



HAL
open science

L'accentuation lexicale en anglais : des suffixés au sein de l'École de Guierre au sous-système des emprunts et à l'influence des langues sources

Pierre Fournier

► **To cite this version:**

Pierre Fournier. L'accentuation lexicale en anglais : des suffixés au sein de l'École de Guierre au sous-système des emprunts et à l'influence des langues sources. Linguistique. Université Sorbonne Paris Nord, 2024. tel-04589072

HAL Id: tel-04589072

<https://hal.science/tel-04589072>

Submitted on 27 May 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ SORBONNE PARIS NORD

HDR LETTRES ET SCIENCES HUMAINES ET SOCIALES

L'accentuation lexicale en anglais : des suffixés au sein de l'École de Guierre au sous-système des emprunts et à l'influence des langues sources

Volume 1

Synthèse de la recherche

Dossier présenté devant l'Université Sorbonne Paris Nord par

Fournier Pierre

Maître de Conférences à l'Université Sorbonne Paris Nord

en vue de l'Habilitation à Diriger des Recherches

le 15 février 2024

Garant : Monsieur le Professeur Grégory Furmaniak

JURY :

M. Grégory FURMANIAK, Professeur à l'Université Sorbonne Paris Nord

M. Olivier GLAIN, Professeur à l'Université Jean Monnet de Saint-Etienne

Mme Sylvie HANOTE, Professeure à l'Université de Poitiers

Mme Sophie HERMENT, Professeure à l'Université d'Aix-Marseille

M. Anthony SABER, Professeur à l'Université Sorbonne Paris Nord

TABLE DES MATIÈRES

Table des illustrations	7
Glossaire	9
Remerciements	11
Introduction	13
Chapitre 1 - Formation au sein de l'École de Guierre	15
1.1 Initiation à la recherche.....	15
1.2 Les positions théoriques de l'École de Guierre.....	18
1.3 Apprentissage et travaux au sein du groupe « morphophonologie ».....	29
1.4 Le Colloque de Villeteuse sur l'anglais oral.....	37
Chapitre 2 - Résultats notables de la thèse	45
2.1 Résultats notables.....	45
2.2 Travaux post-thèse.....	58
Chapitre 3 - La prononciation des emprunts en anglais	67
3.1 Les emprunts à l'italien.....	69
3.2 Les emprunts au français.....	81
3.3 Les emprunts aux autres langues.....	100
3.3.1 Les emprunts au japonais.....	100
3.3.2 Les emprunts au gaélique écossais.....	103
3.3.3 Les emprunts à l'arabe.....	108
Chapitre 4 - Tentatives de formalisation du sous-système des emprunts en anglais	119
4.1 Les différents degrés d'influence des langues sources.....	122
4.2 Les propriétés du sous-système des emprunts en anglais.....	125
4.3 Les sous-systèmes de l'anglais.....	135
Bilan et perspectives	141
Références	147
Annexes	155
Inédit	165

TABLE DES ILLUSTRATIONS

Figures

- Figure 1** - Représentation de la structure métrique de l'adjectif *momentous* (A.M.S. McMahon 2002 : 123).....20
- Figure 2** - Conceptualisation du sous-système des emprunts avec des exemples.....133
- Figure 3** - Représentation des sous-systèmes de l'anglais (2022 : 248).....139

Tableaux

- Tableau 1** - Degré de réduction phonétique de *was like* accompagné des mesures de durée moyenne.....41
- Tableau 2** - Comparaison de la position de la syllabe accentuée en italien et en anglais dans les emprunts avec une structure en <VCV#> (C = <t, d, n, s, z>).....71
- Tableau 3** - Réalisation phonétique du digraphe consonantique <ch> dans l'*OED*.....107
- Tableau 4** - Réalisation phonétique du digraphe consonantique <ch> dans *LPD* et *EPD*.....107
- Tableau 5** - Transcriptions phonétiques arabes extraites de *LPD*.....117
- Tableau 6** - Les trois voyelles libres étrangères principales de l'anglais.....128
- Tableau 7** - Propriétés des emprunts français et des autres emprunts observables au sein du système anglais.....132

Graphiques

- Graphique 1** - Accentuation des verbes dissyllabiques en anglais britannique, américain et australien.....31
- Graphique 2** - Accentuation des verbes dissyllabiques en fonction de leur structure morphologique en anglais britannique.....32
- Graphique 3** - Durée moyenne de *was like* en position 1-2 de la chaîne de répétition.....41
- Graphique 4** - Durée moyenne de *was like* dans la chaîne de répétition.....42
- Graphique 5** - Accentuation des emprunts italiens en <-VIV#> et <-VrV#> en anglais.....72
- Graphique 6** - Taux de correspondance accentuelle des emprunts en <-VIV#> et <-VrV#> en italien et en anglais.....73
- Graphique 7** - Analyse du taux de correspondance accentuelle entre l'italien et l'anglais britannique et américain dans les 308 mots du corpus.....76
- Graphique 8** - Taux de correspondance accentuelle entre les mots italiens et les réalisations des locuteurs anglophones.....78

Graphique 9 - Taux de correspondance accentuelle entre les mots italiens et les réalisations des locuteurs anglophones selon leur niveau de compétence en italien.....	79
Graphique 10 - Accentuation des emprunts français et anglais britannique et américain dans <i>EPD</i>	87
Graphique 11 - Accentuation des emprunts français en anglais britannique et américain avec la distinction nom commun / nom propre dans <i>EPD</i>	88
Graphique 12 - Accentuation des emprunts français en anglais britannique et américain dans <i>EPD18</i> et <i>LPD3</i>	90
Graphique 13 - Accentuation des emprunts français en anglais britannique et américain dans <i>EPD18</i> et <i>LPD3</i> avec la distinction nom commun / nom propre.....	91
Graphique 14 - Accentuation des emprunts français en anglais britannique et américain dans <i>EPD14</i> et <i>LPD1</i> avec la distinction nom commun / nom propre.....	92
Graphique 15 - Position de l'accent dans les emprunts français en anglais britannique et américain dans les mots simples et selon le nombre de syllabes.....	98
Graphique 16 - Position de l'accent en anglais britannique et américain selon le poids syllabique (calculé en mores) de la syllabe finale.....	99
Graphique 17 - Étude de la correspondance accentuelle entre les mots japonais et les emprunts correspondants en anglais.....	102
Graphique 18 - Accentuation des emprunts arabes dissyllabiques dans l' <i>OED</i>	111
Graphique 19 - Accentuation des emprunts arabes au minimum trisyllabiques dans l' <i>OED</i>	112
Graphique 20 - Accentuation des emprunts arabes dans <i>LPD</i> et <i>EPD</i>	113
Graphique 21 - Accentuation des emprunts français en anglais britannique dans <i>LPD3</i> , <i>EPD17</i> , <i>EPD4</i> et <i>EPD1</i>	156
Graphique 22 - Accentuation des emprunts portugais en anglais britannique et américain dans l' <i>OED</i>	159
Graphique 23 - Accentuation des emprunts espagnols en anglais britannique et américain dans l' <i>OED</i> et <i>LPD</i>	162

GLOSSAIRE

/0/ = syllabe inaccentuée

/1/ = syllabe portant l'accent principal (ou primaire)

/2/ = syllabe portant l'accent secondaire

/-1/ = accent sur la syllabe finale

/(-)10/ = accent sur la syllabe pénultième

/(-)100/ = accent sur la syllabe antépénultième

/(-)1000/ = accent pré-antépénultième

NC = nom commun

NP = nom propre

AUS = anglais australien

GB = anglais britannique

IR = anglais irlandais

SC = anglais écossais

US = anglais américain

EPD = Cambridge English Pronouncing Dictionary

LPD = Longman Pronunciation Dictionary

MCQ = Macquarie Dictionary

OED = Oxford English Dictionary (en ligne)

REMERCIEMENTS

Je tiens tout d'abord à remercier le Professeur Grégory Furmaniak d'avoir accepté d'encadrer ce travail et je le remercie également pour sa disponibilité et son écoute. Ses conseils avisés ont joué un rôle fondamental dans la structure de cette synthèse.

Je suis également très honoré que les Professeurs Olivier Glain, Sylvie Hanote, Sophie Herment et Anthony Saber aient accepté de faire partie de mon jury d'Habilitation.

Je remercie également le Professeur Jean-Michel Fournier qui a un joué un rôle déterminant dans mon parcours de chercheur et qui m'a initié à la phonologie de l'anglais dès la troisième année de licence à l'Université de Tours.

Que soient également remerciés les co-auteurs avec lesquels j'ai travaillé tout au long de ces années, et en particulier Quentin Dabouis, grâce auquel les sous-systèmes ont pu prendre forme. Je remercie tout particulièrement Rose-May Pham-Dinh de son travail de relecture « très » précis et de m'avoir rappelé que les virgules avaient une fonction dans la langue. Je remercie également Stephan Wilhelm de sa relecture rigoureuse de mon inédit et de ses commentaires avisés.

Rose-May Pham-Dinh et Viviane Arigne m'ont formé, dès mon arrivée à l'Université Paris 13 en 2010, aux spécificités et aux exigences du métier d'enseignant-chercheur tant sur le plan de la recherche, que sur celui de l'enseignement ou de l'administratif. Qu'elles en soient remerciées.

J'adresse également mes remerciements au laboratoire Pléiade, et notamment à sa directrice Françoise Palleau, pour l'aide et le soutien apporté lors de la préparation de ce travail.

Enfin, un grand merci à Amélie et Théo pour leur soutien et merci de m'avoir laissé le temps nécessaire pour mener à bien ce projet !

Introduction

Le titre de cette synthèse « L'accentuation lexicale en anglais : des suffixés au sein de l'École de Guierre au sous-système des emprunts et à l'influence des langues sources » illustre parfaitement les thématiques sur lesquelles portent mes recherches et montre également mon évolution en tant que chercheur. De jeune doctorant au sein de l'École de Guierre travaillant sur un sujet certes vaste sur les suffixés en *-ous*, mais restrictif en termes d'approche, je me suis ensuite intéressé à d'autres problématiques liées aux emprunts et aux langues auxquelles l'anglais a emprunté, ce qui m'a ouvert des perspectives de recherches illimitées. Cette synthèse retrace mon parcours et propose des réflexions portant sur la pédagogie, l'enseignement ou l'encadrement de la recherche qui, loin d'être des digressions, ont également nourri mes propres problématiques. La synthèse est organisée autour de quatre thématiques à l'intérieur desquelles l'évolution de ma pensée est retranscrite en termes essentiellement chronologiques. La succession des étapes franchies au sein de chaque partie montre la maturation de mes réflexions.

Dans la première partie, je décris mes premiers pas en tant que chercheur au sein de l'École dite de Guierre, également appelée cadre théorique de la « morphophonologie ». Je présente les spécificités de cette École par rapport aux cadres théoriques phonologiques majeurs existant au niveau international. Les travaux menés en collaboration avec les chercheurs du groupe « morphophonologie » de Tours en parallèle de ma thèse y figurent également, de même que mon rôle de co-organisateur, depuis 2014, du Colloque de Villetaneuse sur l'anglais oral, rôle qui m'a permis d'explorer différents aspects de la recherche, mais également de travailler ponctuellement sur des thématiques nouvelles.

Les résultats notables de ma thèse sur l'accentuation et la prononciation des suffixés en *-ous* en anglais contemporain sont présentés dans la deuxième partie, ainsi que les travaux s'inscrivant dans la continuité de ce travail.

La troisième partie, consacrée aux résultats obtenus dans le cadre de mes travaux sur la prononciation des emprunts débute par une présentation des sources et des rencontres qui ont été déterminantes dans mon cheminement intellectuel et mon orientation scientifique. Je décris ensuite les résultats notables obtenus dans ce domaine en structurant cette partie selon les différentes langues sources étudiées (le français, l'italien et le japonais entre autres). Si les années de parution des articles peuvent parfois, en raison des délais de publication, sembler rompre l'ordre chronologique, l'organisation de chaque partie repose bien sur l'évolution de mes réflexions théoriques et méthodologiques.

La dernière partie de la synthèse présente une tentative de formalisation du fonctionnement des emprunts. La formalisation du sous-système des emprunts en anglais ne fut pas immédiate et je ne mesurai la portée réelle de mes travaux qu'au fur et à mesure de mes tentatives pour en rendre compte. Ma première tentative, somme toute maladroite a posteriori, visait à rendre compte des capacités de transmission des propriétés phonologiques des langues sources et je mis en évidence trois fonctionnements différents concernant les emprunts français, italiens et japonais. Cependant, je constatai dans un deuxième temps que les emprunts possédaient des propriétés intrinsèques inhérentes au système de l'anglais et relevant de plusieurs strates linguistiques (morphologique, phonologique, sémantique, segmentale, etc). Je distinguai ainsi le sous-système des emprunts de celui des emprunts français et du reste du lexique. Enfin, la dernière étape de la formalisation en sous-systèmes est le fruit de ma collaboration avec Q. Dabouis, Maître de Conférences à l'Université Clermont Auvergne. La somme de tous nos travaux sur des sujets connexes nous a permis de formuler l'hypothèse selon laquelle le système anglais serait composé de plusieurs sous-systèmes possédant chacun des caractéristiques linguistiques, distinctes de celles des autres sous-systèmes, et de proposer un embryon de modèle théorique.

Chapitre 1 : Formation au sein de l'École de Guierre

Je présente dans cette première partie mes débuts dans le monde de la recherche et mon apprentissage. Après avoir retracé les spécificités de l'École de Guierre par rapport aux autres écoles phonologiques présentes au niveau international, je me focalise ensuite sur les travaux menés en parallèle de ma thèse en collaboration avec les autres chercheurs évoluant au sein de ce cadre. Je termine cette partie par mon implication au sein du comité d'organisation du Colloque de Villeteuse sur l'anglais oral qui me permit de parfaire ma formation scientifique.

1.1 Initiation à la recherche

Mes premiers pas en linguistique anglaise remontent à la licence d'anglais LLCER que je suivis à l'Université de Tours. Je pense tout particulièrement à la troisième année de cette licence durant laquelle j'assistai aux cours de linguistique de G. Col, J.-M. Fournier, F. Toupin et J.-P. Régis entre autres. Cette formation développa mes connaissances d'angliciste et de linguiste et me donna l'occasion d'entrevoir la pluralité des domaines au sein de la linguistique. J'intégrai par la suite le Master « Langues et Représentations » au sein de la même université et j'y reçus une solide formation en énonciation, syntaxe, analyse du discours, morphologie, phonologie, lexicologie, pragmatique, sémantique et traitement automatique des langues, tant en linguistique générale qu'en linguistique anglaise. L'organisation de ce master permettait en effet de recevoir une formation générale dans les principaux domaines de la linguistique, tout en se spécialisant dans une langue donnée, l'anglais pour ma part. En plus des séminaires des collègues cités précédemment, je suivis également ceux dispensés par N. Garric, S. Osu ou encore T. Gallèpe, qui développèrent mon intérêt pour la linguistique. Après un mémoire réalisé en analyse du discours sous la direction de N. Garric, sur le thème de la dramatisation du discours des animateurs de talk-shows en France, je revins vers la linguistique anglaise et m'inscrivis en thèse de phonologie sous la direction du professeur J.-M. Fournier. Sa présentation remarquable en troisième année de licence du système accentuel de l'anglais, dont l'organisation reposait sur une hiérarchisation de paramètres relevant des domaines syntaxiques, morphologiques et segmentaux, fut déterminante dans mon choix.

La tradition universitaire associe très souvent les domaines de la phonétique et de la phonologie à tel point qu'ils paraissent indissociables l'un de l'autre, tant au niveau de l'enseignement qu'à celui de la recherche. J.-M. Fournier se réclame du courant théorique baptisé l'École de Guierre, du nom de son fondateur L. Guierre, premier professeur de

phonologie anglaise dans le monde universitaire français. Il se définit bien plus volontiers comme un grapho-morphophonologue que comme un phonéticien-phonologue. Je reviendrai plus en détail sur ce statut et sur les spécificités de l'École de Guierre dans la suite de ce chapitre.

J'intégrai donc en 2007 le LLL (Laboratoire Ligérien de Linguistique), U.M.R. regroupant notamment les enseignants-chercheurs des Universités de Tours et d'Orléans, pour y effectuer une thèse sur l'accentuation et la prononciation des suffixés en *-ous* en anglais contemporain. L'axe « morphophonologie » du laboratoire était alors en plein essor et regroupait plusieurs enseignants-chercheurs et doctorants travaillant tous sur la morphophonologie de l'anglais, ce qui conduisit à des échanges instructifs et à des collaborations fructueuses tout au long de ma thèse. Les réunions mensuelles du laboratoire me permirent de présenter régulièrement l'avancée de mes travaux et les discussions qui s'ensuivaient avec les chercheurs chevronnés de l'axe ne manquèrent pas de m'inciter à toujours faire preuve d'une plus grande rigueur dans mon travail de linguiste. En plus de nos travaux respectifs, la mise en place d'une thématique commune à l'ensemble des membres du laboratoire favorisait les collaborations. C'est ainsi que dès le doctorat, mes centres d'intérêt, loin de se focaliser uniquement autour de la prononciation des suffixés en *-ous*, furent multiples et me permirent d'explorer d'autres sujets au sein de la sphère de la phonologie de l'anglais, comme par exemple l'accentuation des verbes préfixés.

Si mon travail de thèse en tant que tel eut un retentissement plutôt limité, en raison de son ancrage profond dans la sphère de la morphophonologie de l'anglais, école théorique loin d'être dominante dans le paysage de l'époque de la phonologie en France et à l'étranger, il fut toutefois extrêmement formateur sur de nombreux plans.

La constitution du corpus fut exigeante mais riche d'enseignements. Plusieurs mois furent ainsi nécessaires à la création d'un corpus de suffixés en *-ous* basé sur des données dictionnaires. Les dictionnaires utilisés furent les suivants : l'*Oxford English Dictionary* (version en ligne) (dorénavant *OED*), le *Longman Pronunciation Dictionary* (dorénavant *LPD*) et le *Cambridge English Pronouncing Dictionary* (dorénavant *EPD*). La version en ligne de l'*OED* constitue la plus grande base de données de l'anglais à ce jour et les techniques d'extraction d'items étant encore loin d'être aussi performantes que celles qui existent aujourd'hui, je fus contraint d'extraire manuellement tous les suffixés en *-ous* de la base de données. La méthode actuelle consistant à entrer « *ous » dans la barre de recherche pour obtenir tous les mots se terminant par *-ous* ne permettait en effet pas d'extraire les termes dérivés correspondant à des sous-entrées. La constitution du corpus fut donc une étape longue, d'autant plus que les données phonétiques des deux autres dictionnaires de prononciation mentionnés ci-dessus furent ajoutées à celles du premier, car en 2007, le manque supposé de

fiabilité des transcriptions phonétiques proposées par l'*OED* ne permettait pas d'envisager une étude aboutie. Cette étape me familiarisa toutefois avec le fonctionnement des bases de données dictionnaires et j'en perçus toutes les subtilités, comme par exemple l'interprétation des données étymologiques proposées par l'*OED* ou celle des différents codages utilisés pour indiquer la variation phonétique dans *LPD* et *EPD*.

