

Université

de Strasbourg



ÉCOLE
DOCTORALE
DES HUMANITÉS
(ED 520)

UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE DES HUMANITES (ED 520)

UR 1339 LiLPa (Linguistique, Langues et Parole)

THÈSE présentée par :

Kossi Seto YIBOKOU

soutenue le : 28 novembre 2019

pour obtenir le grade de : **Docteur de l'Université de Strasbourg**
Discipline/ Spécialité : Sciences du langage (Didactiques des langues)

Apprentissage informel de l'anglais en ligne : quelles conséquences sur la prononciation des étudiants français ?

THÈSE dirigée par :

Mme Béatrice Vaxelaire

Professeure, Université de Strasbourg

Mme Denyze Toffoli

Professeure, Université Toulouse 3

RAPPORTEURS :

Mme Cécile Fougeron

Directeur de Recherche, Université Sorbonne-Nouvelle

M. Geoffrey Sockett

Professeur, Université Paris-Descartes

AUTRES MEMBRES DU JURY :

Mme Shona Whyte

Professeure, Université Côte d'Azur

M. Grégory Miras

Maitre de conférences, Université de Rouen

À toi...

Remerciements

Je voudrais exprimer ma plus grande gratitude envers mes directrices de thèse, Madame Béatrice Vaxelaire et Madame Denyze Toffoli. Ce travail n'aurait certainement pas abouti sans leur accompagnement scientifique et moral. Je les remercie pour leur soutien, leurs encouragements ainsi que pour la confiance qu'elles m'ont accordée.

Je remercie sincèrement les membres de mon jury : Monsieur Geoffrey Sockett, Professeur à l'Université Paris-Descartes ; Madame Cécile Fougeron, Directeur de Recherche à l'Université Sorbonne-Nouvelle ; Madame Shona Whyte, Professeure à l'Université Côte d'Azur et Monsieur Grégory Miras, Maître de conférences à l'Université de Rouen ; d'avoir accepté d'évaluer cette recherche.

Je tiens à témoigner toute ma reconnaissance envers les personnes suivantes, ayant contribué de près ou de loin à la réalisation de cette thèse :

À Monsieur le professeur Rudolph Sock pour ses aides, ses précieux conseils, sa bonne humeur et sa capacité à toujours débloquer les situations « sans issues ».

À mes collègues du Département de Linguistique Appliquée et de la Didactique des Langues : Julia Putsche, Peggy Candas, Laurence Schmoll, Laurent Perrot, Julie Dittel, Anissa Hamza et Jean-Paul Meyer, pour leur aide, leur patience et leur compréhension. Je n'oublie pas non plus mes anciens enseignants de ce département : Yannick Lefranc, Laurent Kashema et Nathalie Gettliffe.

À mes locuteurs qui ont bien voulu « prêter leur voix à la science » et aux personnes qui m'ont aidé dans leur recrutement.

À mon comité de suivi de thèse, Madame Nicole Poteaux et Marion Bechet, j'exprime ma profonde gratitude pour leurs conseils.

À Jérémie Beaudet, Marine Le Mené, Alexia Jingang, Jean-Baptiste Dutilleul, Claire Galand, Élodie Lang, Bruno Oberlé, Delphine Charuau et Emmanuelle Dantan, pour avoir consacré du temps à la relecture critique et constructive de la thèse. Toute erreur subsistant dans ce document n'est imputable qu'à moi-même.

À mes collègues de LiLPa Vosges : Rodrigo, Pablo, Chang, Seo, que je tiens à remercier pour chaque moment de discussions constructives.

À toute la famille de l'Institut de Phonétique de Strasbourg, je voudrais leur adresser toute ma reconnaissance. J'ai beaucoup appris pendant les années passées dans ce cadre bienveillant et enrichissant : Fayssal Bouarourou, Saïd Bouzidi, Camille Fauth, Thomas Jauriberry, Hasna Zaouali, Tomoki Koya, Noé Xiu, Naïma Derbal, Seynabou Fall, Nadjia Gaci, Fadila Azzag, Christiane Bosso, Xuelu Zhang, Marie Didiënne et Cindy Ceolho.

Je pense également à mes anciens enseignants de l'Université du Ghana, Legon : Prof Yennah, Prof Asaah, Dr. Chachu, Dr. Badasu, Dr. Azanku, Dr. Kuto, Dr. Koffi.

À toute la famille Pfortner, qui m'a accueilli à bras ouverts depuis tant d'années et qui m'a également transmis l'amour pour Strasbourg : Nicole, Jacques, Corentin, le grand Rudi et Zoé. Je remercie également les familles Yibokou, Kudabor, Gbesemete, Shianti, Agbeshi, Okudzeto et Tutu-yeboah pour leur soutien moral.

Merci infiniment aux familles Galand, Kumielan et Polis pour leur soutien chaleureux.

À mes amis Mahaut, Nicolas et Pâris ; Aline, Steeve et Maël ; Véro, Doro et Lénaïc ; Mosumula, Mr. Shabba et Anike ; Monica Fierro, Marisa et Tommy, Samah, Chloé Faucompré, François Roech, Marion Cargill, Damy, Sophie et Lenny Lamboni, Yohan Solon, Riwad, les amis du RU, Zacharia Otieno, Julien Rentz, Abdoulaye, les Michelsbourg, Quentin Burghart, Kamal, Yao Sonka, Paul Kutanyi; Ben Quenum, Les Tetteh-Klu, Victoria Sabbah, Sena, IGYMS et Gospel Rhymes.

À ma famille qui, de loin, m'a soutenu sur tous les plans, je pense très fort à Miranda et Emefa.

Sommaire

Remerciements.....	ii
Sommaire	iv
Table des abréviations	vii
Introduction générale	9
1 Contexte et problématique de la thèse	12
1.1 Types d'apprentissage : formel, non formel et informel.....	13
1.2 L'apprentissage informel des langues.....	17
1.3 L'apprentissage informel de l'anglais en ligne	18
1.4 Objectifs de cette étude.....	30
1.5 Questions de recherche et Hypothèses.....	32
1.6 Conclusion	33
2 Autour de la langue anglaise	35
2.1 L'importance de la langue anglaise	35
2.2 Le français et l'anglais : des influences réciproques	38
2.3 L'anglais dans le monde : les différents types d'anglais.....	44
2.4 Différences phonétiques et phonologiques entre l'anglais et le français	60
2.5 Conclusion	65
3 Théorie des systèmes complexes et dynamiques	67
3.1 Les origines de la théorie des SCD.....	68
3.2 La théorie des SCD aujourd'hui.....	70
3.3 Caractéristiques de la théorie des SCD	72
3.4 La théorie des SCD dans le domaine de l'apprentissage des langues	84
3.5 Limites de la théorie des SCD en didactique des langues.....	93
3.6 Conclusion	99
4 Théorie de l'imitation vocale.....	102
4.1 Imitation : cas général.....	103
4.2 Imitation du point de vue neurophysiologique : le rôle des neurones miroirs.....	105
4.3 Imitation générale et imitation vocale.....	109
4.4 Imitation vocale volontaire et involontaire.....	113
4.5 Imitation directe (non-médiée) et indirecte (médiée).....	126
4.6 Des modèles d'acquisition phonétique et phonologique	129
4.7 Conclusion	136
5 Viabilité et perturbation en production-perception de la parole	139
5.1 Quelques définitions : fiabilité, variabilité, invariance et viabilité.....	140
5.2 La théorie Quantique : une théorie de l'invariance	142
5.3 La théorie de la Variabilité Adaptative.....	144
5.4 La théorie de la viabilité en production-perception de la parole	147
5.5 Perturbation du système de production de la parole.....	157
5.6 Théorie de la viabilité en phonétique : implications pour l'apprentissage des L2.....	166
5.7 Conclusion	168
6 Cadre méthodologique	170
6.1 Recrutement des participants	170

6.2	Matériel de collecte de données audio.....	171
6.3	La collecte de données.....	172
6.4	Élaboration du support pour la prise de sons.....	172
6.5	Enregistrement des participants.....	178
6.6	Segmentation et mesures acoustiques des sons cibles	181
6.7	Test de perception	196
6.8	Questionnaire.....	198
6.9	Conclusion	201
7	Résultats de la prononciation en vitesse normale	203
7.1	La prononciation du /t/ intervocalique	204
7.2	La rhoticité.....	210
7.3	Le yod-dropping	218
7.4	La prononciation de la voyelle de LOT	224
7.5	La prononciation de la voyelle de GOAT.....	227
7.6	La prononciation des graphies <i> et <y>	230
7.7	La prononciation de la graphie <i>	233
7.8	La prononciation de la graphie <ei> dans leisure	237
7.9	La prononciation de la graphie <e> dans evolution	238
7.10	Nasalisation vocalique	240
7.11	La place de l'accent principal.....	245
7.12	Le nombre de syllabes accentuées	249
7.13	Résultats globaux des mots à vitesse normale	254
7.14	Conclusion	256
8	Résultats de la prononciation en vitesse rapide	258
8.1	La prononciation du /t/ intervocalique	259
8.2	La rhoticité.....	264
8.3	Le yod-dropping	272
8.4	La prononciation de la voyelle de LOT	277
8.5	La prononciation de la voyelle de GOAT.....	280
8.6	La prononciation des graphies <i> et <y>	283
8.7	La prononciation de la graphie <i>	286
8.8	La prononciation de la graphie <ei> dans leisure	289
8.9	La prononciation de la graphie <e> dans evolution	290
8.10	Nasalisation vocalique	291
8.11	La place de l'accent principal.....	295
8.12	La voyelle de la pénultième syllabe et le nombre de syllabes accentuées	298
8.13	Résultats globaux des mots à vitesse rapide	303
8.14	Conclusion	308
9	Résultats du questionnaire et du test de perception.....	310
9.1	Informations biographiques.....	311
9.2	Temps d'apprentissage et temps d'utilisation de l'anglais.....	312
9.3	Séjours et leur durée en pays anglophones.....	315
9.4	Nationalité et accent des enseignants.....	316
9.5	Mots inconnus.....	318
9.6	Imitation d'accent.....	320

9.7	Pratiques d'exposition à l'anglais via les séries télévisées.....	322
9.8	Résultats du test de perception	338
9.9	Triangulation de données.....	341
9.10	Conclusion	353
10	Synthèse de la recherche et conclusion générale	356
10.1	Retour sur les questions et hypothèses de départ.....	356
10.2	Discussion des résultats en lien avec le cadre théorique	360
10.3	Apports de notre travail de recherche.....	365
10.4	Limites et perspectives de ce travail de recherche	367
	Conclusion générale	372
	Bibliographie	375
	Table des matières	396
	Table des figures	401
	Tables des tableaux.....	404
	Index des notions	405

Table des abréviations

ACEDLE	Association des Chercheurs et Enseignants Didacticiens des Langues Étrangères
ACP	Analyse en Composantes Principales
AIAL	Apprentissage Informel de l'Anglais en Ligne
AII	<i>Acoustic Image Imprinting</i> (Impression d'images acoustiques)
BBC	<i>British Broadcasting Corporation</i>
BNC	<i>British National Corpus</i>
CAH	Classification Ascendante Hiérarchique
CALL	<i>Computer Assisted Language Learning</i> (Apprentissage des Langues Assisté par Ordinateur)
CAPES	Certificat d'Aptitude au Professorat de l'Enseignement du Second degré
CECRL	Cadre Européen Commun de Référence pour les Langues (Conseil de l'Europe, 2001)
CNRS	Centre National de la Recherche Scientifique
CNRTL	Centre National de Ressources Textuelles et Lexicales
COCA	<i>Corpus Of Contemporary American</i> (Corpus de l'anglais américain contemporain)
CRAL	Centre de Ressources et d'Apprentissage des Langues
CRL	Centre de Ressources en Langues
CTL	<i>Classroom-Trained Learners</i> (Apprenants formés en classe)
DDE	Disque Dur Externe
EEG	Électro-Encéphalographie
EMLL	<i>Extra-Mural Language Learning</i> (apprentissage extra-mural des langues)
EPD	<i>English Pronouncing Dictionary</i> (Dictionnaire de prononciation en anglais)
FASIL	<i>Fully Autonomous Self-Instructed Learners</i> (Apprenants autoformés entièrement autonomes)
GA	<i>General American</i> (Anglais standard américain)
HHOLD	<i>How I met your mother, House, One tree hill, Lost, Desperate Housewives</i> (Corpus composé des transcriptions d'une saison de chacune de ces séries télévisées américaines)
IDLE	<i>Informal Digital Learning of English</i> (apprentissage informel de l'anglais par le numérique)
IPA	<i>International Phonetic Alphabet</i> (Alphabet Phonétique International)
IPS	Institut de Phonétique de Strasbourg
IRMf	Imagerie par Résonance Magnétique fonctionnelle
L1	Langue(s) première(s) (maternelles)
L2	Langue(s) seconde(s) (appries après les L1)
LanSAD	LANGues pour Spécialistes d'Autres Disciplines
LE	Langue Étrangère
LLBC	Language Learning Beyond the Classroom (Apprentissage des langues au-delà de la salle de classe)
LPD	<i>Longman Pronunciation Dictionary</i>
NEQ	Notions En Question
NLM	<i>Native Language Magnet</i>

OCLL	<i>Out-of-Classroom Language Learning</i> (apprentissage des langues en dehors de la salle de classe)
OED	<i>Oxford English Dictionary</i>
OILE	<i>Online Informal Learning of English</i> (Apprentissage Informel de l'Anglais en Ligne)
PAM	<i>Perceptual Assimilation Model</i>
RAL	Recherche en Acquisition des Langues
RP	<i>Received Pronunciation</i>
RTATL	Rapport Temps d'Articulation-Temps de Locution
RVA	Reconnaissance, Validation et Accréditation
SERIES	<i>Scholars Exchanging and Researching on International Entertainment Series</i>
SBE	<i>Standard British English</i> (Anglais britannique standard)
SCD	Systèmes Complexes et Dynamiques
SLM	<i>Speech Learning Model</i>
SOAP	<i>Corpus of American Soap Operas</i> (Corpus de séries télévisées américaines)
SSBE	<i>Standard Southern British English</i> (Anglais standard du sud)
TIC	Technologies de l'Information et de la Communication
UK	<i>United Kingdom</i> (Royaume-Uni)
VF	Version Française
VN	Vitesse Normale
VNVR	Vitesse Normale et Vitesse Rapide
VO	Version Originale
VOIP	<i>Voice Over Internet Protocol</i> (Voix sur protocole Internet)
VOST	Version Originale Sous-Titrée
VOSTA	Version Originale Sous-Titres Anglais
VOSTF	Version Originale Sous-Titres Français
VOT	<i>Voice Onset Time</i> (délai d'établissement du voisement)
VR	Vitesse Rapide
VTT	<i>Voice Termination Time</i> (délai d'arrêt du voisement)

Introduction générale

L'hégémonie de la langue anglaise dans le monde semble aujourd'hui un fait indéniable (Guo & Beckett, 2007). Omniprésent sur tous les continents, l'anglais est utilisé tant au plan international que national : en politique, en commerce, dans le domaine de la culture et des sciences, etc. Par conséquent, l'anglais est devenu la langue la plus enseignée, apprise et parlée dans le monde entier (Svartvik et Leech, 2016). Avec la recrudescence des objets connectés (ordinateurs, tablettes et téléphones portables) et le progrès des affordances du web 2.0 (réseaux sociaux ; sites d'informations, de vidéos, de musique et de jeux), l'exposition à cette langue s'intensifie et ne cesse de croître, touchant ainsi à toutes les strates de la société, des plus jeunes aux plus âgés, pour qui l'accès à l'information en ligne devient aujourd'hui plus aisé qu'il y a quelques années. Les retombées sont donc nombreuses : la connexion à Internet étant pratiquement gratuite, les apprenants de langues choisissent à leur guise les ressources en ligne pour des pratiques diverses et variées, créant des situations d'apprentissage de la langue, en dehors de la classe, beaucoup plus importantes quantitativement que ce qui se passe au sein de la classe de langues (Sockett, 2014).

L'« apprentissage informel de l'anglais en ligne » (AIAL) est un domaine de recherche sur les apprentissages incidents de l'anglais au travers des pratiques en contextes informels sur Internet (Toffoli et Sockett, 2010). Dans ce contexte, l'AIAL ne se déroule pas selon un emploi du temps imposé, ni commandité par une instance officielle, et le but de l'apprenant n'est pas d'apprendre, mais de pouvoir utiliser la langue et d'échanger avec d'autres personnes. À l'époque où l'enseignant et le manuel constituaient le seul moyen dont disposait l'apprenant pour être en contact avec la langue de manière formelle, les voyages dans les pays anglophones permettaient d'apprendre la langue de manière informelle (Toffoli et Sockett, 2015). Mais de nos jours, l'exposition à l'anglais authentique se fait également au travers des activités en ligne. Aussi, si traditionnellement l'anglais britannique est la première référence pour l'enseignement, actuellement, l'anglais américain devient de plus en plus accessible comme le remarque Carrier : « *American English magazines, books, Internet traffic, movies, television, and music, have inundated the world market and can be found in the remotest places on the planet* » (2014, p. 106).

En effet, en France, quelques chiffres médiométriques montrent qu'entre 2011 et 2016, la moitié des recettes des guichets de salles de cinéma provient de films américains. Entre 2005 et 2017, les

Français passent en moyenne trois heures et demie par jour devant la télévision (Statista, 2019). En 2014, la France occupait le huitième rang des pays où l'on regarde le plus la télévision (Richter, 2015) tandis qu'en 2016, 46 % des jeunes français (de 18 à 34 ans) optent pour les films en VOST¹ et qu'en 2012, plus de 80 % des moins de 35 ans préfèrent les séries américaines aux séries françaises (Statista, 2012). Les films et séries constituent par ailleurs l'un des médias les plus documentés² dans le monde. Les publications sur les séries télévisées se sont multipliées au cours des dernières années, non seulement sur le plan de la vulgarisation mais aussi au niveau de la recherche scientifique. Tout d'abord, sur la toile, nombreux sont les journalistes ou blogueurs qui publient des billets sur le sujet et que l'on retrouve dans les titres : « Pourquoi les séries télévisées envahissent-elles tous les écrans ? » (Séry, 2010) ; « Pourquoi les séries télé ont-elles tant de succès ? » (Van Dievort & Duchâteau, 2014) ; « Séries : pourquoi elles rendent les téléspectateurs complètement accros » (Baudin & Chuc, 2014) ; « Séries, elles ont pulvérisé le petit écran » (Armati, Desplanques, & Poitte, 2011) ; « Séries TV, un phénomène de société » (Chauveau, 2013) ; « Comment les séries ont envahi notre quotidien » (Corbillé, 2014) ; « "Mentalist", "NCIS", "Les Experts": pourquoi les séries américaines cartonnent en France » (Jost, 2013) ; « Depuis quand les Français sont-ils accros aux séries TV ? » (Boutet, 2014), pour en citer quelques-uns.

Au niveau des études scientifiques, un nombre important de travaux universitaires de diverses disciplines, notamment des mémoires et thèses de doctorat, se sont pleinement penchés sur le phénomène sériel³ (Laborde, 2017). En plus d'ouvrages, de numéros de revues, de colloques ou encore de journées d'études, il est possible de dénombrer une centaine de thèses soutenues et en cours qui portent uniquement sur les séries télévisées⁴. On retrouve ainsi des thèses portant sur les séries télévisées en sociologie, en sciences de l'information et de la communication, en littérature, en sciences du langage, en lettres, en histoire, en sciences politiques et en didactiques des langues étrangères (Park, 2013). Certains titres de thèses évoquent leur sujet au détour de formules aussi attrayantes que ceux des billets de vulgarisation et évoquent plus ou moins leur contenu : « La "blockbusterisation" des séries télévisées américaines » (David, en cours)⁵ ; « Les séries télévisées :

1 Version Originale Sous-Titrée.

2 C'est-à-dire qui fait l'objet de recherches scientifiques, à travers des écrits variés (livres, articles, magazines, etc.).

3 Dans la lignée des travaux de Combes (2013), nous utilisons cet adjectif en référence aux pratiques de visionnages de séries télévisées.

4 www.theses.fr

5 <http://www.theses.fr/s78916>

une lutte sans fin » (Lifschutz, 2015)⁶ ; « Apocalypse et fin du monde dans les séries télévisées américaines » (Melquiond, en cours)⁷. Est également apparu au sein du CNRS⁸, la constitution du réseau SERIES. (*Scholars Exchanging and Researching on International Entertainment Series*) et sa revue numérique *TV series* comportant pléthore de travaux transdisciplinaires sur les séries ainsi que la création en 2010 de la revue francophone « Télévision »⁹ par François Jost (Combes, 2013).

Nous situant pour notre part dans la discipline de la didactique des langues étrangères, nous nous intéressons plus particulièrement aux répercussions que peuvent avoir ces médias sur les compétences en langue (ici l'anglais) des apprenants qui s'adonnent à ces activités extramurales. Quelles sont les conséquences de ces pratiques sur l'acquisition de compétences linguistiques, et plus spécifiquement sur la prononciation d'apprenants français ? Si, il y a quelques années, des chercheurs ont réfuté l'impact des médias sur la prononciation (Trudgill, 1986 ; Chambers, 1998 ; Labov, 2001), aujourd'hui cette hypothèse paraît invalidée non seulement par des observations courantes mais aussi par des études empiriques (Stuart-Smith *et al.*, 2011). Chercher à établir des liens entre les médias, surtout les séries télévisées, et l'acquisition phonético-phonologique est ainsi l'objectif que se fixe ce travail de thèse. Il se propose d'analyser la prononciation d'un échantillon d'étudiants français qui pratiquent l'apprentissage informel de l'anglais en ligne.

6 <http://www.theses.fr/2015LYO20086>

7 <http://www.theses.fr/s122703>

8 <http://www.series.cnrs.fr/>

9 http://www.cnrseditions.fr/collection/297_television

1 Contexte et problématique de la thèse

If all learning were to be represented by an iceberg, then the section above the surface of the water would be sufficient to cover formal learning, but the submerged two-thirds of the structure would be needed to convey the much greater importance of informal learning. Coffield (2000, p. 1).

« Est-il possible d'apprendre en dehors de l'école ? De quelles façons ? La culture savante s'acquiert-elle seulement par l'enseignement à l'école ? » (Jacobi, 2018, p. 1). Si depuis quelques années les acteurs du monde éducatif s'interrogent sur les conséquences de l'apprentissage en dehors des milieux scolaires, force est de constater l'intérêt grandissant que des chercheurs portent à ce domaine méconnu (Werquin, 2010). En effet, l'apprentissage explicite et formel a beaucoup et longtemps été documenté dans les recherches scientifiques mais son « contraire », que l'on qualifie souvent d'implicite, de non formel ou d'informel a finalement été peu étudié. La représentation selon laquelle l'apprentissage informel est une activité dénuée de toute valeur éducative est bien répandue et cette activité est souvent associée aux pratiques telles la lecture (livres, romans, presse), l'écoute de la musique, le visionnage de vidéos (films, séries), la visite de lieux particuliers (musées, sites historiques), les jeux, etc. De plus, avec l'avènement d'Internet, les mœurs ont changé et l'on voit aujourd'hui des répercussions sur la vie de ceux qui les pratiquent, notamment les jeunes adultes. Du visionnage de vidéos, de l'écoute de la musique, à la recherche en ligne, en passant par les nombreux réseaux sociaux, ces pratiques bouleversent le quotidien de ces individus, en termes de vie sociale, mais également en termes d'acquisition de savoirs et savoir-faire. Le peu de recherches consacrées à ce domaine, que certains auteurs nommeraient « l'école parallèle » (Jacobi, 2018) ou l'éducation non formelle/informelle (Unesco, 2012) ont montré que le rôle et les impacts de ces pratiques sont loin d'être négligeables. De surcroît, il semblerait que l'on acquiert plus de connaissances et de compétences en dehors de la salle de classe qu'au sein de celle-ci (Jacobi, 2018).

Notre thèse se situe ainsi dans le contexte de l'apprentissage informel de langues, et plus spécifiquement concernant la langue anglaise, au travers d'activités en ligne. Elle a pour principal objectif l'établissement de liens entre les facteurs de prononciation et les pratiques informelles de jeunes français. Ce chapitre sera par conséquent consacré à la présentation du contexte et de la problématique de notre recherche. En premier lieu, nous introduirons les types d'apprentissage qui existent dans le domaine de l'éducation en général, à savoir le formel, le non formel et

l'informel. En guise de précision terminologique, ces trois notions seront présentées selon des définitions que l'on retrouve dans la littérature, et on évoquera les débats actuels sur leur emploi. Il s'agira par exemple de confronter le choix de l'utilisation des mots « éducation » / « apprentissage » ou de l'expression « apprentissage informel » / « apprentissage en contexte informel ». En second lieu, après avoir présenté l'apprentissage informel de langues, nous aborderons le sujet central de ce travail de thèse qui est l'apprentissage informel de l'anglais en ligne (AIAL). Nous définirons ensuite les objets d'étude de ce domaine, puis nous résumerons quelques travaux conduits en France et synthétiserons quelques autres menés hors de la France, en vue de mieux situer notre contexte de recherche. En dernier lieu, nous développerons la problématique de notre travail, en présentant l'objectif, les questions et les hypothèses de recherche.

1.1 Types d'apprentissage : formel, non formel et informel

Cette partie est consacrée à la définition de trois notions importantes qui seront fréquemment utilisées dans de cette étude. En guise de précision terminologique, nous ferons également le point sur les débats actuels concernant l'emploi de ces termes, puis nous nous positionnerons selon les diverses propositions discutées. Dans la sphère de l'éducation en général, trois types d'apprentissages sont habituellement identifiés. Il s'agit de l'apprentissage formel, de l'apprentissage non formel et de l'apprentissage informel, qui constituent trois expressions difficiles à utiliser (Hart, 2013). Dans les paragraphes qui suivent, nous fournirons les définitions de ces trois types d'apprentissage en fonction des instances éducatives officielles au sein desquelles elles sont utilisées (Conseil de l'Europe, UNESCO), suivies des définitions et points de vue d'autres auteurs. Ainsi, l'« apprentissage formel » caractérise :

Un apprentissage dispensé dans un contexte organisé et structuré, spécifiquement consacré à l'apprentissage, et qui débouche généralement sur la délivrance d'une qualification, en général sous la forme d'un certificat ou d'un diplôme ; il comprend les systèmes d'enseignement général, de formation professionnelle initiale et d'enseignement supérieur. (Conseil de l'Union Européenne, 2012, p. 5)

L'« apprentissage non formel » concerne pour sa part :

Un apprentissage dispensé sous forme d'activités planifiées (en termes d'objectifs d'apprentissage et de temps d'apprentissage), reposant sur une certaine forme de ressources ou d'accompagnement (relations étudiant-professeur, par exemple) ; il peut consister en des programmes d'acquisition d'aptitudes professionnelles, d'alphabétisation

des adultes et de formation de base pour des jeunes en décrochage scolaire. L'apprentissage non formel consiste très souvent en des formations en entreprise par lesquelles les employeurs mettent à jour et améliorent les compétences de leurs travailleurs (par exemple dans le domaine des TIC), de formations structurées en ligne (par exemple à l'aide de ressources didactiques en libre accès) et de cours proposés par des organisations de la société civile pour leurs membres, leur groupe cible ou le grand public. (*ibid.*)

Enfin, l'« apprentissage informel » désigne :

Un apprentissage découlant des activités de la vie quotidienne liées au travail, à la famille ou aux loisirs et il n'est ni organisé ni structuré en termes d'objectifs, de temps ou de ressources. Il peut posséder un caractère non intentionnel de la part de l'apprenant. Les acquis de l'apprentissage informel peuvent être des aptitudes acquises par des expériences personnelles et professionnelles, des compétences en gestion de projets ou en informatique acquises au travail, des langues apprises et des aptitudes interculturelles acquises durant un séjour à l'étranger, des compétences informatiques acquises en dehors du travail, des aptitudes acquises dans le cadre d'activités bénévoles, culturelles ou sportives, par le biais du travail auprès des jeunes ou d'activités à domicile (garde d'un enfant, par exemple). (*ibid.*)

Les définitions de l'UNESCO¹⁰ (2012) sont quasi identiques à celles du Conseil de l'Europe. Elles stipulent en effet que les apprentissages non formel et informel peuvent aboutir à des certifications ou d'autres formes de reconnaissance, au travers d'un processus de Reconnaissance, Validation et Accréditation (RVA) des acquis (UNESCO, 2012, p. 8). Nous soulignons ainsi, quoique les éléments dans les définitions soient assez limpides, qu'il existe une certaine porosité entre ces trois formes en termes de lieux de pratique, d'intentions et de types de connaissances acquises. Aussi, selon ces définitions, le non formel serait plus proches du formel que de l'informel, en raison de la possibilité de certification. Ces notions nous importent particulièrement car ils constituent l'un des composants-clés de cette thèse, surtout l'apprentissage informel.

Pour ces auteurs, l'apprentissage informel est aujourd'hui si important qu'il doit être considéré comme fondamental, nécessaire et valorisé, parfois directement dans le travail. Ensuite, si les définitions des termes que nous venons de présenter proviennent de documents officiels et sont communément utilisées, il n'est pas rare de rencontrer différents mots ou expressions annexes dans certains ouvrages. Par exemple, alors que certains auteurs emploient les mots « éducation » et « apprentissage » de manière interchangeable (Hart, 2013; UNESCO, 2012), d'autres font une nette distinction entre ces deux termes qu'il convient d'explicitier. Tout d'abord, Garnier pense que

¹⁰ Nous ne les reprenons pas, sauf pour un point de précision. Les définitions de l'Unesco sont consultables à l'adresse suivante : <http://www.unesco.org/new/en/unesco/resources/online-materials/publications/unescodoc-database/>.

« l'éducation informelle a fait fortune dans la littérature des sciences de l'éducation pour désigner des processus d'apprentissage » (Garnier, 2018, p. 67). Quant à Meunier (2018), elle considère que le terme « éducation » renvoie au lieu et au cadre où l'apprentissage se déroule (de la maternelle à l'université) et que l'apprentissage se réfère au processus « apprendre ». Elle constate que :

[...] en consultant différents écrits qui cherchent à distinguer éducation formelle, non formelle et informelle, nous avons vite compris qu'il y régnait une certaine confusion quant au sens conféré à chacune de ces notions. En outre, éducation et apprentissage y sont confondus : ainsi parlera-t-on sans discernement d'éducation ou d'apprentissage tout comme on emploiera l'expression « apprentissage informel » en lieu et place « d'éducation informelle ». (Meunier, 2018, p. 20)

Toujours selon Meunier, un apprentissage serait les résultats ou les bénéfices de l'éducation : « [...] cette distinction est nécessaire dans la mesure où nous employons le mot 'apprentissage' pour désigner un résultat. Nous nommons 'apprentissage' ce qui est obtenu après que l'apprenant ait été soumis à des environnements qui le suscitent » (*ibid.*, p. 20). Plus loin, l'auteure se demande si un apprentissage peut être qualifié de formel, non formel ou informel (*ibid.*). Cette question pourrait trouver sa réponse dans cette citation de Garnier :

L'une des difficultés sémantiques de l'usage du mot informel provient du fait que cette notion ne peut s'appliquer de la même façon au processus (l'apprentissage comme *learning*) et aux produits (le savoir ou l'apprentissage comme résultat). Il est contestable de prétendre assimiler ces deux sens en considérant que les apprentissages informels ne peuvent produire que des savoirs informels. (Garnier, 2018, p. 70)

Il est important à ce niveau de préciser que nous avons remarqué une contradiction dans les propos de certains auteurs qui essaient de faire la distinction entre « éducation » et « apprentissage » (préférant finalement le mot éducation), mais qui finissent tout de même par les utiliser à plusieurs reprises de manière interchangeable (Berry & Garcia, 2016; Garnier, 2018; Hart, 2013). C'est notamment l'exemple de Garnier (2018, p. 81) qui écrit : « L'éducation informelle est un apprentissage implicite et autonome qui peut fort bien se faire sur la base de supports formels ». Pour nous, l'usage ici du verbe « est » signifie que ce sont des synonymes et qu'il définit l'éducation par l'apprentissage. Dans le même ordre d'idée, l'utilisation des adjectifs « formel », « non formel » et « informel » pour qualifier l'apprentissage serait erronée selon certains auteurs. En effet, ces derniers considèrent qu'un apprentissage ne peut être formel, non formel ou informel, et qu'il convient à la place de qualifier plutôt la situation, le contexte, ou encore le cadre dans lesquels ces apprentissages se déroulent (Jacobi, 2018 ; meunier, 2018 ; Brougère, 2007). Ainsi, pour Silva &

Brougère : « l'apprentissage n'est pas en lui-même informel ou formel : c'est la situation que l'on peut considérer ainsi » (2016, p. 58). Brougère (2017, p. 4) s'explique davantage à ce propos et argumente que « l'apprentissage informel est pour nous une commodité de langage qui pose problème car ce n'est pas à proprement parler l'apprentissage qui est informel, c'est-à-dire dénué de toute mise en forme éducative, mais la situation dans laquelle l'apprentissage est réalisé ». D'abord, il nous semble que ce débat terminologique émane d'une « perte d'énergie » liée à la traduction de l'anglais vers le français de l'expression *Informal Learning* concernant l'apprentissage informel : « Au rôle et à l'impact de ces médias [...], les anglo-saxons ont donné le nom de *informal education*, expression que nous avons préféré traduire en français par éducation non formelle » (Jacobi & Schiele, 1990, p. 83). Apparemment, l'expression en français désigne autre chose, comme nous venons de le voir chez Brougère, ainsi que chez Toffoli : « Dans le contexte francophone, un positionnement terminologique plus juste me ferait pencher pour l'expression apprentissage en contexte informel plutôt qu'apprentissage informel tout court » (Toffoli, 2018, p. 113). Néanmoins, certains auteurs préfèrent garder le mot apprentissage comme le justifie Werquin (2010, p. 44) : « En français, le terme apprentissage est utilisé ici comme substantif du verbe apprendre et, de ce point de vue, correspond en tous points au terme anglais *learning* ». Il se pourrait que le domaine d'étude ou de recherche pris en compte soit également à l'origine de ce débat et il serait intéressant de voir comment les sciences de la nature ou les disciplines technico-pratiques définissent l'apprentissage informel. Ensuite, nous pensons que la valeur « réductrice » attribuée à l'aspect informel peut aussi empêcher certains auteurs d'utiliser l'expression « apprentissage informel », comme nous pouvons le relever dans la citation de Brougère (« dénué de toute mise en forme éducative ») ainsi que chez Jacobi :

Nommer ce secteur par sa traduction littérale - l'éducation informelle -, à cause de la valeur habituellement péjorative de l'épithète informel, pourrait laisser croire que les diverses composantes de ce secteur de la diffusion des savoirs et de la culture sont faibles, inorganisées et dépourvues de moyens conséquents. Or, comme on le verra, c'est tout le contraire : cette école parallèle est par nature différente de l'autre. (2018, p. 3)

Nous comprenons et partageons les arguments soulevés concernant le terme informel, mais nous pensons également que tous les points évoqués ne sont pas totalement erronés dans la mesure où l'apprentissage informel est justement par nature inorganisée (*cf.* Pour un périmètre des définitions de l'AIAL, p. 19).

Enfin, un autre débat terminologique similaire à celui abordé ci-dessus concerne la distinction entre l'apprentissage non formel et l'apprentissage informel. En effet, alors que le Conseil de l'Europe, l'UNESCO et certains auteurs en font une nette distinction, précisant de ne pas confondre l'apprentissage informel et non formel (UNESCO, 2012, p. 8), d'autres pensent qu'il n'y a pas lieu de parler d'« apprentissage informel ». Pour eux, il faudrait plutôt englober ces deux termes sous l'appellation générique d'« apprentissage non formel » (Meunier, 2018 ; Jacobi, 2018) :

Quant à la dénomination « éducation non formelle », elle traduit le concept anglo-saxon *informal education*, quoiqu'elle ne le recouvre pas systématiquement [...]. Dans le monde francophone, l'emploi de l'expression « éducation informelle » plutôt qu'« éducation non formelle » est donc soit injustifié, soit lié à une partition linguistique raffinée au sein du non formel. (Meunier, 2018, p. 18)

En vue de nous positionner dans cette discussion, nous partageons les arguments avancés par Meunier et admettons la distinction entre « apprentissage » et « éducation ». Cependant, nous conservons l'idée qu'il est nécessaire de distinguer le non-formel de l'informel et de continuer à utiliser ces deux termes. C'est pourquoi nous continuerons d'utiliser les expressions « apprentissages formel, non formel et informel » telles que définies par le Conseil de l'Europe. De ce fait, nous emploierons également l'expression « apprentissage en contexte formel, non-formel et informel » afin de nous référer à « l'espace » (physique ou virtuel) dans lequel le processus « apprendre » a lieu, et l'expression « apprentissage informel des langues » pour désigner le champ de recherche qui étudie les conséquences de ce phénomène sur le développement langagier. À présent que les précisions terminologiques sont établies, nous nous consacrerons dans les deux sections suivantes à la présentation du domaine de l'apprentissage informel des langues en général, et sur l'apprentissage informel de l'anglais en ligne en particulier.

1.2 L'apprentissage informel des langues

Ainsi que nous venons de le voir, les trois types d'apprentissage (formel, non-formel et informel) dans l'éducation ont été considérés indépendamment du contexte de la discipline ou de la matière. Nous soulignons cependant que les définitions données plus haut correspondent et s'appliquent également au domaine de l'enseignement/apprentissage des langues. L'apprentissage informel de langues (AIL) concerne ainsi tout apprenant de n'importe quelle langue qui s'engage dans des activités de la vie quotidienne, sans l'intention explicite d'acquérir des connaissances ou compétences langagières nouvelles. Il s'agit dès lors, par exemple, du cas des enfants qui

apprennent à parler leur L1 avant la scolarisation grâce à leurs environnements immédiats (comme avec les membres de leur foyer) ou éloignés (tel qu'avec d'autres individus de la communauté au sein de laquelle cette langue est parlée). De même, nous pouvons inclure les adultes en situation d'immigration qui interagissent avec les locuteurs de la nouvelle communauté et qui, par conséquent, s'approprient naturellement la langue. L'apprentissage formel (et non formel éventuellement) commence lorsque ces personnes s'inscrivent (administrativement et pédagogiquement) dans des établissements éducatifs spécialisés.

Depuis les années 1970, une attention particulière est d'ailleurs portée aux apprentissages de L2 se déroulant aussi bien dans la salle de classe qu'en dehors de celle-ci : « [research] has paid attention not only to learning in the classroom but also to so-called “natural,” “informal,” or “untutored” learning which takes place outside classrooms » (Littlewood, 2004, p. 504). La grande majorité de ces études porte sur la classe de langues mais, depuis peu, l'intérêt pour l'AIL regagne en vigueur, expliquant que des chercheurs s'intéressent de plus en plus à son influence sur l'acquisition langagière. Bon nombre de dénominations qualifiant ce domaine d'études sont ainsi apparues, en particulier dans le monde anglophone. Des expressions comme par exemple *Extra-Mural Language Learning* (EMLL), *Out-of-Classroom Language Learning* (OCLL) ou encore *Language Learning Beyond the Classroom* (LLBC), sont employées pour désigner les apprentissages ayant lieu en dehors de la salle de classe (Benson, 2009 ; Benson & Reinders, 2011). L'AIL représente de ce fait un domaine d'études spécifique en analysant ce type d'apprentissage qui se déroule au travers d'activités diverses et variées, pratiquées dans la plupart des cas pour le loisir et le plaisir, et qui aboutissent *in fine* à des acquisitions langagières.

1.3 L'apprentissage informel de l'anglais en ligne

L'appellation « apprentissage informel de l'anglais en ligne » est circonscrite dans une logique similaire à l'acception de la notion d'apprentissage informel (au sens générique), mais aussi à celle de l'apprentissage informel des langues en particulier. Ce terme est employé pour la première fois¹¹ par Denyze Toffoli et Geoffrey Sockett, lorsqu'ils ont mené en 2009 une étude sur les pratiques informelles en ligne au sujet d'étudiants spécialistes de disciplines autres que l'anglais. Depuis, la

¹¹ *Online Informal Learning of English* (OILE) était le terme original et qui a fait l'objet de la première publication.

publication de cette étude en 2010 a impulsé le coup d'envoi à de nombreuses autres recherches dans le domaine, y compris notre recherche doctorale.

1.3.1 Pour un périmètre des définitions de l'AIAL

Pour commencer, une première question attire notre attention : comment peut-on définir l'AIAL ? Pour y répondre, nous nous reposerons en conséquence sur les définitions de la littérature scientifique et verrons que ce champ de recherche présente la particularité d'être non seulement nouveau, mais également vaste, justifiant de faire appel à plusieurs points de vue sur le sujet. Ainsi, selon Sockett (2011) :

We have defined informal language learning as emerging from a communicative intention, which implies that it does not take place according to a set timetable or in the context of an organised learning activity and indeed that the learner is not necessarily aware that learning is taking place (paragr. 8).

Nous retenons en premier lieu de cette définition que l'AIAL est caractérisé par une intention communicative. Ensuite, qu'il est déterminé par la notion de non-officialité, c'est-à-dire qu'il représente un type d'apprentissage de l'anglais impliquant des pratiques choisies par les apprenants eux-mêmes. Ce dernier ne relève en effet pas d'un devoir de classe ou d'un devoir-maison, ni d'un cours ou d'une leçon, encore moins d'une quelconque activité prescrite par un enseignant. De ce fait, l'apprenant choisit son lieu, son rythme et ses activités : il n'y a pas de calendrier fixe ou encore de progression établie au préalable. Les apprentissages sont inorganisés et les activités pratiquées sont effectuées selon la volonté de l'apprenant, ses envies et ses disponibilités. Cette notion de non-programmation est cruciale car, à la différence des apprentissages formel et non-formel, l'AIAL ne suit aucun programme ou syllabus scolaire (ni même personnel). L'absence d'intentionnalité revient alors à considérer les pratiques de l'AIAL comme étant impromptues et souvent dénuées de toute conscience d'action¹². C'est un apprentissage incident (Rieder, 2003) puisque l'apprenant n'a pas conscience du fait qu'il assimile des savoirs. Ce point constitue l'une des différences majeures entre l'AIAL et l'OCLL, étant donné que cet autre domaine de l'apprentissage informel considère l'acquisition comme gouvernée par une intention d'apprendre. De surcroît, l'AIAL se réfère le plus souvent aux activités de la vie quotidienne liées au travail, à la famille ou aux loisirs. Nous soulignons par ailleurs la présence de

¹² Il peut y avoir de la conscience mais elle n'est pas orientée vers un objectif explicite d'apprentissage de la langue

l'expression « en ligne », comme indiqué dans l'acronyme, et qui se réfère à toute activité se déroulant au travers d'objets connectés à Internet.

Tableau 1 : Classification de l'AIAL, adapté de Lee (2019, p. 115).

	apprentissage formel	apprentissage formel non obligatoire	apprentissage informel	
			Extracurriculaire	AIAL
Formalité	Structuré, Certification	Structuré ; certification obligatoire	Semi-structuré ; sans certification	Non-structuré ; sans certification
Lieu	En classe	En dehors de la classe	En dehors de la classe	Partout
Pédagogie	Dirigée	Dirigée	Autodirigée	Non-dirigée
Type de contrôle	Contrôlé	Contrôlé	En autonomie	En autonomie

Dans une tentative de comparaison des types d'apprentissage en fonction de leurs caractéristiques, Lee (2019, p. 115) se repose pour sa part sur quatre facteurs fondamentaux proposés par Benson (2011), afin de mieux situer l'IDLE¹³ (synonyme de l'AIAL). Le premier facteur (indiqué dans le tableau *supra*), la formalité, se réfère ainsi à la structuration de l'apprentissage et son aboutissement à une certification. Le deuxième désigne de son côté le lieu où se déroule l'apprentissage, c'est-à-dire en classe ou en dehors de la classe. Le troisième, concerne quant à lui, la pédagogie sous-jacente à l'apprentissage, à savoir la présence d'instructions formelles et institutionnelles ou non. Le dernier facteur nommé « locus de contrôle », est similaire à la pédagogie mais se concentre en particulier sur le degré d'autonomisation de l'apprentissage. Ceci traduit ainsi l'idée que la manière de travailler de l'apprenant change en fonction de la situation : il est autonome dans les contextes informels alors qu'il est « contrôlé » dans le cadre formel (*cf.* tableau 1 ci-dessus). Cette représentation des formes d'apprentissage nous aide dès lors à mieux comprendre le terrain de l'AIAL, même si certains aspects méritent d'être reconsidérés ou redéfinis, comme le précise Lee (2019, p. 115) : « *Although Benson himself has admitted that this is "a somewhat rudimentary framework" (p. 15), this conceptualization nevertheless helps us understand English learning in diverse IDLE contexts* ». En effet, nous nous posons des questions auxquelles le tableau ci-dessus ne nous permet pas de répondre. Si l'on prend par exemple le cas du lieu, dans quelle catégorie pourrait en conséquence se placer une situation où un apprenant regarde un épisode d'une série télévisée pendant une pause en classe ? Pour nous ce contexte relève d'une

¹³ *Informal Digital Learning of English.*

activité rassemblant les spécificités de l'extramural à l'exception du contexte (même si, par définitions, l'AIAL peut se dérouler en tout lieu, notamment grâce aux technologies mobiles). C'est pour cette raison que nous avons remplacé dans la colonne AIAL, l'expression « en dehors de la classe » par « partout » dans le tableau ci-dessus.

Aussi, faut-il rajouter le fait que le champ de l'AIAL considère les apprenants comme des utilisateurs¹⁴ de la langue ayant pour objectif de communiquer, et non l'apprentissage explicite de la langue. Ces usagers sont souvent confrontés à diverses situations-problèmes qui prennent la forme de tâches devant être accomplies à des fins bien identifiées (*cf.* CECRL, Conseil de l'Europe, 2001). Cette perspective actionnelle constitue l'une des forces de l'AIAL car elle entraîne une démarche active et de recherche chez les apprenants. Leur statut de locuteurs non natifs n'est ainsi pas considéré comme un frein à leur participation à des communautés de pratiques autour des thèmes qui les intéressent à titre personnel. La réutilisation de structures auxquelles ils sont exposés se trouve donc être analogue à celle des locuteurs natifs (Sockett et Kusyk 2013, paragr. 42). Dans cette perspective, l'apprentissage linguistique est plutôt assimilable à un développement langagier puisqu'il est dynamique, non-linéaire, et résulte de l'interconnexion et de l'interaction entre une multitude de facteurs. Par ce dernier point, il est important de signaler que l'AIAL s'inscrit d'office dans le courant épistémologique constructiviste et socioconstructiviste issu de la psychologie de l'apprentissage (Piaget, 1962, Vygotsky, 1978). En effet, les personnes impliquées dans les activités de l'AIAL sont en permanence en train de construire et de déconstruire les savoirs et savoir-faire, au moyen d'interactions avec d'autres utilisateurs (par exemple les réseaux sociaux).

Ayant présenté les définitions de l'apprentissage informel en ligne et les composantes qui permettent de l'appréhender, nous allons à présent proposer notre propre définition de cette notion. En se fondant sur les informations issues du tableau ci-dessus ainsi que sur les définitions de Toffoli et Sockett (2015), nous considérons l'apprentissage informel de l'anglais en ligne comme étant toute activité pratiquée sur Internet au travers de laquelle un apprenant s'approprie des éléments langagiers (de l'anglais) de manière incidente ou consciente. Ces activités se pratiquent en autonomie, sans structuration ou direction, sans limite liée au lieu et ne vise aucune

¹⁴ Qui n'apprennent pas pour apprendre mais qui utilisent la langue dans une visée bien spécifiée. Par exemple : chercher des lignes de scripts pour coder son profil en ligne sur un blogue ou chercher la signification d'un mot afin de comprendre un passage d'une série en anglais...

certification. En revanche, si nous prenons en compte la différence entre apprentissage et acquisition (Krashen, 1976), et si nous considérons que la finalité ultime de ces deux processus devrait être l'appropriation¹⁵ d'éléments langagiers, permettant aux apprenants de poursuivre leur objectif de communication, nous préférons l'expression « Appropriation Informelle de l'Anglais en Ligne ». En effet, selon Krashen et Terrel (1983), l'apprentissage serait un processus conscient gouverné par des règles scolaires (grammaticales) alors que l'acquisition serait inconsciente, naturelle et libre. Le rapprochement de ces deux terminologies avec celles du formel et de l'informel nous fait alors dire que l'acquisition peut être considérée comme de l'apprentissage en contexte informel. Dans ce cas, nous pourrions également utiliser l'expression « Acquisition de l'Anglais en Ligne » au lieu de « Apprentissage Informel de l'Anglais en Ligne » puisqu'en principe, la rencontre entre l'acquisition et l'apprentissage abouti à l'appropriation. Mais nous souhaitons rappeler que cette dernière peut aussi être le produit de pratiques informelles ou formelles. C'est donc pour expliciter cette différence que nous adoptons l'expression « Appropriation Informelle » par opposition à « Appropriation Formelle ». Aussi préférons-nous le terme « en ligne » pour se référer à toutes les activités nécessitant une connexion à Internet : streaming ou téléchargement de vidéos, lecture ou jeux vidéo (Schmoll, 2017). Pour nous, le mot numérique (*digital*) que l'on retrouve dans l'acronyme *IDLE* serait trop restrictif puisque certaines ressources numériques n'ont pas nécessairement besoin d'une connexion Internet (DVD, consoles de jeux, radio, télévision, etc.). Même si nous aurions préféré l'expression « Appropriation Informelle de l'Anglais en Ligne », nous conservons l'utilisation de l'appellation « Apprentissage Informel de l'Anglais en Ligne » pour des raisons de cohérence.

1.3.2 Les objectifs de l'AIAL

S' « il est largement admis que l'école n'est plus le seul lieu légitime de la diffusion des savoirs culturels, scientifiques et technologiques » et que « le seul milieu scolaire ne détient plus le monopole de l'éducation » (Meunier, Belleville & Grant, 2018, paragr. 4), alors nous assistons à un changement de paradigme sur le plan éducatif en général et de l'enseignement en particulier. En

15 L'appropriation est pour nous un stade au-dessus de l'acquisition et de l'apprentissage qui sont des processus évolutifs. En d'autres termes, l'appropriation est une assimilation/maitrise presque parfaite d'éléments langagiers spécifiques (prononciation, grammaire, lexique, orthographe). On peut également utiliser les mots assimilation, accommodation ou acclimatation/acclimatement.

effet, « l'école fait partie d'une société qui laisse dorénavant de plus en plus de place à d'autres milieux, contextes et lieux d'éducation, dits non formels » (*ibid.*). De ce fait, les failles et échecs du système scolaire, notamment dans le domaine de l'apprentissage des langues et les atouts de l'apprentissage informel inciteraient des chercheurs à considérer un mélange de ces sphères d'apprentissage, voire un basculement vers une approche purement informelle. Pendant que certains auteurs cherchent comment circonscrire ces trois types d'apprentissage, parce que complémentaires (Meunier *et al.*, 2018), d'autres pensent à la reconnaissance de l'apprentissage non formel et informel au même titre que le formel, comme le souligne Werquin : « Une meilleure intégration de la reconnaissance des apprentissages non formels et informels dans les cadres de certification existants renforcerait aussi sa place dans une stratégie cohérente et globale de formation tout au long de la vie » (Werquin, 2010, p. 11).

Ainsi, le champ d'étude de l'AIAL s'installant au sein de la didactique des langues, permet de réfléchir sur les moyens et les conditions d'intégrations et d'exploitation de ces pratiques, en direction d'apprenants et d'enseignants d'anglais mais aussi de décideurs. Les résultats escomptés de ces manœuvres sont l'amélioration des conditions d'apprentissage des acteurs, en s'adaptant au rôle changeant de l'enseignant de langues et d'apporter un meilleur accompagnement des apprenants. Tels sont les objectifs que se fixe le champ de l'AIAL, car : *It is therefore important to continue to measure and analyse these changes in order to get a better grasp of the learner motivations so as to organise the formal teaching of English by making the best of them* (Socket, 2011, paragr. 1). En termes d'applications pédagogiques, des réflexions sont nécessaires en vue d'analyser comment l'accès à l'anglais par Internet peut être utilisé comme outil en classe et quels peuvent être les effets sur le développement des compétences linguistiques (Kusyk et Sockett, 2012, paragr.75). Cependant, il est important de reconnaître les limites d'une telle approche et d'en être conscient. Ainsi, le statut de l'apprenant en tant qu'utilisateur de la langue peut représenter à la fois un danger et un potentiel pour la salle de classe, car l'apprenant informel a de réelles compétences à apporter au cours mais peut être confronté à des difficultés nouvelles. Par exemple, le risque de transposition du registre familial vers celui du registre académique et formel ou encore le désintérêt pour les tâches formelles (Sockett et Kusyk 2013, paragr. 42). Il peut également y avoir des problèmes d'évaluation de ces activités ou d'intrusion dans la vie privée des apprenants. En somme, le champ de l'AIAL semble se donner pour objectifs généraux le recensement des

pratiques informelles, la compréhension de celles-ci, l'étude de leurs impacts sur l'acquisition langagière en général et les moyens de leur exploitation dans la salle de classe au bénéfice des apprenants. Dans la section qui suit, nous discuterons de quelques études importantes dans le domaine parce qu'elles nous aident dans notre problématisation mais aussi, elles contribuent à une meilleure compréhension de notre sujet.

1.3.3 AIAL en France et ailleurs : quelques études

En France, la première recherche dans le cadre de l'AIAL (selon les définitions données ci-dessus) date de 2009 et est menée par Toffoli et Sockett (2010). Les auteurs ont inventorié les différentes activités pratiquées et la fréquence de ces dernières auprès de 222 étudiants spécialistes d'autres disciplines que l'anglais, de l'Université de Strasbourg. Ils ont trouvé que la compréhension orale et écrite, la production écrite et l'interaction écrite constituent les activités les plus pratiquées en ligne en anglais. La production et l'interaction orale sont rarement citées, malgré la disponibilité des technologies VOIP¹⁶. En effet, 90 % des sondés sont régulièrement impliqués dans la CO (au moins une fois par mois), contre 25 % pour la CE. Ces activités comprennent en majorité le visionnage de vidéos et l'écoute de la musique en anglais. D'autres recherches ont été menées à la suite de celle de 2010 et ont permis d'explorer davantage le champ de l'AIAL. Elles ont surtout montré l'étendue du phénomène et comment ces activités sont pratiquées (temps et modalités d'exposition), l'acquisition de *chunks* provenant des séries télévisées, la perception d'enseignants d'anglais par rapport à ces pratiques (Sockett, 2011 ; Kussyk & Sockett, 2012 ; Toffoli & Sockett, 2015 ; Sockett, 2014 ; Kussyk, 2017). En dehors de la France, des études ont également analysé les pratiques informelles des apprenants d'anglais. On retrouve des travaux sur : l'analyse des pratiques d'étudiants à travers les smartphones en Slovénie (Jurkovič, 2019) ; les stratégies des lycéens dans les tâches de compréhension écrite en Norvège (Brevik & Hellekjær, 2018) ; les pratiques de fan-fiction, de fan-traduction et de fan-doublage en Russie et en Espagne (Shafirova et Cassany, 2019). Certaines ont montré le lien entre l'acquisition du vocabulaire anglais et les pratiques de jeux vidéo des adolescents suédois (Sundqvist, 2019) ; les différences de performances linguistiques entre les apprenants « scolaires » et les apprenants « naturels » au Brésil (Cole & Vanderplank, 2016). D'autres aussi examinent la perception

¹⁶ Voice Over Internet Protocol.

d'étudiants autrichiens concernant l'utilité des pratiques de l'AIAL dans la salle de classe (Trinder, 2017) ; les rapports entre l'acquisition linguistique, la quantité et la diversité des activités de l'AIAL en Corée du sud (Lee & Lee, 2019) ; l'acquisition de l'italien par des tunisiens à travers des programmes diffusés sur la chaîne italienne *Rai Uno* (Narcy-Combes et Boughnim, 2011). On pourrait également ajouter d'autres pays dont la Malaisie (Tan, Ng, & Saw, 2010), l'Inde (Mitra, Tooley, Inamdar, & Dixon, 2003), la Chine (Sun, Franklin, & Gao, 2017) et le Japon (Casanave, 2012). Ces études¹⁷ ont montré que les apprenants s'exposent à des pratiques informelles variées qui résultent en l'acquisition incidente du vocabulaire et de la grammaire, à travers des facteurs comme la motivation et le désir de communiquer, la fréquence d'exposition, l'attention et les stratégies cognitives.

En revanche, la question de l'influence de l'AIAL sur la prononciation des apprenants a été soulevée à maintes occasions, mais aucun de ces travaux n'est consacré à l'analyse phonétique des productions orales des sondés. En effet, même si les travaux de Narcy-Combes et Boughnim (2011), Cole & Vanderplank (2016), de Kusyk (2017) et de Lee (2017) ont enregistré des productions orales d'apprenants pendant la collecte des données, celles-ci n'ont pas fait l'objet d'analyses phonétiques ou phonologiques ; seuls les aspects sur l'expression orale, c'est-à-dire l'aisance à communiquer, ont été exploités. Ces aspects de prononciation qui manquent dans les recherches citées jusqu'à présent ont toutefois été traités ailleurs et dans d'autres contextes, c'est ce que nous présentons dans la section suivante. Il s'agira de discuter des impacts de l'exposition aux médias sur l'acquisition d'éléments de prononciation, non seulement chez les natifs mais aussi chez les non-natifs.

1.3.4 L'influence des médias sur la prononciation

Eu égard à l'influence des médias sur la prononciation, certains chercheurs considèrent l'acquisition phonético-phonologique comme étant le pur produit d'interactions *in praesentia* et restent sceptiques voire réfutent tout impact des médias, qualifiés de sources non-interactives, sur l'acquisition phonétique (Trudgill, 1986 ; Chambers, 1998 ; Labov, 2001). D'ailleurs, l'impact des interactions réelles (physiques) sur l'apprentissage phonologique a été considérablement

¹⁷ Pour plus de détails, voir le numéro 1 du volume 23 de *Language Learning & Technology*, 2019 : <https://scholarspace.manoa.hawaii.edu/handle/10125/58892> (consulté le 23/04/2019).

documenté et il a été montré, par exemple, que lors d'interactions directes, les locuteurs s'imitent mutuellement sur les plans gestuel, lexical, syntaxique, rythmique, phonétique, phonologique, etc. (Schmale *et al.*, 2015). Nous discutons ce point dans le chapitre sur l'imitation (section 4.1, page. 103). Ainsi dans cette lignée, il y a environ deux décennies, Chambers (1998) a publié un article intitulé « *TV makes people sound the same* », dans lequel il voulait montrer que la télévision n'influence pas l'acquisition phonético-phonologique :

[...] we are perpetually surprised to find that very often the people we are discussing these linguistic changes with – our students, colleagues in other departments, audiences at lectures, newspaper reporters, dinner-party guests – know exactly why the changes are taking place. It's because of television, they say. It's the mass media – the movies and the radio, but especially television. Television is the primary hypothesis for the motivation of any sound change for everyone, it seems, except the sociolinguists studying it. The sociolinguists see some evidence for the mass media playing a role in the spread of vocabulary items. But the deeper reaches of language change – sound changes and grammatical changes – the media have no significant effect at all (Chambers, 1998, p. 124).

Ainsi Chambers soutient l'idée selon laquelle les changements phonologiques ou d'accents, au sein d'une communauté d'individus, sont la conséquence exclusive d'interactions humaines, et que la télévision n'impacte pas la prononciation des téléspectateurs. Pourtant, l'influence des médias, notamment de la télévision sur l'acquisition d'un répertoire lexical, est largement admise : « *unlike sound changes and grammatical changes, these lexical changes based on media are akin to affections* » (p. 126). Il poursuit en affirmant que « *The fact that these language changes are spreading at the same historical moment as the globalization of mass media should not be construed as cause and effect. It may be that the media diffuse tolerance towards accent and dialect. [...] But the changes themselves must be conveyed in face-to-face interactions among peers* » (p. 129). Pour l'auteur, la télévision faciliterait le discernement et la compréhension d'autres types d'accents, mais ne permet pas à un téléspectateur d'adopter les accents auxquels il est exposé. Aussi, Chambers donne plusieurs exemples issus de recherches empiriques sur lesquels il se fonde pour réfuter l'impact des médias sur la prononciation, ce qu'il appelle « *linguistic science-fiction* » (p. 125). Parmi ces évidences, il présente le fait que la communauté afro-américaine ne parle jamais comme leurs homologues blancs malgré une forte consommation des médias (p. 126), ou encore un enfant (Vincent) issu d'une famille de sourds, qui n'a jamais réussi à parler comme à la télévision, à laquelle il a été régulièrement exposé par les parents (p. 127).

Nonobstant, contrairement aux arguments avancés ci-dessus, nombre de recherches attestent un impact positif de la télévision sur l'accent. Ainsi, les résultats de Mitterer & McQueen (2009), sur la perception de mots anglais prononcés avec des accents australiens et écossais par des apprenants néerlandais, montrent la disposition des apprenants à imiter des accents à partir de sources enregistrées non-interactives. Sanchez, Miller & Rosenblum (2010) ont également trouvé que les sources audio-visuelles ont influencé le VOT (négatif)¹⁸ des consonnes occlusives sonores des participants dans deux tests d'imitation phonétique. Ensuite, Stuart-Smith *et al.* (2007 ; 2011) ont travaillé sur deux groupes de locuteurs écossais ayant effectué des tests d'imitation (perception-catégorisation, production) dans deux conditions différentes. Le groupe contrôle était en interaction directe avec un locuteur natif du Standard British English (SBE), tandis que le groupe témoin devait regarder la vidéo de la séance du premier groupe. Dans les deux conditions, tous les participants ont révélé une imitation phonétique, malgré quelques petites différences (elle était plus marquée chez le groupe en interaction directe). Ils en concluent que l'imitation phonétique, par le biais des médias, est bien possible, mais qu'elle est moindre qu'en interaction directe :

Sometimes experiencing speech without the possibility for interaction, e.g. engaging with a favourite Television show, is related to language change. [...] It is also clear that many kinds of rapid perceptual learning are possible from non-interactive recorded speech. Exposure to non-interactive recorded speech can trigger short-term shifts in speech perception and production. Speakers can also show short-term shifting even towards a virtual interlocutor. (Stuart-Smith et al., 2011, p. 1914).

Dans une autre étude, Stuart-Smith et ses collègues (2013) ont attesté que des visionnages actifs et motivés de séries télévisées accélèrent le changement d'accent dans une langue. Ils ont trouvé que non seulement les jeunes de Glasgow (Écosse) qui regardent régulièrement la série *East Enders* s'attachent émotionnellement aux personnages de la série, mais également qu'ils s'approprient des éléments de l'accent londonien (le *Cockney*). Dans cette étude, deux facteurs de prononciation ont été mis en évidence. Les locuteurs de Glasgow commencent à utiliser le /f/ et le /v/ à place de la fricative dentale sourde /θ/ et fricative dentale sonore /ð/ respectivement (think : /fɪŋk/ au lieu de /θɪŋk/ ; this : /vɪs/ au lieu de /ðɪs/) et à remplacer l'approximante latérale /l/ par la voyelle

¹⁸ Voice Onset Time négatif (le délai d'établissement du voisement) des consonnes occlusives sonores représente l'intervalle de temps entre le début des pulsations périodiques régulières glottique de la consonne et le relâchement de l'occlusion supra-glottique de cette consonne (Lisker et Abramson, 1964).

postérieure fermée /ʊ/ du mot *good* (milk : /mʊk/ ou lieu de /milk/). Stuart-Smith et Ota (2014) ont rapporté des impacts de médias, en particulier de la télévision, sur le changement linguistique en cours dans quelques villes japonaises. Ota et Takano (2014) ont aussi repéré une similitude entre le système intonatif (plat) de deux dialectes distincts parlés dans deux villes géographiquement éloignées (Hokkaido à l'extrême nord et Kagoshima à l'extrême sud)¹⁹, et celui du japonais parlé dans les médias, qui provient de la capitale Tokyo, au centre (*op.cit.*: 143-145). Ils ont également trouvé que les jeunes de ces régions s'identifient aux personnages des dessins animés, les imitent (y compris pour les aspects phonético-phonologiques) et y trouvent une « signification sociale » (*op.cit.* : 151). Toutefois, les auteurs reconnaissent que les séries télévisées ne constituent guère le seul facteur responsable du changement d'accent : elles ne représenteraient qu'un facteur parmi d'autres. Le débat quant à l'influence des médias sur les changements linguistiques continue de susciter de vives réactions entre sociolinguistes. Nguyen et Delvaux (2015) affirment à la lumière des théories de l'imitation phonétique que les moyens audiovisuels favorisent aussi l'imitation phonétique et faciliteraient ainsi l'acquisition de la prononciation, alors qu'Androutsopoulos (2014) constate que :

Despite a fair amount of relevant scholarship in sociolinguistics and neighbouring fields, the role of media in processes of linguistic change is not yet fully understood. [...] The predominant position in variationist sociolinguistics, i.e. that the media play no role in systemic language change [...] is increasingly perceived as unsatisfactory (Androutsopoulos, 2014, p. 3).

Dans la même lancée des travaux sur l'impact des médias sur la prononciation, Maeda (2009) a conduit une étude en vue de savoir quelle variété d'anglais les apprenants japonais préfèrent apprendre et quelle sont celle(s) qui influencent réellement leur prononciation. Elle a enregistré 65 étudiants de deux universités de la région de Kanto (Kanagawa et Tokyo), pendant la lecture de courtes histoires et a fait passer un questionnaire sur leur exposition et attitudes face aux variétés d'anglais. Des phonéticiens (américains, britanniques et japonais) ont jugé les caractéristiques des participants, au regard des variétés d'anglais américain, britannique, australien, japonais ou mixte. Une majorité d'élèves voulait apprendre la prononciation de l'anglais américain, mais une grande minorité désirait apprendre l'anglais britannique, et seulement un petit nombre voulait être facilement et simplement compris. Les résultats indiquent que la prononciation de la plupart des

¹⁹ Les auteurs pensent qu'il est difficile que les habitants de ces régions aient des contacts réguliers (Ota et Takano, 2014, p. 144).

élèves est influencée beaucoup plus par l'anglais américain que britannique. Certains mots sont plus prononcés avec des éléments de la RP (*new*) alors que d'autres ont été confondus (*tune* est prononcé comme *turn*). Des influences du japonais et quelques traces de l'anglais australien ont été également notées.

Navrátilová (2013) a examiné la prononciation des élèves tchèques du second degré, qui ont été enregistrés pendant la lecture de dialogues, afin d'explorer les influences de l'exposition à l'*input* anglais (visionnage et écoute) sur leur accent. Selon l'auteure, en république tchèque, la majorité des films diffusés sur les écrans sont américains et les étudiants des écoles primaires et secondaires sont exposés à la langue anglaise, non seulement en classe mais aussi pendant leur temps libre, durant lesquels ils regardent la télévision ou écoutent de la musique. De plus, ils seraient plus exposés aux séries américaines que britanniques et regardent les séries exclusivement en VO²⁰ grâce à un niveau suffisant leur permettant de comprendre sans les sous-titres. Les sujets ont été soumis à un test de perception par des professionnels en vue de catégoriser les accents des sondés. Les résultats ont montré que les participants prononcent avec un mélange des deux accents (RP et GA) mais avec toutefois une légère prédominance de l'accent américain. Ce mélange de prononciation est également remarqué au sein des mots contenant les mêmes phonèmes cibles. Par exemple, *somebody* et *lot* sont prononcés à la britannique tandis que *God* est prononcé à l'américaine et les auteurs attribuent ces résultats à la notion de fréquence d'exposition, car *God* serait plus fréquent dans les séries. Les résultats du questionnaire évoquent également une exposition à environ 45 minutes de cours d'anglais (par semaine) contre environ 150 heures en dehors de la classe.

Rindal et Piercy, (2013) examinent la prononciation de l'anglais chez les adolescents norvégiens et trouvent après analyse auditive, une utilisation mixte de caractéristiques linguistiques de l'anglais américain et de l'anglais britannique, et de quelques éléments de prononciation formant un accent L2 hybride et variable. Cependant, l'anglais américain est la prononciation dominante et les auteures suggèrent une influence des médias américains. Bien que la majorité des apprenants visent un accent natif, une minorité déclare vouloir éviter les accents natifs et utiliser une variété neutre d'anglais. Elles soulignent néanmoins une forte variabilité dans la prononciation des

²⁰ Précisons que la république Tchèque fait partie des pays qui ne pratiquent pas le doublage notamment en raison du coût financier.

participants, qui contribueraient à la diversité croissante dans le développement de l'anglais comme langue mondiale. Selon les auteurs, ces résultats devraient encourager les enseignants d'anglais en vue de répondre aux besoins des apprenants.

À travers ces quelques exemples d'études, nous voyons des influences des médias sur la prononciation à divers degrés, malgré les hypothèses de Chambers (1998). Il est sûr qu'à ce niveau, ces études s'accordent et révèlent des résultats cohérents. Se pourrait-il que depuis 1998, il y ait eu des évolutions linguistiques considérables ? Avec l'avancée des recherches en AIAL, notamment son impact sur la prononciation, nous devrions assister sous peu à des résultats qui confirment ou infirment les postulats avancés par les chercheurs. Les méthodologies et mesures adaptées à ce type d'hypothèses nous éclaireront sur les éventuelles influences des médias sur la prononciation. Ce point nous amène à la présentation des objectifs de notre étude.

1.4 Objectifs de cette étude

Cette recherche trouve ses racines dans deux sources. Tout d'abord, suite à une présentation des résultats de l'étude de 2009 (Toffoli et Sockett, 2010), des enseignants ont fait part de leur perception de la prononciation de l'anglais de leurs apprenants, qui tendrait de plus en plus vers l'accent américain. Cette perception est le fondement pour une recherche qui se donne pour objectif d'exploiter le lien entre la prononciation des apprenants et leurs pratiques informelles. Un nombre d'enseignants d'anglais en France et au-delà partagent leurs impressions et expriment leur fascination pour ce phénomène. Des remarques telles « *[students] use current idiomatic expressions with appropriate pronunciation* » et « *demonstrate better pronunciation in general* » sont des exemples exprimés par certains enseignants (Toffoli & Sockett, 2015, p. 14). Ainsi, le besoin de conduire une recherche explorant plus précisément ce phénomène s'est imposé dans les études qui ont suivi celle de 2009, comme nous pouvons le remarquer dans les passages suivants : « *In the future, the effect of OILE on oral production, which is attested by some teachers [...], should also be investigated empirically* » (Kusyk et Sockett, 2012, p.16), ou « nous savons encore peu de choses sur bien des aspects du phénomène AIAL, comme son impact sur l'expression orale » (Sockett & Kusyk, 2013, p. 89). Dans d'autres cadres de recherche, certains auteurs ont également formulé des impressions semblables quant au lien entre la prononciation des étudiants aujourd'hui et les médias. Après avoir réalisé des analyses acoustiques et perceptives des voyelles postérieures de

l'anglais britannique chez deux locuteurs francophones, Exare (2009) postule que « l'exposition massive de nos jeunes étudiants (à travers le cinéma et la chanson notamment) à des variétés d'anglais comme l'anglais américain (rhotique) a certainement une incidence sur cette propension aux réalisations rhotiques » (Exare, 2009, paragr. 27). Dans la même logique de ces hypothèses, Rindal (2010) a examiné sur la production et la perception de l'anglais américain et britannique des apprenants norvégiens et pense que :

It is difficult to avoid the impression that learners' pronunciation is influenced by spoken media, seeing as there is limited access to AmE²¹ elsewhere. More research is needed to investigate this suggested link between spoken media and L2 pronunciation (Rindal, 2010, p. 256).

De plus, comme nous venons de le présenter dans les paragraphes précédents, l'influence des médias sur la prononciation a été montrée dans d'autres pays et d'autres contextes, mais pas encore en France. En effet, si l'AIAL permet d'acquérir du vocabulaire et de la grammaire, il resterait à montrer son lien avec la prononciation. Le choix de notre sujet de thèse émane donc de ces remarques et observations.

À la lumière des éléments présentés *supra*, cette thèse a pour objectif de répondre aux questions relatives à l'influence de l'AIAL sur la prononciation, vérifier les remarques et hypothèses des enseignants et des chercheurs, et enfin, compléter les études menées jusqu'alors dans le domaine de l'AIAL quant à ses influences sur l'acquisition langagière. Elle consiste en une étude exploratoire portant sur l'analyse de productions orales de 20 étudiants spécialistes de disciplines autres que l'anglais, et vise l'examen des différentes relations entre leurs activités informelles, surtout de visionnage de séries télévisées, et leurs prononciations. Afin de lever toute ambiguïté, voire toute comparaison avec d'autres recherches, nous tenons toutefois à préciser ce que cette thèse ne cherche pas à accomplir. Tout d'abord, elle ne se destine ni à classer les accents par ordre d'importance ni de supériorité. Ensuite, cette étude n'a pas pour objectif de mettre au point des méthodes, pratiques ou leçons pour l'amélioration de la prononciation. Aussi, nous ne cherchons pas à inciter les apprenants à regarder des séries, mais plutôt à exploiter leurs pratiques déjà existantes. Nous ne jugeons aucunement l'accent des locuteurs en termes de bonne ou de mauvaise prononciation, mais prenons ces facteurs en compte dans l'élaboration de nos questions de recherche et hypothèses que nous présentons ci-après.

²¹ American English.

1.5 Questions de recherche et Hypothèses

En se reposant sur les constats relatifs aux liens entre les pratiques de l'AIAL et la prononciation des étudiants français, nous avons formulé deux questions principales de recherche qui guident notre travail. Ces questions sont les suivantes :

1. Quelles prononciations adoptent les étudiants français qui pratiquent l'AIAL ?
2. Quelles peuvent être les origines de ces éléments de prononciation ?

Les objectifs de notre recherche nous permettent de nous interroger sur les types de prononciation que les étudiants français adoptent lorsqu'ils parlent anglais. Il s'agit plus précisément d'inventorier les caractéristiques de prononciation spécifiques de ces étudiants, en les comparant à celles de deux grands modèles d'accent les plus couramment rencontrés : l'anglais britannique (RP) et l'anglais américain (GA). Le choix de ces deux accents de référence est fondé sur le fait qu'ils sont les plus documentés et les plus connus dans le monde (*cf.* chapitre suivant). Parallèlement, la deuxième question centrale a pour objectif d'explorer la provenance des éléments de prononciation observés chez les locuteurs, en examinant leurs différents contextes de contact avec la langue anglaise.

Pour répondre à ces questions de recherche, nous avons formulé une série d'hypothèses que nous présentons ci-dessous et que nous allons soumettre à validation (ou non) à la fin de ce travail.

1. Nos locuteurs parleraient l'anglais avec des mélanges de caractéristiques de l'accent britannique, de l'accent américain et des éléments phonético-phonologiques de la langue française. En effet, ces locuteurs sont en contact avec l'anglais britannique, l'anglais américain, le français et leur L1 ;
2. Les étudiants français imiteraient des accents pour des objectifs bien spécifiques, à savoir pour des besoins de l'optimisation de leur accent.
3. Les étudiants qui sont exposés en permanence aux activités informelles, avec des *inputs* de l'accent américain, auraient plus d'éléments de prononciation de l'accent américain. Ceux qui y sont le moins exposés auraient plus d'éléments de prononciation de l'accent britannique.
4. Les apprenants conserveraient leur « accent français », malgré la forte exposition aux *inputs* multimédias anglais, du fait de leurs pratiques informelles.
5. Les apprenants qui sont les plus exposés à de l'*input* de l'anglais américain reconnaîtraient mieux les mots prononcés avec cet accent que ceux qui s'y exposent le moins.
6. Les locuteurs qui s'adonnent énormément aux activités de visionnage de séries télévisées américaines produiraient davantage les mêmes sons caractéristiques de cet accent en vitesse normale et rapide que ceux qui s'y exposent le moins.
7. Les séries télévisées constitueraient les activités dominantes, parmi les pratiques de l'AIAL de nos locuteurs, et influeraient ainsi plus l'intégration de l'accent.

1.6 Conclusion

Pour clore ce chapitre sur le contexte de notre recherche, il est essentiel de rappeler les différents points abordés jusqu'ici. Tout d'abord, nous avons commencé par la présentation et la définition des trois types d'apprentissage que l'on retrouve dans la sphère de l'éducation en général, et avons fait un tour d'horizon des débats concernant des précisions terminologiques. Nous avons vu que le choix des mots et expressions peut être problématique et en même temps porteur de sens, en fonction des auteurs, du domaine d'étude et de la langue. Ainsi, « éducation » ou « apprentissage », « apprentissage informel » ou « apprentissage en contexte informel », « l'informel » ou « le non formel », désignent des conceptions différentes et dépendent de la signification que chaque auteur leur assigne. Ce débat, nous le précisons, prend source du fait que le domaine est récent et mérite des travaux en vue de trouver un consensus pour l'emploi des différents termes et expressions. De plus, parce que le champ de l'apprentissage informel est un concept flou et mal défini, et qui manque de fondements théoriques (Coffield, 2000), nous avons opté pour l'expression « Appropriation informel de l'anglais en ligne » qui rassemble les processus d'apprentissage et d'acquisition. Par la suite, nous avons passé en revue notre terrain de recherche qui est l'apprentissage informel de l'anglais en ligne, en le situant au sein du domaine global de l'apprentissage informel de langues. Nous avons alors fourni un cadre définitoire et présenté quelques résultats quant aux pratiques informelles, notamment en ligne, et leurs impacts sur l'acquisition langagière en France mais aussi à l'étranger. Dans la partie décrivant la problématique, nous avons exposé les objectifs du travail au regard des études portant sur l'impact des médias sur la prononciation. Sur le fondement des objectifs de ce champ de recherche, nous avons formulé des questions et des hypothèses qui guident ce travail et auxquelles nous essayerons d'apporter des réponses. Ces réflexions sur l'intégration et la valorisation des pratiques de l'AIAL ont fait l'objet d'un très récent numéro²² spécial de *Language Learning and Technology* intitulé *CALL in the Digital Wilds* et montrent l'ampleur de ce phénomène, ses tenants et aboutissants, et ses perspectives d'évolution. L'expression *Digital Wilds*, que nous pouvons traduire par « jungle numérique » est un exemple parfait de métaphore qui témoigne de la complexité, de la diversité, du dynamisme et de l'émergence des pratiques en ligne. Enfin, ayant évoqué tout au long de ce

22 Voir <https://scholarspace.manoa.hawaii.edu/handle/10125/58892> (consulté le 15/05/2019).

chapitre que les apprenants sont de plus en plus nombreux à être exposés à divers types d'*inputs* de la langue anglaise, cette thèse se propose d'examiner quelques questions liées à ces expositions : Quelle est l'importance de l'anglais pour ces apprenants dans leur vie quotidienne ? À quel(s) type(s) d'accents sont-ils exposés et quelles différences existent-il entre les systèmes phonético-phonologiques de ces accents et celui du français ? Quelles difficultés et quelles répercussions sur la langue française ? Telles sont les interrogations que nous nous proposons d'examiner dans le chapitre suivant concernant la langue anglaise dans sa globalité.

2 Autour de la langue anglaise

American English magazines, books, Internet traffic, movies, television, and music, have inundated the world market and can be found in the remotest places on the planet. Other dialects of English such as Australian, British, Canadian, New Zealand and South Africa English, and even small sub-dialects like Cockney English, are being swamped by their exposure to American English and they are changing to conform (Carrier, 2014, p. 106).

Dans ce chapitre, nous esquissons une présentation de la langue anglaise dans sa globalité. Tout d'abord, nous parlerons de son importance en tant que langue internationale de premier plan, ainsi que de quelques facteurs responsables de son essor. Puis, nous aborderons la question de l'influence des langues, au travers des notions d'anglicisme (mots français d'origine anglaise) et de gallicisme (mots anglais d'origine française). Ensuite, nous passerons en revue les variétés d'anglais dans le monde, avant de mentionner les divers accents présents dans les îles britanniques et aux États-Unis. Ici nous en profiterons pour éclairer et justifier l'utilisation de quelques termes qui prêtent à confusion : l'anglais britannique/américain et la RP/le GA. Par la suite, nous nous attarderons sur les différences linguistiques entre la RP et le GA, sur les plans grammatical, lexical, orthographique et phonétique. Enfin, nous terminerons par une comparaison des systèmes phonétique et phonologique de l'anglais et du français.

2.1 L'importance de la langue anglaise

Aujourd'hui, parmi les quelques 6800 langues dénombrées dans le monde, il y en a 5 qui sont parlées par près de la moitié de la population mondiale : le chinois, l'anglais, l'espagnol, le russe et l'hindi (Hachten & Scotton, 2016). Le chinois, le russe et l'hindi viennent en tête de liste grâce au nombre de locuteurs natifs. L'anglais et l'espagnol (tout comme le français et le portugais) occupent leurs positions grâce à leurs ex-colonies, situées pour la plupart sur le continent américain, en Afrique et en Asie. Outre ces raisons historiques, l'anglais prend une plus grande importance grâce à son caractère international, une langue parlée sur tous les continents et presque dans tous les pays. Il est parlé soit comme une langue maternelle, seconde ou officielle, soit tout simplement étrangère, et ce pour diverses raisons. Svartvik et Leech (2016) en présentent quelques-unes :

The reason is not that the language is easy, beautiful or superior in linguistic qualities. Most people who want to learn it do so because they need it to function in the world at large. Young people, finding it both practical and cool, are attracted by things they can do with English, such as listening to music, watching films and surfing the web. For scientists and scholars, English is a necessity for reaching out to colleagues around the globe, publishing results from their research and taking part in international conferences. For tourists, English is the most useful tool for getting around and communicating with people all over the world (p. 2).

Les raisons qui confèrent une place si importante à la langue anglaise sont nombreuses et nous ne pouvons prétendre toutes les citer ici, mais nous en relevons quelques-unes. Tout d'abord le Royaume-Uni et les États-Unis rassemblent à eux deux les plus puissants centres financiers du monde, dans le domaine des affaires internationales et du commerce. Sur le plan de la diplomatie, de la politique et des organisations internationales, l'anglais est indéniablement la langue de travail la plus utilisée. Dans le monde du divertissement, les films, les chansons, les jeux, etc. sont en grande majorité en anglais et ce dernier constitue la langue la plus utilisée sur la toile devant le chinois (Pimienta, Prado, & Blanco, 2009). Enfin, des mots et termes anglais sont très utilisés dans d'autres langues et dans beaucoup de domaines.

D'un point de vue historique, l'anglais contemporain a intégré de nombreux éléments d'autres langues étrangères. Étant à l'origine une langue germanique, il a subi l'influence des langues nordiques et romanes, sur le plan linguistique et culturel, ce qui fait d'elle une langue relativement facile à apprendre pour les européens, puisque les apprenants y retrouvent, souvent, étymologiquement, des mots dont l'origine proviendrait de leurs langues ou d'une langue proche. Par conséquent, l'anglais est devenu aujourd'hui une langue flexible, dynamique, vivante et en perpétuelle évolution, une langue qui s'enrichit chaque jour de nouveaux mots, liés à des domaines divers de la vie quotidienne. Il intègre aussi des mots d'autres langues comme par exemple l'arabe (*alcohol, cotton, orange, saffron*), le japonais (*Origami, Kimono, Tofu, Emoji*) et l'hindi (*Avatar, bungalow, guru, karma*), etc. Cette malléabilité permet aux utilisateurs de créer à leur tour de nouveaux mots. Par exemple, les mots et expressions suivants, créés à partir d'une combinaison de deux mots, ont récemment fait leur entrée dans le célèbre dictionnaire Oxford English Dictionary (OED²³): *Selfie, Whovian, Screenager, OMG, Chillax, Bromance, Burkini, Hangry, Beer o'clock, Webisode*). C'est également l'une des langues qui a le plus de variétés parlées et écrites dans le

²³ Pour la définition, il suffit de saisir les mots/expressions concernés sur <https://www.oed.com/> (abonnement exigé pour certaines entrées) ou <https://www.lexico.com/en/definition/whovian> (gratuit).

monde. Qu'ils soient natifs ou non, ces locuteurs se retrouvent sans *a priori* dans leur manière de l'utiliser et de la parler, malgré sa diversité. Nous reviendrons plus en détail sur ce point.

Précisons également que l'anglais est la langue étrangère la plus enseignée et la plus apprise dans le monde actuellement (Hachten & Scotton, 2016, p. 75 ; Malakolunthu & Rengasamy, 2017, p. 82). Au niveau scolaire, dans les 28 pays membres de l'Union Européenne, en 2012, « l'anglais était toujours la langue étrangère la plus couramment étudiée au niveau du premier cycle du secondaire, avec 96,7% des élèves qui l'apprennent, loin devant le français (34,1%), l'allemand (22,1%) et l'espagnol (12,2%) », comme le montre la figure ci-dessous (Bourgeois & Gheoghiu, 2014, paragr. 1).

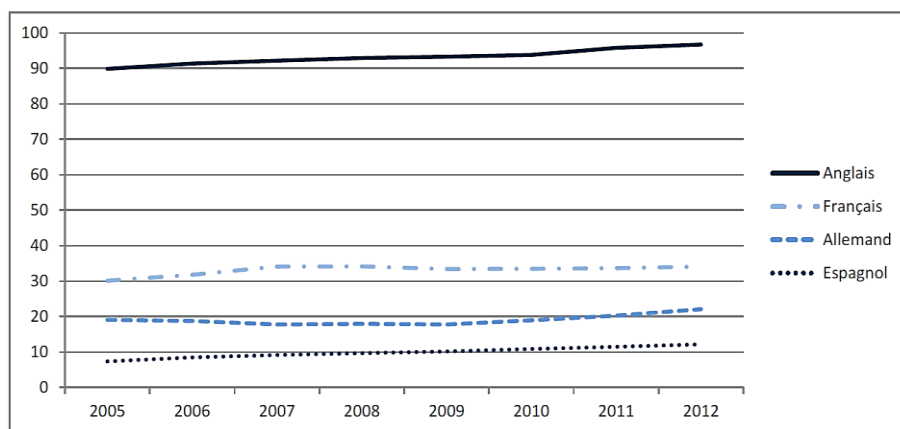


Figure 1 : apprentissage de l'anglais, du français, de l'allemand et de l'espagnol au premier cycle du secondaire dans l'UE, entre 2005 et 2012 (en %) (Bourgeois & Gheoghiu, 2014, paragr. 1).

Aussi, « l'importance de l'anglais en tant que langue étrangère dans l'UE est également confirmée par son *leadership* dans presque tous les États membres » (*ibid.*). En effet, en plus des quelques six-cent mille définitions que compte le *Oxford English Dictionary*, l'anglais est devenu la langue la plus parlée au monde grâce à sa popularité (Crystal, 2012). Plus de 370 millions de personnes ont l'anglais comme leur L1 ou leur seule langue et entre 500.000 et 1 milliard de personnes l'ont comme leur L2 (Carrier, 2014, p. 106). De ce fait, l'anglais influencerait beaucoup d'autres langues du monde comme le souligne Carrier (2014, p. 106).

Nous venons de voir que la langue anglaise est aujourd'hui considérée comme la *lingua franca* dans le monde sur plusieurs aspects. Les facteurs qui lui ont permis de se hisser à cette position d'influence sont nombreux et variés, comme explicités ci-dessus. Nous avons également vu que l'anglais s'est beaucoup enrichi de mots étrangers. Il serait intéressant de se demander s'il

continuera dans cette lancée en tant que langue la plus utilisée au monde, ou si son hégémonie prendra fin un jour. On peut aussi se demander quelle sera la langue qui le détrônera. Dans les paragraphes suivants, nous allons nous focaliser sur l'influence qu'il exerce sur le français et réciproquement, en vue de rester dans la logique de notre sujet, qui concerne les étudiants français.

2.2 Le français et l'anglais : des influences réciproques

L'hégémonie de la langue anglaise (Sockett, 2011), comme nous en avons discuté précédemment, a entraîné l'adoption de mots anglais dans bon nombre de langues dans le monde et cette évolution continue et s'amplifie dans le temps. Le français n'est pas épargné. Dans beaucoup de sphères, notamment les sciences, la technologie, l'informatique, le sport, les affaires, l'éducation, pour en nommer certains, l'utilisation de termes techniques anglais devient incontournable. Ce phénomène affecte aussi le monde médiatique, politique et de plus en plus le parler ordinaire dans la société française. Force est de noter que l'usage de ces mots anglais est ubiquiste chez les jeunes qui s'y exposent tous les jours au travers de diverses activités (jeux, musique, films, etc.), pour lesquelles les technologies de l'information et de la communication (TIC) sont les principaux vecteurs (ordinateurs portables, tablettes, téléphones portables, consoles de jeux, etc.). En Europe, jusqu'à l'adhésion le 1^{er} janvier 1973 de la Grande-Bretagne à la CEE²⁴, le français était la langue de communication dominante entre les pays membres. Mais aujourd'hui l'anglais est devenu la langue la plus utilisée en matière de communication entre personnes ne partageant pas la même langue. C'est un renversement de situation que Buffon²⁵ dénonçait avec véhémence. Dans le but de monter les influences entre l'anglais et le français, nous commencerons par les anglicismes dans la langue française, puisque c'est le premier aspect linguistique où se manifeste l'hégémonie de la langue anglaise. Nous traiterons ensuite les gallicismes en anglais.

2.2.1 L'anglicisme en français

« Anglicisme » est un terme linguistique qui réfère aux emprunts de la langue anglaise par une autre langue, notamment le français dans notre cas. Bogaards (2008, p. 20) le définit comme : « un

²⁴ Communauté Économique Européenne, ex- Union Européenne.

²⁵ Ancien secrétaire général du Comité international pour le français, langue européenne. voir Buffon (2003) au lien suivant http://www.langue-francaise.org/Articles_Dossiers/Dos_match_langue_europe.php, consulté le 13/16/2019

élément du français qui a été emprunté à l'anglais ». Il y aurait plusieurs causes (calques, mots inexistants, mauvaise traduction, styles, mode, raccourcis, etc.) et il se manifeste dans divers domaines de la linguistique (sémantique, lexical, syntaxique, morphologique, phonétique, orthotypographique, etc.). Nous ne développerons pas ces points, mais nous expliciterons deux situations : Une première où les mots anglais sont utilisés sans aucune modification (ex. *brunch*) et une deuxième où ils subissent des modifications (ex. *checker*, *liker*). Nous présentons quelques exemples de mots anglais dans leur orthographe d'origine.

Tableau 2 : Quelques exemples d'anglicismes en français²⁶.

Anglicisme (mots français empruntés à l'anglais)	
Marketing	Coach
Thriller	Flyer
Best-seller	Cutter
Gag	Check-list
Brainstorming	Week-end
(de)briefing	Baby-sitting
Live	Punch line

Les défenseurs de la langue de Molière, dont l'académie française, luttent farouchement contre ces usages et proposent des équivalents français à ces anglicismes (Walsh, 2013). En effet, c'est depuis la publication du livre « *Parlez-vous français ?* » par René Étiemble (1964), dans lequel il dénonce la menace de la langue française par l'anglo-américain, que les gouvernements successifs ont mis en place des mesures de protection pour lutter contre ce phénomène. Ainsi, Jacques Toubon, alors ministre de la culture (1993-1995) a fait voter une loi (Loi Toubon n° 94-665 du 4 août 1994 suivant la loi Bas-Lauriol) visant à protéger et à enrichir la langue française. L'objectif était de rendre obligatoire l'emploi des équivalents français pour les termes étrangers au sein des administrations et services publics :

Dans la désignation, l'offre, la présentation, le mode d'emploi ou d'utilisation, la description de l'étendue et des conditions de garantie d'un bien, d'un produit ou d'un service, ainsi que dans les factures et quittances, l'emploi de la langue française est obligatoire. Les mêmes dispositions s'appliquent à toute publicité écrite, parlée ou audiovisuelle (Article 2).

Les événements scientifiques ne sont pas épargnés :

²⁶ Pour plus de détails sur les anglicismes, consulter Bogaards (2008) et Male (2010, 2015).

Tout participant à une manifestation, un colloque ou un congrès organisé en France par des personnes physiques ou morales de nationalité française a le droit de s'exprimer en français. Les documents distribués aux participants avant et pendant la réunion pour en présenter le programme doivent être rédigés en français et peuvent comporter des traductions en une ou plusieurs langues étrangères. Lorsqu'une manifestation, un colloque ou un congrès donne lieu à la distribution aux participants de documents préparatoires ou de documents de travail, ou à la publication d'actes ou de comptes rendus de travaux, les textes ou interventions présentés en langue étrangère doivent être accompagnés au moins d'un résumé en français (Article 6).

Par contre, il y a des exceptions, comme par exemple les entreprises qui ont besoin de se vendre à l'international (Article 5). Il s'en suit des publications d'auteurs dans la même logique de protection du français, dans lesquelles des milliers d'équivalents français sont proposés en vue de remplacer les mots/termes/expressions étrangers, comme celle de Voirol dont nous soumettons ci-dessous un extrait de la quatrième de couverture :

Toutes les langues doivent quelque chose aux autres : le français a nourri l'anglais, l'anglais a enrichi le français. Mais depuis un demi-siècle, la " balance linguistique " est devenue lourdement déficitaire au détriment du français. [...] Résister à l'anglomanie, ce n'est pas se laisser aller à l'anglophobie (ou à l'américanophobie). C'est défendre le droit de nous exprimer dans notre langue, avec des mots à nous, et d'être compris de tous ceux dont le français est la langue maternelle ou d'usage (Voirol, 2006).

Tout récemment (2015), dans une lettre ouverte adressée aux Français, la secrétaire d'État chargée du Développement et de la Francophonie Annick Girardin demande de ne plus répondre « ASAP » mais « dès que possible », de ne plus « checker » les mails mais de les « vérifier » et de ne plus « débriefer » mais de faire un point²⁷. Aussi, d'autres pays « francophones » comme le Canada, ont aussi mis en place des politiques de protection et d'encadrement de la langue française contre les emprunts linguistiques étrangers plus spécifiquement l'anglicisme. La création du Grand Dictionnaire Terminologique par l'Office Québécois de la Langue Française a été (il l'est toujours) une occasion de fixer les règles d'emploi et les interdits face aux anglicismes. Sur le site du Grand Dictionnaire Terminologique²⁸, on peut retrouver les équivalents des anglicismes (plus de trois millions de termes) et on peut également les trier par domaine professionnel (plus de 200), mais aussi dans d'autres langues (italien, portugais, espagnol, roumain, galicien, catalan et le latin)²⁹.

27 Lettre téléchargeable en ligne sur : <https://www.france-canada.info/annick-girardin-denonce-les-anglicismes-du-monde-du-travail-francais/>

28 <http://www.granddictionnaire.com/> (consulté le 16/06/2019).

29 Ce site nous paraît plus pertinent et efficace pour se documenter car accessible à tous, comparativement aux versions papiers. À notre connaissance il n'existe pas d'équivalents français (de France) de cet outil en ligne.

Ces personnes parviendront-elles à sauver la langue française qui s'enrichit de plus en plus de mots anglais ? Réussiront-elles à décourager les Français dans l'utilisation des anglicismes ? Les équivalents proposés sont-ils vraiment utilisés ? Combien de Français disent « vacancelle » pour weekend, ou encore « syndrome d'épuisement professionnel » pour *burnout* ? Ce revirement semble difficile si l'on considère que naturellement, les langues s'influencent mutuellement et s'acclimatent (en termes d'acclimatation et d'acclimatement cf. Louis-Jean Calvet, 2010, 2013). Encore faut-il se rappeler que « le vocabulaire français dérive du latin à 80% » (Durand, 1999, p. 40). Même les politiques linguistiques ne pourraient inverser la situation, car nombreux sont les Français qui utilisent les anglicismes (Bogaards, 2008). Pourtant, une étude menée par Walsh (2014), montre que la majorité des Français, généralement les jeunes, perçoivent positivement l'utilisation des anglicismes. Certains auteurs pensent qu'ils constituent un enrichissement pour la langue française (Didienne, 2019). Ce phénomène touchant bon nombre de domaines professionnels en France, va à l'encontre des craintes des politiques de défense qui ont été mises en place. Tous les corps de métier sont ainsi concernés : les politiciens, les journalistes, les informaticiens, les commerciaux, les sportifs, les artisans, les enseignants et les étudiants prennent les mots ailleurs et en font les leurs, comme nous pouvons le retrouver dans la citation suivante :

Le Petit Robert définit le sens figuré d'emprunter comme un procédé en deux phases « Prendre ailleurs et faire sien ». La première phase, « prendre ailleurs », est celle qui pose problème à ceux qui mettent en question l'emploi des anglicismes en français (l'Académie française) ; la seconde phase, « faire sien », rend compte des procédés linguistiques utilisés dans la fabrique des anglicismes (Saugera, 2017, paragr. 1).

Si aujourd'hui, l'anglicisme est répandu et attesté en France comme nous venons de le voir, il serait intéressant, d'un point de vue phonétique et didactique, de voir comment ces mots, modifiés ou non, sont prononcés et utilisés par les français (Pillot-Loiseau *et al.*, 2010). Nous nous demandons si ces mots sont prononcés selon les règles de la prononciation de l'anglais ou du français et s'ils ont gardé leur sens d'origine.

Nous venons de faire un bref tour d'horizon sur l'anglicisme en français et sur les avis partagés concernant son utilisation dans la société française. Ainsi, l'anglicisme en France reste toujours en débat et divise la société (Bogaards, 2008). Comme nous l'avons évoqué, le français n'est pas la seule langue à être influencée par d'autres. Nous allons maintenant traiter la situation inverse dans le monde anglophone, c'est-à-dire le gallicisme.

2.2.2 Le gallicisme en anglais

Il est à rappeler que tout au long de son évolution, la langue anglaise a également subi des influences d'autres langues, notamment de la langue française. Ces emprunts lexicaux du français ou gallicisme impliquent l'utilisation de mots français avec ou sans modifications de formes et/ou de sens, comme le précisent Blanchard & Leven, (2007) :

It has become obvious to us that the English language is incorporating French vocabulary at an ever-increasing speed, and that, as it does so, some words are retaining their original French meanings, and some are not. Others have taken on totally new meanings and then been returned to France, where they have completely replaced the words from which they were derived (Blanchard & Leven, 2007, p. vii).

Parallèlement au cas de l'anglicisme, le gallicisme aurait plusieurs causes (calques, mots inexistants, mauvaise traduction, styles, mode, etc.) et il se manifeste dans plusieurs domaines de la linguistique (sémantique, lexical, syntaxique, morphologique, phonétique, ortho-typographique, etc.). Nous présentons dans le tableau suivant quelques exemples de mots français dans leur orthographe originelle.

Tableau 3 : Quelques exemples des gallicismes en anglais.

Gallicismes (mots anglais empruntés au français)	
Boutique	Bourgeoisie
Déjà-vu	Coup d'état
Café	Crème de la crème
Chef d'œuvre	Croissant
Foie gras	Début
Cliché	Flambé
Ballet	Sabotage

Le gallicisme remonte au 11^e siècle, lors de l'occupation normande de la grande Bretagne par Guillaume le conquérant (Male, 2010). Pendant trois siècles, la langue française a été parlée et enseignée dans la société anglaise, notamment à la cour, par les aristocrates, les nobles et les élites. Aujourd'hui encore, beaucoup de domaines comme la justice, la mode, la cuisine, l'administration, les sciences, etc. en gardent un certain nombre (*ibid.* p. 38), comme le précise Svartvik & Leech :

The greater part of the English legal vocabulary comes from the language of the conquerors. Although from 1362 English was established as the official language spoken in the court of justice, a curious mongrel, known as Law French survived, and was officially abandoned in 1731 by an Act of parliament. Many of the following words are today familiar to audiences

round the world who watch American movies or British television series focusing on legal proceedings (Svartvik & Leech, 2016, p. 38).

Les auteurs fournissent quelques exemples ci-dessous (Svartvik & Leech, 2016) :

Mostly French in origin are: Legal roles, such as advocate, attorney, bailiff, coroner, defendant, judge, jury, plaintiff. Legal actions, processes and institutions: bail, bill, decree, evidence, fine, forfeit, jail, inquest, penalty, petition, plea, proof, punishment, ransom, sentence, suit, summons, verdict. The names of crimes: arson, assault, embezzlement, felony, fraud, larceny, libel, perjury, slander, treason, trespass. French word order is preserved in attorney general, court-martial, fee-simple, heir apparent, letters patent - with the adjective following the noun. The cry Oyez! Oyez! Oyez! has probably puzzled many people, especially if pronounced 'O yes'. Used in the past when an official or town crier called for silence, it is the Anglo-Norman form of the Old French Oiez! 'Hear!' Interestingly, Oyez! is still today the opening cry of the Marshal of the United States Supreme Court, where the phrase is sounded to bring the courtroom to order (ibid.).

Nous n'avons pas connaissance des lois officielles l'interdisant, il semblerait même au contraire que le gallicisme soit bien vu et qu'il soit très apprécié dans le monde anglophone. Dans un livre intitulé « *Say Chic: A collection of French words we can't live without* », Françoise Blanchard et Jeremy Leven (Blanchard & Leven, 2007) ont recensé des termes et expressions françaises couramment utilisés en anglais américain. Les termes et expressions sont souvent accompagnés de leur prononciation (transcription phonétique à l'aide de l'alphabet latin) et de leurs significations en anglais américain, et des différences de sens entre l'anglais et le français le cas échéant. Par exemple : au naturel [oh nah char *rel*/ oh nah tyoo *rel*] (p. 1), au pair [oh *pear*] (p. 3), Bête noire [bet nwar] (p. 4), bidet [bee *day*] (p. 5), boutique [boo *teek*] (p. 6). Ces exemples nous donnent une idée de la perception des gallicismes en anglais. Ainsi, certains linguistes estiment à plusieurs milliers de mots le répertoire lexical anglais provenant du français. Il y aurait au moins 25% de mots anglais contemporain qui sont d'origine française (Svartvik & Leech, 2016 : 37). Ce chiffre est variable selon les chercheurs, leurs méthodes de calculs et les facteurs pris en considération. Ainsi, Bogaards (2008, p. 20) estime à 2500 les mots d'origine française (sans compter les modifications). Faudrait-il alors imaginer que l'on parle français plutôt qu'anglais, « cette langue faite à 70% de mots français mal prononcés et plus ou moins déformés » (Buffon, 2003) ? L'anglais parlé aujourd'hui ne serait-il en fait que du français « revisité » ? À en croire Male (2010, 2015), certains anglicismes en français étaient à une époque donnée, des gallicismes en anglais. Pour tout résumer, nous soumettons cette citation à l'appréciation du lecteur.

It has become obvious to us that the English language is incorporating French vocabulary at an ever-increasing speed, and that, as it does so, some words are retaining their original

French meanings, and some are not. Others have taken on totally new meanings and then been returned to France, where they have completely replaced the words from which they were derived, in an impressive display of foreign exchange (Blanchard & Leven, 2007, p. vii).

Nous venons de voir que l'anglais et le français s'influencent mutuellement, à différentes époques, à divers degrés et sous différentes conditions. Nous allons maintenant aborder brièvement les différents types d'anglais parlés dans le monde. Nous en profiterons pour discuter des divers accents au Royaume-Uni et aux États-Unis. Nous introduirons par la suite les deux accents qui nous intéressent et qui constituent notre objet d'étude dans ce travail.

2.3 L'anglais dans le monde : les différents types d'anglais

Il n'est pas question ici de relater l'histoire de la langue anglaise, ni de remonter à ses origines, ni de développer l'historique de sa diffusion. Pour des lectures sur ces thématiques, nous recommandons au lecteur de consulter *World Englishes* de Melchers & Shaw (2013) et *English: One Tongue, Many Voices* de (Svartvik & Leech, 2016). Néanmoins les zones géographiques (pays, villes, etc.) où il est parlé ainsi que ses différentes variétés que l'on peut rencontrer dans les quatre coins du globe nous intéressent.

D'un point de vue générique, l'anglais provient de l'Angleterre et s'est vite diffusé dans le monde, pour diverses raisons que nous venons d'aborder plus haut (colonisation, mondialisation, Internet, etc.). Aujourd'hui, la langue anglaise est parlée par plusieurs millions de personnes dans le monde, réparties dans trois grandes zones. Nous préférons ce terme à celui de « *circles* » utilisés par certains auteurs (Kachru, 1992), car le mot renvoie à l'idée de cloisonnement/enfermement, même si c'est l'idée de concentricité qui est mis en avant. Or, l'idée de « zones » illustre pour nous cette perméabilité et variabilité à laquelle la langue anglaise est soumise. Mais pour des questions de précision, nous revenons tout de même sur l'idée des cercles, telle que rencontrée dans la littérature. Selon Kachru (1992, voir Tableau 4 ci-dessous) l'anglais serait parlé dans un cercle interne (*inner circle*), un cercle extérieur (*outer circle*) et un cercle élargi (*expanding circle*). Le premier cercle contiendrait un infime nombre de pays où l'anglais est la langue première de la majorité des habitants (Royaume-Uni, États-Unis, Australie). Dans le deuxième cercle, un nombre assez important de personnes auraient l'anglais comme L2 ou langue (semi)officielle (Pakistan, Inde, Nigéria). Dans la plupart des cas, ces personnes possèdent déjà une ou plusieurs langue(s) première(s), différentes de l'anglais. Enfin le troisième cercle englobe tous les autres pays où

l'anglais est appris et utilisé comme langue étrangère (Chine, Égypte, Indonésie). Cette dernière catégorie retient notre attention du fait de la variabilité qui subsiste dans ces pays. La façon dont l'anglais est utilisé, en tant que langue étrangère varie selon les pays. Le statut de l'anglais dans les pays nordiques (Danemark, Norvège, Finlande et Suède) et en Allemagne, par exemple, n'est pas comparable au statut de cette langue dans un pays comme la France où l'anglais est moins pratiqué (Roussel & Gaonac'h, 2017). Faudrait-il revisiter ce système de classement proposé par Kachru (1992), en y ajoutant une quatrième zone ? Ceci mérite des recherches approfondies quant aux critères de classement.

Tableau 4 : les trois catégories de pays où l'anglais a un statut particulier. Adapté de Melchers & Shaw (2013, p. 8).

<i>Inner Cercle</i>	<i>Outer Cercle</i>	<i>Expanding Circle</i>
<i>Most people have English as first language: UK, USA, Caribbean, Canada, Australia, New Zealand</i>	<i>People need English for secondary education, politics, law, business inside the country: India, "Anglophone" Africa, Philippines, Singapore, Bangladesh?</i>	<i>People need English for communication in business, politics, education, etc., primarily with speakers of other languages from outside the country: most European and East Asian countries</i>

Au sein de ces trois zones on distingue une pluralité d'anglais. En raison de cette multitude de zones (cercles) où l'anglais est parlé et de cette pléiade de types d'anglais, cette langue est devenue de loin la langue la plus « hétérogène ». On pourrait également parler de dialectes ou de variétés d'anglais, car, malgré les différences dues à de nombreux facteurs, ils possèdent tous un socle commun : « l'anglais ». C'est pour cette raison que le terme de *World Englishes* est plutôt employé pour rendre compte des variétés d'anglais parlées dans le monde (Jenkins, 2003 ; Kachru, Kachru, & Nelson, 2009 ; Melchers & Shaw, 2013). Dans les sections suivantes, nous allons développer deux types d'anglais que sont l'anglais britannique et l'anglais américain.

2.3.1 L'anglais britannique / américain vs. la RP / le GA

Parmi les types d'anglais présentés ci-dessus, nous nous intéressons à deux d'entre eux en particulier : l'anglais parlé au Royaume-Uni et l'anglais parlé aux États-Unis. En effet, nous utilisons ces deux types comme accents de référence (témoins) auxquels les apprenants sont le plus exposés dans le cadre scolaire et aussi parce qu'ils sont considérés comme étant « neutres » et plus « compréhensibles » (Wells, 1982). Pour nous, le fait qu'ils soient les plus connus et documentés

nous procure une bibliographie assez riche pour la recherche. Il y a aussi la diversité au sein de ces zones, qui peut être source de richesse mais également de complexité. Pour la compréhension de la suite de ce travail, nous souhaitons apporter quelques points de précision quant aux termes que nous utilisons pour désigner les types d'accents. Par « anglais britannique » ou « anglais américain », nous entendons l'ensemble de tous les types d'anglais (dialectes et accents), parlés sur les îles britanniques et dans tous les États-Unis d'Amérique. Nous distinguons, de ce fait, l'anglais britannique de la *Received Pronunciation* (RP) et l'anglais américain du *General American* (GA). En effet, il n'est pas rare de rencontrer dans des ouvrages les termes Anglais britannique et RP, utilisés comme des synonymes. Il en est de même pour l'anglais américain et le GA. Le cas du français peut être considéré comme semblable. Par français, qu'il soit standard ou non, nous entendons l'ensemble de la langue française parlée sur tout le territoire français (y compris les départements, régions et collectivités d'outre-mer), ainsi que le français parlé dans les dizaines de pays francophones dans le monde. Mais, si nous voulons entrer dans les détails, en prenant en considérations les accents français, nous aurons autant de variétés que de zones géographiques : les accents du midi, alsacien, breton, parisien, le « Chti », ou encore les accents suisse, belge, québécois, congolais, malgache, sénégalais, etc. Remarquons que malgré toutes ces variétés, le terme « les français du monde », par référence à l'expression *World Englishes* (cf. Figure 2 ci-dessous), n'est pas encore entré dans les usages dans la sphère francophone, ou plus précisément en France. L'utilisation du terme au pluriel changerait-il la donne ? Ou plutôt, qu'apporterait-il, sur le plan politique, de dire « les français du monde » ? Dans quel sens le maintien du terme au singulier reflète-t-il le protectionnisme et l'élitisme du français parisien ? Pour l'instant nous n'avons pas de réponses à ces questions. Notre argument est de montrer simplement que la RP et le GA constituent des sous-variétés (standards) parmi tant d'autres. Le tableau ci-dessous illustre mieux le point de vue que nous défendons quant aux variétés d'anglais qui existent dans le monde. Nous développerons davantage notre propos sur trois variétés, dans ce qui suit.

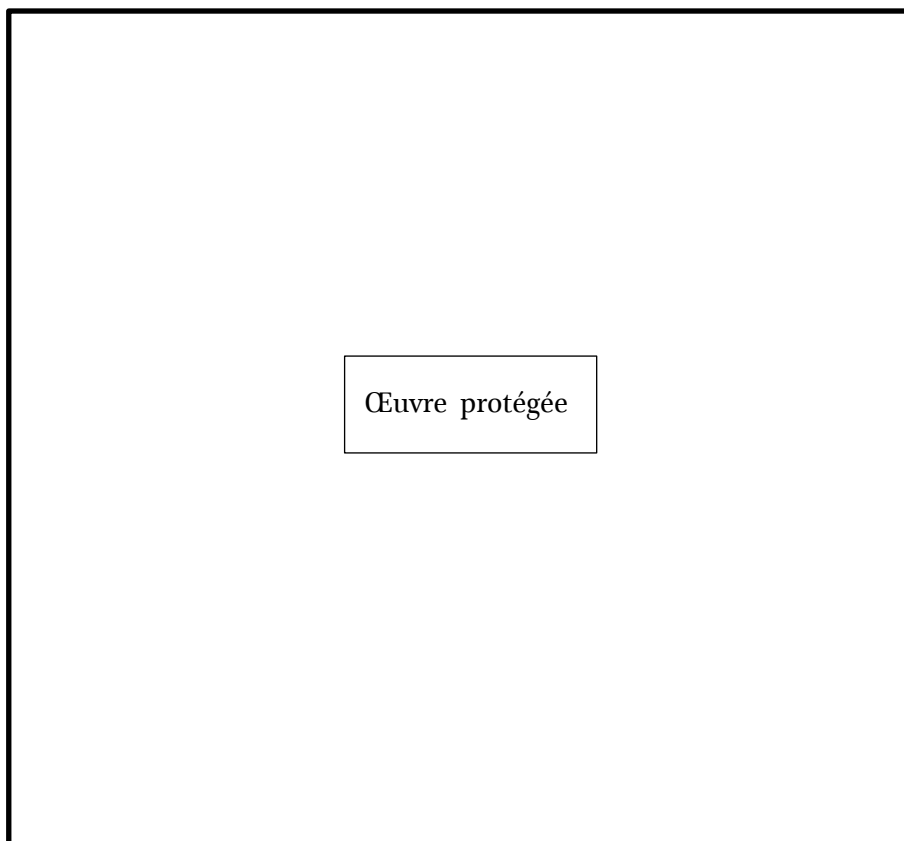


Figure 2 : le cercle des « anglais » du monde de McArthur (1998, p. 97 cité dans Jenkins, 2003, p. 18).

2.3.2 Diversité linguistique dans les îles britanniques

Les Anglais ne parlent pas tous le même anglais, contrairement à ce que l'on pourrait croire. L'anglais parlé au Royaume-Uni et en Irlande est loin d'être homogène. Dans la première édition du livre « *English accents and dialects* » (Hughes & Trudgill, 1979), les auteurs avaient traité onze (11) accents et dialectes. La toute dernière édition (5) présente, au bout de 30 ans, 30 variétés d'anglais parlées sur ces territoires (Hughes, Trudgill, & Watt, 2012). Selon les auteurs, ce chiffre ne couvre que les grandes variétés connues jusqu'à ce jour (p. vii). Nous rappelons, comme le soulignent les auteurs, que ces variétés diffèrent, tant au niveau de la prononciation, que de la grammaire et du lexique. Aussi, faut-il préciser que ces variétés portent le nom de leurs milieux (régions, villes, villages, etc.) et peuvent varier selon le statut social des locuteurs. Dans ce qui suit, nous nous contentons uniquement de présenter brièvement les aspects phonétiques notoires de deux grandes variétés qui sont la RP et le Cockney. Il nous semble importants de les présenter parce que le premier est plus connu, documenté et « neutre » alors que le deuxième est celui qui

est parlé dans la capitale, Londres, et donc plus diffusé et accessible aux étrangers. Les autres accents sont simplement mentionnés à titre d'exemple³⁰.

2.3.2.1 La Received Pronunciation (RP)

La RP est le nom assigné à l'anglais parlé aux Royaume-Uni (Wells, 1982). Cette appellation est beaucoup critiquée par certains linguistes qui se refusent à l'utiliser, car elle suppose que les autres accents ne sont pas acceptés (*received*). Mais Hughes, Trudgill et Watt (2012) assurent que « *received* » aujourd'hui devrait être compris dans le sens d'une variété qui est largement comprise par beaucoup de personnes. Nous partageons cet argument en tant que didacticien, mais emploierons tout de même le terme RP pour des raisons de simplicité et pour rester dans la lignée des travaux existants. Notons que les termes suivants, moins connotés, sont également employés : « *Standard British English* » (SBE) (Modiano, 1996, p. 11), « *Standard English English* » (SEE) (Melchers & Shaw, 2013, p. 49) ou « *Standard Southern British English* » (SSBE) (Hughes, Trudgill, & Watt, 2012, p. 3). Cet anglais dit standard est également appelé *British Broadcasting Corporation* (BBC), car il était auparavant l'accent utilisé pour la diffusion des informations à la radio et à la télévision. C'est également l'anglais de prestige qui était parlé par la classe aisée et éduquée, parmi les aristocrates et à la cour royale.

Nous traitons la RP avant les autres variétés pour plusieurs raisons. Tout d'abord, en raison de son caractère non-localisé, c'est-à-dire qui n'est associé à aucune ville ou région particulière. Il est parlé seulement par une infime minorité (environ 5%) de certaines couches sociales, à travers tout le royaume. Selon Hughes, Trudgill, & Watt (2012, p. 3), il est impossible, de s'appuyer sur la prononciation d'un locuteur RP pour dire sa provenance géographique. En même temps, ces auteurs (op.cit, p. 40-1) précisent que la RP admet aussi des variations phonétiques selon sa localisation. Par exemple, il semblerait que certains locuteurs de la RP utilisent la plosive glottale /ʔ/ à la place de /t/ dans certains contextes : /'mætə/ devient /'mæʔə/ pour le mot *matter*. Serait-ce une influence du Cockney, dont nous parlerons ci-dessous ? Ensuite, elle est la norme (modèle de prononciation) qui est utilisée dans l'enseignement de l'anglais langue étrangère en Angleterre et dans la majorité des pays étrangers qui privilégient la prononciation britannique. Un apprenant de la RP aurait donc de fortes chances d'être compris partout au Royaume-Uni, et au-delà. Pour cette

³⁰ Pour de plus amples détails sur les autres accents, consulter Brulard *et al.* (2015) et Hughes, Trudgill, & Watt (2012).

raison, il est devenu l'accent le plus étudié et documenté de tous. Hughes, Trudgill, & Watt (2012, p. 4) expliquent que : « *Oddly, and misguidedly, many people believe it to be the accent that is closest to the standard written form of English, as though the connection between spelling and pronunciation were somehow more direct for RP than for other accents* ». Par conséquent, la RP constitue la prononciation de référence à laquelle tous les autres accents sont comparés. Nous allons procéder de même : les variétés que nous citons ci-dessus reposent sur la comparaison avec le système phonétique de la RP, en utilisant les symboles de l'API³¹.

Tableau 5 : Les sons de la RP, classés par catégories.

Voyelles	Monophthongues	Brèves	/ɪ/-/kɪt/ (kit) /ɛ/-/drɛs/ (dress) /æ/-/træp/ (trap) /ɒ/-/lɒt/ (lot) /ɔ/-/fɒt/ (foot) /ə/-/ə'gen/ (again) /ʌ/-/strʌt/ (strut)
		Longues	/i:/-/fli:s/ (fleece) /ɑ:/-/pɑ:m/ (palm) /ɔ:/-/fɔ:s/ (force) /u:/-/gu:s/ (goose) /ɜ:/-/nɜ:s/ (nurse)
	Diphthongues	Centrifuges	/eɪ/-/feɪs/ (face) /aɪ/-/praɪs/ (price) /aʊ/-/maʊθ/ (mouth) /ɔɪ/-/tʃɔɪs/ (choice) /əʊ/-/gəʊt/ (goat)
		Centripètes	/ɪə/-/nɪə/ (near) /ʊə/-/kjʊə/ (cure) /ɛə/-/skwɛə/ (square)
	Triphtongues ³²	/aɪə/-/'taɪə/ (tyre) /eɪə/-/'leɪə/ (layer) /aʊə/-/'taʊə/ (tower) /əʊə/-/sləʊə/ (slower)	
Consonnes	Plosives	/p/-/paɪp/ (pipe) /b/-/bentʃ/ (bench) /t/-/taʊn/ (town) /d/-/dʌk/ (duck) /k/-/kamd/ (kind)	

³¹ L'alphabet phonétique international (API) est un ensemble de symboles (lettres, signes diacritiques) mis au point par l'association phonétique internationale, utilisé dans la transcription phonétique des sons du langage parlé et de toutes les langues du monde. <https://www.internationalphoneticassociation.org/>

³² Certains spécialistes y ajoutent des triphtongues mais attirons l'attention sur le fait que cet avis n'est pas partagé par tous dans la communauté des phonéticiens. Certains considèrent que ce sont des diphthongues qui se terminent par un schwa (Hughes, Trudgill, & Watt, 2012, p. 53-54).

	/g/-/gəʊ/ (go)
Fricatives	/f/-/fɪʃ/ (fish) /v/-/'væmpaɪə/ (vampire) /s/-/sɪt/ (sit) /z/-/zen/ (zen) /θ/-/θæŋks/ (thanks) /ð/-/ðen/ (then) /h/-/haʊs/ (house)
Nasales	/n/-/nəʊz/ (nose) /m/-/maɪ/ (my) /ŋ/-/'daɪvɪŋ/ (diving)
Affriquées	/tʃ/-/tʃi:t/ (cheat) /dʒ/-/'dʒʌdʒ/ (judge)
Latérales	/l/-/leɪk/ (lake)
Semi-voyelles	/w/-/weɪv/ (wave) /j/-/'jɒd/ (jod)
Approximantes	/ɹ/-/red/ (red)

2.3.2.2 Les accents de Londres

Le cockney est connu en tant qu'accent communément parlé dans la capitale du Royaume-Uni. Toutefois, certains auteurs y ajoutent l'Estuary (Wells, 1982) et le Mid-Indian (Hughes *et al.*, 2012). Les caractéristiques de ces deux derniers ne font pas l'objet de discussion dans cette partie, bien qu'ils partagent des points communs. Nous nous concentrerons sur le Cockney dont nous présentons ici quelques aspects phonétiques.

- La deuxième voyelle de *city* et de *happy* est réalisée en /i/ et non en /ɪ/ ;
- Le son /h/ est souvent absent dans les mots (/ɒt / à la place de /hɒt/ pour le mot *hot*). Quand il est présent, il est réalisé dans des syllabes accentuées ;
- Les fricatives labio-dentales /f/ et /v/ sont réalisées à la place des dentales /θ/ et /ð/. Ainsi, *think*, *Cathleen*, *path* deviennent /fɪŋk/ / 'kæfli:n / / pa:f / (au lieu de /θæŋk/, /'kæθli:n/, / pa:θ/) et *than*, *another*, *breathe* deviennent /væn/, /ə'nʌvə/, /bri:v/ (à la place de /ðæn/, /ə'nʌðə/, /bri:ð/). Ce phénomène serait bien connu dans la prononciation de certains francophones lorsqu'ils parlent anglais. Nous pourrions nous demander si l'accent français a plutôt influencé le *Cokney* ;
- Les mots comme *paws* et *pause*, *bored* et *board* sont réalisés différemment. Pour ces derniers, la RP utilise la même voyelle /ɔ:/ (/pɔ:z/, /pɔ:z/, /bɔ:d/, /bɔ:d/) ;
- La laterale /l/ est très souvent vocalisée, c'est-à-dire réalisée comme une voyelle. Ainsi, les mots *Paul*, *mall*, *milk*, *table*, *well*, etc. sont prononcés /pɒʊ/, /mɒʊ/, /mɪɔk/, /'tæɪbʊ/, /wɛʊ/ au lieu de /pɔ:l/, /mɔ:l/, /mɪlk/, /'teɪbəl/, /wɛl/ ;
- La terminaison <-ing> est réalisée /ɪn/ ou parfois /ɪnk/ au lieu de /ɪŋ/ ('nʌfɪn/ ou /'nʌfɪnk/ pour /'nʌfɪŋ/ en RP) ;
- Les plosives sourdes /p t k/ sont beaucoup plus aspiré qu'en RP. Parfois il y a une certaine affrication pour le /t/, qui devient un /ts/ ([tsɪ] pour [thi:]) ;
- L'approximante labio-dentale /ʊ/ est produite à la place de /ɹ/ (/vɪ:d / pour /rɪ:d/) ;

- Les diphtongues /eɪ/, /əʊ/, /aɪ/ et /aʊ/ sont produites /æɪ/, /ʌɪ/, /ɑɪ/ et /æə/ respectivement.

2.3.2.3 Autres variétés répertoriées au Royaume-Uni

À côté de ces deux variétés, les suivantes sont également répertoriées : les accents de Norwich, de Bristol, de Southampton, de Pontypridd, de Walsall, de Leicester, de Bradford, de Hull, de Liverpool, de Manchester, de Middlesbrough, de Carlisle, d'Edinburgh (rhotique³³), d'Aberdeen, de Belfast, de Dublin et de Galway (Hughes *et al.*, 2012, p. 73). À noter que les caractéristiques les plus significatives concernent les oppositions vocaliques. Nous allons à présent voir de l'autre côté de l'Atlantique avec l'anglais aux États-Unis.

2.3.3 Diversité linguistique aux États-Unis

De même que dans le cas de l'anglais parlé dans les îles britanniques, il existe de nombreuses variétés linguistiques aux États-Unis. Elles sont cependant moins nombreuses et présentent une assez grande uniformité (Viel, 1981, p. 33), en comparaison avec les anglais britanniques. L'anglais américain étant en général une variété rhotique, la différence la plus notable que nous ayons pu souligner est la non-rhoticité. Cette prononciation non-rhotique est adoptée par certains états du Sud et les vieux ports de l'Est, qui ont su conserver un lien avec leurs ancêtres anglais. Parmi ces dialectes et accents, nous pouvons citer celui de New York, Boston, Philadelphie, New England, Nouvelle-Orléans, Mobile, Savannah, Charleston, etc. Ces derniers possèdent un statut particulier dans la mesure où ils partagent des caractéristiques du GA et en même temps certaines autres de la RP. Ainsi, sur le plan consonantique, ces accents sont non-rhotiques : les mots au sein desquels les r post-vocaliques apparaissent, sont prononcés comme en RP (car_/kɑː/, park_/pɑːk/, yard_/jɑːd/, etc.). En ce qui concerne les voyelles, certaines sont plutôt réalisées à la RP, notamment celles du champ lexical *BATH* (GA = /bæθ/, Boston = /bɑːθ/). La non-rhoticité a aussi entraîné la réalisation des diphtongues de la RP (*cf.* exemples de la RP). Toutefois, les tap et le yod-dropping sont présents.

³³ Un accent rhotique est un accent dans lequel le son /r/ est prononcé partout dans les mots où il se trouve. Par exemple en anglais américain qui est rhotique, le mot car est prononcé /kɑːr/ au lieu de /kɑː/ en anglais britannique qui est non-rhotique (*cf.* méthodologie pour plus de détail).

2.3.3.1 Le General American (GA)

Le GA ou le *Network English* est aux États-Unis, ce que la RP représente au Royaume-Uni. Aujourd'hui, ils auraient les mêmes statuts, mais leur histoire est tout à fait différente. La RP a une histoire bien particulière, elle a pris source dans une région géographique définie alors que le GA a émergé d'un ensemble d'accents parlés par des personnes ayant des origines différentes (les Britanniques, les Allemands, les Néerlandais, les Afro-américains, les Latino-hispaniques, etc.). Le GA est l'accent le plus parlé au centre du pays. La majorité des caractéristiques de la RP peuvent également lui être appliquées. Nous les résumons en ces quelques points :

- Une langue standard parlée sur tout le territoire américain ;
- Une norme de prononciation enseignée aux apprenants étrangers ;
- Un accent de radiodiffusion ;
- Une langue parlée par la majorité des américains ;
- Une langue connaissant aussi des variations ;
- Une langue énormément étudiée et documentée ;
- Une langue de prestige et sans connotation ;
- Une langue ne possédant pas de caractéristiques régionales marquées.

Tableau 6 : Les sons du GA.

Voyelles	Monophthongues	Brèves	/ɪ/-/kɪt/ (kit) /ɛ/-/dres/ (dress) /æ/-/træp/ (trap) /ʊ/-/fʊt/ (foot) /ə/-/ə'gen/ (again) /ʌ/-/strʌt/ (strut)
		Longues	/i:/-/fli:s/ (fleece) /ɑ:/-/pɑ:m/ (palm) /ɔ:/-/fɔ:s/ (force) /u:/-/gu:s/ (goose) /ɜ:/-/nɜ:s/ (nurse)
	Diphthongues	Centrifuges	/eɪ/-/feɪs/ (face) /aɪ/-/praɪs/ (price) /aʊ/-/maʊθ/ (mouth) /ɔɪ/-/tʃɔɪs/ (choice) /oʊ/-/goʊt/ (goat)
		Centripètes	/ɪə/-/nɪə/ (near) /ʊə/-/kjʊə/ (cure) /ɛə/-/skwɛə/ (square)
	Triphthongues ³⁴		/aɪə/-/'taɪə/ (tyre)

³⁴ Certains spécialistes y ajoutent des triphthongues mais attirons l'attention sur le fait que cet avis n'est pas partagé par tous dans la communauté des phonéticiens. Certains considèrent que ce sont des diphthongues qui se terminent avec un schwa (Hughes, Trudgill, & Watt, 2012).

		/eɪə/-/'leɪə/ (layer) /aʊə/-/'taʊə/ (tower) /əʊə/-/'sləʊə/ (slower)
Consonnes	Plosives	/p/-/'paɪp/ (pipe) /b/-/'bentʃ/ (bench) /t/-/'taʊn/ (town) /d/-/'dʌk/ (duck) /k/-/'kɑnd/ (kind) /g/-/'gəʊ/ (go)
	Fricatives	/f/-/'fɪʃ/ (fish) /v/-/'væmpaɪə/ (vampire) /s/-/'sɪt/ (sit) /z/-/'zen/ (zen) /θ/-/'θæŋks/ (thanks) /ð/-/'ðen/ (then) /h/-/'haʊs/ (house) /ɹ ³⁵ /-/'wɛr/ (where)
	Nasales	/n/-/'nəʊz/ (nose) /m/-/'maɪ/ (my) /ŋ/-/'daɪvɪŋ/ (diving)
	Affriquées	/tʃ/-/'tʃi:t/ (cheat) /dʒ/-/'dʒʌdʒ/ (judge)
	Battue	/ɹ/-/'sɪrɪ/ (city)
	Latérales	/l/-/'leɪk/ (lake)
	Semi-voyelles	/w/-/'weɪv/ (wave) /j/-/'jɒd/ (jod)
	Approximantes	/ɹ/-/'red/ (red)

Dans le tableau 4, nous pouvons remarquer que l'inventaire phonémique du GA n'est pas aussi différent de celui de la RP. Le GA possède deux consonnes de plus que la RP : le *tap* (/ɹ/) et le /ɹ/ réalisé par certains américains dans les mots commençant par wh - *when, where, which, while* et *why* (Brulard, Carr, Durand, & Navarro, 2015, p. 121-122). Sur un plan vocalique le GA ne réalise pas le /ɒ/ de /lɒt/ (*lot*). La voyelle de *GOAT* est plutôt /oʊ/. Les diphtongues centripètes (/ɪə/ de /nɪə/ (*near*), /ʊə/ de /kjʊə/ (*cure*), /eə/ de /skwɛə/ (*square*) sont inexistantes en raison de la rhoticité. En effet, en plus de l'ajout de r à la fin des voyelles et là où la RP utilise une diphtongue, le GA emploie une voyelle à laquelle un r est ajouté. *Fear, poor, square, stir* deviennent /fɪr, pʊr, skwɛr/ en RP et /fɪr, pɔ:ɹ, skwɛr/ en GA. Les triphthongues reçoivent également un r (/aɪər/ de /'taɪər/ (*tyre*), /eɪər/ de /'leɪər/ (*layer*), /aʊər/ de /'taʊər/ (*tower*) et le /oʊər/ de /sləʊər/ (*slower*).

³⁵ Fricative labio-vélaire sourde qu'on retrouve dans *where, which, while, when, why*. Le RP utilise /w/ à la place.

2.3.3.2 Autres variétés répertoriées aux États-Unis

Dans le livre *The Atlas of North American English : Phonetics, Phonology and Sound Change*, William Labov, Sharon Ash et Charles Boberg (2006) font l'inventaire des accents présents aux États-Unis d'Amérique. Nous en évoquons certains, sans trop de précision. Il s'agit de l'accent de l'ouest (San Francisco, Los Angeles, Las Vegas), de l'est (New England), des Mid-Atlantic States (New York, Baltimore, Philadelphie), du Midland (Columbus, Kansas city, Pittsburgh, Cincinnati et St. Louis), du sud (Charleston, New Orleans, Atlanta) et du nord (Chicago, Detroit, Syracuse). Nous avons déjà mentionné plus haut qu'outre certains accents non-rhotiques (New York, Boston, Philadelphie, New England, Nouvelle-Orléans, Mobile, Savannah, Charleston), les locuteurs des autres états ont un accent rhotique et parlent de manière pratiquement homogène, se rapprochant du GA.

Nous venons de voir les différentes variétés qui sont présentes dans les îles britanniques et aux États-Unis, avec leurs spécificités sur le plan phonétique (les anglais britanniques et les anglais américains). Avant cela, nous en avons profité pour éclairer et justifier l'utilisation des quelques termes dans cette portion de notre travail, notamment la distinction entre l'anglais britannique/américain et la RP/ le GA. Ces derniers, nous l'avons souligné, représentent des standards de prononciation pour la plupart des populations et également une norme sur laquelle se repose l'enseignement/apprentissages de l'anglais L2. Quel serait le sentiment des apprenants d'anglais face à cette multitude de variétés ? En sont-ils conscients ? Si oui, sur quels fondements choisissent-ils la variété qu'ils souhaiteraient apprendre ? Et sinon, comment réagissent-ils lorsqu'ils le découvrent ? Telles sont les questions que nous nous posons, d'un point de vue de l'enseignement/apprentissage de l'anglais L2 ou étrangère.

2.3.4 Différences linguistiques entre la RP et le GA

Linguistiquement, la RP et le GA diffèrent sur bon nombre d'aspects en dehors de la seule prononciation. Certains sont bien connus du grand public tandis que d'autres le sont moins. Parmi ces différences, nous ne traiterons que les quatre les plus documentées : les niveaux grammatical, lexical, orthographique et celui qui nous intéresse le plus et sur lequel nous nous attarderons, le niveau phonético-phonologique. Nous les présentons dans l'ordre cité.

2.3.4.1 Niveau grammatical

Sur le plan de la grammaire et de la syntaxe, la RP et le GA sont similaires, à l'exception de quelques points de divergences (Darragh, 2000). Ces différences grammaticales sont remarquables, mais connus des apprenants, et dans une certaine mesure des natifs anglophones. La raison principale est que depuis des décennies, ces différences grammaticales entre les deux accents ont été négligées dans les livres et les manuels de cours. Selon Janicki (1977), cela s'explique par trois raisons. D'abord, ces différences sont plus difficilement décelables que les autres. Ensuite, les différences de prononciation et de vocabulaire sont beaucoup plus nombreuses. Enfin les différences grammaticales relèveraient largement des préférences individuelles. Même si, à en croire Darragh et Janicki, ces différences sont peu nombreuses, il faut toutefois se rappeler qu'elles couvrent une variété de champs grammaticaux et syntaxiques. C'est aussi peut-être pour cela qu'elles passent inaperçues. Quoiqu'il en soit, il semblerait que le GA suivrait une certaine régularisation, dans sa grammaire et sa syntaxe. Le GA serait non seulement plus progressif dans le sens où il admet des évolutions linguistiques, mais également, contiendrait des caractéristiques informelles à l'écrit et à l'oral (Rohdenburg & Schlüter, 2009, p. 111). Si ces affirmations sont prouvées, cette facilité de la grammaire du GA pourrait pousser (inconsciemment ou non) les apprenants d'anglais, à préférer l'apprentissage de l'anglais américain et confirmerait l'hypothèse, selon laquelle l'anglais américain est plus facile à comprendre que la RP (Rindal, 2010, 2014). En somme, la difficulté à assigner une différence (cela ne se limite pas à la grammaire) à une variété, en particulier, peut se résumer avec cette citation :

Although many, few of the grammatical differences between British and American are great enough to produce confusion and most are not stable because the two varieties are constantly influencing each other, with borrowing both ways across the Atlantic and nowadays via the Internet. When a use is said to be British, that statement does not necessarily mean that it is the only or even the main British use or that the use does not occur in American also, but only that the use is attested in British sources and is more typical of British than of American English. (Algeo, 2006, p. 2)

Nous présentons quelques domaines (les plus saillants) avec un ou deux exemples précis de ces différences. La plupart d'entre eux est tirée d'Algeo (2006), Darragh (2000), Davies (2005), Modiano (1996), Schur, Ehrlich, & Ehrlich (2007) et Svartvik & Leech (2016).

Tableau 7 : quelques exemples de différences grammaticales entre la RP et le GA.

Aspects grammaticaux	<i>British English</i>	<i>American English</i>
Le passé de certains verbes (irréguliers)	<i>Spell, spelt Learn, learnt Burn, burnt</i>	<i>Spell, spelled Learn, learned Burn, burned</i>
Le tiret	<i>Oil-painting Blood-pressure After-effect</i>	<i>Oil painting Blood pressure After effect</i>
L'heure	<i>6:30 10/02/2019 (jj/mm/aaaa) See you next Monday</i>	<i>6.30 10/02/2019 (mm/jj/aaaa) See you Monday next</i>
Prépositions	<i>He lives in the same street She arrived at ten past two Write to him as soon as possible</i>	<i>He lives on the same street She arrived at ten after two Write him as soon as possible</i>
Accord sujet-verbe (nom collectif)	<i>The committee are going to issue a statement The government are considering the proposal</i>	<i>The committee is going to issue a statement The government is considering the proposal</i>
L'article défini	<i>Onto grounds Members of staff In future</i>	<i>Onto the grounds Members of the staff In the future</i>
Adjectifs et adverbes	<i>It's really good He drives really fast It's awfully nice</i>	<i>It's real good He drives real fast It's awful nice</i>
Les temps	<i>He has just come He's gone to the bank She has already left</i>	<i>He just came He went to the bank She already left</i>
Insertion de <i>that</i>	<i>She said that she's good</i>	<i>She said she's good</i>
Choix des verbes	<i>Have a shower We've got a car We will go</i>	<i>Take a shower We have a car We shall go</i>
Position des adverbes	<i>She has probably arrived I can hardly do it I must certainly call</i>	<i>She probably has arrived I hardly can do it I certainly must call</i>
Style ³⁶	<i>I suppose one might be inclined, in certain cases that is, to perhaps find this proposal to be worthy of consideration</i>	<i>What a great idea!</i>

2.3.4.2 Niveau lexical

Sur un plan général et indépendamment des accents, la langue anglaise présente une certaine homogénéité lexicale, à quelques exceptions près. Ces différences peuvent affecter la compréhension des locuteurs non-natifs qui les rencontrent pour la première fois. Il peut arriver la

³⁶ Exemple pris de Modiano, 1996, p. 131.

même chose chez des natifs anglais, qui ne connaissent pas ces différences. Deux cas de figures peuvent alors prévaloir. Le premier est lorsque qu'il existe un mot qui désigne deux choses différentes. Le deuxième cas concerne deux mots différents qui réfèrent à une même chose. Il faut préciser qu'aujourd'hui, le mélange de deux types de vocabulaires est observé dans les deux accents (Darragh, 2000).

Tableau 8 : quelques exemples de différences terminologiques entre la RP et GA.

Anglais Britannique	Anglais Américain	Traduction Française
Autumn	Fall	Automne
Motorway	Highway	Autoroute
Flat	Apartment	Appartement
Trousers	Pants	Pantalon
Sweet	Candy	Bonbons
Queue	Line-up	File d'attente
Pavement	Sidewalk	Chaussée/trottoir
Lift	Elevator ³⁷	Ascenseur
Bin	Trash / Garbage	Poubelle
Film	Movie	Film
Garden	Yard	Jardin

2.3.4.3 Niveau orthographique

Cette sous-partie expose quelques exemples de mots orthographiés avec de légères différences dans les deux variétés. Bien que cet aspect soit moins pointé du doigt, il comporterait deux facettes en didactique des langues. Pour un apprenant, cela signifie d'une part qu'il peut facilement reconnaître l'orthographe d'un mot à travers un autre. D'autre part, cela peut prêter à confusion et induire en erreur, lorsqu'il s'agit d'écrire dans une variété définie. Sur le plan cognitif, tout se passe comme si l'apprenant devrait apprendre deux mots différents et savoir les utiliser au moment opportun. Aussi, faut-il signaler que certains sont prononcés d'une façon similaire et que d'autres sont soumis à des facteurs phonético-phonologiques différents que nous présentons dans la section suivante (voir section 2.3.4.4).

Tableau 9 : exemples de mots indiquant quelques différences orthographiques entre la RP et GA.

Anglais Britannique	Anglais Américain	Traduction Française
Colour/honour	Color/honor	Couleur/honneur
Centre /theatre	Center /theater	Centre/théâtre
Organise (to)/realise	Organize (to)/realize	Organiser/se rendre compte
Travelling	Traveling	Voyager/orthographe

³⁷ Il est attesté que les Britanniques commencent à utiliser le mot elevator à la place de lift.

Licence/defence	License/defense	License/défense
Catalogue/dialogue	Catalog/dialog	Catalogue/dialogue
Disc	Disk	Disque
Aeroplane	Airplane	Avion
Mum	Mom	Mère
Tyre	Tire	Pneu
Sceptical	Skeptical	Sceptique
Aluminium	Aluminum	Aluminium
Jewellery	Jewelry	Bijoux

2.3.4.4 Niveau phonétique et phonologique

Certes, les différences entre la RP et le GA se situent à plusieurs niveaux - grammaire, vocabulaire, orthographe - mais la différence la plus notable, faisant souvent l'objet d'études, réside dans la prononciation. Nous résumons ci-après les différences rencontrées dans la littérature. Certaines sont indiquées dans les tableaux contenant les sons de la RP et ceux du GA. D'autres seront encore plus détaillées dans le chapitre méthodologie (voir section 6.6).

Généralement, les études qui visent à comparer les accents suivent les systèmes de différences systémique, distributionnelle, lexicale et réalisationnelle (Cruttenden, 2014, p. 86-87). Une différence systémique concerne celle rencontrée dans l'inventaire phonémique. Ainsi, le contraste entre *Sam* et *psalm* ; *putt* et *put*, *sin* et *sing* dans la RP n'est pas présent dans les autres accents, qui prononcent ces mots de la même façon (par exemple dans l'accent écossais). Quant à la différence distributionnelle, elle réfère aux différentes possibilités phonotactiques d'un phonème, et se réalise seulement sous certaines conditions. Par exemple, les locuteurs des accents rhotiques prononcent le <r> postvocalique (*car*, *more*), contrairement aux locuteurs des accents non rhotiques. Les différences lexicales concernent uniquement des phonèmes dans certains mots particuliers. Ceci relève des choix des locuteurs même si ces phonèmes existent dans leur accent. C'est ainsi que la RP utilise /ɒ/ dans *off*, *cloth* et *cross*, tandis que /ɔ:/ est plutôt utilisé en *Cockney*. Enfin, une différence réalisationnelle advient lorsque deux accents utilisent deux réalisations phonétiques différentes d'un même phonème. C'est notamment le cas avec le r et le l, réalisés tantôt en approximant alvéolaire [ɹ], en *trill* alvéolaire [r], en *tap* rétroflexe [ɻ], en *tap* alvéolaire [ɹ], en approximant labiodental [ʋ] ou en *trill* uvulaire [ʀ] et en clair [l] ou en sombre [ɫ] selon les accents. La citation suivante résume la situation :

Comparing RP and the somewhat constructed 'average' accent General American, which may be referred to as the two reference accents, we find that the vowel systems differ quite

substantially. The most striking difference is that American English has fewer diphthongs, generally lacking centring ones and having a monophthong in words such as goat. (Melchers & Shaw, 2013, p. 18)

De notre point de vue, les frontières sont assez ambiguës et se croisent. De ce fait, nous n'avons pas voulu suivre explicitement ces catégorisations. Pour notre part, nous avons rassemblé les différences les plus saillantes dans un seul tableau (ci-dessous), tout en indiquant le nom ou la nature du phénomène étudié (rhoticité, *yod-dropping*). Le cas échéant, nous opposons uniquement les sons qui sont différents dans les deux accents (/ɑ/ ou /ɒ/). Cet ensemble se subdivise en deux sous-catégories à savoir les éléments segmentaux (voyelles et consonnes) et suprasegmentaux (accentuation, intonation). À ces aspects, nous pouvons ajouter des phénomènes comme le *vocal fry* (ou *creaky voice*) et le *uptalk* (*high terminal*), qui sont spécifiques au parler des Américains.

Tableau 10 : quelques exemples de différence de prononciation entre la RP et le GA.

	Éléments de différence	Exemples de mots	Prononciation RP	Prononciation GA
Consonnes	Rhoticité	Car	/kɑ:/	/kɑɪ/
	/t/ intervocalique	City	/'sɪti/	/'sɪri/
	Yod-dropping	Tune	/tju:n/	/tu:n/
	Dark l ³⁸	Light	[laɪt]	[lɑɪt]
	Retroflexion	Red	/ɪɛd/	/ɹɛd/
Voyelles	Nasalisation vocalique	Hand	/hænd/	/hæ̃d/
	/ɛ/ ou /i:/	Zebra	/'zɛbrə/	/'zi:brə/
	/ɪ/ ou /aɪ/	Vitamin	/'vɪtəmm/	/'vaɪtəmm/
	/ɒ/ ou /ɑ:/	Lot	/lɒt/	/lɑ:t/
	/ɒ/ ou /ɔ/	Cloth	/klɒθ/	/klɔ:θ/
	/əʊ/ ou /oʊ/	Goat	/gəʊt/	/goʊt/
	/ɑ:/ ou /æ/	Past	/pɑ:st/	/pæst/
Les Diphtongues	Fear	/fɪə/	/fɪr/	
Supra-segmentales	Accentuation	Rotate	/rəʊ'teɪt/	/'rɔʊteɪt/
	Nombre de syllabes	Secretary	/'sekɹət'ɹi/	/'sekɹətəri/

Nous venons d'exposer sommairement les différents types d'anglais parlés dans le monde et nous avons vu qu'il en existe une multitude. Nous nous sommes ensuite attardé sur les diversités linguistiques au Royaume-Uni et aux États-Unis, avec un passage en revue des deux normes linguistiques dans ces deux régions, la RP et le GA, ainsi que les dissemblances majeures existantes

³⁸ En règle générale, en RP, le /l/ clair est produit devant les voyelles, souvent en début (*like* [laɪk]) et à l'intérieur de mot (*dislike* [dɪs'laɪk]) tandis que le /l/ sombre se retrouve en fin de mot (*pill* [pɪl]). En revanche, le /l/ est sombre dans toutes les positions en GA.

(grammaticales, lexicales, orthographiques et phonético-phonologiques). De la même manière, nous allons voir quelques nuances principales entre l'anglais (en général) et le français, sur le plan de la prononciation uniquement. Les transcriptions reposent sur les symboles phonétiques de l'IPA.

2.4 Différences phonétiques et phonologiques entre l'anglais et le français

Travaillant avec des locuteurs français, la prise en compte du système phonétique de cette langue nous paraît importante. En effet, une comparaison des deux langues ferait certainement ressortir des éléments permettant de mieux comprendre et expliquer les productions de nos locuteurs francophones cibles. Nous utilisons le terme « comparaison », mais un tel projet est bien ambitieux dans la mesure où il s'agit de deux systèmes distincts, donc techniquement non-comparables. Sans entrer dans les détails articulatoire et acoustique du système phonétique du français, nous allons essayer de discuter les quelques principaux points de divergence (et de convergence s'il y en a) entre les deux langues.

