* sans énoncé citant (F0 en demi-tons sur base 100)

8.4 Prototypes de grilles évaluatives de critères discriminants mélodiques

Locuteurs natifs		Apprenants		Barème	Appréciation
L1v1	9			$x \leq 8$	Excellent
		DID0168-S001R	7	$6 \leq x < 8$	Très bon
L5	4	DID0153-S001R	4	$4 \leq x < 6$	Bon
		DID0076-S001R	2	$2 \leq x < 4$	Moyen
		DID0074-S001R	0	$0 \leq x < 2$	Passable
		DID0135-S001R	-2	$x < 0$	Non acceptable

Tableau 96 – Grille évaluative des critères qualitatifs de la qualité mélodique de l'exclamative en *How* sans son énoncé citant

Locuteurs natifs		Apprenants		Barème	Appréciation
L1v1	10			$8 \leq x$	Excellent
		DID0068-S001R	6	$6 \leq x < 8$	Très bon
L5	5	DID0210-S001R	5	$4 \leq x < 6$	Bon
		DID0076-S001R	3	$2 \leq x < 4$	Moyen
		DID0074-S001R	1	$0 \leq x < 2$	Passable
		DID0135-S001R	-1	$x \leq 0$	Non acceptable

Tableau 97 – Grille évaluative des critères qualitatifs de la qualité mélodique de l'exclamative avec son énoncé citant

CHAPITRE IV – Expérimentation de création de profils et de critères d'évaluation prosodiques d'apprenants

Locuteurs natifs		Apprenants		Barème	Appréciation
				$9 < x$	Excellent
L1v1	9	DID0003-S001R	9	$7 < x \leq 9$	Très bon
L1v2	8				
		DID0168-S001R	7	$5 < x \leq 7$	Bon
		DID0153-S001R	5	$3 < x \leq 5$	Moyen
		DID0068-S001R	3	$1 < x \leq 3$	Passable
		DID0056-S001R	-1	$x \leq 1$	Non acceptable

Tableau 98 – Grille évaluative des critères qualitatifs de la qualité mélodique du syntagme intonatif neutre S9

Locuteurs natifs		Apprenants		Barème	Appréciation
L1v1	12			$9 \leq x$	Excellent
L1v2					
L5	7	DID0108-S001R	7	$7 \leq x < 9$	Très bon
		DID0153-S001R	5	$5 \leq x < 7$	Bon
		DID0076-S001R	4	$3 \leq x < 5$	Moyen
		DID0074-S001R	2	$2 \leq x < 3$	Passable
		DID0062-S001R	1	$x < 2$	Non acceptable

Tableau 99 – Grille évaluative des critères qualitatifs de la qualité mélodique du ton \vee sur une syllabe devant frontière mineure

Locuteurs natifs		Apprenants		Barème	Appréciation
				$8 < x$	Excellent
L1v1	7	DID0108-S001R	7	$6 < x \leq 8$	Très bon
L1v2	6	DID0210-S001R	5	$4 < x \leq 6$	Bon
L5	4	DID0068-S001R	4	$2 < x \leq 4$	Moyen
		DID0076-S001R	2	$1 < x \leq 2$	Passable
		DID0078-S001R	1	$x \leq 1$	Non acceptable

Tableau 100 – Grille évaluative des critères qualitatifs de la complexité tonale sur syllabe accentuée de *nobody*

8.5 Prototypes de grilles évaluatives de critères discriminants rythmiques

Locuteurs natifs		Apprenants		Barème	Appréciation
L1v2	11			$x \geq 8$	Excellent
L1v1	10				
L5	9				
		DID0108-S001R	7	7	Très bon
		DID0135-S001R	6	6	Bon
		DID0062-S001R	5	5	Moyen
		DID0168-S001R	4	4	Passable
		DID0210-S001R	3	$x \leq 3$	Non acceptable

Tableau 101 – Grille évaluative des critères qualitatifs de la variable RYT_so4 (paire rythmique « so + adverbe » de la Phrase 4

Locuteurs natifs		Apprenants		Barème	Appréciation
L1v2	13			$9 < x$	Excellent
L1v1	12				
L5	11				
		DID0108-S001R	9	$7 < x \leq 9$	Très bon
		DID0168-S001R	6	$5 < x \leq 7$	Bon
		DID0153-S001R	5	$4 < x \leq 5$	Moyen
		DID0078-S001R	4	$3 < x \leq 4$	Passable
		DID0068-S001R	3	$x \leq 3$	Non acceptable

Tableau 102 – Grille évaluative des critères qualitatifs de la variable RYTMEL_so4 établie à partir de la variable RYT_so4 et d'une éventuelle remontée mélodique sur dernière syllabe inaccentuée

Locuteurs natifs		Apprenants		Barème (%)	Appréciation
L1	100		100	$83,33 < x \leq 100$	Excellent
L5	100				
			83,33	$66,66 < x \leq 83,33$	Très bon
		DID0068-S001R	66,66	$50 < x \leq 66,66$	Bon
		DID0062-S001R	50	$33,33 < x \leq 50$	Moyen
		DID0024-S001R	33,33	$16,66 < x \leq 33,33$	Passable
		DID0056-S001R	16,66	$x \leq 16,66$	Non acceptable
		DID0074-S001R	0		

Tableau 103 – Grille évaluative de respect du patron accentuel sur polysyllabes en contexte (pourcentage de réussite)

9 Conclusion du Chapitre IV

Ce Chapitre IV est destiné à vérifier l'Hypothèse 3 rappelée ci-dessous :

H3 — Le paramètre de rapidité d'élocution étant un indicateur déterminant de niveau des apprenants, la catégorisation de ceux-ci selon ce paramètre devrait permettre d'organiser la variation inter-locuteurs des résultats des autres variables prosodiques afin de trouver des critères d'évaluation fiables.

et de répondre à la Question de Recherche 1 :

QR1 — Étant donné la forte variabilité des réalisations prosodiques natives et d'apprenants, des outils spécifiques doivent être utilisés ou créés afin de l'analyser. La validation des Hypothèses H1, H2 et H3 devrait permettre de créer des profils prosodiques de non natifs permettant d'orienter la création de traits discriminants évaluatifs.

Les Chapitres II et III ont chacun conclu à la validation des hypothèses H1 et H2 concernant respectivement. H3 a été validée en début de ce chapitre par la création d'une grille méthodologique de classement des apprenants selon la durée de lecture du texte *The Selfish Giant*. La vérification de la validation de QR1 a donc pour but la création de traits prosodiques d'apprenants anglicistes francophones à visée évaluative via celle de leurs profils prosodiques.

Après avoir brièvement rappelé les buts à atteindre et les enjeux introduisant l'analyse de la prosodie des apprenants par profilage (section 4.2), ce Chapitre IV a présenté l'expérimentation STAT-NN1 de sélection des fichiers d'apprenants à étudier spécifiquement (section 4.3), puis il a procédé à l'examen des considérations prévalant à l'élaboration de la grille d'analyse conduisant à la sélection des variables et des mesures prises en compte (expérimentations COR-AC-NN, PER-THEM-NN, et PER-LEX-NN dans la section 4.4). Ensuite, trois phases ont marqué la description des résultats acoustiques ou perceptifs obtenus grâce à ces quatre séries d'expérimentation, l'analyse des données qui en résulte, puis leur traitement. Ces étapes découlent l'une de l'autre et sont rappelées ci-dessous :

- La description de diverses variables de l'interlangue selon leur catégorisation en groupes temporels créés sur la base de la vitesse de lecture selon des archétypes de conformité : expérimentation STAT-NN2 (section 4.5),
- La détermination de profils prosodiques d'apprenants selon un profilage par archétypes stéréotypiques : expérimentation STAT-NN3 (section 4.6),

- La détermination de critères prosodiques distinctifs (*criterial features*) dans l'optique de créer un barème évaluatif : expérimentation CRIT-LEC-NN (section 4.7).

La section 4.7 a listé les traits prosodiques distinctifs pressentis et les prototypes de grilles d'évaluation créées à l'issue de ce chapitre. Ainsi, elle valide de fait H4. Elle est le fruit de la méthodologie, rebaptisée « profilage », mise en œuvre depuis le début de cette thèse et prend en compte les quatre grands domaines prosodiques. Elle a permis non seulement de catégoriser mais encore de hiérarchiser les variables prosodiques quantitativement et qualitativement.

Parmi les constatations résultant des points mentionnés, on soulignera que la caractérisation traditionnelle des apprenants par la vitesse de lecture ne correspond pas à la hiérarchie qualitative en matière prosodique. Sans doute les plus lents, représentés par D9, sont-ils peu performants et se situent-ils globalement au bas du classement prosodique. Il faut pourtant noter que les meilleurs résultats prosodiques appartiennent essentiellement aux catégories Q1 et Q3 et que les apprenants rapides de D1 sont rarement les meilleurs. Il en résulte que la catégorisation par rapidité de lecture ne fédère pas suffisamment les autres domaines prosodiques pour être discriminante à elle seule. Bref, les apprenants semblent devoir être hiérarchisés et évalués selon les quatre grands domaines prosodiques que sont la temporalité, le registre, la mélodie et le rythme, et non selon le seul critère de rapidité d'élocution.

Les expérimentations de ce chapitre ont en effet montré que le niveau de maîtrise prosodique variait selon le domaine considéré. Si les « bons » le sont souvent dans de nombreux domaines (sauf en rapidité de lecture, comme on a pu le constater), ce n'est pas le cas des plus faibles puisqu'on a pu observer une variation des plus faibles selon la variable ou le domaine étudié (les « mauvais » ne sont pas toujours les mêmes). Ainsi, l'étude du segment à l'emphase réduite S9 représente un critère différent des autres : c'est parfois la seule variable source de bonus chez les apprenants faibles. Cela indique le caractère non révélateur de ce critère par rapport aux autres dans les circonstances données : la lecture du texte *The Selfish Giant* par des étudiants en début de première année de licence. De même, *all* est fort probablement le relais de *so* dans des circonstances et à un niveau différents. Dans les grilles évaluatives, le barème de notation dépend du niveau d'exigence requis pour chaque variable, lequel est déterminé en fonction du niveau de départ des apprenants et de la progression

CHAPITRE IV – Expérimentation de création de profils et de critères d'évaluation prosodiques d'apprenants

pédagogique de la formation. Ici, les repères évaluatifs ont été établis d'après ceux déterminés par analyse acoustique ou perceptive de la cohorte d'apprenants de départ, ainsi que par analyse des productions natives.

Aussi, il faut être très prudent quant aux résultats de cette recherche : il ne s'agit là que de prototypes modalisateurs destinés à offrir une base générale tangible dans un profilage évaluatif devant être précisé dans un esprit de transférabilité. L'objectif à court terme est la création de repères distinctifs à un niveau donné pour la lecture d'un texte, dont le style de parole est d'une typologie particulière puisqu'il s'agit d'un conte. Le niveau d'acceptabilité est à déterminer par la suite car il dépend de très nombreux facteurs suivant les conditions d'évaluation et des objectifs assignés à l'opération évaluative.

On a vu que les grilles évaluatives incluent les natifs anglophones. En effet, ils forment le sommet de la pyramide à laquelle ils sont intégrés. Ces grilles sont présentées de manière à former un continuum des réalisations natives et non natives au lieu de les catégoriser distinctement. Il est le prolongement des analyses effectuées aux Chapitres II et III sur la variation des réalisations natives. On observe un chevauchement entre celles des natifs les moins expérimentés et les non natifs les plus expérimentés.

On peut constater que ce chapitre contribue à apporter des réponses aux questions qui sont à la base de cette thèse. Ces réponses ouvrent à leur tour la voie à d'autres questions destinées à affiner la compréhension des phénomènes. Toutes ces questions et remarques prennent source dans les choix méthodologiques mis en œuvre dans cette recherche. D'un point de vue « technique », alors que les variables de rapidité, de registre et de rythme sont ordinairement basées sur des corrélats acoustiques et perceptifs, comme c'est le cas ici, il m'a fallu établir un profilage particulier pour rendre compte de la qualité mélodique de manière simple mais rigoureuse et objective. L'analyse comparative des Chapitres II et III m'a permis d'établir une méthodologie évaluative basée sur un système de forfait initial sur lequel se greffent des points de bonus et de malus. On a d'ailleurs pu se demander si le niveau d'un apprenant était davantage déterminé par ses points forts ou ses points faibles. L'analyse sur tableur *Excel* m'a conduite à trouver des réponses plus claires et tranchées en choisissant de privilégier les points de malus, c'est-à-dire d'échec.

Au regard des résultats obtenus, on peut se demander ce qui caractérise encore plus précisément un apprenant de haut niveau ou faible, où se situent les points forts, quels sont les

traits ou domaines prosodiques les plus révélateurs d'un niveau déterminé, les profils les plus prometteurs. Ces points feront l'objet d'études ultérieures.

Ce chapitre pourrait finaliser cette étude. Pourtant, l'opération de profilage se doit d'aboutir à une évaluation des résultats obtenus afin d'en apprécier l'adéquation aux objectifs de départ. Il est donc nécessaire de poursuivre le profilage pour vérifier sa pertinence et son résultat par une validation de bon aloi en confrontant les résultats du classement obtenu par variables semi-automatiques à ceux que donneraient des évaluateurs traditionnels par perception. C'est l'objectif que se fixe le Chapitre V.

CHAPITRE V — Vérification de la fiabilité du système d'évaluation prototypique semi-automatique par variables

1 Résumé du Chapitre V

Après les quatre premiers chapitres, dont l'objectif était de parvenir à créer un système évaluatif semi-automatique de la prosodie anglaise des apprenants francophones, ce cinquième chapitre a pour fonction de vérifier le degré d'adéquation de la hiérarchie opérée par l'application de variables avec celui par perception classique (Question de Recherche 5). C'est ce qu'explique la section 5.2.

Ce chapitre V se propose d'atteindre son objectif en suivant une méthodologie en trois grandes phases :

- Étude perceptive : profilage par catégorisation perceptive des réalisations non natives à visée comparative (expérimentation PER-LEC-NN) (section 5.3.1), profilage hiérarchique puis catégoriel des réalisations non natives par variables en symétrie avec le profilage perceptif (section 5.3.2), analyse des résultats obtenus selon le profilage perceptif (section 5.3.3),
- Étude par variables : le profilage par variables (section 5.4),
- Comparaison des résultats obtenus selon les deux méthodes : par variables et par perception (expérimentation STAT-NN6) (section 5.5).

Pour les quinze fichiers représentatifs de l'ensemble du corpus non natif, l'analyse comparative des résultats en trois catégories de classement conclut à :

- une validation à 56,83% par la totalité des 68 experts,
- une validation à 59,10% par les 40 experts francophones,

et pour les experts profilés selon la durée du test de perception et celle de leur expérience dans le domaine de l'enseignement :

- une validation à 59,74% par les onze experts francophones profilés (meilleur score),
- une validation à 59,18% par les sept experts anglophones profilés,
- une validation à 59,52% par les 18 experts profilés.

On remarque en outre la prépondérance des variables de certains domaines sur d'autres.

Le processus de profilage des variables évaluatives aboutit à un passage obligé final : la vérification de l'adéquation des résultats aux objectifs de départ, c'est-à-dire ici la comparaison des résultats obtenus d'une part par évaluation semi-automatique avec d'autre part ceux obtenus par évaluation perceptive classique. Cette adéquation prend la forme de la Question de Recherche QR2, rappelée ci-dessous.

QR2 — À quel degré de fiabilité les variables à la base des profils prosodiques d'apprenants et de leurs grilles évaluatives sont-elles validées par comparaison de leurs résultats à ceux d'un test perceptif effectué par des experts ?

Ainsi doivent être validées l'efficacité et la pertinence des résultats de la recherche effectuée au cours de cette thèse.

2 Objectif, enjeu et profilage de la validation d'un classement semi-automatique par adéquation avec un classement perceptif classique

La ligne directrice de cette thèse est donnée par son objectif : trouver des critères d'évaluation pouvant discriminer qualitativement la prosodie anglaise non native. Les chapitres précédents ont décrit et mis en œuvre divers processus méthodologiques destinés à atteindre cet objectif. Dans le Chapitre II sur les profils, on a vu que, selon Chapman *et al.*, l'aboutissement du modèle de référence *CRISP-DM* devait pouvoir se traduire selon la formule ([1999] 2000 : 51) :

RESULTS = MODELS + FINDINGS

Au stade présent de cette recherche, on pourrait sans doute pouvoir la traduire selon la formule suivante en soulignant que le profilage de cette étude atteint son stade ultime.

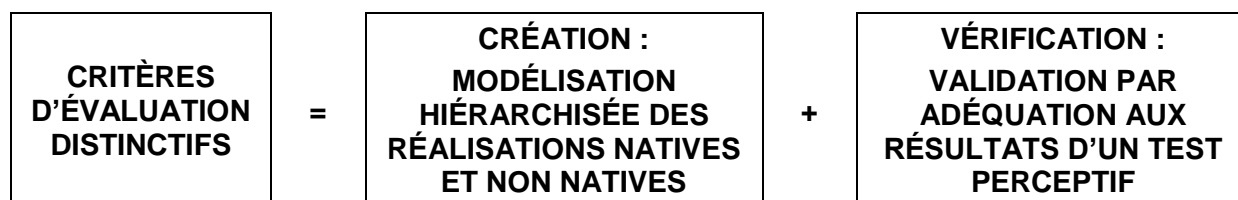


Figure 197 – Modélisation de la dernière phase du profilage des traits distinctifs évaluatifs

La hiérarchie des réalisations se vérifie selon deux modes :

- Mode continu individuel : le profil personnalisé d'un fichier son d'apprenant permet de le classer hiérarchiquement selon ses attributs propres parmi d'autres individus sur un axe évaluatif ; le fichier son se définira par sa place au sein de ce continuum par rapport à celles qu'occupent les autres fichiers ;
- Mode catégoriel : un profilage distributif catégorisera un fichier son par sa conformité avec un groupe catégoriel dont il possèdera les traits distinctifs.

Ce Chapitre V tente de vérifier la pertinence des variables suprasegmentales, qui sont le fondement des traits distinctifs évaluatifs pour non natifs dégagés au Chapitre IV selon ces deux modes hiérarchiques. Il s'agit à présent de valider les résultats obtenus par application de ces variables en termes de hiérarchisation des réalisations non natives pour les quinze apprenants (expérimentation STAT-NN3).

L'enjeu de cette analyse est double :

- Vérifier la validité du résultat de l'évaluation par variables essentiellement semi-automatiques appliquées aux réalisations des quinze apprenants par confrontation avec son degré de similarité avec les résultats d'une perception classique ;
- Vérifier le degré d'importance des quatre domaines prosodiques regroupant les variables par rapport au classement opéré par perception.

Afin de comparer les résultats à ceux d'une évaluation traditionnelle, un test de profilage perceptif doit être mis en place. Il faut aussi s'assurer de l'adéquation entre les résultats provenant des deux sources. En effet, les données obtenues via variables et via perception doivent être analysables comparativement : pour que leurs résultats soient validés, les expérimentations doivent en premier lieu être faisables.

Ayant fixé les objectifs à atteindre, on peut à présent s'orienter vers la mise en place d'un profilage perceptif des réalisations non natives, ensuite vérifier son adéquation au profilage par variables, puis procéder à l'analyse comparative des résultats obtenus.

3 Étude perceptive (expérimentations PER-LEC-NN et STAT-NN4)

Cette section détaille le profilage perceptif des réalisations non natives (expérimentation PER-LEC-NN), la transformation de ces résultats sous forme de variables puis leur analyse (expérimentation STAT-NN4).

3.1 Profilage perceptif des réalisations non natives à visée comparative (expérimentation PER-LEC-NN)

Afin de pouvoir évaluer la validité des grilles expérimentales, il est nécessaire d'établir une correspondance entre les résultats de l'évaluation des réalisations non natives par variables et ceux de celle opérée traditionnellement par des enseignants évaluateurs. Dans cette visée est présentée l'expérimentation PER-LEC-NN avec centrage sur le profilage du support sonore et les biais inhérents. Ensuite est exposé le profilage des participants au test, suivi des modalités d'administration de ce test.

3.1.1 Profilage du support sonore à évaluer et biais

On a vu que l'évaluation par grilles critériées reposait sur l'analyse de 25 variables suprasegmentales représentant les quatre domaines prosodiques et répartis tout au long du texte original du *Selfish Giant*. Il est difficile de trouver des volontaires parmi les enseignants d'anglais francophones ou anglophones si le temps passé pour accomplir le test est supérieur à 20 minutes. C'est la raison pour laquelle j'ai opté pour la concision et procédé à un découpage raccourcissant le texte de départ. Le profilage a été effectué en deux phases.

3.1.1.1 Version 1 du profilage du support sonore

Dans sa phase 1, le profilage avait pour visée une durée comprise dans une fourchette de 60 à 90 secondes pour chaque fichier sonore. Pour cela, j'ai opéré un élagage du texte correspondant à deux impératifs :

- Garder un maximum de traits distinctifs pour que les évaluateurs puissent les repérer subjectivement et baser leur jugement sur ces points déterminants,
- Préserver la cohérence diégétique en conservant les articulations nécessaires à la compréhension du conte et des choix interlocutifs et prosodiques effectués par l'apprenant dans sa lecture.

Le résultat obtenu est présenté dans le tableau ci-dessous.

CHAPITRE V — Vérification de la fiabilité du système d'évaluation prototypique semi-automatique par variables

Domaine prosodique	Variables évaluées par grilles	Présence dans le test perceptif	Domaine prosodique	Variables évaluées par grilles	Présence dans le test perceptif
Vitesse de lecture	DUR_T	texte tronqué	Mélodie	MEL_Excl	oui
	DUR_ABCp	oui		MEL_P5	oui
	DUR_ABC	oui		MEL_S9	oui
	DEB_T	texte tronqué		MEL_s/\	oui
	DEB_ABCp	oui		MEL_s\	oui
	DEB_ABC	oui		MEL_no	oui
Nombre de critères	6 variables (sur 14 points)	6 variables (sur 14 points)	Nombre de critères	6 variables (sur 6 points)	6 variables (sur 6 points)

Domaine prosodique	Variables évaluées par grilles	Présence dans le test perceptif	Domaine prosodique	Variables évaluées par grilles	Présence dans le test perceptif
			Rythme	RYT_lex	3 polysyllabes
Registre	REG_Excl	oui		RYT_so4	oui
	REG_P5	oui		RYT_so22	oui
	REG_so4	oui		RYT_all12	oui
	REG_so22	oui		RYT_all16	oui
	REG_S9	oui		RYT_them	oui
	REG_no	oui		RYT_syll	non
Nombre de critères	6 variables (sur 6 points)	6 variables (sur 6 points)	Nombre de critères	7 variables (sur 12 points)	6 variables (sur 8 points)

Tableau 104 – Vérification de la similarité des critères distinctifs évalués par 24 variables et par test perceptif PER-LEC-NN (Phase 1)

Le tableau précédent montre que les critères évaluatifs critériés ont tous été préservés intégralement dans la version tronquée à l'exception :

- De la vitesse de lecture puisque le texte est écourté,
- Du rythme en deux points :
 - suppression de la variable RYT_syll ;
 - RYT_lex est évalué sur trois polysyllabes au lieu de six ; les mots *conversation, limited, determined* sont absents du montage.

L'objectif étant de comparer le résultat du traitement évaluatif à partir des grilles critériées d'une part, et de la perception d'enseignants évaluateurs d'autre part, il est peu important que le texte soit entier car le style de lecture est préservé par une longueur suffisante : 159 mots dans la version évaluée perceptivement sur les 426 mots du texte de départ, c'est-à-dire

37,3%. Par ailleurs, la variable rythmique RYT_syll n'avait pas donné les résultats positifs attendus lors de sa mise en application sur les réalisations des apprenants. De plus, les polysyllabes *conversation* et *limited* avaient été jugés peu discriminants, leur patron accentuels ayant été reproduit à 60% pour le premier et 53,33% pour le second. Seul *determined* (13,33% de réussite) semble vraiment discriminant dans ce corpus.

Il faut ajouter à ces biais celui inhérent au montage lui-même. Les coupures ont été opérées au moment des pauses et, bien que les durées aient été appréciées subjectivement, un soin particulier m'a guidée pour garder une grande homogénéité dans le style de lecture.

Somme toute, ces biais constatés dans le montage de chaque fichier son, bien que réels, me semblent pourtant négligeables par rapport à leur intérêt dans la facilitation du travail évaluatif des enseignants volontaires. Le bilan des fichiers sons de l'expérimentation PER-LEC-NN dans sa Phase 1 fait état de 24 variables mesurées sur 34 points au lieu des 25 variables des grilles critériées mesurées sur 38 points.

Pour classer les fichiers dans une optique évaluative, il était nécessaire de trouver un angle d'attaque. Évaluer leur qualité prosodique les uns par rapport aux autres signifiait les écouter tous plusieurs fois, ce qui aurait été trop chronophage et dissuasif. Prendre comme point de repère le niveau qu'un étudiant doit ou devrait avoir à son entrée à l'université était encore plus subjectif étant donné la très grande hétérogénéité constatée chez les spécialistes de langue tout autant que chez les spécialistes d'autres disciplines (LANSAD). Un niveau B1 ou B2 selon l'échelle du *Cadre européen commun de référence pour les langues* l'aurait été tout autant étant donné la très grande habitude qu'un évaluateur doit avoir pour pouvoir juger la qualité prosodique uniquement. Aussi, j'ai opté pour un système interne où c'est l'un des fichiers qui servirait à l'étalonnage des autres. Une difficulté était de sélectionner ce fichier de « repère » à l'organisation du test pour représenter le niveau « neutre » à partir duquel les quatorze autres fichiers seraient étalonnés. On a vu dans le Chapitre IV que les quinze lectures avaient été évaluées selon un « bilan de réussite » et un « bilan d'échec », la soustraction de leurs données respectives résultant en un « bilan global ». Le fichier « repère » a été sélectionné en tenant compte de sa place centrale dans chacun de ces bilans : c'est le fichier DID0153-S001R.

Bilan positif (réussite)		Bilan négatif (échec)		Bilan global	
67,41	DID0168-S001R (S10)	4,16	DID0168-S001R (S10)	63,24	DID0168-S001R (S10)
50,89	DID0210-S001R (S1)	14,28	DID0210-S001R (S1)	36,60	DID0210-S001R (S1)
42,55	DID0078-S001R (S2)	17,41	DID0108-S001R (S6)	24,10	DID0108-S001R (S6)
41,51	DID0108-S001R (S6)	23,95	DID0024-S001R	3,27	DID0078-S001R (S2)
31,10	DID0135-S001R	25,74	DID0003-S001R	-1,78	DID0003-S001R
25	DID0018-S001R	25,74	DID0068-S001R	-7,58	DID0068-S001R
23,95	DID0003-S001R	29,91	DID0046-S001R	-9,37	DID0024-S001R
23,80	DID0153-S001R	35,71	DID0153-S001R	-11,90	DID0153-S001R
18,15	DID0068-S001R	39,28	DID0078-S001R (S2)	-15,47	DID0135-S001R
14,58	DID0024-S001R	46,57	DID0135-S001R	-19,04	DID0046-S001R
10,86	DID0046-S001R	50,74	DID0018-S001R	-25,74	DID0018-S001R
7,29	DID0074-S001R (S11)	51,33	DID0062-S001R	-51,19	DID0056-S001R
4,16	DID0056-S001R	55,35	DID0056-S001R	-51,33	DID0062-S001R
3,12	DID0076-S001R (S15)	62,20	DID0076-S001R (S15)	-59,07	DID0076-S001R (S15)
0	DID0062-S001R	69,79	DID0074-S001R (S11)	-62,5	DID0074-S001R (S11)

Tableau 105 – Résultats des trois bilans de réussite, échec et bilan global par analyse semi-automatique à partir des traits discriminants

Un pré-test sur une enseignante francophone volontaire a montré que l'expérimentation projetée ne pouvait être opérée sérieusement en moins d'une heure car il fallait souvent réécouter le fichier témoin. De plus, les fichiers étaient trop longs pour pouvoir établir une comparaison effective (« on peut comparer une phrase, deux ou trois maxi, mais il y a trop de variables sur des extraits aussi longs : des passages peuvent être très réussis et d'autres catastrophiques pour le même lecteur, comment comparer ? »). Le résultat qui en découlait était « vraiment très (trop) subjectif » et aléatoire. Cette première tentative m'a conduite à ne pas poursuivre l'administration de ce pré-test, mais à procéder directement à une seconde étape dans l'opération de profilage.

3.1.1.2 Version 2 du profilage du support sonore

La seconde phase du profilage a eu comme objectifs de réduire encore davantage la longueur des fichiers, le nombre de variables testées, et d'opérer un classement des variables pour faciliter la prise de décision évaluative.

La réduction de la longueur du fichier a signifié supprimer les phrases où ne figuraient aucune variable, ce qui a fortement réduit la dimension diégétique. En outre, le nombre de variables testées a été réduit lui aussi, comme le spécifie le tableau suivant.

Domaine prosodique	Variables évaluées par grilles	Présence dans le test perceptif	Domaine prosodique	Variables évaluées par grilles	Présence dans le test perceptif
Vitesse de lecture	DUR_T	5 phrases / 26	Mélodie	MEL_Excl	oui
	DUR_ABCp	2 tronçons / 3		MEL_P5	oui
	DUR_ABC	2 tronçons / 3		MEL_S9	non
	DEB_T	5 phrases / 26		MEL_s\	non
	DEB_ABCp	2 tronçons / 3		MEL_s\	non
	DEB_ABC	2 tronçons / 3		MEL_no	non
Nombre de critères	6 variables (sur 14 points)	6 variables (10 points / 14)	Nombre de critères	6 variables (sur 6 points)	2 variables (2 points / 6)

Domaine prosodique	Variables évaluées par grilles	Présence dans le test perceptif	Domaine prosodique	Variables évaluées par grilles	Présence dans le test perceptif
Registre	REG_Excl	oui	Rythme	RYT_lex	1 polysyllabe / 6
	REG_P5	oui		RYT_so4	oui
	REG_so4	oui		RYT_so22	oui
	REG_so22	oui		RYT_all12	oui
	REG_S9	non		RYT_all16	oui
	REG_no	non		RYT_them	oui
				RYT_syll	non
Nombre de critères	6 variables (sur 6 points)	4 variables (4 points / 6)	Nombre de critères	7 variables (sur 12 points)	6 variables (6 points / 12)

Tableau 106 – Vérification de la similarité des critères distinctifs évalués par 18 variables et par test perceptif PER-LEC-NN (Phase 2)

Malgré sa longueur tronquée, le texte garde au domaine temporel ses six variables mesurées sur dix points au lieu de quatorze, le registre se réduit à quatre variables au lieu de six, mesurées sur quatre points au lieu de six regroupant la Phrase 5 contenant l'exclamative et les extraits avec les paires « *so* + adjectif ou adverbe ». La mélodie est encore davantage restreinte puisqu'elle ne rend compte que de l'exclamative avec et sans son discours citant (Phrase 5). Le rythme, quant à lui, est davantage représenté : six variables au lieu de sept, même si le nombre de points de mesure décroît de moitié (six au lieu de douze).

Alors que la totalité des variables trouvées au départ du Chapitre IV est au nombre de 28, le profils établis statistiquement comprennent 25 d'entre elles, et à présent, le test perceptif se basera sur 18 variables, dont certaines sont amputées d'une partie des points étudiés plus ou moins importante, comme nous l'avons constaté ci-dessus. Ne pouvant

CHAPITRE V — Vérification de la fiabilité du système d'évaluation prototypique semi-automatique par variables

comparer que ce qui est comparable, on se doit de vérifier les hiérarchies créées à partir des 25 et des 18 variables en les comparant.

	Bilan global 25 variables			Bilan global 18 variables	
	63,2	DID0168-S001R (S10)		79,2	DID0168-S001R (S10)
	36,6	DID0210-S001R (S1)		31,3	DID0210-S001R (S1)
	24,1	DID0108-S001R (S6)		20,8	DID0108-S001R (S6)
	3,2	DID0078-S001R (S2)		2,08	DID0068-S001R
	-1,7	DID0003-S001R		1,25	DID0024-S001R
	-7,5	DID0068-S001R		-4,17	DID0078-S001R (S2)
	-9,3	DID0024-S001R		-20,8	DID0046-S001R
	-11,9	DID0153-S001R		-31,3	DID0003-S001R
	-15,4	DID0135-S001R		-35,4	DID0018-S001R
	-19	DID0046-S001R		-37,5	DID0135-S001R
	-25,7	DID0018-S001R		-39,6	DID0153-S001R
	-51,1	DID0056-S001R		-47,9	DID0076-S001R (S15)
	-51,3	DID0062-S001R		-66,7	DID0056-S001R
	-59	DID0076-S001R (S15)		-70,8	DID0062-S001R
	-62,5	DID0074-S001R (S11)		-83,3	DID0074-S001R (S11)
	126	Empan total		163	Empan total
	126 / 5 = 25,14	Empan de chacune des 5 catégories de classement		163 / 5 = 32,5	Empan de chacune des 5 catégories de classement

Tableau 107 – Hiérarchisation de la qualité prosodique obtenue par application du profilage par 25 et par 18 variables

J'ai gardé la catégorisation en cinq groupes de niveaux colorisés pour une meilleure visualisation du vert foncé (*much higher ability*) au rouge (*much lower ability*) en passant par le vert clair (*higher ability*) et l'orangé (*lower ability*). L'équivalence de niveau prosodique avec le fichier père (*similar to File 0*) se traduit par une absence de couleur.