La dimension morphologique importante inhérente à mon travail sur les suffixés en *-ous* me conduisit à adopter une grande rigueur vis-à-vis des concepts fondamentaux tels que les bases, les suffixes, les terminaisons ou encore la composition. En effet, en tant que chercheur me revendiquant de l'École de Guierre, je considérais que l'interaction entre les niveaux morphologique et phonologique était fondamentale dans le traitement de l'accentuation lexicale en anglais. Il était donc indispensable d'avoir une maîtrise parfaite des outils morphologiques pour cette étude. La dimension morphologique de l'École de Guierre est indissociable de la sphère sémantique. Par exemple, le concept de séparabilité, qui a un impact considérable notamment sur l'accentuation des préfixés, repose essentiellement sur des critères morphosémantiques (J.-L. Duchet 1994 : 24 ; J.-M. Fournier 2010 : 11). Je considérai dans ma thèse que la séparabilité des préfixes (exemples : *co-*, *dis-*, *mis-*, *ex-*) ou des bases savantes (exemple : *bio-*) était avérée dès lors que le sens de l'élément séparable associé à celui de ce que l'on qualifiera simplement de mot, était compositionnel (c'est-à-dire sens de l'élément A + sens du mot B). Ce statut séparable a un impact considérable sur l'accentuation de l'élément A, notamment lorsqu'il s'agit d'un préfixe. Un accent secondaire est en effet attribué à cet élément séparable, indépendamment de l'accentuation du mot. Concernant le suffixe *-ous*, je constatai qu'il pouvait apparaître en tant qu'élément séparable (exemple : *famous* vient de *fame*) ou inséparable (exemple : *generous*), ce qui modifiait considérablement l'impact du suffixe sur la position de l'accent.

L'accentuation lexicale représenta une partie essentielle de mon étude, car son objectif consistait à démontrer l'influence du suffixe *-ous* sur la position de l'accent primaire. Pour mener à bien une telle étude, il fut nécessaire de comprendre la logique accentuelle de l'anglais dans toute sa complexité. L'accentuation lexicale mobilise des paramètres fondamentaux tels que la syllabe, les terminaisons neutres et contraignantes ou la structure segmentale, autant de concepts que je définirai dans la partie présentant les spécificités de l'École de Guierre.

Le dernier pan de ma thèse qui se révéla formateur concerne les règles graphophonématiques de l'anglais, plus communément appelées « règles de graphie-phonie ». Le cadre théorique de la « morphophonologie » accorde une place essentielle à la graphie qui permet, grâce à un système hiérarchisé de règles, de prédire le passage de l'écrit à l'oral,

principalement au niveau des syllabes accentuées. Il s'agissait dans mon travail de rendre compte de la prononciation de l'intégralité des voyelles accentuées des plus de 3000 items en -ous à partir de leur environnement graphique. J'utilisai pour ce faire l'algorithme permettant de déterminer la valeur d'une voyelle accentuée mis au point dans les années 1990 par J.-M. Fournier, dont il fait une description complète dans son ouvrage (2010 : 141). Le fonctionnement de cet algorithme, dont l'organisation illustre une hiérarchisation des règles déterminée à partir de trois paramètres (la nature de la voyelle, le contexte droit et le rang) selon deux variables (la voyelle accentuée hors contexte et en contexte), sera détaillé dans les sous-parties suivantes. Ce travail fut extrêmement bénéfique pour ma compréhension globale du système des règles de graphie-phonie. À cet algorithme hiérarchisant les règles de correspondance graphie-phonie de l'anglais, j'ajouterai l'algorithme permettant de déterminer la position de l'accent principal, également mis au point par J.-M. Fournier (2010 : 84), qui me permit de structurer mon étude accentuelle des suffixés en -ous selon un système de contraintes syntaxiques, morphologiques et segmentales.

Ce travail de thèse, à la croisée entre étude quantitative (corpus de plus de 3000 termes suffixés en -ous) et traitement qualitatif (notamment dans la détection de sous-systèmes identifiés à partir de critères essentiellement morphosémantiques), s'inscrit pleinement dans les problématiques développées par l'École de Guierre. Je ne perçus la portée pédagogique innovante des deux algorithmes de J.-M. Fournier que bien plus tard, lorsque je commençai moi-même à les utiliser dans le cadre de mes enseignements aux niveaux licence et MEEF. De plus, mon travail sur les différents domaines linguistiques (morphologie, lexicologie, sémantique et phonologie principalement) prépara le terrain pour mes recherches futures, dont l'aboutissement fut la mise en évidence des sous-systèmes de l'anglais et de l'interaction entre les niveaux linguistiques.

La section suivante détaille les fondements de l'école dont je me revendique, à savoir l'École de Guierre, et expose les différences théoriques notables par rapport aux courants théoriques majeurs au niveau international.

1.2 Les positions théoriques de l'École de Guierre

En 1968, *The Sound Pattern of English* (SPE dorénavant), publié par N. Chomsky et M. Halle, allait révolutionner le domaine de l'accentuation lexicale. Ces deux chercheurs américains élaborèrent un système de règles permettant de rendre compte de la position de l'accent lexical dont l'argument principal était selon eux le poids syllabique. En d'autres termes,

la position qu'ils défendaient, selon laquelle les syllabes lourdes attirent l'accent, avec un poids syllabique calculé à partir de la strate phonétique, occupa, à la suite de leurs travaux, une place très importante dans le monde de la phonologie. La primauté de l'oral sur l'écrit constitue l'une des différences fondamentales de cette approche avec les positions défendues par l'École de Guierre. De nombreux adjectifs suffixés en *-ous* figurent dans les inventaires proposés par *SPE* (1968 : 80), raison pour laquelle je m'intéressai à la pensée générativiste lors de ma thèse :

I	II	III
<i>magn'animous</i>	<i>de'sirous</i>	<i>mo'mentous</i>
<i>po'lygamous</i>	<i>poly'hedrous</i>	<i>a'morphous</i>
<i>'rigorous</i>		<i>poly'androus</i>
<i>pre'cipitous</i>	<i>so'norous</i>	<i>tre'mendous</i>
<i>ca'lamitous</i>	<i>de'corous</i>	<i>stu'pendous</i>

Pour résumer la position générativiste originelle de *SPE* (1968), l'accentuation des adjectifs non dérivés obéit à la règle suivante : si la syllabe pénultième est lourde, l'accent tombe sur la syllabe pénultième. Une syllabe est considérée comme lourde lorsqu'elle remplit l'une des deux conditions suivantes : soit elle contient une voyelle longue (adjectifs de la colonne II), soit la voyelle de la syllabe pénultième est suivie d'un agrégat consonantique, se situant donc à la frontière entre les deux dernières syllabes du terme (adjectifs de la colonne III). Les adjectifs des colonnes II et III ont une syllabe pénultième lourde et c'est la raison pour laquelle, selon les auteurs de *SPE*, ils sont accentués sur la syllabe pénultième. Lorsque la syllabe pénultième est légère, l'accent remonte sur la syllabe antépénultième (adjectifs de la colonne I) en vertu du principe de rétraction accentuelle. La syllabe pénultième des adjectifs de la colonne I est légère car elle se compose d'une voyelle courte suivie par une seule consonne. La position générativiste, basée sur le poids syllabique, prend donc comme point de départ l'oral et ne confère pas de réel statut à la composante graphique. La pensée générativiste évolua ensuite au fil des décennies et plusieurs théories dérivées des principes générativistes de 1968 virent le jour.

Ainsi, la phonologie métrique (segmentale) apparut par l'intermédiaire des travaux de M. Liberman & A. Prince (1977), M. Halle & J.R. Vergnaud (1987) et B. Hayes (1980). Dans ce cadre théorique, les syllabes s'organisent en pieds et un pied est généralement composé de deux syllabes. Deux structures métriques accentuelles sont attestées :

→ *left-oriented binary foot* = syllabe accentuée + syllabe inaccentuée

→ *right-oriented binary foot* = syllabe inaccentuée + syllabe accentuée

L'anglais se caractérise par un accent situé sur la syllabe la plus à gauche du pied et A.M.S. McMahon (2002) schématise le rapport de force existant entre les syllabes d'un pied de la façon suivante : S (pour *Strong*) représente la syllabe accentuée qui est en position dominante par rapport à la syllabe inaccentuée symbolisée par W (pour *Weak*). Par conséquent, le processus de la suffixation au sein de la structure se caractérise par un déplacement du rapport de force entre les syllabes, qui se schématise de la façon suivante dans *momentous* :

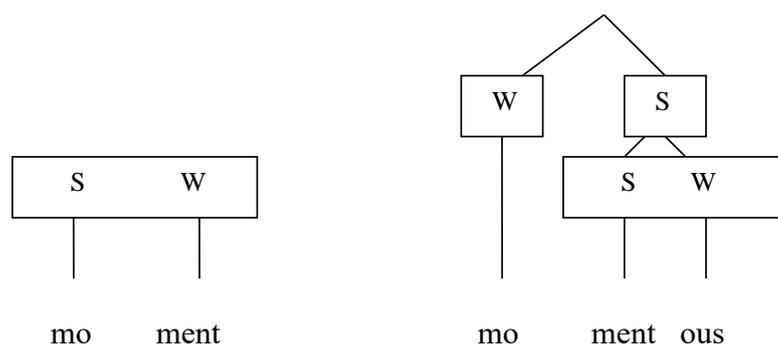


Figure 1 : Représentation de la structure métrique de l'adjectif *momentous* (A.M.S. McMahon 2002 : 123).

Ce schéma de la structure métrique de *momentous* illustre parfaitement l'influence du suffixe *-ous* lors du phénomène de dérivation. Le rapport de force entre les syllabes est en effet modifié. La syllabe forte du dérivant *moment* (<mo>¹) se transforme en syllabe faible après l'ajout du suffixe *-ous*. Le déplacement accentuel provoque alors une altération du pied et une modification du statut de la syllabe <ment> qui occupe la position la plus gauche du pied, car elle reçoit l'accent. L. Burzio (1994) adopte un système de représentation du pied basé sur l'utilisation des parenthèses :

stu(p'endou)s / a(n'omalou)s

La représentation de L. Burzio présente l'avantage de formaliser le comportement des pieds dissyllabiques et trisyllabiques, mais aussi de mettre en évidence les éléments intervenant dans le processus de l'assignation de l'accent lexical. La consonne <s> du suffixe *-ous* est dite extramétrique (concept inventé par B. Hayes (1982)), selon la représentation de L. Burzio, car

¹ Le statut des consonnes intersyllabiques est discutable dans cette représentation lorsque l'on sait que les syllabes accentuées retiennent les consonnes en position coda. Par conséquent, j'opterais pour ma part pour une structure syllabique en <mom> pour illustrer le statut S de la syllabe accentuée dans *moment* et pour <mo> dans *momentous* lorsque la syllabe a un statut W, rapport de force modifié en conséquence du déplacement de l'accent.

elle n'est pas prise en compte dans le calcul métrique du pied. Si tel était le cas, le rapport de force au sein du pied pourrait s'en trouver modifié en raison du statut de syllabe lourde potentiellement attribuable au suffixe *-ous*. En considérant que la consonne finale est extramétrique, les chercheurs qui se revendiquent de cette approche respectent la hiérarchie du poids syllabique et le rapport de force opérant au sein du pied avec une syllabe accentuée se trouvant à gauche.

Je mentionnerai également, à titre d'illustration, le concept d'ambisyllabité formulé par H. Giegerich (1992) qui permet de préciser le statut incertain des consonnes intersyllabiques, tel que formulé dans les approches précédentes, et d'affiner la notion de syllabe lourde dans son interaction avec l'accent lexical. En effet, la définition stricte du concept de syllabe lourde d'origine, à savoir une syllabe avec une voyelle longue et / ou un agrégat consonantique évolua considérablement au fil du temps pour s'adapter aux réalités du système accentuel anglais. H. Giegerich considère ainsi que toute syllabe accentuée est par définition lourde et que les exemples de la colonne I cités dans *SPE* ont une syllabe pénultième légère, mais une syllabe antépénultième lourde qui peut de fait recevoir l'accent. Il utilise la notion d'ambisyllabité pour clarifier sa position, concept qui consiste à attribuer un statut spécifique aux consonnes situées à la frontière entre la syllabe accentuée et la syllabe inaccentuée suivante. Dans l'adjectif '*generous*, la consonne /n/ est ainsi considérée par H. Giegerich comme ambisyllabique, car elle appartient à la fois à la syllabe pénultième /ner/ et à la syllabe antépénultième /dʒen/. Sa position évolua encore par la suite et il finit par considérer que toute syllabe fermée était lourde et pouvait à ce titre prétendre recevoir l'accent.

La phonologie lexicale (P. Kiparsky & I.S. Yang 1982, T. Mohanan 1982) est le courant se rapprochant le plus des positions théoriques défendues par l'École de Guierre. En effet, la phonologie lexicale postule l'existence de deux niveaux distincts au sein desquels des interactions morphologiques et phonologiques en chaîne se produisent. Les suffixes dits de Classe I ont une influence déterminante sur la position de l'accent /1/ et ils interviennent dans la construction de structures « base liée + suffixe » ou « dérivant + suffixe ». Les suffixes de Classe II correspondent globalement aux suffixes dits neutres et les adjectifs en *-ous* correspondant à des formations de type II sont principalement obtenus par dérivation. Le problème de cette classification réside dans le fait qu'elle ne permet pas de rendre compte des comportements hybrides générés par le suffixe *-ous*. En effet, je constatai durant ma thèse que celui-ci pouvait relever de la Classe I ou de la Classe II, en raison tant de son influence hybride sur la position de l'accent lexical que de la pluralité des structures au sein desquelles il était attesté.

J'eus parfois l'impression en tant que jeune chercheur au sein de l'École de Guierre que certains concepts créés au sein des écoles théoriques les plus prestigieuses et les plus reconnues à l'échelle mondiale tentaient en réalité de s'adapter a posteriori aux phénomènes accentuels et essayaient en quelque sorte de rendre compte de tous les phénomènes, même si cela revenait parfois à rendre compte des exceptions. J'expose ici en quelques pages les positions théoriques concurrentes liées à mon travail sur les mots suffixés en *-ous*, qui pourraient paraître hors-sujet au sein de cette synthèse, mais dont la compréhension fut essentielle à mon développement de jeune chercheur, car il me semble que c'est avant tout en comprenant les autres (cadres théoriques) que l'on se comprend soi-même. Je pus ainsi avoir une vision d'ensemble du phénomène accentuel en anglais et des théories visant à en rendre compte, ce qui me permit d'affiner mon positionnement au sein de l'École de Guierre.

À la sortie de l'ouvrage de N. Chomsky et M. Halle, L. Guierre, dont les idées étaient radicalement différentes de celles des générativistes, entreprit de démontrer dans son travail de thèse (1979) et dans les recherches qui s'ensuivirent que l'hypothèse du poids syllabique ne pouvait à elle seule rendre compte de l'accentuation lexicale en anglais. Ses travaux ne conduisirent pas à l'élaboration immédiate d'un cadre théorique, car telle n'était pas son intention initiale. La conception de ce courant théorique est le fruit du travail de trois générations de chercheurs. Les héritiers directs de L. Guierre, à savoir A. Deschamps, J.-L. Duchet, J.-M. Fournier ou encore I. Trevian, développèrent ses idées initiales et les doctorants de cette deuxième génération le firent à leur tour. Au sein de ce cadre théorique, l'accentuation lexicale n'est pas uniquement déterminée par le poids syllabique mais par une interaction entre des paramètres syntaxiques, morphologiques et segmentaux. Le paramètre segmental ne se limite pas à la seule notion de poids syllabique, mais recourt également à la prise en compte du nombre de syllabes. La composante graphique occupe une place essentielle au sein de ce cadre théorique, à l'inverse des approches purement générativistes. La structure morphologique d'un terme est également fondamentale car elle exerce une influence directe sur la position de l'accent. Les domaines de la graphie et de la morphologie sont tellement intrinsèques au modèle théorique développé par L. Guierre que les héritiers dont je fais partie peuvent légitimement se revendiquer comme des grapho-morphophonologues.

Dans son traitement du système accentuel de l'anglais, L. Guierre adopta une approche empirique et numérisa pour ce faire la douzième édition du dictionnaire de D. Jones, l'*English Pronouncing Dictionary* (1963). Les 34 000 mots constituant son corpus lui permirent ainsi de tester l'efficacité des règles accentuelles, car il n'était selon lui pas possible d'envisager de traiter le phénomène sans prendre en compte l'intégralité du lexique. L'une des différences

fondamentales entre les approches guierrienne et générativiste réside d'ailleurs dans le traitement des exceptions. Les listes d'exceptions présentées dans les travaux de L. Guierre (exceptions aux règles accentuelles ou aux règles graphophonématiques) sont exhaustives en raison du traitement empirique des phénomènes, tandis que celles dressées dans l'ouvrage de N. Chomsky et M. Halle sont sélectionnées pour souligner la présence d'un conflit phonologique, sans pour autant tenter de déterminer l'efficacité de la règle présentée ni de quantifier la proportion d'exceptions. De plus, ce sont les mêmes exemples qui sont souvent réutilisés dans les ouvrages générativistes pour justifier certains phénomènes phonologiques, ce qui suggère un manque d'objectivité malvenu en science.

L'approche empirique privilégiée par L. Guierre et sur laquelle repose la quasi-totalité des études menées au sein de ce cadre, nécessite un traitement quantitatif de données principalement extraites de dictionnaires de prononciation. Les deux dictionnaires privilégiés à ce jour sont le *LPD* et l'*EPD*. Il n'est cependant pas rare de voir d'autres sources dictionnairiques être utilisées. Cette primauté donnée aux dictionnaires sur les données orales n'est pas sans rappeler la dichotomie saussurienne *langue vs parole*. Les dictionnaires de prononciation reflètent l'état de la langue à un moment donné et ne peuvent donc prétendre rendre compte de l'intégralité des réalisations possibles par les locuteurs anglophones. Ils sont toutefois représentatifs du système de la langue et leur utilisation permet de faire émerger des tendances. Ce débat entre les données dictionnairiques de l'École de Guierre et l'utilisation des « vraies » données orales prônée par les autres courants théoriques, au risque de perdre la dimension de représentativité inhérente à la langue, est passionnant et alimenta bien des colloques consacrés à l'anglais oral en France.

Il existe au sein de l'École de Guierre quatre principes fondamentaux qui structurent le système accentuel de l'anglais et qui me servent de base théorique pour chacune de mes études.

- 1- Le premier principe stipule que chaque unité lexicale a un accent principal, noté /1/. Si en raison des contraintes rythmiques qui sont détaillées dans les lignes suivantes, un second accent est nécessaire, il sera obligatoirement de degré secondaire (noté /2/).
- 2- Le deuxième principe stipule que les séquences de deux syllabes accentuées sont impossibles au sein d'une même unité lexicale. Selon le principe rythmique de l'anglais caractérisé par l'alternance de syllabes accentuées et inaccentuées, il y a

nécessairement au moins une syllabe inaccentuée entre deux syllabes accentuées au sein d'une unité lexicale².

- 3- Le troisième principe stipule que l'une des deux premières syllabes d'une unité lexicale est nécessairement accentuée. Les contraintes rythmiques de l'anglais imposent la présence d'un accent (de quelque degré qu'il soit) sur l'une des deux premières syllabes.
- 4- Le dernier principe stipule qu'il est impossible de trouver un accent secondaire à droite d'un accent primaire³. Cela signifie que l'accent le plus à droite est de degré /1/. Ce principe paraît logique si l'on considère que la détermination de l'accent primaire s'effectue principalement à partir de la fin du mot. L'impact des terminaisons sur la position de l'accent traduit l'influence de la logique romane sur le système phonologique de l'anglais. Le système accentuel de l'anglais est en réalité hybride car il résulte de la fusion des logiques germanique et romane d'accentuation. Ce point fera l'objet d'un développement lorsque l'accentuation des mots d'origine française sera abordée dans la suite de la synthèse.

La graphie occupe une place essentielle au sein du système. La composante graphique fait l'objet d'un traitement métalinguistique car certaines séquences graphémiques ont un poids phonologique différent de celui que l'on pourrait attendre, mais parfaitement cohérent au vu de la logique du système. Dans le cadre de l'École de Guierre, la notion de l'agrégat consonantique est la plus emblématique de ce méta-système. L'agrégat consonantique se définit comme tout bloc d'au moins deux consonnes, mais cette définition préliminaire doit être affinée, car la nomenclature C₂ représente un fonctionnement graphophonologique basé sur des réalités phonétiques observables plus qu'un fonctionnement purement graphique. Par exemple, le graphème <x> appartient à la catégorie des agrégats consonantiques car il a un effet entravant sur la voyelle accentuée, tandis que les séquences <Ch>⁴ n'en font pas partie parce qu'elles ont un poids phonologique équivalent à celui d'une seule consonne. Lorsqu'un agrégat consonantique est situé juste à droite d'une voyelle accentuée, celle-ci est entravée⁵ (sauf si la voyelle accentuée est un digraphe). L'influence de la graphie sur la réalisation des voyelles sous

² Ce principe n'est en réalité pas toujours respecté notamment au sein des emprunts dissyllabiques qui attestent parfois des successions de syllabes accentuées de type /21/, voire /12/.