Tout d'abord, le système phonétique de l'anglais est plus « complexe » que celui du français. Cette « complexité » se traduit sur plusieurs plans. En effet, l'inventaire phonétique de l'anglais comporte 44 sons (24 consonnes et 20 voyelles) contre 36 (20 consonnes et 16 voyelles) en français : il possède par conséquent plus de sons que le français, soit 8 de plus (4 voyelles et 4 consonnes). D'office, cela présuppose qu'un apprenant français devrait s'adapter à de nouvelles « cibles » perceptives et articulatoires en vue de pouvoir percevoir et prononcer les sons de l'anglais. Selon Viel (1981), les spécificités articulatoires permettant de discriminer les sons les uns par rapport aux autres sont moins contrastées en anglais qu'en français. Cela fait de l'anglais l'une des langues possédant le plus de sons difficiles à prononcer. Comme le souligne Viel (1981, p. 11) : « On s'accorde volontiers à reconnaître la simplicité de la grammaire, mais cette simplicité serait largement compensée par la difficulté de la prononciation ». Nous nous permettons toutefois d'émettre des réserves quant à cette notion de difficulté de prononciation de sons, de façon absolue. Tout semble dépendre de la « distance » ou du degré de la différence du système phonologique entre L1 et L2.

Tableau 11 : les sons du français (adapté de Vaissière, 2006, p. 12).

	Son	Transcription du mot	Mot
Consonnes			
Occlusives	/p/	/pã/	Pan
	/t/	/tã/	Temps
	/k/	/kã/	Camp
	/b/	/bã/	Banc
	/d/	/dã/	Dent
	/g/	/gã/	Gant
Fricatives	/f/	/fã/	Faon
	/s/	/sã/	Sang
	/ʃ/	/ʃã/	Chant
	/v/	/vã/	Vent
	/z/	/zã/	Zan
	/ʒ/	/ʒã/	Gens
	/ʁ/	/ʁã/	Rang
Nasales	/m/	/mã/	Ment
	/n/	/nã/	Nan
	/ɲ/	/aɲo/	Agneau
Latérales	/l/	/lã/	Lent
Voyelles			
Fermées orales	/i/	/pil/	Pile
	/y/	/pyl/	Pull
	/ø/	/pø/	Peu
	/u/	/pul/	Poule
	/o/	/po/	Peau
	/e/	/ekɔl/	École
	/ə/	/lə/	Le
Ouvrées orales	/ɛ/	/nɛʒ/	Neige
	/œ/	/pœʁ/	Peur
	/ɔ/	/pɔl/	Paul
	/a/	/pat/	Patte
	/ɑ/	/pat/	Pâte
Nasales	/ã/	/pãt/	Pente
	/õ/	/põ/	Pont
	/ɛ̃/	/bɛ̃/	Brin
	/œ̃/	/bœ̃/	Brun
Semi-voyelles / consonnes			
Spirantes	/w/	/lwi/	Louis
	/ɥ/	/lɥi/	Lui
	/j/	/fij/	Fille

Les différences entre les deux systèmes peuvent se résumer par les points suivants :

- En plus des consonnes /p, t, k, b, d, g, f, s, ʃ, m, n, ɲ, l, ʒ, R/ qui existent aussi en français, l'anglais compte les consonnes suivantes : /θ, ð, ŋ, h, dʒ, tʃ/ ;

- Contrairement au français, les plosives sourdes anglaises /p, t, k/ sont plus aspirées en positions initiales et en syllabes accentuées ([p^h], [t^h], [k^h] comme dans *pin, tin, kin, pertain, contain* etc. mais pas dans *skin, spin, tomato*) ;
- Le /h/ est prononcé partout en anglais (house_/haʊs/, home_/həʊm/, etc.) sauf à de rares exceptions (hour_/aʊə/, honor_/ˈɒnə/, honest_/ˈɒnest/), alors qu'il n'est jamais réalisé en français (heure_/øʁ/, honneur_/ɔ̃nøʁ/, honnête_/ɔ̃nɛt/) ;
- Il y a deux /l/ en anglais (*light*_/[laɪt]/ et *pill*_/[p^hɪl]/) alors qu'il n'en existe qu'un en français (*lire*_/[liʁ]/) ;
- Il vaut mieux partir de l'idée qu'aucune voyelle anglaise n'est identique à une voyelle française. En effet, certaines voyelles dans les deux langues se ressemblent perceptivement (/u, o, ɔ, a, ε, i/) mais n'ont pas de valeurs formantiques identiques voire proches. Pour la voyelle /ε/ par exemple, la moyenne des trois premiers formants est : F₁=580 Hz, F₂= 1799 Hz et F₃=2605 Hz pour les hommes anglophones (Hillenbrand *et al.*, 1995, p. 3103) et : F₁=530 Hz, F₂=1728 Hz et F₃=2558 Hz pour les hommes francophones (Calliope, p. 84). La différence entre ces deux réalisations semble reposer davantage sur une différence du F₁ ;
- Les diphtongues n'existent plus en français ;
- L'anglais ne possède pas les voyelles françaises /y, ø, œ/ et les voyelles anglaises /ʌ, ɒ, æ, ɜ, ɪ, ɪ, ɑ, ʊ/ n'existent pas en français ;
- Le contraste voyelles courtes/longues n'entraîne pas d'oppositions phonologiques en français, tandis qu'il est phonologique en anglais (*sit/sɪt/* vs. *seat/si:t/*) ;
- La place des syllabes accentuées varie en anglais alors qu'elle est le plus souvent sur la dernière syllabe du groupe rythmique en français ;
- Si nous considérons le français comme étant rhotique, puisque les <r> sont partout prononcés, l'articulation est différente : une approximante alvéolaire en anglais (/ɹ/) contre une uvulaire en français (/ʁ/).

Aussi, sur le plan suprasegmental, les deux langues se comporteraient de la même manière, à un niveau général, mais différent considérablement dans des variables isolées et simples. Dans une étude que nous résumons, Grosjean et Deschamps (1975) ont analysé des variables temporelles de discours issus d'interviews radiophoniques de 60 locuteurs anglophones et francophones (30 de chaque langue). Les variables étudiées sont la vitesse d'élocution et le rapport temps d'articulation-temps de locution (RTATL), sur un plan global et la vitesse d'articulation, la fréquence et la durée des pauses (silencieuses, remplies, syllabes allongées, répétitions et faux départs), sur un plan détaillé. Ces chercheurs ont trouvé que les deux langues ont des valeurs moyennes presque identiques : 175,54 mots/min (soit un RTATL de 83,15%) pour l'anglais et 173,6 mots/min (soit un RTATL de 84,45%) pour le français. Mais à un niveau plus détaillé, la différence la plus importante est que les « Anglais » articulent plus vite. Ainsi, il présente 9,5 syllabes/suite sonore contre 12 syllabes/suite sonore pour le français. Les « Français » réaliseraient plus de répétitions liées aux erreurs grammaticales (accord de genre et de nombre des articles) alors que Les « Anglais »

auraient davantage recours aux substitutions lexicales. Est-ce une preuve que la grammaire française est plus difficile ? Aussi, les « Français » feraient plus de faux départs que les « Anglais » et les pauses remplies seraient plus courtes. Les auteurs expliquent ces résultats par le fait que l'anglais comporte plus de mots grammaticaux et lexicaux plus courts (mono et dissyllabiques) mais effectue davantage de courtes pauses. À contrario, le Français a tendance à effectuer moins de pauses mais qui sont plus longues, malgré une longueur plus importante des suites sonores. Il en résulte que l'Anglais fournirait une quantité d'informations plus conséquentes que le Français avec le même nombre de mots. Il semblerait donc que les Anglais, pour des raisons physiologiques spécifiques, auraient besoin de respirer plus souvent que les Français, mais restent toutefois plus concis. Ils en concluent qu'il y aurait un mécanisme de compensation selon la tâche réalisée et selon chaque locuteur. Des différences individuelles notables ont été également constatées. Il serait toutefois intéressant de reconduire cette étude aujourd'hui, afin de vérifier s'il y a eu évolution dans la manière de parler l'une ou l'autre langue, sur le plan de la vitesse d'élocution et de la quantité d'informations fournie, en fonction des locuteurs des deux langues.

Ensuite, une autre différence entre les deux langues se situe au niveau de l'orthographe. En effet, l'anglais est une langue dont la prononciation ne correspond pas, dans la grande majorité des cas, à la manière dont il est orthographié. Le français non plus mais, à quelques exceptions, il possède une certaine cohérence entre la prononciation des mots et l'orthographe. Cette relation entre la prononciation et l'orthographe est appelée transparence grapho-phonémique (Deschamps, 1994). L'antagonisme entre écrit et prononciation en anglais s'observe principalement à deux niveaux. D'une part, certains sons ne correspondent guère à l'orthographe (*ewe* se prononce /ju:/) et d'autre part, un même graphème (ou orthographe) peut se prononcer différemment selon le mot. <ou> se réalise /ʌ/ dans *tough* (/tʌf/), /əʊ/ dans *though* (/ðəʊ/), /ɔ:/ dans *ought* (/ɔ:t/), /u:/ dans *ouzel* (/ʊ:zəl/ et /aʊ/ dans *bough* (/baʊ/). Aussi, la voyelle /i:/ admet des graphies diverses : <ee> (*see* /si:/), <e> (*obese* /əʊ'bi:s/), <ea> (*meat* /mi:t/), <ie> (*believe* /bi'li:v/), <ei> (*receive* /ri'si:v/), <ey> (*key* /ki:/), <i> (*police* /pə'li:s/), <ay> (*quay* /ki:/), <eo> (*people* /'pi:pəl/) et <ae> (*Caesar* /'si:zə/). Ou encore, la voyelle /ɛ/ représentée par des graphies <e> (*bed* /bed/), <ea> (*breakfast* /'brek fəst/), <a> (*many* /'meni/), <u> (*bury* /'beri/), <ei> (*leisure* /'leɪzə/), <eo> (*leopard* /'lepəd/), ainsi que <ay> (*says* /sez/) et <ai> (*said* /sed/). En revanche, en français, on note très peu de ces discordances et elles relèvent plutôt des habitudes

liées aux accents régionales, voire historique³⁹. Ainsi prononcera-t-on les graphies <un> de brun en /bʁœ/ ou /bʁɛ / ; <o> en /pɔl/ ou /pɒl/ (Paul) et <a> en /pɑt/ ou /pæt/ (pâte) ; <eu> en /ʒɒn/ ou /ʒœn/ (jeune) ; <ai> et <es> en /le/ ou /le/ (lait, les).

Dans le même ordre d'idée, pour confirmer ces discordances grapho-phonémiques, Trenité (1920) dénonce de façon amusante les incohérences de la prononciation anglaise à travers un long poème de 238 vers ; recensant environ 800 irrégularités grapho-phonémiques notoires en anglais. Le poème intitulé « *The chaos* » est apparu pour la première fois dans l'appendice de son livre *Drop Your Foreign Accent* (Trenité, 1920) et a été publié depuis chez plusieurs éditeurs. En voici quelques extraits (le début).

Tableau 12 : Extraits du poème de Trenité, *The Chaos*.

Vers	Transcription large (IPA, version RP)
Dearest creature in creation	'diərist 'kri:tʃə m kri(:)'eɪʃən
Studying English pronunciation,	'stʌdɪŋ 'ɪŋɡlɪʃ prə'nʌnsi'eɪʃən,
I will teach you in my verse	aɪ wɪl ti:tʃ ju: m maɪ vɜ:s
Sounds like corpse, corps, horse and worse.	səʊndz laɪk kɔ:ps, kɔ:, hɔ:s ənd wɜ:s.
I will keep you, Susy, busy,	aɪ wɪl ki:p ju:, 'su:zi, 'bɪzi,
Make your head with heat grow dizzy;	meɪk jɔ: hed wɪð hi:t grəʊ 'dɪzi;
Tear in eye, your dress you'll tear;	tiə m aɪ, jɔ: dres ju:l tiə;
Queer, fair seer, hear my prayer.	kwiə, feə 'si:ə, hiə maɪ preə.
Pray, console your loving poet,	preɪ, kən'səʊl jɔ: 'ləvɪŋ 'pəʊt,
Make my coat look new, dear, sew it!	meɪk maɪ kəʊt lʊk nju:, diə, səʊ ɪt!
Just compare heart, hear and heard,	dʒʌst kəm'peə hɑ:t, hiə ənd hɜ:d,
Dies and diet, lord and word.	daɪz ənd 'daɪət, lɔ:d ənd wɜ:d.
Sword and sward, retain and Britain	sɔ:d ənd swɔ:d, ri'teɪn ənd 'brɪtn
(Mind the latter how it's written).	(maɪnd ðə 'lætə haʊ ɪts 'rɪtɪn).
Made has not the sound of bade,	meɪd hæz nɒt ðə səʊnd ɒv beɪd,
Say-said, pay-paid, laid but plaid.	seɪ-sed, peɪ-peɪd, leɪd bʌt plæd.
Now I surely will not plague you	nəʊ aɪ 'ʃʊəli wɪl nɒt pleɪg ju:

Si nous revenons à un contexte d'enseignement et d'apprentissage des langues secondes et étrangères, force est de constater que ces discordances constituent la plus grande source de difficultés de prononciation pour les non-natifs (Krashen, 2013). Ceux-ci devront adopter une norme, notamment ceux qui apprennent la langue dans un contexte formel et qui passent inévitablement par l'orthographe. Selon Roussel & Gaonac'h (2017), les apprenants qui s'exposent aux activités informelles, sans passer tout de suite par l'orthographe, adoptent une meilleure

39 Non systématique mais liées l'aperture et au couple antériorité-postériorité.

prononciation. Plusieurs études ont montré les effets de l'orthographe sur la lecture et la prononciation, mais les récents travaux en neurosciences de la cognition montrent l'intervention des facteurs neurophysiologiques, environnementaux et individuels (Miras, 2016). En se reposant sur les quelques exemples de (non)correspondances phonie-graphie dans les deux langues, nous concluons que l'anglais est moins transparent que le français. Ceci n'est pas sans conséquences pour les apprenants, particulièrement en matière de prononciation.

2.5 Conclusion

Dans ce chapitre, nous avons parlé de l'importance de l'anglais en tant que langue mondiale, avec ses variétés connexes. Nous avons ensuite circonscrit le terrain en comparant les deux variétés les plus connues, à savoir la RP et le GA. Enfin, après avoir abordé les notions d'anglicisme et de gallicisme, nous avons également comparé l'anglais et le français, sur un plan phonétique et phonologique.

Au regard de tous ces éléments, nous pouvons faire deux observations : d'une part la difficulté de couvrir toutes les différences dans les variétés étudiées, surtout parce que certaines sont subtiles ; d'autre part, la multiplicité de facteurs intervenant dans la catégorisation de ces différences, induisant ainsi un système complexe à étudier et une variabilité accrue. Certes, tous ces facteurs constituent la matière à étudier, surtout dans un monde linguistique en pleine évolution, où les langues s'influencent mutuellement (la RP sur le GA et sur le français, et inversement). Nous nous posons ainsi quelques questions ouvertes par rapport au contexte d'enseignement et d'apprentissage des langues et de l'anglais en particulier. Quelle variété faut-il enseigner ou apprendre ? Comment s'en tenir à un seul accent lorsque nous sommes de plus en plus exposés à diverses variétés que celui que nous avons appris ? Comment juger/évaluer la prononciation des apprenants d'anglais L2, si toutes les variétés standard sont tout autant éloignées les unes des autres ? Peut-on plutôt dire que ce sont les normes qui sont déconnectées de la réalité ? La question de la norme et de l'enseignement-apprentissage des langues se pose ici. Dans ce cadre, il est pertinent de justifier le terrain d'étude de l'AIAL par des fondements théoriques, capables de soutenir et d'expliquer comment l'apprentissage se déroule chez l'individu. Nous allons aborder, dans l'ensemble suivant, notre cadre théorique qui se compose de trois théories globales : la théorie des systèmes complexes et dynamiques, la théorie de l'imitation

phonético-phonologique et la théorie de la viabilité en production-perception de la parole. Outre la présentation de ces dernières, nous expliciterons pourquoi nous les avons choisies et en quoi elles sont pertinentes pour notre travail. Sans plus attendre, nous commençons par la présentation de notre première théorie principale, à savoir la théorie des systèmes complexes et dynamiques.

3 Théorie des systèmes complexes et dynamiques

From a complexity theory point of view, not only do we get a more variegated portrayal of language using patterns, we also get a different, more emic, or learner-centered account of their development.

Learning is not the taking in of linguistic forms by learners but the constant adaptation of their linguistic resources in the service of meaning-making in response to the affordances that emerge in the communicative situation, which is in turn affected by the learners' adaptability (Larsen-Freeman and Cameron, 2008, p. 135).

Dans ce chapitre, nous présentons une théorie centrale à notre travail qui nous permet à la fois de mieux comprendre le processus d'apprentissage informel de l'anglais de nos participants et de mieux expliquer nos résultats. Cette théorie, nommée théorie des Systèmes Complexes et Dynamiques (SCD), est un ensemble de théories issues des sciences dites dures qui réunit la théorie de la systémique, du chaos, de la complexité, des systèmes dynamiques et de la cybernétique, pour n'en citer que quelques-unes. Cette théorie stochastique (liée au chaos) a changé notre regard sur le monde, qui s'est alors détourné d'une conception déterministe des phénomènes qui nous entourent, vers une vision plus holistique et extensive de la manière dont nous vivons. D'une façon synthétique, cette théorie stipule que nous vivons dans un environnement qui est un système complexe et dynamique (SCD), dans lequel tous les éléments sont liés par une relation d'interdépendance évoluant sur des trajectoires non-linéaires, autour de multiples attracteurs qui donnent naissance à des phénomènes émergents imprévisibles et dont les causes sont multiples et variées. Nous utilisons cette théorie dans notre travail, car, à l'instar de nombreux chercheurs, nous considérons l'apprentissage en général comme répondant aux critères d'un SCD. En didactique des langues, l'apprenant peut être alors considéré comme un élément, déjà complexe, évoluant dans un autre système complexe qu'est l'apprentissage. En effet, le système scolaire et extrascolaire, la motivation, l'affect, les objectifs, le passé, la personnalité, les stratégies d'apprentissage, etc. constituent des facteurs que l'apprenant va devoir gérer ou subir de façon concomitante, en vue de délinéer un chemin qu'il choisira (ou pas), en fonction de tous les autres facteurs internes (objectifs personnels, capacité) et externes (environnement, ressources, politiques éducatives) (de Bot, Lowie & Verspoor, 2007, p. 11). Dans notre cas plus précisément, nous considérons le processus d'apprentissage phonétique et phonologique comme étant inscrit dans un SCD. L'apprenant de L2 est exposé à une masse de facteurs interconnectés : les activités

formelles (cours), informelles (musique, films, séries, jeux, lecture, etc.), sa motivation, ses intérêts et ses interactions avec d'autres personnes. Il s'agit, pour nous, d'utiliser cette théorie dans une approche explicative des questions que nous nous posons à l'égard de notre problématique, plutôt que de vouloir chercher à suivre une démarche méthodologique spécifique ayant pour objectif de trouver les causes d'un problème donné. Toutefois, des pistes de solutions ne seront pas exclues dans cette perspective, même si les objectifs principaux sont de comprendre, de décrire et d'expliquer les faits émergents au travers des pratiques de l'AIAL de nos participants.

Dans ce chapitre, nous allons tout d'abord commencer par un rappel historique des origines de la théorie et évoquer rapidement ses applications dans les nombreuses disciplines et domaines aujourd'hui. Ensuite, nous exposerons les caractéristiques principales de la théorie des SCD nécessaires à sa compréhension, avec des définitions et exemples tirés de la vie courante, mais aussi en linguistique appliquée. La partie centrale de ce chapitre sera consacrée à l'emploi de la théorie des SCD dans la sphère de la linguistique appliquée, au sein de laquelle nous aborderons quelques notions capitales pour le traitement de nos données et l'analyse de nos résultats (interlangue, différences individuelles et variabilité). Suite à cela, nous discuterons des limites d'ordre méthodologique de l'utilisation de la théorie en didactique des langues et enfin, nous discuterons de pistes de solutions méthodologiques proposées par des auteurs spécialistes du domaine. Nous commençons la section suivante par les origines de la théorie des SCD.

3.1 Les origines de la théorie des SCD

Au cours du XX^e siècle, il y a eu trois révolutions scientifiques majeures qui ont changé la façon dont nous percevons le monde. Ces révolutions sont liées à la découverte et à la mise au point de lois et principes qui ont été théorisés par des chercheurs : il s'agit de la Relativité de Einstein, de la Physique Quantique de Max Planck et du Chaos de Poincaré (Yoccoz, 1995). La troisième théorie, la théorie du chaos, dont est issue la théorie des SCD, sera l'objet central des discussions menées dans ce chapitre et sera abordée tout au long de cette thèse. Nous commençons par un bref aperçu historique des origines de la théorie des SCD. Ce rappel nous semble d'une importance capitale pour la compréhension globale de la théorie, car, comme nous le verrons par la suite, son utilisation et son interprétation varient énormément dans la littérature (Aubin & Dalmedico, 2002). Au départ, les SCD constituent une théorie mathématique appliquée en physique, en chimie

ou encore en biologie, et dont les principes permettent d'étudier le comportement (fonctionnement, interactions, évolutions) des SCD. Il est souvent de coutume de présenter l'auteur qui a initié une théorie, mais il nous sera pour ce qui concerne la théorie des SCD très difficile d'en attribuer la « paternité » à un auteur spécifique pour plusieurs raisons. Tout d'abord parce que de nombreux chercheurs ont travaillé sur la théorie de manière simultanée et consécutive, non seulement dans des domaines scientifiques variés (sciences exactes), mais également dans divers pays (aux États-Unis, en Europe et en Russie). Ensuite, depuis ses origines, la théorie a évolué dans le temps et a connu des améliorations majeures. Par conséquent, elle se compose de plusieurs sous-théories qui traitent des aspects spécifiques issus de concepts divers. Il émerge alors des dénominations différentes en fonction de ces aspects, parmi lesquelles on peut trouver : la théorie des systèmes complexes, la théorie des systèmes dynamiques, la théorie du chaos, la théorie ergodique, la systémique, etc. Cependant, il faut noter que la théorie remonte à une date antérieure. Ainsi, la majorité des ouvrages dans le domaine s'accordent pour attribuer les prémises de la théorie des SCD à Henri Poincaré⁴⁰, qui a travaillé sur la stabilité du système solaire grâce à l'analyse qualitative des équations différentielles. Il a donc été le premier à avoir évoqué certaines caractéristiques comme la linéarité et la non-linéarité ou l'imprédictibilité (Abraham & Shaw, 1992).

D'une façon générale, même si la théorie des SCD était déjà utilisée auparavant, elle n'a connu son essor que vers les années 1970⁴¹ avec divers travaux scientifiques dans des disciplines variées. On peut notamment citer les travaux de Lorenz sur la météorologie et les attracteurs étranges (1963), ceux de Ludwig von Bertalanffy en biologie dans les années 1950, ceux d'Esther Thelen et Linda Smith en psychologie du développement (1996) ou encore ceux de Roberto Harari en psychanalyse pour l'étude des pulsions (2005). Nous nous attardons sur les travaux de Lorenz, car ceux-ci renferment des métaphores qui se prêtent particulièrement bien à la compréhension du système d'apprentissage des langues. Le météorologue Edward Lorenz (1963) a mis en avant le fait que les systèmes complexes ont une sensibilité particulière aux conditions initiales. En effet, cherchant à étudier les fluctuations météorologiques au travers de simulations informatiques, il mit au point une équation différentielle qui lui donna des résultats « étonnants » : un changement

⁴⁰ Mathématicien français vers la fin du XIXe siècle (1880).

⁴¹ Cette date est variable selon les auteurs.

minime des données initiales engendrait des conséquences énormes et différentes à chaque fois. Il publia alors ses résultats dans un article intitulé Prédicibilité : le battement d'ailes d'un papillon au Brésil provoque-t-il une tornade au Texas ? (Lorenz, 1972). Malheureusement, cette métaphore de l'effet papillon fut mal interprétée et Lorenz dut expliquer qu'il ne s'agit pas d'une relation directe de cause à effet. Il s'agit plutôt d'une accumulation de plusieurs battements d'autres animaux qui peuvent modifier l'état du vent qui souffle (direction, intensité, etc.) et qui par la suite peut causer une tornade. Cette image signifie qu'un petit changement peut avoir des conséquences importantes. Lorenz admet aussi qu'il n'est pas envisageable de prédire la météo sur une longue période à cause de l'état chaotique du climat, dont les variables changent perpétuellement et sont inévitables. Il suffit d'un changement infime dans les conditions atmosphériques pour que même les prédictions météo du jour ne s'exécutent pas.

Depuis Poincaré jusqu'à aujourd'hui, beaucoup d'auteurs dans d'autres disciplines se réfèrent à la théorie des SCD. C'est ce dont nous allons discuter dans la section suivante, avec son application⁴² en sciences humaines et sociales (SHS) et notamment en sociologie (avec les mouvements de flux), puis en linguistique et en didactique des langues.

3.2 La théorie des SCD aujourd'hui

Aujourd'hui, la théorie des SCD est appliquée dans nombreuses disciplines et matières qui s'établissent en distinction aux thèses de réductionnisme⁴³ et de déterminisme⁴⁴. Outre les disciplines scientifiques de prédilection (mathématiques), elle élargit ses champs d'application à d'autres domaines de recherche :

Les résultats sont donc applicables à toutes sortes de systèmes (économiques, organiques, chimiques, biologiques, psychologiques ou sociaux). Les indicateurs boursiers, le climat, les taux de croissance de populations, le rythme cardiaque, tous reflètent des systèmes complexes qui traversent des phases chaotiques (Dewaele, 2002, paragr. 20).

42 Il ne s'agit pas vraiment d'une application *stricto sensu* mais plutôt d'une adaptation, car nous verrons plus loin que la théorie des SCD ne peut pas être « appliquée » en SHS.

43 Courant épistémologique qui repose sur la simplification et l'élimination de certaines notions ou variables fondamentales à la compréhension de phénomènes naturels.

44 Courant épistémologique qui attache une valeur importante aux principes de causalité, de lois physiques et mathématiques comme une suite déterminée d'événements. En d'autres termes, tout est calculable et tout ce qui arrive doit arriver et ne peut arriver autrement.

La théorie est également utilisée en psychologie du développement, dans les études de comportements humains (cognition et personnalité), mais aussi en linguistique, dans le cadre de l'analyse du discours et du développement langagier (Kusyk, 2017). C'est notamment grâce aux travaux du sociologue, philosophe et épistémologue, Edgar Nahoum Morin, sur la pensée complexe et la complexité humaine (1982, 1994, 2000, 2006), que la théorie est connue dans les SHS en France. En effet, l'auteur considère que la recherche ne devrait pas se limiter à des études cloisonnées au sein de disciplines distinctes, ni par une connaissance spécifique, mais devrait émerger d'un ensemble de connaissances interconnectées, permettant de mieux cerner la complexité d'un phénomène. Il considère la pensée complexe comme étant régie par trois grandes catégories de faits : (1), l'ordre, le désordre et l'organisation, (2) l'auto-organisation/régulation et (3) l'itération. Selon Abdelmalek (2004), Edgar Morin admet qu'il est difficile d'appréhender les différentes composantes sans connaître le « tout »⁴⁵, ou de comprendre le « tout » sans connaître ses différentes composantes, car un système est plus que la somme de ses parties (Morin, 2005). En effet, mathématiquement, le « tout » est égal à la somme de ses parties, ni plus ni moins. Mais dans un SCD, le « tout » est autre chose ou quelque chose de différent par rapport à la somme de ses parties, car il faut également prendre en compte les interactions qui existent entre les différents éléments du système. Dans les études de comportements sociaux, les effets dégagés pour l'ensemble des individus sont distincts de ceux observés pour chaque individu pris séparément. Les experts de l'organisation de masse⁴⁶ soutiennent que ce principe décrit la synergie qui existe entre les individus travaillant ensemble dans un effort coopératif. Collectivement, ils sont capables d'atteindre un résultat supérieur à celui de toutes les personnes constituant le groupe travaillant individuellement. Aujourd'hui, bon nombre de recherches prennent en compte cet aspect intrinsèque et fondamental aux SCD.

Nous rappelons que la théorie des SCD regroupe plusieurs autres et est utilisée aujourd'hui dans diverses disciplines comme le montre la figure ci-dessous. On peut déjà remarquer l'ordre imbriqué de sous-systèmes dans d'autres sous-systèmes, qui sont à leur tour incorporés au sein d'un plus grand système, gouverné par deux phénomènes centraux : l'autorégulation et l'émergence. Il est également intéressant de voir que si nous inscrivons chaque composant et sous-

⁴⁵ Le tout représente ici un système.

⁴⁶ Exemples de la société en général : le peuple, l'entreprise, la foule, les élèves, etc.

composant dans des cercles de tailles variables, nous obtiendrons une forme fractale⁴⁷, caractéristique des SCD. Dans les sections qui suivent, nous allons présenter des caractéristiques principales de cette théorie, en nous référant au domaine de l'apprentissage des langues.

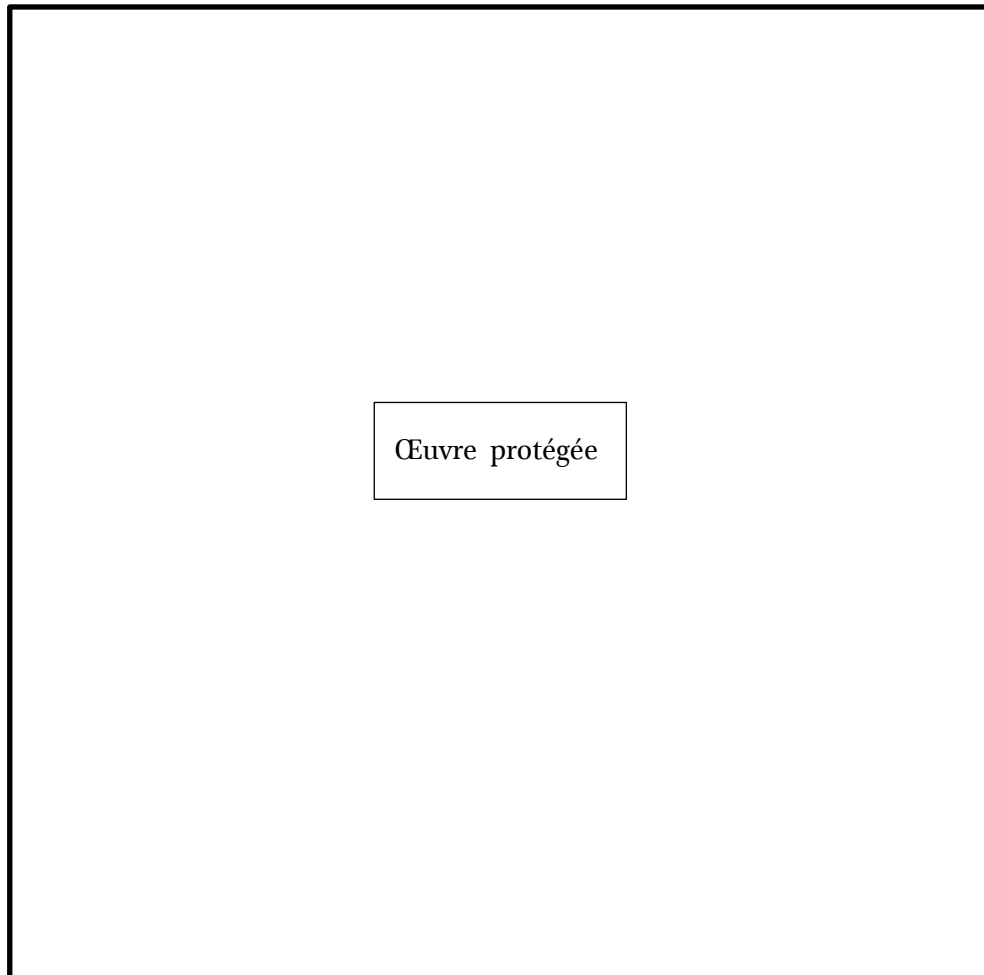


Figure 3 : quelques exemples de théories et disciplines qui composent la théorie des SCD.⁴⁸

3.3 Caractéristiques de la théorie des SCD

Dans cette section, nous définissons des éléments caractéristiques nécessaires à la compréhension de la théorie des SCD à l'aide d'une terminologie spécifiquement dédiée. Ici, il est à entendre par caractéristiques les constituants qui forment des propriétés de fonctionnement d'un

⁴⁷ Les fractales (de fractions) sont des éléments géométriques naturels ou artificiels composés de plus nombreux autres éléments identiques et répétitifs mais miniatures et interconnectés, et dont les différentes composantes sont elles-mêmes des systèmes complexes. On peut ainsi dessiner un millier de petits carrés d'un millimètre dans un grand carré d'un mètre carré. Pour voir un exemple : <https://fr.wikipedia.org/wiki/Fractale>

⁴⁸ Source : https://commons.wikimedia.org/wiki/File:Complex_systems_organizational_map.jpg (Consulté le 26/06/2019). Image sous licence Creative Commons (CC), Hiroki Sayama (2010).

système spécifique. Plus un système est complexe, plus ses composants sont nombreux et deviennent eux-mêmes des SCD. Nous verrons bien que déjà, dans le processus de définition et d'explication de ces termes, il existe une certaine complexité inhérente puisque les composants sont interconnectés, se recourent et qu'une tentative de vouloir les définir de manière isolée reste une tâche difficile⁴⁹, voire « impossible ». Dans cette section, nous expliquons d'abord les concepts de façon générale, avec des exemples tirés de la littérature et dans un second temps, nous faisons le lien avec des exemples du domaine de la didactique des langues (enseignement/apprentissage). Même si tous les autres termes sont d'importance égale, nous commençons ici par le concept clé de complexité.

3.3.1 Complexité et interconnectivité

Une entité ou un objet complexe peut être défini comme un ensemble composé d'éléments interagissant et s'influçant entre eux⁵⁰, et dont les formes évolutives sont difficiles à prévoir et à cerner. Un système complexe est d'autant plus difficile à décrire quand ses composants sont eux-mêmes des entités complexes. Par exemple, le corps humain est considéré comme un système complexe à part entière, composé d'autres organes complexes (le cerveau en est un, par excellence). Le propre de tels systèmes est le fait qu'ils entretiennent en leur sein des relations de connectivités locales et concomitantes. On distingue ainsi deux sortes de systèmes complexes et dynamiques dans la littérature (Yoccoz, 1995). Le premier est dit fermé lorsqu'il n'admet aucune intégration de variables externes. Le deuxième système est dit ouvert et communique avec l'environnement dans lequel il évolue. Les deuxièmes sont majoritaires alors que les premiers sont rares, malgré leur autosuffisance. Nous pouvons citer comme exemples de systèmes complexes : le réseau Internet, les réseaux ferroviaires, les colonies d'insectes, les troupeaux de moutons, les colonies d'oiseaux, les bancs de poissons, la galaxie, la météorologie, les fractales ou encore la société humaine. Ces exemples illustrent bien la variabilité dans les dimensions de ces systèmes : ce n'est pas la taille qui importe mais le nombre et les connexions entre ses composants. Il faut aussi souligner que complexité ne veut pas dire complication : un système peut être complexe sans pour autant être compliqué : il y a de la régularité dans les systèmes complexes. Il existe des

49 Certains concepts se fondent sur la définition d'autres pour être compréhensibles.

50 D'où l'étymologie latine *cum plexus* qui veut dire « tissés ensemble ».

systèmes complexes simples et comme le définit Morin (2005) la complexité est quelque chose qui n'est pas simple, mais qui n'est pas forcément compliqué. Eu égard à cette notion de **non** simplicité, certains auteurs préfèrent utiliser le terme simplicité qui désigne un « ensemble de solutions trouvées par les organismes vivants pour que, malgré la complexité des processus naturels, le cerveau puisse préparer l'acte et en anticiper les conséquences » (Berthoz, 2009, p. 12). Ainsi, l'originalité du vivant est de trouver des solutions qui résolvent le problème de la complexité par des mécanismes qui ne sont pas simples, mais simplexes (*Ibid.*).

Dans le domaine de l'enseignement-apprentissage des langues, le processus d'apprentissage est en soi un système complexe, puisque beaucoup de facteurs entrent en jeu et s'influencent mutuellement. Impliquant des connexions synaptiques et neuronales pour le traitement des informations par le cerveau, l'acquisition de compétences langagières est souvent le résultat d'une coordination entre divers facteurs internes et externes (âge, genre, motivation, parcours, expositions aux *inputs*, pratique, etc.). C'est pourquoi il est « difficile » de prédire les résultats d'un apprentissage. À ce point, nous tenons à préciser que le terme « difficile à prédire » ne signifie pas une « impossibilité à prédire » (Yoccoz, 1994)⁵¹. Grâce à une perspective de simplicité, on peut tout à fait modéliser le système apprentissage avec tous ses composants représentés par des liens d'interconnectivité.

51 Pour plus d'informations sur la prédictabilité des SCD, consulter : Batty & Torrens, 2001 ; Martin, Hofman, Sharma, Anderson, & Watts, 2016 ; Shahinian, 2019 ; Torres & Aguilar, 2018 ; Vidmer, Zeng, Medo, & Zhang, 2015.

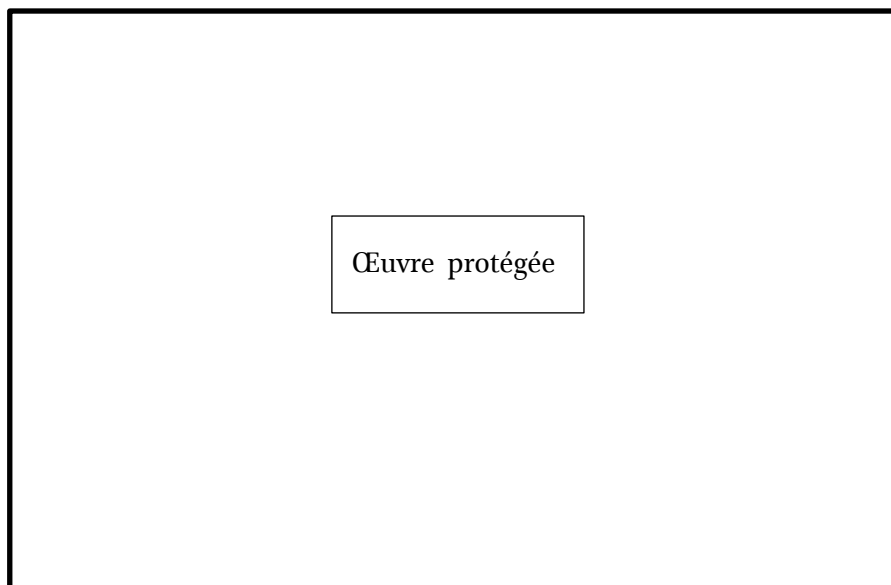


Figure 4 : illustration d'un système complexe (réseau Internet) avec les connexions entre les divers constituants.⁵²

3.3.2 Non-linéarité et linéarité d'une trajectoire

On parle de trajectoire ou cheminement pour se référer à une ligne ou une courbe décrite par une entité en mouvement (réelle ou imaginaire) dans un système dynamique. Une trajectoire linéaire est celle qui suit un cheminement en ligne droite, car elle ne connaît aucune modification de direction dans le temps. À l'inverse, une trajectoire non linéaire est constituée de ramifications formant des courbes (ou des angles) de tailles et directions variées. Au sein d'un système non linéaire, on peut rencontrer une trajectoire linéaire (par morceau) ou encore circulaire, représentée par le fait qu'un objet peut effectuer plusieurs fois les mêmes tours avant de changer de piste. Dans ce dernier cas, on parle de boucle ou d'itération lorsqu'une variable repasse sur une trajectoire qu'elle a déjà empruntée, en raison de l'influence d'autres variables. Ces boucles de rétroaction permettent également d'avoir plus d'informations sur les futures trajectoires de la variable en question. La non-linéarité des trajectoires dans un système complexe augmente la nature aléatoire, chaotique et donc l'imprévisibilité de celui-ci (Abraham & Shaw, 1992). Un même *input* aura des répercussions différentes en fonction des états et de leur environnement.

Les trajectoires qu'emprunte chaque apprenant de langue dans son développement sont individuelles et rarement (ou presque jamais) linéaires. L'apprentissage d'une L2 se déroule sur un

⁵² Source : https://commons.wikimedia.org/wiki/File:Social_Network_Analysis_Visualization.png (Consulté le 26/06/2019). Image sous licence Creative Commons (CC), Adam Majewski (2011).

temps assez long, sur des trajectoires la plupart du temps non linéaires et parsemées de boucles de rétroaction (Larsen-Freeman, 1997). Il requiert de ce fait que les différentes compétences soient travaillées simultanément et consécutivement : elles ne se fondent pas sur un ordre établi d'apprentissage de la grammaire en premier, puis du vocabulaire et enfin de la prononciation. La trajectoire de l'apprenant de langues évolue sous un régime multidimensionnel de va-et-vient entre ces différentes composantes (et d'autres encore), qui sont assujetties à des facteurs internes et externes au système.

3.3.3 État initial et état final

Un état initial (ou condition initiale) peut être considéré comme le point de départ d'un élément sur une trajectoire et est intrinsèque à chaque variable d'un système. Ce point de départ peut constituer n'importe quel moment : par exemple l'instant t où le chimiste introduit un composant chimique dans une solution afin d'en étudier la réaction, ou celui où le physicien augmente la vitesse d'un corps en mouvement pour le calcul de son énergie cinétique. Théoriquement, les conditions finales constituent le point d'arrivée d'un élément d'un système. Ce point d'arrivée peut être un état de repos d'un corps en mouvement : un troupeau de moutons qui s'arrête ou encore la bourse financière qui redevient stable. En pratique, elles n'existent pas vraiment puisque chaque élément est inlassablement en mouvement. En revanche, on note des moments d'accalmie où les « oscillations » s'arrêtent pendant un temps déterminé. Ce temps de pause est assez variable (des minutes, des heures, des jours, voire des années) et dépend également d'autres facteurs internes et externes au système. Notons qu'entre ces deux états, on assiste régulièrement à des phases intermédiaires qui peuvent être également variables dans l'espace et dans le temps. Elles consistent en des moments successifs de stabilité entre l'état initial et l'état final. À ce moment-là, on peut parfaitement décrire le comportement du système, parce que chaque élément du système laisse apparaître ses caractéristiques intrinsèques. Les liaisons qui existent entre les diverses caractéristiques du système sont étudiées et analysées. Les conditions initiales sont difficiles à déterminer, mais on peut identifier des caractéristiques qui lui sont propres à un instant donné. Un système complexe et dynamique connaît plusieurs conditions initiales et finales en fonction du changement du comportement dans le temps (Abraham & Shaw, 1992).

Les trois états précédemment décrits se retrouvent également dans un système tel que l'apprentissage d'une L2 et sont différents en fonction de chaque apprenant (*background*, L1, activités scolaires et extrascolaires formelles ou informelles, motivation, personnalité, etc.). Parmi ces derniers, l'état initial est très important pour le chercheur, car il constitue le facteur qui pourrait lui permettre de pouvoir « prédire » les trajectoires de développement L2 des apprenants. Ici plus précisément, l'état initial d'une variable consiste en l'ensemble de caractéristiques qu'il possède à un instant précis. Par exemple, les états initiaux d'un apprenant d'anglais seront constitués des moments où il a commencé à prendre des cours d'anglais, à regarder des séries télévisées en anglais, à visiter des pays anglophones, à avoir des conversations en anglais, etc. Ainsi, le début de l'apprentissage des différentes compétences de la langue apprise peut être considéré comme un constituant d'un état initial d'un apprenant. Les cycles successifs que parcourt l'apprenant dans un système scolaire (formel), avec des cours donnés par les enseignants, constituent des états intermédiaires et l'état « final » représente le moment où l'apprenant devient capable de produire des *outputs* compréhensibles à l'oral et à l'écrit (Swain, 1995 ; Krashen, 1998). Bien évidemment, le cycle continue puisque l'apprenant poursuit son développement à mesure qu'il pratique la langue. Chaque apprenant de langue dispose ainsi d'un état initial qui lui est propre et qui diffère pour chaque variable. Cela pourrait également être des moments de début et de fin d'une trajectoire (activité), qui est déterminés par un autre SCD (par exemple, la période du Baccalauréat à la fin du cycle universitaire). C'est la multiplicité de ces conditions qui engendre les différences dans les résultats que les chercheurs observent en étudiant différents systèmes d'apprentissage. Enfin, nous avons déjà soulevé le fait qu'un système complexe et dynamique est très sensible aux conditions initiales et qu'un changement minime est capable de provoquer des conséquences notables.

3.3.4 Dynamisme et stabilité

Le dynamisme d'un objet est sa caractéristique d'être en perpétuel mouvement dans un « espace de phases » et très souvent sous l'impulsion de forces internes ou externes. Un corps dynamique possède des propriétés cinématiques (et cinétiques⁵³) et qui lui permettent d'évoluer de façon constante. Pour ce faire, il a besoin d'énergie afin d'assurer ses mouvements. Les cellules

⁵³ Relatif au mouvement et par extension à la vitesse issu d'un mouvement.

vivantes, le temps (météorologie), les pendules, l'évolution d'une population, etc. constituent quelques exemples d'entités dynamiques que nous connaissons. En revanche, lorsqu'un corps n'est plus en mouvement, il est qualifié de stable (qui ne bouge pas). Dans la logique de la théorie des SCD, dynamisme et stabilité constituent deux aspects d'une même question, puisque les éléments constitutifs d'un système connaissent ces deux états entre un moment et un autre. Lorsqu'une condition de stabilité intervient à un instant donné, la position d'une variable peut être prévisible, de même que le comportement du système dans lequel la variable évolue (Yoccoz, 1995).

En didactique des langues, l'apprentissage fonctionne de façon analogue : l'apprenant connaît des moments de stabilité et des moments de dynamisme constants. Trois catégories de situations peuvent alors se présenter pour l'apprenant : une phase de stabilité et deux phases dynamiques. La première est dynamique et se produit lorsque l'apprenant réalise des progrès où il/elle intègre les éléments appris (ce progrès passe aussi par des interlangues). La deuxième est stationnaire (plateaux) et consiste en des périodes où il/elle n'assimile plus « rien » (non perceptible). Dans ce cas, l'apprenant utilise uniquement des éléments de langage déjà intériorisés, on parle de nativisation⁵⁴ langagière (Anderson, 1983). Enfin, la troisième situation, également dynamique a lieu lorsque l'apprenant de langues tend à oublier des éléments qu'il/elle a déjà appris (attrition), ou stagne dans son apprentissage (fossilisation⁵⁵) parce qu'il/elle ne pratique plus la langue. Généralement, ce cas de figure s'étend dans le temps et est également lié à des facteurs socio-psycho-affectivo-cognitivo-motivationnels. Il faut tout de même préciser que souvent, en réalité, il y a des choses qui sont apprises mais elles sont juste imperceptibles sur le moment. Les enseignants et même les apprenants ne réalisent pas les progrès opérés. Il suffirait par exemple de (ré)activer les connaissances pour se rendre compte des assimilations (affiner la perception des sons ou systématiser l'utilisation d'une structure grammaticale).

3.3.5 Prédicibilité et imprédictibilité

Lorsqu'une variable d'un système suit un cheminement linéaire, il n'y a nul doute que cette trajectoire est prédictible. Mais lorsque ce même élément se retrouve sur une piste non-linéaire, il devient très difficile de prédire les trajectoires, du moins sans outils de mesures adéquats (formules

⁵⁴ Nous y reviendrons plus loin.

⁵⁵ Blocage ou stagnation au niveau de l'apprentissage (voir section La théorie des SCD et la notion d'interlangue page 21).

d'équation différentielle). L'imprédictibilité des éléments évoluant dans un système complexe est une conséquence évidente de la nature dynamique et non-linéaire des trajectoires et surtout des attracteurs chaotiques. Nous avons toutefois souligné que la prévisibilité d'un système complexe et dynamique est tout à fait possible pendant les conditions de stabilité, en raison de la régularité émergente. Lorsque les informations sur les états initiaux sont suffisamment connues, le chercheur peut prédire le développement en L2 d'un apprenant, mais les chances de réussite de prédiction y sont très maigres (Yoccoz, 1994). Toutefois, comme nous le verrons plus loin, certains procédés comme la rétrodiction et l'identification des attracteurs peuvent faciliter la prédictibilité de formes linguistiques émergentes chez un apprenant.

3.3.6 Autorégulation / auto-organisation

Tout SCD rencontre des phases d'organisation et de chaos au cours de son évolution. Ces deux phases peuvent être causées par des organismes aussi bien internes qu'externes au système. Lorsque les conditions initiales changent en raison de l'intégration ou de la sortie d'une composante, le système subit un bouleversement, qui peut être un changement de direction (par exemple des oiseaux), de quantité (par exemple l'augmentation du CO₂ dans l'air ou la diminution de la pression atmosphérique), etc. Le système se lance dans sa course, dans son dynamisme, les trajectoires reprennent leurs cours, à la recherche d'une autre condition de stabilité. On parle alors d'autorégulation ou retour à une situation normale/originelle. Lorsque l'on laisse tomber une goutte d'eau au milieu d'une bassine contenant déjà une quantité d'eau et dont la surface est calme (plane), on assiste immédiatement à la naissance d'ondes (vagues) qui se propagent depuis cet endroit jusqu'aux rebords de la bassine. Ici, la goutte d'eau, externe au système, constitue l'élément perturbateur et créateur du désordre. Après la phase « chaotique », la surface de l'eau retrouve son état initial (phase d'accalmie). Nombre de phénomènes naturels (catastrophes) sont dotés d'un système d'autorégulation, puisqu'ils ne nécessitent pas l'intervention d'un agent extérieur pour leur fonctionnement.

L'apprenant connaît certes des conditions d'auto-organisation dans son développement. La réception des éléments linguistiques comme un nouvel élément grammatical, lexical, ou de prononciation, etc. peut être, au début, chaotique pour l'apprenant, mais avec le temps tous ces éléments sont maîtrisés et procurent des capacités/compétences nécessaires à la réalisation de

tâches langagières. L'autorégulation nécessite ainsi un temps d'assimilation et d'adaptation durant lequel les divers éléments linguistiques interagissent et mûrissent.

3.3.7 Attracteur et répulsif

« Région de l'espace des phases vers laquelle converge, au fur et à mesure que le temps s'écoule, le point représentatif d'un système dynamique, quel que soit son état initial », telle est la définition du mot attracteur dans le dictionnaire Larousse en ligne (entrée Attracteur⁵⁶). Cette notion est cruciale dans les tentatives de prédiction de trajectoires dans un système dynamique, car elle permet d'entrevoir une certaine régularité, un certain ordre dans le chaos. Ayant pour vocation d'attirer les autres trajectoires dynamiques autour de lui, un attracteur est, nous l'avons déjà évoqué, très sensible aux conditions initiales, et peut se déplacer de points en points. De même qu'il peut y avoir plusieurs états dans un système, les attracteurs peuvent également être nombreux, de formes et puissances variables dans l'espace-temps. On distingue donc des attracteurs fixes⁵⁷, quasi-mobiles et chaotiques ou étranges (Abraham & Shaw, 1992). Les SCD développent souvent des attracteurs étranges, rendant ainsi le système encore plus imprévisible. Un fleuve peut être considéré comme un parfait exemple d'attracteur, car il attire les mouvements de ruissellement des cours d'eau⁵⁸. La figure ci-dessous (Figure 5) illustre un attracteur étrange de Lorenz (1963) et représente un ensemble de lignes reconstitué artificiellement à partir des différentes positions occupées par les multiples éléments d'un système. Connaître les attracteurs d'un système complexe et dynamique trouve son intérêt dans le fait qu'il est impossible d'étudier le comportement individuel de la multitude de variables qui le compose. Lorsqu'ils sont identifiés, ces attracteurs donnent des informations plus ou moins précises sur ce qui émergera du chaos.

Parmi les différents éléments (internes et externes) qui composent le système d'apprentissage, la motivation peut constituer un attracteur fort. En effet, on peut considérer que toutes les activités de l'AIAL que pratiquent les apprenants suivent le chemin des loisirs et du désir qu'ils ont à comprendre les discours et à parler la langue. De même qu'un attracteur constitue une force

56 <https://www.larousse.fr/dictionnaires/francais/attracteur/6300>

57 Un exemple d'attracteur fixe simple est le degré 37 en Celsius de la température corporelle, car quelles que soient les fluctuations, la température redevient normale (stable) (Abraham & Shaw, 1992, p. 209).

58 La même chose est applicable pour les lacs et les océans.

motrice dans un SCD, l'inverse est aussi possible : on parle de répulsif⁵⁹ (Sockett, 2014). Souvent, on utilise la métaphore des vallées pour représenter l'attracteur et les sommets des montagnes comme des répulsifs. Un répulsif sera donc un élément qui éloigne les autres éléments autour de lui, qui les repousse vers d'autres trajectoires. D'ailleurs, un attracteur peut freiner voire même inhiber le développement d'une L2 alors qu'un répulsif peut le relancer (l'apprentissage par exemple). Sockett donne l'exemple des sous-titres : « *the use of native language subtitles is a good example of this type of attractor as it first facilitates access to authentic documents but can later prove to be a distraction from language processing* (2014, p. 22) ». Selon Roussel et Gaonac'h (2017), les sous-titres peuvent en effet constituer un couteau à double tranchant pour l'apprenant de langues : ils peuvent aider à la compréhension d'une LE, comme ils peuvent aussi la dégrader (p. 77). Sur le plan de la prononciation, les sous-titres aident à la création d'une image phonologique du mot entendu, mais tout dépend d'autres facteurs dont le niveau de l'apprenant, le mode de visionnage, la charge cognitive inhérente à chaque tâche (p. 82).

59 Traduit de l'anglais repeller.

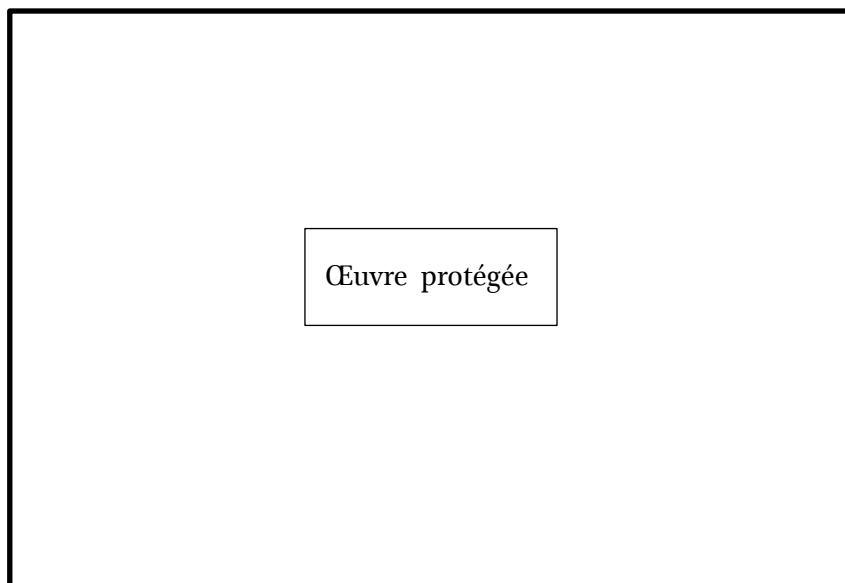


Figure 5 : attracteur étrange de Lorenz (1963)⁶⁰.

3.3.8 Émergence et coadaptation

Si nous appliquons tous les éléments caractéristiques précédemment présentés aux items qui composent un SCD, nous devons nous attendre à une apparition soudaine (ou progressive) d'un phénomène quelconque (structure, comportement, objet, concept, etc.) et dont les propriétés étaient inexistantes au niveau de l'ensemble comme à celui des éléments individuels. On parle d'émergence. Il faut retenir deux types d'émergence : la première consiste à obtenir de l'ordre à partir du chaos et la seconde consiste en du désordre issu d'une organisation. Ce phénomène est : « un processus ouvert de régénération permanente de soi (d'autopoïèse), sans qu'un état final en soit prédéterminé, ni même souhaitable » (Toffoli, 2018, p.26). Prenant en compte l'apprentissage des langues, Ellis (2002) considère l'environnement dans lequel évolue l'apprenant comme facteur d'émergence. La langue constitue alors : « [...] *A dynamical system that emerges within a social context through interactions of cognitive, neurological, and environmental systems and subsystems across nested timescales* » (Kohnert, 2007, p. 13).

Quant à la coadaptation, elle représente des événements au cours desquels émergent des phénomènes étranges et inattendus, à partir d'interactions complexes et fusionnelles des éléments du système. Il s'agit de situations dans lesquelles un élément du système incorpore ou accommode

⁶⁰ Repéré sur : https://commons.wikimedia.org/wiki/Lorenz_attractor#/media/File:Lorenz_attractor.svg (consulté le 16 février 2019). Source : image sous licence Creative Commons (CC), Dschwen (2006).

des propriétés d'un ou plusieurs autres éléments. Dans ce cas précis, les actions de chaque apprenant influent sur tous les autres actions de personnes avec lesquelles il est en contact. Deux camarades de classe qui sont proches évoluent ensemble tout en s'influçant mutuellement et simultanément. Les structures langagières sont également de nature émergente et les nombreux facteurs qui constituent les activités de l'AIAL en témoignent. Les *outputs* des apprenants de langues « jaillissent » de manière inattendue à partir des *inputs* multiples et variés auxquels ils s'exposent. Nul ne peut prédire avec précision les résultats d'un apprentissage à proprement dit, ni la description totale de son fonctionnement. L'émergence est le résultat de l'interconnectivité des différents sous-systèmes, sous l'impulsion des attracteurs. La notion d'accommodation est un exemple de coadaptation dans la mesure où un natif et un non-natif adaptent mutuellement leur discours en vue d'une meilleure intercompréhension : les deux systèmes changent simultanément pour une cause commune. La notion d'émergence est tellement centrale dans l'étude des SCD qu'un courant épistémologique a même vu le jour : il s'agit de l'émergentisme (Miras & Narcy-Combes, 2014). Nous en profitons pour évoquer à ce sujet la journée Notions en Questions (NEQ) sur les courants émergentistes en didactique des langues et dont est issu un numéro spécial de *Recherches en Didactique des langues et cultures : les Cahiers de l'Acedle*⁶¹ coordonné par Miras (2017)⁶².

Cette section sur la présentation des éléments caractéristiques d'un SCD nous permet d'en comprendre le fonctionnement. Nous retenons qu'un système est dit complexe et dynamique lorsque ses éléments constitutifs sont interconnectés et évoluent dans l'espace-temps suivant des trajectoires très souvent non-linéaires autour d'attracteurs ou de répulseurs. Il est sensible aux conditions initiales et admet des phases de stabilité et d'instabilité en fonction des interactions se déroulant en son sein et est par ailleurs assujéti à l'émergence de phénomènes le plus généralement imprédictibles. Les conséquences de ces changements proviennent de nombreuses causes, à la fois incertaines, itératives et rétroactives, et qui n'ont jamais les mêmes conséquences. Dans la section suivante nous allons aborder spécifiquement la théorie des SCD dans le domaine de la didactique des langues. C'est la partie qui nous intéresse le plus afin de soutenir nos résultats, car l'apprenant est considéré comme un sous-système complexe évoluant au sein d'un plus grand

61 Association des chercheurs et enseignants didacticiens des langues étrangères.

62 Numéro disponible sur : <https://journals.openedition.org/rdlc/1082>

SCD, où de multiples facteurs interagissent, en vue de faire émerger des structures linguistiques⁶³ (Cameron & Larsen-Freeman, 2007). Suite à cela, nous traiterons des notions d'interlangue et de variabilité/différence individuelle, avant d'aborder, pour finir, la question des limites liées à l'emploi de la théorie en SHS de manière générale et en didactique des langues/linguistique appliquée en particulier.

3.4 La théorie des SCD dans le domaine de l'apprentissage des langues

Les premiers travaux qui se reposent sur la théorie des SCD en linguistique appliquée datent d'environ une décennie et sont notamment relevés dans les publications de chercheurs tels que Cameron & Larsen-Freeman, 2007 ; van Geert, 2009 ; de Bot *et al.*, 2007 ; van Dijk, Verspoor, & Lowie, 2011 ; Sockett, 2014 ; Kusyk, 2017 ; Toffoli, 2018. En effet, ces chercheurs s'accordent pour considérer l'apprentissage des langues comme étant un SCD, car il est régi par les interconnexions et interactions de nombreux facteurs psychologiques et sociaux. L'apprentissage n'est plus considéré comme étant un processus linéaire et programmable en structures scolaires, mais comme « un ensemble complexe de processus largement déterminé par le social d'une part, mais aussi la (neuro)psychologie de l'individu et donc, de ce fait, non planifiable de façon universelle » (Narcy-Combes & Miras, 2012, p. 42). De même, Cameron & Larsen-Freeman (2007) considèrent l'apprenant, la communauté d'apprentissage et l'apprentissage des langues comme des systèmes complexes imbriqués les uns dans les autres. L'apprenant de langue évolue ainsi dans un environnement complexe et dynamique, où plusieurs facteurs intrinsèques et extrinsèques sont interconnectés et interagissent conjointement. Pour sa part, Toffoli (2018, p. 28) considère l'apprenant de langue et les dispositifs d'apprentissage (servant de contexte) comme deux SCD qui sont circonscrits au sein du système d'apprentissage des langues, qui lui-même est déjà complexe et dynamique. Cette vision de l'apprentissage va dans le même sens et correspond exactement au fonctionnement des SCD, dans la mesure où, très souvent, un système complexe est composé de sous-systèmes, eux aussi complexes. La théorie des SCD est utilisée dans ce champ, surtout en

⁶³ Elles peuvent être phonétiques, morphologiques, syntaxiques, sémantiques, etc.

didactique des langues, comme une métaphore⁶⁴ (Dewaele, 2002, Toffoli, 2018), en vue de décrire, comprendre et « prédire » le développement langagier des mécanismes qui sous-tendent les *outputs* en L2 (Larsen-Freeman & Cameron, 2008).

Dans cette conception, le terme développement (ou processus) est préférable à celui d'apprentissage ou d'acquisition, car dans le cadre des SCD, l'apprenant ne fait pas que progresser (de Bot & Larsen-Freeman, 2011). Celui-ci connaît également des stades de stagnation ou de stagnation, selon la nature des éléments qui « perturbent » le système et la manière dont le système se réorganise. Les activités que fait chaque apprenant sont diverses et fournissent des *inputs* variables en termes de quantité et de qualité, si bien que certaines expositions seront fructueuses tandis que d'autres ne le seront pas. On trouve souvent dans la littérature sur l'apprentissage aussi bien de la L1 que de la L2 que les pratiques de lecture améliorent celles de l'écriture (Reuter, 1995 ; Giguère, Giasson & Simard, 2002). En revanche, selon ces auteurs, les pratiques d'écriture ne montrent pas de véritables avancées sur la lecture, comme le précisent Giguère, Giasson & Simard (2002) : « il est possible qu'un élève soit capable de reconnaître un mot en lecture sans nécessairement être capable d'orthographier correctement ce mot en écriture » (p. 24), car « la reconnaissance d'un mot peut s'effectuer seulement à partir de quelques traits distinctifs ; il est aussi possible de se servir des indices contextuels pour identifier le mot » (*ibid.*). L'influence de la lecture sur les capacités de rédaction paraîtrait plus facile que l'inverse et serait unidirectionnelle puisque : « À la différence de la lecture, l'orthographe suppose une connaissance exacte de chacune des lettres formant le mot ainsi que de l'ordre dans lequel ces lettres sont placées » (*ibid.*). D'autre part, des recherches montrent également qu'une exposition massive aux *inputs* sonores améliorerait la prononciation en L2 (Matter, 2006 ; Moradi & Shahrokhi, 2014). Pourtant, des études ont aussi mis en évidence l'intervention d'autres variables comme par exemple, la correspondance entre l'orthographe (la graphie) et la prononciation de phonèmes (Saletta, 2015 ; Mok *et al.*, 2018). Cette notion de correspondance grapho-phonémique, nous l'avons vu, a des influences variées selon les langues qui sont prises en considération. Nous voyons par-là que la notion de cause à effet n'est pas systématique dans un système complexe, car une conséquence peut avoir plusieurs causes et qu'une cause peut engendrer plusieurs effets. Aussi, apprendre une langue n'est pas uniquement caractérisé par une accumulation « continue » de

64 Nous discutons le choix de ce terme dans la section suivante.

savoirs et de savoir-faire. Il y a des moments au cours desquels des apprentissages ne sont pas visibles et d'autres où des choses apprises sont oubliées. De ce fait, le développement d'une L2 est considéré comme un processus complexe et dynamique, au sein duquel l'apprenant est soumis à une masse de variables interconnectées et qui évoluent dans le temps, faisant émerger des comportements linguistiques plus ou moins prédictibles (Cameron & Larsen-Freeman, 2007 ; de Bot, 2008 ; van Geert, 2009).

En effet, la diversité des activités dans lesquelles les apprenants s'engagent (Toffoli & Sockett, 2010), les multiples facteurs intervenants dans l'acquisition phonético-phonologique (Piske, MacKay & Flege, 2001) et la complexité liée aux développements d'une langue seconde (de Bot & Larsen-Freeman, 2011), rendre plus difficiles les recherches dans le domaine de la linguistique appliquée. Le cadre de ce nouveau paradigme, la théorie des SCD, décrit l'acquisition de la langue comme étant alimentée par diverses ressources linguistiques « bricolées ensemble en temps réel » alors que les apprenants négocient leur identité linguistique selon les interlocuteurs et leurs activités respectives (Boula de Mareüil, 2010 ; de Bot & Larsen-Freeman, 2011 ; Moyer, 2013). L'apprentissage des langues est un processus considéré comme long, désordonné et beaucoup moins prévisible qu'on ne le pensait. Tous les mécanismes psychologiques, cognitifs et sociaux opérationnels sont en mouvement, leur équilibre changeant continuellement en réponse à l'environnement (Lowie, Caspi, van Geert & Steenbeek, 2011 ; Van Geert, 2009). Dans cette dynamique, il est difficile d'espérer que les apprenants de langue puissent s'approprier les éléments qu'on leur enseigne selon un rythme défini et qui est bien évidemment différent des leurs, ni d'ailleurs prétendre à ce que tous les apprenants progressent au même rythme. Il aurait fallu plusieurs années aux pédagogues et aux didacticiens pour se rendre compte de l'erreur de traiter tous les apprenants de la même manière et que la différenciation pédagogique serait une clé pour pallier ce problème.

Dans la lignée des travaux en apprentissage des langues, Dewaele (2002) se sert des principes de la théorie des SCD afin d'expliquer l'usage différencié des pronoms d'adresse « vous » et « tu » dans les interlangues en français par des apprenants non natifs. En effet, l'auteur considère le rapport entre l'usage sociolinguistique du vouvoiement et celui du tutoiement comme relevant d'un système extrêmement complexe et qui génère des conséquences variées. Trente femmes et 23 hommes âgés de 21 à 65 ans venus d'horizons divers et étudiant à l'université de Londres au

Birkbeck College ont participé à cette étude (Dewaele, 2002). Certains participants avaient d'autres L1 à côté de l'anglais : espagnol, italien, néerlandais, allemand, lingala, frioulan, ouro, farsi, créole mauritien, et arabe. Ils ont effectué des activités simulées d'interviews en binômes, qui ont généré de la parole spontanée en français. Le corpus ainsi enregistré et analysé comporte 85000 mots dont 771 occurrences de « vous » et de « tu », en position sujet. Ensuite, deux conversations en situations réelles (client et fournisseur) ont également été enregistrées et ont généré 1769 mots dont 26 occurrences de « vous » et de « tu », toujours en position sujet. Les résultats montrent que les natifs et les non natifs utilisent les marqueurs de vouvoiement et de tutoiement de manières très différentes. Les natifs utilisent plus le « tu » et son usage a davantage d'influence sur les situations de communication. Toutefois, ces observations n'ont pas pu être confirmées statistiquement et l'auteur reconnaît que bien d'autres éléments interviennent dans le système linguistique de l'apprenant qui est « dans un état de flux permanent. Le modèle que le langage présente est le même modèle fractal que celui d'autres systèmes complexes, dynamiques, non-linéaires » (paragr. 46). Dans cet article, la discussion des résultats en rapport avec la théorie des SCD nous intéresse tout particulièrement, car cela s'applique à bon nombre d'aspects de la didactique des langues. Que ce soit dans l'apprentissage de la grammaire, du lexique ou encore de la phonétique au sens large, l'intervention et l'interaction entre facteurs externes et internes sont loin d'être négligeables. Parmi les facteurs externes, l'auteur cite « le contexte politique ou idéologique de l'époque où l'apprentissage se déroule » (paragr. 47). Quant aux facteurs internes, l'auteur évoque la mémoire, l'affect et la motivation et souligne qu'il « existe en outre des facteurs internes encore mal connus qui peuvent expliquer le progrès non-linéaire en langue maternelle [...] comme en langue étrangère » (paragr. 49). Dans la sous-section suivante, nous expliquerons en détails la notion d'interlangue et nous montrerons en quoi elle constitue une manifestation intégrante des SCD au regard du développement langagier chez l'apprenant.

3.4.1 La théorie des SCD et la notion d'interlangue

Cette sous-section a pour objectif de montrer que l'interlangue de l'apprenant est une manifestation systémique au sein du processus d'apprentissage. Elle permet également de porter un regard différent sur la notion et de mieux comprendre son émergence, son fonctionnement et

son évolution dans le temps. Connue sous diverses dénominations⁶⁵, la notion d'interlangue en didactique des langues (Selinker, 1972) se réfère à un mélange de structures linguistiques (syntaxe, vocabulaire, prononciation) se situant entre la L1 et la langue cible (L2). C'est un concept central en psycholinguistique de l'acquisition et il est notamment associé à la notion de fossilisation (Selinker, 1972 ; Dewaele, 2003). La fossilisation serait selon Selinker (1972), une forme langagière non grammaticale de la L2 et correspondrait à une structure psychologique latente dans le cerveau. Elle serait un blocage au niveau de l'apprentissage et serait directement reliée à l'interlangue de l'apprenant. Selon Larsen-Freeman (1997, p. 152), « *the absence of learning in a language, fossilization, occurs when the learners' grammar system becomes closed and settles down to a fixed point attractor* ». L'hypothèse de l'interlangue postule que la parole en L2 est peu conforme à ce que l'on entend des locuteurs natifs de la langue cible (Selinker *et al.*, 1975, p. 140). Ainsi, Besse & Porquier (1991) définissent l'interlangue comme :

[...] la connaissance et l'utilisation « non-natives » d'une langue quelconque par un sujet non natif et non équilingue, c'est-à-dire un système autre que celui de la langue cible mais qui, à quelque stade d'apprentissage qu'on l'appréhende, en comporte certaines composantes. C'est ce que nous avons précédemment appelé grammaire intériorisée par l'apprenant, et que nous appellerons ici interlangue (Besse & Porquier 1991, p. 216).

Outre les caractéristiques qui gouvernent l'interlangue selon Selinker⁶⁶, Galligani en propose trois autres qui sont « celles de dynamique (instabilité, variabilité), de système (systématicité) et de spécificité (perméabilité) » (Galligani, 2003, paragr. 8). On voit que les Recherches en acquisition des Langues (RAL) concernant l'interlangue s'accordent toutes sur certains aspects notoires de ce phénomène. Il s'agit de la progressivité, de la complexification et du caractère transitoire du concept, comme le souligne Galligani (2003) :

Quelles que soient les terminologies qui opposent les diverses conceptions de l'interlangue, on retrouve partout l'idée d'une structuration progressive des connaissances d'apprenants en langue étrangère, d'une complexification par laquelle la compétence intermédiaire se rapproche de l'objectif fixé au départ par les apprenants, à savoir la langue cible (Galligani, 2003, paragr. 5).

65 « Système approximatif, compétence transitoire, dialecte idiosyncrasique, système intermédiaire, interlangue, système approximatif de communication, langue de l'apprenant ou système approché » (Besse et Porquier 1991, p. 216).

66 « First, the stability over time of certain errors and other surface forms in learner-language systems. Second, the mutual intelligibility that appears to exist among speakers of the same IL. Third, the phenomenon of backsliding or the regular reappearance in bilingual speech of fossilized errors that were thought to be eradicated. And fourth, the systematicity of the IL at one particular point in time » (Selinker *et al.*, 1975, p. 141 cité par Galligani, 2003, paragr. 8).

Ces termes nous rappellent sans doute les caractéristiques d'un SCD et nous verrons tout de suite que l'interlangue en est un. Bien que certaines causes de ce phénomène soient documentées, il n'en reste pas moins que d'autres sont encore mal connues. En effet, certains auteurs postulent qu'elle est liée au phénomène de la nativisation, qui désigne un processus par lequel un apprenant de langue construit son apprentissage en se reposant sur des éléments langagiers de sa /ses L1 (Demaizière & Narcy-Combes, 2005 ; Andersen, 1983). Or, les transferts et les interférences depuis la L1 n'expliquent pas toutes les interlangues. Nous pouvons citer le fait que certains apprenants ont des difficultés à prononcer des sons d'une L2, alors que ces derniers existent dans leur L1. Par exemple, il a été attesté que les hispanophones (Catalans plus précisément) parviennent difficilement à réaliser le son /z/ du français qui est systématiquement produit comme /s/, alors que le /z/ appartient au système phonique du catalan (Racine, 2015). D'autres causes pourraient inclure la surgénéralisation, la distorsion, la perception, etc. (Dewaele, 2002). Mais si l'interlangue de l'apprenant de langue est souvent considérée comme le maillon faible de la chaîne dans le processus d'apprentissage et que celui-ci doit y remédier au plus vite, afin d'atteindre une production proche de celle des natifs, certains auteurs pensent au contraire que l'interlangue est un passage obligatoire pour atteindre un niveau de langue « idéal ». Dans une étude longitudinale sur l'acquisition des articles définis en anglais par des apprenants tchèques et slovaques, Young (1996) utilise la théorie des SCD pour expliquer des interlangues dans ses résultats. Il remarque des variations libres et aléatoires, suivies des variations de plus en plus systématiques dans les stades ultérieurs. Dans une lignée analogue, selon Dewaele (2002) :

L'affirmation qu'une interlangue ne se développe pas de façon linéaire peut paraître surprenante. L'effort que l'apprenant fait pour s'approprier une langue étrangère n'est-il donc pas constant ? Un effort d'apprentissage donné ne se traduit-il pas par un progrès concomitant ? La réponse aux deux questions est « non » (Dewaele, 2002, paragr. 47).

Ce passage montre de façon explicite que l'interlangue était considérée comme un phénomène linéaire et les efforts d'apprentissage constants et simultanés. Nonobstant, l'explication que donne l'auteur pour appuyer sa réponse nous interpelle au vu de ce que l'on sait désormais sur l'interlangue :

L'interlangue d'un apprenant (ainsi que la langue maternelle utilisée régulièrement) est donc un système dynamique auto-organisateur, en interaction constante avec son environnement, qui gagne progressivement en complexité. Au début de son développement, l'interlangue est caractérisée par un certain degré de chaos ou de variation libre (Dewaele, 2002, paragr. 47).

Nous l'avons déjà évoqué, l'interlangue est un phénomène complexe, dynamique, et qui se déroule dans le temps, en interaction avec des variables de son environnement. En effet, nous pouvons considérer les solutions de remédiation à l'interlangue⁶⁷ comme erronées et inadaptées, car les productions intermédiaires qui surviennent au cours du développement de la L2 de l'apprenant proviennent de l'émergence de nombreux facteurs du système d'apprentissage. Ce passage de l'état initial jusqu'à l'émergence d'une production finale s'effectue en plusieurs étapes (états) et ceci de manière assez chaotique. Une possibilité de corriger la fossilisation peut rester vaine (Luzar, 2012). Il s'agit là d'un changement important de paradigme par rapport à l'idée prédominante selon laquelle « l'interlangue est un système truffé d'erreurs⁶⁸, qui évolue sur un chemin prédéterminé vers un état mal défini et fossilisé » (Moyer, 2013, p. 49). Ce regard plutôt négationniste porté sur la formation de l'interlangue a évolué grâce aux travaux en acquisition des langues secondes, qui montrent leur importance en tant que socle pour l'appropriation de la L2. Aujourd'hui, nous savons que la notion d'interlangue est déterminante dans le processus de développement de l'apprenant, car elle devient « un instrument de communication authentique » pour ce dernier (Dewaele, 2002, paragr. 48) et lui permet d'évoluer vers des compétences proches des natifs et compréhensibles par ces derniers.

La théorie des SCD aide à comprendre comment l'interlangue se forme et quelles en sont les interférences. Cependant, caractériser l'interlangue des apprenants demeure une tâche rude (Dausendschön-Gay, 2003), car ce sont des systèmes instables et qui évoluent dans le temps. Pour ce faire, le chercheur se doit de l'analyser en tant que tel (c'est-à-dire comme un SCD) avec la prise en compte des caractéristiques inhérentes en passant notamment par l'inventaire des facteurs en jeu et leurs interactions, l'étude des états initiaux et intermédiaires, l'identification des attracteurs éventuels et l'observation de la nature des comportements émergents.

Nous venons de montrer que le phénomène de l'interlangue, longtemps documenté en didactique de langues et qui était perçu comme un aspect négatif dans le processus

67 Par exemple : les pratiques de répétitions effrénées en vue de corriger la prononciation des apprenants ou l'analyse contrastive des systèmes linguistiques de la L1 et de la L2 peuvent permettre d'identifier les interférences dans l'apprentissage, mais n'aident pas à les expliquer.

68 « Inscrites dans la problématique des contacts de langue, les notions d'Interlangue et de communication exolingue sont parfois considérées, comme le rapporte Coste (1989, p. 22), comme des 'objets-tabous', des 'marécages langagiers' ou encore des 'bas quartiers de la communication' par 'les puristes' et les vrais défenseurs de la langue » (Rosen et Porquier, 2003, paragr. 2).

d'apprentissage, constitue une pierre angulaire pour l'apprenant et aussi pour l'enseignant de langues. Aussi avons-nous évoqué son aspect complexe et dynamique, parce qu'assujéti à l'influence d'autres sous-systèmes de son environnement et réciproquement. Ses caractéristiques chaotiques nous amènent à le considérer comme un SCD à part entière dans notre travail. Dans la sous-section suivante, nous présentons deux autres notions liées aux SCD en didactique des langues : les différences individuelles et la variabilité.

3.4.2 La théorie des SCD, les différences individuelles et la variabilité

Cette sous-section rend compte des phénomènes de différences individuelles et de la variabilité en acquisition de langues au regard des SCD. Ces deux notions ont des caractéristiques inhérentes aux SCD qui sont chaotiques, émergentes et imprévisibles. Chaque apprenant de langues est unique, comme le précise Sockett (2014, p. 23) : « *the experience of each learner is unique in terms of individual differences and repertoires of input and interaction* ». Il évolue différemment dans un système langagier qui est complexe et dont il n'est pas le seul acteur. Beaucoup d'agents évoluant également dans le système impactent son développement et inversement. Nous mentionnons⁶⁹ ici les différences individuelles et les variabilités reconnues dans la littérature, au regard du sujet nous intéresse, c'est-à-dire la prononciation. Ainsi, Moyer (2013, p. 51-80) identifie trois types de différences individuelles qui sont :

- les différences individuelles intrinsèques (aptitudes, mémoire, préférence hémisphérique, talent et expérience musicaux, styles et stratégies d'apprentissage, différences entre les sexes) ;
- les facteurs socio-psychologiques (identité, acculturation et égo linguistique, motivation, attitudes et investissement de l'apprenant) ;
- l'expérience et l'exposition (durée du séjour, utilisation de la L2, dominance linguistique, attrition de la L1).

Quant aux facteurs responsables de la variabilité, nous pouvons citer : l'âge du locuteur, la classe sociale, l'âge du début de l'apprentissage/l'acquisition, les instructions officielles (écoles), la personnalité, les attitudes envers la langue et les locuteurs natifs de cette langue, la fréquence d'exposition et d'usage, (Hughes, Trudgill & Watt, p. 40-42).

Il est à souligner qu'à notre connaissance, il n'existe pas une liste exhaustive (canonique) de différences individuelles ou de variabilités. Chaque auteur propose sa classification, mais il y a un

69 Sans les définitions ou exemples. Pour plus de détails le lecteur peut consulter (Flege, Munro, & MacKay, 1995 ; Hughes, Trudgill, & Watt, 2012 ; Moyer, 2013 ; Piske, MacKay, & Flege, 2001).

certain consensus concernant bon nombre d'entre ces facteurs. Par exemple, concernant la catégorisation de Moyer, nous aurions mentionné les deux types de motivation à savoir celle intrinsèque et celle extrinsèque. Même si les chercheurs s'accordent pour dire que les différences individuelles sont aux fondements des variabilités dans la prononciation, nous pensons que certaines variabilités impactent également les différences individuelles, notamment celles qui sont extrinsèque (c'est-à-dire socioculturelles).

Ces caractéristiques de la différence individuelle empêchent non seulement toute sorte de généralisation concernant les profils des apprenants, mais rendent également difficile la description exacte de la manière dont chaque apprenant réagit à de l'*input* langagier. En fait, il y a beaucoup de variations dans le processus et les résultats, même parmi les apprenants qui ont été exposés à la langue au même âge. Les différences individuelles, tant intrinsèques qu'extrinsèques, expliquent cette variation et sont considérées comme des attracteurs de divers degrés pour chaque apprenant de L2 (Ellis, 2004 ; Dörnyei, 2009) De ce point de vue, les apprenants de langues qui suivent le même nombre de cours, avec les mêmes ressources et les mêmes enseignants n'auront jamais des compétences langagières identiques, car leurs capacités cognitives et les activités qu'ils pratiquent en dehors de la classe sont tout autant diversifiées. Par conséquent, les notions de complexité, d'états, de trajectoires, de phases, de non-linéarité, d'autorégulation, de prédictibilité et d'émergence que nous avons présentées plus haut (voir section 3.3), génèrent une variabilité chez les individus qui évoluent dans un même système : une variabilité qui serait inhérente au développement de la L2 dans une lignée des théories des SCD (Ellis, 2004). Selon Moyer (2013) :

[T]here is a growing interest in individual differences as key to second language development and language learning is viewed as chaotic, messy, and far less predictable than previously assumed. All operative psychological, cognitive, and social mechanisms are in flux, their balance continually shifting in response to the environment. [...] There is plenty of variation in both process and outcome, even among those exposed to the target language at the same time. Individual differences, both intrinsic and extrinsic, account for this variation (p. 49).

Malgré cette différence individuelle et la variabilité qui y est associée, on est capable d'identifier, à un instant donné, des phases de stabilité et des attracteurs permettant ainsi de comprendre et « prédire » une trajectoire et donc le développement langagier des apprenants de la L2. Pour sa part, Dewaele (2002) pense que la théorie des SCD est « étrange, et imprévisible, en état d'émergence, mais trace un chemin spécifique que les systèmes dynamiques suivront à travers le temps. C'est ce modèle qui reflète simultanément la rigidité et la stabilité de la langue » (paragr.

51). Ainsi, nous devons nous interroger sur la pertinence explicative de l'acquisition des langues secondes en tant que champ d'exploration de ces traits, car les différences individuelles et les variations doivent être au centre de la recherche sur l'acquisition du langage (Larsen-Freeman, 2009, p. 585). Kussyk (2016, p. 212) rappelle que : « *this variability should not be dismissed as data anomalies; [...] and not let individual trajectories become hidden through group averaging* ». La quête de modèles universels de développement du langage a guidé la recherche linguistique pendant des décennies, y compris dans le domaine de l'acquisition des langues secondes, qui a supposé un continuum d'acquisition caractérisé par des contraintes, pour la plupart définies en termes cognitifs. À l'heure actuelle, en revanche, on s'intéresse de plus en plus aux différences individuelles en tant que facteurs clés du développement de la langue seconde (Dörnyei & Skehan, 2003 ; Smith & Hayes-Harb, 2011 ; Moyer, 2013).

Au terme de cette section, nous comprenons mieux le développement de l'apprenant de langues sous toutes ses facettes, notamment son évolution dans le système apprentissage. Nous retenons aussi que la théorie des SCD nous permet de saisir les notions abordées concernant les interlangues, les différences individuelles et la variabilité en tant que phénomènes naturels et de porter un regard distinct sur leur émergence, fonction et utilité. Toutefois, on peut se demander si l'utilisation de la théorie des SCD en SHS a des limites et si oui, quelles sont-elles. C'est ce qui sera discuté dans la section suivante.

3.5 Limites de la théorie des SCD en didactique des langues

Les bénéfices de l'utilisation de la théorie des SCD en linguistique sont nombreux, nous l'avons vu, car ils permettent aux chercheurs du domaine de porter un regard nouveau sur l'apprentissage et sur son fonctionnement. En revanche, ces pratiques ne sont pas sans conséquences et les limites sont visibles aussi bien sur les plans épistémologiques que méthodologiques. Cette partie est consacrée à quelques limites liées à l'utilisation de la théorie des SCD aux recherches en linguistique en général et en didactique des langues en particulier. Nous soulèverons d'abord deux problèmes majeurs qui sont l'importation de théories des sciences exactes aux SHS dans un sens large et celui de la méthodologie de recherche en didactique des langues. Enfin, nous présenterons quelques pistes de solutions proposées par des auteurs et qui se prêtent également au contexte de l'AIAL.

3.5.1 Problèmes d'importation de théories d'une science à une autre

Dans les sciences exactes, les théories sont le plus souvent utilisées pour justifier les protocoles expérimentaux. Dans cette perspective, la méthodologie se doit de suivre des règles (im)posées. Une sorte de formule mathématique dans laquelle les divers constituants sont remplacés par leurs valeurs, dans le but d'obtenir des résultats. Par exemple, en mathématiques, les équations différentielles sont utilisées dans les études de systèmes complexes et dynamiques, afin de déterminer les trajectoires de particules (la météorologie par exemple). Cependant, contrairement à certaines théories issues des sciences de la nature et formelles, la théorie des SCD est plutôt utilisée comme une métaphore au sein des SHS (notamment en linguistique). C'est particulièrement le cas de la théorie des SCD dans les recherches en développement des L2. Dans le même ordre d'idées, Dewaele (2002) propose des mises en garde quant à une place accordée à la théorie des SCD en linguistique. La première évoque l'insuffisance de discernement du sujet :

[...] L'attrait de la TCC⁷⁰ en linguistique appliquée doit cependant être nuancé. Sokal et Bricmont (1999) ont démontré que la tentation que ressentent les scientifiques d'aller puiser des idées dans d'autres disciplines est risquée et qu'il en résulte souvent des erreurs dues à un manque de compréhension profonde du sujet. L'usage de la perspective de la TCC en linguistique appliquée est donc périlleux. L'obstacle majeur est le fait que la TCC décrit l'émergence de l'ensemble d'un système qui n'est ni reflété dans une des parties ni dirigé par une instance particulière (Dewaele, 2002, paragr. 52).

La deuxième mise en garde concerne le danger du transfert de théories issues des sciences dures en linguistique comme le précise l'auteur :

Un deuxième obstacle concerne le danger inhérent à l'importation de la théorie d'une discipline à une autre. L'exercice est risqué, surtout s'il s'agit d'un passage des sciences exactes aux sciences sociales. Les systèmes décrits par les théoriciens de la TCC n'émanent pas d'une volonté et d'une liberté de choix humains. Les intentions communicatives des êtres humains peuvent difficilement se réduire en équations (*ibid.*, paragr. 53).

Enfin, l'auteur aborde un point très important sur lequel nous nous attarderons, car il concerne les démarches méthodologiques à adopter lorsque l'on souhaite appliquer la théorie des SCD en linguistique.

Finalement, l'usage de la TCC en linguistique appliquée doit se soumettre à la même rigueur méthodologique que toute autre matière. Il faut donc des recherches en laboratoires, des constructions de modèles et de la recherche qualitative. Herdina et Jessner [...] ont présenté la TCC aux linguistes comme une métaphore séduisante et utile,

70 Théorie du Chaos et de la Complexité (TCC). L'auteur utilise ce terme à la place de la théorie des SCD.

tout en se gardant de présenter des données quantitatives. Il est cependant nécessaire d'appliquer la TCC à des données concrètes pour échapper au reproche de construire des modèles infalsifiables (*ibid.*, paragr. 53).

Dans ce dernier passage, nous pouvons relever deux points de discussion. D'une part, il pose des recommandations méthodologiques tout à fait intéressantes, en suggérant « des recherches en laboratoires, des constructions de modèles et de la recherche qualitative ». Toutefois, nous nous demandons si les recherches en laboratoires et les constructions de modèles dépeignent la vraie nature des interactions entre les différents facteurs chez un apprenant de langue. En effet, à ces questions, nous pensons qu'il y a un risque de contrôler ou d'isoler des facteurs pourtant cruciaux au fonctionnement d'un système. Rappelons que dans un SCD, tous les éléments, aussi minimes soient-ils, ont leur place et contribuent à l'émergence de structures linguistiques. D'autre part, le fait que l'auteur parle de rigueur méthodologique et de données concrètes nous interpelle, car comme nous venons de le préciser, certains aspects du comportement humain ne sont tout simplement pas mesurables. Et donc cette conception de rigueur nous semble utopique. On pourrait plutôt parler d'adaptation de la théorie des SCD ou même garder l'idée de métaphore. Cependant, l'utilisation de la théorie des SCD en didactique représente plutôt des images de fonctionnements qui sont assignées à des concepts et qui permettent d'interpréter des résultats au travers d'analyses quantitatives ou qualitatives des phénomènes étudiés. Aucun calcul ne peut ou n'est effectué dans le sens d'une application de formules, car il est très difficile voire quasi-impossible de mesurer certains aspects avec précision⁷¹ au niveau de l'individu. Par exemple, il existe des outils de mesure de la motivation, de l'anxiété, de l'affect, etc. mais la plupart repose sur des questionnaires psychométriques le plus souvent avec des échelles de Likert. Les plus récents utilisent des mesures quasi-expérimentales comme les battements de cœur, les pulsations et la transpirations (voir Dewaele, 2003; Dörnyei & Chan, 2013; Gardner & MacIntyre, 1993) . Les techniques en électroencéphalographie, en imagerie par résonance magnétique fonctionnelle ou même en oculométrie pourraient être d'une grande utilité mais rares sont les études en didactiques des langues qui se sont penchées sur le sujet à notre connaissance (*cf.* Puma, 2016). Pour nous, la difficulté liée aux mesures de ces éléments constitue également une limite de l'adaptation de la théorie des SCD en didactique. Puisque nous ne pouvons pas déterminer avec précision les trajectoires, les attracteurs, les éléments stables ou mobiles, etc. comme dans les

⁷¹ Par des formules mathématiques et non pas seulement par des questionnaires.

autres disciplines en sciences dures, nous ne pouvons qu'émettre des hypothèses probables sur des faits visibles quant aux différents facteurs constitutifs. Néanmoins, ces limites ne signifient pas une impossibilité d'utiliser la théorie des SCD dans les champs de la recherche en didactique des langues. Elles représentent plutôt un potentiel assurément énorme d'études et de défis que les chercheurs en développement des L2 se doivent de relever.

3.5.2 Problèmes de méthodologies de recherche

Souvent, bon nombre de recherches suivent des méthodologies plus ou moins classiques. Parmi les plus répandues, nous en retenons deux qui sont fréquemment utilisées dans les méthodes expérimentales et quasi-expérimentales⁷², et également dans les SHS. La première est le « groupe témoin - groupe expérimental ». Elle consiste à sélectionner à un instant t , une cohorte homogène⁷³ d'apprenants divisée en deux groupes. Un des groupes (groupe test) est généralement soumis à une expérimentation, ou a pratiqué une activité langagière plus que l'autre. Le groupe contrôle (ou témoin) n'est pas soumis à l'élément dont on teste les effets. Les résultats obtenus en mesurant le niveau de performance des deux groupes dans un domaine défini permettraient de dire si l'expérimentation a des conséquences sur le développement langagier des apprenants. L'importance d'utiliser un groupe contrôle est de pouvoir comparer les effets de l'expérimentation ou de l'activité s'il y en a. La deuxième est le modèle « pré-test, post-test ». Il consiste à mesurer des phénomènes chez un groupe d'apprenants avant d'administrer un traitement, suivi d'un post-test sur la même mesure après le traitement. Les résultats des deux moments sont alors comparés en vue d'en tirer une conclusion. Même si nous les présentons séparément, les modèles pré-test / post-test peuvent être utilisés avec ou sans groupe témoin. Ces deux modèles sont souvent utilisés simultanément. Dans ce type de configuration, le chercheur décide d'appliquer un traitement⁷⁴ sur des critères bien définis et en s'appuyant sur des hypothèses, en vue d'en mesurer les conséquences. Il évalue dans un premier temps (t_1) le niveau de performances spécifiques de deux groupes. Dans un second temps (t_2), il applique le traitement uniquement chez un groupe (groupe test). Pendant ce temps, le groupe témoin ne reçoit aucun traitement. Enfin dans un troisième

72 Elles proviennent des sciences dures et sont utilisées pour tester l'efficacité d'un médicament par exemple.

73 L'homogénéité sert à contrôler des facteurs tels que l'âge, le sexe, la provenance, la quantité d'activités des apprenants etc. Mais il est toutefois à noter que ceci est difficile à obtenir en sciences humaines et sociales.

74 Par *traitement* nous entendons un dispositif, un outil, une méthode, etc.

temps (t_3), il réévalue le niveau de performance des deux groupes, puis analyse et compare les résultats de t_1 et t_3 des deux groupes et tire des conclusions par rapport à ses hypothèses de départ. Ainsi, il est relativement plus aisé de tirer des conclusions dans ce genre d'expérimentation car il s'agit de vérifier l'impact direct d'un seul facteur : soit l'expérimentation a des résultats positifs, soit ils sont négatifs (Dörnyei, 2007). Les hypothèses formulées sont donc validées ou infirmées. Les autres variables pouvant influencer les résultats sont ainsi négligées. Les méthodologies classiques peuvent être efficaces dans de nombreux domaines scientifiques notamment en sciences dures, mais beaucoup moins en didactique des langues, lorsque nous reconnaissons que les apprenants sont des individus complexes à part entière, évoluant, changeant et sous l'influence de nombreux facteurs interconnectés (Bertin, Gravé, & Narcy-Combes, 2010, p. 2). Ces méthodologies classiques posent donc des problèmes de validité interne dans les RAL : l'isolation et le contrôle des variables ne sont donc pas souhaités comme le soulignent Narcy-Combes & Miras (2012) : « [...] *la linéarité, et la structuration ont cédé la place à des conceptions qui mettent en avant les ruptures et le besoin d'ajustements permanents qui dépendent des contextes et des individus et varient avec le temps* (p. 25). [...] *les objets de recherche ne sont jamais des items isolés et/ou isolables* » (p. 28). Quelques propositions de remédiations méthodologiques sont présentées dans la sous-section suivante.

3.5.3 Pistes de solutions méthodologiques dans le cadre de la théorie des SCD

Outre les méthodologies classiques dont nous venons de discuter *supra*, d'autres, issues des SHS sont utilisées dans le domaine de l'AIAL et comprennent notamment des questionnaires, des interviews/entretiens et des observations. Si les outils statistiques constituent des moyens fiables pour montrer des corrélations entre diverses variables, ils ne permettent pas d'instaurer des liens de causalité entre ces variables, ni même de dépendre avec précision les interactions entre ces facteurs chez un apprenant de langue qui s'engage dans les activités de l'AIAL. Ainsi, Sockett (2014) préconise des méthodes mixtes alliant statistiques et analyses qualitatives, mais aussi des études longitudinales (p. 23). Même avec l'adoption de ces solutions, le chercheur en RAL se voit toujours confronté à des problèmes d'ordres méthodologiques : d'une part, comment peut-il étudier un seul élément si tous les autres y sont interconnectés ? Et d'autre part, comment arrive-t-il à déduire un lien de causalité entre ces différents facteurs ? Des chercheurs comme de Bot & Larsen-Freeman

(2011) reconnaissent la difficulté de ces questions, et écartent les méthodes traditionnelles. Ils optent pour la prise en compte de tous les facteurs qui constituent le système et proposent quelques pistes de solutions.

Tout d'abord, les études qui se prédestinent à appliquer la théorie des SCD peuvent procéder par ciblage d'événements précis. Cette solution consiste en l'identification des tendances ou des *patterns* en vue de comprendre le système, mais les résultats ne pourraient pas être universellement généralisés, chaque système fonctionnant différemment. Dans cette logique, aucune variable n'est isolée ni contrôlée, tout l'écosystème reste intègre dans son état le plus naturel. Ce faisant, le chercheur se doit de rester ouvert aux diverses explications et explorer des points focaux du système, puisqu'il lui est impossible d'étudier tout le système (de Bot & Larsen-Freeman, 2011, p. 18-19). L'identification des attracteurs, qui signalent la trajectoire des autres éléments, constitue un point de départ sûr pour conduire des recherches en AIAL.

Ensuite, le chercheur peut s'engager dans une méthode rétrodictive d'explication de comportements langagiers (Dörnyei, 2014). Cette méthode « historique » appelée rétrodition (Grignon, 2008) et qui consiste à remonter dans le temps depuis un événement, est une sorte d'étude longitudinale mais dans le sens inverse. En effet, dans une étude longitudinale, le chercheur part d'un instant précis et suit l'évolution de l'expérimentation sur une période relativement longue, afin de voir émerger un résultat. En revanche, dans une logique rétrodictive, le chercheur identifie un fait à un instant identifié puis remonte le temps sur une période relativement longue, en vue de l'identification des causes possibles. Ce dernier procédé consiste à retracer et à décrire les trajectoires du système, avec pour objectif d'identifier les éléments des phases précédentes qui auraient engendré ces trajectoires. Le chercheur part du moment où un phénomène émerge et regarde quels éléments du système ont interagi entre eux et quelles sont les trajectoires suivies, pendant une période de temps bien déterminée. Par exemple, pour un apprenant chez qui émergent les caractéristiques de l'accent américain, le chercheur va rassembler toutes les activités susceptibles d'être sources d'*input* (conversations, films, séries, musique, jeux, lecture, etc.), retracer comment ces activités sont pratiquées dans l'espace-temps (fréquence, quantité, supports, modes, etc.) et déterminer les périodes de départ de ces activités (depuis cinq, dix, quinze ans, etc.). Il faut toutefois noter que cette méthode ne permet pas au chercheur de prédire exactement le développement langagier de l'apprenant, mais plutôt de le comprendre et

de le décrire, comme le précise Kusyk (2016, p. 211) : « [R]esearchers may be able to describe in great detail the developmental trajectory of a language learner without however being able to predict with certitude the exact nature of his or her subsequent system phases. »

Enfin, un dernier travail méthodologique dans une perspective de SCD consiste en l'étude des causalités mutuelles ou la coadaptation que nous avons abordé dans la section Émergence et coadaptation (page 82). Selon Kusyk (2016, p. 212), la coadaptation est : « [...] *the reciprocal causality that occurs when one factor changes and influences another* ». Comme plusieurs éléments peuvent changer simultanément dans un SCD, il est important non seulement d'analyser ces changements mais également d'identifier ceux qui sont dus à d'autres facteurs du système. Ici, le chercheur distingue les changements qui proviennent d'un seul élément de ceux qui proviennent de plusieurs causes, et les interactions entre ces causes. L'intérêt de ces démarches est notamment de : « [...] *study patterns that emerge from interactions rather than focusing solely on single variables, and to investigate real people in their human contexts and interaction, rather than aggregating and averaging across individuals as happens in experimental and quantitative studies* » (de Bot & Larsen-Freeman, 2011).

Dans cette section, nous avons abordé certaines des limites relatives à la théorie des SCD à la linguistique appliquée en général, et à la didactique des langues en particulier. Nous avons soulevé deux problèmes majeurs que sont l'importation de théories des sciences exactes aux SHS et la méthodologie des recherches. Face à ces difficultés, certains chercheurs en L2 préconisent quelques pistes de remédiation méthodologiques dans le cadre d'études se référant à la théorie des SCD, car, contrairement aux études qui ont tendance à isoler et contrôler des variables pour mieux simplifier et modéliser des phénomènes, la logique afférente aux théories des SCD est de prendre en compte toutes les variables présentes dans leur environnement écologique.

3.6 Conclusion

Dans ce chapitre, nous avons abordé une théorie fondamentale pour notre travail : la théorie des SCD. Considérant la nature de l'AIAL, le cadre épistémologique des SCD (Sockett, 2014 ; Kusyk, 2017 et Toffoli, 2018) a souvent été évoqué. Le choix de cette théorie est déterminant, car elle nous permet de mieux appréhender le développement de l'apprenant en tant qu'entité à part entière, évoluant dans un système d'apprentissage qui est complexe et dynamique. En effet, nous avons

considéré l'apprenant comme baignant dans une multitude de facteurs internes et externes, dans une relation d'interconnectivité et dont l'émergence de ses structures linguistiques (grammaire, vocabulaire et phonologie) est difficile à prédire. Après une présentation de la théorie, nous avons lié ses caractéristiques inhérentes au domaine de l'apprentissage des langues. Ainsi, nous avons présenté la théorie des SCD en didactique comme une métaphore permettant de décrire et d'expliquer le cheminement de l'apprenant de langues, dans une logique de non-linéarité, de dynamisme, d'attraction, d'imprédictibilité, d'autorégulation et enfin, d'émergence. Selon ce courant épistémologique, l'émergence d'une L2 serait le résultat d'un parcours complexe, où plusieurs facteurs (*input*, motivation, autonomie, attention, etc.) entrent en jeu et sont en perpétuelle interaction (Dörnyei, MacIntyre & Henry, 2015 ; Toffoli & Perrot, 2017). Les performances réelles de chaque individu apprenant une L2 résulteraient de l'évolution permanente de ces facteurs. Le développement langagier se manifeste sur une base de va-et-vient, où le cheminement de l'apprenant n'est jamais linéaire. Aussi, dans un tel système, de petits changements des conditions initiales peuvent mener à des résultats radicalement différents. Un comportement chaotique peut donc être extrêmement difficile à prévoir. Pour une quantité d'entrée⁷⁵ donnée, ces systèmes peuvent produire des changements de sortie nettement supérieurs ou inférieurs à la somme des entrées, voire des résultats nuls, en fonction de l'état actuel du système ou des valeurs de ses paramètres. Nous avons aussi vu que parce qu'un apprenant de langue est unique et évolue dans un environnement spécifique, les différences individuelles sont inévitables et engendrent des variabilités inter et intra caractéristiques des SCD. En se fondant sur ces notions, il apparaît clairement que la notion d'interlangues en est aussi une manifestation évidente et les regards portés à son encontre deviennent plus souples et plus réalistes. Enfin, nous avons fait état de quelques limites de l'adaptation de la théorie des SCD en linguistique appliquée, en évoquant les problèmes liés à l'importation en SHS des théories issues des sciences dures, notamment des défis méthodologiques, auxquels nous avons proposé des pistes de solutions. Il est alors question pour le chercheur de conduire soit des études longitudinales rétrodictives, soit d'exploiter les attracteurs, soit de repérer des tendances, soit de procéder à l'identification des phénomènes de coadaptation. Ces méthodes peuvent bien sûr être combinées pour plus de robustesse.

⁷⁵ *Input*.

Dans ce travail, la théorie des SCD nous aide à comprendre des productions orales des participants au regard des diverses activités qu'ils pratiquent de manière formelle et surtout informelle. Nous retenons la métaphore de Toffoli (2018) qui envisage « l'appropriation langagière non pas comme le résultat de quelques variables isolées, mais plutôt comme un nuage de vapeur en mouvement perpétuel, changeant aussi bien en taille qu'en consistance, et qui, parfois, produit des gouttelettes » (p. 160). Pour ce faire, nous allons utiliser un nombre de concepts et combiner plusieurs sources de données en vue de récolter un ensemble d'informations susceptibles de nous faciliter la compréhension des liens entre les nuages de vapeurs et les gouttelettes (*ibid.*). L'adoption d'une méthodologie mixte nous permettra d'examiner la multitude de facteurs qui entourent les apprenants et de pouvoir dégager des tendances fortes et caractéristiques de leur développement L2. Dans cette visée systémique, l'identification des trajectoires éminemment non-linéaires, des états initiaux et des attracteurs, nous semble d'une importance essentielle. Ainsi, nous prendrons en compte non seulement leur différence individuelle liée aux variations mais également leur interlangue qui découle du processus d'apprentissage.

Dans le chapitre suivant, nous présenterons une autre théorie fondamentale pour notre travail, car elle nous permet de comprendre l'émergence de caractéristiques de prononciation chez nos participants. Il s'agit de la théorie de l'imitation vocale qui explique de manière détaillée l'acquisition des traits phonétiques et phonologiques chez les humains. Parallèlement, nous en profiterons pour présenter quatre modèles d'acquisition phonétique et phonologiques en L2 : le modèle de l'*Acoustic Image Imprinting* (AII) de Neufeld, le modèle du *Speech Learning Model* (SLM) de Flege, le modèle du *Perceptual Assimilation Model* (PAM) de Best et enfin le modèle *Native Language Magnet* (NLM) de Kuhl.

4 Théorie de l'imitation vocale

Dès leur découverte, on s'est demandé si les neurones miroirs pouvaient être à la base de l'imitation. [...] Le développement progressif du système des neurones miroirs a constitué une composante clé dans l'apparition et l'évolution de la capacité humaine de communiquer, d'abord par les gestes, puis par des mots. (Rizzolatti & Sinigaglia, 2011, p. 151).

Cette deuxième partie de notre cadre théorique sera consacrée au phénomène de l'imitation, que nous allons tenter de traiter de la façon la plus complète qui soit. D'abord, nous définirons ce terme, même s'il admet des connotations différentes, voire contradictoires parfois et nous montrerons les domaines dans lesquels on le retrouve. Ensuite, le rôle des neurones miroirs rendant possible le processus d'imitation sera aussi discuté d'un point de vue neurophysiologique, au travers notamment des travaux de Rizzolatti et ses collaborateurs. Nous verrons que dans l'acception qui sera la nôtre, l'imitation peut être le résultat d'un acte conscient ou inconscient, et peut être assujettie aux notions de fréquence d'exposition et d'usage (exemplaires). Nous distinguerons par la suite l'imitation dans le comportement général et l'imitation vocale, tout en abordant la différence entre l'imitation phonétique et phonologique. Dans la même lignée, nous mettrons en évidence un point central de cette partie : l'imitation vocale médiée. Pour finir, nous présenterons quatre modèles d'acquisition phonétique et phonologique liés à l'imitation vocale et qui tentent d'expliquer comment les mécanismes d'appropriation de la prononciation sont mis en œuvre tant en L1 qu'en L2. Il s'agit de : l'*Acoustic Image Imprinting* (AII) de Neufeld, le *Speech Learning Model* (SLM) de Flege, le *Perceptual Assimilation Model* (PAM) de Best et le *Native Language Magnet* (NLM) de Kuhl. Tout au long de ce travail et spécifiquement dans ce chapitre, le lecteur remarquera que nous faisons un travail de distinction et de précision terminologique, pour des raisons de clarté, car les termes que nous employons sont non seulement polysémiques, mais ils sont aussi utilisés différemment par différents chercheurs dans la littérature qui traitent de la question de l'imitation. Nous présenterons les diverses définitions existantes, suivies de celle(s) que nous retenons et les cadres dans lesquels nous les employons. Dans la section qui suit, nous nous consacrons à une présentation générale du phénomène d'imitation.

4.1 Imitation : cas général

Étymologiquement dérivé du latin⁷⁶ « *imitatio* »/« *imitari* » et signifiant « copie »/« copier », l'imitation est une action qui consiste à reproduire « quelque chose » et le résultat qui en découle. Le e-Larousse⁷⁷ fournit des définitions et même si la première ne définit pas vraiment le mot imitation (« Action d'imiter une personne, un son, un bruit, etc. ; résultat de cette action »), celles qui concernent les champs de la psychologie et des spectacles nous paraissent pertinentes et se rapprochent des notions que nous aborderons dans ce chapitre. En effet, le mot désigne : la « reproduction d'une mimique, d'un geste, d'une posture, d'une situation vécue » en psychologie et l'« action de reproduire la gestuelle, la diction, les tics d'un acteur, d'un chanteur, d'une personnalité, d'un homme politique, etc. ; sketch qui s'y rapporte » en spectacle. L'imitation⁷⁸ représente un phénomène central dans la vie des êtres vivants. Cet acte « primitif » continue de conditionner les manières de faire. Les humains et les animaux ont toujours copié leurs pairs. Le concept d'imitation est connu depuis fort longtemps et est considéré comme un levier de l'apprentissage en général et un facteur de développement personnel. D'ailleurs, c'est vers les années 1960 que ce concept a connu son essor (notamment en France) grâce aux travaux de recherches et d'expérimentations dans la sphère de l'apprentissage social, particulièrement en psychologie cognitive (Raynal & Rieunier, 1998). Ainsi, d'après le psychologue Albert Bandura (1995), l'apprentissage social⁷⁹ est important pour chaque individu et s'effectue essentiellement par imitation. Le terme « vicariant » est ainsi employé pour désigner les mécanismes qui résultent de l'imitation par observation et qui font partie intégrante de cet apprentissage social. Bandura distingue l'apprentissage par imitation, c'est-à-dire la reproduction « fidèle » et « intégrale » du comportement d'une personne, du *modelage* qui, au contraire, implique une réadaptation active (cognitive) des faits observés de la part de l'imitateur. Pour Vygotsky (1978, 1986), l'apprentissage du langage, fruit d'interactions sociales, est dans une certaine mesure fondé sur l'imitation, alors que Bruner (1983, 1987) pense que cet apprentissage par observation est une conséquence naturelle

76 Entrée imitation dans le dictionnaire étymologique en ligne. <https://www.etymonline.com/search?q=imitation> (consulté le 02 juin 2019).

77 <https://www.larousse.fr/dictionnaires/francais/imitation/41662#definition> (consulté le 02 juin 2019).

78 Les termes « convergence », « alignement » ou « accommodation » sont également utilisés comme synonymes du mot imitation. Nous les emploierons donc de manière interchangeable tout au long de ce travail.

79 Apprentissage fondé sur l'observation des comportements de personnes qui nous entourent, de la société.

de l'interaction entre les individus et l'environnement culturel dans lequel ils vivent. Quant à Piaget (1968), il pense que la genèse de la connaissance chez l'individu est le résultat d'une construction à partir de procédures cognitives. Ainsi, Piaget stipule que l'imitation, bien que socle de tout apprentissage, ne constitue pas une condition *sine qua non* de cet apprentissage ; mais qu'elle est plutôt modelée par une adaptation cognitive. Le développement intellectuel de l'enfant, ainsi que la fonction sémiotique se fonderait primordialement sur l'observation des adultes qui l'entourent. L'enfant emmagasinerait, traiterait et réorganiserait ces informations en vue d'une utilisation précise en cas de besoin, dans le seul but de communiquer et d'interagir avec son environnement (Vygotsky, 1978). Bruner (1983) confirme ce postulat : « Il ne s'agit pas bien sûr pour l'enfant d'avoir les mêmes représentations que l'adulte, mais plutôt d'utiliser le même système de signes que lui, à la fois pour communiquer et pour construire sa représentation » (p. 290). Dans la même perspective, Winnykamen (1990, p. 105) considère que « L'activité imitative consiste en l'utilisation intentionnelle des actions observées d'autrui, en tant que source d'information en vue d'atteindre son propre but ». Cette conception de l'imitation-modélisation, impliquant le jeu et l'étayage⁸⁰, se retrouve dans la plupart des travaux sur le développement de l'enfant. De ce fait, il est donc important de préciser l'aspect cognitif, voire métacognitif de l'imitation, parce qu'elle est souvent associée au behaviourisme de Skinner (stimulus-réponse), qui laisse présumer l'absence de toute activité réflexive (Raynal & Rieunier, 1998, p. 285). Bien que l'existence d'une imitation néonatale soit contestée par certains chercheurs, d'autres l'ont mise en évidence (Williamson et coll., 2010, 2011). Dans une étude pionnière, Meltzoff & Moore (1977), ont observé des phénomènes d'imitation de comportement chez des bébés de seulement quelques jours ou semaines. Ces nouveau-nés ont reproduit des expressions faciales (ouverture de la bouche, protrusion des lèvres et de la langue, etc.) présentées par un adulte (Figure 6 ci-dessous). Meltzoff et collègues (2005, 2010) ont également postulé que les enfants imitent dans le but de communiquer et ils comprennent l'intention du modèle, tout comme les parents imitent les enfants avec les mêmes objectifs. Parmi les différentes formes d'apprentissage, l'apprentissage par habitude (les habitudes), l'apprentissage par association (l'observation et l'imitation de modèles) et l'apprentissage par instruction et médiation (le cas des apprentissages scolaires chez

80 De l'anglais *scaffolding* et développé par Bruner (1983), le concept de l'étayage consiste en l'aide, l'accompagnement et le suivi qu'un adulte (parent, enseignant) apporte à un enfant, afin qu'il mène à bien des tâches difficiles auxquelles il/elle se confronte.

les jeunes enfants), ont été mis en évidence (Raynal & Rieunier, 1998, p. 285). Enfin, les caractéristiques de l'imitation peuvent être nombreuses et observables à différents niveaux (Pickering & Garrod, 2004), comme indiqué plus haut :

- Comportemental ou gestuel (le regard, le rire, la posture, les expressions faciales) ;
- Lexical (la syntaxe, la grammaire, le vocabulaire) ;
- Phonétique/phonologique (voyelles et consonnes, opposition entre phonèmes).

Seul le troisième point nous intéresse et fera l'objet de notre travail. Dans les paragraphes qui suivent, les types, les caractéristiques et les conséquences de l'imitation seront discutés à la lumière de la littérature existante sur l'acquisition phonético-phonologique, appuyés par quelques exemples d'études. D'abord, dans la toute prochaine section, nous présentons les mécanismes neurophysiologiques qui sous-tendent le phénomène de l'accommodation en général, à travers le rôle particulier des neurones miroirs.

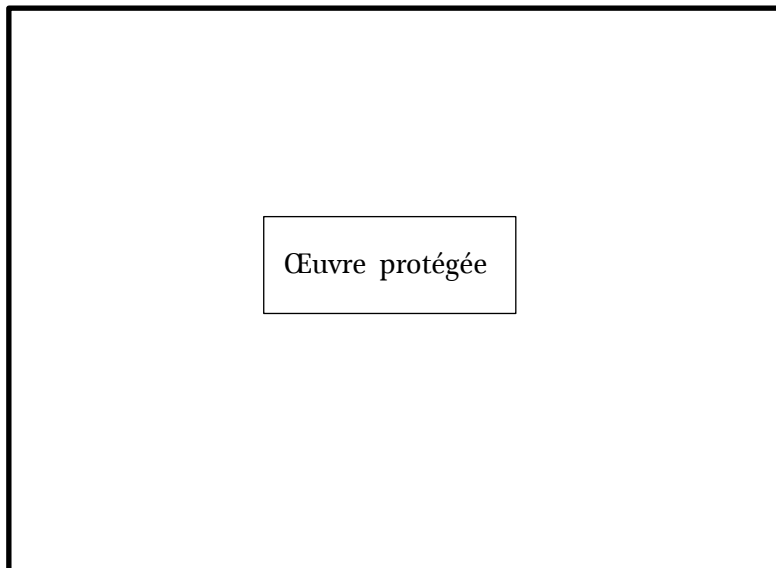


Figure 6 : un bébé de moins d'un mois imitant les expressions faciales d'un adulte. (Meltzoff *et al.*, 2009, p. 206.)

4.2 Imitation du point de vue neurophysiologique : le rôle des neurones miroirs

Dans cette section, nous explicitons le fonctionnement neuronal qui sous-tend le phénomène de l'assimilation en général. Physiologiquement, les neurones miroirs, situés dans la zone de Broca du cerveau, seraient responsables des activités d'imitation et de certaines expressions émotionnelles (Rizzolatti *et al.*, 1996 ; Rizzolatti, 2005). Découverts pour la première fois en 1990 chez le singe par le neurophysiologiste italien Giacomo Rizzolatti et ses collègues de l'Université de

Parme en Italie, les neurones miroirs seraient impliqués dans les fonctions motrices et dans l'apprentissage en général (gestes et langage). En principe, chaque neurone ne répond qu'à un seul type d'action et non à un autre geste. Par exemple : un neurone activé par un mouvement de flexion des doigts ne réagira pas à l'extension du même organe. Cette fonction est dite « sélective ». Les neurones miroirs sont aussi sélectifs et sont baptisés ainsi, car ils s'activent simultanément pendant les processus de perception et de l'exécution motrice d'une action, tout comme lorsque nous nous regardons dans un miroir. Cette découverte est survenue (fortuitement) lorsque ces chercheurs étudiaient le comportement d'un macaque. Ils ont alors remarqué que des cellules situées dans le cortex pré-moteur frontal (Aire F5) répondaient différemment aux autres cellules (dites canoniques). Ces neurones s'activaient uniquement lorsque le singe effectuait des gestes moteurs vers un but défini et lorsqu'il observait les autres singes effectuer le même acte (Rizzolatti *et al.*, 1996). En revanche, lorsque le singe exécutait un geste simple et sans but précis, les neurones miroirs ne s'activaient pas, mais ce sont plutôt les neurones canoniques qui étaient impliqués dans l'activation (Rizzolatti, 2005).

Chez l'homme, la preuve de l'existence d'un système miroir est apportée par des techniques d'Électro-Encéphalographie (EEG) et d'Imagerie par Résonance Magnétique fonctionnelle (IRMf) et implique plusieurs zones du cerveau : « fronto-pariétal » (cortex pré-moteur_Aire 44 de Broca et lobe pariétal inférieur_Aire 40 de Brodman) et « limbique » (insula et gyrus frontal inférieur). Le système miroir est plus développé et plus complexe chez l'homme et se traduit par : l'acte moteur précis (moteur), l'observation (visuel), l'écoute (auditif), mais aussi la pensée (quand nous pensons qu'une personne effectue la même action). Ces observations et découvertes ont conduit les chercheurs à associer la responsabilité des neurones miroirs au phénomène de l'imitation. Sur le plan de la perception – production de la parole, l'imitation joue un rôle prépondérant dans l'émergence et le développement des systèmes phonétique et phonologique chez l'humain (Rizzolatti & Craighero, 2005). Ce phénomène est fondé sur les processus sensori-moteurs établis depuis la naissance et serait également à l'origine des mécanismes d'acquisition phonétique et phonologique du langage (Garnier, Lamalle & Sato., 2013). Le développement langagier par imitation phonétique implique non seulement des processus perceptivo-moteurs mais également des compétences cognitives et sociales :

Imitation is often thought of as a cognitively undemanding (primitive cognitive function), rather elementary form of behaviour. Recent works, across a variety of sciences, showed, however, that it is not true (p.7). [...] It is a cognitive faculty which evolved later from the mirror system following the acquisition of new matching properties by mirror neurons. (Rizzolatti & Buccino, 2005, p. 9)

De manière métaphorique, ces neurones miroirs fonctionnent comme une caméra qui filme et enregistre des événements, en vue d'une utilisation ultérieure. Les éléments de la vidéo ainsi enregistrés sont classés en unités d'images et de sons fonctionnelles, selon des critères bien spécifiques (ressemblance/dissimilitude), et subissent des tris, des réorganisations, voire des computations au niveau du langage humain. Pour Nguyen et Delvaux (2015, p.49) :

The discovery of the mirror system in area F5 of the premotor cortex in the macaque's brain, and the role this system has been established to play in both the execution and recognition of grasping actions, has given rise to the idea that imitation, taken as a general sensory-motor ability, is one of the factors that, in the course of evolution, made our brains "ready for language".

Aussi, de nombreuses études ont utilisé des techniques d'IRM dans la découverte des zones du cerveau qui s'activent pendant l'imitation (Reiterer *et al.*, 2013). Il s'avère que ces zones sont variables, ce qui conduit à des résultats contradictoires dans la littérature portant sur la question. Par conséquent, les interprétations sont nombreuses et les facteurs mis en avant sont variables : l'âge, la L1, la provenance, le « talent » et la différence individuelle des sujets (Reiterer *et al.*, 2013 ; Dörnyei & Skehan, 2003). Tout dernièrement, Carey & McGettigan (2017) ont utilisé (et recommandent) l'IRM au niveau du cerveau (activités neuronales) et au niveau articulatoire-acoustique (conduit vocal), en vue de comprendre les processus en jeu pendant la convergence vocale, et les interactions entre ces lieux. Ainsi, Fadiga et collaborateurs (2002) ont montré que l'écoute de sons/suites de sons spécifiques provoque une réaction motrice visible au niveau de l'excitation des muscles de la langue. Cette excitabilité des muscles articulatoires n'est possible que pour des sons familiers ; aucune activation n'est enregistrée lorsque les sujets écoutent des sons inconnus ou qui ne relèvent pas des sons humains (non linguistiques). Une autre étude en IRMf montre qu'il existe une activation des zones de contrôle moteur dévolues au langage, lors de l'écoute passive de phonèmes d'une langue connue, ou même lors de l'observation de mouvements des lèvres liés à la production de sons du langage. Selon Watkins *et al.* (2003) : « [...] *Speech perception, either by listening to speech or by visual observation of speech-related lip movements, enhanced excitability of the motor units underlying speech production, particularly those in the left*

hemisphere » (Watkins *et al.*, 2003, p. 992). Aussi, les neurones miroirs seraient également responsables des gestes anticipatoires. Dans la production de la parole par exemple, cette observation explique pourquoi nous arrivons à compléter le discours d'autrui par des mots ou expressions manquants chez ce dernier. Ou encore comment nous réussissons à prévoir que quelqu'un ait besoin d'un stylo et à le lui tendre, sans que cette personne ne formule explicitement sa demande.

Des recherches impliquant des animaux, comme par exemple le perroquet, ont montré le rôle et l'importance des neurones miroirs et d'autres neurones du cerveau dans le processus d'imitation (Dautenhahn & Nehaniv, 2002 ; Pepperberg & Wilkes, 2004). En conséquence, Pepperberg (2007) a travaillé sur l'imitation phonologique d'un perroquet (*Psittacus erithacus*) et affirme que : « *a parrot not only has phonological awareness but also demonstrates true imitation rather than mimicry* » (p.1). Le langage humain se serait développé à partir d'une forme précoce d'imitation faciale et vocale chez l'*Homo erectus*, avec un système miroir semblable à celui trouvé chez le singe (Nguyen et Delvaux, 2015, p.49). Ce fait confirme que le langage humain est originellement purement gestuel et est devenu linguistique au fil du temps. Ceci expliquerait également l'existence de certains sons spéciaux comme les clics en langue zoulou ; un ensemble de consonnes qui servait à appeler les animaux. En somme, le cerveau humain aurait évolué pour rendre le langage possible et c'est ce qui expliquerait l'évolution culturelle qui a pris le pas sur l'évolution biologique, pendant la transition des formes de communication de base vers les langues à part entière. S'appuyant sur les dernières recherches en neurosciences, en linguistique et en primatologie, Arbib (2012) présente dans un livre⁸¹ une version actualisée d'une théorie qui permet de comprendre l'importance évolutive des neurones miroirs, dans la reconnaissance des actions chez les humains. Ce n'est que chez l'homme que ces réseaux de neurones ont évolué pour permettre l'« imitation complexe » qui soutient la percée vers le langage. Cela explique aussi pourquoi nous pouvons apprendre la langue des signes aussi facilement que nous pouvons apprendre à parler. L'auteur examine comment les mécanismes cérébraux qui ont rendu possible l'émergence de langues à part entière sont encore actifs dans la manière dont les enfants acquièrent le langage aujourd'hui. Enfin l'auteur explique comment le cerveau a évolué pour « produire le langage » (Arbib, 2012). Pour clore cette partie, nous finissons par « une mise en

81 Intitulé How the Brain Got Language: The Mirror System Hypothesis.

garde » assez parlante et qui nous semble primordiale à saisir concernant les neurones miroirs et l'action d'imitation :

Toutefois, la capacité d'imitation n'est déterminée ni par la richesse du patrimoine moteur ni par la simple présence du système des neurones miroirs. Ce système est une condition nécessaire, mais non suffisante pour procéder à une imitation. Cela vaut non seulement pour la capacité d'apprentissage par imitation qui [...] nécessite l'intervention d'aires corticales extérieures au système des neurones miroirs, mais aussi pour la capacité de répéter des actes exécutés par un tiers [...] Autrement, la perception de n'importe quel acte moteur se traduirait immédiatement par sa reproduction. Ce qui, heureusement pour nous, n'est pas le cas (Rizzolatti & Sinigaglia, 2011, p. 160-161).

Nous venons de voir que les neurones miroirs jouent un rôle capital dans la mise en œuvre des actions quotidiennes comme le sentiment d'empathie, les gestes anticipatoires et imitatifs. Cette dernière fonction nous intéresse en particulier puisqu'elle nous permet de comprendre les processus complexes inhérents à tout acte imitatif et plus globalement, de comprendre l'apprentissage par imitation. Nous comprenons désormais qu'à chaque exposition à de l'*input*⁸² quelconque, un processus d'imitation se met en place sous diverses formes : les neurones miroirs captent et stockent pour nous, de l'information pour un emploi ultérieur. Dans la section suivante, nous introduisons la différence à faire entre la notion générale d'imitation et un type d'imitation spécifique, l'imitation vocale.

4.3 Imitation générale et imitation vocale

L'imitation vocale est au cœur de notre présente étude. À ce titre, nous souhaitons apporter ici quelques éclaircissements sur les notions d'« imitation » et d'« imitation vocale ». Le terme imitation désigne simplement le fait d'imiter quelqu'un ou quelque chose de façon globale (Fagard & Lockman, (2010). Elle peut être immédiate ou différée et couvre un panel d'aspects que nous avons énumérés précédemment (notamment comportemental, lexical ou phonétique). Cependant, le terme de convergence vocale (remarquons l'ajout du mot vocale) se réfère à une catégorie bien spécifique de l'imitation propre à la manière d'imiter le parler d'un individu (accent et prononciation), une sous-catégorie qui prend uniquement en compte le langage (Pardo *et al.*, 2012). Cela serait donc par simplification que certains auteurs emploieraient le terme « imitation »

⁸² Un input est une « entrée » de données, un ensemble d'informations auxquelles l'apprenant est exposé. Par opposition, un output est une « sortie » de données et rassemble toute production (orale et écrite) venant de l'apprenant. Les deux mots input et output existent aussi en français, par conséquent nous ne les mettons pas en italique.

pour ce qui devrait désigner « imitation vocale ». Précisons enfin que le phénomène opposé de l'imitation est appelé divergence, et peut être aussi consciente ou inconsciente.

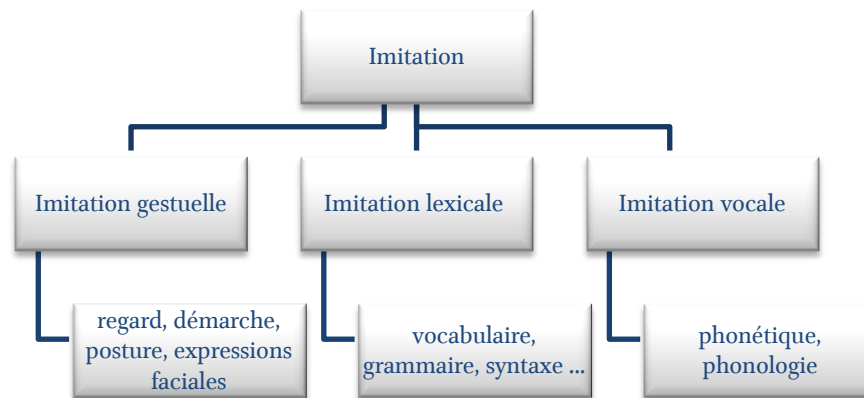


Figure 7 : les types d'imitation et leurs sphères d'influence.

4.3.1 Imitation vocale

L'idée selon laquelle la convergence vocale serait aux fondements de toute acquisition de la phonétique et de la phonologie semble dorénavant un fait indéniable. De nombreux travaux l'ont amplement montré (Piske, MacKay & Flege, 2001 ; Pardo, 2006). Parler une langue avec un accent spécifique ne peut relever que de l'ordre de l'imitation vocale (Markham, 1997). Les théories centrées sur les phénomènes d'imitation vocale stipulent que deux individus quelconques qui s'engagent dans une conversation ou interaction auraient tendance à échanger mutuellement des phonèmes (Pardo & Jay, 2010 ; Pardo *et al.*, 2012). Stuart-Smith et ses collègues (2011) définissent la convergence vocale comme : « *shifts in speech production towards the accent of an interlocutor* » (p. 1914), alors que Pardo *et al.* (2012) la qualifient de : « *the ways in which a talker adjusts speaking style to become more similar to an interacting partner* » (p. 190). Remarquons dans ces définitions que les moyens par lesquels se déroulent les interactions, pourtant centraux, ne sont pas explicités. Ces interactions, supposant la présence physique des agents, pourraient-elles aussi inclure les conversations téléphoniques ou par visioconférence ? Les moyens par lesquels l'imitation vocale serait possible seront traités plus loin (*cf.* section 4.5, p. 126). Dans le même ordre d'idées, Markham (1997) apporte, à l'aide de nombreuses études, des éléments permettant de cerner l'acquisition des caractéristiques phonético-phonologiques des langues et dialectes. Le point central des arguments

de Markham repose sur le postulat suivant : l'acquisition et l'utilisation des caractéristiques phonétiques et phonologiques passent par l'imitation. Il pose que :

Given that the acquisition of language involves an assimilation of input, resulting in cognitive structures which describe and can reproduce the input, either as a true copy or as a functionally equivalent rendering, acquisition is regarded in this thesis as being a strongly imitative phenomenon (Markham, 1997, p. 39).

L'imitation vocale faciliterait sur le plan articulatoire-acoustique, l'apprentissage du vocabulaire dès le plus bas âge. Elle permettrait d'enrichir le réseau lexical et entraînerait souvent l'acquisition phonétique et phonologique (Hodges *et al.*, 2015). Ainsi, les travaux sur la lecture répétée de Han & Chen (2010), sur les exercices des productions orales de Hamada & Koda (2010), sur l'écoute de modèles phonologiques de Webb & Chang (2012), sur la *subvocalisation*⁸³ de Zahar *et al.*, (2001) et sur le *shadowing*⁸⁴ de Commander & Guerrero (2013) confirment tous les conséquences positives de l'alignement vocal. Pour conclure cette section, nous considérons l'imitation vocale comme étant tout processus de convergence au cours duquel des éléments liés à la voix sont « copiés » par un individu. Cette copie concerne les aspects phonétiques mais aussi phonologiques. Nous présentons la différence entre ces deux notions dans la section suivante.

4.3.2 Imitation phonétique et imitation phonologique

Nous venons de mettre en évidence ce qui distingue l'alignement général et l'alignement vocal ainsi que les éléments qui caractérisent ce dernier. Nous présentons au sein de section les deux types d'alignement vocal : phonétique et phonologique. Certains chercheurs, sur la base de la différence entre la phonétique et la phonologie (Vaissière, 2015, p. 21), font la distinction entre la convergence phonétique et la convergence phonologique (D'Imperio, Cavone & Petrone, 2014). Selon eux, la première concerne essentiellement les aspects physiques des sons, c'est-à-dire les consonnes et les voyelles, sans tenir compte du sens qu'ils portent. Quant à la seconde, elle se rapporte non seulement aux sons par rapport aux fonctions qu'ils remplissent au sein d'une langue donnée (c'est-à-dire qui permettent d'opposer deux sons - ex. /l/ - loue et /r/ - roue), mais recouvre également des aspects prosodiques (durée, intensité, accentuation, intonation d'un son). Par exemple, Markham (1997) a étudié la convergence phonétique chez les enfants et chez les adultes

83 Lecture silencieuse : les participants prononcent dans leur tête ce qu'ils écoutent tout en regardant le texte.

84 Répétition immédiatement suite à l'écoute d'un énoncé.

et a trouvé que le phénomène y est plus important chez les enfants. Il en déduit que le développement du système phonétique est distinct du système phonologique. D'après son étude, le système phonologique des adultes serait déjà établi, d'où la non-nécessité d'imiter et, de plus, il s'agit de leur L1. Les résultats d'une étude similaire dans un contexte de L2 auraient été intéressants. Kail (2015) soutient le postulat de Markham en suggérant que : « *les bébés sont meilleurs que les adultes pour acquérir une L2 parce que le processus d'engagement neuronal est incomplet* » (p.42). Si l'école de Prague, représentée par Troubetzkoy (1939) insistait sur la séparation entre la phonétique et la phonologie, Malmberg (1990) propose son point de vue :

Sans l'analyse linguistique des systèmes et des unités fonctionnelles, l'expérimentateur ne saurait que faire. Et sans l'analyse physique et physiologique de tous les faits de prononciation, le linguiste ignorerait la nature concrète des oppositions établies. Les deux genres d'études sont interdépendants et se complètent (p.110).

Nous partageons l'avis de Malmberg et estimons qu'effectivement, dans la parole courante, non seulement les aspects phonétiques sont d'une grande importance mais également les facteurs phonologiques, en particulier lorsque l'on parle de prononciation, et de surcroît, de l'accent. Il est fréquent de rencontrer des ambiguïtés dans la littérature puisque certains auteurs utilisent l'imitation phonétique pour englober les deux aspects alors que d'autres emploient plutôt le terme imitation phonologique. Ceci est dû au fait que les frontières ne sont pas totalement étanches comme nous venons de le voir. Pour éviter ce débat terminologique, des auteurs utilisent tout simplement les termes « imitation vocale » (Pardo *et al.*, 2012), « imitation de parole » (Gregory, Dagan & Webster, 1997) ou « imitation verbale » (Over & Gattis, 2010 ; Hodges *et al.*, 2015). Dans notre travail, nous utiliserons les expressions imitation « vocale » ou imitation « phonéto-phonologique » pour englober les aspects phonétiques et phonologiques. Lorsque, pour des raisons de détail, ces deux phénomènes sont distincts, nous le préciserons (c'est-à-dire imitation phonétique ou imitation phonologique). Comme tout acte imitatif, la convergence vocale peut être volontaire (consciente) ou involontaire (inconsciente) et c'est ce que nous allons aborder dans la prochaine section.

4.4 Imitation vocale volontaire et involontaire

L'imitation vocale se caractérise par un nombre de facteurs agissant parfois simultanément. Il peut s'agir notamment de : l'intelligibilité (l'effort de se faire comprendre par son interlocuteur) ; la

relation (amitié, admiration, etc.) ; l'identité et le style (volonté de s'identifier à une communauté de locuteurs, exemples : le *Uptalk*⁸⁵ ou le *creaky voicing*⁸⁶, etc.) ou encore l'origine géographique. Ces facteurs sont aussi gouvernés par la notion de différence individuelle observée chez les humains (Pardo *et al.*, 2012 ; Reiterer *et al.*, 2013). Contrairement à certains chercheurs (notamment des années 1960), pour qui l'imitation est un acte conscient et intentionnel, nous tenons à rappeler qu'elle est également inconsciente (D'Imperio, Cavone & Petrone, 2014). Ainsi, pour Evans (2002, p. 96), « [...] imitation *refers to the conscious use of a variety which is not the speaker's usual vernacular* ». Notons dans cette citation que l'auteur considère l'imitation comme un fait conscient. Pour sa part, Markham (1997) utilise les termes *d'impersonation* et *d'adaptation* pour désigner respectivement la convergence inconsciente et consciente :

Adaptation occurs when the phonetic/phonological mechanism is directed to behave in accordance with the phonetic characteristics of the input. This is at times called 'accent convergence' [...] Impersonation occurs where a speaker uses a speech-style, speaker-style, dialect, accent, or language for entertainment or expressive effect (Markham, 1997, p. 50).

Nous approfondissons la convergence vocale volontaire et involontaire dans les paragraphes suivants.

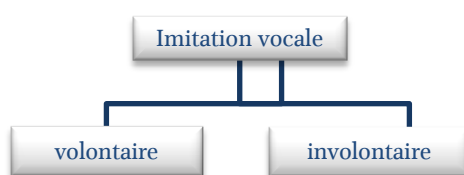


Figure 8 : les deux sous-catégories d'imitation vocale (phonétique et phonologique).

4.4.1 Imitation vocale volontaire

L'imitation volontaire est un acte délibéré et requiert des efforts de conscience (Zajac, 2015). Il convient de noter que la plupart des études sur l'imitation (phonétique et phonologique) en laboratoire s'inscrivent dans ce cadre, car il y a présence de stimuli et des consignes explicites d'imitation vocale sont présentées aux participants. L'imitateur devrait volontairement parler comme une autre personne, comme dans le cas des imitateurs professionnels (certains comédiens)

85 Terminal à haute montée : caractéristique de certains accents anglais qui consiste à prononcer une phrase déclarative comme s'il s'agissait d'une question (Voir Di Gioacchino & Jessop, 2011 ; Shokeir, 2008).

86 Voix craquée : caractéristique d'une voix causée par des vibrations apériodiques des cordes vocales (Grivicic & Nilep, 2004 ; Mendoza-Denton, 2011 ; Yuasa, 2010).

qui reproduisent les manières, y compris la voix de personnes modèles (Garnier, Lamalle & Sato, 2013). Nous pensons notamment aux humoristes et chanteurs imitateurs français les plus connus : Michaël Gregorio, Laurent Gerra, Gérald Dahan, Yves Lecoq, Thierry Garcia, Sébastien Cauet, Patrick Sébastien, Jean-Baptiste Guégan, Stéphane Benjelloun⁸⁷, etc. Nous nous attardons sur ce métier assez particulier et qui capte notre attention par les efforts particuliers, non seulement cognitifs (cela demande énormément d'attention, d'écoute et de capacités fines de perception auditive, permettant de discriminer les sons de la parole), mais aussi articulatoires, (production de la parole) réalisés par les imitateurs. Sous un angle physiologique, ces personnes, également qualifiées d'athlètes ou de contorsionnistes⁸⁸ de la voix⁸⁹, possèderaient une corde vocale plus longue que l'autre et seraient capables de modifier et de manier tout leur système de résonance (glotte, épiglotte, bouche), ainsi que leur système musculaire (larynx, pharynx, cou et visage). Le travail des imitateurs consisterait à reproduire des éléments de voix de personnes, indépendamment du timbre : ils sont capables de rentrer dans les aigus comme dans les graves. Le travail d'imitateur vocal demande beaucoup de travail, d'effort et de temps. Il constitue un système complexe dans lequel plusieurs facteurs (coordination de gestes articulatoires, attention, perception, préparation, etc.) entrent en jeu pour faire émerger le résultat que nous entendons. De même, dans le contexte des imitateurs professionnels, le processus d'imitation vocale ne fonctionne pas parfaitement : l'imitation « exacte » serait donc impossible. En vue d'illustrer nos propos, nous présentons deux exemples d'études sur l'imitation phonétique volontaire.

Schlichter (1995) a réalisé une expérience d'imitation vocale, portant sur l'analyse de la prosodie et du timbre vocalique, dans laquelle elle a comparé les productions d'Yves Lecoq⁹⁰, imitant deux figures politiques (Jacques Delors et Édouard Balladur). L'exercice consistait à écouter des extraits⁹¹ d'enregistrements de ces derniers et à les reproduire, aussi fidèlement que possible. L'auteur a constaté que même si l'imitateur a reproduit des discours se rapprochant énormément de ceux des individus imités, quelques points des deux facteurs étudiés divergent toutefois : les triangles vocaliques, ainsi que certains éléments prosodiques ne coïncident pas avec ceux des personnages

87 Extrait sur Youtube : <https://www.youtube.com/watch?v=ZmqDNklP2A> (consulté le 21/10/2018).

88 C'est aussi le cas des ventriloques.

89 Lorsqu'ils écoutent leur voix enregistrée, ils la percevraient sans déformation, contrairement à la plupart des personnes (Homeovox, s. d.).

90 Imitateur professionnel (voir Imitation vocale volontaire, page 10).

91 Discours enregistrés dans des conditions d'énonciation identiques (émissions de radio).

politiques imités. Il y avait chez l'imitateur des phénomènes de sous- et de sur-articulation concernant le débit de parole et les gestes articulatoires, ce qui aurait entraîné une modification des cavités de résonances, principales responsables du timbre, et donc des différences de fréquences émises sur le plan acoustique (p. 2). D'un point de vue prosodique, Yves Lecoq distinguerait bien les indices indispensables à l'identification de chaque imité. Il contrôle habilement les pauses, les liaisons, la vitesse d'élocution ainsi que la courbe mélodique. Pour imiter Jacques Delors, Yves Lecoq modélise son discours en exagérant les écarts de fréquences vers le haut et en parlant un peu plus vite que d'habitude. Pour imiter Édouard Balladur, il adapte sa vitesse d'élocution en parlant lentement et accentue l'aspect monotone (p. 74). Nous précisons qu'il n'est guère évident de disposer dans une imitation de tous les éléments nécessaires à une identification, qui de toute manière n'est possible que lorsque l'on connaît le modèle. Il est alors primordial pour l'imitateur d'avoir une capacité auditive exceptionnelle l'aidant à détecter la qualité exacte du timbre (clair, enroué), la mélodie (modulée, monotone), la hauteur de la voix (grave, aigue), etc. Les imitateurs professionnels peuvent prêter attention à ces facteurs mais n'arriveraient jamais à produire une imitation exacte (p. 75). L'auteure, partant de ces différences anatomiques (cavités) propres à chaque individu, parle alors d'empreinte vocale⁹² (*vocal print*) en référence aux empreintes digitales (*finger print*). Elle conclut son travail en ces termes :

Une fois de plus, nous constatons que les fréquences vocaliques des imitations divergent de celles obtenues pour la voix naturelle de l'imitateur. Cette constatation prouve qu'il y a un changement dans la configuration du conduit vocal d'Yves Lecoq lorsqu'il imite. Modifiant inconsciemment ses cavités de résonance, il privilégie ainsi d'autres harmoniques qui donnent naissance à des variations de fréquences des deux premiers formants. L'adaptation sera différente selon le locuteur (Schlichter, 1995, p. 71).

⁹² Ce terme est vivement contesté aujourd'hui par nombre de chercheurs, car aucune étude scientifique n'a encore démontré son existence.

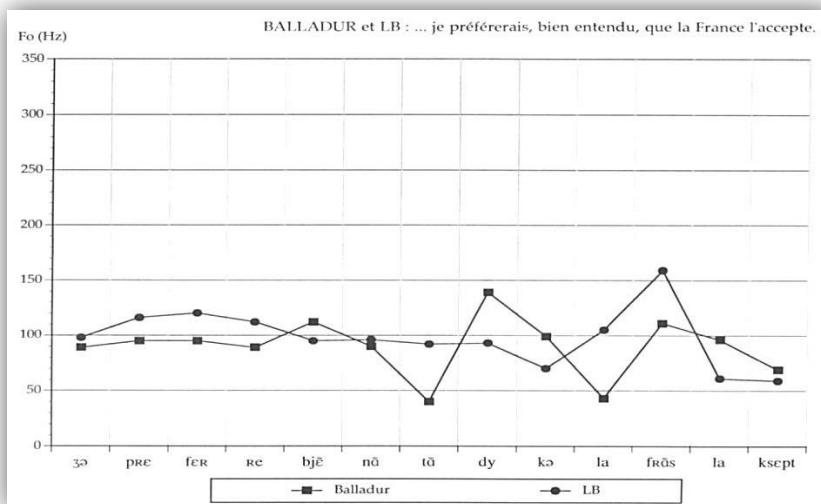


Figure 9 : points de convergence et de divergence dans les contours intonatifs de l'imité (Balladur) et de l'imitateur (Yves Lecoq-LB).

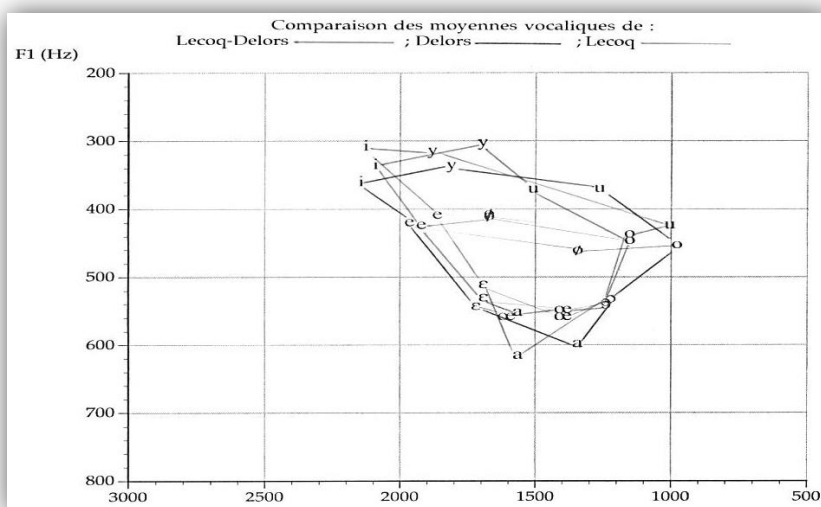


Figure 10 : points de convergence et de divergence dans la superposition des triangles vocaliques de l'imité (Delors) et de l'imitateur (Lecoq = voix naturelle, Lecoq-Delors = imitation).