Dans le tableau de la catégorisation avec 18 variables, l'empan général s'est accru fortement (de 126 à 163 points). On remarque pourtant une robustesse dans le classement du haut et du bas du tableau malgré quelques glissements au centre. Le fichier DID0168-S001R (S10) reste le meilleur (*much higher ability*) et DID0210-S001R (S1) et DID0108-S001R (S6) sont toujours les seuls à être qualifiés de *higher ability*. Dans le bas du tableau les lectures DID0056-S001R, DID0062-S001R et DID0074-S001R (S11) restent les moins performantes (*much lower ability*). Par contre, le fichier DID0076-S001R (S15) se soustrait des plus faibles pour intégrer la catégorie supérieure *lower ability*. Les fichiers DID0003-S001R et DID0153-

S001R le rejoignent aussi, en glissant de la catégorie *similar to File 0*, ce qui double l'effectif de la catégorie *lower ability* (six éléments sur quinze, c'est-à-dire 40% de l'effectif total au lieu de 20%). La catégorie *similar to File 0* perd son caractère central en ce qui concerne la répartition des fichiers les uns par rapport aux autres. Le fichier repère n'est plus le DID0153-S001R mais le DID0024-S001R. Ce dernier conserve ses qualités par sa moyenne alors que DID0153-S001R est rétrogradé dans la catégorie *lower ability*. Étant donné la stabilité de sa catégorisation, il semble que le fichier DID0024-S001R soit un repère plus robuste que le premier. Il reste préférable au fichier DID0078-S001R (S2), ayant une moyenne de -4,17 et fortement déterminé par la rapidité de lecture. En fait, 80% des fichiers occupent toujours la même catégorie qu'avec un classement par 25 variables.

3.1.2 Profilage des participants au test

Afin d'évaluer la qualité prosodique et ainsi que le présente le questionnaire, ce test cible comme participants des enseignants d'anglais en langue étrangère, voire seconde. Les 68 auditeurs-experts sont tous anglicistes. Trois catégories sont déterminées : des enseignants anglicistes francophones, des anglophones natifs et des enseignants d'anglais natifs d'autres langues que le français ou l'anglais.

3.1.2.1 Enseignants anglicistes francophones

Les enseignants anglicistes francophones sont au nombre de 40. Parmi eux se trouvent des enseignants de l'enseignement secondaire affectés dans l'enseignement supérieur ou retraités, des enseignants-chercheurs (maîtres de conférence ou professeurs des universités). Certains sont doctorants. Leur nombre d'années passées à enseigner l'anglais s'échelonne de 4 à 43 ans. Les spécialistes oralistes chercheurs en équipe d'accueil ou en unités mixtes de recherche sont au nombre de treize (33% de l'effectif).

3.1.2.2 Anglophones natifs

Au nombre de 25, les anglophones natifs sont tous enseignants à l'exclusion d'un seul : il s'agit d'un chercheur en phonétique anglaise. Ce sont des lecteurs et, comme précédemment, des enseignants de l'enseignement secondaire affectés dans l'enseignement supérieur ou retraités, des enseignants-chercheurs (maîtres de conférence ou professeurs des universités). Sept d'entre eux sont spécialistes de l'oral en anglais dans des équipes d'accueil ou en unités mixtes de recherche (29,16%).

3.1.2.3 Enseignants d'anglais natifs d'autres langues que le français ou l'anglais

Les trois enseignants d'anglais de langue maternelle non anglaise ni française (arabe, persan, chinois mandarin) poursuivent des études de linguistique axées sur la phonétique.

3.1.3 Modalités d'administration du test

L'objectif visé par ce questionnaire est de recueillir les données évaluatives nécessaires renseignées par le plus grand nombre de participants correspondant au profil recherché. En conséquence, les points étudiés sont le contenu du questionnaire, sa transmission aux éventuels participants (c'est-à-dire le repérage d'éventuels participants et l'ergonomie de l'interface) et bien sûr la récupération des données.

3.1.4 Recueil des données

Comme pour chacune des étapes mentionnées précédemment, l'opération de recueil des données s'est déroulée manuellement. Le faible nombre de données à recopier n'a pas rendu la tâche fastidieuse. Au contraire, son mode d'intégration manuelle sur tableur a permis de vérifier la pertinence de chacune d'entre elles ou leur absence. Ce contact privilégié avec les données, presque « physique », permet à mon sens d'émettre des hypothèses auxquelles on ne songerait pas si l'on obtenait d'emblée un fichier complet et que l'on tentait de vérifier les données instantanément.

La difficulté est de se créer des garde-fous qui conditionnent l'absence d'erreur sans perte de temps. Je l'ai surmontée en codant chaque réponse évaluative selon l'échelle présentée ci-dessous en apposant le nombre correspondant à la réponse donnée à côté de la croix colonne par colonne, ce qui m'a permis de limiter mes vérifications de recopiage. La nature du fichier texte m'a permis de coder les réponses sur le support même que l'on m'envoyait en gardant les informations originales intactes.

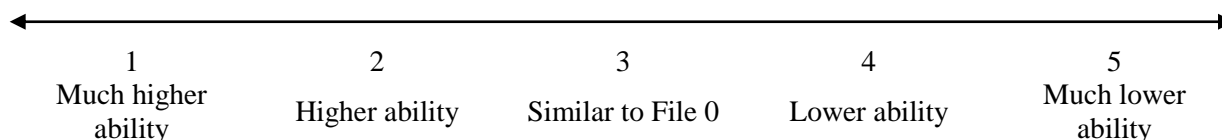
3.1.4.1 Profilage du contenu du questionnaire

Il a été demandé à chacun des enseignants volontaires de compléter un questionnaire sur leur appréciation du niveau prosodique des apprenants. On se souvient que, afin de favoriser l'émergence de points communs entre les données qualitatives correspondant aux variables des quinze apprenants dans les plans factoriels, une catégorisation grossière avait été effectuée en trois ou quatre catégories de préférence, parfois cinq lorsque les données l'exigeaient. Ce questionnaire n'est pas destiné à regrouper les résultats, mais au contraire à

les différencier afin de mieux les hiérarchiser. La finalité est donc de classer chaque enregistrement dans une des cinq catégories. C'est ce que demande l'énoncé suivant.

You are going to listen to 15 first year students reading a few sentences from *The Selfish Giant* by Oscar Wilde. Your aim is to grade their suprasegmental level. Please focus on fluency, intonation and register, word stress and rhythm. For each of them, tick the corresponding grade.

Dans cette optique, l'échelle est constituée de cinq graduations, ce qui permet d'atteindre une certaine précision sans trop complexifier la tâche.



La position centrale est occupée par une moyenne de bon aloi : un fichier « repère », qui évite de recourir à des notions floues, telles que :

- le niveau suprasegmental d'un apprenant de niveau « moyen » à son entrée à l'université. Je ne fais pas de distinction entre spécialistes de l'anglais et spécialistes d'autres disciplines (LANSAD) étant donné que cette différence se retrouverait davantage dans les habitudes de notation des enseignants des diverses UFR ou services communs (Maison des Langues), plutôt que dans les résultats des étudiants eux-mêmes. En effet, on trouve parmi les étudiants LANSAD d'excellents anglicistes. De plus, le niveau des spécialistes se répartit selon une fourchette très large.
- un niveau B1 ou B2, niveaux requis ou à atteindre dans de nombreux tests, certifications ou dossiers de candidature. Si certains enseignants sont aguerris à la tâche d'évaluer un niveau global, il est beaucoup plus aléatoire d'effectuer une évaluation suprasegmentale en lecture, d'autant que les critères distinctifs pour la langue anglaise sont limités, qu'ils proviennent du *Cadre européen commun de référence* ou d'autres sources (exemple : grilles évaluatives de certaines certifications).

Il aurait été intéressant de préciser indépendamment chaque domaine évalué (temporel, mélodique, rythmique, de registre), mais cela aurait complexifié la tâche en augmentant considérablement sa durée. Deux écueils auraient alors été prévisibles :

démotivation des enseignants évaluateurs, voire impossibilité d'accomplir la tâche pour certains d'entre eux non formés à dissocier ces éléments. Par ce test, c'est la correspondance globale avec l'évaluation par variables qui est attendue.

Aussi, c'est l'échelle de différenciation sémantique d'Osgood qui a été sélectionnée pour quantifier les informations d'ordre qualitatif et aboutir à un score numérique, qui pourra être comparé aux résultats obtenus par évaluation semi-automatique. L'échelle d'Osgood s'apparente par ses échelons aux grilles critériées auxquelles elles doivent être comparées ensuite. Cependant, dans une certaine mesure, la présentation utilisée se réfère aussi à l'échelle de Likert par sa symétrie autour de l'axe étiqueté comme « neutre » et représenté par l'échelon 3. Cette zone est non marquée : elle est comparable à celle utilisée dans le tableur *Excel* délimitant les valeurs plus ou moins extrêmes représentées par les autres échelons. Cette position de repli, non marquée, sert aussi de modalité neutre « ni-ni » pour tout auditeur incapable d'orienter son jugement vers un des extrêmes. Elle sert de « repère » à l'organisation du test. Nous avons vu précédemment que c'est un montage du fichier DID0024-S001R qui avant été sélectionné pour représenter ce niveau « neutre » lorsque l'évaluation portait sur 18 variables.

La page de réponse est constituée par un tableau où l'axe des ordonnées représente le nom des fichiers et celui des abscisses les degrés d'appréciation qualitative. Le Fichier 0 « repère », rose vif, se trouve au sommet du tableau accompagné d'une croix rose vif elle aussi, correspondant à la mention *similar to File 0* à l'intersection des deux axes. Dans un premier temps, l'auditeur a la possibilité de réécouter le fichier témoin aussi souvent qu'il le souhaite afin d'évaluer les fichiers par rapport à lui. Dans un second temps, il a loisir de réécouter les fichiers qu'il a évalués afin de parfaire leur disposition dans les cases appropriées. Ces deux tâches ne peuvent être accomplies facilement que si l'évaluateur a accès sur un même écran et en même temps aux fichiers à écouter et à la grille de réponses. Le pré-test a montré qu'une présentation des fichiers par ordre entièrement aléatoire devait être évitée. En effet, il était important que l'évaluateur ait conscience dès le départ des extrêmes dont il devait tenir compte : aussi, j'ai placé le fichier ordonné comme le moins bon en seconde position et le meilleur en troisième. Les autres fichiers ne pouvaient se trouver qu'en position comparable ou intermédiaire. Un agencement aléatoire aurait eu comme conséquence de devoir réagencer l'ensemble des évaluations une fois un extrême trouvé. Le surcroît de travail et le temps perdu occasionné par une ergonomie déficiente aurait découragé un grand nombre de participants et faussé les résultats obtenus.

Le questionnaire comprend en outre une demande de renseignements sur métadonnées : le nombre d'années d'enseignement à l'actif de l'enseignant, ainsi que le temps passé à faire le test.

3.1.4.2 *Sélection de l'ergonomie de l'interface*

De nombreux sites proposent de recueillir des données décisionnelles en ligne sous forme de formulaire dans tableaux sur tableurs. L'écoute de fichiers sons limite leur nombre mais deux d'entre eux avaient particulièrement retenu mon attention :

- *SoSci Survey*⁹³, plateforme de questionnaires en ligne, d'utilisation performante et flexible, gratuite pour les projets de recherche non subventionnés par des entreprises,
- *Labguistic*⁹⁴, plateforme payante dédiée aux expérimentations linguistiques en perception de la parole.

Leur prise en main effective ainsi que leur maîtrise complète par l'évaluation de leurs biais aurait nécessité un temps non négligeable et non prévisible. Il aurait aussi fallu vérifier que la présentation du questionnaire sur une seule et même page soit possible. La durée du test aurait pu être calculée automatiquement d'après le temps de connexion, mais l'expérience m'a montré que l'on peut facilement être dérangé pendant cette activité, ce qui m'amène à privilégier une durée mesurée par l'évaluateur lui-même.

Ayant à plusieurs reprises conduit des expérimentations à partir d'un site que j'administre moi-même, j'ai préféré opter pour une sécurité avérée qui me permettait de garder l'ergonomie que je souhaitais par une mise en place rapide. Une autre raison m'a conduite à préférer organiser l'expérimentation de bout en bout sans avoir recours à une plateforme extérieure. Quelques montages avaient été effectués à partir de fichiers du projet *Longdale*. Bien que non reconnaissables ni utilisables, il était nécessaire que soit effacée toute trace de l'ensemble de l'expérimentation une fois celle-ci effectuée. C'est ce que j'ai pu faire à la fin de cette expérimentation en ligne, qui a durée 24 jours. Cela m'a aussi permis de rétablir les pages concernées le lendemain pendant 35 minutes pour permettre à une dernière enseignante qui venait de me contacter de participer au test. Néanmoins, à deux reprises, l'interface non contraignante a posé problème dans le recueil des données : lorsque les

⁹³ [<https://www.soscisurvey.de>], dernièrement consulté en 2017.

⁹⁴ [<http://www.labguistic.com/modules/fmcontent/content.php?topic=static&id=002&page=presentation>], dernièrement consulté en 2017.

auditeurs ont coché deux cases de degré évaluatif adjacentes au lieu d'une. Il s'agit de l'auditeur anglophone IZ (*lower* et *much lower* pour l'apprenant A) et de l'auditeur arabophone ML (*similar to File 0* et *lower* pour l'apprenant C). Dans les deux cas, ne pouvant faire une moyenne des résultats, j'ai opté pour l'option la plus centralisée : *lower* pour IZ et *similar to File 0* chez ML. Il m'a semblé que ce choix est celui qui trahirait le moins l'interprétation finale des résultats statistiques en cas d'erreur.

La perte apparente en ergonomie a été compensée par un ciblage plus performant, réduisant le temps nécessaire pour effectuer une évaluation de qualité, tout en favorisant la vérification des prises de décisions. Sur les 68 participants au test, seulement trois ont rencontré des difficultés avec l'interface retenue. L'éventualité d'une erreur possible, bien que légère pour deux auditeurs, est largement compensée par la quantité des données obtenues grâce au nombre relativement élevé des volontaires participants.

3.1.4.3 Recherche de participants au test et recueil des données

Étant donné le caractère confidentiel de certains fichiers, la spécificité professionnelle recherchée chez les éventuels participants et l'allègement nécessaire du questionnaire, trois angles d'attaque ont été favorisés dans le recrutement de volontaires experts-évaluateurs, leur implication solidaire dans le projet expérimental et la personnalisation du lien .

L'appel à participation d'enseignants d'anglais a été opéré de vive voix ou par messagerie nominale à des collègues de mon entourage personnel, professionnel ou en liaison avec ma recherche. Afin d'élargir le champ des volontaires, j'ai contacté les membres de la société savante dont je suis membre, la *Société des Anglicistes de l'Enseignement Supérieur (SAES)*⁹⁵. J'ai en outre bénéficié de la solidarité de certains collègues, notamment ceux membres de l'*Association des anglicistes pour les études de langue orale dans l'enseignement supérieur, secondaire et élémentaire (ALOES)*, association fille de la SAES, lesquels ont relayé le message à des personnes correspondant au profil recherché.

Les raisons invoquées pour faire appel à la bonne volonté de collègues, que je ne connaissais que peu ou pas du tout, étaient expliquées dans un courriel où j'exposais le bien-fondé de mes travaux de recherche. Ainsi, malgré le caractère rébarbatif de la tâche et l'indisponibilité des collègues, j'ai pu obtenir une bonne participation d'anglicistes qui ne me connaissaient pas mais se sentaient solidaires d'un projet de recherche qu'ils jugeaient utile à

⁹⁵ [<http://saesfrance.org/>] site dernièrement consulté en 2017.

la communauté des enseignants et chercheurs anglicistes et qu'ils m'encourageaient à poursuivre.

Cet enthousiasme n'aurait pas pu être soulevé par une démarche reposant sur l'utilisation d'une plateforme privilégiant l'anonymat des participants. En effet, la filtration de la communication par son caractère nominal a permis de personnaliser la participation, de tenir compte de la traçabilité des données, d'offrir la possibilité de contacter chaque participant pour le remercier individuellement et, si besoin était, de demander des précisions et compléments sur des données absentes du questionnaire.

À chaque étape du processus expérimental, j'ai privilégié non pas l'anonymat d'une plateforme, mais au contraire la relation personnalisée pour permettre une collaboration interactive, participative et solidaire.

3.2 Profilage des réalisations non natives par variables en symétrie avec le profilage perceptif

On a vu au Chapitre IV des exemples de grilles évaluatives qui sont l'aboutissement de l'application de 25 variables quantitatives sur les réalisations non natives. À partir de la hiérarchisation des fichiers créée numériquement dans le fichier *Excel*, une catégorisation prend le pas pour regrouper les résultats avec l'objectif de créer des échelles de valeur qualitatives et évaluatives. Une moyenne des résultats pour chacun des quatre grands domaines prosodiques permet alors de constituer un profil de chaque apprenant par domaine, ce qui relève d'une démarche non seulement évaluative, mais trouve des applications décisives dans une modulation possible de l'enseignement prosodique. À partir de ce profilage, une moyenne des résultats détermine un profil prosodique global de chaque apprenant. C'est ce résultat-là qui devrait être confronté à l'appréciation donnée par les enseignants évaluateurs aux prestations non natives.

Pourtant, pour raison de faisabilité, le profilage PER-LEC-NN de la section précédente ne prend en considération qu'un nombre restreint de variables, dont certaines sont tronquées. C'est donc aux résultats de ce second profilage que ceux de l'expérimentation PER-LEC-NN devront être confrontés. Si les phases successives des profilages mis en œuvre dans cette thèse sont validées, cela doit provenir d'une correspondance plus ou moins étroite entre les résultats des classements des réalisations non natives par évaluateurs professionnels avec ceux obtenus par application de variables, la correspondance attendue étant celle avec les 18 variables. Examinons la méthodologie utilisée pour extraire cette série de résultats à partir des données.

Cette approche numérique a été élaborée de la manière suivante. On a vu que, pour chacun des grands domaines prosodiques, une moyenne des points de bonus et une autre des points de malus avaient été calculées, donnant un poids égal à chaque domaine indépendamment du nombre de variables prises en compte. Un bilan des moyennes de bonus et de malus est alors mis en place, puis les points de malus sont soustraits des points de bonus. On obtient alors un nombre qui peut être positif ou négatif suivant les données de départ. Pour les 18 variables correspondant à celles testées dans l'expérimentation PER-LEC-NN, la moyenne pour chaque fichier s'étalonne de 79,1 à -83,3. L'échelle résulte donc en une graduation positive de 162,5 points, constituée de cinq catégories de 32,5 points. À partir de là, un histogramme peut très simplement catégoriser chaque fichier en délimitant chacune des catégories, représentée comme on l'a vu par un chiffre allant de cinq pour le niveau maximal à un pour le niveau minimal. Pour obtenir un regroupement des fichiers non plus en cinq catégories mais en trois, il suffit d'amalgamer les catégories 1 et 2 d'une part, et 4 et 5 d'autre part. Il ne restera que trois catégories.

3.3 Analyse des résultats obtenus selon le profilage perceptif (expérimentation STAT-NN4)

La présente analyse se base sur les résultats de deux sortes d'approches : l'évaluation traditionnelle par experts et l'évaluation semi-automatique par mesures acoustiques ou perceptives de remplacement. Pour vérifier la correspondance entre ces deux séries de résultats, deux approches sont possibles :

- Celle qui vérifie le degré de conformité des deux séries catégorielles et le présente sous forme de pourcentage ;
- Celle qui consiste à confronter les résultats des classements individuels.

Considérons tout d'abord les classements de réalisations opérés par voie traditionnelle, c'est-à-dire par évaluation d'experts de l'anglais en tant que langue étrangère ou seconde, puis ensuite les classements obtenus par variables. Les divers résultats et classements de réalisations obtenus selon l'approche choisie devraient servir de base de comparaison pour apprécier et affiner ceux générés par approche semi-automatique en renseignant sur la hiérarchisation à privilégier par la suite comme référence parmi les domaines et parmi les variables.

Mais déjà au départ, de par la diversité des experts volontaires, plusieurs profils doivent être considérés selon la langue maternelle des évaluateurs, mais aussi selon le respect des consignes du test ainsi que par une sélection plus fine des évaluateurs. Ainsi, quatre grands groupes peuvent être distingués : la totalité des 68 experts, les 40 experts de langue maternelle française, les 25 experts de langue maternelle anglaise et les 18 experts sélectionnés par profilage spécifique. Ces groupes prennent en compte les cinq niveaux d'appréciation par rapport au fichier repère. Afin de pouvoir comparer les résultats de cette étude à ceux d'autres études, une réduction à trois termes sera opérée ensuite et les groupes aux effectifs les plus et moins nombreux seront étudiés. Considérons les résultats de chaque groupe un à un.

3.3.1 Classement par l'ensemble des évaluateurs (5 catégories)

On a là l'ensemble des 68 participants au test perceptif PER-LEC-NN. L'histogramme ci-dessous montre la répartition des votes pour chacun des quatorze fichiers évalués. Le quinzième étant le fichier repère, il était d'emblée donné de valeur 3, c'est-à-dire moyenne, puisque cette évaluation est relative. J'ai repris les codes couleurs que j'avais utilisés dans le fichier *Excel* : gradation du vert soutenu = 5 (meilleur résultat : *much higher general prosodic ability*) au rouge = 1 (plus faible résultat : *much lower general prosodic ability*). Le blanc reste le repère neutre = 3.

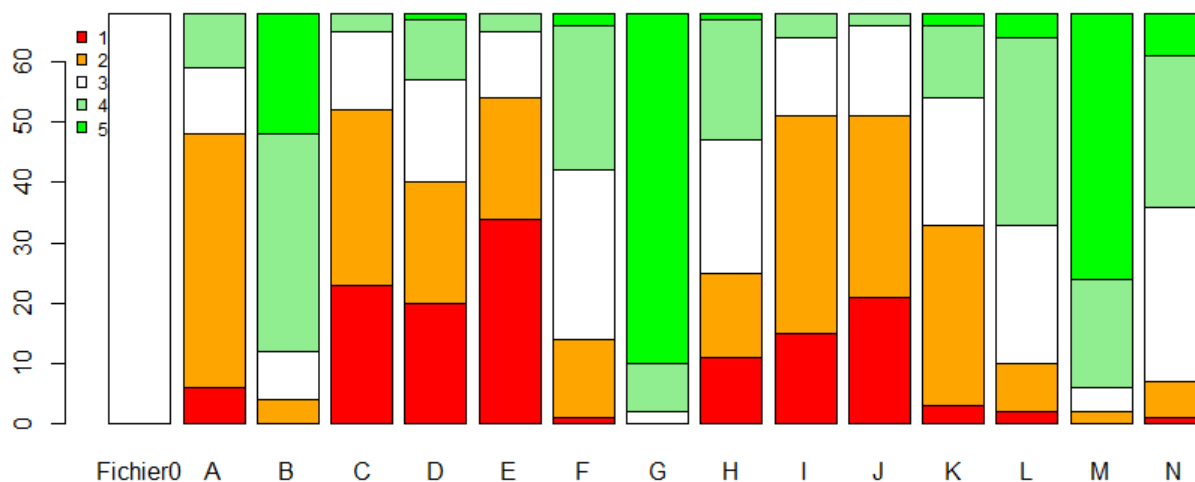


Figure 199 – Résultats de l'évaluation des quinze fichiers d'apprenants (abscisse) par perception des 68 évaluateurs (ordonnées) selon le code : 1 = *much worse*, 2 = *worse*, 3 = *similar to File 0*, 4 = *better*, 5 = *much better*

Ce graphique permet de visualiser d'emblée les codes évaluatifs attribués à chacune des réalisations apprenantes. On perçoit nettement que le meilleur résultat est celui du fichier G (aucune couleur chaude et la plus grande proportion de vert soutenu), suivi de près par le fichier M (vert soutenu de moindre proportion et apparition de coloration orange), puis le B

(moins de couleurs froides et un peu plus d'orangé). Le résultat le plus faible se remarque par son plus haut niveau de rouge, lequel donne aux couleurs chaudes la plus grande importance, et l'absence complète de vert soutenu. Les autres résultats présentent un camaïeu plus ou moins homogène.

L'analyse par composantes multiples, représentée par le graphique ci-dessous, offre une gradation qualitative plus facilement perceptible car le plan se présente sur deux dimensions. Dans la partie positive de la Dimension 1, la partie haute à droite regroupe les deux meilleurs résultats et l'on peut voir que le fichier G est le plus excentré. Avec le fichier M qui le suit, ils se démarquent fortement du reste des autres fichiers sur la Dimension 1. Dans la partie négative de la Dimension 1, toujours sur la partie haute mais à gauche, six fichiers figurent très proches les uns des autres : E, C, J, I, A, D. Leur position sur le plan n'est pas facilement discernable. Pourtant on voit clairement que E est le plus excentré, suivi de C. Les extrêmes se situent dans la partie la plus positive de la Dimension 2. La partie basse du plan, quant à elle, montre qu'à part le Fichier 0, les plus bas dans la Dimension 2 sont N, F et L. Ils sont aussi les plus proches de l'axe de la Dimension 1, ce qui leur donne un caractère « moyen » par rapport à l'ensemble des quinze fichiers. Le Fichier 0 quant à lui occupe une position relativement centrale par rapport à l'emplacement des autres fichiers car il apparaît qu'il y a pratiquement autant de fichiers situés à sa gauche qu'à sa droite sur l'axe de la Dimension 1. Ce plan offre une hiérarchisation qualitative des fichiers les uns par rapport aux autres qui peut être précisée.

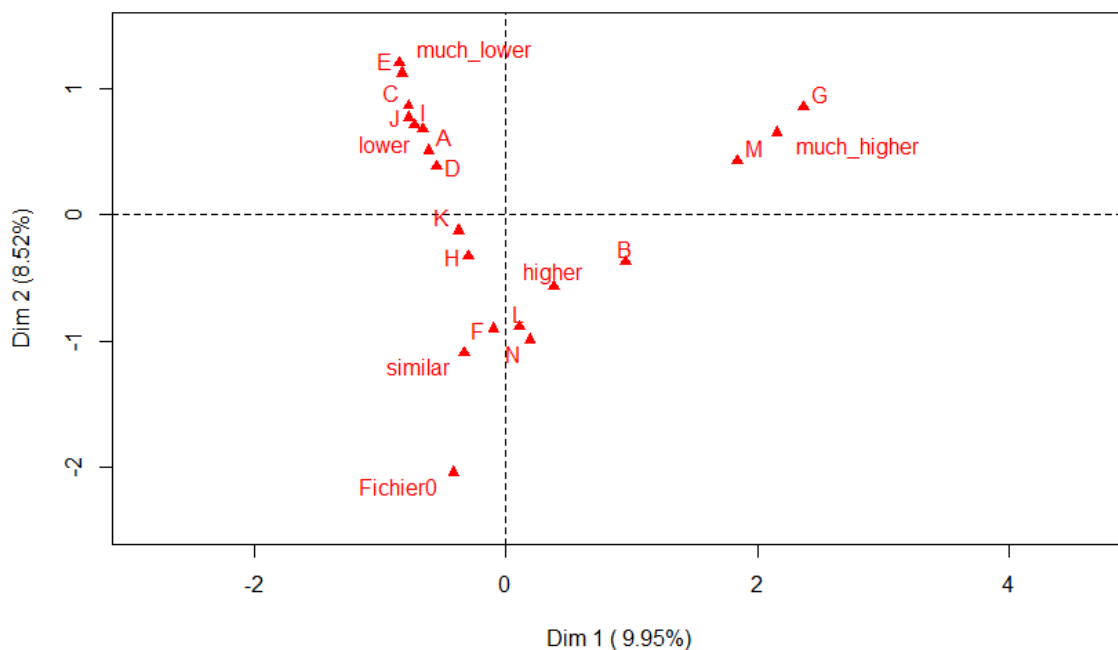


Figure 200 – Plan factoriel de l'analyse des correspondances multiples de l'évaluation perceptive des quinze apprenants par 68 évaluateurs spécialistes

Ci-dessous, l'arbre d'inférence conditionnelle nous renseigne davantage en regroupant les productions similaires hiérarchiquement.

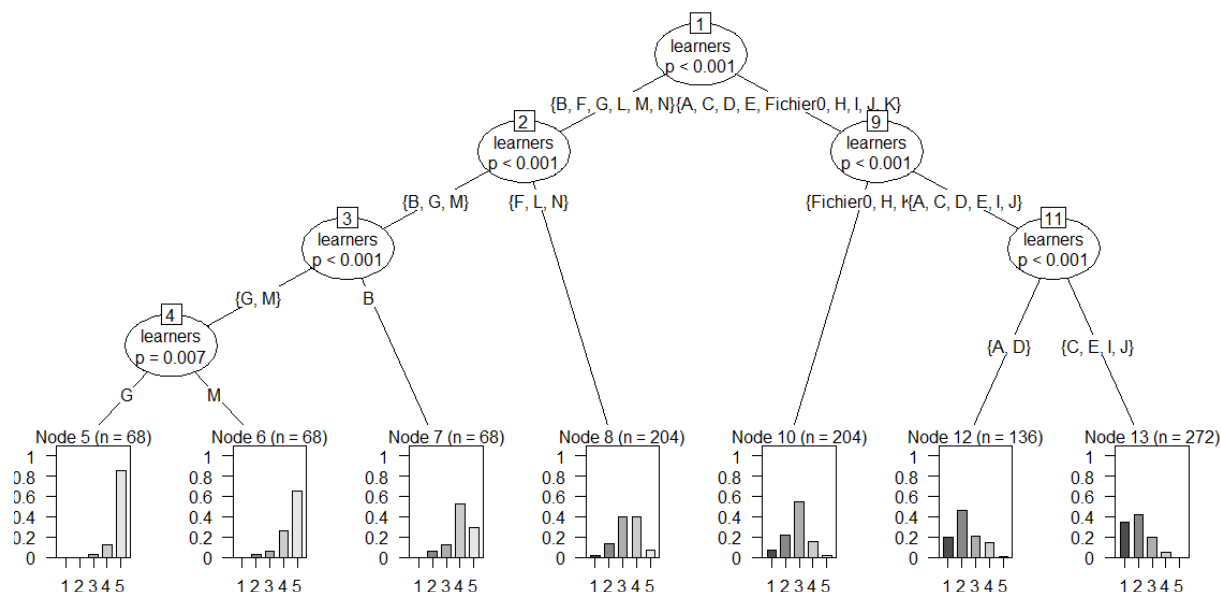


Figure 201 – Arbre d'inférence conditionnelle catégorisant et ordonnant les prestations des apprenants selon l'ensemble des 68 évaluateurs spécialistes par perception auditive

L'arbre présente une catégorisation en sept groupes et un classement des fichiers d'apprenants de gauche (meilleure qualité) à droite (qualité la plus faible). On remarque que ce classement place le Fichier 0 au centre de la hiérarchie *ex æquo* avec deux autres fichiers, ce qui conforte le choix de départ d'en faire le fichier repère pour sa valeur moyenne. Le classement d'ensemble est le suivant :

$$G > M \gg B \gg \gg F, L, N \gg \gg \gg \text{Fichier 0, H, K} \gg \gg A, D \gg C, E, I, J$$

Figure 202 – Classement perceptif catégoriel des fichiers d'apprenants par arbre d'inférence selon les 68 experts

Dans chaque même niveau, le classement est opéré non pas par résultat mais par ordre alphabétique. Le signe « > » permet de visualiser le degré de niveau séparant les catégories : « >>> » signifiera que les catégories sont très proches alors que « >>>>> » en montrera le plus grand éloignement.

Alors que les fichiers sont présentés par ordre alphabétique dans chaque catégorie, la moyenne des résultats offre davantage de précision sur la hiérarchie en jeu parmi les fichiers.

$$G > M > B > N > L > F > \text{Fichier 0} > H > K > A > D > I > C > J > E$$

Figure 203 – Classement perceptif des fichiers d'apprenants par moyenne des résultats attribués par les 68 experts

3.3.2 Classement par les évaluateurs de langue maternelle française (5 catégories)

Cette sous-section est destinée à vérifier une éventuelle modification du classement obtenu suivant la langue maternelle des experts évaluateurs, et tout d'abord des experts de langue maternelle française.

On remarque dans l'histogramme ci-dessous⁹⁶ que les évaluateurs francophones se montrent plus conciliants par rapport aux fichiers B, G et L, et un peu avec F. Par contre, ils sont plus exigeants avec C, E, H et dans une moindre mesure D.