³ Il existe quelques exceptions à ce principe, que l'on rencontre notamment en anglais américain ou dans le domaine de la composition.

⁴ C'est-à-dire toute consonne suivie de <h> (exemples : <gh>, <ph>, <th>).

⁵ Je décris simplement ici l'effet principal de l'agrégat consonantique sur la voyelle accentuée. Il existe évidemment des exceptions et des cas particuliers tels que les sous-classes (exemples : <-ight#>, <-ange#>, <-Vste#>).

accent est telle au sein du système que les gémées consonantiques (exemples : <tt>, <rr>, <ll>) sont considérées comme des agrégats consonantiques, car elles ont un effet entravant sur la voyelle accentuée, même si paradoxalement ces gémées correspondent à un seul phonème. L'agrégat consonantique est d'ailleurs l'un des concepts-clés qui se retrouve dans toutes les théories mondiales rendant compte du fonctionnement du système accentuel de l'anglais. Sa présence a un effet déterminant sur le poids d'une syllabe qui est considérée comme lourde. Et dans le cas de l'accentuation des substantifs, la présence d'un agrégat consonantique en position préfinale (c'est-à-dire à la frontière entre les deux dernières syllabes du terme), conditionne le poids de la syllabe (ici lourde) en position pénultième et qui tend donc à retenir l'accent. J'eus donc ainsi durant ma période de thèse une phase d'apprentissage nécessaire du système métalinguistique de l'École de Guierre, particulièrement visible au sein des règles graphophonématiques qui occupaient une place essentielle dans mon travail sur les suffixés en *-ous*. Ces considérations théoriques sur le statut et l'influence de l'agrégat consonantique font immanquablement écho à l'une des notions fondamentales présente au cœur de toutes les théories phonologiques, à savoir le concept de syllabe.

Si les principes des approches générativistes développées précédemment démontrent le statut résolument phonétique de la syllabe, l'École de Guierre adopte une position différente en prenant pour fondement la syllabe phonologique. La définition même de la phonologie en tant que discipline rendant compte du fonctionnement de la langue orale en tant que système implique que l'accentuation anglaise relève d'un système de règles et que la conception de la syllabe en tant qu'entité phonologique est plus appropriée. Les domaines de la phonétique et de la phonologie relèvent en réalité de niveaux différents. Le niveau phonologique permet de justifier l'organisation du système en utilisant la syllabe phonologique comme entité de base, tandis que la syllabe phonétique n'est que la matérialisation orale des principes phonologiques appliqués au niveau précédent. En d'autres termes, les phénomènes phonétiques propres à l'oral telles que les élisions ou les assimilations relèvent davantage de la strate phonétique que de la strate phonologique. Il faut donc veiller à utiliser l'entité théorique relevant de son niveau d'analyse. Les cas de dérivation avec déplacement accentuel obtenu par l'ajout d'une terminaison contraignante illustrent parfaitement la réalité des deux conceptions.

ocean + *-ic* → *oceanic*

['əʊʃ^ən] + [-ɪk] → [,əʊʃɪ 'æɪnɪk]

La dérivation *ocean* → *oceanic* est fréquemment utilisée comme exemple dans les ouvrages relevant de l'École de Guierre. Le dérivant *ocean* ['əʊʃ^ən] a deux syllabes phonétiques et trois syllabes phonologiques. Le nombre de syllabes phonétiques est calculé à partir du nombre de

voyelles prononcées, tandis que le nombre de syllabes phonologiques est basé sur le nombre de voyelles graphiques (à l'exception du <e#> muet). Le recours à une analyse dérivationnelle permet de démontrer le statut résolument dissyllabique de la séquence <ea> dans cette configuration non accentuée. Lorsque la terminaison contraignante *-ic* est ajoutée, l'adjectif dérivé *oceanic* est réalisé [ˌəʊʃɪ'ænik]. La différence accentuelle entre le dérivant et le dérivé, provoquée par l'ajout du suffixe contraignant *-ic*, fait apparaître quatre syllabes, tant au niveau phonétique qu'au niveau phonologique. Il est donc logique de considérer qu'à un niveau phonologique sous-jacent, <ea> est dissyllabique et que la réduction en [-ʃn] n'est qu'une conséquence phonétique prévisible en surface de l'absence d'accent. J'adoptai ainsi lors de mon travail de thèse une conception résolument phonologique de la syllabe. Ma conception évolua toutefois au cours de mes recherches ultérieures, car je me rendis compte en travaillant sur l'accentuation des emprunts que la syllabe définie en tant qu'entité phonétique permettait de mieux restituer les phénomènes phonético-phonologiques se produisant à l'interface langue source - langue cible.

L'interaction fondamentale entre les niveaux morphologique et phonologique au sein du cadre nécessita l'approfondissement de concepts morphologiques clés lors ma thèse et notamment de celui de terminaisons contraignantes. Celles-ci jouent un rôle fondamental dans l'assignation de l'accent, dans la mesure où elles bloquent la relation accentuelle dérivationnelle dérivant / dérivé en interdisant toute relation isomorphique permettant d'associer un schéma accentuel systématique pour chaque terminaison concernée. En d'autres termes, lorsque je constatai une différence accentuelle entre un dérivant et un dérivé en *-ous*, cela signifiait que cette différence était provoquée par la présence d'une terminaison contraignante. La notion de terminaison contraignante fut donc au cœur de mon travail de thèse, le suffixe *-ous* se retrouvant au sein de plusieurs terminaisons relevant de cette catégorie. L'École de Guierre élargit la portée de la contrainte en utilisant plus volontiers le terme de « terminaison contraignante » que celui de « suffixe contraignant », car il apparaît très clairement que les structures déterminant un tel système de fonctionnement au niveau accentuel dépassent parfois le cadre strict de la suffixation. Par exemple, si *-ious* est une terminaison contraignante correspondant à un suffixe adjectival en $V(C_0(e))$ présente dans *tremendous* dépasse les limites morphologiques de la stricte suffixation. De même, si *-ation* est un suffixe dans *organization*, ce n'est pas le cas dans *nation*. Par conséquent, pour des raisons d'unification du fonctionnement du système accentuel de l'anglais, le terme de « terminaison » est préféré à celui de « suffixe », car il est plus en adéquation avec la variété des structures rencontrées. L'influence des terminaisons

contraignantes sur l'accentuation est telle que je remarquai qu'elle s'étendait même aux termes n'ayant pas de dérivant associé. Par exemple, *obvious* n'a pas de dérivant attesté mais relève toutefois de la même règle accentuelle que *gracious*, dérivé de *grace*, car ils partagent la même terminaison contraignante. Les terminaisons sont classées en trois catégories, selon qu'elles sont associées à un accent final, pénultième ou antépénultième. En réalité, les terminaisons contraignantes « n'imposent » pas un schéma accentuel. Il s'agit là d'un raccourci, car telle n'est pas exactement leur fonction. Leur fonction principale consiste à bloquer la relation accentuelle entre le dérivant et le dérivé. Il en résulte un schéma accentuel spécifique associé à chaque terminaison (J.-M. Fournier 1998). Dans certains cas (exemple : *tune / tunic*), il n'est pas possible d'observer de différence accentuelle entre le dérivant et le dérivé. Ceci ne doit cependant pas être interprété comme la marque d'une absence de contrainte. Cela signifie dans l'exemple ci-dessus que le schéma accentuel associé à la terminaison contraignante *-ic* et la règle accentuelle des monosyllabes dans *tune* ont des résultats accentuels concordants, mais relevant de deux logiques distinctes.

La notion de suffixe neutre a un impact fondamentalement différent sur l'accentuation. En effet, la présence d'un suffixe neutre ne modifie pas le lien existant entre un dérivant et son dérivé. On parle alors d'isomorphisme tant accentuel que phonétique au niveau de la réalisation de la voyelle accentuée (exemple : *fame* → *famous*). Certains mots possèdent des suffixes généralement qualifiés de neutres mais n'ont pas de dérivant (exemple : *generous*). J'ai donc considéré dans ma thèse que des items tels que *generous* relevaient des mêmes règles accentuelles permettant de rendre compte des mots non suffixés en anglais.

Concernant la préfixation, l'École de Guierre identifie généralement deux catégories : les préfixes séparables (*re-* dans *readmit*) et les préfixes inséparables (*re-* dans *retain*). La distinction a un impact direct sur l'accentuation, car les préfixes séparables, considérés comme des unités lexicales à part entière véhiculant un contenu sémantique plus important que les unités non séparables, reçoivent un accent secondaire, indépendamment de l'accentuation de la base. Le degré secondaire de l'accentuation traduit toutefois leur dépendance vis-à-vis de la base. Les préfixes inséparables ne sont pas automatiquement accentués⁶, mais ils ont cependant un impact essentiel sur la logique accentuelle. En effet, l'héritage germanique est particulièrement visible au sein du système dans ce contexte, car les préfixés autres que substantifs reçoivent un accent sur la première syllabe de la base, selon la logique germanique d'accentuation qui consiste à ne pas prendre en compte les préfixes dans le calcul accentuel.

⁶ Ils peuvent l'être mais dans ce cas, il s'agit d'un accent secondaire dicté par les spécificités rythmiques de l'anglais, comme dans *represent*.

Par conséquent, la préfixation et la contrainte d'ordre syntaxique (c'est-à-dire la catégorie grammaticale) sont deux éléments déterminants dans l'accentuation de milliers d'items en anglais. L'influence de la préfixation est telle qu'elle est déterminante par extension pour des structures considérées comme opaques. Il est par exemple peu probable que les locuteurs anglophones identifient le caractère préfixé du verbe *divorce*. Une recherche étymologique montre cependant qu'il l'est, et cette particularité morphologique, même opaque, détermine la position de l'accent lexical, qui tombe sur la base en vertu de la logique germanique d'accentuation.

Le recours à l'étymologie pour déterminer la structure morphologique des items m'amène à replacer les principes de l'École de Guierre par rapport à la dichotomie saussurienne synchronie / diachronie. Mon étude au sein de ce cadre théorique fut menée en synchronie, car elle visait à rendre compte des régularités tant accentuelles que graphophonématiques au sein des suffixés en *-ous* en anglais contemporain. Le recours à l'étymologie, parfois nécessaire pour clarifier la structure morphologique de certains items, ne fut toutefois pas prétexte à explorer ces phénomènes d'un point de vue diachronique. Les bases de données dictionnaires me servant de sources pour sélectionner les mots du corpus étaient d'ailleurs la version en ligne de l'*OED* et les deux dernières versions de *LPD* et *EPD*, ce qui garantissait un traitement synchronique de données contemporaines. Cependant, tout un pan de recherches mené par l'École de Guierre travaille principalement sur l'accentuation et les correspondances graphie-phonie dans une perspective diachronique. Le travail de thèse titanesque mené par J. Castanier (2016) sur l'évolution accentuelle du lexique anglais en utilisant comme sources les dictionnaires majeurs de plusieurs périodes en est une parfaite illustration. Si la première partie de mes recherches fut axée sur l'anglais contemporain, j'eus par la suite recours à des analyses diachroniques (P. Fournier 2016a), notamment dans le cadre des mots d'origine française, afin d'évaluer l'impact du facteur temporel sur l'accentuation des emprunts en anglais.

Le dernier point notable concernant les spécificités de l'École de Guierre est la variété d'anglais servant de référence. Si pendant très longtemps les recherches, tant sur l'accentuation que sur les rapports graphophonématiques, furent menées uniquement sur la variété d'anglais britannique, des études sur l'anglais américain ou sur l'anglais australien (M. Martin 2011) virent le jour par la suite. Si mon travail de thèse investit uniquement la sphère britannique, les recherches que je menai par la suite inclurent très souvent la variété d'anglais américain (P. Fournier 2013 notamment).

J'adoptai les fondements de l'École de Guierre lors de mon travail de thèse en raison de l'ancrage profondément guierrien des membres du groupe « morphophonologie » de l'Université de Tours. Il ne m'était pas alors possible d'émettre de réels jugements sur la pertinence des positions développées au sein de ce cadre, car je ne disposais pas de suffisamment d'expérience dans le domaine de la recherche pour avoir le recul nécessaire. Si mes perspectives théoriques et méthodologiques ont évolué depuis ma thèse, je suis néanmoins resté très attaché aux fondements mêmes de ce courant. L'hypothèse d'une interaction entre les niveaux morphologique, syntaxique et segmental est la plus à même de rendre compte du fonctionnement du système accentuel de l'anglais, et les corpus de données dictionnaires permettent un traitement quantitatif des faits accentuels. Si ces données dictionnaires font l'objet de critiques récurrentes de la part des chercheurs issus des autres approches en raison des problèmes de représentativité qu'elles soulèvent, je considère pour ma part qu'elles sont représentatives de la langue et que les dictionnaires ne sauraient répertorier l'intégralité des prononciations attestées en anglais contemporain. Je me distingue toutefois aujourd'hui du jeune chercheur que j'étais sur la conception de la syllabe. Travaillant dorénavant essentiellement dans le domaine des emprunts, je considère la conception phonétique de la syllabe préférable à celle de la conception phonologique lorsqu'il s'agit d'étudier les faits phonologiques se produisant à l'interface langue source - langue cible.

La partie suivante détaille les résultats marquants de mes travaux menés en parallèle de mon doctorat en collaboration avec les membres de l'équipe « morphophonologie » du laboratoire LLL de Tours, travaux qui furent extrêmement formateurs tant sur un plan théorique que méthodologique.

1.3 Apprentissage et travaux au sein du groupe « morphophonologie »

En parallèle à mes recherches de doctorat, les réunions mensuelles du groupe « morphophonologie » au sein du Laboratoire Ligérien de Linguistique permirent de dégager des thématiques communes à tous les membres de l'équipe. Ces réunions instaurées par le responsable de l'équipe, J.-M. Fournier, furent très enrichissantes car elles me permirent de développer mes connaissances théoriques et méthodologiques au contact de chercheurs plus chevronnés. De 2009 à 2012, la thématique retenue pour ces travaux collectifs était la création d'une base de données dictionnaires avec un terrain concret d'application orienté vers l'accentuation des dissyllabes préfixés en anglais contemporain. Depuis la numérisation de la 12^e édition d'*EPD* (1964) de D. Jones réalisée par L. Guierre, l'analyse statistique et

quantitative des faits accentuels est fondamentale au sein du cadre théorique de référence, et le projet de création d'une base de données revêtait donc une importance toute particulière pour les chercheurs du groupe. Durant ces trois années, nous participâmes à cinq conférences regroupant des chercheurs internationaux en phonétique et phonologie. Les principaux résultats de ces trois années consacrées à ce projet sont résumés ci-dessous.

En 2009, lors de la session « posters » du *PAC Workshop*⁷, nous démontrâmes l'importance d'un tel outil dans la mise en évidence des paramètres morphophonologiques qui régissent la logique du système de l'anglais oral. Notre objectif consistait à créer une base de données regroupant toutes les informations nécessaires à un traitement quantitatif et exhaustif des faits accentuels, et ce dans trois variétés d'anglais : l'anglais britannique, l'anglais américain et l'anglais australien, variété sur laquelle travaillait M. Martin (2011). Les informations essentielles figurant dans la base de données sont les suivantes : la graphie, la transcription phonétique britannique (accompagnée du schéma accentuel numérique et des éventuelles variantes), les transcriptions phonétiques américaine et australienne (avec les mêmes informations), la catégorie grammaticale, le statut lexical (dérivant / dérivé), la structure morphologique, les informations sémantiques et les données fréquentielles. Des accords furent passés entre la direction du laboratoire et les éditeurs détenteurs des droits d'exploitation des trois dictionnaires suivants : *EPD* (17^e édition), *LPD* (3^e édition) et *Macquarie Dictionary* (4^e édition). En enrichissant les bases de données xml de ces trois sources avec des données morphologiques, phonologiques, sémantiques et fréquentielles, nous obtenions ainsi un outil permettant de travailler sur des sous-systèmes encore inexploités de l'anglais. Plusieurs notions présentées lors de cette conférence restaient toutefois à affiner, notamment les données relatives à la fréquence. Cette première présentation servit à exposer les bases de notre travail futur sur la base de données.

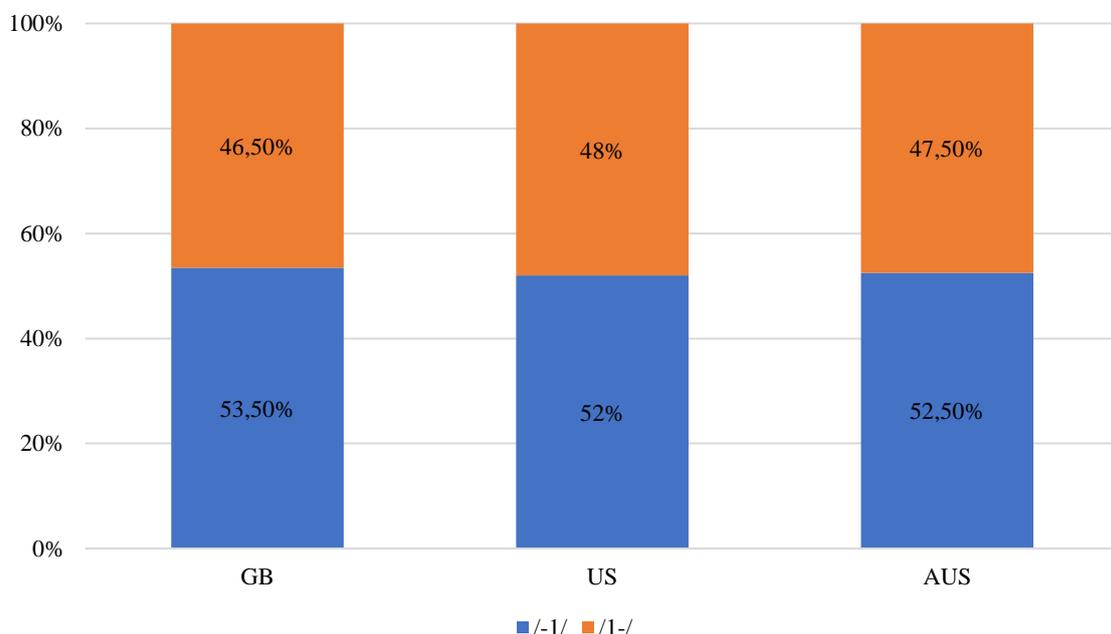
En 2010⁸, une communication effectuée dans le cadre du *PAC Workshop* nous permit de présenter un exemple concret d'application de la base de données. Loin encore d'être entièrement au point, cette base de données permettait tout de même déjà d'obtenir les données brutes extraites des dictionnaires, données que nous passâmes de longues heures à trier et à compléter, afin de mettre en évidence le comportement accentuel des verbes dissyllabiques de

⁷ Abasq, V., Fournier, J.-M., Fournier, P. & M. Martin (2009), "A Dictionary Database of Contemporary English", poster présenté au PAC Workshop : "Models, Variation & Phonological Corpora", Université d'Aix-en-Provence (10-11 septembre).

⁸ Fournier J.-M., Fournier, P., Girard, I. & M. Martin (2010), "Structure, Variation, Usage and Corpora: the Case of Word Stress Assignment in Disyllabic Verbs", communication présentée lors du PAC Workshop : "The Phonology of English: Usage, Varieties and Structure", Université Montpellier III (13-14 septembre).

l'anglais. L'étape de nettoyage des données brutes extraites de la base nous permet de constituer un corpus de 2550 verbes. La répartition des schémas accentuels au sein des verbes dissyllabiques est la suivante (seule la première variante mentionnée est prise en compte) :

- en anglais britannique, 46,5% des verbes sont accentués en /10/ contre 53,5% en /01/ (dans les deux dictionnaires).
- en anglais américain, 48% des verbes sont accentués en /10/ contre 52% en /01/ (dans les deux dictionnaires).
- en anglais australien, 47,5% des verbes sont accentués en /10/ contre 52,5% en /01/.