Dans ce deuxième exemple, il ne s'agit pas de l'imitation vocale d'une personne par un professionnel, mais celle de caractéristiques phonétique et phonologique propres à un ensemble de personnes, d'une communauté linguistique entière. C'est également cet aspect qui nous intéresse et que nous cherchons à mettre en avant dans notre étude : l'imitation d'accent. Ayant travaillé sur les caractéristiques des consonnes voisées (/b/, /d/ et /g/) et non voisées (/p/, /t/ et

/k/) de l'alsacien, Steiblé (2014) a mis en place une expérimentation d'imitation volontaire d'accent, dans laquelle des locuteurs français⁹³ non alsaciens (7 femmes et 7 hommes de 20 à 30 ans) ont imité leurs homologues alsaciens. Ils devaient reproduire telles qu'ils le percevaient, des phrases (18 phrases répétées 12 fois) en français contenant des consonnes de l'alsacien. Trois conditions distinctes de productions étaient utilisées : initiale de mot, intervocalique et finale de mot (exemples : la date est juste, l'abattoir est fermé le matin, le crabe a des pinces). Rappelons au passage que l'alsacien est une langue alémanique (germanique) dans laquelle l'opposition consonnes voisées-non/voisées n'existe pas ou peu visiblement contrairement au français. Les Alsaciens prononceraient les consonnes /b/, /d/ et /g/ comme /p/, /t/ et /k/⁹⁴ en jouant sur les paramètres *lenis* et *fortis*. Les résultats de l'expérimentation ont montré que les imitateurs étaient capables de percevoir et produire le dévoisement, caractéristique de l'accent alsacien. Certains locuteurs ont « parfaitement » imité les schémas propres à l'alsacien, en reproduisant des trajectoires articulatoires adéquates (les bons *lenis* et *fortis*). D'autres n'avaient pas adopté les gestes articulatoires qui conditionnent les patterns⁹⁵ temporels du dévoisement (VOT, silence acoustique, VTT) en alsacien (Steiblé, 2014, p. 229). Au-delà des nombreuses différences individuelles notées, des cas d'exagérations ont également été observés : les VOT et les silences consonantiques sont plus longs que ceux produits par les Alsaciens. L'auteure remarque que :

[...] si on ne peut juger de la perception de locuteurs qui imiteraient « mal » l'accent alsacien en français, il est au moins certain que ce qui est produit a été perçu, à tout le moins suffisamment pour conduire à la reproduction de certains paramètres articulatoire-acoustiques. En effet, les locuteurs les plus performants sont également ceux qui sont en contact depuis longtemps avec le dialecte ou le français d'Alsace, ou ceux qui ont un contact récent mais quotidien avec ces langues ou variantes phonétiques (Steiblé, 2014, p. 228).

Cette citation nous paraît d'une grande importance, car elle évoque la notion de *l'input* bien connue en didactique des langues. En effet, les travaux des chercheurs Krashen & Terrel (1983) misent sur la quantité de *l'input* : le temps d'exposition (contact long) et la fréquence (contact quotidien) seraient déterminants pour un apprentissage de langues « réussi ». Les auteurs

93 Ils ont déclaré connaître l'alsacien et savent l'imiter.

94 La thèse de Steiblé a montré que le trait de dévoisement des consonnes en alsacien n'est pas seulement lié aux vibrations des cordes vocales mais concerne aussi les éléments intersegmentaux comme le VOT (Voice Onset Time ou délai d'établissement du voisement), la durée du silence acoustique et le VTT (Voice Termination Time ou délai d'arrêt du voisement).

95 Peut se traduire par patrons, mais nous préférons garder le terme utilisé par l'auteur.

développent un modèle appelé *Input hypothesis* dans lequel ils placent la compréhension de l'*input* au centre de toute acquisition. Selon leur postulat, l'exposition massive à de l'*input* compréhensif conduirait systématiquement l'apprenant de langue vers une appropriation de la langue. Pourtant, Ellis (2002) a montré que la quantité de l'*input* n'est pas une condition suffisante. La fréquence d'utilisation de la langue et l'attention portée aux *inputs* seraient également nécessaires. Pour sa part, Swain (2018) propose l'argument inverse en soutenant le modèle de l'*output hypothesis*, dans lequel la production par l'apprenant (output) serait plutôt un facteur déterminant pour l'acquisition.

Les résultats de Schlichter (1995) et Steiblé (2014) confirment le fait que l'imitation d'accents ne peut être exacte. Le phénomène de l'imitation vocale se manifeste à plusieurs niveaux : certains sons/traits seraient plus facilement imitables que d'autres. Cela dépendrait aussi de plusieurs facteurs tels que la proximité des langues, le « don » qu'ont certaines personnes à imiter les sons, etc. Des études ont aussi montré que les locuteurs « perdent » leurs acquis et reviennent sur leur production naturelle (originelle) en condition perturbée lors des tâches d'imitation vocale volontaire (Markham, 1997). D'ailleurs, l'imitation vocale consciente peut avoir les mêmes caractéristiques que le mimétisme⁹⁶ observé chez les animaux : se faire passer pour une personne afin d'échapper aux dangers, bénéficier de faveurs, cacher son identité, être bien perçu, etc. Pour finir, nous rappelons que vouloir imiter consciemment la voix d'une personne n'est pas un acte facile à reproduire et qu'il requiert des capacités fines de perception et de production. Cela nécessite des efforts d'attention à des éléments vocaux et tout un appareillage de couplage en vue de « sonner » comme cette personne. À cause de tous ces facteurs, nous avons également vu que l'imitation vocale volontaire n'est jamais parfaite. Voyons dans la section suivante, ce qui caractérise l'imitation vocale inconsciente.

4.4.2 Imitation vocale involontaire

Quant à l'imitation vocale involontaire ou inconsciente, elle représente une situation dans laquelle un individu « parle » comme un autre, sans le savoir, ni le vouloir, surtout après un « contact » de durée variable (Pardo, 2006 ; Pardo, Gibbons, Suppes, & Krauss, 2012 ; Pardo, Jay, &

96 Nous utilisons ce terme pour désigner les stratégies volontaires ou involontaires, naturelles ou artificielles qu'emploient certains animaux pour échapper aux dangers ou pour attraper leurs proies. Le Larousse le définit également comme la « Reproduction machinale, inconsciente, de gestes et d'attitudes des gens de l'entourage ».

Krauss, 2010). L'adage qui dit « qui se ressemble s'assemble » serait inversé dans le cas de l'imitation ? C'est-à-dire, qui s'assemble se ressemble ? Ce type d'imitation vocale peut également advenir dans des conditions expérimentales (en laboratoire), mais en l'absence de toute consigne explicite. Dans ce cas, l'objectif de l'étude est initialement inconnu, c'est-à-dire que les participants ignorent complètement qu'ils participent à une étude sur l'imitation vocale. Ainsi, pour Steiblé (2014, p. 41), l'imitation involontaire « est un phénomène que l'on constate en situation normale de parole : dans le cadre d'interactions spontanées, la présence d'une imitation d'accent semble assez naturelle, et émerge sans aucune consigne particulière ». Dans la majorité des cas, ce phénomène se manifeste naturellement, sans aucun effort de reproduction, et serait le résultat de la fréquence et du temps d'exposition aux exemplaires⁹⁷. L'imitation vocale involontaire est difficilement quantifiable par nature car il est difficile de savoir à quel moment elle se produit et à quelle fréquence d'exposition ; cela se déroule à l'insu de l'observation du chercheur (Sato, Grabski, Granjon, Schwartz, & Nguyen, 2010). Seuls les résultats de l'imitation vocale involontaire sont mesurables, à condition d'avoir des données de pré-test. Kappes et collaborateurs (2009) ont également mis en évidence des degrés d'imitation involontaire au travers des répétitions des non-mots (*nonword*). Une fois établie, l'imitation involontaire devient difficilement contrôlable, et s'ancre dans les manières de parler. L'une des principales caractéristiques de l'imitation involontaire est que, lorsque les conditions de discours sont perturbées, les éléments étudiés restent stables. En production de parole, suite à une augmentation de la vitesse d'élocution par exemple, la durée de certaines voyelles et consonnes se réduit. Les phénomènes sont robustes à la perturbation et entraînent des effets compensatoires, en fonction de la résistivité des gestes articulatoires (Vaxelaire, 2007).

Parallèlement, afin de répondre à l'interrogation « *Does listening to and observing the speaking interlocutor influence phoneme production ?* », Gentilucci & Bernardis (2007) ont mené une étude dans laquelle des participantes devaient reconnaître et, ensuite, répéter la chaîne de phonèmes /aba/ présentée visuellement, acoustiquement et audiovisuellement par des acteurs. L'étude consiste en deux tests d'imitation vocale involontaire. Dans l'expérience 1, un acteur a présenté la chaîne de phonèmes. La cinématique des lèvres et les spectres vocaux des participantes (14 femmes volontaires de 22 à 25 ans) ont été comparés à ceux d'une condition contrôle de lecture.

⁹⁷ Nous traitons ce point dans la session suivante.

L'expérience 2 était menée auprès d'un nouvel échantillon composé de 10 femmes volontaires âgées de 21 à 23 ans. Des actrices et acteurs ont présenté la chaîne de phonèmes. La cinématique des lèvres et le spectre vocal des réponses des participantes aux acteurs (et non des actrices) ont été comparés à ceux des actrices (condition contrôle), puis à ceux des acteurs. Dans les deux expériences, toutes les participantes étaient naïves quant au but de l'expérience. Parmi les chaînes de phonèmes utilisés (*/aga/*, */ada/* et */aba/*⁹⁸), seul le logatome */aba/* faisait l'objet de traitement, car il permet d'obtenir des mouvements de lèvres plus visibles et donc plus aisément enregistrables par la caméra. Les autres logatomes (*/aga/* et */ada/*) servaient uniquement de distractions afin d'éviter tout automatisme. Gentilucci & Bernardis (2007) ont utilisé la notion bien connue selon laquelle les mouvements labiaux des hommes sont habituellement plus grands et que leurs formants de voix sont plus bas que ceux des femmes. Ils ont posé l'hypothèse que les femmes percevraient et imiteraient des acteurs masculins présentant visuellement et acoustiquement des chaînes de phonèmes. Par conséquent, les auteurs s'attendaient à une augmentation des paramètres cinématiques des lèvres et à une diminution des formants de la voix chez les participantes : une hypothèse qui a été validée puisque c'est exactement ce qui s'est produit. Les résultats des deux expériences montrent que la cinématique des lèvres dans les présentations visuelles et les spectres vocaux dans les présentations acoustiques ont changé par rapport aux conditions de contrôle, ces derniers s'approchant plus des valeurs des acteurs que des actrices. La variation de la cinématique des lèvres a également induit des changements dans les formants de la voix, mais seulement dans la présentation visuelle. Les deux caractéristiques de la cinématique des lèvres et des spectres vocaux auraient tendance à être machinalement imitées, lors de la répétition d'une chaîne de phonèmes présentée par un interlocuteur qui parle de façon visible et/ou audible. L'utilisation de l'imitation conduirait à une tendance automatique et inconsciente de l'auditeur à interagir étroitement avec son interlocuteur (p. 5). Ces chercheurs concluent que les résultats sont conformes à l'idée que les circuits résonants sont activés par le système miroir, qui relie l'observation à l'exécution des gestes du bras et de la bouche : « [...] *imitation is related to an automatic process of resonance, which expresses an automatic and unconscious tendency of the perceiver to interact closely with a visible and audible interlocutor* » (p.

98 Aussi utilisés pour des expériences portant sur l'effet McGurk (voir chapitre précédent : Perturbation non écologique ou « artificielle »).

7). Nous pouvons d'ores et déjà retenir de cette étude que l'imitation vocale involontaire est bien possible au travers des moyens audio-visuels et que l'interaction est importante pour qu'elle puisse se mettre en place.

S'appuyant sur la théorie de l'accommodation de la parole ([*speech*] *communicative accommodation theory* – Giles & Ogay, 2007), Pardo et collaborateurs (2012) ont montré que les étudiants qui partageaient la même « chambre » (colocation) s'imitent phonétiquement mutuellement, de manière involontaire au fil du temps. Selon la théorie invoquée par ces auteurs, il y a convergence ou divergence lorsque deux personnes s'engagent dans une conversation : « convergence refers to the ways in which a talker adjusts speaking style to become more similar to an interacting partner, whereas divergence refers to changes in speaking style that result in reduced similarity to a partner » (Pardo et al., 2012, p. 190).

Deux tests ont été élaborés et ont permis de collecter les données de ce travail : un test de production (lecture de mots) et autre de perception. Pour le test de production, cinq couples d'élèves américains et coréens (N=10) venus de différents États (New York, New Jersey, Tennessee, Floride, et Séoul) ont été enregistrés, à des intervalles étalés sur une année académique : la première fois dès leur inscription en août au *Columbia College* et avant leur installation en chambre universitaire, la seconde en octobre, la troisième en décembre et la dernière à la fin de l'année académique en janvier. Le test de perception implique un jugement de prononciation des productions des locuteurs par 30 étudiants du même établissement. Aucun participant ne présente de troubles auditifs ou de parole. Les voyelles étudiées se trouvent dans des mots-cibles (hV⁹⁹t/d, exemple : *hot*) placés dans des phrases porteuses (*say X¹⁰⁰ again*). Ils ont également dû prononcer deux fois deux phrases contenant des mots (*dark suit, greasy, oily rag, and wash water*) avec les mêmes voyelles cibles et dont la prononciation varie selon les dialectes des différentes régions des États-Unis. Les résultats ont montré qu'il y a eu convergence phonétique mais à différents degrés selon les quatre étapes : « *The degree and patterns of convergence varied across pairs, phrases, and measures.* » (p. 196). Grâce aux informations fournies par un questionnaire rempli par les locuteurs, les auteurs ont également observé une corrélation entre le degré d'amitié et la convergence. Les

99 V = voyelle.

100 X = mot étudié.

résultats du test de perception sont venus corroborer ceux issus de l'analyse acoustique. Ces uniques analyses n'ont pas permis aux auteurs d'en tirer des conclusions suffisantes :

At this point, the most effective way to assess these relationships is to rely on the judgments of ordinary listeners. Because ordinary perception integrates multiple dimensions simultaneously, a carefully designed perceptual similarity test provides a global assessment of phonetic convergence without committing to a single acoustic–phonetic attribute that might not be used consistently by every talker on every occasion [...] a complete understanding of phonetic convergence is unlikely to result from acoustic analyses alone (p. 196).

En somme, nous retenons que l'imitation vocale involontaire est un processus par lequel des éléments phonétiques, phonologiques et prosodiques sont enregistrés dans le répertoire phonique de l'individu qui est exposé à ces *inputs*, à son insu et grâce au système miroir. Ces quelques études présentées ci-dessus montrent amplement les conséquences sur l'acquisition sonore et notre manière de communiquer. Dans la section suivante, nous présentons une autre théorie qui explique le fonctionnement de l'imitation vocale involontaire à partir de la notion de catégorisation (prototypes ou exemplaires).

4.4.3 Théorie de l'exemplaire

La convergence vocale involontaire fonctionnerait par catégorisation. Catégoriser reviendrait à ranger un certain nombre d'items phonético-phonologiques dans une même classe (ou catégorie). Il existe plusieurs conceptions de la notion. Les plus importantes sont la théorie classique, la théorie du prototype et la théorie de l'exemplaire. Selon la conception classique de catégorisation, une catégorie est définie par un ensemble de propriétés nécessaires et suffisantes. Un item « X » sera catégorisé comme un « A », si « X » possède les caractéristiques qui définissent la catégorie « A ». Par exemple, la catégorie « voyelle ouverte » sera définie comme la classe des voyelles à grande aperture. Ainsi, une voyelle qui sera produite avec une grande aperture serait systématiquement considérée comme appartenant à la catégorie « voyelle ouverte ». Cette conception est souvent critiquée car elle ne prend pas en compte les items sur un continuum d'appartenance : il n'y aurait pas de solutions intermédiaires. Tout se passe comme si tous les exemplaires d'une catégorie étaient équivalents et qu'ils représentaient de la même manière la catégorie (Ettlinger, 2007). Or, en phonétique, un son est défini par plusieurs traits (aperture, nasalité, antériorité/postériorité, etc.).

La théorie du prototype¹⁰¹ part des failles de la théorie classique et pose l'hypothèse selon laquelle nos catégories mentales s'organiseraient plutôt autour d'une moyenne abstraite des exemples rencontrés (*Ibid.*, p. 405). Tous les items d'une catégorie seraient représentés à partir d'un exemple modèle (le prototype) facilement reconnaissable. Certains membres d'une catégorie sont plus représentatifs que d'autres. Une chaise est plus fréquemment citée comme meuble qu'une table de nuit (*Ibid.*). Dans la catégorie « apprenant », « élève » serait le prototype de « étudiant », « lycéen », « collégien », etc. Le prototype phonétique serait donc le meilleur exemplaire d'une catégorie sonore dans un répertoire donné. Or, cette approche ne fonctionne pas en phonétique car chaque son est unique et donc difficilement représentable par un autre. Même si l'on considère qu'en français, certains sons sont plus fréquents que d'autres (Wioland, 1991), il serait erroné de dire que /œ/ est le prototype de /e, ø, ə/. Le terme archiphonème¹⁰² serait plus adapté (/œ/ est l'archiphonème de [œ] et [ø]) (*ibid.*).

La théorie de l'exemplaire nous intéresse particulièrement car elle va à l'encontre des deux premières théories (quoique parfois confondue avec la théorie du prototype, comme nous le verrons plus bas¹⁰³) et rejette l'idée que les sujets construisent une représentation abstraite des exemplaires d'une catégorie. Elle suppose que nous stockons en mémoire l'ensemble des exemples d'une catégorie. Par conséquent, pour catégoriser un nouvel exemple, nous pourrions le comparer aux exemples de la catégorie stockée en mémoire (Ettlinger, 2007). La théorie de l'exemplaire est introduite pour la première fois en psychologie en tant que théorie de « catégorisation » et, de manière générale, part de l'idée que les éléments que nous apprenons sont comparés entre eux, puis classés selon des similarités avec des catégories déjà existantes dans notre répertoire (Schweitzer *et al.*, 2015). Elle est utilisée en sciences du langage pour l'étude de la catégorisation d'items langagiers (vocabulaire, phonèmes, etc.). Pour Ettlinger (2007, p. 686), « *Exemplar theory's application to linguistics is based on the idea that each of the sounds of the language represents a category which is represented by a cloud of the remembered tokens of that category* ». La théorie stipule que nous sommes capables de reproduire les phonèmes d'une langue selon la fréquence de

101 Initiée dans les années 1970 à partir de la notion d'« air de famille » par Eleanor Rosch, psycholinguiste à l'université de Berkeley.

102 Un archiphonème est une unité phonologique regroupant les particularités distinctives de deux ou plusieurs phonèmes (Wioland, 1991).

103 En fait, ce sont deux théories alternatives et concurrentes qui font des prédictions souvent équivalentes (Ettlinger, 2007). Elles ont été reprises par de nombreux linguistes et étendues à d'autres domaines.

leur écoute (exposition). Plus nous écoutons et percevons un phonème, plus nous serons en mesure de le reproduire (Hulstijn, 2003).

Dans ce modèle, lorsque nous écoutons des exemples de phonèmes, ceux-ci sont d'abord stockés et organisés dans notre mémoire à court terme¹⁰⁴ en catégories de « nuages¹⁰⁵ ». Le risque d'oubli serait ainsi élevé. Pour que ces éléments passent de la mémoire à court terme à la mémoire à long terme, la répétition doit être augmentée. Les nuages contenant des unités sonores phonétiquement similaires sont rapprochés, tandis que ceux qui diffèrent sont (davantage) éloignés. De ce mécanisme, il ressort deux conséquences : d'une part, il sera difficile de reproduire un son auquel nous ne sommes pas beaucoup exposés, et d'autre part, il serait quasi impossible de reproduire un phonème mal perçu ou perçu différemment de la façon dont les natifs le produisent, car : « ... *a nonnative prosodic category is categorized when listeners perceive it as an exemplar of some L1 prosodic category* » (So & Best, 2014, p. 199). La théorie de l'exemplaire est utilisée en phonétique dans le cadre de la perception et de la production de la parole. Par exemple, la catégorie du son /i/ consiste en l'ensemble des *tokens* ou exemples des /i/ écoutés et stockés par un auditeur. Elle se base également sur les valeurs de F1 et de F2 (premier et deuxième formant). Toutes les voyelles dont le F1 avoisine 285 Hz (+/- 5 %) sont placées dans une même zone dans le triangle vocalique, et ainsi de suite. Cette théorie conditionne la perception des sons du langage et serait également un des vecteurs des changements vocaliques (Ettlinger, 2007) et de l'acquisition phonologique d'une langue étrangère, proche des locuteurs natifs (Bybee, 2003). À ces problèmes de perception, s'ajoutent ceux liés aux habitudes articulatoires. Nous produisons ce que nous percevons mais la perception est une condition nécessaire toutefois insuffisante. Il faut noter que pour produire d'une L2, les muscles articulatoires doivent être entraînés. L'articulation de la fricative dentale *th* nécessite des entraînements répétés et réguliers en vue d'une maîtrise parfaite de ce son étranger (pour un natif français), car, bien entendu, ce son n'existe pas dans le répertoire phonétique français. La convergence phonétique devient possible lorsque les zones de stockage de phonèmes d'un non-natif coïncident avec, ou se rapprochent de celles d'un natif. La voyelle /a/ a presque les mêmes valeurs formantiques pour tous les individus, avec quelques différences d'écart entre les femmes, les hommes et les enfants (voix grave ou aigue). Selon certains chercheurs

104 Nous conservons l'opposition « court » et « long » pour des raisons de simplicité, mais en réalité il s'agit de la mémoire de travail. (Voir Baddeley, 1986) pour plus de détails).

105 Métaphore signifiant catégorie, classe, regroupement, etc.

(Goldinger, 1996), les effets des exemplaires diminuent avec l'âge et seraient limités¹⁰⁶ pour chaque langue et chaque locuteur (Goldinger, 1996, p 17).

La théorie de l'exemplaire est parfois associée à celle du *Usage-based*¹⁰⁷ (Ellis, 2002 ; Tomasello, 2003) car elles sont toutes deux assujetties au facteur articulation, comme le confirme Pavlík (2016, p. 111) :

Usage-based theory (Exemplar theory) tries to model linguistic phenomena in terms of usage patterns, frequency of occurrence, categorization, language variation, and language change. [...] it speculates that surface forms are not generated from underlying representations, but every word (or phrase) in the lexicon consists of a cloud/cluster of phonetic exemplars (variants) containing detailed information about the context in which they are used.

Cette notion de fréquence peut prêter à confusion si le type n'est pas spécifié. À ce niveau, il est important de préciser la différence entre la fréquence d'exposition et la fréquence d'usage. La première représente la quantité et la régularité d'écoute, de visionnage, d'observation, etc. de l'*input*, alors que la deuxième concerne l'utilisation de l'*output*. En effet, la théorie de l'exemplaire considèrerait l'exposition à l'*input* comme point de départ. Un mot écouté une seule fois mais utilisé à plusieurs reprises dans le temps a une forte probabilité d'être ancré dans la mémoire à long terme, contrairement à un mot écouté plusieurs fois dans le temps, mais qui ne sera utilisé qu'une seule fois. Dans le cadre de l'imitation vocale involontaire, selon la situation de communication dans laquelle se trouve le locuteur, les exemplaires les plus représentatifs qui sont stockés dans la mémoire à long terme surgissent d'eux-mêmes. Catégoriser un nouvel objet repose ainsi sur l'évaluation de sa similitude avec les exemplaires stockés. Le locuteur ne choisit pas les exemplaires à utiliser : le stockage dans la mémoire à long terme facilite la restitution involontaire lors de la production de la parole. En d'autres mots, la quantité des exemplaires à laquelle un locuteur s'expose au travers des médias pourrait faciliter l'acquisition de formes phonétiques et phonologiques. Ce phénomène pourrait être à l'origine de l'imitation vocale involontaire dans le cadre de l'exposition aux médias, aspect que nous traitons dans la section suivante.

106 Le nombre de fois qu'un auditeur écoute une voyelle est limité à environ dix millions de fois au cours de sa vie (Goldinger, 1996, p. 17).

107 Apprentissage fondé sur l'usage de la langue, à partir de la fréquence des exemplaires et d'autres compétences cognitives (lecture d'intentions et recherche de patterns).

4.5 Imitation directe (non-médiée) et indirecte (médiée)

La convergence en général et vocale en particulier peuvent aussi être directes ou indirectes. Par « directe », nous entendons un acte qui a lieu en présentiel, c'est-à-dire en face à face, sans aucun intermédiaire. Par opposition, l'accommodation indirecte se déroule au travers d'un média. Le téléphone, les applications de visioconférence, la radio, la télévision, etc. en sont quelques exemples. Ainsi, les études sur la convergence en général portent en majorité sur des situations directes parce que les liens sont plus faciles à établir que dans les situations indirectes. Les difficultés liées à l'étude de la convergence indirecte résident dans le fait que celle-ci est différée, ses conséquences ne sont donc pas mesurables « à chaud ». Nous situant dans le cadre de l'imitation à travers les médias, et en particulier la télévision, il nous semble important de mentionner une étude qui avait relevé il y a cinquante ans, l'impact de la télévision sur le comportement. Dans leur étude, Bandura, Ross & Ross (1963) ont mis en place une situation expérimentale (pré-test, test, post-test) en vue d'examiner le comportement des enfants qui s'exposent aux actes violents et agressifs dans la vie réelle et à travers les films et dessins animés. Ces auteurs ont ainsi apporté les premières preuves de l'imitation à travers la télévision et ont conclu en ces termes :

Filmed aggression, not only facilitated the expression of aggression, but also effectively shaped the form of the subjects' aggressive behavior. The finding that children modeled their behavior to some extent after the film characters suggest that pictorial mass media particularly television, may serve as an important source of social behavior (p.9).

Ils ont également trouvé que l'imitation de comportement était plus importante suite à l'exposition télévisuelle que dans les autres conditions :

In fact, a possible generalization of responses originally learned in the television situation to the experimental film may account for the significantly greater amount of aggressive gun play displayed by subjects in the film condition as compared to subjects in the other conditions (ibid.).

Si les études sur l'imitation vocale directe sont très nombreuses dans la littérature (Trudgill, 1986 ; Kuhl *et al.*, 2008 ; Pepperberg & Wilkes, 2004 ; Pickering & Garrod, 2004 ; Chambers, 1998 ; Kail, 2015), la convergence phonético-phonologique indirecte est rarement mise en évidence. Il a été longtemps supposé que l'imitation vocale, plus précisément l'imitation d'accents, n'est possible qu'au travers d'interactions directes (présence physique d'interlocuteurs), mais comme nous

l'avons montré au sein du chapitre I, l'accommodation vocale indirecte est elle aussi attestée. D'une part, sur l'impossibilité de l'imitation vocale médiée, nous pouvons par exemple citer les travaux de Kail (2015) : « [...] si le vocabulaire peut être acquis par exposition à des programmes télévisuels, en revanche, les aspects plus complexes du langage comme la phonétique ou la grammaire ne peuvent pas l'être. » (p. 43). Kuhl *et al.* (2008), confirment cette hypothèse : « *Intriguingly, recent work testing infants' perception of a foreign-language phonological contrast suggests that live interaction with a speaker of the language gives rise to learning where equivalent exposure via video fails to* ». (p. 979) Dans une autre étude plus récente, Kuhl (2010) montre l'importance de l'interaction sociale dans l'apprentissage-acquisition des langues étrangères. Elle étudie deux groupes de nourrissons de neuf mois : l'un a suivi 12 séances de chinois mandarin au travers d'interactions naturelles (présentielles donc) avec un locuteur chinois, et l'autre groupe a reçu les mêmes informations linguistiques pendant les mêmes durées, mais transmises par les médias. Les résultats ont montré que le groupe en interaction physique a acquis des éléments linguistiques, contrairement au groupe audiovisuel. Elle rappelle que : « [...] *referential information is present in both the live and televised conditions, but it is more difficult to pick up via television, and is totally absent during audio-only presentations. Gaze following is a significant predictor of receptive vocabulary* » (Kuhl, 2010, p. 720). Ces résultats sont fort intéressants, mais il nous semble que les enfants (surtout de cet âge) imitent mieux en présentiel, car ils ont besoin de suivre les mouvements des adultes pour comprendre¹⁰⁸. Aussi, les informations présentées dans les vidéos engendrent probablement de la surcharge cognitive¹⁰⁹, puisqu'il y a beaucoup d'animations (Chanquoy, Tricot & Sweller, 2007 ; Amadiou & Tricot, 2006). Ils seraient donc incapables de sélectionner des informations parmi toutes celles présentées (souvent simultanément) à la télévision. Les adultes parviennent à effectuer des tris d'informations selon leurs besoins.

108 C'est pourquoi ils regardent souvent la main plutôt que l'objet qui est pointé.

109 Leur cerveau n'est pas encore entraîné à gérer plusieurs informations (animations) à la fois et leurs mouvements moteurs sont encore très lents.

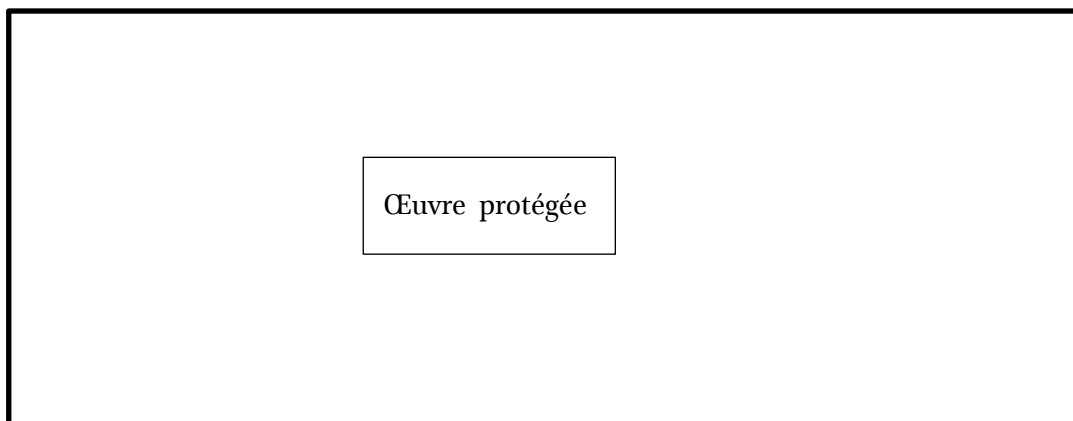


Figure 11 : conditions expérimentales de l'étude de Kuhl : à gauche l'exposition interactive directe, à droite l'exposition indirecte¹¹⁰ (Kuhl, 2010, p. 720).

Plusieurs études récentes ont montré le contraire des observations faites par les recherches que nous venons de mentionner : l'interaction directe n'est pas la seule condition de l'acquisition du langage. Nous renvoyons le lecteur au chapitre 1 au sein duquel plusieurs recherches ont montré les influences des médias sur la prononciation, notamment celles de Stuart-Smith *et al.*, 2007, 2011, 2013 ; Stuart-Smith et Ota (2014) ; Ota et Takano (2014) ; Sanchez, Miller & Rosenblum (2010), pour n'en citer que quelques-unes.

Si l'imitation phonétique audio-visuelle est prouvée dans les conditions expérimentales, pour des unités de sons plus petites, il serait intéressant de mener des recherches en conditions non-expérimentales, c'est-à-dire en l'absence de stimuli. Nous pouvons imaginer les mêmes processus dans le cadre de l'apprentissage informel de l'anglais en ligne (Sockett, 2014 ; Toffoli & Sockett, 2015), où un apprenant, après le visionnage d'une série, imite un accent particulier lorsqu'il interagira avec un locuteur anglophone. Par conséquent, un auditeur peut imiter phonétiquement et phonologiquement après un temps d'exposition aux médias, car l'imitation vocale est un « *phenomenon that could be facilitated by audio-visual congruence* » (Nguyen & Delvaux, 2015 : 47). Dans la section suivante, nous allons nous intéresser à quelques théories/modèles de l'acquisition phonétique et phonologique qui expliquent comment l'apprenant acquiert les sons d'une L2. Parmi les modèles existants, nous en présenterons quatre, qui, à nos yeux, sont les plus connus et les plus représentatifs.

¹¹⁰ Remarquons la luminosité de la pièce et celle de l'écran, la taille de l'écran et la distance entre l'écran et les yeux de l'enfant.

4.6 Des modèles d'acquisition phonétique et phonologique

Il existe plusieurs théories de l'acquisition de la matière phonique dans la littérature. Quatre modèles, quelque peu différents dans leur approche mais complémentaires, voire interconnectés méritent amplement d'être mentionnés. Les modèles que nous présentons¹¹¹ sont pour nous d'une importance capitale, car ils expliquent les mécanismes sous-jacents de l'apprentissage phonéto-phonologique en L2. Ils constituent les seules théories de prononciation existantes (à notre connaissance) qui se focalisent explicitement sur l'acquisition du langage L2 et permettraient de prédire l'apprentissage de la prononciation. Même si la prononciation (production) est l'aspect central, la perception tient évidemment une place primordiale et comme nous l'avons évoqué, la perception auditive est une condition nécessaire pour une imitation vocale réussie. Aussi, plusieurs notions et termes abordés *supra* viennent faciliter leur compréhension et l'on pourra aisément entrevoir la passerelle qui relie ces modèles et l'imitation vocale. Nous commençons par le premier modèle qui est l'*Acoustic Image Imprinting* (AII).

4.6.1 Acoustic Image Imprinting

Le modèle de l'*Acoustic Image Imprinting* (AII), initié par Neufeld (1977), comprend deux hypothèses majeures. La première considère que malgré la variabilité interindividuelle dans l'apprentissage d'une langue, les adultes sont capables d'atteindre une prononciation proche des natifs, et que ce potentiel d'acquisition concerne toutes les L2. En revanche, l'auteur s'interroge sur la manière de mettre en œuvre des méthodes permettant d'atteindre cet objectif : « *the problem was how to access this potential and how to sufficiently motivate our students to exploit it* » (Neufeld, 1977, p. 47). Nous pouvons tout de suite comprendre que ce modèle discrédite l'hypothèse selon laquelle il existerait une période critique pour l'apprentissage d'une L2, c'est-à-dire que les enfants seraient plus aptes à mieux apprendre une langue que les adultes. La deuxième hypothèse part de l'idée que chaque son est représenté par une image abstraite :

We assume that, to produce native-like sounds in a language, we must have internalized an auditory representation (the phonetic image) for the sound patterns of that language; acoustic image is gradually strengthened over time and through exposure. In addition to this

¹¹¹ Nous nous sommes beaucoup inspiré du modèle de présentation de Markham (1997).

input, image sphere appears to exist in a parallel image which represents the phonetic shapes of the individual's output (Neufeld, 1977, p. 57).

Toujours selon l'auteur, lorsqu'un apprenant prononce un son d'une L2, très souvent il déforme cette « image » associée au modèle acoustique du son natif. En vue de corriger cette représentation erronée, les apprenants devraient être encouragés à écouter d'abord l'*input*, plutôt qu'à produire des sons L2, afin qu'une image acoustique correcte puisse être générée, comme Neufeld le précise : « [...] *our students be thoroughly introduced or sensitized to the sound patterns of the target language before*¹¹² *attempting to produce them* » (p. 47). Il continue en suggérant que : « *By delaying oral production, we hope to minimize contamination of the "acoustic image imprint"* » (*ibid.*). La notion du temps de latence entre l'exposition à de l'*input* et la phase de production orale est cruciale dans ce postulat de Neufeld. Il fait une analogie avec la façon dont les bébés et les jeunes enfants acquièrent, par imitation, les formes linguistiques : « *unlike the baby, who has heard his language long before attempting to imitate it [...] this undesirable interference could be minimized if the child were to refrain from speaking until he had had sufficient time to imprint the new sound patterns* » (p. 57). En effet, bon nombre de chercheurs en acquisition de L2 s'accordent sur cette phase d'assimilation de l'*input* avant toute production langagière. En revanche, cet aspect n'est pas appliqué dans la salle de classe, où l'apprenant est « forcé » à produire dans l'immédiat c'est-à-dire même en tout début d'apprentissage. Autrement, le modèle préconise également que les apprenants évitent les modèles non natifs, car ceux qui sont potentiellement capables d'atteindre une prononciation de type natif (et ils ne sont pas tous considérés comme capables) ne peuvent réussir que dans de bonnes conditions d'apprentissage. Toutefois, le modèle a été beaucoup critiqué parce que difficile à tester et aussi en raison du nombre limité d'études. Aussi, bien que le modèle de Neufeld n'ait pas pour but d'expliquer les détails de l'apprentissage sonore, l'hypothèse selon laquelle tous les sons peuvent être appris à un degré semblable à celui de la langue maternelle présente un intérêt considérable, à condition que la possibilité de générer une représentation acoustique non déformée soit fournie. L'affirmation selon laquelle cette réussite n'est pas possible pour certains apprenants (Neufeld, 1977, p. 46) est également importante. Cela montre que chaque apprenant est différent et évolue dans un système dynamique et complexe, où il n'y a pas que des progrès mais aussi de la stagnation, de la fossilisation, de l'attrition, etc. Cette

112 Souligné par l'auteur lui-même.

différence entre apprenants est aussi présentée dans ce modèle comme inhérente uniquement en termes de limitations psychosociologiques et non à des limites naturelles ou linguistiques. Néanmoins, le fait que ce modèle propose éviter les sons non-natifs est critiquable, car cela revient à privilégier une prononciation standard d'une langue. Cette conception de l'apprentissage la prononciation mise en avant par le modèle de l'*Acoustic Image Imprinting* (AII) est largement démodée, comme nous l'avons vu avec la notion de *World Englishes* (cf. section 2.3) Outre ce fait, les hypothèses de l'AII nous semblent pertinentes et ont tout leur intérêt dans le domaine de l'enseignement/apprentissage des langues. La section suivante est consacrée au second modèle que nous souhaitons présenter ici, le *Speech Learning Model* (SLM).

4.6.2 Speech Learning Model

Le *Speech Learning Model* (SLM), comme son nom l'indique, est un modèle bien connu d'apprentissage de la parole, développé par James Emil Flege (1992). Il repose sur la différence des systèmes phonétiques entre la L1 et la L2 en vue de prédire l'apprentissage de la prononciation en L2. Son objectif premier est de prendre en compte les changements tout au long de la vie dans l'apprentissage de la production et de la perception de segments phonétiques. Le SLM a également proposé une série d'hypothèses générales fondées sur des données probantes et solides, collectées sur une période de 30 ans. Tout d'abord, l'auteur pose que les bilingues ne peuvent pas séparer complètement leurs sous-systèmes phonétiques L1 et L2. Ensuite, il considère que les capacités qui sous-tendent l'acquisition réussie de la parole en L1 restent intactes pendant toute la vie. Il réfute ainsi l'idée selon laquelle l'apprentissage de la parole en L2 serait limité par une période critique (Flege, 2003, p. 327). Ainsi, il a montré que des adultes sont capables d'apprendre des sons L2 aussi bien que les enfants. Ce modèle met également l'accent sur la façon dont les individus apprennent, ou n'apprennent pas, à produire et à percevoir des segments phonétiques (voyelles et consonnes) dans une L2 ; la perception y occupe une place capitale. Selon l'auteur, un apprenant de L2 aurait besoin de distinguer trois catégories de sons : nouveau, similaire et identique. Les sons identiques en L1 et L2 ne posent pas de problème de perception-production. Lorsqu'un /t/ a les mêmes caractéristiques acoustiques et est réalisé avec les mêmes gestes articulatoires dans deux langues, les apprenants appliqueraient la loi de l'économie d'énergie. Ils ne produiront aucun effort en vue de créer une nouvelle catégorie de sons propres à la L2, mais vont tout simplement produire le /t/

déjà existant dans leur L₁, ce qui nous semble une évidence. Ensuite, les catégories des sons similaires sont théoriquement les plus difficiles pour les apprenants de L₂ car ils (les apprenants) ont tendance à calquer la production de ces sons L₂ sur celle des sons de la L₁. La classification d'équivalence est le processus par lequel les apprenants de L₂ perçoivent un son L₂ comme une instance d'une catégorie L₁, bloquant la formation d'une nouvelle catégorie phonétique. Ceci conduirait à la réalisation du son de la L₂ qui dévie de la cible native. Les sons similaires (proches ou voisins) sont par définition ceux qui partagent des points phonétiques identiques, mais divergent sur quelques caractéristiques (formants et durée). Le /i/ français et le /i:/ anglais sont des exemples classiques de sons similaires entre les deux langues. Dans la même lignée, il est communément admis que les francophones produisent souvent un /i/ français à la place du /i/ anglais lorsqu'ils parlent en anglais (Herry-Bénit, 2011, p. 78). Nous rappelons que deux causes possibles peuvent expliquer ce phénomène : soit ils ne perçoivent pas la différence entre les deux sons et concevraient donc ces deux sons comme identiques (voir point précédent), soit ils percevraient la différence, mais auraient des difficultés au niveau articulatoire (nous le verrons avec le modèle présenté dans la section suivante). Enfin, pour le troisième type de sons (identiques), l'auteur stipule qu'ils seraient mieux assimilés, puisqu'ils n'existent pas dans le répertoire phonique de la L₁. Toutefois cette condition ne saurait être remplie que lorsque l'apprenant perçoit les éléments phonétiques caractéristiques (durée, hauteur, etc.) nécessaires à la prononciation de ces sons nouveaux. Il faut ainsi entendre qu'il existerait une corrélation étroite entre la perception et la production. Cependant, nous pensons que certains sons défient cette « règle ». La fricative dentale *th* anglais (le voisé /ð/ et le non voisé /θ/) pour des apprenants francophones ou la voyelle antérieure fermée /y/ français pour les apprenants anglophones sont des exemples de sons nouveaux (n'existant pas dans leur répertoire phonique de base). Pourtant, la prononciation de ces sons poserait problème à ces locuteurs. La parole est un geste moteur qui nécessite un temps d'apprentissage et d'entraînement, la perception étant un prérequis. La corrélation entre perception et production constitue une des critiques majeures portées à l'égard de ce modèle et que le prochain modèle de la sous-section suivante essaie de rectifier.

4.6.3 Perceptual Assimilation Model

Le *Perceptual Assimilation Model* (PAM) est élaboré par Catherine Best et ses collègues (Best 1995 ; Best, McRoberts & Goodell, 2001). Le modèle de Best est, à bien des égards, analogue au SLM, dans la mesure où il a également pour ambition d'expliquer le comportement de l'apprenant dans l'acquisition de sons de la L2, en tenant compte de la perception de la relation entre les sons des L1 et L2. La principale dissemblance réside dans le fait que le PAM fait des hypothèses explicites sur l'articulation. Tout comme le SML de Flege, le PAM prend aussi en compte la relation entre la L1 et la L2 mais ne traite pas les nouveaux sons comme étant parfaitement maîtrisables. Sur l'exemple du <th>, Best reconnaît l'importance de nouvelles catégories de sons, mais explique que les obstacles articulatoires seraient à la source des difficultés de prononciation des sons nouveaux. Nous prenons un exemple dans la description initiale du PAM dans Best (1995) :

[...] Because the universal phonetic domain and native phonological space are defined by the spatial layout of the vocal tract and the dynamic characteristics of articulatory gestures, those distal properties provide the dimension within which similarity is judged. For a native listener of a language that has no dental stop but does have bilabial, alveolar, and velar stops, the tongue tip constriction of the dental stop is straightforwardly closer in native phonological space to the alveolar place than to the others (p. 193).

Le PAM propose que l'auditeur assimile les sons en tant que sons linguistiques ou non linguistiques¹¹³. Les sons linguistiques sont ensuite assimilés dans les catégories L1 existantes ou comme des sons non catégorisés, qui formeront une nouvelle catégorie à leur tour (*cf.* sous-section 4.4.3). Les deux types d'assimilations correspondent respectivement aux classifications du SLM, à savoir les sons identiques, similaires et nouveaux. Le PAM introduit également l'idée de bons exemples, non idéaux et déviants de catégories natives. Ici, le son L2 est assimilé à une catégorie native, mais n'est pas nécessairement entendu comme le bon exemple de cette catégorie. Contrairement au SLM, cette approche facilite considérablement l'explication de l'exemple de l'assimilation du [θ] anglais par le /s/ ou /f/ français. L'accent est mis sur la perception des contrastes de L2 par les apprenants L2, plutôt que sur la perception d'un seul son L2. Best a proposé différents types d'assimilation pour tenir compte de la façon dont les composantes d'une

¹¹³ Best et collaborateurs (1988) ont trouvé que les auditeurs américains de leur étude ont relativement su discriminer les clics du zoulou mais ils ne les percevaient toutefois pas comme un son linguistique. Les clics sont des consonnes particulières de la langue zoulou parlée en Afrique du Sud. Pour en savoir plus consulter Ladefoged (1996) ou consulter des exemples disponibles sur : <http://www.phonetics.ucla.edu/course/chapter6/zulu/zulu.html> (consulté le 06/11/2018).

paire de contrastes L2 sont assimilés. Ces sons sont alors considérés comme appartenant à deux catégories distinctes, à des exemplaires non idéaux d'une catégorie L1 existante, à des exemplaires plus idéaux de la même catégorie, à de nouveaux sons (non catégorisés), etc. Les catégories L2 distinctes qui ne sont pas assimilées de façon perceptible par une catégorie L1 seront bien discriminées, même en l'absence d'expérience préalable. Cependant, il semble que la plupart des sons de parole L2 soit assimilés perceptiblement par une catégorie L1, du moins au début. Dans ce cas, les sons de parole de la L2 seront discriminés avec plus de précision s'ils sont assimilés par deux sons de parole L1 distincts que s'ils sont assimilés par une seule catégorie de sons de parole L1. Le PAM prédit que la précision de la discrimination peut également être influencée par le degré de similitude acoustico-articulatoire des sons de parole L2, avec les sons de parole L1. Plus précisément, le PAM prédit qu'une paire de sons de parole L2 perçus différemment pour une catégorie L1 sera mieux discriminée qu'une paire de sons de parole L2 ayant la même qualité d'adéquation qu'un son de la catégorie L1. Par exemple, les locuteurs japonais sont susceptibles d'assimiler l'anglais /r/ et /l/ comme de mauvais exemples d'un seul phonème japonais (/r/ ou peut-être /w), et de mal discriminer le contraste /r/-/l/ (Best, 1995). Kartushina & Frauenfelder (2013, p. 2118) résumant les trois types d'assimilation en ces termes :

Three patterns of L2 contrast assimilation are distinguished: Two-Category (TC) assimilation, the two L2 sounds map onto two different L1 categories; Category- Goodness (CG) assimilation, both L2 sounds map onto one L1 category, but one L2 sound is a better exemplar of this L1 category than the other; and Single-Category (SC) assimilation, both L2 sounds are equally good exemplars. Because in SC and CG assimilations both L2 sounds assimilate to one native category, these L2 contrasts are claimed to be more difficult to acquire than the L2 contrasts that show the TC assimilation pattern.

4.6.4 Native Language Magnet

Le principal objectif du modèle du *Native Language Magnet* (NLM) mis au point par Patricia Kuhl et ses collègues (Davis & Kuhl, 1994 ; Kuhl, 2010 ; Kuhl *et al.*, 2008 ; Kuhl & Meltzoff, 1996) est de rendre compte de la transition du traitement perceptif auditif au traitement perceptif spécifique au langage. Ayant fait l'objet de modifications au cours des années, le modèle est passé de son appellation originelle *Perceptual Magnet Effect* –PME (Davis & Kuhl, 1994) à sa version révisée, le Native Language Magnet theory Expanded (NLM-e - Kuhl *et al.*, 2008). Du NLM avec ses trois phases de développement au NLM-e « révisé » à cinq étapes clés, Kuhl *et al.*, 2008, proposent :

[...] *A model of early speech perception termed the NLM model, which focused on infants' native phonetic categories and how they could be structured through ambient language experience (p. 982). [...] NLM-e predicts an association between infants' early perception of native language phonetic units and later language development, an association that differs for native and non-native perception. [...] NLM-e predicts strong linkages between the perceptual representations formed through experience with language and vocal imitation [...] A distinction can be drawn between NLM-e and motor theories, and also between NLM-e and the hypothesized 'mirror neurone' system, a neural system that reacts to actions produced by others as identical to the same actions produced by oneself (p. 985).*

La perception des sons fonctionnerait comme un aimant qui attire tous les exemplaires sans distinction. Cet effet magnétique, qui empêche les exemplaires de se classer en catégories rendrait l'opposition phonético-phonologique difficile :

NLM assumes that speech perception involves general auditory mechanisms that process acoustic rather than specifically phonetic information. In NLM, native prototypes have magnetlike effects, in which the nearby perceptual space is "shrunk," making it more difficult to discriminate phonetic variation around prototypes than around non-prototypes, or poor exemplars, of the same category (Best, McRoberts, & Goodell, 2001, p. 3).

Les auteurs proposent que les premières expériences conditionnent la perception des propriétés acoustiques des sons de la parole. Selon eux, les nourrissons trient perceptiblement les unités de la taille d'un segment en catégories en fonction de la récurrence (fréquence d'exemplaires) des caractéristiques qu'ils ont détectées en entrée vocale. Il en résulterait une mise en correspondance spécifique à la langue, entre les catégories développées pour les sons de parole L1 et l'entrée phonétique qui anime cet aspect crucial de l'acquisition du langage. Kuhl (2010) avance que la cartographie¹¹⁴ perceptive des sons ambiants de la parole crée un « réseau complexe », ou filtre, à travers lequel le langage est perçu. Nous pouvons tout à fait établir le lien avec la notion du crible phonologique de Troubetzkoy (1939). En effet, les résultats traditionnels sur la perception de la parole chez les non-natifs suggèrent que les adultes ont des difficultés à distinguer les oppositions des segments qui ne sont pas employés phonologiquement dans leur propre langue, alors que les enfants y arrivent. La perception que les adultes ont des contrastes de la parole serait fortement influencée par leur expérience du système phonologique de leur langue maternelle (Best, McRoberts & Goodell, 2001). Une explication traditionnelle de ce phénomène a été une version perceptive du concept de « filtre phonologique » (Troubetzkoy, 1939).

¹¹⁴ Mappage ou *mapping* : ensemble d'images associées à chaque son pour former des catégories (voir l'AIJ de Neufeld).

Le NLM précise que l'harmonisation perceptive avec les catégories L₁ peut plus tard façonner la perception des sons de parole L₂. Des effets d'interférence peuvent survenir en raison de la difficulté inhérente à la séparation fonctionnelle des différentes catégories. Un engagement neuronal envers les mappages des catégories L₁ influencera plus tard le traitement des sons vocaux L₂. Il est important de noter que le NLM propose que les contraintes sur la perception des sons de parole L₂ proviennent d'expériences antérieures et non d'une perte de plasticité, résultant d'une maturation neuronale normale. Cependant, selon la NLM, l'apprentissage perceptif des adultes n'est pas impossible. Kuhl *et al.* (2008) a suggéré que l'influence de l'expérience antérieure peut être minime pour les enfants qui apprennent deux langues simultanément dans leur petite enfance, du moins si « deux cartographies différentes¹¹⁵ » sont acquises pour les sons de parole L₁ et L₂. La meilleure façon pour les apprenants adultes d'une L₂ de contourner les effets de l'interférence L₁ peut être de récapituler l'expérience de la parole L₁, c'est-à-dire de recevoir des signaux acoustiques exagérés, de multiples instances (exemplaires) par de nombreux orateurs et une expérience d'écoute en masse.

Nous arrivons à la fin de ce chapitre. Nous venons de présenter quatre modèles qui expliquent et prédisent l'acquisition de la prononciation en général. Nous retenons que bien qu'ils diffèrent sur certains points, leurs hypothèses de fonctionnement vont dans le même sens, à savoir l'influence de la L₁ sur la L₂ et le phénomène de l'imitation vocale qui tous deux sous-tendent l'acquisition de la matière phonique.

4.7 Conclusion

L'imitation est un facteur primordial du développement cognitif et social chez l'enfant et chez l'adulte. Facilitée par des capacités perceptivo-motrices, elle permettrait de comprendre, de communiquer et d'interagir avec l'environnement dans lequel nous évoluons. Ainsi, nous avons traité du rôle que jouent les neurones miroirs dans l'apprentissage par imitation et dans la compréhension des gestes d'autrui. Il s'agit d'un mécanisme qui nous permet non seulement d'anticiper nos actions, mais aussi d'assurer la communication/interaction avec les agents sociaux de notre entourage. La découverte de ces neurones spéciaux joue un rôle prépondérant dans la

¹¹⁵ Nous recommandons une vidéo de Patricia Kuhl sur Ted (The linguistic genius of babies) : www.ted.com/talks/patricia_kuhl_the_linguistic_genius_of_babies/discussion?quote=1150 (consulté le 01/11/2018).

compréhension du développement de la parole et du langage ainsi que dans l'acquisition de la L2. Dans ce chapitre, la distinction que nous avons établie entre l'imitation générale et l'imitation vocale montre que ce phénomène couvre des domaines de comportements humains divers. La notion d'imitation vocale, qu'elle soit volontaire ou non, serait gouvernée par des mécanismes qui sous-tendent les aspects phonétiques et phonologiques de la perception-production de la parole. D'ailleurs, nous avons noté que des deux types, l'imitation involontaire est la plus durable (ancrage dans la mémoire à long terme) et que l'imitation volontaire peut, à terme, devenir elle aussi un réflexe, et donc involontaire. De plus, nous avons montré un aspect central de notre travail de thèse qui consiste en la mise en évidence de la convergence vocale médiée. Nous considérons ainsi que les apprenants qui sont exposés aux *inputs* audiovisuels imitent (in)volontairement des aspects phonético-phonologiques. Aussi, les divers modèles théoriques notoires de catégorisation (prototypes, exemplaires) et d'acquisition phonétique et phonologique (AII, SLM, PAM et NLM) viennent éclairer notre conception et compréhension de l'acquisition de la matière phonique en L1 et L2. Ces théories, nous l'avons vu, utilisent des approches différentes, mais sont complémentaires et leur validité est sujette à des facteurs tels que le type d'expérimentation, la méthodologie employée, les sujets étudiés, etc. Nous soutenons le modèle de l'*Acoustic Image Imprinting* concernant le concept d'association d'images acoustiques et le fait que les adultes peuvent aussi atteindre un niveau de langue proche des natifs. *A contrario*, l'idée de l'enseignant natif reste discutable et nous ne la partageons pas. À propos du *Speech Learning Model*, le fait qu'il existerait trois catégories de sons (similaires, identiques et nouveaux) lorsqu'on apprend une L2 nous semble recevable, mais les postulats selon lesquels les sons similaires poseraient plus de problèmes aux apprenants et qu'une perception auditive correcte garantirait une prononciation adéquate ne sont pas tout à fait justifiés (nous avons tenté de le montrer par des exemples). Ensuite, comme le propose le *Perceptual Assimilation Model*, nous partageons l'idée de l'importance des facteurs articulatoires dans l'acquisition de la prononciation des sons nouveaux. Enfin, le *Native Language Magnet* nous semble le modèle le plus approprié car il intègre diverses notions clés que sont l'imitation, la perception, l'articulation, la fréquence et l'exemplaire.

Savoir si les participants de notre étude imitent vocalement et savoir quelles sont les formes que peuvent prendre ces imitations nous permettra de comprendre leurs diverses stratégies et ainsi, d'approfondir les interprétations de nos résultats. L'objectif du prochain chapitre sera

d'analyser les soubassements théoriques qui expliquent comment des interlocuteurs arrivent à se comprendre face aux multiples facteurs liés à la production de la parole. Nous aborderons les concepts de la viabilité et de la perturbation au regard de la variabilité et de la complexité inhérentes à l'acquisition de la prononciation en L2.

5 Viabilité et perturbation en production-perception de la parole

Ainsi, la variabilité omniprésente en production-perception de la parole, largement responsable de l'absence apparente de pertinence de certains comportements phonétiques, s'estompe lorsqu'elle est replacée dans la perspective de la viabilité des patterns, quels que soient leurs aspects. (Sock, 1998, p. 169).

Ce chapitre sera consacré à la théorie de la viabilité en phonétique (initiée par Sock, 1998, 2001) et la notion de perturbation en production et perception de la parole (Vaxelaire, 1993, 2007 ; Sock & Vaxelaire, 2004, 2010). D'autres théories et concepts connexes, permettant de mieux saisir le noyau de la théorie de viabilité et la perturbation, seront également abordés. Il s'agira des théories Quantique (Stevens, 1985, 2002 ; Stevens et Blumstein, 1978) et de la Variabilité Adaptative (Lindblom, 1983, 1987, 1990). Parallèlement, nous définirons les termes importants utilisés par ces différents auteurs et que nous reprenons dans cette section, en vue d'une conceptualisation théorique et tâchant de montrer les liens que ces termes tissent entre eux. Nous avons résumé, repris et paraphrasé des parties d'études, afin d'assurer, autant que possible, l'exactitude et la fidélité des propos des auteurs. Afin qu'elles correspondent au mieux à notre recherche, les propos des auteurs sont parfois reformulés dans des exemples plus larges, mais l'idée directrice reste identique, nous l'espérons. Nous y ajoutons des analogies ou des métaphores, qui même si elles s'écartent un peu des propos des auteurs, permettent de mieux comprendre notre problématique de recherche. À la fin de ce chapitre, nous verrons les implications de la théorie de la viabilité et de la notion de perturbation dans le champ de la didactique de langues, et nous poserons les prémisses de son application à l'enseignement/apprentissage des langues.

5.1 Quelques définitions : fiabilité, variabilité, invariance et viabilité

Avant d'aborder la théorie de la viabilité¹¹⁶, et pour une meilleure compréhension de la suite de ce chapitre en particulier et de cette thèse en général, nous définissons quelques termes importants que sont : variabilité, invariance, fiabilité et viabilité. Ces mots couvrent des aspects certes différents mais nous verrons très vite qu'ils sont très voisins, de par les mécanismes qu'ils désignent. Les définitions que nous fournissons ci-après viennent des dictionnaires en ligne¹¹⁷. Nous citerons également des exemples qui correspondent au domaine de la production et de la perception de la parole.

Parmi les définitions du mot variabilité trouvées sur le site web du CNRTL¹¹⁸, nous en gardons deux qui semblent correspondre à notre sujet. La première est : « Aptitude d'un être vivant ou d'un ensemble d'êtres vivants à subir des modifications dans sa forme ou dans ses fonctions, sous l'influence de facteurs externes ou internes » et la seconde est : « Ensemble des différences que peuvent présenter les membres d'une même population ». Nous retenons les termes « modifications » et « différences » pour caractériser la variabilité, qui est assujettie à des changements au fil du temps. Une réalisation phonétique variable est celle qui subit des modifications articulatoires pouvant entraîner des déviations quant aux sons « cibles » (nous y reviendrons plus loin). À l'évidence, l'invariance est l'état d'un objet qui est constant, fixe, identique et stable dans le temps. À cet effet, une réalisation phonétique dont les valeurs ne varient pas, constitue une invariante phonétique. La notion de fiabilité est souvent définie comme le caractère de ce qui est fiable, c'est-à-dire à qui l'on peut s'y fier, faire confiance. Ainsi, en production de la parole, il y a fiabilité lorsqu'une production est celle qui admet le moins de variations possibles sur un ensemble de répétitions. À *contrario*, une production peut être considérée comme non fiable lorsque, sur plusieurs répétitions, les mesures obtenues montrent des variabilités erratiques trop nombreuses. Par conséquent, il serait difficile de tirer des

¹¹⁶ Remarquons que viabilité et fiabilité forment une paire minimale sur le plan de la prononciation : le /f/ et le /v/ sont ici en opposition phonologique.

¹¹⁷ <http://www.cnrtl.fr/definition>, <https://www.larousse.fr/dictionnaires/francais>

¹¹⁸ Centre National de Ressources Textuelles et Lexicales. <https://www.cnrtl.fr/definition/variabilit%C3%A9> consulté le 18/07/2019.

conclusions concernant les productions non-fiables. Admettons qu'un locuteur réalise la voyelle /i/ sur 10 répétitions et que les valeurs formantiques¹¹⁹ présentent une forte variabilité (écart-type élevé), avoisinant les valeurs d'autres voyelles (ex. /e/), la production de cette voyelle serait qualifiée de non fiable. Rappelons que, même si la fiabilité implique des mouvements articulatoires plus ou moins précis, les consonnes, notamment les occlusives sourdes /p t k/, permettent une délimitation plus facile et sont donc plus fiables (Sock, 1998).

Le dictionnaire du CNRTL¹²⁰ définit le mot viabilité comme étant le : « Caractère de ce qui est apte à vivre, à se développer » alors que le e-Larousse¹²¹ le définit comme : « Qui est susceptible de durer, de continuer d'exister ». De ces deux définitions, nous pouvons retenir que la viabilité désigne la capacité d'un élément qui présente des conditions nécessaires pour rester en vie, de continuer d'exister et qui peut être mené à son terme, qui peut aboutir. Dans notre cas, il peut s'agir par exemple de toute production phonétique qui résiste aux conditions tant internes qu'externes de perturbations (par exemple variations de la vitesse d'élocution, voix chuchotée, *bite blocks*, reconstruction linguale après glossectomie, etc.). Nous apporterons des éléments d'explication complémentaires à ces notions au fur et à mesure que nous progressons dans ce chapitre.

Ayant défini les termes qui nous intéressent, il nous semble fondamental d'évoquer brièvement deux autres théories existantes, qui auraient certains liens avec la théorie de la viabilité repose, afin de consolider ses fondements et ses principes. Ces deux théories sont : la théorie Quantique et la théorie de la Variabilité Adaptative. Nous les présentons avant d'aborder celle qui nous intéresse. La première théorie permet de comprendre la seconde et enfin, les deux permettent à leur tour de mieux appréhender la théorie de la viabilité. Nous commençons par la présentation de la théorie Quantique.

119 Concernant la fiabilité ou exactitude des mesures acoustiques (formants, Fo, Intensité, les durées), Sock précise que « la notion de formants comme paramètres perceptifs est controversée [...] et ce particulièrement à cause de la difficulté bien connue de leur extraction fiable » (Sock, 1998, p. 161).

120 <https://www.cnrtl.fr/definition/viabilit%C3%A9> consulté le 18/07/2019.

121 <https://www.larousse.fr/dictionnaires/francais/viable/81784> consulté le 18/07/2019.

5.2 La théorie Quantique : une théorie de l'invariance

La théorie Quantique est mise au point par Stevens (1985 ; 2002) et Blumstein (1989), et stipule l'existence d'invariances dans les structures acoustiques de la parole. « Stevens et Blumstein (1978) tentent de montrer l'existence de propriétés acoustiques invariantes permettant la caractérisation du lieu d'articulation de la consonne, quel que soit le contexte vocalique » (Sock, 1998, p. 135). Cette théorie de l'invariance pose un cadre phonétique dans lequel des traits de segments phonologiques qui structurent l'acquisition du lexique, seraient soumis à une organisation hiérarchique (*op. cit.*). Le signal produit par un locuteur serait fortement structuré et contiendrait des patterns immuables (modèles à reproduire) qu'un locuteur devrait apprendre, indépendamment des contextes phonétiques et des langues (*ibid.*, p. 136). D'après la théorie Quantique, le passage du niveau articulatoire au niveau acoustique s'effectuerait d'une façon non-linéaire. Cette non-linéarité créerait trois types de zones dans la correspondance entre l'espace articulatoire et l'espace acoustique. Dans la première et la troisième zone (I et III, *cf.* Figure 1Figure 12), une phase de stabilité est observée. En revanche, la deuxième zone (II) serait plutôt assujettie à une instabilité, puisqu'un changement drastique de paramètres acoustiques interviendrait à la suite d'une minime variation sur le plan articulatoire. La figure ci-dessous illustre l'idée de zones de variabilité et d'invariance soutenue par Stevens.

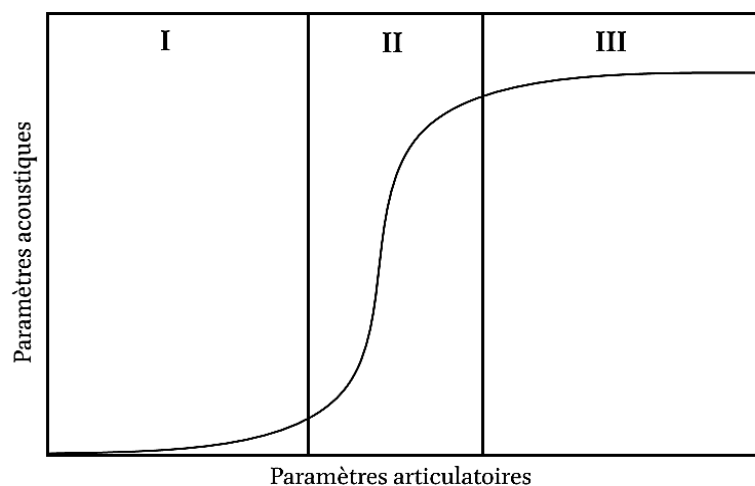


Figure 12 : zones de stabilité et d'instabilité selon la Théorie Quantique. La courbe représente la non-linéarité du passage de paramètres articulatoires aux paramètres acoustiques (Stevens, 1989 *in* Sock, 1998, p. 137)

Bien que la théorie Quantique¹²² suppose une invariance articulatoire dans la production de la parole, ses auteurs n'ont pas réussi à tenir une position assez ferme quant au lien qui existerait entre la théorie Quantique et la théorie de l'invariance. Dans le même ordre d'idées, Ladefoged (1983) ne partage pas les fondements de cette théorie « étant donné que [ces auteurs] posent que les similitudes entre les langues du monde soient le fait de facilités articulatoires » et que l'existence des voyelles /i a u/ dans la majorité des langues est due aux productions extrêmes du tractus vocal (cité par Fauth, 2012, p. 21). Il manquerait à la modélisation de la théorie Quantique une prise en considération de la variabilité en tant que fait incontournable et inhérent à la production de la parole. Sock constate que « le lien entre une théorie de l'invariance acoustique et la Théorie Quantique n'est pas clairement défini chez Stevens [...] qui souligne ce fait en disant qu'il n'y avait rien d'intrinsèque à la Théorie » (Sock, 1998, p. 136). La théorie Quantique de Stevens et Blumstein (1978) serait une théorie phonétique qui prône l'invariance articulatoire, c'est-à-dire ne prenant pas en considération la variabilité dans la production de la parole. Partant ainsi des insuffisances et incohérences de ce dernier concept et en admettant les variations inhérentes à la parole, Lindblom (1983) propose un cadre conceptuel nommé théorie de la Variabilité Adaptative, que nous verrons dans ce qui suit.

¹²² Tout comme la théorie motrice de Liberman que nous nous gardons de développer dans ce travail ; celle-ci est assez éloignée de notre cadre théorique. Pour plus de détails, voir Liberman et Mattingly, 1985, 1989.

5.3 La théorie de la Variabilité Adaptative

Faisons tout d'abord un point sur la notion de la variabilité en phonétique, car nous verrons qu'elle est capitale pour notre problématique et nous l'utiliserons tout au long de ce travail. Rappelons que la variabilité est avant toute chose une affaire de différence individuelle à laquelle chaque personne est sujette, comme le précise Jusczyk (2000) : « [...] *the acoustic characteristics of the same word produced by different talkers may vary considerably. They may actually closely resemble the acoustic characteristics of a different word (as opposed to the same word) produced by another talker* » (p.7 cité dans Lapteva, 2011, p. 15). La variabilité est autant inter-individuelle qu'intra-individuelle. La première signifie que le même énoncé oral produit par différents locuteurs a des représentations acoustiques différentes. La seconde fait référence aux variations entre les énoncés oraux produits par un même locuteur. En règle générale, il est impossible de trouver la copie exacte du même mot ou de la même phrase produite par la même personne : il est bien connu que le même mot produit deux fois par la même personne présentera des différences acoustiques plus ou moins remarquables (Lapteva, 2011). La variabilité concerne la façon dont les objets de la parole, soient les phonèmes et les sons (au niveau cognitif ou physique) varient dans leur comportement ou dans leur rendu. En phonétique expérimentale, les courants théoriques en production de la parole modélisent la variabilité en tant que processus qui relève de plusieurs niveaux. Au niveau phonétique, les objets allophoniques extrinsèques planifiés varient systématiquement en fonction des processus physiques issus de la mécanique de l'appareil vocal ou de l'aérodynamique du processus de production (Tatham & Morton, 2011). Il existe aussi des variations complémentaires au niveau phonétique en raison du fait que les processus de production ne sont pas parfaits, c'est-à-dire qu'une certaine quantité de comportements aléatoires ou inattendus est intrinsèque aux processus aérodynamiques, mécaniques et acoustiques. En effet, Fauth (2012) rappelle que l'importance de la variation articulatoire et ses conséquences acoustiques ne sont pas corrélés. Elle donne l'exemple de la voyelle /a/ qui admet une variabilité articulatoire plus importante que les autres voyelles (ayant donc une zone de viabilité plus large), car sa réalisation est capable de supporter beaucoup plus de contraintes réalisationnelles, afin d'aboutir à un résultat acoustique stable et fiable (*ibid.*, p. 22).

Tableau 13 : sources de variabilité en phonétique (adapté de Tatham & Morton, 2011, p. 10).

Variabilité	Source ¹²³
Phonologique	Volontaire (ex. l'alternance entre le /l/ clair ou palatalisé ([l], [ɫ]) et le /l/ sombre ou velarisé ([ɫ̥], [ɫ̥̥]) - souvent spécifique à une langue.
Phonétique	Déterminé par la coarticulation (ex. nasalisation des voyelles entre des consonnes nasales en anglais - /mæn/ -> [mæ̃n]) - fondamentalement universel au langage-
Physique	Aléatoire - non systématique dans tous les sons de la parole.

Nous venons de voir quelques caractéristiques de la variabilité inhérentes à la production de la parole, ainsi que les sources éventuelles de cette variabilité. Penchons-nous à présent sur la théorie de la Variabilité Adaptative dans le domaine de la production-perception de la parole.

À travers sa théorie de la Variabilité Adaptative, Lindblom (1983, 1987, 1990) rejette l'hypothèse, émise par les défenseurs de la théorie Quantique et de la théorie motrice, selon laquelle le lien entre les paramètres articulatoires et acoustiques serait gouverné par la présence d'une certaine invariance (Stevens et Blumstein, 1978). La notion d'invariance serait en outre applicable à la compréhension générale de l'auditeur et non au niveau de production de la parole. Ainsi peut-on souligner que les théories de l'invariance pourraient éventuellement se rationaliser dans le domaine de la perception globale uniquement (Lindblom, 1987). La théorie de la Variabilité Adaptative part du postulat selon lequel il existerait deux « forces » rivales qui définissent le fonctionnement du système moteur de production de la parole. Le besoin de distinctivité perceptive de la part de l'auditeur et la demande du locuteur d'économie articulatoire serait le phénomène clef dans toute communication linguistique :

C'est le concours de ces deux forces qui serait le responsable principal des phénomènes de variabilité dans les données obtenues en production de la parole. Ces phénomènes de variabilité peuvent être représentés dans la dichotomie d'hyper- et d'hypo-articulation ; ils correspondent à ces possibilités de compensation dans l'interaction entre le locuteur et l'auditeur (Sock, 1998, p. 137).

Dans le même ordre d'idée, Tatham & Morton (2016) confirment la notion des deux forces et avancent que la parole est conçue pour communiquer efficacement et elle est produite par un

123 On pense qu'une certaine variabilité est associée aux émotions dans l'expression orale, mais cette variabilité n'a pas encore été entièrement modélisée. Certains chercheurs tentent actuellement de constituer des bases de données de variantes dans le signal acoustique (Brulard, Carr, Durand, & Navarro, 2015).

dispositif biologique qui fonctionne selon le principe de l'économie d'effort (2016, p. 222). La Variabilité Adaptative concerne l'hypothèse selon laquelle les locuteurs seraient capables d'ajuster la façon dont ils produisent la parole en fonction des exigences de la situation communicative, ce qui inclut le fonctionnement du système de perception de la parole. Selon la perspective fondatrice de la théorie de Lindblom : « *Adaptive variability is all about the hypothesis that speakers are able to adjust how they produce speech according to the requirements of the communicative situation, which includes how perception works* » (Tatham & Morton, 2016, p. 222). En effet, lors d'une interaction orale, le locuteur va « adapter » la « variabilité » de son système de production de la parole afin de se faire comprendre au mieux par son interlocuteur, mais tout en évitant de déployer trop d'énergie, lors des mouvements articulatoires. Afin d'aboutir à ce compromis, le locuteur va mettre en place un mécanisme d'équilibre lui permettant d'atteindre son double objectif : optimiser l'énergie articulatoire et de garantir une compréhension maximale. Cette situation idéale que nous pouvons qualifier, par métaphore, de « point de patinage » (*Ideal cases* – Lindblom 1987) est représentée par la diagonale sur la figure suivante.

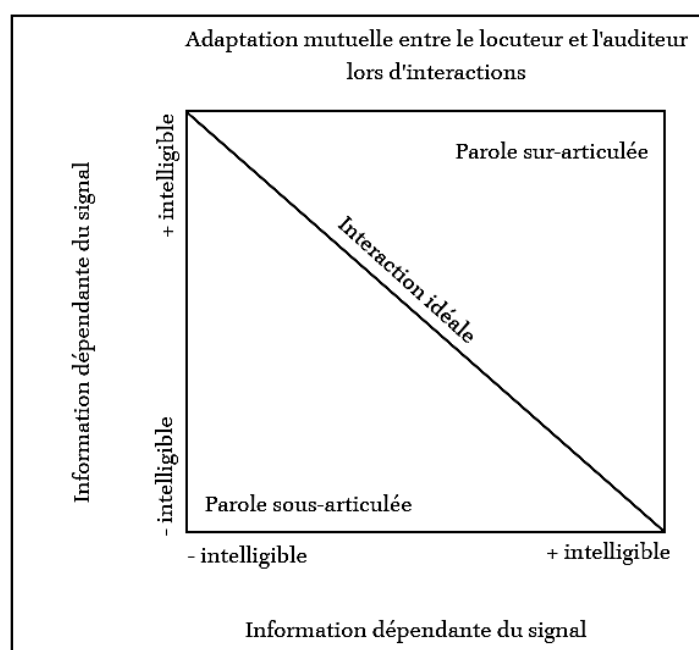


Figure 13 : le cas idéal selon la théorie de la variabilité, entre hypo-/ hyper-articulation et informations contextuelle / informations du signal (Lindblom 1987 cité par Sock, 1998, p. 138).

Sur cette figure, les interlocuteurs s'engageant dans une interaction vont mettre en place des stratégies de communication réciproque afin de leur permettre de mieux se comprendre. Par ces

mécanismes, ils vont alors chercher à tendre vers une situation idéale, en ajustant les niveaux de production et de perception de la parole. Parallèlement à l'idée d'équilibre entre l'hypo et hyper articulation, Tatham & Morton (2016) ayant appliqué la théorie de la Variabilité Adaptative au phénomène de la coarticulation en production de la parole, proposent également la notion de point de patinage, qui serait « un juste milieu » entre des valeurs maximales et des valeurs minimales. La coarticulation, dans la lignée de la théorie de la Variabilité Adaptative, entraîne potentiellement une réduction du contraste phonologique perçu (Fougeron, D'Alessandro & Lancia., 2018). La diminution est scalaire¹²⁴ en fonction de l'endroit où se situe l'événement sur le continuum hyper-/hypo-phonique. Ce continuum caractérise l'intervalle entre la coarticulation résistante et la coarticulation intrinsèque. Dans la terminologie de Tatham & Morton, ce continuum va de l'amélioration maximale jusqu'à la contrainte maximale, en se fixant sur la valeur optimale de l'énoncé donné, une fois que les contraintes phonologiques ont été évaluées. Pour la stricte délimitation des processus phonologiques et phonétiques, la diminution du contraste est essentiellement une altération de la « valeur de déclenchement », de telle sorte qu'il y ait une diminution apparente du contraste phonologique. Cela signifie en fait une différence entre l'énoncé planifié et l'énoncé rendu (Tatham & Morton, 2016, p. 222).

La théorie de la Variabilité Adaptative (Lindblom, 1987) concerne les stratégies d'ajustement entre l'hyper et l'hypo articulation mises en œuvre par des interlocuteurs afin de se rapprocher d'une compréhension idéale et réciproque, tout en produisant le moindre effort possible. Elle est importante pour notre travail, car non seulement elle rend compte de la gestion de la compréhension mutuelle entre des communicants mais également nous aide à mieux conceptualiser la théorie de la viabilité que nous développons dans la section suivante.

5.4 La théorie de la viabilité en production-perception de la parole

Nous en venons au point central de cette section qui est la théorie de la viabilité en production-perception de la parole. Comme nous l'avons évoqué plus haut, Sock (1998) initia cette théorie en phonétique, afin de rendre compte de l'intelligibilité du discours humain, notamment en

¹²⁴ Suivant une échelle (0, 1, 2, 3 ...).

phonétique et en phonologie. Toutes les voix, trajectoires et structures complexes inhérentes à la parole, seraient gérées par la notion de viabilité, afin de garantir (ou non) la compréhension entre interlocuteurs. Ces possibilités font que les membres d'une communauté définie se comprennent malgré les différences inter et intra-individuelles de prononciation. Nous avons tous des manières bien différentes d'articuler, pourtant nous nous faisons comprendre. L'idée fondatrice est « d'expliquer l'évolution d'un système, modulée par des dynamiques non déterministes spécifiques, ainsi que par des contraintes de viabilité, afin de mettre au jour l'interaction des diverses composantes auto-organisées du système sous-jacent à sa régulation » (Sock, 2001, p. 285). Partant du concept de viabilité en mathématique¹²⁵ (voir Aubin et Cellina, 2012 ; Aubin, Bayen, & Saint-Pierre, 2011), les fondements de la théorie de la viabilité en production-perception de la parole s'alignent sur celle de la Variabilité Adaptative, et donc vont à l'encontre des théories Quantique et motrice (voir *supra*). Sock soutient le principe du *satisficing*¹²⁶ et opte pour « l'adoption d'une solution sous-optimale mais toutefois acceptable, au sens où elle se situe au-dessus d'un seuil critique de satisfaction » (p.167). Considérant qu'un agent régulateur veille sur le bon fonctionnement d'un système et qu'un élément contrôleur détermine comment le régulateur doit fonctionner, Sock (1998) argumente que le contrôle serait dans ce cas un facteur extrinsèque au système alors que toute régulation devrait être un facteur intrinsèque au système de production de la parole. Il discute ensuite le fonctionnement du système dans le cadre de ces deux notions : celles de la prescription¹²⁷ et celle de la proscription¹²⁸, en proposant « d'adopter un changement d'orientation et de passer d'une logique prescriptive à une logique proscriptive » (Sock, 1998, p. 166). Pour lui, « on ne considèrera plus que tout ce qui n'est pas permis est interdit mais plutôt que tout ce qui n'est pas interdit est permis » (*ibid.*). Autrement dit, « les trajectoires ou les comportements que l'on observe dans les systèmes dynamiques ne sont pas des processus dictés

125 La théorie de la viabilité conçoit et développe des méthodes mathématiques et algorithmiques pour étudier l'adaptation aux contraintes de viabilité des évolutions régies par des systèmes complexes sous l'incertitude que l'on retrouve dans de nombreux domaines impliquant des êtres vivants, de l'évolution biologique à l'économie, des sciences de l'environnement aux marchés financiers, de la théorie du contrôle et de la robotique aux sciences cognitives. Il s'agit de recherches interdisciplinaires couvrant des domaines qui se sont traditionnellement développés de manière isolée (Aubin, Bayen, & Saint-Pierre, 2011).

126 Le terme *satisficing*, est composé des mots *satisfying* et *sufficing* et désigne le fait de vouloir obtenir une solution qui soit à la fois satisfaisante et suffisante, sans trop chercher l'optimalité qui coûterait cher en termes de temps et d'énergie/ressource.

127 Contrôle et prescrit le comportement du système.

128 Régule et proscrie le fonctionnement du système.

ou prescrits » (Bouarourou, 2014, p. 103), mais plutôt des possibilités qui assurent la survie des cibles articulatoires, et cela pour le maintien des oppositions linguistiques. Tant que les productions sont viables, l'on peut se passer de l'optimalité et de la prescription.

La théorie tente de répondre à la question suivante : comment réussissons-nous à comprendre nos interlocuteurs, malgré les diversités contextuelles et la forte variabilité sous-jacentes aux productions linguistiques ? Il faut comprendre par ceci que du moment où une trajectoire¹²⁹ articulatoire, empruntée en vue de produire un son, n'est pas interdite, nous pouvons la suivre, à condition d'aboutir à la cible, ou plus précisément de se placer dans une zone de viabilité. Si l'idée d'une invariance stricte en production-perception est toutefois écartée, force est de remarquer qu'il existe certains domaines de régularités et de possibles articulatoires-acoustiques qui caractérisent les discours de tous les jours. Selon la théorie de la viabilité phonético-phonologique, la quête d'une trajectoire optimale, en vue d'atteindre une cible articulatoire « parfaite » importe peu. Ce qui compte le plus dans ce cadre conceptuel, c'est que quel que soit le degré de variabilité des réalisations du locuteur, les rendus acoustiques se placent dans une zone de perception viable, permettant ainsi à l'auditeur de comprendre le message véhiculé.

Les notions de trajectoire et de cible sont tellement centrales dans cette théorie que l'on peut se demander s'il est possible que différentes trajectoires mènent à une même cible. Ainsi Sock s'interroge à ce sujet : « Mise à part cette stratégie consistant à garder les phases relativement stables, existe-t-il d'autres trajectoires que les classes phonétiques auraient pu emprunter en restant toujours dans le cadre de la viabilité de l'opposition linguistique ? » (p. 145). L'émergence de catégories phonétiques acceptables n'est que la satisfaction de l'opposition des classes et du maintien de ces oppositions (Sock, 1998 cité in Bonnot, 2001, p. 283). De ce fait, Sock a montré que dans les langues à quantité à opposition simple (VC vs. VVC) et les langues à distribution complémentaire (VCC vs. VVC), la stabilité et variabilité en *timing* de la quantité (vocalique et consonantique) étaient liées à une question de contraintes linguistique et biomécanique¹³⁰. Il fait remarquer que « [...] du point de vue de la préservation du voisement - qu'il soit

129 Par trajectoire, nous entendons le chemin qu'empruntent les organes de la parole (mouvements articulatoires), afin d'atteindre un son que l'on désire prononcer (cible).

130 Fonctionnement inhérent aux organes de la parole, que nous ne pouvons intentionnellement contrôler. Par exemple, l'activité des plis vocaux ne s'arrêtent jamais brusquement, mais ont un délai d'arrêt progressif, c'est-à-dire l'amortissement.

phonologiquement pertinent ou phonétiquement réalisé- le choix d'augmenter la proportion de la phase vocalique dans le domaine VC¹³¹ ne correspond pas à une trajectoire viable » (p.145). Il conclut en ces termes : « Ces constats nous amènent à dire que maintenir les phasages relativement stables est une contrainte principalement linguistique : il s'agit de n'emprunter que les trajectoires viables pour l'opposition linguistique » (*op. cit.*). Sock avance que : « la variabilité omniprésente en production-perception de la parole, largement responsable de l'absence apparente de pertinence de certains comportements phonétiques, s'estompe lorsqu'elle est replacée dans la perspective de la viabilité des patterns, quels que soient leurs aspects » (Sock, 1998, p. 169).

Dans la lignée des travaux de Gaskell & Marslen-Wilson (1998) et Coenen *et al.* (2001) qui utilisent le terme de « viabilité contextuelle », Snoeren (2005), ayant travaillé sur le phénomène de l'assimilation liée aux variations phonologiques en production et perception de la parole, pense que la notion de la viabilité serait associée à la compatibilité avec le changement phonologique assimilatoire (p. 74). Il pose alors que les temps de réalisation de l'assimilation étaient plus limités pour les sons-cibles suivant les consonnes assimilées générées dans des contextes viables, c'est-à-dire compatibles avec le changement phonologique (p.77). Cette idée de compatibilité se retrouve également chez Bouarourou (2014) qui pense qu'« il s'agirait plutôt d'appréhender de tels comportements comme des processus visant à écarter ce qui n'est pas compatible avec la survie et la reproduction (ou la répétition) du système dynamique en question » (p. 103). Selon Snoeren, le contexte phonologique serait utilisé de façon régressive par l'auditeur afin d'évaluer les possibilités de changement phonologique dans les formes de surface (p. 90).

Nous retenons que la notion de la viabilité serait gouvernée par des phénomènes de compatibilité et d'adaptation des systèmes de production et ceux de la perception de la parole, en dépit de la présence de variabilité dans les productions des locuteurs. Ce cadre nous intéresse particulièrement car il permet d'appréhender les mécanismes qui sous-tendent la compréhension du discours entre l'émetteur et le récepteur d'un message linguistique à l'égard de ce qui est permis et acceptable et de ce qui relève de l'interdit. Nous pouvons donc admettre qu'il devrait exister des zones à l'intérieur desquelles des sons ou catégories de sons peuvent évoluer garantissant ainsi leurs viabilité, et des zones au-delà desquelles la compatibilité et l'adaptation ne seraient plus

¹³¹ Voyelle-Consonne.

effectives, mettant en danger la viabilité des productions linguistiques. Nous abordons ces espaces de viabilité dans la section suivante.

5.4.1 Zones de viabilité ou espaces de possibilités fiables

Une zone de viabilité est un espace « virtuel » délimité qui renferme des sons viables. La taille de cet espace est variable en fonction des sons qui sont produits et aussi en fonction des locuteurs. Celle-ci circonscrit des caractéristiques articulatoires et acoustiques de productions linguistiques. Par exemple, le triangle vocalique est une représentation spatiale des premiers et deuxièmes formants des voyelles et donne des indications sur leur articulation (postériorité vs. antériorité, aperture). Elle permet surtout de visualiser des ensembles de sons et leurs distributions les uns par rapport aux autres. Pour illustrer des zones de viabilité, nous nous repons sur une étude de Hillenbrand *et al.* (1995), sur la production et la représentation de 10 voyelles (/ æ, ɑ, ɔ, ʌ, ɜ, i, ɪ, u, o et ε/) de l'anglais américain, par 46 hommes, 48 femmes et 64 enfants. Cette étude n'a pas de lien direct avec la théorie de viabilité¹³², mais la variabilité observée nous permet de la retenir afin de montrer schématiquement la notion de zones de productions viables et non viables (figure 3). Sur cette graphique, nous pouvons remarquer trois configurations importantes. La première est la variabilité dans la production des voyelles. En effet, il n'existe pas un point précis de production dans cet espace¹³³ pour tous les locuteurs, mais toute une zone, dont la taille varie selon chaque voyelle. Ces zones, représentées par les ellipses (figure 3) sont obtenues en fonction des écart-types des moyennes des valeurs formantiques de chaque voyelle pour tous les locuteurs. La deuxième remarque est le chevauchement remarquable des voyelles ; certaines zones étant littéralement incluses dans d'autres. Ceci s'expliquerait par le fait qu'en anglais certaines voyelles possèdent des propriétés acoustiques proches dans l'espace vocalique (ex. /æ/, /ɑ/, /ε/ et /e/ ; /i/ et /ɪ/). Nous pouvons nous demander si l'auditeur réussira à faire la différence entre les éléments vocaliques sur le plan perceptif ? Il nous semble que tout dépendra du contexte et de l'existence d'oppositions linguistiques dans la langue en question. Enfin, la troisième remarque concerne les productions qui sont à l'extérieur des ellipses et qui se chevauchent également. À ce stade, nous considérons,

¹³² Du moins cela n'a pas été annoncé.

¹³³ Remarquons que cette espace vocalique ne correspond ni au trapèze ni au triangle classiques que l'on connaît et qui montrent les voyelles cardinales.

par analogie, que les ellipses représentent les zones de viabilité : les sons se situant à l'intérieur sont viables et ceux de l'extérieur sont non viables.

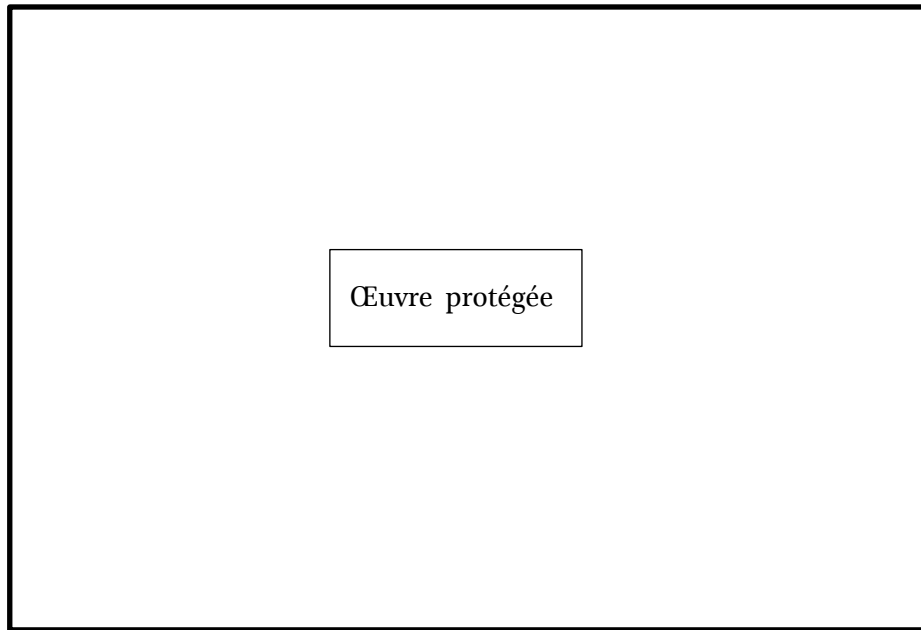


Figure 14 : espace illustrant la variabilité dans la production de 10 voyelles / æ, a, ɔ, ʌ, ɜ, i, I, u, o et ε / de l'anglais américain par 46 hommes, 48 femmes et 64 enfants (Hillenbrand *et al*, 1995, p. 3104).

Suite aux explications ci-dessus, nous avons adapté une illustration de Bechet (2011) afin de mieux conceptualiser le postulat de la théorie de la viabilité que nous résumons en ces mots : à une multitude de productions articulatoires correspondent une cible acoustique et à une multitude de signaux acoustiques correspondent une cible perceptive (Bechet, 2011). Nous pouvons donc conclure qu'à une multitude de productions articulatoires correspondent une seule cible perceptive¹³⁴. Ainsi, la description que donne la théorie de la viabilité à travers ce schéma, répond à la question précédemment posée, à savoir comment l'auditeur identifie chaque son malgré la variabilité des productions. Quelles que soient la nature des données acoustiques, tant que les différentes productions des locuteurs se placent dans une zone de viabilité (sons viables), l'auditeur identifiera la cible. Les sons situés en dehors de cette zone de viabilité engendreraient des quiproquos perceptifs, surtout lorsqu'ils traversent les frontières de la zone d'un autre son voisin qui partiellement permet d'assurer l'opposition linguistique d'un message spécifique.

¹³⁴ Nous rappelons également l'importante place de l'environnement.

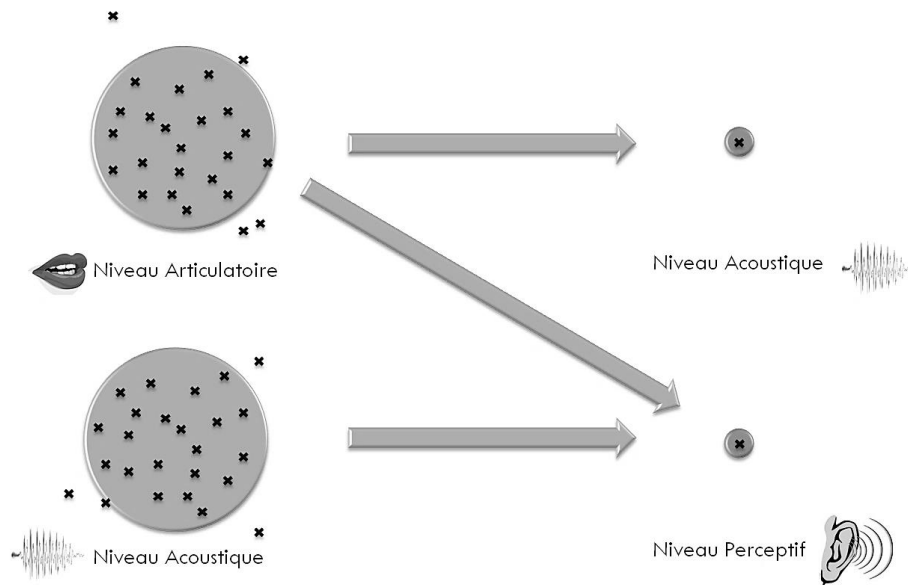


Figure 15 : représentation schématique de la théorie de la viabilité en production-perception de la parole de Sock (1998), adapté de Bechet (2011, p. 432).

Sur un plan plus général, et en prenant des exemples de phonèmes, la notion de viabilité permettrait de rationaliser le maintien d'une opposition phonologique. Ainsi, peut-on obtenir la plosive bilabiale sonore /b/ sans que les deux lèvres ne soient en contact ? C'est-à-dire, peut-on emprunter une trajectoire articuloire différente de celle de la (bi)labialité et de l'explosion, en vue d'obtenir ce son¹³⁵ ? Il nous semble qu'une tentative de production d'un /b/ sans la labialité ou l'explosion résulterait en un /v/ (approximante labiodentale voisée) ou /β/ (fricative bilabiale voisée) (Ladefoged & Maddieson, 1996). Deux cas de figures peuvent alors se présenter. D'une part, si dans une langue, les deux sons ne sont pas dans une relation oppositive (contraste phonologique) et que les locuteurs peuvent se comprendre, alors /v/ ou /β/ peuvent être considérés comme des sons viables, et donc des allophones de /b/. D'autre part, si dans une langue, ces deux sons forment une paire minimale, il y a opposition phonologique et dans ce cas /v/ ou /β/ seraient non viables par rapport à /b/. Il va de même pour les deux mots « fiabilité » et « viabilité » dont nous avons fait allusion plus haut : /v/ et /f/ forment deux traits distincts et permettent d'opposer les deux mots (ils se distinguent par le voisement). Si un locuteur réalise un /b/

¹³⁵ Ladefoged & Maddieson (1996) ont montré que l'on peut produire le son /b/ en créant l'occlusion grâce au contact entre la lèvre supérieure et les dents inférieures. Cet exemple montre que l'on peut effectivement produire certains sons avec deux voire plusieurs gestes articuloires.

(biabilité¹³⁶) à la place du /v/ (viabilité), il se pourrait qu'il ne soit pas compris¹³⁷. Il aurait non seulement utilisé une trajectoire différente, ce qui n'est pas grave en soi¹³⁸, mais aussi il aurait « raté » sa cible : sa production ne serait tout simplement pas viable. En revanche, si nous prenons les allophones de <r> (/r, ɾ, ʀ, ʁ, ʁ̥ et ʁ̥/), il y a une très forte chance que certains, voire la majorité soit des sons viables en français, malgré les différentes manières de réalisation. Le <r> du mot <rien> prononcé sous un de ces allophones sera toujours compris quelle que soit la variabilité dans la réalisation : ils font partie d'une même zone de viabilité en français. Dans ce cas précis, on notera que certains auditeurs français identifieront un petit accent lorsque le mot est réalisé par /ɾ/, /ʁ̥/, /r/ et peut être /ʁ̥/. Les autres ne poseront pas difficulté, car ils sont utilisés dans certains accents de l'hexagone. En anglais, l'emploi des sons /t/, /tʰ/, /ɾ/, /t̥/, /d/, /ʔ/ comme allophones pour le <t> intervocalique dans le mot *city*, sera toujours compris par des locuteurs des divers accents qui les utilisent, ainsi que par toute personne connaissant ses réalisations possibles. Par conséquent, tous ces sons, bien que articulés différemment peuvent être considérés comme viables dans ce contexte.

Un aspect de prononciation où la théorie de la viabilité porterait sens serait la production-perception de la parole chez le bébé et ses parents. En effet, il est admis que les adultes, particulièrement les parents peuvent reconnaître leur langue maternelle dans le babillage des nourrissons (jusqu'à six mois) et comprendre entièrement leur discours déjà à partir d'un an (Roulin, 2006, p. 106-107). Or, il semblerait aussi que même si le discours des enfants finit par s'aligner à celui des parents, les deux diffèrent radicalement au début de la production de leur parole. Ainsi, Smith (1973) a montré que les enfants, au cours de l'acquisition phonéto-phonologique, produiraient des sons différents des cibles des adultes. Il montre tout d'abord que le discours des enfants renfermerait une certaine régularité (feet, finger et fork sont réalisées [wi:t][wiŋə][wɔ:k]), des irrégularités (*lorry, light, like* sont prononcés [lɔli:][dait][gaik]) ainsi que des sons étrangers à l'inventaire phonémique des parents (*slug, Smith, neck* sont produits en [ɫg], [mis], [ŋɛk]). Ces sons et séries de sons produits au départ se concrétiseront (ressembleront aux sons des adultes) lorsque l'enfant grandit (Smith, 1973, p. 2-4). L'auteur souligne que non

136 Ce mot n'existe pas en français.

137 En l'absence bien sûr de tout contexte pouvant aider à la compréhension.

138 Selon les principes de la viabilité, plusieurs trajectoires peuvent mener à une cible, grâce à de la compensation et ou d'un réajustement, surtout suite à une perturbation.

seulement les parents comprennent parfaitement le discours des enfants, mais également la capacité des enfants à comprendre leurs propres discours au travers d'enregistrements. La question que nous nous posons est de savoir comment les parents arrivent-ils à identifier ces sons, pourtant mal réalisés, et à les comprendre en plus ? Peut-on présupposer que dans la perspective de la viabilité, ces sons sont viables ? Comment les bébés assurent-ils les oppositions phonético-phonologiques, en produisant des sons entièrement différents ? Nous serions enclins à affirmer qu'il y a une certaine viabilité dans leur production, puisqu'ils se font comprendre. Toutefois, il semblerait que ceci soit lié aux contextes connus, aux répétitions, aux imitations des sons et de l'intonation par l'enfant (Kail, 2015). L'habitude et le contexte de communication joueraient un rôle capital dans l'intercompréhension entre ces deux agents.

Nous venons de voir des cas particuliers où des trajectoires articulatoires non viables pouvaient générer un discours compréhensible, grâce au contexte de communication. Il serait toutefois intéressant de conduire des études de production-perception de la parole dans l'objectif de déterminer dans quelles mesures des trajectoires articulatoires engendreraient de l'incompréhension d'un énoncé, malgré la présence de facteurs contextuels. Ces éléments de réponses devraient permettre d'ajouter d'autres facteurs pertinents pour la modélisation de la théorie de viabilité. Pour l'instant, dans la sous-section suivante, nous présentons les prémisses pour la modélisation de la théorie de la viabilité selon Sock (1998).

5.4.2 Modélisation de la production-perception de la parole dans le cadre de la théorie de la viabilité

Il faut rappeler qu'auparavant, certains chercheurs traitaient séparément la production de la parole et la perception de la parole, comme si ces deux aspects de la parole formaient des champs isolés. L'articulatoire et l'acoustique n'auraient ainsi peu de lien. Mais de nos jours, nous savons qu'une telle approche n'est pas satisfaisante : la production et la perception sont intimement liées à tel point qu'il est dorénavant impossible de conduire des études articulatoires tout en négligeant le versant acoustico-perceptif, et vice versa. Dans ce cadre de la non-séparation de l'articulatoire de l'acoustique, les travaux de Sock (1998, 2001) ne se sont pas limités à la théorisation pure et ne pouvaient être complets sans une approche modélisatrice. En effet, l'auteur propose de modéliser la production-perception de la parole au regard de la théorie de la viabilité, en s'appuyant sur des

recherches conduites dans le domaine de la biologie, notamment sur la modélisation de systèmes autonomes simple, artificiel ou vivant, de Varela et Bourgine (1992). De plus, il convient de ne pas oublier le rôle primordial que joue le système nerveux central dans la coordination des processus sensori-moteurs. La douleur est un exemple concret de ce processus : lorsque l'on se blesse sur une partie du corps (musculaire – morsure de la langue par exemple), un signal est envoyé au cerveau. Ce dernier analyse les signaux et renvoie cette sensation de douleur éprouvée. L'inverse serait les réactions physiques (transpiration, battements cardiaques, convulsions, etc.) en raison d'un sentiment de frayeur par exemple. Dans la même lignée, en phonétique, Sock (1998) nous rappelle que « tout comme les actions motrices ont des conséquences sensorielles, les actions sensorielles ont aussi des conséquences motrices » (p. 173). Il illustre ce mécanisme par le schéma suivant (figure 5).

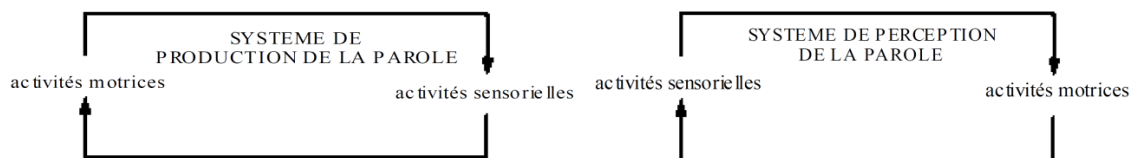


Figure 16 : (à gauche) visualisation de l'organisation du système de production de la parole en tant qu'unité ayant sa propre clôture opérationnelle et (à droite) visualisation de l'organisation du système de la perception de la parole en tant qu'unité ayant sa propre clôture opérationnelle (Sock, 1998, p. 174).

En intégrant les paramètres neuronaux au schéma précédent, Sock (1998) obtient un modèle représentatif de production-perception de la parole illustré ci-dessous (figure 6). Selon ce schéma, le système de production-perception de la parole fait émerger des événements articulatoire-acoustiques, grâce à une interconnexion cohérente entre l'environnement et les dynamiques neuronales. Les trajectoires entre ces deux entités fonctionnent dans les deux sens et s'effectuent de manière concomitante. En conséquence, le système fait émerger des patterns articulatoire-acoustiques robustes et assure, en son sein, un comportement dynamique. Parmi un ensemble de trajectoires articulatoire-acoustiques possibles, le système doit déterminer celles qui sont viables linguistiquement dans le temps et dans l'espace. Tout ce passe comme si le système était doté d'une capacité de filtrage, quant aux trajectoires viables et celles qui ne le seront pas. Selon Sock : « un système autonome doit deviner, à tout moment, un ensemble de solutions qui soient toutes viables en éliminant les autres » (1998, p. 177).

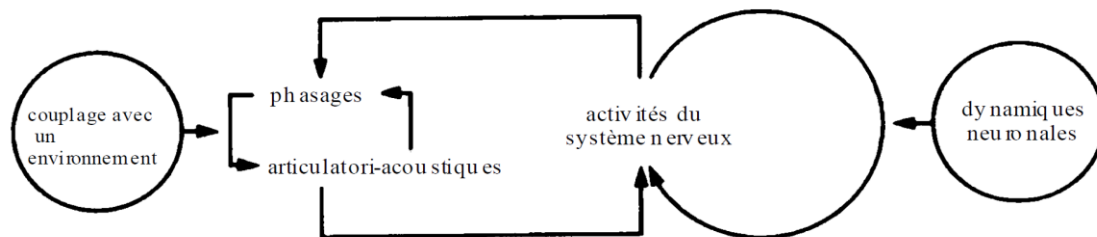


Figure 17 : catégories phonétiques, émergeant sous forme de phasages articulatoire-acoustiques, correspondent à la manifestation d'une certaine cohérence obtenue par la clôture de toutes les matrices d'interconnexion du système de production-perception de la parole (Sock, 1998, p. 175).

La théorie de la viabilité admet que la variabilité est un phénomène inhérent à la production-perception de la parole et que celle-ci découle de la complexité du système dans lequel s'inscrit toute production de la parole. Ce cadre stipule une adaptation malgré la variabilité sous-jacente, en vue de préserver la survie du message au travers des oppositions linguistiques. Cependant, il existerait une zone dite de viabilité au sein de laquelle toutes les productions de la parole seraient perçues comme viables, c'est-à-dire intelligibles. Dans ce contexte, la viabilité serait circonscrite dans un fonctionnement systémique et mobiliserait l'interaction d'un nombre de facteurs agissant simultanément de façon stable et instable. Il s'agit des gestes, de la perception, d'activités du système nerveux et de l'environnement au sein duquel ils œuvrent. Grâce à la modélisation de ce concept de viabilité, nous comprenons mieux que les systèmes de production et de perception de la parole sont intimement liés au travers une spécification mutuelle. Dans la toute prochaine section, nous proposons de traiter cette notion d'in/stabilité articulatoire à travers le cadre conceptuel de la perturbation en production-perception de la parole.

5.5 Perturbation du système de production de la parole

Maintenant que nous avons tenté d'explicitier les soubassements de la théorie de la viabilité et son importance en production-perception de la parole, nous pouvons introduire un concept clé connexe, qui est la perturbation en production de la parole. En effet, tout facteur pouvant entraîner de la désorganisation dans le système de la production de la parole est considéré comme étant un perturbateur. Celui-ci affecte souvent l'identité des marqueurs linguistiques d'énoncés, notamment phonético-phonologiques. Cette notion, traitée d'une manière exhaustive et claire par

Vaxelaire (1996, 2007), montre qu'il existe des facteurs internes et externes responsables de déviations éventuelles dans le système linguistique. Vaxelaire en distingue deux types majeurs présentés comme suit :

a) La première est constituée par l'environnement externe, sous forme de source d'événements indépendants, au sens où ces événements ne sont pas déterminés par l'organisation du système lui-même. Le meilleur exemple de ce cas est celui d'un environnement « bruité », les discordances audio-visuelles (l'effet McGurk), lors d'une tentative de communication linguistique

b) La deuxième est constituée par le système lui-même, sous forme de changements d'états, suite aux modifications des interactions structurelles de ses propres composantes. On peut citer, dans cette catégorie de perturbations, celle qui nous intéresse au premier chef ici, la vitesse d'élocution, mais aussi d'autres, telles les productions sous ou sur-articulées (hypo-hyper speech), les pathologies de la production-perception de la parole (Vaxelaire, 2007, p. 33).

Quelle que soit la catégorie des perturbateurs (internes ou externes), deux cas de figure peuvent se présenter. En premier lieu, les effets d'une perturbation sont « conservateurs » lorsque, dans un système, seules les relations entre les composantes subissent des modifications par leurs actions. Les composantes conservent une stabilité relative dans leur fonctionnement phonétique. Ainsi, malgré la perturbation, certains gestes articulatoires en vue de produire un son spécifique restent relativement inchangés. On parle alors de la résistivité¹³⁹ des patrons spatiotemporels de tels gestes. C'est notamment le cas de la nasalité, où le contrôle moteur des mouvements du vélum, lors de l'augmentation de la vitesse d'élocution, subit des compressions temporelles (durées absolues), sans que la stratégie utilisée (timing relatif) dans la nasalisation elle-même ne soit atteinte (Vaxelaire, 1995).

En deuxième lieu, les constituants du système lui-même connaissent des modifications de la part des perturbateurs. Ce sont là des effets dits « innovateurs » provoqués par la perturbation. Dans ce cas précis, les trajectoires articulatoires inhérentes aux organes responsables de la production de la parole subissent des changements internes. Des phases d'instabilité sont donc observées. Sous l'effet de l'augmentation de la vitesse d'élocution par exemple, les données articulatoire et acoustique montrent des trajectoires différentes pour les consonnes simples et les consonnes doubles/gémées (VC et VCC), dans le cadre d'une opposition phonétique ou phonologique.

139 Lorsque les gestes articulatoires résistent aux perturbations auxquelles ils sont soumis.

5.5.1 Perturbateur écologique ou naturelle

Un perturbateur écologique est celui qu'un « locuteur est susceptible de s'imposer à lui-même dans le cadre d'une conversation normale » (Steiblé, 2014, p. 38). Nous pouvons rajouter que c'est un objet ou une situation qui entraîne la modification de la parole normale, soit en contexte naturel (conversation en famille), soit en contexte expérimental (demandé par un investigateur). Parmi les facteurs internes définis plus haut, la vitesse d'élocution, les productions sous- ou sur-articulées (hypo-hyper speech) et les pathologies de la production-perception de la parole (Bechet, 2011 ; Fauth, 2012 ; Hirsch, 2007) constituent des exemples concrets de perturbateurs, mais « [...] la variation de la vitesse d'élocution est retenue comme moyen naturel de perturbation du système de production de la parole » (Vaxelaire, 2007, p. 27). Cette variation peut s'opérer à divers degrés, plus ou moins lente ou plus ou moins rapide. Tout dépend de la capacité de diction et de perception de chaque individu, car un débit de parole lent pour un locuteur A, peut être perçu comme rapide par un locuteur B. Steiblé (2014) inclut le fait de parler une L2 et l'imitation d'accents comme des perturbateurs écologiques (nous y reviendrons ultérieurement). Il est communément admis que la vitesse de parole se divise en deux aspects reliés. Il y a la vitesse d'élocution proprement dite et la vitesse d'articulation. Les deux correspondent au nombre de syllabes réalisées par seconde, à la seule différence que la durée des pauses est prise en compte dans la première. La durée d'élocution d'un locuteur est égale au temps d'articulation plus la durée totale de toutes les pauses (remplies, silencieuses, hésitations, faux départs, etc.¹⁴⁰). L'augmentation de la vitesse d'élocution nous intéresse particulièrement, car non seulement elle représente une situation réelle et naturelle de la vie quotidienne, mais aussi parce qu'elle permet de tester la résistivité ou la robustesse des phénomènes articulatoire-acoustiques. Le locuteur choisit de varier sa vitesse d'élocution (ajustement) en fonction de son interlocuteur, pour des besoins d'intercompréhension ou tout simplement pour s'adapter aux contextes dans lesquels il se trouve. Ainsi, toutes les activités de production de la parole, impliquant des contraintes, notamment de temps (du *Speed Dating* à une conversation de dernière minute pendant le départ du train), peuvent être considérées comme des exemples de situations réelles où un locuteur pourrait être contraint à s'exprimer rapidement.

¹⁴⁰ Consulter Grosjean & Deschamps (1975) pour plus de détails.

5.5.2 Perturbation non écologique ou « artificielle »

Dans une acception large, est considéré comme perturbateur non écologique ou artificiel, tout facteur entraînant des dysfonctionnements de la parole, sous l'effet d'un agent externe. Contrairement aux perturbateurs naturels qui sont internes au système, les perturbateurs non-écologiques proviennent dans la majeure partie des cas de sources non habituelles. Précisons que même si ces perturbateurs sont plutôt utilisés en phonétique, à des fins expérimentales (exemple du *bite-block*¹⁴¹ ou du placement sur la langue d'un palais artificiel comportant des électrodes), ils peuvent également intervenir dans des situations naturelles, comme par exemple lorsque l'on parle avec une cigarette ou un *chewing-gum* dans la bouche. Il s'agit également de situations génératrices de bruits souvent assourdissants, comme les boîtes de nuit ou des concerts musicaux, où les locuteurs sont obligés de parler très fort pour se faire entendre.

Un autre type de perturbateur non écologique se compose des discordances audio-visuelles bien connues du type Effet McGurk-Macdonald (McGurk & Macdonald, 1976 ; Skipper, van Wassenhove, Nusbaum, & Small, 2007). Ce dernier nous intéresse particulièrement dans notre travail qui concerne le visionnage des séries télévisées. L'effet McGurk-Macdonald est un phénomène d'illusion perceptivo-moteur qui implique une interférence entre l'audition et la vision, sous l'impulsion de l'interprétation de signaux provenant du cerveau. Alors que la perception de la parole était pendant longtemps considérée un mécanisme auditif, des études ont montré que la vision y contribue considérablement (Dubois, 2009). Ainsi, la perception auditive est accentuée dans une communication lorsque les sujets visualisent le mouvement des lèvres de la personne qui parle. Cette lecture oro-faciale¹⁴² en générale et labiale en particulier permet de mieux comprendre un message, car les mouvements du visage et des lèvres donnent des indices complémentaires quant aux émotions exprimées et aussi aux sons produits. C'est l'une des raisons pour laquelle nous nous sentons obligés de parler plus fort et plus distinctement lorsque notre interlocuteur ne nous voit pas. C'est aussi un élément qui aide les malentendants lorsqu'une personne s'adresse à eux. Ce phénomène est valable aussi bien en présentiel qu'en distanciel (à

141 Qui empêche les gestes de mastication. C'est-à-dire le blocage des mouvements de la mandibule par insertion d'un objet léger entre les molaires.

142 Il a d'ailleurs été montré par des expériences de neuro-imagerie que les aires prémotrices du langage parlé (aire de Broca) interviennent dans la lecture oro-faciale (Roulin, 2006).

travers une vidéo). Il existerait en définitive une flexibilité dans la correspondance audio et visuelle. L'effet McGurk-Macdonald consiste à présenter une information conflictuelle à des sujets afin d'analyser leur réaction. On leur présente un extrait vidéo d'une personne prononçant par exemple¹⁴³ le son /ba/. Mais la sortie du son est doublée pour que la bande sonore joue le son /ga/. Les résultats montrent que de façon générale les sujets disent entendre le son /da/. Ce mélange d'informations acoustique et visuelle serait un compromis acoustique entre les sons /ba/ et /ga/ interprété par le cerveau en /da/ (Skipper, van Wassenhove, Nusbaum, & Small, 2007). Des études similaires ont été conduites avec des vidéos muettes ou non modifiées. Les sujets ont été capables de deviner les sons prononcés. Pour nous, ces données indiquent que le visionnage de vidéos en VO faciliterait l'imitation phonético-phonologique, et donc l'acquisition (si elle a lieu) d'*input*, et que les vidéos doublées seraient sources de perturbations pour la compréhension du message.

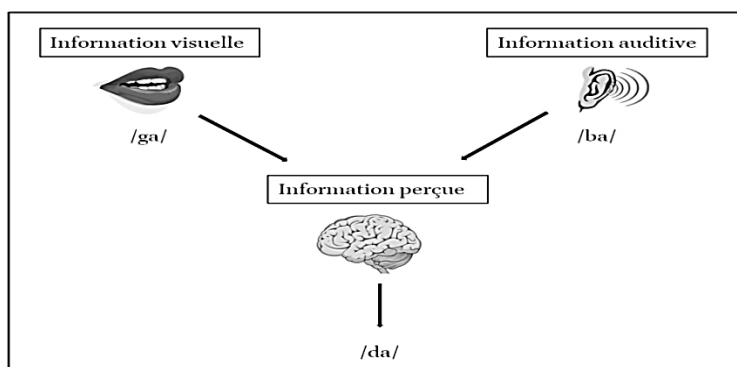


Figure 18 : illustration de l'effet McGurk-Macdonald.

5.5.3 Conséquences des perturbations sur la production de la parole

Suite aux travaux de classification de perturbateurs en phonétique de Vaxelaire (1995, 2007), nous présentons dans le tableau suivant un récapitulatif des types de perturbateurs, quelques exemples et leurs conséquences possibles. Nous indiquons également des auteurs ayant confirmé leurs conséquences par des études.

Tableau 14 : exemples de perturbateurs et leurs conséquences sur la production-perception de la parole.

Perturbateurs internes et externes : exemples	Quelques conséquences	Études
---	-----------------------	--------

¹⁴³ On utilise aussi les syllabes VCV (aba + aga =>ada), les mots mail + deal => nail et les phrases « my bab pope me poo brive + my gag koke me koo grive => my dad taught me to drive ».

Bruit	Augmentation d'amplitude, de durée, de hauteur, d'intensité de la voix ; variation des formants ; hyper articulation	Summers, Pisoni, Bernacki, Pedlow, & Stokes, 1988 ; Le Prell & Clavier, 2017
Illusions perceptives/audiovisuelles	Fusion et combinaison articulatoires (Effet McGurk)	MacDonald, 2018 ; McGurk & MacDonald, 1976 ; Tiberghien, Roulin, & Beauvois, 1992
Contraintes de temps	Augmentation du débit de parole, vitesse d'élocution, vitesse d'articulation	Howell, 2008 ; Lively, Pisoni, Summers, & Bernacki, 1993
État d'ivresse	Augmentation de la durée des segments/mots (débit de parole lente), augmentation de la fréquence fondamentale, hypo-articulation, ratage de cibles, omissions de phonèmes/mots, ellipses, etc.	Andrews, Cox, & Smith, 1977 ; Behne, Rivera, & Pisoni, 1991 ; Hollien, <i>et al.</i> , 2001
Discussion avec des enfants	Hyper-articulation, débit lent, schémas intonatifs atypiques	Cooper, Fecher, & Johnson, 2018 ; Honig, 2017 ; Suttora <i>et al.</i> , 2017
Pathologies ¹⁴⁴ (divisions palatines, bégaiement, thyroïdectomie, glossectomie, Dysarthrie)	Altération au niveau de réalisations consonantiques et vocaliques	Bechet, 2011 ; Fauth, 2012 ; Hirsch, 2007 ; Fougeron <i>et al.</i> , 2010 ; Zaouali (2019)
Vire langues ¹⁴⁵	« Mélange » et erreurs de prononciation des sons/phonèmes cibles	Goldrick & Blumstein, 2006 ; Goldstein, Pouplier, Chen, Saltzman, & Byrd, 2007
Répétition	Augmentation progressive de la vitesse d'élocution ; une plus ou moins meilleure articulation	Lambert, Kormos, & Minn, 2017 ; Rogalsky <i>et al.</i> , 2015
Lecture	Réduction de la vitesse, ratage de cibles, confusion de phonèmes	Mok, Lee, Li, & Xu, 2018 ; Saletta, 2015 ; Zhang & Damian, 2012
Parler en mangeant	Difficultés d'articulation de certains sons, problèmes de résonance	(Hantke <i>et al.</i> , 2016 ; Hiiemae & Palmer, 2003)
Langues secondes	Difficultés d'articulation de certains sons (inexistants dans la L1), prosodie	Delvaux, Huet, Piccaluga, & Harmegnies, 2013 ; Flege, 2003
Imitation phonétique	Hypo ou hyperarticulation, sous ou surexagération	Steiblé, 2004 ; Marc Sato, Grabski, Granjon, Jean-Luc Schwartz, & Nguyen, 2010 ; Pardo, 2006 ; Pardo, Gibbons, Suppes, & Krauss, 2012 ; Pardo, Jay, & Krauss, 2010)

144 Voir les Actes des 4e Journées de Phonétique Clinique, mai 2011 (Sock, Vaxelaire, & Fauth, 2014).

145 De l'anglais Tongue *twisters* il désigne un groupe de mots difficiles à articuler, assemblés dans un but ludique ou pour servir d'exercice d'élocution (<https://www.larousse.fr/dictionnaires/francais/virelangue/186851> consulté le 18/07/2019).

De ce qui précède, nous voyons que les sources et les conséquences des perturbateurs évoqués peuvent être nombreuses et variées. Cependant, force est de constater que les messages que nous formulons dans ces conditions d'instabilité demeurent généralement compréhensibles. À quoi serait due cette préservation des énoncés linguistiques en dépit des perturbations ? Dans la sous-section suivante, nous présentons un de ces facteurs¹⁴⁶ à travers la notion de la résistivité et de ses caractéristiques.

5.5.4 Résistivité dans le cadre de la perturbation en production de la parole

Pour clore cette partie sur la perturbation, nous évoquons quelques notions importantes pour la compréhension globale des mécanismes sous-jacents à la production-perception de la parole. Nous avons vu que les deux catégories de perturbateurs, quelle que soit leur nature (externes ou internes, écologiques ou non-écologiques) peuvent parfois entraîner la désintégration du système linguistique et que ces perturbateurs peuvent être conservateurs ou innovateurs (Sock & Vaxelaire, 2001 ; Vaxelaire, 2006). Dans les deux cas, toutes les trajectoires articulatoires qui ne subissent pas de modifications après une perturbation, c'est-à-dire qui se maintiennent stables dans le système, sont dites résistantes. Comme le souligne Vaxelaire (2007) : « [...] toute résistivité spatio-temporelle des gestes linguistiques face à un facteur perturbateur serait indicative de leur pertinence linguistique » (p. 26) et qu'« il deviendra possible de saisir le taux de résistivité de ces gestes, ainsi que les stratégies déployées par les articulateurs pour préserver le message linguistique » (*ibid.*, p. 60). Dans le cas contraire (absence de résistivité), le système cherche le moyen de compenser les effets provoqués par la perturbation, en empruntant une trajectoire différente. Lorsqu'après compensation la nouvelle trajectoire réussit à atteindre la cible, on parle de réajustement. Tout ceci témoigne de la flexibilité¹⁴⁷ du système, capable de se réorganiser après des phases de turbulences, comme le précise Vaxelaire (2007) :

La résistivité ou la robustesse de l'organisation spatio-temporelle des gestes du conduit vocal implique souvent des réajustements moteurs touchant à des structures spécifiques du conduit vocal. C'est ainsi qu'il nous est possible de mettre au jour les éventuelles

146 Nous avons déjà montré que le contexte et l'habitude constituent également des facteurs déterminants pour la compréhension des messages.

147 On parle également de plasticité ou d'élasticité.

stratégies de compensation motrice, souvent révélatrices du degré de flexibilité des structures articulatoires impliquées dans la production d'une tâche phonétique ou phonologique déterminée (p. 32).

Au terme de ces explications, il est à retenir que la résistivité d'un système de production de la parole dépend de la manière dont il se réorganise, à travers des stratégies de compensation ou de réajustement adopté par un individu. Cette autorégulation implique souvent l'emprunt de voix articulatoires différentes de celles qui sont généralement utilisées. Nous verrons dans ce qui suit le lien entre ce phénomène articulatoire et la notion de la viabilité.

5.5.5 Lien entre la théorie de viabilité en phonétique et la notion de perturbation

Comme nous l'avons fait remarquer, de manière explicite ou implicite tout au long de cette partie, il existe un lien étroit entre la notion de perturbation et la théorie de la viabilité (Sock & Vaxelaire, 2001). Face à la variabilité des trajectoires articulatoires inhérentes aux productions d'un locuteur, sous les effets de la perturbation, le système linguistique tend à résister ou à se réajuster. Toutes ces configurations concourent à un objectif précis qui est celui de garantir des réalisations qui seraient circonscrites dans une zone de viabilité, permettant à l'auditeur de saisir les contrastes phonétiques et phonologiques. En effet, un son qui conserve son identité linguistique, même après une perturbation, est un son viable. Pour sa part, un système non viable est celui qui perd son identité linguistique après une perturbation, et ce malgré les tentatives de compensations et de réajustement, comme le souligne Vaxelaire (2007) :

Il arrive parfois que le système de production-perception de la parole ne puisse plus adopter différents états cohérents, ou que son organisation n'arrive plus à garder une certaine stabilité, face à une perturbation interne ou externe : cela amène inéluctablement la désintégration du système et, en conséquence, la perte de l'identité de ses propriétés émergentes. Nous aurons ainsi affaire à des catégories sensori-motrices linguistiquement non viables (Vaxelaire, 2007, p. 34).

Nous tenons à préciser que cette situation, même si elle semble binaire (viable / non viable), n'est pas figée, car un système non viable à un instant t , peut le devenir à un autre moment. Et lorsqu'un son « rentre dans son domaine de viabilité, il y restera pour toujours » (Sock, 1998, p.

178). Ainsi, il semblerait que ce passage d'un état non viable à un état viable soit l'un des facteurs à l'origine des changements linguistiques¹⁴⁸.

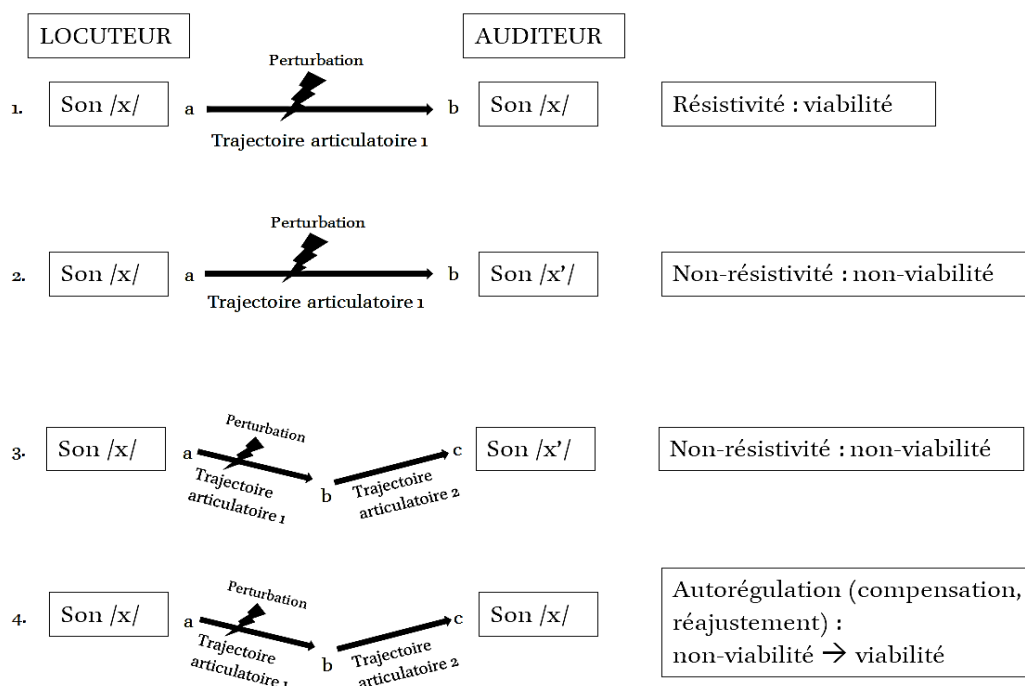


Figure 19 : schéma représentant la relation entre viabilité et perturbation.

La figure ci-dessus (g) montre le lien qui existe entre la théorie de la viabilité et la notion de perturbation en production-perception de la parole. Nous avons résumé en quatre catégories les différentes situations qui peuvent se présenter. Le son /x/ se réfère à un son/phonèmecible, pouvant être remplacé par n'importe quel son et le son/x'/ représente un autre son quelconque mais différent de /x/ (cf. symboles phonétiques de l'IPA, 2005). Quant aux trajectoires, elles ne sont qu'un exemple parmi les nombreuses possibilités empruntables en vue d'atteindre la cible. Par exemple, pour la trajectoire articulaire du son /i/¹⁴⁹, le mouvement de la langue est constitué par son déplacement de sa position initiale¹⁵⁰ (a) vers l'avant de la cavité buccale (b). Ainsi, dans le cas 1, nous assistons à la production du son cible malgré la perturbation. La trajectoire a-b est donc

148 En général, un changement linguistique est un processus par lequel les systèmes phonétique, morphologique, grammatical, morphologique, sémantique, etc. d'une langue donnée subissent des variations dans le temps. Au niveau phonétique, les changements vocaliques, consonantiques et prosodiques seraient responsables de la formation des accents. Les autres causes sont notamment le facteur économique (principe du moindre effort), le contact linguistique, et l'environnement culturel (cf. Labov, 2001 ; Ettliger, 2007 ; Lakkaraju, Swarup, & Gasser, 2012 pour plus d'informations sur les changements phonétiques).

149 /i/ est une voyelle antérieure fermée non arrondie (palatale).

150 État de repos.

résistante et nous pouvons déduire que ce système est viable. À l’opposé, dans le cas 2, la perturbation a entraîné la réalisation d’un son différent /x’/. Nous avons affaire à la non-résistivité de la trajectoire a-b et donc à un système non viable. Quant au cas 3, le système cherche une autre trajectoire possible (b-c), après la perturbation de a-b, en vue de produire le son cible /x’/. Or, il se produit un « ratage » de cible puisque le son obtenu au final est tout de même différent de /x/ (/x’/). La même conclusion que dans le cas 1 est retenue, c’est-à-dire un système non viable, même après qu’il ait tenté de se réajuster. Le dernier cas est similaire au troisième mais avec un résultat différent : la solution de remplacement après la perturbation (b-c) a permis d’obtenir le son cible /x/ grâce à une stratégie d’autorégulation (compensation, réajustement). Un système non viable au départ deviendrait viable par la suite.

En définitive, un système de production de la parole qui résiste aux conditions de perturbation est assurément viable. Ce cadre est capital pour nous, dans une logique d’enseignement/apprentissage d’une L2, car il peut déterminer les facteurs intervenant dans la communication de la chaîne parlée. Comme nous venons de le voir à travers ce qui précède, les notions de viabilité et de perturbation sont liées par un mécanisme plus ou moins complexe mais qui nous permet de mieux appréhender le fonctionnement des messages linguistiques. Sans tarder, nous allons voir quelles pourraient être les implications de la théorie de la viabilité et la notion de perturbation en production-perception de la parole, pour le domaine de l’apprentissage des L2.

5.6 Théorie de la viabilité en phonétique : implications pour l’apprentissage des L2

La théorie de la viabilité en production-perception de la parole peut être rentabilisée dans le domaine de l’enseignement/apprentissage des L2, dans la mesure où elle peut permettre non seulement d’évaluer la résistivité ou le réajustement de la parole face aux perturbateurs (écologiques et non écologiques), mais aussi d’appréhender comment les apprenants se font comprendre par les natifs ou non-natifs¹⁵¹ d’une langue donnée. Nous avons mentionné la L2 et l’imitation d’accents comme étant des perturbateurs écologiques de parole : la prise de parole en public notamment devant les natifs et le fait de s’efforcer à imiter l’accent d’un locuteur peuvent

¹⁵¹ Le problème de compréhension touche aussi bien aux interactions non-natifs/natifs qu’aux interactions entre non-natifs.

constituer des situations anxiogènes pour les apprenants. Or, il faut rappeler que l'appropriation d'une L2 nécessite des efforts considérables, aussi bien au niveau articulatoire que perceptif. Nous savons aussi que pour la langue première (L1), ces deux niveaux s'acquièrent dès la naissance (bien avant pour la perception). Du point de vue perceptif, bien que n'importe quel nouveau-né discrimine quasiment tous les contrastes phonétiques de toutes les langues du monde, cette capacité s'amenuise progressivement vers la fin de la première année : l'enfant se « spécialise » alors dans sa L1. Quant à l'articulatoire, les bébés acquièrent, à travers le gazouillis et le babillage, les gestes articulatoires de leurs L1, par imitation des gestes oraux, avec leurs conséquences sonores, de leur entourage. À partir d'un certain âge, excepté les langues apparentées, l'apprenant de L2 doit s'approprier tous les aspects articulatoire-acoustiques de la L2. Pour ce faire, il doit apprendre à distinguer perceptivement les sons qui sont viables de ceux qui ne le sont pas. De même, il doit apprendre à produire des sons viables pour se faire comprendre. La théorie de la viabilité permettrait ainsi de rationaliser le cheminement de l'apprenant qui veut rendre intelligible son discours. En effet, il doit réaliser une trajectoire articulatoire qu'il n'a pas appris à faire dès l'enfance, afin d'atteindre une cible viable (à défaut d'être celle d'un natif) pour se faire comprendre. Cet équilibre articulatoire-acoustique de la part du locuteur permettra ensuite à l'auditeur de percevoir les sons viables. N'ayant pas l'habitude de parler la L2, les productions articulatoires (qui, rappelons-le, sont des tâches aérodynamiques et musculaires) de l'apprenant vont fortement varier, surtout au début de l'apprentissage. Nous avançons l'hypothèse que tant que cette variabilité peut s'adapter d'un point de vue articulatoire-acoustique et reste dans une zone de viabilité, l'apprenant se ferait comprendre. Il faut toutefois rappeler que cette viabilité doit tout autant concerner les aspects segmentaux que suprasegmentaux (prosodiques). Ces éléments soulèvent d'ailleurs des interrogations quant à la primauté du segmental sur le suprasegmental ou vice versa. Y a-t-il compréhension lorsque seulement un des aspects est respecté ? Ceci étant, nous nous demandons si la notion de *nativeness*¹⁵² n'est pas remise en cause dans ce cadre. Doit-on toujours exiger des apprenants d'avoir des productions proches de celles des natifs ? Ou, doit-on tout simplement considérer leur prononciation comme telle, du moment où leur discours est intelligible, c'est-à-dire dont les trajectoires articulatoire-acoustiques sont viables ? Ces questions

¹⁵² La notion de *nativeness* (cf. Gut, 2009 ; Waniak-Klimczak & Shockey, 2012) désigne les productions d'un apprenant de L2 (très) proches de celles des natifs. Nous reviendrons sur ce concept dans la discussion des résultats.

rejoignent celles posées par Jenkins et collaborateurs (Jenkins, 2000 ; Jenkins, Baker, & Dewey, 2017) qui pensent que les productions orales des apprenants devraient être corrigées seulement lorsqu'elles ne sont pas compréhensibles, c'est-à-dire, pour nous, pas « viables ». En d'autres termes, toute prononciation qui n'entrave pas la distinctivité phonétique et phonologie devraient être considérée comme acceptable au vue de l'influence de la L1. C'est ainsi qu'indépendamment des accents adoptés par les apprenants, l'objectif de communication devrait être accompagné d'un travail¹⁵³ de viabilité en production-perception de la parole. Produire des sons se situant dans les zones de viabilité assurerait une meilleure intelligibilité. Le travail du didacticien serait alors d'aider les apprenants à produire des sons viables, en priorité.

5.7 Conclusion

La variabilité dans la production-perception de la parole est inhérente à chaque individu. Elle se manifeste sous plusieurs formes et à divers degrés, sous l'influence d'une multitude de facteurs externes et internes au locuteur-auditeur. Dans ce chapitre, nous avons essayé d'expliquer comment la variabilité se manifeste dans les divers contextes phonétiques et phonologiques, afin de garantir une meilleure intelligibilité du discours de l'énonciateur. La parole serait ainsi conditionnée par un principe de moindre effort, tout en gardant un seuil de comportement acceptable et compatible avec la survie et la reproduction du système de production et de perception de la parole (Sock, 1998). Pour une meilleure compréhension entre le locuteur et l'auditeur, un compromis est souvent trouvé entre ces deux entités en fonction du contexte. Ainsi, nous avons soutenu que quelle que soit la nature des facteurs perturbateurs, nous pouvons assister aussi bien à une stabilité qu'à une instabilité du système de production-perception de la parole. Lorsque le système maintient sa stabilité malgré les perturbations, il est dit résistant. En revanche, si des phases d'instabilités apparaissent, le système trouve le moyen de compenser, voire se réajuste en empruntant une autre trajectoire, afin d'assurer son fonctionnement habituel. Ceci étant, la théorie de la viabilité en production-perception de la parole tente d'expliquer ces mécanismes sous-jacents à la variabilité phonético-phonologique. Elle pose que toute trajectoire articulatoire permettant d'aboutir à une cible articulatoire-acoustique est permise et donc viable. Une cible devient viable lorsqu'elle permet de maintenir les oppositions phonétique et

¹⁵³ Nous fournissons des détails dans la partie discussion de ce manuscrit.

phonologique dans une langue donnée. Certains aspects contribuant à la compréhension du discours linguistique peuvent ainsi être saisis dans le cadre de la théorie de la viabilité en production-perception de la parole : un son viable malgré la variabilité est un son fiable. Pour notre part, le lien que nous avons établi entre la viabilité et l'apprentissage des langues est crucial. Ce lien entre viabilité et l'apprentissage des langues, à partir de nos données expérimentales, devrait nous permettre de mettre au jour les différentes stratégies que les apprenants peuvent déployer pour produire des sons viables, tout en prenant en compte la variabilité dans leurs productions. Nous verrons plus tard que la théorie de la viabilité et la perturbation en production-perception de la parole s'avèrent particulièrement pertinentes dans le cadre de notre recherche.

6 Cadre méthodologique

La nature des données mais aussi celle de leur mode de collecte ou de production, de sériation ou de constitution en corpus précontraignent le choix d'une méthodologie destinée à les exploiter. (Beacco, 2011, p. 32).

La méthodologie est la fondation de toute recherche car sur elle repose en partie la validité des résultats et des conclusions qui en découlent. Parce qu'elle permet de réduire les biais et de garantir la scientificité du travail de recherche, les protocoles expérimentaux se doivent d'être rigoureux et soigneusement bâtis. Ce faisant, ayant évoqué plus haut les démarches méthodologiques recommandées dans le cadre des études impliquant des sujets complexes (longitudinales, rétrodictives, mixtes), nous avons opté pour une approche multiple et mixte de collecte de données. Nous traiterons dans ce chapitre trois grandes parties qui sont la méthodologie utilisée pour : la collecte, le traitement et l'analyse des données. La première décrit le recrutement et le profil des participants, le support qui a servi aux enregistrements audio, les modalités des enregistrements et les mesures acoustiques des éléments étudiés, effectuées à la lumière de la littérature sur la différence phonético-phonologique entre la RP et le GA. Dans la deuxième partie nous décrivons le test de perception, qui a été élaboré en vue de savoir si les participants distinguent les deux accents. Enfin, la dernière partie est consacrée au questionnaire rempli par les participants, dans laquelle nous présentons les questions qui leur ont été posées, concernant les activités auxquelles ils s'adonnent en anglais.

6.1 Recrutement des participants

Avec la volonté de rester dans la lignée des travaux sur l'AIAL, cette étude s'est fixé l'objectif de mobiliser uniquement des étudiants spécialistes d'autres disciplines (Candas, 2009), qui ne prennent que quelques heures de cours d'anglais (généralement entre une et deux heures par semaine). Ils font donc partie du secteur des Langues pour Spécialistes d'Autres Disciplines (LanSAD) que Poteaux (2014) définit comme « l'enseignement des langues étrangères aux étudiants qui ne se spécialisent pas en langues mais dans d'autres disciplines » (paragr. 1). Ce public est particulier et diffère des spécialistes en anglais, dans la mesure où les cours de phonétique, de phonologie ou de littérature anglaise par exemple ne leur sont pas dispensés. Dans

cette lignée, les participants de cette étude sont des étudiants inscrits dans divers cursus à l'Université de Strasbourg. La majorité provient des Centres de Ressources en Langues (CRL), car ces derniers constituent l'endroit idéal pour rencontrer un public LanSAD (Toffoli et Sockett, 2015). Nous nous ne les connaissions pas et nous les avons sélectionnés au hasard, à la sortie de séances de cours. Ensuite, nous leur avons proposé de participer à une étude sur la lecture à voix haute en anglais. Notre objectif final relatif à la prononciation ne leur était pas communiqué. À celles et ceux qui ont souhaité y participer et qui ont accepté de fournir leur adresse électronique, nous leur avons envoyé un mail explicatif (déroulement, durée du test, lieu, etc.) et un Doodle¹⁵⁴ afin de fixer une date à laquelle ils souhaiteraient passer le test. Quelques collègues enseignants ont aussi collaboré lors du recrutement, en relayant l'annonce dans leur classe. Grâce à cette intervention, une dizaine de volontaires (11) ont pu se joindre au groupe.

Au total, 32 étudiants ont répondu à notre appel et ont été enregistrés. 12 d'entre eux ont été écartés, étant donné que le français n'est ni leur L1, ni leur L2. Ce fait est imprévisible en raison de la méthodologie suivie, car il fallait contrôler, à ce niveau, les caractéristiques des locuteurs. En fin de compte, nous n'avons retenu que 20 étudiants (N=20) pour les analyses, dont 11 femmes et 9 hommes. Cette disparité homme-femme est un constat, qui aujourd'hui est assez courant dans les effectifs des étudiants dans les universités françaises, notamment dans les sciences humaines et sociales. La méthodologie de recrutement employée (*random picking*¹⁵⁵) a également contribué à cet écart, car la population féminine, outre le nombre, a été plus enclin à se porter volontaire et à honorer leur engagement à venir passer le test¹⁵⁶.

6.2 Matériel de collecte de données audio

Nous avons enregistré tous les locuteurs dans le studio insonorisé de l'Institut de Phonétique de Strasbourg (IPS), à l'aide d'un enregistreur audio numérique Marantz Professional® (PMD661 MKII) et d'un microphone Sennheiser® (e845S). La nécessité de la prise de sons dans des conditions anéchoïques¹⁵⁷, couplée avec l'utilisation du matériel professionnel est cruciale. Elle permet

154 Outil de création de sondages et de planificateur de réunion en ligne : <https://doodle.com/fr/>

155 Recrutement aléatoire de participants sans aucun critère de sélection prédéfini et qui est propre à la recherche exploratoire.

156 Nous tenons à souligner que parmi les personnes qui ont répondu positivement à l'annonce, seules quelques-unes se sont présentées aux rendez-vous.

157 En l'absence d'échos et de bruits.

d'éviter des signaux bruités (interférences et parasites sonores), pouvant rendre (encore plus) difficiles les analyses acoustiques. De ce fait, les téléphones portables ou tout autre appareil électronique fonctionnant sur des ondes magnétiques, n'étaient pas autorisés dans la salle d'enregistrement. Même la liste des mots, imprimée sur papiers A4, a été plastifiée en vue d'amenuiser ces bruits. Ceci nous a permis d'obtenir des fichiers audio de très bonne qualité, enregistrés en format *mono* (fréquence de 44,1 kHz) et qui sont stockés directement sur une carte SDHC¹⁵⁸ Class 10¹⁵⁹ Samsung 32 Go (formaté en FAT32¹⁶⁰). Tous les fichiers sont enregistrés en extension *wav* (sans perte de qualité).

6.3 La collecte de données

La collecte des données s'est déroulée en trois grandes étapes, les conditions expérimentales étant les mêmes pour tous les participants. Dans un premier temps, les participants sont enregistrés, pendant la lecture d'éléments sur un support (détails ci-dessous). Dans un second temps, ils ont effectué un test de perception, et enfin dans un troisième temps, ils ont rempli un questionnaire. Le temps global nécessaire pour accomplir ces trois phases est d'environ trois quarts d'heure, mais dépend évidemment des capacités de chacun. Il est réparti comme suit : 12 à 18 minutes pour les enregistrements audio, 10 à 15 minutes pour le test de perception et 5 à 10 minutes pour le remplissage du questionnaire.

6.4 Élaboration du support pour la prise de sons

Avant d'entamer l'enregistrement des participants, par soucis de clarté, il nous semble important de commencer par la présentation du support ayant servi à la collecte des données audio. Les enregistrements audio reposent sur une liste composée de trois parties distinctes : une série de mots placés dans une phrase porteuse, une autre composée de blocs de mots lexicalisés (Sockett & Kussyk, 2012) et un texte court. Le corpus oral ainsi obtenu à partir des enregistrements représente une durée totale de 6 heures 37 minutes de paroles (durée moyenne par locuteur : 16 minutes 72 secondes, écart-type : 2,06). Les pauses silencieuses de plus de 20 secondes ne sont pas comptabilisées.

¹⁵⁸ Secure Digital High Capacity.

¹⁵⁹ Vitesse de séquence d'écrire minimale supérieure ou égale à 10 Mo par seconde.

¹⁶⁰ File Allocation Table : table d'allocation de fichier utilisant un système d'adresse 28 à 32 bits.

6.4.1 Les mots

Cette première partie de la liste est élaborée de façon à ne contenir que des mots dont la prononciation diffère manifestement selon les deux accents de référence. L'objectif premier de cette étude est d'observer les éléments phonético-phonologiques qu'adoptent nos participants lorsqu'ils prononcent ces mots. Par conséquent, nous avons consulté la littérature sur la différence de prononciation entre la RP et le GA (Cruttenden, 2014 ; Hughes, Trudgill & Watt, 2012 ; Melchers & Shaw, 2013 ; Roach, 2009 ; Svartvik & Leech, 2016). Des dictionnaires de prononciation, notamment pour les transcriptions phonétiques (EPD)¹⁶¹, (LPD)¹⁶² et le *Macmillan Online Dictionary*¹⁶³, et des bases de données audio contenant des productions de divers locuteurs natifs anglophones (Brulard *et al.*, 2015 ; Hughes *et al.*, 2012 ; Wells, 1982) ont également fait l'objet de recherches.

Selon les différences de prononciation rencontrées dans la majorité des ressources consultées, nous avons pu constituer 10 catégories, selon des phénomènes phonétiques et phonologiques que nous présentons ci-dessous. Le détail de ces aspects se trouve dans la partie « segmentation et mesures acoustiques des sons cibles » :

- la prononciation du /t/ intervocalique ;
- rhoticité ;
- le yod-dropping ;
- la prononciation de la voyelle de *LOT* ;
- la prononciation de la voyelle de *GOAT* ;
- la prononciation des graphies <i> et <y> ;
- la prononciation de la graphie <ei> et <e> ;
- la nasalisation vocalique ;
- la place de l'accent principal ;
- le nombre de syllabes accentués ;
- l'intonation.

Nous avons ainsi recensé une centaine de mots. Mais, nous devons opérer un choix pour des raisons de temps d'expérimentation. Nous avons donc gardé les plus saillants et couramment rencontrés dans la littérature. De ce fait, la liste finale ne contient que 24 mots dans lesquels nous avons inventorié 10 sons ou série de sons cibles, dont la prononciation diffère de façon marquée

¹⁶¹ *Cambridge English Pronouncing Dictionary* (Jones, 2011).

¹⁶² *Longman Pronunciation Dictionary* (Wells, 2008).

¹⁶³ <http://www.macmillandictionary.com/>

entre les accents RP et GA. Nous avons recherché les occurrences de ces mots dans le corpus SOAP¹⁶⁴, afin de vérifier que nos locuteurs les connaissent *a priori*, de par leurs habitus articulatoires dans le langage parlé, principalement dans les séries télévisées. Les résultats seront présentés ultérieurement (*cf.* résultats du questionnaire section 9.7 p. 322). Parfois, nous nous référons à ces sons en utilisant les termes de phénomènes ou caractéristiques. Par exemple, un tap /ɾ/ est une des caractéristiques phonétiques de l'accent GA. Dans notre étude, un son peut se retrouver dans plusieurs mots. Ainsi, le tap /ɾ/ apparaît dans *city*, *vitamin* et *party*. De même, un mot peut contenir deux sons qui différencient ces deux accents : la prononciation des lettres <i> et <t> dans *vitamin*. Chaque mot est placé dans la phrase porteuse *it is **** please* afin d'assurer la production de la totalité du mot et donc du son étudié (*It is hot please*). Une phrase porteuse (ou phrase-cadre) est utilisée en phonétique expérimentale pour simuler « l'effet acoustique d'un discours continu » (Rossi, 2007, p. 171) et par conséquent garantir un contexte segmental et prosodique identique aux sons, aux mots ou encore aux séquences cibles étudiées. Il n'est ni nécessaire, ni souhaitable que cette phrase porteuse soit sémantiquement correcte. Ceci a été utilisé simplement afin d'éviter de présenter des mots isolés à nos participants. Le tableau ci-dessous présente les mots cibles dans la phrase porteuse, avec leurs transcriptions respectives (première partie de la liste).