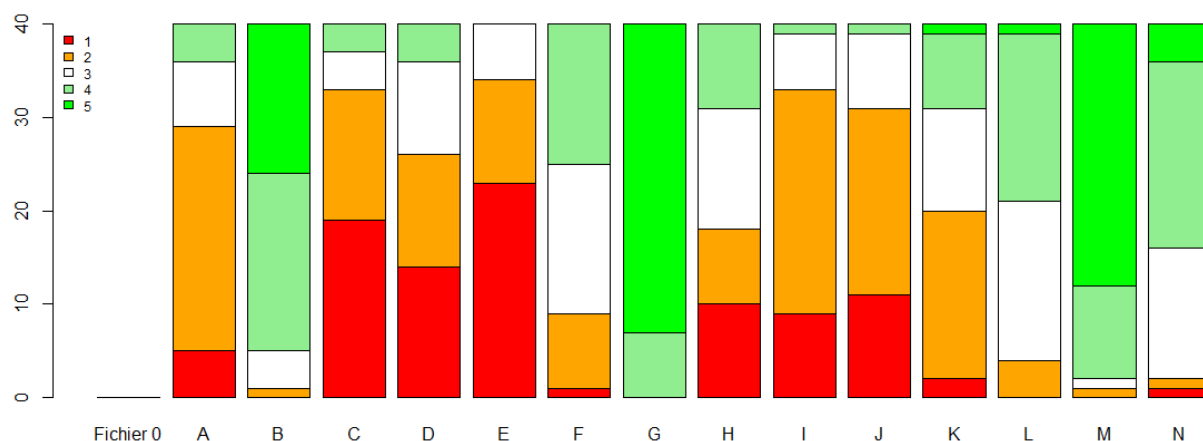


Figure 204 – Résultats de l'évaluation des quinze apprenants (abscisse) par perception des 40 enseignants francophones (ordonnées) selon le code : 1 = *much worse*, 2 = *worse*, 3 = *similar to File 0*, 4 = *better*, 5 = *much better*

Le plan factoriel de l'analyse des correspondances multiples qui suit renseigne sur les implications de l'histogramme précédent en montrant l'écart de distance moyen par rapport à la valeur moyenne des résultats. G et M sont perçus comme très proches qualitativement. Il en est de même des fichiers H et K. Par contre, les prestations des apprenants faibles dans le coin supérieur gauche du plan sont sensiblement perçus de la même manière par l'ensemble des évaluateurs, indépendamment de leur langue maternelle. Par contre, les fichiers F, L et N sont moins regroupés, c'est-à-dire qu'ils sont perçus comme davantage dissemblables par les francophones.

⁹⁶ Le script de cette rubrique sur les évaluateurs francophones a été réalisé par Maëlle Amand. J'ai transposé ce script pour l'adapter aux autres groupes d'évaluateurs.

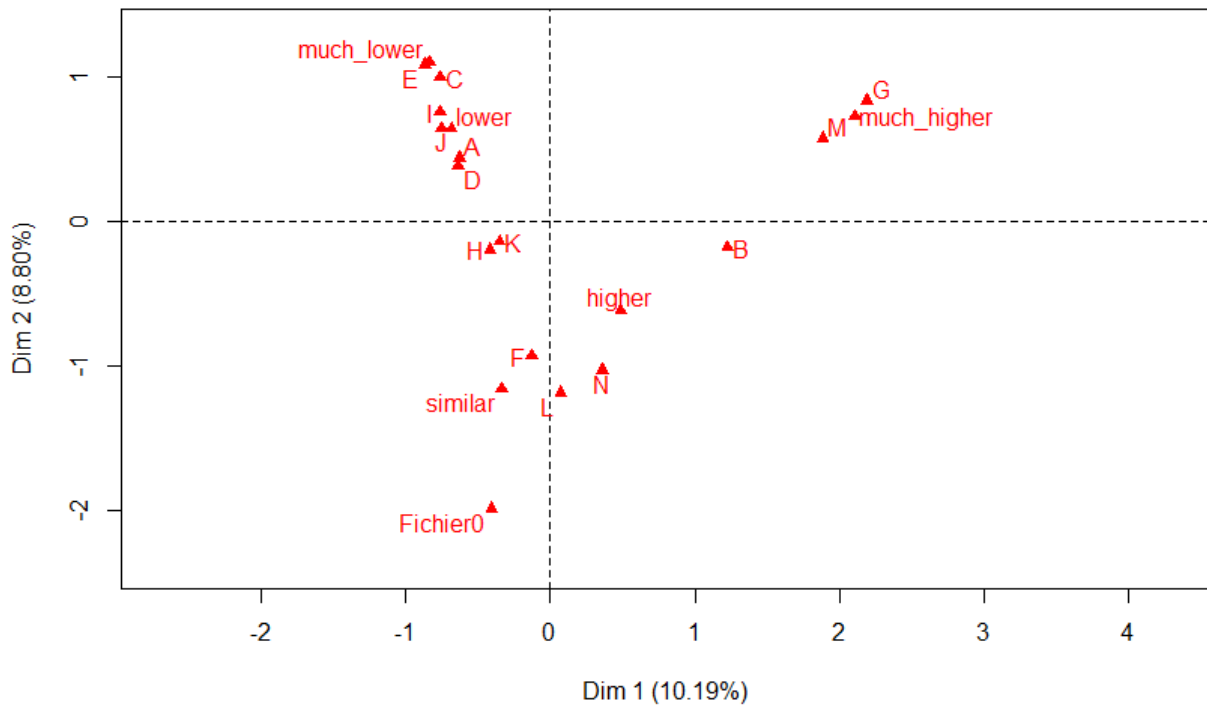


Figure 205 – Plan factoriel de l'analyse des correspondances multiples de l'évaluation perceptive des quinze apprenants par 40 experts francophones

On peut à présent voir dans l'arbre d'inférence conditionnelle des experts francophones les effets de la modification de perception. La catégorisation s'opère en sept groupes. Les fichiers G et M sont perçus de qualité analogue et F rejoint la catégorisation du Fichier 0 alors que H et K leur sont inférieurs. Les fichiers I et J, qui étaient catégorisés comme ceux de plus faible qualité prosodique, sont à présent dissociés de C et E pour rejoindre A et D.

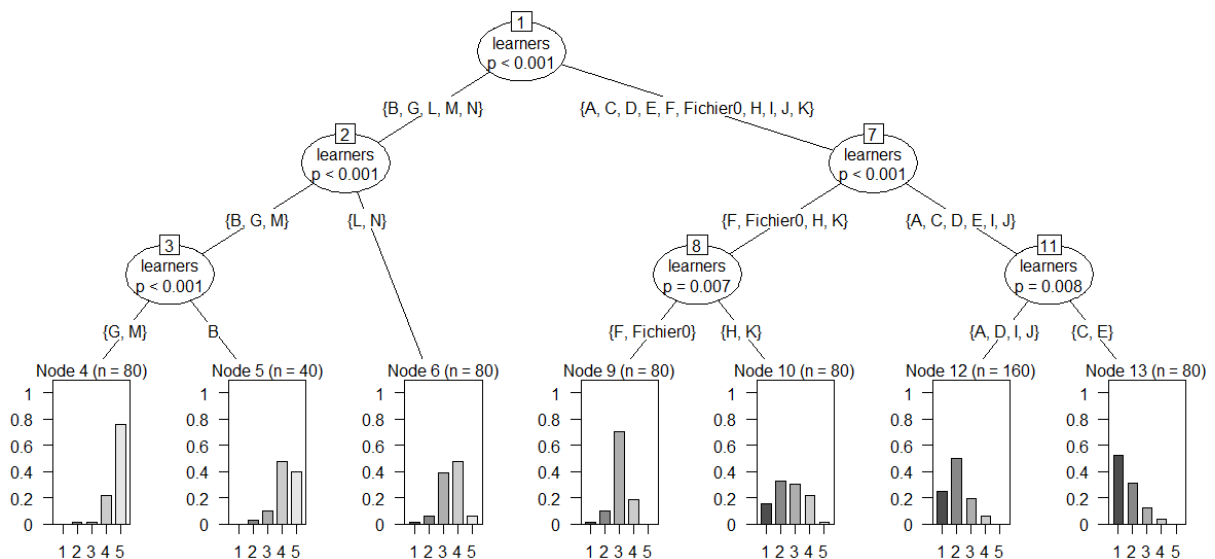


Figure 206 – Arbre d'inférence conditionnelle catégorisant et ordonnant les prestations des quinze apprenants par perception auditive selon les 40 enseignants francophones

La hiérarchisation qualitative découlant de ces graphiques figure ci-dessous.

G, M > B >> L, N >>> F, Fichier 0 > H, K >> A, D, I, J > C, E

Figure 207 – Classement perceptif catégoriel des fichiers d'apprenants par arbre d'inférence selon les 40 enseignants francophones

Le classement à partir de la moyenne des résultats attribués à chacun des quatorze fichiers sons est le suivant :

G > M > B > N > L > F > Fichier 0 > K > H > A > D > I = J > C > E

Figure 208 – Classement perceptif des fichiers d'apprenants par moyenne des résultats attribués par les 40 experts francophones

Les histogrammes qui leur font suite détaillent les votes évaluatifs de chaque expert francophone. On peut ainsi constater les similarités entre les résultats des fichiers G et M d'une part, et C et E d'autre part, ou encore I, J et A ; D, bien que proche de A, lui est plus différent visuellement. On comprend que le fichier B soit esseulé : il est dissemblable des autres.

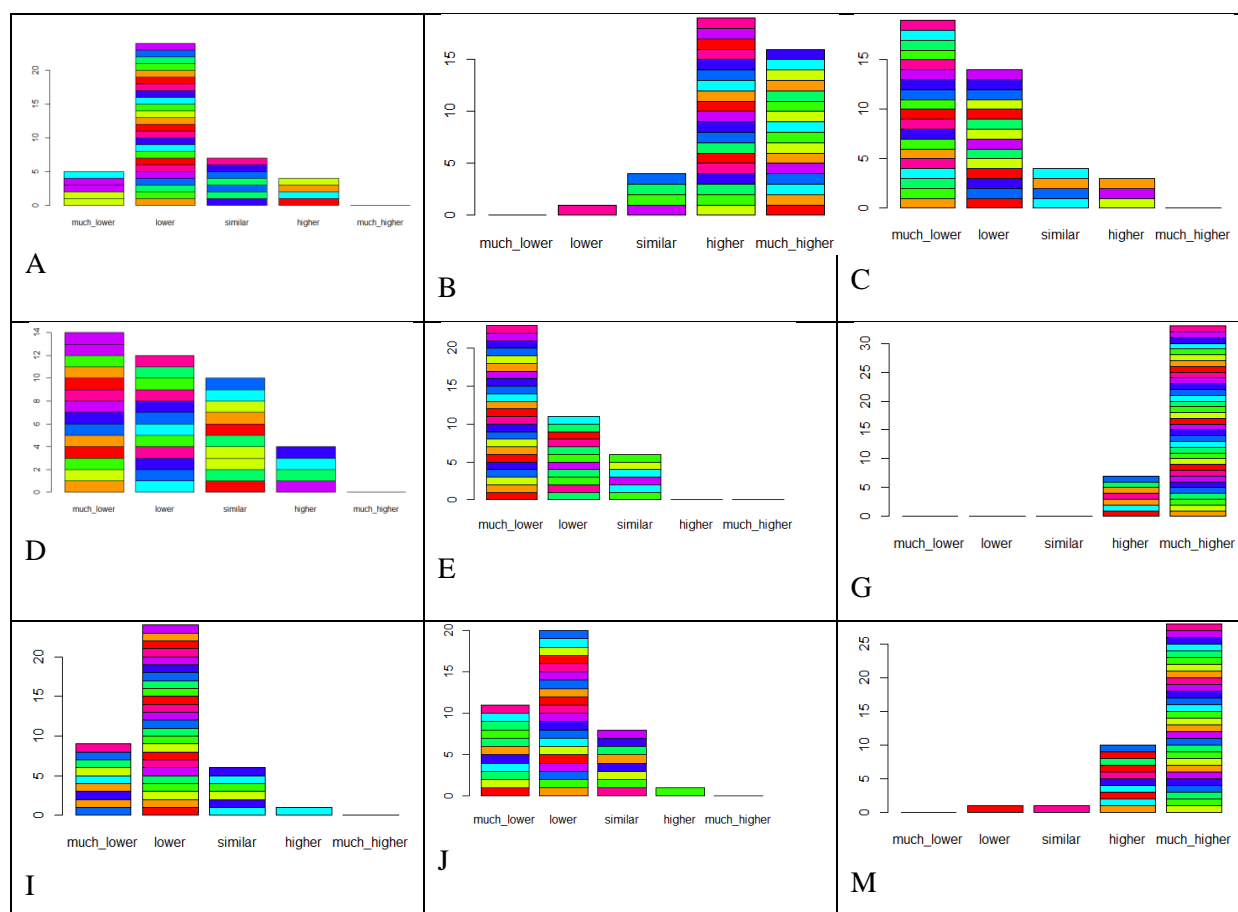


Figure 209 – Détail de l'évaluation perceptive des 14 apprenants par chacun des 40 experts francophones

À l'issue de cette section focalisée sur les résultats d'évaluation perceptive des 40 experts anglicistes francophones, on peut affirmer qu'il existe des différences d'appréciation entre ces résultats et ceux de la totalité des 68 experts, qui prennent forme dans les

regroupements de fichiers effectués. On doit néanmoins reconnaître que cette différence de catégorisation ne remet pas en question la hiérarchisation des fichiers les uns par rapport aux autres : aucun fichier n'est interverti.

3.3.3 Classement par les évaluateurs de langue maternelle anglaise (5 catégories)

Il est à présent nécessaire de vérifier une possible divergence des résultats des experts anglophones natifs par rapport à ceux des francophones. Pour cela, les opérations suivies sont renouvelées avec le contingent des 25 anglophones.

L'histogramme ci-dessous confirme les appréciations précédentes en montrant une similarité des proportions de couleurs vertes entre G, M et puis B. À l'autre extrême du paradigme, on pourrait amalgamer J, E et I par leur fort taux de couleurs chaudes, notamment de rouge, et leur faible taux de vert. Il est beaucoup moins aisé de catégoriser les apprenants moins extrêmes.

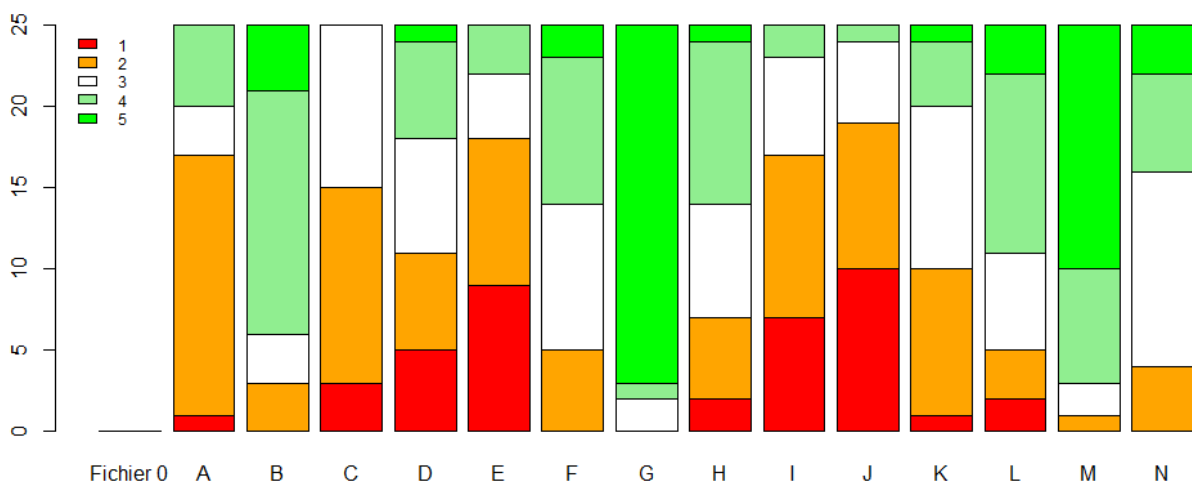


Figure 210 – Résultats de l'évaluation des 14 apprenants (abscisse) par perception des 25 enseignants francophones (ordonnées) selon le code : 1 = *much worse*, 2 = *worse*, 3 = *similar to File 0*, 4 = *better*, 5 = *much better*

À l'exception des extrêmes, on a une nouvelle fois constaté la difficulté de catégoriser hiérarchiquement la valeur prosodique des apprenants en se fiant uniquement aux figures précédentes. L'arbre d'inférence conditionnelle ci-dessous présente une classification claire en quatre catégories. On y voit que chacune des quatre catégories prend en compte l'importance d'une des variables par rapport aux autres. La variable *much lower* ne prenant jamais l'avantage par rapport aux autres, il ne figure que quatre catégories en tout.

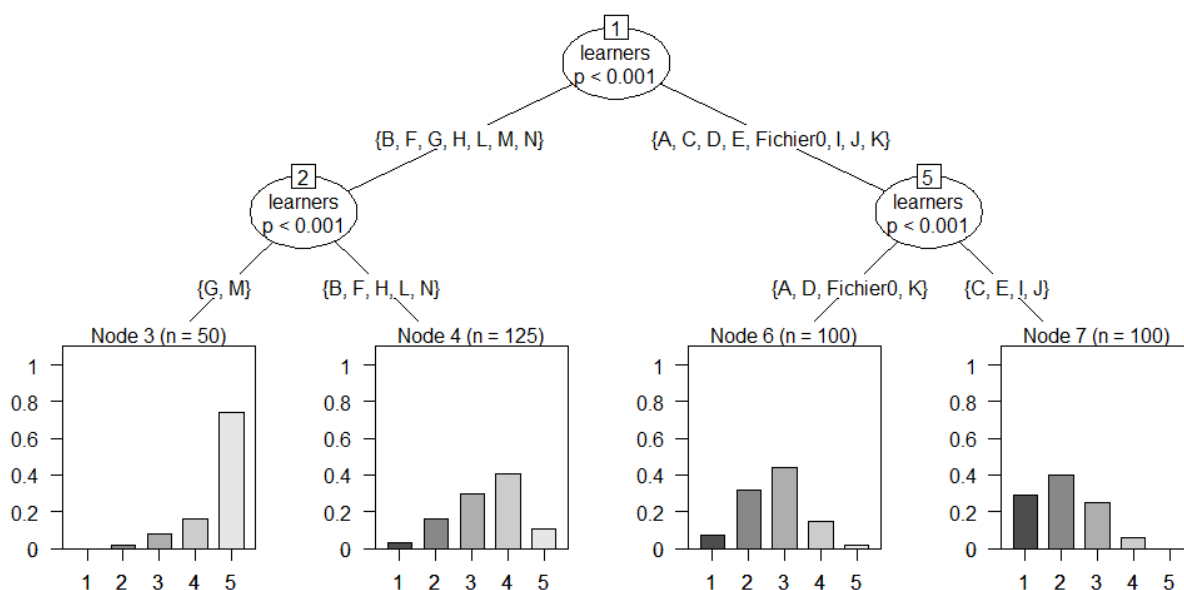


Figure 211 – Arbre d'inférence conditionnelle catégorisant et ordonnant les prestations des 15 apprenants par perception auditive selon les 25 évaluateurs spécialistes anglophones natifs

Le classement hiérarchique qualitatif obtenu ainsi par les experts anglophones est le suivant.

$$G, M > B, F, H, L, N \gg A, D, \text{Fichier } 0, K > C, E, I, J$$

Figure 212 – Classement perceptif catégoriel des fichiers d'apprenants par arbre d'inférence selon les 25 spécialistes anglophones natifs

Pour plus de précision sur la hiérarchie des fichiers à l'intérieur de chaque catégorie, voici celle obtenue grâce aux moyennes des résultats des anglophones natifs.

$$G > M > B > L > F = N > H > \text{Fichier } 0 > K > D > A > C > I > E > J$$

Figure 213 – Classement perceptif des fichiers d'apprenants par moyenne des résultats attribués par les 25 experts anglophones

Voyons comment les résultats de l'arbre d'inférence conditionnelle peuvent être expliqués par le nombre de votes donné par chacun des évaluateurs anglophones aux quatorze apprenants. M et G ont en commun leur fort nombre de votes dans la section *much higher*, comme précédemment. À l'autre extrême, E et J se rejoignent très fortement par l'accent mis sur les valeurs *lower* et *much lower*, et I et C marquent la prépondérance de la catégorie *lower*. B, F, H et L ont en commun la prépondérance de la variable *higher*, bien que F se caractérise par un même taux sur *higher* et *similar*. C'est ce point qui l'apparente à N car ce dernier regroupe le plus de votes sur *similar*, ce qui lui fait rejoindre cette catégorie. Les votes sur A, D et K sont davantage répartis sur les appréciations *lower* et moyennes. L'arbre montre que le fichier repère Fichier 0 est catégorisé dans ce groupe.

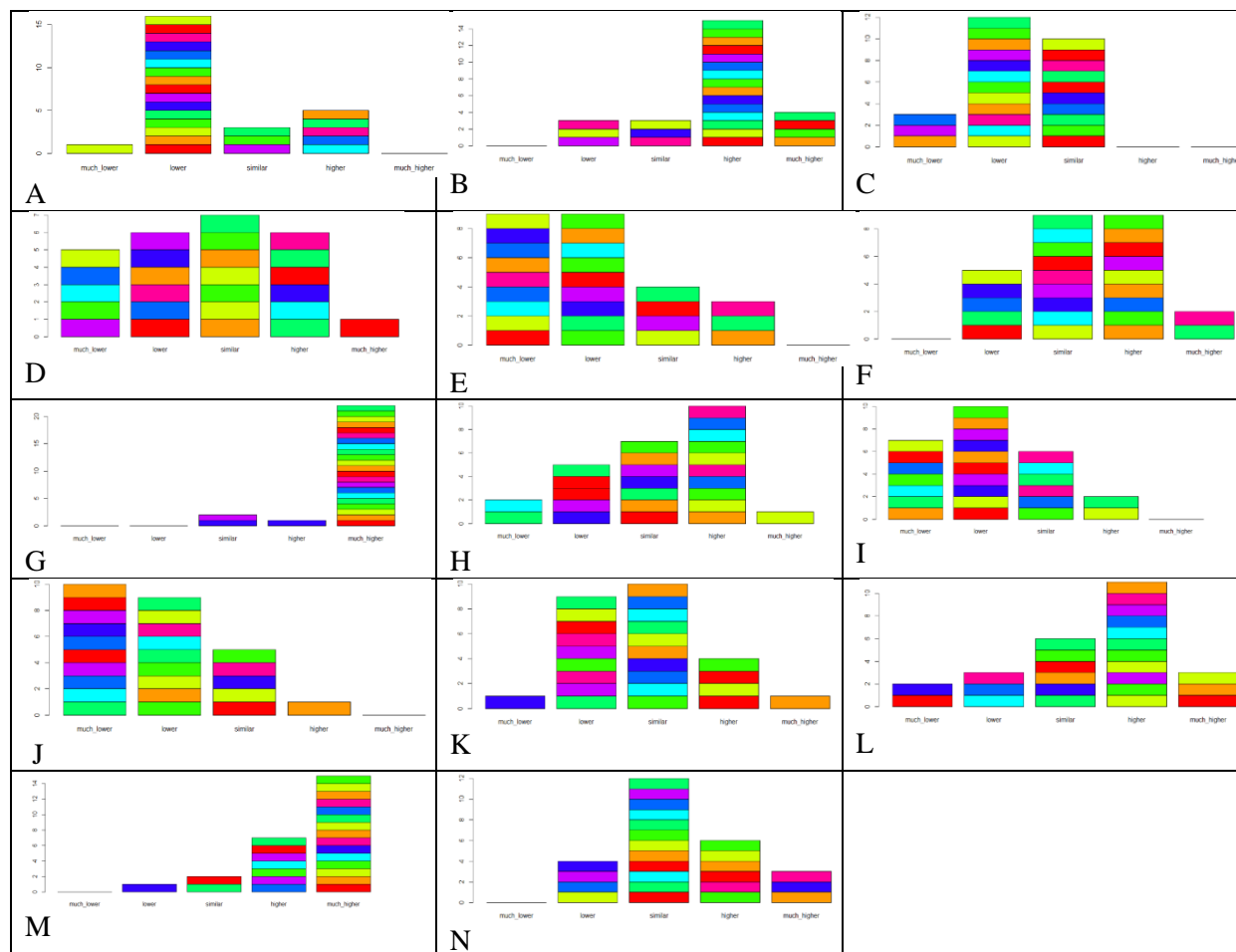


Figure 214 – Détail de l'évaluation perceptive des 14 apprenants par chacun des 25 experts anglophones

3.3.4 Classement par l'ensemble des évaluateurs ayant plus de 30 ans d'expérience et dont le test a duré au moins 20 minutes (5 catégories)

Jusqu'à présent, les experts ont été recrutés en fonction de leur statut et de leur bonne volonté dans leur participation. D'autres profilages sont possibles et parmi eux, il m'a semblé intéressant de cibler ceux qui devraient faire preuve de la plus grande expertise dans cette tâche évaluative. En cela, j'ai suivi la logique de profilage utilisée jusqu'à présent selon laquelle un faible nombre de données bien ciblées devrait être plus révélateur qu'un nombre qui lui serait largement supérieur, et qui équivaudrait au résultat d'un très grand nombre de données.

Pour cela, j'ai utilisé les deux critères que les participants devaient fournir dans leur fiche de réponse, c'est-à-dire la durée effective du test et leur degré d'expérience dans le domaine de l'enseignement en nombre d'années. En m'appuyant sur les réponses obtenues, j'ai sélectionné les experts profilés selon ces deux critères :

- Participants reconnaissant avoir consacré au moins 20 minutes à effectuer le test (durée requise pour effectuer le test dans des conditions correctes),
- Participants ayant plus de 30 ans d'expérience d'enseignement (nombre d'années correspondant à un degré d'expertise maximal dans le domaine).

Ainsi, la liste des experts a été fortement réduite à un total de 18 évaluateurs, dont onze francophones natifs et sept anglophones natifs. L'histogramme suivant en représente les résultats.

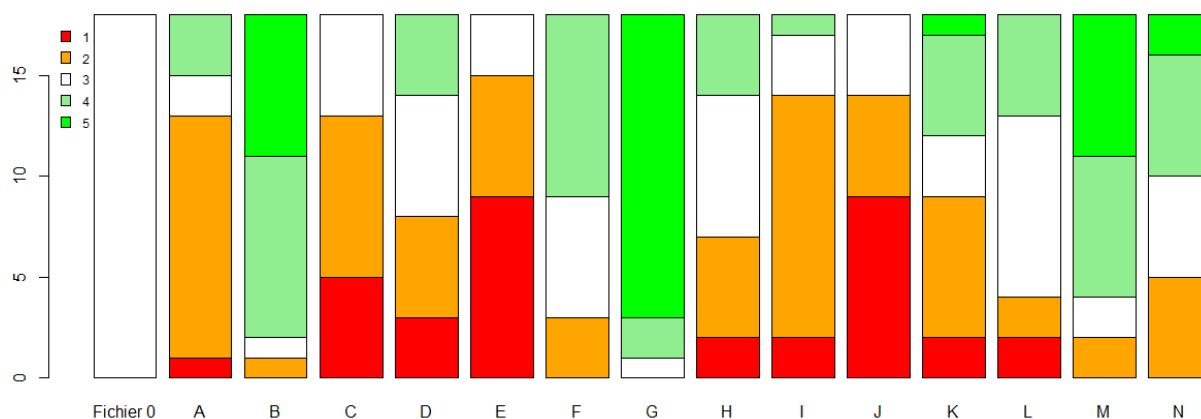


Figure 215 – Résultats de l'évaluation des 15 apprenants (abscisse) par perception des 18 experts profilés (ordonnées) selon le code : 1 = *much worse*, 2 = *worse*, 3 = *similar to File 0*, 4 = *better*, 5 = *much better*

On y remarque la singularité de G, dont la proportion de vert vif (légende 5) se distingue très clairement de B et M. À l'autre bout du paradigme, en tenant compte de la proportion de rouge vif et l'absence ou faible présence de vert pâle, on constate que E et J sont analogues, suivis de C, puis de I, puis de A.

Cette interprétation visuelle de l'histogramme est confirmée par l'arbre d'inférence conditionnelle ci-dessous. Il établit cinq catégories de niveaux et place G à part, devant B et M. À l'autre extrême de l'échelle, C, E, I et J sont regroupés, alors que A est déjà considéré comme plus moyen et rejoint les fichiers D, H, K, L, ainsi que le Fichier repère, dans une catégorie plus moyenne.

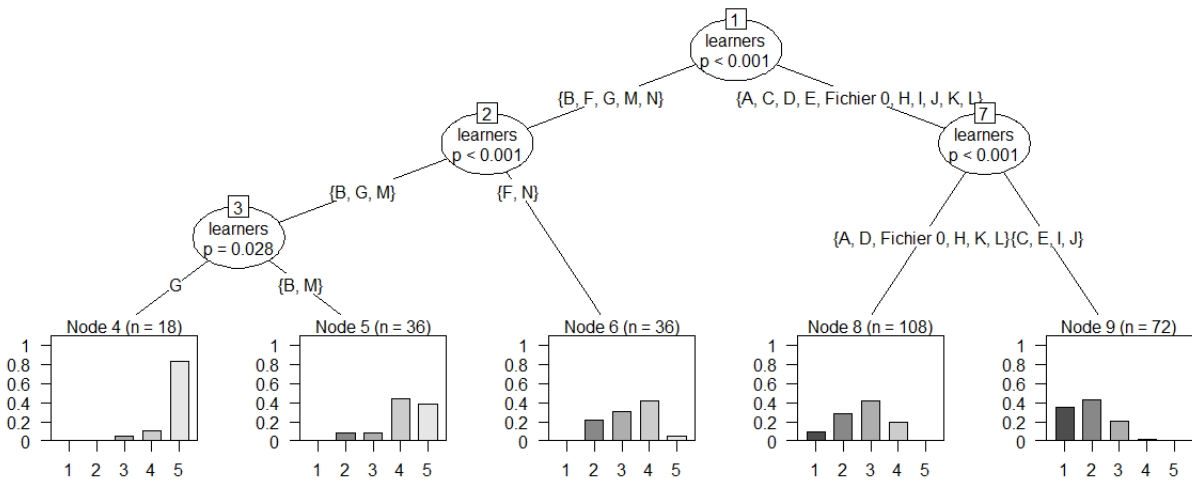


Figure 216 – Arbre d’inférence conditionnelle catégorisant et ordonnant les prestations des 15 apprenants par perception auditive selon les 18 experts profilés

Le classement qualitatif des fichiers sons selon cet arbre d’inférence conditionnelle est détaillé ci-dessous.

$$G > B, M \gg F, N \gg \gg A, D, \text{Fichier } 0, H, K, L \gg C, E, I, J$$

Figure 217 – Classement perceptif catégoriel des fichiers d’apprenants par arbre d’inférence selon les 18 experts profilés

Cet agencement considère le fichier repère « Fichier 0 » relativement faible tout en lui conservant sa place moyenne parmi les autres candidats. Ci-dessous est précisée la hiérarchie numérique à l’intérieur de chaque catégorie.

$$G > B > M > F > N > \text{Fichier } 0 > L > K > H > D > A > I > C > J > E$$

Figure 218 – Classement perceptif des fichiers d’apprenants par moyenne des résultats attribués par les 18 experts profilés

3.3.5 Classement des groupes d’évaluateurs à l’effectif maximal et minimal selon trois catégories évaluatives

Jusqu’à présent ont été considérées les cinq catégories évaluatives présentées lors de l’administration du test PER-LEC-NN : *much higher*, *higher*, *similar to File 0*, *lower* et *much lower*. La finesse de cette classification désavantage ses résultats, que l’on ne peut comparer à ceux obtenus par ailleurs dans d’autres recherches évaluatives. Afin de pouvoir les homologuer, il est intéressant de les profiler en trois catégories. C’est l’objet de cette section.

Étant donné le profilage déjà effectué en cinq catégories, j’ai opté pour la manière la plus simple de réduire les résultats en trois catégories en conservant la catégorie centrale et en amalgamant les catégories périphériques. J’applique alors la fonction =SI(G8=3;3;SI(G8>3;4;2)) dans le tableau *Excel* des résultats du test PER-LEC-NN, G8 étant un exemple de cellule du tableau). Et j’obtiens les catégories suivantes :

- Le résultat *Similar to File 0* reste identique à lui-même,
- Les résultats qui lui sont supérieurs (*higher* et *much higher*) sont amalgamés,
- Les résultats qui sont inférieurs à *Similar to File 0* (*lower* et *much lower*) sont amalgamés.

Le biais que présente cette répartition est que la valeur *Similar to File 0* risque à présent d'être davantage sous-représentée dans cette nouvelle version que s'il n'y avait eu que trois choix disponibles dès le départ. Néanmoins, son grand avantage est la facilité avec laquelle elle a pu être créée. Procédons à l'examen des profils en trois catégories de la totalité des 68 évaluateurs puis du groupe le plus restreint (18 évaluateurs profilés).

3.3.5.1 Classement selon trois catégories de niveaux par l'ensemble de 68 évaluateurs

Bien qu'il soit réduit à trois niveaux de maîtrise représentés par trois couleurs, l'histogramme ci-dessous semble représenter une fidèle image de celui avec cinq niveaux. On observera cependant la faible place couverte par les zones blanches représentant une évaluation moyenne, résultat du biais souligné en amont.