Graphique 1⁹ : Accentuation des verbes dissyllabiques en anglais britannique, américain et australien.

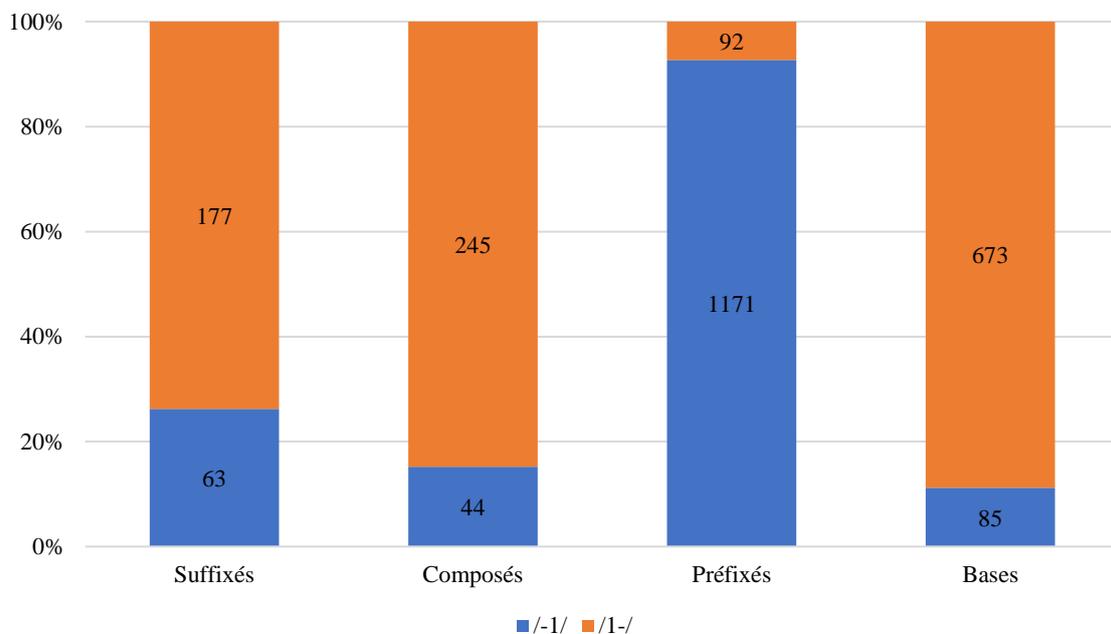
Les résultats du graphique 1 montrent que la proportion de schémas en /01/ et en /10/ est globalement équivalente et que l'accentuation des verbes dissyllabiques en anglais n'est pas soumise à des considérations variétales, car on constate des pourcentages équivalents dans les trois variétés testées. L'étude se poursuit en étudiant le taux de correspondance accentuelle

⁹ Ce graphique reprend les résultats de l'étude en question et est réalisé par mes soins dans le cadre de cette synthèse, comme d'ailleurs tous les autres graphiques de ce travail. Je me suis rendu compte au fur et à mesure des années que les tableaux que j'utilisais systématiquement pour présenter mes résultats ne mettaient pas suffisamment en valeur les données et les faits notables. J'ai donc opté depuis pour une représentation en histogramme, qui est mieux adaptée à la présentation des données relatives aux faits accentuels.

existant entre les deux dictionnaires d'anglais britannique et américain (*LPD* et *EPD*). Mais l'intérêt majeur de cette étude résida dans le fait que, pour la première fois, une étude quantitative basée sur la structure morphologique des verbes vit le jour et tenta de corréler structure morphologique et schéma accentuel (/10/ ou /01/). Nous distinguâmes ainsi quatre types de structure morphologique au sein des verbes dissyllabiques en anglais :

- les verbes préfixés : *accuse, appease, become, detect, retain*.
- les verbes composés : *air-cool, air-dry, benchmark, blackmail, bobtail*.
- les verbes suffixés : *anguish, blemish, castrate, facet, fatten*.
- les bases verbales : *answer, argue, borrow, budget, canoe*.

Les résultats de l'analyse de la corrélation entre structure morphologique et accentuation en anglais britannique¹⁰ furent à la fois très révélateurs et prometteurs (graphique 2).



Graphique 2 : Accentuation des verbes dissyllabiques en fonction de leur structure morphologique en anglais britannique.

Les résultats majeurs du graphique 2 sont les suivants :

- les verbes composés sont massivement accentués en /10/ (85% contre 15% de verbes accentués en /01/).
- les verbes suffixés sont massivement accentués en /10/ (74% contre 26% en /01/).

¹⁰ Seule la variété d'anglais britannique fut étudiée à ce stade car la première partie de l'étude avait démontré que l'accentuation des verbes dissyllabiques n'était pas sujette à des différences variétales.

- les verbes correspondant à des bases sont massivement accentués en /10/ (89% contre 11% en /01/).
- les verbes préfixés sont massivement accentués en /01/ (93% contre 7% en /10/).

Si les résultats concernant l'accentuation des verbes suffixés, composés et bases ne furent pas une réelle surprise, car, de par les propriétés morphologiques spécifiques de ces items, l'accent tombe de façon assez prévisible sur la première syllabe, les données concernant les verbes préfixés furent très révélatrices d'un comportement spécifique de ces unités. Les résultats de cette étude établirent avec certitude que, d'une part, la logique accentuelle des verbes (et possiblement d'autres parties du discours) était corrélée à la structure morphologique de ces items et que, d'autre part, la préfixation avait un effet fondamental sur l'accentuation des verbes dissyllabiques. Dans les verbes préfixés, l'accent principal n'est que très rarement attesté sur le préfixe et tombe surtout sur la base, selon la logique germanique d'accentuation qui, comme on l'a rappelé, consiste à ne pas prendre en compte le préfixe dans le calcul de l'accent et à toujours accentuer la base. La logique germanique d'accentuation était certes connue, mais aucun traitement quantitatif n'avait encore démontré son réel impact sur l'accentuation des verbes. Les résultats de cette étude montrent également qu'en l'absence de préfixe, les verbes sont massivement accentués en /10/, c'est-à-dire sur la première syllabe de la base. En croisant les données globalement concordantes de trois sources dictionnairiques et en étudiant le phénomène au sein de trois variétés d'anglais, les résultats statistiques avaient le mérite d'être significatifs.

Une étude des données fréquentielles extraites du *COCAE (Corpus of Contemporary American English)* apporta une précision de taille sur la perception qu'avaient les phonologues anglophones de l'accentuation des verbes dissyllabiques. Dans de nombreux ouvrages généraux traitant d'accentuation (P. Roach 1983, H. Giegerich 1992, P. Carr 1999), il est dit que les verbes dissyllabiques sont principalement accentués en /01/ en anglais. Les verbes généralement utilisés pour illustrer ce phénomène sont les suivants : *to receive, to accept, to decide, to arrive, to prefer, to subsist, to contain*. Mais ce constat ne s'appuie sur aucune étude quantitative. En réalité, si l'on regarde la tranche des 10% de verbes les plus fréquents en anglais, ils sont effectivement presque tous accentués en /01/. Par conséquent, cela pourrait laisser penser par extension que tous les verbes sont accentués en /01/. Or, nos données contredisent cette affirmation, en démontrant une proportion quasi équivalente de schémas en /10/ et en /01/. Les verbes en /10/ sont globalement moins fréquents au sein du lexique que les verbes en /01/ et la distribution des schémas accentuels est régie par des paramètres morphologiques.

Ces résultats nous semblèrent tellement concluants que nous décidâmes de présenter nos travaux lors du 19^e *Manchester Phonology Meeting*¹¹ en 2011, place forte de la phonologie en Europe et dans le monde. Cette conférence regroupe les grands spécialistes mondiaux de la discipline. Notre poster, dont le format même nous incita à ne présenter que les éléments les plus marquants, reprit les résultats du congrès PAC de 2010, à savoir qu’il existe une corrélation entre la structure morphologique des verbes dissyllabiques et la position de l’accent primaire. Les verbes préfixés sont massivement accentués en /01/ et les verbes non préfixés sont généralement accentués en /10/. Concernant nos trouvailles sur la distribution fréquentielle des verbes, nous insistâmes sur le fait que sur les 20% des verbes les plus fréquents en anglais, 70% environ sont accentués en /01/, et que la quasi-totalité de ces verbes sont préfixés. Nous parvînmes ainsi à faire notre entrée dans un paysage phonologique au sein duquel l’argument principal permettant de justifier l’accentuation lexicale demeure le poids syllabique et toutes les notions qui en sont dérivées, et à nous y implanter de façon durable.

L’année 2012 fut l’occasion de présenter une nouvelle étude statistique sur l’accentuation des paires de multicatégoriels noms / verbes (aussi appelés pluricatégoriels) préfixés lors de deux conférences majeures. La première conférence à laquelle nous présentâmes nos résultats fut celle du *PAC Workshop* à l’Université de Toulouse¹². Les pluricatégoriels (noms / verbes) avaient été écartés des précédentes conférences en raison des difficultés majeures que ces items présentent. Les trois logiques accentuelles mises en évidence par J.-M. Fournier (2010) à leur propos sont les suivantes :

- logique verbale : le verbe et le nom sont accentués en /01/ (exemple : *control*). Ce comportement s’applique pour 60% des paires.
- logique d’alternance : le verbe est accentué en /01/ et le nom est accentué en /10/ (exemple : *record*). Cela correspond au comportement de 30% des paires.
- logique nominale : le nom et le verbe sont accentués en /10/ (exemple : *access*). Cela correspond à 10% des cas.

La logique verbale correspond à la logique germanique qui, on l’a dit, consiste à ne pas prendre en compte le préfixe dans le calcul de l’accent. Si cette règle s’applique généralement aux items préfixés autres que les substantifs, on remarque qu’au sein des multicatégoriels, la logique

¹¹ Descloux, E., Fournier, P., Martin, M. & S. Vanhoutte (2011), “A Dictionary Database of Contemporary English: When the Tool Meets the Needs. The Particular Case of Word Stress Assignment in Disyllabic Verbs”, poster présenté lors du 19^e Manchester Phonology Meeting (MFM) (19-21 mai).

¹² Abasq, V., Dabouis, Q., Descloux, E., Fournier, J.-M., Fournier, P., Girard, I., Martin, M. & S. Vanhoutte (2012), “Multicategorical Prefixed Words Stress Behaviour: Variation and Frequency”, communication présentée lors du Colloque International PAC 2012, “The Phonology of Contemporary English: Variation and Change”, Université Toulouse II-Le Mirail (29 février-2 mars).

verbale s'étend dans 60% des cas aux substantifs associés, qui sont donc également accentués en /01/ (c'est-à-dire sur la première syllabe de la base). La logique nominale consistant à accentuer en /10/ tant le substantif que le verbe associé est moins influente, car seulement 10% des paires adoptent ce comportement. La logique germanique d'accentuation est donc prédominante au sein de cette catégorie.

Après avoir effectué une extraction de tous les verbes dissyllabiques préfixés de la base de données du LLL et vérifié dans les données du *COCAE* si ces verbes pouvaient également relever de la catégorie des substantifs, le corpus fut encore affiné. En effet, nous ne conservâmes dans le corpus final de multicatégoriels que les paires pour lesquelles le verbe et le nom avaient une fréquence supérieure à 0,5 pour 1 million. Cette fréquence d'utilisation, dont les données furent également extraites du *COCAE*, était fondamentale à l'étude, dans la mesure où elle permettait d'exclure les cas pour lesquels l'utilisation nominale d'un verbe relevait du cas particulier, voire de l'exception. Ainsi, les 216 paires de multicatégoriels du corpus d'analyse correspondaient réellement à des paires pour lesquelles l'utilisation tant verbale que nominale était attestée en anglais contemporain.

Exemples : *access, collapse, comfort, defeat, divide, export, incline, mistrust, outlaw, permit, preview, progress, recall, subject, transfer, upset.*

L'étude montra que la distribution des schémas accentuels mentionnée dans les ouvrages était exacte, avec 60% des paires adoptant la logique verbale, 13% adoptant la logique nominale et 27% relevant de la logique d'alternance accentuelle (c'est-à-dire /01/ pour le verbe et /10/ pour le nom). Les résultats démontrèrent également que la variation inter-dictionnaire était très faible concernant cette catégorie d'items, avec des résultats concordants dans *LPD* et *EPD*. La variation inter-dialectale était également très faible, puisque très peu de différences accentuelles étaient constatées au sein des données de *LPD*, *EPD* et *MCQ*. Les résultats concernant la variation intra-dialectale laissaient penser que le phénomène de variation était plus élevé au sein de la logique d'alternance accentuelle entre le nom et le verbe. Cette relative instabilité traduisait vraisemblablement un conflit entre les influences des logiques verbale et nominale. À l'inverse, la variation était très faible au sein des paires relevant de la logique verbale.

Notre participation à la conférence de Manchester¹³ cette année-là reprit les principales conclusions de l'étude sur l'accentuation des multicatégoriels préfixés que je viens de présenter.

¹³ Abasq, V., Dabouis, Q., Descloux, E., Fournier, J.-M., Fournier, P., Girard, I., Martin, M. & S. Vanhoutte (2012), "Stress in Prefixed Disyllabic Verb / Noun Pairs", poster présenté lors du 20^e Manchester Phonology Meeting (MFM) (24-26 mai).

Toutefois, nous eûmes entretemps la possibilité de tester l'influence d'un facteur fondamental dans la distribution de l'accent, à savoir la réalisation vocalique des voyelles inaccentuées. L'hypothèse de départ fut donc formulée ainsi : la présence d'une voyelle pleine en syllabe inaccentuée entraîne-t-elle un schéma accentuel de type alternant (/01/ pour le verbe et /10/ pour le nom) ? Cette problématique prenait tout son sens dans le contexte d'une conférence telle que celle de Manchester, où le paramètre du poids syllabique et des concepts associés demeure prédominant dans les interventions présentées.

Les résultats concernant les substantifs au sein des paires présentant un schéma alternant semblaient confirmer cette hypothèse, avec presque 100% de voyelles pleines en contexte inaccentué, mais les résultats concernant les verbes l'invalidèrent, avec seulement 50% de voyelles pleines en contexte inaccentué. En revanche, l'hypothèse fut validée au sein de la logique d'accentuation nominale (nom et verbe en /10/). En effet, même si 50% des items ont une voyelle pleine en contexte inaccentué seulement, cela ne déclenche pas une accentuation de type alternant.

La deuxième hypothèse que nous testâmes fut la suivante : la proportion de voyelles pleines est-elle liée à la catégorie grammaticale, se traduisant par une proportion importante de voyelles pleines en contexte inaccentué dans les noms et des voyelles plutôt réduites en contexte inaccentué dans les verbes ? Les résultats de notre étude montrèrent que l'hypothèse ne se vérifiait pas, car lorsque les verbes et les noms ont des accentuations identiques, en vertu de l'application de la logique verbale ou logique nominale, la proportion de voyelles pleines / voyelles réduites en contexte inaccentué est identique dans les deux catégories.

La dernière hypothèse que nous formulâmes était relative à la présence notable de variantes accentuelles au sein des paires : la présence de voyelles pleines en contexte inaccentué entraîne-t-elle une augmentation des cas de variation accentuelle au sein du corpus ? Si les résultats concernant la réalisation de la voyelle au niveau de la seconde syllabe semblaient aller dans ce sens, les résultats concernant la réalisation des voyelles au sein de la première syllabe ne permettaient pas de l'affirmer car 87% des items avec une variation accentuelle en /10/ ont en réalité une voyelle réduite au sein de la première syllabe des variantes principales de type /01/. Même si la réalisation des voyelles en contexte inaccentué n'était globalement pas déterminante dans la distribution des schémas accentuels au sein de cette catégorie, l'existence de variantes accentuelles semblait en revanche avoir un effet sur la réalisation des voyelles inaccentuées, et induire une augmentation significative de la proportion de voyelles pleines en contexte inaccentué.

Si ces travaux ne donnèrent pas lieu à des publications regroupant l'ensemble des participants du groupe « morphophonologie » du laboratoire, ils furent extrêmement formateurs sur un plan personnel, car le partage avec des enseignants-chercheurs titulaires et beaucoup plus expérimentés me permit d'une part, de compléter mon bagage théorique relatif au système accentuel de l'anglais, et d'autre part, de cultiver un traitement méthodologique rigoureux des données. Je fus ensuite recruté en tant que Maître de conférences au sein du département d'anglais LLCER de l'Université Paris 13 en 2012, ce qui, sans me couper complètement de mon laboratoire d'origine, ne me permit plus de collaborer aussi étroitement avec mes collègues de Tours, car j'intégrai alors un nouveau laboratoire de recherche (CRIDAF¹⁴). Des collaborations ponctuelles se poursuivirent néanmoins avec certains collègues du laboratoire de Tours, concernant d'autres thématiques que je développerai dans la suite de la synthèse.

La publication conjointe de 2023¹⁵ sur les spécificités de l'École de Guierre constitue, de mon côté, un aboutissement de mes années de formation. Ce chapitre figure au sein d'un ouvrage publié par la prestigieuse maison d'édition *Edinburgh University Press*, ce qui montre que les idées de notre cadre théorique commencent à émerger et à s'installer de manière durable au niveau international. Je ne développerai pas ici les idées avancées dans ce travail car il s'agissait de présenter les positions défendues par notre modèle théorique en regard du domaine de l'accentuation lexicale, spécificités sur lesquelles je suis revenu tout au long de cette partie.

La dernière section du premier chapitre est consacrée au Colloque de Villetaneuse sur l'anglais oral que je coorganise depuis 2012. J'ai décidé d'y consacrer une sous-partie, car d'une part, cette conférence est historiquement associée à l'École de Guierre, et, car d'autre part, en tant que coorganisateur, cette manifestation me permit de présenter mes travaux s'inscrivant dans cette tradition guierrienne, mais également de rencontrer et de collaborer avec des chercheurs venus d'autres horizons.

1.4 Le Colloque de Villetaneuse sur l'anglais oral

Mon recrutement en tant que Maître de conférences au sein du département d'anglais LLCER de l'Université Paris 13, à la suite de mes deux années d'ATER au sein de la même institution, me permit dans un premier temps de faire partie du comité d'organisation du

¹⁴ CRIDAF (EA 453) : Centre de Recherches Interculturelles sur les Domaines Anglophones et Francophones.

¹⁵ Dabouis, Q., Fournier, J.-M., Fournier, P. & M. Marjolaine (2023), "English Word Stress and the Guierrian School", *New Perspectives on English Word Stress*, E. Yamada, A. Przewozny, J.-M. Fournier & N. Ballier (éds.), Edinburgh University Press, 53-82.

Colloque de Villeteuse sur l'anglais oral, puis d'en être le co-organisateur. Lors du mouvement de 2012, l'organisation du Colloque sur l'anglais oral était une condition *sine qua non* du recrutement d'un enseignant-chercheur titulaire en phonétique / phonologie. Je fis donc tout d'abord partie du comité d'organisation du colloque de 2012¹⁶, ce qui me donna une première expérience dans la mise sur pied de manifestations scientifiques d'une telle ampleur ; puis je pris en charge l'organisation locale des colloques de 2014¹⁷, 2016¹⁸, 2018¹⁹ et 2022²⁰. Ce colloque sert également de support aux manifestations scientifiques de la société savante ALOES (Société des Anglicistes Oralistes de l'Enseignement Supérieur, Secondaire et Élémentaire) depuis les années 1990. Les colloques sont donc conjointement organisés par l'Université Sorbonne Paris Nord via le laboratoire Pléiade (dans la version en cours) et la société savante ALOES. Cette manifestation est réputée au sein du paysage des études portant sur l'anglais oral en France à plus d'un titre. Il s'agit tout d'abord de l'une des toutes premières conférences consacrées à l'anglais oral dans les universités françaises, puisque sa première édition remonte à 1982. Tout au long de ces quarante années, elle a accueilli des invités prestigieux comme par exemple A.C. Gimson, J.C. Wells, P. Roach ou encore H. Giegerich. Enfin, les colloques de Villeteuse sur l'anglais oral ont pendant longtemps eu une orientation morphophonologique, transmise par le biais des présentations de Lionel Guierre et de tous ses héritiers directs et indirects.