Tableau 15 : Les 24 mots étudiés avec leur transcription phonétique selon les deux accents.

Mots	Transcription RP	Transcription GA
more	/mɔː/	/mɔːr/
band	/bɑːnd/	/bænd/
city	/'sɪti/	/'sɪri/
vitamin	/'vɪtəmm/	/'vaɪrəmm/
organisation	/,ɔːgənəɪ'zeɪʃ(ə)n/	/,ɔːrgənɪ'zeɪʃ(ə)n/
joke	/dʒɔk/	/dʒoʊk/
hot	/hɒt/	/hɑːt/
tune	/tjuːn/	/tuːn/
bird	/bɜːd/	/bɜːd/
donate	/dəʊ'neɪt/	/'doʊneɪt/
leisure	/'leɪzə(r)/	/'liːzə/
party	/'pɑːti/	/'pɑːri/
military	/'mɪlɪtri/	/'mɪlɪteri/

164 Le corpus SOAP (*Corpus of American Soap Operas*) contient 100 millions de mots provenant de 22 000 transcriptions des séries américaines depuis les années 2000 (Davies, 2011). Ce corpus peut être consulté sur le site <https://corpus.byu.edu/corpora.asp>, de *Brigham Young University* (BYU) qui abrite également le *British National Corpus* (BNC) et le *Corpus Of Contemporary American English* (COCA).

dark	/dɑ:k/	/dɑ:rk/
hand	/hɑ:nd/	/hænd/
dynasty	/'dɪnəsti/	/'dɑ:nəsti/
legalisation	/,li:gəlar'zeɪf(ə)n/	/,li:gəli'zeɪf(ə)n/
road	/rəʊd/	/rɒd/
top	/tɒp/	/tɑ:p/
new	/nju:/	/nu:/
actor	/'æktə/	/'æktə-/
necessary	/'nesəs(ə)ri/	/'nesə,seri/
locate	/ləʊ'keɪt/	/'loʊ,keɪt/
evolution	/,i:və'lju:ʃ(ə)n/	/,evə'lʊʃ(ə)n/

6.4.2 Les phrases

La deuxième partie de la liste est composée de 10 phrases issues des agglomérats de 4 mots (blocs lexicalisés) de Sockett (2011) et des phrases de Kusyk et Sockett (2012). Dans la première étude, Sockett (2011) a détecté à l'aide du logiciel de concordance *Wordsmith 5*, une liste de 50 agglomérats de quatre mots les plus fréquents¹⁶⁵ dans un corpus de scripts de séries télévisées. Le corpus ainsi obtenu a été nommé HHOLD, car il comporte une saison chacune des séries suivantes : *House*, *How I Met Your Mother*, *One Tree Hill*, *Lost* et *Desperate Housewives*. Ces blocs de 4 mots ont été ensuite comparés avec ceux du corpus BNC. Dans la seconde étude, Kusyk et Sockett (2012) ont réutilisé 42¹⁶⁶ blocs de 4 mots de l'étude de 2011, dans le but de vérifier leur reconnaissance par des étudiants (N=45) de l'Université de Strasbourg. Ces blocs ont été placés dans des contextes pour assurer leur compréhensibilité. Par exemple *I don't want to* donne lieu à *I don't want to go to the store* et *Can I get you* donne lieu à *Can I get you something to drink?* Dans cette deuxième étude, ces phrases ont été enregistrées par un natif américain. Ensuite, les étudiants devaient noter s'ils connaissaient (ou avaient déjà rencontré) ou non ces phrases dans des séries télévisées, et indiquer ce qu'elles signifient pour ensuite fournir leur traduction. Nous avons choisi ces 10 phrases¹⁶⁷, car en plus de porter du sens pour les participants, leur d'apparition dans les séries télévisées est fréquente. Suite aux études citées précédemment (Sockett, 2011 ; Kusyk & Sockett, 2012), la prononciation de ces phrases nous intéresse particulièrement, car elles contiennent de nombreux phénomènes phonético-phonologiques, dont les analyses nous permettront d'évaluer l'accent de nos locuteurs. Ainsi, dans ces phrases, nous retrouvons des

165 The most frequently occurring four-word clusters.

166 Les huit autres phrases ont été écartées, car elles avaient de doublures ou étaient dépourvues de verbes.

167 Elles ont été choisies de manière aléatoire.

éléments phonético-phonologiques pouvant varier suivant les accents GA et RP, tels que l'accentuation, l'assimilation, la liaison, en plus des éléments étudiés dans la première partie de la liste de mots. Nous prenons le cas de la phrase 1, *What are you doing?*, dans laquelle le premier mot se termine par un /t/ (*what*) suivi de *are*. Ici, la liaison entre ces deux mots se réalise avec une battue en GA (/wʌr_ɑ ju: 'du:ɪŋ/) et avec un /t/ en RP (/wɒt_ɑ ju: 'du:ɪŋ/). De plus, on notera la rhoticité dans la prononciation du mot *are*. On retrouvera le même cas de figure dans la phrase 5, *get out et here* (/geɪ əʊt əv hɪə/_RP, ɡeɪəʊt əv hɪr_GA). La liste complète des phrases se trouve dans le tableau ci-dessous.

Tableau 16 : les 10 phrases utilisées dans la deuxième partie de la liste de mots.

What are you doing?
I don't want to go to the store.
What the hell is he doing here?
What are you talking about?
Get out of here!
That has nothing to do with you.
You have no idea how important this is.
Come on, come on.
Give it to me please.
Can I get you something to drink?

6.4.3 Le texte

Enfin, dans la troisième partie de la liste, nous avons choisi un passage des fables d'Ésope¹⁶⁸ intitulé « *the north wind and the sun* » pour sa richesse en phénomènes phonétiques (contrastes phonémiques, par exemple *cloak* /kləʊk/ en RP et /kloʊk/ en GA). Pour Low (2014, p. 56), ce passage « *contains all the phonemes of the English language at least once* ». C'est un texte communément utilisé dans les travaux en phonétique, notamment dans les analyses des langues parlées dans le monde (apprentissage de L2, variétés et dialectes). Dans le *Handbook of the International Phonetic Association (IPA, 2005)*, il est traduit en plusieurs langues, souvent accompagné de sa transcription phonétique. La version que nous avons utilisée est la même pour le GA et le RP, selon l'association, à l'exception de deux éléments¹⁶⁹ : le double « ll » dans « *traveller* » (*traveler* en GA et *traveller* en RP) et la conjugaison du verbe « *to shine* » (*shined* en GA et *shone* en RP). Toutefois, de manière informelle, en écoutant des enregistrements du passage de

168 IPA, 2005, p. 39-44.

169 Le texte est déjà uniformisé en vue de réduire ces différences.

locuteurs américains en ligne, nous avons constaté des modifications concernant l'accent GA (tableau 3). Nous pensons que ceci pourrait signifier que la version GA proposée par L'IPA peut sembler inhabituelle pour les locuteurs du GA. Ainsi, il existerait plusieurs versions et réécritures du passage (y compris celles destinées aux enfants) selon les auteurs¹⁷⁰ (Eliot and Jacobs ; JBR version ; Townsend, etc.). Les transcriptions orthographique et phonétique en RP, en GA et la version GA en ligne (sans transcription phonétique) du texte se trouvent ci-dessous.

Tableau 17 : transcription orthographique et phonétique RP de the north Wind and the Sun (IPA, 1999, p. 44).

Transcriptions orthographique RP	Transcriptions phonétique (large)
<p>The North Wind and the Sun were disputing which was the stronger, when a traveller came along wrapped in a warm cloak. They agreed that the one who first succeeded in making the traveler take his cloak off should be considered stronger than the other. Then the North Wind blew as hard as he could, but the more he blew the more closely did the traveler fold his cloak around him; and at last the North Wind gave up the attempt. Then the Sun shone out warmly, and immediately the traveler took off his cloak. And so the North Wind was obliged to confess that the Sun was the stronger of the two.</p>	<p>ðə nɔ:θ wɪnd ænd ðə sʌn wɜ: dɪs'pju:tɪŋ wɪtʃ wɜz ðə 'strɒŋgə, wɛn ə 'træv(ə)lə kem ə'ləŋ ræpt ɪ ə wɔ:m kləʊk. ðeɪ ə'grɪ:d ðæt ðə wʌn hu: fɜ:st sək'sɪ:dɪd ɪn 'meɪkɪŋ ðə 'trævlə teɪk hɪz kləʊk ɒf fʊd bi: kən'sɪdəd 'strɒŋgə ðæn ðɪ 'lðə. ðen ðə nɔ:θ wɪnd blu: æz hɑ:d æz hi: kʊd, bʌt ðə mɔ: hi: blu: ðə mɔ: 'kləʊsli dɪd ðə 'trævlə fəʊld hɪz kləʊk ə'raʊnd hɪm; ænd æt lɑ:st ðə nɔ:θ wɪnd geɪv ʌp ðɪ ə'tempt. ðen ðə sʌn fɒn aʊt 'wɔ:mli, ænd ɪ'mɪ:diətli ðə 'trævlə tʊk ɒf hɪz kləʊk. ænd səʊ ðə nɔ:θ wɪnd wɜz ə'blɑ:dʒd tu: kən'fes ðæt ðə sʌn wɜz ðə 'strɒŋgə ɒv ðə tu:.</p>

Tableau 18 : transcription orthographique et phonétique GA de the north Wind and the Sun (IPA, 1999, p. 44).

Transcriptions orthographique GA	Transcriptions phonétique (large)
<p>The North Wind and the Sun were disputing which was the stronger, when a traveler came along wrapped in a warm cloak. They agreed that the one who first succeeded in making the traveler take his cloak off should be considered stronger than the other. Then the North Wind blew as hard as he could, but the more he blew the more closely did the traveler fold his cloak around him; and at last the North Wind gave up the attempt. Then the Sun shined out warmly, and immediately the traveler took off his cloak. And so the North Wind was obliged to confess that the Sun was the stronger of the two.</p>	<p>ðə 'nɔ:θ ,wɪnd ən ðə 'sʌn wə dɪs'pju:tɪŋ 'wɪtʃ wɜz ðə 'strɒŋgə, wɛn ə 'trævələ ,kem ə'ləŋ 'ræpt ɪ ə 'wɔ:ɪm 'kləʊk. ðe ə'grɪd ðæt ðə 'wʌn hu 'fɜ:st sək'sɪdəd ɪn 'meɪkɪŋ ðə 'trævələ 'tek ɪz 'kləʊk ,ɒf fʊd bi kən'sɪdəd 'strɒŋgə ðæn ðɪ 'əðə. ðen ðə 'nɔ:θ ,wɪnd 'blu əz 'hɑ:d əz ɪ 'kʊd, bət ðə 'mɔ: hi 'blu ðə 'mɔ: 'kləʊsli dɪd ðə 'trævələ 'fəʊld hɪz 'kləʊk ə'raʊnd ɪm; ,æn ət 'læst ðə 'nɔ:θ ,wɪnd ,geɪv 'ʌp ðɪ ə'tempt. 'ðen ðə 'sʌn 'fʌnd ,aʊt 'wɔ:ɪmli ənd ɪ'mɪ:diətli ðə 'trævələ 'tʊk ,ɒf ɪz kləʊk. ən 'səʊ ðə 'nɔ:θ ,wɪnd wɜz ə'blɑ:dʒd tə kən'fes ðæt ðə 'sʌn wɜz ðə 'strɒŋgə əv ðə 'tu:.</p>

¹⁷⁰Pour plus de détails, voir le lien <https://fablesfaesop.com/the-north-wind-and-the-sun.html>

Tableau 19 : transcription orthographique GA de *the north Wind and the Sun* (utilisée en GA) (Lapteva, 2011, p. 167).

Transcriptions orthographique GA	Transcriptions phonétique (large) ¹⁷¹
<p>The North Wind and the Sun were arguing one day about which of them was the stronger, when a traveller came upon wrapped up in an overcoat. They agreed that the one who could make the traveller take his coat off would be considered stronger than the other one. Then the North Wind blew as hard as he could, but the harder he blew, the tighter the traveller wrapped his coat around him; and at last the North Wind gave up trying. Then the Sun began to shine hot, and right away the traveller took his coat off. And so the North Wind had to admit that the Sun was the stronger than he was.</p>	<p>ðə nɔrθ wɪnd ænd ðə sʌn wɜr 'ɑrgjuɪ wʌn dei ə'baʊt wɪf'ʌv ðem wʌz ðə 'strɔŋər, wen ə 'trævələr keɪm ə'pʌn ræpt ʌp ɪn ən 'oʊvər,kəʊt. ðeɪ ə'grɪd ðæt ðə wʌn hu kʊd meɪk ðə 'trævələr teɪk hɪz kəʊt ɔf wʊd bi kən'sɪdərd 'strɔŋər ðæn ði 'ʌðər wʌn. ðen ðə nɔrθ wɪnd blu æz hɑrd æz hi kʊd, bʌt ðə 'hɑrdər hi blu, ðə 'tʌtər ðə 'trævələr ræpt hɪz kəʊt ə'raʊnd hɪm; ænd æt læst ðə nɔrθ wɪnd geɪv ʌp 'traɪɪŋ. ðen ðə sʌn brɪ'gæn tu ʃaɪn hət, ænd raɪt ə'weɪ ðə 'trævələr tʊk hɪz kəʊt ɔf. ænd soʊ ðə nɔrθ wɪnd hæd tu əd'mɪt ðæt ðə sʌn wʌz ðə 'strɔŋər ðæn hi wʌz</p>

6.5 Enregistrement des participants

Le déroulement de l'enregistrement se décompose en un temps de présentation, suivi d'une phase de lecture en vitesse normale et pour terminer, d'une phase de lecture en vitesse rapide, de la même liste. Les locuteurs étaient installés sur une chaise à quelques centimètres du microphone. Pendant la prise de sons, ils pouvaient demander une pause quand ils le voulaient afin de se rafraîchir. En raison des répétitions, il nous semble important que les locuteurs puissent prendre ces pauses pour se rafraîchir (boire de l'eau, café, jus) et prendre de la respiration, afin de garantir une meilleure qualité de la voix. Nous y avons veillé, car il est arrivé pendant le premier enregistrement, que le locuteur n'arrêtait pas de tousser, puisqu'il avait la gorge sèche. Ils pouvaient également demander à sortir si besoin. Suite à ces dispositions, les enregistrements se sont déroulés dans les conditions les plus sereines, calmes et conviviales possibles.

6.5.1 Présentation orale

En vue d'obtenir de la parole spontanée de la part des participants, qui pourrait éventuellement servir de contrôle, nous les avons invités à se présenter brièvement en anglais. La consigne donnée était de se présenter sans se stresser, de manière basique et classique, comme s'ils venaient à

¹⁷¹ Transcription automatique réalisée sur le site <https://tophonetics.com/>

discuter avec un anglophone qu'ils rencontrent pour la première fois. La majorité a fourni les éléments suivants : nom, prénom(s), âge, études, nationalité, ville de naissance et loisirs. D'autres, se sentant plus à l'aise (à l'oral) sont allés plus loin, en expliquant et en détaillant certaines de ces informations, notamment les loisirs.

6.5.2 Vitesses d'élocution

Chaque participant a lu tous les items du support, une fois en vitesse normale, puis une autre fois en vitesse rapide (Vaxelaire, 2007). Lors de la première modalité, nous leur avons demandé de parler posément et de façon naturelle. Alors que dans la deuxième modalité, ils devaient parler aussi vite que possible, mais tout en gardant une prononciation audible et bien articulée. Aucune contrainte de temps n'a été donnée, car nous sommes conscient que certaines personnes parlent plus vite que d'autres. Nous avons considéré cette différence liée aux aspects anatomo-physiologiques, parce que dans ce contexte, les locuteurs se trouvent certainement dans une situation de perturbation de parole, car l'anglais n'est pas leur L1, et qu'ils ne la maîtrisent pas forcément. Le but de cet exercice est de perturber la parole de nos participants, en appliquant la vitesse d'élocution rapide, afin de pouvoir tester la résistivité des gestes articulatoires (Vaxelaire, 1995 ; 2007). En effet, nous estimons que les productions qui sont entièrement acquises (ancrage définitif) résisteraient face à cette perturbation et resteraient identiques à celles obtenues durant la vitesse normale, pour toutes les répétitions. Cette notion de perturbation (*cf.* théorie de la viabilité) est liée et souvent utilisée dans la théorie de la viabilité (Sock, 2001 ; Sock & Vaxelaire, 2001), où des cibles articulatoires sont réajustées sous certaines conditions, en vue d'assurer leur compréhension. En ces termes, un son peut être produit différemment - d'un point de vue articulatoire-acoustique - mais tout en étant parfaitement viable (compréhensible). Toutefois, il arrive souvent de noter chez nos locuteurs, des phénomènes de réduction vocalique et de neutralisation de certains phonèmes (Exare, 2009 ; Fougeron, 2013 ; Fougeron *et al.*, 2007). En revanche, une prononciation variée dans les deux modalités peut être due à plusieurs facteurs ; nous le verrons durant la présentation des résultats. Notre rôle, qui est celui de veiller au respect des deux modalités, a été très difficile dans cet exercice particulier, car les phénomènes de liaison et de réduction vocalique sont sensibles au style de parole (Fougeron, Goldman, Dart, Guélat & Jeager, 2001). En effet, certains locuteurs parlent plus vite que d'autres. Par conséquent, il nous a

alors fallu à plusieurs reprises imposer une pause en vue de les « rappeler à l'ordre ». Ce qui n'était pas sans conséquence sur la parole de cette modalité, devenant sous-articulée et peu audible.

6.5.3 Nombre de répétitions

Pour chaque modalité (vitesse rapide et vitesse normale), les participants ont effectué cinq répétitions pour les deux premières parties de la liste. En prenant l'exemple de la première partie de la liste, tous les 24 mots sont lus une première fois. Puis une deuxième fois, puis une troisième, jusqu'à la cinquième et ainsi de suite avec la deuxième partie (les phrases). En raison de sa longueur, seulement deux répétitions sont effectuées pour le texte. Le nombre total de répétitions par locuteur s'élève ainsi à $284(((24 \times 5) + (10 \times 2) + 2) \times 2)$, soit 5680 pour tous nos locuteurs. Il est à noter que les phrases et le texte sont plus compliqués, car ils contiennent plusieurs mots et donc plusieurs éléments à analyser. L'ordre des phrases est randomisé pour chaque répétition et pour chaque locuteur, afin d'éviter l'effet de liste. Étant donné qu'un même locuteur, sur plusieurs répétitions, ne produira jamais des sons identiques, les répétitions nous serviront à délimiter les zones de productions possibles¹⁷² et de pouvoir obtenir une moyenne selon les valeurs acoustiques. Ainsi, les cinq répétitions nous permettent non seulement d'obtenir un nombre relativement suffisant d'occurrences, afin de calculer une moyenne de valeurs, mais aussi de prendre en compte la variabilité dans les productions des locuteurs.

6.5.4 Recodage et anonymisation de données

En vue de faciliter les importations et les analyses dans Praat (Boersma & Weenink, 2017), les fichiers audio ont été sectionnés à l'aide du logiciel Audacity¹⁷³, selon chaque partie de la liste, la vitesse d'élocution et les participants. Nous rappelons que les enregistrements ont été effectués sur un seul fichier audio, par locuteur. Ensuite, avant le commencement des analyses des données recueillies, nous avons codifié les noms et prénoms de participants en adoptant simplement les expressions : locuteur 1, locuteur 2, locuteur 3, ..., locuteur 20. Les fichiers ainsi sectionnés ont été renommés pour faciliter les repérages sans recourir aux personnes, même si nous pouvons les

¹⁷² En termes d'une zone acoustique au sein de laquelle les productions orales sont viables / compréhensibles. Cf théorie de la viabilité de Sock, 2001.

¹⁷³ Logiciel gratuit d'enregistrement et d'édition audio : <https://www.audacityteam.org/>

identifier (sur les questionnaires). Ce procédé sert également à préserver leur anonymat pour le lectorat plus large.

6.6 Segmentation et mesures acoustiques des sons cibles

Comme nous l'avons évoqué précédemment, les mesures acoustiques sont réalisées à l'aide du logiciel Praat (Boersma & Weenink, 2017) version 6.0.28 (23 mars 2017), fonctionnant avec une extension Fricative Script¹⁷⁴ (voir dossier Script_Praat en Annexe B en ligne sur Seafile). Ces mesures portent sur les sons cibles se trouvant dans les mots et qui sont classés par catégories. Les éléments segmentés et mesurés pour chaque son cible, concernent notamment les formants (les quatre premiers), la fréquence fondamentale, la durée et l'intensité (surtout concernant les voyelles). Afin de faciliter la visibilité des éléments étudiés et de garantir la fiabilité de nos analyses, nous avons modifié les paramètres par défaut du logiciel Praat : le nombre de formants est limité à 5 pour couvrir l'étendu des voix « féminines » et « masculines » (maximum 5500 Hz) et la taille de l'interface est réglée à 0,04 à la place d'une valeur initiale de 0,025 (Zaouali, 2019). Nous fournissons si possible des informations articulatoires afin d'expliquer les mécanismes de production des sons cibles, sur la base d'éléments acoustiques notables (Sock, 1998). Lorsque les données acoustiques ne permettent pas l'identification précise d'un son, nous complétons par une série d'écoutes, validée par des phonéticiens. Dans les paragraphes suivants, nous détaillons ces sons cibles, selon les mots dans lesquels ils se trouvent (catégorie) et les mesures les concernant (d'abord les éléments segmentaux puis suprasegmentaux). Les différences de prononciation entre le GA et le RP que nous évoquons reposent sur des évidences tirées de la littérature existante. Pour chaque point évoqué, nous donnons un bref rappel historique sur les divergences, avant de détailler nos procédures d'analyse. Nous donnerons également des exemples d'images illustrant les différences de prononciation en fonction des analyses acoustiques (oscillogrammes et spectrogrammes). Une liste plus exhaustive sur les illustrations et les tableaux contenant des valeurs des analyses acoustiques se trouve dans les annexes (Annexe A 1, p. 1-35).

¹⁷⁴ Tous les scripts que nous avons utilisés sont en ligne et proviennent de Cédric Gendrot : http://lpp.in2p3.fr/Equipe/cedric_gendrot/scripts.htm

6.6.1 Éléments consonantiques

Sans aller dans le détail acoustique (mesures de VOT, VTT¹⁷⁵ par exemple), nous avons analysé trois éléments consonantiques les plus saillants faisant la différence entre la RP et le GA : le /t/ intervocalique (*city, party et vitamin*), la rhoticité (*more, organisation, dark, party, bird, actor et leisure*) et le yod-dropping (*tune et new*). Nous avons choisi plus de mots contenant le /r/ en raison des différents environnements phonétiques, soit les sons qui l'entourent (<a, i, o, u> pour les voyelles et <g, k, t, d, s> pour les consonnes).

6.6.1.1 La prononciation du /t/ intervocalique

Lorsque l'on parle de l'accent américain, l'une des premières choses à laquelle nous pensons est probablement la prononciation du phonème /t/, qui sonne comme un /d/ (Ladefoged & Johnson, 2011, p. 34). En effet, il est vrai que dans certains contextes, le /t/ intervocalique (entre deux voyelles) devient un *tap* ou un *flap* en anglais américain (Ladefoged & Maddieson, 1996). Soit /V₁/ une voyelle accentuée, /V₂/ une voyelle inaccentuée et /t/ une plosive alvéolaire sourde, le contexte du /t/ intervocalique présentant une structure /V₁tV₂/ devient un tap¹⁷⁶ (Ladefoged & Johnson, 2011, p. 46). Par exemple, cette structure se retrouve dans les mots et les liaisons de mots suivants : *water, letter, pretty, what is, get over*, etc ('wɑ:tə, 'lɛtə, 'prɪri, wɑt_ɪz, gɛt_ʊvər). Nous rappelons que lorsque V₂ est accentuée comme dans les mots *italic, pretend, attach* (/ɪ'tæɪlɪk/, /prɪ'tɛnd/ et /ə'tæɪf/), le *tap* ne s'applique pas. D'autres conditions phonétiques donnent également lieu à l'apparition d'un *tap* dans l'accent américain. Il s'agit notamment des contextes suivants : (1) /rV/ (exemples : *garter_* /'gɑrə/, *barter_* /'bɑrə/), (2) /lV/ (exemples : *faculty_* /'fækəlri/, *faulty_* /'fɔlri/) et (3) /Vt/ (exemples : *battle_* /'bæɪl/, *little_* /'lɪl/, etc.) où V étant une voyelle en position inaccentuée. Cependant, il est important de souligner, sans entrer dans le détail phonologique et articulatoire¹⁷⁷, que les conditions évoquées *supra* ne produisent pas toujours un *tap*. Des allophones peuvent être réalisés à la place du *tap* : un coup de glotte (*important, mountain, certain, button_* /ɪm'pɔɪtənt, 'maʊntən, 'sɜːən, 'bʌtən/) ou tout simplement un /t/ silencieux au profit de la consonne d'avant, habituellement dans le cas de /n/ (par exemple *twenty*,

¹⁷⁵ Voice Onset Time (délai d'établissement du voisement), Voice Termination Time (délai d'arrêt du voisement).

¹⁷⁶ Une battue.

¹⁷⁷ Pour plus de détails, voir (Ladefoged & Johnson, 2011 ; Ladefoged & Maddieson, 1996 ; Roach, 2009).

wanted, *center* et *winter* /'twɛni, 'wʌnəd, 'sɛnə et 'wɪtə/ et les mots commençant par *inter* (*international*, *interrupt* /,ɪnə'næʃənəl, ,ɪnə'ɪɹpt/). Dans notre travail, nous nous focalisons uniquement sur le premier cas, soit le contexte où /t/ intervocalique (/VɪtV₂/, V₁ accentuée et V₂ inaccentuée).

Du point de vue articulatoire, un *tap* ou un *flap* fait partie de la famille des rhotiques et est un son produit par un rapide contact (de l'ordre de 20 ms) entre un articulateur mobile (ici la pointe de la langue) et un articulateur fixe (ici l'alvéole) (Ladefoged & Johnson, 2011, p. 175 ; Jauriberry, 2016). Nous utilisons les deux termes de manière interchangeable, puisque les explications en vue de les différencier restent ambiguës. Il faut aussi remarquer que dans le tableau des symboles phonétiques de l'IPA, les deux mots sont utilisés comme des synonymes (*tap* ou *flap*) et sont représentés par un seul symbole /ɾ/¹⁷⁸. De surcroît, il semblerait que ces sons possèdent un nombre d'allophones et soient réalisés de plusieurs manières. C'est peut-être la raison pour laquelle, du point de vue perceptif, nous entendons parfois un /d/, parfois un /ɾ/¹⁷⁹ ou des fois, des /t/ et /d/ rétroflexes (/ɽ/ et /ɖ/). En effet, à notre niveau, la seule hypothèse que nous pouvons avancer, concernant la différence de réalisations, est le fait que la direction du mouvement de contact varie selon le type de la voyelle située après le /t/. Ainsi, nous estimons que la pointe de la langue frappe les alvéoles en direction des dents, dans le cas des voyelles antérieures (*pretty*, *city*, *what if*) ou en direction du palais, en présence des voyelles centrales ou postérieures (*better*, *ghetto*). Quoiqu'elle paraisse évidente, cette hypothèse nécessite des recherches en phonétique articulatoire (IRM par exemple) pour être confirmée.

Dans notre liste, concernant le /t/ intervocalique, nous avons analysé trois mots, bien que d'autres /t/ intervocaliques aient été identifiés (par exemple dans *it is* /ɪrɪz/). Il s'agit de *city*, *vitamin* et *party*. Lors de la segmentation de ces mots, nous avons isolé le son produit entre les deux voyelles. La segmentation du troisième mot était particulièrement compliquée, car celui-ci présente un cas spécial, à cause de la présence de <r>, qui est parfois confondu avec le *tap* sur le signal (/ˈpɑrri/). Mais, lorsque ces deux sons sont séparés, nous segmentons à partir de la fin du /r/ (disparition de sa structure formantique clairement définie) jusqu'au début de la voyelle suivante

178 Nous n'utilisons que ce symbole même si le système américain utilise plutôt le /t/ avec un signe diacritique du voisement en-dessous de celui-ci (/ɽ/).

179 C'est pourquoi certains chercheurs le placent parmi la famille des rhotiques (Ladefoged & Johnson, 2011 ; Jauriberry, 2016).

(apparition d'une structure formantique clairement définie de cette voyelle). Nous avons alors mesuré la durée des segments vocaliques et consonantiques, et analysé le spectre du signal acoustique, en prenant en compte la présence ou non d'un silence acoustique, d'une explosion, d'une aspiration, etc. L'objectif est d'identifier un /t/ « normal » d'une battue ou d'un autre allophone, grâce aux mesures du VTT, de la tenue consonantique et du VOT, sans forcément entrer dans les analyses temporelles détaillées (Sock, 1998).

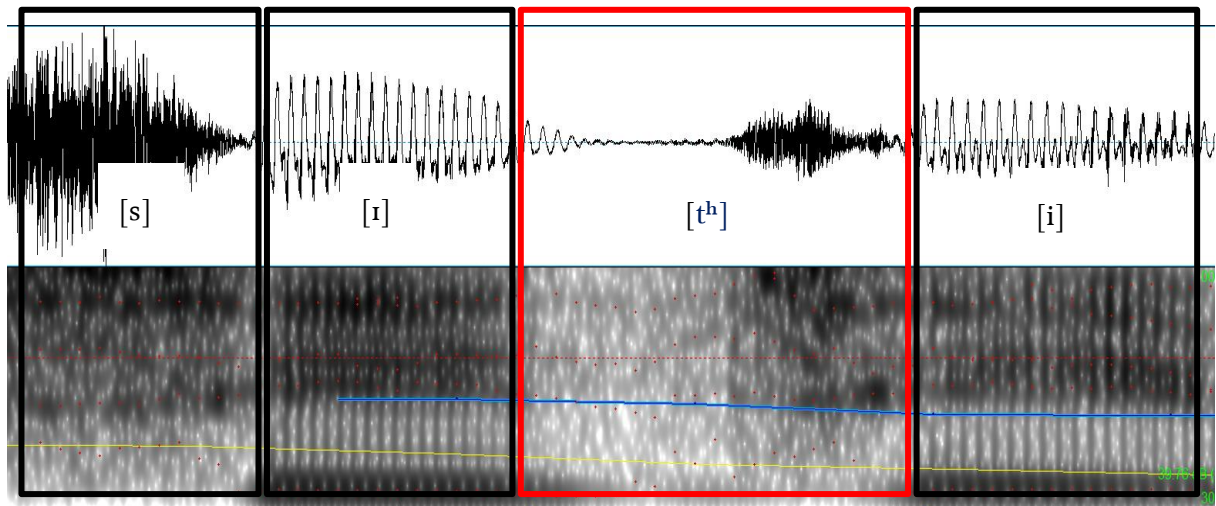


Figure 20 : un exemple de prononciation du /t/ aspiré dans le mot city (/sɪ[tʰ]i/) (encadré rouge)

6.6.1.2 La rhoticité

Le mot « rhotique » provient du grec ancien *rhô* qui représente la lettre <r> de l'alphabet latine, dont le symbole est représenté par un ρ . La rhoticité représente une manière bien spécifique de prononcer le graphème <r> partout dans les mots. On parle alors de rhotacisation (ou rétroflexion) et les sons (ou phonèmes) ayant subi ces transformations sont dits rhotacisés. L'anglais américain, le GA¹⁸⁰ pour être plus précis, est connu comme étant une variété rhotique de l'anglais (Ladefoged & Johnson, 2011). Il peut alors s'agir de la prononciation soit d'un /r/ en position prévocallique, c'est-à-dire qui vient avant une voyelle dans une même syllabe (*red, road, arrive, etc.*). Soit d'un /r/ post-vocalique¹⁸¹ où le /r/ figure après une voyelle dans une syllabe (*part, dark, car*). Deux annotations sont possibles dans les transcriptions phonétiques : la voyelle précédée d'un r (/er/) ou le signe diacritique du r greffé à la voyelle (/ɚ, ɝ/). Certains auteurs justifient l'utilisation séparée de ces deux systèmes de représentation pour rendre compte des différentes réalisations

¹⁸⁰ Et aussi l'accent écossais et irlandais (Hughes, Trudgill, & Watt, 2012, p. 63).

¹⁸¹ Éventuellement pré-consonantique.

articulatoires (Ladefoged & Johnson, 2011, p. 41). De ce fait, le premier est utilisé lorsqu'il y a une nette démarcation entre la voyelle et le /r/ (*marble* /'mɑ:bəl/, *where* /wer/, *nor* /nɔ:r/) et le deuxième lorsque la voyelle et le r sont produits simultanément en une seule et même entité (*bird* /bɜ:d/, *fur* /fɜ:/, *nerd* /nɜ:d/ ainsi que les terminaisons en *er*, *or* : *looser* /lu:sə/, *actor* /'æktə/). Par contre, dans une variété non-rhotique comme le RP¹⁸², le /r/ post-vocalique n'est pas prononcé mais uniquement le /r/ pré-vocalique comme dans *red*, *horrible*, *arrange*. Il est plutôt remplacé par un système d'allongement de la voyelle dans les transcriptions phonétiques. Les deux points sont alors ajoutés à la voyelle pour illustrer cet allongement (/hɑ:d/). En ce qui concerne les terminaisons en <er>, <or>, etc. le RP utilise un schwa (/ˈlu:sə/, /'æktə/). Quant à Durand *et al.* (2015, p. 130), la notation de la voyelle suivie de /r/ (əɹ) serait plus appropriée mais admettent néanmoins que d'autres variantes soient également acceptables (ex. *nurse* = /nəs/, /nɪs/, /nɜ:s/ ou /nəɪs/ etc.).

Nous avons analysé les mots *more*, *party*, *bird*, *leisure*, *dark*, *actor*, et *organisation* dans le but de détecter la présence ou non du /r/ vocalique dans le signal acoustique, lorsque celui-ci est présent dans l'orthographe. Selon Ladefoged & Johnson, 2011, p. 230 ; cette présence du /r/ est attestée sur le signal acoustique par une chute¹⁸³ de F₃ (éventuellement du F₄). Toujours, selon eux (p.231), cette fluctuation formantique (diminution) commence dès la production d'une voyelle « colorée en r¹⁸⁴ », comme dans *bird* et *dark*. Néanmoins, il semblerait que la baisse de F₃ dans la perception des rhotiques ne soit pas une condition *sine qua non* et « qu'un F₃ abaissé ne peut pas être retenu comme la propriété qui unifierait le groupe de rhotiques » (Jauriberry, 2016, p. 65).

182 Et aussi en Australie, Nouvelle Zélande, Afrique du Sud, Afrique de l'Ouest, le sud des États-Unis (New York, new England) et les Caraïbes (Hughes, Trudgill, & Watt, 2012, p.63).

183 (Stuart-Smith *et al.*, 2015) ont montré qu'il peut également y avoir une hausse de F₃ pour les /r/ post-vocaliques.

184 /r/-colored vowel.

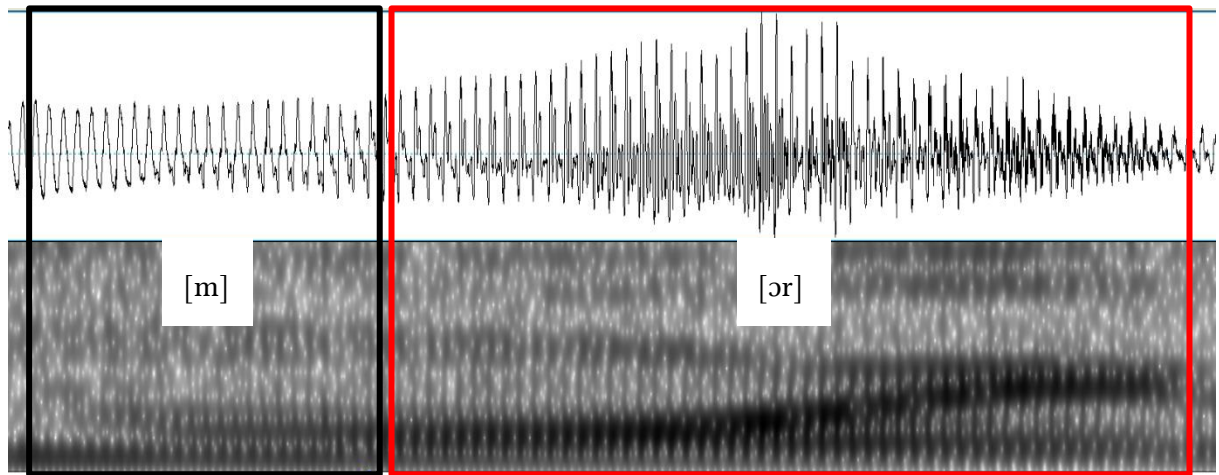


Figure 21 : un exemple de production rhotique dans le mot *more* (/mɔr/) (encadré rouge)

6.6.1.3 Le *yod-dropping*

En RP, la majorité des mots contenant les graphèmes <ew> et les variantes¹⁸⁵ de <u>, dans une syllabe accentuée, sont prononcés en [ju:] (Cruttenden, 2014, p.84). Par exemple, la transcription phonétique des mots *dupe*, *due*, *knew*, *new*, *tune*, *tutor*, *tuesday*, *tuna*, etc. sera /dju:p/, /dju:/, /nju:/, /tju:n/, /'tju:tə/, /'tju:zdeɪ/, /'tju:nə/, etc. Il y a pourtant quelques exceptions avec les mots *rude* (/ru:d/), *rule* (/ru:l/), *clue* (/klu:/), *Blew* (/blu:/), *jew* (/dʒu:/) et *crew* (/kru:/) (Beal, 2014). En GA en revanche, la semi-voyelle /j/ a disparu de cette structure. L'abandon du /j/ fait que la prononciation tombe directement sur le /u:/ (/ju:/ devient ainsi /u:/). Les premiers exemples ci-dessus deviennent cependant /du:p/, /du:/, /nu:/, /tu:n/, /'tu:tə/, /'tu:zdeɪ/, /'tu:nə/. Tout comme en RP, il existe également en GA des exceptions, soit la présence du /j/ et celles-ci adviennent après une consonne (bi)labiale, vélaire et après /f/ et /v/ (*beauty*, *pure*, *cute*, *few*, *view* etc. (Wells 1982 : 247).

Le phénomène que nous venons de décrire est connu sous le nom de *Yod-dropping* ou *Yod-deletion*¹⁸⁶ (Wells 1982 ; Glain, 2012 ; Beal, 2014). Il est à noter que les réalisations sont assez variables et seraient assujetties à l'environnement phonétique, aux items lexicaux, à la variété de l'accent et aux locuteurs. Aussi, nous tenons à signaler que le *Yod-dropping* n'est pas uniquement caractéristique au GA. Les deux réalisations sont attestées en Grande Bretagne et aux États-Unis (Brulard *et al.*, 2015). C'est ainsi que dans l'est de l'Angleterre, une surgénéralisation est attestée ; soit une disparition complète du /j/ dans tous les mots, y compris les exceptions citées ci-dessus.

¹⁸⁵ eu, ui, ue comme dans *zeugma*, *suit* et *suet* respectivement.

¹⁸⁶ Yod est la dixième lettre de l'alphabet hébreu qui représente l'alphabet latin <j> .

Pew, beauty, view, few, queue, music, human sont prononcés respectivement comme suit /pu:/, /'bʊ:ti/, /vʉ:/, /fʉ:/, /ku:/, /'mu:zik/ et /'hu:mən (Hughes *et al.*, 2012, p. 69 ; Roach, 2009c, p. 162).

Dans notre analyse, nous avons cherché à détecter la présence ou non du yod dans les mots *tune* et *new*. Sa combinaison avec le /u/ (/ju/) devrait laisser apparaître sur le spectrogramme, une structure formantique similaire à celle d'une diphtongue (en RP). À contrario, son absence (/u/) présentera une structure analogue à celle d'une voyelle « pure » (monophthongue) et dont les caractéristiques sont généralement davantage retrouvées en GA. Outre ces deux cas de figure, il est possible de constater une palatalisation dans les mots comme *tune, dune, etc.*, qui résulterait d'une coarticulation, plus précisément une affrication entre le /t/, le /d/ et le /j/, formant ainsi un /tʃ/ et un /dʒ/ (/tʃu:n/ au lieu de /tju:n/ et /dʒu:n/ au lieu de /dju:n/). L'amplitude et la durée sur le signal acoustique peuvent être révélatrices du type de réalisation (Glain, 2012). Aujourd'hui, les dictionnaires LPD et EPD prennent en considération ces diverses prononciations dans les transcriptions phonétiques avec, si possible, le pourcentage de leur réalisation selon les locuteurs et leur accent.

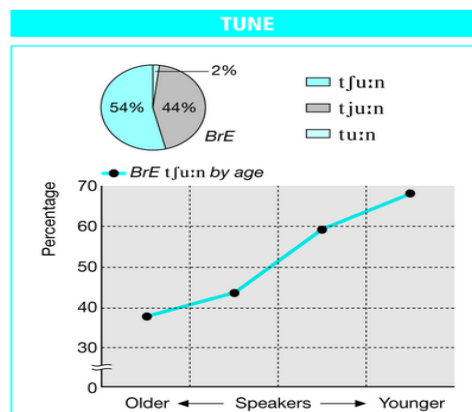


Figure 22 : différentes formes de réalisations du mot *tune* en anglais britannique dans le LPD (Wells, 2008).

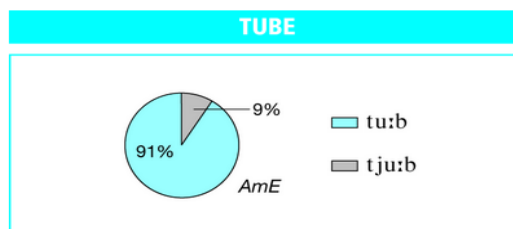


Figure 23 : pourcentage de réalisations du mot *tube* en anglais américain dans le LPD (Wells, 2008).

6.6.2 Éléments vocaliques

Les aspects phonético-phonologiques qui diffèrent substantiellement entre la RP et le GA résident dans le système vocalique (Melchers & Shaw, 2013, p. 18). Les paragraphes suivants seront consacrés à quelque unes de ces voyelles, notamment les plus saillantes. Il s'agit des voyelles <o> de *LOT* (*hot* et *top*), <o> et <oa> de *GOAT* (*joke* et *road*), <i> et <y> de *SIT* (*vitamin*, *dynasty*, *organisation* et *legalisation*), de <ei> et <e> (*leisure* et *evolution*) et de la nasalisation vocalique (*hand* et *band*).

6.6.2.1 La prononciation de la voyelle de *LOT*

La voyelle postérieure ouverte arrondie¹⁸⁷ /ɒ/ n'existe pas en GA. Les mots dans lesquels le RP utilise cette voyelle sont plutôt prononcés avec un /ɑ/ ou parfois avec un /ɔ/ en GA (Cruttenden, 2014, p. 87). Alors, les mots de la catégorie lexicale *LOT/CLOTH* créé par Wells comme *lot*, *hot*, *top*, *spot*, *slot*, *dog*, *coffee* sont prononcés /lɑ:t/, /hɑ:t/, /tɑ:p/, /spɑ:t/, /slɑ:t/ et /'kɑ:fi/ en GA au lieu de /lɒt/, /hɒt/, /tɒp/, /spɒt/, /slɒt/, /dɑg/ et /'kɒfi/ en RP. Nous avons pris deux de ces mots (*top* et *hot*) et y avons analysé la voyelle cible. Les mesures formantiques, qui nous ont également permis de déduire des informations articulatoires, ont été capitales dans la détection de la voyelle produite par nos participants. Sur le plan articulatoire, ces deux voyelles partagent des mêmes critères de production à savoir la postériorité et l'ouverture, à l'exception de l'arrondissement des lèvres. /ɒ/ est arrondie alors que /ɑ/ ne l'est pas. Nous devons nous attendre à ce que, pour /ɒ/, l'arrondissement entraîne une baisse des formants, notamment de F2 et F3 (Dohls, 2010, p. 108).

6.6.2.2 La prononciation de la voyelle de *GOAT*

Certains mots anglais contenant les graphèmes <o> ou <oa> connaissent aussi une différence de prononciation notable, dans l'accent GA ou dans l'accent RP. Considérons la suite de mots suivants *no*, *go*, *hotel*, *load*, et *goat*; alors que le RP utilise le phonème /əʊ/, le GA opte pour /oʊ/. La particularité qui distingue ces deux diphtongues réside dans la production de la voyelle de départ. On relèvera, que dans la transcription phonétique de ces graphèmes, le premier commence par la voyelle centrale /ə/ et que le deuxième débute par la voyelle postérieure mi-fermée arrondie /o/. Conformément, nous avons considéré les deux mots *joke* et *road*. Les structures

¹⁸⁷ Légèrement arrondie (Brulard, Carr, Durand, & Navarro, 2015, p. 100 ; Roach, 2009a, p. 14).

spectrographiques et formantiques visibles dans le signal acoustique, ainsi que les valeurs de ces formants nous ont permis de faire la lumière sur le type de voyelle commençant la diphtongue. Malgré la variabilité inter-locuteur dans la réalisation du schwa (Fougeron *et al.*, 2007), la structure formantique de cette voyelle centrales, peu importe le locuteur¹⁸⁸, suivrait la formule 1X-3X-5X. Cela signifie que les valeurs de F3 font cinq fois celles de F1, et celles de F2 en font le triple. Concrètement, cela veut dire que si pour un locuteur, le F1 est de 500Hz, le F2 en ferait 1500 Hz et le F3 2500 Hz (Ladefoged & Johnson, 2011, p. 193). Tandis que, pour un /o/, les deux premiers formants seront rapprochés dans les basses fréquences (F1 aux alentours de 350 et 400 Hz, F2 entre 800 et 900 Hz) et un F3 qui se retrouve dans les hautes fréquences (F3 ~2200-2700 Hz)¹⁸⁹ (Calliope, 1989, p. 84).

6.6.2.3 La prononciation des graphies <i> et <y> dans certains mots

Pour savoir quelle réalisation phonétique a été effectuée pour les graphies <i> et <y> dans les mots *vitamin* et *dynasty*, nous avons examiné la voyelle de la première syllabe, qui est prononcée différemment dans les deux variétés (ici entre le /v/ et le /t/, le /d/ et le /n/). Nous trouvons que la monophthongue /ɪ/ de KIT est utilisée en RP (/ˈvɪtəmm/, /ˈdɪnəsti/) alors que la diphtongue /aɪ/ de PRICE (/ˈvaɪrəmm/, /ˈdaɪnəsti/) est plutôt dominante en GA (Brulard, Carr, Durand, & Navarro, 2015, p. 132). L'identification de la voyelle produite s'est déroulée en deux temps. Premièrement, il s'agit de détecter s'il y a présence d'une monophthongue ou d'une diphtongue. Dans ce cas, nous avons mesuré la durée et relevé la structure formantique de la voyelle. Dans des conditions normales¹⁹⁰, une monophthongue devrait avoir une durée plus courte et une structure formantique plus stable qu'une diphtongue. Cette dernière serait plus longue et devrait présenter une structure dont la trajectoire des trois premiers formants reste très variable. Deuxièmement, afin de déterminer avec plus de précision la qualité des voyelles produites, nous avons comparé les valeurs formantiques prélevées avec celles qui ont été indiquées dans la littérature (Ladefoged & Johnson, 2011, p. 193). Ce procédé permet « d'affirmer » que la voyelle produite est bien un [ɪ] et non un [i:]

188 Que ce soit une femme, un homme ou un enfant. Pour plus de détails, voir (ladefoged, 1996, p.193).

189 Ces chiffres sont des exemples et donc ne sont pas généralisables ; les valeurs formantiques varient selon les locuteurs.

190 Ceci n'est pas aussi simple dans la réalité : une diphtongue peut être monophthonguée et visa versa.

un [e] ou un [ə] pour la monophthongue ; un [aɪ] et non un [aʊ], un [eɪ] ou un [ɔɪ] pour la diphtongue.

Toujours concernant la différence de prononciation de <i> en RP et en GA, la même analyse acoustique évoquée précédemment est effectuée sur les mots *organisation* et *legalisation*. Les mesures portent principalement sur la prononciation du <i> dans l'antépénultième syllabe, mais nous en avons profité pour regarder la présence de rhoticité dans le mot *organisation*. En effet, hormis les différences d'orthographe (le s et le z), les deux types d'accent utilisent une prononciation distincte dans les terminaisons <-isation>/<-ization>. Pour la RP, ces deux mots prononcés /,ɔ:gənəɪ'zeɪʃ(ə)n/ et /,li:gələɪ'zeɪʃ(ə)n/ alors que chez les locuteurs du GA, on retrouve plutôt /,ɔrgənɪ'zeɪʃ(ə)n/ et /,li:gəlɪ'zeɪʃ(ə)n/. Les formes réduites /,ɔrgənə'zeɪʃ(ə)n/ et /,li:gələ'zeɪʃ(ə)n/ sont autant souvent utilisées en GA. Nous avons opté pour le <s> à la place du <z> dans ces mots afin de garder la cohérence dans l'utilisation du RP comme modèle orthographique, mais encore pour éviter de suggérer ou d'imposer la prononciation américaine. Nonobstant, il faut souligner que si la forme britannique peut également amener à prononcer avec la RP, ceci est moins flagrant, puisque, les deux prononciations ont été attestées en RP (Modiano, 1996, p. 108).

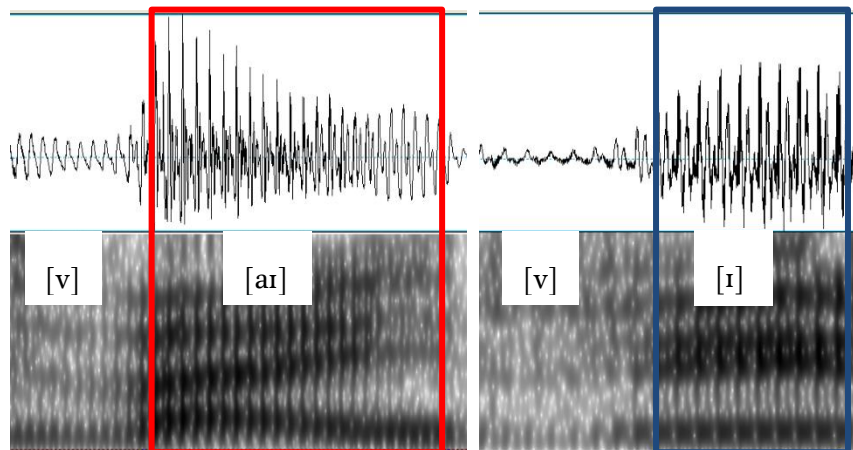


Figure 24 : un exemple de pronociation de <i> dans la syllabe <vi> du mot *vitamin* : [ai] (/ˈv[ai]təmin/) dans l'encadré rouge et [ɪ] (/ˈv[ɪ]təmin/) dans l'encadré bleu.

6.6.2.4 La pronociation de la graphie <ei> et <e>

Nous avons analysé la première voyelle des mots *leisure* et *evolution* en vue de détecter le type de voyelle qui est produite. Tandis que la pronociation /ˈleɪzə/ et /i:vəˈljʊ:fən/¹⁹¹ est dominante en RP, /ˈli:zə/ et /evəˈlu:fən/ est plus souvent rencontrée en GA. En effet, il semblerait que les deux pronociations soient indiquées au sein des deux accents. Seules les variantes dominantes sont prises en considération dans notre travail. Remarquons l'apparition de deux éléments phonétiques dans chacun des mots : /e/ vs /i:/ et /ə/ vs /ə:/ dans le mot *leisure* et /i:/ vs /e/ et /ju:/ vs /u:/ dans *evolution*. Pour ces mots, les mêmes procédés d'analyse cités précédemment ont été appliqués, à savoir la différence entre /e/ et /i:/, et le *yod-dropping*.

6.6.2.5 Nasalisation vocalique

Il semblerait que les Américains parlent du nez, à savoir qu'ils nasalisent les sons (notamment les voyelles). Or, de manière générale, la nasalisation vocalique (les voyelles nasales) n'existe pas en anglais, contrairement au français qui en possède quatre comme dans les mots *pain*, *brun*, *blanc* et *long*, (/ɛ̃, œ̃, ɑ̃, ɔ̃/ respectivement). D'où provient cette impression de nasalisation perçue chez les Américains ? Viel (1981) nous fournit un élément de réponse dans la citation suivante :

Les Britanniques croient souvent que les Américains ont une pronociation « nasale ». Il y a bien une légère nasalisation des sons en américain, et surtout en américain substandard, mais

¹⁹¹ Selon le LPD, 85% des Britanniques utiliseraient /i:vəˈljʊ:fən/ contre 15% pour /evəˈljʊ:fən/. Nous n'avons pas les chiffres quant aux différentes pronociations de *leisure* en GA mais deux autres pronociations, minoritaires, sont aussi attestées. Il s'agit de /ˈleɪzə/ et /leɪzə/.

l'effet produit aux oreilles des Britanniques est dû principalement à une forte tension pharyngale¹⁹² (Viel, 1981, p.34-35).

Du point de vue articulatoire, un son pharyngalisé est produit avec une rétraction de la racine de la langue vers la paroi du pharynx, créant un rétrécissement à cet endroit (Ladefoged & Maddieson, 2009). La création de ce conduit pharyngonasal est à l'origine d'une tension pharyngale, causée par la pression subglottique qui fait décoller le vélum. La pharyngalisation entrainerait aussi, dans une certaine mesure, une nasalisation. C'est probablement cette impression de nasalité dont parle Viel, puisque, plus la langue recule vers le pharynx, plus l'air s'accumule au bout du vélum et finit par s'échapper par la cavité nasale (Feng & Castelli, 1996). La nasalité implique un geste de décollage / abaissement du vélum, qui laisse passer l'air à travers la cavité nasale. Lorsque le vélum est rabattu contre la paroi du pharynx (blocage), l'air passe uniquement par la cavité buccale et le son est dit oral. Réciproquement, lorsqu'il est décollé de la paroi, une partie de l'air s'échappe par le nez et le son est qualifié de nasal (Marchal, 2007, p. 83). On distingue ainsi des voyelles nasales (ɛ̃, œ̃, ã, õ̃) et des consonnes nasales (/m/ de **malade**, /n/ de **note**, /ŋ/ de **parking** et /ɲ/ de **montagne**) (Calliope, 1989, pp. 89-90). Notons de plus que /d/ et /n/ ainsi que /m/ et /p/ partagent le même lieu d'articulation excepté que dans la réalisation de /n/ et /m/, le vélum est abaissé. Puisqu'en anglais il n'existe pas de voyelle nasale, la nasalisation vocalique qui est associée à l'anglais américain résulterait de l'influence de l'environnement phonétique (Janicki, 1977, pp. 34-35). Dès lors, une voyelle orale a tendance à être nasalisée lorsqu'elle est entourée de consonnes nasales (citées *supra*). Ceci s'explique par deux phénomènes. Premièrement, le vélum ne se colle pas au pharynx pendant la production de la voyelle mais reste décollé jusqu'à la production de la consonne nasale. On parle alors de nasalisation anticipée (/an/ devient /ã/). Deuxièmement, la nasalisation dite retardée intervient lorsqu'après la production d'une consonne nasale, le vélum, « n'ayant pas le temps » de s'accoler à la paroi du pharynx, reste décollé (/na/ devient /nã/). La nasalisation, qu'elle soit anticipée ou retardée, est plus gouvernée par des aspects phonétiques (notamment articulatoires) que par des règles phonologiques (Cohn, 1993) et n'est pas contrastive dans ce cas¹⁹³. Par conséquent, elle n'induit pas d'opposition phonologique en anglais, autrement, il n'y aurait pas de changement de sens. Ainsi pour le mot

192 C'est nous qui soulignons.

193 Certains auteurs comme Malécot (1960) affirment que la nasalisation entraîne une opposition phonologique en anglais notamment dans les paires /cap-camp/, /hit-hint/, /cat-can't/, /buck-bunk/, /lick-link/, /back-bank/. Pour nous, ce n'est pas la même situation, car il s'agit plutôt de la présence d'une consonne nasale.

kindness par exemple [kamdn̩s] ne s'oppose pas à [kãmdn̩s]. Au pire des cas, ce qui risque d'arriver est la prononciation d'un mot qui n'existe pas en anglais, ou qui ne soit pas reconnue par des natifs, mais qui au contraire, peut l'être pour des non-natifs en raison des habitudes de prononciation communes. Cependant, il nous faut mentionner que la nasalisation vocalique est contrastive dans des langues comme le français, où elle permet de distinguer deux phonèmes dans des paires minimales (paix /pɛ/ vs. pain /pɛ̃/, las /la/ vs. lent /lã /).

Pour savoir si nos locuteurs nasalisent les voyelles anglaises ou non dans le contexte décrit ci-dessus, nous avons examiné les mots *hand* et *band*. Les mesures acoustiques portent principalement sur les quatre premiers formants. Contrairement à une voyelle orale, une voyelle nasalisée devrait présenter une augmentation de F3 et F4 et une diminution de l'amplitude de F1, avec les deux premiers formants autour de 300 et 1000 Hz respectivement¹⁹⁴ (Feng & Castelli, 1996). Aussi, puisque la nasalisation entraîne une division du flux d'air provenant des poumons, dont une partie passe par les cavités nasales, le débit d'air qui sort par la bouche devrait être moindre par rapport à une voyelle orale. Ceci s'observe sur l'amplitude de signal acoustique (Oostendorp, Ewen, Hume, & Rice, 2011, pp. 619, 1419). L'aspect du spectrogramme indique visiblement si l'élément qui se situe entre la consonne de début et celle de départ, forme un ensemble ou deux entités séparées. Une voyelle nasalisée présente un bloc entier et dans le cas échéant, la voyelle et la consonne (/æ/ et /n/ dans notre cas) sont clairement démarquées.

6.6.3 Prosodie – Intonation

La littérature relative aux différences prosodiques entre le GA et la RP reste variable (Brulard, Carr, Durand, & Navarro, 2015, p. 138). Alors que certains phonéticiens affirment qu'il existe de fortes différences entre ces deux accents, d'autres en pensent le contraire (Janicki, 1977, p. 45). Toutefois, nous pouvons retenir au moins deux différences, sur lesquelles les chercheurs s'accordent. Dans un premier temps, il semblerait que les schémas intonatifs (la mélodie des énoncés) des Américains soient assez monotones et que la RP utilise davantage une variabilité intonative beaucoup plus riche (*rise and fall*). Dans un deuxième temps, il paraît que les Américains utilisent très souvent les *high rise*¹⁹⁵ : un schéma intonatif ascendant à la fin des

194 Encore une fois, nous signalons que ces valeurs sont variables selon chaque locuteur.

195 Un phénomène souvent connu sous le nom de *Uptalk*.

énoncés affirmatifs. Nous rappelons que le *high rise* est normalement employé dans les énoncés interrogatifs. Dans la phrase « *Do you speak English?* », le mot *English* reçoit la marque de l'intonation montante (*rise*), tandis que dans « *I do speak English* », le dernier mot connaît un schéma plutôt descendant (*fall*). En plus des caractéristiques évoquées précédemment, l'analyse des phrases et du texte porte également sur des aspects phonétiques comme l'assimilation (dévoisement, voisement), l'affrication (/tʃ/ et /dʒ/) et surtout les processus de liaison (Adda-Decker *et al.*, 2012 ; Fougeron & Delais-Roussarie, 2014). De ce fait, nous avons mesuré la durée, l'intensité et la fréquence fondamentale (cette dernière étant la plus importante) en vue d'analyser les schémas intonatifs concernant la place de l'accent principal et le nombre de syllabes accentuées. Nous les présentons dans les deux sous-sections suivantes.

6.6.3.1 La place de l'accent principal

En anglais, les mots de deux syllabes ont une règle fixe d'accentuation. Les noms dissyllabiques sont en générale accentués sur la première syllabe et les verbes dissyllabiques sont accentués sur la deuxième (Roach, 2009, p. 77). Comme dans la majorité des règles, il y a des exceptions dans les deux cas, mais nous ne les évoquerons pas ici (consulter Cruttenden, 2014 pour les détails). Nous nous concentrons sur les verbes, car la RP et le GA ont également un point de divergence notamment avec certains¹⁹⁶ des verbes finissant par *-ate*. Alors que la RP respecte cette règle d'accentuation, en plaçant l'accent sur la deuxième syllabe des mots *rotate*, *donate* et *locate* (/ləʊ'keɪt/, /rəʊ'teɪt/ et /dəʊ'neɪt/), c'est plutôt la première syllabe qui est accentuée chez la plupart des locuteurs du GA (/ˈləʊkeɪt/, /ˈrəʊteɪt/ et /ˈdəʊneɪt/).

Afin de savoir laquelle des deux syllabes porte l'accent, nous avons mesuré trois éléments acoustiques sur chaque syllabe des mots *locate* et *donate*. Celle qui est accentuée devrait présenter une plus importante valeur d'au moins un des éléments suivants : la durée, l'intensité et la fréquence fondamentale. La dernière reste l'indicateur le plus révélateur.

196 Le GA, comme la RP, respecte cette règle pour certains verbes. Plus de détails dans le EPD et LPD.

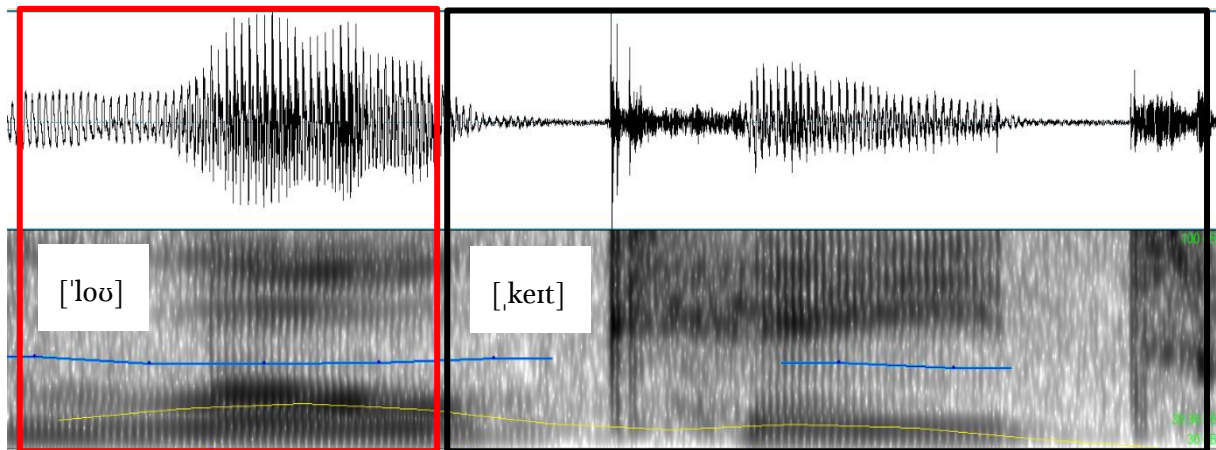


Figure 25 : un exemple d'accentuation sur la première syllabe dans locate (/['ləʊ],keɪt/) (encadré rouge)

6.6.3.2 Le nombre de syllabes accentués

Une autre différence notable entre le GA et la RP, portant sur l'accentuation, est la prononciation des mots terminant par <-ary> (*extraordinary, secretary*); <-ory> (*laboratory, explanatory*); <-erry> (*strawberry, Cranberry*); <-ury> (*canterbury, Glastonbury*) ou <-ony> (*testimony, ceremony*). Dans les mots de cette catégorie, la voyelle de la pénultième syllabe est réduite (utilisation du schwa) voire totalement absente en RP ('sekrətəri ou 'sekrətri), alors qu'en GA, cette même syllabe se voit affectée un accent secondaire ('sekrəteri) (Gynésy 2000 pp. 87 ; Cruttenden, 2014 p.62). Pour Durand et Navarro (2015, p. 136), ce phénomène s'explique par le fait qu'en RP ces mots ne représentent qu'un seul pied¹⁹⁷ alors qu'ils ont deux pieds en GA. Pour tous nos participants, nous avons examiné l'accentuation des mots *necessary* et *military*. Pareillement dans le cas de l'identification des syllabes accentuées (session précédente), nous avons pris des mesures portant sur la durée, l'intensité et la Fo, des syllabes concernées. Au reste, nous avons procédé à l'identification de la nature de la voyelle, si elle existe, dans la syllabe étudiée. Ceci nous permet de pouvoir affirmer qu'il s'agit d'un [ə], d'un [e] ou d'une autre voyelle.

Nous venons de passer en revue onze différences phonético-phonologiques entre la PR et le GA, avec les éléments sous-jacents et la manière dont nous les avons mesurés. Nous avons également précisé que certaines sont difficiles à cerner et que d'autres sont tout simplement subtiles. À présent, dans les sessions suivantes, nous allons décrire les deux dernières méthodes de collecte de données.

¹⁹⁷ Pour ces auteurs, « un pied métrique trochaïque [est] défini comme une unité qui se compose, pour l'anglais, d'une syllabe accentuée et des syllabes inaccentuées la suivant » (Brulard et al., 2015, p. 136).

6.7 Test de perception

Nous décrivons dans cette section le processus de collecte des données du test de perception qui a suivi immédiatement les enregistrements audio. Ce dernier s'est déroulé cette fois-ci dans la salle informatique de l'IPS. Les locuteurs sont installés devant un ordinateur, muni d'un casque audio à travers lequel ils écoutent les mots. Ils doivent appuyer sur des touches du clavier pour saisir leurs réponses. Toutes les instructions pour mener à bien le test sont expliquées en français et défilent sur l'écran de l'ordinateur. L'objectif principal du test est de vérifier si nos participants connaissent ou arrivent à reconnaître la différence de prononciation entre les deux types d'accents de référence. De plus, nous avons opté pour ce test en vue de pouvoir effectuer un lien, éventuellement, entre la perception et la production orale des participants. Celui-ci consiste à écouter des mots et à choisir l'accent avec lequel ils sont prononcés. Dans le but de concevoir ce test, nous avons téléchargé dans la base de données du dictionnaire sonore en ligne du MacMillan Dictionary¹⁹⁸ 17 mots parmi ceux de la première partie de notre liste de mots. La prononciation de ces mots a fait l'objet de vérifications dans les deux dictionnaires de prononciation, le EPD et le LPD, afin de s'assurer qu'ils soient conformes aux accents étudiés. Dans une volonté de réduire la durée de l'épreuve, nous avons seulement choisi quelques mots de la liste, mais de manière à ce que toutes les catégories étudiées soient représentées. Par conséquent, les mots suivants sont retenus : *actor, band, bird, city, dark, dynasty, joke, leisure, military, more, top, tune, donate, evolution, legalization, necessary, road*. Chaque mot a donc été prononcé par des locuteurs natifs américains et des locuteurs natifs britanniques, avec bien évidemment l'accent RP et l'accent GA respectivement. Pour déterminer le taux d'erreur permettant de s'assurer que les locuteurs ne répondent pas au hasard, nous y avons inséré deux fois la même prononciation de deux mots, parmi les 34 occurrences (17 x 2). Ainsi, le mot *road* est apparu deux fois. Mais la deuxième occurrence n'est qu'une copie de la première prononciation du même locuteur natif américain, et non de la prononciation du natif britannique. Le mot *dynasty* a également connu la même manipulation avec cette fois, deux occurrences du locuteur natif britannique.

Le test de perception est réalisé sur un programme informatique conçu avec le langage Python, par un de nos collègues doctorant Bruno OBERLÉ. L'interface utilisateur est en noir (écran de MS

¹⁹⁸ <https://www.macmillandictionary.com/>

DOS). Au commencement, nous avons explicité toutes les consignes nécessaires au bon déroulement du test. Au début, les participants effectuent un essai avec deux exemples, en guise d'entraînement, leur permettant de s'habituer à la tâche et aussi pour s'assurer que tout est compris, avant de commencer le test. Le script affiche tout de même un rappel des instructions avant chaque écoute et randomise automatiquement les fichiers audio à chaque écoute. Le nombre de mots restant à écouter y est également visible. Le sujet a le contrôle total de l'exercice : il décide quand il peut commencer et quand il veut effectuer une pause. Le script utilisé est écrit sur mesure afin de répondre aux fonctionnalités spécifiques dont nous avons besoin, lesquels nous permettent d'extraire les données présentées dans la Figure 27 : exemple d'un fichier de résultats généré automatiquement par le script du test de perception. C'est pour toutes ces raisons que nous n'avons pas voulu utiliser des logiciels de test de perception existants sur le marché (Perceval, PsychoPy, etc.), dont les fonctionnalités sont limitées.

La question principale à laquelle les participants doivent répondre est : « s'agit-il d'une prononciation ? » (1) américaine, (2) britannique et (3) ni l'un ni l'autre et (4) je ne sais pas. Ils doivent appuyer sur les touches 1, 2, 3 ou 4 du clavier d'ordinateur pour indiquer leurs réponses. Le degré de certitude sur une échelle de Lickert doit être obligatoirement fourni. La certitude 1 correspond à « je suis très sûr.e », la 2 correspond à « je suis plutôt sûr.e » et enfin, la certitude 3 indique « pas du tout sûr.e ». Lorsqu'ils répondent : (3) ni l'un ni l'autre ou (4) je ne sais pas, la certitude est automatiquement zéro (0)¹⁹⁹. Aussi, ils doivent justifier les éléments qui leur ont permis d'effectuer leurs choix, mais cette option est facultative et ils sont libres d'écrire en anglais ou en français. Ils ont également l'opportunité de réécouter les mots autant de fois que possible. De la sorte, leurs réponses sont stockées en temps réel dans un dossier automatiquement créé à cet effet, en format .csv, que l'on peut ouvrir, soit avec Ms Excel ou Libre/ Open Office Calc. Les fichiers de résultats comprennent : le nom du participant, le nom des fichiers, la réponse du participant, la réponse correct (solution), justesse (vrai ou faux), nombres d'écoutes (tentatives), explication (les raisons de leur choix), le degré de certitude, le temps total en millisecondes (par fichier audio) et le détail des temps (selon le nombre de tentatives). Lorsque l'utilisateur choisit le numéro 4 (je ne

199 Ceci ne suit pas une logique d'ordre, sinon la certitude suivante devrait être plutôt 4, nous avons opté pour zéro comme nulle.

sais pas) alors que la réponse est soit 1 (américain), soit 2 (britannique), le champ « justesse » y est laissé vide (Figure 26 ci-dessous).

```

Il vous reste 2 mots à écouter.
Appuyer sur Entrée pour commencer l'écoute du prochain mot...
- Entrez 1 si c'est AMÉRICAIN
- Entrez 2 si c'est BRITANNIQUE
- Entrez 3 si c'est AUTRE CHOSE
- Entrez 4 si VOUS NE SAVEZ PAS
- Appuyer sur Entrée pour réécouter
Votre choix: 1
Quel(s) élément(s) vous ont incité à faire ce choix (optionel)? d
Entrez 1 si vous êtes tout à fait sûr.e de votre réponse.
Entrez 2 si vous êtes plutôt sûr.e de votre réponse.
Entrez 3 si vous n'êtes pas du tout sûr.e de votre réponse.
Votre choix: 2

```

Figure 26 : exemple d'interface du test de perception auditive créée avec Python.

nom du participant:								
nom du fichier	réponse participant	réponse correct	justesse	nombres d'écoute	explication	certitude	temps total	détail des temps
band_ga.mp3	américain	américain	VRAI	1	DIND	1	926	926
band_rp.mp3	britannique	britannique	VRAI	1	BAND	1	1315	1315
bird_ga.mp3	américain	américain	VRAI	1	RID	1	1317	1317
bird_rp.mp3	britannique	britannique	VRAI	2		1	4919	3791+1128
city_ga.mp3	américain	américain	VRAI	1	DI	1	593	593
city_rp.mp3	britannique	britannique	VRAI	1	TI	1	1113	1113
dark_ga.mp3	américain	américain	VRAI	1	ARK	1	800	800
dark_rp.mp3	britannique	britannique	VRAI	1	AK	1	879	879
donate_ga.mp3	américain	américain	VRAI	1		2	3525	3525
donate_rp.mp3	britannique	britannique	VRAI	1		2	3557	3557
dynasty_rp1.mp3	britannique	britannique	VRAI	1		2	1201	1201
dynasty_rp2.mp3	britannique	britannique	VRAI	1	DI	1	1175	1175
more_ga.mp3	américain	américain	VRAI	1	MOR	1	1206	1206
more_rp.mp3	britannique	britannique	VRAI	1	O	1	1065	1065
military_rp.mp3	britannique	britannique	VRAI	1	TRI	1	989	989
necessary_ga.mp3	américain	américain	VRAI	1	SARY	1	769	769
military_ga.mp3	américain	américain	VRAI	1	TARY	1	660	660
necessary_rp.mp3	britannique	britannique	VRAI	1	SRI	1	1341	1341

Figure 27 : exemple d'un fichier de résultats généré automatiquement par le script du test de perception.

6.8 Questionnaire

Au départ, le questionnaire était conçu pour être diffusé sur *Google Forms*, pour permettre la facilité de certaines manipulations de base (recueil, organisation et analyse, graphiques, exportation en format Excel, etc.) et de gagner du temps. Mais, ayant testé le questionnaire papier sur deux étudiants pilotes, nous avons abandonné la solution en ligne pour diverses raisons. Pour commencer, le format papier laissait le libre choix d'ajouter des réponses, soit dans les marges, soit au verso. Par conséquent, nous avons prévu une dizaine de ligne pour renseigner les séries regardées alors qu'un locuteur devait en écrire une vingtaine. Ensuite, les retours écrits sur le

papier, nous ont permis de modifier et d'améliorer certaines questions dans la version définitive. Il a de plus permis d'obtenir des informations complémentaires qui sont très utiles pour nos analyses. Cela a été une sorte d'entretien pendant lequel nous avons pris des notes. Enfin, notre présence a permis aux locuteurs de poser des questions sur des aspects mal compris et d'avoir des éclaircissements, une opportunité qui n'aurait pas été certaine avec un questionnaire en ligne. Il était donc évident que le questionnaire papier allait nécessiter beaucoup plus de travail, car en plus des fonctions qu'offre *Google forms*, nous avons dû reporter manuellement les données récoltées dans un fichier Excel.

Les locuteurs ont donc rempli ce questionnaire (voir Annexe A 4, p. 184) visant à non seulement collecter des informations identitaires, mais également à indiquer tout renseignement portant sur leurs pratiques informelles de l'anglais. La première partie du questionnaire comporte 7 questions. Dans les trois premiers champs, les locuteurs ont renseigné leur nom et prénom, l'âge et leur genre. Puis, les deux champs suivants ont permis de connaître leurs études et leurs niveaux. Enfin, leur langue maternelle (L1) ainsi que les langues parlées et / ou en cours d'apprentissage (L2) ont été renseignées dans les deux derniers champs. Afin de garantir l'anonymat des participants, des initiaux dérivés de leurs noms et prénoms leur ont été assignés (voir section Anonymisation des locuteurs). La deuxième partie du questionnaire comporte 25 questions (avec des sous-questions) et a été constituée de façon à obtenir des informations les plus exhaustives possibles sur les pratiques en anglais et les activités de visionnage de séries. Les types de questions posées et éventuellement les réponses fournies ont permis de dégager trois grandes parties à savoir :

- le contact avec la langue anglaise,
- les séries télévisées,
- l'imitation et adoption d'accent.

6.8.1 Contact avec la langue anglaise

Dans cette partie, les étudiants ont renseigné le nombre d'années d'apprentissage de l'anglais dans un contexte formel en général et la durée dans un CRL ou équivalent à l'université (au jour de l'enregistrement). Ils ont également choisi une fourchette de temps pendant laquelle ils utilisent l'anglais en dehors des CRL (exemple : entre 4h et 8h) et ensuite ils ont renseigné les activités dans lesquelles ils s'engagent (exemple : écoute de la musique en anglais). S'ils communiquent oralement en anglais, ils ont évoqué : de quel pays proviennent leur(s) interlocuteur(s), la relation

(famille ou ami-e-s), la durée des discussions et par quels moyens (exemple : visioconférence). Ensuite, ils devaient annoncer s'ils ont effectué des séjours dans des pays anglophones, notamment aux États-Unis et au Royaume-Uni, et indiquer quelle en était la durée. Enfin, lorsque les locuteurs s'en souviennent, ils devaient indiquer la nationalité (exemple : indienne, française) et l'accent (exemple : britannique) de leurs enseignants d'anglais (contact formel) depuis le collège jusqu'à leur niveau actuel (licence, master).

6.8.2 Les séries télévisées

Ici les participants devaient répondre aux questions relatives au visionnage de séries télévisées. L'objectif étant d'indiquer s'ils regardent des séries américaines, britanniques ou autres. Si la réponse était oui, ils devaient renseigner depuis combien de temps et estimer (en moyenne) le nombre total d'heures de visionnage par semaine ou par mois. Par la suite, ils devaient indiquer le mode de visionnage, soit en version originale (VO), soit en version originale sous-titrée anglais (VOSTA), soit en version originale sous-titrée français (VOSTF), soit en version française, et estimer si possible, en pourcentage, la proportion des modes dans le cas où ils alternent (par exemple : 20 % VO, 70 % VOSTA et 10 % VOSTR). Nous leur avons également demandé dans quel but ils regardent les séries ainsi que les aspects qui les intéressent dans ces dernières. Ils devaient aussi indiquer le(s) support(s) utilisé(s) (téléchargement sur disque dur, streaming, etc.) Enfin, ils devaient lister le titre des séries qu'ils regardent (américaines, britanniques et autres).

6.8.3 L'adoption d'accent

Les dernières questions permettent de recueillir les impressions personnelles des sondés par rapport à l'accent qu'ils adoptent lorsqu'ils parlent anglais. Ils devaient dire s'ils imitent sciemment un type d'accent lorsqu'ils parlent anglais et fournir la raison si possible. Ensuite, ils devaient dire s'ils pensent parler avec un accent américain, si quelqu'un le leur en a déjà fait part et si oui, d'où provenait cet accent, s'ils le savent. Pour finir, nous avons voulu vérifier si nos locuteurs connaissaient tous les mots de la liste, y compris leur prononciation, avant l'enregistrement audio. Ainsi, nous leur avons demandé de lister les mots qu'ils ne connaissaient pas, sur la liste qui a servi aux enregistrements. Ce questionnement nous permet de comparer la production de ces mots inconnus avec ceux qu'ils estiment mieux connaître, du point de vue de leur prononciation.

6.9 Conclusion

Nous venons de présenter en détail le protocole expérimental de ce travail. Il nous faut signaler que c'est un processus complexe, dans lequel nous avons tenu à bien expliciter les démarches méthodologiques. Ainsi, la collecte des données s'est effectuée en trois grandes phases que sont : les enregistrements audio au cours desquels les productions orales de nos 20 participants ont été recueillies, le test de perception pour savoir s'ils (re)connaissent les accents utilisés dans la prononciation des mots et aussi un questionnaire qui renferme des informations personnelles et sur les habitudes de leurs pratiques informelles de l'anglais. Nous avons passé en revue un nombre d'éléments nécessaires à une meilleure compréhension des données analysées. Dans la première phase, nous avons parlé : du recrutement et du déroulement de l'enregistrement des locuteurs, de la construction et de la composition des différentes parties constitutives du support ayant servi aux prises des productions orales, des sons étudiés et leur analyse acoustique (segmentation et mesures). Durant cette dernière section, nous avons expliqué les choix des éléments de différences entre l'accent britannique et américain, en nous reposant sur la littérature existante. Dans la deuxième phase, l'élaboration, le fonctionnement, le déroulement et les types de résultats du test de perception ont été présentés. Enfin, la troisième phase a fait l'objet d'un débroussaillage des types de questions posées aux participants. Celles-ci recouvrent des informations personnelles concernant les activités liées à la pratique informelle de l'anglais. Aussi avons-nous justifié le choix des prononciations dans les deux modalités d'élocution à savoir l'adoption de la vitesse normale et rapide. La vitesse d'élocution rapide, rappelons-le, est fondée sur le fait qu'elle constitue un perturbateur écologique par excellence et nous permet de tester la résistivité du système de perception et de production des participants à cette étude.

Dans les chapitres prochains, il sera question de présenter, d'analyser et de discuter les résultats obtenus, tout d'abord les uns après les autres et ensuite procéder à une triangulation consistant à croiser les trois sources de données. Nous traiterons premièrement les résultats de la prononciation des mots en contexte de vitesse normale (chapitre 7), deuxièmement les mêmes résultats en situation de vitesse rapide (chapitre 8). Troisièmement, nous combinerons les résultats du questionnaire avec ceux du test de perception (chapitre 9) et quatrièmement la triangulation où nous présenterons des corrélations entre quelques variables importantes au

regard de nos questions et hypothèses de recherches (chapitre 10). Ce dernier contiendra également une discussion générale des résultats.

7 Résultats de la prononciation en vitesse normale

Are those who sound native expected to “fool some of the people some of the time, all the people some of the time, or all the people all the time?” [...] not all L2 learners strive for native-like-ness. Some see it as an unrealistic target, or prefer to hold onto their discernibly foreign accent for its symbolic value.
(Moyer, 2013, p. 18).

Dans ce chapitre, nous présentons les résultats des sons et séquences de sons, contenus dans des mots placés dans la phrase porteuse (*it is *** please*), à la vitesse normale. Les présentations suivent l'ordre tel qu'il est décrit dans le cadre méthodologique (section. 6.4) : les consonnes, puis les voyelles et enfin les éléments suprasegmentaux (accentuation et nombre de syllabes accentuées). La présentation est identique pour chaque élément étudié. Premièrement, nous décrivons les résultats obtenus pour chaque mot, ensuite nous résumons les résultats généraux concernant les mots qui contiennent ces éléments. Deuxièmement, nous illustrons ces résultats à l'aide de graphiques. Les présentations sont suivies d'une brève discussion au cours de laquelle nous essayons de comprendre et d'expliquer les différentes réalisations, en passant par des questionnements et des hypothèses. À la fin des présentations, nous résumons les résultats globaux pour tous les éléments étudiés ainsi que les pourcentages de prononciation de chaque locuteur. Cette présentation est également suivie d'une discussion générale.

Afin de faciliter la lecture des résultats et les graphiques correspondants, une charte graphique est établie²⁰⁰. Un code couleur²⁰¹ est respecté tout au long de ce chapitre et des suivants, concernant les résultats des analyses acoustiques. De ce fait, le rouge foncé correspond aux productions proches de la RP et le bleu foncé représente les réalisations proches du GA. Les mêmes couleurs sont représentées, mais en plus claires, lorsque deux ou plusieurs réalisations sont attestées dans un système de prononciation particulier. Par exemple, le rouge clair indique une production possible de la RP. Ensuite, nous utilisons le vert pour des productions que nous jugeons proches ou équivalentes au système phonétique du français (ou des langues parlées par certains

²⁰⁰ La version papier de ce tapuscrit est également en couleur, afin que les graphiques correspondent aux différents codes de couleurs évoqués. Ce choix de mélange est simplement fait afin de mieux distinguer les différentes parties des diagrammes pour une meilleure lisibilité.

²⁰¹ Aucune signification symbolique n'est assignée aux couleurs utilisées.

locuteurs²⁰²) et les couleurs vert clair pour se référer aux réalisations autres que celles des précédents systèmes phonétiques identifiés, et que nous nommons « Autres » (avec un A). Parallèlement, les diagrammes en secteurs illustrent les pourcentages des différentes productions pour un élément étudié et les histogrammes sont utilisés pour montrer la production de chaque locuteur. Les valeurs acoustiques (formants, durées des segments, fréquence fondamentale et intensité) ayant servi à l'identification des sons, ainsi que des exemples d'illustrations des diverses productions évoquées (oscillogrammes et spectrogrammes) se trouvent en annexes (Annexe A 1, p. 1-35).

Enfin, nous rappelons que dans la présentation des résultats, nous utilisons deux signes différents pour se référer à deux types de productions bien particulières. En premier lieu, nous mettons les symboles phonétiques entre barres obliques (/ /) pour indiquer un phonème, c'est-à-dire un son général qui peut être décliné en allophones²⁰³. Par exemple, /l/ et /r/ sont deux phonèmes distincts en anglais, car *load* (/ləʊd/) et *road* (/rəʊd/) ont deux sens différents. En deuxième lieu, nous utilisons les crochets ([]) pour indiquer les productions phonétiques fines qui représentent des allophones, c'est-à-dire les différentes variétés d'un phonème particulier. Que le mot *load* soit réalisé avec [ɫ] ou [l] ne change pas le sens du mot en question. (Idem pour le mot *road* s'il est prononcé avec [r], [ɾ] ou [ʀ]). Nous utilisons aussi les barres obliques pour des sons qui ne sont pas issus de nos résultats, par exemple lorsque nous citons les dictionnaires. En somme, les // représentent un son ou phonème en général et les [] indiquent une réalisation spécifique.

Nous présentons les résultats des éléments segmentaux consonantiques analysés dans les sections suivantes et dans l'ordre suivant : le /t/ intervocalique, la rhoticité et le *yod-dropping*²⁰⁴.

7.1 La prononciation du /t/ intervocalique

Nous commençons par la présentation des résultats des /t/ intervocaliques contenus dans les mots *city*, *vitamin* et *party*. Nous obtenons un total de 300 occurrences (cinq répétitions par locuteur). Parmi ces répétitions, nous observons huit types de productions différentes que nous décrivons brièvement, avec des exemples de mots dans lesquels ces sons apparaissent. Il s'agit du

202 Voir les résultats du questionnaire, section 9.1 p. 304.

203 Ils n'entraînent donc pas d'oppositions de sens.

204 Nous plaçons la semi-consonne /j/ dans les résultats des consonnes, car c'est elle qui « subit » le changement et pas la voyelle /u:/ qui la suit.

/t/ aspiré [t^h] que l'on retrouve en initiale de mot et dans les syllabes accentuées (*time* - [t^ham], *spectator* - [spek't^hertə]), du /t/ sans aspiration [t], de l'affriquée [tʃ] (*match* -[mætʃ]), du coup de glotte [ʔ], du /t/ voisé [t̚]²⁰⁵, du [d] (*deep* - [di:p]), du [r] (*city* - [siri]) et enfin de la double articulation [tʂ], qui consiste à prononcer un /s/ immédiatement à la suite d'un /t/.

7.1.1 City

Dans la prononciation du mot *city*, nous avons identifié sept productions différentes (Figure 28). Sur 100 itérations, le [t^h] est produit à 37 %, le [r] à 23 %, le [d] à 14 %, le [tʃ] à 12 %, le [tʂ] à 6 %, le [t̚] à 5 % et le [t] à 3 %. Les locuteurs 1, 3, 13 et 17 ne réalisent que des [t^h], alors que les locuteurs 8, 9 et 16 réalisent exclusivement le son [r], et le locuteur 20 qui ne réalise que des [tʃ]. Les autres mélangent au moins deux sons. Les locuteurs 5, 6, 14, 15 et 18 réalisent un mélange de deux sons avec respectivement [t^h] et [tʂ] ; [t^h] et [tʃ] ; [t^h] et [tʃ] ; [t^h] et [d] ; [t^h] et [tʂ]. De même que les locuteurs 2, 12 et 19 qui utilisent à la fois [d] et [r] ; [tʃ] et [tʂ] ; [d] et [r]. Les locuteurs 7, 10 et 11 mélangent trois productions avec respectivement²⁰⁶ [t̚], [d] et [r] ; [t^h], [d] et [tʂ] ; [t], [t̚] et [d]. Enfin, le locuteur 4 est le seul à mélanger quatre sons ([t], [t̚], [d] et [r]).

²⁰⁵ Voir les travaux de Steiblé (2014) sur les consonnes occlusives du français par certains Alsaciens.

²⁰⁶ À travers ce chapitre et les autres sur les résultats acoustiques, la part de chaque son pour chaque locuteur se trouve sur les histogrammes. Par exemple : le locuteur 2 produit quatre [r] et un [d].

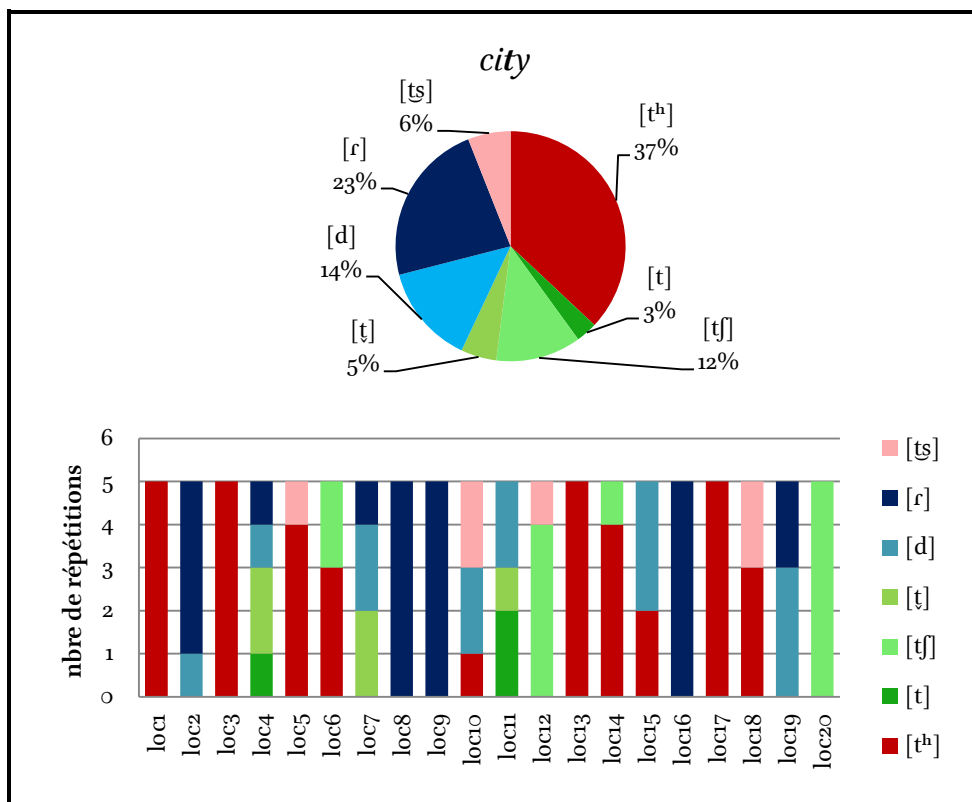


Figure 28 : pourcentages et répartition par locuteur de la production du <t> de *city*.

7.1.2 Vitamin

Le mot *vitamin* donne lieu à sept réalisations différentes, dont six présentées précédemment. Un (1) coup de glotte [ʔ] a émergé et le [t͡ʃ] n'est pas produit pour ce mot. Ainsi, [tʰ] est réalisé avec une majorité de 60 % sur les 100 itérations, suivi du [r] (16 %), du [d] (8 %), du [t] et du [ʈ] (6 % chacun), du [t͡s] (3 %) et enfin du [ʔ] (1 %). Pendant que les locuteurs 1, 3, 5, 7, 12, 15, 17, 18 et 20 sont homogènes dans leurs productions du [tʰ], le locuteur 9 ne réalise que des [r]. Les locuteurs 2, 6, 11, 13, 14 et 16 mélangent deux sons respectivement : [d] et [r] ; [tʰ] et [t͡s] ; [t] et [r] ; [tʰ] et [t] ; [tʰ] et [t͡s] ; [tʰ] et [ʈ]. Trois productions sont mélangées par les locuteurs 4, 8, 10 et 19 avec respectivement [t], [d] et [r] ; [ʔ], [ʈ] et [r] ; [tʰ], [ʈ] et [d] et enfin [ʈ], [d] et [r].

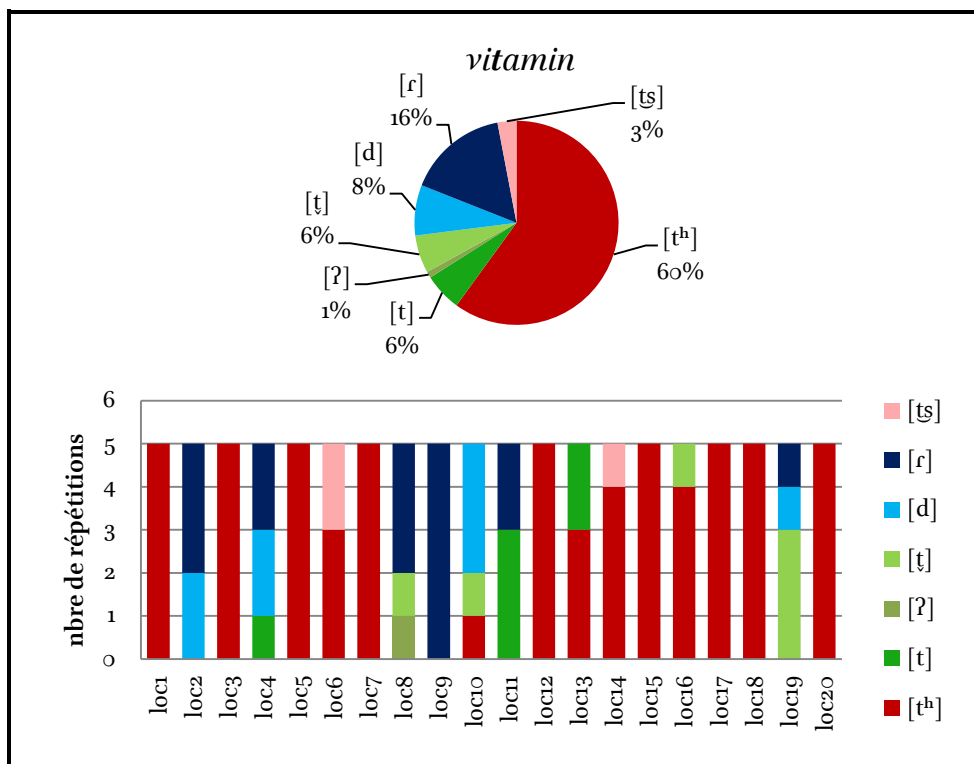


Figure 29 : pourcentages et répartition par locuteur de la production du <t> de *vitamin*.

7.1.3 Party

Les mêmes sons qui sont présentés dans l'analyse du mot *city* se retrouvent dans la prononciation du mot *party*. Le [tʰ] vient toujours en tête avec 45 % des productions sur 100 répétitions, suivi du [r] (34 %), du [d] (9 %), du [tʃ] (6 %), du [t] (3 %), du [tʂ] (2 %) et enfin du [ʔ] (1 %). Un schéma similaire concernant les mélanges se dessine aussi pour ce mot. Les locuteurs 4, 11, 16, 18 et 20 mélangent respectivement trois réalisations, à savoir [tʰ], [d] et [r] ; [t], [d] et [r] ; [tʰ], [d] et [r] ; [tʰ], [t] et [r] ; [tʰ], [t] et [r]. Les locuteurs 6, 7, 8, 10, 13 et 15 en mélangent deux : [tʰ] et [tʂ] ; [d] et [r] ; [d] et [r] ; [tʰ] et [tʂ] ; [tʰ] et [t] ; [tʰ] et [d]. Les autres sont constants dans leurs réalisations avec : des [r] pour les locuteurs 2, 5, 9 et 19 ; des [tʰ] pour les locuteurs 1, 3, 14 et 17. Le locuteur 12 produit uniquement des [tʃ].

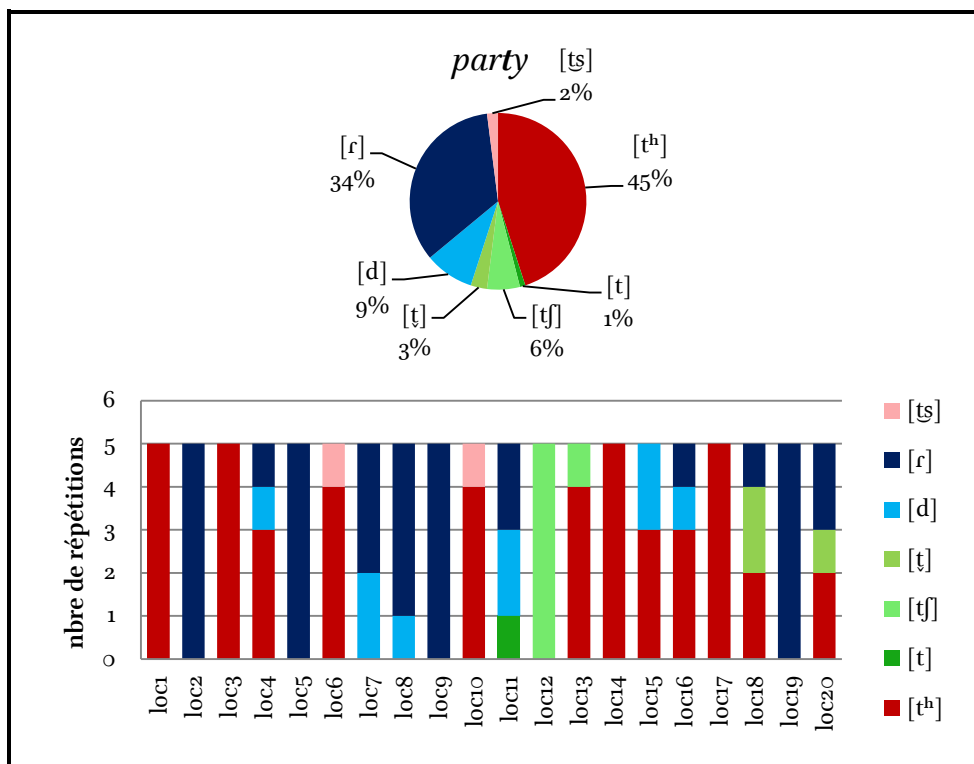


Figure 30 : pourcentages et répartition par locuteur de la production du <t> de *party*.

7.1.4 Résultats du /t/ intervocalique

Pour les trois mots présentés, le /t/ intervocalique est ainsi réalisé avec une dominance du [tʰ] à 48 %. Le [r] suit avec 24 %, soit la moitié du précédent. [d] est produit à 10 %, suivi du [tʃ] à 6 %, du [t̥] à 5 %, du [t͡s] à 4 %, du [t] à 3 % et enfin du [ʔ] qui n'est produit qu'une seule fois. Comme le montre la Figure 31, le schéma est beaucoup plus compliqué avec des mélanges allant de deux à cinq sons différents. Seuls les locuteurs 1, 3 et 17 pour [tʰ], le locuteur 9 pour [r] gardent des réalisations homogènes pour toutes leurs répétitions (Figure 31).

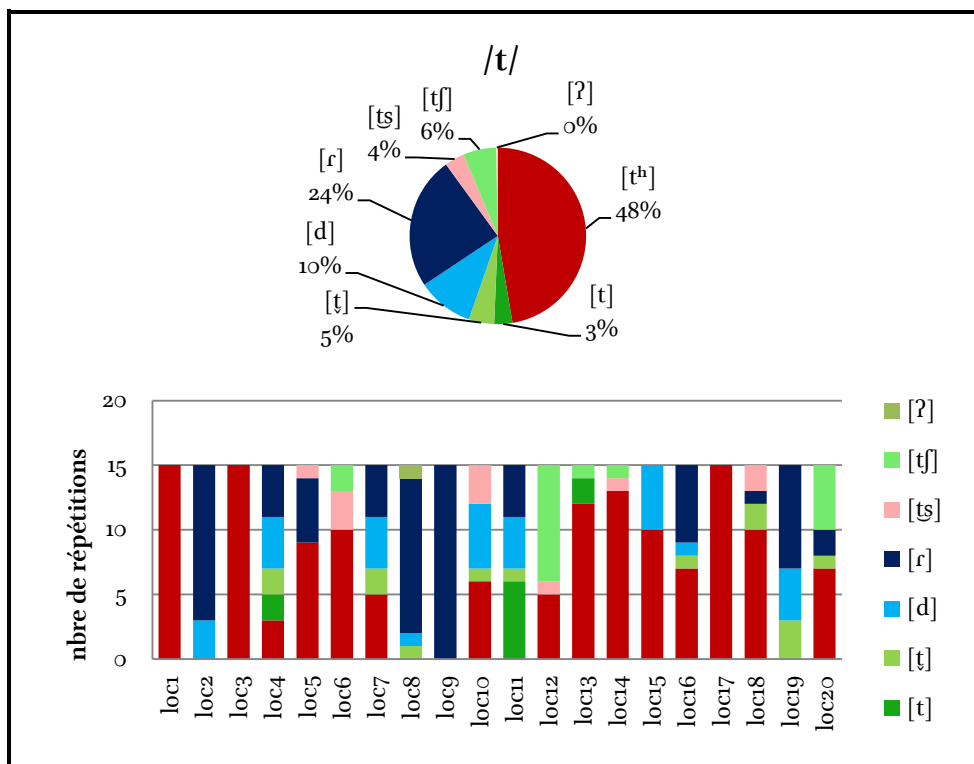


Figure 31 : pourcentages et répartition par locuteur de la prononciation du /t/ intervocalique.

7.1.5 Discussion

Le /t/ intervocalique est le phonème qui a généré le plus de variations dans les productions de nos locuteurs (huit allophones). Même si les caractéristiques de l’accent RP et GA sont prédominantes, les autres réalisations occupent une place non négligeable.

Le fait que les mots *city* et *party* soient quelques fois prononcés avec un [tʃ], contrairement à *vitamin*, peut s’expliquer par l’environnement phonétique, qui se traduit par l’influence des sons voisins sur celui étudié ([ɪ] de <y> et le [ə] de <a>). Ce phénomène aurait également joué un rôle dans l’émergence de certains sons « Autres ». Le [ɪ] inciterait à produire un [tʃ], c’est-à-dire que le son [tʃ] intervient uniquement lorsque le /t/ intervocalique est suivi de la voyelle [ɪ]. En effet, ceci ne se produit pas dans le cas de *vitamin*, où c’est la graphie <a> qui suit le /t/ intervocalique. Nous pouvons aussi lier la production du son [tʃ] à l’influence de la L1 notamment celle du parler de jeunes de banlieues²⁰⁷. Nous pouvons également considérer que certaines réalisations sont le produit d’un « ratage » de cible. Le cas du [d] peut être retenu, puisque ce son est proche acoustiquement et articulatoirement du *flap*. Soit il s’agit d’une battue mal réalisée, soit une

²⁰⁷ Voir Lehka-Lemarchand (2007) pour plus de détails sur l’accent de banlieues.

difficulté de perception qui fait que les locuteurs percevraient un /d/ à la place d'une battue. La même hypothèse peut être émise pour le /t/ voisé ([t̥]). Quant à la production du coup de glotte, nous savons seulement qu'il est généralement utilisé dans l'accent londonien (Cockney) et n'avons pas d'hypothèse sur son utilisation ici.

7.2 La rhoticité

Afin d'étudier la rhoticité dans la prononciation de nos locuteurs, nous analysons les mots *more*, *organisation*, *dark*, *party*, *bird*, *actor* et *leisure*. L'idée est de voir si la graphie <r> retrouvée dans ces mots est prononcée ou non. La somme des répétitions s'élève à 700, soit 100 par mot.

7.2.1 More

La prononciation du mot *more* est effectuée avec 70 % de rhoticité (/mɔɪ/) et 30 % de non rhoticité (/mɔ:/). Les locuteurs 1, 3, 4, 5, 6, 12, 13, 16, 17, 19 et 20 réalisent des [ɔɪ] pour toutes leurs répétitions tandis que les autres (locuteurs 2, 7, 11, 14, 15 et 18) produisent des [ɔ:] (Figure 32).

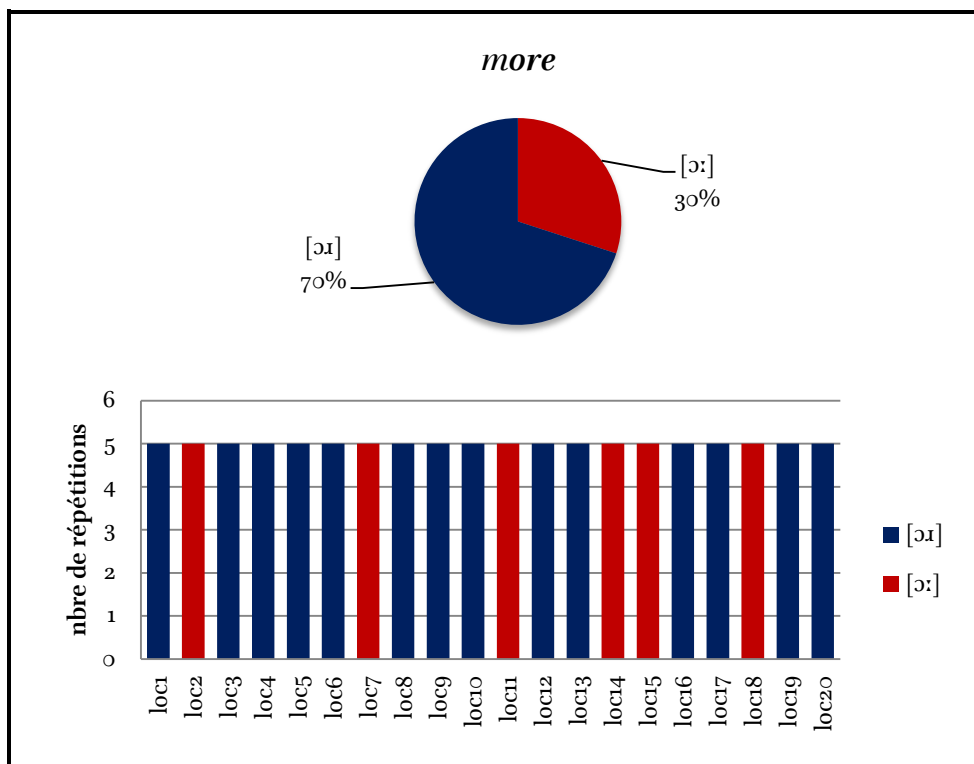


Figure 32 : pourcentages et répartition par locuteur de la production du <or> de *more*.

7.2.2 Organisation

La première syllabe du mot *organisation* est produite à 52 % avec le son [ɔ:] et à 48 % avec le son [ɔɪ]. Les locuteurs 13, 17 et 20 réalisent des mélanges avec respectivement [ɔ:] et [ɔɪ] ; [ɔ:] et [ɔɪ] et [ɔ:] et [ɔɪ]. Les locuteurs 1, 2, 6, 7, 11, 12, 14, 15 et 19 ne produisent que des [ɔ:] et les locuteurs 3, 4, 5, 8, 9, 10, 16 et 18 réalisent des [ɔɪ] pour toutes leurs répétitions (Figure 33).

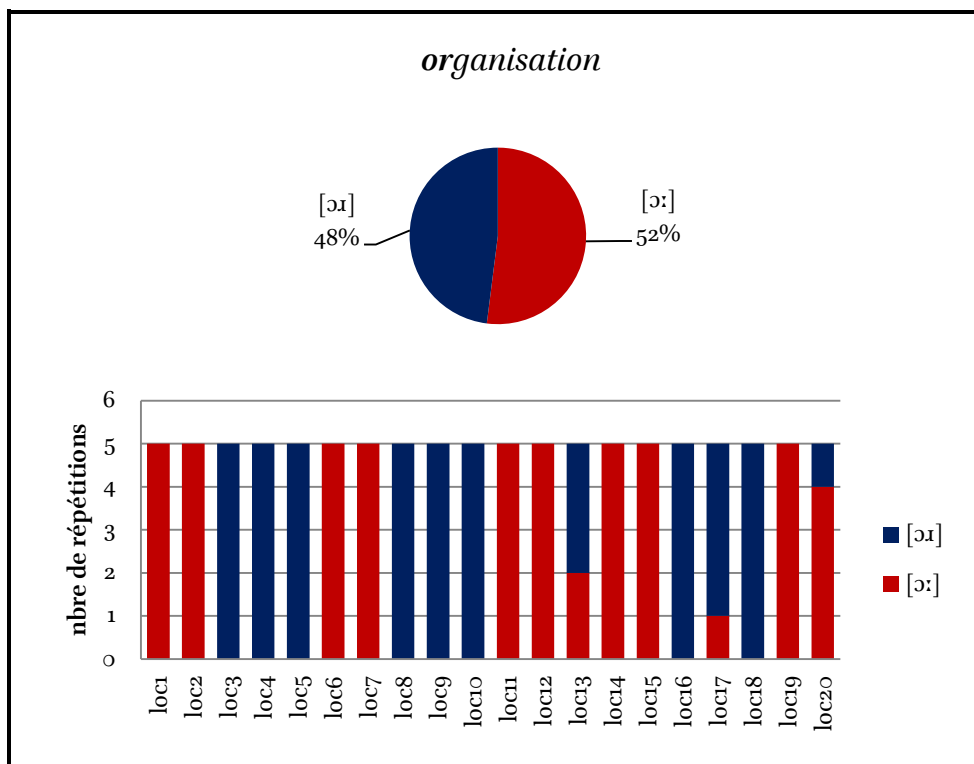


Figure 33 : pourcentages et répartition par locuteur de la production du <or> de *organisation*.

7.2.3 Dark

Le mot *dark* est prononcé avec une dominance [ɑ:] (63 %) et avec [ɑ:] à 37 %, pour les 100 répétitions. Seul le locuteur 5 produit un mélange de [ɑ:] et [ɑ:]. Les locuteurs 1, 2, 7, 10, 11, 14 et 15 produisent des [ɑ:], alors que les locuteurs 3, 4, 6, 8, 9, 12, 13, 16 et 20 réalisent des [ɑ:] (Figure 34).

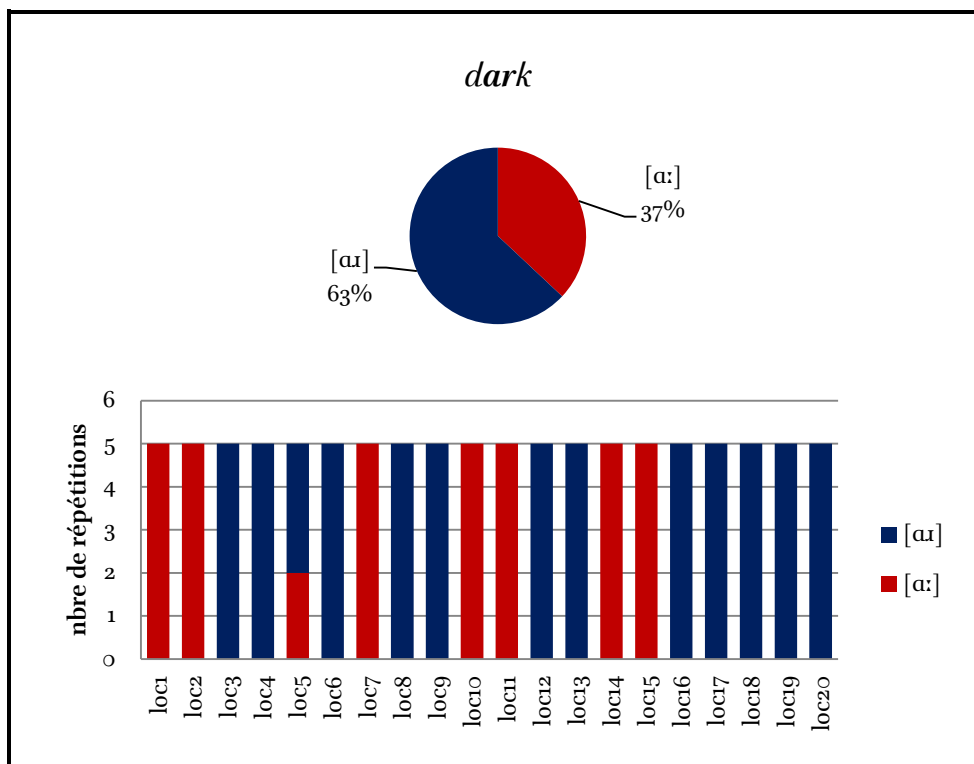


Figure 34 : pourcentages et répartition par locuteur de la production du <ar> de *dark*.

7.2.4 Party

Sur 100 répétitions, la première syllabe du mot *party* est prononcée à 69 % avec un [ɑ:] et à 31 % avec un [ɑ:]. Pendant que les locuteurs 1, 7, 10, 11, 14 et 15 réalisent des [ɑ:], les autres (2 à 7, 8, 9, 12, 13, 17 à 20) ne produisent que des [ɑ:], à l'exception du locuteur 16 qui réalise un mélange de [ɑ:] et [ɑ:] (Figure 35).

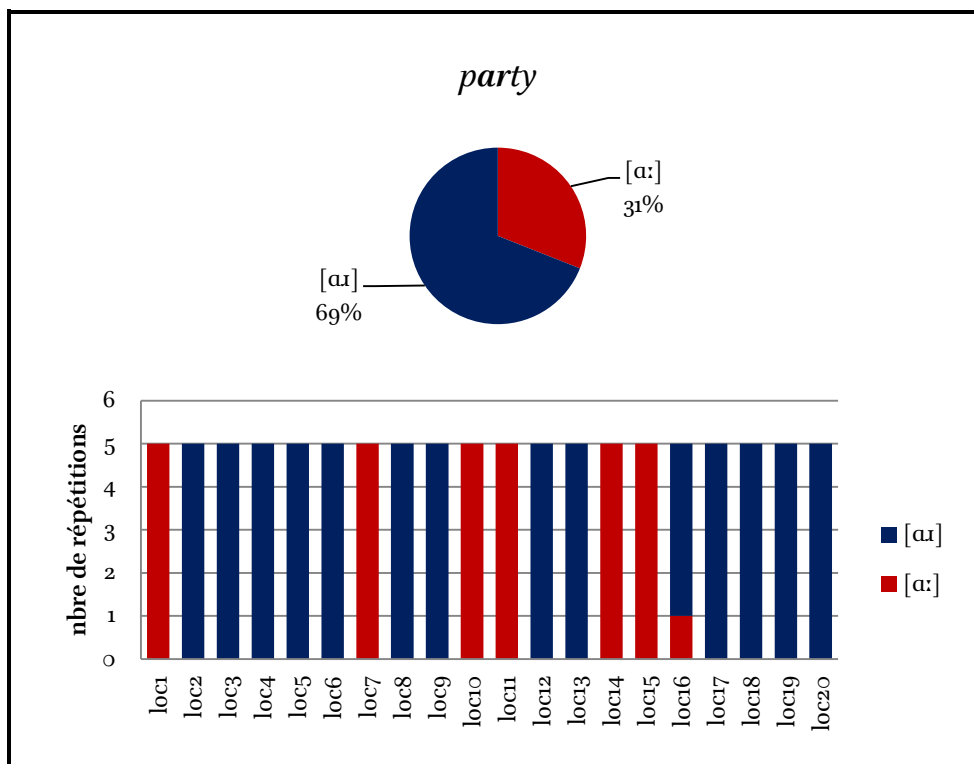


Figure 35 : pourcentages et répartition par locuteur de la production du <ar> de *party*.

7.2.5 Bird

Trois productions sont constatées dans la prononciation du mot *bird*. Il s'agit de [ɜ:] qui est réalisé à 66 %, du [ɜɪ] qui est réalisé à 23 % et du [ɜ:] qui est réalisé à 11 % sur les 100 répétitions. Les locuteurs 1 à 9, 17 à 19 ne produisent que des [ɜ:]. Les locuteurs 11 à 14 réalisent des [ɜɪ] et les locuteurs 15 et 20 produisent des [ɜ:]. Le locuteur 10 mélange [ɜ:] et [ɜ:], de même que le locuteur 16 avec les sons [ɜ:] et [ɜɪ].

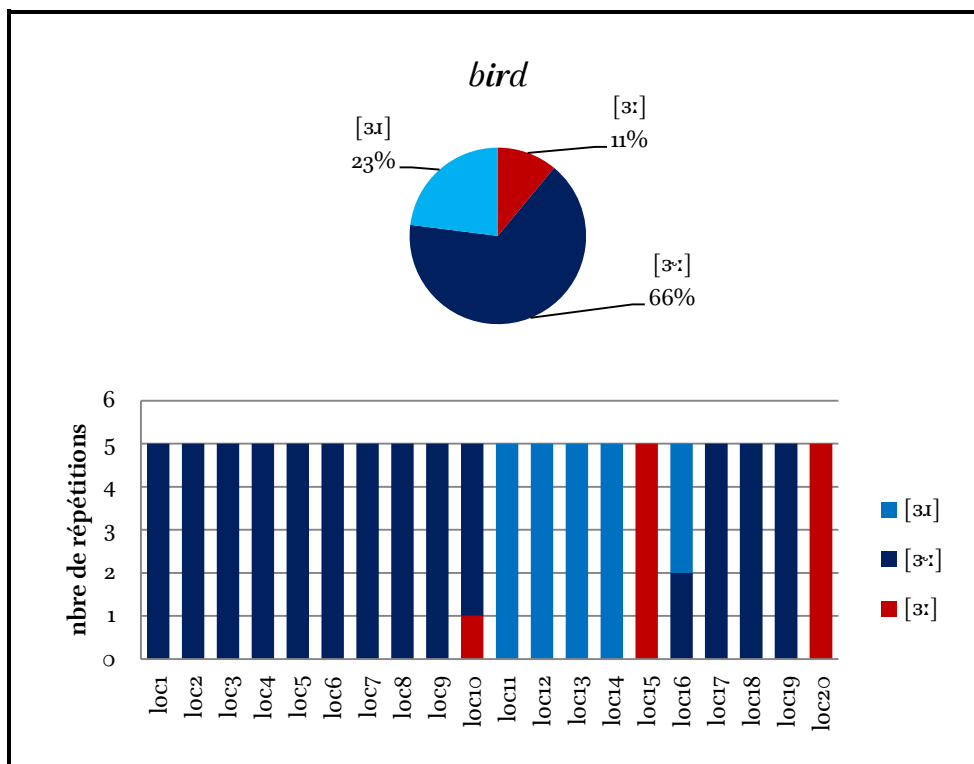


Figure 36 : pourcentages et répartition par locuteur de la production de <ir> de *bird*.

7.2.6 Actor

La deuxième syllabe du mot *actor* est produite à 55 % avec une prédominance de [ə], à 25 % avec [ə] et à 20 % avec [ɔ]. Ce dernier son que l'on retrouve plutôt dans le mot *more* (/mɔ:/^{RP}) est produit entièrement²⁰⁸ par les locuteurs 2, 6, 14 et 15. Les locuteurs 10, 11, 12, 18 et 20 ne produisent que des [ə], tandis que les locuteurs 1, 3, 4, 5, 7, 8, 9, 13, 16, 17 et 19 ne réalisent que des [ə].

²⁰⁸ Nous utilisons le mot « entièrement » pour se référer à une prononciation « homogène », c'est-à-dire lorsque le même son est produit pour les cinq itérations du mot.

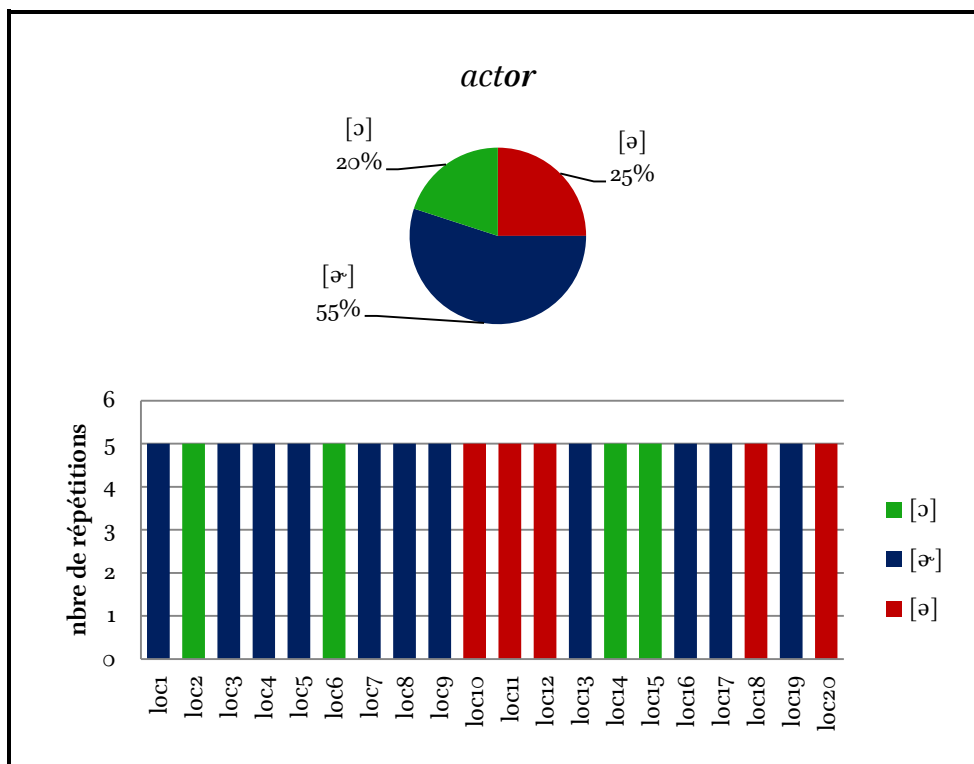


Figure 37 : pourcentages et répartition par locuteur de la production du <or> de *actor*.

7.2.7 Leisure

Les productions rhotiques dominent à 69 % dans la prononciation de la dernière syllabe du mot *leisure*. Les réalisations non rhotiques atteignent les 31 %. Les locuteurs 1, 3, 4, 5, 7, 8, 9, 11, 12, 13, 14, 16 et 19 produisent uniquement des [ə]. Les locuteurs 2, 6, 10, 15, 18 et 20 réalisent des [ə]. Seul le locuteur 17 mélange [ə] et [ə].

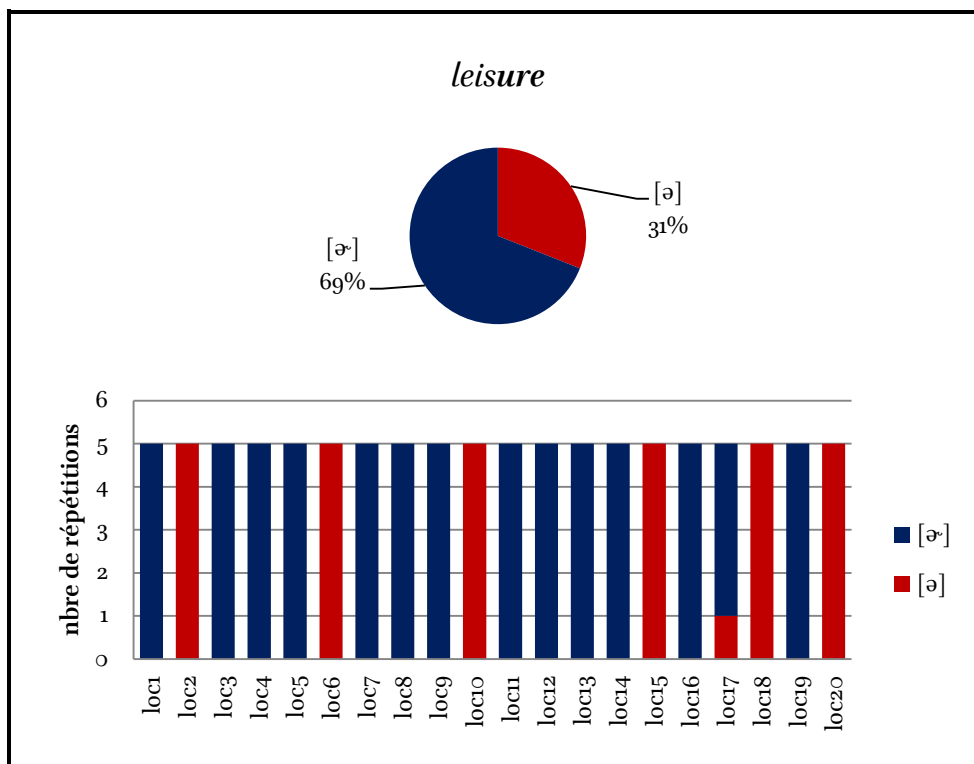


Figure 38 : pourcentages et répartition par locuteur de la production du <ur> de *leisure*.

7.2.8 Résultats d'ensemble pour la rhoticité

Les résultats pour l'ensemble de ces mots permettent de mieux visualiser la part globale de rhoticité et de non rhoticité. Ainsi, sur un total de 700 répétitions (35 par locuteur), la rhoticité domine avec 66 %. Remarquons la binarité dans les productions : elles sont soit rhotiques soit non rhotiques.

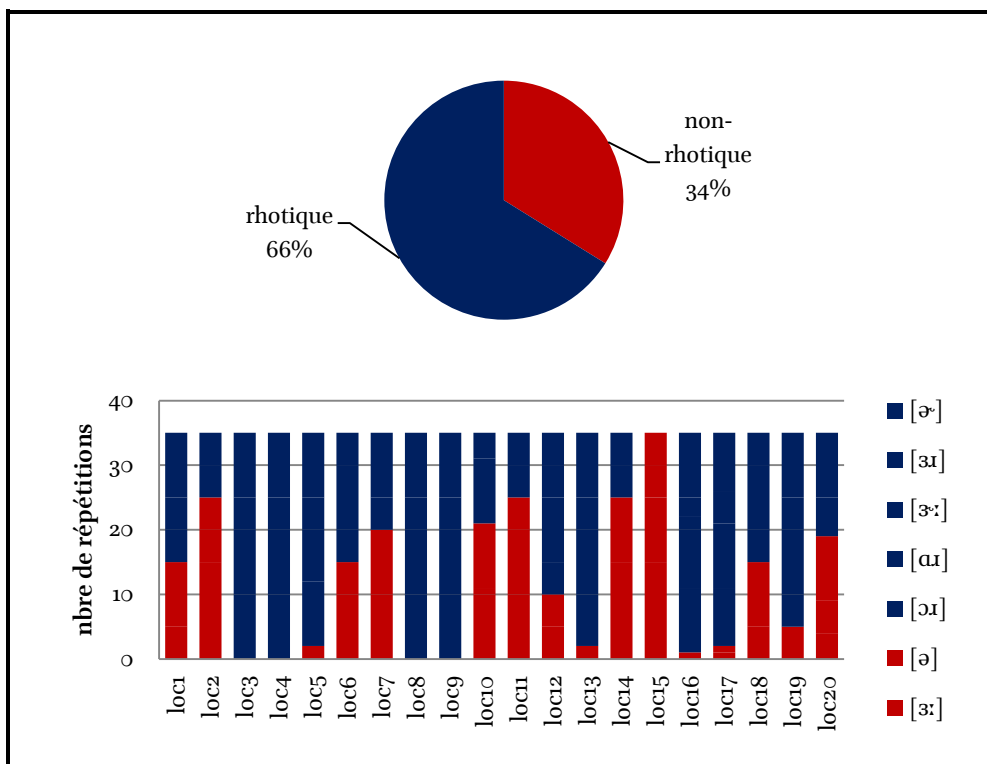


Figure 39 : pourcentages et répartition par locuteur de la rhocité.

7.2.9 Discussion

La prononciation du mot *actor* avec le son [ɔ] sur la deuxième syllabe peut laisser penser à l'influence de la graphie <or>. Dans ce cas, nous devons nous attendre à une production rhotique due à la présence du <r>, c'est-à-dire /'æktɔr/. Mais, force est de constater une production non rhotique de la part de ces locuteurs. Nous n'avons pas d'explications pour ce phénomène. Concernant la prononciation du mot *bird*, bien que le LPD et l'EPD indiquent seulement les deux premiers sons ([ɜ:] pour la RP et [ɜ:] pour le GA), la troisième réalisation [ɜ] est également attestée dans la littérature, notamment chez certains Écossais (Jauriberry, 2016).

7.3 Le yod-dropping

Dans la méthodologie, nous avons prévu d'analyser les mots *tune* et *new* pour l'élément *yod-dropping*. Mais nous rajoutons par la suite le mot *evolution* dans les résultats, car celui-ci contient également ce phénomène au niveau de <lu>. Ce mot connaît donc également une double prononciation : l'une avec un [u:] et l'autre avec un [ju:].

7.3.1 Tune

Sur un total de 100 répétitions (20 locuteurs x 5 répétitions), le mot *tune* est produit fortement avec [ju:] (88 %), suivi du [u:] (10 %) et seulement à 2 % avec un [y], comme dans le mot français tu (/ty/) (Figure 40). Il y a une forte dominance de la production proche de la RP chez tous les locuteurs, à l'exception des locuteurs 3 et 12 (avec une production entièrement en [u:]). Seul le locuteur 2 mélange [y] et [ju:].

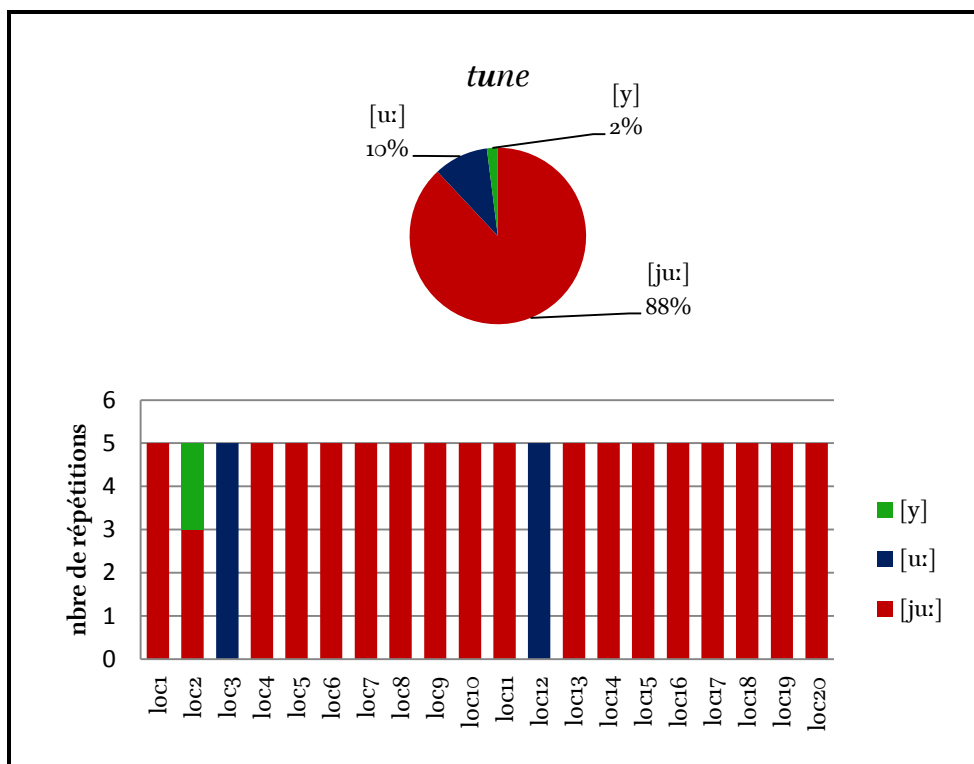


Figure 40 : pourcentages et répartition par locuteur de la production du <u> de *tune*.

7.3.2 New

Le mot *new* ne connaît que deux réalisations : une dominance de [ju:] (67 %) et l'autre avec du [u:] (33 %). Au niveau individuel, les locuteurs 1, 2, 4, 5, 6, 8, 13, 15, 16, 19 et 20 réalisent toutes les itérations du mot avec une prononciation proche de la RP [ju:]. La prononciation proche du GA est réalisée entièrement par les locuteurs 3, 7, 9, 12 et 17. Les locuteurs 10, 11, 14 et 18 réalisent le mot avec des mélanges de [ju:] et de [u:] (Figure 41).

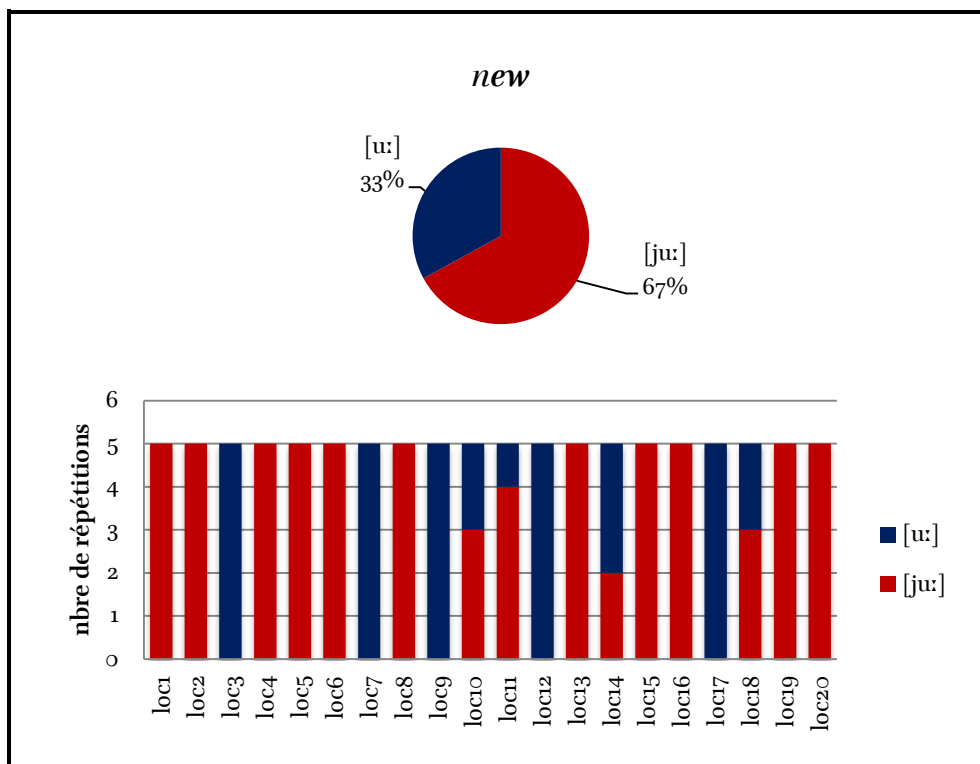


Figure 41 : pourcentages et répartition par locuteur de la production du <ew> de *new*.

7.3.3 Evolution

Contrairement aux mots *tune* et *new* dans lesquels la suite de sons [ju:] domine, le mot *evolution* est produit à 53 % avec [u:], à 37 % avec un [y] et seulement 10 % avec [ju:]. Le locuteur 11 utilise [y] et [u:]. Les locuteurs 5 et 7 utilisent chacun la prononciation avec un [ju:]. Les locuteurs 1, 6, 10, 13, 14, 15 et 19 produisent entièrement [y] tandis que les locuteurs 2, 3, 4, 8, 9, 12, 16, 17, 18 et 20 produisent entièrement le mot avec un [u:] (Figure 42).

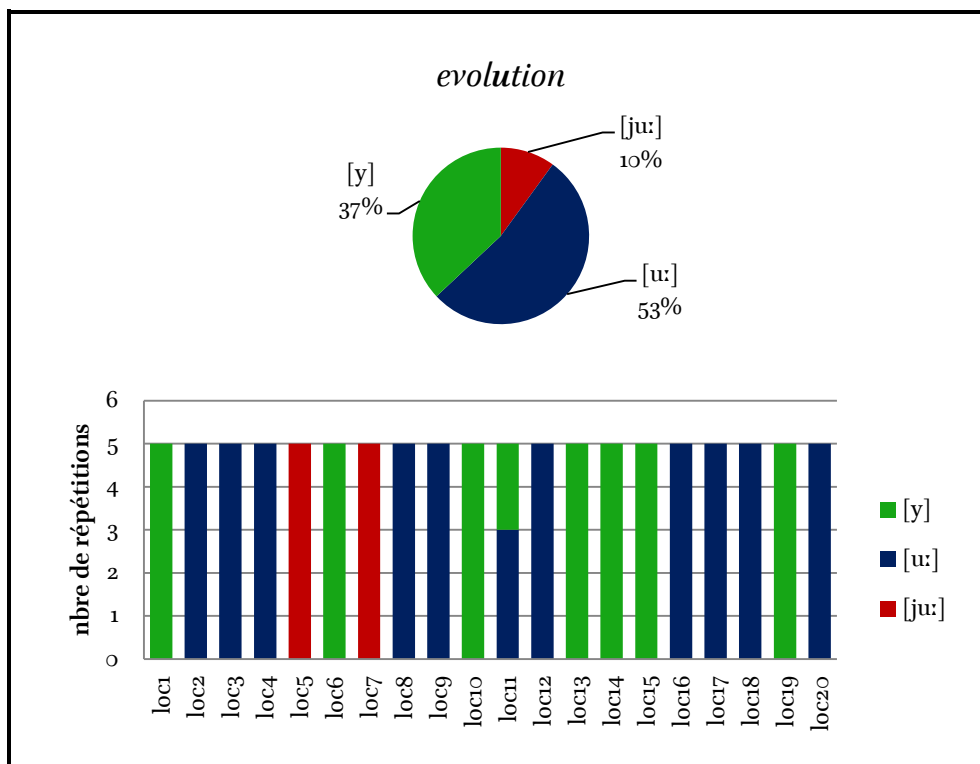


Figure 42 : pourcentages et répartition par locuteur de la production du <u> de *evolution*.

7.3.4 Résultats d'ensemble pour le yod-dropping

En réunissant la production des sons étudiés dans les trois mots *tune*, *new* et *evolution*, nous pouvons noter la prédominance de la réalisation du [ju:] à 55 %, celle du [u:] à 32 % et celle du [y] à 13 % sur les 300 répétitions. Le détail montre que le locuteur 5 réalise ces trois mots (15 répétitions) avec un [ju:] alors que les locuteurs 3 et 12 adoptent entièrement une prononciation en [u:]. Les autres mélangent deux (locuteurs 1, 4, 6, 7, 8, 9, 13, 15 à 20) ou trois réalisations (2, 10, 11 et 14). Les productions en [y] sont à 13 % (sur les 300 occurrences), notamment en raison des mots *tune* et *evolution*. Le mot *new* n'est pas produit avec un [y] (Figure 43).

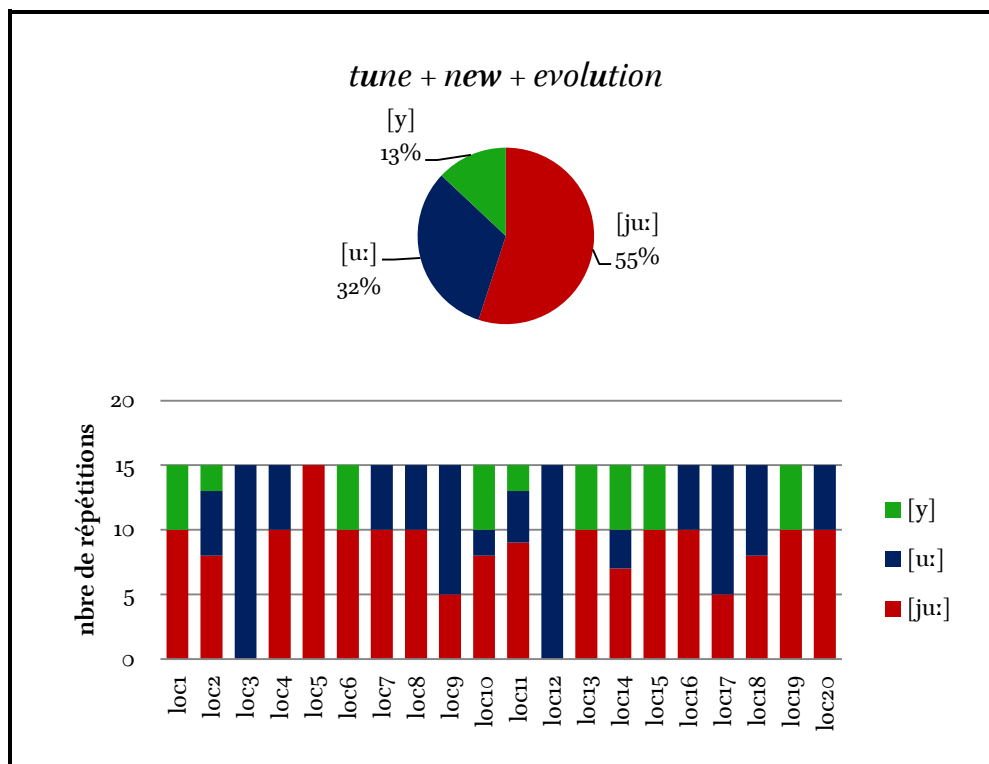


Figure 43 : pourcentages et répartition par locuteur du yod-dropping.

7.3.5 Discussion

Les résultats de la prononciation des mots *tune*, *new* et *evolution* suggèrent une influence de la lecture des graphies <u> et <ew>. Outre la prononciation traditionnelle et la plus répandue qui est celle de la réalisation en [ju:], les locuteurs ont tendance à utiliser le son [y] comme en français, à cause de la graphie <u> et aussi parce que ces mots existent en français (*tune* et *évolution*). Cependant, la part des productions de [y] dans les mots *tune* et *evolution* ne pourrait être uniquement justifiée par l'influence de la lecture, qui aurait systématiquement entraîné une répartition quasi-équivalente pour les deux mots, ce qui n'est pas le cas. En effet, *tune* ne connaît que deux répétitions en [y], contre 39 pour *evolution*. La raison devrait se trouver ailleurs et nous l'ignorons. Une hypothèse probable est que le mot *tune* est plus fréquemment rencontré que le mot *evolution* et ceci sous sa prononciation britannique. Cette hypothèse serait notamment validée puisque le mot *tune* apparaît 564 fois contre seulement 23 pour *evolution* dans le corpus SOAP. Quant au mot *new*, la lecture de <ew> pourrait également être mise en cause, du fait de la forte dominance de la prononciation en [ju:], et expliquerait pourquoi il n'y a pas eu présence du son [y], puisque le mot n'en contient pas. Mais, là encore, ce postulat ne semble pas entièrement valide, parce que le mot est pourtant réalisé à 33 % avec un [u:], alors qu'il n'y a pas de présence de

la graphie <u>. Ce pourcentage n'est pas négligeable et ne peut donc être imputable qu'à l'effet de fréquence d'exposition.

Nous venons d'exposer les résultats des trois éléments segmentaux consonantiques que sont le /t/ intervocalique, la rhoticité et le yod-dropping. Les productions proches du GA ont été dominantes à 47 %, suivies par celles de la RP (36) et en enfin celles des « Autres » (13 %). Toutefois, il faut remarquer que les caractéristiques de la RP ont occupé la première place uniquement pour les éléments du /t/ intervocalique et du yod-dropping. Les spécificités du GA viennent en tête notamment grâce à la rhoticité, à l'exception du mot *organisation*. Le résumé de ces résultats se trouve dans le Tableau 20.

Tableau 20 : résumé des éléments segmentaux consonantiques.

Phénomène étudié	Mots	% de caractéristiques de la RP	% de caractéristiques du GA	% de caractéristiques « Autres »
/t/ intervocalique	<i>City</i>	37	23	40
	<i>Vitamin</i>	60	16	24
	<i>party</i>	45	34	21
	<i>total</i> ²⁰⁹	48	24	28
Yod-dropping/rhoticité	<i>More</i>	30	70	0
	<i>Organisation</i>	52	48	0
	<i>Dark</i>	37	63	0
	<i>Party</i>	31	69	0
	<i>Bird</i>	11	66	23
	<i>Actor</i>	25	55	20
	<i>Leisure</i>	31	69	0
	<i>total</i>	31	63	6
Yod-dropping	<i>Tune</i>	88	10	2
	<i>New</i>	67	33	0
	<i>Evolution</i>	10	53	37
	<i>total</i>	55	32	13
Total des éléments consonantiques (sur 1300 répétitions)		40	47	13

Nous avons aussi noté une variabilité dans toutes les productions à deux niveaux : entre les locuteurs (inter-locuteur) et au sein des productions d'un même locuteur (intra-locuteur). Enfin, nous présentons les résultats des éléments segmentaux vocaliques dans les prochaines sections. Il

²⁰⁹ Total des répétitions des mots analysés. Par exemple ici il s'élève à 300 (100 fois 3).

s'agit de la prononciation de la voyelle de LOT, du GOAT²¹⁰, de la prononciation des graphies <i>, <y>, <ei> et <e> et enfin la nasalisation vocalique.

7.4 La prononciation de la voyelle de LOT

L'étude de la qualité de la voyelle <o> dans les mots *hot* et *top* montre qu'en plus des deux types de réalisations connus (/ɒ/ pour la RP et /ɑː/ pour le GA), il y a une troisième production qui émerge. Il s'agit du son /ɔ/ comme dans le mot français *pote* (/pɔt/).

7.4.1 Hot

Sur un total de 100 répétitions, le son [ɒ] est largement produit pour le mot *hot* (60 %). Le son [ɑː] est réalisé à 20 % ainsi que le son [ɔ] pour le même mot. Les locuteurs 4, 7, 8 et 9 réalisent complètement le son [ɑː] et les locuteurs 2, 11, 14 et 20 ne réalisent que le son [ɔ]. Le son [ɒ] est partout réalisé par les locuteurs 1, 3, 5, 6, 10, 12, 13, 15, 16, 17, 18 et 19. Notons qu'il n'y a aucun mélange (Figure 44).

²¹⁰ Nous utilisons la notation du système lexical de Wells (1982) et repris par Brulard, Carr, Durand, & Navarro (2015). Voir chapitre Autour de l'anglais pour des rappels (section 2.3.2.1, p. 39).