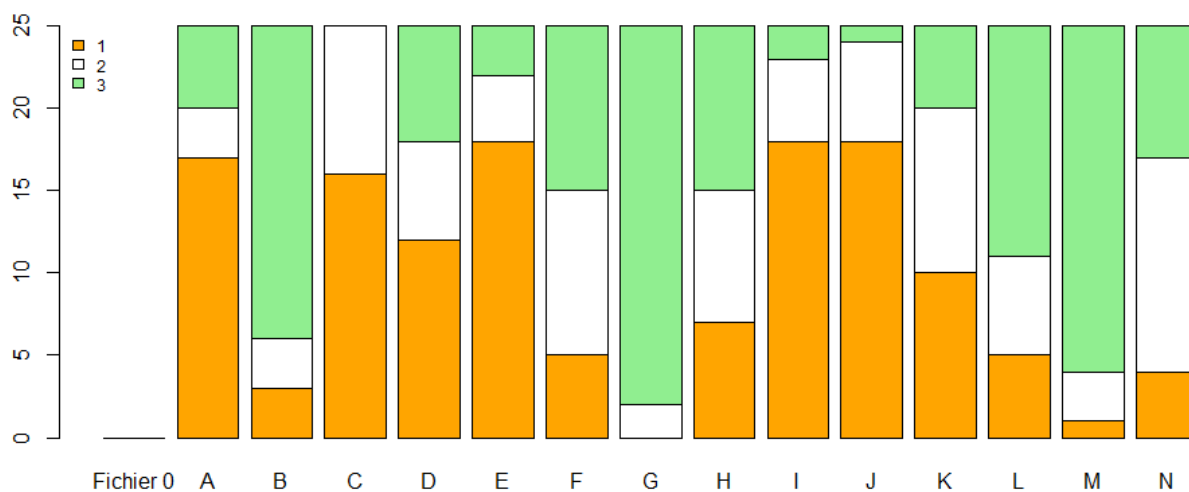


Figure 219 – Résultats de l'évaluation des 15 apprenants (abscisse) par perception des 68 experts (ordonnées) selon le code : 1 = worse, 2 = similar to File 0, 3 = better

On obtient aussi l'arbre d'inférence conditionnelle et les hiérarchies suivantes pour le profilage des 68 évaluateurs.

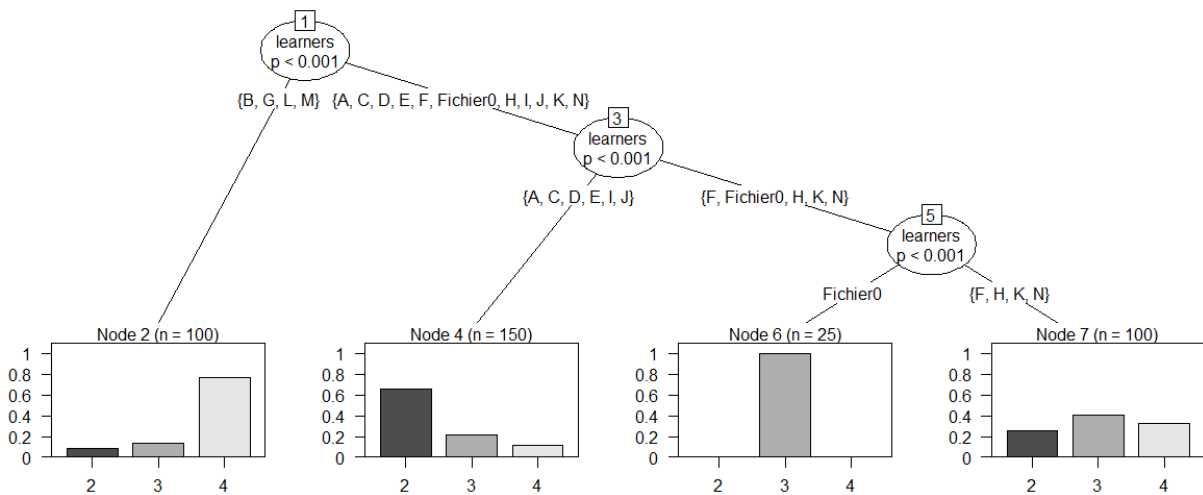


Figure 220 – Arbre d’inférence conditionnelle de l’évaluation perceptive en trois catégories des 15 apprenants par les 68 experts

Cette catégorisation en quatre blocs ne reflète pas la hiérarchisation des niveaux mais davantage des groupes d’apprenants dans un objectif de remédiation. En effet, elle se subdivise en deux branches : les apprenants ayant déjà acquis un niveau assez élevé (nœud 2) et ceux dont le niveau est à parfaire (nœud 3). Celui-ci se subdivise à son tour en deux groupes : le nœud 4 des plus faibles d’une part, et le nœud 5 de ceux qui sont de niveau moyen d’autre part. Ce dernier nœud se ramifie une fois de plus en fonction de l’hétérogénéité des lacunes (nœud 7) et l’homogénéité du niveau (nœud 6 pour le Fichier 0).

On semble obtenir la hiérarchie de classement en trois groupes fondamentaux suivante : niveau élevé (G, M, B, L), niveau moyen (N, F, Fichier 0, H, K) et niveau faible (D, A, C, I, J, E). De par l’homogénéité de son unique résultat, puisqu’il est le fichier repère, le Fichier 0 est perçu comme une classe à part bien qu’il corresponde à un niveau moyen.

$$G > M > B > L > N > F > \text{Fichier 0} > H > K > D > A > C > I > J > E$$

Figure 221 – Classement perceptif des fichiers d’apprenants en trois catégories par arbre d’inférence selon les 68 experts profilés

Malgré son intérêt didactique, il ne m’apparaît pas souhaitable d’utiliser cet arbre d’inférence pour rendre compte d’une évaluation hiérarchique des fichiers. Pour obtenir un classement reflétant le niveau attribué, basons-nous uniquement sur les moyennes des résultats par fichier.

$$G > M > B > N > L > F > \text{Fichier 0} > H > K > D > A > C > I > J > E$$

Figure 222 – Classement perceptif des fichiers d’apprenants par moyenne des résultats attribués par l’ensemble des 68 experts

On constate que la comparaison des deux hiérarchie met en exergue le fait que les fichiers N et L soient intervertis. Les votes qu’ils reçoivent respectivement figurent ci-dessous.

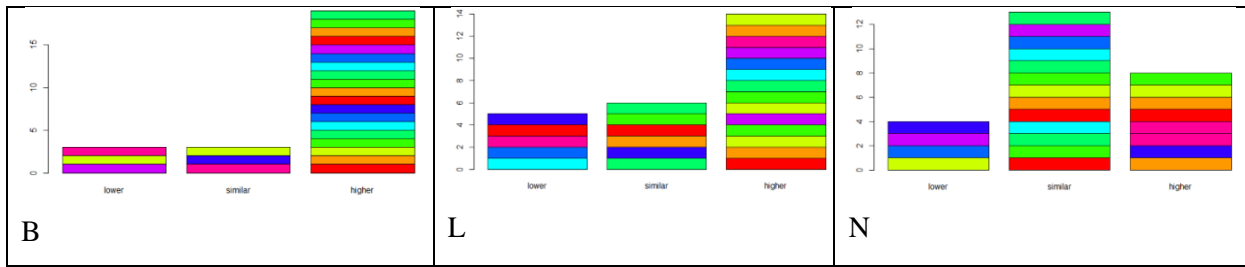


Figure 223 – Détail de l'évaluation perceptive des apprenants en trois catégories par les 68 experts

Le fichier B est sans conteste le meilleur des trois avec 19 bons résultats, trois moyens et trois faibles. L est d'un profil similaire bien qu'atténué, avec quatorze bons résultats, six moyens et cinq faibles. N semble être le plus faible des trois en ne totalisant que huit bons résultats, treize moyens et quatre faibles.

3.3.5.2 Classement selon trois catégories de niveaux par le groupe des 18 évaluateurs profilés

Ci-dessous figurent les résultats des quatorze fichiers sons d'apprenants évalués par les 18 experts profilés selon la durée du test et celle de leur activité d'enseignement.

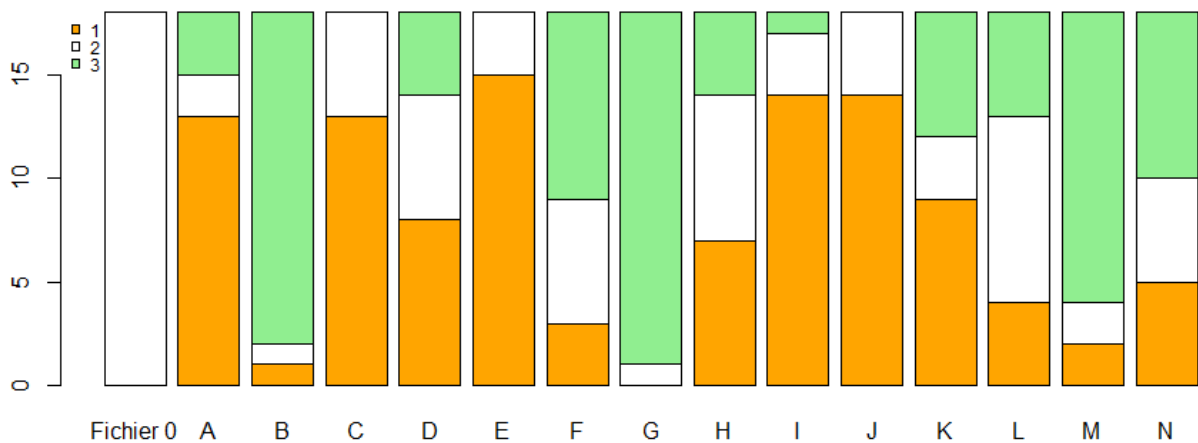


Figure 224 – Résultats de l'évaluation des 15 apprenants (abscisse) par perception des 18 experts profilés (ordonnées) selon le code : 1 = worse, 2 = similar to File 0, 3 = better

L'impression donnée par cet histogramme par rapport au précédent (évaluation des 68 experts en trois catégories) est une exacerbation des résultats, comme si les évaluateurs s'étaient montrés plus sévères avec les apprenants de moindre niveau et davantage bienveillants avec ceux de niveau élevé. Ce phénomène est clairement visible chez les apprenants qu'ils jugent les meilleurs B et G, et les apprenants C, D, E, I, J qu'ils placent en bas de la hiérarchie qualitative.

Ici encore, la catégorisation par arbre d'inférence aboutit à quatre groupes aux effectifs homogènes.

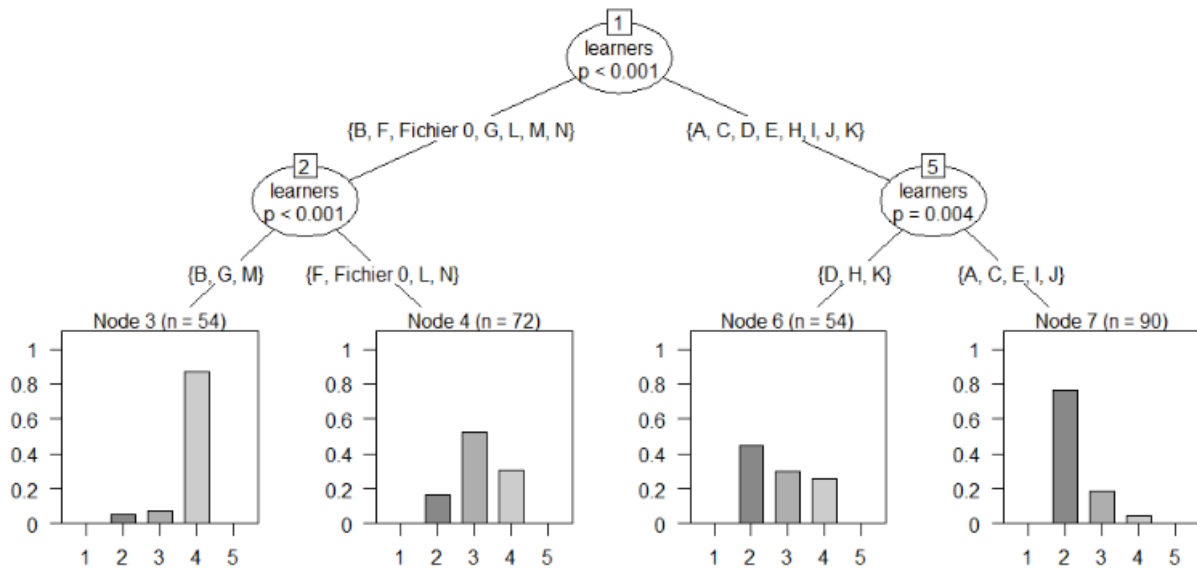


Figure 225 – Arbres d'inférence conditionnelle de l'évaluation perceptive en trois catégories des 15 apprenants par les 18 experts profilés

Ainsi que l'indiquent les histogrammes insérés dans les cases catégorielles, le nœud 3 comprend les meilleurs fichiers (B, G et M) repérés par une graduation 4 très élevée, le nœud 4 les fichiers de niveau moyen (graduation 3 très élevée). Les nœuds suivants regroupent les fichiers de niveau faible, le nœud 7 se spécialisant dans les fichiers particulièrement faibles (graduation 2 très élevée). Ici le fichier repère est classé dans les fichiers de niveau moyen. Le classement catégoriel fourni est le suivant. L'agencement hiérarchique des fichiers dérive du classement par moyenne des résultats ci-dessous.

$$G > B > M > F > N > L > \text{Fichier 0} > H = K > D > A > C = I > J > E$$

Figure 226 – Classement perceptif des fichiers d'apprenants par arbre d'inférence en trois catégories selon les 18 experts profilés

$$G > B > M > F > N > L > \text{Fichier 0} > H = K > D > A > C = I > J > E$$

Figure 227 – Classement perceptif des fichiers d'apprenants par moyenne des résultats attribués par les 18 experts profilés

3.4 Synthèse des classements obtenus par profilage perceptif

Voici récapitulés ci-dessous les classements évaluatifs catégoriels obtenus précédemment par les quatre groupes d'experts pris en compte et suivant le nombre de catégories évaluatives considérées. Les moyennes des résultats ont permis de réagencer chaque fichier selon sa place réelle et non selon son ordre alphabétique à l'intérieur de la catégorie définie par l'arbre d'inférence conditionnelle.

<i>Groupes d'experts avec nombre de réponses dans le questionnaire</i>	<i>Nombre de catégories dans l'arbre</i>	<i>Classements catégoriels obtenus par arbres d'inférence conditionnelle avec hiérarchie rétablie à l'intérieur des catégories grâce aux moyennes des résultats sur tableur</i>
Total des 68 (5 catégories)	7	G > M >> B >>> N, L, F >>>> Fichier 0, H, K >>> A, D >> I, C, J, E
40 natifs francophones (5 catégories)	7	G, M > B >> N, L >>> F, Fichier 0 > K, H >> A, D, I, J > C, E
25 natifs anglophones (5 catégories)	4	G, M > B, L, F, N, H >> Fichier 0, K, D, A > C, I, E, J
18 experts profilés (5 catégories)	5	G > B, M >> F, N >>> Fichier 0, L, K, H, D, A >> I, C, J, E
Total des 68 (3 catégories)	4	G > M > B > N > L > F > Fichier 0 > H > K > D > A > C > I > J > E ⁹⁷
18 experts profilés (3 catégories)	4 ok	G, B, M > F, N, L, Fichier 0 >> H = K, D > A, C = I, J, E

Tableau 108 – Récapitulatif du classement évaluatif catégoriel des 15 apprenants par les quatre groupes d'experts étudiés avec regroupement coloré des zones de stabilité

Ces divers classements, obtenus suivant les profilages d'experts considérés en fonction du nombre maximal d'évaluateurs, de leur langue maternelle ou du niveau d'expertise maximale, offrent une optique évaluative intéressante tant en ce qui concerne les points communs que les aspects divergents. Une étude comparative visuelle regroupe les zones de stabilité en les colorant en vert.

Parmi les similarités, on constate que le Fichier 0, repère de l'expérimentation PER-LEC-NN, occupe toujours une place centrale sur toutes les échelles évaluatives : le nombre de candidats à sa droite et à sa gauche sont sensiblement les mêmes (de 5 à 7 à sa gauche). Les extrêmes des échelles sont aussi occupés par les mêmes candidats dont le meilleur est unanimement G, alors que M et B le suivent. Ensuite viennent N, L et F en ordre dispersé, sauf dans le classement des 18 experts profilés, où L se trouve après le Fichier 0. Ainsi, dans tous les classements, sauf celui des 18 experts profilés, on obtiendra un bloc regroupant les fichiers 0, H et K en ordre variable. S'ensuivent A et D, puis en dernier I, C, J et E. Les classements réagencés hiérarchiquement à l'intérieur des catégories selon leur valeur propre et non plus leur ordre alphabétique offrent un agencement hiérarchique particulièrement stable.

⁹⁷ La hiérarchie utilisée ici est celle de la moyenne des résultats PER-LEC-NN fournis sur tableur étant donné que la catégorisation par arbre d'inférence conditionnelle ne privilégie pas le niveau qualitatif des réalisations prosodiques.

Malgré ces similarités fortes, on perçoit deux divergences de perception selon les groupes d'experts. La première concerne le nombre de catégories dégagées pour classer les réalisations des apprenants : alors que les francophones classent les réalisations en sept groupes, tout comme l'ensemble des 68 spécialistes dont ils sont majoritaires, les anglophones se satisfont de quatre groupes. Les experts profilés selon le respect des consignes et le degré d'expertise dans l'appréciation d'un niveau prosodique considèrent, quant à eux, cinq catégories. Sans entrer dans des considérations hors sujet et peu utiles ici sur le bien-fondé des diverses approches perceptives, on constate néanmoins l'affinité de la répartition en cinq catégories avec celle considérée dans l'approche ergonomique des variables qualitatives pour sa clarté et sa facilité d'utilisation. Par ailleurs, la perte d'information occasionnée par la réduction des catégories à trois au lieu de cinq se répercute sur le nombre de groupes de classement qui chute à quatre quel que soit le nombre d'experts du groupe.

La deuxième différence concerne le classement attribué aux apprenants les uns par rapport aux autres. En tête du classement, le fichier M est parfois associé au meilleur candidat G (groupes 68 et 18 experts) et parfois au bon candidat B (natifs 40 francophones et 25 natifs anglophones). On a aussi vu que les jugements sur L, bien que très stables dans l'ensemble, pouvaient varier selon le groupe d'évaluateurs. Il serait sans doute éclairant de vérifier ce qui justifie les alternances de classement de ces fichiers.

Malgré ces quelques divergences mineures, l'expérimentation STAT-NN4 a montré qu'une fois de plus la perception fournissait une base évaluative stable et robuste. Essayons à présent de vérifier la validité du profilage par variables.

4 Étude des résultats obtenus selon le profilage par variables (expérimentation STAT-NN5)

On a vu ci-dessus que le profilage par perception dépendait d'une catégorisation mise en place par des évaluateurs. Le profilage semi-automatique des réalisations prosodiques des apprenants, objet de cette thèse, se décline lui aussi en domaines subdivisés en variables, ainsi qu'on l'a vu au chapitre précédent. La validation d'un profilage doit se vérifier à l'aune de la pertinence des domaines et variables pris en compte, mais aussi de la finesse de la catégorisation adoptée. C'est pourquoi cette section est destinée à fournir des données précises qui pourront ensuite être confrontées aux résultats perceptifs obtenus précédemment. Pour cela, il est nécessaire d'obtenir ici aussi un classement hiérarchique de qualité

prosodique comme base de comparaison. Le Chapitre IV a présenté une analyse des données par taux de réussite ou d'erreurs par variable pour les six réalisations d'apprenants servant de base aux profils non natifs. Il s'agit ici de la compléter en étudiant l'ensemble des 15 fichiers de lecture d'apprenants afin de pouvoir comparer les résultats obtenus avec ceux précédemment dégagés de l'expérimentation STAT-NN4. Pour cela, procédons à une hiérarchisation quantitative puis qualitative des 15 fichiers sons.

On se rappelle que l'on avait ôté trois variables des 28 créées au départ étant donné qu'elles reprenaient des moyennes de variables déjà prises en compte. C'est à partir de ces 25 variables que l'on aurait dû calibrer les quatre grands domaines prosodiques, puis affiner le profilage par la prise en compte des variables. Pourtant l'expérimentation PER-LEC-NN, pour des raisons de faisabilité, avait requis le raccourcissement des fichiers à évaluer, réduisant les variables effectivement testées à 18. Pour ces raisons, seuls les résultats à 18 variables peuvent être raisonnablement comparés à ceux obtenus grâce à l'expérimentation PER-LEC-NN. C'est pourquoi il est nécessaire de les établir ainsi que le classement qui en résulte. Les données et classements à 25 variables gardent néanmoins tout leur intérêt car ils permettront d'éclairer les jugements comparatifs par extrapolation des données. Ainsi sera préservée une meilleure possibilité de cerner les phénomènes en jeu.

Dans la hiérarchisation quantitative par appartenance des fichiers à l'une des cinq ou des trois catégories de classement, il est intéressant de constater que les trois meilleurs apprenants restent dans les mêmes catégories (un dans la catégorie 5, et deux dans la catégorie 4) ; le plus faible reste au bas de l'échelle dans la catégorie 1. Si certains fichiers restent dans la même catégorie quel que soit le nombre de variables, ce n'est pas le cas de tous. Aussi, il est utile d'établir une comparaison basée sur le taux de correspondance des catégories de niveau entre les deux approches (classique et semi-automatique) visant spécifiquement les variables figurant dans l'extrait de texte utilisé pour l'évaluation de l'expérimentation PER-LEC-NN. Ces données catégorielles pourront être confrontées à celles obtenues par perception et éventuellement éclairées par elles.

Voici les classements obtenus par bilan global des quatre grands domaines prosodiques après avoir calculé le taux de réussite et d'erreurs (points de bonus desquels sont soustraits les points de malus) dans chacun d'entre eux. Pour faciliter la compréhension, le tableau suivant mentionne les résultats en double codage : les codes utilisés pour l'expérimentation PER-LEC-NN en lettres et ceux des profilages du Chapitre IV en S suivi

d'un nombre. Le Fichier 0 est noté « 0 ». Il s'agit en fait du fichier S5. La correspondance des noms de fichiers figure en Annexe pour pouvoir établir des liens entre toutes les parties de la thèse.

Nombre de variables	Bilan	Classements selon système de bonus/malus Selon les deux dénomination de fichiers (PER-LEC-NN et expérimentations du Chapitre IV)
18 variables	Général	B > M > G > D > 0 > H > J > L > C > N > E > I > F > A > K S10>S1>S6>S12>S5>S2>S9>S4>S3>S8>S14>S15>S13>S7>S11
	Durée	C = H = M > B = L > N > 0 = G > J > A = D = F = K > E > I S3=S2=S1>S10=S4>S8>S5=S6>S9>S7=S12=S13=S11>S14>S15
	Registre	B > 0 = D = I = M > G = J > L > F = H = N > C = E > A = K S10>S5=S12=S15=S1>S6=S9>S4>S13=S2=S8>S3=S14>S7=S11
	Mélodie	B = G > D = M > 0 = E = J > H = I > A = C = F = K = L = N S10=S6>S12=S1>S5=S14=S9>S2=S15>S7=S3=S13=S11=S4=S8
	Rythme	B > G = H > 0 > D = E = N > A = L = M > C = I = J > F > K S10>S6=S2>S5>S12=S14=S8>S7=S4=S1>S3=S15=S9>S13>S11
25 variables	Général	B > M > G > H > L > D > 0 > E > J > N > C > F > A > I > K S10>S1>S6>S2>S4>S12>S5>S14>S9>S8>S3>S13>S7>S15>S11
	Durée	C = H = M > B = L > N > 0 = G > J > A = D = F = K > E > I S3=S2=S1>S10=S4>S8>S5=S6>S9>S7=S12=S13=S11>S14>S15
	Registre	B > G = M > 0 = D > J = L > I > H > E = F = K = N > A = C S10>S6=S1>S5=S12>S9=S4>S15>S2>S14=S13=S11=S8>S7=S3
	Mélodie	G > B > E > M > D > L > 0 = J = N > H > F > A = I = K > C S6>S10>S14>S1>S12>S4>S5=S9=S8>S2>S13>S7=S15=S11<S3
	Rythme	B > E > G = H = M = N > 0 > A = D = J > C = I = L > F > K S10>S14>S6=S2=S1=S8>S5>S7=S12=S9>S3=S15=S4>S13>S11

Tableau 109 – Hiérarchisation évaluative des 15 apprenants selon le système de bonus-malus global et par domaine prosodique selon 18 et 25 variables et les deux systèmes de codage des fichiers utilisés

Lorsque l'on compare les bilans généraux, on remarque que le nombre de variables n'influe pas sur les apprenants en début et en fin de classement : les fichiers B, M et G gardent leur places respectives en début de classement et le fichier K reste dernier. Par contre, les fichiers intermédiaires pourront être décalés de plusieurs places.

Les deux histogrammes suivants détaillent ces résultats sous forme visuelle et l'on peut clairement voir les altérations occasionnées par la non prise en compte de variables importantes. Dans chaque histogramme, les apprenants sont hiérarchisés par classement du bilan général (moyenne des bilans par domaine).

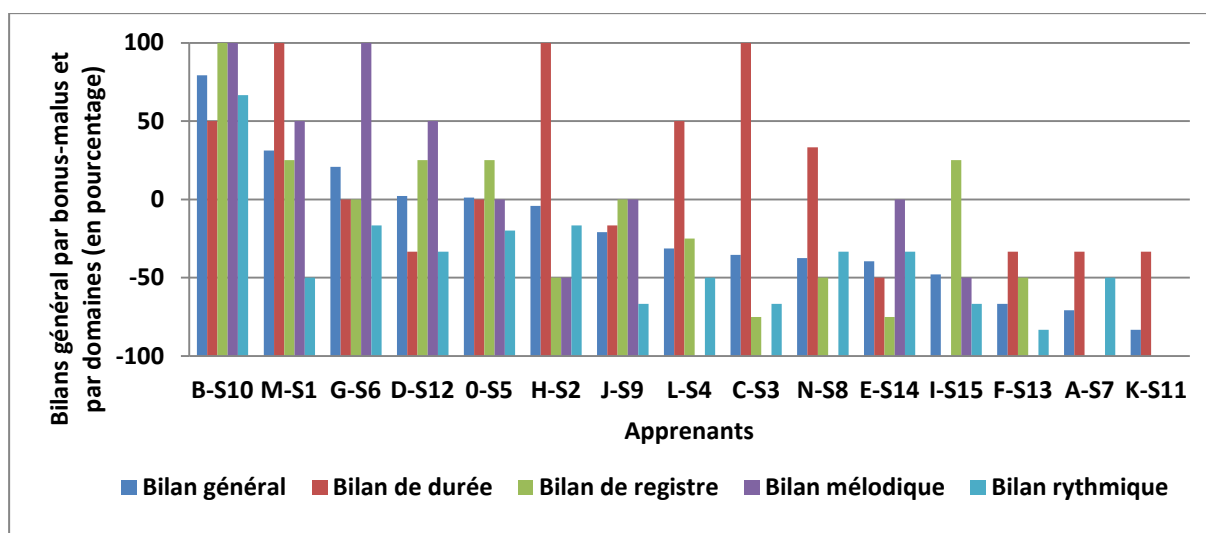


Figure 228 – Bilan général et par domaine prosodique des 15 apprenants selon le système de bonus/malus sur 18 variables

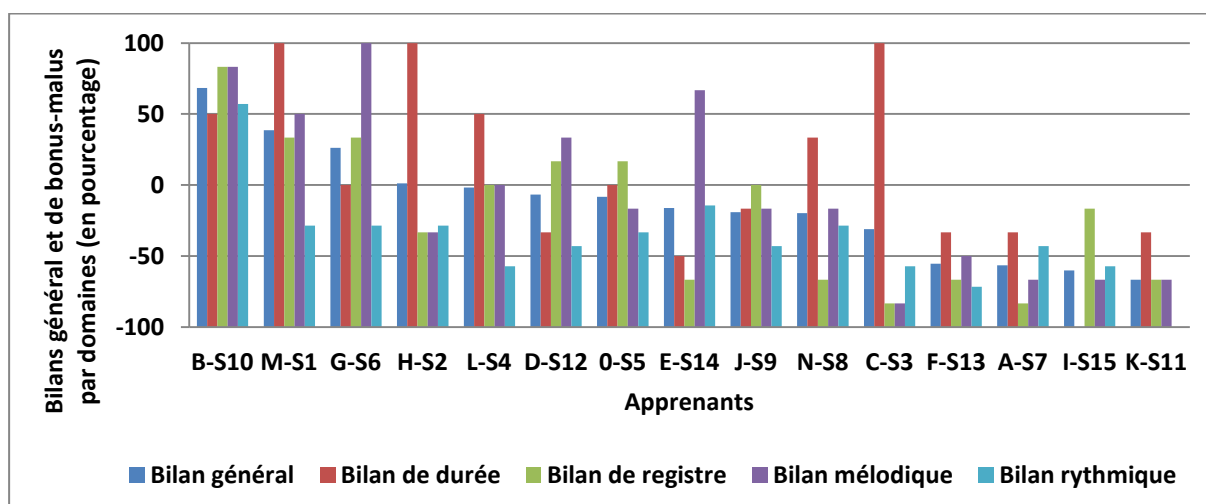


Figure 229 – Bilan général et par domaine prosodique des 15 apprenants selon le système de bonus/malus sur 25 variables

Ayant obtenu les données nécessaires, on peut à présent s'orienter vers une analyse comparative des données perceptives et par variables.

5 Analyse comparative des résultats obtenus selon les deux méthodes : par variables et par perception (expérimentation STAT-NN6)

On a constaté précédemment la robustesse de la hiérarchisation perceptive des fichiers non natifs par la stabilité observée malgré les divers profilages d'évaluateurs. La moyenne des résultats de cette évaluation a été traduite numériquement afin d'objectiver la comparaison

entre les deux types d'évaluation : classique (par perception) et semi-automatique (par application de variables). Procédons à une comparaison des résultats sous forme de pourcentage de correspondance entre les deux formes de classement.

Profilage des évaluateurs	Nombre de catégories	Pourcentage de correspondance entre les classements perceptifs et par variables pour les 18 variables testées
68 experts (totalité des évaluateurs)	5	23,42
	3	56,83
40 experts francophones	5	24,64
	3	59,10
25 experts anglophones	5	21,43
	3	53,71
3 experts d'autres langues	5	23,80
	3	52,38
18 experts (totalité des experts profilés spécifiquement)	5	27,78
	3	59,52
11 experts francophones profilés spécifiquement	5	29,22
	3	59,74
7 experts anglophones profilés spécifiquement	5	25,51
	3	59,18

Tableau 110 – Moyenne générale de correspondance entre les classements perceptif et semi-automatique en fonction du nombre de catégories retenues et du profilage des experts évaluateurs (en pourcentage)

On voit que l'adéquation entre les résultats des deux systèmes évaluatifs reste très faible lorsque cinq catégories sont testées. Elle se vérifie à 23,42% pour l'ensemble des évaluateurs, et grimpe jusqu'à 29,22% dans le cas des experts francophones profilés spécifiquement.

C'est par contre en regroupant les catégories de classement pour n'en totaliser que trois que l'on obtient des résultats satisfaisants. Les plus intéressants figurent en gras dans le tableau ci-dessus. La conformité des deux modes évaluatifs par le plus grand nombre d'évaluateurs (68) est de 56,83%. Elle dépasse 59% dans le cas des évaluateurs francophones (59,10%), et des experts profilés selon leur durée de test et de l'engagement professionnel d'enseignant : onze experts francophones profilés (59,74% : meilleur score), sept experts anglophones profilés (59,18%), et l'ensemble des 18 experts profilés (59,52%).

Bien sûr, un profilage à 18 variables est bien moins représentatif d'un niveau que s'il consistait en davantage de variables, surtout si celles-ci ne sont pas les plus représentatives, mais il est sans doute éclairant de vérifier ce qui justifie les alternances de classement de ces fichiers. Ainsi parmi les divergences que dévoile le tableau suivant, ainsi que les histogrammes précédents, on constate que le critère de rapidité de lecture ne correspond pas à

la qualité prosodique maximale : l'apprenant C-S3 en est un excellent exemple. De même, l'apprenant G est d'une rapidité de lecture moyenne : il atteint 0 sur l'échelle graduée de 100 à -100, alors que sur la même échelle, l'apprenant B reçoit la graduation 50, et M la graduation maximale (100).

Source		Classement
Perception (5 catégories)	68 experts	G > M >> B >>> N, L, F >>>> Fichier 0, H, K >>> A, D >> I, C, J, E
	18 experts profilés	G > B, M >> F, N >>> Fichier 0, L, K, H, D, A >> I, C, J, E
Semi-automatisme	Bilan général 18 variables	B > M > G > D > 0 > H > J > L > C > N > E > I > F > A > K
	Bilan général 25 variables	B > M > G > H > L > D > 0 > E > J > N > C > F > A > I > K

Tableau 111 – Classements individuels des fichiers non natifs sur axe évaluatif continu

Il est en outre intéressant de vérifier les caractéristiques du niveau le plus élevé. On s'aperçoit que B et G reçoivent 100% de bonus dans le domaine mélodique, et M 50%. Lorsqu'on examine les résultats des 25 variables, la différence qualitative devient alors flagrante : G est le seul apprenant à recevoir 100% de bonus dans le domaine mélodique ; B arrive second en en recevant 83,3% et M 50%. Dans le domaine de la vitesse de lecture, que ce soit sur 18 ou 25 variables, M réussit un parcours sans faute avec 100% de bonus (ce qui « compense » en quelque sorte ses maladrotes mélodiques), B engrange un bilan mitigé de 50% et G 0%.