Avec des collègues linguistes de l'université et des représentants de l'ALOES, je fus donc chargé de coorganiser ces colloques, dont les thématiques et les invités prestigieux me permirent de suivre l'actualité phonétique et phonologique du moment et d'être exposé à toute une diversité d'approches méthodologiques et théoriques. J'appris donc énormément lors de ces manifestations. J'y présentai certains de mes travaux s'inscrivant directement dans la lignée de la pensée morphophonologique, mais ces conférences furent également le lieu d'échanges et d'ouvertures sur d'autres approches. Je pense notamment à la présentation faite lors du colloque

¹⁶ 16^e Colloque de Villeteuse sur l'anglais oral : « Multicultural Spoken English ? / L'anglais oral mondialisé ? » (30-31 mars 2012).

¹⁷ 17^e Colloque de Villeteuse sur l'anglais oral : « Spoken English and the Media / L'anglais parlé et les médias » (4-5 avril 2014).

¹⁸ 18^e Colloque de Villeteuse sur l'anglais oral : « Phonetics and Forensics: Speaker, Register and Dialect Identification / Les experts : caractérisation des locuteurs, des variétés et des styles de parole » (1-2 avril 2016).

¹⁹ 19^e Colloque de Villeteuse sur l'anglais oral : « La qualité de voix en anglais : de la production d'un type de phonation à la perception d'une caractéristique sociale / Voice Quality in English: From the Production of a Phonation Type to the Perception of a Social Variable » (6-7 avril 2018).

²⁰ 20^e Colloque de Villeteuse sur l'anglais oral : « Phonetics / Phonology: interface, regularities, realisations - Phonétique / Phonologie : interface, régularités, réalisations » (1-2 avril 2022).

de 2016²¹ avec mon collègue Y. Fuchs, spécialiste des quotatifs en anglais et auteur d'une thèse sur le sujet en 2012, dont je restitue ci-dessous les résultats notables.

Notre décidâmes de travailler sur le comportement phonétique et phonologique des quotatifs en interaction, ce qui nous permit d'associer nos deux champs de compétence. Cela représentait une avancée méthodologique majeure pour moi que de travailler sur des corpus oraux enregistrés à partir de productions de locuteurs britanniques ou irlandais, n'ayant pour ma part analysé que des données dictionnaires. De plus, le recours au logiciel d'analyse *Audacity* fut également une découverte. Le groupe des quotatifs (I. Buchstaller & I. Van Alphen 2012) s'est récemment élargi car si, historiquement, *say* et *think* étaient principalement utilisés pour introduire des segments représentatifs de paroles ou de pensées au discours direct, *go* et *be like* constituent dorénavant des alternatives crédibles. Ce paradigme élargi offre désormais la possibilité aux locuteurs anglophones d'utiliser plusieurs formes au sein d'une même séquence. Cette évolution au sein de la catégorie des quotatifs a été énormément documentée avec l'identification de facteurs sociolinguistiques, grammaticaux, sémantiques, pragmatiques et interactionnels qui permettraient d'anticiper l'apparition d'une forme ou d'une autre (S. Romaine & D. Lange 1991, S. Tagliamonte & R. Hudson 1999, J. Dailey O'Cain 2000, G. Andersen 2000). Peu d'études s'étaient en revanche intéressées à l'enchaînement des quotatifs au sein d'une séquence dialoguée. Y. Fuchs (2014) mit en évidence le fait que certains locuteurs alternaient les quotatifs, tandis que d'autres utilisaient le même quotatif pour introduire des paroles rapportées et il établit un lien entre répétition du quotatif et renforcement de la cohésion discursive. Notre étude de 2016 visait à tester la validité d'une telle hypothèse sur un plan phonétique en observant le degré de réduction du marqueur *be like* en interaction, à la suite d'études ayant déjà établi une corrélation entre réduction phonétique du marqueur et degré de cohésion textuelle (N. Ballier 2013). Notre choix se porta sur *be like* en raison de sa productivité au sein du corpus et de sa grande capacité de réduction phonétique présentant plusieurs stades distincts.

Reprenant le principe de la fonction cohésive de la répétition (D. Tannen 2007), Y. Fuchs (2014) démontra que lorsque la cohésion en conversation spontanée était menacée par des accidents dits conversationnels, les locuteurs anglophones s'appuyaient sur le mécanisme de répétition des quotatifs pour renforcer la cohésion discursive ainsi menacée. Plus la cohésion

²¹ Fournier, P. & Y. Fuchs (2016), « Répétition, cohésion et réalisation phonétique : le cas des quotatifs à l'oral », 18^e Colloque de Villeteuse sur l'anglais oral : « Phonetics and Forensics: Speaker, Register and Dialect Identification / Les experts : caractérisation des variétés, des locuteurs et des styles de parole », Université Paris 13 (1-2 avril).

était forte et plus les quotatifs avaient tendance à être réduits phonétiquement. Notre étude visait donc à répondre aux deux problématiques suivantes : la répétition d'un même quotatif contribue-t-elle à renforcer la cohésion discursive d'un texte ? Et si la réduction phonétique est liée au degré de cohésion, les multiples répétitions d'un même quotatif s'accompagnent-elles alors d'un plus grand degré de réduction phonétique ?

À partir des données du *Cambridge Student Corpus* (CSC ci-dessous) enregistré en 2010 et 2011, nous avons isolé toutes les occurrences du marqueur *be like* qui représentait plus de 50% des quotatifs du corpus. Nous avons toutefois écarté les formes *is like* ou *were like* car minoritaires par rapport au marqueur *was like* qui est massivement utilisé par les anglophones pour introduire des récits d'expériences à la première ou troisième personne du singulier. Seules les séquences de répétition de *was like* et les séquences de répétitions suivies d'alternance dites mixtes (exemple : *was like + said + was like*) ont été conservées. Le logiciel *Audacity*, qui permet une précision à la milliseconde près, a été utilisé pour réaliser quatre types de mesures sur les 92 occurrences du corpus :

- la durée totale du segment < *was like* + pause éventuelle >
- la durée du segment <*was like*> sans pause
- la durée du segment <*like*>
- la durée de la pause éventuelle (mathématiquement déduite des autres mesures).

Nous avons effectué les mesures séparément et procédé ensuite à la mise en commun de nos résultats, en procédant au calcul de la moyenne arithmétique à partir de nos deux mesures. Tous les segments de l'étude ont ensuite été transcrits phonétiquement en double aveugle, puis nous avons procédé à une mise en commun de nos transcriptions avec des données sensiblement équivalentes concernant la réduction ou l'absence de réduction de la voyelle de *like* et assez similaires concernant les réalisations allophoniques et les assimilations. Le codage des répétitions a été effectué de la façon suivante : chaque occurrence du quotatif a reçu un chiffre symbolisant sa position dans la chaîne de répétitions.

Exemple : I couldn't go so then I I **was like (1)** I **was like (2)** mum I can't go to school she **was like (3)** why I **was like (4)** I'm just so upset it's like I can't go

CSC (2011)

Si la chaîne de répétitions est interrompue par un accident conversationnel ou par une alternance avec un autre quotatif, la numérotation repart à 1, c'est-à-dire à une première occurrence de *was like*. Les données concernant les mesures et les différentes transcriptions phonétiques de *was like* observées, en allant des réalisations les moins réduites vers les réalisations les plus réduites en passant par les réalisations intermédiaires, sont répertoriées dans le tableau ci-dessous.

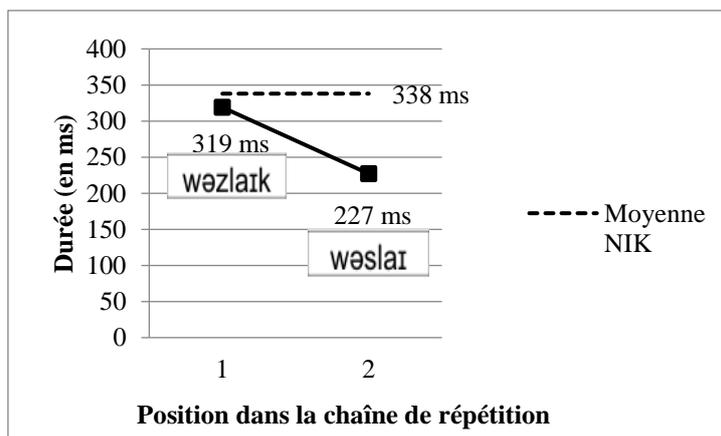
- réduit	... réduction ...	+ réduit
wəzlɑik wəzlɑig wəzlɑiʔ wəslɑik	wəzlɑʔ wəzɑik wəzɑg wəzə	sɑg slɑig sɑiʔ zɑʔ sɑ
Moyenne = 391 ms	...	Moyenne = 202 ms

Tableau 1 : Degré de réduction phonétique de *was like* accompagné des mesures de durée moyenne.

Sans surprise, plus la forme phonétique d'un quotatif est réduite, plus sa durée moyenne de réalisation diminue. L'exemple ci-dessous extrait du corpus nous permet de mettre en évidence le fait que le quotatif se situant en seconde position au sein de la chaîne de répétition adoptait une forme plus réduite, mais aussi plus brève que celui chargé d'établir la cohérence discursive et situé en première position de la chaîne.

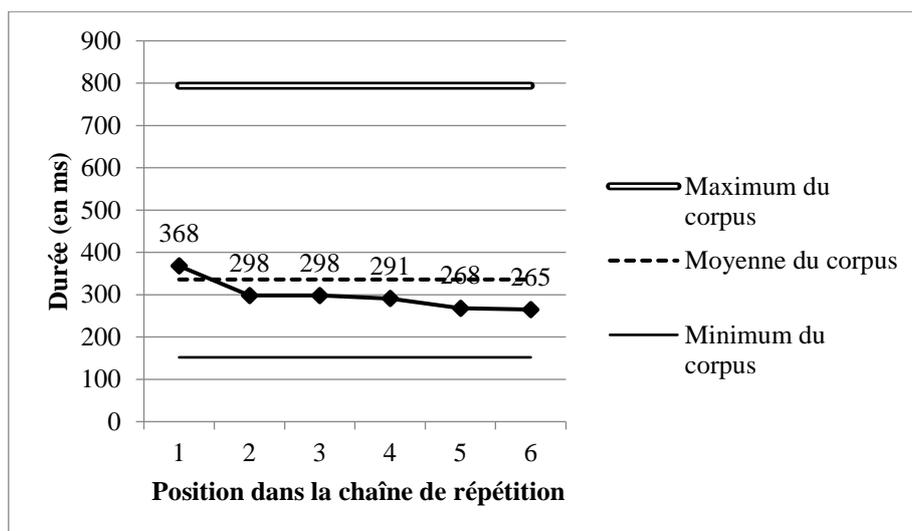
Exemple : he **was like (1)** they didn't see me <rire> I **was like (2)** ha so you admit it

CSC (2011)



Graphique 3 : Durée moyenne de *was like* en position 1-2 de la chaîne de répétition.

Le graphique 3 montre que la position dans la chaîne de répétition a un effet sur la réalisation phonétique mais également, par incidence, sur la durée du quotatif. Et cela se vérifie non seulement sur les deux premières occurrences mais également sur l'ensemble du corpus, comme le montre le graphique 4.



Graphique 4 : Durée moyenne de *was like* dans la chaîne de répétition.

Notre codage initial dut subir un léger réajustement, car il suivait une logique strictement linéaire sans tenir compte des incidents interactionnels tels que des silences ou des interruptions, ce qui ne nous permettait pas d'obtenir des résultats significatifs. En revanche, si la numérotation repart à 1 dès lors qu'un accident conversationnel survient, nous avons constaté que le degré de réduction et la durée d'un quotatif étaient bien corrélés à sa position dans la chaîne de répétition. Dès lors qu'un accident conversationnel survient, la première occurrence du quotatif *was like* est moins réduite et plus longue, ce qui permet de rétablir la cohésion discursive et de réaffermir la position du locuteur par rapport à ce qu'il avance. En conclusion, nous avons clairement démontré que la prononciation d'un quotatif était liée au besoin qu'ont les locuteurs de renforcer ou non la cohésion discursive d'un texte. Ces résultats préliminaires seraient sans nul doute enrichis par une analyse phonétique réalisée à l'aide d'un spectrogramme et portant sur tout l'inventaire des quotatifs qu'offre dorénavant l'anglais. Si notre étude initiale date de 2016, une publication reprenant ces conclusions devrait voir prochainement le jour. Cette étude pionnière du comportement phonétique des quotatifs en anglais me permit d'explorer un autre univers que celui au sein duquel j'évoluais, de me former à l'analyse de corpus oraux et d'ouvrir de belles perspectives de recherches futures en collaboration avec les collègues de mon laboratoire.

Les colloques de Villetaneuse sur l'anglais oral furent ainsi très enrichissants dans mon parcours de chercheur et me donnèrent également l'occasion d'aborder un aspect différent de la recherche : le travail d'édition scientifique. Je fus ainsi en charge de la préparation de trois

numéros dérivés des thématiques du Colloque sur l'anglais oral²² en collaboration avec des collègues de l'ALOES. Cette tâche de diffusion de la recherche et de mise en valeur du travail des autres chercheurs fut extrêmement gratifiante et me permit de développer encore davantage la rigueur et l'esprit critique indispensables aux enseignants-chercheurs, ainsi qu'une réelle capacité à travailler en équipe.

Cette première partie a présenté les différentes étapes de mon initiation à la recherche. De mes balbutiements de jeune chercheur sur le traitement des suffixés en *-ous*, à quelques incursions enrichissantes, telles que mon étude sur les quotatifs en collaboration avec Y. Fuchs dans le cadre du Colloque de Villetaneuse sur l'anglais oral, en passant par des travaux collectifs très formateurs en association avec l'équipe « morphophonologie » de Tours, j'ai pu bénéficier de fructueuses collaborations qui m'ont permis de développer mes compétences méthodologiques et théoriques. Le deuxième chapitre de cette synthèse présente les résultats notables obtenus dans le cadre de ma thèse, consacrée à l'accentuation et la prononciation des mots suffixés en *-ous* en anglais contemporain, et dont la soutenance fut organisée en décembre 2011 à l'Université de Tours. Les travaux post-thèse s'inscrivant dans le prolongement direct de ces problématiques figurent également dans ce deuxième chapitre.

²² Fournier, P., Herment, S. & H. Le Priault (éds.) (2019), *Voice Quality in English*, *Anglophonia* 27 (<https://journals.openedition.org/anglophonia/1869>).

Fournier, P. & I. Girard (éds.) (2018), *Multicultural Spoken English*, *Corela* HS-24 (<https://journals.openedition.org/corela/5084>).

Le Priault, H. & P. Fournier (éds.) (2016), *Approches pluridisciplinaires en phonologie anglaise*, *Anglophonia* 21 (<http://anglophonia.revues.org/694>).

Chapitre 2 : Résultats notables de la thèse

Les résultats notables de mon travail de thèse sont présentés dans la première partie de ce chapitre. Les travaux post-thèse qui s'inscrivent directement dans la lignée des conclusions formulées dans le cadre de ma thèse sont exposés dans la seconde partie.

2.1 Résultats notables

L'objectif de mon travail étant résolument tourné vers la composante phonologique (position de l'accent lexical et réalisation des voyelles sous accent entre autres), je dus apprendre d'une part à décoder et interpréter toutes les subtilités de l'art de la présentation des données phonologiques au sein des dictionnaires, et d'autre part, à déchiffrer les données de l'*OED*, dont les conventions diffèrent de celles de *LPD* et d'*EPD*. La constitution du corpus de thèse fut donc une étape très formatrice tant sur le plan méthodologique que sur le plan théorique.

Mon travail fut le premier à s'appuyer sur les transcriptions fournies par l'*OED* (dans sa version en ligne) pour mener une étude sur les régularités du système accentuel de l'anglais. En effet, la fiabilité des données proposées par ce dictionnaire a longtemps été questionnée en raison tout d'abord de conventions de transcription distinctes de celles des dictionnaires servant de référence aux concours de recrutement, mais également en raison des nombreuses parties non actualisées du dictionnaire tant sur le plan lexical, qu'étymologique, sémantique ou phonétique. Les données phonologiques des parties non rénovées sont donc anciennes et ne proposent pas la transcription de l'anglais américain. Cependant, les travaux de remise à jour des données de l'*OED* dans sa version en ligne me permirent d'utiliser les données accentuelles pour mon étude. Un test de concordance accentuelle entre les trois sources dictionnaires fut toutefois mené lors de la constitution du corpus, afin de déterminer la fiabilité des données de l'*OED*, et les résultats montrèrent un taux de correspondance accentuelle supérieur à 99% pour les mots suffixés en *-ous* présents dans les trois sources (502 cas de correspondance accentuelle sur 507 items concernés). Ce résultat me parut plus que convaincant et j'intégrai ainsi les transcriptions phonologiques de l'*OED* à mon étude.

Je ne me doutais pas à l'époque que la « simple » étape de constitution du corpus se révélerait aussi longue, ardue et formatrice dans de nombreux secteurs interdépendants de la linguistique. La partie traitant des critères de constitution de mon corpus dictionnaire de plus

de 3000 termes suffixés en *-ous* me permit de parfaire mes connaissances dans plusieurs domaines. La dimension lexicographique d'une telle tâche m'obligea à travailler sur la valeur sémantique du suffixe *-ous* en synchronie et en diachronie, en analysant les données fournies par l'*OED* et en recoupant ces données étymologiques et sémantiques avec des données morphologiques. Je devais en effet m'assurer de ne traiter que des items suffixés en *-ous*. Les relations syntactico-sémantiques des structures morphologiques furent également étudiées en profondeur afin de ne conserver que des formes pertinentes dans mon analyse. Par exemple, le caractère séparable du préfixe *un-* dans la construction *ungracious* est aisément repérable tant au niveau morphologique qu'au niveau sémantique, avec un sens strictement compositionnel (I. Plag 2003) : le sens de *ungracious* résulte de l'addition du sens du préfixe *un-* et de celui de l'adjectif *gracious*. Les exemples ci-dessous fonctionnent sur le même modèle.

asynchronous = *a-* + *synchronous*

disadvantageous = *dis-* + *advantageous*

extemporaneous = *ex-* + *temporaneous*

intercartilaginous = *inter-* + *cartilaginous*

L'identification de telles structures reposant sur des considérations morphologiques, sémantiques et syntaxiques – l'interaction entre un élément séparable non autonome et une structure autonome traduisant selon J.-M. Fournier (2010) une relation de nature syntaxique – était nécessaire, afin de ne pas travailler sur des formes redondantes. En effet, la position de l'accent primaire est déterminée à partir de la structure autonome, l'élément séparable ne jouant aucun rôle dans l'assignation de l'accent principal. J'observai ainsi un isomorphisme phonétique entre les constructions syntaxiques et les structures autonomes impliquées dans de telles constructions. Pour simplifier, il n'était pas nécessaire de conserver *ungracious* dans le corpus d'analyse car *gracious* y figurait déjà et puisque les règles permettant de justifier l'accentuation et la prononciation de la voyelle accentuée (accent principal) de *ungracious* étaient identiques à celles de *gracious*. Les structures avec une base savante séparable en lieu et place du préfixe furent également inventoriées et retirées du corpus final (exemple : *gynandromorphous* est formé à partir de *gyno-* et de l'adjectif *andromorphous* qui figure déjà dans le corpus). La distinction entre les concepts de préfixation et de base savante initiale (traduction personnelle de l'expression *initial combining form*) sera abordée dans la section relative à la composition savante. Je fus parfois confronté à des problèmes méthodologiques dont les implications théoriques dépassaient le niveau de compétence du jeune chercheur que j'étais. Mais je dus me positionner très tôt sur ces problématiques, car un travail de constitution de corpus ne saurait tolérer d'approximations. L'aide de mon directeur de recherches fut

précieuse dans ce processus. Sachant par exemple que les adjectifs *cephalous* ou *petalous* étaient attestés dans le corpus, mais qu'il existait dans le même temps des dizaines de composés savants en *-cephalous* et *-petalous*, comment dès lors interpréter des formations telles que *acephalous* ou *apetalous* ? Devais-je considérer ces deux formations comme des constructions syntaxiques composées du préfixe séparable *a-* et des adjectifs *cephalous* et *petalous* ou bien comme des compositions savantes du type « préfixe *a-* + base savante liée *-cephal-* + suffixe *-ous* » ?

préfixe *a-* + adjectif *cephalous* OU

préfixe *a-* + base savante liée *-cephal-* + suffixe *-ous* ?