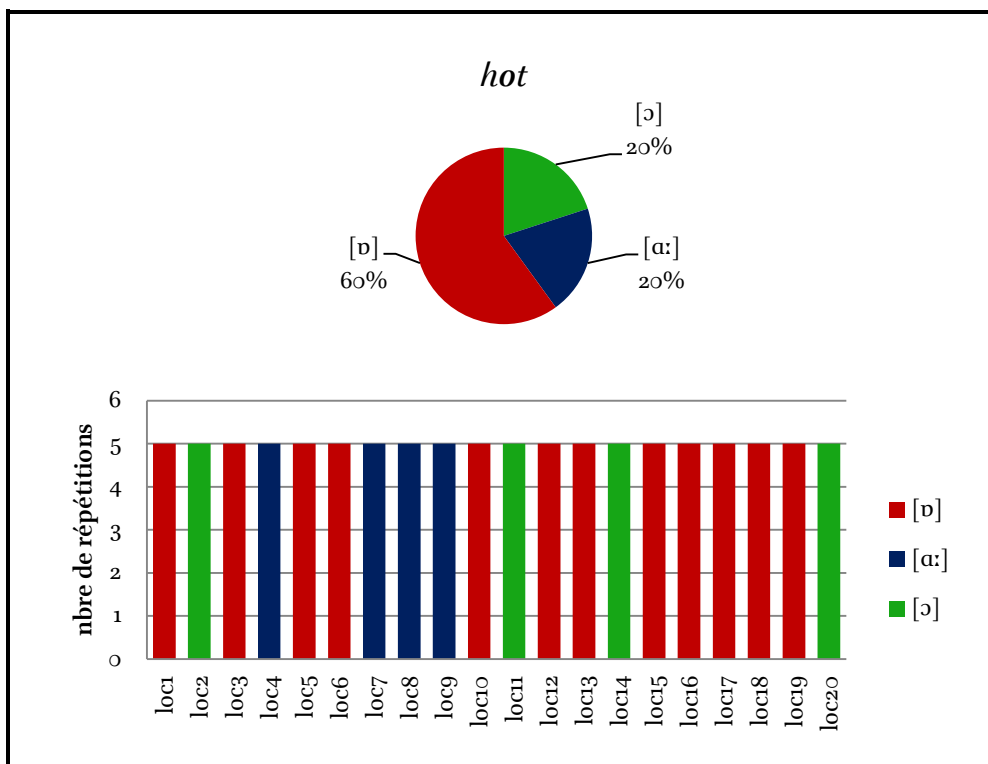


Figure 44 : pourcentages et répartition par locuteur de la production du <o> de *hot*.

7.4.2 Top

De manière assez intéressante, les mêmes résultats du mot *hot* (pourcentages globaux et répartition par locuteur) sont obtenus pour la prononciation du mot *top*, toujours sur 100 répétitions, à une exception près : le locuteur 8 produit cette fois un [ɒ] pour toutes ses répétitions (Figure 45).

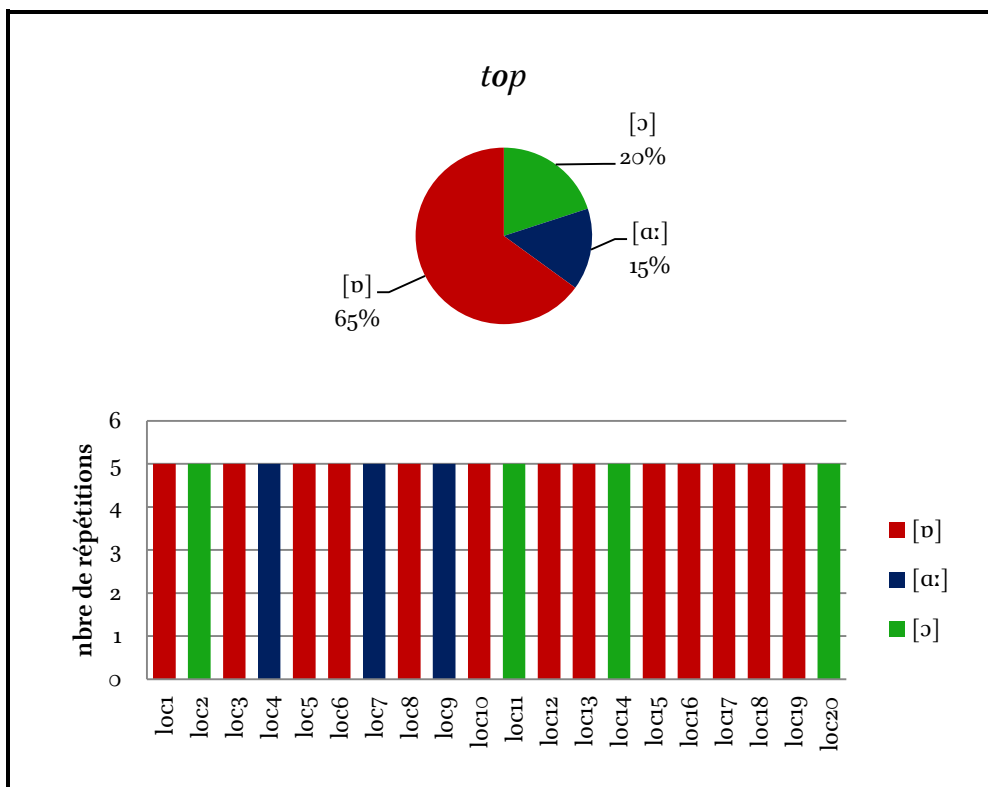


Figure 45 : pourcentages et répartition par locuteur de la production du <o> de *top*.

7.4.3 Hot et top

Sur les 200 cas de productions, les pourcentages généraux et la répartition par locuteur restent inchangés sauf pour le locuteur 8 qui produit des [ɑ:] pour le mot *hot* et des [ɒ] pour *top*, comme nous l'avons indiqué précédemment (Figure 46). Enfin, il semble y avoir une stabilité dans la production de cette voyelle chez chaque locuteur, puisqu'il n'y a pas eu de mélanges au sein de la prononciation de ces deux mots.

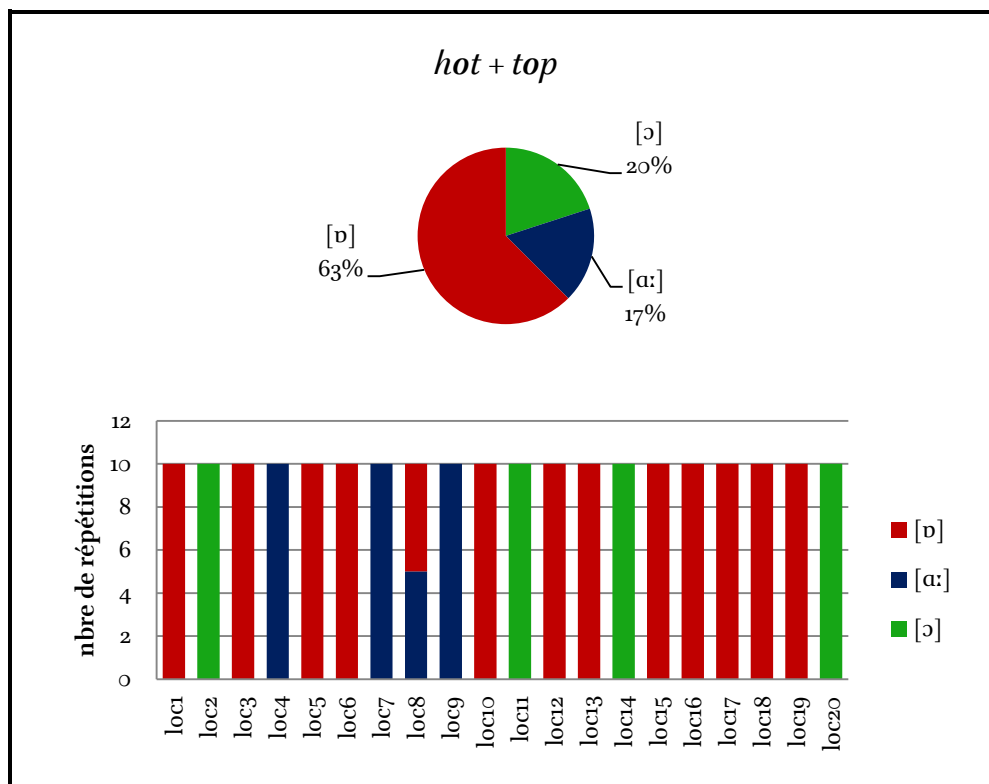


Figure 46 : pourcentages et répartition par locuteur de la production du <o> de *hot* et *top*.

7.5 La prononciation de la voyelle de GOAT

En plus de la série de sons /əʊ/ utilisée en RP et celle de /oʊ/ en GA, les sons /ɔ/ (/ɔm/ de homme) et /o/ (/po/ de pot) sont identifiés dans les analyses acoustiques des mots *joke* et *road*.

7.5.1 Joke

Le mot *joke* est produit avec une dominance de la série de sons [oʊ] (75 %) pour les 100 occurrences. Les sons et séries de sons [əʊ], [ɔ] et [o] sont respectivement produits à 10 %, 8 % et 7 % par l'ensemble des locuteurs. Les résultats individuels montrent les répartitions suivantes. Les locuteurs 4 et 7 produisent des [əʊ], le locuteur 2 produit des [ɔ] et le locuteur 20 produit des [o]. Le locuteur 11 est le seul à mélanger [ɔ] et [o]. Les autres locuteurs (1, 3, 5, 6, 8, 9, 10, 12 à 19) ont une réalisation complètement en [oʊ] (Figure 47).

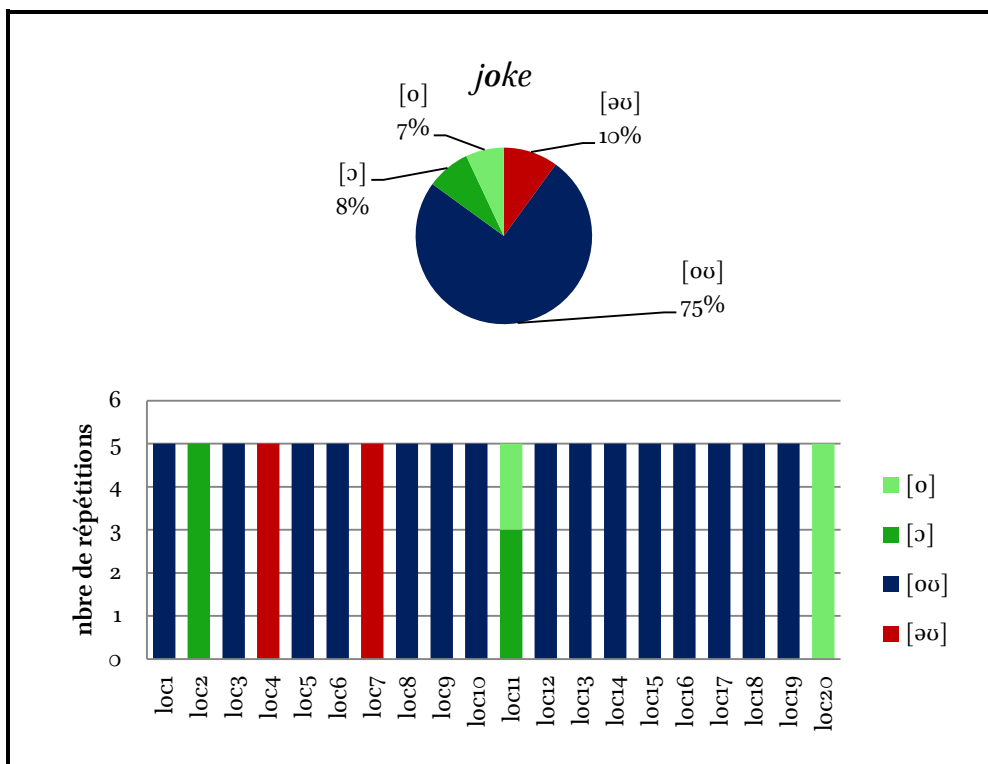


Figure 47 : pourcentages et répartition par locuteur de la production de <o> de *joke*.

7.5.2 Road

Un schéma presque similaire s'observe pour le mot *road* avec une forte dominance de la diphtongue [ou] (80 %), suivie de [o] (9 %), de [ɔ] (6 %) et de [əu] (5 %), sur les 100 répétitions. Les locuteurs 1, 3 à 6, 8, 9, 10, 12 à 17, 19 et 20 ne produisent que des [ou]. Les locuteurs 2 et 7 gardent les mêmes sons ([ɔ] et [əu] respectivement). Le locuteur 18 produit cette fois des [o] et le locuteur 11 reste sur le mélange avec [ɔ] et [o] (Figure 48).

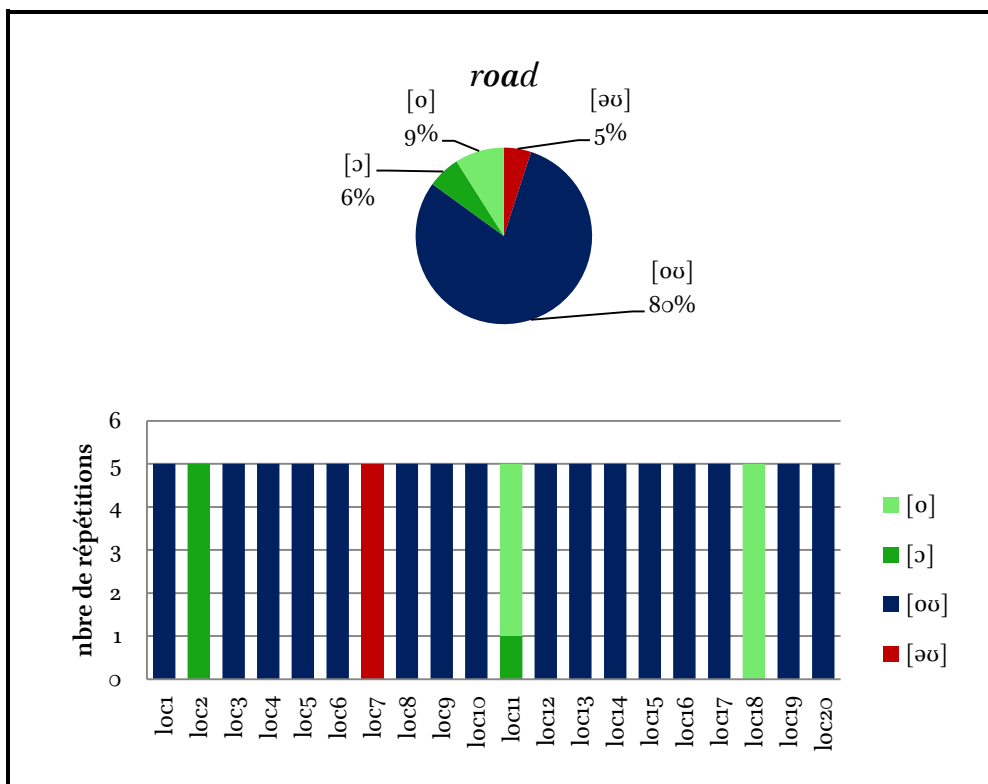


Figure 48 : pourcentages et répartition par locuteur de la production de <oa> de *road*.

7.5.3 Joke et road

Au total, 78 % des 200 occurrences des deux mots *joke* et *road* sont produites avec un [ou]. Le son [o] est réalisé avec 8 % et les sons [ou] et [ɔ] sont *ex-aequo* avec 7 % des réalisations. En fin de compte, les locuteurs 1, 3, 5, 6, 8, 9, 10, 12 à 17, 19 gardent leur production totale avec un [ou]. Rien n'a changé pour les locuteurs 2 et 7 avec respectivement [ɔ] et [əu]. Viennent ensuite les locuteurs qui procèdent à un mélange : [əu] et [ou] pour le locuteur 4 ; [ɔ] et [o] pour le locuteur 11 et enfin [ou] et [o] pour les locuteurs 18 et 20 (Figure 52).

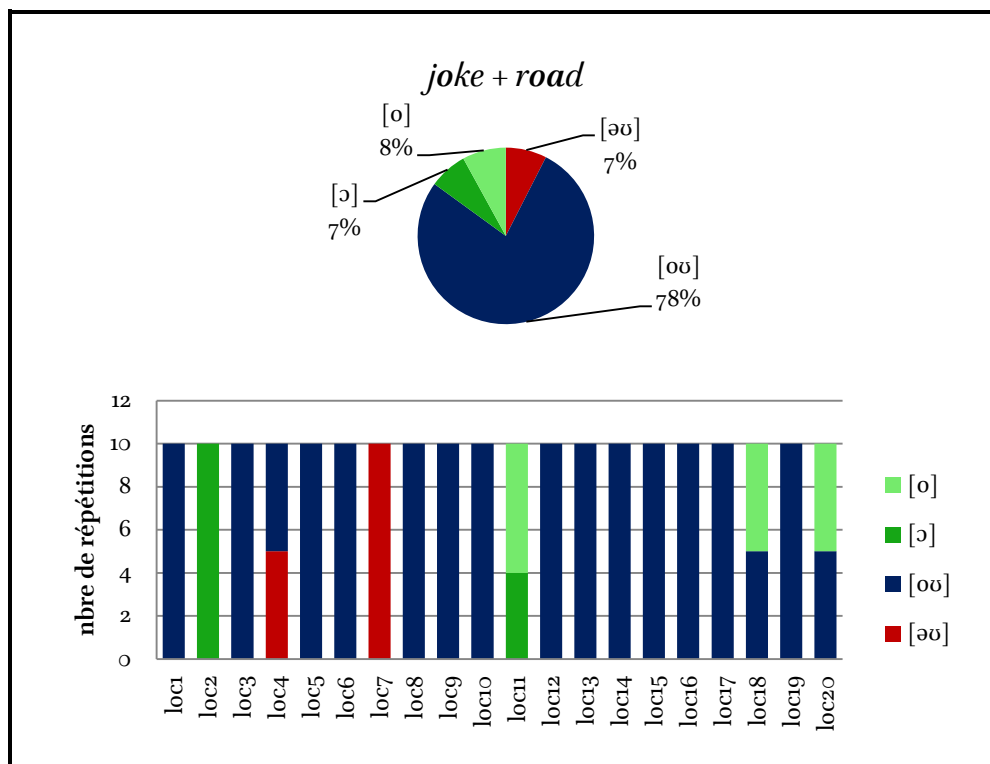


Figure 49 : pourcentages et répartition par locuteur de la production de <o> et <oa> de *joke* et *road*.

7.5.4 Discussion

La graphie <o> dans le mot *joke* (ainsi que *hot* et *top*) pourrait voir sa prononciation se justifier en [o] et en [ɔ] dans une certaine mesure, mais pas celle de la graphie de <oa> de *road*. En revanche, la suite des deux voyelles de la graphie <oa> de *road* pourrait inciter à la prononciation avec la diphtongue [oʊ]. L'hypothèse de l'influence de l'orthographe pourrait ainsi être écartée pour la production en [ɔ] du mot *road*. Il en va de même pour les deux mots (*joke* et *road*), puisque l'on aurait pu s'attendre à ce que l'un ou l'autre mot soit prononcé avec des sons différents. Nous retenons également la part de la réalisation en [oʊ] pour ces deux mots (75 % et 80 %), malgré les différences de graphies des voyelles étudiées. Enfin, il convient de noter la quasi-régularité dans la prononciation de la majorité des locuteurs.

7.6 La prononciation des graphies <i> et <y>

La première syllabe des mots *dynasty* et *vitamin* est prononcée avec une monophthongue /ɪ/ en RP (/ˈdɪnəsti/, /ˈvɪtəmm/) et plutôt avec une diphtongue /aɪ/ en GA (/ˈdɑməsti/, /ˈvaɪrəmm/).

7.6.1 Dynasty

Dans nos résultats, la graphie <y> de la première syllabe du mot *dynasty* est réalisée à 54 % avec [ai], et à 46 % avec [ɪ] sur 100 occurrences. Ce dernier est produit à cent pour cent par les locuteurs 1, 5, 7, 11, 13, 14, 15 et 18 pendant que [ai] est réalisé entièrement par les locuteurs 2, 4, 6, 8, 9, 10, 12 et 19. Les locuteurs 3, 16, 17 et 20 mélangent [ɪ] et [ai] (Figure 50).

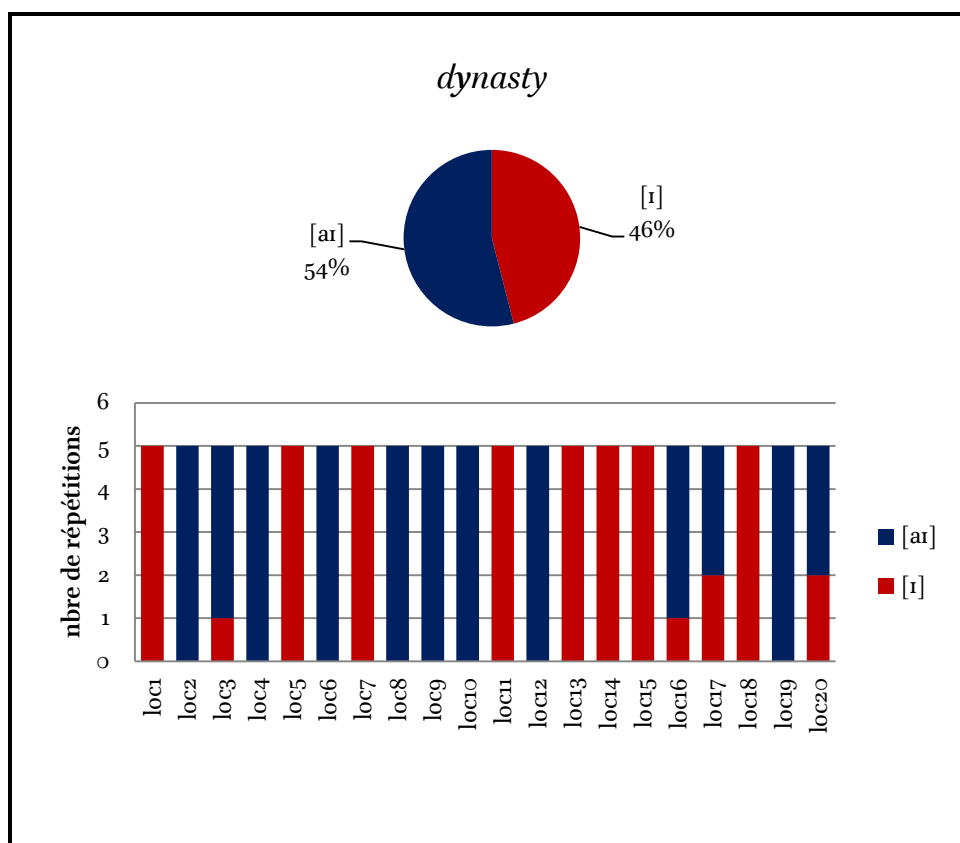


Figure 50 : pourcentages et répartition par locuteur de la production du premier <y> de *dynasty* en VR.

7.6.2 Vitamin

La graphie de la première syllabe du mot *vitamin* est majoritairement prononcée avec un [ɪ] (62 %) et avec un [ai] (38 %). Les locuteurs 1, 4, 5, 7, 8, 9, 11, 13, 14, 15, 16 et 20 réalisent toutes les répétitions avec le son [ɪ] tandis que les locuteurs 2, 3, 10, 12, 18 et 19 produisent entièrement le son [ai]. Cependant, les locuteurs 6 et 17 réalisent chacun un mélange des deux sons (Figure 51).

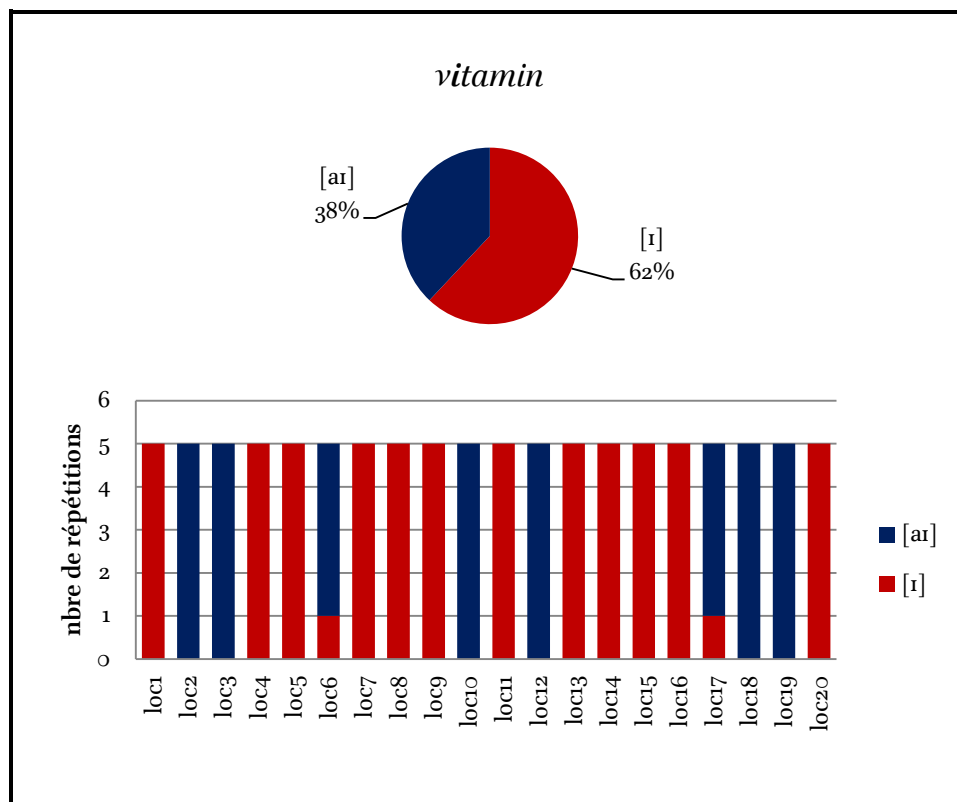


Figure 51 : pourcentages et répartition par locuteur de la production de <i> de *vitamin*.

7.6.3 Dynasty et vitamin

Les premières voyelles dans les mots *dynasty* et *vitamin* (<y> et <i>) sont produites à 54 % avec un [i] et à 46 % avec un [ai] sur un total de 200 répétitions. Le son [i] est entièrement réalisé par les locuteurs 1, 5, 7, 11, 13, 14 et 15. Le son [ai] est également produit totalement par les locuteurs 2, 10, 12 et 19. Les locuteurs 3, 4, 6, 8, 9, 16, 17, 18 et 20 mélangent respectivement [i] et [ai] (Figure 52).

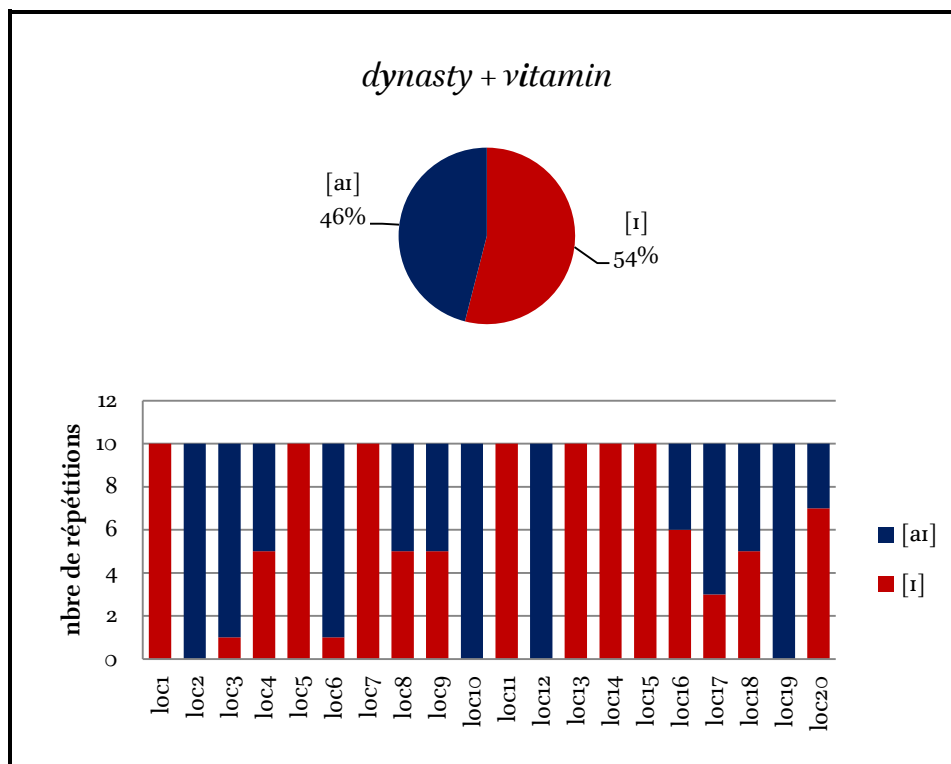


Figure 52 : pourcentages et répartition par locuteur de la production de <y> et <i> de *dynasty* et *vitamin*.

7.7 La prononciation de la graphie <i>

La voyelle de la troisième syllabe des mots *organisation* et *legalisation* (<i>) est produite avec soit la diphtongue [aɪ], soit la monophthongue [ɪ] en RP et est plutôt réalisée avec un [ə] en GA.

7.7.1 Organisation

Sur les 100 répétitions du mot *organisation*, le son [aɪ] est réalisé 76 fois et le son [ɪ] 9 fois, soit un total de 84 % de prononciation proche de la RP. La prononciation proche du GA n'arrive qu'à 16 %. Les locuteurs 13, 15 et 16 produisent toutes leurs répétitions avec un [ə]. Le locuteur 10 mélange [aɪ] et [ə] et le locuteur 12 réalise des [aɪ]. Les autres locuteurs (1 à 9, 11 et 14, 17 à 20) ne produisent que des sons [ɪ] (Figure 53).

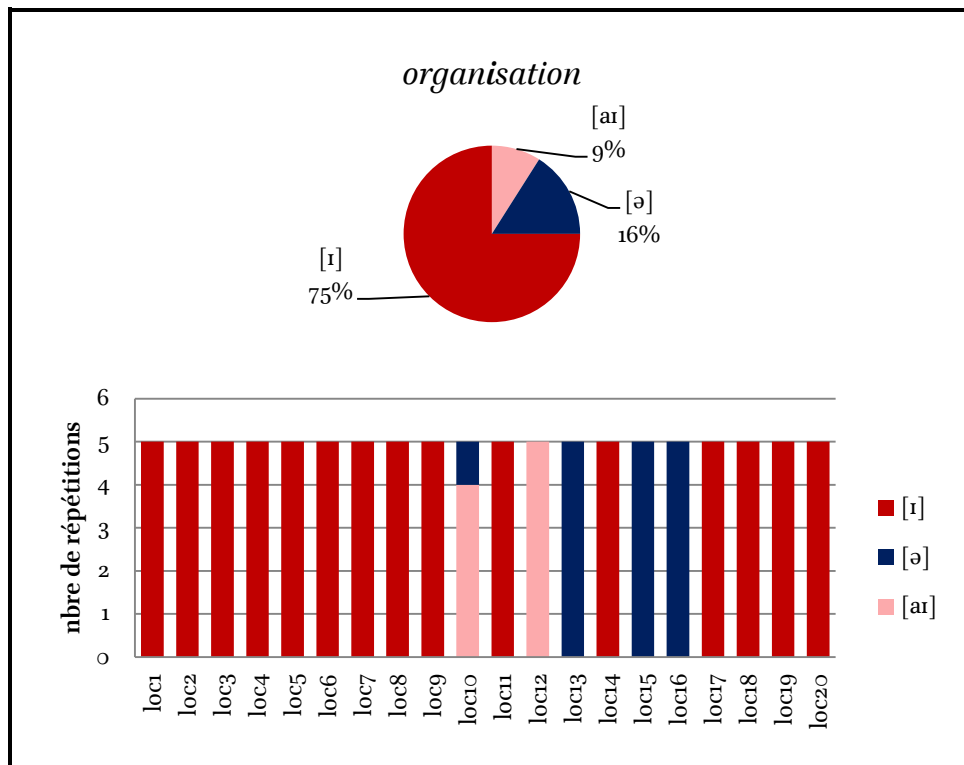
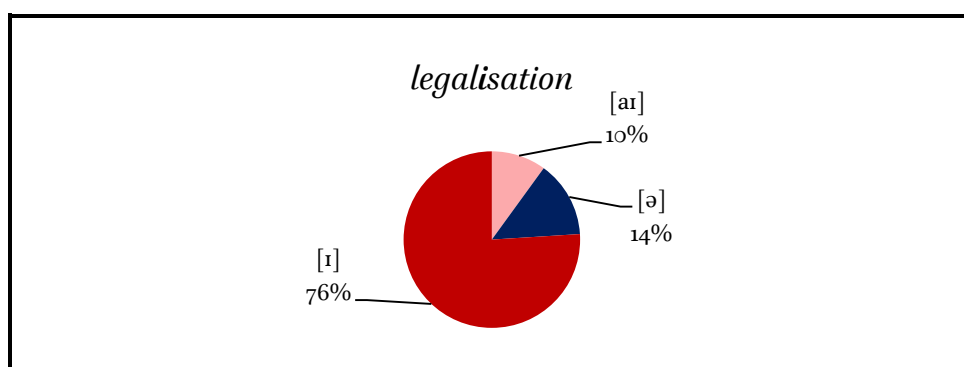


Figure 53 : pourcentages et répartition par locuteur de la production du premier <i> de *organisation*.

7.7.2 Legalisation

Des pourcentages similaires que précédemment sont obtenus pour le mot *legalisation*. [ɪ] est produit à 76 %, [aɪ] à 10 % (86 % proche de la RP) et [ə] à 14 % (proche du GA). Les locuteurs 10 et 12 ne produisent que des [aɪ]. Les locuteurs 17 et 18 ne réalisent que des [ə]. Pendant que le locuteur 8 mélange des [ə] et [ɪ], les autres locuteurs ne réalisent que des [ɪ] (Figure 54).



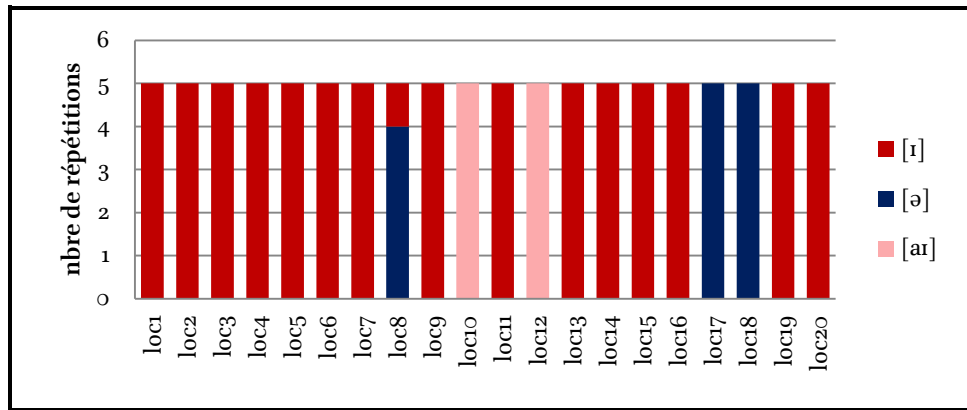


Figure 54 : pourcentages et répartition par locuteur de la production du premier <i> de *legalisation*.

7.7.3 Organisation et legalisation

Nous retrouvons le même schéma discuté plus haut concernant les pourcentages pour les deux mots *organisation* et *legalisation*. Sur 200 répétitions, [ɪ] est produit à 76 %, [aɪ] à 9 % (85 % proche de la RP) et [ə] à 15 % (proche du GA). Les locuteurs 1 à 7, 9, 11, 19 et 20 sont constants dans leurs productions du son [ɪ] pour toutes les occurrences. Il en va de même pour le locuteur 12 avec le son [aɪ]. Le locuteur 10 réalise [aɪ] et [ə]. Un mélange de [ə] et de [ɪ] est observé chez les locuteurs 8, 13, 15, 16, 17 et 18 (Figure 55).

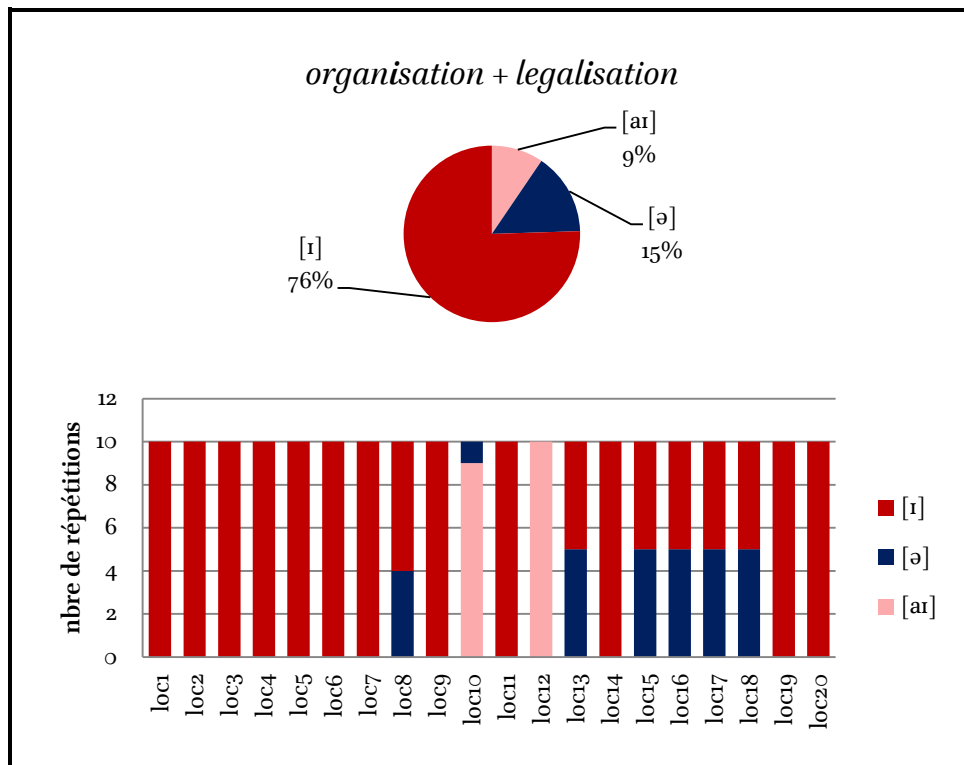


Figure 55 : pourcentages et répartition par locuteur de la production du premier <i> de *organisation* et *legalisation*.

7.7.4 Discussion

La prononciation de la lettre <i> constitue une des difficultés majeures pour les apprenants d'anglais L2, car aucune règle phonologique n'explique s'il faudrait prononcer [ɪ] ou [aɪ] (Ginésy, 2000 ; Herry-Bénil, 2011). À cette difficulté existante s'ajoute le fait que le <i> est parfois prononcé [ɪ] ou [aɪ] selon la norme prise en considération. Le mélange des deux prononciations surtout chez un même locuteur pourrait être le résultat de l'exposition aux deux variétés d'anglais et cela est tout à fait concevable. Qu'en est-il des prononciations uniformes ? Une hypothèse probable concernant les locuteurs qui adoptent une production homogène (soit [ɪ] ou soit [aɪ]), pourrait être le fait qu'ils choisissent la voie de la facilité. Sans vouloir se mettre en danger, ils préfèrent prononcer les deux mots avec le même son. Nous pouvons le remarquer chez 11 locuteurs. Sinon, est-il possible qu'ils prononcent ces mots comme ils les ont appris ou entendu ? Cela revient-il à considérer qu'ils n'ont jamais rencontré l'autre variété ? C'est moins probable, surtout lorsque les deux sons peuvent être utilisés en RP, alors que le GA conserve le [aɪ] partout (LPD, EPD entrée *vitamin* et *dynasty*).

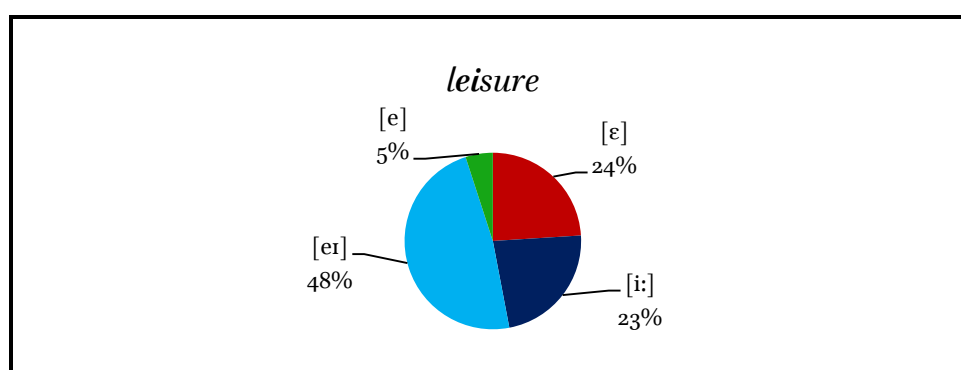
Aussi, le fait que ces mots existent en français (*vitamine* et *dynastie*) pourrait justifier le fort taux de leur production avec un [ɪ]. Les mêmes remarques sont valables pour les mots *organisation*

et *legalisation* qui existent également en français (organisation et légalisation). C'est un constat non négligeable mais qui n'explique pas la part du [aɪ] qui est réalisé à 46 %, soit presque la moitié des répétitions. D'autres facteurs interviendraient, mais que nous ne pouvons expliquer, sinon remettre l'hypothèse de la fréquence d'exposition. Toutefois, un dernier constat mérite d'être pointé. Nous constatons que le mot *vitamin* est produit avec plus de [ɪ] (62 %) tandis que le mot *dynasty* est plus réalisé avec un [aɪ] (54 %). Nous pourrions supposer que la graphie <y> incite plus nos locuteurs à prononcer ce mot avec un [aɪ]. Mais, là également nous restons prudents quant aux hypothèses que nous avançons.

7.8 La prononciation de la graphie <ei> dans *leisure*

La suite de voyelles de la première syllabe du mot *leisure* connaît une seule prononciation en RP, celle de /ɛ/ ('lɛzə). En plus de ce dernier (/ɛ/), le GA emploie deux autres réalisations différentes pour la graphie <ei> : /i:/ ('li:zə) et /eɪ/ ('lerzə) (LPD entrée *leisure*). Outre ces trois sons, /e/ est également identifié dans les résultats, faisant au total quatre productions différentes, rien que pour la première syllabe de ce mot²¹¹.

Sur les 100 occurrences, le son [eɪ] occupe presque la moitié de la production totale (48 %). Les sons [ɛ] et [i:] sont respectivement réalisés à 24 % et 23 %. Le son [e] est seulement produit cinq fois (5 %). La répartition individuelle montre que les locuteurs 2, 6, 7, 11, 13, 14, 15, 18 et 19 ne produisent que des [eɪ]. Les locuteurs 1, 8, 9 et 16 ne réalisent que des [i:] et les locuteurs 4, 10, 12 et 17 utilisent un [ɛ] pour toutes leurs répétitions. Enfin, le locuteur 5 mélange [ɛ] et [eɪ] tandis que le locuteur 20 mélange [ɛ] et [i:] (Figure 56).



²¹¹ La deuxième syllabe est aussi prononcée différemment dans les deux variétés en raison de la rhoticité (/zə/ pour la deuxième syllabe en GA /zə/ en RP).

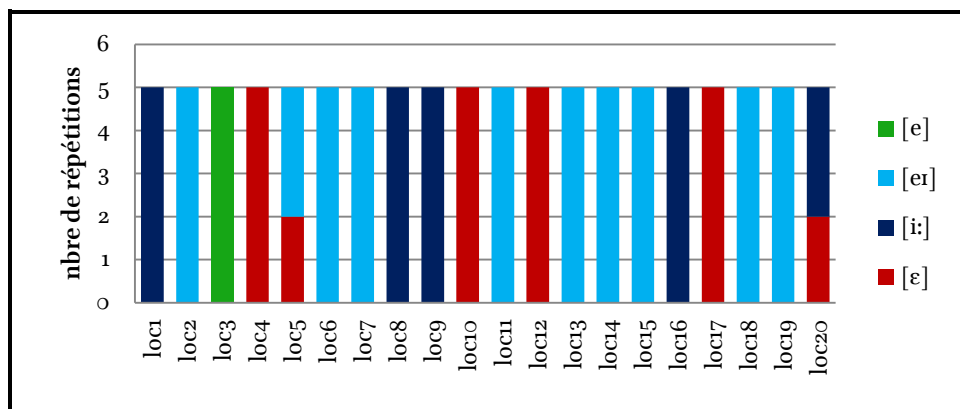


Figure 56 : pourcentages et répartition par locuteur de la production de <ei> de *leisure*.

7.9 La prononciation de la graphie <e> dans evolution

Le mot *evolution* se prononce de deux façons dans les deux variétés de référence. La première syllabe représentée par la lettre <e> se prononce /i:/ (,i:və'ljʊ:fən) ou /ɛ/ (,ɛvə'ljʊ:fən²¹²). Cependant, le *Longman Pronunciation Dictionary* (LPD) indique que la prononciation avec un /ɛ/ est généralement utilisée en GA (à 85 %). Seulement 15 % des Américains prononcent le mot avec un /i:/ (LPD entrée *evolution*). En revanche, dans nos analyses acoustiques, une troisième réalisation apparaît dans les résultats, il s'agit de la réalisation de la voyelle <e> en [e], comme dans le mot français dé. L'émergence de ce son peut être considérée comme étant l'influence de la L1. Sur les 100 répétitions, le son [e] est majoritairement produit (61 %), suivi par le [i:] à 29 % et enfin le [ɛ] à 10 %. Le son [e] est réalisé entièrement par les locuteurs 2, 3, 5, 6, 12, 13, 17, 18 et 19. Quant au son [i:], il est complètement réalisé par les locuteurs 4, 7, 15 et 16 alors que son [ɛ] n'est réalisé entièrement que par le locuteur 1. Le locuteur 10 mélange des [i:] et [e] ; le locuteur 11 mélange des [ɛ] et [e] ; le locuteur 14 mélange des [i:] et [e] et enfin le locuteur 20 mélange les sons [i:], [ɛ] et [e] (Figure 57).

212 Il est à noter que le /ju:/ devient un /u:/ en GA.

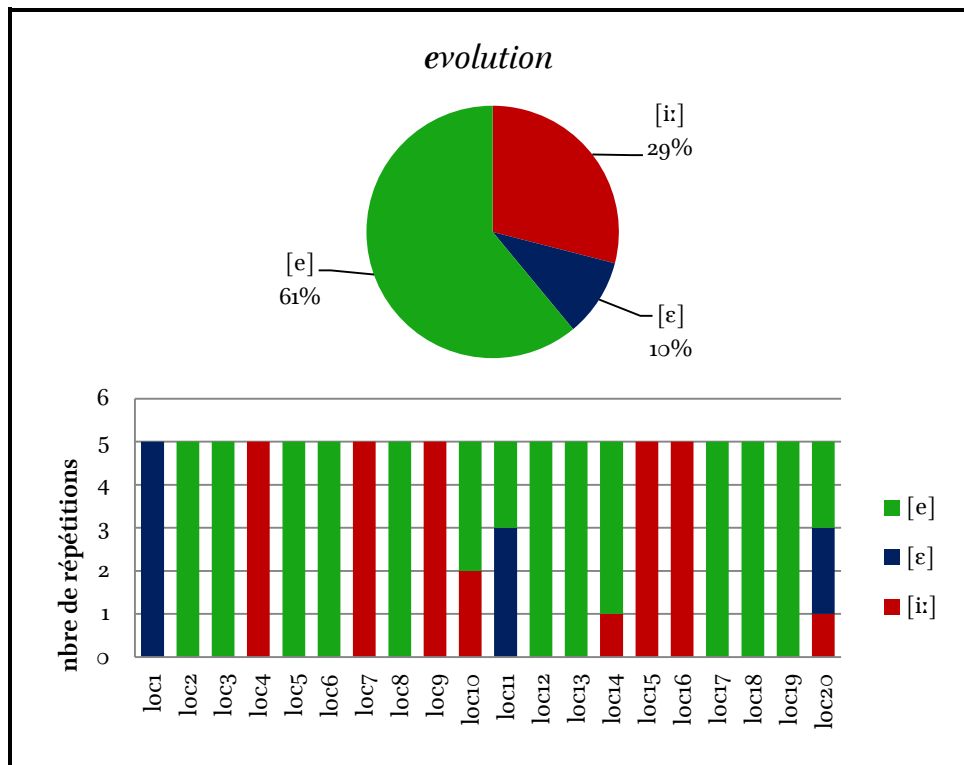


Figure 57 : pourcentages et répartition par locuteur de la production de <e> de *evolution*.

7.9.1 Discussion

L'influence de l'orthographe du mot *leisure* sur ses différentes réalisations ne pourrait être exclue, si nous considérons que la graphie <ei> incite à la réalisation dominante /eɪ/ (notons la ressemblance entre son et graphie). Si tel est le cas, pouvons-nous considérer la même chose pour la prononciation GA ? Nous nous demandons si la prononciation américaine par les natifs américains en /'leɪʒə/ ne serait pas aussi influencée par la graphie <ei>. En revanche, elle ne nous permet pas non plus d'expliquer les productions « Autres ». Sinon, en quoi la graphie <ei> suggère-t-elle les réalisations en /i:/ et en /ɛ/? Il faut rappeler que la graphie <ei> possède plusieurs prononciations (au moins quatre) selon les mots qui les contiennent. Tantôt réalisée en /eɪ/ comme dans *eight* (/eɪt/) et *weigh* (/weɪ/), tantôt en /i:/ comme dans *ceiling* (/ˈsi:lɪŋ/) et *seize* (/si:z/), tantôt en /aɪ/ comme dans *height* (/haɪt/) et *either*²¹³ (/ˈaɪðə) ou même parfois en /ɛ/ comme dans certains noms de villes anglaises (*Leister* (/ˈlestə/) et *Leicestershire* (/ˈlestəʃɪə/), cette graphie peut constituer une difficulté supplémentaire pour les apprenants d'anglais. Les mêmes

²¹³ Il est également à noter que *either* se prononce à 84 % avec un [i:] et à 16 % avec un [aɪ] en anglais américain alors qu'il est prononcé à 87 % avec un [aɪ] et à 13 % avec un [i:] en anglais britannique (LPD). Aussi, l'EPD souligne que la prononciation américaine avec un [i:] gagne de plus en plus du terrain chez les jeunes britanniques.

remarques que précédemment ont été faites pour le mot *evolution* (ainsi que *legalisation*²¹⁴), qui subit également les conséquences de « l'irrégularité » dans la prononciation de certains mots anglais, avec notamment la graphie <e> qui peut être prononcée tantôt en /ɛ/, en /i:/ ou en /ɪ/.

7.10 Nasalisation vocalique

Les analyses acoustiques des mots *band* et *hand* ont fait émerger quatre sortes des productions pour la série de graphies <an>. D'une part, il y a deux réalisations nasalisées [ã] et [æ̃] comme dans les mots français paon et pain, et d'autre part, deux productions non nasalisées /an/ et /æn/. Dans la présentation de ces derniers, nous ne prenons en compte véritablement que les voyelles /a/ et /æ/.

7.10.1 Band

La prononciation du mot *band* avec la voyelle nasalisée [æ̃] vient en tête avec 47 % sur les 100 occurrences. Elle est suivie par le son oral²¹⁵ [æ] à 38 %, puis par la nasale [ã] à 10 % et enfin par le son oral [a] à seulement 5 %. Cette dernière réalisation ([a]) n'est effectuée que par le locuteur 16. Les locuteurs 1 et 8 réalisent des [ã] pour toutes leurs répétitions. Les locuteurs 2, 3, 4, 5, 9, 10, 11, 15 et 20 ne produisent que des [æ̃], ainsi que les locuteurs 7, 12, 13, 14, 17, 18 et 19 pour le son [æ]. Seul le locuteur 6 mélange [æ] et [æ̃] (Figure 58).

²¹⁴ Prononcé à 45 % avec [i:], à 43 % avec [e] et 12 % avec [ɛ]. Nous ne présentons pas ces résultats, car il n'y a pas de différence entre les deux accents concernant la première syllabe : [i:] est utilisé partout aussi bien en RP qu'en GA.

²¹⁵ Aussi non nasal, par opposition à nasal.

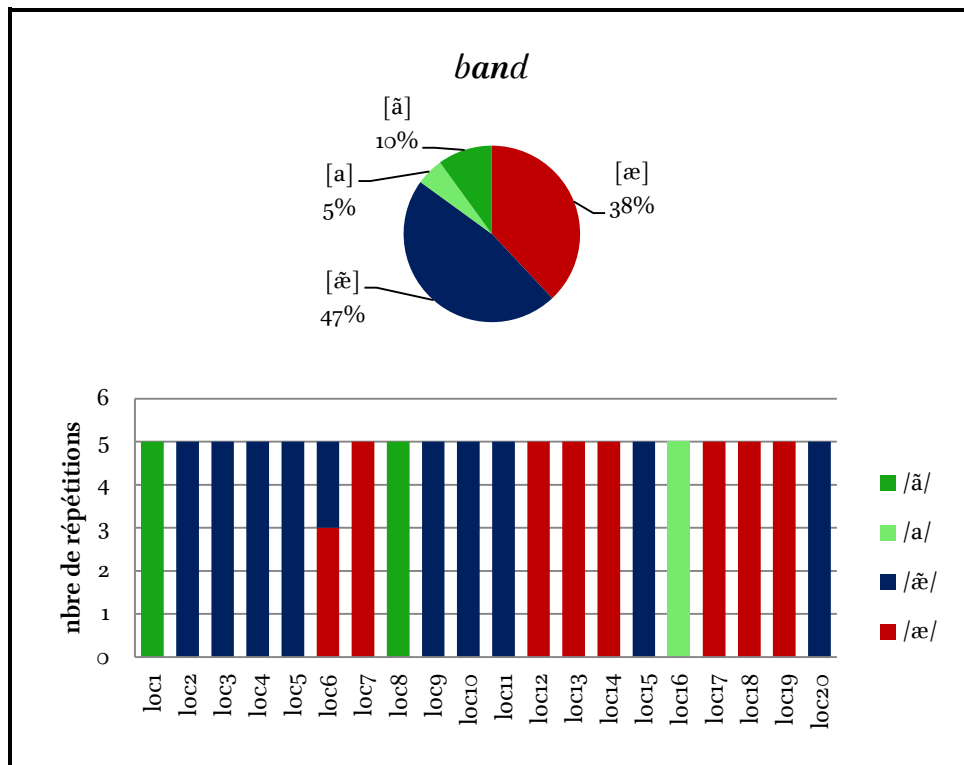


Figure 58 : pourcentages et répartition par locuteur de la production de <an> de *band*.

7.10.2 Hand

Pour ce qui concerne le mot *hand*, le locuteur 6 mélange toujours les réalisations, mais avec des [æ] et [ã] pendant que le locuteur 14 réalise cette fois un mélange de [æ] et de [ã]. Autrement, ce sont exactement les mêmes pourcentages que précédemment qui sont obtenus chez les autres locuteurs pour le mot *band* (sur 100 répétitions). Les locuteurs 1 et 8 réalisent des [ã] pour toutes leurs répétitions. Les locuteurs 2, 3, 4, 5, 9, 10, 11, 15 et 20 ne produisent que des [ã], ainsi que les locuteurs 7, 12, 13, 14, 17, 18 et 19 pour le son [æ]. Par conséquent, les pourcentages globaux passent à 50 % pour le [ã], 35 % pour le [æ], 10 % pour le [ã] et enfin 5 % pour le [a] (Figure 59).

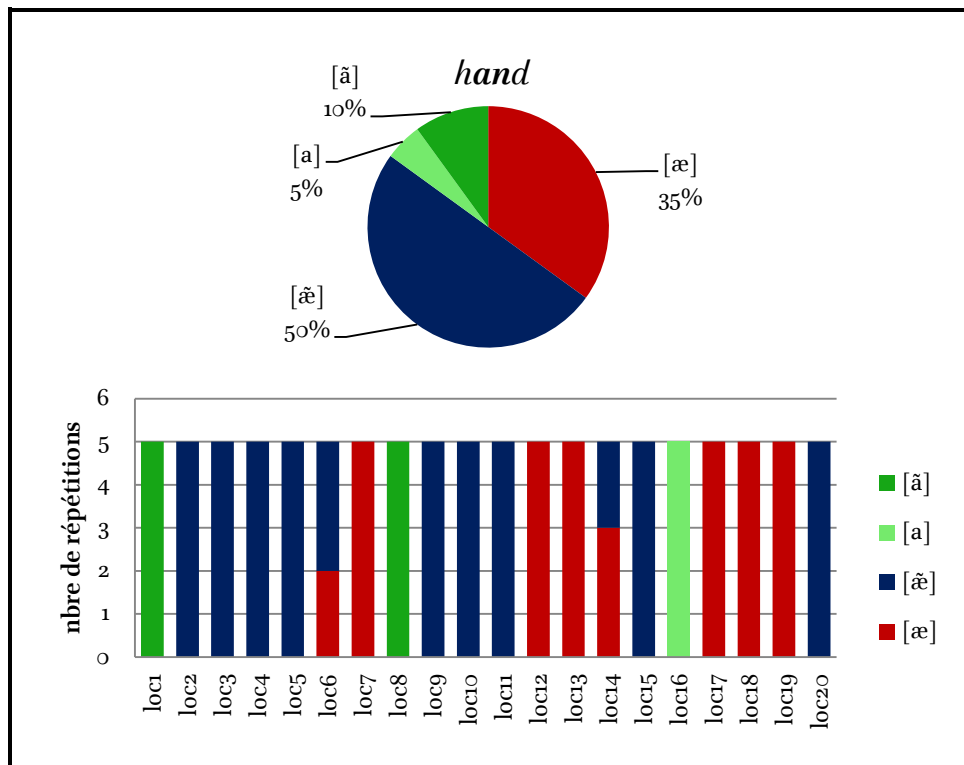


Figure 59 : pourcentages et répartition par locuteur de la production de <an> de *hand*.

7.10.3 Band et hand

Après avoir rassemblé les résultats des deux mots, nous remarquons des productions homogènes et des pourcentages identiques pour les 200 répétitions, sauf pour les locuteurs 6 et 14 qui mélangent respectivement des [æ] et [ã]. En fin de compte, la réalisation du [a] stagne sur 5 % ainsi que celle du [ã] à 10 %. [ã] reste en tête avec 49 % suivi du [æ] à 36 % (Figure 60).

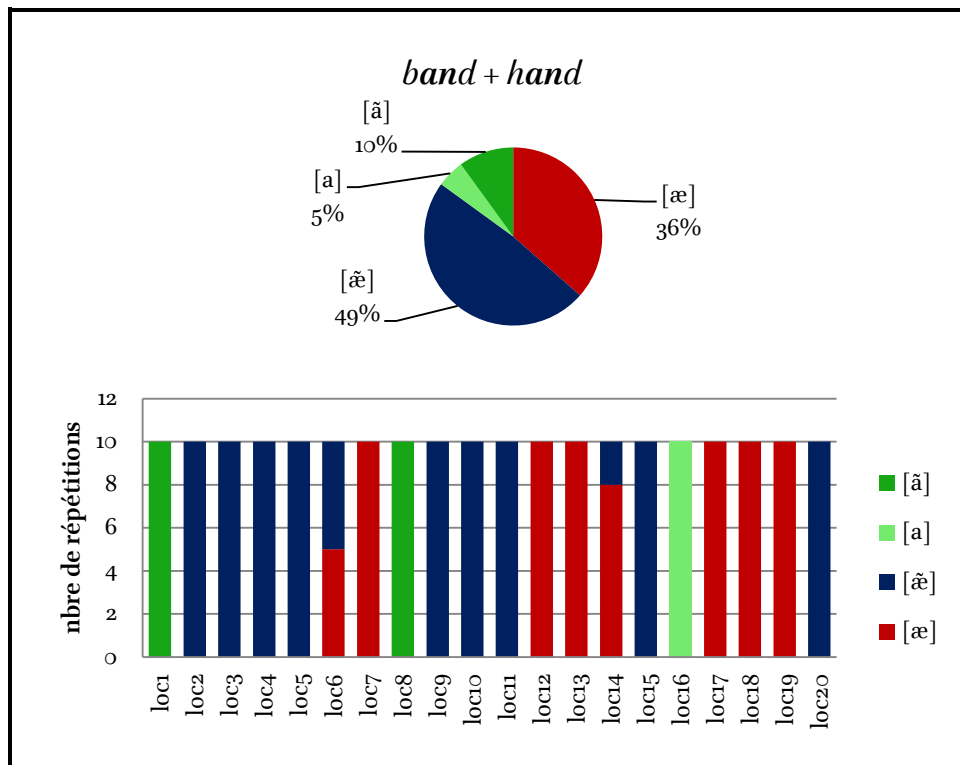


Figure 60 : pourcentages et répartition par locuteur de la production de <an> de *band* et *hand*.

7.10.4 Discussion

La prononciation des mots *band* et *hand* avec la voyelle nasalisée [ã] émane sans nul doute de la prononciation des mots français contenant les mêmes graphies (<an>). Nous voyons encore une fois la notion de nativisation²¹⁶ à l'œuvre. Cette caractéristique bien particulière des apprenants à calquer la prononciation L2 sur celle de la L1 peut, dans une certaine mesure, justifier aussi les réalisations en [a]. Concernant la réalisation des sons [æ̃], là également il n'y a aucun doute qu'elle provient de l'exposition à l'accent américain et nous pouvons le dire avec certitude, car, même si la nativisation était à contribution, rien ne laisse suggérer la réalisation d'un tel son à partir de la graphie <an>. S'il avait été question de la graphie <in>, nous n'aurions pas pu faire cette déduction. Un autre point important qui mérite d'être soulevé concerne la transcription phonétique de ces deux mots dans les dictionnaires de prononciation EPD et LPD. Par conséquent, *hand* et *band* sont transcrits avec la même voyelle /æ/, alors qu'à l'écoute, l'on entend une différence claire du timbre vocalique entre la prononciation GA et RP. Bien qu'il n'y ait pas de documentations scientifiques

²¹⁶ En didactique des langues, ce terme désigne un processus par lequel un apprenant de langue construit son Apprentissage en se reposant sur des éléments langagiers de sa /ses L1 (ce que nous appelons aussi calquage). Voir Demaizière et Narcy-Combes (2005), Andersen (1983).

sur ce phénomène, la prononciation britannique aurait bien pu être transcrite avec un /ɑː/ (/bɑːnd/ et /hɑːnd/) et la prononciation américaine plutôt avec un /æ/ (/bænd/ et /hænd/). Tout comme c'est le cas pour les mots *dance*, *grant*, *plant* et *branch*, qui sont transcrits avec un /ɑː/ en RP et avec un /æ/ en GA (/dɑːnts/ vs /dænts/ ; /graːnt/ vs /grænt/ ; /plɑːnt/ vs /plænt/ et /brɑːntʃ/ vs /bræntʃ/). Nous émettons ainsi l'hypothèse que nos locuteurs qui prononcent les mots *band* et *hand* avec un /a/ (en dans une certaine mesure /ã/) se reposent plutôt sur la prononciation britannique des mots *dance*, *grant*, *plant*, *branch*, etc. Ce manque de transparence entre les graphies et leurs prononciations constituent un facteur qui induit en erreur énormément d'apprenants en L2. Que la graphie <a> soit prononcée tantôt /a/, /ɑː/ ou /æ/ ne relève pas de règles phonologiques en tant que telles, mais plutôt d'habitudes de prononciation assez variées des divers natifs anglophones. L'apprenant d'anglais qui s'expose à ces variations doit non seulement effectuer des efforts pour assigner une image constante sur un son, mais doit aussi trier et choisir une seule prononciation parmi celles disponibles, en vue d'en faire un modèle unique de prononciation qu'il devra adopter dans des situations de communication précises. Connaître les différentes façons de prononcer un mot selon les divers accents anglais peut aider l'apprenant à mieux se faire comprendre lorsqu'il interagit avec des locuteurs natifs spécifiques.

Cette section sur la présentation des résultats des éléments segmentaux vocaliques montre encore une fois une variabilité pour tous les locuteurs. Certains sons ou séquences de sons se rapprochent plus de la RP alors que d'autres ressemblent plus aux caractéristiques du GA. Les productions « Autres » restent toutefois significatives. Le tableau suivant résume la portion de chaque élément en fonction des mots dans lesquels ils se trouvent. Comme nous l'indiquons ci-dessous (Tableau 21), les éléments caractéristiques de la RP dominent largement les productions (55 %). La part de ceux du GA arrive à 29 % (grâce aux mots *dynasty*, *leisure*, *band* et *hand*), suivie des « Autres » (16 %). Ces résultats confirment le fait que les voyelles anglaises seraient plus difficiles à assimiler que les consonnes. Les sections suivantes exposent les résultats des deux éléments suprasegmentaux, qui sont : la place de l'accent principal et le nombre de syllabes prononcées.

Tableau 21 : résumé des éléments segmentaux vocaliques.

Phénomène étudié	Mots	% de caractéristiques de la RP	% de caractéristiques du GA	% de caractéristiques « Autres »

La voyelle de LOT	Hot	60	20	20
	Top	65	15	20
	total	63	17	20
La voyelle de GOAT	Joke	75	10	15
	Road	80	5	15
	total	78	7	15
La graphie <y> et <i>	Dynasty	46	54	0
	Vitamin	62	38	0
	total	54	46	0
La graphie <i>	Organisation	75	16	9
	Legalisation	76	14	10
	total	76	15	9
La graphie <ei>	Leisure	24	71	5
La graphie <e>	Evolution	29	10	61
Nasalisation vocalique	band	38	47	15
	hand	35	50	15
	total	36	49	15
Total des éléments vocaliques (sur 1200 répétitions)		55	29	16

7.11 La place de l'accent principal

Deux types de prononciation sont en principe attestés dans la littérature concernant les mots *donate* et *locate* (cf. chapitre méthodologie section 6.6.3.1). Il s'agit de l'accentuation de la première syllabe en GA (/ˈləʊkət/ et /ˈdəʊnət/) et de la deuxième en RP (/ləʊˈkət/ et /dəʊˈnət/). Mais, nous verrons ci-dessous l'émergence d'une troisième prononciation (plate) qui consiste en la non-accentuation d'aucune des deux syllabes, ou du moins, à un même niveau d'accentuation de ces dernières (voir les indications acoustiques en Annexe A 1, p. 1-35).

7.11.1 Donate

L'accentuation de la deuxième syllabe (syll₂) est produite à 38 % pour le mot *donate*, celle de la première syllabe (syll₁) à 45 % et celle de l'accentuation plate à 17 %, sur 100 répétitions. Ainsi, les locuteurs 7, 8, 12, 13, 19 et 20 n'accentuent que la syll₂ pour la totalité des répétitions. Les locuteurs 3, 9, 10, 11, 15 et 18 accentuent uniquement la syll₁. Seul le locuteur 16 produit entièrement la prononciation plate. Les autres locuteurs (1, 2, 4, 5, 6, 14 et 17) procèdent à un mélange, avec respectivement des syllabes plates, des syll₂ et des syll₁ (Figure 61).

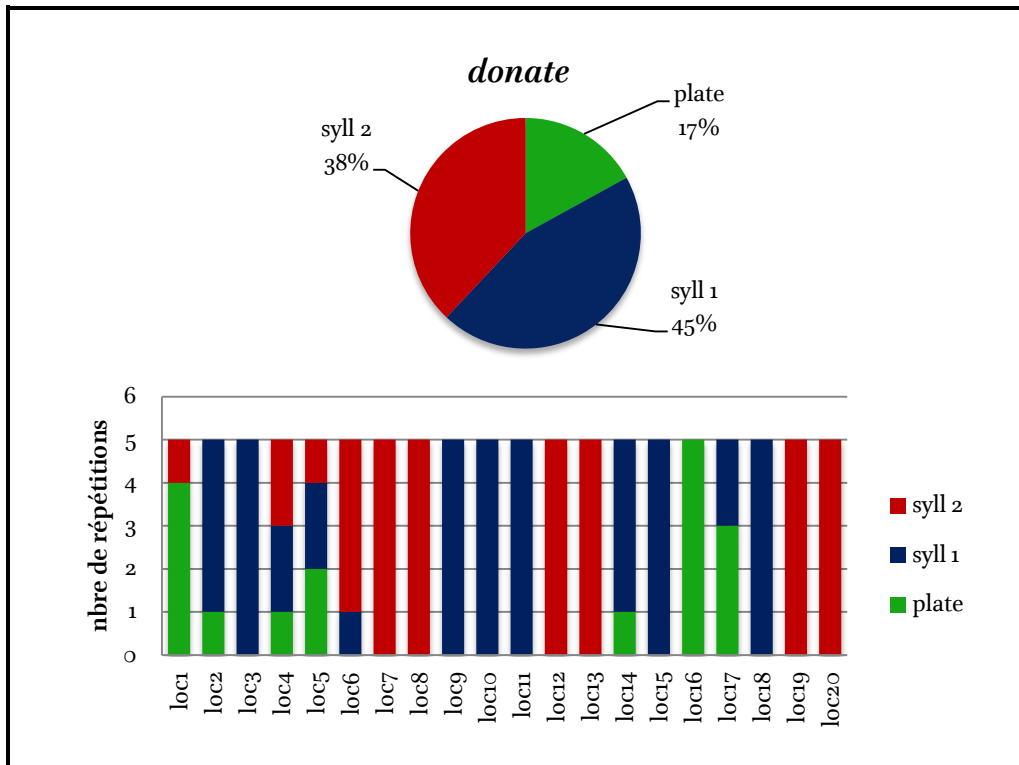


Figure 61 : pourcentages et répartition par locuteur de l'accentuation de *donate*.

7.11.2 Locate

Concernant le mot *locate*, la syll1 est accentuée par les locuteurs 3, 9, 10, 11, 14, 16 et 18 et la syll2 par les locuteurs 7, 8, 12, 13 et 19, pour toutes les répétitions. Des mélanges des syll1, syll2 et de la prononciation plate sont effectués chez les locuteurs 1, 2, 4, 5, 6, 15 et 20. Les pourcentages d'ensemble pour ce mot s'élèvent à 51 % de syll1, 34 % de syll2 et 14 % de la prononciation plate, sur les 100 répétitions (Figure 62).

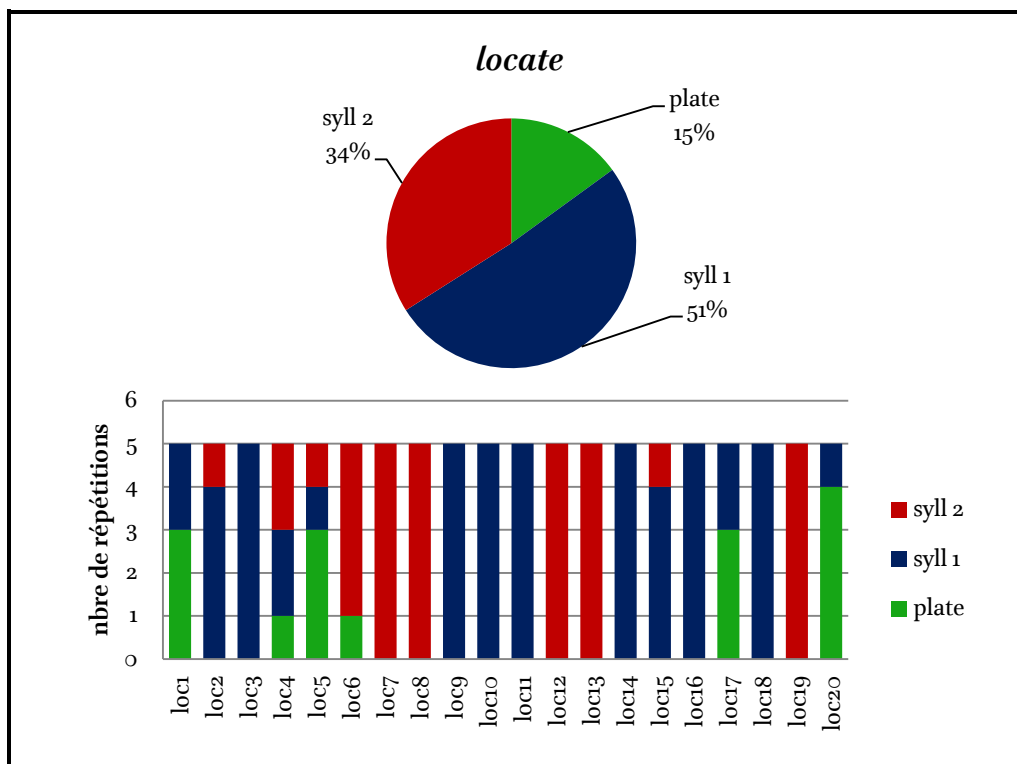


Figure 62 : pourcentages et répartition par locuteur de l'accentuation de *locate*.

7.11.3 Donate et locate

Les pourcentages pour les deux mots (sur 200 occurrences) placent l'accentuation de la première syllabe en tête avec 48 %, suivis par celle de la deuxième syllabe à 36 % et l'accentuation plate à 16 %. Nous remarquons une constance dans la production des locuteurs 3, 9, 10, 11 et 18 qui accentuent la première syllabe et les locuteurs 7, 8, 12, 13 et 19 qui accentuent la deuxième syllabe, pour les deux mots (10 répétitions chacune). Ils représentent la moitié de l'échantillon. L'autre moitié mélange deux ou trois accentuations (locuteurs 1, 2, 4, 5, 6, 14, 15, 16, 17 et 20). Cette constance se voit également dans les pourcentages globaux de la production des deux mots : il n'y a pas une grande différence pour l'élément étudié (syll1 45 % contre 51 %, syll2 38 % contre 34 %, plat 17 % contre 15 %) (Figure 63).

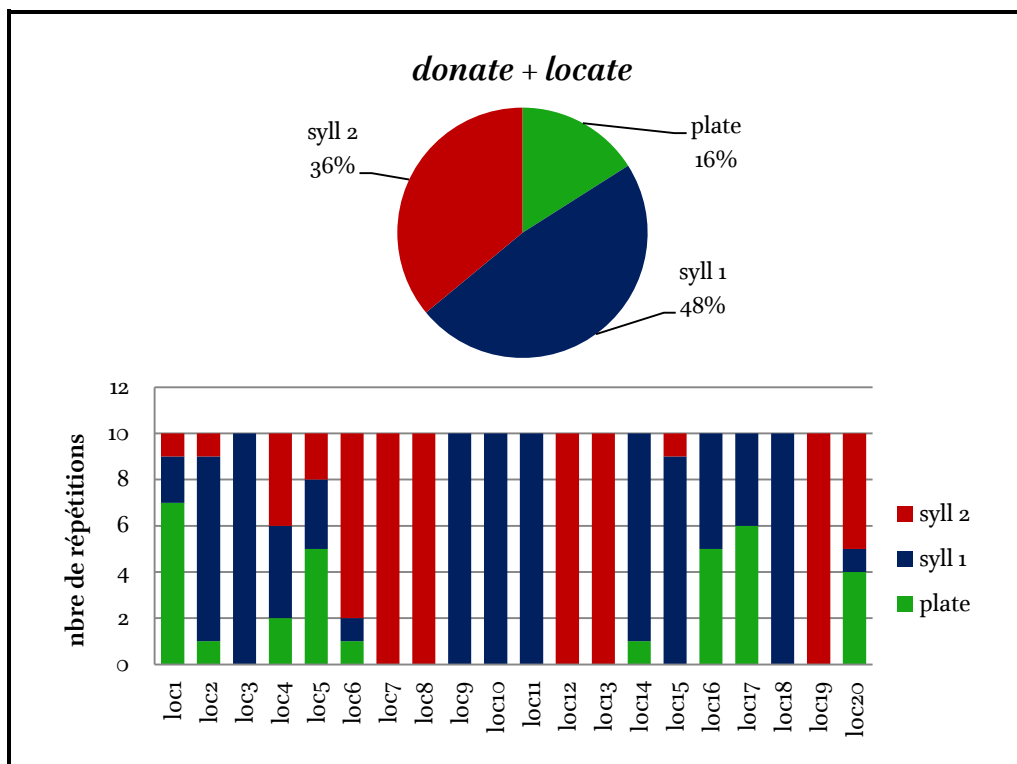


Figure 63 : pourcentages et répartition par locuteur de l'accentuation de *donate* et *locate*.

7.11.4 Discussion

Les deux mots *donate* et *locate* ont des similitudes avec le français, car ces mots existent sous plusieurs formes (*donner/donataire, locataire/location, etc.*). La RP aurait ainsi hérité du système d'accentuation concernant certains mots d'origine française (*cf.* section 2.2.2, p. 42), soit un système d'accentuation qui consiste à placer la « proéminence²¹⁷ » sur la dernière syllabe du groupe rythmique. L'on devrait s'attendre à ce que les locuteurs francophones accentuent systématiquement tous ces mots sur la deuxième syllabe, de par le phénomène de la nativisation. Or, la deuxième syllabe n'est accentuée qu'à 36 % par nos locuteurs. L'accentuation plate (16 %) pourrait également s'expliquer par ce même système. En effet, certains locuteurs produiraient toute la phrase avec une accentuation uniforme et ne placeraient la proéminence que sur le mot de la fin (*please*). Même si les entrées dans le LPD et l'EPD indiquent que les deux mots sont parfois accentués sur la deuxième syllabe en GA, comme en RP, ils précisent que l'accentuation de la deuxième syllabe est seulement pratiquée à 12 % en GA (EPD, p. 147 ; 293, LPD, p. 244 ; 468). Ceci

²¹⁷ Selon Pagel, Madeleni & Wioland, (2012, p. 37-38), en français : « La dernière syllabe prononcée de chaque mot phonétique jouit d'un statut privilégié. Les syllabes non finales sont prononcées de manière 'monotone' alors que la dernière syllabe est prononcée avec une énergie articulatoire très marquée à ne pas confondre avec l'énergie acoustique qui caractérise la syllabe accentuée dans nombre de langues ».

n'explique pas le pourcentage élevé quant à l'accentuation de la première syllabe chez nos locuteurs (48 %). Nous supposons qu'ils ont sûrement entendu cette prononciation quelque part, notamment par l'exposition au GA. Nous évoquons les travaux de Fournier (2016) sur l'accentuation des emprunts français en anglais, car certains de ses résultats sont contradictoires des nôtres. En effet, Fournier a analysé un corpus de mots empruntés au français et qui ont les mêmes terminaisons françaises (<-aire>, <-euse>, <-eur>, <-eux>, <-que>, <-ier>, <-ais>, <-eau> et <-aux>). En se fondant sur les dictionnaires EPD (2016a), et EPD + LPD (2016b), il étudie l'accentuation de ces mots en anglais américain et en anglais britannique. Dans l'article de 2016 (parag. 27), l'auteur a trouvé que les Américains accentuent plus la syllabe finale des emprunts français que les britanniques (73 % contre 58 %). Dans l'autre étude, l'anglais américain respecterait mieux l'accentuation finale du français que l'anglais britannique, qui aurait modifié l'accentuation d'origine afin de correspondre au système d'accentuation de l'anglais (Fournier, 2016, p. 424, 432). Ses résultats sont différents des nôtres notamment pour deux raisons : la première tient au fait que les mots *donate* et *locate*, quoiqu'ils soient des emprunts français, ne rentrent pas dans les catégories étudiées par Fournier (<-ate>). Deuxièmement *donate* et *locate* sont des verbes, alors que Fournier a uniquement pris en compte les noms propres et communs. Toutefois, l'auteur reconnaît les limites liées au choix de ces terminaisons et propose de prendre en compte la seule origine française des mots comme critère principal dans les futures études (2016b, p. 434).

7.12 Le nombre de syllabes accentuées

Les mots *military* et *necessary* sont prononcés soit en trois (en RP), soit en quatre syllabes (en RP et en GA). Dans le premier cas, c'est uniquement la première syllabe qui porte l'accent (/ˈmɪlɪtri/, /ˈnɛsəri/). Dans le deuxième cas, la première syllabe garde l'accent primaire et la troisième syllabe reçoit un accent secondaire. Celle-ci est alors produite avec une voyelle centrale /ə/ en RP (/ˈmɪlɪtəri/ ou /ˈmɪlɪ˞ri/, /ˈnɛsə˞ri/) et avec une voyelle ouverte /ɛ/ en GA (/ˈmɪlɪtəri/, /ˈnɛsəri/). Nous n'analysons pas l'accentuation de la première syllabe de ces mots, car elle ne présente aucune variation selon les deux accents.

7.12.1 Military

Le mot *military* est prononcé de quatre manières différentes, parmi lesquelles nous retrouvons les trois réalisations connues. Il s'agit des sons [ɛ] pour le GA et du [ə] et du [-] pour le RP. Pointons que l'absence de voyelle est représentée par le symbole [-] afin de rester cohérent avec les autres notations. Elle indique également que le mot est prononcé en trois syllabes et pas quatre. La quatrième production consiste en un [ɑ]. Sur un total de 100 répétitions, [ə] domine avec 55 % des productions, suivi par le son [ɛ] à 25 %, ensuite du [-] à 15 % et enfin du [ɑ] à 5 %. Les locuteurs 1, 3, 5, 7, 8, 11, 12, 14 et 17 réalisent entièrement des [ə]. Tandis que les locuteurs 4, 6 et 18 mélangent respectivement [ə] et [ɛ], le locuteur 19 mélange [ə] et [ɛ]. Les locuteurs 10, 15 et 16 n'ont produit aucune voyelle. Les locuteurs 2, 9 et 20 produisent des [ɛ] dans toutes leurs répétitions (Figure 64).

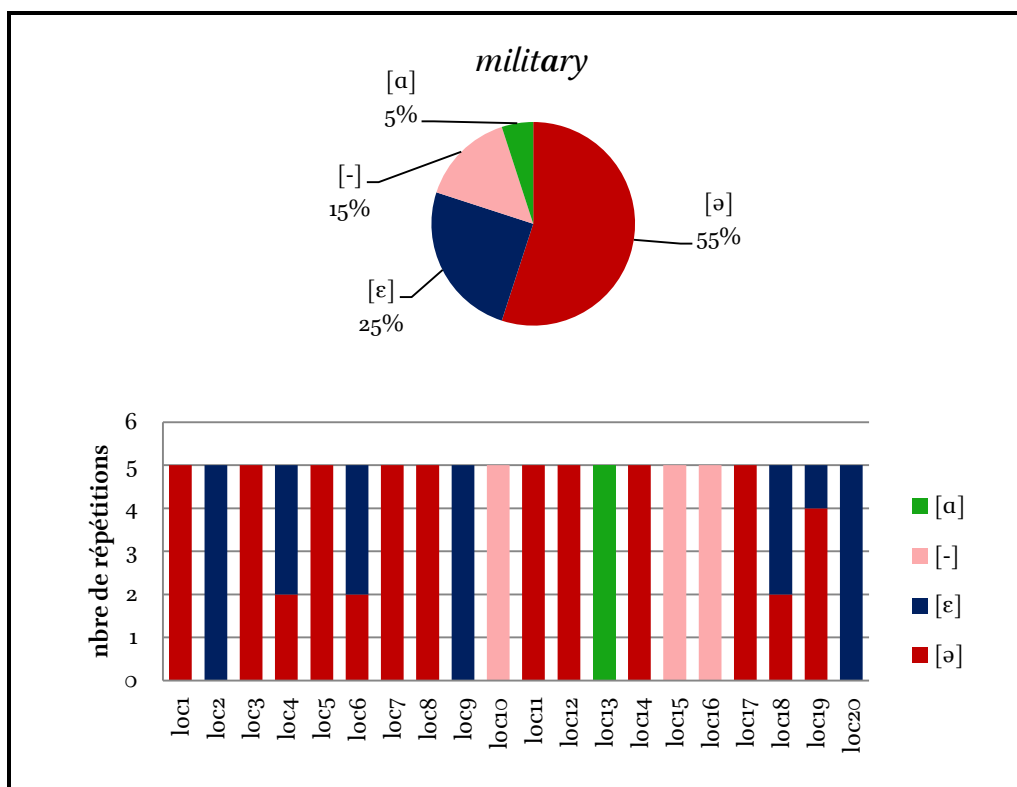


Figure 64 : pourcentages et répartition par locuteur de la prononciation de <ary> de *military*.

7.12.2 Necessary

Le mot *necessary* n'a connu que deux prononciations, à savoir les sons [ə] qui sont produits à 45 % et le son [ɛ] à 55 % (Figure 65). Remarquons l'absence des sons [-] et [ɑ] pour ce mot ainsi que l'absence de mélange. Le son [ɛ] est réalisé pour toutes les répétitions par les locuteurs 2 à 6, 9, 10, 15, 16, 19 et 20 et [ə] est produit par les autres locuteurs (1, 7, 8, 11 à 14, 17 et 18).

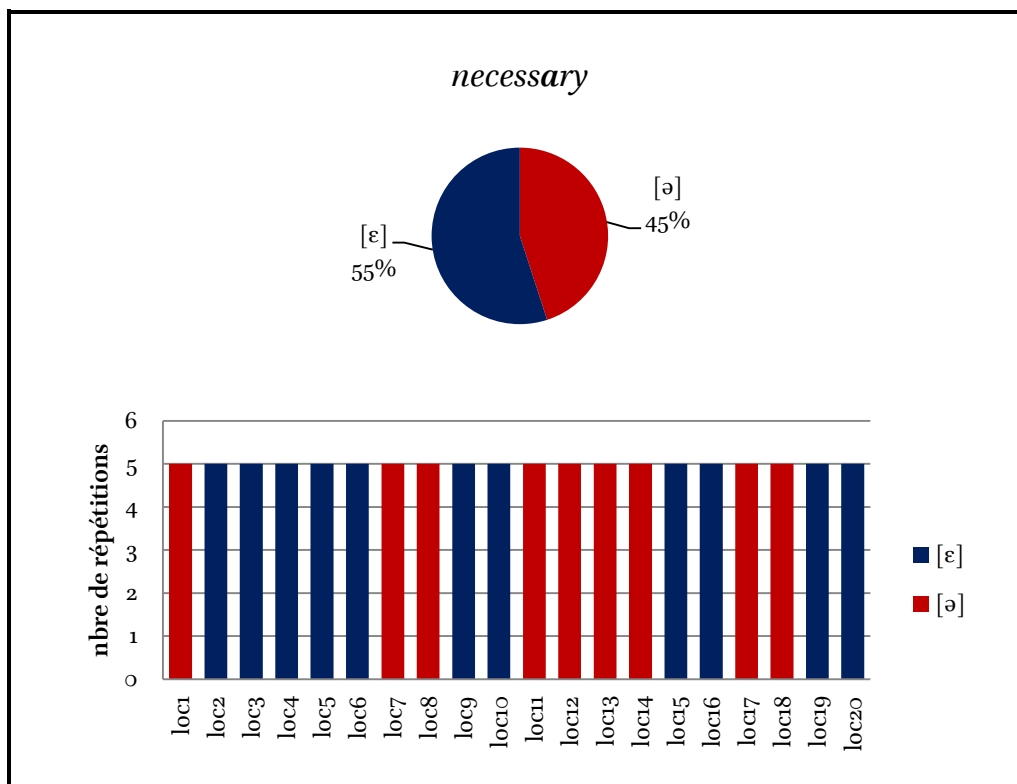


Figure 65 : pourcentages et répartition par locuteur de la prononciation de <ary> de *necessary*.

7.12.3 Military et necessary

Pour les deux mots, [ə] concerne la moitié des 200 répétitions (50 %). Le son [ɛ] est réalisé à 40 %, [-] à 7 % et [ɑ] à seulement 3 %. Il y a plus de variations dans la production du mot *military*. Le décompte total montre des mélanges : le locuteur 13 produit [ə] et [ɑ] ; les locuteurs 10, 15 et 16 produisent chacun [ɛ] et [-]. Les locuteurs 3, 4, 5, 6, 18 et 19 réalisent respectivement des [ə] et [ɛ]. Les locuteurs 1, 7, 8, 14 et 17 restent constants dans toutes leurs productions avec des [ə], ainsi que les locuteurs 2, 9 et 20 avec des [ɛ] (Figure 66).

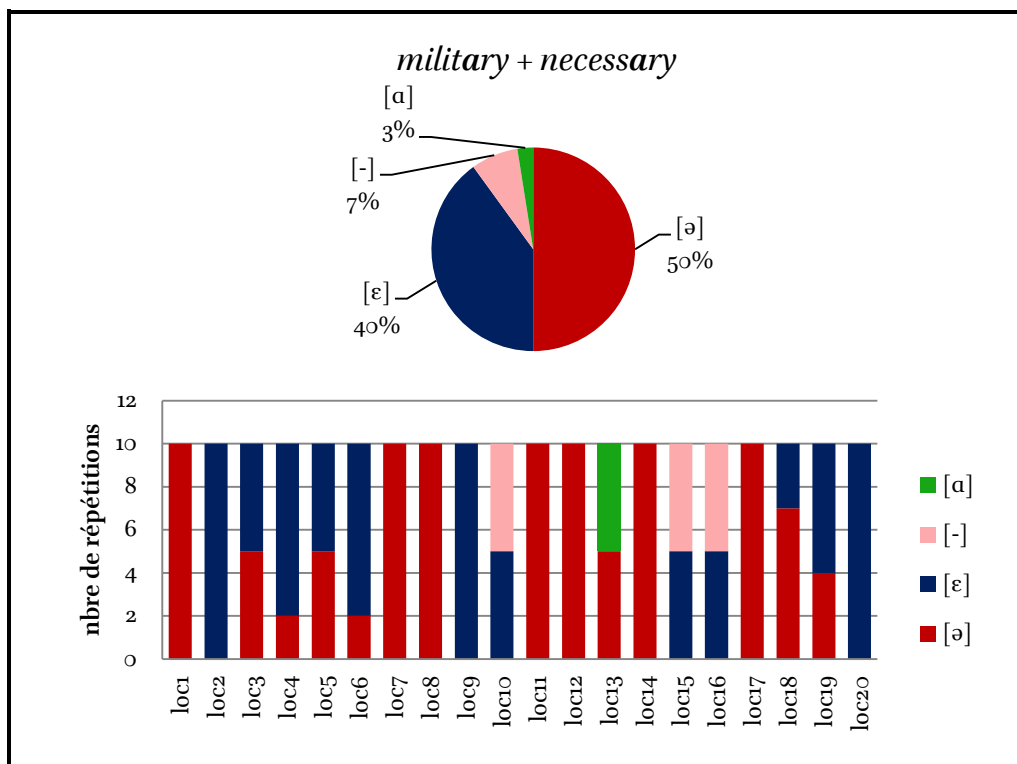


Figure 66 : pourcentages et répartition par locuteur de la prononciation de <ary> de *military* et *necessary*.

7.12.4 Discussion

Alors que le mot *necessary* est prononcé avec seulement les voyelles [ε] et [ə], le mot *military* connaît deux réalisations en plus de ces sons. La réalisation sans aucune voyelle ([-]), c'est-à-dire en trois syllabes, est la prononciation dominante en RP à côté de celle réalisée avec un [ə]. Contrairement à la réalisation sans voyelle de *military*, celle réalisée avec [a] laisse penser à une influence de l'orthographe. Néanmoins, il est intéressant de voir que le mot *necessary* n'a pas connu le même sort. Il aurait pu être produit avec un [a] puisqu'il contient également la graphie <a>. Ou alors, il aurait pu être produit sans aucune voyelle comme il est souvent prononcé en RP ('nesəsri). Aucun élément ne nous permet d'expliquer cette différence de traitement, si ce n'est l'attribution à la fréquence d'exposition : la façon dont ces locuteurs ont entendu la prononciation. Cette fréquence d'exposition pourrait également être à l'origine de l'écart de réalisation de [ε] dans les deux mots : *necessary* étant réalisé avec plus de [ε] que *military*. En effet, le corpus SOAP indique que le mot *necessary* est plus fréquent (3951 occurrences) que le mot *military* (546 occurrences). Mais, nous n'avons pas d'information sur le type de prononciation qui revient le plus fréquemment, c'est-à-dire la prédominance de la prononciation américaine ou britannique. Toutefois, nous pouvons spéculer qu'ils sont plus exposés à la version américaine, si nous

considérons la quantité des *inputs* de l'accent GA dans les médias. Une autre tentative d'explication réside dans la similitude de ces mots avec militaire et nécessaire (/mili'tɛʁ/ et /nese'sɛʁ/). Nous aurions pu penser qu'il y ait calquage de prononciation de la dernière syllabe de ces mots sur celle de l'anglais, c'est-à-dire que la réalisation en [ɛ] serait une influence du /ɛʁ/ français. Mais encore une fois cette explication ne peut être valide, puisque sinon le mot *military* aurait été prononcé avec plus de [ɛ] que de [ə]. Le postulat de la notion de fréquence d'exposition reste la plus recevable. Un dernier point mérite d'être discuté. Nous observons que la durée des voyelles [ə] est proche, voire parfois égale à celle des [ɛ], comme si ces locuteurs voulaient accentuer cette voyelle. Tout laisse à croire qu'il y a eu « ratage » de cible : ils auraient voulu prononcer un [ɛ] mais la qualité de la voyelle ne serait pas atteinte. Cela pourrait provenir d'un problème de perception dans la mesure où ils perçoivent la différence de durée de la voyelle [ɛ] dans les deux mots, mais n'arrivent pas à réaliser correctement la bonne voyelle. Nous pouvons dire qu'en dépit de la différence phonétique entre [ɛ] et [ə], ces deux sons se situent dans une même zone de viabilité, puisque les deux prononciations existent et sont compréhensibles.

Nous remettons encore une fois le constat sur la variabilité dans la production de nos locuteurs, concernant les syllabes accentuées et le nombre de syllabes prononcées, sans oublier la qualité de la voyelle produite. Le résumé de ces résultats se trouve dans le tableau ci-dessous et montre que globalement les caractéristiques du GA dominant celles de la RP, grâce au placement de l'accent dans la prononciation des mots *necessary*, *donate* et *locate* (Tableau 22).

Pour finir, nous présentons dans ce qui suit les résultats globaux depuis le début de ce chapitre, à savoir tous les mots et éléments étudiés et leur pourcentage pour l'ensemble des locuteurs.

Tableau 22 : résumé des éléments suprasegmentaux.

Phénomène étudié	Mots	% de caractéristiques de la RP	% de caractéristiques du GA	% de caractéristiques « Autres »
<i>Place de l'accent</i>	<i>donate</i>	38	45	17
	<i>locate</i>	34	51	15
	<i>total</i>	36	48	16
<i>Nombre de syllabes prononcées et qualité de la voyelle</i>	<i>military</i>	55	25	20
	<i>necessary</i>	45	55	0
	<i>total</i>	57	40	3
Total des éléments suprasegmentaux (sur 400 répétitions)		43	44	13

7.13 Résultats globaux des mots à vitesse normale

Nous rassemblons ici tous les éléments étudiés et présentés *supra* pour tous les locuteurs. Le nombre total de productions analysées s'élève à 2900, soit 145 productions par locuteur. Nous pouvons remarquer tout de suite que les pourcentages pour les deux accents de référence sont assez proches : les caractéristiques de l'accent britannique représentent 44 % et 45 % pour celles de l'accent américain. Les caractéristiques restantes sont mises sous l'étiquette « Autres » (11 %).

Les pourcentages de productions par locuteur sont assez variables. Le locuteur 1 prononce 56 % de caractéristiques proches de la RP, 29 % proches du GA et 15 % « Autres ». La part des caractéristiques « Autres » du locuteur 2 est un peu plus élevée avec 22 %, suivie par celles proches de la RP à 30 % et celles proches du GA à 48 %. Le locuteur 3 ne produit que 7 % de caractéristiques « Autres », 28 % de celles de la RP et une large part de celles du GA (65 %). Le locuteur 4 prononce également 7 % de caractéristiques « Autres », mais celles de la RP atteignent 34 % et celles du GA 59 %. Pour sa part, le locuteur 5 produit 45 % d'éléments caractéristiques de la RP, suivi de ceux du GA à 47 % et seulement 8 % de caractéristiques « Autres ». Le locuteur 6 réalise au total 49 % des éléments de l'accent RP contre 40 % de l'accent GA ; les caractéristiques « Autres » sont produites à 11 %. Quant au locuteur 7, il ne réalise que 4 % d'éléments « Autres », 27 % pour ceux du GA contre un fort pourcentage de caractéristiques de la RP (69 %). Nous obtenons chez le locuteur 8, 56 % des éléments du GA, 32 % pour ceux de la RP et 12 % pour les autres productions. Il est intéressant de remarquer que le locuteur 9 est le seul à n'avoir pas produit des caractéristiques « Autres ». De plus, il produit une forte dominance des caractéristiques du GA (83 %) contre seulement 17 % de celles de la RP. Les aspects du GA et de la RP sont très proches chez le locuteur 10 (45 % et 43 % respectivement), les autres réalisations sont à 12 %. Concernant le locuteur 11, il réalise les propriétés de la RP à 44 %, celles du GA à 32 % et celles des « Autres » à 24 %. Pour le locuteur 12, la part des attributs de l'accent britannique occupe 48 %, celle de l'accent américain 42 % et 10 % pour celle des autres productions. Le locuteur 13 pour sa part, réalise 51 % de ces répétitions avec les marques de l'accent britanniques, celle de l'accent américain sont à 37 % et les autres productions sont à 12 %. Les propriétés de l'accent RP sont produites à 58 % chez le locuteur 14, celles de l'accent GA arrivent à 27 % et celles des autres réalisations à 15 %. Une dominance de 63 % des caractéristiques de l'accent RP est produit chez le

locuteur 15 ; 30 % pour celles du GA et uniquement 7 % « Autres ». Ces pourcentages se rapprochent de ceux du locuteur 7. Une ressemblance des prononciations du locuteur 8 est également observée chez le locuteur 16 qui produit 54 % des propriétés du GA, 34 % de celles de la RP et 12 % des productions « Autre ». 47 % des spécificités de l'accent GA, 45 % de celles de la RP et 8 % celles des autres réalisations sont observés chez le locuteur 17. Un schéma presque identique est observé chez le locuteur 18 avec 48 % des caractéristiques de la RP, 42 % de celles du GA et 10 % « Autres ». Le locuteur 19 produit aussi un schéma similaire : 47 %, 41 % et 12 % respectivement pour les caractéristiques du GA, de la RP et « Autres ». Enfin, 42 % des particularités de l'accent RP, 39 % de celles du GA et 19 % « Autres », sont observés chez le locuteur 20.

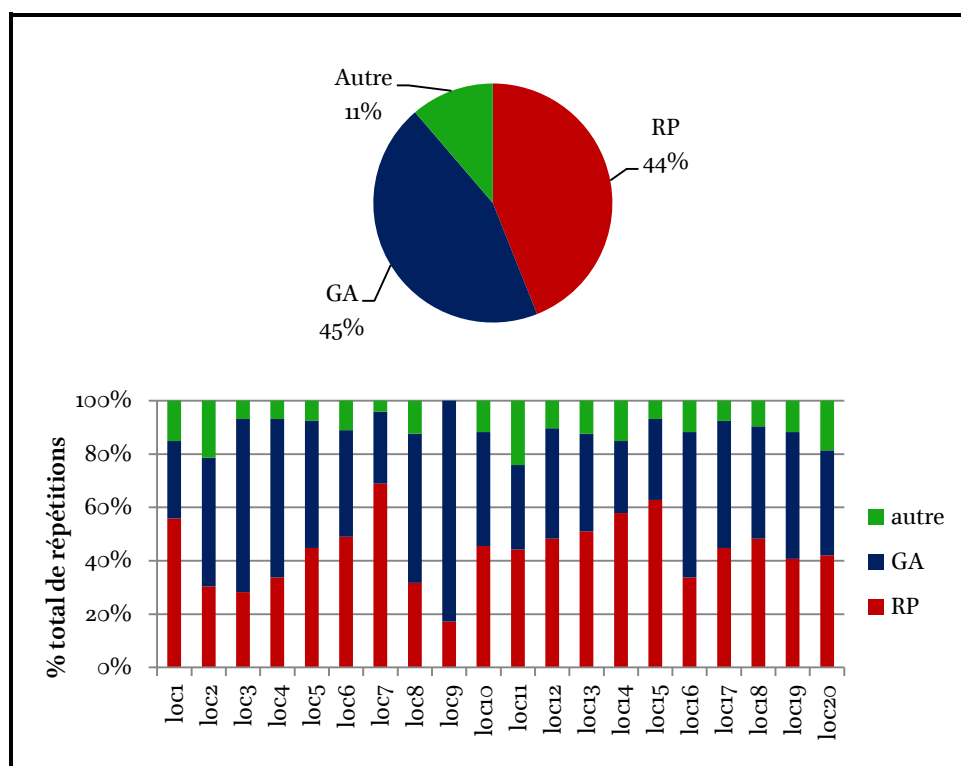


Figure 67 : pourcentages généraux pour tous les éléments étudiés et répartition par locuteur.

7.13.1 Discussion

Certains locuteurs se ressemblent de par les pourcentages de leurs productions. C'est par exemple le cas des locuteurs 5 et 17 qui produisent exactement les mêmes pourcentages (47 % GA, 45 % RP, 8 % « Autres »). Il en va de même pour les locuteurs 12 et 18 qui sont à 48 % RP, 42 % GA et 10 % « Autres ». Parfois, ils ont des pourcentages identiques pour une caractéristique spécifique (par exemple les locuteurs 4 et 16 ont chacun 34 % pour les caractéristiques RP mais les autres

éléments sont assez différents). Nous précisons que ces ressemblances peuvent concerner aussi bien des données au sein d'un seul type d'élément étudié que dans d'autres types.

Certes, la lecture aurait un effet indéniable sur la prononciation. Mais faute de preuves tangibles, nous ne pouvons qu'émettre des hypothèses. Toutefois, nous rappelons que bien des réalisations ne peuvent s'expliquer par l'influence des graphies, ce que nous avons largement montré. Ainsi, nous pouvons dire que ces prononciations relèvent de l'imitation inconsciente (*cf.* section 4.4.2) comme c'est le cas chez les natifs et correspondent aux prononciations des mots telles qu'ils les entendent et les « perçoivent ». Il se pourrait qu'au cours des répétitions des mots, nos locuteurs se rappellent toutes les variétés de prononciation auxquelles ils ont été exposés, durant les différentes activités dans lesquelles ils s'engagent (formelles, non formelles et informelles). Cela pourrait éventuellement justifier les mélanges des productions. Les « ratages » de cibles et les mélanges de prononciation pourraient également provenir de situations d'insécurité linguistique, la peur de ne pas pouvoir bien prononcer les mots, ou tout simplement une difficulté articulatoire du moment. Enfin, le mélange d'accents observé dans les productions orales de nos locuteurs peut vraisemblablement provenir non seulement de l'exposition à l'accent transatlantique (ou *Mid-Atlantic accent*²¹⁸), mais également à d'autres accents du monde entier (*cf.* section 2.3 p. 44). Cette exposition à des variétés mixtes expliquerait la forte variabilité dans les productions et le fait que certains mots soient prononcés en RP lorsque d'autres sont produits en GA.

7.14 Conclusion

Nous venons de présenter dans cette partie les résultats de la prononciation de nos 20 locuteurs. Dix éléments de prononciation²¹⁹ contenus dans 24 mots²²⁰ ont été analysés pour le mode vitesse normale, concernant les 2900 répétitions. En somme, trente-cinq (35) différentes productions sont recensées. Tout d'abord, les résultats ainsi obtenus ont été largement décrits, illustrés au travers de deux types de graphiques (les diagrammes en barres et les secteurs) et

²¹⁸ Cet accent particulier de l'anglais est une variété codifiée qui consiste à utiliser les caractéristiques phonético-phonologiques de l'anglais britannique et de l'anglais américain (voir Modiano, 1996 et la Bouff, 2008).

²¹⁹ /t/ intervocalique, rhoticité, *yod-dropping*, voyelle de LOT et de GOAT, <i>/<y>, <ei>/<e>, nasalisation, accent principal et nombre de syllabes accentuées.

²²⁰ *City, vitamin, party, more, dark, leisure, actor, bird, organisation, band, hand, top, hot, joke, road, military, necessary, locate, donate, legalisation, tune, new, vitamin* et *dynasty*.

commentés pour l'ensemble des locuteurs. Ensuite, partant de ces analyses, nous avons pu identifier trois types de productions : celles qui se rapprochent de l'accent britannique (RP), ensuite celles qui sont proches de l'accent américain (GA) et enfin celles qui ne relèvent ni de l'un, ni de l'autre (« Autres »). Nous avons également soulevé les mélanges des éléments de la RP et ceux du GA uniquement et avons postulé une possible exposition à l'accent transatlantique et autres accents. Deux causes principales ont été identifiées pour les productions : d'une part il s'agirait de l'influence de la graphie et de l'autre, de la fréquence d'exposition aux *inputs* médiatiques (notamment sériels). Pour nous, il est important de préciser que ces productions sont proches des accents de référence pris en considération. Nous ne disons pas que ces productions sont exactement celles de la RP ou du GA. Concernant l'émergence des réalisations « Autres », nous avons argumenté sur leurs causes possibles, en mettant en avant l'impact de la graphie et l'influence de la L1. Dans la même lancée, nos résultats montrent que les éléments de prononciation proches de la RP et du GA sont dominants et occupent presque les mêmes pourcentages. Nous précisons toutefois que les caractéristiques du GA dépassent celles de la RP dans seulement 23 occurrences. Quant à la part des productions « Autres », elle est moindre face aux deux autres types de réalisations, mais elle demeure non négligeable avec 11 %, soit 327 occurrences sur les 2900. Aussi avons-nous montré les aspects homogènes et hétérogènes dans les différentes prononciations. En effet, certains locuteurs produisent les mêmes sons pour toutes leurs répétitions tandis que d'autres en mélangent parfois plusieurs. Rappelons que certains mots sont plus enclins à être prononcés avec des propriétés de la RP que d'autres (*yod-dropping*). Certains mots attirent plus de caractéristiques du GA (rhoticité) alors que d'autres sont plus réalisés avec les caractéristiques « Autres » (/t/ intervocalique). Ce dernier point nous amène à soulever un aspect très important qui découle de ces phénomènes d'homogénéité et d'hétérogénéité. Il s'agit de la variabilité individuelle. D'un côté, cette variabilité est évidemment inter-locuteur²²¹, puisque chaque locuteur réalise des sons différents des autres locuteurs. De l'autre côté, la variabilité est intra-locuteur, à savoir le fait qu'un même locuteur produit des sons différents au sein d'un même mot. Dans le tout prochain chapitre, nous allons présenter les résultats de la prononciation des mêmes mots en vitesse rapide.

221 Dans le sens de « entre » les différents locuteurs. À ne pas confondre avec un interlocuteur, une personne avec qui l'on parle.

8 Résultats de la prononciation en vitesse rapide

Il arrive parfois que le système de production-perception de la parole ne puisse plus adopter différents états cohérents, ou que son organisation n'arrive plus à garder une certaine stabilité, face à une perturbation interne ou externe : cela amène inéluctablement la désintégration du système et, en conséquence, la perte de l'identité de ses propriétés émergentes. (Vaxelaire, 2007, p. 34).

Nous présentons dans ce chapitre la deuxième partie des résultats des sons et séries de sons placés dans la phrase porteuse (*it is *** please*), en (condition de) vitesse rapide (VR). Comme dans le chapitre précédent sur la vitesse normale (VN), les présentations des données suivent l'ordre décrit dans le chapitre méthodologie : d'abord les consonnes, ensuite les voyelles et enfin les éléments suprasegmentaux (accentuation et nombre de syllabes accentuées). Cette partie est aussi structurée de la même manière. Nous commençons par la description des résultats obtenus pour chaque mot. Ensuite, nous résumons les résultats généraux concernant les mots qui contiennent les éléments étudiés. Puis, nous illustrons ces résultats à l'aide de deux graphiques : le premier montre les pourcentages des productions possibles pour un phénomène étudié et ceci pour chaque mot ; le deuxième représente la part de production de chaque locuteur, pour les éléments identifiés. Tout à la fin de la présentation de ces éléments, nous résumons encore les résultats globaux pour tous les éléments étudiés, ainsi que la part de prononciation pour chaque locuteur. Contrairement au chapitre précédent dans lequel la présentation des phénomènes étudiés est immédiatement suivie d'une discussion au cours de laquelle nous essayons de comprendre et d'expliquer les différentes réalisations, en passant par des questionnements et des hypothèses, celui-ci ne présente qu'une seule section discussion. La raison est que les explications concernant les causes possibles de ces productions et les hypothèses sous-jacentes s'appliquent également pour les productions en VR. Ici nous présentons les résultats de la VR à l'échelle globale et individuelle, puis selon la même procédure, nous faisons une comparaison en rappelant les pourcentages des résultats de la VN. Finalement, nous illustrons les résultats de la VR à l'aide des figures correspondantes. Alors que la discussion sur la VN a pour objectif d'expliquer la provenance des productions, celle de la VR a pour objectif de décrire les écarts de prononciation entre la VN et la VR. Pour des raisons d'espace, nous ne reprenons pas les graphiques concernant les résultats de la VN. Toutefois, nous ferons un renvoi vers les graphiques montrant les

pourcentages pour la VN (par exemple comparaison²²² : Figure x, page y), le lecteur pourrait se référer au chapitre précédent pour plus de clarté et de précision. Nous rappellerons les références correspondantes en indiquant les numéros et titres des graphiques ainsi que les numéros de page des sections concernées. Ainsi, la section discussion que nous réservons à la fin de ce chapitre est destinée aux différences majeures entre ces deux situations de prononciation. Par conséquent, les rappels sur les différences de prononciation entre le RP et GA ne sont plus repris dans ce chapitre (contrairement à ce que nous faisons au début de chaque élément étudié dans le chapitre précédent).

La même charte graphique est appliquée dans ce chapitre. Pour rappel, le rouge foncé correspond aux productions proches de la RP et le bleu foncé représente les réalisations proches du GA. Les mêmes couleurs sont représentées, mais en plus claires, lorsque deux ou plusieurs réalisations sont attestées dans un système de prononciation particulier. Par exemple : le rouge clair indique une autre production possible de la RP. Ensuite, nous utilisons le vert pour des productions que nous jugeons proches ou équivalentes au système phonétique du français (ou d'autres langues parlées par certains locuteurs²²³) et les couleurs vert clair pour se référer aux réalisations « Autres ». Parallèlement, les diagrammes en secteurs (camemberts) illustrent les pourcentages des différentes productions pour un élément étudié et les colonnes (histogrammes) sont utilisées pour montrer la production de chaque locuteur. Les notations // représentant un son ou phonème en général et les [] indiquant une réalisation spécifique des locuteurs, ne changent pas (voir chapitre précédent). Pour commencer, nous présentons les résultats des éléments segmentaux consonantiques analysés, à savoir le /t/ intervocalique, la rhoticité et le *yod-dropping*²²⁴ dans l'ordre, dans les trois prochaines sections.

8.1 La prononciation du /t/ intervocalique

Nous commençons par la présentation des résultats des /t/ intervocaliques contenus dans les mots *city*, *vitamin* et *party*.

²²² Nous emploierons l'abréviation « cp. » pour « comparaison » et « p. » pour page, dans la suite des présentations.

²²³ Voir les résultats du questionnaire (section 9.1, p. 304).

²²⁴ Nous plaçons la semi-consonne /j/ dans les résultats des consonnes, car c'est elle qui subit le changement et pas la voyelle /u:/ qui la suit.

8.1.1 City

La prononciation du mot *city* a généré six sons dont le [r] (34 %), le [t^h] (30 %), le [d] (19 %), [t] (10 %), [tʃ] (5 %) et le [tʂ] (2 %). Les locuteurs 9, 16 et 19 ne produisent que des [r], les locuteurs 3, 5, 6, 10 et 13 ne produisent que des [t^h] et le locuteur 12 produit uniquement des [tʃ]. Les autres locuteurs mélangent entre deux sons : [t^h] et [tʂ], [r] et [d], [t] et [tʂ], [d] et [t] (locuteurs 1, 2, 4, 7, 8, 11, 14, 15, 17 et 20) et trois sons : [d], [t] et [t^h] (locuteur 18). Il est intéressant de noter le nombre de locuteurs qui mélangent le [d] et le [r] (Figure 68). D'ailleurs, nous avons considéré la production du [d] comme étant un possible « ratage de cible »²²⁵ de celle de [r]. En effet, nous avons montré que ces deux sons sont similaires sur les plans acoustique, articulatoire et perceptif.

Lorsque nous comparons ces chiffres avec ceux obtenus en VN, nous constatons deux choses : d'une part qu'il y a une augmentation de la production des sons [r], [d], [tʃ] et [t] et une diminution des sons [t^h] et [tʂ] ; d'autre part nous notons la non-production du son [t] (cp. Figure 28, p. 206). Aussi, en VN les locuteurs ont mélangé jusqu'à quatre sons contre seulement trois en VR (locuteur 4). La VN a donc généré plus de mélanges que la VR, pour ce mot. Sur le plan individuel, les locuteurs 3 et 13 concernant le son [t^h], les locuteurs 9 et 16 concernant le [r], gardent les mêmes productions dans les deux modalités (VN et VR). Puis les locuteurs 5, 6, 10, 12 et 19 produisent des sons homogènes en VN mais réalisent des mélanges en VR. La situation inverse est constatée chez les locuteurs 1, 8, 17 et 20 qui réalisent des mélanges en VN mais sont plutôt constants en VR. Enfin, les locuteurs 2, 4, 7, 11, 14, 15 et 18 mélangent les diverses productions dans les deux situations.

²²⁵ Prononciation du [r] déviée en [d].

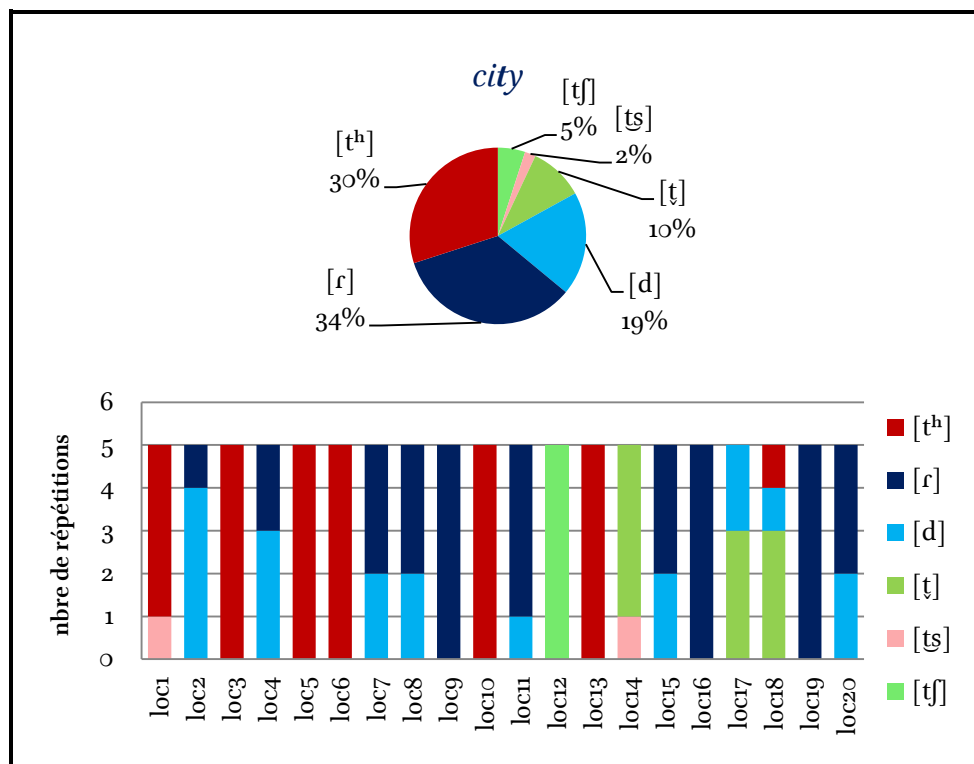


Figure 68 : pourcentages et répartition par locuteur de la production du <t> de *city*.

8.1.2 Vitamin

Pour la prononciation du mot *vitamin*, nous constatons également la production des six sons. Il s'agit du [tʰ] (50 %), du [r] (32 %), du [d] (8 %), du [t] (7 %), du [ʈ] (2 %) et enfin d'une seule production du coup de glotte (1 %). Notons que contrairement au mot *city*, les sons [tʃ] et [ʈ͡] ne sont pas produits. Pendant que les locuteurs 3, 5, 6, 7, 12, 14, 16 et 17 ne prononcent le mot qu'avec des [tʰ], les locuteurs 8 et 19 prononcent le mot avec des [r], le locuteur 13 est le seul à prononcer qu'avec des [t]. Un mélange de deux sons : [tʰ] et [t], de [r] et [d], de [r] et [tʰ] est réalisé par les locuteurs 1, 2, 4, 9, 10, 15 et 20 et de trois sons [t], [d] et [r] et de [ʈ], [t] et [d] est produit par les locuteurs 11 et 18. Aussi, nous remarquons la même combinaison des sons [r] et [d] chez quatre locuteurs (2, 4, 9 et 11).

La comparaison de ces résultats avec ceux de la VN montre qu'en VR, la production du [tʰ] a augmenté (de 50 à 60 %) ainsi que celle du [r] (16 à 32 %), pendant que celle des autres sons reste relativement similaire (cp. Figure 29, p. 207). Le [d] stagne à 8 %, le [t] passe de 6 à 7 %, le [ʈ] passe de 2 à 6 %. Enfin le [ʈ͡] garde sa proportion de 1 % et nous tenons à préciser qu'il n'est pas produit par la même personne : en VN il est réalisé par le locuteur 8 alors qu'en VR, il est produit par le locuteur 18. D'un point de vue individuel, pendant que les locuteurs 3, 5, 7, 12 et 17 sont constants

dans leur production du [t^h], les locuteurs 2, 4, 10 et 11 réalisent des mélanges dans les deux conditions. Par ailleurs, les locuteurs 1, 9, 15, 18 et 20 sont constants en VN alors qu'ils réalisent des mélanges en VR. Quant aux locuteurs 6, 8, 13, 14, 16 et 19, ils produisent des mélanges en VN et deviennent constants en VR.

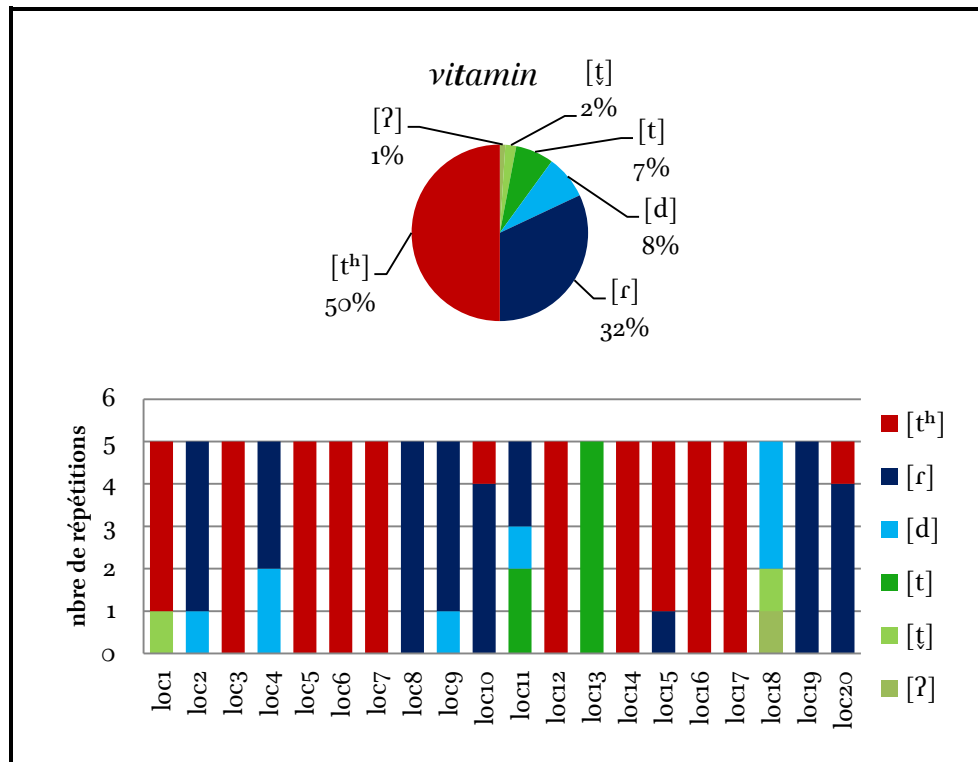


Figure 69 : pourcentages et répartition par locuteur de la production du <t> de *vitamin*.

8.1.3 Party

La graphie <t> dans le mot *party* est prononcée de six façons différentes : il y a le [r] (47 %), le [t^h] (34 %), le [d] (9 %), le [tʃ] (5 %), le [t̥] (3 %), et enfin le [t] (2 %). Cette dominance de la prononciation avec un [r] se reflète chez les locuteurs 2, 5, 8, 9, 15, 18 et 19 et 20 qui l'ont produit à 100 %. Les locuteurs 1, 3, 6, 10, 13 et 14 ont également produit des [t^h] et le locuteur 12 ne réalise que des [t̥]. Seulement les locuteurs 4, 7, 11 et 17 ont mélangé deux sons ([t^h] et [d] ; [r] et [d] ; [t^h] et [r]) et trois sons ([r], [d] et [t] ; [t^h], [d] et [t̥]). On note ainsi une production plus homogène pour le mot *party*.

Comparativement à la VN, on assiste à moins de productions différentes en VR (cp. Figure 30, p. 208). Ainsi il y a sept sons produits contre six en VR : [t̥] n'est pas produit. On note également une augmentation de la production du [r] qui passe de 34 à 47 % et une chute de la production du [t^h] de 45 à 34 %, ainsi qu'une diminution de [t̥] qui passe de 6 à 5 %. Les sons [d] et [t̥] stagnent sur

9 % et 3 % respectivement. Ainsi, les locuteurs 1, 3 et 14 pour le [t^h], 2, 5, 9 et 19 pour le [r] et 12 pour le [tʃ], maintiennent des productions identiques en VN et en VR, pendant que les locuteurs 4, 6, 7, 11 et 16 réalisent des mélanges de différents sons dans les deux modalités. Quant aux locuteurs 8, 10, 13, 15, 18 et 20 ils réalisent des mélanges en VN mais sont constants en VR, le locuteur 17 réalise l'inverse puisqu'il est constant en VN mais réalise des mélanges en VR.

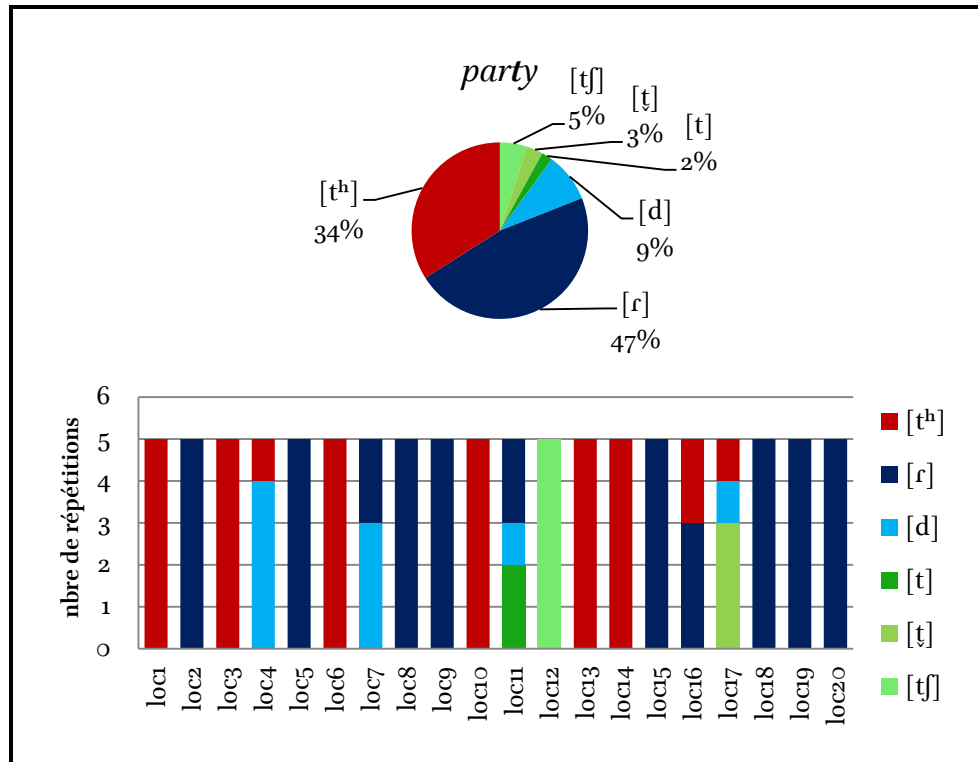


Figure 70 : pourcentages et répartition par locuteur de la production du <t> de *party*.

8.1.4 Résultats du /t/ intervocalique

Pour l'ensemble des mots *city*, *vitamin* et *party*, les sons [t^h] et [r] sont produits à pourcentage égal, soit 38 %. Le [d] arrive à la troisième place avec 12 % des productions, suivi par le [t] (5 %), le [tʃ] et le [t̥] (3 % respectivement) et enfin le [t̥s] (1 %). Les locuteurs 3 et 6 n'ont produit que des [t^h] pendant que le locuteur 19 est le seul à n'avoir produit que des [r]. Les autres ont mélangé deux sons (2, 5, 8, 9, 10, 12, 13 et 16), trois sons (1, 4, 7, 11, 14, 15, 17 et 20) cinq sons (18).

Comparé à la VN, nous constatons en VR une diminution de la production des sons suivants : le [t^h] (48 à 38 %), le [tʃ] (6 à 3 %) le [t̥s] (4 à 1 %) et une augmentation du [r] (24 à 38%), le [d] (10 à 12 %). Les sons [t̥], [t] et [ʔ] stagnent à 5, 3, 1 % respectivement (cp. Figure 31, p. 209). De cette comparaison, nous retenons que globalement les aspects proches de l'accent GA (le [r] et dans une

certaines mesures le [d]) sont plus produits en VR qu'en VN, et que les caractéristiques de la RP (le [t^h] et aussi le [tʂ]) sont moins produites en VR.

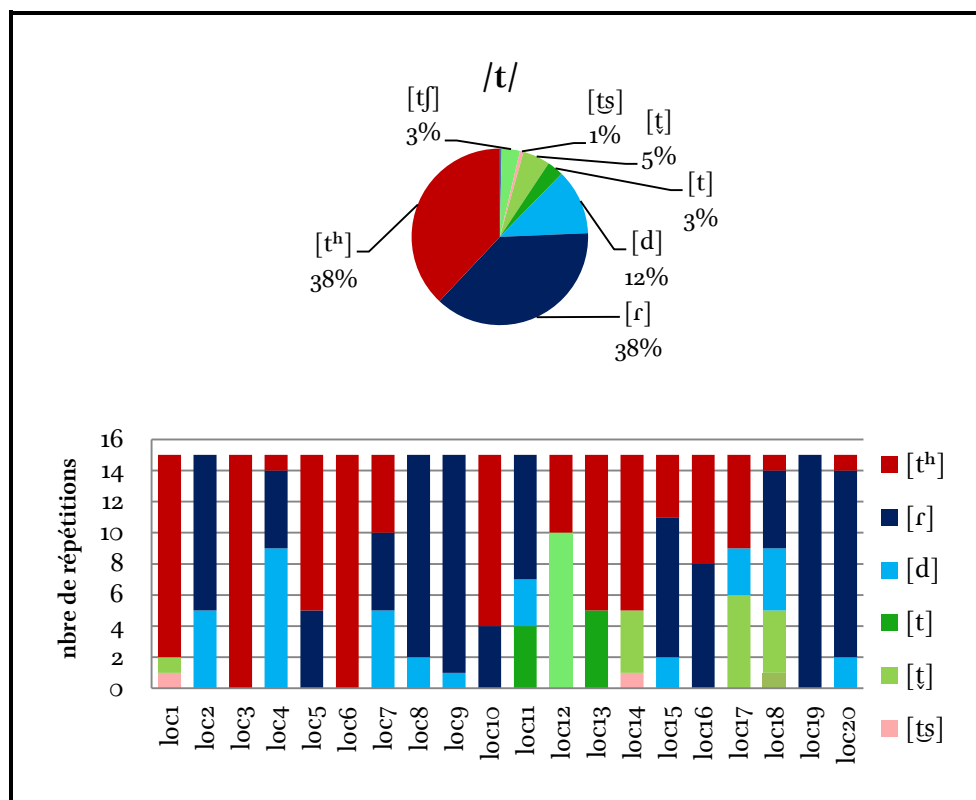


Figure 71 : pourcentages et répartition par locuteur de la prononciation du /t/ intervocalique.

8.2 La rhoticité

Afin d'étudier la rhoticité dans la prononciation de nos locuteurs, nous analysons les mots *more*, *organisation*, *dark*, *party*, *actor*, *actor* et *leisure*.

8.2.1 More

Dans la prononciation du mot *more*, la rhoticité est observée à 55 % et la non-rhoticité à 45 % sur l'ensemble des 100 occurrences. Pendant que les locuteurs 1 et 8 mélangent le [ɔɪ] et le [ɔː], les autres locuteurs prononcent le mot de manière homogène : [ɔɪ] pour les locuteurs 3, 4, 5, 6, 9, 12, 13, 16, 17 et 19 et [ɔː] pour les locuteurs 2, 7, 10, 11, 14, 15, 18 et 20.

Alors que la prononciation du mot *more* avec la rhoticité est dominante en VN, elle chute de 15 % en VR (de 70 à 55 %). Aussi, contrairement à la VN où tous les locuteurs sont homogènes (c'est-à-dire qu'ils ne produisent que des [ɔɪ] ou des [ɔː]), nous remarquons des mélanges de production en VR (loc 1 et 8). Ensuite, nous notons que certains locuteurs gardent les mêmes productions dans

les deux modalités : les locuteurs 3, 4, 5, 6, 9, 12, 13, 16, 17 et 19 pour le [ɔɪ] et les locuteurs 2, 7, 11, 14, 15 et 18 pour le [ɔ:] (cp. Figure 32, p. 211). Enfin, nous notons que les locuteurs 8, 10 et 20 ont des productions rhotiques en VN et des productions non rhotiques en VR.

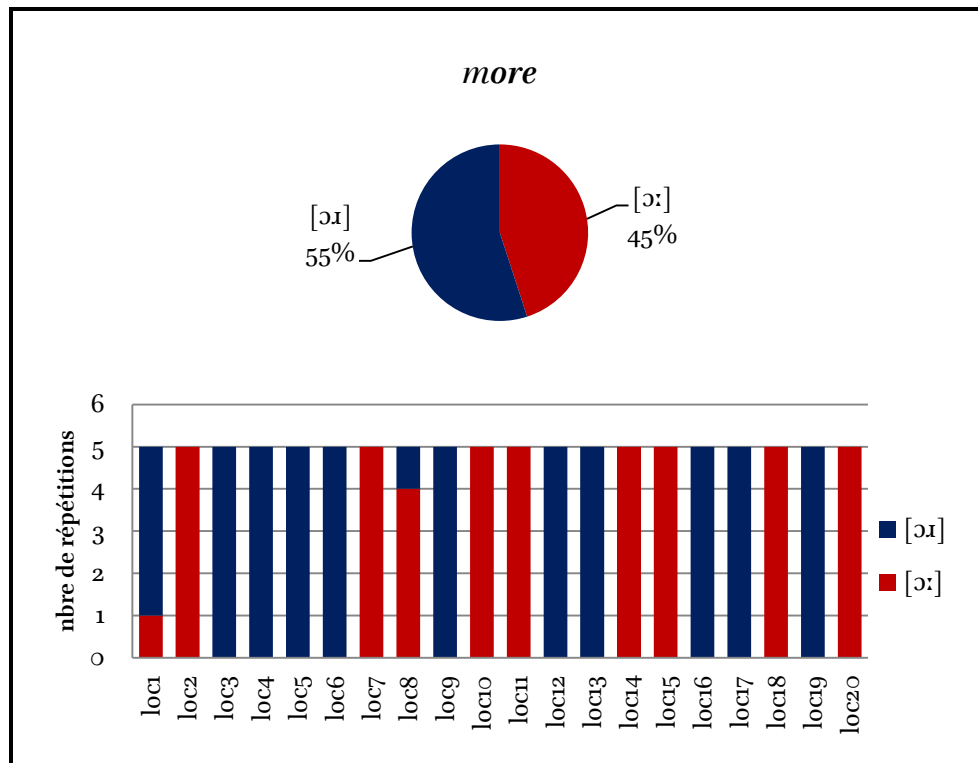


Figure 72 : pourcentages et répartition par locuteur de la production du <or> de *more*.

8.2.2 Organisation

La prononciation de la première syllabe du mot *organisation* est réalisée avec une dominance de la production non-rhotique [ɔ:] (58 %), suivi par la production rhotique [ɔɪ] (42 %). Il y a une seule occurrence de la production de [o] comme dans le mot français *pot* (/po/). Nous le comptons parmi les productions non-rhotiques, car il n'y a pas présence de [ɪ], c'est juste un changement dans le timbre de la voyelle. Concernant les productions homogènes, les locuteurs 2, 7, 8, 10, 12, 14, 15, 17 et 20 ne produisent que des [ɔ:] tandis que les locuteurs 4, 5, 9, 16, 18 et 19 ne produisent que des [ɔɪ]. Les autres procèdent à un mélange (locuteurs 1, 3, 6, 11 et 13).

Les proportions de la production des deux sons [ɔ:] et [ɔɪ] restent similaires dans les deux modalités : 52 % en VN contre 58 % en VR pour la non-rhoticité ; 48 % en VN contre 42 % en VR pour la rhoticité (cp. Figure 33, p. 212). Les locuteurs 2, 7, 12, 14 et 15 pour la non-rhoticité et les locuteurs 4, 5, 9, 16 et 18 pour la rhoticité conservent des productions homogènes dans les deux

modalités. Les autres changent et mélangent des productions entre les deux modalités (locuteurs 1, 3, 6, 8, 10, 11, 17, 19 et 20).

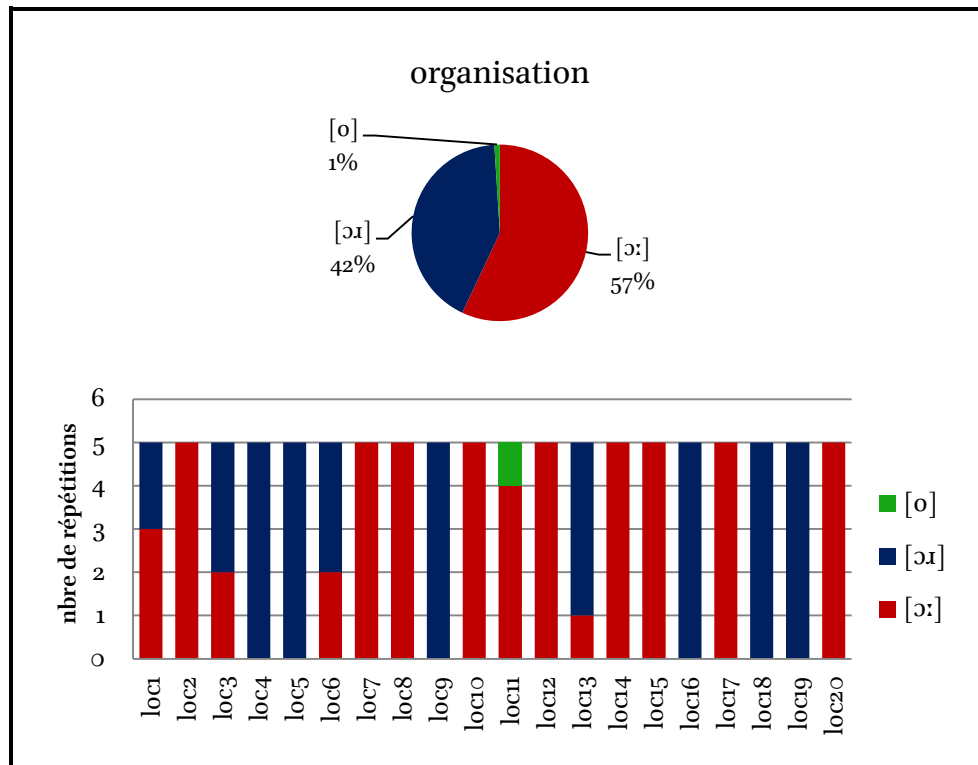


Figure 73 : pourcentages et répartition par locuteur de la production du <or> de *organisation*.

8.2.3 Dark

nombre de répétitions

La rhoticité domine dans la prononciation du mot *dark* avec 52 % de [ɔɪ] et 5 % de [ɔɪ], la non-rhoticité est produite à 43 % ([ɔ:]). Comme dans le cas précédent le locuteur 2 produit un son différent de celui qui est attendu : un [ɔ] au lieu d'un [ɔɪ]. Seuls les locuteurs 5 et 12 mélangent les deux sons [ɔɪ] et [ɔ:], tandis que les autres ont des productions homogènes. En effet, les locuteurs 1, 6, 7, 8, 10, 11, 15 et 20 ne produisent que des [ɔ:] pendant que les locuteurs 3, 4, 9, 13, 14, 16, 17, 18 et 19 ne produisent que des [ɔɪ]. Le locuteur 2 ne réalise que des [ɔ].

Même si la prononciation rhotique du mot *dark* domine dans les deux modalités, la production en [ɔɪ] a diminué en VR, passant de 63 % à 52 %. De manière évidente, la production de la non-rhoticité a augmenté de 37 % en VN à 43 % en VR. Les locuteurs 1, 7, 10, 11 et 15 gardent les mêmes productions non-rhotiques, de même que les locuteurs 3, 4, 9, 13, 16, 17, 18 et 19 avec leurs productions rhotiques, dans les deux modalités. Le locuteur 2 garde aussi les mêmes mélanges en

VN et VR. En revanche, les locuteurs 2, 6, 8, 12, 14 et 20 changent de productions entre les deux modalités (cp. Figure 34, p. 213).

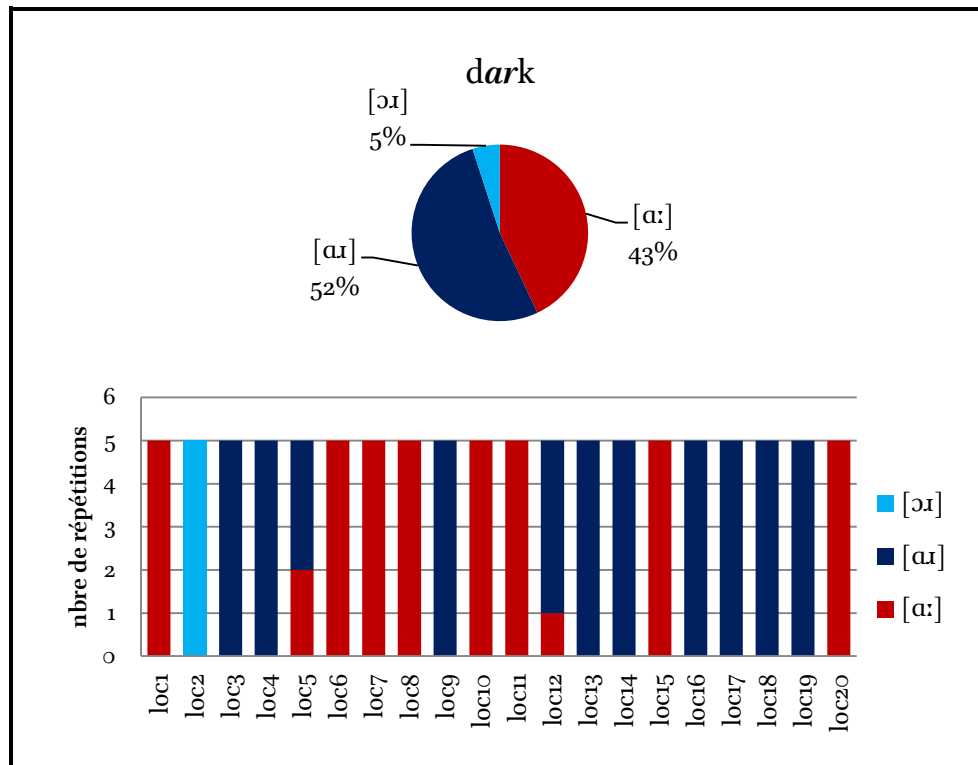


Figure 74 : pourcentages et répartition par locuteur de la production du <ar> de *dark*.

8.2.4 Party

La prononciation rhotique de la première syllabe du mot *party* représente les trois quarts de la production totale, la non-rhoticité occupe le quart restant. Cette dominance de prononciation avec un [ɑ:] se reflète chez les locuteurs 2, 3, 4, 5, 6, 8, 9, 12, 13, 14, 16, 18, 19 et 20, qui le produisent de façon homogène. Néanmoins, une production stable du [ɑ:] est observée chez les locuteurs 1, 10, 15 et 17, tandis que les locuteurs 7 et 11 mélangent les deux séries de sons.

La rhoticité domine dans les deux modalités mais avec une légère augmentation en VR (69 % en VN contre 75 en VR). La tendance inverse est notée pour la part de la production non-rhotique (31 % en VN contre 25 en VR). Les mêmes pourcentages de prononciation rhotique sont maintenus dans les deux modalités par les locuteurs 2, 3, 4, 5, 6, 8, 9, 12, 13, 18, 19 et 20, quant aux locuteurs 1, 10 et 15, ils gardent eux aussi les mêmes proportions de leur production non-rhotique (cp. Figure 35, p. 214). Les locuteurs 7, 11 et 16 mélangent toujours les deux séries de sons alors que les locuteurs 14 et 17 changent complètement de production d'une modalité à l'autre (rhotique en VN et non rhotique en VR pour le locuteur 17, et inversement pour le locuteur 14).

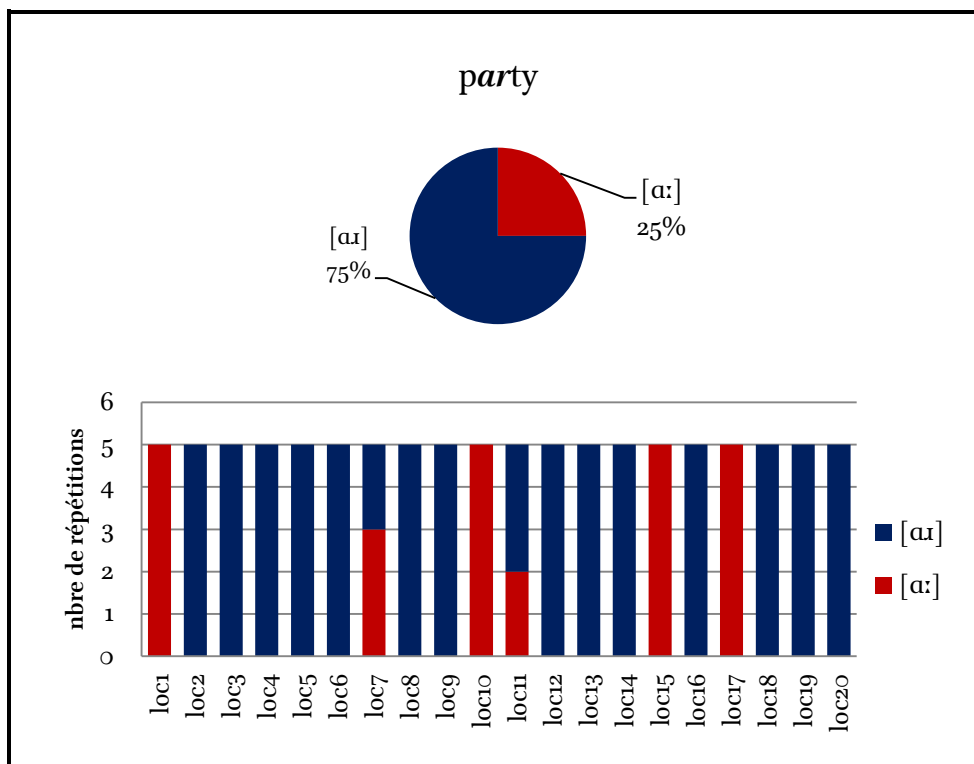


Figure 75 : pourcentages et répartition par locuteur de la production du <ar> de *party*.

8.2.5 Bird

Nous avons expliqué précédemment qu'il existe deux types de productions rhotiques : le premier consiste à prononcer une voyelle et le /r/ simultanément (ex : /ɜ:/), et le deuxième à prononcer une voyelle et le /r/ de manière consécutive (ex : /ɜr/). D'après nos résultats, uniquement les deux types de prononciation rhotique (rhotique₁, rhotique₂) sont constatés, mais avec une écrasante dominance du premier type rhotique₁. Ainsi le [ɜ:] est produit à 95 % et le [ɜr] à seulement 5 % (locuteur 14) sur la totalité des itérations. La forme non-rhotique [ɜ] n'est pas observée.

Outre l'absence de la prononciation non-rhotique, la production du [ɜ:] est dominante en VR plutôt qu'en VN. Nous constatons aussi qu'il n'y a pas de mélange de productions en VR (cp. Figure 36, p. 215). En effet, les locuteurs 10, 11, 12, 13, 15, 16 et 20 qui ont des productions non-rhotique ([ɜ]) et rhotique₂ ([ɜr]) en VN basculent tous en rhotique₁ ([ɜ:]) en VR.

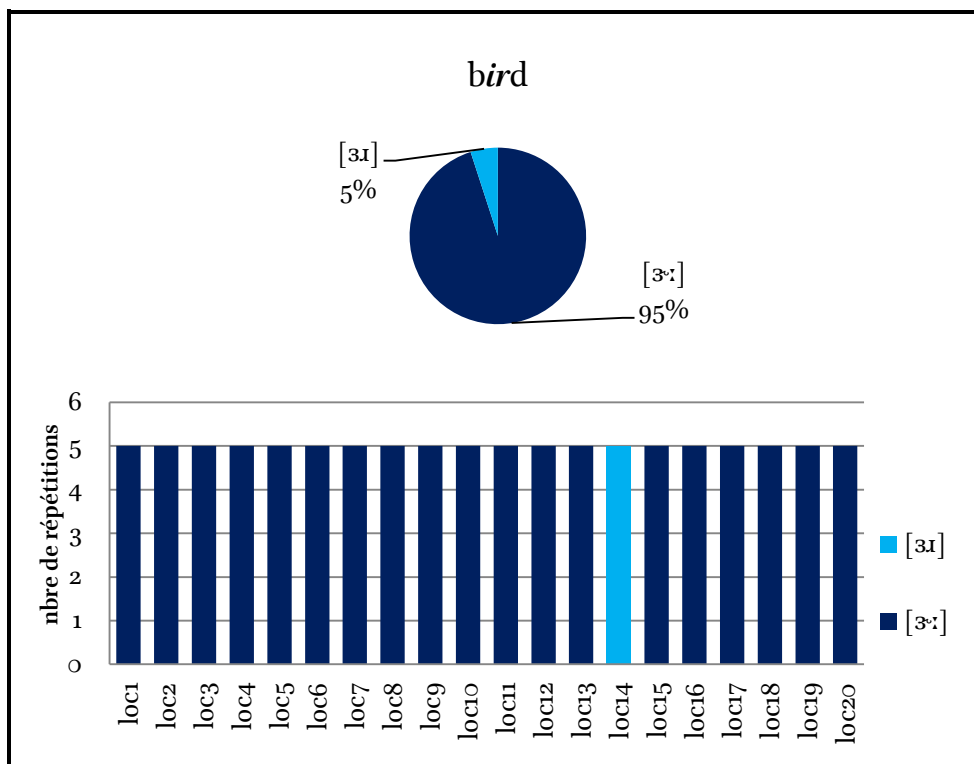


Figure 76 : pourcentages et répartition par locuteur de la production de <ir> de *actor*.