À l'autre bout de l'échelle, les fichiers les plus faibles se caractérisent par leur grande lenteur, la faiblesse de leur registre, leur mélodie et leur rythme. Les histogrammes précédents montrent bien que les meilleurs apprenants totalisent un pourcentage élevé de réussite dans plusieurs domaines alors que les plus faibles sont en échec dans la plupart des cas. La rapidité de lecture ne peut masquer à elle seule les autres déficiences, même si elle contribue à faire illusion.

6 Synthèse

L'objectif de ce dernier chapitre était de valider le processus de profilage à finalité évaluative de cette thèse via ses résultats en les comparant à ceux obtenus par méthode

classique perceptuelle d'experts anglicistes et ce par deux approches de classification : celle consistant à ordonner chaque fichier le long d'un axe évaluatif continu de la meilleure qualité à la plus faible, l'autre étant de catégoriser les fichiers par rapport à des modèles distributifs.

L'analyse comparative des résultats conclut à un faible degré de validation lorsque l'on opère sur une base distributive en cinq catégories puisqu'elle affiche 23,42% validés pour l'ensemble des 68 évaluateurs, son seuil maximal étant de 29,22% dans le cas des experts francophones sélectionnés par profilage spécifique. Néanmoins, lorsque les groupes catégoriels sont réduits à trois, la conformité des deux modes évaluatifs atteint 56,83% avec 68 experts (effectif maximal), 59,10% avec les 40 experts francophones, 59,18% avec les sept experts anglophones profilés, 59,52% avec l'ensemble des 18 experts profilés et culmine à 59,74% avec les onze experts francophones profilés.

Le résultat de ce mode de classement en trois catégories peut être mis en perspective par rapport aux résultats de la recherche de Tortel (2009). Neuf paramètres rythmiques (métriques) trouvés dans la littérature lui avaient permis de prédire l'appartenance d'apprenants anglicistes à l'une des trois catégories de niveaux de maîtrise langagière (voir Chapitre II). Sa prédiction a abouti à une validation à 66,74% par la combinaison à quatre variables %V, ΔV , cvV, nPVI_V. Elle préfère néanmoins la fiabilité à 62,30% de la combinaison à deux variables ΔC , cvC pour des raisons d'ordre didactiques. Les résultats de 56,83% et plus de 59% atteints par la présente recherche sont comparables quantitativement, bien que les effectifs de non natifs testés et d'évaluateurs soient plus faibles. En cela, la Question de Recherche QR2, rappelée ci-dessous, est validée.

QR2 — À quel degré de fiabilité les variables à la base des profils prosodiques d'apprenants et de leurs grilles évaluatives sont-elles validées par comparaison de leurs résultats à ceux d'un test perceptif effectué par des experts ?

Le classement individuel des apprenants les uns par rapport aux autres contribue à éclairer les raisons de leur rang global par celui qu'ils obtiennent dans chaque domaine prosodique. Ainsi, une fois de plus, on constate que la rapidité de lecture n'est pas un critère de bonne qualité prosodique, mais par contre qu'il est un critère de mauvaise qualité prosodique lorsqu'elle est trop faible. La qualité mélodique est prépondérante pour déterminer la meilleure qualité prosodique. Il existe donc une hiérarchie des domaines prosodiques entre eux en fonction de certains critères conditionnels.

Conclusion

1 Rappel des objectifs

L'objectif de cette thèse, affiché dans son titre « Élaboration de critères d'évaluation prosodique pour anglicistes francophones à partir de la notion de « profil » et du phénomène d'emphase », est centré sur l'évaluation prosodique non native via apprenants anglicistes francophones. Ce pari hasardeux, motivé par la nécessité de trouver des points d'ancrage tangibles utilisables dans une évaluation automatisée de la prosodie, domaine « erratique » par excellence, n'aurait pu aboutir sans l'application d'une méthodologie à l'épreuve d'une telle gageure. Il a pris la forme d'un parcours du combattant, mais est relaté dans cette thèse à la manière d'un journal de voyage, qui ne doit pas détourner de la rigueur méthodologique qui la caractérise.

2 Hypothèses, Questions de Recherche et structure de la thèse

Après avoir abordé objectifs et enjeux propres à l'évaluation prosodique dans son Introduction, ce travail de recherche aborde méthodiquement des angles d'attaque propres à trouver la « quadrature du cercle » ; en fin de Chapitre I, il a présenté les hypothèses et questions de recherche le balisant ainsi que sa démarche générale, lesquelles ont toutes été validées, et H2 approfondie en H2'. Je les rappelle ci-dessous :

H1 — Étant donné la forte variabilité des réalisations prosodiques natives et d'apprenants, des outils spécifiques doivent être utilisés afin de l'analyser. La création de « profils prosodiques » devrait permettre de mieux organiser et hiérarchiser les réalisations inter-locuteurs en fonction de regroupements des variables. (expérimentation validée en fin de Chapitre II)

H2' — Le phénomène d'emphase favorise la variabilité des réalisations prosodiques. Aussi, son repérage permet de limiter l'étude à des entités concentrant la variation maximale **ou au contraire en la réduisant pour cibler des entités prosodiquement stables**. Le phénomène d'emphase devrait alors permettre de mieux comprendre, organiser et hiérarchiser le degré de variabilité des phénomènes prosodiques intra-locuteurs en graduant la variation pour

l'analyser plus précisément ou au contraire pour l'exclure des analyses. (expérimentation validée en fin de Chapitre III)

H3 — Le paramètre de rapidité d'élocution étant un indicateur déterminant de niveau des apprenants, la catégorisation de ceux-ci selon ce paramètre devrait permettre d'organiser la variation inter-locuteurs des résultats des autres variables prosodiques afin de trouver des critères d'évaluation fiables. (expérimentation validée en début de Chapitre IV)

QR1 — Étant donné la forte variabilité des réalisations prosodiques natives et d'apprenants, des outils spécifiques doivent être utilisés ou créés afin de l'analyser. La validation des Hypothèses H1, H2 et H3 devrait permettre de créer des profils prosodiques de non natifs permettant d'orienter la création de traits discriminants évaluatifs. (expérimentation validée dans le Chapitre IV)

QR2 — À quel degré de fiabilité les variables à la base des profils prosodiques d'apprenants et de leurs grilles évaluatives sont-elles validées par comparaison de leurs résultats à ceux d'un test perceptif effectué par des experts ? (expérimentation validée dans le Chapitre V)

La méthodologie utilisée dans cette thèse, aussi appelée « profilage », se base sur le profilage *CRISP-DM*, et sa modélisation, appliquée à la présente recherche, se traduit sous la forme de la spirale ci-dessous.

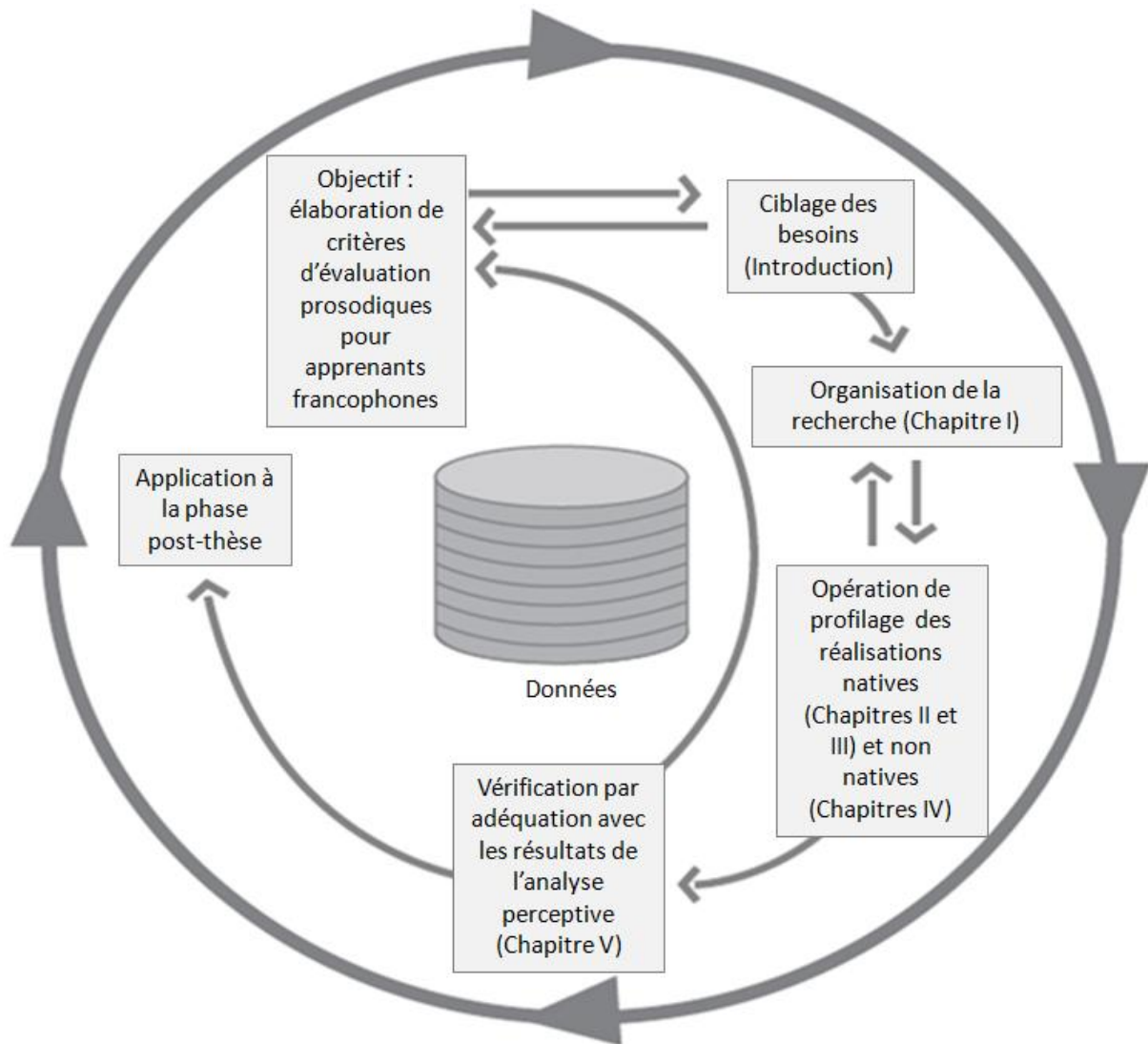


Figure 230 – Modélisation de la recherche de critères distinctifs évaluatifs chez les non natifs anglicistes francophones selon le profilage CRISP-DM (à partir du CRISP-DM Process Guide and User Manual)

De manière plus détaillée, le graphique ci-dessous schématise sommairement les différentes phases d'expérimentation de cette thèse ainsi que leur agencement en vue de l'aboutissement final.

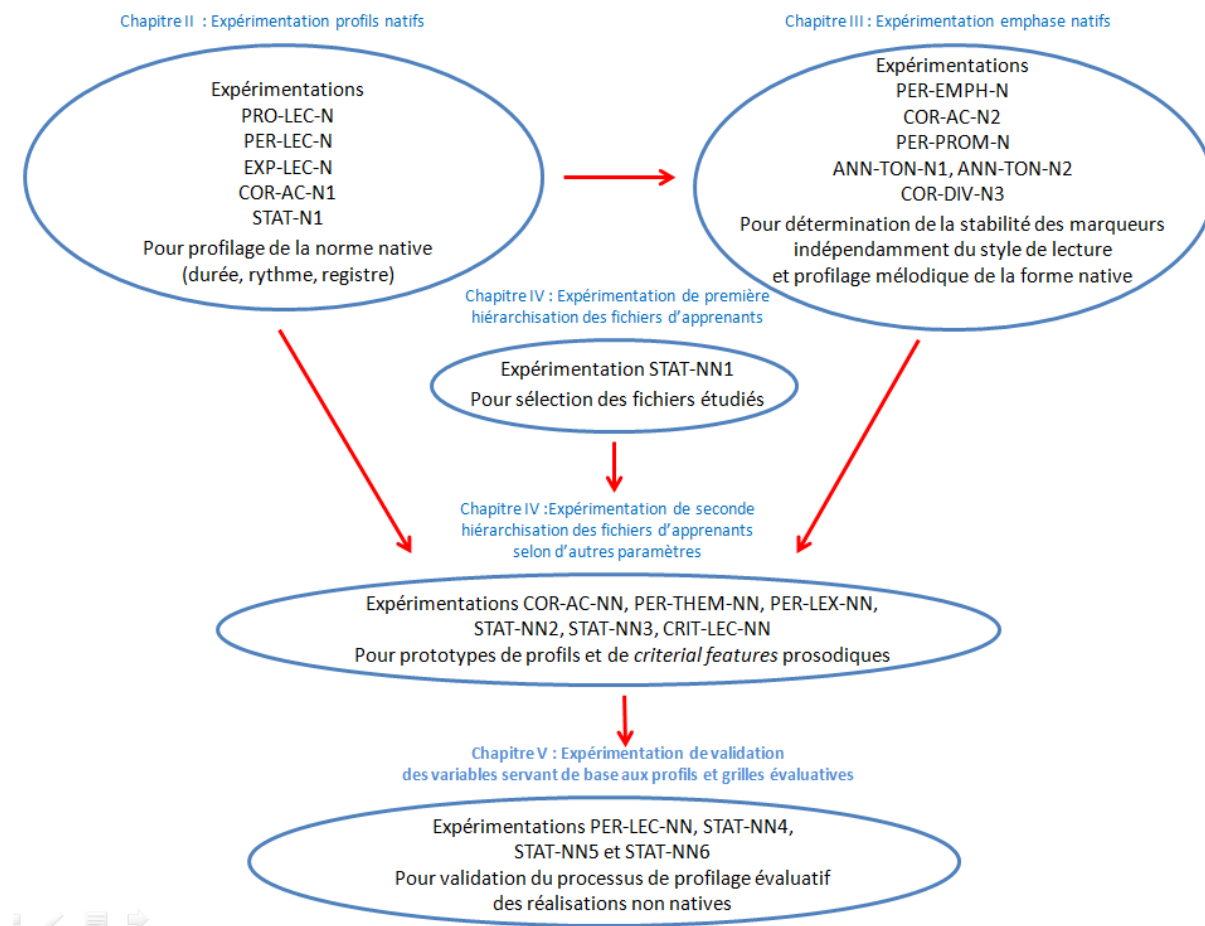


Figure 231 – Organisation des expérimentations sur les natifs et les non natifs pour déterminer les traits discriminants en prosodie de l'interlangue

3 Résultats obtenus

En liaison avec les objectifs assignés, ce travail de recherche sur l'évaluation de la production prosodique en lecture a contribué à fournir les avancées suivantes :

- Des grilles prototypiques modélisatrices de traits distinctifs évaluatifs prosodiques chez les anglicistes francophones,
- Des profils prosodiques de natifs anglophones et d'anglicistes francophones,
- Une validation des résultats ci-dessus par une conformité sur trois catégories évaluatives à 56,83% par perception de 68 experts, et à 59,10% par perception de 18 experts profilés en fonction de leur durée de test et de leur durée d'enseignement,
- Une meilleure compréhension des règles sous-tendant la prosodie native et non native (la connaissance prosodique révèle le fonctionnement transcendant de la

grammaire d'une langue par ses correspondances avec la situation interlocutive),

- Une base évaluative permettant de distinguer et hiérarchiser le niveau prosodique par domaines, pour enclencher une remédiation appropriée avec modulation possible de l'enseignement prosodique,
- Plusieurs axes méthodologiques dont ceux basés sur les techniques de « profilage » ou sur l'utilisation de l'emphase comme outils d'investigation permettant de mettre à jour des influences extralinguistiques, de type social, interlocutif ou subjectif sur la variabilité prosodique.
- Quelques indications sur la prépondérance de certains domaines prosodiques pour situer la qualité des réalisations d'un apprenant à un niveau spécifique d'acquisition prosodique.

Ce travail constitue une étude pionnière ayant le mérite de poser des jalons méthodologiques transférables à d'autres styles de parole que la lecture et de proposer des résultats modélisateurs perfectibles mais tangibles trouvant une application directe en évaluation.

4 Considérations sur ce travail de recherche et perspectives futures

Cette recherche est ancrée dans une étude sur la lecture et apporte une réponse modélisatrice limitée au phonostyle du conte, et peut-être uniquement du seul conte *The Selfish Giant*. Pourtant, à chaque étape, j'ai gardé à l'esprit la transférabilité de ces résultats à l'évaluation de la parole non guidée pour faciliter une transposition dans ce sens. La caractérisation de certaines variables s'inscrit dans cette optique : elles pourront servir de base à une étude ultérieure dans cette direction. Ce travail reste néanmoins modeste par les résultats obtenus. On a vu qu'ils étaient limités à un support de lecture particulier et le plus souvent à une seule occurrence. Cela est extrêmement réduit. On a aussi pu constater les différences dans les résultats obtenus suite à l'étude d'occurrences multiples d'éléments semblant pourtant analogues. Aussi, une grande prudence s'impose. Barsalou met en garde contre une recherche non aboutie (1999: 665) :

Skill results from compiling simulations for most of the plans in a domain through extensive experience (cf. Anderson 1993; Logan 1988; Newell 1990). [...]

Expertise is achieved when an agent can almost always simulate what is about to occur, rarely stopping to revise the simulation.

Le travail de recherche effectué ici doit être poursuivi car il ne propose que des prototypes modélisateurs, non des grilles évaluatives de référence ayant réussi leur validation face à de multiples confrontations avec la réalité du terrain sur les supports variés.

Outre les résultats obtenus conformément à l'objectif visé et les perspectives nouvelles qui s'ouvrent en linguistique et didactique, ce travail de recherche se distingue par le cheminement méthodologique suivi tout au long de son parcours, fruit d'une approche délibérément transversale, la seule qui m'a semblé être à la hauteur de l'enjeu pour atteindre l'objectif fixé. Alors que profils et profilage bénéficient actuellement d'un grand succès tous azimuts comme outils et méthodologie d'investigation, la notion d'emphase est essentiellement utilisée comme phénomène à décrire plus que comme outil d'identification et de sélection. Profilage et emphase ont été décisifs pour trouver la cohérence dans la diversité et les critères internes d'organisation caractérisant les réalisations. Il importait de trouver des repères tangibles dans la foison de corrélats acoustiques que constitue la parole d'apprenant, ou même celle de natifs. Une norme de guidage était nécessaire, j'ai choisi de créer celle de « bon lecteur ». De la réalité physique quantifiable, je suis parvenue à extraire une réalité qualitative. Cette valeur devenant prototypique dans le cas du corpus étudié, elle a pu rendre compte à son tour des corrélats acoustiques qui en sont l'essence. La question était de savoir, parmi cette multitude de données acoustiques, lesquelles étaient pertinentes dans la construction d'un profil prosodique. L'emphase favorisant la variabilité des réalisations phonétiques, leur repérage a permis de cibler l'étude sur les entités concentrant la variation maximale, d'établir des catégories pour classer les occurrences rencontrées, puis de les hiérarchiser les unes par rapport aux autres. Aussi, ce n'est qu'après avoir « fait surgir » les zones propices à une étude fructueuse que j'ai véritablement pu m'orienter vers une analyse *corpus-driven*. La modélisation acoustique de la perception auditive résulte de l'application des deux notions de profils et d'emphase en s'affranchissant des limites d'ordre subjectif tout en leur rendant leur juste place dans la modélisation de groupes, soit en les sélectionnant, soit en les écartant. J'ai œuvré dans le sens d'une méthodologie de recherche appliquée à tous les niveaux et ne pense en aucun cas prêter le flanc à une éventuelle accusation de « picorage »⁹⁸

⁹⁸ « Le décret du 3 juillet 1996 relatif à l'enrichissement de la langue française recommande l'emploi du terme « picorage » plutôt que l'expression anglo-saxonne *cherry picking* (page 137 ; LEHU, J.-M. (2012). *L'encyclopédie du marketing commentée & illustrée*. Paris : Eyrolles.)

(ou *cherry-picking*⁹⁹), qui, en minimisant certains paramètres suite à une orientation du choix des données étudiées, favoriserait un biais remettant en cause la validité des données obtenues. Chacun de ces paramètres est présenté ci-dessous afin d'établir les hypothèses générales dont il fait l'objet, les données requises pour constituer les profils, les résultats attendus à l'issue de l'analyse (hypothèses opérationnelles), ainsi que la grille d'analyse qui est l'aboutissement de cette étape du parcours, qui stipule comment est opéré le recueil de données et la sélection des éléments à mesurer. Le protocole de sélection vise à constituer une grille d'analyse pour caractériser ces paramètres, les comparer entre eux et déterminer dans quelle mesure ils sont révélateurs d'un profil prosodique ou susceptibles de constituer des critères d'évaluation. Dans l'idéal, il aurait été préférable que les mesures prises manuellement puissent être confrontées à des mesures prises par un autre linguiste afin de situer les écarts inter-annotateurs. C'est plutôt la diversité des outils utilisés qui a tenu lieu de confrontation des mesures.

Sans être une apologie de la subjectivité, cette thèse a pris en compte ce paramètre pour explorer des domaines inaccessibles autrement. De nombreuses études prennent en compte la subjectivité du locuteur, de l'auditeur, voire de l'annotateur ou de l'analyste. Il est beaucoup plus rare de remettre en cause la subjectivité du chercheur prosodiste, surtout lorsqu'il s'appuie sur des analyses acoustiques sur corpus. Pourtant, il s'agit là de comprendre le monde de l'humain, les interactions avec son environnement, afin de pouvoir répondre de manière optimale aux défis auxquels l'humain est confronté. Pour ma part, mon but n'a pas été d'éradiquer cette subjectivité, bien au contraire, mais de la cadrer et de l'utiliser pour concevoir des hypothèses, corroborer des conclusions, poursuivre un raisonnement, c'est-à-dire l'amener à frayer un chemin dans le non accessible pour pouvoir ensuite l'appuyer par démonstration objective. C'est ainsi que j'ai systématiquement utilisé des repères sensoriels (couleurs, diagrammes et figures, schémas, tableaux) comme points d'appui tangibles pour m'aider à accéder à une compréhension plus fine des phénomènes, destinée à toujours trouver de nouvelles pistes, de les poursuivre ou d'en abandonner d'autres¹⁰⁰. En ce sens, je participe à une certaine réhabilitation de la subjectivité en essayant d'éprouver en quelque sorte la

⁹⁹ "Cherry picking, the selective culling of evidence to support a claim, is a common problem in political debate and social science research. [...] Cherry picking is compounded as people have a tendency to more carefully scrutinize data which conflict with their preconceived assumptions and to readily accept data that support their position." (page 37; KLASS, G. M. (2012). *Just Plain Data Analysis: Finding, Presenting, and Interpreting Social Science Data*. Plymouth: Rowman & Littlefield Publishers, Inc.)

¹⁰⁰ Voir en annexe le récapitulatif des étapes suivies dans un de mes nombreux fichiers de travail (Mesures_apprenants10.xlsx), feuille après feuille, de la prise de mesures à leur interprétation.

« sensualité » des données en faisant corps avec l'objet de recherche, ce qui est loin d'être le cas de l'application de statistiques à l'aveugle sans démarche cognitive ni même perceptuelle. En effet, les présents travaux prennent appui sur le paradigme de l'énaction¹⁰¹, à tous les niveaux et surtout pour l'étude de l'emphase et des profils prosodiques : l'individu est un tout qu'il faut relier à son environnement, ses expériences passées, ses modes de pensée et profils d'apprentissage. L'objectif n'est pas de rechercher un modèle dans l'intelligence artificielle, mais de comprendre le vivant en faisant émerger la conscience. Aussi, il m'a semblé nécessaire de me démarquer des démarches fondées sur des phénomènes d'auto-organisation et de dépasser leurs restrictions afin de trouver une correspondance externe mais en liaison avec l'objet d'étude. C'est ce que je me suis efforcée de faire, très modestement, via la subjectivité de la perception, qui enregistre malgré l'individu ses interactions avec l'environnement dans lequel il évolue, et dont on retrouve les traces en prosodie.

Toujours sur le plan méthodologique, j'ai émis l'hypothèse qu'un nombre d'angles d'attaque accru permettrait de percer ou d'affiner la compréhension des phénomènes prosodiques tout en offrant un élargissement des options de catégorisation. Ainsi, on a vu que l'étude phonétique se déclinait ordinairement selon trois approches : phonétique articulatoire, phonétique perceptive et phonétique acoustique ; la catégorisation prosodique en acoustique s'articulait ordinairement selon divers domaines selon les écoles. J'ai formulé l'hypothèse qu'un éclairage nouveau pouvait être apporté par les phénomènes de profils-profilage et d'emphase en réconciliant la prosodie avec sa nature, sa raison d'être. L'utilisation de ces phénomènes comme outils d'investigation a fait émerger un agencement prosodique non exploré systématiquement et en a offert une réorganisation complète qu'il a été possible de catégoriser. À partir de là, de nouveaux axes d'étude ont été dégagés perceptivement puis analysés acoustiquement. Les variables obtenues ont alors créé de nouvelles hiérarchies quantitatives et qualitatives d'où sont extraits les prototypes modalisateurs de grilles évaluatives. Le principe fondamental a été d'augmenter le degré de granularité permettant de structurer le phénomène prosodique. Je dois aussi mentionner qu'à certains moments de mon étude le manque de granularité dans les possibilités d'analyse m'a fait créer des catégories considérées comme peu orthodoxes selon les canons de l'étude prosodique. Alors que j'ai tenté de réconcilier perception et acoustique autant que faire que peut, dans certains cas je me suis focalisée sur l'aspect purement acoustique en me dégageant complètement de l'aspect

¹⁰¹ Le paradigme de l'énaction, proposé par Maturana, Varela *et al.*, est destiné à concevoir la cognition en privilégiant la manière dont les organismes vivants et esprits humains s'organisent eux-mêmes en interaction avec leur environnement.

perceptif (démarche cohérente par rapport à l'objectif final d'automatisation de l'évaluation). C'est la raison pour laquelle je n'ai pas poursuivi les analyses que j'avais faites avec *Prosogramme* (Mertens 2004), basé sur le phénomène perceptif. Au contraire, je souhaitais m'en éloigner pour accéder à la catégorisation d'une granularité plus fine en fonction des indices que l'acousticien constate visuellement sur un spectrogramme et numériquement dans ses mesures sans que la différence soit perceptible, du moins consciemment : c'est en quelque sorte le domaine du « subliminal », défini comme *not recognized or understood by the conscious mind, but still having an influence on it*¹⁰². Cette catégorisation plus fine permet d'accéder, selon moi, à une meilleure compréhension du caractère qualitatif de la prosodie en favorisant une analyse plus proche du continuum acoustique.

Outre son implication dans le domaine méthodologique, les travaux de cette thèse sur les réalisations natives et non natives attirent l'attention sur deux autres phénomènes. La plupart des études considèrent la dichotomie catégorielle entre natifs et non natifs. Les données obtenues ici montrent une interconnexion entre les deux groupes car les valeurs observées se répartissent le long d'un axe avec intersection. La prosodie s'organise donc comme un continuum aussi bien quantitatif (mesures acoustiques) que qualitatif (le meilleur apprenant avait été pris pour un natif anglophone par une évaluatrice francophone). Une deuxième constatation est qu'il existe une gradation qualitative à l'intérieur-même du groupe des natifs, les meilleurs lecteurs se trouvant pour la plupart dans les groupes professionnels et matures.

En ce qui concerne les traits distinctifs et leurs variables, on a vu que leur nombre accroît la fiabilité de leur résultat. Pourtant, on a remarqué la prépondérance de certains domaines sur d'autres et une hiérarchie des variables opérée entre elles. Un niveau élevé en prosodie anglaise ne peut tolérer de mauvais résultats dans aucun domaine, mais le meilleur apprenant se doit de privilégier le domaine mélodique. Il en est de même pour les natifs et il est particulièrement intéressant de remarquer que la variable mélodique de la première syllabe de *nobody* bénéficie d'une suprématie sur les autres. Les résultats natifs et non natifs collectés jusqu'à présent sur cette variable me portent à croire que c'est elle qui fournit le plus sûr trait discriminant de qualité prosodique chez les non natifs de haut niveau et chez les natifs eux-mêmes : la meilleure réalisation hyper-complexe est celle d'un narrateur en voix off de stature

¹⁰² [<http://dictionary.cambridge.org/us/dictionary/english/subliminal>], consulté en 2017.

internationale¹⁰³, les suivantes dans l'ordre hiérarchique décroissant sont celles de lecteurs matures très expressifs. Ce point, comme beaucoup d'autres pistes, mérite une investigation plus poussée encore pour obtenir une reconnaissance à part entière. Outre la découverte de traits discriminants, il me semble important d'orienter une recherche future selon une approche conditionnelle plus fine établissant une hiérarchie des domaines, des variables et des critères entre eux selon le niveau de maîtrise prosodique.

¹⁰³ Christopher Kent [<http://www.christopherkent.co.uk/>], consulté en 2017.

Bibliographie

Cette bibliographie référentielle se limite aux sources fondamentales mentionnées mais non présentées intégralement dans le texte de cette thèse et faisant l'objet d'une référence directe.

1 Classement par auteur

ABERCROMBIE, D. (1967). *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

ABERCROMBIE, D. (1968). Paralanguage. *British Journal of Disorders of Communication*, vol. 3 (55-59).

ALBRESPIT, J. (2011). *Construire l'énoncé en anglais : Voix, négation, exclamation, interrogation*. Université de Toulouse-Le Mirail : Presses Universitaires du Mirail.

ALLEN, W. S. (1954). *Living English speech : stress and intonation practice for the foreign student*. London ; New York : Longmans.

ARNAULD, A., LANCELOT, C. ([1660] 1803). *Grammaire générale et raisonnée de Port-Royal*. Paris : Perlet.

[http://books.google.fr/books?id=Qe8FAAAAQAAJ&pg=PP15&hl=fr&source=gb_s_selected_pages&cad=3#v=onepage&q&f=false]

AUSTIN, J. L. (1962). *How to do things with Words*. Oxford : Oxford University Press.

AVANZI, M., LACHERET-DUJOUR, A., VICTORRI, B. (2008). ANALOR : Un outil d'aide pour la modélisation de l'interface prosodie - grammaire. *Cahiers du CERLICO*, 2008, 21 (27-46).

AVANZI, M., LACHERET-DUJOUR, A., VICTORRI, B. (2008). ANALOR. A Tool for Semi-Automatic Annotation of French Prosodic Structure. Actes de la conférence ANALOR. *A Tool for Semi-Automatic Annotation of French Prosodic Structure*. Campinas : Brésil (119-122).

[http://hal.archives-ouvertes.fr/docs/00/33/46/56/PDF/avanzi_et_alii-SP08.pdf]

BAILEY, K. M. (1999). *Washback in Language Testing* (RM-99-04, TOEFL-MS-15). *Educational Testing Services*, ETS Research Memorandum (RM-99-04.pdf).

[https://www.ets.org/research/policy_research_reports/publications/report/1999/icmj]

- BALLIER, N., MARTIN, Ph., AMAND, M. (2016). Variabilité des syllabes réalisées par des apprenants de l'anglais. *Actes de la conférence conjointe JEP-TALN-RECITAL 2016*, volume 1 : JEP 2016, Paris, France (732-740).
[<https://jep-taln2016.limsi.fr/actes/Actes%20JTR-2016/V01-JEP.pdf>]
- BALLIER, N., MARTIN, Ph. (2015). Speech annotation of learner corpora. Dans Granger S., Meunier F. & Gilquin G. (dir.). *The Cambridge Handbook of Learner Corpus Research*, CUP (107-134).
[<http://www.cambridge.org/us/academic/subjects/languages-linguistics/applied-linguistics-and-second-language-acquisition/cambridge-handbook-learner-corpus-research>]
- BALLIER, N., DELAIS-ROUSSARIE, E., HERMENT, S., TORTEL, A. (2014). Modélisation de l'intonation interlangue : le cas des questions. *XXXe Journées d'Études sur la Parole (JEP 2014)*, Juin 2014, Le Mans, France. *Actes des XXXe Journées d'Études sur la Parole* (740-748).
[https://scholar.google.fr/citations?view_op=view_citation&hl=fr&user=1DqF0BgAAAAJ&citation_for_view=1DqF0BgAAAAJ:SeFeTyx0c_EC]
- BARKER, F., POST, B., SCHMIDT, E., McCARTHY, M. (2011). Identifying criterial aspects of pronunciation in L2 English across CEFR levels: Implications for language learning. *The Impact of Applied Linguistics: Proceedings of the 44th Annual Meeting of the British, Scitsiugnill Press*.
- BARSALOU, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22 (577-609, 649-660).
- BAVOUX, C. (1997). Idiolecte. Dans Moreau, M.-L. (dir.) *Sociolinguistique : Les concepts de base*. Sprimont : Mardaga (265).
- BIBER, D., JOHANSSON, S., LEECH, G., CONRAD, S., FINEGAN, E. (1999). *Longman Grammar of Spoken and Written English*. Harlow : Longman.
- BOGAARDS, P. (1994). *Le vocabulaire dans l'apprentissage des langues étrangères*. Paris : Didier.
- BOISSEAU, M., HAMM, A. (2015). Introduction. Dans Boisseau, M., Hamm, A. (dir.) *Saillance : La saillance en langue et en discours*, Volume 2. Besançon : Presses universitaires de Franche-Comté (7-13).
- BOLINGER, D. (1989). *Intonation and Its Uses: Melody in Grammar and Discourse*. Stanford : Stanford University Press.
- BOTTINEAU, D. (2008). Language and enaction. Dans Stewart, J., Gapenne, O., Di Paolo, E. (dir.) *Enaction: towards a new paradigm for cognitive science*, MIT (1-67).
- BOULONNAIS, D. (2006). So et les phénomènes d'expansion contextuelle. Dans Girard-Gillet G. (dir.) *Aux Marges du texte, texte et co-texte*. CIEREC, Travaux 128, Université de Saint-Etienne.
- CALAS, F. (2011). *La stylistique : Méthode et commentaires*. Paris : Armand Colin.