La distinction peut à première vue paraître minime, mais l'enjeu était de taille car si j'optais pour la première option, de telles formations devaient être retirées du corpus, en vertu du principe d'isomorphisme. Je décidai de retenir la seconde formulation, organisée autour du concept de composition savante, compte tenu de facteurs sémantiques et diachroniques. Je détaillerai les principaux résultats de mon chapitre de thèse consacré à la composition savante dans les pages à venir.

Ma thèse mit en évidence la formidable diversité des structures morphologiques au sein desquelles est attesté le suffixe *-ous* en anglais. I. Trevian (2003 : 8) définit le suffixe comme « un élément morphémique affixé à la finale d'un lexème attesté » et la terminaison comme un « élément final dont le retrait nous laisserait sans radical synchroniquement identifiable ». Ainsi, selon lui, *-ous* est un suffixe dans *poisonous*, construit à partir du dérivant *poison*, mais a un statut de terminaison dans *generous*, car il n'y a pas de dérivant attesté. Cette distinction purement basée sur le caractère séparable de *-ous* en contexte, au détriment de toute considération morphologique hors contexte, ne me parut pas adaptée à mon sujet. En effet, mon étude étymologique des données de l'*OED* montra que de nombreux termes suffixés en *-ous* n'avaient en réalité pas de dérivant. Pour autant, il s'agissait tout de même de structures suffixées dont l'opacité rendait l'identification certes parfois difficile, mais dont l'impact sur l'accentuation restait indéniable, au même titre que les structures séparables. J'optai donc pour une classification strictement morphologique en suffixes séparables et suffixes inséparables, car l'identification des structures séparables me permit de travailler sur les concepts fondamentaux de terminaison contraignante et de suffixe neutre. Le concept de terminaison fut donc utilisé dans mon travail pour qualifier les structures contraignantes dépassant les strictes limites de la suffixation et ayant un rôle au niveau phonologique.

J'illustre ici le concept de terminaison contraignante avec le terme *momentous*, car le suffixe *-ous* ne permet pas à lui seul de justifier la différence accentuelle existant entre *moment* et *momentous*.

'moment + -ous → mo'mentous

La combinaison de l'agrégat consonantique <nt> et du suffixe adjectival *-ous* est la condition nécessaire à la différence accentuelle entre le dérivant et le dérivé. La terminaison contraignante « C₂ + suffixe adjectival en *-ous* » bloque ainsi la référence au dérivant et impose un mode de calcul direct de l'accent. La séquence contraignante *-ntous* dépasse le cadre strict de la suffixation et relève donc du concept de terminaison. La présence de l'agrégat consonantique est fondamentale dans ces structures car je mis en évidence le fait que son absence faisait basculer le suffixe *-ous* dans la catégorie des suffixes neutres, comme l'exemple ci-dessous le démontre.

'poison + -ous → 'poisonous

L'absence de différence accentuelle dans l'exemple ci-dessus entre le dérivant et le dérivé démontre le caractère neutre du suffixe *-ous*. L'interaction morphologie - phonologie étant au cœur du système de l'École de Guierre, j'eus tôt fait de remarquer que les structures morphologiques avaient un impact considérable sur le plan phonologique et notamment sur la position de l'accent principal. Je démontrai ainsi que le suffixe *-ous* se trouvait à la croisée de deux principes accentuels fondamentaux en anglais, à savoir la notion de contrainte et celle d'isomorphisme.

Le suffixe *-ous* est présent dans mon corpus au sein de nombreuses terminaisons contraignantes.

→ *-inous* : *aluminous, bituminous, gelatinous, platitudinous*.

→ *-ulous* : *acidulous, fabulous, fantabulous, tuberculous*.

→ *-eous* : *advantageous, calcareous, courageous, instantaneous*.

→ *-ious* : *avaricious, delicious, efficacious, judicious*.

→ *-uous* : *ambiguous, conspicuous, perspicuous, superfluous*.

→ C₂ + suffixe adjectival : *asbestous, disastrous, enormous, tremendous*.

Les terminaisons *-eous*, *-ious* et *-uous* relèvent de la règle $-\{i,e,u\}+V(C_0(e))$, plus communément appelée « règle des mots en *-ion* » dans les ouvrages généraux de phonologie. J'appliquai une méthode rigoureusement identique en vérifiant systématiquement le caractère contraignant de chaque terminaison, afin de tester la fiabilité de la classification générée par les héritiers directs de L. Guierre. Si en testant le caractère contraignant d'une terminaison, un seul cas de dérivation me permettait d'observer une différence accentuelle entre un dérivant et son

dérivé, j'en attribuais la différence à la terminaison testée, prouvant ainsi qu'elle était contraignante et qu'elle bloquait la référence au dérivant, en empêchant l'application du principe d'isomorphisme. Par exemple, la relation 'cartilage → ,carti'laginous permet d'observer une différence accentuelle imputable à l'influence de *-inous*, qui est donc une terminaison contraignante à laquelle on associe une accentuation antépénultième. L'accentuation antépénultième associée à la terminaison contraignante s'applique par extension à tous les suffixés n'ayant pas de dérivant (exemple : *serotinous*), aux termes pour lesquels *-inous* dépasse les limites de la stricte suffixation (exemples : *glutinous*, *resinous*) ou encore aux relations dérivationnelles ne permettant pas de constater de différence accentuelle en raison de l'application de principes accentuels distincts avec des résultats concordants (exemples : *pre'stige* → *pre'stiginous*, *'resin* → *'resinous*). Je procédai ainsi à une recherche systématique des dérivants de tous les mots en *-ous* de mon corpus.

Les terminaisons contraignantes *-inous*, *-ulous*, *-eous*, *-ious*, *-uous* sont associées à une accentuation antépénultième. Les trois dernières terminaisons ont un statut dissyllabique en raison de la conception phonologique de la syllabe, comme je l'ai mentionné précédemment. Le système accentuel de l'anglais se situant à un niveau de représentation phonologique, il est donc parfaitement cohérent de considérer la syllabe phonologique comme la seule entité permettant de traiter les faits accentuels. Les cinq terminaisons associées à une accentuation antépénultième montrent une parfaite régularité du système avec un nombre extrêmement faible d'exceptions (9 au total), parmi lesquelles : *'alkalinous* et *'spirituous* en /1000/ ; *scarla'tinous* en /(-)10/.

La terminaison C₂ + suffixe adjectival *-ous* est quant à elle associée à une accentuation pénultième. Toutefois, au contraire des autres terminaisons contraignantes, l'accentuation des adjectifs relevant de cette structure fut plus compliquée à traiter en raison du nombre plus important d'exceptions accentuées en /100/ (8 au total). On citera par exemple *'cavernous*, *'chivalrous*, *'hazardous*, *'rivalrous* ou *'tyrannous*.

J'explorai plusieurs pistes tentant de rendre compte du nombre anormalement élevé d'exceptions au sein de cette structure, de la résistance du principe isomorphique des structures avec C₂ à la structure segmentale de l'agrégat consonantique, en passant par une réminiscence de la logique germanique consistant à accentuer la première syllabe de la base (en excluant les préfixes du calcul de l'accent). Je dus me résoudre à l'évidence : l'activité de chercheur consiste à émettre des hypothèses puis à adopter une démarche de travail permettant de les vérifier sur un objet donné, mais les résultats ne s'avèrent pas systématiquement concluants. La piste d'une

conservation du principe isomorphique est sans doute la plus plausible pour rendre compte des exceptions présentes au sein de ce sous-corpus.

Mon travail de thèse comprenait également tout un chapitre sur la dérivation neutre. J'entrepris de démontrer de façon systématique toutes les relations dérivationnelles attestées au sein de mon corpus et d'observer le lien accentuel et phonétique unissant les dérivants et leurs dérivés. Le suffixe *-ous* se révèle neutre dans de nombreux adjectifs relevant d'un mode de dérivation par juxtaposition :

'joy → *'joyous*
'bluster → *'blusterous*
'cancer → *'cancerous*

ou par substitution :

'virus → *'virous*
'arsenic → *'arsenous*
i'dolatry → *i'dolatrous*.

Je démontrai dans mon travail que le mode de dérivation n'exerçait pas d'influence sur le lien existant entre le dérivant et le dérivé. Les cas d'isomorphisme accentuel et phonétique sont attestés au sein des deux types de dérivation. La présence d'un suffixe neutre implique que le calcul de l'accentuation et la réalisation de la voyelle sous accent sont déterminés à partir de la structure du dérivant et que l'ajout d'un tel suffixe ne modifie en rien ces deux paramètres.

Le système des règles graphophonématiques ou règles de correspondance graphie-phonie permettant de déterminer la prononciation des voyelles accentuées à partir de la graphie, dont l'un des précurseurs fut A. Wijk (1966), fut ensuite utilisé pour observer la réalisation des voyelles sous accent dans les suffixés en *-ous*. Ce système de règles connut plusieurs évolutions, à l'initiative notamment des travaux de A. Deschamps (1994), et fut hiérarchisé par J.-M. Fournier (2010). L'algorithme de J.-M. Fournier permet ainsi de déterminer la valeur d'une voyelle accentuée à partir de son environnement graphique et sa hiérarchie repose sur trois paramètres intervenant à deux niveaux distincts. Les trois paramètres fondamentaux sur lesquels reposent les règles sont les suivants : la nature de la voyelle, le contexte à droite et le rang. Les deux niveaux distincts au sein desquels opèrent ces trois paramètres sont celui de l'environnement immédiat de la voyelle accentuée et celui la syllabe accentuée en contexte. J'utilisai cette méthode algorithmique rendant compte de la hiérarchie des règles graphophonématiques sur les suffixés en *-ous* de mon corpus et force fut de reconnaître l'extrême régularité des règles en anglais.

Le fonctionnement de l'algorithme se résume de la façon suivante : une règle calculée à partir du lien graphie-phonie permet de déterminer une « valeur ». La valeur d'une voyelle correspond à une série de prononciations appelées « qualités ». Les valeurs et les qualités des voyelles monographe et digraphe de l'anglais sont consignées dans un tableau de correspondance permettant de faire le lien entre les deux niveaux (J.-M. Fournier 2010). Les règles de graphie-phonie furent extrêmement efficaces pour justifier la prononciation des voyelles accentués des suffixés en *-ous*.

En premier lieu, je remarquai que le concept de motivation savante décrit dans les travaux de J.-M. Fournier sur les adjectifs suffixés en *-ic* (1990) pouvait expliquer certaines des exceptions présentes au sein de mon corpus. La motivation savante se définit comme le lien entretenu entre un dérivant et son dérivé afin de traduire l'origine scientifique de ce dernier. En effet, certains items présentant une terminaison contraignante qui interdit de fait toute référence au dérivant, tant au niveau accentuel qu'au niveau de la prononciation de la voyelle accentuée, conservent tout de même un lien phonétique avec le dérivant. Ce lien sert à marquer l'héritage savant de l'item désigné. Les suffixés en *-inous* suivants illustrent parfaitement le concept de motivation savante :

'chitinous, lum'baginous, 'prasinous, tor'pedinous, u'redinous

Tous ces adjectifs devraient présenter des voyelles entravées sous accent, en vertu de l'application de la règle de Luick²³, particulièrement efficace au sein des mots longs. Or, ils reproduisent tous la prononciation de la voyelle accentuée de leur dérivant respectif :

'chitin [aɪ], lum'bago [eɪ], 'prasine [eɪ], tor'pedo [i:], u'redine [i:]

Alors que le mode de calcul direct imposé par la présence de terminaisons contraignantes et l'application des règles graphie-phonie en vigueur devraient déclencher l'apparition de voyelles entravées, il est évident au vu des exemples ci-dessus que les voyelles libres des dérivés sont héritées des dérivants et que la conservation de ces voyelles libres traduit l'origine savante des dérivés. Par conséquent, la motivation savante prend ici le pas sur le concept de terminaison contraignante. Ce constat marqua le point de départ de mon intérêt pour le comportement phonologique des unités « empruntées ». En effet, les dérivés ci-dessus ne sont pas à proprement parler des emprunts, car ils sont le fruit d'un processus morphologique issu de l'anglais, mais leurs dérivants respectifs sont clairement issus des langues classiques (latin et / ou grec) et ils appartiennent à des lexiques scientifiques très spécialisés. Cet héritage classique et scientifique a un impact considérable sur la façon dont sont prononcés ces items. Je reviendrai

²³ La règle de Luick déclenche la valeur entravée de la voyelle sous accent dans les mots longs (trissyllabiques et plus) qui sont accentués au minimum sur la syllabe antépénultième (exemples : *elephant, paradise*).

sur ces considérations plus longuement lors de la partie consacrée à mes recherches sur les emprunts.

Le point qui selon moi constitua une avancée significative au sein de mon cadre théorique fut le traitement des composés savants, tant au niveau de la conception même du terme de composition savante qu'à celui des régularités accentuelles émanant de cette classe. La définition du terme *neoclassical compounds* proposée par I. Plag (2003) se retrouve globalement à l'identique dans tous les courants théoriques anglo-saxons. Il correspond à des formations composées de lexèmes d'origine grecque ou latine permettant de générer du lexique en anglais. Ce type de formation lexicale se justifie par le développement rapide de la technologie et de la science, qui suscite l'apparition de nouveaux concepts et nécessite un nouveau vocabulaire. Les chercheurs français utilisent généralement le terme « composé savant » pour faire référence à ce type d'items. Dès 1979, L. Guierre remarque que les composés savants sont composés de quasi-morphèmes qui ont des propriétés morphologiques, sémantiques, phonologiques et distributionnelles qui les distinguent du reste du lexique. C'est en partie grâce à cet argument avancé par L. Guierre, mais aussi en raison du fait que l'accentuation des composés savants est peu abordée dans les ouvrages majeurs de phonologie anglaise rédigés par les chercheurs de l'École de Guierre, que je décidai de traiter la composition savante en profondeur.

Je remarquai que si d'un point de vue catégoriel, la dénomination « composition savante » faisait consensus parmi les chercheurs héritiers de l'École de Guierre et les morphophonologues anglo-saxons, le traitement morphologique et phonologique de ces items donnait lieu à de grandes disparités. La tradition morphophonologique française utilise un système de représentation alphabétique pour symboliser les éléments entrant en interaction. L. Guierre (1975 : 117) décrit ces structures savantes de la façon suivante. L'élément A correspond à une base savante qui peut compter jusqu'à trois syllabes, l'élément B représente la seconde base savante, qui est principalement monosyllabique, et l'élément C correspond aux suffixes. Sa conception de l'élément B évolua rapidement, puisque quelques années plus tard, il mentionna la présence massive de bases savantes dissyllabiques au sein de ses données (1979 : 764). Ce système de représentation alphabétique fut repris dans les travaux de J.-L. Duchet (1994 : 54) et M. Ginésy (2000 : 56). I. Trevian précisa les contours segmentaux des constructions savantes en mentionnant la présence quasi-systématique de la voyelle épenthétique <o>, une structure syllabique de type -CVC- ainsi que la présence de graphies héritées du grec (exemples : <Ch>, <ae>, <oe>) (2003 : 183).

Le terme le plus fréquemment utilisé pour qualifier les deux éléments entrant en interaction au sein de l'École de Guierre est celui de quasi-morphème, la tradition anglo-saxonne optant pour celui de *combining form*. Je préférerais quant à moi reprendre le terme de « base savante liée » utilisé par J.-M. Fournier et dont la dénomination reflète tant le contenu lexical inhérent à de tels éléments que leur capacité à entrer en interaction pour former des composés. H. Marchand (1969 : 129) démontra très tôt que de telles structures se situaient en réalité à la frontière entre affixation et composition et qu'il était difficile de trancher. Les morphologues, dont L. Bauer résume parfaitement la position (1983 : 213), considèrent que les *combining forms* se distinguent des affixes en ce qu'ils véhiculent un contenu sémantique plus important et qu'ils ont une plus grande liberté compositionnelle. On peut en effet retrouver certains de ces éléments tant à gauche (*initial combining forms*) qu'à droite (*final combining forms*) : *log(o)- / -logue*, *phil(o)- / -phile*. La tradition morphophonologique française, dont l'un des objectifs consiste à rendre compte des régularités accentuelles des structures morphologiques, prend en considération la frontière entre l'élément B et le suffixe C. L'élément B ne peut donc en aucun cas être considéré comme un suffixe, contrairement à ce que certaines entrées de l'*OED* pourraient laisser suggérer. Le traitement accentuel réservé dans la littérature à cette catégorie de mots dépend en grande partie de la conception morphologique qu'en ont les chercheurs.

Les sources dictionnaires, contraintes par leur mode de présentation alphabétique des entrées, attribuent un comportement accentuel à chaque séquence <-BC> (exemples : lorsque les séquences *-vorous* et *-phagous* sont attestées, l'accent tombe sur la syllabe précédant de telles séquences). La frontière entre la base et le suffixe n'est pas prise en compte et cela ne permet pas une économie de traitement, puisqu'il leur faut mentionner toutes les suffixations possibles avec la base *-vor-* ou *-phag-* dans leurs données. Les sources dictionnaires ne mentionnent que les structures les plus fréquentes et les plus régulières sur un plan accentuel. L. Guierre (1979 : 740), quant à lui, démontre une propriété des quasi-morphèmes, à savoir leur invariance accentuelle. Les structures ABC sont généralement accentuées sur la dernière syllabe de l'élément A lorsque l'élément B est monosyllabique.

Exemple : *hete'ro-+-nym-+-ous*.

A B C

Il mentionne toutefois les spécificités segmentales de certaines bases B qui semblent justifier une accentuation sur la syllabe pénultième, comme par exemple la présence d'une voyelle digraphe au sein de la base B.

Exemple : *dino-+-'saur-+-us*

A B C

J.-L. Duchet montra par la suite (1994 : 56-57) que certains suffixes ne modifiaient pas la structure accentuelle du dérivant et que ces composés savants suffixés relevaient en fait de la règle spécifique aux composés savants AB non suffixés qui privilégie une accentuation sur la première syllabe de l'élément A.