8.2.6 Actor

La forme non-rhotique de la deuxième syllabe du mot *actor* domine les productions avec [ə] (48 %) et [ɔ] (5 %), suivie des deux formes rhotiques [ɜː] (29 %) et [ɜɪ] (10 %). Nous assistons également à l'émergence d'une production sans voyelle du mot ([ækt]²²⁶) (8 %). Cette prononciation représenté par le son [-] est réalisée entièrement par le locuteur 11 et partiellement par le locuteur 7. Autrement dit, les locuteurs 2, 3, 4, 8, 10, 12, 15, 18 et 20 sont homogènes dans leur prononciation en [ə], de même que les locuteurs 5, 9, 16, 17 et 19 pour [ɜː] et finalement pour les locuteurs 13 et 14 avec le [ɜɪ].

La prononciation en [ɜː] est dominante en VN mais elle passe de 55 % à 29 %, c'est plutôt le [ə] qui devient dominant en VR en passant de 25 % à 48 %. Le [ɔ] a également diminué de 20 % en VN à 5 % en VR. Nous constatons l'apparition de [ɜɪ] et de la production sans voyelle en VR (cp. Figure 37, p. 216). Ensuite, nous remarquons que les locuteurs 5, 9, 16, 17 et 19 et les locuteurs 10, 12, 18 et 20 gardent respectivement les mêmes proportions de leurs productions homogènes dans les deux

²²⁶ Contrairement à la production du [ə], cette prononciation se manifeste sur le signal acoustique par une absence totale de la voyelle. Aucune indication formantique ou temporelle n'est relevée ni sur le spectrogramme ni sur l'oscillogramme (voir Annexe A 1, p. 30).

modalités. Aussi, parmi les locuteurs 2, 6, 14 et 15, seul le locuteur 6 a maintenu ses productions en [ɔ]. En revanche, les locuteurs 2, 3, 4 et 8 adoptent la production en [ə] plutôt qu'en VR, pendant que les locuteurs 1 et 7 mélangent deux sons en VR et non en VN.

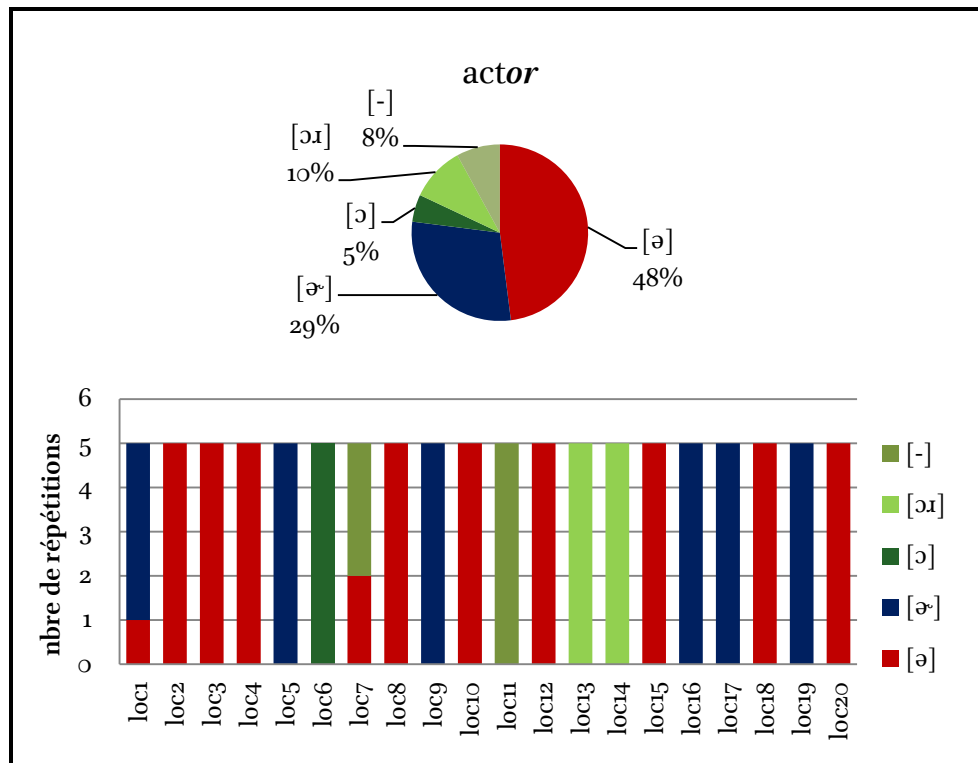


Figure 77 : pourcentages et répartition par locuteur de la production du <or> de *actor*.

8.2.7 Lecture

Nous observons une dominance de la prononciation rhotique dans la dernière syllabe du mot *leisure*. Ainsi, le [ə̃] est produit à 61 % contre 39 pour la prononciation en [ə]. Seul le locuteur 4 mélange les deux sons. Les autres sont homogènes dans leur prononciation du [ə] (locuteurs 2, 6, 7, 10, 12, 15 et 18) et du [ə̃] (locuteurs 1, 3, 5, 8, 9, 11, 13, 14, 16, 17, 19 et 20).

D'une part, la prononciation rhotique du mot *leisure* domine toujours dans les deux modalités (VN et VR) mais elle connaît une légère diminution en VR (de 69 à 61 %). D'autre part, les locuteurs 1, 3, 5, 8, 9, 11, 13, 14, 16 et 19 gardent les mêmes proportions de productions rhotiques et les locuteurs 2, 6, 10, 15 et 18 conservent les mêmes productions non rhotiques dans les deux modalités (cp. Figure 38, p. 217). En outre, concernant les locuteurs 7, 12 et 20 basculent de non-rhotique en VN à rhotique en VR, néanmoins l'inverse n'est pas constaté. Enfin, le locuteur 17 a mélangé les deux séries de sons en VN et non en VR, au contraire du locuteur 4 qui ne mélange rien en VN mais plutôt en VR.

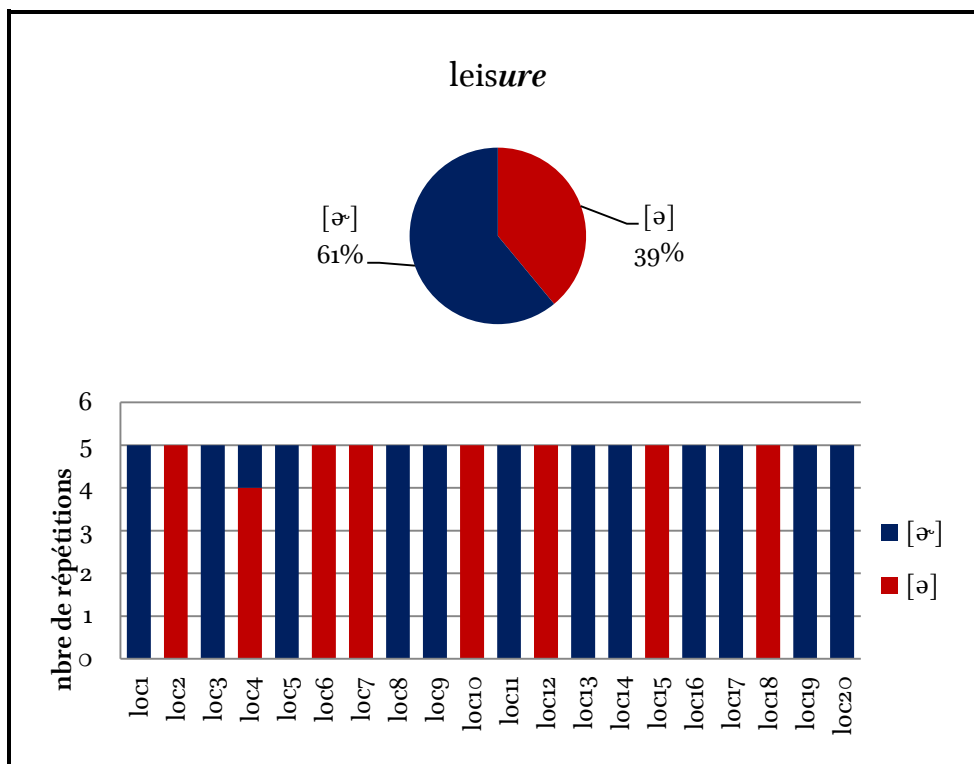


Figure 78 : pourcentages et répartition par locuteur de la production du <ur> de *leisure*.

8.2.8 Résultats d'ensemble pour la rhoticité

Les résultats de l'ensemble de ces mots permettent de mieux visualiser la part complète de rhoticité et de non rhoticité. Ainsi, sur un total de 700 répétitions (35 par locuteur), la rhoticité domine à 61 %. Remarquons la binarité dans les productions : soit elles sont rhotiques soit elles sont non rhotiques. Nous rappelons que ces pourcentages ne prennent en compte que l'aspect rhotique/non-rhotique, et non spécifiquement la différence entre la RP et le GA, puisque certaines productions qui ne relèvent ni de la RP ni du GA ([ɔ], [ɔɪ] et [-]). Les résultats sont similaires dans l'ensemble concernant les deux modalités : la rhoticité domine toujours, mais avec une légère diminution en VR (alors qu'elle est de 66 % en VN). Par ailleurs, les locuteurs 3, 4, 8 et 9 ont des productions entièrement rhotiques en VN alors qu'en VR, ce sont les locuteurs 9, 16 et 19 qui adoptent une articulation totalement rhotique. Les autres mélangent les deux types de productions (cp. Figure 39, p. 218).

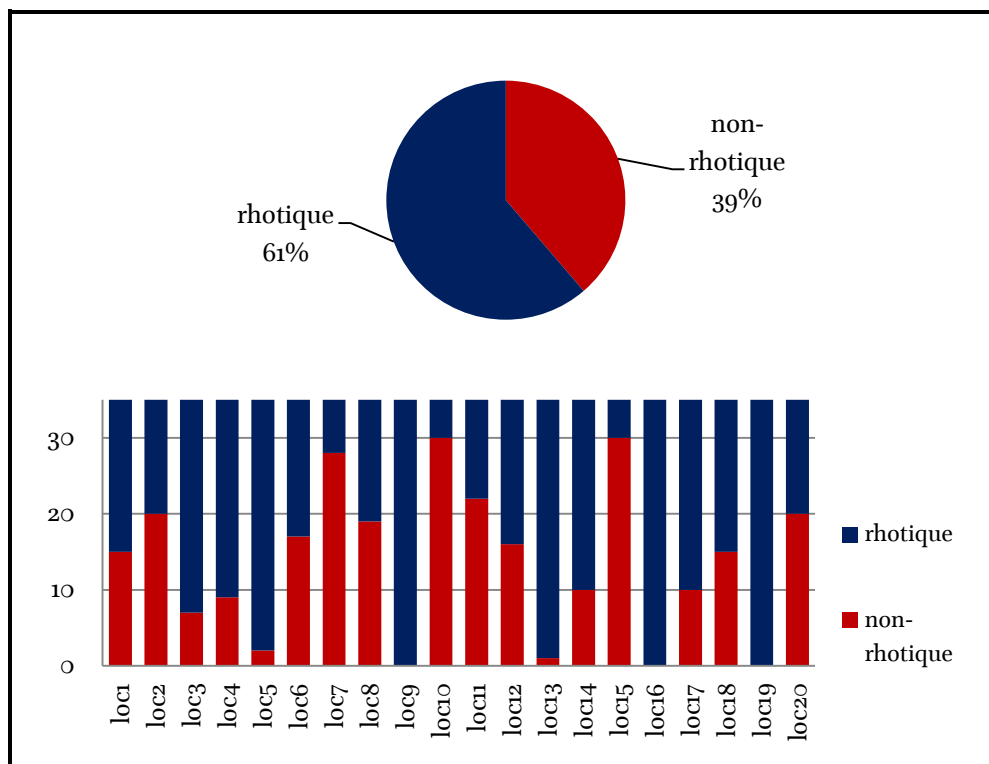


Figure 79 : pourcentages et répartition par locuteur de la rhocité.

8.3 Le yod-dropping

Les mots *tune*, *new* et *evolution* sont sujets à analyse quant au phénomène du *yod-dropping*.

8.3.1 Tune

La série de sons [ju:] domine la prononciation du mot *tune* à 85 %, suivie par le [u:] (8 %), le [ɜ:] (3 %) et enfin le [y] (2 %). Notons l'apparition d'un nouveau son que l'on retrouve dans le mot *bird* (/bɜ:d^{RP}) et qui est produit par le locuteur 2. Le locuteur 3 ne prononce que des [u:] alors que les locuteurs 9, 10 et 13 mélangent deux sons ([ju:], [u:] et [y]). Les autres sont homogènes dans leur production du [ju:].

Nous constatons des résultats similaires dans les deux modalités : [ju:] domine toujours (88 % en VN contre 85 % en VR), [u:] passe de 10 % en VN à 8 % en VR et [y] reste inchangé. Quant aux productions individuelles, nous remarquons tout d'abord que les locuteurs 9 et 13 sont homogènes dans leur production en [ju:] en VN mais mélangent le [ju:] et le [u:] en VR. Puis, il y a le locuteur 2 qui mélange [ju:] et [y] en VN mais adopte le son [ɜ:] partout en VN. Ensuite, nous observons que le locuteur 12 change d'une production [ju:] en VN en une production [u:] en VR. Finalement, le

locuteur 3 conserve la même production en [u:], pareil pour la prononciation en [ju:] par les locuteurs 1, 4 à 8, 11, 14 à 20 (cp. Figure 40, p. 219).

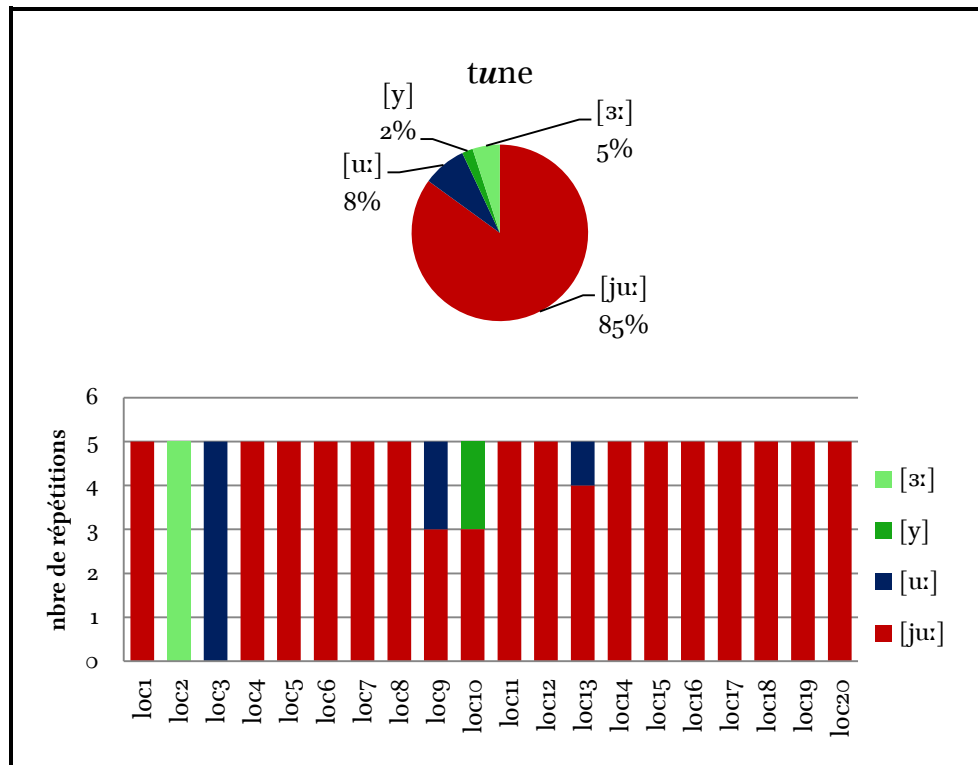


Figure 80 : pourcentages et répartition par locuteur de la production du <u> de *tune*.

8.3.2 New

Tous les locuteurs sauf un (loc 3) adoptent la prononciation en [ju:] pour ce mot. À part les locuteurs qui ont déjà une prononciation en [ju:] (1, 2, 4, 5, 6, 8, 13, 15, 16, 19 et 20), seul le locuteur 9 conserve sa prononciation en [u:]. Les autres basculent de [u:] (et du mélange) en VN à [ju:] en VR (locuteurs 3, 7, 10, 12, 14, 17 et 18) (cp. Figure 41, p. 220).

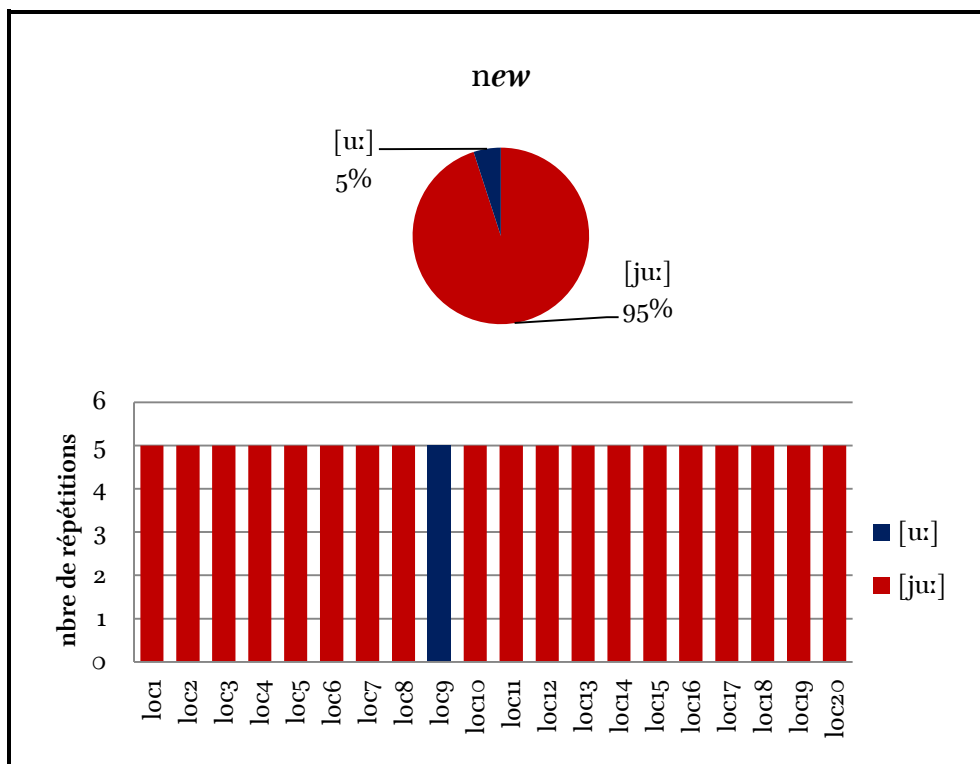


Figure 81 : pourcentages et répartition par locuteur de la production du <ew> de *new*.

8.3.3 Evolution

Concernant la prononciation du mot *evolution*, nous constatons la dominance du son [u:] (46 %), puis du son [y] (35 %), du [-] (15 %) et enfin du [ju:] (4 %). Ici également nous relevons l'absence de voyelle (représentée par [-]) entre le /l/ et /t/, c'est-à-dire que le <u> n'est pas produit (/i:və'ljʌn/). Ainsi, nous notons une prononciation homogène chez les locuteurs 1, 4, 9, 12, 15, 17, 18 et 20 pour le [u:], les locuteurs 3, 8 et 16 pour le [-], les locuteurs 6, 7, 13, 14 et 19 pour le [y]. Les autres mélangent soit [ju:] et [u:] (loc 5) soit [u:] et [y] (loc 2, 10 et 11).

Bien qu'il y ait une légère réduction de la proportion du [u:] en VR, ce son domine toujours la prononciation du mot *evolution* (de 53 % en VN à 46 % en VR). La part du [y] et du [ju:] passe de 37 en VN à 35 % en VR ; de 10 en VN à 4 % en VR respectivement. Rappelons que la production sans voyelle n'est pas observée en VN. En effet, les locuteurs 4, 9, 12, 17, 18 et 20 gardent des productions homogènes en [u:] dans les deux modalités, ainsi que les locuteurs 6, 13, 14 et 19 pour le [y] et le locuteur 11 pour le mélange de [u:] et de [y]. Les autres changent de sons entre la VN et la VR (cp. Figure 42, p. 221).

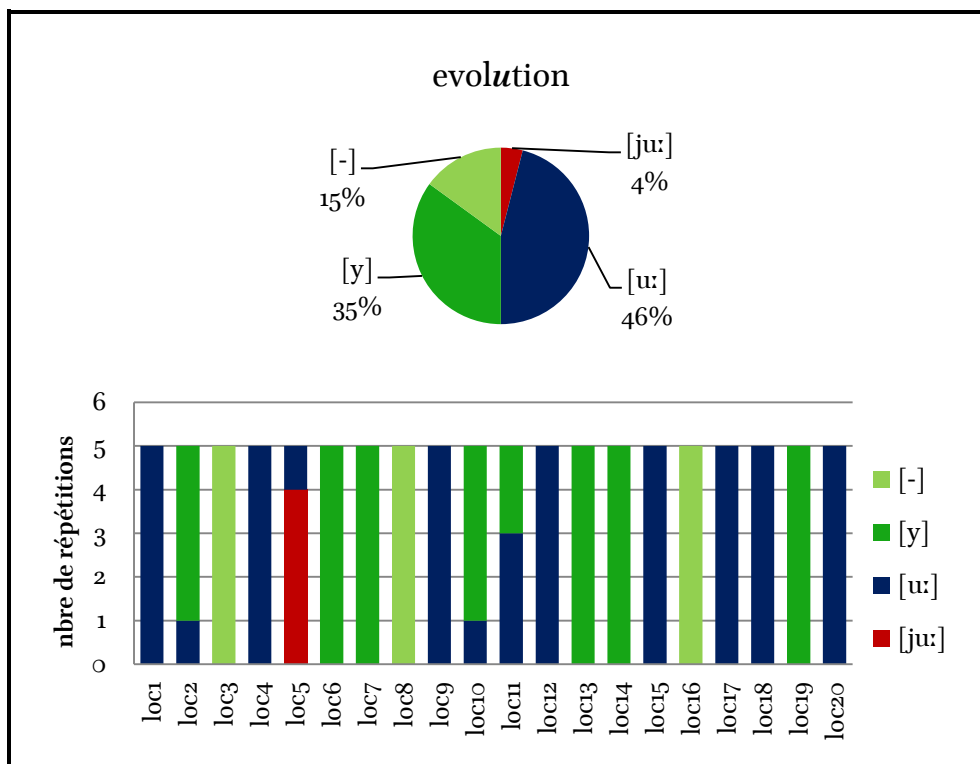


Figure 82 : pourcentages et répartition par locuteur de la prononciation du <u> dans *evolution*.

8.3.4 Résultats d'ensemble pour le yod-dropping

En rassemblant la production des sons et séries de sons étudiés dans les mots *tune*, *new* et *evolution*, nous pouvons pointer la dominance de la production en [ju:] qui augmente de 55 % en VN à 61 % en VR. S'ensuit celle du [u:] avec une réduction de 32 % en VN à 20 % en VR et celle du [y] de 13 % en VN à 12 % en VR. Les résultats individuels montrent que les locuteurs mélangent deux sons (locuteurs 1, 4 à 9, 14 à 20), trois sons (locuteurs 3, 10, 11, 12 et 13) et quatre sons (locuteur 2) en VR, alors que le mélange ne concerne que trois sons en VN. Nous constatons ainsi que l'augmentation du nombre de sons mélangés en VR est facilitée par l'apparition des deux nouveaux sons [ɜ:] et [-] (cp. Figure 43, p. 222).

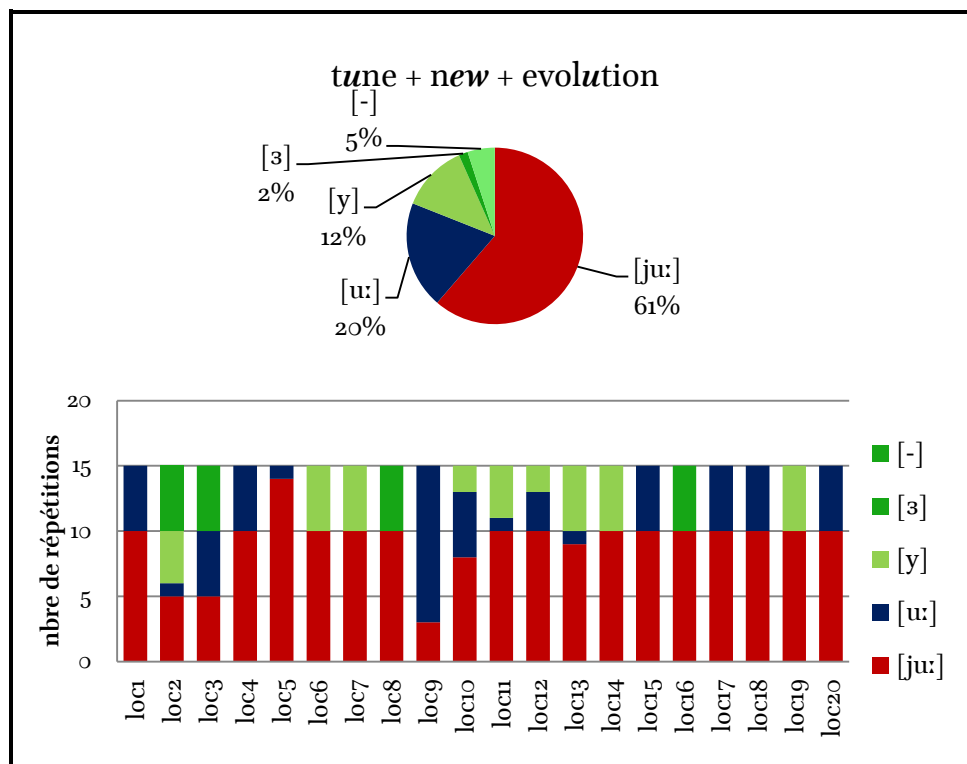


Figure 83 : pourcentages et répartition par locuteur du yod-dropping.

Les résultats globaux des éléments segmentaux consonantiques montrent une dominance de ceux de la RP (61 %), suivis de ceux du GA (37 %) et des « Autres » (10 %). En revanche, en prenant en compte les résultats par groupes d'éléments, on remarque une production équivalente de la RP et du GA pour le /t/ intervocalique (38 % chacun) ; une dominance des éléments proches du GA pour la rhoticité (61 %) et enfin une dominance des caractéristiques proches de la RP pour le yod-dropping (61 %). Nous constatons une omniprésence de la variabilité par rapport à la prononciation des mots au sein de chaque groupe étudié (Tableau 20). D'autre part, la comparaison des résultats dans les deux modes de production indique que globalement les caractéristiques proches de la RP dominant VR (53 %), alors qu'ils occupent la deuxième position en VN. Les éléments « Autre » et ceux proches du GA diminuent entre les deux modalités (13 % en VN et 10 % en VR ; 47 % en VN, 37 % en VR, respectivement) (cp. Tableau 20, p. 222).

Tableau 23 : résumé des éléments segmentaux consonantiques.

Phénomène étudié	Mots	% de caractéristiques de la RP	% de caractéristiques du GA	% de caractéristiques « Autres »
/t/ intervocalique	<i>City</i>	30	34	36
	<i>Vitamin</i>	50	32	18
	<i>party</i>	47	34	19

	total	38	38	24
Yod-dropping rhoticité	<i>More</i>	45	55	0
	<i>Organisation</i>	58	42	0
	<i>Dark</i>	43	57	0
	<i>Party</i>	25	75	0
	<i>Actor</i>	100	0	0
	<i>Actor</i>	61	39	0
	<i>Leisure</i>	39	61	0
	<i>total</i>	39	61	0
Yod-dropping	<i>Tune</i>	85	8	7
	<i>New</i>	95	5	0
	<i>Evolution</i>	4	46	50
	<i>total</i>	61	20	19
Total des éléments consonantiques		53	37	10

Nous avons aussi noté une variabilité dans toutes les productions à deux niveaux : entre les locuteurs (inter-locuteur) et au sein des productions d'un même locuteur (intra-locuteur). Finalement, nous présentons les résultats des éléments segmentaux vocaliques dans les prochaines sections. Il s'agit de la prononciation de la voyelle de *LOT*, du *GOAT*, de la prononciation des graphies <i>, <y>, <ei> et <e> et enfin la nasalisation vocalique.

8.4 La prononciation de la voyelle de LOT

L'étude de la qualité de la voyelle <o> dans les mots *hot* et *top* montre des réalisations binaires : /ɒ/ pour la RP et /ɑ:/ pour le GA.

8.4.1 Hot

Le mot *hot* est prononcé avec une dominance de la voyelle [ɒ] (80 %) contre 20 % pour la voyelle [ɑ:]. Aucun autre son n'est observé. Toutes les productions sont homogènes : tandis que les locuteurs 8, 9, 19 et 20 ne réalisent le mot qu'avec des [ɑ:], les autres ne produisent que des [ɒ].

La comparaison montre que la production en [ɒ], déjà dominante en VN, augmente en VR (de 60 % en VN à 80 % en VR) (cp. Figure 44, p. 225). Nous remarquons l'absence du son [ɔ] en VR, alors qu'il est produit à 20 % en VN. En effet, les locuteurs qui prononcent le mot avec ce dernier son ([ɔ]) en VN, produisent plutôt des [ɒ] (locuteurs 2, 11 et 14) et des [ɑ:] (locuteur 20) en VR. Dans la même lignée, les locuteurs 4, 7, 8 et 9 produisent de [ɑ:] en VN mais des [ɒ] en VR, le

contraire étant observé chez les locuteurs 8, 9, 19 et 20 qui produisent des [v] en VN mais des [ɑ:] en VR. Les autres conservent les mêmes sons dans les deux modalités.

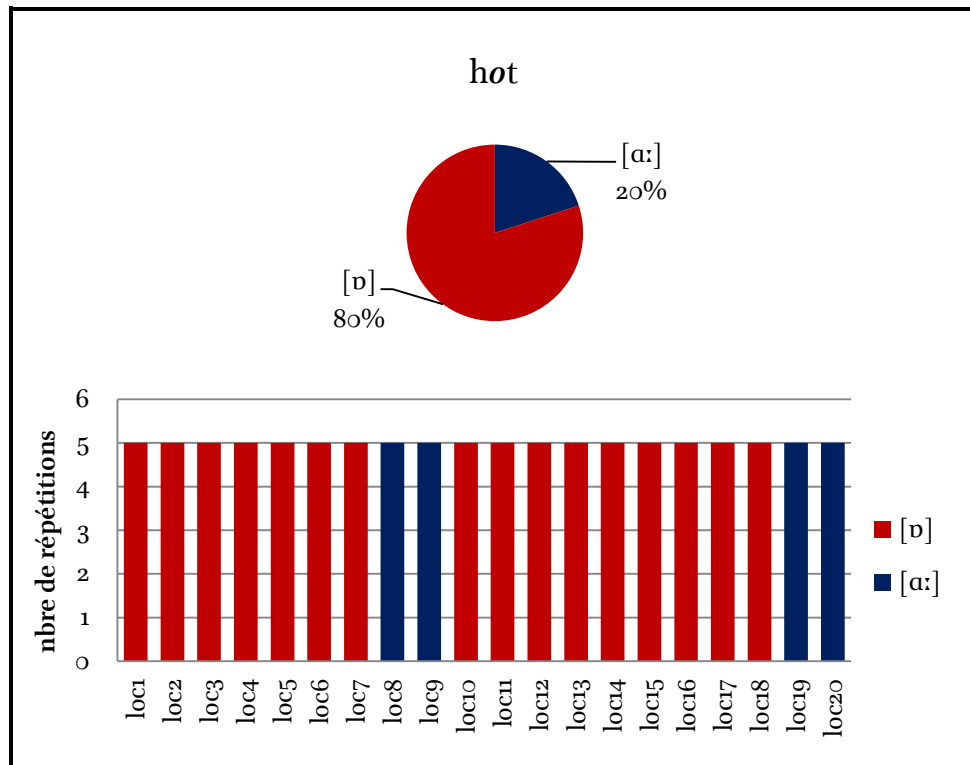


Figure 84 : pourcentages et répartition par locuteur de la production du <o> de *hot*.

8.4.2 Top

Deux sons sont employés dans la prononciation du mot *top* : le son [v] domine à 89 % et le son [ɑ:] est produit à seulement 11 %. Les locuteurs 7, 9, 11 et 19 mélangent les deux sons alors que les autres ne produisent que des [v].

Comme dans le cas de *hot*, il y'a aucun autre son à part [v] et [ɑ:] et la [v]. Cette dernière domine toujours les productions et augmente de 65 % en VN à 89 % en VR. Les locuteurs ayant adopté le [ɔ] et le [ɑ:] en VN change leur prononciation en [v] en VR. Signalons qu'il n'y a pas de mélange en VN mais plutôt en VR par les locuteurs 7, 9, 11 et 19 (cp. Figure 45, p. 226).

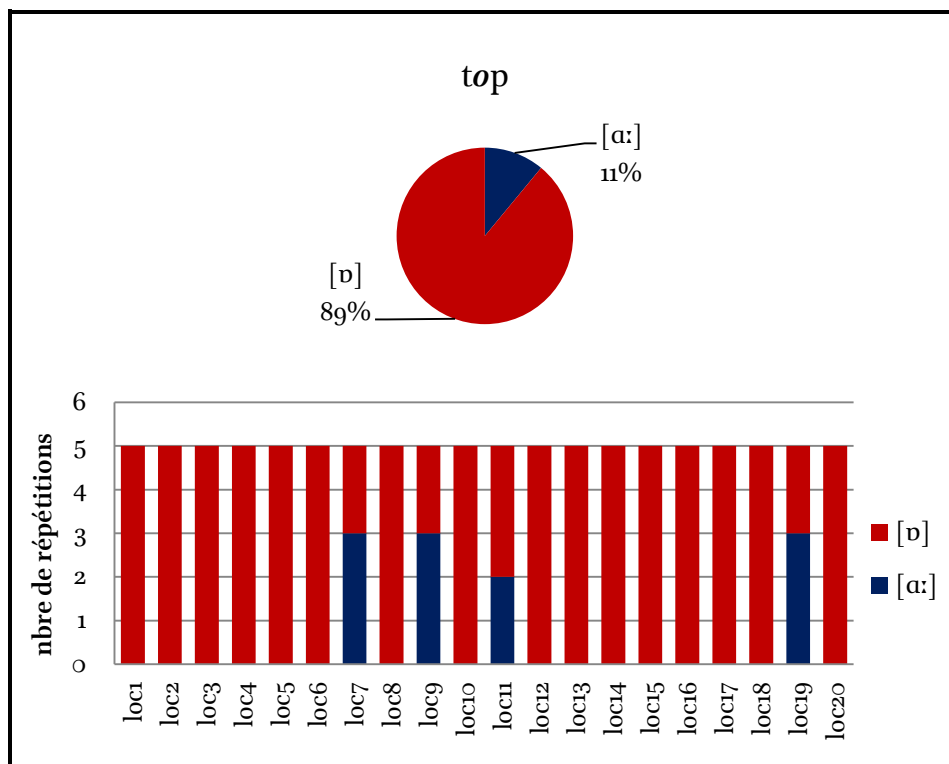


Figure 85 : pourcentages et répartition par locuteur de la production du <o> de *top*.

8.4.3 Hot et top

Comme conséquence évidente de l'augmentation du son [v] dans les deux mots et l'absence de sons « Autres », le son [v] domine toutes les productions à 85 % sur les 200 occurrences. Globalement, tous les locuteurs prononcent les deux mots avec un [v] excepté les locuteurs 7, 8, 9, 11, 19 et 20 qui mélangent [v] et [ɑ:].

Évoluant de 63 % en VN à 85 % en VR, la production des deux mots en [v] reste la dominante. Il y a une légère diminution de la voyelle [ɑ:] qui passe de 17 % en VN à 15 % en VR. On remarque également que les productions en VN sont plus homogènes que celles de la VR : le mélange est réalisé par seulement un locuteur en VN contre cinq en VR (cp. Figure 46, p. 227). Néanmoins, comme nous l'avons déjà indiqué, il y a un son de moins en VR ([ɔ]).

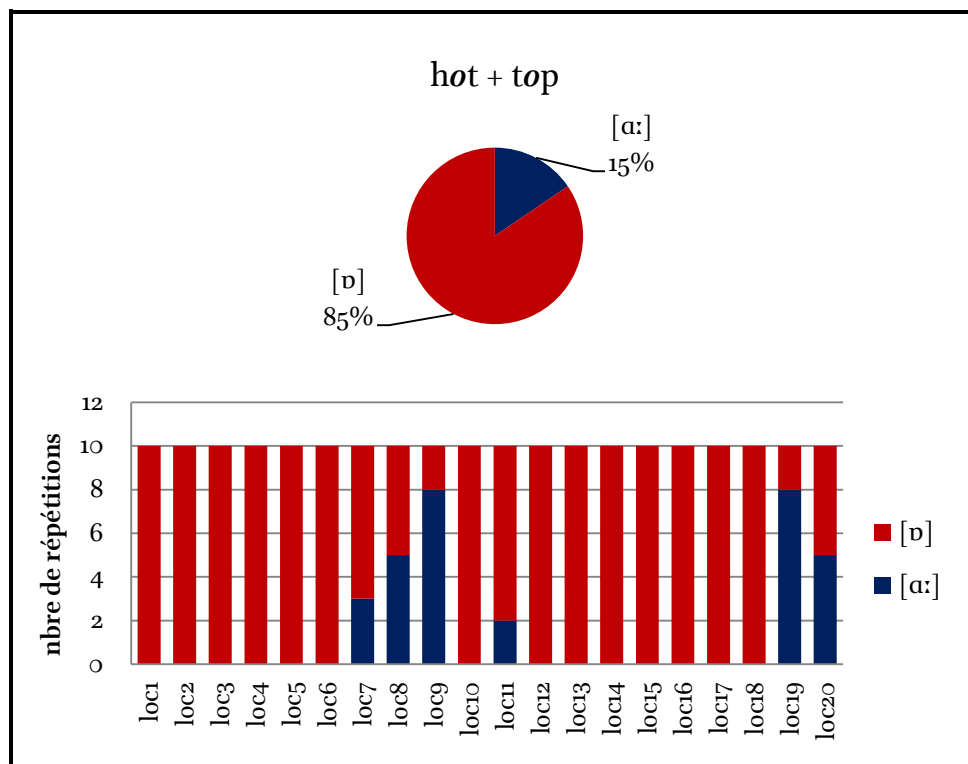


Figure 86 : pourcentages et répartition par locuteur de la production du <o> de *hot et top*.

8.5 La prononciation de la voyelle de GOAT

En plus de la série de sons /əʊ/ utilisée en RP et celle de /oʊ/ en GA, les sons /ɔ/ (/ɔm/ de homme) et /o/ (/po/ de pot) sont identifiés dans les analyses acoustiques des mots *joke* et *road*.

8.5.1 Joke

Les mêmes sons qui sont employés dans la prononciation du mot *joke* en VN se retrouvent aussi en VR mais à des proportions différentes. La suite de sons [oʊ] est produite avec une dominance de 52 %, suivie de [o] (40 %), du [ɔ] (7 %) et enfin du [əʊ] qui n'est réalisé qu'une seule fois (locuteur 12). Une production homogène est réalisée par : les locuteurs 1, 7, 8, 9, 14, 15, 16, 17 et 19 pour le son [oʊ], les locuteurs 3, 4, 5, 6, 13, 18 et 20 pour le son [o] et enfin le locuteur 2 pour le son [ɔ]. Des mélanges sont réalisés par les locuteurs 10 ([o] et [oʊ]), 11 ([o] et [ɔ]) et 12 ([o], [oʊ] et [əʊ]).

Nous notons une chute dans les résultats de la dominante [oʊ] qui passe de 75 % en VN à 52 % en VR et aussi celle de la suite de sons [əʊ] de 10 % à 1%. Nous constatons également une ascension de la production de [o] passant de 8 % en VN à 40 % en VR, pendant que le son [ɔ] reste stable (7 itérations) dans les deux modalités (cp. Figure 47, p. 228). Les résultats individuels montrent que les locuteurs 1, 8, 9, 14, 15, 16, 17 et 19 maintiennent les mêmes résultats de [oʊ] en VN

qu'en VR, de même pour les locuteurs 11 et 20 pour les sons [o] et le mélange [o]/[ɔ] respectivement. Enfin, les locuteurs 3, 4, 5, 6, 13 et 18 produisent plutôt des [o] en VR alors qu'ils réalisent des [əʊ] et [oʊ] en VN.

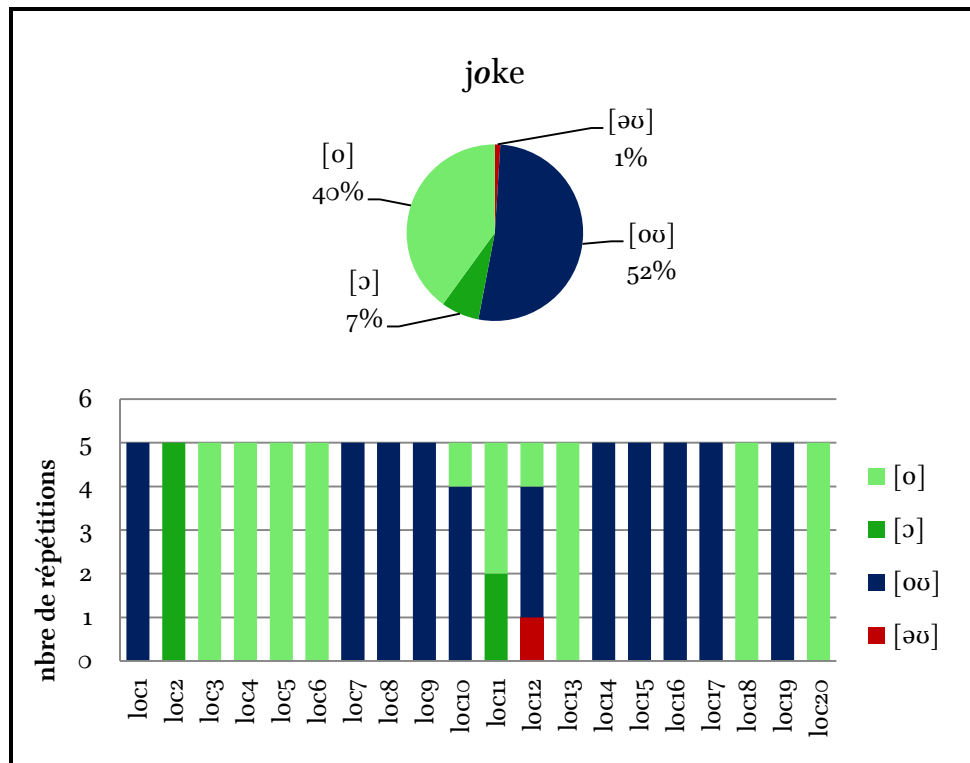


Figure 87 : pourcentages et répartition par locuteur de la production du <o> de *joke*.

8.5.2 Road

La production du [o] est dominante dans la prononciation du mot *road* (49 %) mais elle est suivie immédiatement par celle du [oʊ] (46 %). Le son [ɔ] n'est produit qu'à cinq reprises. Des productions analogues sont observées pour les locuteurs 1, 9, 12, 13, 15, 16, 17 et 19 ([oʊ]), les locuteurs 3 à 8, 11, 18 et 20 ([o]), pendant que les locuteurs 2, 10 et 14 mélangent deux et trois sons ([o] et [ɔ] ; [o] et [oʊ] ; [o],[ɔ] et [oʊ]).

Concernant la comparaison des deux modalités, la production de [oʊ] domine en VN (80 %), alors que le [o] l'emporte en VR (49 %). Remarquons que la diphtongue [əʊ] n'est cette fois pas produite. La diminution de la production du son [oʊ] est visible chez les locuteurs 3, 4, 5, 6, 8 et 20 qui produisent des [oʊ] en VN mais basculent dans la production du [o] en VR. Hormis les locuteurs 1, 9, 12, 13, 15, 16, 17 et 19 qui préservent les mêmes productions dans les deux modalités, le locuteur 7 passe de la production [əʊ] en VN en [o] pour la VR (cp. Figure 48, p. 229). Nous

constatons également que seul le locuteur 11 mélange les sons ([o] et [ɔ]) en VN contre trois locuteurs (2, 10 et 14) qui mélangent des sons en VR.

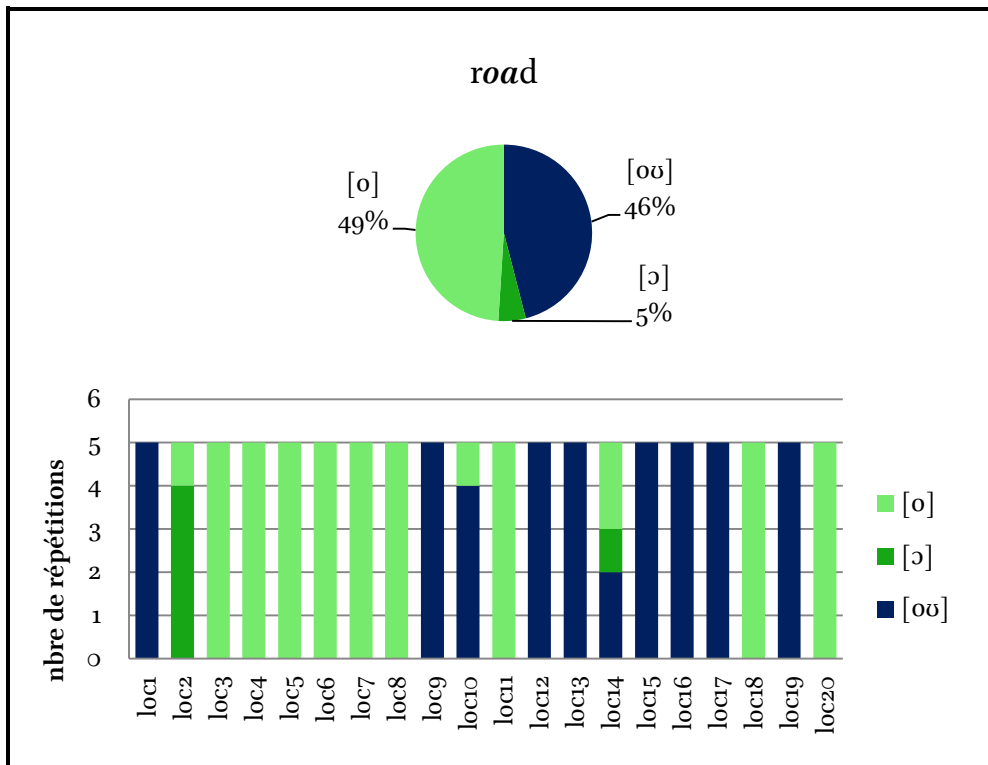


Figure 88 : pourcentages et répartition par locuteur de la production du <oa> de *road*.

8.5.3 Joke et road

En dépit de la légère augmentation de la production du son [o] dans le mot *road*, le son [ou] domine dans la totalité des productions des deux mots (49 %), suivi du [o] (45 %), du [ɔ] (6 %) et enfin le [əu] reste sur une seule production. Dans l'ensemble, il y a des mélanges de deux à trois sons chez les locuteurs 2, 7, 8, 10 et 14 pendant que les locuteurs 1, 9, 15, 16, 17 et 19 ne produisent que des [ou] et les locuteurs 3, 4, 5, 6, 18 et 20 ne réalisent que des [o].

Les mêmes sons et le même nombre se retrouvent dans les deux modalités. Les différences majeurs concernent la chute de la production de la diphtongue [ou] en VR (49 % contre 78 % en VN), du son [ɔ] et celui la diphtongue [əu] et l'augmentation de la production de [o]. C'est ainsi que les locuteurs 1, 9, 15, 16, 17 et 19 conservent la même proportion de la diphtongue [ou] dans les deux modalités. Enfin, à part les locuteurs 2, 11 et 14 qui produisent des [ɔ] et des [əu], les autres ont une prononciation plutôt en [o] pour les deux mots (cp. Figure 49, p. 230).

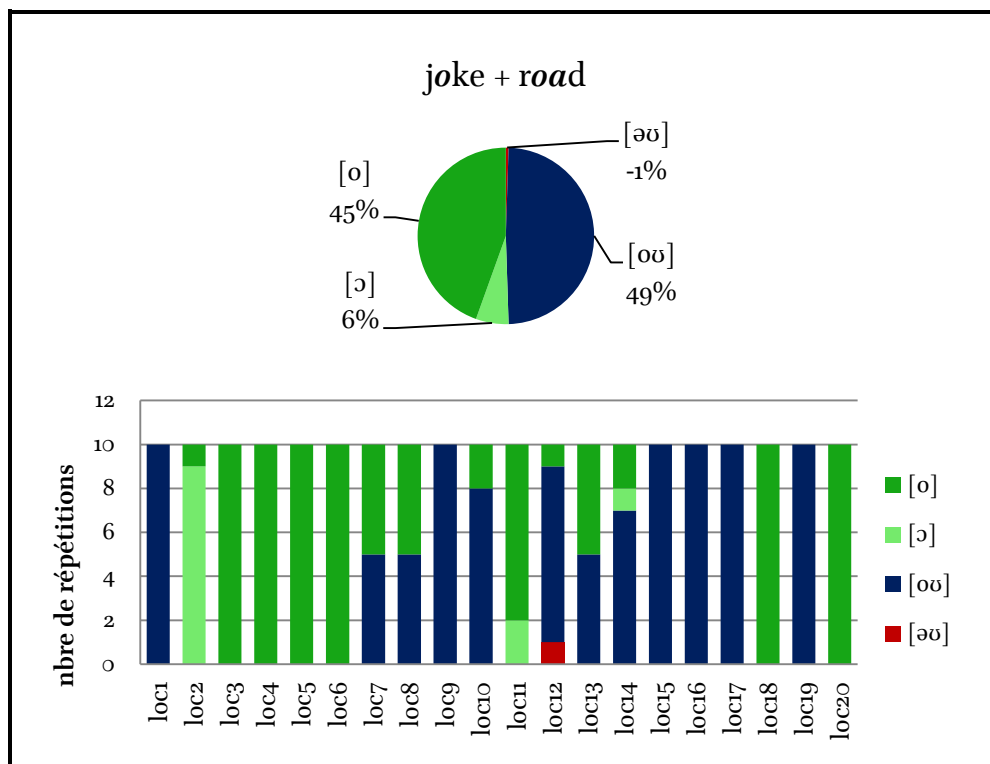


Figure 89 : pourcentages et répartition par locuteur de la production de <o> et <oa> de *joke* et *road*.

8.6 La prononciation des graphies <i> et <y>

La première syllabe des mots *dynasty* et *vitamin* est prononcée avec une monophthongue /ɪ/ en RP (/ˈdɪnəsti/, /ˈvɪtəmm/) et plutôt avec une diphtongue /aɪ/ en GA (/ˈdɑməsti/, /ˈvaɪrəmm/).

8.6.1 Dynasty

La graphie <y> de la première syllabe du mot *dynasty* est prononcée de quatre manières différentes par nos locuteurs. Outre la diphtongue [aɪ] qui domine à 63 % et le [ɪ] (15 %), nous assistons à la production de deux nouveaux sons : le [i] (21 %) et le [ɔɪ] (juste une seule fois). Des productions homogènes sont constatées chez les locuteurs 1, 11, 13 et 14 pour le son [i], les locuteurs 3, 4, 6, 8, 9, 10, 12, 16, 17, 19 et 20 pour le [aɪ] et les locuteurs 5, 7 et 15 pour le [ɪ]. Les locuteurs 2 et 18 produisent des mélanges de [aɪ] et [ɔɪ] ; [aɪ] et [i] respectivement.

En comparant les deux modalités, nous apercevons en dehors de l'apparition de [ɔɪ] et [i], que le [aɪ] reste supérieur, avec toutefois une augmentation dans sa production en VR (de 54 à 63 %). [ɪ] diminue évidemment de 46 à 15 %. Les résultats de la comparaison entre la VN et la VR montrent que les locuteurs 4, 6, 8, 9, 10, 12 et 19 gardent leur [aɪ] et les locuteurs 5, 7 et 15 maintiennent le [ɪ] dans les deux systèmes. Aussi, pendant que les locuteurs 1, 11, 13 et 14 adoptent

une production complète en [i] en VR, les autres changent de productions entre les deux modalités (cp. Figure 50, p. 231).

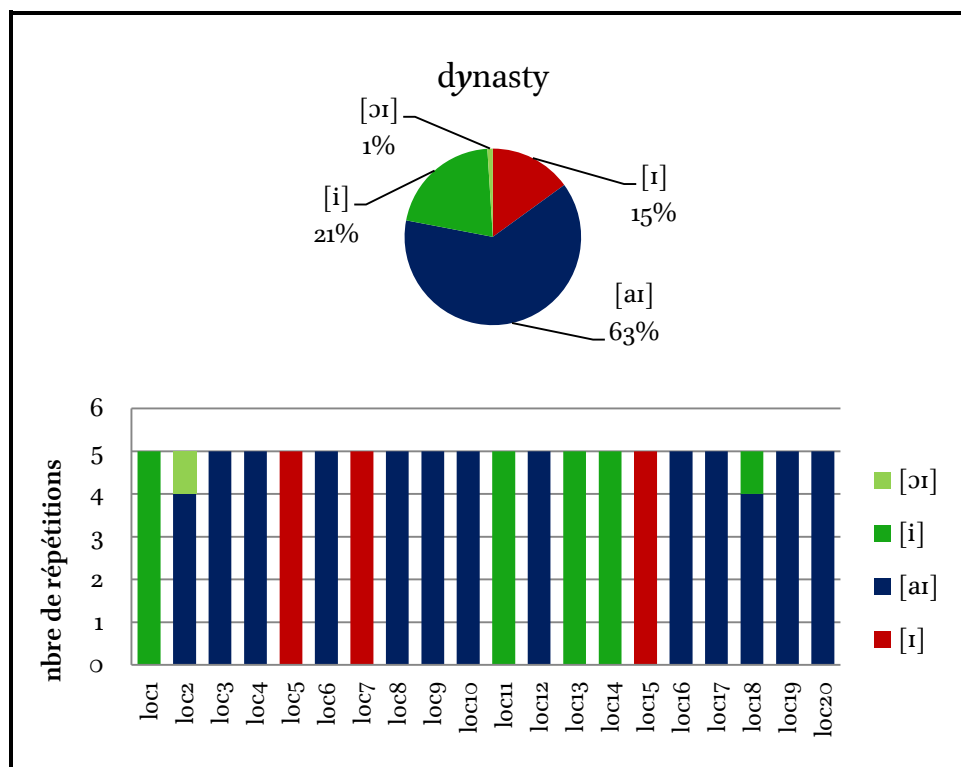


Figure 90 : pourcentages et répartition par locuteur de la production du premier <y> de *dynasty*.

8.6.2 Vitamin

En ce qui concerne la prononciation de la première syllabe du mot *vitamin*, la suite de sons [ai] domine de peu avec 38 %, s'en suivent les sons [ɪ] et [i] (36 % et 26 % respectivement). Quant au niveau individuel, nous recensons des productions homogènes chez les locuteurs 1, 4, 11, 14 et 20 (pour [i]), les locuteurs 2, 3, 10, 12, 17, 18 et 19 (pour [ai]), les locuteurs 5, 7, 8, 9, 13, 15 et 16 (pour [ɪ]). Le locuteur 6 est le seul à mélanger les sons [ai], [ɪ] et [i].

Comme dans le cas de *dynasty*, nous remarquons que le [i] n'est pas produit en VN, mais qu'il émerge en VR. Aussi, le [ɪ] domine en VN tandis que le [ai] domine en VR. Par conséquent la production du [ɪ] est moins importante en VR (de 62 à 36 %). Ensuite, nous notons que les locuteurs 2, 3, 10, 12, 18 et 19 (pour le [ai]), les locuteurs 5, 7, 8, 9, 13, 15 et 16 (pour le [ɪ]) conservent les mêmes productions dans les deux modalités, alors que les autres changent de production d'une modalité à l'autre, notamment les locuteurs 1, 4, 11, 14 et 20 qui basculent vers le [i] en VR. Enfin, il n'y a qu'un locuteur (6) qui mélange en VR contre deux (6 et 17) en VN (cp. Figure 51, p. 232).

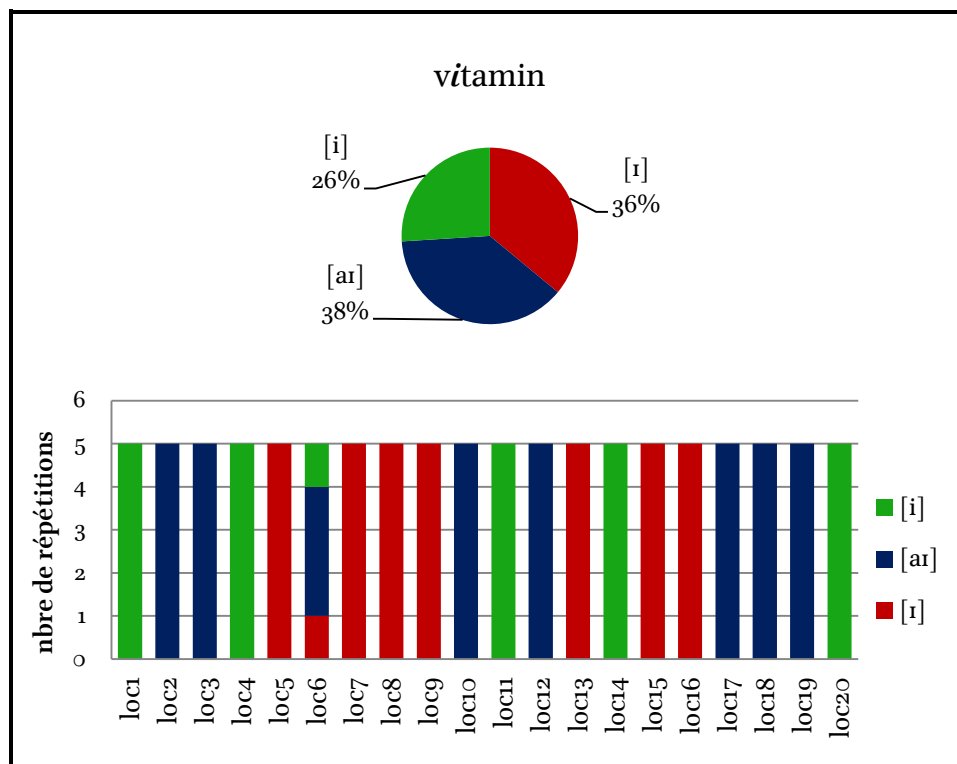


Figure 91 : pourcentages et répartition par locuteur de la production de <i> de *vitamin* en VR.

8.6.3 Dynasty et vitamin

La prononciation de la première syllabe des deux mots montre les résultats suivants : [ai] domine sur la moitié des productions (50 %), suivi du [ɪ] (25 %), du [i] (24 %) et enfin du [ɔɪ] (toujours une seule occurrence). Quant aux réalisations homogènes, les locuteurs 1, 11 et 14 ne produisent que des [i], les locuteurs 3, 10, 12, 17 et 19 ne produisent que des [ai] et enfin les locuteurs 5, 7 et 15 produisent des [ɪ]. Les autres réalisent des mélanges de deux (locuteurs 2, 4, 8, 9, 13, 16, 18 et 20) à trois sons (locuteur 6).

Ainsi, le [ai] domine plutôt en VR tandis que le [ɪ] est moitié moins présent (54 à 25 %) en raison de l'apparition du son [i] et de la montée du [ai]. Enfin nous notons la conservation des mêmes productions et proportions chez les locuteurs 5, 7 et 15 (pour le [ɪ]) ; 10, 12 et 19 (pour le [ai]) et 8, 9 et 16 (pour le mélange [ai] et [ɪ]) (cp. Figure 52, p. 233).

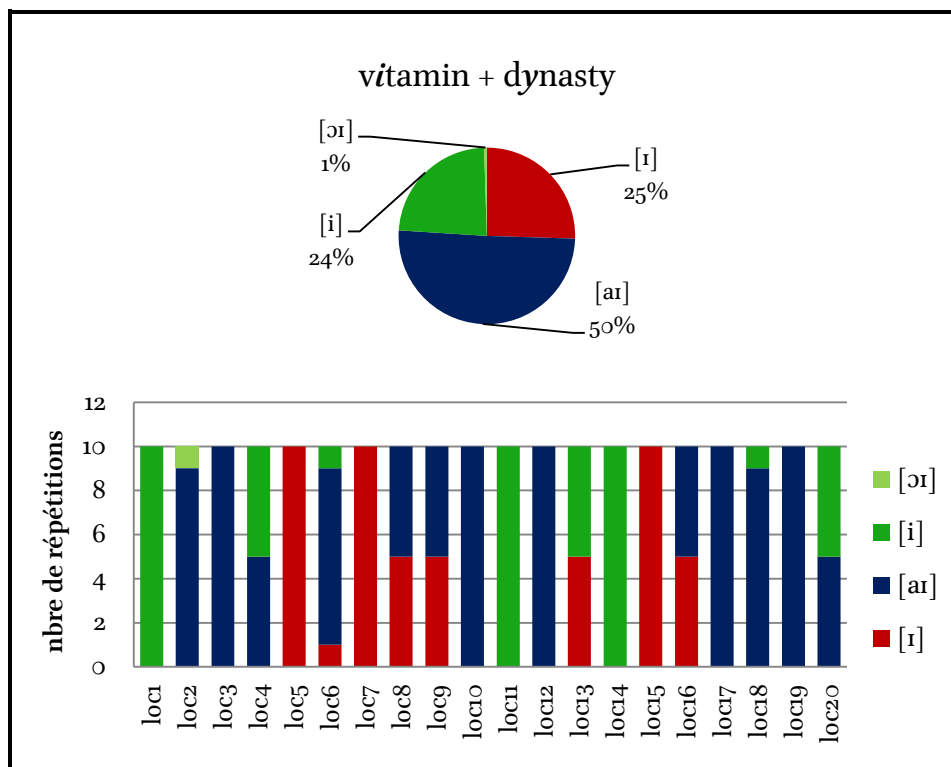


Figure 92 : pourcentages et répartition par locuteur de la production de <y> et <i> de *dynasty* et *vitamin*.

8.7 La prononciation de la graphie <i>

La voyelle de la troisième syllabe des mots *organisation* et *legalisation* (<i>) est produite avec soit la diphtongue [aɪ], soit la monophthongue [ɪ] en RP et est plutôt réalisée avec un [ə] en GA.

8.7.1 Organisation

Considérant la prononciation du mot *organisation*, nous remarquons que la voyelle entre le <n> et le <s> n'est pas produite (troisième et quatrième syllabe). C'est le cas des mots *actor* et *evolution* que nous représentons par le symbole [-] ($\text{ɔ:gəʃn}[-]'\text{zeɪʃən}^{\text{RP}}$). Ainsi, le [ə] vient en tête des productions avec 55 %, suivi du [-] (28 %), du [ɪ] (11 %) et enfin du [aɪ] (6 %). En effet, les sons et série de sons suivants (entre parenthèses) sont réalisés de manière homogène par les locuteurs 1, 5 et 15 ([aɪ]), les locuteurs 3, 4, 7, 9, 11 et 20 ([ə]), les locuteurs 16 et 19 ([ɪ]) et enfin le locuteur 12 ([aɪ]). Les autres locuteurs mélangent les différentes productions : [ə] et [-] (locuteurs 2, 6, 8, 14, 17 et 18), [ə] et [aɪ] (locuteur 10) et enfin [ə] et [ɪ] (locuteur 13). Notons par ailleurs l'apparition du mélange de [ə] et [-] (six locuteurs).

Les résultats de la comparaison des deux modalités montrent une inversion de tendances : d'une part, le [ɪ] est dominant à 75 % en VN alors qu'il est moins fréquent en VR (11 %) et d'autre

part, le [ə] apparaît plus souvent en VR (55 %) qu'en VN (16 %). Le [aɪ] est plus souvent réalisé en VN (9 %) qu'en VR (6 %). Rappelons que la production de [-] contribue en partie à ces résultats (cp. Figure 53, p. 234). Par conséquent, outre les locuteurs 19 et 12 qui conservent leurs [ɪ] et [aɪ] respectivement dans les deux modalités et le locuteur 16 qui ne produit que des [ə] en VN (mais des [ɪ] en VR), les autres locuteurs passent de la production des [ɪ] en VN vers celle de [ɪ] et/ou [ə] en VR.

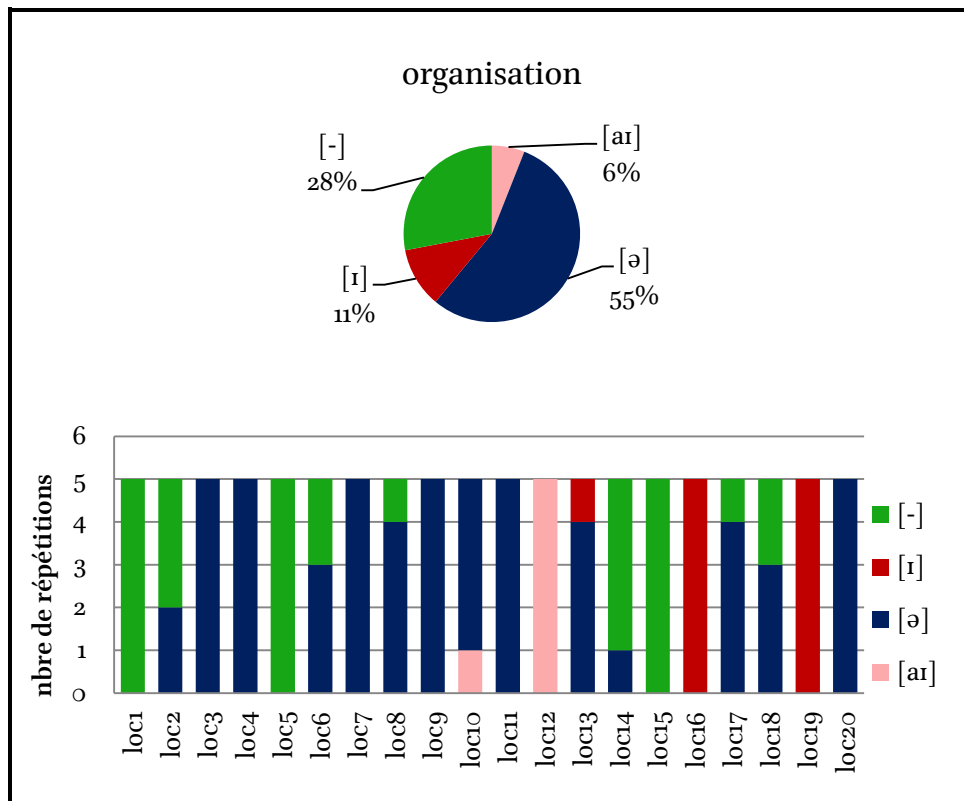


Figure 93 : pourcentages et répartition par locuteur de la production du premier <i> dans organisation.

8.7.2 Legalisation

La présence de [-] est également constatée dans le cadre de la prononciation du mot *legalisation*. Ainsi, nous observons une dominance de [ɪ] (45 %), suivi de près par le [ə] (31 %), de [-] (15 %) et enfin de [aɪ] (9 %). L'homogénéité des productions se retrouve notamment chez les locuteurs 1 et 18 (pour le [-]), les locuteurs 2 à 7, 11, 13 et 19 (pour le [ɪ]), le locuteur 10 (pour le [aɪ]) et en fin de compte les locuteurs 8, 9, 15, 16, 17 et 20 (pour le [ə]). En revanche, les locuteurs 12 et 14 mélangent des [aɪ] et [-], des [ə] et [-] respectivement.

En comparant ces résultats avec la VN, nous retenons : une moindre production de [ɪ] (76 à 45 %) et celle de [aɪ] (10 à 9 %), l'augmentation évidente de [ə] (14 à 31 %) et l'apparition du [-] en

VR. Le [ɪ] reste toutefois dominant dans les deux modalités (cp. Figure 54, p. 235). Les locuteurs 2 à 7, 11, 13 et 19 conservent le [ɪ] dans les deux modalités, de même que les locuteurs 10 (pour [aɪ]) et 17 (pour [ə]). Les autres (1, 12, 14 et 18) adoptent le [-] en VR.

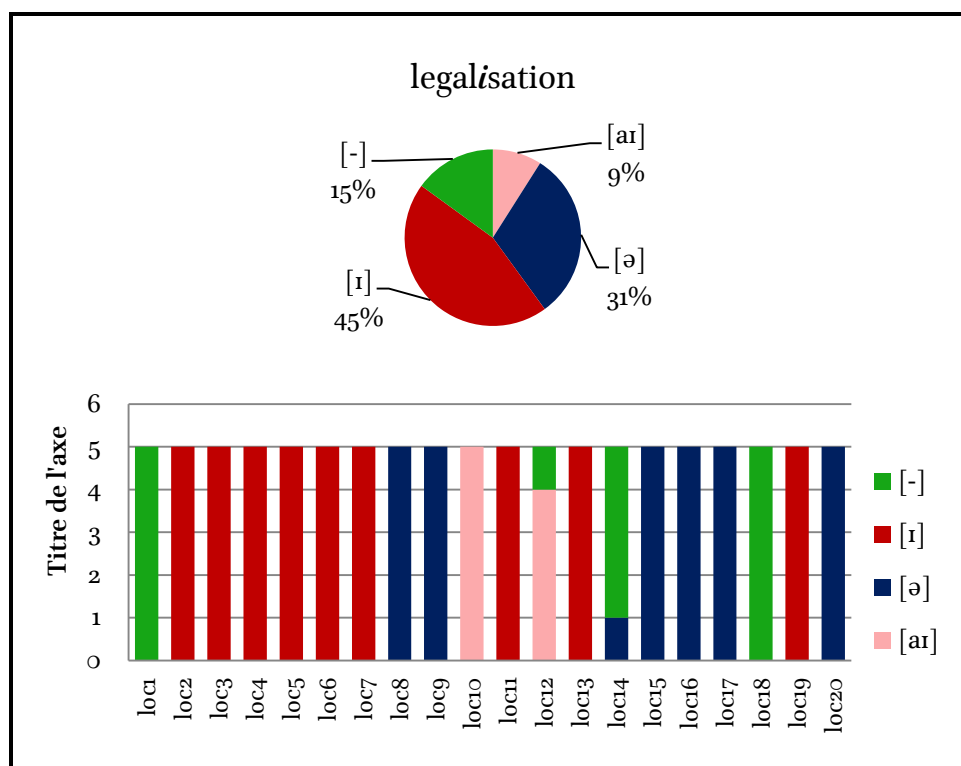


Figure 94 : pourcentages et répartition par locuteur de la production du premier <i> de *legalisation*.

8.7.3 Organisation et legalisation

D'une façon globale, le [ə] domine dans la prononciation de la troisième syllabe des mots *organisation* et *legalisation*. Il est produit à 43 % suivi du [ɪ] (28 %), du [-] (22 %) et du [aɪ] (7 %). En dehors des locuteurs 1 ([-]), 9 ([ə]), 19 ([ɪ]) et 20 ([ə]) qui sont constants dans leur production, tous les autres combinent deux sons (locuteurs 3, 4, 5, 7, 8, 10 à 18) et trois sons/série de sons (2 et 6).

Mis à part l'apparition de [-] en VR, le [ɪ] perd sa position (76 % en VN à 28 % en VR) au profit du [ə] (15 % en VN à 43 % en VR), le [aɪ] passe de 9 % en VN à 7 % en VR. Par contre, nous constatons des productions plus homogènes en VN qu'en VR. Seuls les locuteurs 16 (pour [ə] et [ɪ]) et 19 (pour [ɪ]) maintiennent les mêmes productions dans les deux vitesses (cp. Figure 55, p. 236).

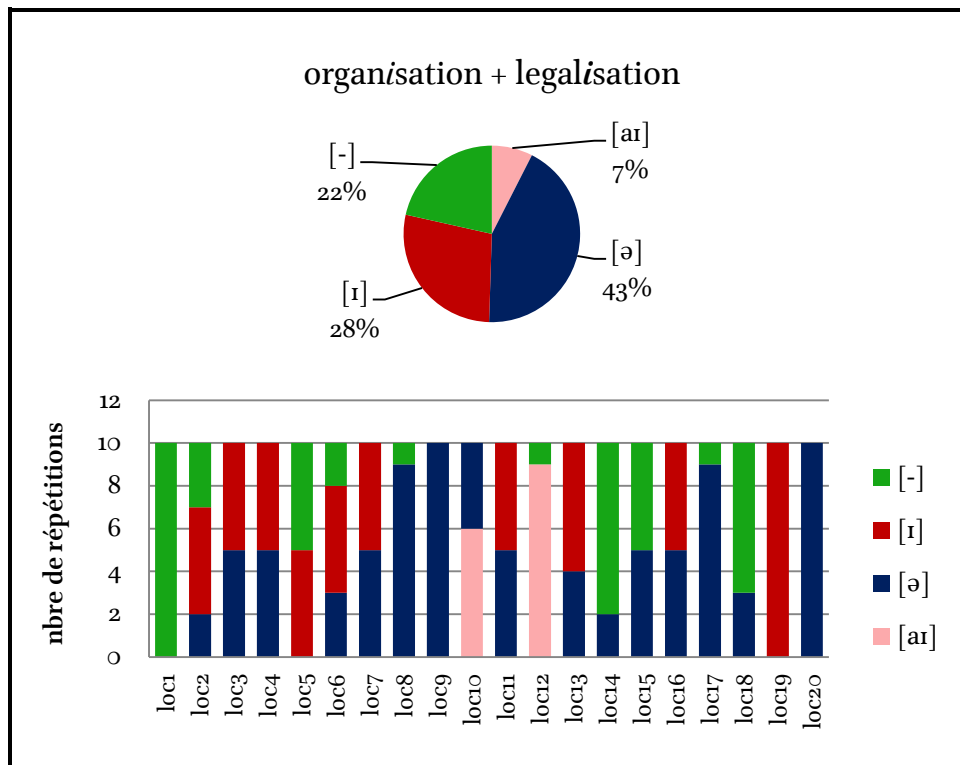


Figure 95 : pourcentages et répartition par locuteur de la prononciation de *organisation* et *legalisation*.

8.8 La prononciation de la graphie <ei> dans *leisure*

La suite de voyelles dans la première syllabe du mot *leisure* (<ei>) est produite en [eɪ] (en tête avec 34 %), en [i:] (25 %), en [ɛ] (23 %) et en [e] (18 %). Quant aux détails, nous notons que les locuteurs 2, 6, 7, 13, 15 et 19 ne produisent que des [eɪ], les locuteurs 1, 8, 9 et 16 que des [i:], les locuteurs 10, 12 et 17 que des [ɛ] et le locuteur 18 que des [e]. Les autres combinent deux des quatre sons : [e] et [ɛ] (locuteurs 3, 4 et 5), [e] et [eɪ] (locuteur 14) et finalement [e] et [i:] (locuteur 20).

Ainsi, nous constatons que [eɪ] domine dans les deux modes de production, mais avec une chute en VR (48 à 34 %). [ɛ] réduit aussi de 24 à 23 % tandis que [i:] et [e] grimpent légèrement en VR (23 à 25 %, 5 à 18 % respectivement). Pendant que les locuteurs 1, 8, 9 et 16 (pour le [i:]), les locuteurs 2, 6, 7, 13, 15 et 19 (pour le [eɪ]), les locuteurs 10, 12 et 17 (pour le [ɛ]) conservent les mêmes productions dans les deux modalités, les locuteurs 3, 4, 11 et 14 passent des productions homogènes en VN à des mélanges en VR. Les locuteurs 5 et 20 changent de mélanges entre les deux modalités. Notons que non seulement les locuteurs utilisent les mêmes sons/séries de sons dans les deux modalités, mais également, les réalisations sont beaucoup plus constantes en VN qu'en VR : seuls deux locuteurs mélangent les productions en VN contre six en VR (cp. Figure 56, p. 238).

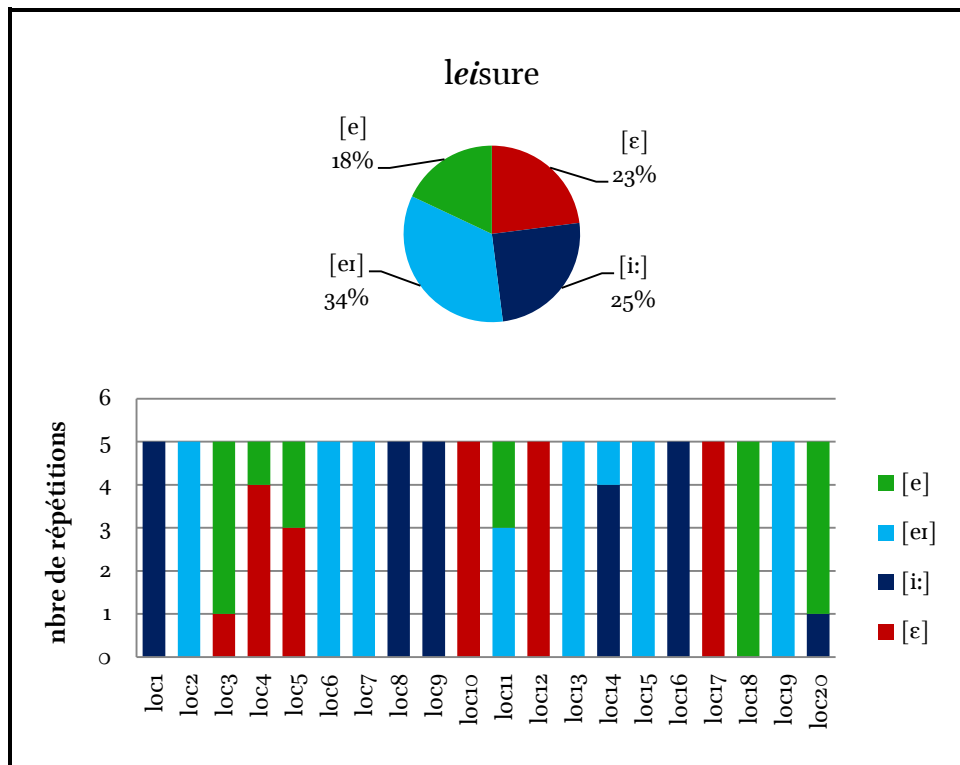


Figure 96 : pourcentages et répartition par locuteur de la production de <ei> de *leisure*.

8.9 La prononciation de la graphie <e> dans evolution

Le mot *evolution* se prononce de deux façons dans les deux variétés de référence (RP et GA). La première syllabe représentée par la lettre <e> se prononce /i:/ (,i:və'ljʊ:ʃən) ou /ɛ/ (,ɛvə'ljʊ:ʃən)²²⁷. Cependant, le LPD indique que la prononciation avec un /ɛ/ est généralement utilisée en GA (à 85 %). Seuls 15 % des Américains prononcent le mot avec un /i:/ (LPD entrée *evolution*). En revanche, dans nos analyses acoustiques, une troisième réalisation apparaît dans les résultats, il s'agit de la réalisation de la voyelle <e> en [e], comme dans le mot français « dé ».

En ce qui concerne la production de la graphie <e> dans le mot *evolution*, [e] vient en tête avec 75 %, suivi de [ɛ] (15 %) et de [i:] (10 %). Quant aux productions individuelles, il est intéressant de noter qu'il n'y a aucun mélange. Tous les locuteurs restent constants : les locuteurs 10 et 15 ne produisent que des [i:], les locuteurs 9, 16 et 17 que des [ɛ] et les autres ne réalisent que des [e].

Dans les deux systèmes de productions, le [e] domine avec une augmentation en VR (61 à 75 %), suivi du [ɛ] (de 10 % en VN à 15 % en VR). Le [i:] diminue de 29 % en VN à 10 % en VR (cp. Figure 57, p. 239). Ici également, les locuteurs adoptent les mêmes sons à savoir [e], [ɛ] et [i:]. De

²²⁷ Il est à noter que le /ju:/ devient un /u:/ en GA.

plus, les locuteurs 2, 3, 5, 6, 8, 12, 13, 15 et 18 gardent des productions identiques en VN et en VR. Ensuite, les locuteurs 1, 4, 7, 14 et 20 basculent en [e] en VR. Et enfin, les locuteurs 9 et 16 changent de [i:] en VN à [ɛ] en VR.

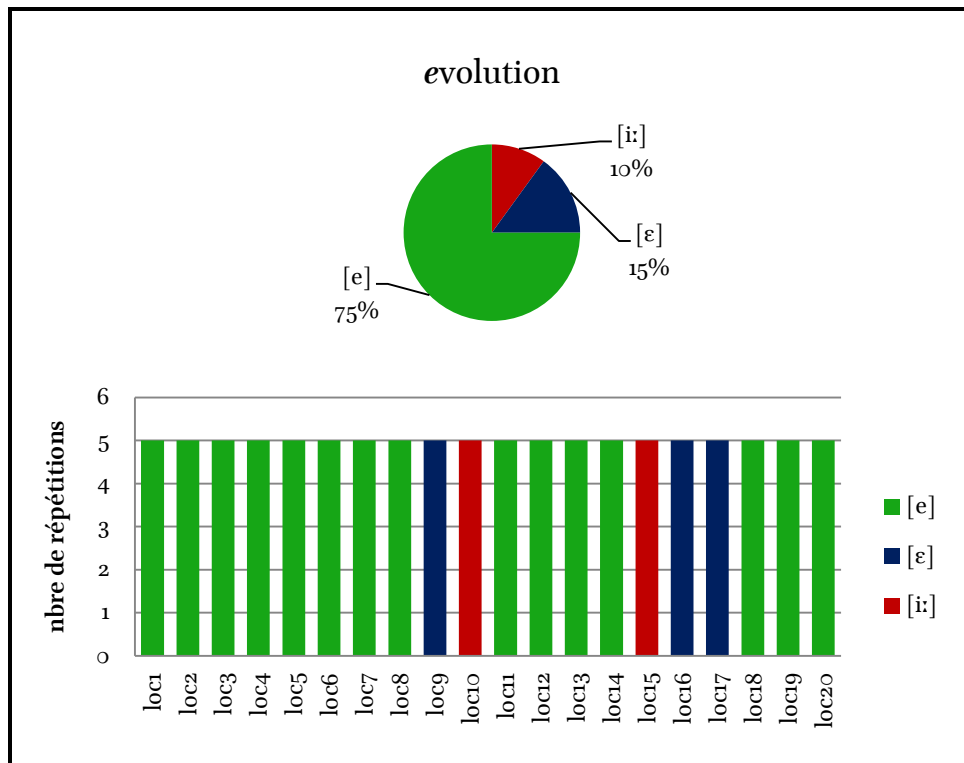


Figure 97 : pourcentages et répartition par locuteur de la production de <e> de *evolution*.

8.10 Nasalisation vocalique

Dans la modalité VR, les analyses acoustiques des mots *band* et *hand* n'ont fait émerger que trois sortes des productions pour la série de graphies <an>. D'une part, il y a deux réalisations nasalisées [ã] et [æ̃] comme dans les mots français paon et pain, et d'autre part, une production non nasalisées /æn/. Dans la présentation de ces derniers, nous ne prenons en compte véritablement que les voyelles /a/ et /æ/.

8.10.1 Band

La prononciation du mot *band* avec la voyelle nasalisée [æ̃] vient en tête avec 74 % des occurrences. Elle est suivie par le son [ã] (16 %) et le son [æ] (10 %). Pendant que les locuteurs 2, 4, 5, 6, 7, 9, 10, 11, 12, 15, 17 à 20 (pour le [æ̃]), 3 et 13 (pour le [æ]) 14 et 16 (pour le [ã]) adoptent des productions homogènes, les locuteurs 18 combinent des [æ̃] et des [ã].

Par conséquent, la comparaison des deux conditions de production montre que le [ǣ] domine toujours et connaît une forte augmentation en VR (de 47 % en VN à 74 % en VR). Le son [ã] augmente aussi mais légèrement, en passant de 10 en VN à 16 % en VR, alors que le [æ] diminue de la VN à la VR (38 à 10 %). Enfin, nous remarquons l'absence de la production du son [a] en VR (pourtant présent en VN) (cp. Figure 58, p. 241). En effet, d'un côté, les locuteurs 2, 4, 5, 9, 10, 11, 15 et 20 (pour le [ǣ]) et le locuteur 13 (pour le [æ]) sont constants dans leurs productions concernant les deux modalités. De l'autre côté, nous assistons à un changement de production entre les deux modalités : les locuteurs 7, 12, 17, 18 et 19 passent de [æ] en VN à [ǣ] en VR ; le locuteur 14 passe de [æ] en VN à [ã] en VR ; le locuteur 3 passe de [ǣ] en VN à [æ] en VR ; le locuteur 16 passe de [a] en VN à [ã] en VR ; les locuteurs 1 et 8 passent de [ã] en VN à un mélange de [ã] et de [ǣ] en VR.

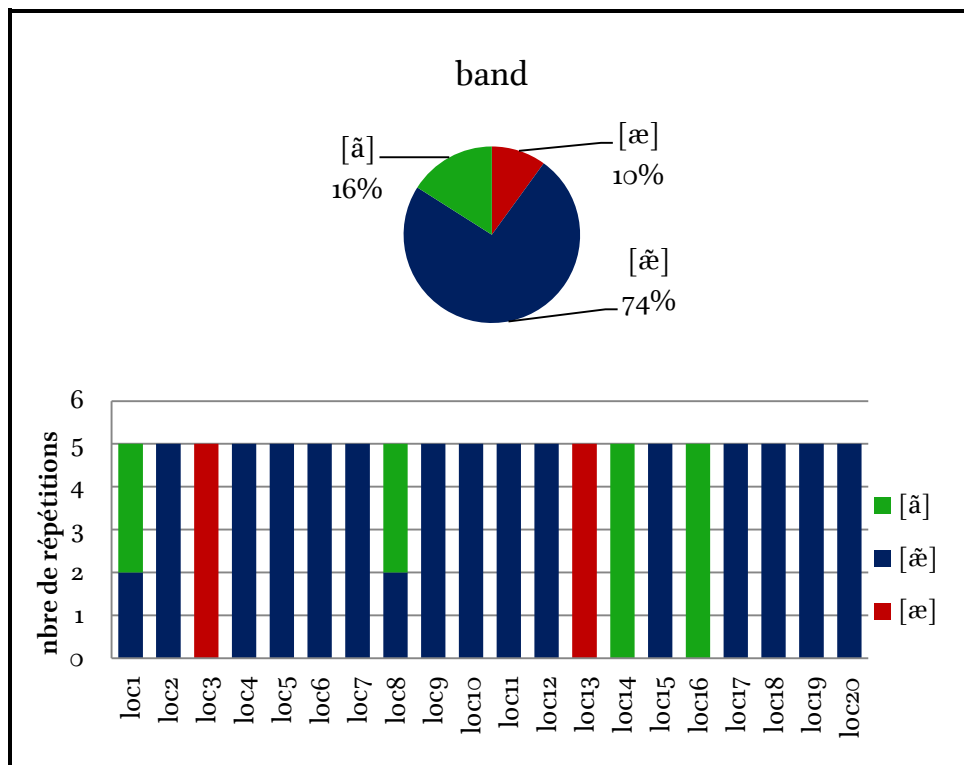


Figure 98 : pourcentages et répartition par locuteur de la production de <an> de *band*.

8.10.2 Hand

La production de la voyelle nasalisée [ǣ] est plus importante (70 %) que les autres ([ã] 25 %, [æ] 5 %). Tous les locuteurs sont constants dans la prononciation du mot avec [ǣ] (locuteurs 2, 4 à 7, 9 à 12, 15, 17 à 20), avec [ã] (locuteurs 1, 3, 8, 14 et 16) et avec [æ] (locuteur 13).

Les sons [ǣ] et [ã] augmentent de 50 % et 10 % en VN à 70 % et 25 % en VR respectivement. En revanche le son [æ] chute de 35 % en VN à 5 % en VR. Signalons l'absence du son [a] en VR mais

qui est pourtant produit en VN (cp. Figure 59, p. 242). Pour les productions individuelles, les locuteurs 1 et 8 (pour [ã]), 2, 4, 5, 9, 10, 11, 15 et 20 (pour [æ̃]) et 13 (pour [æ]) conservent les mêmes productions dans les deux conditions. Les locuteurs 7, 12, 17, 18 et 19 basculent de la production de [æ] en VN à celle de [æ̃] en VR, de même que les locuteurs 3 et 16 qui passent de [æ̃] et [a] en VN à [ã] en VR. Enfin les locuteurs 6 et 14 mélangent le [æ̃] et le [æ] en VN mais produisent entièrement des [æ̃] et des [ã] respectivement en VR.

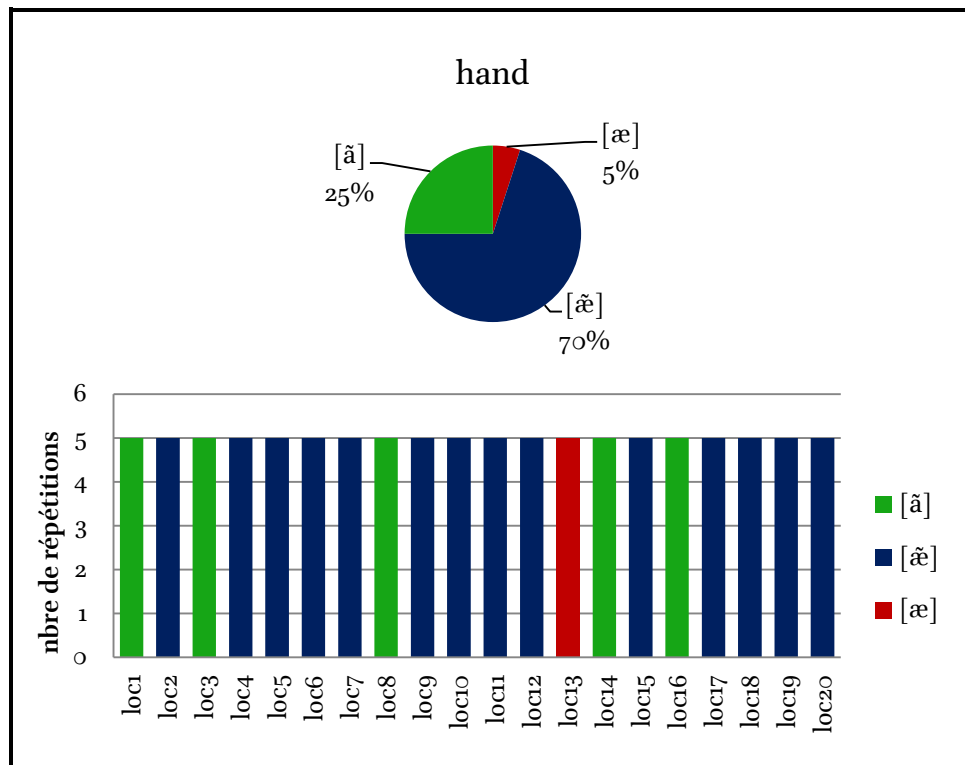


Figure 99 : pourcentages et répartition par locuteur de la production de <an> de *hand*.

8.10.3 Band et hand

Au final, la voyelle nasalisée [æ̃] domine la prononciation des deux mots *hand* et *band* à 72 %, suivi du [ã] (21 %) et du [æ] (7 %). Le son [a] n'est pas produit. Les locuteurs 2, 4, 5, 6, 7, 9, 10, 11, 12, 15, 17 et 20 ne produisent que des [æ̃] pendant que les locuteurs 14 et 16 ne réalisent que des [ã] et le locuteur 13 produit des [æ]. Les locuteurs 1 et 8 réalisent des mélanges de [ã] et [æ̃] de même que le locuteur 3 avec le [ã] et le [æ].

Les résultats de la comparaison des deux modes de production montrent une prédominance du son [æ̃] (49 % en VN et 72 % en VR). Le [ã] augmente aussi en VR (10 en VN à 21 % en VR) alors que le son [æ] diminue en VR (36 en VN à 7 % en VR) (cp. Figure 60, p. 243). Les résultats individuels indiquent que les locuteurs 2, 4, 5, 9, 10, 11 et 15 conservent des productions identiques

[æ] dans les deux modalités, ainsi que le locuteur 13 avec le [æ]. Ensuite, les locuteurs 7, 12, 17, 18 et 19 produisent uniquement des [æ] en VN et des [æ̃] en VR, de même que le locuteur 16 qui ne produit que des [a] en VN et des [ã] en VR. Les locuteurs 1, 3 et 8 sont homogènes en VN (pour [ã] et [æ̃]) mais produisent des mélanges en VR ([ã] et [æ̃] ; [ã] et [æ] ; [ã] et [æ̃] respectivement). Nous notons la situation inverse chez les locuteurs 6 et 16 qui produisent des mélanges en VN ([æ] et [æ̃]) alors qu'ils sont homogènes en VR ([æ̃], [ã] respectivement).

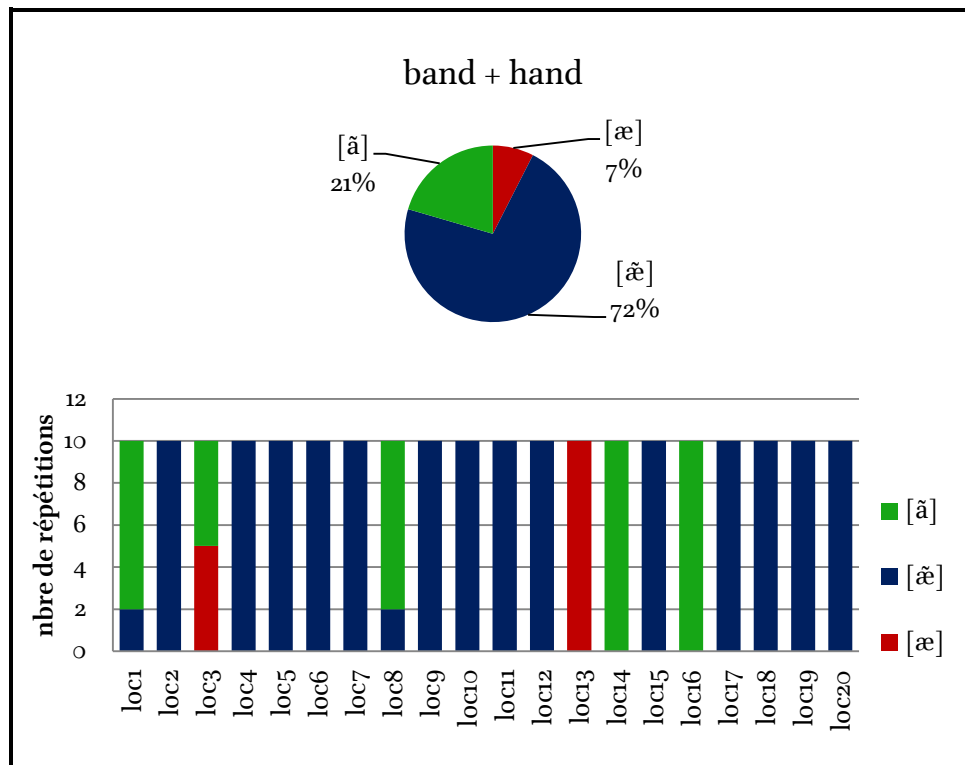


Figure 100 : pourcentages et répartition par locuteur de la production de <an> de *band et hand*.

Cette section sur la présentation des résultats des éléments segmentaux vocaliques montre une fois de plus une variabilité pour tous les locuteurs. Certains sons ou séries de sons se rapprochent plus de la RP alors que d'autres ressemblent plus aux caractéristiques du GA. Les productions « Autres » restent toutefois non négligeables. Le tableau suivant résume la portion de chaque élément en fonction des mots dans lesquels ils se trouvent. Contrairement aux résultats de la VN, où les caractéristiques de la RP dominent largement les productions (55 %), les éléments proches du GA viennent en tête en VR (41 %) et ceux de la RP ne sont produits qu'à 27 %. Il est fort intéressant de pointer que les productions « Autres » ont doublé en passant de 16 % en VN à 32 % en VR (cp. Tableau 21, p. 244). Nous présentons dans les sections suivantes les résultats des deux

éléments suprasegmentaux, à savoir la place de l'accent principal et le nombre de syllabes accentuées, en prenant en compte la voyelle de la pénultième syllabe.

Tableau 24 : résumé des éléments segmentaux vocaliques en VR.

Phénomène étudié	Mots	% de caractéristiques de la RP	% de caractéristiques du GA	% de caractéristiques « Autre »
La voyelle de LOT	<i>Hot</i>	80	20	0
	<i>Top</i>	89	11	0
	total	85	15	0
La voyelle de GOAT	<i>Joke</i>	1	52	47
	<i>Road</i>	0	46	54
	total	0	49	51
La graphie <y> et <i>	<i>Dynasty</i>	15	22	63
	<i>Vitamin</i>	36	38	26
	total	25	50	25
La graphie <i>	<i>Organisation</i>	11	55	34
	<i>Legalisation</i>	45	31	24
	total	28	43	29
La graphie <ei>	<i>Leisure</i>	23	59	18
La graphie <e>	<i>Evolution</i>	10	15	75
Nasalisation vocalique	<i>band</i>	10	74	16
	<i>hand</i>	5	70	25
	total	7	72	21
Total des éléments vocaliques (sur 1200 itérations)		27	41	32

8.11 La place de l'accent principal

Deux types de prononciation sont en principe attestés dans la littérature concernant les mots *donate* et *locate* (cf. section 6.6.3.1 p. 194). L'accentuation de la première syllabe en GA (/ˈloʊkɛt/ et /ˈdoʊneɪt/) et de la deuxième en RP (/ləʊˈkɛɪt/ et /dəʊˈneɪt/). Mais, nous verrons ci-dessous l'émergence d'une troisième prononciation (plate) qui consiste en la non-accentuation d'aucune des deux syllabes, ou du moins, à un même niveau d'accentuation de ces dernières (voir les indications acoustiques en Annexe A 1, p. 1-35).

8.11.1 Donate

Concernant la prononciation du mot *donate*, la première syllabe (syll1) est accentuée avec une dominance de 54 %, suivie de la deuxième syllabe (syll2) (44 %) et de l'accentuation plate (2 %). D'un point de vue individuel, sont constants dans leurs productions les locuteurs 3, 6, 9, 10, 14, 16, 17

et 18 pour la syll1 et les locuteurs 1, 4, 7, 8, 13, 19 et 20 pour la syll2. Enfin, des mélanges sont réalisés par les locuteurs 2, 11, 12 et 15 (syll1 et syll2) et le locuteur 5 (syll1, syll2 et plate).

En ce qui concerne la comparaison des deux modalités, nous remarquons que l'accentuation de la syll1 domine dans tous les cas : 45 % en VN et 54 % en VR (cp. Figure 61, p. 246). La syll2 augmente aussi légèrement de 38 % en VN à 44 % en VR, alors que l'accentuation plate chute de 17 % en VN à 2 % en VR. Quant aux mêmes productions dans les deux situations, nous notons que les locuteurs 3, 9, 10 et 18 conservent l'accentuation de la syll1, les locuteurs 7, 8, 13, 19 et 20 accentuent la syll2 et les locuteurs 2 et 5 gardent l'accentuation mixte. Les locuteurs 11, 12 et 15 ont des productions homogènes en VN mais finissent avec des mélanges en VR. Inversement, les réalisations des locuteurs 1, 6, 14 et 17 sont des mélanges en VN mais deviennent des productions homogènes en VR.

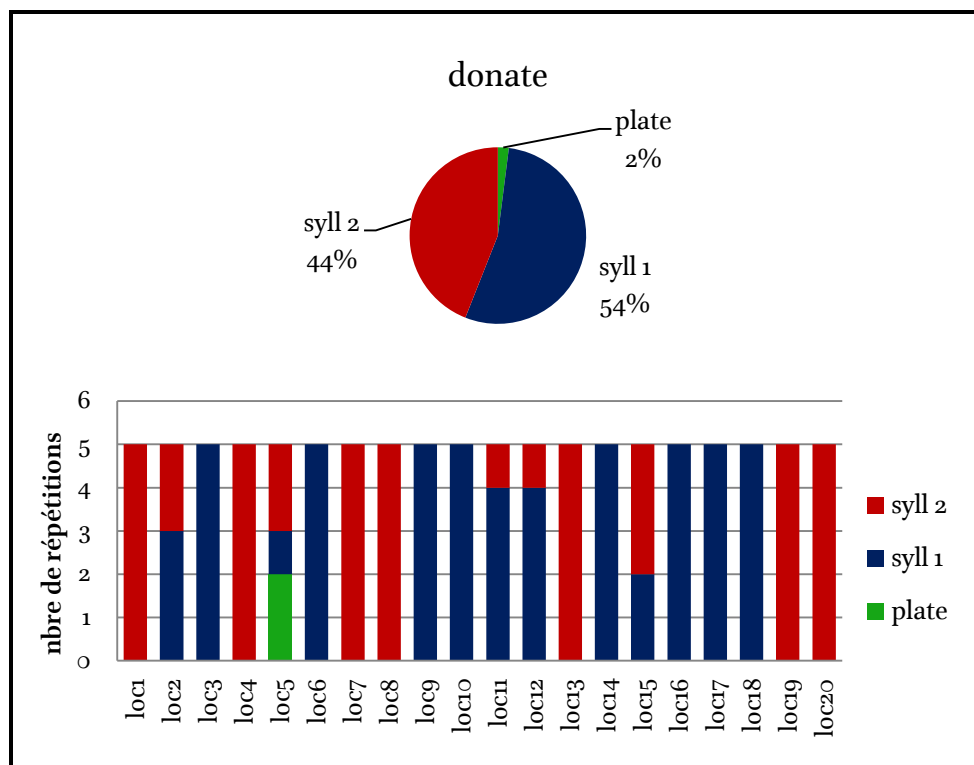


Figure 101 : pourcentages et répartition par locuteur de l'accentuation de *donate*.

8.1.1.2 Locate

Nous constatons que dans la prononciation du mot *locate*, la syll1 est accentuée avec une prédominance de 57 %, suivie de la syll2 (36 %) et l'accentuation plate (7 %). Par conséquent, nous notons que les locuteurs 2, 3, 9, 10, 11, 14, 16, 17 et 18 sont constants dans l'accentuation de la syll1 et

les locuteurs 1, 4, 7, 8, 13, 15 et 19 sont également réguliers dans l'accentuation de la syll2. Seuls les locuteurs 5, 6, 12 et 20 réalisent des mélanges de la syll1 et syll2, de la syll2 et plate.

Parallèlement aux résultats du mot *donate*, l'accentuation de la syll1 domine également la prononciation du mot *locate* (51 % en VN, 57 % en VR) (cp. Figure 62, p. 247). Celle de la syll2 passe de 34 % en VN à 36 % VR tandis que l'accentuation plate va de 15 % en VN à 7 % en VR. Tout d'abord, nous constatons que les locuteurs 3, 9, 10, 11, 14, 16 et 18 pour la syll1 et les locuteurs 7, 8, 13 et 19 pour la syll2, conservent des productions identiques dans les deux modalités. Puis, nous notons que le locuteur 12 a des productions régulières en VN mais mélange plutôt en VR et les locuteurs 1, 2, 4, 15 et 17 réalisent des mélanges en VN mais leurs productions sont homogènes en VR. Enfin, les locuteurs 5, 6 et 20 réalisent des mélanges de différents sons dans les deux modalités.

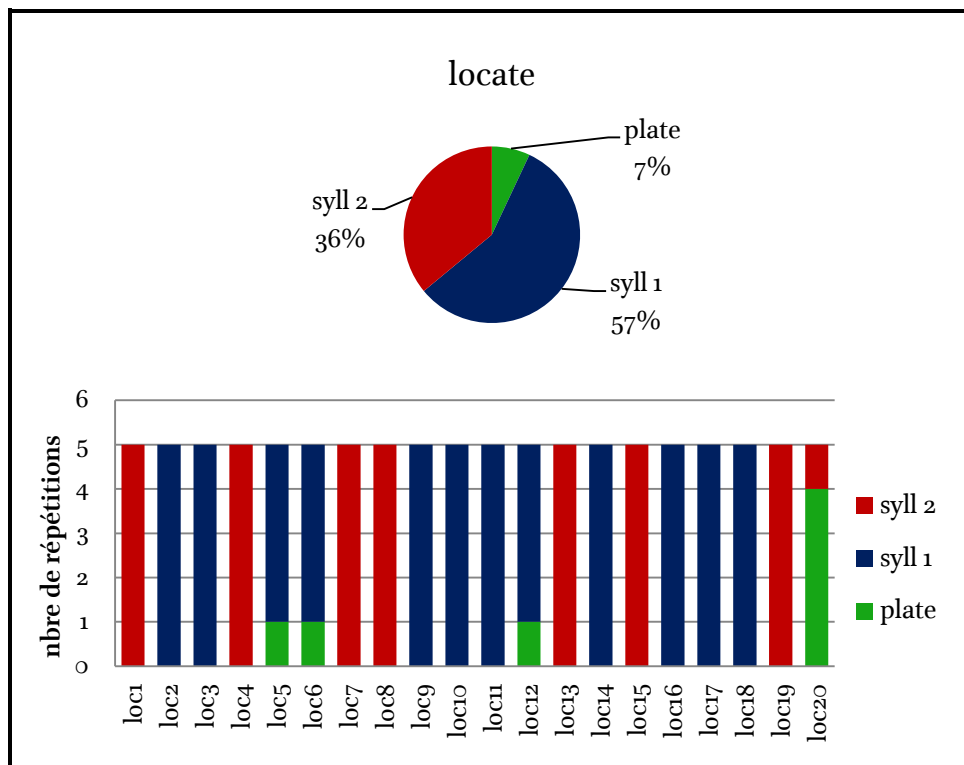


Figure 102 : pourcentages et répartition par locuteur de l'accentuation de *locate*.

8.11.3 Donate et locate

La conséquence évidente de l'augmentation de l'accentuation de la syll1, dans la prononciation des deux mots *donate* et *locate*, se traduit par sa dominance sur les autres productions (56 %). S'en suivent l'accentuation de la syll2 (40 %) et de l'accentuation plate (4 occurrences). Tandis que les locuteurs 3, 8, 10, 14, 16, 17 et 18 n'accentuent que la syll1 et les locuteurs 1, 4, 7, 8, 13 et 19 que sur la syll2, les autres (2, 5, 6, 11, 12, 15 et 20) procèdent à une combinaison des trois productions.

De manière générale, la syll1 garde sa position de dominance dans les deux situations de production (48 % en VN, 56 % en VR) suivie de la syll2 (36 % en VN, 40 % en VR) et l'accentuation plate (16 % en VN, 4 % en VR). En effet, les locuteurs 3, 9, 10 et 18 pour la syll1 et les locuteurs 7, 8, 13 et 19 pour la syll2, conservent des réalisations identiques dans les deux modalités (cp. Figure 63, p. 248). Pendant que les locuteurs 11 et 12 sont constants en VN et non en VR, les locuteurs 2, 5, 6, 15 et 20 effectuent le contraire, à savoir qu'ils réalisent des mélanges en VN mais deviennent constants en VR. Des mélanges des diverses productions sont réalisés par les locuteurs 1, 4, 14, 16 et 17. Signalons enfin qu'aucune accentuation homogène plate n'est constatée dans les deux situations.

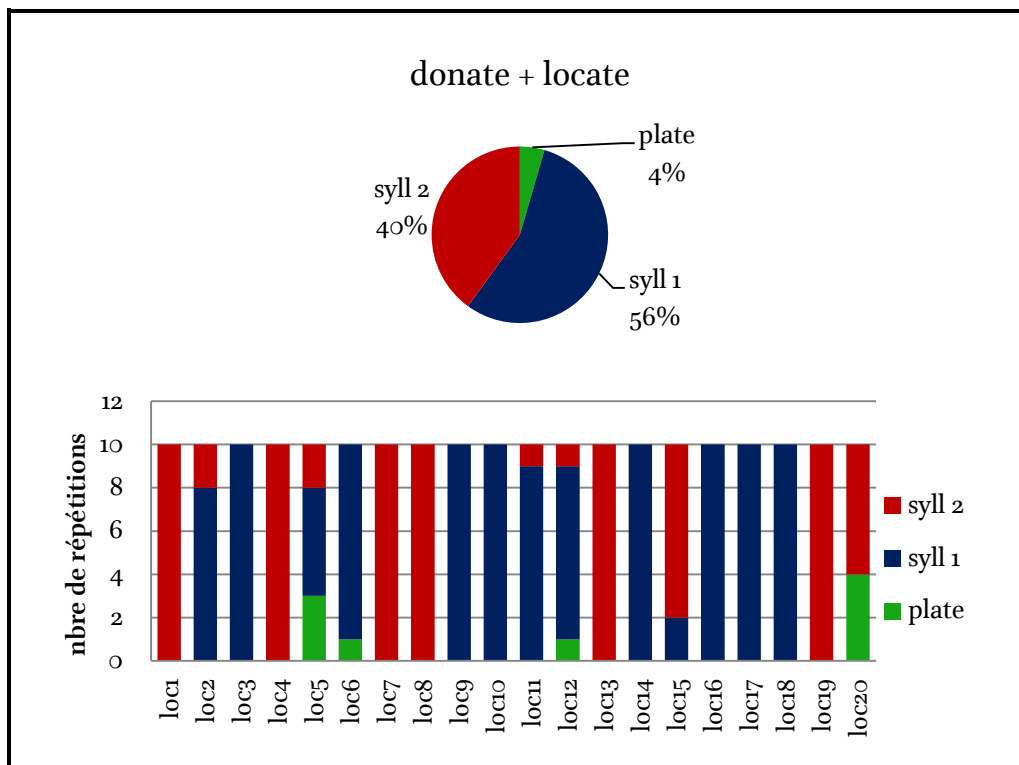


Figure 103 : pourcentages et répartition par locuteur de de l'accentuation de *donate* et *locate*.

8.12 La voyelle de la pénultième syllabe et le nombre de syllabes accentuées

Les mots *military* et *necessary* sont prononcés soit en trois (en RP) ou en quatre syllabes (en RP et en GA). Dans le premier cas, c'est uniquement la première syllabe qui porte l'accent (/ˈmɪlɪtri/, /ˈnesəri/). Dans le deuxième cas, la première syllabe garde l'accent primaire et la troisième syllabe reçoit un accent secondaire. Celle-ci est alors produite avec une voyelle centrale /ə/ en RP (/ˈmɪlɪtəri/ ou /ˈmɪlɪ˞ri/, /ˈnesəsəri/) et avec une voyelle ouverte /ɛ/ en GA (/ˈmɪlɪtəri/, /ˈnesəseri/).

Nous n'analysons pas l'accentuation de la première syllabe de ces mots, car elle ne présente aucune variation selon les deux accents.

8.12.1 Military

Le mot *military* est prononcé de quatre manières différentes, parmi lesquelles nous retrouvons les trois réalisations connues. Il s'agit des sons [ɛ] pour le GA et du [ə] et du [-] pour le RP. Nous rappelons que l'absence de voyelle est représentée par le symbole [-] afin de rester cohérent avec les autres notations. Ce symbole indique également que le mot est prononcé en trois syllabes et pas quatre. La quatrième production consiste en un [ɑ].

La voyelle [ə] domine la prononciation de la pénultième syllabe du mot *military* (40 %), suivie du [ɛ] (35 %), du [-] (15 %) et du [ɑ] (10 %). Remarquons que tous les locuteurs sont constants dans leurs réalisations : les locuteurs 1, 3, 5, 7, 8, 11, 12 et 17 ne produisent que des [ə], les locuteurs 2, 4, 6, 9, 18, 19 et 20 que des [ɛ], les locuteurs 10, 15 et 16 que des [-] et enfin les locuteurs 13 et 14 ne réalisent que des [ɑ].

Ainsi, nous observons que malgré une légère diminution en VR, le [ə] domine toujours les productions (55 % en VN 40 % en VR). Le [ɛ] augmente de 25 % en VN à 35 % en VR, le [ɑ] augmente aussi de 5 % en VN 10 % en VR, alors que le [-] reste stable dans les deux conditions (15 %). Pour ce qui est des productions individuelles, nous constatons que les locuteurs 4, 6, 18 et 19 mélangent des [ɛ] et des [ə] en VN mais ne produisent que des [ɛ] en VR. Le locuteur 14 change également de production en passant des [ə] en VN à des [ɑ] en VR. Hormis ces changements, il est intéressant de noter que tous les autres locuteurs conservent les mêmes productions dans les deux vitesses (cp. Figure 64, p. 250).

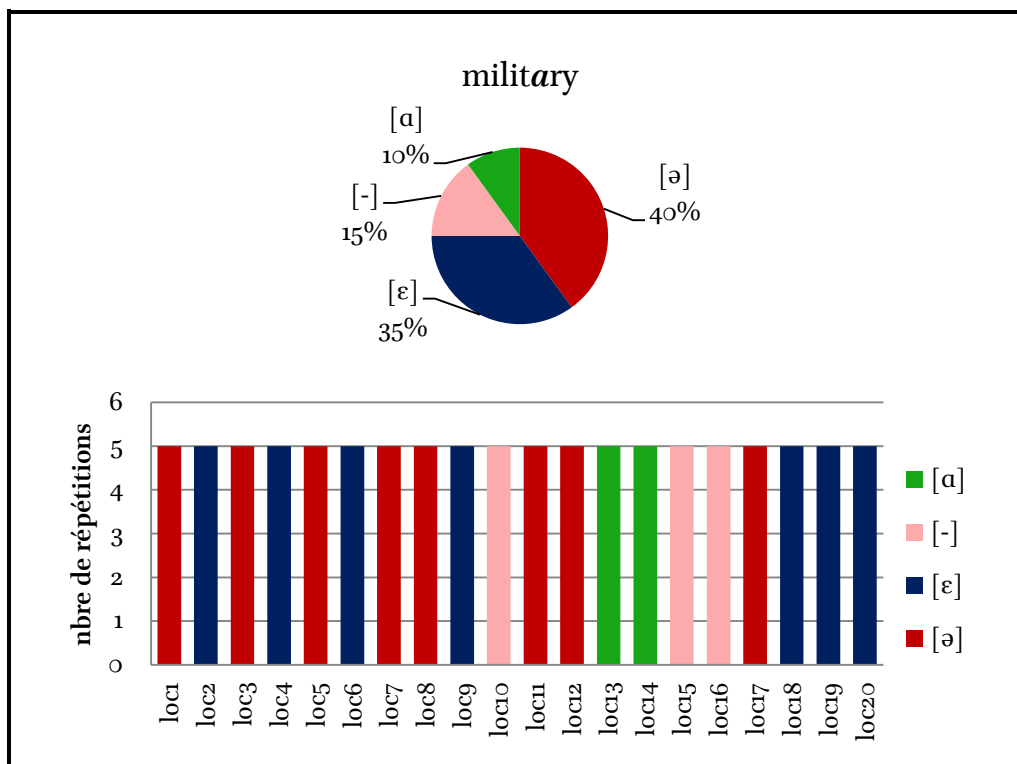


Figure 104 : pourcentages et répartition par locuteur de la prononciation de <ary> de *military*.

8.12.2 Necessary

Contrairement à la prononciation du mot *military*, le mot *necessary* n'a connu que deux prononciations, à savoir le son [ə] produit avec une dominance de 60 % et le son [ε] à 40 %. Remarquons l'absence des sons [-] et [a] dans la prononciation de ce mot. Quatre locuteurs (3, 12, 14 et 18) réalisent des mélanges de [ə] et de [ε] pendant que les autres locuteurs ne produisent que des sons homogènes : locuteurs 2, 4, 6, 9, 10, 15, 16, 17, 19 et 20 pour le son [ε] et les locuteurs 1, 5, 7, 8, 11, et 13 pour le son [ə].

Outre les quatre locuteurs précédemment évoqués (3, 12, 14 et 18) et qui combinent les deux sons en VR, alors qu'ils ont des productions homogènes en VN ([ε], [ə], [ə], [ə] respectivement), le locuteur 17 bascule de [ə] en VN à [ε] en VR tandis que le locuteur 5 change ses productions de [ε] en VN à [ə] en VR. Pour finir, les locuteurs 2, 4, 6, 9, 10, 15, 16, 19 et 20 pour le son [ε] et les locuteurs 1, 7, 8, 11 et 13 pour le son [ə], conservent des réalisations identiques dans les deux modalités (cp. Figure 65, p. 251).

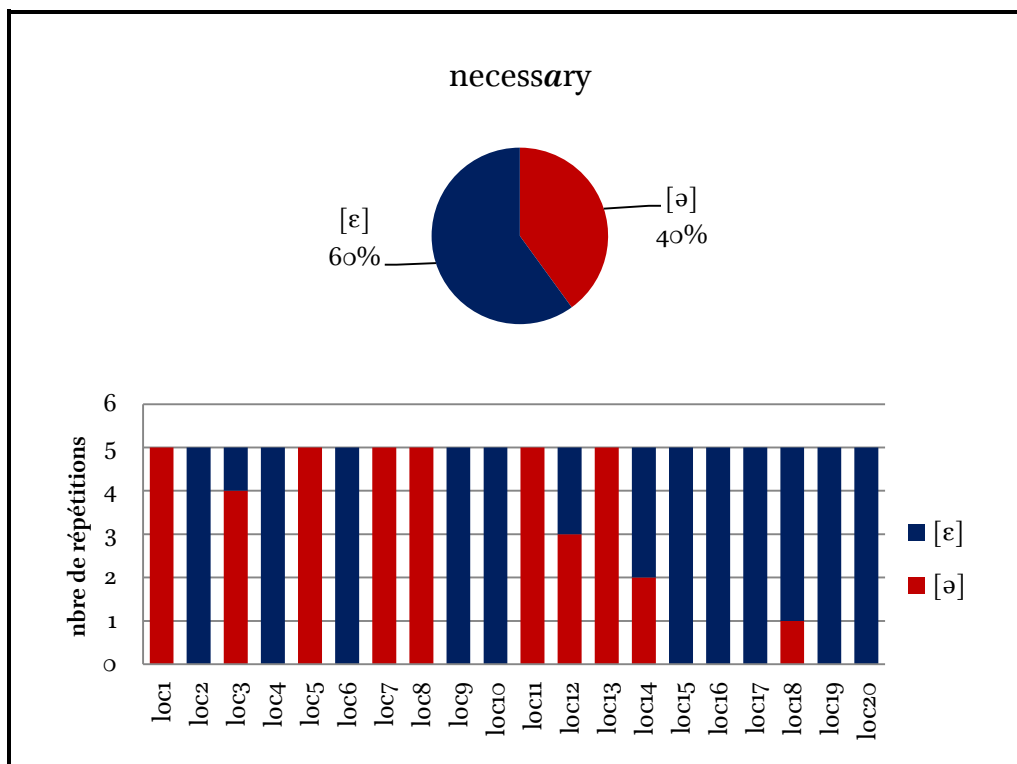


Figure 105 : pourcentages et répartition par locuteur de la prononciation de <ary> de necessary à VR.

8.12.3 Military et necessary

Les résultats de la prononciation des deux mots *military* et *necessary* indiquent que le son [ɛ] est majoritairement produit à 47 %, suivi du [ə] (40 %), du [-] (8 %) et du [ɑ] (5 %), sur les 200 occurrences. Nous précisons qu'il y a plus de variations dans la prononciation du mot *military* que *necessary*, en raison de [-] et [ɑ]. Les résultats par locuteur montrent que les locuteurs 1, 5, 7, 8 et 11 pour le [ə] et les locuteurs 2, 4, 6, 9, 19 et 20 pour le [ɛ], sont constants dans leurs productions et que les autres locuteurs effectuent des combinaisons des différents sons.

En comparant les deux modalités, nous remarquons une inversion de situation dans la production globale : le son [ə] domine en VN (50 % contre 40 % en VR) alors qu'en VR c'est le [ɛ] qui l'emporte (47 % contre 40 % en VN). [-] passe quant à lui de 7 % en VN à 8 % en VR, [ɑ] grimpe de 3 % en VN à 5 % en VR. Concernant les productions individuelles, elles montrent que les locuteurs 1, 7, 8 et 11 pour le [ə], les locuteurs 2, 9 et 20 pour le [ɛ], les locuteurs 10, 15 et 16 pour le mélange [-] et [ɛ], et le locuteur 13 pour le mélange [ɑ] et [ə], conservent les mêmes productions en VN qu'en VR. Pendant que les locuteurs 4, 5, 6 et 19 basculent de mélanges aux productions homogènes, les locuteurs 12, 14 et 17 font le contraire (de productions homogènes aux mélanges).

Enfin, les locuteurs 3 et 18 mélangent divers sons dans les deux conditions de parole (cp. Figure 66, p. 252).

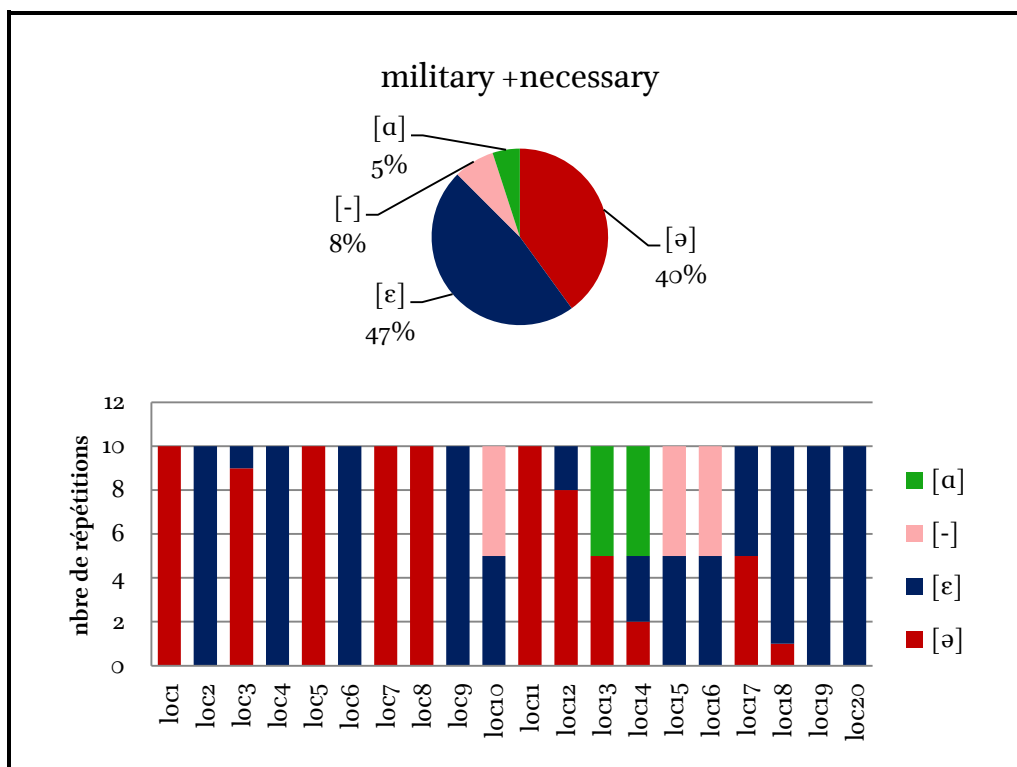


Figure 106 : pourcentages et répartition par locuteur de la prononciation de <ary> de military et necessary.

Il n'y a pas de schéma prédéfini quant à la dominance des divers éléments ; tout est variable, mais dans l'ensemble, tout comme en VN, les particularités proches du GA dominent la production des éléments suprasegmentaux (52 % en VN, 44 % en VR), suivis par celles de la RP (43 % en VN, 44 % en VR) et celle des autres productions (13 % en VN, 5 % en VR). On note donc une augmentation des productions proches du GA, une stagnation de celles de la RP et une diminution des productions « Autres » en VR (cp. Tableau 22, p. 253). Pour finir, nous présentons dans cette dernière section, les résultats globaux depuis le début de ce chapitre, à savoir tous les mots et éléments étudiés et leur pourcentage pour chaque locuteur.

Tableau 25 : résumé des éléments suprasegmentaux.

Phénomène étudié	Mots	% de caractéristiques de la RP	% de caractéristiques du GA	% de caractéristiques « Autres »
<i>Place de l'accent</i>	<i>donate</i>	44	54	2
	<i>locate</i>	36	57	7
	<i>total</i>	40	56	4
<i>Nombre de syllabes prononcés</i>	<i>military</i>	55	35	10
	<i>necessary</i>	40	60	0

<i>et qualité de la voyelle</i>	<i>total</i>	40	55	5
Total des suprasegmentaux (sur 400 occurrences)	éléments (sur 400)	44	52	5

8.13 Résultats globaux des mots à vitesse rapide

Nous présentons les résultats globaux ainsi qu'individuels de tous les mots étudiés concernant la modalité VR. Nous rappelons que comme pour la VN le nombre total de répétitions s'élève à 2900, soient 145 itérations par locuteur. Ainsi, les caractéristiques proches du GA viennent en tête avec 46 %, suivies des caractéristiques proches de la RP (36 %) et des caractéristiques « Autres » (18 %).

Les productions par locuteurs, sur les 145 occurrences donnent des résultats assez variés. Ainsi, le locuteur 1 produit 47 % de caractéristiques de la RP, 29 % de caractéristiques du GA et 24 % de caractéristiques « Autres ». Le locuteur 2 produit 26 % de caractéristiques de la RP, 48 % de caractéristiques du GA et 26 % de caractéristiques « Autres ». Le locuteur 3 réalise 36 % de spécificités de la RP, 41 % de spécificités du GA et 23 % de spécificités « Autre ». Quant au locuteur 4, il réalise 30 % de particularités de la RP, 46 % de particularités du GA et 24 % de caractéristiques « Autres ». Les caractéristiques de la RP sont produites à 42 %, les caractéristiques du GA à 37 % et les caractéristiques « Autres » à 21 % par le locuteur 5. Alors que le locuteur 6 réalise 37 % des caractéristiques de la RP, 43 % des caractéristiques du GA et 20 % des caractéristiques « Autres », le locuteur 7, lui, produit 55 % des particularités de la RP, 28 % des particularités du GA et 17 % des particularités « Autres ». En ce qui concerne les locuteurs 8, 9 et 10, ils produisent respectivement 41 %, 7 % et 55 % des propriétés de la RP ; 41 %, 92 % et 42 % des propriétés du GA et 18 %, 1 % et 3 % des propriétés « Autres ». Les caractéristiques « Autres » sont produites respectivement à 30 %, 14 % et 25 %, celles du GA à 35 %, 41 % et 34 % et celle de la RP à 35 %, 45 % et 41 % par les locuteurs 11, 12 et 13. Le locuteur 14 réalise 29 % de caractéristiques de la RP, 36 % des caractéristiques du GA et 35 % de caractéristiques « Autres », alors que le locuteur 15 produit 57 % de spécificités de la RP, 39 % de spécificités du GA et 5 % de spécificités « Autres ». Concernant les locuteurs 16 et 17, ils réalisent chacun 61 % de propriétés du GA mais respectivement 26 % et 32 % de propriétés de la RP, 14 % et 7 % de propriétés « Autres ». Les locuteurs 18 et 19 réalisent respectivement 26 % et 14 % de propriétés « Autres », 26 % et 15 % de celles du GA et 49 % et 71 %

de celles de la RP. Enfin, le locuteur 20 produit 50 % de caractéristiques du GA, 29 % de propriétés de la RP et 21 % de spécificités « Autres » (cp. Figure 67, p. 255).

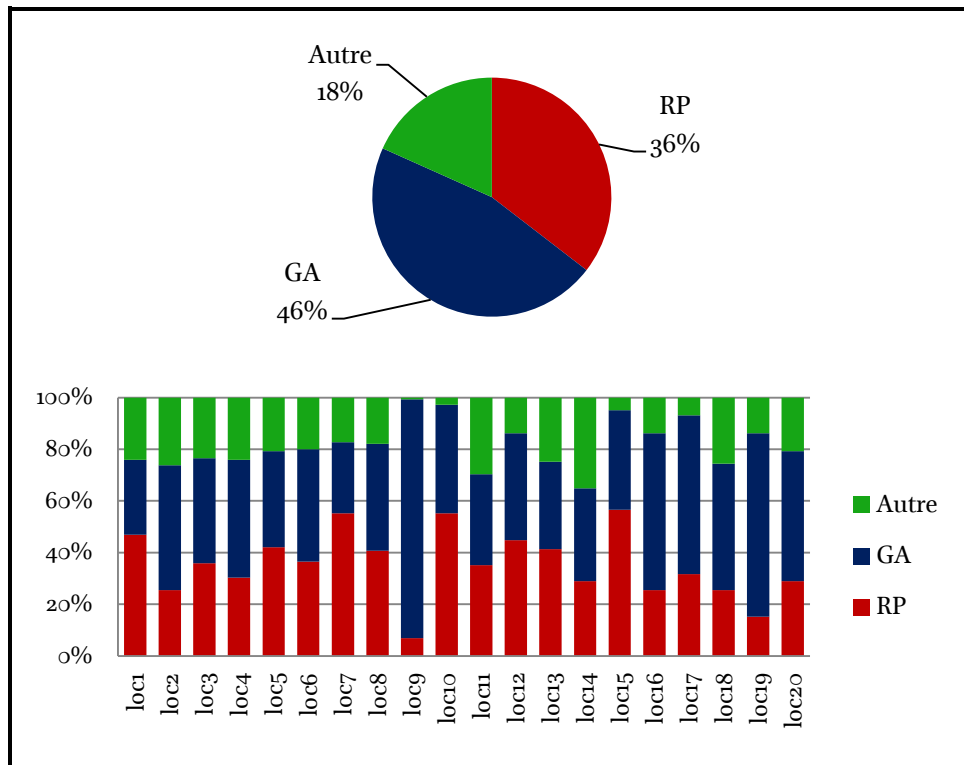


Figure 107 : pourcentages généraux pour tous les éléments étudiés et répartition par locuteur.

Les résultats de la comparaison indiquent que les caractéristiques proches du GA dominent dans les deux modalités (45 % en VN, 46 % en VR). Elles sont suivies de celles de la RP qui connaissent une légère diminution en VR (36 % contre 44 %VN). En revanche, nous observons que les caractéristiques « Autres » restent toujours en dernière position même si elles augmentent de 11 % en VN à 18 % en VR (Figure 108).

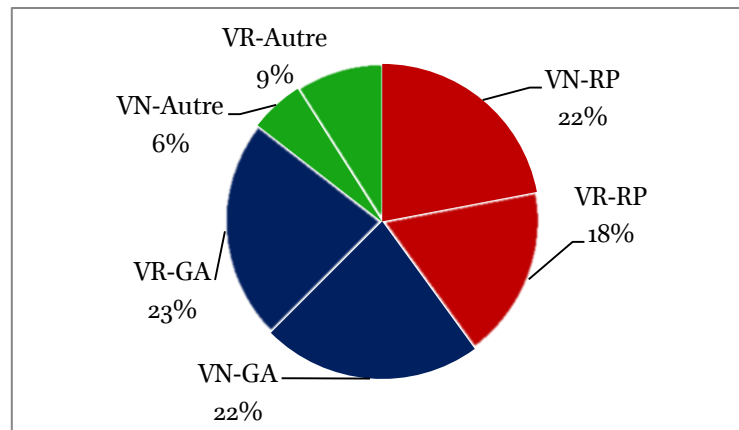


Figure 108 : comparaison VN/VR

Enfin, nous résumons dans le tableau suivant les résultats déjà présentés plus haut et qui représentent les pourcentages d'ensemble de chaque locuteur pendant les deux modes de productions (Tableau 26). Dans la section suivante, nous allons discuter ces résultats généraux par rapport à la comparaison de ces deux modes, en montrant les conséquences de l'augmentation de la vitesse de parole.

Tableau 26 : Comparaison des productions en % des caractéristiques proches de la RP, du GA et « Autres », en VN et en VR par locuteur.

	% RP		% GA		% « Autres »	
	VN	VR	VN	VR	VN	VR
loc1	56	47	29	29	15	24
loc2	30	26	48	48	21	26
loc3	28	36	65	41	7	23
loc4	34	30	59	46	7	24
loc5	45	42	48	37	8	21
loc6	49	37	40	43	11	20
loc7	69	55	27	28	4	17
loc8	32	41	56	41	12	18
loc9	17	7	83	92	0	1
loc10	46	55	43	42	12	3
loc11	44	35	32	35	24	30
loc12	48	45	41	41	10	14
loc13	51	41	37	34	12	25
loc14	58	29	27	36	15	35
loc15	63	57	30	39	7	5
loc16	34	26	54	61	12	14
loc17	45	32	48	61	8	7
loc18	48	26	42	49	10	26
loc19	41	15	48	71	12	14
loc20	42	29	39	50	19	21

8.13.1 Discussion générale

À la lumière des résultats présentés ci-dessus, nous retenons plusieurs points qui méritent d'être discutés et que nous avons classés en cinq catégories les plus pertinentes. Globalement, les productions des locuteurs sont caractérisées par une variabilité complexe, non seulement sur le plan inter-individuel mais également intra-individuel, comme nous avons pu le constater en VN. Toutefois, en VR, ces variabilités se manifestent différemment en fonction des locuteurs et des sons/séries de sons étudiés. Les cinq aspects que nous présentons concernent principalement la comparaison des deux systèmes de productions et sont consacrés aux effets qu'entraîne la VR, accompagnés d'exemples précis. Ces conséquences sont autant valables au niveau de l'ensemble des sujets qu'au niveau de l'individu. Premièrement, la VR a engendré une diminution dans le nombre de production de certains sons. Par exemple, les sons [t^h] et [tʃ] dans *city*, la rhoticité dans *more* et le [u:] dans *tune*. Deuxièmement, les sons [ɒ] dans *hot*, [aɪ] dans *vitamin* et [e] dans *evolution* ont vu le nombre de leur production augmenter en VR. Il faut noter que la diminution ou l'augmentation n'entraîne pas systématiquement un changement de position de dominance. La production d'un son peu avoir diminué ou augmenté mais elle conserve une majorité ou une minorité sur l'ensemble. C'est le cas du [ə] dans *military* qui domine les productions malgré une diminution en VR. Troisièmement, l'augmentation de la vitesse d'élocution n'a eu aucun effet sur le [ɔ] dans *joke*, le [y] dans *tune* et l'absence de voyelle dans *military* ([]). Ces derniers restent stables dans les deux modes (même pourcentage). Il faut remarquer que la majorité de ces sons proviennent des productions « Autres ». Quatrièmement, on remarque en VR l'émergence du son [i] dans la prononciation de *dynasty* et *vitamin*, l'absence de voyelle à la fin du mot *actor* ([ækt]) ainsi que dans la troisième syllabe d'*organisation* (,ɔ:ɹgən'zeɪʃən^{RP}). Et enfin, on assiste en VR à la disparition des sons [tʃ] dans *party*, [əʊ] dans *road* et [a] dans *band* et *hand*, pourtant existants en VN.

Parallèlement, nous avons évoqué les productions homogènes pour qualifier les productions d'un même type de sons ou séries de sons aux mêmes pourcentages dans les deux modalités. Les mélanges, eux, concernent des productions de différentes sortes de sons ou séries de sons. Ces phénomènes dépendent bien évidemment des locuteurs et des sons/séries de sons pris en considération. Certains mots subissent plus de perturbations que d'autres. C'est le principe : lorsque c'est hétérogène, il y a variation. Par exemple, nous avons noté moins de productions

différentes et plus d'homogénéité en VN qu'en VR, pour les mots *party, necessary, organisation*. À l'inverse, les mots *locate, hot et top, road* sont produits avec plus de variations en VN qu'en VR. Globalement, pour tous les locuteurs et pour tous les mots, il y a eu plus de variabilité de VR qu'en VN.

D'autres conséquences importantes de la VR valent la peine d'être soulignées. Tout d'abord il y a le phénomène de la réduction vocalique qui consiste en une diminution de la durée d'une voyelle (Bürki, Fougeron, Gendrot & Frauenfelder, 2011 ; Fougeron, 2013 ; Lancien, Nicolas & Fougeron, 2018). Ensuite le phénomène de monophthongaison qui consiste à réaliser une diphtongue supposée en une monophthongue (*joke* et *road* : [ɔ] ou [o] au lieu de [ou] et [əu]). Puis il y a la non-production de voyelles (*legalisation, organisation, actor...*) et enfin le phénomène de centralisation vocalique, où lorsque la production des voyelles tend vers le centre du triangle vocalique, c'est-à-dire vers le [ə] ([ɪ] => [ə] : [ɔ:gənə'zeɪʃən]^{RP} au lieu de [ɔ:gənɪ'zeɪʃən]). En VR, les gestes articulatoires sont réalisés plus vite, permettant ainsi d'atteindre la cible plus rapidement. Ainsi le [ə] et [-] renforcent cette hypothèse : les locuteurs produisent des [ə] ou ne produisent rien du tout. Ces constats sont toutefois prévisibles car cela a été montré dans la littérature scientifique (Harmegnies & Poch-Olivé, 1992). D'ailleurs, l'apparition des nouveaux sons s'expliquerait par la nativisation (exemple du son [i] dans *vitamin* et *dynasty*, alors qu'il est absent en VN) ou le « ratage » de cibles. Certaines productions pourraient aussi s'expliquer par le fait qu'en VN, ils disposent d'assez de temps pour réfléchir à la prononciation des mots (une sorte d'hyperarticulation/hypercorrection s'installe) alors qu'en VR ils « n'ont pas le temps d'y penser » et donc prononcent « naturellement ». Nous pensons que le fait de trop réfléchir avant de prononcer en VN (pour avoir la bonne prononciation) pourrait aussi être à l'origine de certains types de productions « Autres ». Cette question d'attention portée à la prononciation en VN pour bien prononcer pourrait expliquer pourquoi on assiste à la production de plus de sons distincts en VR, puisque les locuteurs ne réfléchissent pas avant de prononcer. Le « naturel » (sons « Autres ») revient alors qu'il est évité en VN (ils prononcent selon le premier son qui leur vient en tête). En outre, ceux qui conservent les mêmes sons dans les mêmes proportions, dans les deux modalités ont des productions robustes, c'est-à-dire qui résistent à la perturbation (ou éventuellement réajustent leur articulation en vue d'obtenir les mêmes sons). Conséquemment, ceux qui changent de productions entre la VN et la VR ont des productions non stables : nous pouvons alors parler de

la non-résistivité ou non-ancrage solide de ces sons (Vaxelaire, 2007). Pour finir, nous voyons combien les résultats montrent la variabilité émergente et nous permettent de mieux comprendre la différence entre les pourcentages d'ensemble et ceux des éléments isolés. En effet, dans le résumé du tableau ci-dessous (Tableau 27), on remarque que les proportions individuelles des éléments consonantiques, vocaliques et suprasegmentaux changent beaucoup selon les deux modes, mais les pourcentages d'ensemble (de tous les éléments) nous donnent un aperçu concis et précis.

Tableau 27 : Comparaison des productions en % des caractéristiques proches de la RP, du GA et « Autres », en VN et en VR par groupes d'éléments étudiés.

	% de caractéristiques de la RP		% de caractéristiques du GA		% de caractéristiques « Autres »	
	VN	VR	VN	VR	VN	VR
Totaux						
éléments consonantiques	40	52,5	47	37,5	13	10
éléments vocaliques	55	27	29	41	16	32
éléments suprasegmentaux	43	44	44	52	13	5
tous les éléments étudiés	44	36	45	46	11	18

8.14 Conclusion

Ce chapitre est la suite logique du précédent et est consacré à la présentation des résultats de la prononciation des mots en condition de vitesse rapide pour nos 20 locuteurs. Le nombre d'éléments de prononciation reste le même (24 mots, 2900 répétitions, 35 différentes productions). Ici également, les résultats obtenus ont été soigneusement décrits, illustrés au travers de deux types de graphiques (les diagrammes en barres et les secteurs) et commentés pour l'ensemble des locuteurs. Par conséquent, les productions analysées se rapprochent soit de l'accent britannique (RP), soit de l'accent américain (GA) ou soit ne relèvent ni de l'un, ni de l'autre (« Autres »). Notons également que certains mots sont plus susceptibles d'être prononcés avec des propriétés de la RP (*yod-dropping*), d'autres attirent plus de caractéristiques du GA (rhoticité) pendant que d'autres sont plus réalisés avec les caractéristiques « Autres » (/t/ intervocalique).

De plus, nous avons procédé à une comparaison et avons montré les différences qui existent dans les productions entre les deux modalités d'élocution. Suite à la confrontation de ces données, nous constatons une variabilité inter et intra-individuelle beaucoup plus importante en condition de vitesse rapide. En fonction des sons/séries de sons pris en considération, certains locuteurs gardent les mêmes productions à la même proportion dans les deux systèmes, tandis que d'autres y

mélangent différents sons/séries de sons à divers pourcentages. Les résultats de la comparaison indiquent que les caractéristiques proches du GA dominant dans les deux modalités (45 % en VN, 46 % en VR). Elles sont suivies par celles de la RP qui connaissent une diminution en VR (36 % contre 44 %VN) et enfin des caractéristiques « Autres », qui restent toujours en dernière position même si elles augmentent de 11 % en VN à 18 % en VR. En outre, nous avons identifié cinq changements induits par la vitesse rapide de parole : il s'agit de la diminution dans la production de certains sons ; de l'augmentation d'autres sons/séries de sons ; de la phase de stabilité où il n'y a eu aucun effet ; de l'émergence de nouveaux sons et de la disparition d'autres sons et séries de sons. Concernant l'émergence des réalisations « Autres », nous avons argumenté sur leurs causes possibles, en mettant en avant toujours l'impact de la L1 d'une part et l'influence de l'augmentation de la vitesse de production d'autre part. Aussi avons-nous noté le fait que les productions homogènes relevaient de la résistivité du système de production des sons/séries de sons concernées. Ces résultats sont très intéressants pour nous, car ils montrent que malgré la perturbation certains locuteurs sont capables de maintenir les mêmes stratégies articulatoires, c'est-à-dire un système qui est viable. Il nous semble que c'est aussi une preuve que ces productions sont ancrées dans la prononciation de ces locuteurs.

Ayant présenté et discuté des résultats de la prononciation de nos locuteurs en VN et en VR, nous poursuivrons dans le chapitre suivant avec la présentation des résultats du questionnaire. Puis, nous présenterons les résultats du test de perception qui a pour objectif de savoir si les locuteurs distinguent les deux accents. Nous exposerons ensuite, les retombées de la triangulation dans laquelle nous croiserons les trois sources de données. Il s'agira de rassembler des variables importantes des productions orales dans les deux conditions de parole, du test de perception et du questionnaire, en vue d'établir des tests statistiques de corrélation.

9 Résultats du questionnaire et du test de perception

Une enquête récente du CSA²²⁸ montre, en effet, que nous sommes le seul pays d'Europe dans lequel deux séries américaines sont en tête des audiences des fictions (« Mentalist » et « Dr House ») et où six séries du top 10 sont américaines. En Italie, les dix premières sont italiennes (avec « Maria de Nazaret » en tête, qui ne nous est pas parvenu), en Espagne « Aquila Roja », en Allemagne « Tatort ». Hormis cette dernière, la plupart des séries qui triomphent chez nos voisins ne sont jamais passées en France. « Mentalist », « NCIS », Les Experts : pourquoi les séries américaines cartonnent en France ? (Jost, 2013).

Nous présentons dans ce chapitre les résultats du questionnaire que nous avons distribué à nos 20 participants, en vue de connaître des informations sur leurs pratiques d'exposition à l'anglais, après la prise de sons et le test de perception (cf. section 6.3). La première partie du questionnaire comprend sept champs non numérotés et permet aux sondés de renseigner nom et prénom, âge, genre, études, année d'étude, L1 et L2. La deuxième partie est composée de 25 questions principales dont une quinzaine sont des questions imbriquées (exemple : si oui, pourquoi ?). À la fin des questions, nous laissons au répondant la possibilité d'ajouter des commentaires. Une copie du questionnaire est consultable dans les annexes (Annexe A 4, p. 184). Ensuite, les thématiques traitées dans le questionnaire sont classées en trois catégories, à savoir : contact(s) avec la langue anglaise, séries télévisées et adoption d'accent. Par la suite, ces thématiques sont analysées en sous-thématiques, permettant de mieux visualiser les diverses composantes qui forment le système sous-jacent des pratiques des participants. Nous présenterons dans cette section les informations biographiques, le temps d'apprentissage et d'utilisation de l'anglais, les séjours effectués et leur durée dans les pays anglophones, la nationalité et l'accent des enseignants d'anglais, les mots inconnus, l'imitation d'accent et les pratiques d'exposition aux *inputs*. Dans cette dernière partie consacrée aux activités sérielles, nous décomposerons les pratiques en fonction des modes de visionnage de séries, les raisons du visionnage de celles-ci, les intérêts pour les séries, les supports utilisés pour le visionnage, les séries télévisées visionnées et enfin les accents dominants dans les séries. Nous terminerons la présentation des résultats par une discussion globale, reprenant quelques aspects notoires des résultats, sous forme d'essais d'explication et de propositions. Rappelons que le cœur de ce travail demeure l'analyse des

228 Conseil Supérieur Audiovisuel.

productions orales de ces locuteurs, mais l'objectif du questionnaire est de pouvoir procéder à des triangulations des différentes données, afin de mieux comprendre le parcours de chacun. Ces informations nous seront très utiles dans la description des liens éventuels qu'ont ces pratiques sur la prononciation des locuteurs, notamment les séries télévisées. Dans la même lancée, nous présenterons également les résultats du test de perception et ceux de la triangulation effectuée avec des variables des trois sources de données, à savoir les analyses de la prononciation, le test de perception et le sondage. Enfin, il est important pour nous de préciser que les résultats que nous présentons sont le reflet des représentations de ce qu'ils font et non d'une réalité objective. Ils ne sont donc pas mesurés et nous ne disposons pas de moyens afin de procéder à des vérifications. Toutefois, nous signalons dans la discussion les incohérences que nous identifions dans les réponses fournies. Dans la section suivante, nous commençons la présentation des résultats du questionnaire, par les données biographiques de nos locuteurs.