- CARLETTA, J. (2006). The NITE XML Toolkit: data model and query language. *Language Resources and Evaluation*, 2006, 39 (4), (313-334).
- CARR, Ph. (2013). *English Phonetics and Phonology*. Oxford : Blackwell/Wiley.
- CAUVIN, É. (2014). « Plaidoyer pour une politique linguistique LANSAD intégrant la phonologie anglaise ». Dans Daval, R., Hilgert, É., Nicklas, Th., Thomières, D. (dirs.) *Mélanges en hommage à Pierre Frath*. Reims : Éditions et Presses universitaires de Reims (371-402).
- CAUVIN, É. (2013). Intonational phrasing as a potential indicator for establishing prosodic learner profiles. Dans Granger S., Gilquin G. & Meunier F. (dir.). *Corpora and Language in Use. Actes du colloque Learner Corpus Research 2011, "20 years of learner corpus research: looking back, moving ahead"*, 15-17 septembre 2011, Louvain-la-Neuve. Louvain : Presses Universitaires de Louvain (75-88).
- CAUVIN, É. (2012). Le CECRL et l'évaluation de la compétence phonologique: point de vue à partir d'un corpus sur l'anglais. *Les Langues Modernes*, n°1-2012. Paris : APLV (33-40).
- CAUVIN, É. (2012). Interlocution et variation prosodique dans la lecture à haute voix : de l'influence de la subjectivité des lecteurs sur celle des auditeurs. Dans Douay C. & Roulland, D. (dir.) *L'interlocution comme paramètre*. Rennes : Presses Universitaires de Rennes (279-302).
- CELCE-MURCIA, M., BRINTON, D. M., GOODWIN, J. M. (1996). *Teaching Pronunciation*. Cambridge : Cambridge University Press.
- CHAPMAN, P. *et al.* ([1999] 2000). CRISP-DM consortium (eds). *CRISP-DM 1.0 Step-by-step data mining guide* (6).
- CORDER, S. P. (1981). *Error Analysis and Interlanguage*. Oxford : Oxford University Press.
- CRUTTENDEN, A. ([1962] 2014). *Gimson's Pronunciation of English, 8th edition*. London & New York : Routledge.
- CRUTTENDEN, A. ([1986] 1997). *Intonation*. Cambridge University Press.
- CRYSTAL, D. ([1980] 2008). *A Dictionary of Linguistics and Phonetics*. Malden, Oxford & Victoria : Blackwell Publishing.
- CRYSTAL, D. ([1982] 1992). *Profiling Linguistic Disability*. San Diego : Singular Publishing Group.
- CRYSTAL, D., QUIRK, R. (1964). *Systems of Prosodic and Paralinguistic Features in English*. The Hague : Mouton & co.
- CYR, P. ([1996] 1998). *Les stratégies d'apprentissage*. Paris : CLE International.
- DAUER, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, 1983, 11 (51-62).
- DE KETELE, J.-M. (1980). *Observer pour éduquer*. Berne : Peter Lang.

- DE LOOZE, C. (2010). *Analyse et interprétation de l'empan temporel des variations prosodiques en français et en anglais*. Thèse de doctorat sous la direction de D. Hirst. Aix-en-Provence : Université de Provence.
- DI CRISTO, A. (2013). *La prosodie de la parole*. Bruxelles : Solal Editeurs.
- DI CRISTO, A. (2004). La prosodie au carrefour de la phonétique, de la phonologie et de l'articulation formes-fonctions. *Travaux Interdisciplinaires du Laboratoire Parole et Langage d'Aix-en-Provence (TIPA)*, 2004, 23 (67-211).
[<https://hal.archives-ouvertes.fr/hal-00285554/document>]
- DIDEROT, D., D'ALEMBERT, J. (LE ROND) (dir.) ([1751] 1772). *Encyclopédie, ou Dictionnaire raisonné des sciences, des arts et des métiers*. Paris : chez Briasson : David l'aîné : Le Breton : Durand.
- DOUAY, C. (2000). *Éléments pour une théorie de l'interlocution — Un autre regard sur la grammaire anglaise*. Rennes : Presses Universitaires de Rennes.
- DOWNING, A., LOCKE, Ph. ([1992] 2006). *English Grammar : A University Course*. London: Routledge.
- DURAND J. (2005). Chapitre 3 - Les primitives phonologiques : des traits distinctifs aux éléments. Dans Nguyen, N., Wauquier-Gravelines, S. & Durand, J. (dir.), *Phonologie et phonétique : forme et substance*. Paris : Hermes.
- English Profile (2011). *EnglishProfile : Introducing the CEFR for English, Version 1.1*. Cambridge : UCLES/CUP.
- FADDEN, L. (2006). Prosodic Profiles of Suspects' Speech. Actes du colloque *Conference on Language, Psychology and the Law*, University of Leicester, UK. July 10-12, 2006.
- FERRAGNE, E. (2013). Automatic suprasegmental parameter extraction in learner corpora. Dans Ballier, N., Díaz-Negrillo, A. & Thompson, P. (dir.) *Automatic Treatment and Analysis of Learner Corpus Data*. John Benjamins (151-168).
- FÓNAGY, I. ([1983] 1991). *La vive voix : Essais de psycho-phonétique*. Paris : Payot.
- FOURNIER, J.-M. (2010). *Manuel d'anglais oral*. Paris : Ophrys.
- GERADTS, Z., SOMMER, P. (2008). D6.7c: *Forensic Profiling*, FIDIS Deliverables, 6 (7c).
[http://www.fidis.net/fileadmin/fidis/deliverables/fidis-wp6-del6.7c.Forensic_Profiling.pdf]
- GILBERT, E. (1993). La théorie des Opérations Énonciatives d'Antoine Culioli. Dans COTTE P. et al. *Les théories de la grammaire anglaise en France*. Paris : Hachette (69-74).
- GRAMMONT, M. (1937). *Le Vers Français*. Paris : Delagrave.
- GRANGER, S., RAYSON, P. (1998). Automatic profiling of learner texts. Dans Granger, S. (dir.) *Learner English on Computer*. London : Longman (119-131).

- GREENBAUM, S. (1996). *Oxford English Grammar*. Oxford: Oxford University Press.
- GUIERRE, L. ([1984] 1992). *Drills in English Stress-Patterns: Ear and speech training drills and tests for students of English as a foreign language*. Paris : Armand Colin-Longman.
- GUT, U. (2009). *Non-native speech: A Corpus-based Analysis of Phonological and Phonetic Properties of L2 English and German*. Francfort : Peter Lang.
- HAWKINS J. A., FILIPOVIĆ, L. (2012). *English Profile Studies 1. Criterial Features in L2 English: Specifying the Reference Levels of the Common European Framework*. Cambridge : Cambridge University Press (1-12).
- HAWKINS J. A., BUTTERY P. (2010). Criterial Features in Learner Corpora: Theory and Illustrations. *English Profile Journal*, 1, e5 (40-62).
- HERMENT, S., BALLIER, N., DELAIS-ROUSSARIE, E., TORTEL, A. (2014). Modelling interlanguage intonation: the case of questions. Actes du colloque international *Speech Prosody* (7 : 2014 mai 20-23 : Dublin, Irlande) (492-496).
- HERMENT, S., LOUKINA, A., TORTEL, A., HIRST, D., BIGI, B. (2012). AixOx, a multi-layered learners' corpus: automatic annotation. Actes du colloque *4th International Conference on Corpus Linguistics - Language, corpora and applications: diversity and change (CILC 2012, Jaén, Espagne, 22-24 mars 2012)*.
- HERMENT-DUJARDIN S., HIRST, D. J. (2003). L'emphase en anglais : critères acoustiques et analyse statistique. Tentative de définition. *Interfaces Prosodiques*, Mar 2003, Nantes, France (51-56).
[<http://hal.archives-ouvertes.fr/docs/00/36/99/33/PDF/1982.pdf>]
- HERMENT-DUJARDIN, S. (2001). *L'emphase dans le discours spontané anglais : corrélats acoustiques et prosodiques*. Thèse de doctorat sous la direction de D. Hirst, *Laboratoire Parole et Langage*. Aix-en-Provence : Université de Provence.
- HERMENT, S. (article non publié). Is intensity a relevant criterion in the perception of spontaneous speech? The case of emphasis in English.
- HILDEBRANDT, M., BACKHOUSE, J. (dir.) (2005). *D7.2: Descriptive analysis and inventory of profiling practices*. Technical report, *FIDIS (Future of Identity in Information Society)* deliverable.
[www.fidis.net]
- HIRST, D. (2007). A Praat plugin for Momel and INTSINT with improved algorithms for modelling and coding intonation. Proceedings of the *16th International Congress of Phonetic Sciences (ICPhS 2007 Proceedings)*.
[<http://www.icphs2007.de/conference/Papers/1618/1618.pdf>]
- HORGUES, C. (2010). *Prosodie de l'accent français en anglais et perception par les auditeurs anglophones*. Thèse sous la direction d'A. Deschamps. Université Paris 7.

- HORNER, David (2010). *Le CECRL et l'évaluation de l'oral : Anglais*. Paris : Belin.
- HUART, R. (2010). *Nouvelle grammaire de l'anglais oral*. Paris : Ophrys.
- HUDDLESTON, R., PULLUM, G. K. (2002). *The Cambridge Grammar of the English Language*. Cambridge : Cambridge University Press.
- HUSSON, F., LÊ, S., PAGÈS, J. ([2009] 2016). *Analyse de données avec R*. Rennes : Presses Universitaires de Rennes.
- INKOVA, O. (2012). « Avant-propos ». Dans Inkova O. (dir.) *Saillance : Aspects linguistiques et communicatifs de la mise en évidence dans un texte*. Presses universitaires de Franche-Comté. Vol. 1 (9-19).
- JEAN-DAUBIAS, St. (2011). *Ingénierie des profils d'apprenants*. Habilitation à Diriger des Recherches, Université Claude Bernard Lyon 1.
- JESSEN, M. (2010). Forensic speaker identification by experts. Dans Coulthard, M., Johnson, A. (dir.) *The Routledge Handbook of Forensic Linguistics*. Routledge (378-394).
- KERBRAT-ORECCHIONI, C. ([1999] 2009). *L'énonciation: de la subjectivité dans le langage*. Paris : Armand Colin.
- KHALIFA, J.-C. (2004). *Syntaxe de l'anglais : théories et pratique de l'énoncé complexe aux concours*. Paris : Ophrys.
- KINGDON, R. (1958). *The groundwork of English intonation*. London : Longman.
- KLASS, G. M. (2012). *Just Plain Data Analysis: Finding, Presenting, and Interpreting Social Science Data*. Plymouth: Rowman & Littlefield Publishers, Inc.
- LACHERET, A. (2011). Annotation prosodique de corpus oraux, *Journées PFC*, Paris, décembre 2011.
[\[www.projet-pfc.net/bulletins-et-colloques/doc_download/284-lacheret-annotation-prosodiques-de-corpus-oraux.html\]](http://www.projet-pfc.net/bulletins-et-colloques/doc_download/284-lacheret-annotation-prosodiques-de-corpus-oraux.html)
- LACHERET, A. (2007) : Prosodie du discours, une interface à multiples facettes. *Nouveaux cahiers de linguistique française, Interface Discours-Prosodie*, 28 (7-40).
- LANDRAGIN, F. (2015). Sur les aspects multicritères et multidimensionnels de la saillance. Dans Boisseau, M., Hamm, A. (dir.) *Saillance : La saillance en langue et en discours, Volume 2*, Annales Littéraires de l'Université de Franche-Comté. Presses universitaires de Franche-Comté (15-29).
- LANDRAGIN, F. (2012). La saillance : questions méthodologiques autour d'une notion multifactorielle. *Faits de langues*. Peter Lang, 2012 (15-31).
[\[https://halshs.archives-ouvertes.fr/halshs-00690831/document\]](https://halshs.archives-ouvertes.fr/halshs-00690831/document) (version préliminaire HAL)

- LANDRAGIN, F. (2011). De la saillance visuelle à la saillance linguistique. Saillance : Aspects linguistiques et communicatifs de la mise en évidence dans un texte. *Annales Littéraires de l'Université de Franche-Comté*. Presses Universitaires de Franche-Comté (67-84).
- LANDRAGIN, F. (2006). Saillance. Dans Godard, D., Roussarie, L. Corblin, F. (dir.) *Sémanticopédie : dictionnaire de sémantique*. GDR Sémantique & Modélisation, CNRS.
[<http://www.semantique-gdr.net/dico/>]
- LAPAIRE, J.-R., ROTGÉ, W. ([1991] 2002). *Linguistique et grammaire de l'anglais*. Toulouse : Presses Universitaires du Mirail.
- LAPAIRE, J.-R., ROTGÉ, W. ([1991] 1998). *Linguistique et grammaire de l'anglais*. Toulouse : Presses Universitaires du Mirail.
- LARREYA, P., RIVIÈRE, C. (2010). *Grammaire explicative de l'anglais*. Paris : Pearson.
- LAUFER, B. (2005). Lexical Frequency Profiles : From Monte Carlo to the Real World, A Response to Meara. *Applied Linguistics*, 26/4. Oxford : Oxford University Press (582-588).
- LAUFER, B., NATION, P. (1995). Vocabulary Size and Use : Lexical Richness in L2 Written Production. *Applied Linguistics*, Volume 16, Issue 3 (307–322).
- LAVER, J., HUTCHESON, S. (1972). *Communication in Face to Face Interaction: selected readings*. Harmondsworth : Penguin Books.
- LE GRÉZAUZE, E. (2012). *The interface Between the POS/semantic categorization and the prosodic features of the marker SO: a Multilevel study*. Mémoire de master sous la direction de N. Ballier. Université Paris Diderot - Paris 7.
- LEHU, J.-M. (2012). *L'encyclopédie du marketing commentée & illustrée*. Paris : Eyrolles.
- LÉON, P. (2011). *Phonétisme et prononciation du français*. Paris : Armand Colin.
- LOTHE, E. (1982). *La parole et la voix*. Hamburg : H. Buske Verlag (233).
- MARTIN, Ph. (2009). *Intonation du français*. Paris : Armand Colin.
- MARTINEZ, L., BERKHOUT, C. (2009). Poser une question de recherche. *Exercer : la revue française de médecine générale*, 2009, 89 (143-146).
- MATURANA, H. R., VARELA, F. J. (1980). *Autopoiesis and Cognition*. Dordrecht, Holland: D. Reidel.
- McCARTHY, M. (2010). Spoken fluency revisited. *English Profile Journal*, 1, e4.
- MERTENS, P. (2004). Le Prosogramme : une transcription semi-automatique de la prosodie. *Cahiers de l'Institut de Linguistique de Louvain*, 30(1), (7-25).

- MILANOVIC, M. (2002). *Évaluation de compétences en langues et conception de tests*. Strasbourg : Conseil de l'Europe.
[\[http://www.coe.int/t/dg4/linguistic/webpublications/listpubl2_standards_fr.asp?toPrint=yes&\]](http://www.coe.int/t/dg4/linguistic/webpublications/listpubl2_standards_fr.asp?toPrint=yes&)
- MONTLIVAUT, C. M. Guyon (1828). *Grammaire générale et philosophique ; précédée d'un coup-d'œil sur la nature et le mécanisme des langues*. Paris : Pihan Delaforest.
- O'CONNOR, J. D., ARNOLD, G. F. ([1961] 1973). *Intonation of colloquial English*. London : Longman.
- PASSY, P. (1890). *Étude sur les changements phonétiques et leur caractères généraux*. Paris : Didot.
- PEETERS, B. (1992). *Diachronie, phonologie, et linguistique fonctionnelle*. Collection « Bibliothèque des cahiers de l'Institut de linguistique de Louvain ». Louvain-la-Neuve : Peeters.
- POPIN, J. (1998). *La ponctuation*. Paris : Nathan.
- POTTIER, B. (1973). *Le langage : de Ferdinand de Saussure à Noam Chomsky*. Paris : C.E.P.L.-Denoël.
- QUIRK, R., GREENBAUM, S., LEECH, G., SVARTVIK, J. ([1985] 2005). *A Comprehensive Grammar of the English Language*. Harlow : Longman.
- QUIVY, M., TARDIEU, C. ([1997] 2002). *Glossaire de didactique de l'anglais*. Paris : Ellipses.
- RIEGEL, M., PELLAT, J.-C., RIOUL, R. ([1994] 2013). *Grammaire méthodique du français*. Paris : Presses Universitaires de France (667).
- ROACH P. *et al.* (2011). *Cambridge English Pronouncing Dictionary – 18th Edition (CEPD18)* on CD-ROM.
- ROACH P. ([1983] 2009). *English Phonetics and Phonology: a practical course*. Cambridge : Cambridge University Press.
- ROBERT, J.-P., ROSEN, E. (2010). *Dictionnaire pratique du CECR*. Paris : Ophrys.
- ROGERS, H. (2000). *The sounds of language: an introduction to phonetics*. London : Pearson Education Limited.
- ROGGERO, J. (1977). *Introduction à la grammaire anglaise*. Marseille : CRDP.
- SEGUIN, J.-P. (2000). Le point d'exclamation dans Le Père de Famille de Diderot : Une ponctuation musicale ? Dans Dürrermatt J. (dir.) *La ponctuation*. Poitiers : UFR de langues et littératures de l'Université de Poitiers.
- SELINKER, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, 10:3 (209-241).

- SELINKER, L., LAMENDELLA, J. T. (1981). Updating the Interlanguage Hypothesis. Actes du V^{ème} Congrès de / Proceedings of the 5th Congress of l'Association Internationale de Linguistique Appliquée, Savard J. G. et Laforge L. Québec (dir.). Les Presses de l'Université Laval (402-424).
- SHEARER, C. (2000). The CRISP-DM Model: The New Blueprint for Data Mining, *Journal of Data Warehousing*. Vol. 5, No. 4, 2000 (13-22).
- SIMON, A.-C., AUCHLIN, A., AVANZI, M., GOLDMAN, J.-P., *et al.* (2009). Les phonostyles : une description prosodique des styles de parole en français. Dans Abecassi, M., Ledegen, G. (dir.) *Les voix des Français. Volume 2, En parlant, en écrivant*. Peter Lang Publishing.
- TARDIEU, C. (2013). Testing et certification. Recherches en didactique des langues et des cultures : *Les Cahiers de l'Acedle*, volume 10, numéro 2, 2013 – « Apprendre les langues autrement » (237-251).
- TORTEL, A., HIRST, D. (2010). Rhythm metrics and the production of English L1/L2. *Speech Prosody 2010 Conference Proceedings*, Chicago.
- TORTEL, A. (2009). *Évaluation qualitative de la prosodie d'apprenants français : apport de paramétrisations prosodiques*. Thèse sous la direction de D. Hirst. Aix-en-Provence : Université de Provence.
- TRAGER, G. L. (1958). Paralanguage: A First Approximation. *Studies in Linguistics*, Volume 13 (1-12).
- TROUVAIN, J., GUT, U. (2007) (dir.). *Non-Native Prosody : Phonetic Description and Teaching Practice*. Berlin : Mouton de Gruyter.
- VAISSIÈRE, J. (2006). *La phonétique*. Paris : Presses Universitaires de France.
- VAISSIÈRE, J. ([1983] 2016). Language-Independent Prosodic Features. *Prosody: Models and Measurements*, Springer Verlag (53-65).
- VARELA, F. J., THOMPSON, E. T., ROSCH, E. (1993). *The Embodied Mind: Cognitive Science and Human Experience*. MIT Press.
- VOGEL, I. (2016). Chaire internationale du Laboratoire d'excellence « Fondements Empiriques de la Linguistique », Labex EFL à Paris (juin et juillet 2016).
- WACEWICZ, S., ŻYWICZYŃSKI, P. (2012). Human Honest Signalling and Nonverbal Communication. *Psychology of Language and Communication*, 2012, Volume 16, Numéro 2 (113-130).
- WELLS, J. C. ([1990, 2000] 2008). *Longman pronunciation dictionary*. Harlow : Pearson-Longman.
- WELLS J. C. (2006). *English Intonation: An Introduction*. Cambridge : Cambridge University Press.
- WICHMANN, A. (2004). The intonation of *Please*-requests: a corpus-based study. *Journal of Pragmatics*, Volume 36, Issue 9, September 2004 (1521-1549).

WICHMANN, A. (2000). *Intonation in texts and discourse: Beginnings, middles and ends*. Harlow : Longman.

WILDE, O. ([1888] 1995). The Selfish Giant. Dans *The Happy Prince and Other Tales*. London : Everyman's Library.

2 Sitographie et logiciels

Association des Anglicistes pour les études de Langue Orale dans l'Enseignement Supérieur, secondaire et élémentaire (ALOES)

[<https://aloesfrance.wordpress.com/>]

ANCET, J., ANCET, L. *La Caminata*. Du rythme : de l'importance du rythme dans la danse d'improvisation qu'est le tango.

[<https://sites.google.com/site/annecytango/chroniques/du-rythme>], consulté en 2017.

ENGLISH [Corpus]. *Laboratoire parole et langage - UMR 7309 (LPL) (2008)*. ORTOLANG (Open Resources and TOols for LANGuage)

[www.ortolang.fr], [<https://hdl.handle.net/11403/sldr000731/v2>]

Association of language testers in Europe (ALTE) (2011). *Manual for Language Test Development and Examining: For Use with the CEFR*. Language Policy division, Council of Europe.

[www.coe.int/t/dg4/linguistic/ManualLanguageTest-Alte2011_EN.pdf]

BARRETT, G. R.

[<https://www.amazon.com/Feel-Good-Meditations-Positive-Affirmations/dp/B004YXDIZ8>], consulté en 2016.

[<http://ia310828.us.archive.org/1/items/TheSelfishGiant-ByOscarWilde-NarratedByGrantRaymondBarrett/01SelfishGiant.mp3>], consulté en 2010.

The Cambridge American English Dictionary

[<https://dictionary.cambridge.org/dictionary/american-english/>]

The Cambridge British English Dictionary

[<https://dictionary.cambridge.org/dictionary/english/>]

The Cambridge English Pronouncing Dictionary – 18th Edition (CEPD18) sur CD-ROM (Roach *et al.* 2011).

CHAN, K. L. (2016). Power Language Index. *INSEAD*.

[https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&cad=rja&uact=8&ved=0ahUKEwjy2Yzd_8nVAhUJB8AKHSYvCQcQFggyMAI&url=http%3A%2F%2Fwww.kailchan.ca%2Fwp-content%2Fuploads%2F2016%2F09%2FKC_PLIpresentationatINSEAD_25May2016_v2.pdf&usq=AFQjCNFttsIGmEli32oyvePcX6ebnsZpWQ], consulté en 2017.

CHAPMAN, P. (1999). The CRISP-DM User Guide. *Brussels SIG Meeting*. NCR Systems Engineering Copenhagen.

[<http://lyle.smu.edu/~mhd/8331f03/crisp.pdf>], consulté en 2017.

- COBB, T., *Compleat Lexical Tutor v.8.*
[<http://www.lextutor.ca/>], consulté en 2016.
- Conseil de l'Europe (2000). *Un Cadre européen commun de référence pour les langues : Apprendre, enseigner, évaluer.* Division des Politiques Linguistiques, Strasbourg.
[http://www.coe.int/t/dg4/linguistic/Source/Framework_FR.pdf], consulté en 2011.
- Conseil de l'Europe ([2001] édition restructurée). *Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer.* Unité des Politiques linguistiques, Strasbourg.
[<https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwiz1-XL5dLVAhVIWxQKHap8Dm8QFggoMAA&url=https%3A%2F%2Frm.coe.int%2F16802fc3a8&usg=AFQjCNELdQBWhFlk9Y5vL8S1YfRMMy2D83w>]
- Council of Europe ([2001] restructured version). *Common European Framework of Reference for Languages: Learning, teaching, assessment.* Unité des Politiques linguistiques, Strasbourg.
[https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=5&cad=rja&uact=8&ved=0ahUKEwjW07_x0NLVAhVJ2hoKHel0AP0QFghHMAQ&url=https%3A%2F%2Fwww.coe.int%2Ft%2Fdg4%2Flinguistic%2FSource%2FFramework_EN.pdf&usg=AFQjCNG_66ByIABYGiNsCBiE-QDYJpHG7Q]
- Le Dictionnaire de français Larousse*
[<http://www.larousse.fr/dictionnaires/francais/emphase/28913>], consulté en 2016.
[<http://www.larousse.fr/dictionnaires/francais/exclamation/32013#qKVBYy0E55vcSPRY.99>], consulté en 2015.
- Encyclopaedia Universalis*
[<http://www.universalis.fr/encyclopedie/ivan-fonagy/>], consulté en 2017.
[<http://www.universalis.fr/dictionnaire/geolecte/>], consultation en 2016.
- Programme *English Profile*
[<http://www.englishprofile.org/>], dernièrement consulté en 2017.
- Logiciel *FormatFactory 3.2.1* (Chen Jun Hao, 2007)
[<http://www.pcfreetime.com/formatfactory/index.php?language=en>]
- Le grand dictionnaire terminologique* de l'Office québécois de la langue française
[<http://www.granddictionnaire.com/>]
- HUVER, E. (2014). CECR et évaluation : interprétations plurielles et logiques contradictoires, *Les Cahiers du GEPE.*
[<http://www.cahiersdugepe.fr/index2652.php>]
- KENT, Ch. (British actor, voiceover artist)
[<http://www.christopherkent.co.uk/>], consulté en 2017.

Plateforme *Labguistic*

[<http://www.labguistic.com/modules/fmcontent/content.php?topic=static&id=002&page=presentation>], dernièrement consulté en 2017.

Corpus *LeaP*

[<https://sourceforge.net/projects/leapcorpus/>]

Projet *LONGDALE (Longitudinal Database of Learner English)*

[<http://www.clillac-arp.univ-paris-diderot.fr/projets/longdale>], dernièrement consulté en 2017.

Projet *Longdale-Charliphonia*

[<http://www.clillac-arp.univ-paris-diderot.fr/projets/charliphonia>], dernièrement consulté en 2017.

The Longman Dictionary of Contemporary English

[<https://www.ldoceonline.com/>]

The Longman Pronunciation Dictionary (2008). CD-ROM based on the 18th edition of the *Cambridge English Pronouncing Dictionary*. © Cambridge University Press 2011, Version 1.0.

The Macmillan Dictionary

[<https://www.macmillandictionary.com/>]

The Merriam-Webster's Online Dictionary

[<https://www.merriam-webster.com/>]

Logiciel tableur *Microsoft Excel (Microsoft Office, 2007)*.

The Oxford English Dictionary

[<https://en.oxforddictionaries.com/>]

Le Petit Robert (CD-ROM du *Petit Robert*, version électronique du *Nouveau Petit Robert*, version 2.1. Dictionnaires *Le Robert*, VUEF, 2001).

Logiciel *Praat* (Boersma et Weenik, 1996)

[<http://www.fon.hum.uva.nl/praat/>], dernièrement consulté en 2017.

Script *ProsodyPro - A Praat script for large-scale systematic analysis of continuous prosodic events* (Xu)

[<http://www.homepages.ucl.ac.uk/~uclyyix/ProsodyPro/>], consulté en 2016.

Script *Prosogramme - Pitch contour stylization based on a tonal perception model* (script *Praat*) (Mertens, 2004)

[<https://sites.google.com/site/prosogram/home>]

Alphabet *SAMPA*

[<http://www.phon.ucl.ac.uk/home/sampa/>]

Plateforme *SoSci Survey*

[<https://www.soscisurvey.de>], dernièrement consulté en 2017.

Société des Anglicistes de l'Enseignement Supérieur (SAES)

[<http://saesfrance.org/>], site dernièrement consulté en 2017.

Logiciel *SPPAS* (Bigi, 2012)

[<https://www.ortolang.fr/market/tools/sldr000800/>]

Le Trésor de la Langue Française Informatisé

[<http://atilf.atilf.fr/>]

Logiciel *WaveLab 6.11 - Audio Editing And Mastering Suite* (Steinberg, Goutier, 1995).

Logiciel *WinPitch LTL* (Martin)

[<http://www.winpitch.com/>]

Logiciels *WinPitchW8*, *WinPitchW10* (Martin)

[<http://www.winpitch.com/>]

Site *Word and Phrase . Info*

[<http://www.wordandphrase.info/frequencyList.asp>], dernièrement consulté en 2017.

Annexes

Contenu des annexes :

1. Préambule aux annexes
2. Texte de lecture : *The Selfish Giant*
3. Segmentation des polysyllabes
4. Rappel des 28 variables créées
5. Rappel des codes des expérimentations
6. Correspondance des code des locuteurs
7. Dénomination des expérimentations
8. Expérimentation perceptive PER-LEC-N (Chapitre II)
9. Expérimentation EXP-LEC-N (Chapitre II)
10. Expérimentation perceptive PER-EMPH-N (Chapitre III)
11. Expérimentation perceptive PER-THEM-NN (Chapitre III)
12. Explications du fichier « Mesures_apprenants10.xlsx » (Chapitres IV et V)
13. Expérimentation perceptive PER-LEC-NN (Chapitre V)
14. Scripts R (Chapitre V)
15. Travaux auxquels la thèse a donné lieu

1 Préambule aux annexes

Cette thèse, qui analyse les productions d'apprenants à la fois selon des critères qualitatifs (tests de perception) et des méthodologies quantitatives (telles que l'ACP), a procédé à des mesures manuelles et à des calculs statistiques. Le nombre de figures de la thèse alourdirait considérablement le volume des annexes si tous les fichiers *Excel* qui ont servi à les générer devaient y figurer. J'ai fait le choix de ne donner dans les annexes qu'un seul exemple de fichier commenté qui illustre la manière dont les mesures prises à la main ont été visualisées et interprétées. Il s'agit du fichier qui porte sur l'analyse des variables et qui évalue la possibilité de transformer des séries de mesures en critères interprétatifs (*critical features*). Les procédures transformant les mesures relevées en pondération de la variable (système de malus et de bonus) à partir de la visualisation et des classements ordonnés des

locuteurs et des scores des productions d'apprenants sont décrites dans la thèse et détaillées dans les feuilles du fichier *Mesures_apprenants10.xlsx*. Je ne fais pas figurer le fichier qui a servi de bac à sable pour l'élaboration finale des *criterial features* (dans lequel toutes les mesures donnaient lieu à des représentations des apprenants sous forme d'histogrammes). À l'inverse, figurent intégralement les jeux de données qui ont fait l'objet d'un traitement particulièrement quantitatif au Chapitre V, ainsi que les scripts *R* apparentés. Je n'ai pas inclus les scripts correspondant aux simples visualisations des chapitres précédents.

En toute rigueur, il aurait fallu mettre en regard des mesures manuelles l'ensemble des fichiers sons correspondants. J'ai préféré donner à entendre la partie du corpus qui a fait l'objet du Chapitre V dans le test en ligne. Ce test en ligne figure également.

L'ensemble des fichiers sons serait difficilement mis à disposition de la communauté scientifique, car une partie relève du corpus longitudinal *Longdale*, qui doit donc se plier aux conditions juridiques imposées par les collègues de Louvain porteurs du projet. La consultation en ligne des productions des apprenants (sous forme transcrite) est annoncée, au 22 septembre 2017, comme imminente.

Pour les prises de mesure acoustiques, seules les données relatives à l'emphase ont été générées automatiquement à l'aide de *ProsodyPro* (Xu 2013). L'ensemble des mesures a été pris manuellement pour la détection de F0 sous *Praat* (Boersma et Weenink [1992] 2013) et la plupart du temps vérifié sous *WinPitch* (Martin 2017) pour les parties problématiques.

2 Texte de lecture : *The Selfish Giant*

Il s'agit d'extraits du début du conte d'Oscar Wilde. Ce texte a servi de support à la lecture des enregistrements de natifs anglophones et de non natifs.

From **The SELFISH GIANT**, by **Oscar WILDE**

Every afternoon, as they were coming from school, the children used to go and play in the Giant's garden.

It was a large lovely garden, with soft green grass. Here and there over the grass stood beautiful flowers like stars, and there were twelve peach-trees that in the spring-time broke out into delicate blossoms of pink and pearl, and in the autumn bore rich fruit. The birds sat on the trees and sang so sweetly that the children used to stop their games in order to listen to them. "How happy we are here!" they cried to each other.

One day the Giant came back. He had been to visit his friend the Cornish ogre, and had stayed with him for seven years. After the seven years were over he had said all that he had to say, for his conversation was limited, and he determined to return to his own castle. When he arrived he saw the children playing in the garden.

"What are you doing here?" he cried in a very gruff voice, and the children ran away.

"My own garden is my own garden," said the Giant; "any one can understand that, and I will allow nobody to play in it but myself." So he built a high wall all round it, and put up a notice-board.