Exemple : 'ortho-+-dox

A B

Avec plus de 1000 composés savants suffixés en *-ous* attestés au sein de mon corpus, je disposais d'une base de données substantielle pour mettre en évidence certains des mécanismes morphophonologiques caractéristiques de cette catégorie. Mon analyse des structures morphologiques présentes démontra que le concept de composition savante tel qu'il était perçu par les spécialistes du domaine était en réalité trop restreint et qu'il pouvait être étendu à d'autres structures dont je présente ci-dessous les particularités. La première structure morphologique que j'associé au concept de composition savante fut celle comportant un élément A de nature préfixale. Je distinguai donc les deux structures suivantes au sein de la composition savante :

A + B + C

base + base + -ous

carni- + *-vor-* + *-ous*

zoo- + *-phil-* + *-ous*

préfixe + base + -ous

hyper- + *-troph-* + *-ous*

ana- + *-log-* + *-ous*

Cette extension du concept me paraissait nécessaire, dans la mesure où les composés de type <préfixe + base + -ous> relevaient de la même sphère sémantique et que la distinction entre préfixe et base était floue (exemples : selon les sources, *hyper-* et *ana-* sont considérés comme des préfixes ou des bases) et sujette à variation au sein des sources dictionnairiques. De plus, la finalité de mon travail étant de mettre en évidence les mécanismes sous-jacents permettant de rendre compte d'une éventuelle régularité accentuelle au sein de telles structures, je remarquai rapidement que le statut de l'élément A n'était absolument pas déterminant dans l'assignation de l'accent principal. Je ne fis donc pas la distinction entre préfixe contemporain productif (exemples : *a-*, *co-*, *dis-*) et préfixe historique (*apo-*, *cata-*, *epi-*) au sein des données, car elle n'était pas pertinente au niveau phonologique. De la même façon, j'intégrai les structures avec un élément A de type numéral à la classe des composés savants :

élément numéral initial + base + *-ous*

tri- + -loph- + ous

hexa- + -chaet- + -ous

Je m'aperçus bientôt que la représentation alphabétique ABC ne permettait pas de rendre compte de toutes les structures attestées au sein de la composition savante. En effet, quelques composés savants de mon corpus résultaient d'une interaction entre trois éléments et le suffixe *-ous* (exemples : *a-bio-gen-ous*, *dia-meso-gam-ous*). Je considérai qu'il s'agissait là également de structures correspondant aux composés savants et je les inclus dans la catégorie. Au sein du corpus, je notai que certaines bases initiales possédaient une entrée spécifique dans les bases de données (exemples : *homo-*, *iso-*, *ortho-*) tandis que d'autres n'en avaient pas (exemples : **lipo-*, **lyo-*, **dolicho-*). Cette différence reposait à mon sens sur la capacité de certaines bases à générer des composés. Plus une base est productive au sein de la composition savante et plus ses propriétés morphologiques, sémantiques et phonologiques émergent, lui permettant ainsi d'acquérir un statut institutionnalisé, matérialisé par une entrée spécifique. Pour autant, même si certaines bases peu productives n'ont pas d'entrée associée, elles ont une influence déterminante sur l'accentuation et reçurent en conséquence un traitement identique au sein de mon travail. Je dressai la liste de toutes les bases A et B intervenant dans le domaine de la composition savante en *-ous*. Ce travail, relevant plus de la lexicologie et de la sémantique que de la phonologie, se justifiait toutefois en raison de la dimension morpho-sémantico-phonologique du traitement des composés savants. Ce travail morphosémantique permit donc dans un premier temps d'étendre le concept de composés savants en anglais mais également de mettre en évidence ses limites.

Je pris soin d'identifier avec précision toutes les bases B²⁴ des 938 composés savants du corpus, ce qui ne fut pas sans poser un certain nombre de problèmes méthodologiques tant graphiques que sémantiques. J'eus tôt fait de mettre en évidence le caractère contraignant du suffixe *-ous* au sein de telles structures. Les nombreuses différences accentuelles observables entre les dérivants AB et les dérivés ABC correspondants le démontrent :

'calcifuge + -ous → cal' cifugous

'oxygen + -ous → o'xygenous

'collagen + -ous → col' lagenous

Le caractère contraignant du suffixe *-ous*, en interdisant toute référence au dérivant et en privilégiant ainsi un mode de calcul direct de l'accent, me conduisit à m'interroger sur

²⁴ Je ne parlerai ici que des bases B (et non des bases A) car elles seules ont une influence déterminante sur la position de l'accent.

l'influence de certaines bases B sur la position de l'accent, à la suite des remarques de L. Guierre sur les propriétés accentuelles spécifiques des quasi-morphèmes formulées précédemment. En constatant par exemple que tous les composés savants en *-gen-ous* étaient accentués sur la syllabe qui précédait (c'est-à-dire la syllabe antépénultième) et que tous les composés en *-chaet-ous* étaient accentués sur la base B *-chaet-* (c'est-à-dire la syllabe pénultième), j'en conclus que les accentuations en /(-)10/ et /(-)100/ étaient les deux schémas préférentiels au sein de ces formations. L'accentuation sur la syllabe pénultième ou au contraire le principe de rétraction accentuelle sur la syllabe antépénultième semblaient corrélés à l'influence des bases B sur l'accent. Je décidai donc de proposer une formalisation du système permettant de tester l'hypothèse d'une influence déterminante de telles bases suffixées. Il était donc possible de réaliser une typologie des bases B en fonction de leur influence sur l'accent. Le terme « influence » me parut adéquat, dans la mesure où il mettait en exergue le dynamisme interactionnel du système. La typologie repose sur l'encodage des schémas accentuels que j'effectuai en utilisant la méthode de représentation numérique de l'accentuation adoptée par R. Kingdon (1958). Cette vision certes statique de l'accentuation ne visait dans un premier temps qu'à rendre compte de la pluralité des schémas accentuels associés à chaque base pour pouvoir dans un second temps caractériser son influence dans une perspective plus dynamique. Par exemple, l'accent tombe systématiquement sur la syllabe précédant *-phagous*.

Exemples : *car'pophagous*, *po'lyphagous*, *sa'prophagous*.

Je classai donc *-phag-* dans les bases avec une influence de type <-1> pour indiquer que l'accent principal se situait systématiquement sur la syllabe précédant la base. Les composés en *-chaetous* étant à l'inverse tous accentués sur la base *-chaet-*, je classai cette base dans la catégorie des bases avec une influence de type <+1>, c'est-à-dire avec un accent sur la première syllabe de la base. Cette représentation héritée des travaux de R. Kingdon permet de faire apparaître plus clairement les cas attestés de variation au sein de la composition savante. Certaines bases furent ainsi considérées comme instables du point de vue de leur influence sur l'accentuation car elles généraient des schémas accentuels différents. L'exemple de la base *-spor-* est représentatif des cas de variation au sein du système.

-spor- = <-1> (5) *a'crosporous* / *,angi'osporous* / *,hete'rosporous* / *ho'mosporous* / *zo'osporous*
 = <+1> (3) *,mono'sporous* / *,poly'sporous* / *tri'sporous*

L'influence de la base *-spor-* sur la position de l'accent principal est qualifiée d'instable car dans les cinq premiers exemples, elle est associée à une influence de type <-1> (accent sur la syllabe précédant la base *-spor-*) tandis que les trois derniers cas illustrent une influence de type <+1> (accent sur la base *-spor-*).

Après avoir répertorié toutes les influences par rapport à la position de l'accent de chaque base B du corpus, je dus me positionner sur certains problèmes relevant selon moi du pur statut d'exception. Par exemple, je remarquai au sein des composés en *-genous* les deux comportements suivants :

-gen- = <+1> (1) *bi'genous*

= <-1> (87) *en'dogenous, ho'mogenous, ha'logenous*

Le rapport de force entre les deux comportements est déséquilibré, car je notai une seule influence de type <+1> contre 87 de type <-1>. Je considérai donc que *-gen-* relevait d'une influence de type <-1> et que cette base avait une influence stable sur la position de l'accent, le cas de *bigenous* étant interprété comme une exception. Je poursuivis mon raisonnement en tentant d'uniformiser le traitement accentuel des composés savants suffixés, tout en le reliant à la logique globale qui organise le fonctionnement de l'intégralité du lexique anglais. Le pourcentage important d'accentuation pénultième au sein de cette catégorie n'était pas sans me rappeler l'accentuation des emprunts en anglais et cela constituait selon moi une piste de réflexion. L'influence des bases B sur l'accentuation pouvait matérialiser une conservation de certaines propriétés accentuelles de telles formations dans les langues classiques. Je décidai toutefois lors de ce travail de thèse de relier les principes accentuels des composés savants suffixés aux autres paramètres régissant le système accentuel global.

J'utilisai la terminologie de bases « attractives » (base retenant l'accent), « non attractives » (base ne retenant pas l'accent) et « semi-attractives » (phénomène de variation avéré) pour qualifier les trois types d'influence mis en évidence par mon travail. Pour simplifier les conclusions de ma thèse, le comportement des bases B monosyllabiques, considérées comme non attractives car ne retenant pas l'accent en syllabe pénultième (l'accent est donc sur la syllabe antépénultième), fut relié au fonctionnement de la Règle Normale d'accentuation, qui attribue un accent antépénultième aux mots longs. Cela évitait ainsi d'attribuer une propriété spécifique aux bases B des composés qui, à l'évidence, relevaient d'une logique accentuelle similaire aux autres mots longs du reste du lexique anglais. De plus, il était assez difficile de déterminer si le type <-1> était imputable à une réelle influence des bases qui déclenchaient un phénomène de rétraction accentuelle, ou au contraire à une absence totale d'influence, générant ainsi un accent antépénultième en vertu de l'application de la Règle Normale. La terminologie « non attractive » me parut à l'époque adéquate pour qualifier les deux scénarios. Je formulai ainsi la règle de la façon suivante :

base savante non-attractive + *-ous* → Règle Normale (/(-)100/)

Les bases savantes monosyllabiques permettant de contribuer à une accentuation pénultième furent, quant à elles, qualifiées d'attractives. Leur potentiel d'attraction probablement hérité des propriétés phonologiques des langues classiques est indéniable, car l'accent principal est retenu sur la base en position pénultième et n'est pas soumis au phénomène de rétraction accentuelle constaté dans les composés savants avec des bases non attractives. Je formulai ainsi la règle suivante pour rendre compte de ce phénomène :

base savante attractive + *-ous* → /(-)10/

Enfin, les bases semi-attractives illustrent l'équilibre et le rapport de force existant entre l'application de la Règle Normale (accentuation antépénultième) et l'influence marquée des bases classiques qui retiennent l'accent en syllabe pénultième. La variation accentuelle importante attestée au sein de la composition savante montre que le système oscille entre les deux logiques, même si l'accentuation antépénultième reste tout de même plus attestée au sein du système.

Mon travail de thèse s'acheva sur une tentative de formalisation algorithmique du comportement accentuel des suffixés en *-ous* en anglais contemporain, à la manière des algorithmes proposés par J.-M. Fournier pour rendre compte du système accentuel de l'anglais et de la valeur des voyelles sous accent. Avec le recul, je trouve cette tentative maladroite, car il s'agit d'une représentation algorithmique peu aboutie, reprenant des paramètres développés par mon directeur de thèse. C'est l'expérience de l'enseignement, de l'encadrement et de la recherche qui me permit de comprendre cela bien des années plus tard.

Dans la partie suivante, je présente mes travaux post-thèse qui s'inscrivent directement dans la continuité de mes recherches sur les suffixés en *-ous* et la composition savante.

2.2 Travaux post-thèse

Les deux publications s'inscrivant dans le prolongement direct de mon travail de thèse et des idées du cadre théorique de l'École de Guierre traitent de l'accentuation des composés savants. Si mon article de 2013²⁵ sur le sujet est une reformulation plus adaptée et un approfondissement du chapitre de thèse consacré à l'accentuation des composés savants en

²⁵ Fournier, P. (2013), « Accentuation des composés savants en anglais contemporain : identification morphosémantique et enjeu phonologique », *CORELA* - Numéro 2, Volume 11 (<https://journals.openedition.org/corela/2971>).

- *ous*, celui de 2017²⁶ constitue un véritable prolongement de mon étude initiale. Un bref rappel de mes positions concernant l'accentuation des composés savants s'impose. J'avais démontré dans ma thèse que les composés savants en *-ous* avaient des propriétés morphosémantiques et phonologiques distinctes du reste du lexique anglais. Ainsi, j'avais constaté que l'accentuation pénultième était plus représentée dans cette catégorie d'items que dans le reste du lexique (à l'exception des emprunts). J'avais formulé l'hypothèse selon laquelle l'accentuation pénultième serait en réalité un héritage des principes phonologiques des langues classiques (latin et grec), langues auxquelles l'anglais avait massivement emprunté au sein de la composition savante. J'avais donc ainsi établi une typologie des bases B en fonction de leur influence sur la position de l'accent. L'influence était avérée si l'accent restait sur la base B, dite « attractive », avec une accentuation pénultième. Les bases B non attractives n'exerçaient en réalité aucune influence sur la position de l'accent qui subissait un phénomène classique de rétraction accentuelle sur la syllabe antépénultième en vertu de la Règle Normale. Enfin, les bases « semi-attractives » traduisaient une forme d'instabilité entre influence et attractivité accentuelle de la base (héritage des langues classiques) et application des règles accentuelles du système anglais.

Je me devais donc de tester les hypothèses formulées à partir de la classification des bases suffixées en *-ous*, en observant le comportement des mêmes bases en interaction avec un autre suffixe. Je choisis pour cette étude d'observer le comportement des composés savants suffixés en *-al* car ce suffixe a des caractéristiques sémantiques et segmentales similaires à celles du suffixe *-ous*, mais également parce qu'il est productif au sein de la composition savante. J'eus donc à nouveau en tout premier lieu un travail considérable au niveau morphosémantique pour constituer un corpus représentatif de composés savants suffixés en *-al*.

Je repris le concept élargi de composition savante que j'avais mis en évidence dans mon travail de thèse, car j'avais démontré que les éléments initiaux, préfixes ou bases, ne jouaient pas un rôle essentiel dans l'assignation de l'accent principal. Je repris la liste de toutes les bases B productives avec le suffixe *-ous* identifiées dans ma thèse et entrepris de rechercher tous les cas de suffixation en *-al* sur les mêmes bases au sein des trois sources dictionnaires suivantes : *OED*, *LPD* et *EPD*. L'*OED* retranscrit toutes les variantes orthographiques héritées du passage du système graphique des langues classiques vers le système anglais et je décidai de ne conserver que la forme citationnelle de la base, sans prendre

²⁶ Fournier, P. (2017), « Vers une tentative d'uniformisation du traitement accentuel des composés savants suffixés en anglais contemporain », *Anglophonia* 23, *Nouvelles frontières de la grammaticalité* (<http://journals.openedition.org/anglophonia/1210>).

en compte toutes les variations orthographiques et allomorphiques. En raison de contraintes combinatoires d'une part et de la notion de concurrence suffixale existant entre *-ous* et *-al* d'autre part, il s'avéra que certaines bases productives avec *-ous* ne l'étaient pas avec *-al*. Je ne conservai que les formes propres à l'anglais contemporain, en écartant toutes les formes obsolètes attestées dans l'*OED* ainsi que tous les composés sans transcription phonétique associée. Je parvins tout de même à rassembler 348 composés savants en *-al*, ce qui démontra une nouvelle fois le potentiel de création lexicale de cette catégorie. Les patrons morphologiques les plus représentés furent sans surprise les mêmes que ceux mis en évidence au sein des composés savants en *-ous*.

A + B + C

préfixe + base + *-al*

a- + *-pter-* + *-al*

élément numéral initial + base + *-al*

penta- + *-gon-* + *-al*

base + base + *-al*

bacterio- + *-phag-* + *-al*

Les patrons morphologiques des 348 composés savants se répartissent entre les trois structures proposées ci-dessus, avec un élément A pouvant être un préfixe, une base ou bien un élément de type numéral. Je souhaitais tester l'hypothèse selon laquelle les bases B dites attractives au sein des composés savants en *-ous* exerçaient également une influence déterminante sur la position de l'accent principal lorsqu'elles étaient suffixées en *-al*. La typologie des bases B présentée dans mon travail de thèse était-elle effective en tout type de suffixation ? Il s'agissait d'une tentative de formalisation du système accentuel des composés savants suffixés en anglais, tentative qui se solda par de nombreuses nouvelles questions et peu de conclusions satisfaisantes.

Je repris le système de codage des schémas accentuels effectués à partir de la base B que j'avais mis en place dans mon travail de thèse et qui s'inspirait des travaux de R. Kingdon (1958). Quelques exemples illustrent ce système de codage de l'accentuation pour les bases suffixées en *-al* :

- la base *-pter-* est associée au codage <-1>, car l'accent tombe systématiquement sur la syllabe qui précède la base.

Exemples : *ˌcole'opteral*, *'dipteral*, *he'mipteral*.

- la base *-neur-* est associée au codage <+1>, car l'accent tombe systématiquement sur la seule syllabe de la base B.

Exemples : *ab'neural*, *dermo'neural*, *epi'neural*.

- la base *-mycet-* est associée au codage <+2>, car l'accent tombe systématiquement sur la deuxième syllabe de la base B.

Exemple : *myxomy'cetal*.

Sur les 68 bases savantes productives avec le suffixe *-al*, j'identifiai 52 bases associées à un codage unique par rapport à la position de l'accent principal, ce qui signifiait que ces bases B ne permettaient pas de générer de variation accentuelle. En revanche, 16 bases B étaient associées à plusieurs schémas accentuels, comme par exemple la base *-fug-*.

-fug- = <-1> : *cel'lifugal*, *a'xifugal*, *ba'sifugal*

<+1> : *culici'fugal*, *helio'fugal*

<-1/+1> : *cen'trifugal* / *centri'fugal*

Si la plupart des composés savants en *-fugal* sont accentués sur la syllabe qui précède la base comme dans *cel'lifugal*, *a'xifugal* ou *ba'sifugal* et sont donc associés au codage <-1>, on remarque toutefois la présence du codage <+1> avec un accent principal tombant sur la base dans *culici'fugal* ou *helio'fugal* et un cas avéré de variation accentuelle illustré par le codage <-1/+1> dans *cen'trifugal* / *centri'fugal*. Je comparai ensuite les résultats du codage obtenus pour le suffixe *-al* avec ceux effectués dans mon travail de thèse pour observer le pourcentage de correspondance. La comparaison montre que sur les 68 bases communes aux deux suffixes, 44 d'entre elles sont associées à un schéma accentuel identique, soit une correspondance de 65%. Si le pourcentage de correspondance ne fut pas aussi élevé qu'escompté, je notai toutefois avec satisfaction que les correspondances ne se cantonnaient pas seulement aux cas de codage de type <-1>, qui ne traduisaient aucune réelle influence de la base, mais plutôt une accentuation par défaut sur la syllabe antépénultième, en vertu de l'application de la Règle Normale. Il y avait également des correspondances entre bases associées à un codage de type <+1>, ce qui suggérait des recouvrements au niveau de l'attractivité des bases.

Par exemple, les bases savantes *-flor-*, *-chaet-*, *-gyr-*, *-rhiz-* ou *-hedr-* sont associées à un codage de type <+1>, tant avec le suffixe *-ous* qu'avec le suffixe *-al*, ce qui signifie que ces bases sont accentuées. Cela me permit donc de formuler l'hypothèse selon laquelle ces bases étaient attractives, quel que soit le suffixe impliqué. Cependant, 24 bases savantes (35% des bases du corpus) sont associées à un codage différent, ce qui se traduit par des différences accentuelles, selon qu'elles sont suffixées en *-ous* ou en *-al*. Le comportement de la base *-ton-*, illustré ci-dessous, est particulièrement révélateur :

-ton- + *-ous* = <-1> (6 cas répertoriés)

-ton- + *-al* = <-2> (2 cas répertoriés) et <+1> (9 cas répertoriés)

Le comportement et l'influence de la base *-ton-* sur la position de l'accent seraient donc différents selon le suffixe impliqué dans la formation. Elle serait ainsi non attractive avec le suffixe *-ous* et permettrait l'application de la Règle Normale avec une accentuation sur la syllabe antépénultième. Mais lorsqu'elle est suffixée en *-al*, elle deviendrait attractive avec le codage de type <+1> mis en exergue dans l'exemple ci-dessus et un schéma accentuel sur la syllabe pénultième.

Ce travail de comparaison de l'attractivité des bases fut difficile à mener en raison de plusieurs facteurs, dont celui de la faible productivité de certaines d'entre elles. Le degré de correspondance de 65% n'était certes pas négligeable, mais il ne correspondait qu'à la comparaison de deux suffixes. Qu'en serait-il si l'intégralité des cas de suffixation était prise en compte ? Je décidai donc de clore cette étude en envisageant d'autres pistes permettant de rendre compte de l'accentuation des composés savants suffixés, dont certaines entraient potentiellement en contradiction avec le paramètre de l'attractivité des bases.