9.1 Informations biographiques

Les participants à cette recherche sont des jeunes femmes (11) et des jeunes hommes (9) âgés de 18 à 33 ans ($\bar{x} = 21,9$; $\sigma^{229} = 3,9$)²³⁰. Leur niveau d'étude va de la licence 1 au Master 2 et ils sont inscrits dans diverses filières académiques (sciences humaines et sociales, sciences dures) à l'Université de Strasbourg. Le français est leur première langue maternelle. Certains en possèdent une deuxième parmi lesquelles on retrouve l'alsacien, l'arabe, l'italien, le russe et le thaï. Ils étudient l'anglais en tant que première langue étrangère, à côté de l'allemand, de l'espagnol, de l'italien, du russe et du chinois. Le tableau ci-dessous montre les informations que nous venons de décrire pour chaque locuteur.

Tableau 28 : âge, genre, études et année d'étude des locuteurs.

locuteur	âge	genre	études	année
loc1	27	homme	médecine	L3
loc2	18	femme	sciences du langage	L1
loc3	22	femme	sciences de la vie	L1
loc4	27	femme	sciences du langage	M2

229 L'écart-type correspond à la mesure de dispersion entre les différentes valeurs de la variable concernée. Plus il est petit (tend vers 0) plus la distribution est homogène et tourne autour de la moyenne, c'est-à-dire moins il y a de la variabilité entre les données. Nous indiquons ainsi l'écart-type afin de rendre compte des différences par rapport aux données numériques renseignées dans le questionnaire.

230 Pour rappel nous employons le symbole \bar{x} pour représenter la moyenne et σ pour l'écart-type.

loc5	20	homme	sciences sociales	L1
loc6	21	femme	physique	L2
loc7	21	femme	physique	L2
loc8	18	homme	biologie	L1
loc9	19	femme	arts du spectacle	L1
loc10	25	homme	géologie	M1
loc11	24	homme	physique	L1
loc12	23	femme	lettres modernes	L3
loc13	20	femme	arts du spectacle	L1
loc14	33	homme	sciences du langage	M2
loc15	25	homme	sciences politiques	M2
loc16	19	homme	littérature	L2
loc17	18	femme	physique	L1
loc18	18	femme	physique-chimie	L1
loc19	19	femme	biologie	L1
loc20	22	homme	génie civil	M1

9.2 Temps d'apprentissage et temps d'utilisation de l'anglais

Le temps d'apprentissage de l'anglais en général, qui correspond aux années passées à apprendre cette langue dans un cadre formel, varie de 5 ans à 23 ans ($\bar{x} = 12,3$; $\sigma = 3,9$). À la question de savoir depuis combien de temps ils apprennent l'anglais à l'université, les réponses vont de 7 mois à 48 mois²³¹ ($\bar{x} = 21,4$; $\sigma = 14,3$). Nous constatons tout de suite qu'il y a une corrélation entre l'âge des locuteurs, le temps d'apprentissage de l'anglais en général et le temps d'apprentissage de l'anglais à l'université. En somme, nous observons que les plus âgés détiennent évidemment le plus grand nombre de temps d'apprentissage dans les deux situations (en général et à l'université) (Figure 109).

²³¹ Soit 4 ans. Nous mettons l'unité en mois pour rester cohérent avec le locuteur qui n'apprend l'anglais que depuis 7 mois et aussi pour respecter la valeur moyenne et l'écart-type.

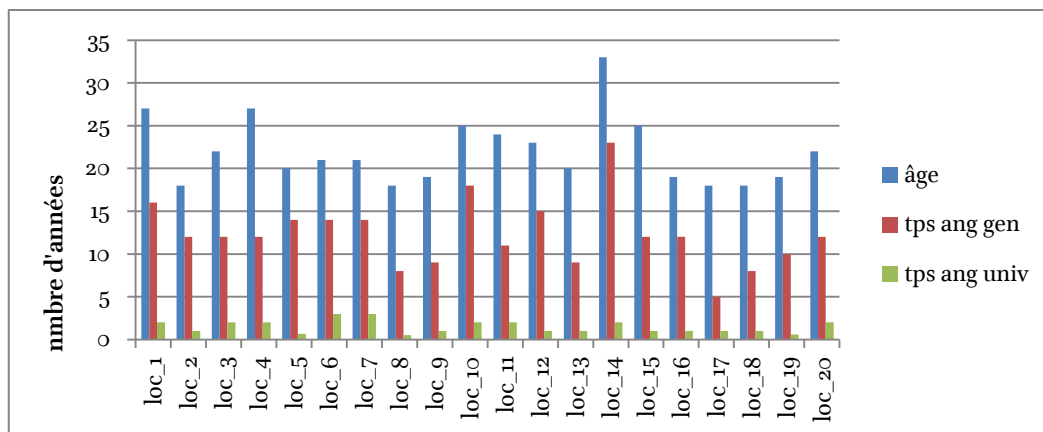


Figure 109 : croisement montrant le lien entre les données sur l'âge, le temps d'apprentissage de l'anglais en général et le temps d'apprentissage de l'anglais à l'université.

Ensuite, nous leur demandons d'estimer le temps qu'ils passent à utiliser l'anglais en dehors des cours à l'université. Les réponses varient de « moins de 4 heures » à « plus de 24h » par semaine. Cinq personnes pratiquent l'anglais pendant moins de 4 heures (locuteurs 2, 8, 9, 13 et 16) ; quatre personnes entre 4 et 8 heures (locuteurs 3, 11, 15 et 17) ; quatre autres entre 8 et 14 heures (locuteurs 5, 12, 18 et 20) ; deux personnes entre 14 et 18 (locuteurs 4 et 10) ; deux autres entre 18 et 24 (locuteurs 6 et 7) et enfin trois personnes pendant plus de 24 heures (locuteurs 1, 14 et 19) (Figure 110).

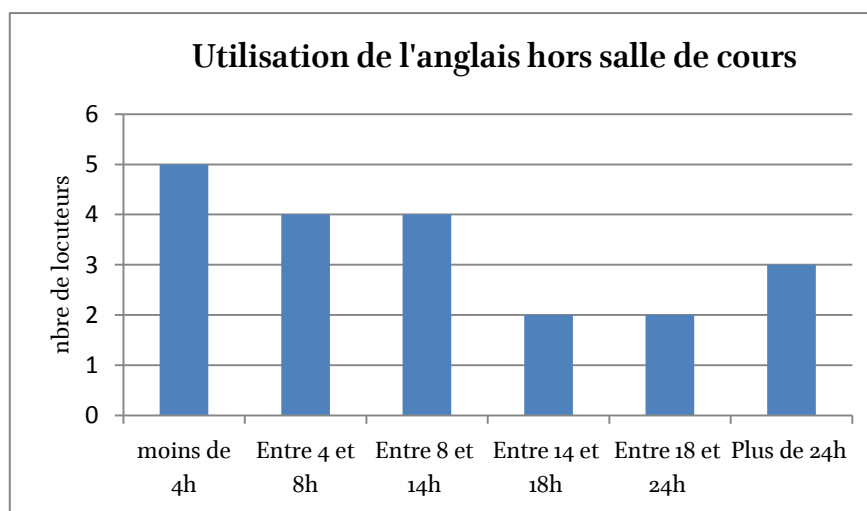


Figure 110 : nombre de locuteurs et le temps d'utilisation par semaine de l'anglais en dehors des cours.

Concernant les locuteurs qui communiquent en anglais, les moyens qui sont privilégiés sont surtout le présentiel (11 occurrences) et la visioconférence (9 apparitions). Certains utilisent le téléphone (appels et SMS). Concernant la communication (orale et écrite) en anglais par semaine,

onze locuteurs utilisent divers moyens pour parler, entre une et six heures ($\bar{x} = 3$; $\sigma = 1,7$), avec de la famille et des amis vivant aux États-Unis, au Royaume-Uni, au Canada, en Grèce et en France.

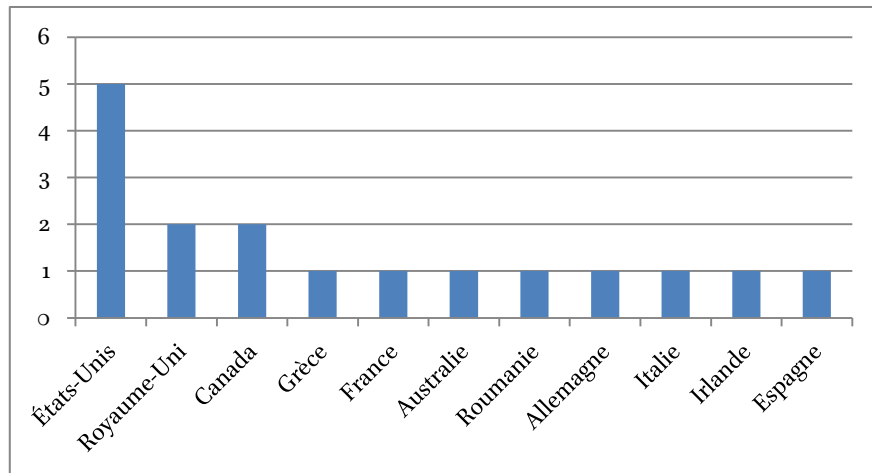


Figure 111 : fréquence des pays d'origine par rapport aux contacts en anglais.

En classifiant les relations qui existent entre les participants et les personnes avec qui ils parlent en anglais, nous relevons quatre catégories d'interlocuteurs : les amis, les colocataires, la famille et les partenaires (Figure 112).

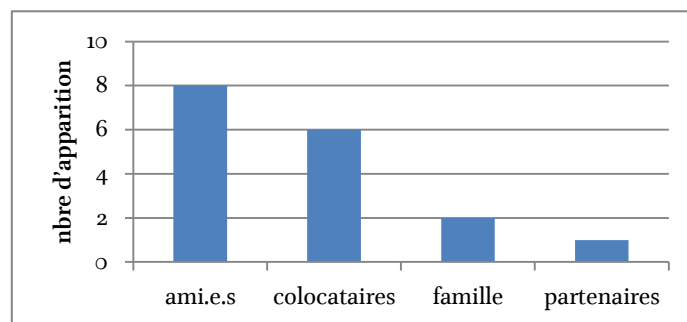


Figure 112 : relations avec les interlocuteurs.

De ce fait, le locuteur 1 passe six heures à parler avec des amis américains et ceci à l'aide de la visioconférence. La même chose pour le locuteur 2 qui y consacre trois heures. Le locuteur 3 parle également trois heures en présentiel avec des amis britanniques et roumains. Un mélange de moyens de communication est utilisé par le locuteur 4 qui parle anglais en présentiel, par téléphone, par SMS et par visioconférence, à raison de deux heures. Le locuteur 8 parle l'anglais en présentiel pendant quatre heures avec des colocataires américains, allemand et irlandais. Le locuteur 9 fait aussi de la visioconférence pendant quatre heures avec une de ces cousines qui habite au Canada. Le locuteur 10 passe environ une heure en visioconférence avec des amis

pendant que le locuteur 11 communique pour deux heures de présentiel avec ses colocataires grec, italien et espagnol. Le locuteur 12 par contre parle une heure avec sa mère, qui est française et enseignante d'anglais vivant en France. Le locuteur 17 fait deux heures de visioconférence avec des amies canadiennes et enfin, le locuteur 19 habite et parle au moins six heures avec son partenaire américain. En somme, huit locuteurs communiquent avec des amis, six locuteurs avec des colocataires, deux locuteurs avec la famille et un locuteur avec son partenaire.

9.3 Séjours et leur durée en pays anglophones

Nous avons demandé aux locuteurs s'ils ont déjà visité les États-Unis d'Amérique (EUA), le Royaume-Uni (RU) ou d'autres pays anglophones. Si la réponse est oui, ils devaient fournir des informations supplémentaires concernant le nombre de visites et la durée de chaque séjour. Les résultats sont illustrés dans ci-dessusci-dessous.

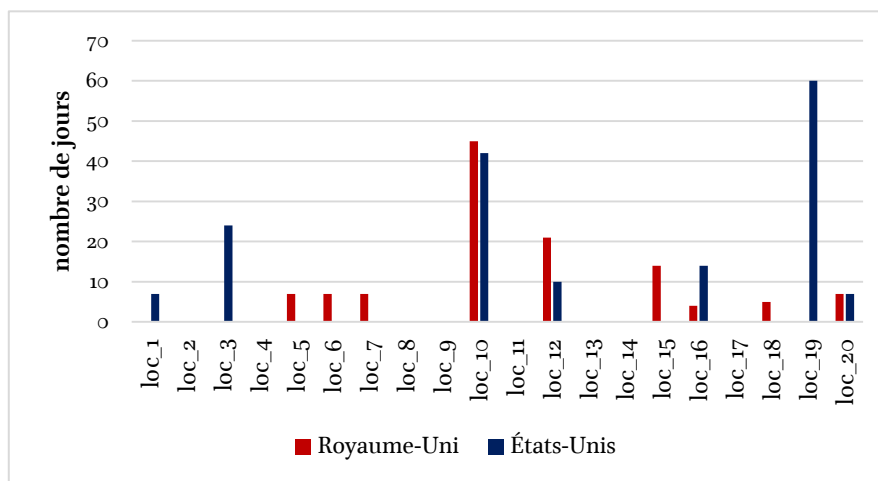


Figure 113 : temps passé dans un pays anglophone.

Ces résultats montrent qu'aucun locuteur n'a vécu au-delà de trois mois dans un pays anglophone. Huit locuteurs n'y sont jamais allés (locuteurs 2, 4, 8, 9, 11, 13, 14, 17). Les 12 autres sont allés au Royaume-Uni ou aux États-Unis entre une et trois fois pour une durée approximative d'un à 60 jours ($\bar{x} = 17,56$; $\sigma = 16,92$). Comme le montre la figure ci-dessous (Figure 113), le temps passé dans les pays anglophones est assez variable pour ces 12 locuteurs. Les locuteurs 1, 3 et 9 ne se sont rendus qu'aux États-Unis pour une durée de 7, 24 et 60 jours respectivement et les locuteurs 5, 6, 7, 15 et 18 se sont uniquement rendus au Royaume-Uni pendant 7, 7, 7, 14 et 5 jours respectivement.

Les deux pays sont visités par les autres locuteurs pendant : 45 jours RU, 42 jours EUA (loc10) ; 21 jours RU, 10 jours EUA (loc12) ; 4 jours, RU 14 jours EUA (loc16) et 7 jours RU, 7 jours EUA (loc20).

9.4 Nationalité et accent des enseignants

Afin d'évaluer les accents auxquels nos participants sont exposés dans le milieu scolaire (formel/non formel), nous leur demandons de fournir la nationalité et l'accent avec lequel les enseignants d'anglais parlent anglais durant les cours, et ceci de la classe de troisième jusqu'à la licence 3 (collège, lycée et université). Cette question est aussi particulière, car elle suppose que les participants s'en souviennent avec précision²³². Comme conséquence évidente, certaines cases sont restées sans réponses lorsque les locuteurs ne se sont pas souvenus. Toutefois, nous remarquons que la nationalité des enseignants a recueilli plus de réponses que la question sur l'accent. Aussi, il arrive que certains locuteurs ne se rappellent que de la nationalité d'un enseignant spécifique, mais pas de son accent et vice versa.

9.4.1 Nationalités des enseignants

Cent-cinquante-sept (157) enseignants depuis la classe de troisième jusqu'à la licence 3 sont mentionnés au total par nos locuteurs. Parmi ce chiffre figurent : 124 Français, 17 Britanniques, neuf Américains, trois Franco-britannique, un Australien, un Autrichien, un Hongrois et un Allemand. Les pourcentages sont exprimés ci-dessous (Figure 114). Si nous prenons la proportion des nationalités par cycle, nous constatons qu'il y a plus d'enseignants français dans les collèges et lycées des personnes interrogées qu'à l'université²³³.

²³² Certains locuteurs ne se souviennent ni de la nationalité ni de l'accent de leurs enseignants.

²³³ Ce qui semble logique, car il faut être Français (ou Européen) pour passer un concours de l'éducation nationale.

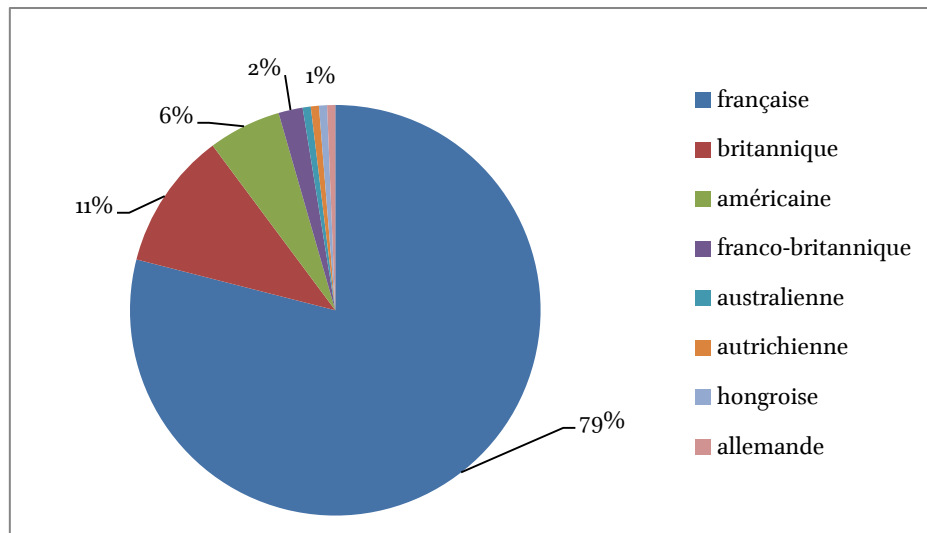


Figure 114 : nationalités des enseignants de la 3^{ème} à la licence 3.

9.4.2 Accents des enseignants

Nous posons la même question par rapport à l'accent de ces enseignants et obtenons 142 réponses. Étonnement la variabilité se réduit de huit nationalités à six accents. Être autrichien, hongrois ou encore allemand ne signifie pas que l'on parle avec ces accents respectifs, puisque les locuteurs ne mentionnent pas l'accent autrichien, hongrois ou allemand. En revanche, l'accent écossais apparaît dans la liste, alors qu'aucun locuteur n'a mentionné la nationalité écossaise. Nous constatons l'apparition en grand nombre de l'accent « français » et « franco-britannique »²³⁴, mentionnés par certains locuteurs. En effet, selon eux, quelques enseignants ont un accent typiquement français lorsqu'ils parlent en anglais (locuteur 2 et 17). Ils rajoutent que parfois les accents sont plus britanniques, mais avec des « touches »²³⁵ françaises (locuteur 14, 16 et 19), d'où le mot « franbritannique »²³⁶. Pour finir cette section, nous précisons que les résultats montrent clairement le fait que, dans la représentation des personnes interrogées, la majorité des enseignants français possède un accent britannique, le deuxième plus grand groupe aurait un accent français. Pour nous, ce type d'exposition peut influencer les éléments de prononciation des locuteurs : ils auraient une tendance à l'imitation des accents des professeurs pendant les cours

²³⁴ Français et britannique.

²³⁵ Synonyme de caractéristiques, mais nous faisons le choix de garder le mot tel écrit par les locuteurs concernés.

²³⁶ Le terme provient de nous.

d'anglais. Nous abordons la question de l'imitation d'accents par les locuteurs dans la section suivante.

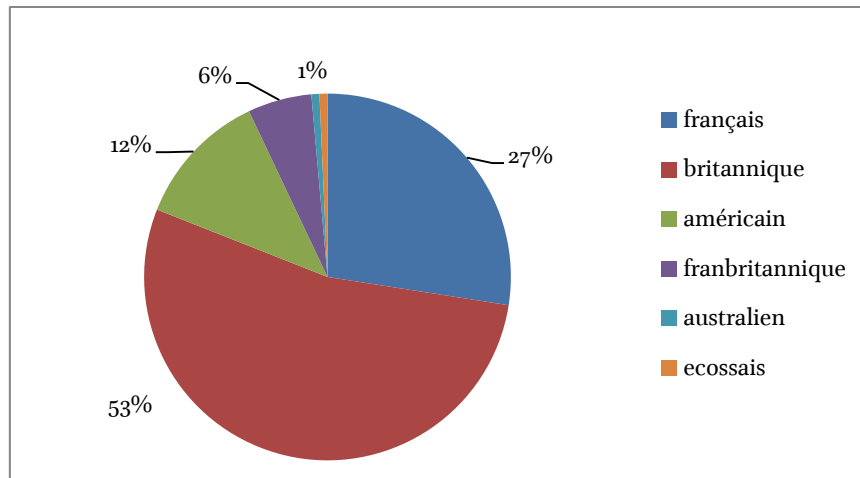


Figure 115 : accents des enseignants de la 3^{ème} à la licence 3.

9.5 Mots inconnus

Lors du remplissage du questionnaire, les locuteurs renseignent les mots de la liste globale (mots, phrases et texte) dont ils ne connaissent pas la prononciation. Le but est de comparer la prononciation de ces derniers avec celle des mots qu'ils connaissent. L'hypothèse que nous émettons est simple : les mots inconnus seront mal prononcés, tandis que ceux qu'ils connaissent seront prononcés avec plus de précision et d'assurance. Sur l'ensemble des mots, *cloak* est visiblement celui qu'ils ne connaissent pas (20 apparitions), suivi de *leisure*, *wrapped*, *tune*, *dynasty* et *vitamin* (Figure 116). Nous rappelons que les mots *cloak* et *wrapped* apparaissent seulement dans le texte, alors que les autres se retrouvent dans la liste des 24 mots.

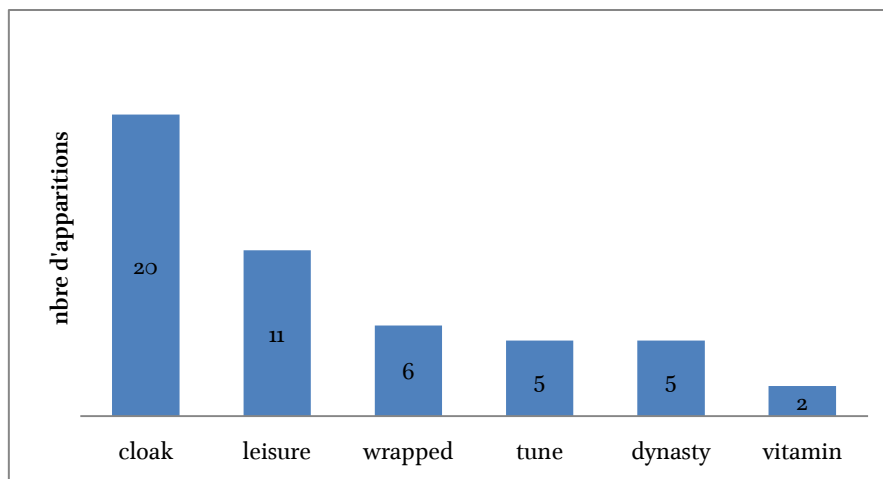


Figure 116 : nombre d'apparitions des mots inconnus mentionnés par les locuteurs.

La vérification de la fréquence de ces mots dans le corpus SOAP donne les résultats suivants : *wrapped* apparaît 1509 fois ; 564 fois pour *tune* ; 131 fois pour *vitamin* ; 120 fois pour *dynasty* ; 80 fois pour *leisure* et 44 fois pour *cloak* (Figure 117). En effet, la méconnaissance du mot *cloak* peut se justifier non seulement par le fait qu'il est le moins fréquent dans le corpus SOAP, mais aussi parce qu'il serait un mot du vieil anglais britannique²³⁷ qui est traduit aujourd'hui par *coat*. Le même constat concernant la fréquence d'apparition peut être établi pour le mot *leisure* (80 occurrences) et nous pouvons faire le lien avec la variabilité dans sa prononciation par nos locuteurs ('leɜə, 'li:ɜə, 'leɪɜə, 'leɜə). Ce mot serait difficile à prononcer (le <ei>) car, de manière non étonnante, il est prononcé de trois façons par les natifs anglophones ('leɜə, 'li:ɜə, 'leɪɜə). Par contre, les mots *wrapped* (le prétérit du verbe *wrap*), *tune* et *vitamin* sont les mots les plus fréquents dans le corpus SOAP, et pourtant, ils apparaissent aux 3^e, 4^e et 5^e positions des mots inconnus (avec respectivement 6, 5 et 5 apparitions) (Figure 116). En définitive, nous pouvons conclure que le classement établi par nos locuteurs est justifié par la fréquence de ces mots dans le corpus SOAP, à l'exception des mots *wrapped*, *tune* et *vitamin*.

²³⁷ Le dictionnaire étymologique en ligne etymonline indique que son usage date du 13e siècle et est dérivé de l'ancien français cloque. Voir <https://www.etymonline.com/search?q=cloak> (consulté le 23/06/2019).

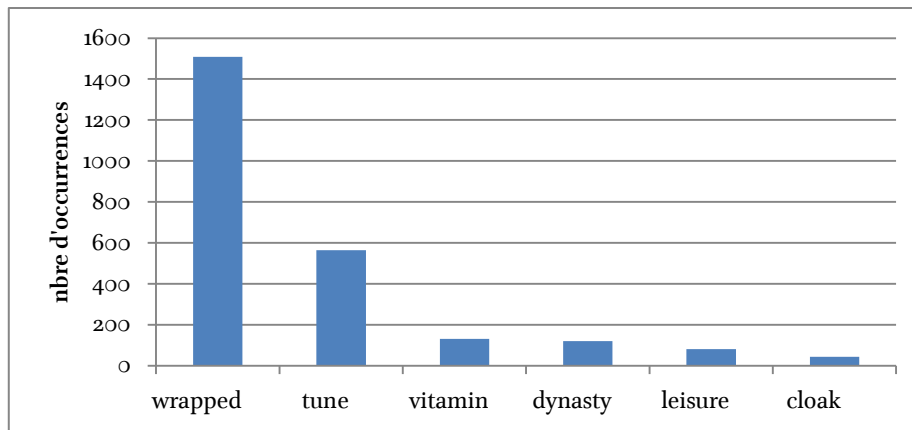


Figure 117 : fréquence d'apparition dans le corpus SOAP, des mots inconnus cités par les locuteurs.

9.6 Imitation d'accent

Cette section est consacrée à la notion de l'imitation d'accent et a pour but de recenser les impressions des locuteurs quant à leur prononciation en anglais. Nous avons posé une série de questions que nous rassemblons par thèmes pour plus de clarté. Premièrement, à la question : « lorsque vous parlez anglais, imitez-vous intentionnellement un type d'accent ? », les locuteurs doivent répondre oui ou non et fournir la raison de leur choix (question ouverte). L'objectif de cette question est de savoir s'ils imitent consciemment un accent spécifique, afin de pouvoir faire un lien avec leurs productions orales.

Tableau 29 : raisons pour lesquelles les locuteurs n'imitent pas un accent particulier.

locuteur1	j'essaie de m'intégrer l'accent de manière passive
locuteur2	j'ai beaucoup de mal à faire la différence des accents
locuteur3	pour ne pas compliquer l'apprentissage
locuteur6	j'écoute différents accents donc je ne peux pas en imiter un
locuteur7	C'est difficile d'en imiter un
locuteur8	car ma pensée va trop vite pour imiter un accent en même temps
locuteur9	c'est devenu normal pour moi de parler avec l'accent
locuteur10	le but est de parler pas de faire le clown
locuteur12	parce que je ne sais pas le faire et je n'en vois pas l'intérêt
locuteur13	l'accent me vient naturellement
locuteur14	c'est déjà assez dur comme ça
locuteur15	cela vient naturellement
locuteur16	c'est un mélange des deux accents
locuteur17	ça ne m'intéresse pas
locuteur18	je parle comme ça me vient, sans forcément faire attention à l'accent
locuteur19	parce que je ne m'en rends pas forcément compte
locuteur20	parce que c'est inutile de le préciser

De manière assez intéressante, uniquement trois locuteurs déclarent imiter l'accent américain. Les raisons sont respectivement : (1) « pour mieux prononcer, parler le mieux possible » (locuteur 4), (2) « j'ai l'impression que cela devient de plus en plus naturel » locuteur 5 et (3) « oui pour ressentir les mots quand je les prononce » (locuteur 11). Alors que la raison 2 nous semble correspondre à la réponse attendue, les deux autres raisons paraissent toutefois peu claires. Pourquoi faudrait-il imiter uniquement le GA pour mieux prononcer ou pour mieux ressentir les mots ? Ces locuteurs considèrent-ils le GA comme le meilleur modèle de prononciation anglaise ? Qu'en est-il de la RP ou d'autres accents ? Ces questions restent celles des représentations et de la construction identitaire (Rindal, 2010) qu'il conviendrait de vérifier dans une future étude, chez des étudiants français.

Concernant les autres locuteurs qui n'imitent aucun accent, les raisons fournies sont diverses et variées (*cf.* Tableau 29). À travers ces raisons, nous pouvons observer l'émergence de grandes tendances que nous résumons par trois mots : difficile, naturel et indifférence. En effet, pendant que certains locuteurs pensent qu'il est difficile pour eux d'imiter un accent (locuteurs 2, 3, 6, 7, 8 et 14), d'autres considèrent l'imitation d'accent comme étant un phénomène plutôt naturel (9, 13, 15, 18 et 19). Ce point nous rappelle l'imitation d'accent inconsciente dont nous avons traité dans le chapitre sur l'imitation (*cf.* section 4.4.2). Pour les derniers locuteurs, ils ne voient pas l'intérêt d'imiter un accent (10, 12 et 17). Nous remarquons que certaines de ces données sont très intéressantes pour des analyses didactiques liées à la psycholinguistique, dans la mesure où elles touchent à d'autres concepts du domaine (charge cognitive, attention, motivation et variation).

Deuxièmement, nous voulons savoir si des personnes de leur entourage (amis, famille, etc.) leur ont déjà dit qu'ils (les locuteurs) parlent anglais avec un accent. Douze locuteurs révèlent qu'ils pensent parler anglais avec un accent américain. Nous les interrogeons ensuite quant à leur perception de leur propre accent. Ils doivent répondre s'ils pensent parler anglais avec un accent américain ou non, et quelle en est l'origine selon eux. Les résultats montrent que douze locuteurs (les mêmes que précédemment) pensent parler avec un accent américain. Les origines selon eux se trouvent dans le tableau ci-dessous (Tableau 30). Globalement, ces locuteurs attribuent leur accent aux médias (films, séries, musique) américains. Seul le locuteur 19 pense que son accent est dû aux « échanges avec une correspondante américaine ».

Tableau 30 : origine des accents adoptés par certains locuteurs, selon eux.

locuteur1	les médias en général
locuteur3	les films, séries et musique américains
locuteur4	l'influence des média américains
locuteur5	les séries
locuteur6	je regarde beaucoup les séries télévisées américaines
locuteur7	je regarde les séries télévisées américaines
locuteur8	la musique et les séries
locuteur9	parce que je regarde beaucoup de séries et films américains
locuteur11	films et musique
locuteur14	enfin, surtout avec un accent français
locuteur15	séries tv
locuteur19	j'ai fait un échange aux USA avec une correspondante

Suite à la question précédemment posée, les locuteurs s'expriment également sur l'accent britannique. Seulement quatre personnes pensent parler aussi avec un accent britannique (locuteurs 1, 2, 8 et 10). Ces dernières évoquent les raisons suivantes : « les cours », « je pense que mon accent englobe les deux », « la musique et les séries » et « l'accent britannique est plus joli (ex : water et wader²³⁸) ». Enfin, quatre participants pensent parler avec un autre accent outre les deux accents mentionnés (locuteurs 1, 5, 8 et 12). Il s'agit de l'accent « français » et les raisons sont respectivement : « l'école », « le français », « les cours » et « le français ». Les réponses que nous venons de présenter sont assez intéressantes car elles confirment largement les études sur l'AIAL et vont dans le sens des hypothèses que nous avons posées. L'impact des médias américains constituent un ressenti de la plupart de nos locuteurs.

9.7 Pratiques d'exposition à l'anglais via les séries télévisées

Nous posons des questions relatives à la consommation de séries télévisées aux locuteurs. Ils estiment depuis combien de temps ils les visionnent, leur quantité (nombre) et la moyenne de temps qu'ils consacrent au visionnage. Les renseignements portent tout d'abord sur les séries télévisées américaines plus spécifiquement, puis sur les autres séries (y compris britanniques). Nous rappelons que le choix des séries télévisées est motivé par les résultats de l'étude de Kusyk et Sockett (2012), qui montrent qu'elles sont les activités les plus pratiquées par les étudiants et qui

²³⁸ Exemple donné par le locuteur 10.

devraient avoir le plus d'impacts sur la prononciation²³⁹. Les sondés fournissent également des informations concernant l'écoute de la musique, ce qui nous permet d'établir des comparaisons avec les informations sérielles. Ainsi, ils peuvent dire s'ils écoutent de la musique en anglais et la moyenne de temps qu'ils y accordent. Les résultats indiquent que tous les locuteurs sauf 14, 16, 18 et 20 écoutent de la musique en anglais à raison d'une heure à six heures par semaine ($\bar{x} = 2,4$; $\sigma = 1,4$). Ils regardent tous²⁴⁰ des séries télévisées américaines depuis un minimum de deux à un maximum de 15 ans ($\bar{x} = 7$; $\sigma = 3,1$) et à raison de 2 à 30 heures de visionnage par semaine ($\bar{x} = 10,4$; $\sigma = 6,6$). Le nombre de séries télévisées américaines visionnées²⁴¹ varie entre 4 et 43, pour un total qui s'élève à 207 ($\bar{x} = 10,3$; $\sigma = 8,3$). Celui des séries britanniques se situe entre un et dix, avec un total de 22 séries ($\bar{x} = 2,7$; $\sigma = 2,9$). D'autres séries sont également visionnées telles que des séries canadiennes, coréennes, etc. (voir section 9.7.5 et la liste des séries en Annexe A 6, p. 189). La Figure 118 illustre le nombre de séries américaines visionnées par chaque locuteur comparé aux autres séries²⁴². Le locuteur 10 visionne beaucoup plus que les autres, c'est aussi parce qu'il fait partie de ces locuteurs qui consultent leur agenda afin de remplir le questionnaire (nous détaillons ce point à la section Séries télévisées visionnées p. 330).

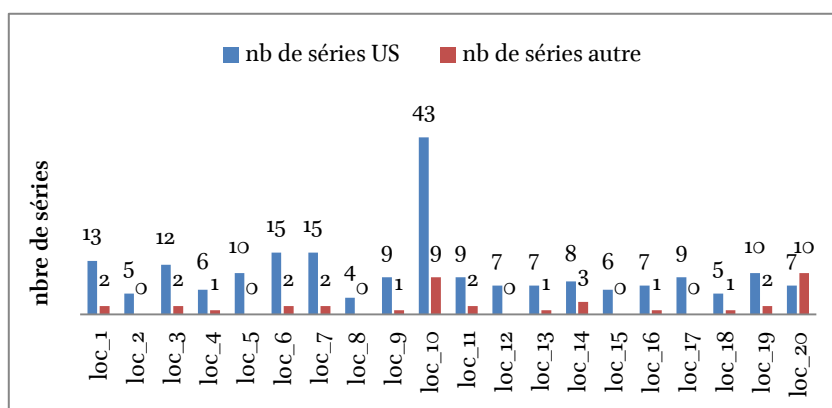


Figure 118 : nombre de séries américaines et autres visionnées par chaque locuteur.

Aussi est-il intéressant de voir la différence entre le volume d'heures que les locuteurs passent à écouter de la musique, comparé à celui du visionnage de séries. Pour l'ensemble des sondés, ils

239 En vue de cadrer notre recherche, nous nous sommes principalement intéressés aux séries télévisées, car elles constituent l'activité la plus pratiquée régulièrement parmi celles de l'AIAL (les films, documentaires, informations, ou vidéos divers de type Youtube, Ted, etc.).

240 Nous espérons rencontrer des locuteurs qui ne regardent pas de séries, en vue de procéder à une comparaison avec ceux qui regardent. Mais il s'avère que tous en regardent.

241 Nous n'avons pas d'informations sur la fréquence des visionnages.

242 Depuis le moment où ils ont commencé à en regarder.

passent 39 heures par semaine sur la musique contre 208 heures par semaine sur les séries américaines. La même remarque est faite concernant l'écart entre le volume d'heures de séries américaines (208), des séries britanniques (43 heures) et des autres séries (12 heures) (Figure 119). Les autres activités pratiquées en anglais demeurent la lecture en ligne (dix locuteurs), les jeux vidéo (trois locuteurs), l'écriture de texte (un locuteur) et le chant (un locuteur). En revanche le temps consacré à ces activités n'est pas renseigné. Ces informations nous sont capitales car elles confirment la position des visionnages des séries américaines comme étant la pratique informelle la plus fréquente.

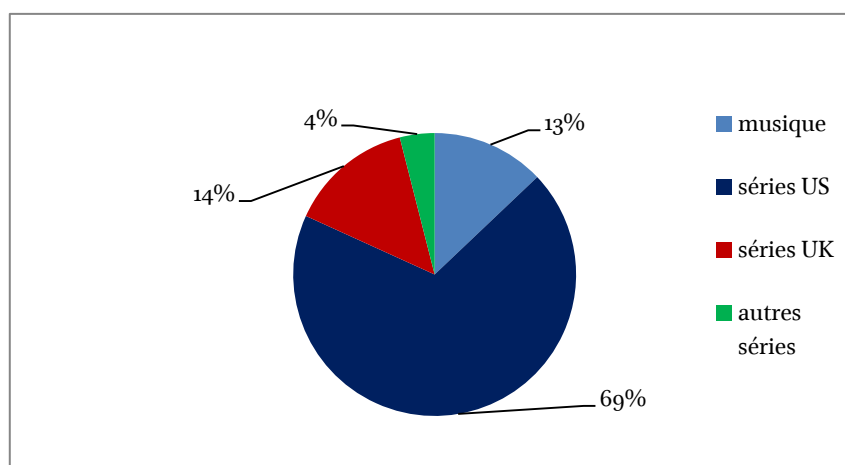


Figure 119 : comparaison des volumes horaires consacrés aux activités citées par les locuteurs.

9.7.1 Modes de visionnage de séries

Les modes de visionnage correspondent à la langue ou aux langues auxquelles les locuteurs sont exposés lorsqu'ils regardent les séries télévisées. Ainsi, nous proposons aux locuteurs de choisir parmi les quatre modes de visionnage connus et qui sont : Version Originale (VO), Version Originale Sous-Titrée en Anglais (VOSTA), Version Originale Sous-Titrée en Français (VOSTF) et la Version Française (VF ou le doublage). À ceci, ils doivent également fournir une estimation en pourcentage pour chaque mode choisi.

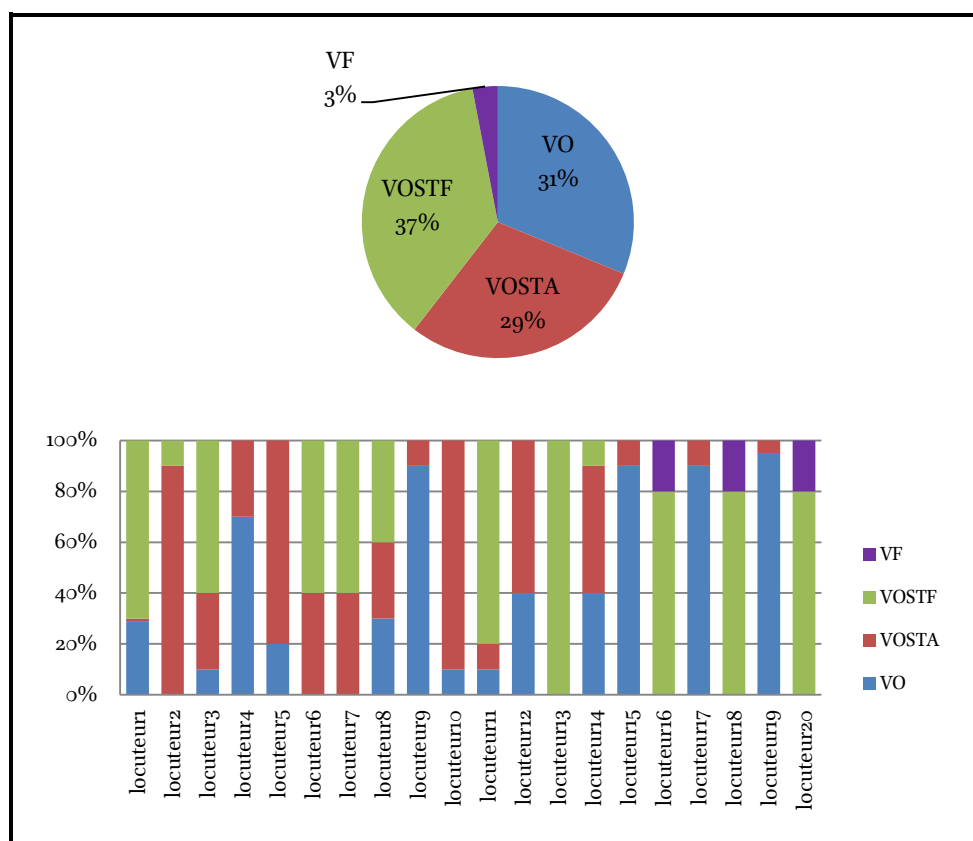


Figure 120 : pourcentages estimés des modes de visionnage.

Les résultats indiquent que la VOSTF est le mode dominant (37 % des modes de visionnages), suivi de la VO (31 %), de la VOSTA (29 %) et enfin la VF 5 (3 % uniquement). Au niveau individuel, les pratiques varient d'un à quatre modes, mais avec des pourcentages différents pour tous (Figure 120). Nous remarquons également que les trois locuteurs 16, 18 et 20 sont les seuls à combiner la VF (20 %) et la VOSTA (80 %). Ces différences de pratiques concernant les modes peuvent indiquer leurs niveaux en compréhension de la langue. En somme, si nous regroupons la part de l'*input* sonore anglais, elle représente 97 % des visionnages. Ce pourcentage montre le volume conséquent de l'exposition à la langue anglaise et les implications que cela peut avoir sur l'imitation d'accent. Les sous-titres seraient un moyen d'accéder plus facilement au sens, en complément aux lacunes qu'ils pourraient avoir au travers de la VO uniquement. Dans la section suivante, nous présentons les raisons pour lesquelles nos locuteurs regardent les séries. Les réponses peuvent nous être utiles afin de comprendre pourquoi tel mode de visionnage est plus prisé.

9.7.2 Raisons de visionnage de séries

Afin de connaître les raisons qui amènent nos locuteurs à visionner des séries télévisées, nous proposons trois choix auxquels ils doivent répondre oui ou non. Il s'agit de : (1) se divertir, (2) améliorer son accent et (3) améliorer sa compréhension orale. Les résultats montrent que le divertissement constitue la première raison pour laquelle ils regardent les séries télévisées (d'après 19 locuteurs). L'amélioration de la compréhension orale vient en seconde position avec 13 réponses positives tandis que la moitié des locuteurs visionnent les séries pour améliorer leur accent (Figure 121). Ensuite, dans le champ à réponses libres, trois participants rajoutent les raisons suivantes : « les doublages sont mal faits » (locuteur 2), « m'enrichir intellectuellement » (locuteur 5) et « recherche pour écrire du jeu de rôle » (locuteur 10).

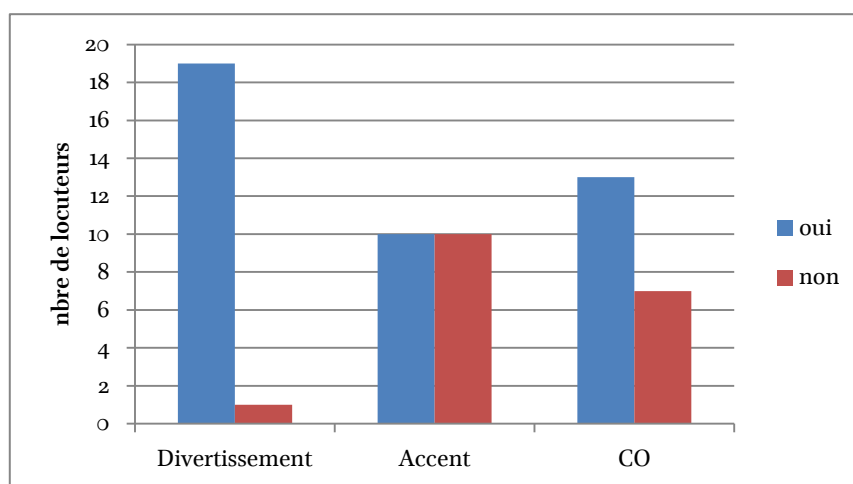


Figure 121 : les raisons de visionnage de séries télévisées.

De la différence entre les deux dernières réponses, nous pouvons déduire que la CO les engage plus que l'amélioration d'accent. Nous pouvons interpréter ces résultats par le fait qu'il est plus important pour eux de pouvoir comprendre les discours en anglais, que de devoir parler eux-mêmes avec les accents auxquels ils sont exposés. De plus, les raisons se rejoignent, puisque le croisement de ces données avec ceux de l'imitation d'accent, nous permet de comprendre pourquoi ils ne souhaitent pas imiter un accent spécifique. Aussi, pour eux, la capacité à comprendre les discours est primordiale, l'accommodation d'accent serait systématique et inconsciente, comme en témoignent les réponses : « c'est normal », « c'est naturel », « je ne fais pas attention », « c'est un mélange d'accent », etc. (cf Imitation d'accent). Dans la même ligne d'analyse, nous notons que la CO et l'amélioration l'accent comme des raisons de visionnage

semblent justifier le choix des modes VOSTA et VOSTFR. Cependant, nous ne voyons aucun lien logique entre ces modes et le divertissement, car il nous semble que le divertissement peut également avoir lieu en visionnant uniquement en VF. Sinon, visionner en VO uniquement tout en se divertissant, suppose d'être capable de comprendre tout le discours en CO et qui permet de comprendre tout le discours en anglais sans avoir besoin de sous-titres.

Maintenant que nous avons dégagé des liens possibles entre l'imitation d'accent, les modes et raisons de visionnage, nous exploitons dans la section suivante, une deuxième catégorisation. Celle-ci constitue les différents intérêts que les locuteurs ont pour les séries et ne sont pas en rapport direct avec les aspects linguistiques, mais plutôt avec le divertissement.

9.7.3 Intérêts pour le séries

Nous demandons aux locuteurs d'exprimer ce qui les intéresse le plus dans les séries. Cette question est ouverte et a pour but d'approfondir la dimension de divertissement que les locuteurs voient dans les séries. Les résultats varient entre une et cinq réponses (chaque locuteur a la possibilité de fournir cinq réponses). Ainsi, les résultats montrent que l'histoire est de loin la première chose qui les intéresse (mentionnée par 18 locuteurs), suivi par l'action, l'univers et les acteurs. Les proportions et les autres centres d'intérêt figurent ci-après (Figure 122).

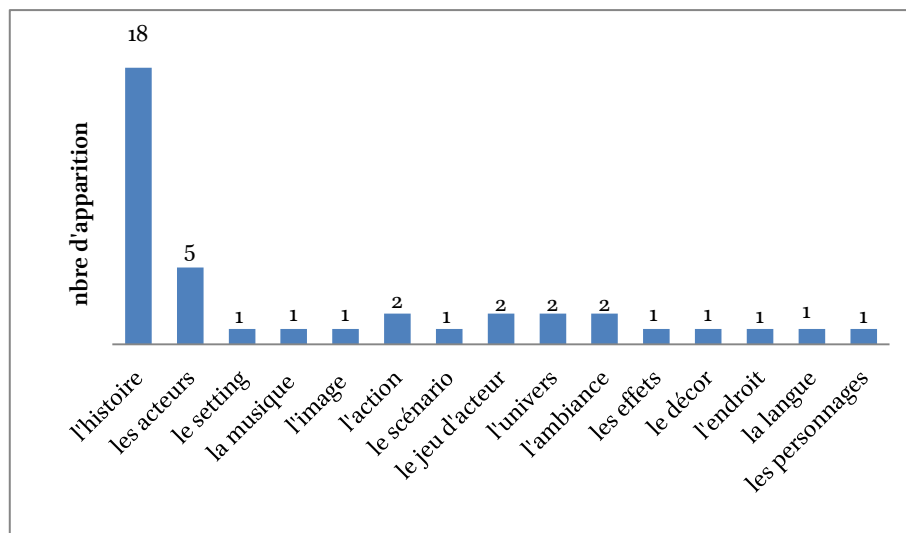


Figure 122 : nombre d'apparition des mots qualifiant les intérêts pour les séries.

En considérant que certains termes mentionnés se rapprochent ou se réfèrent à la même chose, nous procédons à des regroupements afin de réduire la variabilité dans les réponses. Ainsi, en se fondant sur les éléments classiques d'analyse littéraire, nous avons catégorisé ces facteurs selon le

récit, les personnages, le décor, le style et le thème. Nous obtenons ainsi un diagramme légèrement différent du précédent, quant au nombre d'apparition des mots qui désignent les intérêts des locuteurs pour les séries (Figure 123).

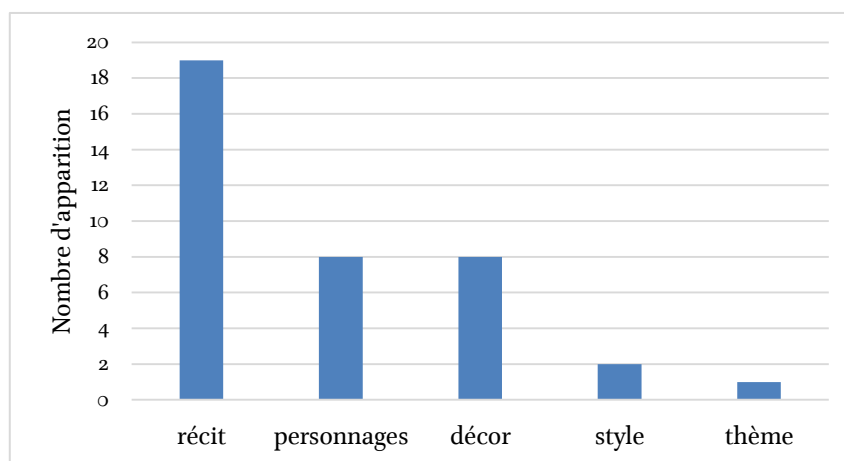


Figure 123 : regroupement en classes des mots qualifiant les intérêts pour les séries.

En définitive, nous retenons que le récit, mais aussi les personnages et le décor constituent les premiers intérêts des locuteurs lorsqu'ils regardent des séries télévisées. Même si la langue est mentionnée une fois, ces résultats indiquent que l'objectif original des visionnages demeure celui du divertissement et non pas pour des aspects linguistiques : une confirmation des études conduites jusqu'à présent dans le domaine de l'AIAL (Kusyk 2016). Nous allons voir dans la section suivante, les supports qui sont utilisés par les locuteurs pour le visionnage des séries.

9.7.4 Supports utilisés pour le visionnage

Quatre propositions sont faites concernant les supports qui sont utilisés pour regarder les séries : (1) sur Internet, (2) téléchargement sur disque dur externe (DDE), (3) au cinéma et (4) à la télévision. La première proposition inclue le visionnage sur YouTube ou des sites de *streaming* sur lesquels les services de téléchargement sont également possibles, gratuitement ou moyennant quelques euros en fonction des séries et des offres. Les deux premiers choix supposent le visionnage sur ordinateurs (portables et fixes) l'un de manière synchrone (*streaming*) et l'autre asynchrone (téléchargement). Quant au troisième choix, il constitue un détracteur, dans la mesure

où les séries ne sont pas diffusées au cinéma. Enfin, nous avons inclus la traditionnelle et toujours usitée méthode de visionnage sur les postes téléviseurs²⁴³.

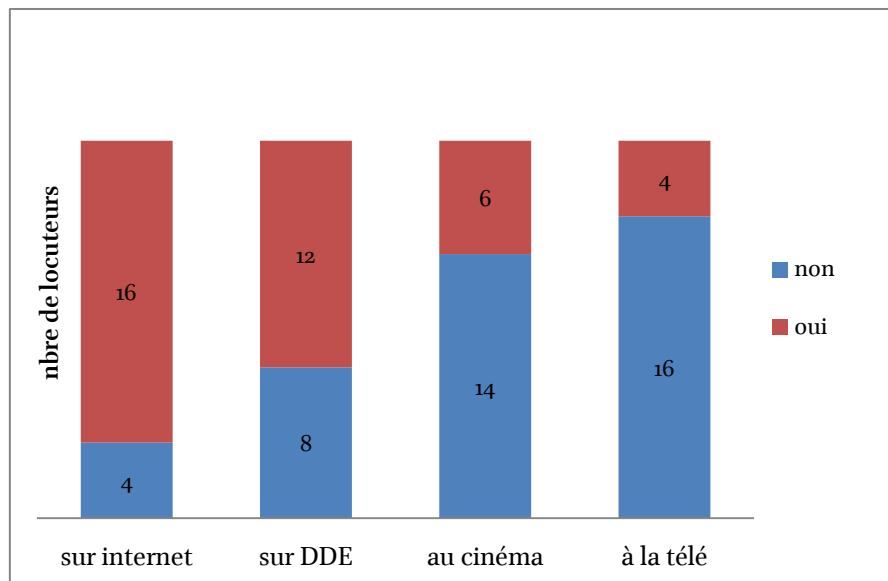


Figure 124 : les supports sur lesquels les séries sont visionnées.

Parmi ces choix, le visionnage sur Internet est l'activité la plus pratiquée, suivie par le téléchargement sur disque dur externe. La télévision demeure le support le moins prisé par nos locuteurs, et ceci pour diverses raisons²⁴⁴. Le cinéma est tout de même choisi par six locuteurs. Il se pourrait qu'il ait eu une confusion probable avec les films. Nous pensons qu'ils auraient choisi le visionnage de séries au cinéma soit par manque d'attention, d'intérêt, d'implication ou par négligence, soit ils ont renseigné tous les types de visionnage comme un ensemble. Dans ce cas, nous devrions extrapoler qu'ils ont également cité des films ou autres vidéos lorsqu'ils ont répondu aux questions relatives aux visionnages. Outre ces supports, nous offrons la possibilité d'en renseigner d'autres, grâce à une case vide. Seul un locuteur mentionne les DVD. Nous pensons qu'il s'agit de coffrets (DVD, *BlueRay*). Par ailleurs, il est intéressant de noter l'asymétrie entre les réponses, concernant Internet et la télévision (Figure 124).

Cette section nous a permis de connaître les supports les plus utilisés par nos locuteurs pour visionner les séries télévisées. Les réponses indiquent une dominance des visionnages sur Internet

²⁴³ Hertzienne, satellite, câble, ADSL et fibre optique (voir <https://clesdelaudiovisuel.fr/Connaitre/Le-paysage-audiovisuel/Comment-recevoir-la-television> pour plus de détails. Consulté le 26/02/2018).

²⁴⁴ Les raisons peuvent se trouver dans Sockett et Kusyk (2012). En effet, les participants à cette étude veulent suivre les séries dès lors qu'elles sont sorties et ne peuvent pas attendre la VF (doublée). Autrement, le coût des abonnements par exemple chez Netflix peuvent également constituer un frein pour les étudiants.

et au téléchargement sur DDE. Ces résultats sont conformes à ceux obtenus dans les études sur l'AIAL (Toffoli et Sockett, 2010, Kusyk et Sockett, 2012, Kusyk, 2017). Regarder des séries seul (ou à deux) ne comporte aucune contrainte de temps ; on peut le pratiquer selon son temps libre et ses envies, tout ceci dans une logique purement divertissante. Dans le chapitre suivant, nous présentons le contenu, c'est-à-dire les séries télévisées qui sont regardées par nos locuteurs.

9.7.5 Séries télévisées visionnées

Les locuteurs sont amenés à lister toutes les séries télévisées qu'ils ont regardées dans le passé et celles qu'ils visionnent actuellement (au jour de la collecte des données). Cette question semble bien évidemment difficile et il est clair qu'ils ne citent que celles dont ils se remémorent. D'ailleurs certains participants effectuent rapidement des recherches en ligne, soit en regardant sur une liste de séries disponibles, soit en consultant un site spécial servant d'une sorte d'« agenda » ou « d'organisateur » qui recense des informations sur les séries visionnées. Ces sites constituent un phénomène nouveau dans la vie des sériephiles. On y trouve toutes les séries visionnées, les saisons et épisodes concernés, et le temps qui y est consacré. Ces plateformes informent également sur les dates de sortie des prochaines séries (mais aussi pour les films des sorties cinéma), la possibilité de gagner des points et d'échanger avec d'autres utilisateurs (**Erreur ! Source du renvoi introuvable.**). Parmi ces organisateurs, nous pouvons citer *TV show tracker 3*, *Sérievore*²⁴⁵, *Hobi*, *Mes séries TV*, *Betaserie*²⁴⁶, *TV show time*²⁴⁷, etc. qui comptent des millions d'abonnés et existent aussi sous forme d'applications mobiles pour plus de facilité de consultation. Ces organisateurs donnent une traçabilité des activités de ces sériephiles et constituent une véritable richesse à exploiter dans le cadre des recherches en AILE et autres (Combes, 2017). L'exemple des renseignements se trouvant sur la figure 18 est tout à fait intéressant dans la mesure où l'utilisateur peut suivre et gérer ses activités sérielles selon plusieurs facteurs (temps, jours, bonus, genres de séries, etc.). Nous ignorons comment ces chiffres sont générés mais leur quantité peut être surprenante. Cependant, il s'avère que les étudiants (du moins ceux qui s'adonnent à ces activités) passent plus de temps à regarder les séries que d'heures en cours d'anglais à l'université, qui ne sont que de deux heures par semaines pour les spécialistes de disciplines non-linguistiques. On

245 <http://www.seriesvore.fr/>

246 <https://www.betaserie.com/>

247 <https://www.tvtime.com/fr>

peut toutefois se demander s'il était imaginable que les étudiants suivent les mêmes quantités horaires pour des cours de langues (458 séances de 30 à 45 minutes par mois²⁴⁸). Cette question reste un débat ouvert, mais il est certain que ces étudiants sont beaucoup plus exposés aux *inputs* télévisuels qu'aux *inputs* provenant de la salle de classe (enseignants et ressources confondus).

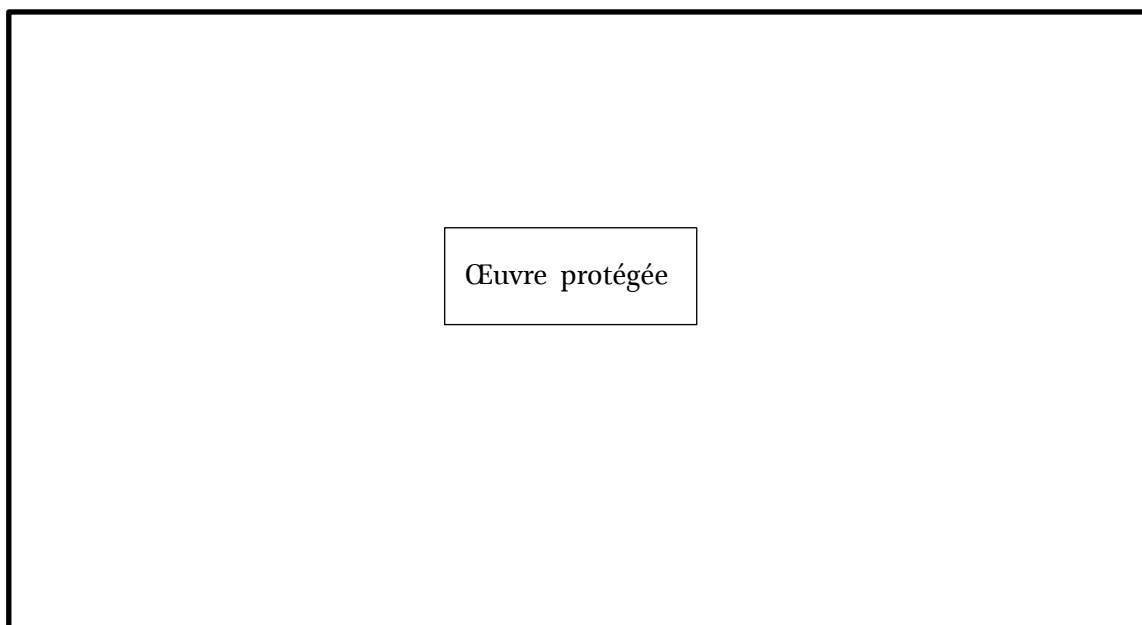


Figure 125 : copie d'écran de la page d'accueil de l'organisateur *Betaseries* d'un locuteur (avec son accord).

Conséquemment, nous retrouvons dans la liste compilée par nos locuteurs, aussi bien des séries plus ou moins récentes (de *Friends*, *X-files* et *Seinfeld* à *Mentalist*). Certains citent également des films (*Iron man*) et des dessins animés (Tortues Ninja). Ainsi, en raison de la variabilité et du nombre, nous ne donnons pas le nom de toutes ces séries télévisées citées dans cette partie. Toutefois, nous donnons des exemples de séries en guise d'illustration. En fait, dans le processus de remplissage, les participants mentionnent d'abord le nom des séries télévisées américaines, suivies des autres séries non-américaines. Mais, il s'avère que certains locuteurs se trompent dans le classement ou bien ils ignorent la provenance des séries, car certaines séries américaines sont citées parmi les autres catégories, notamment britanniques. Alors, nous nous sommes occupé de reclasser toutes les séries mentionnées selon leur nom, leur pays de provenance, leur genre et les accents qu'ont (ou qu'adoptent²⁴⁹) les acteurs. Ces informations se retrouvent dans un tableau en

²⁴⁸ Par référence à la figure 18 où l'utilisateur regarde 458,36 épisodes par mois.

²⁴⁹ Certains acteurs britanniques imitent l'accent américain dans des séries américaines (ou vice versa). Voir l'exemple de *Dr House* interprété par Hugh Laurie.

annexe (Annexe A 8, p. 188). En lien avec la question des accents rencontrés dans ces séries (voir section 9.7.6 ci-dessous), il nous est important d'évoquer brièvement le fait qu'aujourd'hui, il est de plus en plus fréquent de voir plusieurs acteurs avec des dialectes ou accents divers au sein d'une même série. En effet, il arrive souvent que des acteurs américains, britanniques ou australiens se côtoient sur scène, soit en imitant un accent particulier pour des questions d'homogénéité d'accents, soient ils gardent leur accent d'origine. Dans la même lignée, il y a des séries américaines (de production) dans lesquelles l'accent américain est totalement absent (*Game of thrones* par exemple). Ce phénomène pourrait expliquer le fait que certains locuteurs ne connaissent pas l'origine de certaines séries.

Concernant les résultats de la liste, les séries télévisées américaines représentent plus de 80 % (116) de la totalité des 216 séries mentionnées, comme le montre la Figure 126 ci-dessous. Les séries britanniques viennent très loin derrière (22 séries), suivies par celles issues de la collaboration entre les États-Unis et d'autres pays (Canada, Irlande, France).

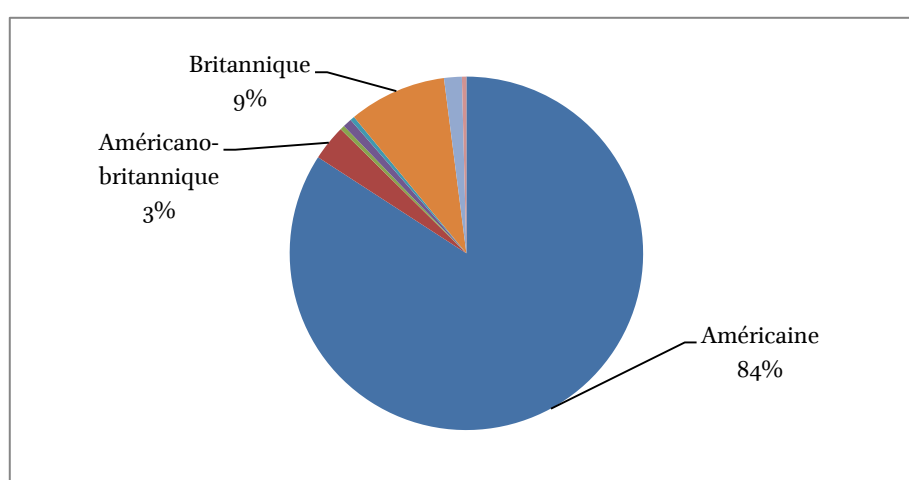


Figure 126 : la part des séries télévisées mentionnées par pays d'origine.

Classer les séries télévisées selon leur nombre d'apparition nous permet de dresser un tableau contenant le nombre de locuteurs qui citent une même série. Ainsi, parmi les trois premières séries les plus regardées, *Games of Thrones* vient en tête avec 14 apparitions, suivie par *Breaking Bad* (9) et *Sherlock* (9). Nous remarquons que la série britannique *Sherlock* occupe la troisième place des séries les plus regardées par le plus grand nombre de locuteurs. En d'autres termes, nous aurions pu nous attendre à ce que les premières séries les plus visionnées soit toutes américaines, en raison de leur quantité. Toutefois, il nous semble judicieux de faire la distinction entre le nombre de fois

qu'une série est mentionnée et la quantité de temps qui lui est allouée. En effet, certains locuteurs ne regardent que quelques saisons ou épisodes. Inversement, certaines séries sont moins mentionnées mais elles sont visionnées dans leur totalité, en termes de saisons, épisodes et nombre d'heures. Nous présentons ci-dessous (Figure 127), les premières séries mentionnées en fonction du nombre de locuteurs qui les visionnent. Concernant *Games of Thrones*, notre classement confirme le fait que c'est la série la plus piratée en 2015 avec 14,4 millions (M) de téléchargements via BitTorrent²⁵⁰. Les autres sont : *The Walking Dead* (6,9 M), suivie par *The Big Bang Theory* (4,4 M), *Arrow* (3,9 M) et *The Flash* (3,6 M)²⁵¹.

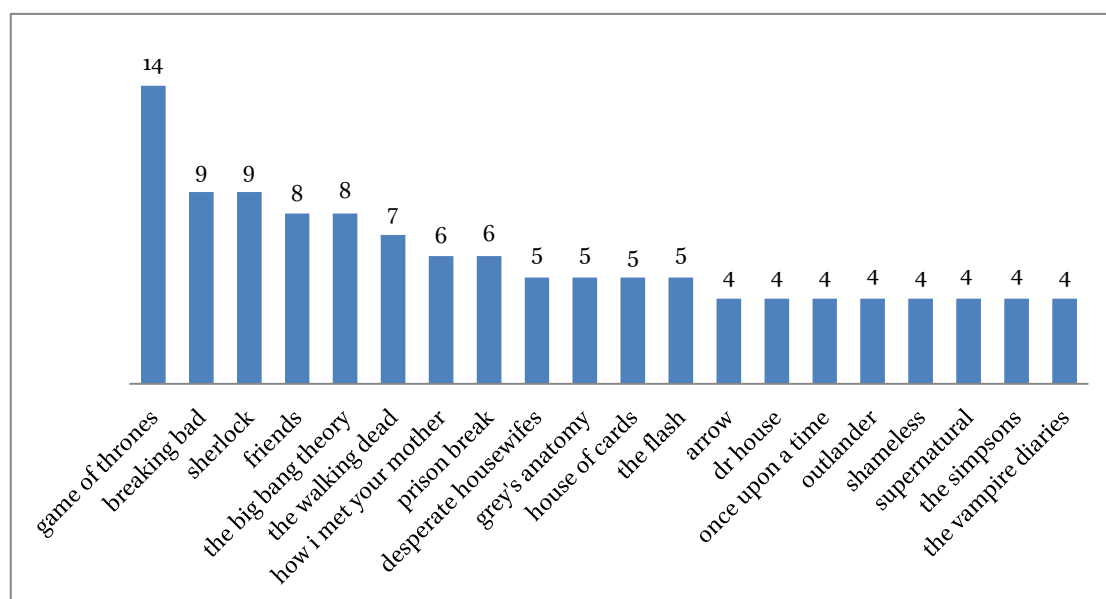


Figure 127 : les 20 premières séries avec le plus haut nombre d'apparition.

Nous venons de passer en revue les différentes séries télévisées que nos participants regardent, après une présentation rapide autour d'outils qui leur permettent d'organiser le visionnage de ces séries. Force est de constater que notre classement en fonction du nombre d'apparition des séries confirme les grandes tendances actuelles en matière de consommation sérielle. Il est à noter que ces pourcentages reflètent évidemment l'énorme quantité de séries américaines sur le marché. Enfin, dans la section prochaine, nous terminons la présentation des résultats par un bref commentaire sur les accents retrouvés dans ces séries.

²⁵⁰ Plateforme ou logiciel permettant de partager (téléversements et téléchargements) des fichiers (séries, films, jeux, logiciels, vidéos, musiques, etc.) entre divers utilisateurs.

²⁵¹ Source : <https://torrentfreak.com/game-of-thrones-most-pirated-tv-show-of-2015/>

9.7.6 Accents dominants dans les séries

Cette dernière section est consacrée à la présentation des accents que l'on peut retrouver dans les diverses séries télévisées mentionnées par nos locuteurs. Parmi ceux-ci, nous pouvons citer les accents américain, l'accent britannique, l'accent australien, l'accent irlandais, l'accent écossais, l'accent indien, canadien, afro-américain, etc. Comme nous l'avons déjà évoqué supra, une série d'origine américaine ne signifie pas que tous les acteurs ou tous les accents sont forcément américains. En effet, la primauté de l'accent américain dans les séries visionnées par les locuteurs est une conséquence évidente de la quantité de séries américaines listées ci-dessus (voir liste complète en Annexe A 6, p. 188). Pour ce qui concerne les séries dans lesquelles l'accent britannique est parlé, nous avons recensé : *The Office*, *Broadchurch* (UK), *Sherlock*, *Black Mirror*^{*252}, *Outlanders*, *Reign*, *Game of thrones**, *Orphan black**, *Penny Dreadful**, *ER**, *Rome*, *Black Sails**, *Taboo*, *Peaky Blinders*, *The crown*, *Only fools and Horses*, *Downtown Abbey*, *Twin peaks**, *The shannara Chronicles**, *The Hour*, *Hit and miss**, *Doctor Who*, *Utopia*, *Big school*, *The IT Crowd*, *Murdoch*, *The A word*, *Misfits**, *Skins*, *The Royals**. Le reste des séries contient uniquement l'accent américain.

Cette section se veut d'une importance capitale pour notre travail, car elle montre quels sont les accents auxquels nos participants sont les plus exposés dans leurs activités de visionnage de séries et subséquemment les conséquences sur le phénomène d'imitation phonético-phonologique (in/consciente). Nous achevons ce chapitre par une discussion générale sur l'ensemble des résultats qui y sont présentés.

9.7.7 Discussion générale sur les résultats du questionnaire

Tout d'abord, les questions de l'âge, du temps d'apprentissage de l'anglais en général et à l'université, ainsi que le temps d'utilisation de l'anglais en dehors de la classe sont primordiales afin de comprendre la durée d'exposition à l'*input* sonore et aux productions orales des participants. Nous pouvons admettre que logiquement, plus les locuteurs sont âgés et s'exposent à de l'*input* sériel, plus ils s'imprègnent de la langue. Ce processus aura comme conséquence un nombre important de caractéristiques de l'accent auquel ils sont le plus exposés. Le contraire

²⁵² L'astérisque (*) indique les séries dans lesquelles il y a un mélange d'accents, mais l'accent britannique reste majoritaire.

pourrait être vrai pour les moins âgés. Les mêmes suppositions sont valables pour les séjours dans les pays anglophones et aussi par rapport au temps de visionnage et le nombre de séries visionnées. Nous reviendrons vérifier ces hypothèses après la présentation des résultats du test de perception et des productions orales.

Ensuite, nous avons rendu compte des renseignements sur les pays dans lesquels nos locuteurs se sont rendus, mais il s'avère que ces informations sont insuffisantes si nous voulons approfondir les effets du contact avec un accent particulier sur leur prononciation. En effet, l'idée d'un pays est vague et nous avons vu qu'il y a rarement une seule variété d'accent. Par exemple, il aurait été intéressant de savoir si les locuteurs qui ont été aux États-Unis étaient plutôt à Boston ou à *New England*, où l'accent dominant est non rhotique, ou encore en Californie, à Los Angeles où l'accent est plutôt rhotique. La même remarque peut être faite sur les villes comme Dublin, Londres, Glasgow ou Edinburgh au Royaume-Uni. Ce point pourrait être à prendre en considération dans nos futures études sur l'appropriation de l'accent.

À propos de la durée des activités informelles, il y aurait une contradiction ou du moins une incohérence dans les estimations des locuteurs. En effet, lorsque nous comparons par exemple le temps d'utilisation de l'anglais en dehors de la salle de classe et l'ensemble des durées consacrées aux diverses activités (communication, musique et séries), nous remarquons qu'il y a une sous-estimation du temps total d'utilisation de la langue. En guise d'exemple, nous mentionnons les locuteurs 8 et 9 qui ont choisi entre 0 et 4h par semaine pour la question du temps d'utilisation de l'anglais en dehors de la classe, et qui ont également renseigné 4h par semaine pour la durée de communication en anglais. Or, plus loin, ils renseignent 1 et 2h par semaine pour l'écoute de la musique, 3 et 30h de visionnage respectivement. Jusque-là, nous n'avons pas pris en compte la durée des autres activités, comme la lecture en ligne ou les jeux vidéo (Schmoll, 2017). Deux explications peuvent justifier ces incohérences. D'une part, il se pourrait que ces locuteurs n'aient pas compris la question, à savoir que le mot « utilisation » implique aussi bien les activités de production que de réception en anglais. Il se pourrait qu'ils aient plutôt compris le mot « production » et pas « réception », d'où les 4h pour les deux questions. D'autre part, nous pensons qu'ils auraient « minimisé » ce temps d'utilisation, afin de ne pas culpabiliser à l'idée qu'ils passent beaucoup trop de temps à regarder des séries. Ce dernier point peut être sous-entendu dans les propos du locuteur 9 lorsqu'il dit : « Betaseries me permet de ne pas oublier et surtout de sauter

des épisodes. Mais j'avoue que ça fait peur de voir tout le temps que j'ai passé à regarder des séries télévisées ». En effet, les renseignements fournis dans le questionnaire par ce même locuteur (9), concernent les temps d'activités, le nombre de séries visionnées, etc. mais les informations que nous avons vues sur le tableau de bord de son organisateur corroborent les incohérences dont nous faisons référence. En effet, ce locuteur a terminé 68 séries, quinze sont en cours, sur une période de six mois et une semaine, alors qu'il n'a renseigné que 11 séries au total, pour une durée approximative de 12 heures par semaine. Cette peur pourrait-elle être à l'origine du fait que certaines autres estimations aient été réduites volontairement ? Ou serait-ce une possibilité ? Nous pensons notamment au temps passé à écouter de la musique ou même à visionner des séries. De même, il aurait fallu nous renseigner en détail sur le fait qu'ils regardent tous les épisodes et saisons ou s'ils regardent juste de manière sporadique. Notons que même si nous avons offert l'opportunité de renseigner les autres activités qu'ils pratiquent, il n'y a pas eu d'informations sur les vidéos qu'ils regardent sur YouTube ou les films qu'ils regardent. Aussi, il aurait été pertinent de savoir dans quel mode sont visionnées les séries les plus mentionnées. En d'autres termes, nous aurions voulu voir si les séries qui sont en tête de liste de visionnage sont plutôt regardées en VO, en VOSTA, en VOSTF ou en VF. Établir un lien entre les séries les plus regardées et le mode de visionnage fréquemment utilisé aurait pu nous livrer plus d'informations sur l'impact du visionnage de séries sur la prononciation.

Certes, nous constatons que certains de nos résultats confirment ceux évoqués dans les précédentes études sur l'AIAL. Ainsi, Kusyk (2017) avait montré que la pauvre qualité des doublages constitue une des raisons pour lesquelles les étudiants préfèrent regarder les séries en VO, au risque de s'engager dans des téléchargements illégaux, puisque les seules versions disponibles sont en anglais. Dans le même ordre d'idée, Sockett et Toffoli (2010) ont montré que les étudiants français ne peuvent pas attendre la sortie de la version doublée, à cause du décalage de temps entre la sortie de la VO et celle de la VF, qui peut prendre des semaines. La volonté de suivre les épisodes au fur et à mesure de leurs sorties et ainsi de rester au diapason constitue un phénomène bien connu aujourd'hui chez tous les sériephiles et sérievores. Aussi, les résultats des locuteurs concernant le divertissement reflètent ceux établis dans les études de l'AIAL, notamment celle de Kusyk (2017), et qui montrent la dominance des activités liées au loisir (p.158). Il est vrai que notre étude repose principalement sur les séries télévisées et non sur l'impact de la musique.

Cependant, la réponse de certains locuteurs qui déclarent ne pas écouter de la musique est étonnante. Nous nous demandons si de nos jours, il est possible que certaines personnes n'écoutent pas de la musique en anglais, tellement elle est omniprésente. Serait-ce parce qu'ils en écoutent mais sans réel effort ou attention comme ils le font avec les séries télévisées ? Se pourrait-il que ces locuteurs préfèrent remplacer le temps d'écoute de la musique par le visionnage de séries ? Ces résultats rejoignent les observations de Kusyk qui pense qu'en général, les activités de loisir requièrent un minimum d'attention sauf la musique (2017, p. 155). Il faudrait approfondir ces questions auprès des étudiants concernés afin d'y apporter une réponse adéquate. Si cela s'avère, nous assistons alors à un phénomène nouveau ou un changement de pratiques (voire de paradigme) qui donne un statut particulier aux séries télévisées, parmi les activités des étudiants aujourd'hui.

De plus, sur la question de l'imitation d'accent, il est intéressant de noter le nombre de locuteurs qui disent ne pas volontairement vouloir imiter un accent en particulier. Les raisons évoquées sont assez révélatrices. Ainsi, ils pensent qu'il ne vaut pas la peine d'imiter, car cela est évident qu'ils les intègrent de manière passive, naturelle et sans effort. En revanche, ces mêmes locuteurs sont persuadés qu'ils parlent avec l'accent américain et que d'autres personnes le remarquent. Bien évidemment, les médias américains, notamment les séries y seraient pour quelque chose. L'inverse est aussi très évocateur, car ceux qui pensent parler avec un accent britannique évoquent les cours, l'école et les enseignants comme cause. Ces réponses montrent quelque part qu'ils se rendent compte de l'intégration inconsciente de l'accent dans leur prononciation. Ou bien, ils n'auraient pas envie d'admettre qu'ils imitent intentionnellement un accent en particulier. Serait-il humiliant de l'avouer ? Il se pourrait que ce soit une question d'imaginaire social et de représentation faite par rapport à un accent spécifique (Moyer, 2013). Ou bien serait-ce un sentiment mitigé de fierté d'être ce que l'on est et en même temps de se sentir proche d'autres personnes ? (ici les Américains). Cela se voit dans les raisons évoquées plus haut : « ça ne m'intéresse pas », « parce que je ne sais pas le faire et je n'en vois pas l'intérêt », « le but est de parler pas de faire le clown », etc. Il faut également retenir que si les premières études portant sur l'AIAL considèrent que l'acquisition langagière aux travers des diverses activités informelles se déroulent de manière inconsciente, nos résultats (points sur l'imitation d'accent) indiquent que

certains locuteurs sont quelque peu conscients de leur *uptake* (assimilations). Aussi, savent-ils attribuer leurs accents aux activités sérielles comme nous l'avons remarqué dans leurs réponses ?

Enfin, au sujet de la nationalité des enseignants d'anglais, nous avons pu constater que le nombre de Français est beaucoup plus élevé au lycée et au collège. Les autres nationalités n'apparaissent qu'à l'université, dès la première année de licence. En effet, les ressortissants hors UE n'ont pas l'autorisation de passer le concours du CAPES²⁵³ ou de l'agrégation, qui sont exigés en France pour enseigner dans le secondaire. En revanche, ces concours n'étant pas exigés à l'université, divers natifs anglophones²⁵⁴ sont présents dans les facultés. Nous imaginons que les mêmes remarques ne seraient plus valables dans quelques années, puisque des enseignants anglophones passent de plus en plus ces concours et postulent dans le secondaire. Tel est le constat que nous avons établi dans les annonces de l'académie de Strasbourg par exemple, où des enseignants étrangers sont de plus en plus recherchés pour des postes de titulaires et de remplaçants. Ces constats sont confirmés et relayés dans les médias (radio et télévision) par rapport au manque d'enseignants dans le secondaire en général et spécifiquement les enseignants d'anglais. D'un point de vue didactique, cela implique que contrairement au passé où les enseignants d'anglais étaient en majorité des Français (ou des Britanniques), les élèves et étudiants seront exposés à d'autres types d'accents dans les salles de classe. De ce point de vue, nous allons voir si déjà les apprenants discriminent les mots prononcés avec l'accent britannique et américain, dans la section suivante.

9.8 Résultats du test de perception

Les résultats du test de perception montrent que globalement tous les locuteurs distinguent l'accent américain de l'accent britannique pour les 34 mots prononcés avec les deux accents. En effet, les pourcentages de réponses correctes sont compris entre 74 % et 100 %, soit un minimum de 25 mots sur les écoutes effectuées par chaque locuteur ($\bar{x} = 29$; $\sigma = 2$)²⁵⁵. Lorsque nous prenons les résultats par accent, nous remarquons qu'il n'y a aucune différence significative quant au nombre de réponses correctes, qui est de 14 mots ($\bar{x} = 14,9$; $\sigma = 1,3$ pour les deux accents). Nous en

253 Certificat d'aptitude au professorat de l'enseignement du second degré qualifie à l'enseignement dans les lycées et collèges.

254 Américains, Britanniques, Néo-Zélandais, Canadiens, Indiens, Australiens, et même des non natifs anglophones.

255 Pour rappel nous employons le symbole \bar{x} pour représenter la moyenne et σ pour l'écart-type.

déduisons que la reconnaissance des mots prononcés avec ces deux accents s'effectue sans grande difficulté.

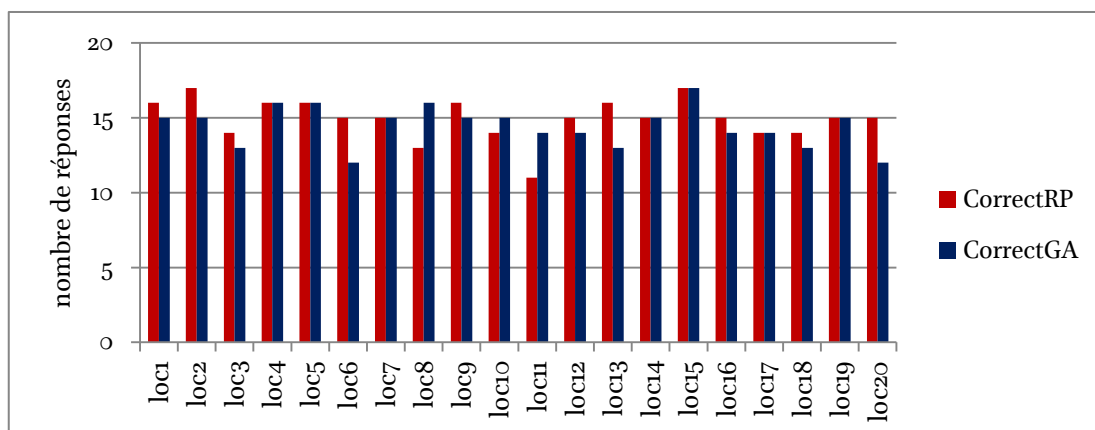


Figure 128 : nombre de réponses correctes par locuteur pour la RP et le GA.

Pour l'ensemble, les locuteurs effectuent une écoute ($\bar{x} = 1,1$; $\sigma = 0,11$) et ils sont généralement très sûrs de leurs réponses ($\bar{x} = 1,3$; $\sigma = 0,19$). Toutefois, lorsque nous prenons les accents séparément, nous remarquons que le nombre d'écoutes est plus élevé et le degré de certitude est moindre concernant les mots prononcés avec l'accent britannique ($\bar{x} = 1,8$; $\sigma = 1,19$ et $\bar{x} = 1,7$; $\sigma = 0,9$ pour la RP contre $\bar{x} = 1$; $\sigma = 0,19$ et $\bar{x} = 1,1$; $\sigma = 0,17$ pour le GA). Ces chiffres dévoilent que les locuteurs écoutent plus de fois les mots prononcés avec l'accent britannique et qu'ils sont moins sûrs par rapport au même accent qu'avec les mots prononcés avec l'accent américain.

Le temps moyen qui s'écoule entre l'écoute et le choix de la réponse est de 2,2 secondes pour l'ensemble des mots ($\sigma = 0,7$). En revanche, pour chaque accent, nous constatons que le temps de réponse est plus long dans le cas de l'accent britannique. Cela révèle que les locuteurs s'attendent plus sur les mots prononcés avec la RP qu'avec le GA, une confirmation des résultats présentés plus haut. Même si les écarts mentionnés jusqu'ici sont peu significatifs statistiquement, celui-ci nous interpelle car il est plus visible. En conséquence, tous les locuteurs, sauf 1 et 17, passent plus de temps à reconnaître les mots prononcés avec l'accent britannique, comme le montre la figure ci-dessous.

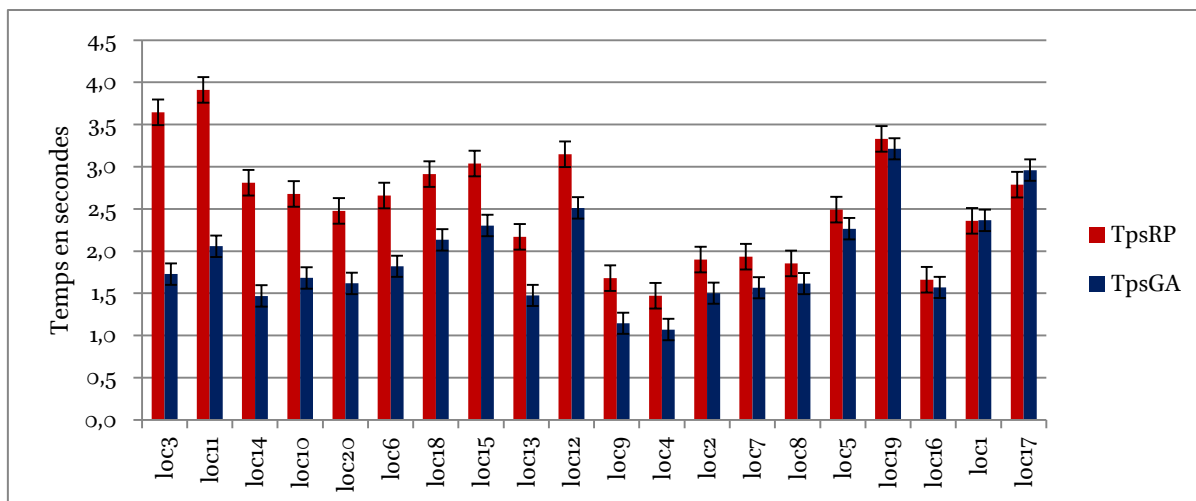


Figure 129 : temps de réponses pour la reconnaissance des mots prononcés avec la RP et le GA.

9.8.1 Discussion

En principe, si nous considérons que les étudiants français sont généralement plus exposés à l'anglais britannique qu'à l'anglais américain, notamment grâce au cadre formel, nous devrions nous attendre à ce qu'ils reconnaissent mieux les mots prononcés avec l'accent RP. Or, ce postulat n'est pas vérifié par nos résultats. Nous pouvons mettre en avant quelques pistes d'explications. D'une part, il se pourrait que l'exposition à l'accent américain aujourd'hui surpasse celle à l'accent britannique et provoque ce type de renversement de tendance. Rindal (2010) montre que l'anglais américain serait plus facile à reconnaître et à parler que l'accent britannique. D'ailleurs, le fait que la RP soit plus proche d'autres accents parlés dans le monde, comme l'anglais australien, néo-zélandais, ou sud-africain par exemple pourrait renforcer cet obstacle de sa reconnaissance. Nous pensons que même en ayant été habitué à l'accent RP, sa diversité et sa proximité avec d'autres types d'accents demanderaient plus de temps de reconnaissance aux locuteurs. De plus, nous soulignons que même si le nombre de réponses correctes, le nombre d'écoutes, le degré de certitude et le temps de réponse constituent des indicateurs robustes de reconnaissance d'accent, il se pourrait que des locuteurs aient besoin de plus de temps afin de choisir les bonnes réponses. Un locuteur pourrait répondre très rapidement parce qu'il ne connaît pas la réponse et qu'il voudrait tout simplement passer à la question suivante. Ces résultats sont donc à prendre avec distance, puisque nous disposons de peu de données approfondies pour les valider (entretien individuel par exemple). Pour l'heure ils nous permettent d'entrevoir les comportements perceptifs des locuteurs à l'égard des deux accents de référence. Par ailleurs, ces résultats que nous

venons de présenter nous serviront dans la triangulation de nos données, que nous abordons dans la section ci-dessous.

9.9 Triangulation de données

Ayant présenté séparément les résultats de la prononciation, du questionnaire et du test de perception, nous allons procéder à une triangulation des résultats de ces trois sources de données. Deux types d'analyses feront l'objet de cette section : le profilage des locuteurs indiquant en quoi ces derniers sont proches ou distants les uns des autres, en fonction des caractéristiques qui leur sont propres et l'exploration des corrélations qui existent entre diverses variables les concernant. L'objectif poursuivi ici est de répondre à nos questions et hypothèses de départ. Nous commençons par le profilage dans la sous-section suivante.

9.9.1 Profilage des locuteurs

Afin de profiler nos locuteurs, nous avons généré dans un premier temps, un dendrogramme ou arbre de classification (figure ci-dessous) à partir d'une Classification Ascendante Hiérarchique (CAH²⁵⁶). Nous accompagnerons la lecture du dendrogramme par un retour aux données car les rapprochements se fondent sur des variables spécifiques qui font que deux individus proches, sur l'arbre, peuvent avoir plus d'affinités avec d'autres individus (éléments partagés).

²⁵⁶ Une méthode statistique basée sur le regroupement d'individus selon des critères spécifiques représentés par des distances interindividuelles.

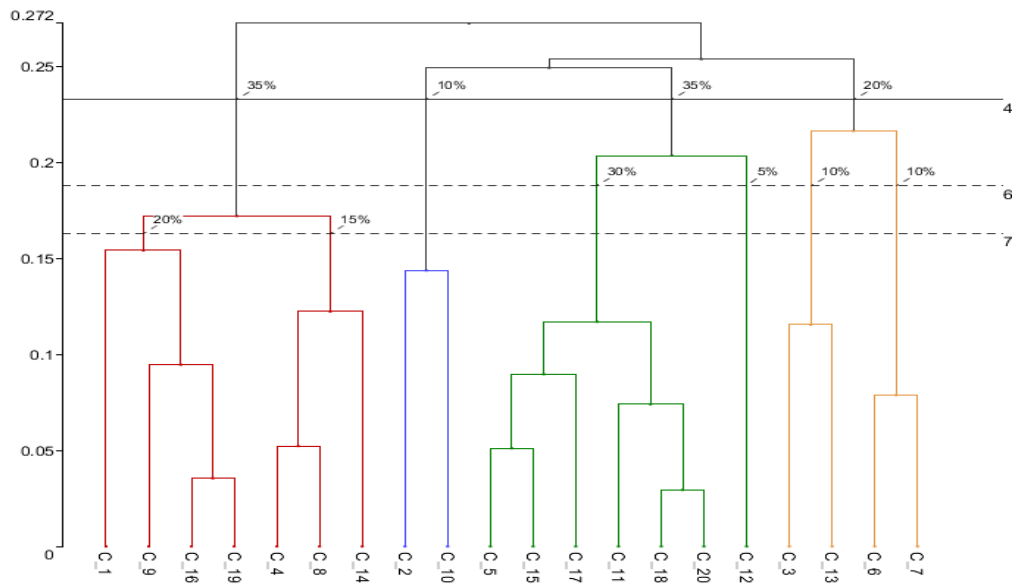


Figure 130 : dendrogramme issu de la Classification Ascendante Hiérarchique.

Cet arbre de classification représente la subdivision de notre échantillon en différentes classes en fonction des liens qu'ils tissent entre eux. Les inscriptions C_x désignent les locuteurs (x pour le numéro du locuteur) et les lignes horizontales représentent les coupures en groupes de ressemblances (4, 6 et 7). Moins il y a de coupures (4), moins nous aurons de groupes, mais plus les relations seront difficiles à analyser (beaucoup trop de variables), et plus il y a de groupes (7), plus la description sera personnalisée (moins de variables). Dans notre présentation, nous allons partir de la ligne pleine (4) jusqu'à la base du diagramme. Ainsi, en partant de la droite du graphique, nous obtenons quatre classes formées : de sept locuteurs (1, 9, 16, 19, 4, 8 et 14, en rouge), du binôme 2 et 10 (en bleu), de sept autres locuteurs (5, 15, 17, 11, 18, 20 et 12, en vert) et enfin de quatre locuteurs (3, 13, 6 et 7, en jaune). Il est important de préciser à ce point, que cette première représentation ne montre qu'une relation basique entre les locuteurs, car certaines variables n'apparaissent pas sur ce schéma, à cause des données manquantes par exemple. Il nous aurait fallu un plan tridimensionnel (3D) voire multidimensionnel pour avoir une meilleure vue. À défaut du 3D, nous avons édité un autre type de graphique à partir d'une analyse qui prend en compte une meilleure représentativité des relations existantes entre les locuteurs. Dans un second temps, nous avons conduit une analyse factorielle, avec des variables du questionnaire, les résultats des productions orales et ceux du test de perception. Nous avons opté pour l'Analyse en Composantes Principales (ACP) qui est une méthode de traitement de données statistiques multivariée

(factorielle) permettant de réduire la quantité de variables corrélées, en vue de rendre les informations importantes plus claires. Elle permet également de mieux représenter la variabilité entre divers facteurs. Afin d’approfondir nos interprétations, nous avons dû manuellement repérer les variables clés dans le fichier des données compilées sous Excel (réponses du questionnaire, les résultats des productions orales et ceux du test de perception).

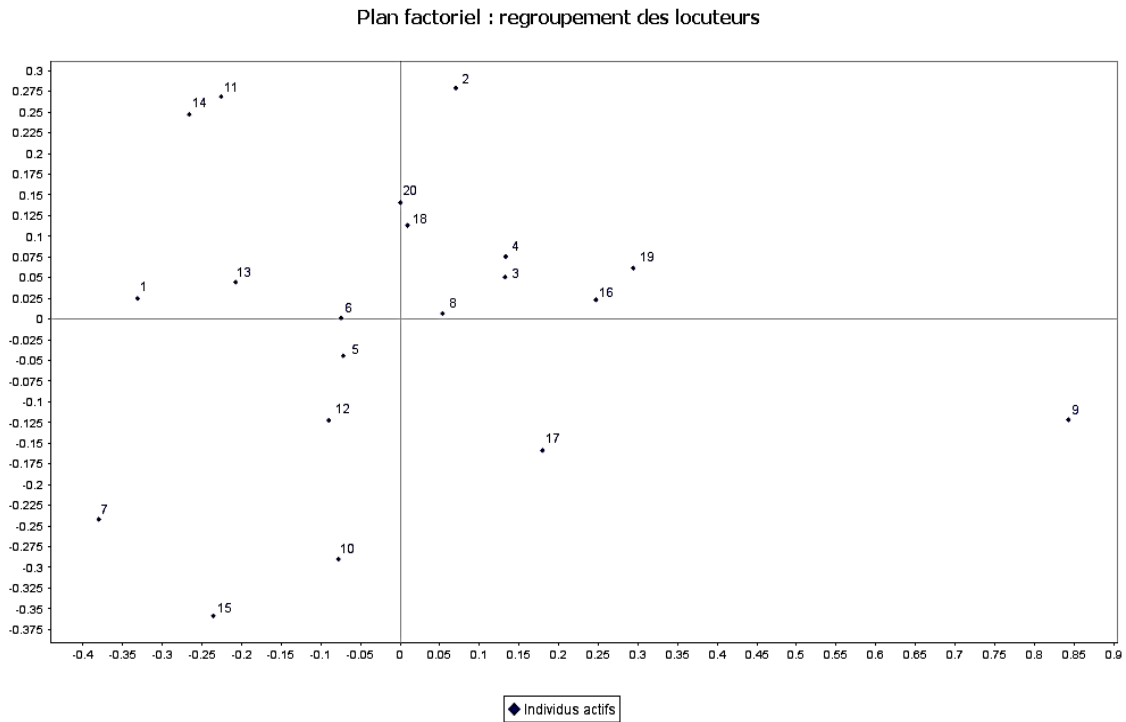


Figure 131 : représentation graphique des relations existant entre les locuteurs.

Ainsi, ce graphique correspondant au plan factoriel ci-dessus, a été obtenu automatiquement en fonction des variables que les locuteurs partagent (ou non) et le degré de rapprochement de ces variables (illustré par la distance entre les points). À première vue, nous remarquons une grande variabilité dans la distribution spatiale des locuteurs. De grandes tendances se dessinent et nous pouvons en retenir quatre. D’abord, nous observons des locuteurs très proches (loc 3 et 4, 18 et 20, 11 et 14, 5 et 6, 16 et 19), ensuite un groupe central formé par les locuteurs 20, 19, 17, 12 et 13 et tous ceux qui se trouvent entre ces derniers (5, 6, 8, 18, 4, 3 et 16). Puis nous avons ceux qui gravitent immédiatement autour de l’agglomérat central (1, 7, 15, 10, 17, 19, 2, 11 et 14) et enfin, un locuteur qui est très excentré des autres (loc9). Remarquons que le rapprochement entre les locuteurs 18 et 20, 16 et 19 sur le dendrogramme est conservé sur ce dernier graphique.

En vue de faciliter la description des profils, nous avons procédé à deux types de codages. Le premier consiste à classer les locuteurs par le pourcentage de certaines variables (quantitatives) et en fonction des valeurs des premiers et troisièmes quartiles²⁵⁷. Nous avons préféré cette méthode statistique en vue non seulement de garantir la validité et afin d'éviter tout découpage fondé sur des valeurs subjectives, mais aussi pour pouvoir comparer les locuteurs entre eux (notre échantillon) et non avec d'autres échantillons d'autres études. En effet, dans la thèse de Kusyk (2017), une durée de sept heures de visionnage de séries télévisées par semaine est considérée comme une participation « fréquente » pendant que deux à trois heures par semaine sont suffisantes pour « un taux fréquent minimum » (Kusyk, 2017, p. 295). Dans d'autres contextes encore, certains pratiquants visionnent à raison d'une à trois heures par jour (Cole et Vanderplank, 2016). Nous pensons que cette question de seuil²⁵⁸ mérite d'être fixée afin de « généraliser » l'usage de certains qualificatifs dans les recherches sur l'AIAL comme le reconnaît Kusyk : « [...] il serait utile de poursuivre une discussion dans de futures recherches sur l'interprétation des participations 'fréquentes' » (p. 295-6). Ainsi, les locuteurs dont les valeurs sont inférieures au premier quartile sont classés dans le groupe « petit », ceux dont les valeurs sont supérieures au troisième quartile appartiennent au groupe « grand » et ceux qui se situent entre les deux sont placés dans le groupe « moyen ». Par exemple, concernant le nombre de réponses correctes de mots prononcés avec la RP (*cf.* test de perception), six locuteurs (3, 8, 10, 11, 17 et 18) sont en dessous du premier quartile ($Q_1=14$) et appartiennent au groupe « petit » (c'est-à-dire qu'ils perçoivent moins bien cet accent), sept locuteurs (1, 2, 4, 5, 9, 13 et 15) sont au-delà du troisième quartile ($Q_3=16$) et sont dans le groupe « grand », tandis que les autres (6, 7, 12, 14, 16, 19, 20) sont assignés au groupe « moyen ».

²⁵⁷ Un quartile est une mesure qui divise une distribution en quatre parties (chacune de 25 %) : ils sont au nombre de trois quartiles Q_1 , Q_2 et Q_3 .

²⁵⁸ Un consensus sur ce qui constitue un grand, moyen, ou petit participant, en fonction de la durée (deux, cinq, dix, trente heures) et de la fréquence (annuel, mensuel, hebdomadaire ou journalier).

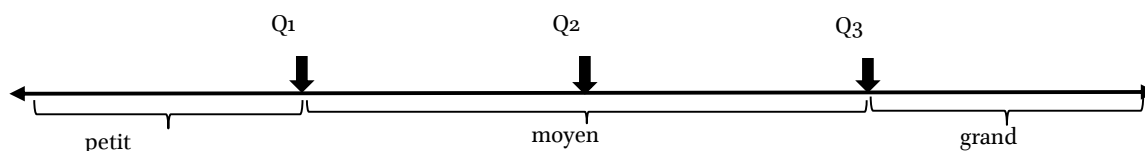


Figure 132 : répartition en groupe selon les quartiles.

Le deuxième codage consiste à classer les locuteurs selon les activités qu'ils pratiquent ou non. Par exemple les locuteurs 5, 6, 7, 13, 14, 15, 16, 18 et 20 ne communiquent pas en anglais en dehors des cours. Sur la base de ces informations, nous avons compilé le tableau ci-après dans l'objectif de faciliter le profilage définitif des locuteurs.

Légende du tableau :

- RPVN : prononciation de caractéristiques proches de la RP en vitesse normale (VN)
- RPVR : prononciation de caractéristiques proches de la RP en vitesse rapide (VR)
- GAVN : prononciation de caractéristiques proches du GA en VN
- GAVR : prononciation de caractéristiques proches du GA en VR
- AuVN : prononciation de caractéristiques « Autres » en VN
- AuVR : prononciation de caractéristiques « Autres » en VR
- GA VNVR : mêmes productions de caractéristiques proches du GA en VN et en VR
- RP VNVR : mêmes productions de caractéristiques proches de la RP en VN et en VR
- Percep GA : perception des mots prononcés avec le GA
- Percep RP : perception des mots prononcés avec la RP
- Recep VO : expo au visionnage en mode VO
- Recep VOSTA : expo au visionnage en mode VOSTA
- Recep VOSTFR : expo au visionnage en mode VOSTFR
- Recep VF : expo au visionnage en mode VO
- +++ : « grand »
- ++ : « moyen »
- + : « petit »
- _ : non applicable

Tableau 31 : Présentation des variables en fonction de la classe des locuteurs (« petit », « moyen » et « grand »).

Locuteurs	RPVN	RPVR	GAVN	GAVR	« Autres » VN	« Autres » VR	GA VNVR	RP VNVR	percep GA	percep RP	recep VO	recep VOSTA	recep VOSTFR
loc1	+++	+++	+	+	+++	+++	+	+++	+++	+++	++	+	++
loc2	+	+	++	++	+++	+++	++	+	+++	+++	+	+++	+
loc3	+	++	+++	++	+	++	+++	++	+	+	++	++	++
loc4	+	++	+++	++	+	+++	+++	+	+++	+++	+++	++	_
loc5	++	++	++	++	++	++	++	+++	+++	+++	++	+++	_
loc6	+++	++	++	++	++	++	+	++	+	++	+	+++	++
loc7	+++	+++	+	+	+	++	+	+++	+++	++	+	+++	++
loc8	+	++	+++	++	++	++	++	++	+++	+	++	++	+
loc9	+	+	+++	+++	+	+	+++	+	+++	+++	+++	++	_
loc10	++	+++	++	++	++	+	++	+++	+++	+	++	+++	_
loc11	++	++	+	+	+++	+++	+	++	++	+	++	++	+++
loc12	++	+++	++	++	++	+	++	++	++	++	++	++	_

loc13	+++	++	++	+	++	+++	+	+++	+	+++	+	+	+++
loc14	+++	++	+	+	+++	+++	+	+	+++	++	++	++	+
loc15	+++	+++	+	++	+	+	+	+++	+++	+++	+++	++	_
loc16	+	+	+++	+++	++	+	+++	+	++	++	+	+	+++
loc17	++	++	++	+++	++	+	++	+	++	+	+++	++	_
loc18	++	+	++	++	++	+++	++	++	+	+	+	++	+++
loc19	++	+	++	+++	++	+	+++	++	+++	++	+++	+	_
loc20	++	++	++	+++	+++	++	++	+	+	++	+	+	+++

Ce tableau montre les variables pour lesquelles les locuteurs sont considérés comme des « grands », « moyens » et « petits » producteurs (prononciation), récepteurs (exposition) ou auditeurs (test de perception). Par exemple, le locuteur 1 est un grand producteur de caractéristiques proches de la RP et « Autres » dans les deux modalités d'élocution, un grand producteur des mêmes éléments pour le même accent, un grand auditeur pour les deux accents de référence, un récepteur moyen du VO et du VOSTFR. À l'opposé, il est un petit producteur d'aspects proches du GA dans les deux modalités de parole, dans la prononciation des mêmes sons en GA et dans la réception du mode de visionnage VOSTA. En continuant avec la classification basée sur la description précédente pour tous les locuteurs, nous obtenons les grandes tendances suivantes :

- Pour les éléments proches de l'accent américain, les grands producteurs sont les locuteurs 3, 4, 8, 9, 16, 17, 19 et 20, et les petits producteurs sont composés des locuteurs 1, 7, 11, 13, 14 et 15 ;
- Concernant les aspects proches de l'accent britannique, les grands producteurs sont les locuteurs 1, 6, 7, 10, 12, 13, 14 et 15, et les petits producteurs sont les locuteurs 2, 3, 4, 8, 9, 16, 18 et 19 ;
- Par rapport aux éléments « Autres », les grands producteurs sont les locuteurs 1, 2, 11, 13, 14, 18 et 20, et les petits producteurs sont constitués des locuteurs 3, 7, 9, 10, 12, 15, 16, 17 et 19 ;
- Les locuteurs 2, 4, 9, 16, 17 et 20 sont des grands producteurs des mêmes éléments du GA mais des petits pour les aspects de la RP, alors que l'inverse s'observe avec les locuteurs 1, 6, 7, 11, 13 et 15 qui sont des grands producteurs en RP et des petits en GA ;
- À propos de la perception, nous retenons les locuteurs 1, 2, 4, 5, 7, 9, 14, 15 et 19 comme étant des grands dans les deux accents et les locuteurs 3, 6, 11, 17, 18 et 20 comme des petits auditeurs des deux accents ;
- Enfin nous considérons les locuteurs 4, 9, 15, 17 et 19 comme étant les plus exposés aux visionnages en VO, les locuteurs 2, 5, 6, 7 et 10 pour la VOSTA, les locuteurs 11, 13, 16, 18 et 20 pour la VOSTFR.

Malgré la difficulté à réaliser un profilage parfait, c'est-à-dire qui prend en compte toutes les variables simultanément, il nous semble important de signaler deux types de profils que nous trouvons intéressants. Le premier concerne les locuteurs 3, 4, 9, 16 et 19 qui sont les plus exposés à

l'anglais, reconnaissent mieux la différence entre les deux accents de référence, prononcent plus avec l'accent américain et les mêmes éléments de cet accent dans les deux modalités de parole. Inversement, ils prononcent moins avec les caractéristiques de l'accent britannique et « Autres ». Le deuxième type de profil est constitué par les locuteurs 16, 18 et 20. Ces derniers s'exposent globalement moins à l'anglais de manière informelle mais ont un niveau relativement élevé dans la production de spécificités semblables à celle de l'accent américain. Le profil de ces locuteurs nous intéresse particulièrement car contrairement aux autres (type 1), leurs seules pratiques d'exposition à de l'*input* de l'anglais informel ne permettent pas d'expliquer leur prononciation dans les deux accents. Un retour sur leurs réponses dans le questionnaire montre surtout qu'ils ont beaucoup de choses en commun : ils ne communiquent pas en anglais, n'écoutent pas de la musique anglophone, ils ne visionnent pas en VO et VOSTA mais seulement en VOSTFR et VF, ils ont tous voyagé aux États-Unis et au Royaume-Uni et seul le locuteur 20 avait un enseignant américain. À ce stade, nous ne disposons pas d'éléments pour expliquer ce constat, mais nous pensons notamment à l'intervention d'autres facteurs que nous n'avons pas pu mesurer.

Nous poursuivons avec le profilage en présentant quelques exemples de variables qui distinguent ou qui unissent certains locuteurs. Même si les locuteurs 3 et 4 partagent énormément de points (groupe « grand » producteurs d'éléments du GA), le locuteur 4 passe plus de temps sur les activités informelles et produit plus d'éléments du GA que le locuteur 3. Aussi, il a eu un meilleur score dans le test de perception que le locuteur 3. Celui-ci n'a jamais eu d'enseignants américain ou parlant avec un accent américain, et n'a jamais visité de pays anglophone. Enfin, contrairement au locuteur 4, le locuteur 3 visionne des séries en mode VOSTFR (à plus de 60 % du temps) et n'imité pas intentionnellement l'accent américain lorsqu'il parle anglais. Concernant les deux locuteurs 18 et 20, ils regardent les séries en VOSTFR et en VF (plus de 80 % du temps) mais le locuteur 20 complète ses visionnages avec du VOSTA. Dans les activités de production et de perception, le locuteur 20 est plus exposé à l'anglais de manière formelle et produit plus de caractéristiques de la RP et « Autres » que le locuteur 18. Dernière présentation mais non la moindre, les deux locuteurs 2 et 15 se situent complètement aux antipodes du graphique, ils ont beaucoup plus d'éléments de divergence que les relations présentées ci-dessus. Le locuteur 15 fait partie des grands producteurs d'éléments proches de la RP, moyen producteur pour le GA et petit producteur pour les autres. Il n'y a pas de différences majeures pour les résultats du test de

perception sauf le fait que le locuteur 15 répond moins vite aux questions. Nous remarquons que le locuteur 15 ne communique pas en anglais en dehors des cours mais qu'il est beaucoup plus exposé à l'anglais américain (90 % VO, 10 % VOSTA) que le locuteur 2. En revanche, le locuteur 2 n'a jamais séjourné dans un pays anglophone, contrairement au locuteur 15 qui est allé deux fois au Royaume-Uni.

En somme, nous retenons qu'il est très difficile de procéder à un profilage concernant 20 individus, mais cela nous a tout de même aidé à mieux catégoriser nos locuteurs par rapport à certaines variables importantes. Nous sommes limité par rapport au profilage des locuteurs car nous n'avons pas pu constituer des groupes au-delà de deux ou trois individus. Comme nous l'avons précisé, cette limite est due à la taille de notre échantillon. Vouloir forcer l'élargissement des groupes nous ferait courir des risques de pertes d'informations ou de répétitions des mêmes facteurs partout. Nous avons vu les trois classifications de producteurs d'un accent spécifique, de récepteurs et d'auditeurs (« petits », « moyens » et « grands »), mais nous n'avons pas été en mesure d'expliquer leurs productions en se fondant uniquement sur les réponses du questionnaire. Parallèlement au profilage, dans la sous-section suivante, nous allons présenter les résultats des tests de significativité menés pour quelques variables, en vue de mieux visualiser les types de relations existantes entre elles.

9.9.2 Corrélations entre les variables

Étant donné la taille de notre échantillon et le fait que nos mesures soient non gaussiennes²⁵⁹, nous avons opté pour un test statistique non paramétrique pour les valeurs de corrélations entre diverses variables sélectionnées. Le choix de ces données est porté sur leur pertinence à apporter des éléments de réponse à nos questions et hypothèses de recherche. Les variables que nous avons écartées sont celles que tous les locuteurs partagent ou qui n'auront de toute manière aucune incidence sur les résultats statistiques (les données manquantes, données biographiques, nombre de séries, les loisirs). Nous nous sommes servi du logiciel de traitements statistiques SPAD²⁶⁰ pour

²⁵⁹ Qui ne suivent pas une distribution en forme de cloche (*cf.* loi normale).

²⁶⁰ SPAD est un logiciel payant d'analyses statistiques créé par la société française Coheris et mis à disposition gratuitement par l'université de Strasbourg (version 9.1), sur la plateforme applis.unistra.fr et accessible via la Connexion bureau à distance).

tous nos tests statistiques et nous avons utilisé le coefficient de corrélation r de Spearman²⁶¹ sur les variables concernant tous les locuteurs, que nous présentons dans le tableau suivant.

Tableau 32 : variables utilisées dans les tests de corrélation pour la triangulation des données.

Production	Perception	Questionnaire
Productions proches de la RP en vitesse normale	Réponses correctes RP Réponses correctes GA	Temps de visionnage de séries américaines par semaine
Productions proches du GA en vitesse normale	Temps de réponse RP Temps de réponse GA	
Productions « Autres » en vitesse normale	Degrés de certitudes RP	% de VO uniquement
Productions proches de la RP en vitesse rapide	Degrés de certitudes GA Nombre d'écoute RP	% de VOSTA
Productions proches du GA en vitesse rapide	Nombre d'écoute GA	% de VOSTFR
Productions « Autres » en vitesse rapide		Accent britannique des enseignants
Mêmes productions proches de la RP en VNVR		Accent américain des enseignants
Mêmes productions proches du GA en VNVR		Accent français des enseignants

Les résultats des analyses montrent des corrélations statistiques qui varient de très significatives pour certaines variables, à aucune relation pour d'autres. Dans ce qui suit, nous expliquerons les relations intervariables tout en présentant les valeurs obtenues pour le coefficient de corrélation de Spearman (r) suivi de la valeur de la probabilité (p). Afin de faciliter la compréhension de ces valeurs, nous rappelons leur signification. La probabilité p , aussi appelée degré de signification, permet d'indiquer s'il y a ou non une corrélation entre deux variables mesurées. Conventionnellement, un seuil de 5 % est utilisé, mais encore une fois, la taille de notre échantillon nous conduit à choisir un taux de 1 %, car il est plus rassurant de considérer qu'il y a moins de 1 % de chance qu'une relation soit due au hasard (risque d'erreur inférieur à 1 sur 100) que 5 % (5 chances sur 100) de se tromper. En adoptant un seuil de significativité de 1 %, nous garantissons la significativité de la relation (à un seuil très fort). Ainsi, lorsqu'une valeur de p est supérieure ou égale à 1 % (0,1), nous concluons que la relation entre deux facteurs n'est pas statistiquement significative. En revanche, si cette valeur est inférieure à 1 %, la relation est considérée comme statistiquement significative et on admet les cas de figure suivants :

²⁶¹ Il est plutôt adapté pour les échantillons de taille faible (moins de 30).

Tableau 33 : signification des valeurs de p.

Seuil de significativité (probabilité)	Relation
> 0.10	non significative
0.05 à 0.10	peu significative
0.01 à 0.05	significative
0.001 à 0.01	très significative
< 0.001	extrêmement significative

Quant au coefficient de corrélation, il indique la puissance et le sens de la relation existante entre deux variables : la corrélation est nulle (ou neutre) lorsque sa valeur est égale à 0 (ou tend plus vers 0) et elle est très forte lorsque sa valeur est égale à plus ou moins 1 (ou tend vers +/-1). Il en résulte quatre grandes tendances résumées dans le tableau suivant :

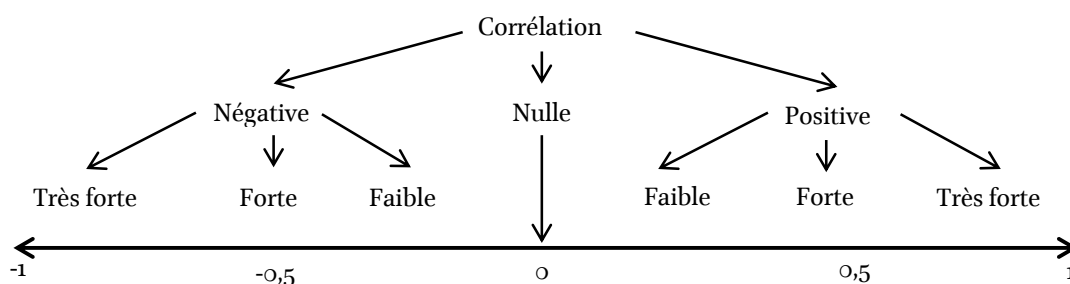


Figure 133 : signification des valeurs de r.

Par corrélation, le lecteur doit comprendre une relation et non une causalité. Il peut y avoir une relation extrêmement significative entre deux variables sans que l'une soit la cause de l'autre. La cause peut se trouver ailleurs, dans d'autres variables internes ou externes. De même, dans la présentation et l'interprétation des résultats de corrélation, nous restons prudent quant à toute explication de cause à effet. Même si des causalités sont possibles voire sous entendues entre des variables, nous éviterons l'emploi des verbes influencer, causer, provenir, engendrer et entraîner, car cela relève de l'interprétation.

Tout d'abord, nous constatons un lien significatif entre les variables du temps de visionnage des séries américaines (TVSA) et la prononciation avec des éléments proches de l'accent américain. Cette corrélation est plus forte en vitesse normale (VN) ($r = 0,529$, $p = 0,008$) qu'en vitesse rapide (VR) ($r = 0,518$; $p = 0,01$). En revanche, le TVSA est négativement lié aux productions proches de la RP (VN : $r = -0,4$; $p = 0,04$ VR : $r = -0,403$; $p = 0,039$) et des autres productions (VN : $r = -0,423$; $p = 0,031$. VR : $r = -0,297$; $p = 0,02$). Ces résultats indiquent que lorsque le TVSA augmente, les productions proches du GA augmentent pareillement, mais les productions de la RP et « Autres »

diminuent. Nous pouvons dire que les locuteurs qui regardent plus les séries télévisées américaines, parlent plus l'anglais avec des caractéristiques de l'accent américain, et moins avec les éléments de la RP et « Autres » (y compris l'accent français). Nous n'avons pas inclus les variables sur les séries télévisées britanniques car tous les locuteurs ne les visionnent pas, et ceux qui s'y adonnent y passent moins de temps comparé aux séries américaines.

Une autre suite de relations significatives est perçue entre les variables de la prononciation et les modes de visionnages des séries. Nous constatons un fort lien entre la production en VR d'éléments de prononciation du GA et le pourcentage de visionnage des séries en VO uniquement ($r = 0,631$; $p = 0,001$), et une relation non significative en VN ($r = 0,217$; $p = 0,179$). Comme dans le cas précédent, nous avons enregistré des valeurs négatives de r pour les deux prononciations « Autres » (VN : $r = -0,422$; $p = 0,032$. VR : $r = 0,460$; $p = 0,021$) et la production RP en VR ($r = 0,421$; $p = 0,032$). En outre, cette variable n'a montré aucun rapport significatif avec les productions en VN de la RP et du GA ($p = 0,417$; $p = 0,179$ respectivement). Ces résultats signifient que les locuteurs qui parlent l'anglais avec plus d'éléments de l'accent américain en vitesse rapide sont ceux qui utilisent le plus le visionnage uniquement en VO. De plus, ces mêmes locuteurs sont enclins à moins prononcer les mots anglais avec des aspects de l'accent britannique et des autres caractéristiques. Nous pouvons postuler que le mode de visionnage en VO seul permettrait de parler l'anglais avec l'accent américain, même en condition de perturbation (augmentation de la vitesse d'élocution dans notre cas). Cependant, partons du fait que le visionnage en VO et la prononciation en VN avec l'accent américain n'aient pas de lien fort, cela pourrait expliquer que ceux qui regardent plus en VO possèdent déjà un bon niveau dans la prononciation de cet accent.

Au niveau de la perception, nous notons que le temps de réponse en vue de reconnaître les mots prononcés avec l'accent américain est fortement associé au nombre d'écoutes de ce même accent ($r = 0,572$; $p = 0,004$). Ceci implique que les étudiants qui écoutent le moins de fois les mots prononcés avec l'accent américain sont nombreux à trouver les réponses correctes en peu de temps. Une relation similaire mais faible est constatée concernant les mots prononcés avec l'accent britannique ($r = 0,417$; $p = 0,034$). Ensuite, un lien négatif et faible existe entre le temps de réponse et le degré de certitude dans l'identification des réponses correctes pour les deux accents (RP : $r = -0,469$; $p = 0,018$. GA : $r = -0,560$; $p = 0,011$), ce qui signifie que plus les locuteurs sont sûrs

d'eux-mêmes dans la reconnaissance des deux accents, moins ils y consacrent de temps. Ce résultat nous semble cohérent.

Le temps de visionnage des séries américaines par semaine est positivement corrélé au mode du visionnage en VO uniquement ($r = 0,309$; $p = 0,092$). Ces résultats indiquent que plus le temps de visionnage est grand, plus le pourcentage de visionnage en VO est important. D'ailleurs, comme évoqué plus haut, nous pouvons également faire le lien entre ces éléments et le niveau de prononciation de l'anglais américain, dans la mesure où nous considérons que les étudiants qui visionnent beaucoup les séries télévisées américaines ont plus tendance à privilégier le mode VO uniquement, en raison de leur niveau suffisant pour comprendre les discours sans les sous-titres. Nos interprétations vont dans le même sens que les remarques de Kusyk et Sockett (2012) selon lesquelles les étudiants français regardent les séries en VO en raison de leur disponibilité (les versions téléchargeables en ligne sont souvent en VO) et du fait qu'ils préfèrent suivre les séries dès leur sortie.

Comme prévu, la quantité d'exposition aux séries est fortement corrélée à la prononciation des mêmes aspects dans les deux conditions d'élocution (vitesse normale - vitesse rapide). Ces valeurs viennent confirmer les informations que nous avons déjà indiquées dans la section 9.9.2 (page 348) concernant les proportions de sons viables. En effet, la variable temps de visionnage de séries américaines en VO/VOSTA est positivement corrélée à la variable des mêmes productions proches du GA en VNVR ($r = 0,71$; $p = 0,008$) et négativement liée à la variable des mêmes productions proches de la RP en VNVR ($r = -0,577$; $p = 0,021$). Ceci indique que les locuteurs qui prononcent le plus les mêmes spécificités de l'accent américain dans les deux vitesses de parole sont aussi ceux qui sont le plus exposés aux *inputs* de l'anglais américain *via* les séries. Inversement, les locuteurs qui pratiquent moins d'activités informelles prononcent davantage selon la norme britannique dans les deux conditions de parole.

Une seule relation significative a été constatée entre les mesures de la prononciation et celles de la perception des mots. Il s'agit d'une corrélation négative plutôt faible entre la prononciation proche de l'accent américain et le temps de réponse pour les mots prononcés avec l'accent britannique ($r = -0,323$; $p = 0,082$). Cela indique que les locuteurs qui prononcent avec des éléments de l'accent du GA sont à même de vite reconnaître les mots prononcés avec l'accent RP. Ces résultats pourraient s'expliquer par le fait que la capacité à produire des sons proches du GA

permettraient aux locuteurs de mieux catégoriser (et donc d'écarter) les prononciations qui ne relèvent pas du même accent. Ces résultats nous semblent contradictoires à première vue, car nous nous attendions à ce que la perception d'un accent spécifique facilite la prononciation de ce même accent. Nous signalons ainsi la confirmation du *Perceptual Assimilation Model* (PAM de Best 1995) qui pose que la perception est une condition nécessaire mais non suffisante pour la production des sons en L2.

Nous venons de présenter les liens entre quelques variables importantes et leur degré de significativité. Alors que certaines variables sont extrêmement liées, d'autres le sont faiblement ou pas du tout. Les relations identifiées entre les variables, couplées avec différents profils liés aux pratiques de chaque locuteur nous donnent une idée de la nature des connexions, certes chaotiques, qui s'établissent entre les différents facteurs du système. Par conséquent, il nous faut reconnaître la difficulté à déterminer les tendances fortes avec précision. Toutefois, les quelques indications de corrélations présentées plus haut nous permettent d'entrevoir des pistes de réponses à nos questions de recherche que nous allons traiter dans le chapitre suivant.

9.10 Conclusion

Nous venons de présenter les résultats du questionnaire que les participants ont rempli à la fin de la phase de collecte des données principales. Dans le chapitre méthodologie, nous avons exposé cette partie en trois catégories, à savoir le contact avec la langue anglaise, les séries télévisées et l'adoption d'accent. Mais, pour des raisons de clarté et de cohérence, ces parties sont éclatées en plusieurs sous-sections traitant d'aspects particuliers présentés dans le questionnaire. Par exemple, nous citons les informations biographiques, les séjours et les durées dans les pays anglophones, le mode et les raisons de visionnage, les séries télévisées visionnées, les accents, les centres d'intérêt, la nationalité et l'accent des enseignants d'anglais, etc. En effet, ces différents points nous ont permis de passer en revue les facteurs clés qui définissent le fonctionnement des pratiques de l'AIAL. Aussi, ces quelques points nous ont aidés à entrevoir la nature complexe et dynamique du système dans lequel les activités sont pratiquées et pour lesquelles les étudiants sont non seulement des acteurs « passifs » mais en même temps des acteurs « actifs ». Passifs car le système leur impose des influences qu'ils ne peuvent contrôler et actifs puisqu'ils exercent en retour des influences sur le système, qui se modifie et se réorganise grâce à l'interaction de tous les facteurs

que nous avons pu identifier (ou non). L'influence mutuelle de ces facteurs est variable et s'observe à divers degrés. Certaines modalités ayant plus d'influences que d'autres (séries télévisées contre musique), d'autres facteurs sont dominants (accent américain contre accent britannique ou encore la nationalité et l'accent des enseignants d'anglais). Nous avons également discuté de quelques aspects concernant les résultats, en vue d'apporter des éléments d'explication qui ne figurent pas directement dans la présentation des résultats du questionnaire. À travers les résultats du questionnaire, nous avons pu observer une variation dans les réponses et donc dans les pratiques de nos locuteurs. Mais, malgré cette différence individuelle, nous avons noté des tendances fortes et des points communs pour certains aspects présentés.

En somme, nous retenons quelques points forts des résultats présentés plus haut. Tout d'abord, nos locuteurs sont des étudiants âgés de 18 à 33 ans et sont des spécialistes d'autres disciplines que l'anglais qui s'engagent dans des pratiques informelles variées en ligne, entre 4 et 24 heures en moyenne par semaine. Ces étudiants sont également exposés à l'anglais de manière formelle, notamment à travers leur scolarité, mais la proportion, en termes de quantité d'heures et de fréquence d'exposition, des activités informelles est beaucoup plus importante. De plus, dans le cadre de leurs études, ces étudiants sont exposés à 80 % à l'accent britannique et français, étant donné la nationalité de leurs enseignants. En moyenne, un apprenant assiste à deux heures de cours d'anglais à l'école contre 10 heures d'exposition à cette langue de manière informelle. Nous avons également noté que les séries télévisées, surtout américaines, constituent la pratique dominante et seraient le facteur qui impacte le plus la prononciation de certains locuteurs. En effet, alors que les locuteurs 2, 3 et 11 imitent consciemment l'anglais américain, les autres affichent des éléments émanant de l'imitation inconsciente. Ensuite, considérant les résultats du test de perception, il s'avère que tous les étudiants savent distinguer les deux accents de référence à 80 %, même si la reconnaissance de l'accent britannique pose un peu plus de difficultés. Enfin, nous avons pu esquisser un profilage de notre échantillon, en catégorisant les groupes d'individus qui se ressemblent et les facteurs qui les rassemblent. Ainsi, les locuteurs 3, 4, 8, 9, 16, 17, 19 et 20 sont beaucoup plus proches concernant les variables de production de l'accent américain et les modes de visionnage en VO que les locuteurs 7, 11, 13, 14 et 15. Aussi avons-nous mis en évidence des relations statistiquement significatives entre certaines variables quantitatives importantes comme le temps d'exposition aux séries télévisées, le mode de visionnage des séries, la prononciation avec

des éléments proches de l'accent américain dans les deux vitesses de production de parole et la capacité à distinguer les deux accents.

Nous allons clore ce chapitre sur les résultats du questionnaire, du test de perception et de la triangulation des données, que nous avons présentés de manière décomposée afin d'en faciliter la compréhension. Ces éléments nous serviront dans le prochain chapitre à synthétiser nos résultats à la lumière de notre contexte d'étude, à répondre à nos questions et hypothèses de recherche et à établir des liens avec les théories employées. Dans ce dernier chapitre, nous présenterons également les apports de notre étude, ses limites et nos perspectives de recherches futures.

10 Synthèse de la recherche et conclusion générale

Conclure c'est toujours quelque part recommencer, reboucler la boucle démarrée tantôt et construire un apprentissage en spirale (Toffoli, 2019, p. 165).

S'inscrivant dans la continuité des travaux sur l'apprentissage informel de l'anglais en ligne, qui se donne pour objectif l'étude de ses conséquences sur le développement langagier des apprenants, notre recherche nous a permis d'analyser des éléments de prononciation d'un échantillon d'étudiants LANSAD et d'émettre des hypothèses quant à leurs origines. Cette partie a pour objectif de discuter des différents résultats obtenus dans ce travail, au regard des questions et hypothèses que nous avons formulées au départ et aussi d'établir un lien avec les divers concepts théoriques mentionnés dans notre état de l'art. Il s'agira alors de confirmer ou d'infirmer ces hypothèses au vu de nos données et d'effectuer un retour sur les trois grandes théories évoquées afin d'examiner la validité conceptuelle de leurs postulats, en prenant en compte notre échantillon. Aussi, nous débattons des contributions de ce travail et aborderons les limites que nous avons identifiées, en suggérant conjointement des pistes d'amélioration pour nos travaux futurs. Nous débutons par la section concernant les questions et hypothèses de départ.

10.1 Retour sur les questions et hypothèses de départ

Nous allons à présent répondre aux deux questions de recherche que nous avons posées, en validant ou en rejetant les hypothèses correspondantes invoquées dans le chapitre I (section 1.5), à la lumière des résultats obtenus. Les réponses seront données à la suite du rappel de chaque question et hypothèse.

À la première question qui est : « quelles prononciations adoptent les étudiants français qui pratiquent l'AIAL ? », nous avons émis les hypothèses suivantes :

1. Nos locuteurs parleraient l'anglais avec des mélanges de caractéristiques de l'accent britannique, de l'accent américain et des éléments phonético-phonologiques de la langue française. En effet, ces locuteurs sont en contact avec l'anglais britannique, l'anglais américain, le français et leur L1 ;

Cette hypothèse est entièrement validée par les résultats de la prononciation des sons et séries de sons étudiés. Comme nous l'avons montré, nos locuteurs produisent des éléments caractéristiques semblables à ceux de la RP et du GA mais aussi des éléments ne ressemblent à

aucun de ces deux accents de référence. Nous avons postulé qu'il s'agissait d'un mélange de spécificités des divers accents auxquels ces locuteurs sont exposés dans le cadre de situations d'apprentissage, aussi formelles qu'informelles. Nous avons également évoqué le fait que ces éléments mixtes, que nous avons appelés « interprononciation » proviendraient du système phonético-phonologique de la langue française ou d'autres langues.

2. Les étudiants français imiteraient des accents pour des objectifs bien spécifiques.

Considérant que seulement trois locuteurs admettent imiter volontairement l'accent américain, nous ne pouvons nous prononcer sur la validité de cette hypothèse. En revanche, les raisons que ces derniers ont mises en avant semblent aller dans le sens de notre postulat (adopter ce qu'ils considèrent être la bonne prononciation).

3. Les étudiants qui sont exposés en permanence aux activités informelles, avec des *inputs* de l'accent américain, auraient plus d'éléments de prononciation de l'accent américain. Ceux qui sont les moins exposés aux *inputs* américains auraient plus d'éléments de prononciation de l'accent britannique.

La forte variabilité relevée parmi nos locuteurs concernant leurs différentes pratiques d'exposition à l'anglais et leurs différentes prononciations ne nous permet pas de confirmer cette hypothèse. En effet, comme l'ont montré nos résultats, tous ceux qui sont les plus exposés aux *inputs* de l'accent américain ne sont pas des « grands producteurs » d'éléments du même accent. L'inverse est également valable, car certains locuteurs sont plus exposés à l'anglais britannique mais prononcent les mots avec beaucoup plus d'aspects proches du GA. En revanche, comme nous l'avons vu, la majorité des « grands » producteurs d'éléments proches du GA sont également de « grands » récepteurs de cet accent (locuteurs 4, 9 et 19) et les « grands » producteurs de caractéristiques proches de la RP sont de « petits » récepteurs de l'anglais américain (locuteurs 1, 6, 7 et 13). En nous reposant sur ces résultats, nous pouvons affirmer que cette hypothèse est partiellement validée.

4. Les apprenants conserveraient leur « accent français », malgré la forte exposition aux *inputs* multimédias anglais, du fait de leurs pratiques informelles.

Cette hypothèse aurait été approuvée si la prononciation des mots avec les éléments « Autres » était prédominante par rapport à ceux de la RP ou du GA. Or, les données montrent que le pourcentage « Autres » ne dépasse pas les 15 % de l'ensemble des productions de chaque locuteur. Ils ont tous produit un faible pourcentage d'éléments avec un « accent français » indépendamment de la quantité d'exposition aux *inputs* de l'anglais. Même ceux qui s'exposent le moins aux *inputs*

anglais prononcent les sons avec une dominance d'éléments proches de l'anglais britannique ou de caractéristiques proches de l'accent américain. Cette hypothèse est donc démentie.

5. Les apprenants qui sont les plus exposés à de l'*input* de l'anglais américain distingueraient mieux les mots prononcés avec le même accent tandis que ceux qui y sont les moins exposés reconnaîtraient mieux les mots prononcés avec l'accent britannique.

Nous avons noté que tous les locuteurs distinguent les mots prononcés avec l'accent britannique au même titre que les mots prononcés avec l'anglais américain, quelle que soit la quantité de leur exposition à l'un ou l'autre des deux accents. Il n'y a donc pas eu de différence significative entre les « grands consommateurs » de séries télévisées américaines et les « petits consommateurs ». En se fondant sur ces données (*cf.* section 9.9.2, p. 348), nous rejetons cette hypothèse dans sa globalité.

6. Les locuteurs qui s'adonnent énormément aux activités de visionnage de séries télévisées américaines produiraient davantage les mêmes sons caractéristiques de cet accent en vitesse normale et rapide que ceux qui s'y exposent le moins.

Pendant la présentation des résultats concernant les productions en vitesse normale (VN) et en vitesse rapide (VR), nous avons relevé que certains locuteurs prononcent plus les mêmes sons que d'autres. Par exemple, les locuteurs 3, 4, 9, 16 et 19 produisent entre 34 et 72 % de caractéristiques proches de l'accent américain alors que les locuteurs 1, 6, 7, 11, 13, 14 et 15 produisent ces mêmes éléments à moins de 18 %. Or le premier groupe a été caractérisé comme étant les « grands » consommateurs d'activités informelles (temps de visionnage, pourcentage de VO, VOSTA, VOSTFR, etc.) et le deuxième groupe faisant partie des « petits » consommateurs pour les mêmes variables. Compte tenu de l'étroite relation entre la quantité d'exposition à l'accent américain des locuteurs et leur prononciation constante de ce même accent, nous pouvons conclure sur la confirmation de notre présente hypothèse. Ainsi, l'exposition massive aux *inputs* de l'accent américain pourrait entraîner la résistivité²⁶² des trajectoires articulatoires chez ces locuteurs (« grands consommateurs ») et inversement. Nous retenons que la perturbation (augmentation de la vitesse d'élocution) a peu d'impacts sur la prononciation d'éléments proches de l'accent GA. Néanmoins, nous soulignons les exceptions constatées chez les locuteurs 15 et 17 qui sont de « grands consommateurs » de séries en VO et VOSTA et pourtant ils ont très peu de productions homogènes de l'accent GA dans les deux conditions d'élocution.

²⁶² Voir Vaxelaire, 2007 (page 162 section 5.5 sur la perturbation).

L'hypothèse émise à l'égard de la deuxième question « quelles peuvent être les origines de ces éléments de prononciation ? » est la suivante :

7. Les séries télévisées constitueraient les activités dominantes, parmi les pratiques de l'AIAL de nos locuteurs, et influeraient ainsi plus l'intégration de l'accent.

Nous avons noté que les séries télévisées américaines constituent environ 70 % de toutes activités de nos locuteurs, les séries télévisées britanniques et l'écoute de la musique sont loin derrière (14% et 13 % respectivement). Aussi, le nombre et la quantité de séries visionnées sont les conséquences de l'intérêt suscité par l'histoire et les actions. Aussi, les résultats du questionnaire montrent que les éléments de prononciation observés dans les productions de nos locuteurs sont issus de sources diverses. Nous avons relevé l'apprentissage formel de l'anglais y compris l'accent des professeurs d'anglais, les discussions en anglais en France ou pendant les séjours en pays anglophones, les séries télévisées, la musique etc. comme étant les expositions qui influenceraient leur prononciation. Toutefois, la comparaison de la quantité de chaque exposition nous permet de déduire que le cadre institutionnel (cours et accents des enseignants prioritairement) constitue la source principale d'assimilation d'éléments de l'anglais britannique. En effet, nous avons noté que la moyenne du temps d'apprentissage formel de l'anglais s'élève à 12 ans (entre 5 à 23 ans), environ 90 % des enseignants d'anglais sont des Français et des Britanniques qui parlent en majorité avec un accent britannique (53 %) ²⁶³ contre seulement 6 % d'enseignants américains et uniquement 6 % des séries télévisées contiennent l'accent britannique. À l'inverse, nous avons noté que près de 80 % de l'*input* provenant des activités informelles sont en anglais américain (séries télévisées, films, musique, etc.). En se fondant sur ces données, nous confirmons cette hypothèse dans la mesure où les apprenants intégreraient l'accent américain en conséquence de leur engouement pour les séries américaines, ou peut-être de manière incidente.

Sur les sept hypothèses que nous avons énoncées, trois sont confirmées, deux sont partiellement validées et deux sont démenties. Ces différents éléments de réponse nous ont permis, non seulement de répondre à nos deux questions de recherches, mais également nous offrent des pistes pour des recherches futures, notamment dans l'exploitation des retombées de cette étude, pour des propositions d'ordres pédagogique et didactique. Nous rappelons que les conclusions que nous tirons concernent uniquement notre échantillon et ne peuvent en aucun cas

²⁶³ Selon la perception des étudiants.

être généralisées. Aussi, gardons-nous à l'esprit que nous nous situons dans un système complexe et dynamique qui implique des changements dans le temps et qui fait qu'aucune condition n'est permanente. Deux variables fortement liées ou une hypothèse qui est validée aujourd'hui ne le seront peut-être plus demain. Pour l'instant, en ce qui nous concerne, il est important de pouvoir établir des liens entre nos résultats avec des différentes théories évoquées et voir en quoi il y a accord ou désaccord. C'est ce que nous abordons dans la section suivante.

10.2 Discussion des résultats en lien avec le cadre théorique

Dans cette partie, la discussion des résultats suit exactement l'ordre de présentation des théories dans l'état de l'art. Il s'agit tout d'abord de la théorie des systèmes complexes et dynamiques, puis de la théorie de l'imitation vocale, et enfin la théorie de la viabilité et perturbation en production-perception de la parole.

10.2.1 Systèmes complexes et dynamiques

Concernant les théories des systèmes complexes²⁶⁴, nous avons vu que dans l'apprentissage des langues, l'émergence de formes linguistiques est le résultat de l'interaction entre différents composants internes et externes à l'apprenant. Ainsi, les réponses au questionnaire indiquent que les différents types de prononciation observés chez nos locuteurs (les caractéristiques proches de l'accent britannique, de l'accent américain et d'autres accents) proviendraient de diverses sources d'exposition à de l'*input* sonore. Il y a sûrement des connexions complexes entre les activités informelles (les voyages dans les pays anglophones, les interactions directes et indirectes en anglais en France, la musique, les séries télévisées, les films et autres vidéos en anglais et les jeux vidéo) et les activités formelles (les enseignants d'anglais, les matériels audiovisuels utilisés dans le cadre des cours). Dans cette visée systémique, nous soutenons que l'interaction entre ces activités est à l'origine de l'émergence de la prononciation mixte en anglais de notre échantillon. Il est par ailleurs important de rappeler que l'interprononciation représentée par les productions « Autres » et les variations inter/intra-individuelles présentes dans nos résultats témoignent de la complexité inhérente au système dans lequel évoluent nos locuteurs. Nous considérons ainsi le développement de la prononciation de nos locuteurs comme étant dynamique, non linéaire,

²⁶⁴ Voir section 3,5, p. 85.

instable et émergeant, sous l'impulsion de ces facteurs externes et internes. Les résultats du profilage en témoignent puisque les locuteurs sont très proches par certaines variables et très éloignés pour d'autres. Ils n'ont donc pas les mêmes conditions initiales et n'auront pas les mêmes conditions « finales ».

Comme nous venons de le mentionner, les diverses possibilités liées à la provenance des caractéristiques d'accents observées chez nos locuteurs confirment le caractère éminemment complexe de leurs pratiques et nous limitent dans toute recherche et élaboration de facteurs de cause à effet. En revanche, la quantité et la fréquence d'exposition de certaines activités laissent percevoir le rôle important qu'elles jouent en tant qu'attracteurs puissants dans l'assimilation d'éléments de prononciation. En effet, nous pouvons en déduire que les séries télévisées américaines représentent l'une des sources qui influencent le plus la prononciation, parce que leur visionnage constitue l'activité la plus pratiquée en termes de volume d'heures. En outre, les relations de corrélation que nous avons présentées vont dans le sens des relations complexes entre les diverses variables issues de la triangulation des données. Comme nous l'avons indiqué plus haut, une relation statistiquement significative n'implique pas forcément un lien de causalité et une absence d'association statistiquement significative ne signifie pas non plus une absence de lien.

10.2.2 Imitation vocale

La théorie de l'imitation vocale²⁶⁵ est sans aucun doute à l'œuvre chez nos locuteurs et il nous semble que les résultats de la prononciation et les réponses qu'ils ont fournies dans le questionnaire le montrent largement. Tout d'abord, nous rappelons qu'à la question « lorsque vous parlez anglais, imitez-vous intentionnellement un type d'accent ? », trois locuteurs (4, 5 et 11) ont répondu affirmativement et leurs raisons sont dans l'intention de bien prononcer les mots en anglais. Ce type de convergence serait à notre avis phonético-phonologique, différée, volontaire et médiée ou non (*cf.* section 4.5, p. 126). Pour les autres locuteurs qui déclarent ne pas imiter intentionnellement un type d'accent, car ils pensent que l'imitation de la prononciation est difficile ou naturelle, nous pouvons mettre en avant un mécanisme de convergence phonético-phonologique différée, involontaire. Des spécificités de l'accent américain se retrouvent dans leur

²⁶⁵ Voir section 4.4, p. 105.

prononciation sans qu'ils aient cherché à imiter cet accent. Dans cette perspective, nous pouvons considérer le mot « naturel » ici comme un synonyme « d'involontaire », c'est-à-dire un mécanisme de convergence dévoué de tout effort de conscience. D'ailleurs, nous rappelons au travers de nos résultats que la prononciation avec des éléments proches de l'accent britannique est considérée comme étant « normale » chez nos locuteurs, grâce au système scolaire (enseignants et ressources audiovisuelles éducatives) dans lequel la prononciation britannique est dominante. Pour ceux qui ont eu des enseignants d'anglais ayant différents accents, y compris la RP et le GA, nous pouvons admettre que ces éléments proviennent également du cadre formel. Cependant, la prononciation proche de l'accent américain des locuteurs qui n'ont pas eu d'enseignants ayant cet accent ne pourrait émaner que de l'alignement vocal inconscient à partir des activités informelles. Si en plus de cette situation, des locuteurs n'interagissent jamais en anglais en présentiel, nous ne pouvons que constater une convergence involontaire médiée (locuteurs 5, 6, 7, 14, 16 et 18). Nous formulons cette hypothèse, car à ce stade, nous ne disposons d'aucun autre facteur nous permettant d'expliquer la présence de caractéristiques proches de l'accent américain dans la prononciation de ces locuteurs.

Concernant les différents modèles d'acquisition de la prononciation présentés dans notre cadre théorique, force est de constater que les résultats de la production et de la perception sont partagés entre plusieurs postulats émis par les théoriciens de ces modèles. En premier lieu, l'hypothèse de l'*Acoustique Image Imprinting* (Neufeld, 1977), selon laquelle l'apprentissage phonéto-phonologique d'une L2 nécessite suffisamment de temps d'exposition à de l'*input* phonique, serait vérifiée chez les locuteurs qui consacrent plus de temps aux activités informelles de l'anglais en ligne (les « grands consommateurs »). Il en résulte que plus les locuteurs passent du temps sur des activités informelles, plus leur cerveau emmagasine des informations qui facilitera la prononciation plus tard. Le deuxième postulat qui part du principe qu'une « bonne » prononciation des sons d'une L2 implique la création d'images correspondantes « correctes » est également observé dans nos résultats. Pour notre part, les productions « Autres » relevées chez les locuteurs constituent des images déformées de sons et séries de sons au regard des deux accents de référence avec lesquels leur prononciation est comparée. Nous précisons que cela ne signifie pas qu'il s'agit d'une prononciation erronée tant que l'intelligibilité est en place.