*TRESPASSERS
WILL BE
PROSECUTED*

He was a very selfish Giant. The poor children had now nowhere to play.

... Then the Spring came, and all over the country there were little blossoms and little birds. Only in the garden of the Selfish Giant it was still winter. The birds did not care to sing in it as there were no children, and the trees forgot to blossom.

"I cannot understand why the Spring is so late in coming," said the Selfish Giant, as he sat at the window and looked out at his cold white garden; "I hope there will be a change in the weather."

But the Spring never came, nor the Summer.

One morning the Giant was lying awake in bed when he heard some lovely music. It sounded so sweet to his ears that he thought it must be the King's musicians passing by.

"I believe the Spring has come at last," said the Giant; and he jumped out of bed and looked out. He saw a most wonderful sight. Through a little hole in the wall the children had crept in, and they were sitting in the branches of the trees. In every tree that he could see there was a little child. ...

3 Segmentation des polysyllabes

The Selfish Giant — Segmentation des polysyllabes

Mots surlignés de jaune : syllabation différente en fonction du dictionnaire.

Case jaune : choix effectué en segmentation.

Phrase	Mot	Longman Pronunciation Dictionary	Cambridge English Pronouncing Dictionary
1	Every	'ev ri — In very formal style occasionally also 'ev ər i (and in compounds too)	'ev.ri
	Afternoon	,ɑ:ft ə 'nu:z̩ ‹,ɑ:ft ə 'nu:z̩ ‹ ʒ,æft-, ʒ-'nʊn, -ən 'u:z̩ ,æft ər-	,ɑ:f.tə'nu:n, [ʒ] ,æf.tə'-
	Garden	'gɑ:d ən 'gɑ:rd ən	'gɑ:d.ə'n, [ʒ] 'gɑ:r-
2	Lovely	'lʌv li	'lʌv.li
3	Over	!əʊv ə !əʊv ə — but before a vowel sound, !əʊv ər !əʊv ər	'əʊ.və', [ʒ] 'əʊ.və'
	Beautiful	'bjʊ:t əf əl -ɪf-, -ɪ fʊl, -ə- 'bjʊ:t-	'bjʊ:t.i.fəl, -tə-, -fʊl, [ʒ] -tə-
	Flowers	'flaʊəz 'flaʊəz	flaʊəz, 'flaʊ.əz, [ʒ] 'flaʊ.ə'
	Peach-trees		
	Spring-time	'sprɪŋ taɪm (springtime)	'sprɪŋ.taɪm (springtime)
	Delicate	'del ɪk ət -ək-, -ɪt	'del.ɪ.kət, '-ə-, -kɪt
	Blossom	'blɒs əm 'blɑ:s-	'blɒs.əm, [ʒ] 'blɑ:səm
	Autumn	'ɔ:t əm 'ɔ:t əm 'ɑ:t-	'ɔ:t.əm, [ʒ] 'ɑ:təm, 'ɔ:-
4	Sweetly	'swi:t li	swit -li
	Children	'tʃɪldr ən 'tʃɪldr- 'tʃɪld ərn	'tʃɪl.drən
	Order	'ɔ:d ə 'ɔ:rd ə	'ɔ:d.ə', [ʒ] 'ɔ:r.d ə'
	Listen	'lɪs ən	'lɪs.ən
5	Happy	'hæp i	'hæp i
	Other	'ʌð ə 'ʌð ə	'ʌð.ə', [ʒ] -ə'
6	Giant	'dʒaɪənt	dʒaɪ.ənt
8	Conversation	,kɒn və 'seɪʃ ən ,kɑ:n və-	,kɒn.və'seɪ.ʃən, [ʒ] ,kɑ:n.və'-
	Limited	'lɪm ɪt ɪd ʒət-, -əd ət əd	'lɪm -ɪ.tɪd, -nɪs, [ʒ] -ɪ.tɪd, -nɪs
	Determined	dɪ 'tɜ:m ɪn də-, -ən -'tɜ:m-d	dɪ'ts:mɪnd, də-, [ʒ] dɪ'ts:-
9	Arrived	ə 'raɪvd	ə'raɪv -d
	Playing	'pleɪ ɪŋ	pleɪ -ɪŋ
10	Doing	'du:ɪŋ	'du:ɪŋ
	Very	'ver i — Some speakers use a casual weak form vər i	'ver.i
	Away	ə 'weɪ	ə'weɪ

11	Anyone	'en i wʌn	'en.i.wʌn
	Understand	ˌʌnd ə 'stænd -ər-	ˌʌn.də 'stænd, [ʊs] -dər-
	Allow	ə 'laʊ	ə'laʊ
	Nobody	'nəʊb əd 'nəʊ ,bɒd 'nɒb əd 'nɒʊ ,bɑ:d , -bʌd i	nəʊ.bə.d i, -bɒd i, [ʊs] 'nɒʊ.bɑ:d i, -bʌd i, -bə.d i
	Myself	maɪ 'self mɪ-, mə- —In British English the forms mɪ-, mə- are on the whole restricted to very casual or non-standard speech.	maɪ'self, mɪ-, [ʊs] mɑ-, mə-
12	(Around)	ə 'raʊnd	ə'raʊnd
	Notice-board	'nəʊt ɪz §-əs 'nɒʊt əs 'notice board	'nəʊ.tɪs.bɔ:d, [ʊs] 'nɒʊ.tɪs.bɔ:rd
13	Trespassers	'tresp əs §'tres pɑ:s, §-pæs 'tres pæs ə/z ər/z	'tres.pəs, [ʊs] -pæs, -pəs -əz, [ʊs] -əz
	Prosecute / Prosecuted	'prɒs ɪ kju:t -ə- 'pra:s- kju:t ɪd -əd kju:t əd	'prɒs.ɪ.kju:t, '-ə-, [ʊs] 'pra:si- -kju:ɪd, [ʊs] -kju:ɪd
14	Very	'ver i —Some speakers use a casual weak form vər i	'ver.i
	Selfish	'self ɪʃ	'sel.fɪʃ
16	Over	ˌəʊv ə ˌoʊv ə	'əʊ.vər, [ʊs] 'oʊ.vər
	Country	'kʌntri i	'kʌn.trɪ i
	Blossom	'blɒs əm 'blɑ:s-	'blɒs.əm, [ʊs] 'blɑ:səm
	Little	'lɪt əl 'lɪt əl	'lɪt.əl, [ʊs] 'lɪt.əl
17	Only	'əʊn li §-i 'oʊn li (!)	'əʊn.li, [ʊs] 'oʊn-
19	Cannot	'kæn ɒt -ət; §kæ 'nɒt, kə- 'kæn ɑ:t kə 'nɑ:t, kə-	'kæn.ɒt, -ət, [ʊs] 'kæn.ɑt, kə'nɑt
	Understand	ˌʌnd ə 'stænd -ər-	ˌʌn.də 'stænd, [ʊs] -dər-
	Coming	'kʌm ɪŋ	kʌm -ɪŋ
	Weather	'weð ə -ər	'weð .ər, [ʊs] -ər
21	Lovely	'lʌv li	'lʌv.li i
22	Sounded	'saʊnd ɪd -əd	saʊnd -ɪd
	Musician	mju 'zɪʃ ən	mju:'zɪʃ.ən
	Passing	'pɑ:s ɪŋ §'pæs- 'pæs ɪŋ	pɑ:s, [ʊs] pæs -ɪŋ
24	Wonderful	'wʌnd ə fəl -fʊl -ər- —In casual American English also 'wʌn ər-	'wʌn.də .fəl, -fʊl, [ʊs] -dər-
25	Sitting	'sɪt ɪŋ 'sɪt ɪŋ	sɪt -ɪŋ, [ʊs] 'sɪt.ɪŋ
	Branches	'brɑ:nʃ ɪz §'bræntʃ-, -əz 'bræntʃ -əz	brɑ:nʃ, [ʊs] bræntʃ -ɪz
26	Every	'ev ri — In very formal style occasionally also 'ev ər i (and in compounds too)	'ev.ri

4 Rappel des 28 variables créées

Domaine	Nom	Code	Caractéristiques
Temporalité	DUR_T	D1	Durée brute du texte <i>The Selfish Giant</i>
	DUR_ABCp	D2	Durée des 3 segments non ponctués ABC avec pauses (P4, 16 et 25)
	DUR_ABC	D3	Durée des 3 segments non ponctués ABC sans pause (P4, 16 et 25)
	DEB_T	D4	Nombre de syllabes par seconde dans le texte <i>The Selfish Giant</i> avec repère moyen de 541 syllabes
	DEB_ABCp	D5	Nombre de syllabes par seconde dans 3 segments non ponctués ABC avec pauses (P4, 16 et 25)
	DEB_ABC	D6	Nombre de syllabes par seconde dans 3 segments non ponctués ABC sans pause (P4, 16 et 25)
Registre	REG_Excl	Re1	Ampleur du registre de l'exclamative (F0max-F0min)
	REG_P5	Re2	Ampleur du registre de P5 (F0max-F0min)
	REG_so4	Re3	Ampleur du registre de P4 (F0max sur <i>so sweet(ly)</i> -F0min P4)
	REG_so22	Re4	Ampleur du registre de P22 (F0max sur <i>so sweet</i> -F0min P22)
	REG_max	Re5	Ampleur maximale des registres Re1-2-3-4-8
	REG_moy	Re6	Moyenne des mesures de registre de l'emphase (Re1-2-3-4-8)
	REG_S9	Re7	Ampleur du registre de S9 (F0mean sur voyelle à hauteur maximale-F0min S9)
	REG_no	Re8	Ampleur du registre de /n@U/ dans <i>nobody</i> P11(F0max-F0min)
Mélodie	MEL_Excl	M1	Qualité mélodique de l'exclamative
	MEL_P5	M2	Qualité mélodique de la Phrase 5
	MEL_S9	M3	Qualité mélodique du syntagme intonatif S9
	MEL_s\	M4	Qualité du ton montant-descendant sur une syllabe
	MEL_s/	M5	Qualité du ton descendant-montant sur une syllabe
	MEL_no	M6	Qualité et complexité tonale réalisée sur la syllabe /n@U/ de <i>nobody</i> (P11)
Rythme	RYT_lex	Ry1	Respect des patrons accentuels canoniques dans les polysyllabes
	RYT_so4	Ry2	Qualité de proéminence par F0 du marqueur <i>so sweet(ly)</i>
	RYT_so22	Ry3	Qualité de proéminence par F0 du marqueur <i>so sweet</i> par F0
	RYT_all12	Ry4	Qualité de proéminence par F0 du marqueur <i>all round</i> par F0
	RYT_all16	Ry5	Qualité de proéminence par F0 du marqueur <i>all over</i> par F0
	RYT_them	Ry6	Absence de proéminence d'un pronom en fin de syntagme intonatif par test perceptif
	RYT_syll	Ry7	Rapport de durée inter-syllabique sur la paire de syllabes adjacentes <i>sat at</i> (P19)
Rythme et mélodie	RYTMEL_so4	RM	Qualité de proéminence par F0 du marqueur <i>so sweet(ly)</i> et de la hauteur mélodique de (<i>sweet</i>) <i>ly</i>

Tableau 112 – Variables utilisées pour l'établissement des profils des apprenants, regroupées par domaine prosodique (rappel du Tableau 73, page 417)

Le rappel du Tableau 73 ci-dessus détermine les 28 variables utilisées dans le profilage des apprenants. Elles sont regroupées par domaine prosodique. Les codes facilitent une visualisation des données dans les graphiques et figures.

5 Rappel des codes des expérimentations

Cette thèse fait mention de l'expérimentation **REP-PIAN-NN**. Il s'agit d'une opération de repérage des apprenants « pianistes ». Bien que partie intégrante de ce travail de recherche, elle est décrite en détail dans l'article « *Intonational phrasing as a potential indicator for establishing prosodic learner profiles* », publié en 2013 (voir bibliographie).

Le rappel de la Figure récapitulative 231 « Organisation des expérimentations sur les natifs et les non natifs pour déterminer les traits discriminants en prosodie de l'interlangue » aide à situer les expérimentations listées dans le tableau qui lui fait suite.

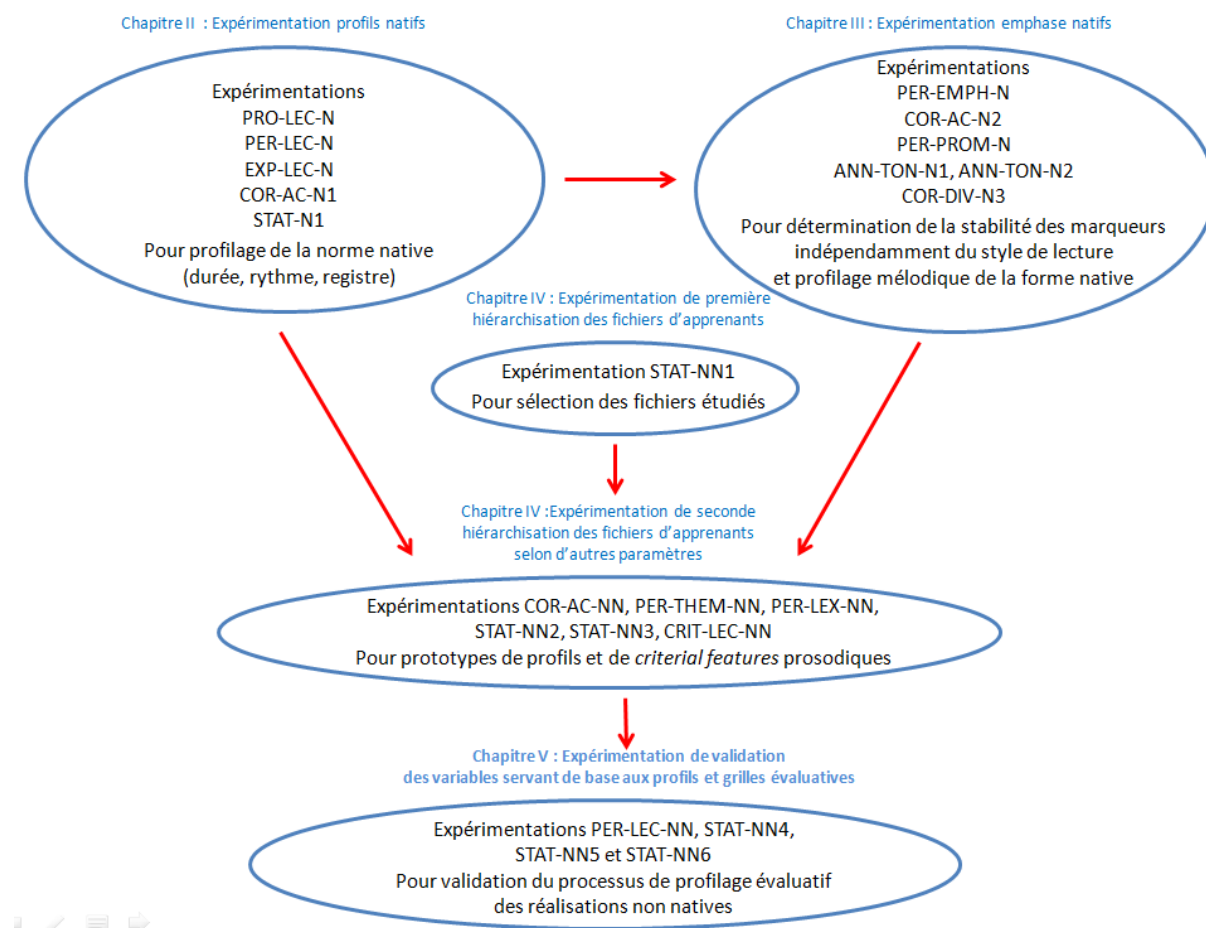


Figure 232 – Organisation des expérimentations sur les natifs et les non natifs pour déterminer les traits discriminants en prosodie de l'interlangue (rappel de la Figure 231, page 535)

<i>Chapitres</i>	<i>Code de l'expérimentation</i>	<i>Indications descriptives</i>
Chapitre II	PRO-LEC-N	Expérimentation portant sur la production de lecture à voix haute par les natifs
	PER-LEC-N	Expérimentation portant sur la perception en lecture chez les natifs
	EXP-LEC-N	Expérimentation portant sur l'expression des natifs enregistrés
	COR-AC-N1	Expérimentation portant sur le premier recueil de corrélats (acoustiques) chez les natifs
	STAT-N1	Première expérimentation par analyse statistique chez les natifs
Chapitre III	PER-EMPH-N	Expérimentation portant sur la perception de l'emphase chez les natifs
	COR-AC-N2	Expérimentation portant sur le deuxième recueil de corrélats (acoustiques) chez les natifs
	PER-PROM-N	Expérimentation portant sur la perception de proéminences chez les natifs
	ANN-TON-N1	Première expérimentation portant sur l'annotation tonale (syllabique) chez les natifs
	ANN-TON-N2	Deuxième expérimentation portant sur l'annotation tonale (syllabique) chez les natifs
	COR-DIV-N3	Expérimentation portant sur le troisième recueil de corrélats (de nature diverse) chez les natifs
Chapitre IV	STAT-NN1	Première expérimentation par analyse statistique chez les non natifs
	COR-AC-NN	Expérimentation portant sur le recueil de corrélats acoustiques chez les non natifs
	PER-THEM-NN	Expérimentation portant sur la perception d'une accentuation du pronom <i>them</i> chez les non natifs
	PER-LEX-NN	Expérimentation portant sur la perception des patrons d'accentuation lexicale réalisés dans les polysyllabes par les non natifs
	STAT-NN2	Deuxième expérimentation par analyse statistique chez les non natifs
	STAT-NN3	Troisième expérimentation par analyse statistique chez les non natifs
	CRIT-LEC-NN	Expérimentation de création de critères distinctifs en lecture chez les non natifs
Chapitre V	PER-LEC-NN	Expérimentation portant sur la perception en lecture chez les non natifs
	STAT-NN4	Quatrième expérimentation par analyse statistique chez les non natifs
	STAT-NN5	Cinquième expérimentation par analyse statistique chez les non natifs
	STAT-NN6	Sixième expérimentation par analyse statistique chez les non natifs

Tableau 113 – Liste des expérimentations de la thèse par chapitre, REP-PIAN-NN exceptée

6 Correspondance des codes des locuteurs

	Âge	Genre	How would you describe your dialect or variety of English?
L1	50+	male	<i>Native language:</i> English ; standard / “BBC”/ Queen’s English - “Received pronunciation” / “Oxford accent” <i>Place where you grew up:</i> South of England
L3	65	male	<i>Native language:</i> English Standard English / R.P. <i>Place where you grew up:</i> All over England, esp Devon & Lincolnshire
L4	31	female	<i>Native language:</i> English <i>Place where you grew up:</i> Rural Northwest Indiana
L5	23	female	<i>Native language:</i> English (American English) <i>Place where you grew up:</i> Colorado, USA
L6	27	female	<i>Native language:</i> RP (South-east middle class) <i>Place where you grew up:</i> Hastings, England
L10	?	male	Grant Raymond Barrett http://ia310828.us.archive.org/1/items/TheSelfishGiant-ByOscarWilde-NarratedByGrantRaymondBarrett/01SelfishGiant.mp3 “Grant holds a B.A. from the University of Michigan with emphasis in German Language, Psychology and Music Performance. He has attended the American Institute For Musical Studies in Graz, Austria and the Flint School Of Performing Arts. He is a Certified Massage Therapist (FSTM 500 Hours) and is excellently trained in Deep-tissue, Swedish, Sports, Shiatsu, Chakra Release and Reflexology.”

Tableau 114 – Caractéristiques principales des locuteurs stéréotypiques

Archétypes stéréotypiques	Code dans l'expérimentation PRO-LEC-N (Chapitre II)	Code dans l'expérimentation PER-LEC-N (Chapitre II)	Code dans l'expérimentation PER-EMPH-N (Chapitre III)	Code dans l'expérimentation CRIT-LEC-NN (Chapitre IV)
« meilleur lecteur »	SG-UP-A1-m-56	L1	Reader A	L1(v1) ou L1v1
	SG-UP-A2-m-56			L1(v2) ou L1v2
✓	SG-UP-M1-m-65	L3		
	SG-PD-He1-f-23	L5		L5
✓	SG-PD-Ka1-f-27	L6	Reader B	
✓	SG-Pro-GRB-m	L10	Reader C	

Tableau 115 – Correspondance des codes des locuteurs natifs participant à plusieurs expérimentations

Profils stéréotypiques	Code dans l'expérimentation PER-LEC-N (Chapitre II)	Catégorisation selon la vitesse de lecture du texte STAT-NN1 (Chapitre IV)	Code dans l'expérimentation PER-LEC-NN (Chapitre V)	Codes Longdale
Archétype stéréotypique	S1	D1	M	DID0210-S001R
Archétype stéréotypique « pianiste »	S2	D1	H	DID0078-S001R
	S3	D1	C	DID0018-S001R
	S4	Q1	L	DID0003-S001R
	S5	Q1	O	DID0024-S001R
Archétype stéréotypique	S6	Q1	G	DID0108-S001R
« pianiste »	S7	Médiane	A	DID0062-S001R
	S8	Médiane	N	DID0135-S001R
	S9	Médiane	J	DID0046-S001R
Archétype stéréotypique	S10	Q3	B	DID0168-S001R
Archétype stéréotypique	S11	Q3	K	DID0074-S001R
	S12	Q3	D	DID0068-S001R
	S13	D9	F	DID0056-S001R
	S14	D9	E	DID0153-S001R
Archétype stéréotypique	S15	D9	I	DID0076-S001R

Tableau 116 – Correspondance des codes des locuteurs non natifs participant à plusieurs expérimentations

7 Dénomination des expérimentations

Les noms utilisés pour déterminer les expérimentations comportent trois, voire quatre parties, séparées par des traits d'union et éventuellement numérotées. La première partie concerne la nature de l'expérimentation (enregistrement de l'expression, test de perception, mesure de corrélats, analyse statistique, création de critères d'évaluation). La deuxième partie est optionnelle ; elle précise la nature ou le support de l'expérimentation : (perception de) la lecture, de l'emphase, du pronom *them*, des patrons d'accentuation lexicale, (corrélats) acoustiques, (création de critères de) lecture. La troisième partie spécifie la population étudiée : les sujets sont natifs ou non natifs. Lorsque plusieurs expérimentations sont similaires quant à leurs caractéristiques définitives, elles sont alors numérotées, ce qui correspond à une éventuelle quatrième partie. Les annexes suivantes exemplifient ce propos.

8 Expérimentation perceptive PER-LEC-N (Chapitre II)

Le questionnaire en ligne se trouve dans le dossier « Quest_PER-LEC-N ».

9 Expérimentation EXP-LEC-N (Chapitre II)

Le questionnaire papier à compléter se trouve ci-dessous et ses résultats figurent dans le fichier *Excel* « Données EXP-LEC-N.xls ».

Questionnaire for Native English Speakers

Identification number: Age of participant: Gender identity: male female

How would you describe your dialect or variety of English?

Native language:

Place where you grew up:

While reading the text, who did you imagine yourself telling the story to?

Did you have previous knowledge of the text/the story? yes no

This questionnaire is meant to assess your previous speaking experiences.
Please circle the number that best describes your situation.

- | | Never | Often |
|---|-------|-------|
| 1. Did you receive any formal training in the art of public speaking? | 1 | 4 |
| Comments: | | |
| 2. Have you ever read fairy tales or other stories to children? | 1 | 4 |
| Comments: | | |
| 3. Have you ever read poems or recitation pieces out loud? | 1 | 4 |
| Comments: | | |
| 4. Have you had any experience with addressing a public audience? | 1 | 4 |
| Comments: | | |
| 5. Have you ever participated in drama or staging of shows? | 1 | 4 |
| Comments: | | |
| 6. Have you ever had any teaching experience? | 1 | 4 |
| Comments: | | |
| 7. Have you had any singing experience? | 1 | 4 |
| Comments: | | |

- | | Like | Dislike |
|---|------|---------|
| 8. What are your feelings about speaking to small groups in your native language? | 1 | 4 |
| 9. What are your feelings about speaking to a large audience in your native language? | 1 | 4 |
| 10. What are your feelings about speaking to any audience in a foreign language? | 1 | 4 |

Comments on speaking in public:

.....

10 Expérimentation perceptive PER-EMPH-N (Chapitre III)

Le questionnaire en ligne se trouve dans le dossier « Quest_PER-EMPH-N »

11 Expérimentation perceptive PER-THEM-NN (Chapitre III)

Le questionnaire en ligne se trouve dans le dossier « Quest_PER-THEM-NN »

12 Explications du fichier « Mesures_apprenants10.xlsx » (Chapitres IV et V)

Le fichier *Excel* « Mesures_apprenants10.xlsx » figure dans le dossier joint. Il est représentatif de ma méthode de travail (et de la manière dont certaines figures ont été générées). Il constitue la première étape des *crierial features* décrits dans le Chapitre IV et montre comment je suis partie des mesures brutes pour premièrement visualiser toute une série de réalisations de profils et deuxièmement pour constituer les classes que j'ai retenues dans mon analyse. Dans un certain nombre de cas, les mesures brutes sont ultérieurement interprétées en classes qualitatives au vu des histogrammes correspondants. Cette annexe décrit ce processus de classification subjectif à partir des mesures effectuées. Il est à noter que je n'ai pas utilisé de classification automatique (*clustering* type *k-means*), mais les mesures figurent et il serait possible, au nom de la reproductibilité, de comparer les résultats en classes obtenues avec du *clustering* et celles qui figurent dans la thèse. Je détaille le fonctionnement du fichier à partir de la caractérisation de ses différentes feuilles accessibles par onglets. Les feuilles sont consacrées à l'étude des différents domaines sur les différentes phrases (initiale P des noms des fichiers) dans l'ordre chronologique d'apparition dans le texte. Cohérente par rapport à ma démarche de m'appuyer sur la subjectivité du chercheur au lieu de m'efforcer de l'occulter, je n'ai aucunement « toiletté » le fichier « Mesures_apprenants10.xlsx » dans son cheminement vers l'objectif assigné, ni ses méandres tentateurs.

Feuille SS_S001R

Cette feuille de calcul contient les fichiers d'apprenants *Charliphonia* de lecture du texte *The Selfish Giant* S001R sélectionnés à l'issue de la première phase de sélection selon leur vitesse de lecture du texte. La feuille sert de suivi du traitement des données de départ ; au départ, l'analyse devait porter sur huit et non pas sur trois locuteurs et devait inclure quatre

fichiers « hors classe » longitudinaux aux extrêmes des valeurs observées, les autres concernant les 1^{er} et 9^{ème} déciles, les 1^{er} et 3^{ème} quartiles et la médiane.

Feuille Dur-vit

Durée et vitesse d'élocution : cette feuille récapitule les mesures de durée (en secondes) pour les différents passages analysés (tronçons). L'ensemble des locuteurs figure dans les lignes qui occupent les trois premières colonnes, les durées des tronçons sont indiquées dans les colonnes.

Pour chaque tronçon, on commence par indiquer les bornes de début et de fin à partir des mots repères (par exemple *that* et *them* pour le tronçon A), puis on mesure le temps de lecture de chaque tronçon et les durées des pauses le cas échéant. Seules les pauses supérieures à 0,1 seconde ont été consignées. En pratique, une pause au maximum a été observée pour ces segments en principe sans pause. Il a alors été possible de calculer le nombre de syllabes par seconde, avec ou sans pause. Des histogrammes permettent de visualiser les mesures effectuées. La dernière colonne récapitule la moyenne des tronçons.

Feuille Nb syll L1&L5

Cette feuille récapitule le compte des syllabes réalisées par les locuteurs natifs servant de modèles aux réalisations des apprenants. Les deux enregistrements du locuteur L1 ont été analysés et figurent des explications qualitatives relatives à certains résultats. Dans deux enregistrements sur trois, tous les mots du texte n'ont pas été lus, ce qui donne lieu à la pondération proposée.

Feuille TronçonsN

Cette feuille récapitule les durées des tronçons pour les natifs, ce qui a été mesuré pour les non natifs dans la feuille numéro 2 (Dur-vit). La locutrice la plus lente a servi de repère de durée inférieur pour les résultats des apprenants. Cette mesure est fondamentale pour indiquer une limite normée inférieure de l'acceptable en termes évaluatifs. La grande vitesse de lecture du « meilleur lecteur » L1 parmi les natifs le disqualifie de pour indiquer une limite normée supérieure de l'acceptable : les résultats de L5 ont été jugés plus adaptés et amplement suffisants.

Feuille P4_SO

La feuille P4_SO mesure et analyse les réalisations du segment contenant *so sweetly* dans la Phrase 4. Le premier tableau consigne les mesures de la F0 en Hertz, le second tableau transcrit ces mesures en demi-tons dans un but comparatif avec la formule ($=12*\text{LOG}(X/100)/\text{LOG}(2)$). Quelques observations complètent les mesures, telles que les algorithmes de détection de *WinPitch* utilisés pour les mesures de F0. À partir des mesures converties en demi-tons de F0 de plusieurs syllabes déterminantes, on a établi des traits critériés fondés sur les différentiels maximum observés.

Les colonnes AP à AS servent à calculer les pondérations en bonus-malus expliquées dans la thèse. Leur somme permet de hiérarchiser les résultats en termes quantitatifs et qualitatifs après catégorisation. Les variables rythmique et mélodique-rythmique sont établies aux colonnes AT et AV sous leur forme numérique desquelles sont dérivées les colonnes suivantes, AU et AW, fournissant la catégorisation qualitative après inspection visuelle des histogrammes (voir BO à BT).

Des essais de visualisation de seuils de niveau de performance font intervenir des camaïeux de couleurs pour la colonne colorée BH et des représentations sous forme d'histogrammes (colonnes BK à BL, ligne 61 à 70). Aux lignes 76 à 88, colonnes AM à AW, un essai d'histogramme à quatre colonnes permet de visualiser la variable RYT_so4 et les trois séries de mesures qui la constituent.

Feuille P4_SO graph1

Cette feuille reprend les données de la précédente et permet de tester la visualisation de la variable RYT_so4 et les séries de mesures qui la constituent limitées à deux.

Feuille P4_SO graph2

Cette feuille reprend les données des précédentes et a servi à générer la Figure 197 de la thèse, qui permet de visualiser les différentes valeurs observables pour les différentes syllabes du tronçon considéré avec les deux variables RYT_so4 et RYTMEL_so4 (Détail des données constitutives des variables RYT_so4 et RYTMEL_so4 par catégorie d'apprenants selon la vitesse de lecture (barème qualitatif)).

Feuille P4_THEM

Cette feuille est laissée à titre d'illustration des limitations rencontrées dans l'analyse acoustique. Dans une démarche exploratoire, les mesures de F0 prises sur *listen to them*, en

Hertz sous *Praat*, puis *WinPitch* si besoin (mesures en vert) on fait l'objet de tentatives de catégorisation de schémas intonatifs (chute sur *them*). Cette piste a été abandonnée car les valeurs numériques représentaient ne correspondaient pas à ma perception des réalisations. Cette contradiction a été résolue par l'adoption d'une autre méthode. Les colonnes O à S récapitulent les perceptions : par des sujets anglophones de P à S et les miennes (colonne O). Les pourcentages de perception (colonne U) ont été transformés en un critère RYT_THEM après inspection de l'histogramme représenté de G à L (ligne 77 à 89).

Feuille P5_EXCL

Cette feuille recueille les analyses de l'énoncé *How happy we are here*. Les colonnes R à X reprennent en demi-tons les mesures effectuées en Hertz aux colonnes D à J. Plusieurs observations sont consignées à la colonne M, la colonne K précise la syllabe ayant servi pour la mesure la plus basse de la F0 de l'énoncé citant.

Les couleurs utilisées aux colonnes R à X ont servi à visualiser les valeurs maximales et minimales et surtout à localiser les syllabes à la F0 la plus et la moins élevée sur l'exclamative et la Phrase 5 entière. Les colonnes AC et AE établissent les soustractions des valeurs maximales de la F0 récapitulées dans les colonnes Z à AB. La colonne AH est le vestige d'une tentative pour faire de la réalisation de *How happy* un trait critérié en distinguant des réalisations de tons montants ou descendants. Cette catégorisation binaire (montant-descendant) a été abandonnée car l'examen des réalisations des natifs a montré la coexistence des deux types de réalisations, plutôt en fonction de l'âge des natifs que de la qualité de leur lecture. Les colonnes AJ à AI récapitulent les différentiels de réalisations entre deux syllabes déterminantes de l'énoncé, envisageant trois couples de syllabes. AO reprend les valeurs de registre sous forme qualitative à partir des données chiffrées reportées aux lignes 78 à 92 (colonnes B à H) et de la visualisation sous forme d'histogramme aux lignes 79 à 92 (colonnes B à H). Les colonnes AP et AQ établissent les registres à partir des valeurs AK et AL, corrigées par un système de pénalisation des réalisations jugées incorrectes.

Les colonnes AU à BA permettent simplement de visualiser les réalisations des syllabes en demi-tons (colonnes R à X).