Aussi, la fréquence d'exposition aux divers exemplaires de mots prononcés avec un accent spécifique se traduit par la quantité d'heures plus ou moins importante que les étudiants consacrent aux activités en ligne pour leurs loisirs. Nous pensons que la fréquence joue un rôle capital dans l'installation de mécanismes neurocognitifs, facilitant la convergence vocale involontaire qui permet une acquisition phonético-phonologique stable et indispensable tout au long de la vie. En effet, l'hypothèse du *Native Language Magnet* (Davis & Kuhl, 1994) à laquelle sont soumis les prototypes de sons entendus et qui expliquerait la capacité de discrimination des sons de la L2, confirme le rôle du classement d'exemplaires selon des critères bien définis. En d'autres termes, les premières expériences d'expositions aux sons natifs permettent une classification phonique correcte et faciliterait la distinction des sons de la L1 en fonction de celles de la L2. Il en résulterait plus tard, une production de sons de la L2 qui serait meilleure grâce à la capacité de discrimination auditive des sons des deux langues.

Quant à la fréquence d'usage d'exemplaires, nos données montrent que les étudiants s'offrent peu d'opportunités de pratiques communicatives/interactionnelles orales qui leur permettraient non seulement d'acquérir les mouvements articulatoires adéquats, mais également de s'habituer à ces gestes cognitivo-moteurs. Ce point rejoint l'un des postulats du modèle *Perceptual Assimilation Model* (PAM de Best, 1995) qui place l'articulation au cœur de l'acquisition de la prononciation. Contrairement au *Speech Learning Model* (SLM de Flege, 1992) qui admet une corrélation directe entre perception et production, le PAM considère que la perception auditive est une condition nécessaire mais non suffisante pour une assimilation « réussie » d'éléments de prononciation d'une L2. Cependant, nos données, notamment celles de la triangulation ont révélé des résultats mitigés en fonction des sons. La perception et la production sont corrélés chez certains et non chez d'autres.

Enfin, les différents types de classification des sons d'une L2 par rapport à ceux de la L1, seraient justifiés dans la prononciation de nos locuteurs, car, comme nous l'avons montré, certains sons subissent plus de modifications que d'autres. Qu'ils soient linguistiques ou non, identiques, similaires ou nouveaux (SLM), ou encore de bons exemples, non idéaux ou déviants de catégories natives (PAM et NLM), les sons d'une langue ne seraient pas tous acquis de la même manière. Par exemple, nous avons relevé que certains sons (ex : /t/ intervocalique) admettent plus de variations que d'autres (rhoticité). Parmi les causes possibles de cette variabilité, nous avons évoqué la nature

et le type du son (difficulté articulatoire ou perceptive), la ressemblance ou dissemblance avec ceux de la L1, la fréquence d'apparition, etc. Encore une fois nous voyons la complexité à l'œuvre au regard des influences multiples sur l'acquisition de la matière phonique, à travers l'imitation vocale, qui, rappelons-le n'est pas qu'un simple acte de mimétisme, opérant à partir de constructions cognitives et motrices.

10.2.3 Viabilité et perturbation

Partant de la notion de l'invariance en production-perception de la parole, nous avons souligné que le concept de Variabilité Adaptative met en lumière la capacité des locuteurs à ajuster des caractéristiques de leur prononciation en fonction des situations de communication. Résultat de la complexité inhérente à la production de la parole, la variabilité inter et intra-personnelle constitue un des constats centraux de notre travail. À cet égard, la théorie de viabilité en production-perception de la parole (Sock, 1998)²⁶⁶ permet de rendre compte de l'intelligibilité du discours humain sur le plan phonétique et phonologie, malgré les différences inter et intra-individuelles au niveau de la prononciation. Globalement, la théorie stipule qu'il existerait une zone de production de sons au-delà de laquelle certains sons/séries de sons ne seraient pas compréhensibles, c'est-à-dire qui ne coïncident pas aux cibles spécifiques (viabiles). Lorsqu'un locuteur produit un son cible compréhensible, nous considérons que la trajectoire (les gestes) articulatoire correspondante est viable, quelle que soit sa nature.

Dans notre travail, nous avons employé le concept de la perturbation en production-perception de la parole (Vaxelaire, 2007) afin de tester la viabilité des trajectoires articulatoires, en se fondant sur des analyses acoustiques de la production de nos locuteurs. Ainsi, à travers l'augmentation de la vitesse d'élocution, qui est un perturbateur naturel de parole, nous avons mis en évidence des cas de viabilité de certaines productions. Comme nous l'avons vu, cette viabilité se traduit par la production des mêmes sons durant la vitesse normale et la vitesse rapide. Ce fait témoigne de la robustesse ou de la résistivité des gestes articulatoires en vue d'une préservation des messages linguistiques (Vaxelaire, 2007). Par exemple, il y a eu résistivité de trajectoires articulatoires (visibles dans les signaux acoustiques) chez les locuteurs qui ont produit des battues dans la prononciation du mot *city*. Nous supposons aussi que la perturbation a entraîné le déploiement

²⁶⁶ Voir section 5.4, p. 140.

des stratégies de compensation ou de réajustement dans la production de nos locuteurs. En revanche, en l'absence de mesures spécifiques²⁶⁷, nous ne pouvons pas garantir la compréhensibilité de la prononciation de nos locuteurs. Nous prenons cette précaution car la production d'un son « raté » peut toutefois générer ou non de la compréhension linguistique comme c'est le cas avec la prononciation du mot *city* prononcé avec diverses trajectoires articulatoires ([t^h] et [tʃ], [r] et [d], [t̥] et [tʃ], [d] et [t̥]). Comme nous l'avons évoqué plus haut, nous rappelons que certains sons sont plus viables et résistent mieux à l'augmentation de la vitesse d'élocution que d'autres et cela dépend autant des locuteurs. En somme, nous avons montré que les résultats de notre travail soutiennent les théories évoquées à bien des égards. Notre recherche est également source de nouvelles données dans le champ de l'acquisition des langues secondes et dans celui de l'apprentissage informel de l'anglais en ligne. Nous traiterons ce point dans la prochaine section.

10.3 Apports de notre travail de recherche

Nous abordons dans cette section les apports de ce travail en termes d'intérêt pour la recherche dans le domaine de l'AIL en général et de l'AIAL en particulier. Cette recherche est ancrée dans la lignée des travaux sur l'AIAL et est la toute première à aborder la question de la prononciation²⁶⁸ des apprenants français en lien avec leurs pratiques informelles. Elle est donc pionnière dans ce champ spécial de l'AIAL et apporte un nombre d'éléments nouveaux de compréhension et d'observation concernant l'adoption d'accents. Dans ce sens, malgré les différents types de prononciations recensés, nous offrons des éléments d'indices empiriques, bien que modestes, que la prononciation des apprenants français contient de plus en plus de caractéristiques phonéto-phonologiques de l'accent américain, comme le percevaient spontanément des enseignants d'anglais (Toffoli & Sockett, 2015b). Aussi, notre étude s'est largement reposée sur les recherches précédentes concernant l'AIAL en matière des questions de l'enquête. Néanmoins, l'adaptation du questionnaire à notre sujet (la prononciation) a généré de nouvelles informations comme celles liées à l'imitation d'accents.

²⁶⁷ Il s'agit notamment du test de perception en direction des enseignants.

²⁶⁸ Kussyk (2017) a analysé la prononciation d'étudiants français et allemands mais d'un point de vue de l'aisance à communiquer et non d'un point de vue acoustique (phonétique).

L'une des forces de notre étude est sans doute l'originalité des procédures méthodologiques dont elle s'est dotée et qui lui confère son caractère scientifiquement robuste. En effet, pendant que certaines études se focalisent sur une seule source de collecte de données (entretiens ou questionnaires par exemple), la nôtre en a rassemblé trois bien distinctes et complémentaires. L'exploitation des résultats de la production orale, du test de perception et de l'enquête nous a permis de procéder à une triangulation de données et de pouvoir mener des analyses statistiques de corrélation. Cela implique le croisement d'informations nécessaires à une meilleure interprétation des résultats et l'assise scientifique des conclusions tirées des observations. Notre étude s'est également servie des analyses acoustiques dans la détection exacte des sons et séries de sons étudiés, au lieu de simples écoutes par l'expérimentateur. De surcroît, notre analyse acoustique incorpore plusieurs éléments segmentaux et suprasegmentaux qui sont en principe difficiles à appréhender (le son /r/ par exemple).

Un autre apport de cette recherche réside dans la capacité des locuteurs à distinguer l'accent américain de l'accent britannique avec une précision. Ainsi, le temps de réponse court, le nombre élevé de réponses correctes, la force du degré de certitude et les quelques justifications des choix, témoignent d'une conscience phonologique bien ancrée envers ces deux accents. L'importance de ce test de perception repose sur l'automatisation des réponses des locuteurs en fonction du nombre d'écoutes et de la durée de leur réaction (grâce au script). Parallèlement, ces résultats suggèrent une fiabilité et une fidélité de réponses des locuteurs concernant les accents adoptés par leurs enseignants d'anglais et ceux qu'ils déclarent imiter. Nous pouvons même supposer que la connaissance des caractéristiques d'un accent particulier pourrait leur permettre de mieux comprendre leurs divers interlocuteurs et ainsi de mieux interagir dans différents contextes de communication (Whyte, 2019). En effet, dans une approche didactique, la compétence en communication peut être considérée comme la cible de l'acquisition d'une langue seconde ou étrangère, et constitue le point focal des évaluations en langues (*Ibid.*)

Notre cadre théorique a apporté des connaissances fondées et solides permettant de mieux comprendre le fonctionnement du système dans lequel l'AIAL se déroule. Elle montre que l'acquisition phonético-phonologique d'une L2 est régie par une démarche systémique, au sein de laquelle divers constituants interagissent de manière chaotique. La prononciation est dans ce sens l'un des éléments qui émerge de l'interconnexion de ces facteurs. Considérant notre échantillon, il

semblerait que les différentes activités issues de pratiques informelles et formelles interviennent toutes dans la prononciation adoptée pour les aspects étudiés. Cependant, l'identification des attracteurs qui auraient plus d'influences sur les types de prononciation nous a permis de cerner ces facteurs : les caractéristiques proches de l'accent américain proviendraient principalement des activités informelles alors que celles qui sont proches de l'accent britannique découleraient essentiellement du cadre formel. Ce contexte nous a également permis de comprendre que les mécanismes de l'imitation vocale seraient le soubassement de l'apprentissage phonéto-phonologique d'une L2 et constitueraient un acte cognitif, moteur, social et culturel. Suivant les données du questionnaire, nous savons que les locuteurs imitent les différents *inputs* de manière consciente et/ou inconsciente et que les expositions directe et indirecte forment des modalités sous-jacentes. Enfin, la théorie de la viabilité et le concept de perturbations en perception-production nous éclairent également sur la façon dont nos productions sont perçues par nos interlocuteurs en termes de compréhensibilité et d'intelligibilité. Nous comprenons mieux comment nos itérations phonétiques résistent aux conditions de perturbations, grâce aux stratégies de compensation-réajustement de gestes articulatoires. Jusque-là utilisée uniquement en phonétique, cette théorie est utilisée pour la première fois dans le domaine de la didactique des langues, au travers de l'augmentation de la vitesse d'élocution, afin de tester la robustesse de trajectoires articulatoires. Pour des chercheurs intéressés par le domaine de l'enseignement de la prononciation, ces trois théories constituent une richesse en concepts par lesquelles ils trouveront des pistes exploitables. Enfin, forte de sa documentation variée, notre étude se nourrit de concepts et recherches pluri- et interdisciplinaires, notamment dans les champs de la phonétique et de la phonologie, des neurosciences, des sciences de l'éducation, des mathématiques, de la physique, etc. Il nous semble que cet aspect a certainement des conséquences sur l'approfondissement et l'articulation de certains concepts notoires sur lesquels nous aurions pu porter plus d'attention. Ce point nous amène à traiter les limites de notre travail dans la section suivante.

10.4 Limites et perspectives de ce travail de recherche

Notre étude étant pionnière dans son domaine, elle comporte certes des limites mais qui peuvent toutefois servir à d'autres chercheurs. Dans les points que nous allons évoquer ci-dessous,

nous discutons ces difficultés et formulons des pistes d'amélioration pour des recherches futures concernant l'intégration de nouveaux facteurs méthodologiques.

Tout d'abord, en augmentant la taille de l'échantillon à un nombre de locuteurs suffisant pour conduire des analyses statistiques plus poussées, nous pourrions déceler des liens statistiquement significatifs plus viables. En effet, des tests de corrélation (stagnations et factoriels) nous permettraient de révéler des relations mieux fondées entre les diverses activités et leurs modalités. Par exemple, croiser les résultats du temps de visionnage des séries américaines simultanément, du nombre de séries visionnées et des pourcentages de productions des éléments de l'accent américain, pourrait nous offrir davantage d'éléments d'interprétation. En faisant cela, nous ne pouvons que gagner en certitude au niveau des conclusions. Toutefois, œuvrant dans une logique systémique, et conscient des risques courus à vouloir contrôler des variables, nous pensons qu'il serait difficile, voire impossible d'isoler telle ou telle activité pratiquée par les participants. Au contraire, nous souhaiterions récolter plus d'éléments de réponses en élargissant l'éventail de questions relatives aux pratiques de l'AIAL.

Considérant l'augmentation du nombre de questions posées et afin de tendre vers une exhaustivité de certaines questions, nous pourrions interroger les participants sur leurs impressions personnelles et les raisons de l'adoption d'éléments de prononciation. Nous imaginons ainsi qu'après la phase de production, ils répondent à des questions du type : vous avez prononcé ce mot avec un accent britannique, savez-vous pourquoi ? Ou, pourquoi cette prononciation vous est-elle venue en premier ? etc. Aussi, il nous paraît intéressant d'explorer les ressentis des participants, en posant des questions liées à la motivation et à l'attachement (Toffoli, 2018), car ces facteurs devraient également jouer un rôle dans l'assimilation d'un type d'accent particulier. D'un point de vue psychosociologique, nous pensons également qu'il pourrait s'y constituer une sorte d'interaction virtuelle avec les acteurs des séries. Comme l'ont montré les études de Stuart-Smith *et al.* (2013) et celles de Takano et Ota (2014)²⁶⁹, il se pourrait que le public français imite vocalement des acteurs qu'ils aiment et auxquels ils désirent « ressembler ». Savoir si les apprenants s'attachent émotionnellement à des personnages de séries ou films pourrait fournir des pistes de compréhension quant à l'imitation vocale qui s'opère chez eux.

²⁶⁹ Les jeunes japonais imitent vocalement et s'identifient aux personnages des dessins animés, ainsi que l'imitation vocale d'acteurs de la série *East Enders* dans le contexte de Glasgow (*cf.* section 1.3.4, p. 18).

Deux tests de perception différents ont été prévus au commencement de cette thèse. Le premier test a été réalisé et consiste à savoir si nos participants reconnaissent la différence entre l'accent américain et l'accent britannique (*cf.* Résultats du test de perception, page 338). Le deuxième qui n'a malheureusement pas abouti, avait un objectif double : d'une part, il s'agissait d'examiner si les productions de nos locuteurs sont intelligibles et compréhensibles par les natifs et non natifs anglophones, et d'autre part, il était question d'identifier des productions de nos locuteurs au regard de l'accent RP et du GA par des auditeurs professionnels. Nous rappelons que l'idée de départ était d'appliquer les fondements de la théorie de la viabilité en perception-production de la parole (Sock, 1998). En d'autres termes, nous voulions explorer les éléments qui rendraient les productions de nos locuteurs (in)intelligibles ou (in)compréhensibles. Cependant, nous avons conduit une étude pilote²⁷⁰ qui a révélé que même si certaines productions se rapprochaient de la RP et d'autres du GA, globalement l'accent français était inévitablement décelable. Les retours de ces experts montrent que le découpage des mots (isolés) et l'absence de contexte (et donc d'éléments prosodiques) ont été déterminants dans l'identification d'accents et le non-aboutissement de ce test. Nous pensons toutefois que contourner ces manquements ou reconduire un autre test en respectant ces principes pourrait constituer une amélioration du test de perception, dans le cadre de prochaines études.

D'autres approches méthodologiques plus classiques comme la constitution d'un groupe témoin et d'un groupe expérimental pourraient faciliter l'établissement de liens entre les diverses variables étudiées. Cette démarche peut constituer un moyen de contrôle naturel mais pourrait être fastidieuse voire impossible à mettre en place, car cela requerrait en amont un nombre successif d'isolation d'autres variables. Ainsi, si nous souhaitons comparer les productions d'un groupe qui regarde les séries télévisées américaines avec un groupe qui n'en regarde pas, rien ne permettrait de statuer sur les origines de ces éléments, au regard des autres activités pratiquées et (non)isolées. Cette approche nécessite une sélection rigoureuse de facteurs concernant le groupe contrôle. Il faudrait alors s'assurer que les participants de ce dernier groupe ne s'engagent pas dans d'autres activités impliquant l'exposition à l'accent américain, comme par exemple l'écoute de la musique (à la radio ou sur Deezer, Spotify ou YouTube), les visionnages de clips vidéo (sur

²⁷⁰ Cinq enseignants d'anglais du CRL de Strasbourg (LEBEL) dont deux Américains, deux Britannique et un Français, ont jugé la prononciation des mots par nos locuteurs.

YouTube), de l'actualité (sur les chaînes d'information), des séjours aux États-Unis ou les interactions avec des personnes parlant l'anglais avec l'accent américain (enseignants, famille et amis), etc. Nous pourrions imaginer une approche multi-groupe, dans laquelle chaque groupuscule serait contrôlé par un nombre fixe et déterminé de facteurs à étudier. Par exemple, nous pourrions comparer la production orale de trois groupes : un premier dont les membres n'ont « jamais » regardé²⁷¹ et ne regardent pas de séries, un deuxième dans lequel les participants regardent modérément et un dernier constitué de locuteurs qui regardent énormément²⁷² des séries télévisées américaines.

Au départ, nous avons prévu d'analyser les dix « phrases » issues de Kusk et Sockett (2012) et le texte *The North wind and the sun* (IPA, 2005), comme annoncé dans le chapitre méthodologie (cf. section 6.6.3 p. 193). Ne les ayant pas inclus dans ce présent travail, nous conservons ces données en vue d'analyser plus tard les relations entre l'articulation de la parole et l'organisation prosodique des phrases (Fougeron, 1998). Cette analyse nous permettra d'identifier des processus de liaison et de réduction, spécifiques à l'anglais britannique et à l'anglais américain. Par exemple, la liaison entre *what are* dans l'interrogation *what are you doing* ? Serait réalisée avec un battue en GA (/wʌr_ɑɪ ju: 'duŋ/) et avec un /t/ en RP (/wɒt_ɑ: ju: 'du:ŋ/). Dans cette même phrase, une réduction vocalique de *are you* (/əjə/) et une intonation montante sur le *doing* seront également attendues. De plus, nous pourrions identifier l'influence d'autres facteurs liés au style, à la syntaxe, au lexique (Fougeron *et al.*, 2001) et du type de langage comme le soulève Miras (2013, paragr. 37) :

Il faut donc, dans la médiation de la perception / production des sons en langue étrangère, apporter un plus grand intérêt aux aspects psychosociaux qui interviennent directement ou indirectement dans la production des sons.

En partant des analyses prosodiques, nous imaginons d'autres méthodologies. Malgré les difficultés inhérentes à l'homogénéité²⁷³ dans les prises de sons et aux analyses acoustiques, enregistrer des locuteurs dans des conditions plus écologiques pourrait générer d'autres types de résultats pertinents. Ainsi, nous pourrions envisager des enregistrements de productions spontanées et contextuelles, qui font non seulement « sens » pour les participants, mais également nous permettent d'obtenir des caractéristiques relativement homogènes et faciles à analyser. Nous

²⁷¹ Seulement en VF.

²⁷² La fourchette entre visionnages « modéré » et « énorme » mériterait d'être déterminée en termes d'heures par semaine. Ceci rejoint le débat du nombre d'heures qui constituerait un visionnage « fréquent » (Kusk, 2017).

²⁷³ C'est-à-dire les mêmes sons étudiés pour tous les participants.

pourrions demander aux locuteurs de raconter un bref extrait d'un épisode d'une de leurs séries télévisées préférées et ensuite examiner des caractéristiques bien identifiées dans leur discours. Nonobstant, il serait souhaitable d'analyser les mêmes sons et séquences de sons chez tous les participants. Autrement, toute démarche susceptible de générer de la parole spontanée mais à *minima* contrôlée, c'est-à-dire exploitable, serait intéressante. À défaut, la simple et traditionnelle description d'images serait une alternative moins risquée. Ce moyen de collecte de données en environnements écologiques²⁷⁴ de production de parole spontanée permettrait *in fine* l'obtention de données non seulement phonético-phonologiques, mais aussi riches en éléments prosodiques.

Comme nous l'avons déjà mentionné, cette recherche est la première dans la lignée des travaux sur l'influence de l'apprentissage informel de l'anglais en ligne sur la prononciation. Ainsi, d'autres études seront nécessaires en vue d'approfondir les tendances relevées et de procéder à une modélisation conceptuelle propre à l'acquisition d'éléments de prononciation au travers des médias. En suivant des recommandations méthodologiques de théoriciens linguistes des systèmes complexes et dynamiques (de Bot & Larsen-Freeman, 2011 ; van Dijk, Verspoor, & Lowie, 2011), nous souhaiterions mener une expérimentation longitudinale afin de mesurer les réelles conséquences de l'AIAL sur la prononciation, en se focalisant uniquement sur les séries télévisées américaines. Cette étude de cas nécessitera de notre part une rigueur non seulement dans les procédures expérimentales mais aussi dans la composition de l'échantillon, concernant des critères de sélection des sujets. Des questionnaires bien élaborés, des entretiens poussés et des enregistrements variés de discours (contrôlés et écologiques) nous permettraient de comprendre et de suivre le développement de la prononciation chez les étudiants. À terme, comme dans le travail de thèse de Kussyk (2017), nous espérons étendre cette étude longitudinale sur la prononciation à d'autres pays ainsi qu'à d'autres langues.

Ayant posé les bases de l'état des lieux de quelques spécificités de la prononciation d'un échantillon d'étudiants français, il nous incombe d'appliquer ces principes à une plus grande échelle afin de mieux analyser les différents profils qui se dessineront et consolider nos connaissances sur les comportements évolutifs des apprenants. En effet, dans notre cohorte, il nous a été difficile de catégoriser les différents profils car certains groupes étaient uniquement composés d'une ou deux personnes, ce qui est insuffisant dans une démarche de profilage. Un

274 Parole spontanée.

projet d'une plus grande envergure, qui fait intervenir plusieurs étudiants de l'université de Strasbourg, voire d'autres universités françaises serait souhaitable. Conjointement, nous désirerons travailler auprès d'étudiants LANSAD (notamment des centres de ressources en langues) afin de nous impliquer davantage dans l'application des résultats de nos recherches à la « médiation » de la prononciation. En nous situant dans les travaux émergentistes et socioconstructivistes, nous préférons le terme de médiation à celui de l'enseignement en raison de l'aspect complexe de l'apprentissage (Lowie, 2017 ; Hilton, 2017, Miras, 2017). Dans cette visée, nous pensons à l'exploitation des pratiques informelles auxquelles s'adonnent les étudiants (musique, jeux vidéo, séries, films, lecture, réseaux sociaux, etc.) afin qu'ils en tirent un maximum de profit, durant leur cheminement vers l'acquisition de la prononciation L2 (Miras, 2013). Il sera question précisément d'examiner des aspects liés à l'intelligibilité et la compréhensibilité des discours de natifs mais également d'étudiants, à travers quelques propositions didactiques que nous présentons dans la section ci-dessous.

Conclusion générale

Se situant dans le domaine de recherche de l'apprentissage informel de l'anglais en ligne, la première préoccupation de notre travail est de dresser un état des lieux de quelques éléments de prononciation d'un échantillon d'étudiants français. Ainsi, ayant comparé les résultats de l'analyse acoustique de ces éléments de prononciation de notre échantillon avec ceux de l'anglais américain et de l'anglais britannique (références), nous avons noté trois grandes tendances. Certaines réalisations des participants se rapprochent plus de l'anglais américain tandis que d'autres se rapprochent plus de l'anglais britannique. Les caractéristiques analysées, que nous avons nommées « Autres » par comparaison à « RP » et « GA », proviendraient vraisemblablement du français et peut-être d'autres accents auxquels ces locuteurs sont exposés. D'autre part, l'augmentation de la vitesse d'élocution, qui a pour objectif de perturber la parole et de tester la résistivité dans la production des sons analysés, a également montré une stabilité dans la production de certains sons. À cet égard, nous posons que la prononciation des mêmes sons dans les deux contextes d'élocution suppose que l'apprenant maîtrise « parfaitement » son articulation et nous faisons le lien avec la notion de l'imitation inconsciente qui permet d'assimiler des éléments de prononciation d'une L2 de manière naturelle et sans effort.

La deuxième préoccupation de notre thèse consiste en l'exploration des influences éventuelles des pratiques informelles sur la prononciation de nos locuteurs. Ainsi, les données du questionnaire que nous avons collectées auprès de notre échantillon nous ont permis d'entrevoir les nombreuses activités pratiquées par ces derniers et d'inférer sur les origines de ces différents éléments de prononciation. Les éléments caractéristiques proches de l'accent britannique sembleraient provenir du cadre formel et celles qui se rapprochent de l'accent américain seraient davantage influencées par les pratiques informelles, notamment les séries télévisées. Même si tous les locuteurs savent distinguer les mots prononcés avec les deux accents, il existe une plus grande variabilité dans la prononciation et dans les réponses du questionnaire. En se reposant sur toutes nos données, nous avons surtout noté des tendances, dont certaines sont très fortes, parmi les éléments qui caractérisent les locuteurs. Ainsi, nous avons pu identifier les « grands » locuteurs d'éléments proches des deux accents et les « petits » locuteurs (*cf.* profilage section 9.9.1, p. 341). Toutefois, nous prenons ces résultats avec précaution étant donné les limites de ce travail (l'interconnexion entre les diverses variables) et nous précisons qu'outre la variabilité, nos conclusions ne sont pas entièrement généralisables à d'autres apprenants, mais uniquement valables pour notre échantillon. Néanmoins, nos résultats confirment certaines données des études sur l'AIAL en France et les études sur l'impact des médias sur la prononciation (Kusyk, 2017 ; Stuart-Smith, 2011). Concernant les pratiques scolaires, nous pensons qu'il serait nécessaire de réintroduire les cours de phonétique en classe de langue afin de permettre aux apprenants de se rendre compte de l'évolution de leur prononciation (Miras, 2013 ; 2019). Les méthodes correctives se reposant sur des techniques perceptives et articulatoires permettront à ces apprenants de travailler les sons et séquences de sons qui leur paraissent difficiles à assimiler.

Toute étude, surtout pionnière comme la nôtre, nécessite des pistes d'amélioration et d'ouverture comme nous l'avons évoqué dans la section sur les limites et perspectives. Nous sommes conscient de l'intervention dans nos résultats d'autres facteurs internes et externes, qui n'ont pas pu être analysés à travers la collecte des données, et qui seraient à prendre en considération dans des études à venir. À notre stade, il est difficile d'affirmer l'influence des médias avec certitude. L'apprentissage informel de l'anglais en ligne étant assujéti à un fonctionnement systémique dans lequel plusieurs facteurs sont en interaction perpétuelle ; des études longitudinales seraient nécessaires pour une meilleure compréhension de ce phénomène. Cette

nature complexe et dynamique du système dans lequel évoluent les apprenants suscite le désir d'approfondir la compréhension des mécanismes sous-jacents à l'accommodation d'éléments de prononciation, comme les facteurs motivationnels, psychosociologiques et émotionnels. Pour l'heure, il est normal que les apprenants continuent à faire ce qu'ils veulent, dans la mesure où ils s'approprient des éléments de prononciation authentique. Les chercheurs et enseignants ne peuvent que constater et étudier ces développements langagiers en fonction des différents accents.

Bibliographie

- Abdelmalek, A. (2004). Edgar Morin, sociologue et théoricien de la complexité: des cultures nationales à la civilisation européenne. *Sociétés*, 86(4), 99-117.
- Abraham, R. H., & Shaw, C. D. (1992). *Dynamics: The Geometry Of Behavior*. Redwood City : Basic Books.
- Adda-Decker, M., Fougeron, C., Gendrot, C., Delais-Roussarie, E., & Lamel, L. (2012). La liaison dans la parole spontanée familière: une étude sur grand corpus. *Revue française de linguistique appliquée, Vol. XVII* (1), 113-128.
- Algeo, J. (2006). *British or American English?: A Handbook of Word and Grammar Patterns*. Cambridge: Cambridge University Press.
- Amadiou, F., & Tricot, A. (2006). Utilisation d'un hypermédia et apprentissage : deux activités concurrentes ou complémentaires ? *Psychologie Française*, 51(1), 5-23.
- Anderson, R. W. (1983). *Pidginization and Creolization as Language Acquisition*. Rowley, MA: Newbury House.
- Andrews, M. L., Cox, W. M., & Smith, R. G. (1977). Effects of alcohol on the speech of non alcoholics. *Central States Speech Journal*, 28(2), 140-143.
- Androutsopoulos, J. (2014). Mediatization and sociolinguistic change. Key concepts, research traditions, open issues. In J. Androutsopoulos (Éd.), *Mediatization and sociolinguistic change* (p. 3-48). Berlin: De Gruyter.
- Arbib, M. A. (2012). *How the Brain Got Language : The Mirror System Hypothesis*. Oxford: Oxford University Press.
- Armati, L., Desplanques, E., & Poitte, I. (2011). Séries, elles ont pulvérisé le petit écran. Consulté 7 juillet 2019, à l'adresse Télérama.fr website: <https://www.telerama.fr/television/series-elles-ont-pulverise-le-petit-ecran,66030.php>
- Aubin, D., & Dalmedico, A. D. (2002). Writing the History of Dynamical Systems and Chaos: Longue Durée and Revolution, Disciplines and Cultures. *Historia Mathematica*, 29(3), 273-339. <https://doi.org/10.1006/hmat.2002.2351>
- Aubin, J.-P., & Cellina, A. (2012). *Differential Inclusions: Set-Valued Maps and Viability Theory* (2^e éd.). Berlin: Springer-Verlag.
- Aubin, J.-P., Bayen, P., & Saint-Pierre, P. P. (Éd.) (2011). Overview and Organization. In *Viability Theory* (p. 1-40).
- Baddeley, A. D. (1986). *Working Memory*. Clarendon Press.
- Bandura, A. (1995). *L'apprentissage social*. Paris: Editions Mardaga.
- Bandura, A., & Walters, R. (1963). *Social Learning and Personality Development*. New York: Holt Rinehart & Winston.
- Bandura, A., Ross, D., & Ross, S. A. (1963). Imitation of film-mediated aggressive models. *The Journal of Abnormal and Social Psychology*, 66(1), 3-11. <https://doi.org/10.1037/h0048687>
- Baudin, J., & Chuc, N. (2014). Séries : Pourquoi elles rendent les téléspectateurs complètement accros. Consulté 7 juillet 2019, à l'adresse FIGARO website: <http://tvmag.lefigaro.fr/le-scan-tele/series/2014/09/19/28005-20140919ARTFIG00228-series-pourquoi-elles-rendent-les-telespectateurs-completement-accros.php>
- Beal, J. C. (2014). *English in Modern Times*. Routledge.
- Bechet, M. (2011). Perturbation de la production des occlusives chez des locuteurs présentant une division palatine ou labio-palatine (Thèse de Doctorat). Université de Strasbourg, Strasbourg.

- Behne, D. M., Rivera, S. M., & Pisoni, D. B. (1991). Effects of alcohol on speech : I. Durations of isolated words, sentences, and passages in fluent speech. *The Journal of the Acoustical Society of America*, 90(4), 2311-2311. <https://doi.org/10.1121/1.401071>
- Benson, P. (2006). Autonomy in language teaching and learning. *Language Teaching*, 40(1), 21-40. <https://doi.org/10.1017/S0261444806003958>
- Benson, P. (2011). *Teaching and Researching Autonomy in Language Learning*. United Kingdom: Pearson Education.
- Benson, P., & Reinders, H. (Éd.). (2011). *Beyond the Language Classroom*. Basingstoke: Palgrave Macmillan.
- Berry, V., & Garcia, A. (2016). Éducation formelle et éducation informelle : Regards croisés sur la notion de compétence (transversale). *Éducation et socialisation. Les Cahiers du CERFEE*, (41). <https://doi.org/10.4000/edso.1636>
- Berthoz, A. (2009). *La simplicité*. Paris: Odile Jacob.
- Bertin, J.-C., Gravé, P., & Narcy-Combes, J.-P. (2010). *Second Language Distance Learning and Teaching : Theoretical Perspectives and Didactic Ergonomics*. Hershey-New York: Information Science Reference.
- Besse, H. & Porquier, R. (1991). *Grammaire et didactique des langues*. Paris: Hatier.
- Best, C. T. (1995). A Direct Realist View of Cross-Language Speech Perception. In W. Strange (Éd.), *Speech perception and linguistic experience : Theoretical and Methodological Issues* (p. 171-204). New York: Timonium.
- Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America*, 109(2), 775-794.
- Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology. Human Perception and Performance*, 14(3), 345-360. <https://doi.org/10.1037//0096-1523.14.3.345>
- Blanchard, F., & Leven, J. (2007). *Say Chic : A Collection of French Words We Can't Live Without*. New York: Scribner.
- Blumstein, S. E. (1989). Theoretical implications of the quantal nature of speech: a commentary. *Journal of Phonetics*, 17, 55-61.
- Boersma, P., & Weenink, D. (2017). Praat : Doing phonetics by computer (Version 6.0.28) [C++]. Consulté à l'adresse <http://www.fon.hum.uva.nl/praat/>
- Bogaards, P. (2008). *On ne parle pas français : La langue française face à l'anglais*. Bruxelles: De Boeck.
- Bonnot, J. P. F. (2001). Évolution des représentations phonologiques ou comment l'esprit vient à la matière. In D. Keller, J. P. Durafour, J. P. F. Bonnot, & R. Sock (Éd.), *Percevoir : Monde et langage. Invariance et variabilité du sens vécu* (p. 271-315). Hayen, Sprimont: Mardaga.
- Bouarourou, F. (2014). *La gémation en tarifit : Considérations phonologiques, études acoustique et articulatoire* (Thèse de Doctorat). Université de Strasbourg, Strasbourg.
- Boula de Mareüil, P. (2010). *D'où viennent les accents régionaux ?* Paris : le Pommier.
- Bourgeois, V., & Gheoghiu, S.-F. (2014). L'anglais, le français et l'allemand demeurent les langues étrangères les plus étudiées au premier cycle du secondaire dans l'UE28 en 2012. p. 3. Consulté à l'adresse <http://ec.europa.eu/eurostat>.

- Boutet, M. (2014). Depuis quand les Français sont-ils accros aux séries TV ? Consulté 7 juillet 2019, à l'adresse La Revue des Médias website: <http://larevuedesmedias.ina.fr/depuis-quand-les-francais-sont-ils-accros-aux-series-tv>
- Brevik, L. M., & Hellekjær, G. O. (2018). Outliers: Upper secondary school students who read better in the L2 than in L1. *International Journal of Educational Research*, 89, 80-91. <https://doi.org/10.1016/j.ijer.2017.10.001>
- Brougère, G. (2007). Les jeux du formel et de l'informel. *Revue française de pédagogie. Recherches en éducation*, (160), 5-12. <https://doi.org/10.4000/rfp.582>
- Brougère, G. (2017). Qu'entendre par jeu dans l'enseignement et l'apprentissage des langues : Diversité des situations et des modalités d'apprentissage. *Recherche et pratiques pédagogiques en langues de spécialité. Cahiers de l'Apliut*, 36(2). <https://doi.org/10.4000/apliut.5652>
- Brulard, I., Carr, P., Durand, J., & Navarro, S. (2015). *La prononciation de l'anglais contemporain dans le monde : Variation et structure*. Toulouse: Presses Universitaires du Midi.
- Bruner, J. S. (1983). Le développement de l'enfant. Savoir-faire, savoir dire (4^e éd.). Paris: PUF.
- Bruner, J. S. (1987). Comment les enfants apprennent à parler. Paris: Retz.
- Buffon, H. L. (2003). Le français une langue pour l'Europe. *Paris Match*, (2799). Consulté à l'adresse http://www.langue-francaise.org/Articles_Dossiers/Dos_match_langue_europe.php
- Bürki, A., Fougeron, C., Gendrot, C., & Frauenfelder, U. H. (2011). Phonetic reduction versus phonological deletion of French schwa: Some methodological issues. *Journal of Phonetics*, 39(3), 279-288. <https://doi.org/10.1016/j.wocn.2010.07.003>
- Bybee, J. (2003). *Phonology and Language Use*. Cambridge: Cambridge University Press.
- Calliope. (1989). La parole et son traitement automatique. Paris: Masson.
- Calvet, L.-J. (2010). *Histoire du français en Afrique*. Paris: Ecriture.
- Calvet, L.-J. (2013). *La sociolinguistique* (8e éd.). Paris: Presses universitaires de France.
- Cameron, L., & Larsen-Freeman, D. (2007). Complex systems and applied linguistics. *International Journal of Applied Linguistics*, 17(2), 226-239.
- Carey, D., & McGettigan, C. (2017). Magnetic resonance imaging of the brain and vocal tract: Applications to the study of speech production and language learning. *Neuropsychologia*, 98, 201-211. <https://doi.org/10.1016/j.neuropsychologia.2016.06.003>
- Carrier, J. (2014). *Tapestry : The History and Consequences of America's Complex Culture*. Algora Publishing.
- Casanave, C. P. (2012). Diary of a dabbler: Ecological influences on an EFL teacher's efforts to study Japanese informally. *TESOL Quarterly*, 46(4), 642-670.
- Chambers, J. K. (1998). TV Makes People Sound the Same. In Laurie Bauer & Peter Trudgill (Éd.), *Language Myths* (p. 123-131). New York: Penguin.
- Chanquoy, L., Tricot, A., & Sweller, J. (2007). *La charge cognitive : Théorie et applications*. Armand Colin.
- Chauveau, A. (2013). Séries TV, un phénomène de société. Consulté 7 juillet 2019, à l'adresse Le Huffington Post website: https://www.huffingtonpost.fr/agnes-chauveau/series-tv-un-phenomene-de_b_4015384.html
- Coenen, E., Zwitserlood, P., & Bölte, J. (2001). Variation and assimilation in German: Consequences for lexical access and representation. *Language and Cognitive Processes*, 16(5-6), 535-564. <https://doi.org/10.1080/01690960143000155>
- Coffield, F. (2000). *The Necessity of Informal Learning*. Policy Press.

- Cohn, A. C. (1993). Nasalisation in English : Phonology or phonetics. *Phonology*, 10(1), 43-81.
<https://doi.org/10.1017/S0952675700001731>
- Cole, J., & Vanderplank, R. (2016). Comparing autonomous and class-based learners in Brazil : Evidence for the present-day advantages of informal, out-of-class learning. *System*, 61, 31-42.
- Combes, C. (2013). *La pratique des séries télévisées : une sociologie de l'activité spectatorielle* (Thesis, Paris, ENMP). Consulté à l'adresse <http://www.theses.fr/2013ENMP0010>
- Combes, C. (2017). Figures de la sériphilie. *Terrains travaux*, N° 31(2), 223-243.
- Commander, M., & Guerrero, M. (2013). Reading as a social interactive process: The impact of shadow-reading in L2 classrooms. *Reading in a Foreign Language*, 25(2), 170-191.
- Conseil de l'Union Européenne. (2012). *Recommandations du Conseil du 20 décembre 2012 relative à la validation de l'apprentissage non formel et informel*. Consulté à l'adresse [https://eur-lex.europa.eu/legal-content/FR/HIS/?uri=celex:32012H1222\(01\)](https://eur-lex.europa.eu/legal-content/FR/HIS/?uri=celex:32012H1222(01))
- Conseil de l'Europe. (2001). Un Cadre Européen Commun de Référence pour les Langues : Apprendre, Enseigner, Évaluer. Didier.
- Cooper, A., Fecher, N., & Johnson, E. K. (2018). Toddlers' comprehension of adult and child talkers : Adult targets versus vocal tract similarity. *Cognition*, 173, 16-20.
- Corbillé, G. (2014, octobre 2). Comment les séries ont envahi notre quotidien. Consulté 7 juillet 2019, à l'adresse [Psychologies.com website: https://www.psychologies.com/Culture/Medias/Articles-et-Dossiers/Comment-les-series-ont-envahi-notre-quotidien](https://www.psychologies.com/Culture/Medias/Articles-et-Dossiers/Comment-les-series-ont-envahi-notre-quotidien)
- Cross, J. (2006). *Informal learning: Rediscovering the natural pathways that inspire innovation and performance*. San Francisco, CA, US: Pfeiffer/John Wiley & Sons.
- Cruttenden, A. (2014). *Gimson's pronunciation of English*. New York: Routledge.
- Crystal, D. (2012). *English as a Global Language*. Cambridge: Cambridge University Press.
- D'Imperio, M., Cavone, R., & Petrone, C. (2014). Phonetic and phonological imitation of intonation in two varieties of Italian. *Frontiers in Psychology*, 5, 1226-1236.
<https://doi.org/10.3389/fpsyg.2014.01226>
- Darragh, G. (2000). *A To Zed, A To Zee A Guide To The Differences Between British And American English*. Madrid: Editorial Stanley.
- Dausendschön-Gay, U. (2003). Communication exolingue et Interlangue. *Linx. Revue des linguistes de l'université Paris X Nanterre*, (49), 41-50. <https://doi.org/10.4000/linx.531>
- Dautenhahn, K., & Nehaniv, C. L. (Éd.). (2002). *Imitation in Animals and Artifacts*. Cambridge, MA, USA: MIT Press.
- David, J. (thèse en cours). *La « blockbusterisation » des séries télévisées américaines*. Consulté à l'adresse <http://www.theses.fr/s78916>
- Davies, C. (2005). *Divided by a Common Language : A Guide to British and American English*. Boston: Houghton Mifflin Company.
- Davies, M. (2011). *Corpus of American Soap Operas: 100 million words*. Consulté à l'adresse <http://corpus.byu.edu/soap/>
- Davis, K., & Kuhl, P. K. (1994). Tests of the perceptual magnet effect for American English /k/ and /g/. *The Journal of the Acoustical Society of America*, 95(5), 2976-2976.
<https://doi.org/10.1121/1.408982>
- de Bot, K. (2008). Introduction: Second Language Development as a Dynamic Process. *The Modern Language Journal*, 92(2), 166-178.

- de Bot, K., & Larsen-Freeman, D. (2011). Researching second language development from a dynamic systems theory perspective. In M. Verspoor, K. D. Bot, & W. Lowie (Éd.), *A Dynamic Approach to Second Language Development: Methods and Techniques* (29e éd., p. 5-23). Amsterdam/Philadelphie: John Benjamins Publishing.
- de Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamics systems theory approach to second language acquisition. *Bilingualism: Language and Cognition*, 10(1), 7-22.
- Delvaux, V., Huet, K., Piccaluga, M., & Harmegnies, B. (2013). Production training in Second Language Acquisition : A comparison between objective measures and subjective judgments. *Interspeech*, 25-29.
- Demaizière, F., & Narcy-Combes, J.-P. (2005). Méthodologie de la recherche didactique : Nativisation, tâches et TIC. *Alsic. Apprentissage des Langues et Systèmes d'Information et de Communication*, (Vol. 8, n° 1). <https://doi.org/10.4000/alsic.326>
- Deschamps, A. (1994). De l'écrit à l'oral et de l'oral à l'écrit : Phonétique et orthographe de l'anglais. Paris: Ophrys.
- Dewaele, J.-M. (2002). Variation, chaos et système en interlangue française. *Acquisition et interaction en langue étrangère*, (17), 143-167.
- Dewaele, J.-M. (2003). Compte rendu – hommage : L'œuvre de L. Selinker. *Linx. Revue des linguistes de l'université Paris X Nanterre*, (49), 153-159. <https://doi.org/10.4000/linx.564>
- Di Gioacchino, M., & Jessop, L. C. (2011). Uptalk-Towards a quantitative analysis. *Toronto Working Papers in Linguistics*, 33(1).
- Didienne, M. (2019). *Néologismes et enrichissement de la langue française* (Mémoire de master). Université de Strasbourg, Strasbourg.
- Dohlus, K. (2010). The Role of Phonology and Phonetics in Loanword Adaptation : German and French Front Rounded Vowels in Japanese. Peter Lang.
- Dörnyei, Z. (2007). *Research Methods in Applied Linguistics: Quantitative, Qualitative, and Mixed Methodologies*. Oxford ; New York, N.Y: OUP Oxford.
- Dörnyei, Z. (2009). Individual differences: Interplay of learner characteristics and learning environment. *Language Learning*, 59, 230–248.
- Dörnyei, Z. (2014). Researching complex dynamic systems: 'Retrodictive qualitative modelling' in the language classroom. *Language Teaching*, 47(1), 80–91.
- Dörnyei, Z., & Chan, L. (2013). Motivation and Vision: An Analysis of Future L2 Self Images, Sensory Styles, and Imagery Capacity Across Two Target Languages. *Language Learning*, 63(3), 437-462. <https://doi.org/10.1111/lang.12005>
- Dörnyei, Z., & Skehan, P. (2003). Individual Differences in Second Language Learning. In C. J. Doughty & M. H. Long (Éd.), *The handbook of second language acquisition* (p. 589-630). Oxford: Blackwell.
- Dörnyei, Z., MacIntyre, P., & Henry, A. (2015). Introduction: Applying complex dynamic systems principles to empirical research on L2 motivation. In Z. Dörnyei, P. MacIntyre, & A. Henry (Éd.), *Motivational Dynamics in Language Learning* (p. 1-7). Bristol: Multilingual Matters.
- Dschwen. (2006). Lorenz attractor, calculated with octave and converted to SVG using a quick hack perl script. Trace starts in red and fades to blue as t progresses. Work in progress. Created by User: Dschwen. Consulté à l'adresse https://commons.wikimedia.org/wiki/File:Lorenz_attractor.svg
- Dubois, C. (2009). Les bases neurophysiologiques de la perception audiovisuelle syllabique : Étude simulatnée en Imagerie par Résonance Magnétique fonctionnelle et en

- Électroencéphalographie (IRMf/EEG) (Thèse de Doctorat). Université de Strasbourg, Strasbourg.
- Durand, C. X. (1999). *La langue française : Atout ou obstacle*. Presses Univ. du Mirail.
- Ellis, N. C. (2002). Frequency Effects In Language Processing: A Review with Implications for Theories of Implicit and Explicit Language Acquisition. *Cambridge University Press*, 24(2), 143-188.
- Ellis, R. (2004). Individual Differences in Second Language Learning. In A. Davies & C. Elder (Éd.), *The Handbook of Applied Linguistics* (p. 525-551). Oxford: Blackwell Publishing Ltd.
- Ettlinger, M. (2007). An exemplar-based model of chain shifts. *Proceedings of the 16th international congress of the phonetic science*, 685-688.
- Exare, C. (2009). La mise en place de trois voyelles postérieures en anglais langue étrangère. Analyse acoustique et perspective de quelques interférences pour des pistes de correction phonétique. *Recherche et pratiques pédagogiques en langues de spécialité. Cahiers de l'Aplut*, (Vol. XXVIII N° 3), 68-84. <https://doi.org/10.4000/apliut.122>
- Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *The European Journal of Neuroscience*, 15(2), 399-402. <https://doi.org/10.1046/j.0953-816x.2001.01874.x>
- Fagard, J., & Lockman, J. J. (2010). Change in imitation for object manipulation between 10 and 12 months of age. *Developmental Psychobiology*, 52(1), 90-99. <https://doi.org/10.1002/dev.20416>
- Fauth, C. (2012). Perturbation de la production de la parole suite à une opération de la glande thyroïde (Thèse de Doctorat). Université de Strasbourg, Strasbourg.
- Feng, G., & Castelli, E. (1996). Some acoustic features of nasal and nasalized vowels : A target for vowel nasalization. *The Journal of the Acoustical Society of America*, 99(6), 3694-3706. <https://doi.org/10.1121/1.414967>
- Flege, J. (1992). Speech learning in a second language. In C, Ferguson, L, Menn, & C, Stoel-Gammon (Éd.), *Phonological Development Models, Research, and Implications* (p. 565-604). New York: Timonium.
- Flege, J. (2003). Assessing constraints on second-language segmental production and perception. *Phonetics and phonology in language comprehension and production: Differences and similarities*, 6, 319-355.
- Fougeron, C. (1998). Variations articulatoires en début de constituants prosodiques de différents niveaux en français (Thèse de Doctorat). Université Paris 3, Paris.
- Fougeron, C. (2013). Renforcements et réductions segmentales : influences de la structure prosodique, de la phonologie, et des contraintes motrices en parole pathologique (Habilitation à Diriger des Recherches). Université de Lyon 2, Lyon.
- Fougeron, C., & Delais-Roussarie, E. (2004). Liaisons et enchaînements « Fais_en à Fez_en parlant ». *Actes des Journées d'Etudes sur la Parole*, 221-224.
- Fougeron, C., Crevier-Buchman, L., Fredouille, C., Ghio, A., Meunier, C., Chevrier-Muller, C., ... Vincent, C. (2010). Developing an Acoustic-phonetic Characterization of Dysarthric Speech in French. *Proceedings of the Seventh conference on International Language Resources and Evaluation*. Présenté à LREC'10, Malte.
- Fougeron, C., D'Alessandro, D., & Lancia, L. (2018). Reduced coarticulation and aging. *The Journal of the Acoustical Society of America*, 144(3), 1905-1905. <https://doi.org/10.1121/1.5068348>

- Fougeron, C., Goldman, J.-P., Dart, A., Guélat, L., & Jeager, C. (2001). Influence de facteurs stylistiques, syntaxiques et lexicaux sur la réalisation de la liaison en français. *Actes de TALN*, 173-182.
- Fournier, P. (2016). L'Accentuation des Emprunts Français dans les Dictionnaires Anglais de Prononciation. *Proceedings of the XVII EURALEX International Congress*, 424-435.
- Galligani, S. (2003). Réflexion autour du concept d'interlangue pour décrire des variétés non natives avancées en français. *Linx. Revue des linguistes de l'université Paris X Nanterre*, (49), 141-152. <https://doi.org/10.4000/linx.562>
- Garnier, B. (2018). L'éducation informelle contre la forme scolaire ? *Carrefours de l'éducation*, (45), 67-91. <https://doi.org/10.3917/cdle.045.0067>
- Garnier, M., Lamalle, L., & Sato, M. (2013). Neural correlates of phonetic convergence and speech imitation. *Frontiers in psychology*, 4, 600.
- Gaskell, M. G., & Marslen-Wilson, W. D. (1998). Mechanisms of phonological inference in speech perception. *Journal of Experimental Psychology. Human Perception and Performance*, 24(2), 380-396.
- Gentilucci, M., & Bernardis, P. (2007). Imitation during phoneme production. *Neuropsychologia*, 45(3), 608-615.
- Giguère, J., Giasson, J., & Simard, C. (2002). Les relations entre la lecture et l'écriture : Représentations d'élèves de différents niveaux scolaires et de différents niveaux d'habileté. *Revue canadienne de linguistique appliquée*, 5(1-2), 23-50.
- Giles, H., & Ogay, T. (2007). Communication Accommodation Theory. In B. B. Whaley & W. Samter (Éd.), *Explaining communication: Contemporary theories and exemplars* (p. 293-310). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Ginésy, M. (2000). *Phonétique et phonologie de l'anglais*. Paris: Ellipses.
- Glain, O. (2012). The Yod/j : Palatalize it or drop it. How Traditional Yod Forms are Disappearing from Contemporary English. *Cercles*, 22, 4-24.
- Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(5), 1166-1183. <https://doi.org/10.1037/0278-7393.22.5.1166>
- Goldrick, M., & Blumstein, S. E. (2006). Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. *Language and Cognitive Processes*, 21(6), 649-683. <https://doi.org/10.1080/01690960500181332>
- Goldstein, L., Pouplier, M., Chen, L., Saltzman, E., & Byrd, D. (2007). Dynamic action units slip in speech production errors. *Cognition*, 103(3), 386-412. <https://doi.org/10.1016/j.cognition.2006.05.010>
- Gregory, S. W., Dagan, K., & Webster, S. (1997). Evaluating the Relation of Vocal Accommodation in Conversation Partners' Fundamental Frequencies to Perceptions of Communication Quality. *Journal of Nonverbal Behavior*, 21(1), 23-43. <https://doi.org/10.1023/A:1024995717773>
- Grignon, C. (2008). Prédiction et rétroédiction. *Revue européenne des sciences sociales. European Journal of Social Sciences*, (XLVI-142), 75-90. <https://doi.org/10.4000/ress.135>
- Grivicic, T., & Nilep, C. (2004). When phonation matters : The use and function of yeah and creaky voice.
- Grosjean, F., & Deschamps, A. (1975). Analyse contrastive des variables temporelles de l'anglais et du français : Vitesse de parole et variables composantes, phénomènes d'hésitation. *Phonética*, 31, 144-184.

- Guo, Y., & Beckett, G. H. (2007). The Hegemony of English as a Global Language: Reclaiming Local Knowledge and Culture in China. *Convergence*, 40(1-2), 117-132.
- Gut, U. (2009). *Non-native Speech: A Corpus-based Analysis of Phonological and Phonetic Properties of L2 English and German*. Peter Lang.
- Hachten, W. A., & Scotton, J. F. (2016). *The World News Prism: Digital, Social and Interactive*. Oxford: John Wiley & Sons.
- Hamada, M., & Koda, K. (2010). The Role of Phonological Decoding in Second Language Word-Meaning Inference. *Applied Linguistics*, 31(4), 513-531. <https://doi.org/10.1093/applin/amp061>
- Han, Z., & Chen, C. A. (2010). Repeated-reading-based instructional strategy and vocabulary acquisition: A case study of a heritage speaker of Chinese. *Reading in a Foreign Language*, 22(2), 242-262.
- Hantke, S., Weninger, F., Kurle, R., Ringeval, F., Batliner, A., Mousa, A. E.-D., & Schuller, B. (2016). I Hear You Eat and Speak: Automatic Recognition of Eating Condition and Food Type, Use-Cases, and Impact on ASR Performance. *PLoS ONE*, 11(5). <https://doi.org/10.1371/journal.pone.0154486>
- Harmegnies, B., & Poch-Olivé, D. (1992). A study of style-induced vowel variability: Laboratory versus spontaneous speech in Spanish. *Speech communication*, 11(4-5), 429-437.
- Hart, S. A. (2013, juin). Apprentissage formel, informel, non formel, des notions difficiles à utiliser... Pourquoi? Consulté 9 juillet 2018, à l'adresse OCE - L'Observatoire compétences-emplois, Université du Québec à Montréal website: <https://oce.uqam.ca/article/apprentissage-formel-informel-non-formel-des-notions-difficiles-a-utiliser-pourquoi/>
- Herry-Bénit, N. (2011). *Didactique de la phonétique anglaise*. Rennes: Presses Universitaires de Rennes.
- Hiiemae, K. M., & Palmer, J. B. (2003). Tongue Movements in Feeding and Speech. *Critical Reviews in Oral Biology & Medicine*, 14(6), 413-429. <https://doi.org/10.1177/154411130301400604>
- Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical society of America*, 97(5), 3099-3111.
- Hilton, H. E. (2017). Enjeux méthodologiques de l'émergentisme pour la recherche en acquisition et en didactique des langues. *Recherches en didactique des langues et des cultures. Les cahiers de l'Acedle*, 14(14-1). <https://doi.org/10.4000/rdlc.1101>
- Hirsch, F. (2007). *Le bégaiement: Perturbation de l'organisation temporelle de la parole et conséquences spectrales*. Université Marc Bloch, Strasbourg 2, Strasbourg.
- Hodges, R., Munro, N., Baker, E., McGregor, K., & Heard, R. (2017). The Monosyllable Imitation Test for Toddlers: influence of stimulus characteristics on imitation, compliance and diagnostic accuracy. *International Journal of Language & Communication Disorders*, 52(1), 30-45. <https://doi.org/10.1111/1460-6984.12249>
- Hollien, H., DeJong, G., Martin, C. A., Schwartz, R., & Liljegren, K. (2001). Effects of ethanol intoxication on speech suprasegmentals. *The Journal of the Acoustical Society of America*, 110(6), 3198-3206. <https://doi.org/10.1121/1.1413751>
- Honig, A. S. (2017). Language insights for caregivers with young children. *Early Child Development and Care*, 187(3-4), 527-541. <https://doi.org/10.1080/03004430.2016.1263917>
- Howell, P. (2008). Effect of Speaking Environment on Speech Production and Perception. *Journal of the human-environment system: JHES*, 11(1), 51-57. <https://doi.org/10.1618/jhes.11.51>
- Hughes, A., & Trudgill, P. (1979). *English Accents and Dialect: An Introduction to Social and Regional Varieties of British English*. London: Edward Arnold.

- Hughes, A., Trudgill, P., & Watt, D. (2012). *English Accents and Dialects* (5^e éd.). London: Hodder Education.
- International Phonetic Association. (2005). *Handbook of the International Phonetic Association: A Guide To The Use Of The International Phonetic Alphabet*. Cambridge: Cambridge University Press.
- Jacobi, D. (2018b). Une introduction à l'analyse de l'éducation non formelle. In D. Jacobi, *Public et Culture. Culture et éducation non formelle* (p. 1-13). Quebec: PUQ.
- Janicki, K. (1977). *Elements of British and American English*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Jauriberry, T. (2016). *Rhotiques et rhoticité en Ecosse: Une étude sociophonétique de l'anglais écossais standard* (Phd). Université de Strasbourg, Strasbourg.
- Jenkins, J. (2000). *The Phonology of English as an International Language: New Models, New Norms, New Goals*. Oxford: Oxford University Press.
- Jenkins, J. (2003). *World Englishes: A resource book for students*. Psychology Press.
- Jenkins, J., Baker, W., & Dewey, M. (2017). *The Routledge Handbook of English as a Lingua Franca*. Routledge.
- Jones, D. (2011). *Cambridge English Pronouncing Dictionary* (18^e éd.; P. Roach, J. Setter, & J. Esling, Éd.). Cambridge: Cambridge University Press.
- Jost, F. (2013). « Mentalist », « NCIS », « Les Experts » : Pourquoi les séries américaines cartonnent en France. Consulté 7 juillet 2019, à l'adresse <http://leplus.nouvelobs.com/contribution/916285-mentalist-ncis-les-experts-pourquoi-les-series-americaines-cartonnent-en-france.html>
- Jurkovič, V. (2019). Online informal learning of English through smartphones in Slovenia. *System*, 80, 27-37. <https://doi.org/10.1016/j.system.2018.10.007>
- Kachru, B. (1992). World Englishes: approaches, issues and resources. *Language Teaching*, 25(1), 1-14. <https://doi.org/10.1017/S0261444800006583>
- Kachru, B., Kachru, Y., & Nelson, C. (2009). *The Handbook of World Englishes*. Oxford: Wiley-Blackwell.
- Kail, M. (2015). *L'acquisition de plusieurs langues*. Paris: PUF.
- Kappes, J., Baumgaertner, A., Peschke, C., & Ziegler, W. (2009). Unintended imitation in nonword repetition. *Brain and Language*, 113(3), 140-151. <https://doi.org/10.1016/j.bandl.2009.08.008>
- Kartushina, N., & Frauenfelder, U. H. (2013). On the role of L1 speech production in L2 perception : Evidence from Spanish learners of French. *InterSpeech*, 2118-2122.
- Kohnert, K. (2007). *Language disorders in bilingual children and adults*. San Diego, CA: Plural Publishing.
- Krashen, S. (1976). Formal and Informal Linguistic Environments in Language Acquisition and Language Learning. *TESOL Quarterly*, 10(2), 157-168. <https://doi.org/10.2307/3585637>
- Krashen, S. (1998). Comprehensible output? *System*, 26(2), 175-182. [https://doi.org/10.1016/S0346-251X\(98\)00002-5](https://doi.org/10.1016/S0346-251X(98)00002-5)
- Krashen, S. (2013). The Effect of Direct Instruction on Pronunciation: Only Evident When Conditions for Monitor Use Are Met? *Gist Education And learning Research Journal*, (7), 271-275.
- Krashen, S., & Terrell, T. D. (1983). *The Natural Approach: Language Acquisition in the Classroom*. New York: Pergamon Press.
- Kuhl, P. K. (2010). Brain Mechanisms in Early Language Acquisition. *Neuron*, 67(5), 713-727. <https://doi.org/10.1016/j.neuron.2010.08.038>

- Kuhl, P. K., & Meltzoff, A. N. (1996). Infant vocalizations in response to speech: Vocal imitation and developmental change. *The Journal of the Acoustical Society of America*, 100(4 0 1), 2425-2438.
- Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society: Biological Sciences*, 363(1493), 979-1000. <https://doi.org/10.1098/rstb.2007.2154>
- Kusyk, M. (2016). An introduction to the Online Informal Learning of English: Perspectives on Second Language Development through leisure habits. In M. Roy, M. Kusyk, G. Schlemminger, & D. Bechmann (Éd.), *Environnements numériques et interactions en langue étrangères: du formel à l'informel, du réel à la réalité* (Multilingualism in Europe, p. 205-216). Berne: Peter Lang.
- Kusyk, M. (2017). Les dynamiques du développement de l'anglais au travers d'activités informelles en ligne : une étude exploratoire auprès d'étudiants allemands et français (Thèse de Doctorat). Université de Strasbourg, Strasbourg.
- Kusyk, M., & Sockett, G. (2012). From informal resource usage to incidental language acquisition: language uptake from online television viewing in English. *ASp. La Revue Du GERAS*, (62), 45-65. <https://doi.org/10.4000/asp.3104>
- Laborde, B. (2017). Television Series : Inventory of Research in France. *Series - International Journal of TV Serial Narratives*, 3(2), 91-100. <https://doi.org/10.6092/issn.2421-454X/7631>
- Labov, W. (2001). Principles of Linguistic Change, Volume 3 : Social Factors (Vol. 2). Oxford: Wiley-Blackwell.
- Labov, W., Ash, S., & Boberg, C. (2006). The Atlas of North American English : Phonetics, Phonology and Sound Change. Berlin: Walter de Gruyter.
- Ladefoged, P. (1983). The linguistic use of different phonation types. In D. Bless & J. Abbs (Eds.), *Vocal fold physiology: Contemporary research and clinical issues* (pp. 351-360). San Diego: College Hill Press.
- Ladefoged, P. (1996). *Elements of Acoustic Phonetics*. University of Chicago Press.
- Ladefoged, P., & Johnson, K. (2011). *A course in Phonetics* (Sixth edition). Boston: Wadsworth.
- Ladefoged, P., & Maddieson, I. (1996). *The Sounds of the World's Languages*. Oxford: Blackwell Publishing Ltd.
- Ladefoged, P., & Maddieson, I. (2009). *Sounds of the World's Languages*. Oxford: John Wiley and Sons Limited.
- Lakkaraju, K., Swarup, S., & Gasser, L. (2012). Consensus under Constraints : Modeling the Great English Vowel Shift. In S. J. Yang, A. M. Greenberg, & Mica Endsley (Éd.), *Social Computing, Behavioral—Cultural Modeling and Prediction* (p. 1-8). https://doi.org/10.1007/978-3-642-29047-3_1
- Lambert, C., Kormos, J., & Minn, D. (2017). Task Repetition And Second Language Speech Processing. *Studies in Second Language Acquisition*, 39(1), 167-196. <https://doi.org/10.1017/S0272263116000085>
- Lancien, M., Nicolas, A., & Fougeron, C. (2018). Effet de la situation de parole sur la variabilité des voyelles en français. *XXXIIe Journées d'Études sur la Parole*, 338-346. <https://doi.org/DOI:10.21437/JEP.2018-39>
- Lapteva, O. (2011). Speaker Perception and Recognition. An Integrative Framework for computational speech processing. Kassel: Kassel University Press GnbH.
- Larsen-Freeman, D. (1978). An ESL index of development. *TESOL quarterly*, 12, 439-448.

- Larsen-Freeman, D. (1997). Chaos/Complexity Science and second Language Acquisition. *Applied Linguistics*, 18(2), 141-165.
- Larsen-Freeman, D. (2009). Adjusting Expectations: The Study of Complexity, Accuracy, and Fluency in Second Language Acquisition. *Applied Linguistics*, 30(4), 579-589. <https://doi.org/10.1093/applin/amp043>
- Larsen-Freeman, D., & Cameron, L. (2008). Research Methodology on Language Development from a Complex Systems Perspective. *Modern Language Journal*, 92(2), 200-213. <https://doi.org/10.1111/j.1540-4781.2008.00714.x>
- Le Prell, C. G., & Clavier, O. H. (2017). Effects of noise on speech recognition: Challenges for communication by service members. *Hearing Research*, 349, 76-89. <https://doi.org/10.1016/j.heares.2016.10.004>
- Lee, J. S. (2017). Informal digital learning of English and second language vocabulary outcomes: Can quantity conquer quality? *British Journal of Educational Technology*, 50(2), 767-778. <https://doi.org/10.1111/bjet.12599>
- Lee, J. S. (2019). EFL students' views of willingness to communicate in the extramural digital context. *Computer Assisted Language Learning*, 1-21. <https://doi.org/10.1080/09588221.2018.1535509>
- Lee, J. S., & Lee, K. (2019). Perceptions of English as an international language by Korean English-major and non-English-major students. *Journal of Multilingual and Multicultural Development*, 40(1), 76-89. <https://doi.org/10.1080/01434632.2018.1480628>
- Lehka-Lemarchand, I. (2007). *Accent de banlieue: approche phonétique et sociolinguistique de la prosodie des jeunes d'une banlieue rouennaise* (Thesis, Université de Rouen). Consulté à l'adresse <http://www.theses.fr/2007ROUEL617>
- Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21(1), 1-36.
- Liberman, A. M., & Mattingly, I. G. (1989). A specialization for speech perception. *Science* 243, 489-494.
- Lifschutz, V. (2015). *Les séries télévisées: une lutte sans fin* (Thesis, Lyon 2). Consulté à l'adresse <http://www.theses.fr/2015LYO20086>
- Lindblom, B. (1983). Economy of Speech Gestures. In P. F. MacNeilage (Éd.), *The Production of Speech* (p. 217-245). https://doi.org/10.1007/978-1-4613-8202-7_10
- Lindblom, B. (1987). Adaptive variability and absolute constancy in speech signals: two themes in the quest for phonetic. *Proceedings of the Eleventh International Congress of Phonetic Sciences*, 3, 9-18. Tallinn, Estonia: ICPhS.
- Lindblom, B. (1990). Explaining Phonetic Variation: A Sketch of the H&H Theory. In W. J. Hardcastle & A. Marchal (Éd.), *Speech Production and Speech Modelling* (p. 403-439). https://doi.org/10.1007/978-94-009-2037-8_16
- Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20(3), 384-422.
- Littlewood, W. (2004). Second language learning. In A. Davies & C. Elder (Éd.), *The Handbook of Applied Linguistics* (p. 501-524). Oxford: Blackwell Publishing Ltd.
- Lively, S. E., Pisoni, D. B., Summers, W. V., & Bernacki, R. H. (1993). Effects of cognitive workload on speech production: Acoustical analyses and perceptual consequences. *The Journal of the Acoustical Society of America*, 93(5), 2962-2973.

- Lorenz, E. N. (1963). Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences*, 20(2), 130-141. [https://doi.org/10.1175/1520-0469\(1963\)020<0130:DNF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2)
- Lorenz, E. N. (1972). Predictability: Does the Flap of a Butterfly's Wings in Brazil Set off a Tornado in Texas. *American Association for the Advancement of Science*. Consulté à l'adresse <https://fermatlibrary.com/s/predictability-does-the-flap-of-a-butterflys-wings-in-brazil-set-off-a-tornado-in-texas>
- Low, E.-L. (2014). *Pronunciation for English as an international language: From research to practice*. New York: Routledge.
- Lowie, W. (2017). Emergentism: wide ranging theoretical framework or just one more meta-theory? *Recherches en didactique des langues et des cultures. Les cahiers de l'Acedle*, 14(14-1). <https://doi.org/10.4000/rdlc.1140>
- Lowie, W., Caspi, T., van Geert, P., & Steenbeek, H. (2011). Modeling development and change. In M. Verspoor, K. de Bot, & W. Lowie (Éd.), *A Dynamic Approach to Second Language Development: Methods and Techniques* (p. 99-121). Amsterdam/Philadelphia: John Benjamins Publishing.
- Luzar, L. (2012). Sur la possibilité de corriger une fossilisation. *Synergies Argentine*, (1), 53-58.
- MacDonald, J. (2018). Hearing Lips and Seeing Voices: The Origins and Development of the 'McGurk Effect' and Reflections on Audio-Visual Speech Perception Over the Last 40 Years. *Multisensory Research*, 31(1-2), 7-18. <https://doi.org/10.1163/22134808-00002548>
- Maeda, M. (2009). What varieties of English pronunciation are Japanese learners learning? *Studies in Language, Kanagawa University*, 32, 1-31.
- Malakolunthu, S., & Rengasamy, N., C. (Éd.). (2017). *Policy Discourses in Malaysian Education: A nation in the making*. London, New York: Routledge.
- Male, T. V. (2010). *Liaisons généreuses : L'apport du français à la langue anglaise*. Paris: Arléa.
- Male, T. V. (2015). *Made in français : Les mots anglais venus du français*. Paris: Serendib.
- Malécot, A. (1960). Vowel Nasality as a Distinctive Feature in American English. *Language*, 36(2), 222-229. <https://doi.org/10.2307/410987>
- Malmberg, B. (1990). *La phonétique* (2^e éd.). Paris: Presses Universitaires de France.
- Marchal, A. (2007). *La production de la parole*. Paris: Hermes Science Publications.
- Markham, D. (1997). *Phonetic imitation, accent, and the learner* (Travaux de l'institut de Linguistic de Lund, Vol. 33). Lund, Sweden: Lund University Press.
- Matter, J. F. (2006). La prononciation authentique en langue étrangère : Un problème négligé. *Revue française de linguistique appliquée, Vol. XI*(1), 21-32.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588), 746-748.
- Melchers, G., & Shaw, P. (2013). *World Englishes*. Routledge.
- Melquiond, A. (En préparation). *Apocalypse et fin du monde dans les séries télévisées américaines* (Paris 10). Consulté à l'adresse <http://www.theses.fr/s122703>
- Meltzoff, A. N., & Moore, M. K. (1977). Imitation of Facial and Manual Gestures by Human Neonates. *Science*, 198(4312), 75-78.
- Meltzoff, A. N., & Moore, M. K. (2005). Imitation et développement humain : les premiers temps de la vie. *Terrain. Anthropologie & sciences humaines*, (44), 71-90. <https://doi.org/10.4000/terrain.2455>
- Meltzoff, A. N., & Williamson, R. A. (2010). The Importance of Imitation for Theories of Social-Cognitive Development. In G. Bremner & T. D. Wachs (Éd.), *The Wiley-Blackwell Handbook of Infant Development* (p. 345-364). <https://doi.org/10.1002/9781444327564.ch11>

- Meltzoff, A. N., Kuhl, P. K., Movellan, J., & Sejnowski, T. J. (2009). Foundations for a New Science of Learning. *Science*, 325(5938), 284-288. <https://doi.org/10.1126/science.1175626>
- Mendoza-Denton, N. (2011). The semiotic hitchhiker's guide to creaky voice: Circulation and gendered hardcore in a Chicana/o gang persona. *Journal of Linguistic Anthropology*, 21(2), 261-280.
- Meunier, A. (2018). L'éducation dans les musées : une forme d'éducation non formelle. In D. Jacobi, *Public et Culture. Culture et éducation non formelle* (p. 15-32). Quebec: PUQ.
- Meunier, A., Belleville, B., & Grant, O. (2018). Penser avec ses mains ou la philosophie du tinkering. *La Lettre de l'OCIM. Musées, Patrimoine et Culture scientifiques et techniques*, (177), 29-35. <https://doi.org/10.4000/ocim.2557>
- Miras, G. (2013). « Enseigner / apprendre » la prononciation autrement : une approche psychosociale musique-parole. *Recherches en didactique des langues et des cultures. Les cahiers de l'Acedle*, 10(10-1). <https://doi.org/10.4000/rdlc.1508>
- Miras, G. (2014). Approche plurielle des liens musique-parole pour la didactique de la prononciation du français comme langue étrangère/seconde (Thèse de doctorat). Université de la Sorbonne Nouvelle, Paris, France.
- Miras, G. (2016). Facteurs environnementaux et représentations sur la lecture-écriture. Une étude transdisciplinaire: musique, parole et langue des signes. *Dossiers d'HEL*, 9, 61-80.
- Miras, G. (2017). Réponse à Joëlle Aden : le corps à la lumière de l'émergentisme linguistique. *Recherches en didactique des langues et des cultures. Les cahiers de l'Acedle*, 14(14-1). <https://doi.org/10.4000/rdlc.1442>
- Miras, G. (2019). De la correction à la médiation : la doxa terminologique en didactique de la prononciation du français comme langue étrangère. *Recherches en didactique des langues et des cultures. Les cahiers de l'Acedle*, 16(16-1). <https://doi.org/10.4000/rdlc.4298>
- Miras, G., & Narcy-Combes, J.-P. (2014). Conséquences sur les pratiques d'une prise en compte intégrée des théories socio-constructivistes et émergentiste. *Travaux & documents*, (46), 15-25.
- Mitra, S., Tooley, J., Inamdar, P., & Dixon, P. (2003). Improving English pronunciation: An automated instructional approach. *Information Technologies & International Development*, 1(1), pp-75.
- Mitterer, H., & McQueen, J. M. (2009). Foreign Subtitles Help but Native-Language Subtitles Harm Foreign Speech Perception. *PLOS ONE*, 4(11), e7785. <https://doi.org/10.1371/journal.pone.0007785>
- Modiano, M. (1996). A Mid-Atlantic Handbook : American and British English. Studentlitteratur.
- Mok, P. P. K., Lee, A., Li, J. J., & Xu, R. B. (2018). Orthographic effects on the perception and production of L2 mandarin tones. *Speech Communication*, 101, 1-10. <https://doi.org/10.1016/j.specom.2018.05.002>
- Moradi, F., & Shahrokhi, M. (2014). The Effect of Listening to Music on Iranian Children's Segmental and Suprasegmental Pronunciation. *English Language Teaching*, 7(6), p128. <https://doi.org/10.5539/elt.v7n6p128>
- Moyer, A. (2013). *Foreign Accent : The Phenomenon of Non-native Speech*. Cambridge: Cambridge University Press.
- Narcy-Combes, J.-P., & Boughnim, A. (2011). Télévision et apprentissage de l'italien dans un contexte plurilingue : Pistes pour l'enseignement des langues étrangères à l'école élémentaire. *Recherches en didactique des langues et des cultures. Les cahiers de l'Acedle*, 8(8-1). <https://doi.org/10.4000/rdlc.2288>

- Narcy-Combes, J.-P., & Miras, G. (2012). 40 ans de modélisation en didactique des langues. *Mélanges CRAPEL*, (33), 25-43.
- Navrátilová, M. (2013). *Aspects of Pronunciation Teaching: The Influence of American Media on Pronunciation of Czech Students* (Masarykova univerzita, Pedagogická fakulta). Consulté à l'adresse https://is.muni.cz/th/385293/pdf_b/
- Neufeld, G. G. (1977). Language Learning Ability in Adults : A Study on the Acquisition of Prosodic and Articulatory Features. *Working Papers on Bilingualism*, (12), 44-60.
- Nguyen, N., & Delvaux, V. (2015). Role of imitation in the emergence of phonological systems. *Journal of Phonetics*, 53, 46-54. <https://doi.org/10.1016/j.wocn.2015.08.004>
- Oostendorp, M. van, Ewen, C. J., Hume, E. V., & Rice, K. (Éd.). (2011). *The Blackwell Companion to Phonology*. UK: Wiley-Blackwell.
- Ota, I., & Takano, S. (2014). The media influence on language change in Japanese sociolinguistic contexts. In J. Androutsopoulos (Éd.), *Mediatization and sociolinguistic change* (p. 171-203). Berlin: De Gruyter.
- Over, H., & Gattis, M. (2010). Verbal imitation is based on intention understanding. *Cognitive Development*, 25(1), 46-55. <https://doi.org/10.1016/j.cogdev.2009.06.004>
- Pagel, D., Madeleni, É., & Wioland, F. (2012). *Le rythme du Français parlé*. Paris: Hachette Fle.
- Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America*, 119(4), 2382-2393.
- Pardo, J. S., Gibbons, R., Suppes, A., & Krauss, R. M. (2012). Phonetic convergence in college roommates. *Journal of Phonetics*, 40(1), 190-197.
- Pardo, J. S., Jay, I. C., & Krauss, R. M. (2010). Conversational role influences speech imitation. *Attention, Perception, & Psychophysics*, 72(8), 2254-2264.
- Park, J.-Y. (2013). L'utilisation des séries judiciaires télévisuelles pour l'enseignement de l'anglais juridique : études des séries américaines et anglaises comme supports pédagogiques (Thèse de doctorat). Paris 8, Paris.
- Pavlík, R. (2016). A usage-based account of /r/-liaison in Standard British English. *Journal of Phonetics*, 54, 109-122. <https://doi.org/10.1016/j.wocn.2015.10.001>
- Pepperberg, I. (2007). Grey parrots do not always 'parrot': the roles of imitation and phonological awareness in the creation of new labels from existing vocalizations. *Language Sciences*, 29(1), 1-13.
- Pepperberg, I. M., & Wilkes, S. R. (2004). Lack of referential vocal learning from LCD video by Grey parrots (*Psittacus erithacus*). *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems*, 5(1), 75-97. <https://doi.org/10.1075/is.5.1.05pep>
- Piaget, J. (1962). The Stages of the Intellectual Development of the Child. *Bulletin of the Menninger Clinic*, 26(3), 120-135.
- Piaget, J. (1964). *Six études de psychologie*. Paris: Folio.
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and brain sciences*, 27(02), 169-190.
- Pillot-Loiseau, C., Amelot, A., & Fredet, F. (2010). Apports de la phonétique expérimentale à la didactique de la prononciation du français langue étrangère. Étape 1 : réflexion autour de l'établissement d'un corpus. *Recherche et pratiques pédagogiques en langues de spécialité. Cahiers de l'Apliu*, (Vol. XXIX N° 2), 75-88. <https://doi.org/10.4000/apliut.723>
- Piske, T., MacKay, I. R., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2 : A review. *Journal of phonetics*, 29(2), 191-215.

- Poteaux, N. (2014). Les langues étrangères pour tous à l'université : regard sur une expérience (1991-2013). *Les dossiers des sciences de l'éducation*, (32), 17-32. <https://doi.org/10.4000/dse.644>
- Racine, I. (2015). La liaison chez les apprenants hispanophones avancés de FLE. *Bulletin VALS-ASLA*, (102), 147-167.
- Raynal, F., & Rieunier, A. (1998). *Pédagogie : dictionnaire des concepts clés. Apprentissage, formation, psychologie cognitive* (2^e éd.). Paris: ESF Editeur.
- Reiterer, S. M., Hu, X., Sumathi, T. A., & Singh, N. C. (2013). Are you a good mimic? Neuro-acoustic signatures for speech imitation ability. *Frontiers in psychology*, 4, 782.
- Reuter, Y. (1995). Les relations lecture-écriture dans le champ didactique. *Pratiques : linguistique, littérature, didactique*, (86), 5-23. <https://doi.org/10.3406/prati.1995.1740>
- Richter, F. (2015). Infographic : Americans Are The World's Biggest TV Lovers. Consulté 7 juillet 2019, à l'adresse Statista Infographics website: <https://www.statista.com/chart/4167/tv-viewing-in-2014/>
- Rindal, U. (2010). Constructing identity with L2 : Pronunciation and attitudes among Norwegian learners of English. *Journal of Sociolinguistics*, 14(2), 240-261.
- Rindal, U. (2014). Questioning English standards : Learner attitudes and L2 choices in Norway. *Multilingua*, 33(3-4), 313-334. <https://doi.org/10.1515/multi-2014-0015>
- Rindal, U., & Piercy, C. (2013). Being 'neutral'? English pronunciation among Norwegian learners. *World Englishes*, 32(2), 211-229. <https://doi.org/10.1111/weng.12020>
- Rizzolatti, G. (2005). The mirror neuron system and imitation. In S. Hurley & N. Chater (Éd.), *Perspectives on Imitation. From Neuro-science to social science* (Vol. 1, p. 55-76). Cambridge, Ma: MIT Press.
- Rizzolatti, G., & Buccino, G. (2005). The Mirror Neuron System and Its Role in Imitation and Language. In S. Dehaene, J.-R. Duhamel, M. D. Hauser, & G. Rizzolatti (Éd.), *From Monkey Brain to Human Brain : A Fyssen Foundation Symposium* (p. 213-233). Cambridge, London: MIT Press.
- Rizzolatti, G., & Sinigaglia, C. (2011). *Les Neurones miroirs*. Paris: Odile Jacob.
- Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3(2), 131-141. [https://doi.org/10.1016/0926-6410\(95\)00038-0](https://doi.org/10.1016/0926-6410(95)00038-0)
- Roach, P. (2009). *English Phonetics and Phonology : A Practical Course*. Cambridge University Press.
- Rogalsky, C., Poppa, T., Chen, K.-H., Anderson, S. W., Damasio, H., Love, T., & Hickok, G. (2015). Speech repetition as a window on the neurobiology of auditory-motor integration for speech : A voxel-based lesion symptom mapping study. *Neuropsychologia*, 71, 18-27. <https://doi.org/10.1016/j.neuropsychologia.2015.03.012>
- Rohdenburg, G., & Schliüter, J. (2009). (2009). *One Language, Two Grammars? : Differences Between British and American English*. New York: Cambridge University Press.
- Rosen, É., & Porquier, R. (2003). Présentation. L'actualité des notions d'Interlangue et d'interaction exolingue. *Linx. Revue des linguistes de l'université Paris X Nanterre*, (49), 7-17.
- Rossi, M. (2007). *Audio*. Lausanne: PPUR.
- Roulin, J.-L. (2006). *Psychologie cognitive, Cours, Documents, Exercices* (2e édition; G. Amy & M. Piolat, Éd.). Rosny-sous-Bois: Bréal.
- Roussel, S., & Gaonac'h, D. (2017). *L'apprentissage des langues*. Paris: Retz.
- Saletta, M. (2015). Literacy transforms speech production. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.01458>

- Sanchez, K., Miller, R. M., & Rosenblum, L. D. (2010). Visual influences on alignment to voice onset time. *Journal of Speech, Language, and Hearing Research*, 53(2), 262–272.
- Sato, M., Grabski, K., Granjon, L., Schwartz, J.-L., & Nguyen, N. (2010). Converging to a common speech code: Automatic imitative and perceptuo-motor recalibration processes in speech communication. *2nd Annual Neurobiology of Language Conference (NLC 2010)*. Présenté à San Diego, Californie, United States. San Diego, Californie, United States.
- Saugera, V. (2017). La fabrique des anglicismes. *Travaux de linguistique*, n° 75(2), 59-79.
- Schlichter, H. (1995). *Analyse acoustique comparée de deux imitations vacoles* (Mémoire de D.E.A). Université des Sciences Humaines de Strasbourg, Strasbourg.
- Schmale, R., Seidl, A., & Cristia, A. (2015). Mechanisms underlying accent accommodation in early word learning: Evidence for general expansion. *Developmental science*, 18(4), 664–670.
- Schur, N. W., Ehrlich, E., & Ehrlich, R. (2007). *British English A to Zed*. New York: Facts On File, Inc.
- Schweitzer, K., Walsh, M., Calhoun, S., Schütze, H., Möbius, B., Schweitzer, A., & Dogil, G. (2015). Exploring the relationship between intonation and the lexicon: Evidence for lexicalised storage of intonation. *Speech Communication*, 66, 65-81. <https://doi.org/10.1016/j.specom.2014.09.006>
- Selinker, L. (1972). Interlanguage. *IRAL-International Review of Applied Linguistics in Language Teaching*, 10(1-4), 209–232.
- Selinker, L., Swain, M., & Dumas, M. (1975). The interlanguage hypothesis extended to children. *Language Learning*, 25(1), 139-152.
- Séry. (2010). *Pourquoi les séries télévisées envahissent-elles tous les écrans?* Consulté à l'adresse https://www.lemonde.fr/a-la-une/article/2010/12/28/pourquoi-les-series-televisees-envahissent-elles-les-e-crans_1458468_3208.html
- Shokeir, V. (2008). Evidence for the stable use of uptalk in South Ontario English. *University of Pennsylvania Working Papers in Linguistics*, 14(2), 4.
- Schmoll, L. (2017). Penser l'intégration du jeu vidéo en classe de langue. *Recherche et pratiques pédagogiques en langues de spécialité. Cahiers de l'Apliu*, 36(2). <https://doi.org/10.4000/apliut.5722>
- Silva, H., & Brougère, G. (2016). Le jeu entre situations formelles et informelles d'apprentissage des langues étrangères. *Synergies Mexique*, 6, 57-68.
- Skipper, J. I., van Wassenhove, V., Nusbaum, H. C., & Small, S. L. (2007). Hearing Lips and Seeing Voices: How Cortical Areas Supporting Speech Production Mediate Audiovisual Speech Perception. *Cerebral cortex (New York, N.Y.: 1991)*, 17(10), 2387-2399. <https://doi.org/10.1093/cercor/bhh147>
- Smith, B. L., & Hayes-Harb, R. (2011). Individual differences in the perception of final consonant voicing among native and non-native speakers of English. *Journal of Phonetics*, 39(1), 115–120.
- Smith, N. V. (1973). *The Acquisition Of Phonology: A Case Study*. Cambridge: Cambridge University Press.
- Snoeren, N. D. (2005). Variations phonologiques en production et perception de la parole: le phénomène de l'assimilation (Thèse de Doctorat). Université Paris 5 - René Descartes, Paris.
- So, C. K., & Best, C. T. (2014). Phonetic influences on English and French listeners' assimilation of mandarin tones to native prosodic categories. *Studies in Second Language Acquisition*, 36(2), 195-221. <https://doi.org/10.1017/S0272263114000047>
- Sock, R. (1998). *Organisation temporelle en production de la parole: Émergence de catégories sensori-motrices phonétiques* (Doctorat d'Etat ès Lettres et Sciences Humaines). Université Stendhal-Grenoble 3, Grenoble.

- Sock, R. (2001). La théorie de la viabilité en production-perception de la parole. In D. Keller, J. P. Durafour, J. P. F. Bonnot, & R. Sock (Éd.), *Percevoir : Monde et langage. Invariance et variabilité du sens vécu* (p. 285-316). Hayen, Sprimont: Édition Mardaga.
- Sock, R., & Vaxelaire, B. (2001). Peut-on travailler sans représentations en production-perception de la parole ? In C. Buridant, G. Kleiber, & J.-C. Pellat (Éd.), *Par monts et par vaux : itinéraires linguistiques et grammaticaux* (p. 379-392). Louvain-Paris: Éditions Peeters.
- Sock, R., & Vaxelaire, B. (2001b). Réflexions sur le timing de la quantité. *Travaux de l'Institut de Phonétique de Strasbourg, TIPS*, 31, 89-126.
- Sock, R., & Vaxelaire, B. (2004). Le diable perceptif dans les détails sensori-moteurs anticipatoires. In R. Sock & B. Vaxelaire (Éd.), *L'anticipation: à l'horizon du présent* (p. 141-158). Liège: Mardaga.
- Sock, R., & Vaxelaire, B. (2010). Spéciale, la parole. In D. Huck & T. Choremi (Éd.), *Parole(s) et langue(s), espaces et temps. Mélanges offerts à Arlette Bothorel-Witz* (p. 141-184). Strasbourg: Université de Strasbourg.
- Sock, R., Vaxelaire, B., & Fauth, C. (Éd.). (2014). *La voix et la parole perturbées, Travaux en Phonétique Clinique* (Actes des 4e Journées de Phonétique Clinique, mai 2011). Mons: Centre International De Phonétique Appliquée De Mons.
- Sockett, G. (2011a). From the cultural hegemony of English to online informal learning: Cluster frequency as an indicator of relevance in authentic documents. *ASp. La Revue Du GERAS*, (59), 5-20. <https://doi.org/10.4000/asp.2469>
- Sockett, G. (2011b). Les processus cognitifs de résolution de problèmes pour l'apprentissage des langues dans des environnements multimédias. Apprentissage informel et réseaux sociaux. *Recherches en didactique des langues et des cultures. Les cahiers de l'Acedle*, 8(8-1). <https://doi.org/10.4000/rdlc.2237>
- Sockett, G. (2012). Le web social – La complexité au service de l'apprentissage informel de l'anglais. *Alsic. Apprentissage des Langues et Systèmes d'Information et de Communication*, (Vol. 15, n° 2). <https://doi.org/10.4000/alsic.2505>
- Sockett, G. (2014). *The Online Informal Learning of English*. Basingstoke: Palgrave Macmillan.
- Sockett, G., & Kusyk, M. (2013). L'apprentissage informel en ligne: nouvelle donne pour l'enseignement-apprentissage de l'anglais. *Recherche et pratiques pédagogiques en langues de spécialité. Cahiers de l'Aplut*, (Vol. XXXII N° 1), 75-91. <https://doi.org/10.4000/aplut.3578>
- Sockett, G., & Kusyk, M. (2015). Online informal learning of English: frequency effects in the uptake of chunks of language from participation in web-based activities. In *Usage-Based Perspectives on Second Language Learning*. Consulté à l'adresse <https://www.degruyter.com/view/books/9783110378528/9783110378528-009/9783110378528-009.xml>
- Sockett, G., & Toffoli, D. (2012). Beyond learner autonomy: A dynamic systems view of the informal learning of English in virtual online communities. *Procedia - Social and Behavioral Sciences*, 34, 212-215. <https://doi.org/10.1016/j.sbspro.2012.02.042>
- Speech Production and Perception. Basingstoke: Palgrave Macmillan.
- Statista. (2012). On television, do you prefer French series or American TV shows. [Sondages]. Consulté 7 juillet 2019, à l'adresse Statista website: <https://www.statista.com/statistics/767152/series-americaaines-francaises-preference-age-la-france/>

- Statista. (2019b, juillet 7). Daily television viewing time in France 2017 | Statistic. Consulté 7 juillet 2019, à l'adresse Statista website: <https://www.statista.com/statistics/381558/daily-television-viewing-time-of-individuals-in-france/>
- Steiblé, L. (2014). Le contrôle temporel des consonnes occlusives de l'alsacien et du français parlé en Alsace (Thèse de doctorat). Université de Strasbourg, Strasbourg.
- Stevens, K. N. (1985). Evidence for the role of acoustic boundaries in the perception of speech sounds. *The Journal of the Acoustical Society of America*, 69(1), 243-255. <https://doi.org/10.1121/1.386489>
- Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics*, 17, 3-45.
- Stevens, K. N. (2002). Toward a model for lexical access based on acoustic landmarks and distinctive features. *The Journal of the Acoustical Society of America*, 111(4), 1872-1891. <https://doi.org/10.1121/1.1458026>
- Stevens, K. N., & Blumstein, S. E. (1978). Invariant cues for place of articulation in stop consonants. *The Journal of the Acoustical Society of America*, 64(5), 1358-1368. <https://doi.org/10.1121/1.382102>
- Stuart-Smith, J., & Ota, I. (2014). Media models, 'the shelf', and stylistic variation in East and West. In J. Androutsopoulos (Éd.), *Mediatization and sociolinguistic change* (p. 127-170). Berlin: De Gruyter.
- Stuart-Smith, J., Pryce, G., Timmins, C., & Gunter, B. (2013). Television can also be a factor in language change : Evidence from an urban dialect. *Language*, 89(3), 501-536.
- Stuart-Smith, J., Smith, R., Rathcke, T., Li Santi, F., & Holmes, S. (2011). Responding to accents after experiencing interactive or mediated speech. *ICPhS XVII*, 1914-1917. Hong Kong.
- Stuart-Smith, J., Timmins, C., & Tweedie, F. (2007). 'Talkin'Jockney'? Variation and change in Glaswegian accent. *Journal of Sociolinguistics*, 11(2), 221-260.
- Summers, W. V., Pisoni, D. B., Bernacki, R. H., Pedlow, R. I., & Stokes, M. A. (1988). Effects of noise on speech production : Acoustic and perceptual analyses. *Journal of the Acoustical Society of America*, 84(3), 917-928. <https://doi.org/10.1121/1.3966660>
- Sun, Y., Franklin, T., & Gao, F. (2017). Learning outside of classroom: Exploring the active part of an informal online English learning community in China. *British Journal of Educational Technology*, 48(1), 57-70.
- Suttora, C., Salerni, N., Zanchi, P., Zampini, L., Spinelli, M., & Fasolo, M. (2017). Relationships between structural and acoustic properties of maternal talk and children's early word recognition. *First Language*, 37(6), 612-629. <https://doi.org/10.1177/0142723717714946>
- Svartvik, J., & Leech, G. (2006). *English : One Tongue, Many Voices*. New York: Palgrave MacMillan.
- Swain, M. (1985). Communicative competence: some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Éd.), *Input in Second Language Acquisition* (p. 235-256). New York: Newbury House.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidelhofer (Éd.), *Principle and Practice in Applied Linguistics: Studies in Honor of H.G. Widdowson* (p. 125-144). Oxford: Oxford University Press.
- Swain, M. (2018). The Output Hypothesis : Just Speaking and Writing Aren't Enough. *Canadian Modern Language Review*. <https://doi.org/10.3138/cmlr.50.1.158>
- Tan, K., Ng, M., & Saw, K. (2010). Online activities and writing practices of urban Malaysian adolescents. *Science Direct System*, 38(1), 548-559.
- Tatham, M., & Morton, K. (2011). *Guide to Speech Production and Perception*. Edinburgh: Edinburgh University Press.

- Tatham, M., & Morton, K. (2016). *Speech Production and Perception*. Basingstoke: Palgrave Macmillan.
- Tiberghien, G., Roulin, J.-L., & Jean-Léon Beauvois (Éd.). (1992). Ecoutez voir ! « L'effet McGurk » ou les illusions audiovisuelles de la parole. In *Manuel d'études pratiques de psychologie* (p. 291-302). Paris: PUF.
- Tiberghien, G., Roulin, J.-L., & Jean-Léon Beauvois (Éd.). (1992). Ecoutez voir ! « L'effet McGurk » ou les illusions audiovisuelles de la parole. In *Manuel d'études pratiques de psychologie* (p. 291-302). Paris: PUF.
- Toffoli, D. (2018). *L'apprenant.e de langue 2020 : profil, dynamiques, dispositifs* (Habilitation à Diriger des Recherches). Université de Lille, Lille.
- Toffoli, D., & Perrot, L. (2017). Autonomy, the Online Informal Learning of English (OILE) and Learning Resource Centres (LRCs): Relationships between Learner Autonomy, L2 Proficiency, L2 Autonomy and Digital Literacy. In Cappellini, M, Lewis, T, & Rivens Mompean, A. (Éd.), *Learner Autonomy and Web 2.0* (p. 198-228). Sheffield: Equinox.
- Toffoli, D., & Sockett, G. (2010). How non-specialist students of English practice informal learning using web 2.0 tools. *ASP. La Revue Du GERAS*, (58), 125-144. <https://doi.org/10.4000/asp.1851>
- Toffoli, D., & Sockett, G. (2015a). L'apprentissage informel de l'anglais en ligne (AIAL) : quelles conséquences pour les centres de ressources en langues ? *Recherche et pratiques pédagogiques en langues de spécialité. Cahiers de l'ApliuT*, (Vol. XXXIV N° 1), 147-165. <https://doi.org/10.4000/apliut.5055>
- Toffoli, D., & Sockett, G. (2015b). University teachers' perceptions of Online Informal Learning of English (OILE). *Computer Assisted Language Learning*, 28(1), 7-21. <https://doi.org/10.1080/09588221.2013.776970>
- Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, Mass.: Harvard University Press.
- Trenité, G. N. (1920). *Drop Your Foreign Accent: Engelsche Uitspraakoefeningen*. Amsterdam: Willink.
- Trinder, R. (2017). Informal and deliberate learning with new technologies. *Elt Journal*, 71(4), 401-412. <https://doi.org/10.1093/elt/ccw117>
- Troubetzkoy, N. S. (1939). *Principes de phonologie*. Paris 2: Klincksieck.
- Trudgill, P. (1986). *Dialects in Contact*. Oxford: Blackwell Publishers.
- UNESCO. (2012). *Lignes directives de L'UNESCO pour la reconnaissance, la validation et l'accréditation des acquis de l'apprentissage non formel et informel*. Hambourg: Institut de l'UNESCO pour l'apprentissage tout au long de la vie.
- Vaissière, J. (2006). *La phonétique*. Paris: Presses universitaires de France.
- Vaissière, J. (2015). *La phonétique* (2^e éd.). Paris: Presses universitaires de France.
- Van Dievort, C., & Duchâteau, J.-P. (2014). Pourquoi les séries télé ont-elles tant de succès ? Consulté 7 juillet 2019, à l'adresse La Libre.be website: <https://www.lalibre.be/debats/ripostes/pourquoi-les-series-tele-ont-elles-tant-de-succes-52e509033570e5b8eee76009>
- van Dijk, M., Verspoor, M., & Lowie, W. (2011). Variability and Dynamic Systems Theory. In M. Verspoor, K. D. Bot, & W. Lowie (Éd.), *A Dynamic Approach to Second Language Development: Methods and Techniques* (p. 55-84). Amsterdam/Philadelphia: John Benjamins Publishing.

- van Geert, P. (2009). Systèmes dynamiques et modélisation dans l'étude de l'acquisition du langage : tendances, transitions et fluctuations. In M. Kail, M. Fayol, & M. Hickman (Éd.), *Apprentissage des langues*. Paris: CNRS Éditions.
- Varela, F., & Bourgine, P. (Éd.). (1992). *Toward a practice of autonomous systems: proceedings of the first European Conference on Artificial Life*. Cambridge, Mass.: MIT Press.
- Vaxelaire, B. (1993). Étude comparée des effets des variations de débit, lent, rapide, sur les paramètres articulatoires, à partir de la cinéradiographie (sujets français) (Thèse de Doctorat Nouveau régime). Université Marc Bloch, Strasbourg 2, Strasbourg.
- Vaxelaire, B. (1994). Variation de geste et débit : contribution à une base de données sur la production de la parole, mesures cinéradiographiques, groupes consonantiques en français. *Travaux de l'Institut de Phonétique de Strasbourg, TIPS*, 24, 109-146.
- Vaxelaire, B. (1995). Étude comparée des effets des variations de débit, lent, rapide, sur les paramètres articulatoires, à partir de la cinéradiographie (sujets français). *Travaux de l'Institut de Phonétique de Strasbourg, TIPS*, 25, 67-93.
- Vaxelaire, B. (2006). La Résistivité spatio-temporelle des gestes linguistiques. Ou perturber le linguistique en augmentant la vitesse d'élocution. In B. Vaxelaire, R. Sock, G. Kleiber, & F. Marsac (Éd.), *Perturbations et réajustements : langue et langage*. Strasbourg: Publications de l'Université Marc Bloch.
- Vaxelaire, B. (2007). *Le geste et la production de la parole. Résultats et implications d'études cinéradiographiques quantitatives* (Habilitation à Diriger des Recherches). Université Marc Bloch, Strasbourg 2, Strasbourg.
- Vaxelaire, B. (2010). Méthodes de recherche en production de la parole : Entreprendre des investigations à partir de la cinéradiographie. In D. Huck & T. Choremi (Éd.), *Parole(s) et langue(s), espaces et temps. Mélanges offerts à Arlette Bothorel-Witz* (p. 149-166). Strasbourg: Université de Strasbourg.
- Viel, M. (1981). *La phonétique de l'anglais*. Paris: PUF.
- Voirol, M. (2006). *Anglicismes et anglomanie* (3e éd.). Paris: CFPJ victoires.
- Vygotsky, L. S. (1978). *Thought and language*. Cambridge, Ma: Harvard University Press.
- Vygotsky, L. S. (1978). *Mind in Society*. Cambridge, Ma: MIT Press.
- Walsh, O. (2014). 'Les anglicismes polluent la langue française'. Purist attitudes in France and Quebec. *Journal of French Language Studies*, 24(3), 423-449. <https://doi.org/10.1017/S0959269513000227>
- Walsh, O. (2015). Attitudes towards English in France. In A. Linn, N. Bermel, & G. Ferguson (Éd.), *Attitudes towards English in Europe* (Language and Social Life, p. 27-54). Berlin, Boston: De Gruyter Mouton.
- Walsh, O. M. (2013). *Linguistic purism in France and Quebec* (Ph.D., University of Cambridge). Consulté à l'adresse <https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.608141>
- Wanick-Klimczak, E., & Shockey, L. R. (2012). *Teaching and Researching English Accents in Native and Non-native Speakers*. Springer Science & Business Media.
- Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, 41(8), 989-994. [https://doi.org/10.1016/s0028-3932\(02\)00316-0](https://doi.org/10.1016/s0028-3932(02)00316-0)
- Webb, S. A., & Chang, A. C.-S. (2012). Second Language Vocabulary Growth. *RELC Journal*, 43(1), 113-126. <https://doi.org/10.1177/0033688212439367>
- Wells, J. (2008). *Longman Pronunciation Dictionary* (1^{re} éd.). Harlow: Pearson Longman.

- Wells, J. C. (1982). *Accents of English*. Cambridge: Cambridge University Press.
- Werquin, P. (2010). Reconnaître l'apprentissage non formel et informel : Résultats, politiques et pratiques. Paris: OECD Publishing.
- Williamson, R. A., & Meltzoff, A. N. (2011). Own and others' prior experiences influence children's imitation of causal acts. *Cognitive Development*, 26(3), 260-268. <https://doi.org/10.1016/j.cogdev.2011.04.002>
- Williamson, R. A., Jaswal, V. K., & Meltzoff, A. N. (2010). Learning the rules: Observation and imitation of a sorting strategy by 36-month-old children. *Developmental Psychology*, 46(1), 57-65. <https://doi.org/10.1037/a0017473>
- Winnykamen, F. (1990). *Apprendre en imitant ?* Paris: PUF.
- Wioland, F. (1991). *Prononcer les mots du français : Des sons et des rythmes*. Consulté à l'adresse <http://phonetiquedufle.canalblog.com/archives/2013/05/03/27067373.html>
- Whyte, S. (2019). Revisiting Communicative Competence in the Teaching and Assessment of Language for Specific Purposes. *Language Education & Assessment*, 2(1), 1-19.
- Yoccoz, J.-C. (1994). Recent Developments in Dynamics. In S. D. Chatterji (Éd.), *Proceedings of the International Congress of Mathematicians* (p. 246-265). Zürich, Switzerland.
- Yoccoz, J.-C. (1995). Introduction to Hyperbolic Dynamics. In B. Branner & P. Hjorth (Éd.), *Real and Complex Dynamical Systems* (p. 265-291). https://doi.org/10.1007/978-94-015-8439-5_11
- Young, R. (1996). Form-function relations in articles in English interlanguage. In R. Bayley & D. Preston, *Second language acquisition and linguistic variation* (p. 135-175). Amsterdam, Philadelphia: John Benjamins Publishing.
- Yuasa, I. P. (2010). Creaky voice : A new feminine voice quality for young urban-oriented upwardly mobile American women? *American Speech*, 85(3), 315-337.
- Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring Vocabulary through Reading: Effects of Frequency and Contextual Richness. *Canadian Modern Language Review*, 57(4), 541-572.
- Zajac, M. (2015). Phonetic convergence in the speech of Polish learners of English (Thèse de Doctorat). The University of Lodz, Lodz.
- Zaouali, H. (2019). *Etude acoustique de la production de la parole chez les patients glossectomisés* (Thèse de Doctorat). Université de Strasbourg, Strasbourg.
- Zhang, Q., & Damian, M. F. (2012). Effects of Orthography on Speech Production in Chinese. *Journal of Psycholinguistic Research*, 41(4), 267-283. <https://doi.org/10.1007/s10936-011-9193-z>

Table des matières

Remerciements	ii
Sommaire.....	iv
Table des abréviations.....	vii
Introduction générale.....	9
1 Contexte et problématique de la thèse	12
1.1 Types d'apprentissage : formel, non formel et informel.....	13
1.2 L'apprentissage informel des langues.....	18
1.3 L'apprentissage informel de l'anglais en ligne.....	19
1.3.1 Pour un périmètre des définitions de l'AIAL.....	19
1.3.2 Les objectifs de l'AIAL.....	23
1.3.3 AIAL en France et ailleurs : quelques études	24
1.3.4 L'influence des médias sur la prononciation	26
1.4 Objectifs de cette étude.....	31
1.5 Questions de recherche et Hypothèses	32
1.6 Conclusion.....	34
2 Autour de la langue anglaise.....	36
2.1 L'importance de la langue anglaise	36
2.2 Le français et l'anglais : des influences réciproques	39
2.2.1 L'anglicisme en français	39
2.2.2 Le gallicisme en anglais	43
2.3 L'anglais dans le monde : les différents types d'anglais.....	45
2.3.1 L'anglais britannique / américain vs. la RP / le GA.....	47
2.3.2 Diversité linguistique dans les îles britanniques.....	48
2.3.3 Diversité linguistique aux États-Unis.....	52
2.3.4 Différences linguistiques entre la RP et le GA	56
2.4 Différences phonétiques et phonologiques entre l'anglais et le français.....	61
2.5 Conclusion.....	66
3 Théorie des systèmes complexes et dynamiques	68
3.1 Les origines de la théorie des SCD.....	69
3.2 La théorie des SCD aujourd'hui.....	71
3.3 Caractéristiques de la théorie des SCD.....	74
3.3.1 Complexité et interconnectivité	75
3.3.2 Non-linéarité et linéarité d'une trajectoire.....	77
3.3.3 État initial et état final	78
3.3.4 Dynamisme et stabilité.....	79
3.3.5 Prédicibilité et imprédictibilité.....	81
3.3.6 Autorégulation / auto-organisation	81
3.3.7 Attracteur et répulsif.....	82
3.3.8 Émergence et coadaptation.....	84
3.4 La théorie des SCD dans le domaine de l'apprentissage des langues	86
3.4.1 La théorie des SCD et la notion d'interlangue	90
3.4.2 La théorie des SCD, les différences individuelles et la variabilité	93
3.5 Limites de la théorie des SCD en didactique des langues.....	96
3.5.1 Problèmes d'importation de théories d'une science à une autre	96
3.5.2 Problèmes de méthodologies de recherche	99
3.5.3 Pistes de solutions méthodologiques dans le cadre de la théorie des SCD.....	100
3.6 Conclusion.....	102
4 Théorie de l'imitation vocale	105
4.1 Imitation : cas général	106
4.2 Imitation du point de vue neurophysiologique : le rôle des neurones miroirs	109
4.3 Imitation générale et imitation vocale	113

4.3.1	Imitation vocale	114
4.3.2	Imitation phonétique et imitation phonologique	115
4.4	Imitation vocale volontaire et involontaire	116
4.4.1	Imitation vocale volontaire	117
4.4.2	Imitation vocale involontaire	123
4.4.3	Théorie de l'exemplaire	126
4.5	Imitation directe (non-médiée) et indirecte (médiée).....	130
4.6	Des modèles d'acquisition phonétique et phonologique	133
4.6.1	Acoustic Image Imprinting.....	134
4.6.2	Speech Learning Model.....	136
4.6.3	Perceptual Assimilation Model	137
4.6.4	Native Language Magnet.....	139
4.7	Conclusion.....	141
5	Viabilité et perturbation en production-perception de la parole.....	144
5.1	Quelques définitions : fiabilité, variabilité, invariance et viabilité	145
5.2	La théorie Quantique : une théorie de l'invariance	147
5.3	La théorie de la Variabilité Adaptative.....	149
5.4	La théorie de la viabilité en production-perception de la parole	153
5.4.1	Zones de viabilité ou espaces de possibilités fiables.....	156
5.4.2	Modélisation de la production-perception de la parole dans le cadre de la théorie de la viabilité. 161	
5.5	Perturbation du système de production de la parole	163
5.5.1	Perturbateur écologique ou naturelle.....	165
5.5.2	Perturbation non écologique ou « artificielle »	166
5.5.3	Conséquences des perturbations sur la production de la parole	167
5.5.4	Résistivité dans le cadre de la perturbation en production de la parole	169
5.5.5	Lien entre la théorie de viabilité en phonétique et la notion de perturbation	170
5.6	Théorie de la viabilité en phonétique : implications pour l'apprentissage des L2.....	173
5.7	Conclusion.....	174
6	Cadre méthodologique	176
6.1	Recrutement des participants	176
6.2	Matériel de collecte de données audio.....	177
6.3	La collecte de données	178
6.4	Élaboration du support pour la prise de sons.....	178
6.4.1	Les mots	179
6.4.2	Les phrases	181
6.4.3	Le texte	182
6.5	Enregistrement des participants.....	184
6.5.1	Présentation orale.....	185
6.5.2	Vitesses d'élocution	185
6.5.3	Nombre de répétitions.....	186
6.5.4	Recodage et anonymisation de données.....	187
6.6	Segmentation et mesures acoustiques des sons cibles.....	187
6.6.1	Éléments consonantiques.....	188
6.6.2	Éléments vocaliques.....	194
6.6.3	Prosodie – Intonation	200
6.7	Test de perception.....	202
6.8	Questionnaire	205
6.8.1	Contact avec la langue anglaise.....	206
6.8.2	Les séries télévisées	207
6.8.3	L'adoption d'accent	207
6.9	Conclusion.....	208
7	Résultats de la prononciation en vitesse normale.....	210
7.1	La prononciation du /t/ intervocalique.....	212
7.1.1	City	212

7.1.2	Vitamin.....	213
7.1.3	Party.....	214
7.1.4	Résultats du /t/ intervocalique.....	215
7.1.5	Discussion.....	216
7.2	La rhoticité.....	217
7.2.1	More.....	217
7.2.2	Organisation.....	218
7.2.3	Dark.....	219
7.2.4	Party.....	220
7.2.5	Bird.....	221
7.2.6	Actor.....	222
7.2.7	Leisure.....	223
7.2.8	Résultats d'ensemble pour la rhoticité.....	224
7.2.9	Discussion.....	225
7.3	Le yod-dropping.....	225
7.3.1	Tune.....	226
7.3.2	New.....	226
7.3.3	Evolution.....	227
7.3.4	Résultats d'ensemble pour le yod-dropping.....	228
7.3.5	Discussion.....	229
7.4	La prononciation de la voyelle de LOT.....	231
7.4.1	Hot.....	231
7.4.2	Top.....	232
7.4.3	Hot et top.....	233
7.5	La prononciation de la voyelle de GOAT.....	234
7.5.1	Joke.....	234
7.5.2	Road.....	235
7.5.3	Joke et road.....	236
7.5.4	Discussion.....	237
7.6	La prononciation des graphies <i> et <y>.....	238
7.6.1	Dynasty.....	238
7.6.2	Vitamin.....	239
7.6.3	Dynasty et vitamin.....	239
7.7	La prononciation de la graphie <i>.....	240
7.7.1	Organisation.....	240
7.7.2	Legalisation.....	241
7.7.3	Organisation et legalisation.....	242
7.7.4	Discussion.....	243
7.8	La prononciation de la graphie <ei> dans leisure.....	244
7.9	La prononciation de la graphie <e> dans evolution.....	245
7.9.1	Discussion.....	246
7.10	Nasalisation vocalique.....	247
7.10.1	Band.....	247
7.10.2	Hand.....	248
7.10.3	Band et hand.....	249
7.10.4	Discussion.....	250
7.11	La place de l'accent principal.....	252
7.11.1	Donate.....	252
7.11.2	Locate.....	253
7.11.3	Donate et locate.....	254
7.11.4	Discussion.....	255
7.12	Le nombre de syllabes accentuées.....	256
7.12.1	Military.....	257
7.12.2	Necessary.....	257

7.12.3	Military et necessary	258
7.12.4	Discussion.....	259
7.13	Résultats globaux des mots à vitesse normale.....	261
7.13.1	Discussion.....	263
7.14	Conclusion.....	264
8	Résultats de la prononciation en vitesse rapide	266
8.1	La prononciation du /t/ intervocalique.....	268
8.1.1	City	268
8.1.2	Vitamin.....	269
8.1.3	Party	270
8.1.4	Résultats du /t/ intervocalique	271
8.2	La rhoticité.....	272
8.2.1	More	272
8.2.2	Organisation.....	273
8.2.3	Dark.....	274
8.2.4	Party	275
8.2.5	Bird	276
8.2.6	Actor.....	277
8.2.7	Leisure	278
8.2.8	Résultats d'ensemble pour la rhoticité	279
8.3	Le yod-dropping.....	280
8.3.1	Tune	280
8.3.2	New.....	281
8.3.3	Evolution.....	282
8.3.4	Résultats d'ensemble pour le yod-dropping.....	283
8.4	La prononciation de la voyelle de LOT.....	285
8.4.1	Hot.....	286
8.4.2	Top.....	286
8.4.3	Hot et top	287
8.5	La prononciation de la voyelle de GOAT.....	288
8.5.1	Joke.....	288
8.5.2	Road	289
8.5.3	Joke et road.....	290
8.6	La prononciation des graphies <i> et <y>.....	291
8.6.1	Dynasty.....	291
8.6.2	Vitamin.....	292
8.6.3	Dynasty et vitamin.....	293
8.7	La prononciation de la graphie <i>	294
8.7.1	Organisation.....	294
8.7.2	Legalisation	296
8.7.3	Organisation et legalisation.....	297
8.8	La prononciation de la graphie <ei> dans leisure	298
8.9	La prononciation de la graphie <e> dans evolution.....	299
8.10	Nasalisation vocalique.....	300
8.10.1	Band	301
8.10.2	Hand.....	302
8.10.3	Band et hand	303
8.11	La place de l'accent principal	305
8.11.1	Donate	305
8.11.2	Locate	306
8.11.3	Donate et locate	307
8.12	La voyelle de la pénultième syllabe et le nombre de syllabes accentuées.....	308
8.12.1	Military	309
8.12.2	Necessary	310

8.12.3	Military et necessary	311
8.13	Résultats globaux des mots à vitesse rapide	313
8.13.1	Discussion générale.....	316
8.14	Conclusion.....	318
9	Résultats du questionnaire et du test de perception	321
9.1	Informations biographiques	322
9.2	Temps d'apprentissage et temps d'utilisation de l'anglais.....	323
9.3	Séjours et leur durée en pays anglophones	326
9.4	Nationalité et accent des enseignants	327
9.4.1	Nationalités des enseignants.....	327
9.4.2	Accents des enseignants	328
9.5	Mots inconnus.....	329
9.6	Imitation d'accent	331
9.7	Pratiques d'exposition à l'anglais via les séries télévisées.....	333
9.7.1	Modes de visionnage de séries.....	336
9.7.2	Raisons de visionnage de séries.....	338
9.7.3	Intérêts pour le séries.....	339
9.7.4	Supports utilisés pour le visionnage.....	341
9.7.5	Séries télévisées visionnées	342
9.7.6	Accents dominants dans les séries	346
9.7.7	Discussion générale sur les résultats du questionnaire	347
9.8	Résultats du test de perception	351
9.8.1	Discussion.....	353
9.9	Triangulation de données.....	354
9.9.1	Profilage des locuteurs.....	354
9.9.2	Corrélations entre les variables	361
9.10	Conclusion.....	366
10	Synthèse de la recherche et conclusion générale.....	369
10.1	Retour sur les questions et hypothèses de départ	369
10.2	Discussion des résultats en lien avec le cadre théorique	373
10.2.1	Systèmes complexes et dynamiques	373
10.2.2	Imitation vocale	375
10.2.3	Viabilité et perturbation.....	377
10.3	Apports de notre travail de recherche.....	379
10.4	Limites et perspectives de ce travail de recherche	381
	Conclusion générale.....	386
	Bibliographie.....	389
	Table des matières.....	410
	Table des figures	415
	Tables des tableaux	418
	Index des notions.....	419

Table des figures

Figure 1 : apprentissage de l'anglais, du français, de l'allemand et de l'espagnol au premier cycle du secondaire dans l'UE, entre 2005 et 2012 (en %) (Bourgeois & Gheoghiu, 2014, paragr. 1).	37
Figure 2 : le cercle des « anglais » du monde de McArthur (1998, p. 97 cité dans Jenkins, 2003, p. 18).	47
Figure 3 : quelques exemples de théories et disciplines qui composent la théorie des SCD.	72
Figure 4 : illustration d'un système complexe (réseau Internet) avec les connections entre les divers constituants.	75
Figure 5 : attracteur étrange de Lorenz (1963).	82
Figure 6 : un bébé de moins d'un mois imitant les expressions faciales d'un adulte. (Meltzoff <i>et al.</i> , 2009, p. 206.)	105
Figure 7 : les types d'imitation et leurs sphères d'influence.	110
Figure 8 : les deux sous-catégories d'imitation vocale (phonétique et phonologique).	113
Figure 9 : points de convergence et de divergence dans les contours intonatifs de l'imité (Balladur) et de l'imitateur (Yves Lecoq-LB).	116
Figure 10 : points de convergence et de divergence dans la superposition des triangles vocaliques de l'imité (Delors) et de l'imitateur (Lecoq = voix naturelle, Lecoq-Delors = imitation).	116
Figure 11 : conditions expérimentales de l'étude de Kuhl : à gauche l'exposition interactive directe, à droite l'exposition indirecte (Kuhl, 2010, p. 720).	128
Figure 12 : zones de stabilité et d'instabilité selon la Théorie Quantique. La courbe représente la non-linéarité du passage de paramètres articulatoires aux paramètres acoustiques (Stevens, 1989 <i>in</i> Sock, 1998, p. 137)	143
Figure 13 : le cas idéal selon la théorie de la variabilité, entre hypo- / hyper-articulation et informations contextuelle / informations du signal (Lindblom 1987 cité par Sock, 1998, p. 138).	146
Figure 14 : espace illustrant la variabilité dans la production de 10 voyelles / æ, ɑ, ɔ, ʌ, ɜ, i, ɪ, u, o et ε / de l'anglais américain par 46 hommes, 48 femmes et 64 enfants (Hillenbrand <i>et al.</i> , 1995, p. 3104).	152
Figure 15 : représentation schématique de la théorie de la viabilité en production-perception de la parole de Sock (1998), adapté de Bechet (2011, p. 432).	153
Figure 16 : (à gauche) visualisation de l'organisation du système de production de la parole en tant qu'unité ayant sa propre clôture opérationnelle et (à droite) visualisation de l'organisation du système de la perception de la parole en tant qu'unité ayant sa propre clôture opérationnelle (Sock, 1998, p. 174).	156
Figure 17 : catégories phonétiques, émergeant sous forme de phasages articulatoire-acoustiques, correspondent à la manifestation d'une certaine cohérence obtenue par la clôture de toutes les matrices d'interconnexion du système de production-perception de la parole (Sock, 1998, p. 175).	157
Figure 18 : illustration de l'effet McGurk-Macdonald.	161
Figure 19 : schéma représentant la relation entre viabilité et perturbation.	165
Figure 20 : un exemple de pronociation du /t/ aspiré dans le mot city (/sɪ[tʰ]i/) (encadré rouge)	184
Figure 21 : un exemple de production rhotique dans le mot more (/m[ɔɹ]/) (encadré rouge)	186
Figure 22 : différentes formes de réalisations du mot tune en anglais britannique dans le LPD (Wells, 2008).	187
Figure 23 : pourcentage de réalisations du mot tube en anglais américain dans le LPD (Wells, 2008).	187
Figure 24 : un exemple de pronociation de <i> dans la syllabe <vi> du mot vitamin : [aɪ] (/ˈv[aɪ]təmin/) dans l'encadré rouge et [ɪ] (/ˈv[ɪ]təmin/) dans l'encadré bleu.	191
Figure 25 : un exemple d'accentuation sur la première syllabe dans locate (/['loʊ]keɪt/) (encadré rouge)	195
Figure 26 : exemple d'interface du test de perception auditive créée avec Python.	198
Figure 27 : exemple d'un fichier de résultats généré automatiquement par le script du test de perception.	198
Figure 28 : pourcentages et répartition par locuteur de la production du <t> de city.	206
Figure 29 : pourcentages et répartition par locuteur de la production du <t> de vitamin.	207
Figure 30 : pourcentages et répartition par locuteur de la production du <t> de party.	208
Figure 31 : pourcentages et répartition par locuteur de la pronociation du /t/ intervocalique.	209
Figure 32 : pourcentages et répartition par locuteur de la production du <or> de more.	211
Figure 33 : pourcentages et répartition par locuteur de la production du <or> de organisation.	212
Figure 34 : pourcentages et répartition par locuteur de la production du <ar> de dark.	213
Figure 35 : pourcentages et répartition par locuteur de la production du <ar> de party.	214
Figure 36 : pourcentages et répartition par locuteur de la production de <ir> de bird.	215
Figure 37 : pourcentages et répartition par locuteur de la production du <or> de actor.	216
Figure 38 : pourcentages et répartition par locuteur de la production du <ur> de leisure.	217

Figure 39 : pourcentages et répartition par locuteur de la rhocité.	218
Figure 40 : pourcentages et répartition par locuteur de la production du <u> de <i>tune</i> .	219
Figure 41 : pourcentages et répartition par locuteur de la production du <ew> de <i>new</i> .	220
Figure 42 : pourcentages et répartition par locuteur de la production du <u> de <i>evolution</i> .	221
Figure 43 : pourcentages et répartition par locuteur du <i>yod-dropping</i> .	222
Figure 44 : pourcentages et répartition par locuteur de la production du <o> de <i>hot</i> .	225
Figure 45 : pourcentages et répartition par locuteur de la production du <o> de <i>top</i> .	226
Figure 46 : pourcentages et répartition par locuteur de la production du <o> de <i>hot et top</i> .	227
Figure 47 : pourcentages et répartition par locuteur de la production de <o> de <i>joke</i> .	228
Figure 48 : pourcentages et répartition par locuteur de la production de <oa> de <i>road</i> .	229
Figure 49 : pourcentages et répartition par locuteur de la production de <o> et <oa> de <i>joke et road</i> .	230
Figure 50 : pourcentages et répartition par locuteur de la production du premier <y> de <i>dynasty</i> en VR.	231
Figure 51 : pourcentages et répartition par locuteur de la production de <i> de <i>vitamin</i> .	232
Figure 52 : pourcentages et répartition par locuteur de la production de <y> et <i> de <i>dynasty et vitamin</i> .	233
Figure 53 : pourcentages et répartition par locuteur de la production du premier <i> de <i>organisation</i> .	234
Figure 54 : pourcentages et répartition par locuteur de la production du premier <i> de <i>legalisation</i> .	235
Figure 55 : pourcentages et répartition par locuteur de la production du premier <i> de <i>organisation et legalisation</i> .	236
Figure 56 : pourcentages et répartition par locuteur de la production de <ei> de <i>leisure</i> .	238
Figure 57 : pourcentages et répartition par locuteur de la production de <e> de <i>evolution</i> .	239
Figure 58 : pourcentages et répartition par locuteur de la production de <an> de <i>band</i> .	241
Figure 59 : pourcentages et répartition par locuteur de la production de <an> de <i>hand</i> .	242
Figure 60 : pourcentages et répartition par locuteur de la production de <an> de <i>band et hand</i> .	243
Figure 61 : pourcentages et répartition par locuteur de l'accentuation de <i>donate</i> .	246
Figure 62 : pourcentages et répartition par locuteur de l'accentuation de <i>locate</i> .	247
Figure 63 : pourcentages et répartition par locuteur de l'accentuation de <i>donate et locate</i> .	248
Figure 64 : pourcentages et répartition par locuteur de la prononciation de <ary> de <i>military</i> .	250
Figure 65 : pourcentages et répartition par locuteur de la prononciation de <ary> de <i>necessary</i> .	251
Figure 66 : pourcentages et répartition par locuteur de la prononciation de <ary> de <i>military et necessary</i> .	252
Figure 67 : pourcentages généraux pour tous les éléments étudiés et répartition par locuteur.	255
Figure 68 : pourcentages et répartition par locuteur de la production du <t> de <i>city</i> .	261
Figure 69 : pourcentages et répartition par locuteur de la production du <t> de <i>vitamin</i> .	262
Figure 70 : pourcentages et répartition par locuteur de la production du <t> de <i>party</i> .	263
Figure 71 : pourcentages et répartition par locuteur de la prononciation du /t/ intervocalique.	264
Figure 72 : pourcentages et répartition par locuteur de la production du <or> de <i>more</i> .	265
Figure 73 : pourcentages et répartition par locuteur de la production du <or> de <i>organisation</i> .	266
Figure 74 : pourcentages et répartition par locuteur de la production du <ar> de <i>dark</i> .	267
Figure 75 : pourcentages et répartition par locuteur de la production du <ar> de <i>party</i> .	268
Figure 76 : pourcentages et répartition par locuteur de la production de <ir> de <i>actor</i> .	269
Figure 77 : pourcentages et répartition par locuteur de la production du <or> de <i>actor</i> .	270
Figure 78 : pourcentages et répartition par locuteur de la production du <ur> de <i>leisure</i> .	271
Figure 79 : pourcentages et répartition par locuteur de la rhocité.	272
Figure 80 : pourcentages et répartition par locuteur de la production du <u> de <i>tune</i> .	273
Figure 81 : pourcentages et répartition par locuteur de la production du <ew> de <i>new</i> .	274
Figure 82 : pourcentages et répartition par locuteur de la prononciation du <u> dans <i>evolution</i> .	275
Figure 83 : pourcentages et répartition par locuteur du <i>yod-dropping</i> .	276
Figure 84 : pourcentages et répartition par locuteur de la production du <o> de <i>hot</i> .	278
Figure 85 : pourcentages et répartition par locuteur de la production du <o> de <i>top</i> .	279
Figure 86 : pourcentages et répartition par locuteur de la production du <o> de <i>hot et top</i> .	280
Figure 87 : pourcentages et répartition par locuteur de la production du <o> de <i>joke</i> .	281
Figure 88 : pourcentages et répartition par locuteur de la production du <oa> de <i>road</i> .	282
Figure 89 : pourcentages et répartition par locuteur de la production de <o> et <oa> de <i>joke et road</i> .	283
Figure 90 : pourcentages et répartition par locuteur de la production du premier <y> de <i>dynasty</i> .	284
Figure 91 : pourcentages et répartition par locuteur de la production de <i> de <i>vitamin</i> en VR.	285
Figure 92 : pourcentages et répartition par locuteur de la production de <y> et <i> de <i>dynasty et vitamin</i> .	286

Figure 93 : pourcentages et répartition par locuteur de la production du premier <i> dans organisation.	287
Figure 94 : pourcentages et répartition par locuteur de la production du premier <i> de <i>legalisation</i>	288
Figure 95 : pourcentages et répartition par locuteur de la prononciation de <i>organisation</i> et <i>legalisation</i>	289
Figure 96 : pourcentages et répartition par locuteur de la production de <ei> de <i>leisure</i>	290
Figure 97 : pourcentages et répartition par locuteur de la production de <e> de <i>evolution</i>	291
Figure 98 : pourcentages et répartition par locuteur de la production de <an> de <i>band</i>	292
Figure 99 : pourcentages et répartition par locuteur de la production de <an> de <i>hand</i>	293
Figure 100 : pourcentages et répartition par locuteur de la production de <an> de <i>band et hand</i>	294
Figure 101 : pourcentages et répartition par locuteur de l'accentuation de <i>donate</i>	296
Figure 102 : pourcentages et répartition par locuteur de l'accentuation de <i>locate</i>	297
Figure 103 : pourcentages et répartition par locuteur de de l'accentuation de <i>donate</i> et <i>locate</i>	298
Figure 104 : pourcentages et répartition par locuteur de la prononciation de <ary> de <i>military</i>	300
Figure 105 : pourcentages et répartition par locuteur de la prononciation de <ary> de <i>necessary</i> à VR.	301
Figure 106 : pourcentages et répartition par locuteur de la prononciation de <ary> de <i>military</i> et <i>necessary</i>	302
Figure 107 : pourcentages généraux pour tous les éléments étudiés et répartition par locuteur.	304
Figure 108 : comparaison VN/VR.	305
Figure 109 : croisement montrant le lien entre les données sur l'âge, le temps d'apprentissage de l'anglais en général et le temps d'apprentissage de l'anglais à l'université.	313
Figure 110 : nombre de locuteurs et le temps d'utilisation par semaine de l'anglais en dehors des cours.	313
Figure 111 : fréquence des pays d'origine par rapport aux contacts en anglais.	314
Figure 112 : relations avec les interlocuteurs.	314
Figure 113 : temps passé dans un pays anglophone.	315
Figure 114 : nationalités des enseignants de la 3 ^{ème} à la licence 3.	317
Figure 115 : accents des enseignants de la 3 ^{ème} à la licence 3.	318
Figure 116 : nombre d'apparitions des mots inconnus mentionnées par les locuteurs.	319
Figure 117 : fréquence d'apparition dans le corpus SOAP, des mots inconnus cités par les locuteurs.	320
Figure 118 : nombre de séries américaines et autres visionnées par chaque locuteur.	323
Figure 119 : comparaison des volumes horaires consacrés aux activités citées par les locuteurs.	324
Figure 120 : pourcentages estimés des modes de visionnage.	325
Figure 121 : les raisons de visionnage de séries télévisées.	326
Figure 122 : nombre d'apparition des mots qualifiant les intérêts pour les séries.	327
Figure 123 : regroupement en classes des mots qualifiant les intérêts pour les séries.	328
Figure 124 : les supports sur lesquels les séries sont visionnées.	329
Figure 125 : copie d'écran de la page d'accueil de l'organisateur <i>Betaseries</i> d'un locuteur (avec son accord).	331
Figure 126 : la part des séries télévisées mentionnées par pays d'origine.	332
Figure 127 : les 20 premières séries avec le plus haut nombre d'apparition.	333
Figure 128 : nombre de réponses correctes par locuteur pour la RP et le GA.	339
Figure 129 : temps de réponses pour la reconnaissance des mots prononcés avec la RP et le GA.	340
Figure 130 : dendrogramme issu de la Classification Ascendante Hiérarchique.	342
Figure 131 : représentation graphique des relations existant entre les locuteurs.	343
Figure 132 : répartition en groupe selon les quartiles.	345
Figure 133 : signification des valeurs de r.	350

Tables des tableaux

Tableau 1 : Classification de l'AIAL, adapté de Lee (2019, p. 115).	20
Tableau 2 : Quelques exemples d'anglicismes en français.	39
Tableau 3 : Quelques exemples des gallicismes en anglais.	42
Tableau 4 : les trois catégories de pays où l'anglais a un statut particulier. Adapté de Melchers & Shaw (2013, p. 8).	45
Tableau 5 : Les sons de la RP, classés par catégories.	49
Tableau 6 : Les sons du GA.	52
Tableau 7 : quelques exemples de différences grammaticales entre la RP et le GA.	56
Tableau 8 : quelques exemples de différences terminologiques entre la RP et GA.	57
Tableau 9 : exemples de mots indiquant quelques différences orthographiques entre la RP et GA.	57
Tableau 10 : quelques exemples de différence de prononciation entre la RP et le GA.	59
Tableau 11 : les sons du français (adapté de Vaissière, 2006, p. 12).	61
Tableau 12 : Extraits du poème de Trenité, <i>The Chaos</i> .	64
Tableau 13 : sources de variabilité en phonétique (adapté de Tatham & Morton, 2011, p. 10).	145
Tableau 14 : exemples de perturbateurs et leurs conséquences sur la production-perception de la parole.	161
Tableau 15 : Les 24 mots étudiés avec leur transcription phonétique selon les deux accents.	174
Tableau 16 : les 10 phrases utilisées dans la deuxième partie de la liste de mots.	176
Tableau 17 : transcription orthographique et phonétique RP de the north Wind and the Sun (IPA, 1999, p. 44).	177
Tableau 18 : transcription orthographique et phonétique GA de the north Wind and the Sun (IPA, 1999, p. 44).	177
Tableau 19 : transcription orthographique GA de <i>the north Wind and the Sun</i> (utilisée en GA) (Lapteva, 2011, p. 167).	178
Tableau 20 : résumé des éléments segmentaux consonantiques.	223
Tableau 21 : résumé des éléments segmentaux vocaliques.	244
Tableau 22 : résumé des éléments suprasegmentaux.	253
Tableau 23 : résumé des éléments segmentaux consonantiques.	276
Tableau 24 : résumé des éléments segmentaux vocaliques en VR.	295
Tableau 25 : résumé des éléments suprasegmentaux.	302
Tableau 26 : Comparaison des productions en % des caractéristiques proches de la RP, du GA et « Autres », en VN et en VR par locuteur.	305
Tableau 27 : Comparaison des productions en % des caractéristiques proches de la RP, du GA et « Autres », en VN et en VR par groupes d'éléments étudiés.	308
Tableau 28 : âge, genre, études et année d'étude des locuteurs.	311
Tableau 29 : raisons pour lesquelles les locuteurs n'imitent pas un accent particulier.	320
Tableau 30 : origine des accents adoptés par certains locuteurs, selon eux.	322
Tableau 31 : Présentation des variables en fonction de la classe des locuteurs (« petit », « moyen » et « grand »).	345
Tableau 32 : variables utilisées dans les tests de corrélation pour la triangulation des données.	349
Tableau 33 : signification des valeurs de p.	350

Index des notions

A

accent, 26, 27, 29, 30, 31, 32, 48, 49, 50, 51, 52, 54, 58, 65, 98, 109, 110, 112, 113, 116, 117, 119, 128, 131, 133, 166, 173, 174, 175, 177, 182, 184, 186, 187, 188, 190, 194, 195, 196, 199, 200, 201, 203, 209, 210, 243, 244, 245, 249, 253, 254, 256, 257, 263, 295, 298, 302, 308, 310, 316, 317, 320, 321, 322, 325, 326, 327, 332, 334, 335, 337, 338, 339, 340, 344, 346, 347, 348, 350, 351, 352, 353, 354, 356, 357, 358, 359, 360, 361, 363, 365, 366, 367, 368, 369, 373
accent rhotique, 54
Acoustic Image Imprinting, 101, 102, 129, 131, 137
acoustique, 60, 107, 111, 115, 122, 130, 142, 143, 144, 145, 151, 152, 158, 161, 167, 174, 179, 180, 182, 185, 187, 189, 190, 193, 201, 248, 260, 269, 365, 366, 372
acquisition, 11, 12, 13, 22, 24, 25, 26, 28, 31, 33, 74, 85, 86, 88, 89, 90, 91, 93, 101, 102, 105, 106, 107, 110, 111, 118, 122, 124, 125, 127, 128, 129, 130, 131, 133, 135, 136, 137, 138, 142, 154, 161, 338, 362, 363, 364, 365, 366, 371, 372
anglais américain, 9, 28, 29, 31, 32, 43, 45, 46, 51, 55, 151, 182, 184, 187, 192, 239, 249, 256, 340, 348, 352, 354, 357, 358, 359, 370, 372
anglais britannique, 9, 28, 29, 31, 32, 35, 45, 46, 51, 54, 187, 239, 249, 256, 319, 340, 357, 358, 359, 370, 372
anglicisme, 35, 40, 41, 42, 65
apprentissage, 9, 11, 12, 13, 14, 15, 16, 17, 20, 21, 22, 23, 25, 33, 37, 54, 55, 64, 65, 67, 69, 72, 73, 74, 75, 77, 78, 80, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 93, 99, 101, 103, 106, 109, 111, 117, 127, 128, 129, 130, 131, 136, 139, 166, 169, 176, 199, 310, 312, 313, 320, 334, 356, 357, 359, 360, 362, 365, 367, 371, 372, 373
apprentissage incident, 9
apprentissage informel de l'anglais en ligne, 9, 11, 13, 21, 22, 33, 128, 356, 365, 371, 372, 373
appropriation, 22, 90, 101, 102, 118, 167, 335
articulatoire, 60, 132, 134, 142, 143, 144, 157, 158, 164, 165, 167, 182, 183, 188, 192, 248, 256, 260, 364
attracteur, 67, 69, 79, 80, 83, 90, 92, 95, 98, 100, 101, 361, 367

C

caractéristiques « Autres », 223, 244, 253, 254, 257, 276, 302, 303, 304, 308, 309, 345

D

différences individuelles, 63, 68, 91, 92, 93, 100, 117
dynamique, 21, 36, 67, 75, 76, 77, 78, 79, 80, 83, 84, 86, 88, 89, 90, 91, 99, 130, 150, 156, 353, 360, 374

E

émergence, 33, 71, 82, 83, 87, 90, 92, 93, 94, 95, 100, 101, 106, 108, 149, 209, 238, 245, 257, 269, 295, 306, 309, 321, 360
émergentisme, 83
exemplaire, 122, 123, 124, 125, 137

F

flap, 183, 209, 386

G

GA, 29, 32, 35, 45, 46, 51, 52, 53, 54, 55, 56, 58, 59, 65, 170, 173, 174, 176, 177, 178, 181, 182, 184, 186, 187, 188, 189, 190, 191, 193, 194, 195, 196, 203, 209, 218, 219, 223, 224, 227, 230, 233, 234, 235, 236, 237, 238, 239, 240, 243, 244, 245, 248, 249, 250, 253, 254, 255, 256, 257, 259, 263, 271, 276, 277, 280, 283, 286, 290, 294, 295, 298, 299, 302, 303, 304, 305, 308, 309, 321, 339, 340, 345, 346, 347, 349, 350, 351, 352, 356, 357, 358, 362, 369, 370, 372
gallicisme, 35, 41, 42, 43, 65
grapho-phonémique, 63, 85

H

hégémonie, 9, 38

I

imitation, 26, 27, 28, 65, 101, 102, 103, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 116, 118, 119, 122, 125, 126, 128, 129, 130, 135, 136, 159, 161, 166, 199, 201, 256, 310, 317, 320, 321, 325, 326, 327, 334, 337, 354, 360, 361, 364, 365, 367, 368, 372
imitation vocale, 102, 109, 110, 111, 112, 114, 118, 119, 121, 122, 125, 126, 128, 129, 137
informel, 9, 15, 16, 17, 22, 23, 176, 201, 324, 347, 354
input, 29, 32, 75, 91, 92, 98, 100, 109, 111, 113, 117, 125, 130, 161, 325, 334, 347, 358, 359, 360, 362
interconnexion, 21, 156, 366, 373
interlangue, 68, 84, 87, 88, 89, 90, 101
interprononciation, 357, 360

N

nasalisation, 145, 158, 173, 188, 191, 192, 193, 224, 277
Native Language Magnet, 101, 102, 134, 137, 363
neurones miroirs, 102, 105, 106, 107, 108, 109, 136
non-formel, 12, 13, 14, 15, 17, 33, 316
non-linéaire, 21, 78, 87, 142

P

perception, 24, 27, 29, 30, 43, 66, 78, 89, 106, 108, 109, 114, 117, 118, 121, 122, 124, 127, 129, 131, 133, 135, 136, 137,

139, 140, 147, 150, 154, 156, 157, 158, 159, 160, 161, 163, 164, 165, 166, 168, 170, 172, 185, 196, 198, 201, 210, 253, 309, 310, 321, 335, 338, 341, 342, 344, 345, 346, 347, 351, 352, 354, 355, 359, 360, 362, 363, 364, 365, 366, 367, 369, 370

Perceptual Assimilation Model, 101, 102, 133, 137, 353, 363

perturbation, 119, 138, 139, 154, 157, 158, 159, 163, 164, 165, 166, 179, 307, 309, 351, 358, 360, 364

pratiques informelles, 12, 22, 24, 30, 32, 33, 199, 201, 354, 357, 365, 367, 372, 373

production, 24, 27, 30, 66, 89, 90, 106, 107, 109, 110, 114, 118, 119, 121, 124, 125, 129, 130, 131, 137, 138, 139, 140, 141, 143, 145, 147, 150, 151, 154, 156, 157, 158, 159, 161, 163, 164, 165, 166, 168, 174, 181, 185, 186, 188, 192, 196, 200, 201, 203, 206, 207, 208, 209, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 241, 242, 243, 247, 250, 251, 253, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 298, 299, 301, 302, 307, 309, 332, 335, 347, 351, 353, 354, 360, 362, 363, 364, 366, 367, 368, 369, 370, 371, 372

prononciation, 11, 12, 22, 25, 26, 27, 28, 29, 30, 31, 32, 33, 41, 43, 47, 48, 50, 51, 52, 54, 55, 58, 60, 63, 64, 65, 76, 79, 81, 85, 88, 90, 91, 92, 101, 102, 110, 112, 121, 128, 129, 130, 131, 133, 136, 137, 138, 140, 148, 154, 162, 167, 171, 173, 175, 179, 181, 182, 184, 186, 188, 189, 190, 191, 193, 195, 196, 197, 200, 201, 203, 204, 205, 207, 209, 210, 214, 215, 216, 218, 219, 220, 221, 222, 224, 225, 226, 227, 230, 233, 236, 237, 238, 239, 240, 243, 245, 246, 249, 250, 251, 252, 253, 256, 258, 259, 260, 261, 262, 264, 265, 266, 267, 268, 269, 270, 272, 273, 274, 275, 276, 277, 278, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 295, 296, 297, 299, 300, 301, 302, 306, 307, 308, 309, 311, 317, 318, 319, 320, 321, 323, 335, 336, 337, 341, 345, 346, 347, 350, 351, 352, 354, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 368, 369, 371, 372, 373, 374, 1

R

réduction vocalique, 179, 307, 370

résistivité, 119, 159, 163, 164, 166, 179, 201, 308, 309, 358, 364, 372

rhoticité, 51, 53, 59, 173, 176, 182, 184, 190, 204, 210, 217, 223, 237, 256, 257, 259, 264, 265, 266, 267, 271, 276, 277, 308, 363

RP, 29, 32, 35, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 58, 59, 65, 170, 173, 174, 176, 177, 181, 182, 185, 186, 187, 188, 189, 190, 191, 193, 194, 195, 196, 203, 209, 215, 218, 219, 223, 224, 227, 230, 233, 234, 235, 236, 237, 240, 243, 244, 245, 248, 249, 250, 252, 253, 254, 255, 256, 257, 259, 264, 271, 272, 276, 277, 280, 283, 286, 290, 294, 295, 298, 299, 302, 303, 304, 305, 307, 308, 309, 321, 339, 340, 344, 345, 346, 347, 349, 350, 351, 352, 356, 357, 362, 369, 370, 372

S

Speech Learning Model, 101, 102, 131, 137, 363

Systèmes Complexes et Dynamiques, 67

T

tap, 51, 53, 58, 174, 182, 183

tap/flap, 182, 183

V

variabilité, 29, 44, 65, 68, 73, 84, 88, 91, 92, 93, 129, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 151, 152, 157, 167, 168, 180, 189, 193, 223, 244, 253, 256, 257, 276, 277, 294, 306, 307, 308, 311, 317, 319, 327, 331, 343, 357, 363, 364, 373

viabilité, 66, 138, 139, 140, 141, 144, 147, 148, 149, 150, 151, 152, 154, 157, 164, 165, 166, 168, 179, 180, 253, 360, 364, 367, 369

visionnage, 12, 24, 29, 31, 32, 81, 125, 128, 160, 199, 200, 310, 322, 323, 324, 325, 326, 327, 328, 329, 333, 334, 335, 337, 344, 345, 346, 349, 350, 351, 352, 353, 354, 358, 361, 368, 370

vitesse d'élocution, 62, 115, 119, 141, 158, 159, 162, 179, 180, 201, 351, 358, 364, 365, 367, 372

vitesse normale, 32, 178, 179, 180, 201, 203, 254, 256, 258, 345, 349, 350, 352, 358, 364

vitesse rapide, 178, 179, 180, 201, 257, 258, 303, 308, 345, 349, 350, 351, 352, 358, 364

voyelle de *GOAT*, 53, 173, 188, 227, 245, 280, 295

voyelle de *LOT*, 173, 188, 224, 245, 256, 277, 295

Y

yod-dropping, 51, 59, 173, 182, 186, 191, 204, 218, 221, 222, 223, 256, 257, 259, 272, 275, 276, 308

Kossi Seto YIBOKOU

**Apprentissage informel de l'anglais en ligne :
quelles conséquences sur la prononciation des
étudiants français ?**

Résumé

Ce travail s'inscrit dans le cadre de l'apprentissage informel de l'anglais en ligne et explore les pratiques liées aux diverses sources d'exposition d'un échantillon d'étudiants de l'Université de Strasbourg. Les données collectées à partir d'un test de prononciation, d'un test de perception et d'un sondage, montrent une variabilité inter et intra-individuelle inhérente à la complexité du système dans lequel évoluent les participants. Eu regard aux accents *Received Pronunciation* et *General American*, les analyses acoustiques d'éléments de prononciation mettent en évidence des productions orales constituées de mélanges de spécificités des deux accents et celles de la langue française. La perturbation de la production de la parole, mise en œuvre au travers de la variation de la vitesse d'élocution, indique une résistivité du système pour certains sons/séquences de sons. Les résultats montrent également que les séries télévisées constituent les activités les plus influentes parmi celles qui favorisent l'imitation vocale et permettent une appropriation phonéto-phonologique.

Mots-clés : apprentissage informel de l'anglais en ligne, variabilité inter et intra-individuelle, systèmes complexes, imitation vocale, perturbation et viabilité

Abstract

This work is part of the online informal learning of English and explores practices related to various sources of exposure of a sample of students from the University of Strasbourg. The data collected, based on a pronunciation test, a perception test and a survey, show inter- and intra-individual variability inherent to the complexity of the system in which participants evolve. With regards to Received Pronunciation and General American accents, acoustic analyses of pronunciation elements highlight oral productions composed of mixtures of characteristics of the two accents and those of the French language. The perturbation of speech production, implemented through fast speech variation, indicates a resistivity of the system for certain sounds/sequences of sounds. Results also show that television series are the most influential activities among those that promote vocal imitation and allow phonetic-phonological appropriation.

Keywords: online informal learning of English, inter- and intra-individual variability, complex systems, vocal imitation, perturbation and viability