Feuille P9_aRRIVED

La feuille P9_aRRIVED consigne les mesures effectuées sur le mot *arrived* afin de mettre en évidence les mouvements mélodiques sur la syllabe accentuée. L'analyse a

débouché sur la proposition de la variable MEL_s \checkmark . Elle cherche à établir l'aptitude des apprenants à réaliser le ton descendant-montant sur une seule syllabe. En substance, elle est l'addition (colonnes AA) des valeurs montante et descendante, obtenues par pondération d'un barème (colonnes X et Y) suggéré par la visualisation des histogrammes. D'abord en Hertz (colonnes F à I), puis calculés en demi-tons (Q à T, récapitulées dans le graphique des lignes 73 à 92, qui est la Figure 130 de la thèse), ces mesures de F0 aboutissent à une caractérisation tonale tendancielle sommaire (colonnes U et V), qui est précisée par des mesures de différentiels entre fin et début de syllabe (colonnes W et X). À partir de là, des formules de calcul établissent un barème (Y et Z) dont la somme sert de base à la variable tonale syllabique \checkmark . Les graphiques des colonnes AD à AI offrent les contours moyens pour chaque catégorie d'apprenants d'après la vitesse de lecture.

Je précise ici comment a été établie la formule de calcul du barème, transformant les réalisations en demi-tons en points du barème. La visualisation par histogrammes des parties descendante (colonnes AB à AI, lignes 94 à 115) et montante (colonnes I à Q, lignes 94 à 109) de la variable pour chaque apprenant se trouve au bas de la feuille. Les valeurs numériques additionnées figurent colonne Y lignes 94 à 108. Le lissage en valeurs de 0 à 6 a été obtenu colonne Z (lignes 94 à 108). Les valeurs en Y ont été classées par ordre décroissant, l'hypothèse indirecte étant que le meilleur score devait être attribué au différentiel le plus important observé sur la somme de la partie descendante et de la partie montante. La formule ainsi codée dans la colonne Y est la résultante des visualisations des histogrammes : elle permet d'établir les seuillages de niveau entre des différences de valeur qui me paraissaient significativement différentes (du moins visuellement). L'échelle de points du barème ainsi obtenu n'est pas linéaire.

Le Graphique des lignes 129 à 140 (Figure 184 de la thèse) récapitule, pour chaque sous-classe de lecteur considérée en fonction de la rapidité de lecture, les valeurs moyennes obtenues en fonction du barème ainsi créé.

Feuille P9_aRRIVED L1&L5

La feuille P9_aRRIVED L1&L5 reproduit la même démarche mais avec les trois fichiers des natifs.

Feuille P9_fall (2)

Cette feuille sert à établir les variables de registre et de mélodie pour le segment S9 (REG_S9 et MEL_S9). Pour la variable de registre, un différentiel est établi à partir des valeurs maximales de F0 observées par F0mean, auxquelles on a retranché les valeurs de F0min. Pour la variable de mélodie MEL_S9, c'est la somme (colonne BC) de la variable de registre et de divers différentiels entre deux valeurs de syllabes en demi-tons, pondérées selon le même principe que précédemment. C'est la conformité aux réalisations attendues (montée sur la syllabe de *saw*) qui guide le système de pénalisation reporté pour les réalisations des syllabes jugées les plus déterminantes dans le segment S9. Le paramétrage est effectué aux colonnes AX à BB à partir des mesures en Hertz de F0 mean (colonnes D à N), après conversion en demi-tons (colonnes AS à AW).

Feuille P11_NOBODY

Cette feuille permet d'obtenir les variables REG_no de registre sur la première syllabe de *nobody* et MEL_no de mélodie tonale sur la même syllabe. REG_no est le différentiel des valeurs maximales et minimales de la F0 en demi-tons (récapitulatif colonnes T, U et V). Après des essais peu concluants (colonnes D à I) de prise de mesures en trois points, je me suis orientée vers une prise de mesures en quatre points (colonnes K, L, M et N). Le calcul en demi-tons du mot entier se trouve dans les colonnes AH à AM et permet d'obtenir le graphique des contours sur le mot entier aux lignes 76 à 88. Les mesures en demi-tons de la syllabe /n@U/ (colonnes AQ à AT) permettent de dégager un essai de patrons tonaux (colonne AU), une représentation graphique (lignes 74 à 88), et des différentiels préliminaires à la constitution d'un barème évaluatif pour la variable mélodique MEL_no (colonnes BF à BJ).

Feuille NO

Cette feuille est destinée à obtenir l'histogramme (colonnes J à O, lignes 3 à 13) de la représentation de la syllabe /n@U/ suivant le pourcentage d'occurrence de chaque contour tonal. Quatre catégories tonales ont été dégagées.

Feuille P12_ALL round

Cette feuille se focalise sur la variable rythmique RYT_all12. Une prise de mesures en Hertz sur les syllabes les plus significatives du segment *high wall all around it* (colonnes D à H) fournit des données en demi-tons (colonnes S à W). La variable résultante (colonne Y) est le différentiel en demi-tons des syllabes *wall* et *all*. De même que précédemment, les divers

graphiques permettent de visualiser les courbes mélodiques du segment par catégorie de vitesse de lecture (colonnes AB à AH, et aussi S à AB aux lignes 81 à 93), de figurer les moyennes des groupes (colonnes S à AB, lignes 81 à 92), et de procéder à l'élaboration d'une catégorisation qualitative de la variable.

Feuille P16_ALL over

La feuille P16_ALL over suit le même principe que la précédente pour créer la variable RYT_all16.

Feuille P19_SAT AT

Cette feuille conduit à l'obtention de la variable RYT_syll. La colonne F calcule le différentiel entre les mesures de durée des syllabes /s{t/ et /{t/ à partir de leurs mesures en seconde. La colonne G en établit la correspondance qualitative grâce à l'histogramme (colonnes P à V). Les histogrammes (colonnes 73 à 100) fournissent des représentations visuelles des données.

Feuille P22_SO

La feuille P22_SO aboutit à l'obtention des variables de registre REG_so22 et de rythme RYT_so22. Les mesures en Hertz du tronçon *It sounded so sweet to his ears* (colonnes D à K) sont calculées en demi-tons (colonnes Z à AG). La variable de registre provient du différentiel entre la F0max sur *so sweet* et la mesure de F0min sur la totalité de la Phrase 22 (colonnes AI à AK). La variable RYT_so22 (colonne AX) provient de la somme des deux notations (colonnes AV et AW) basées sur les différentiels entre la F0max sur *so* et la syllabe qui la précède d'une part et entre la F0max sur *so* et *sweet* d'autre part (colonnes AR et AS). Les graphiques en lignes (colonnes BA à BI, ainsi que Z à AI aux lignes 77 à 89) montrent les contours de F0 sur l'ensemble du segment. L'histogramme (colonnes AR à AY, lignes 74 à 86) récapitule l'étude par la notation évaluative selon le barème adopté. L'histogramme (colonnes BL à BP) permet de guider les choix de seuils de niveaux qualitatifs.

Feuille SO

Cette feuille est un essai de regroupement des variables dans lesquelles le marqueur *so* apparaît dans une optique comparative. Il n'a pas abouti.

Feuille PlacLex

La feuille PlacLex conduit à l'obtention de la variable du respect de l'accentuation lexicale RYT_syll. Elle consigne les résultats de l'expérimentation PER-LEX-NN. Les lignes 4 à 94-97 présentent les six schémas accentuels par apprenants et par catégorie selon la vitesse de lecture (colonnes A à D). Les résultats des deux spécialistes francophones (colonnes E et F) sont analysés (colonnes G). Dans le cas où ils sont peu concluants (« à vérifier »), ils sont alors confrontés au bilan des résultats des anglophones natifs (colonnes H à J) dans la colonne K, fournissant la décision de réussite (« juste ») ou d'échec (« faux »). La justification de certaines décisions est parfois consignée dans la colonne M pour une éventuelle utilisation future. Les colonnes N à V rendent compte du nombre de schémas justes par apprenant et par polysyllabe testé (lignes 95 à 97). Les trois histogrammes au bas de la feuille fournissent un bilan graphique des mots testés (colonnes Y à AC ; lignes 83 à 94), du respect du schéma accentuel par apprenant (lignes 98 à 111), et du respect du schéma accentuel par catégorie d'apprenant (lignes 113 à 125).

Feuille V_TEM1

La feuille V_TEM1 récapitule les données par apprenant et par catégories selon la vitesse de lecture des six variables temporelles calculées et répertoriées dans la feuille Durvit. Elle est la phase de préparation à la version finale figurant dans la feuille suivante : V_TEM2.

Feuille V_TEM2

La feuille V_TEM2 présente les six variables temporelles quantitatives et leur corolaire qualitatif. Pour DUR_T, les extrêmes D1 et D9 représentent respectivement les bonifications et pénalisations accordées. Les histogrammes au bas de la feuille montrent la méthodologie décisionnelle utilisée pour déterminer les seuils dans les cinq autres variables.

Feuille V_REG1, Feuille V_REG2, Feuille V_MEL1, Feuille V_MEL2, Feuille V_RYT1, Feuille V_RYT2, Feuille V_RYT3

Ces feuilles suivent la même logique dans la création des variables de registre, de mélodie et de rythme, à la seule différence que l'élaboration des variables rythmiques s'effectue sur trois feuilles étant donné leur complexité. Alors que V_RYT2 tente d'amalgamer les deux variables basées sur *all* pour les fondre en une seule (essai abandonné, d'où l'effacement des références), V_RYT3 les dissocie.

Feuille V_TOUS

La feuille V_TOUS fournit le tableau comprenant toutes les variables quantitatives et qualitatives mises en œuvre au Chapitre IV (colonnes C à BF). La dénomination des apprenants figure en colonne A dans sa version *Longdale* et en colonne B pour sa correspondance dans l'expérimentation PER-LEC-NN. La ligne 1 mentionne le nom de chaque variable dans sa version quantitative (rose soutenu) et sa version qualitative (rose plus clair). La ligne 2 en synthétise les appellations pour une visibilité accrue dans les figures (plans factoriels et cercles de corrélations). La ligne 19 comptabilise le nombre de colonnes figurant dans le tableau servant à l'analyse statistique et les couleurs utilisées signalent leur domaine prosodique (RYTMEL_so est à en partie mélodique et en partie rythmique).

Feuille varSS

La feuille varSS reproduit la feuille précédente pour n'en garder que les éléments nécessaires à la création du fichier de données de variables d'étudiants (varSS.txt) analysé statistiquement par R. La ligne 18 sert de repère et n'est pas destinée à être reproduite dans le fichier de données varSS.txt.

Feuille V_PROFILS

La feuille V_PROFILS permet d'établir les profils d'apprenants par domaine prosodique à partir du tableau de la feuille V_TOUS ; seule la variable RYTMEL_so (colonnes BI et BJ) n'est pas prise en compte étant donné sa double appartenance. Pour chaque domaine, les points de bonus et de malus sont comptabilisés puis transcrits en pourcentages (colonnes O à R pour la durée, AI à AL pour le registre, AY à BB pour la mélodie et BS à BV pour le rythme). Les colonnes BX à CP récapitulent ces résultats et en dressent le bilan par moyenne du taux de bonification et de pénalisation (colonnes CR et CS). Les profils obtenus figurent sous la forme de trois histogrammes : un profil global à partir des données non ordonnées des colonnes CR et CS (colonnes CU à CY ; lignes 2 à 13), un classement par taux d'échec (colonnes DG à DM ; lignes 2 à 15), un classement par taux de réussite (colonnes DI à DM ; lignes 18 à 30). Un histogramme détaillant ces résultats par domaine pour chaque apprenant figure dans les colonnes CU à DC, lignes 15 à 34.

Feuille V_PROFILS-3

La feuille V_PROFILS-3 reprend la précédente en écartant trois variables de nature à fausser les résultats des profils : les variables RYTMEL_so, ainsi que REG_max et REG_moy, qui ne sont que le fruit de calculs à partir de variables figurant déjà dans le tableau.

Feuille V_PROFILS-3 (2)

La feuille V_PROFILS-3 (2) reprend la précédente en y ajoutant un bilan des deux séries de résultats (colonne CO). Elle hiérarchise les quinze apprenants selon les taux de réussite (colonne CM) et d'erreur (colonne CN), puis par bilan global (colonne CO) : les résultats obtenus servent à établir des essais de catégorisation hiérarchique en fonction des résultats (lignes 20 à 47 ; colonnes AT à BE, BI à BV, CC à CP).

Feuille V_PROFILS-3 PER-LEC-NN2

La feuille V_PROFILS-3 PER-LEC-NN2 est destinée à profiler les résultats précédents afin qu'ils soient compatibles avec ceux de l'expérimentation perceptive PER-LEC-NN2. Pour cela, les variables sont réduites de 25 à 18 pour correspondre aux éléments que les auditeurs entendront. La colonne CF calcule la nouvelle moyenne à partir des quatre domaines. Les lignes 20 à 67 établissent des comparaisons entre les divers résultats.

Feuille V_PROFILS-3 (3)

La feuille V_PROFILS-3 (3) reprend la feuille V_PROFILS-3 (2) pour établir un bilan général et par domaine prosodique des quinze apprenants selon le système de bonus-malus sur 25 variables. À cet effet figure une hiérarchisation des résultats par durée (lignes 22 à 37), par bilan général (lignes 40 à 55 ou lignes 59 à 74). L'histogramme correspondant (colonnes J à W) constitue la Figure 232 de la thèse. La feuille fournit aussi un bilan global par moyenne des quatre domaines prosodiques (colonnes CW à DA).

Feuille V_PROFILS-3 PER-LEC-NN2 (2)

La feuille V_PROFILS-3 PER-LEC-NN2 (2) reprend la feuille V_PROFILS-3 PER-LEC-NN2 pour établir un bilan général et par domaine prosodique des quinze apprenants selon le système de bonus-malus sur les 18 variables testées. Le tableau (colonnes A à H ; lignes 22 à 38) reproduit un bilan global et par domaine des quinze apprenants. La version suivante (colonnes A à H ; lignes 42 à 58) ordonne hiérarchiquement les données du bilan global (colonne D) afin d'en extraire l'histogramme (bilan général et par domaine prosodique

des quinze apprenants selon le système de bonus-malus sur 18 variables : Figure 231 de la thèse). Le tableau dupliqué aux lignes 62 à 78 était destiné à une investigation plus poussée, qui ne voit pas le jour dans cette thèse.

13 Expérimentation perceptive PER-LEC-NN (Chapitre V)

13.1 PER-LEC-NN : Sélection des parties de texte pour évaluation par perception

Surlignage jaune : parties sur lesquelles portent les variables conservées

Surlignage gris anthracite : parties à élaguer

Surlignage gris clair : parties à élaguer sur lesquelles porte un marqueur de variable

13.1.1 Phase 1

From The SELFISH GIANT, by Oscar WILDE

1 Every afternoon, as they were coming from school, the children used to go and play in the Giant's garden.

2 It was a large lovely garden, with soft green grass. 3 Here and there over the grass stood beautiful flowers like stars, and there were twelve peach-trees that in the spring-time broke out into delicate blossoms of pink and pearl, and in the autumn bore rich fruit. 4 The birds sat on the trees and sang so sweetly that the children used to stop their games in order to listen to them. 5 "How happy we are here!" they cried to each other.

6 One day the Giant came back. 7 He had been to visit his friend the Cornish ogre, and had stayed with him for seven years. 8 After the seven years were over he had said all that he had to say, for his conversation was limited, and he determined to return to his own castle. 9 When he arrived he saw the children playing in the garden.

10 "What are you doing here?" he cried in a very gruff voice, and the children ran away.

11 "My own garden is my own garden," said the Giant; "any one can understand that, and I will allow nobody to play in it but myself." 12 So he built a high wall all round it, and put up a notice-board.

13 **TRESPASSERS
WILL BE
PROSECUTED**

14 He was a very selfish Giant. 15 The poor children had now nowhere to play.

16 ... Then the Spring came, and all over the country there were little blossoms and little birds. 17 Only in the garden of the Selfish Giant it was still winter. 18 The birds did not care to sing in it as there were no children, and the trees forgot to blossom.

19 "I cannot understand why the Spring is so late in coming," said the Selfish Giant, as he sat at the window and looked out at his cold white garden; "I hope there will be a change in the weather."

20 But the Spring never came, nor the Summer.

21 One morning the Giant was lying awake in bed when he heard some lovely music.

22 It sounded so sweet to his ears that he thought it must be the King's musicians passing by.

23 "I believe the Spring has come at last," said the Giant; and he jumped out of bed and looked out. 24 He saw a most wonderful sight. 25 Through a little hole in the wall the children had crept in, and they were sitting in the branches of the trees. 26 In every tree that he could see there was a little child. ...

Résultat de la Phase 1: Version retenue pour test pré-expérimental

Every afternoon, the children used to go and play in the Giant's garden. The birds sat on the trees and sang so sweetly that the children used to stop their games in order to listen to them. "How happy we are here!" they cried to each other.

One day the Giant came back. When he arrived he saw the children playing in the garden. "My own garden is my own garden," said the Giant; "and I will allow nobody to play in it but myself." So he built a high wall all round it, and put up a notice-board.

**TRESPASSERS
WILL BE
PROSECUTED**

... Then the Spring came, and all over the country there were little blossoms and little birds. It sounded so sweet to his ears that he thought it must be the King's musicians passing by. Through a little hole in the wall the children had crept in, and they were sitting in the branches of the trees.

13.1.2 Phase 2

Every afternoon, the children used to go and play in the Giant's garden. The birds sat on the trees and sang so sweetly that the children used to stop their games in order to listen to them. "How happy we are here!" they cried to each other.

One day the Giant came back. When he arrived he saw the children playing in the garden. "My own garden is my own garden," said the Giant; "and I will allow nobody to play in it but myself." So he built a high wall all round it, and put up a notice-board.

**TRESPASSERS
WILL BE
PROSECUTED**

... Then the Spring came, and all over the country there were little blossoms and little birds. It sounded so sweet to his ears that he thought it must be the King's musicians passing by. Through a little hole in the wall the children had crept in, and they were sitting in the branches of the trees.

Résultat de la Phase 2: Version retenue pour l'expérimentation PER-LEC-NN

The birds sat on the trees and sang so sweetly that the children used to stop their games in order to listen to them. "How happy we are here!" they cried to each other.

So he built a high wall all round it, and put up a notice-board.

... Then the Spring came, and all over the country there were little blossoms and little birds. It sounded so sweet to his ears that he thought it must be the King's musicians passing by.

13.2 Questionnaire en ligne de l'expérimentation PER-LEC-NN

Voir dossier « Quest_PER-LEC-NN ».

13.3 Fichier Excel d'analyse des données de l'expérimentation PER-LEC-NN

Voir fichier « Données PER-LEC-NN7.xlsx ».

14 Scripts R (Chapitre V)

Voir le dossier « Scripts & données R ».

15 Annexe des travaux auxquels la thèse a donné lieu

15.1 Communications

15.1.1 Colloques de phonétique et phonologie anglaises

- 2017 : Communication intitulée « Médiation, remédiation et évaluation prosodique de l'anglais chez les apprenants francophones à partir de corpus de lecture » présentée aux Journées d'étude *ALOES 2017* « La Phonologie de l'anglais : théorie et terrain : la langue orale de la formation à la pratique pédagogique en classe » (31 mars et 1^{er} avril à Limoges).

- 2016 : Communication présentée au XVIII^{ème} colloque d'anglais oral *ALOES 2016* (1-2 avril à Villetaneuse) sur les experts : caractérisation des locuteurs, des variétés et des styles de parole : « Profils et emphase : nouveaux modes d'organisation prosodique pour catégoriser et hiérarchiser les locuteurs et leurs réalisations acoustiques ».
- 2012 : Communication présentée au colloque *PAC 2012 (Phonologie de l'Anglais Contemporain, 29 fév.-2 mars à Toulouse)* sur les usages, la variation et la structure: « Emphasis and variation in read – or prepared – speech ».
- 2010 : Communication présentée au colloque *PAC 2010 (Phonologie de l'Anglais Contemporain, 12-13 septembre à Montpellier)* sur les usages, la variation et la structure : « Cross-dialect and inter-speaker variation in read speech prosodic typology: EFL compared with other English dialects ».
- 2010 : Communication présentée au 15^{ème} colloque de Villetaneuse sur l'anglais oral : « Interlanguage study : L2 learners' reading performances and phonological competence in prosody ».
- 2010 : Communication présentée à *ALOES 2010 Pre-conference workshop on learner corpora (25 Mars 2010)*, conjointement avec Nicolas Ballier : « Investigating suprasegmental features in the *Charles V* learner corpus ».

15.1.2 Colloques de linguistique

- 2012 : Communication présentée au *Colloque International Jeunes Chercheurs en Linguistique (CIJC 2012)* sur « Les Classifications en Linguistique: Problèmes, Méthodologies, Enjeux », 6-8 juin à Strasbourg: « Defining criterial features for L2 prosody of English ».
- 2011 : Communication présentée au *Colloque sur « L'interlocution comme paramètre : nouvelles données/nouveaux modèles »* (Amiens, 6-7 janvier 2011) : « Interlocution et variation prosodique dans la lecture à haute voix : de l'influence de la subjectivité des lecteurs sur celle des auditeurs ».

15.1.3 Colloques de didactique et acquisition des langues

- 2012 : Communication présentée au colloque *RANACLES 2012* sur « Les vingt ans de *RANACLES* : quel bilan et quel avenir pour les Centres de Langues ? » (*Université de Reims Champagne Ardenne, 29-30 novembre – 1^{er} décembre à Reims*) : « Plaidoyer pour une politique linguistique *LANSAD* intégrant la prononciation ».
- 2011 : Communication présentée au colloque sur « 20 years of learner corpus research: Looking back, moving ahead » (*Université Catholique de Louvain en Belgique, 15-17 septembre 2011 à Louvain-la-Neuve*) : « Intonational phrasing as a potential indicator for establishing prosodic learner profiles ».

15.2 Publications d'articles

15.2.1 Publication en linguistique, phonétique et phonologie

- « Interlocution et variation prosodique dans la lecture à haute voix : de l'influence de la subjectivité des lecteurs sur celle des auditeurs ». Dans Douay Catherine, Roulland Daniel (dir.) (2012) *L'interlocution comme paramètre*. Collection Rivages Linguistiques. Rennes : Presses Universitaires de Rennes, 279-302.

15.2.2 Publications en didactique et acquisition des langues

- « Plaidoyer pour une politique linguistique *LANSAD* intégrant la phonologie anglaise ». Dans Daval René, Hilgert Émilie, Nicklas Thomas, Thomières Daniel (dir.) (2014), *Sens, formes, langage : Contributions en l'honneur de Pierre Frath*. Reims : Éditions et Presses universitaires de Reims, 371-402. ISBN : 978-2-915271-83-6.
- « Intonational phrasing as a potential indicator for establishing prosodic learner profiles ». Dans Granger Sylviane, Gilquin Gaëtanelle & Meunier Fanny (dir.) (2013) *Twenty Years of Learner Corpus Research: Looking back, Moving ahead*. Corpora and Language in Use – Proceedings 1, Louvain-la-Neuve: UCL Presses universitaires de Louvain, 75-88. ISBN: 978-2-87558-199-0.
- « Le *CECRL* et l'évaluation de la compétence phonologique : un point de vue de corpus sur l'anglais ». *Les Langues Modernes*, n°1-2012. *APLV*, Paris, 33-40.

Index

- accent, 93, 237, 257
- accent (définition de Léon), 93
- accent d'insistance, 92
- accent de phrase, 94
- accent démarcatif, 93, 177
- accent étranger, 97, 104, 170
- accent français, 84, 101, 169
- accent lexical, 95
- accent mélodique *Voir* accent nucléaire
- accent nucléaire, 362
- accent primaire, 370
- accent secondaire, 337, 370
- accent tonique, 146, 339, 369
- accented syllables*, 79
- accentuation, 92, 93, 94, 99
- accentuation lexicale, 94, 99
- aisance *Voir* fluency
- aisance (*fluency*) (définition de Gut), 151
- amplitude totale, 246
- analyse ascendante, 127
- analyse descendante, 127
- approche métrique autosegmentale, 78, **79**, 161
- archétype de conformité, 220, 488
- archétype de divergence, 220
- archétype du "meilleur" étudiant en prosodie, 427
- archétype du « mauvais » étudiant en prosodie, 427
- archétype stéréotypique, 207, 220, 415, 488
- auto-évaluation, 60
- autopoièse (principe d'), 51
- backwash effect*, 59
- barème (évaluatif), 56, 445, 489, 571, 573
- barycentre, 208
- bonus *Voir* taux de réussite et d'erreur
- catégorisation de la proéminence, 378
- construct*, 55, 56, 103, 120
- corrélats acoustiques de l'emphase, 245
- corrélats acoustiques de l'emphase, 376
- critères distinctifs, 58
- crierial features*, 59, 61, 149, 329, 489, 555
- data mining* (définition), 127
- débit *Voir* débit d'élocution (définition)
- débit d'élocution (définition), 397
- descripteur (*descriptor*), 57
- descripteurs prosodiques du CECRL, 105
- differentiable feature* *Voir* criterial feature
- dynamic pitch contours*, 79
- École britannique de l'intonation, 46, 78, 112, 146
- effet de retour, 59, 62
- élégance, 154, 172, 181, 378, 440
- emphase (définition d'Herment), 237
- emphase (définition), 224, 225, 227
- emphase (degré), 113, 241, 242, 243, 253, 254, 373, 377
- emphase (dualité de l'), 238, 240
- emphase (mise en scène), 239
- emphase (notion), 226, 237, 376
- emphase (pointage), 253, 254
- emphase (pourcentage), 243
- emphase (repérage), 233, 236, 241, 242, 375
- emphase (utilisation), 223, 378
- emphase prosodique (ma définition), 239
- emphase purement émotionnelle (*purely emotive emphasis*), 287
- énaction, 47, 48, 539
- énoncés interrompus (*interrupted phrases*), 159
- équité (*fairness*), 57
- évaluation critériée (*criterion-referencing*), 61
- évaluation de la performance ou capacité (*proficiency assessment*), 60
- évaluation diagnostique, 59
- évaluation directe et évaluation indirecte, 61
- évaluation du savoir ou du niveau (*achievement assessment*), 60
- évaluation formative (*formative assessment*), 60
- évaluation normative (*proficiency assessment*), 61
- évaluation sommative (*summative assessment*), 60
- extralinguistique, 45

- faisabilité (*practicality* ou *feasability*), 57
- fiabilité, 56
- fidélité (*reliability*), 56
- filled pauses* Voir pauses remplies
- fluency*, 66, 102, 105, 106, 107, 169, 428
- focalisation informationnelle (*information focusing*), 287
- focalisation unique (structure de), 267
- focus, 88, 177, 237
- focus (absence de), 322, 323
- focus étroit (*narrow focus*), 88
- focus large (*broad focus*), 88, 321
- fouille des données (*data mining*), 126
- head*, 79
- hiérarchie accentuelle, 93, 176, 177, 179
- indicateur du profil d'âge, 377
- information ancienne (*given*) Voir structure informationnelle
- information nouvelle (*new*) Voir structure informationnelle
- instructional value* Voir effet de retour
- inter-évaluation, 60
- interlangue, 94, 103, 104, 273, 274, 357
- interlangue (définition), 102
- intonation, 43, 52, 74, 86
- intonation (définition de Gut), 78, 160
- intonation (définition), 78
- intonation (ma définition), 48
- intonation (modélisation), 78
- intonation des exclamatives, 272
- intonation des questions en WH-, 309
- intonation des questions polaires, 312
- intonation phrase*, 80
- isosyllabité, 175, 177, 180
- langage corporel (*body language*), 45
- langue à accent lexical libre, 100
- langue à isochronie accentuelle, 95
- langue isosyllabique, 95
- lignes d'annotations (*tiers*), 243
- locus accentuel, 366, 367
- loi de la dérivation neutre, 369
- macro-profil, 141
- malus Voir taux de réussite et d'erreur
- mélodie, 74, 88, 560, Voir intonation
- mélodie tonale, 408
- mètre (selon Di Cristo), 91
- métrique Voir mètre
- micro-pauses, 69
- micro-profil, 138, 141, 148
- modèle *CRISP-DM*, 124, 195, 205, 533
- nucleus, nuclei*, 79
- objectivité (en évaluation), 57
- oxytonique, 99
- paralangue (définition de Crystal), 44
- paralinguistique, 43, 45, 49, 50, 51, 70
- patron tonal, 342, 343, 407, 473
- pause (définition), 68
- pause pleine Voir pause remplie
- pause remplie, 68
- pause silencieuse, 68
- perception, 75, 370
- perception de l'emphase, 237
- perception différentielle de F0, 76
- perception du rythme, 90
- phonostyle, 71, 129, 536
- pianiste (profil), 440, 441, 442, 443, 444, 561, 564
- polysyllabes, 367
- prehead*, 79
- principe d'économie, 238
- principe d'emphase, 237
- proéminence, 237, 247, 257, 409
- proéminence (définition d'Avanzi *et al.*), 93
- proéminence (degré), 318
- proéminence accentuelle, 394
- profil (approche diachronique), 118
- profil (définition générale), 120
- profil (ma définition), 129
- profil (notion), 117, 119, 128, 138, 148, 227
- profil (utilisation selon Crystal), 147
- profil d'apprentissage Voir style d'apprentissage
- profil de groupe, 120, 158
- profil distributif, 127, 159, 220, 399, 403
- profil extrême Voir archétype stéréotypique
- profil jeune, 277, 377
- profil linguistique (définition de Crystal), 136

Index

- profil linguistique (finalité), 136
- profil mature, 277
- profil médico-légal, 136
- profil non-distributif, 128, 220
- profil personnalisé (*personalised*), 120, 122, 130, 219
- profil prosodique (définition de Crystal), 141
- profil prosodique (ma définition), 129
- profil stéréotypique d'échec, 425
- profil stéréotypique de réussite, 425
- profilage (définition), 121, 186
- profilage (finalité), 121
- profilage (notion), 117, 119, 130, 187
- profilage (processus), 122
- profilage de détection de l'emphase, 247
- profilage de groupe, 123, 219
- profilage de l'interlangue, 186
- profilage de l'intonation, 161
- profilage judiciaire *Voir* profilage phonétique médico-légal
- profilage linguistique, 140
- profilage par analyse acoustique, 195
- profilage par la rapidité de lecture, 411, 446, 447, 467, 489, 529, 530, 531
- profilage par prototype de grille évaluative (syllabe accentuée de *nobody*), 473
- profilage par prototype de grille évaluative (syntagme neutre), 463
- profilage par prototype de grille évaluative (ton \), 467
- profilage par prototypes de grilles évaluatives (exclamative en *How*), 458
- profilage par prototypes de grilles évaluatives (*so*), 478
- profilage phonétique médico-légal, 131
- profilage selon Crystal, 140
- prosodème, 43, 44
- prosodie (CECRL), 105
- prosodie (définition de Crystal), 43
- prosodie (définition de Di Cristo), 43, 44
- prosodie (définition de Vaissière), 42
- prosodie (ma définition), 48
- prosodie (organisation), 51
- prosodie (subjectivité), 112
- prosodie (subliminal), 378
- prosodie interphonologique, 149
- rapidité de lecture, 249, 325, 396
- rythme, 41, 43
- rythme (conclusion des définitions), 92
- rythme (corrélats acoustiques utilisés par Gut), 157
- rythme (définition de Léon), 93
- rythme (hypothèse de Bolinger), 156
- rythme (interlangue), 97
- rythme (profilage de Tortel), 164
- rythme (selon Di Cristo), 91
- rythme (selon Grammont), 89
- rythme (selon Gut), 90, 95
- rythme (selon Tortel), 91
- rythme (selon Vaissière), 93
- rythme (structure rythmique sous-jacente), 91
- rythme du Selfish Giant, 111
- rythme naturel, 90
- rythme produit artificiellement, 90
- rythmicité binaire, 91
- saillance (définition de Landragin), 228
- saillance (définition), 227, 228
- saillance (dualité de la), 240
- saillance cognitive (C-saillance), 230
- saillance discursive, 231
- saillance intrinsèque, 231
- saillance physique (P-saillance), 230
- saillance située (S-saillance), 230, 231
- schéma accentuel, 369, 370, 394
- single focus structure* *Voir* structure de focalisation
 - unique
- sous-profil, 377
- speech rate* (vitesse d'articulation), 69
- stress, 237
- stress group*, 94
- stressed syllables*, 79
- structure informationnelle (*information structure*), 88
- style d'apprentissage (définition de Cyr), 444
- subliminal, 378, 540
- suprasegmental, 43, 160, 245, 247
- syllabe faible, 364, 365
- syllabe forte, 364, 365
- tail*, 79

taux d'articulation (*speech rate*) Voir vitesse
d'articulation
tempo, 43
tempo (définition de Di Cristo), 70
tempo (perception), 71
Test de positionnement (*placement test*), 60
timbre, 47, 177, 247, 301
ton bidirectionnel, 341, 348, 371, 406, 466
ton tridirectionnel, 340, 473
tonetic stress marks, 326
tonicity (définition de Crystal), 146
tonicity (indeterminate), 146
tonicity (stereotyped), 146

tons hypercomplexes ou tridirectionnels, 343
trait critérié Voir trait discriminant
trait critérisé Voir trait discriminant
trait distinctif ou discriminant (*criteria feature*), 59
unité intonative (*intonation phrase, IP*), 46, 159
unstressed syllables, 79
validité (*validity*), 55
vitesse d'articulation, 70, 72
vitesse d'élocution Voir tempo
vitesse de lecture Voir tempo
vitesse de parole Voir tempo
washback effect, 62, Voir effet de retour