Un critère potentiellement déterminant dans l'accentuation des composés savants suffixés est celui de l'existence autonome de la séquence <BC>. En d'autres termes, si la séquence <BC> correspond à un mot attesté dans les dictionnaires, l'élément A est potentiellement séparable d'un point de vue morphosémantique. La frontière entre composition savante et structure séparable n'est toutefois pas si simple à déterminer. Je menai ainsi une recherche morphosémantique sur les structures se situant à la frontière afin de déterminer l'influence potentielle de ce paramètre sur l'accentuation. Le cas de *neural* est intéressant car le lien sémantique entre l'adjectif *neural* et la base *-neur-* suffixée en *-al* est indéniable dans les formations telles que *dermoneural* ou *exoneural*. De plus, tous les termes en *-neural* sont accentués sur la syllabe pénultième, c'est-à-dire sur *-neur-*. Il semblerait donc y avoir une correspondance accentuelle entre l'existence autonome de l'adjectif <BC> et l'accentuation des composés savants correspondants accentués de façon identique. Comment devais-je donc interpréter de tels cas ? Comme des composés savants montrant un phénomène de type interactionnel entre les éléments ou bien comme des structures séparables avec un élément A de type savant ? J'en vins donc à remettre en question l'hypothèse d'une influence des bases savantes B dans de telles structures d'autant plus que d'autres exemples semblaient venir confirmer mes impressions.

Exemple : 'gyral / ,circum'gyral, ,cosmo'gyral, ,inter'gyral, ,intra'gyral

L'existence autonome de l'adjectif *gyral* semble avoir une influence sur l'accentuation des composés savants en *-gyral*, pour lesquels on ne constate aucun cas de rétraction accentuelle au-delà des limites posées par l'adjectif. De plus, je constatai que l'absence de relation

sémantique entre l'adjectif *genal* (relatif à la notion anatomique de « joue ») et les composés savants en *-genal* se traduisait par une rétraction de l'accent dans les composés allant au-delà des limites de l'adjectif. En effet, tous les composés savants en *-genal* sont accentués en /(-)100/, c'est-à-dire sur la syllabe précédant *-genal*. Mais que penser de l'adjectif *podal* et des composés savants en *-podal* ? En effet, les composés savants en *-podal* sont sujets à la variation accentuelle avec des accentuations pénultièmes ou antépénultièmes, traduisant une forme d'instabilité entre les deux logiques de conservation de l'accentuation de l'adjectif *podal* ou au contraire application de la Règle Normale relative aux bases non attractives au sein de la composition savante. Il m'est toujours difficile à l'heure actuelle d'évaluer l'influence d'un tel argument sur la logique accentuelle des composés savants.

Le deuxième paramètre pouvant exercer une influence sur l'accentuation des composés savants semblait lié à la productivité des bases B. Je remarquai dans mes données que plus une base B permettait de générer de formations, et plus les schémas accentuels tendaient à être réguliers. À l'inverse, l'apparition de variantes accentuelles semblait corrélée au faible nombre d'items produits à partir d'une base.

J'interprétei enfin ces phénomènes de variation accentuelle à la lumière du degré d'instruction des locuteurs. Les cas de variation concernent généralement les schémas accentuels en /(-)10/ et en /(-)100/. La proportion d'accentuation sur la syllabe pénultième est plus élevée au sein de la composition savante que pour n'importe quelle autre partie du lexique anglais, à l'exception des emprunts. De plus, ces composés savants relèvent de domaines scientifiques et techniques, ce qui implique que ces items sont partagés par un faible nombre de locuteurs instruits. Les accentuations pénultièmes (c'est-à-dire sur la base B) seraient donc un marqueur du caractère savant du composé et la maîtrise de cette accentuation pénultième serait également synonyme d'appartenance à un microcosme savant. Les locuteurs « moins instruits », dans le sens où ils ne maîtriseraient pas les subtilités sémantiques savantes de ces items, opteraient davantage pour une accentuation antépénultième, caractéristique du comportement phonologique des mots longs en anglais. Cette hypothèse d'une accentuation pénultième permettant d'une part d'identifier l'origine savante de l'item et illustrant d'autre part le degré d'instruction du locuteur est intéressante et ambitieuse, mais reste très difficile à démontrer, car une telle expérience est à la croisée entre phonologie, sémantique et sociolinguistique. Pour autant, elle est séduisante car elle permettrait d'expliquer le phénomène important de variation accentuelle constaté au sein de cette catégorie d'items. Dans cette étude de 2017, je questionnai toutefois la légitimité des schémas accentuels proposés par les sources dictionnaires. Comment les personnes en charge des dictionnaires de prononciation

déterminent-elles l'accentuation de composés savants, dont l'utilisation est réduite à un microcosme scientifique ? Il est fort à parier que ces items intègrent les dictionnaires par l'intermédiaire de locuteurs instruits, maîtrisant le domaine et pour lesquels les accentuations sur la syllabe pénultième sont la marque de l'héritage scientifique des langues classiques. Un test confrontant les données dictionnaires à des données orales permettrait très certainement de révéler une proportion plus élevée d'accentuations antépénultièmes chez les locuteurs anglophones *lambda* par rapport aux schémas accentuels proposés par les dictionnaires. Je terminai cette étude en citant le cas du composé savant *conjugal*, composé savant de par sa formation morphologique mais dont la sphère sémantique ne correspond plus à un domaine savant. Ce terme est accentué en /100/ et relève donc de la Règle Normale. Il fonctionne différemment des deux autres termes construits à partir de la base *-jug-* que sont *extra-jugal* et *subjugal* et qui sont accentués en /(-)10/. L'importante fréquence d'utilisation de *conjugal*, associée à l'argument selon lequel cet adjectif ne relève pas ou plus de la sphère savante sur un plan sémantique, sont autant d'éléments permettant de justifier une régularisation accentuelle sur la syllabe antépénultième caractéristique des mots longs en anglais, tandis que la fréquence très faible et le domaine sémantique restreint des deux autres adjectifs vont de pair avec une accentuation pénultième, reflet de leur héritage savant et de leur faible productivité. Cette étude dont le but était d'uniformiser le comportement accentuel des composés savants souleva en réalité plus de questions qu'elle n'apporta de réponses. Mais elle offrit cependant toute une série de pistes à explorer.

L'ouvrage co-écrit en 2022²⁷ avec ma collègue I. Girard, membre du groupe « morphophonologie », illustre parfaitement notre ancrage profond au sein de l'École de Guierre. Il s'agit avant tout d'un ouvrage pédagogique visant à initier les étudiants de licence à la pratique de la transcription phonétique anglaise. Les points abordés et l'évolution des problématiques au sein de l'ouvrage sont basés sur notre propre expérience en tant qu'enseignants de phonétique-phonologie. Cependant, même si la finalité de l'ouvrage est avant tout pédagogique, l'organisation des chapitres et la façon de présenter les faits phonétiques reflètent notre ancrage scientifique dans la théorie de la « morphophonologie ». En effet, nous accordons une place fondamentale dans l'ouvrage aux correspondances graphie-phonie, tant au niveau vocalique que consonantique et les correspondances majeures sont illustrées par des inventaires le plus souvent exhaustifs. Le traitement réservé à l'accentuation lexicale démontre notre croyance en des problématiques accentuelles profondément corrélées à des paramètres

²⁷ Fournier, P. & I. Girard (2022), *Manuel de transcription phonétique de l'anglais*, Paris : Ellipses.

d'ordre morphologique. Les terminaisons contraignantes, les suffixes neutres et les préfixés sont abordés de façon méthodique afin de montrer aux étudiants que l'accentuation lexicale ne se déduit pas de la prononciation (et donc du poids syllabique), mais qu'elle est le résultat d'une interaction entre des paramètres morphologiques, syntaxiques et segmentaux. C'est la raison pour laquelle je qualifie volontiers cet ouvrage de manuel pédagogique, mais s'inscrivant dans une pédagogie dictée et orientée par des années de recherche en morphophonologie. Il se situe donc selon moi à la croisée entre manuel pédagogique et manifeste de recherche sur les spécificités de mon cadre théorique.

Depuis 2011, la quasi-totalité de mes préoccupations en termes de recherche se focalise autour d'une seule et même question : quelle logique sous-tend la prononciation des emprunts en anglais ? Dans la troisième partie de cette synthèse, j'exposerai tout d'abord les raisons qui m'ont conduit à me questionner sur cette problématique, puis j'organiserai mon propos de la façon suivante. Je présenterai les résultats majeurs de mes recherches en distinguant les langues sources étudiées, car je remarquai très tôt que les études consacrées aux emprunts aux langues telles que le français, l'italien ou le japonais permettaient de faire émerger des phénomènes différents, ou que je conceptualisai comme tels tout du moins à l'époque. Il me fallut plusieurs années pour proposer des tentatives de formalisation du système phonologique des emprunts en anglais, ma première tentative somme toute assez naïve datant de 2016.

Si la troisième partie de ma synthèse présente les résultats marquants de mes études sur les emprunts sans pour autant les relier sur un plan théorique, la quatrième et dernière partie expose mes tentatives de formalisation du système des emprunts en faisant le lien avec les études présentées dans cette troisième partie.

Chapitre 3 : La prononciation des emprunts en anglais

L'étude des composés savants suffixés en *-ous* me permit de mettre en évidence une spécificité accentuelle inhérente à cette catégorie et globalement peu répandue au sein du lexique anglais, à savoir une accentuation sur la syllabe pénultième. Cette accentuation pénultième dans les composés savants, dont l'attribution semble relever tant des caractéristiques phonologiques héritées des langues classiques que d'une volonté du système accentuel d'indiquer l'origine étrangère et savante de l'item, n'est pas sans rappeler l'accentuation des mots étrangers en anglais. Les emprunts sont ainsi massivement accentués sur la syllabe pénultième en anglais et présentent des spécificités graphophonématiques, sémantiques et segmentales, entre autres, distinctes du reste du lexique. L'idée me vint ainsi très vite, à la suite de mon étude sur les bases savantes classiques, de travailler sur la logique accentuelle régissant le système des emprunts en anglais.

De plus, une rencontre allait s'avérer déterminante dans mon cheminement. À l'occasion de la journée des doctorants de l'école doctorale SHS durant laquelle je présentai mon travail de thèse²⁸, une discussion informelle avec un doctorant italien occupant un poste de lecteur à l'Université de Tours allait profondément influencer mes recherches. En effet, lors d'une conversation sur nos travaux respectifs, il évoqua la présence en anglais de nombreux emprunts d'origine italienne et je mentionnai alors la fameuse règle des mots « italiens », qui déclenche un schéma accentuel en /(-)10/ lorsque l'emprunt en question relève d'une structure segmentale spécifique. Il fut intrigué par la nomenclature utilisée pour cette règle et me dit que bien que l'italien soit une langue à accent lexical et que la plupart des items soient effectivement accentués sur la syllabe pénultième, une proportion non négligeable de mots italiens est accentuée sur la syllabe antépénultième et, dans une moindre mesure, sur la syllabe finale. Quelle ne fut pas ma surprise d'apprendre cela, moi qui m'étais toujours imaginé que les mots étaient accentués en /(-)10/ en italien et que c'était la raison pour laquelle la règle des mots « italiens » avait été mise en évidence par les morphophonologues de l'École de Guierre ! Je faisais à cette époque une confusion majeure entre une règle accentuelle déduite à partir des spécificités du système anglais, comme c'était le cas de la règle des mots « italiens », et une correspondance accentuelle éventuelle entre l'accentuation des mots en italien et les emprunts correspondants en anglais. Je n'aurais jamais imaginé un seul instant que cette règle était en réalité une règle accentuelle pédagogique fondée à partir des régularités segmentales de

²⁸ Fournier, P. (2010), « Le suffixe *-ous* en anglais contemporain : à la croisée des principes accentuels », communication effectuée lors des Journées de l'Ecole Doctorale SHS de l'Université de Tours (3 juin).

l'anglais. Ce qui signifie qu'elle ne s'appliquait pas seulement à des mots d'origine italienne et que de plus tous les emprunts italiens en anglais ne relevaient pas de cette règle. Ma naïveté intellectuelle de l'époque fut vite chassée dès lors que je commençai à me pencher d'une part sur le fonctionnement de cette règle des mots « italiens » en anglais et d'autre part à observer les cas de correspondance accentuelle entre l'italien et l'anglais, comme autre dynamique de la logique accentuelle. Ceci marqua le point de départ de mes travaux sur l'accentuation des emprunts.

Enfin, le dernier élément déterminant fut la lecture des travaux de K. Church (1985). Cette citation est extraite de son article :

It is clear (...) that the stress of Italian loans is not dependent on the weight of the penultimate syllable (...) I believe that speakers of English adopt what I like to call a pseudo-foreign accent. That is, when speakers want to communicate that a word is non-native, they modify certain parameters of the English stress rules in simple ways that produce bizarre "foreign sounding" outputs.

La partie de la citation relative à l'inefficacité de la notion de poids syllabique comme moteur de la dynamique accentuelle au sein des emprunts italiens confirma la position que je partageais déjà à l'époque, en vertu de l'influence des principes de l'École de Guierre – bien qu'une telle position affirmée dès 1985 soit déjà en soi révolutionnaire – et la mention d'un schéma accentuel spécifique aux emprunts au sein du système anglais (*a pseudo-foreign accent*) alluma une étincelle. Cela fit d'une certaine façon écho à ce que j'avais mis en évidence de façon moins pertinente au sein des composés savants. Ainsi, le système anglais disposait d'un marqueur de nature phonologique pour indiquer l'origine étrangère d'un item. J'allais ainsi passer les dix années suivantes de ma carrière de chercheur à méditer et à vérifier cette hypothèse.

Ce qui ne manque toujours pas de me surprendre concernant cet article de K. Church, c'est qu'il ne se revendique absolument pas comme un article de recherche en phonétique et phonologie. L'objectif de son étude consistait à concevoir un outil permettant de surmonter l'écueil de l'accentuation lexicale dans les voix de synthèse automatiquement générées. Il tentait ainsi de créer un modèle permettant de placer correctement les accents lexicaux dans les phrases générées, mais il se rendit vite compte que le paramètre du poids syllabique ne permettait que partiellement de le faire, car toute une partie du lexique empruntée aux langues étrangères obéissait à une logique distincte. Une telle intuition dès 1985 me semble encore aujourd'hui remarquable.

S'ajoutait à ces différents facteurs le fait que l'accentuation des emprunts constituait un domaine relativement peu traité au sein de l'École de Guierre, à l'exception de quelques

principes que je développerai dans les sections suivantes (notamment la fameuse règle des mots « italiens », le système des terminaisons contraignantes et les logiques romane et germanique au sein du système anglais). Je décidai de me lancer pleinement dans ce domaine de recherche. Les résultats de mes travaux sur les différentes langues sources sont ordonnés de la façon suivante : les emprunts à l'italien, les emprunts au français et les emprunts aux autres langues sources (au japonais, au gaélique écossais et à l'arabe).

3.1 Les emprunts à l'italien

L'étude de l'accentuation des emprunts italiens en anglais marqua le début de mes recherches dans le domaine et c'est la raison pour laquelle je commence par cette langue source. Le développement de mes travaux sur les emprunts italiens suit un cheminement intellectuel chronologique, mais cette chronologie ne se reflète pas dans les années de parution de mes articles, en raison des délais de publication parfois importants.

L'article paru en 2021²⁹ correspond ainsi, en réalité, à une étude effectuée en 2011. Je testai dans cette étude la réalité et la fiabilité de la règle des mots « italiens » telle qu'elle est formulée par les chercheurs de l'École de Guierre et je tentai de déterminer si elle correspondait à une quelconque réalité basée sur le système accentuel italien. En d'autres termes, les mots empruntés à l'italien et présentant une structure segmentale de type <-VCV#> sont-ils tous accentués en /(-)10/ et cette accentuation pénultième est-elle basée sur une quelconque réalité du système accentuel italien comme la terminologie utilisée pour qualifier la règle pourrait laisser penser ?

La règle des mots « italiens » est formulée de la façon suivante dans les ouvrages majeurs de l'École de Guierre (A. Deschamps *et al.* 2004, L. Guierre 1987, J.-M. Fournier 2010) :

/t, d, n, s, z/ + <a, e*, i, o, u># → /-10/ (*<e> non-muet)

Cette règle présente la particularité d'être à la croisée entre les structures phonétique, segmentale et graphique. En effet, la consonne prévocale au sein de la formation de type <- VCV#> est nécessairement une alvéolaire (à l'exception des consonnes /l/ et /r/) et la voyelle finale est une voyelle monographe prononcée. Les guillemets encadrant le terme « italiens » signifient que les items concernés par cette règle ne sont pas essentiellement d'origine italienne mais proviennent également d'autres langues telles que l'espagnol, le japonais ou encore

²⁹ Fournier, P. (2021), « Transcription et formalisation de l'accentuation des mots « italiens » en anglais contemporain », *Transcrire, écrire, formaliser 2*, F. Toupin & C. Collin (éds.), Collection des Travaux Linguistiques du CerLiCO : Presses Universitaires de Rennes, 53-68.

l'arabe. Cette règle ne doit son appellation qu'au fait que la majorité des emprunts concernés par cette structure segmentale sont d'origine italienne. Elle est basée sur une réalité grapho-phonético-segmentale de l'anglais et absolument pas sur la provenance réelle des items. Cela étant dit, je décidai de vérifier si cette terminologie associée à un schéma accentuel en /(-)10/ était un héritage direct des propriétés phonologiques des mots italiens concernés. Bien qu'elles soient exclues de la règle, je décidai d'inclure les alvéolaires /l/ et /r/ dans mon étude, afin d'observer le comportement accentuel des items présentant de telles propriétés segmentales. Le fait que l'italien soit également une langue avec un accent lexical a guidé mon choix car la comparaison accentuelle entre les deux langues s'en trouve grandement facilitée.

La répartition des schémas accentuels en italien est la suivante :

- 80% des mots sont accentués en /(-)10/.
- 16% des mots sont accentués en /(-)100/.
- 4% des mots sont accentués en /-1/.

L'accent lexical n'est pas fixe en italien et sa position est soumise à des règles.

Mon corpus est extrait de *LPD* et *EPD* à partir d'une recherche par structure segmentale <CV#> finale dans les items au minimum trisyllabiques (exemples : <*ta>, <*da>, <*na>, <*sa>, <*za>, <*la>, <*ra>) dont les données furent croisées avec les descriptions étymologiques proposées par l'*OED* afin de ne conserver que les emprunts directs à l'italien. Ce critère direct de l'emprunt, que j'applique systématiquement dans mes travaux, est fondamental, car il garantit un transfert de l'italien vers l'anglais sans transition par une langue intermédiaire qui pourrait avoir influencé la prononciation de l'emprunt lors de cette phase transitoire. Je notai ensuite tous les schémas accentuels des emprunts italiens en anglais et fis de même avec les mots italiens en utilisant un dictionnaire d'italien standard en ligne³⁰. Malgré une conservation exceptionnelle de la structure segmentale et orthographique lors de l'introduction du mot italien en anglais, je notai quelques modifications engendrées essentiellement par des problématiques concernant la retranscription du genre et du nombre.

Exemples : *broc(c)oli* de l'italien *broccolo*, *calamari* de l'italien *calamaro*, *cannelloni* de l'italien *cannellone*.

Le corpus final se composait ainsi de 140 emprunts italiens, dont 93 comportaient l'une des cinq consonnes mentionnées dans la règle des mots « italiens » (/t, d, n, s, z/) et 47 l'une des deux autres consonnes alvéolaires, à savoir /l/ ou /r/. Les résultats de la comparaison de la

³⁰ *HOEPLI* (version en ligne), dictionnaire d'italien standard.

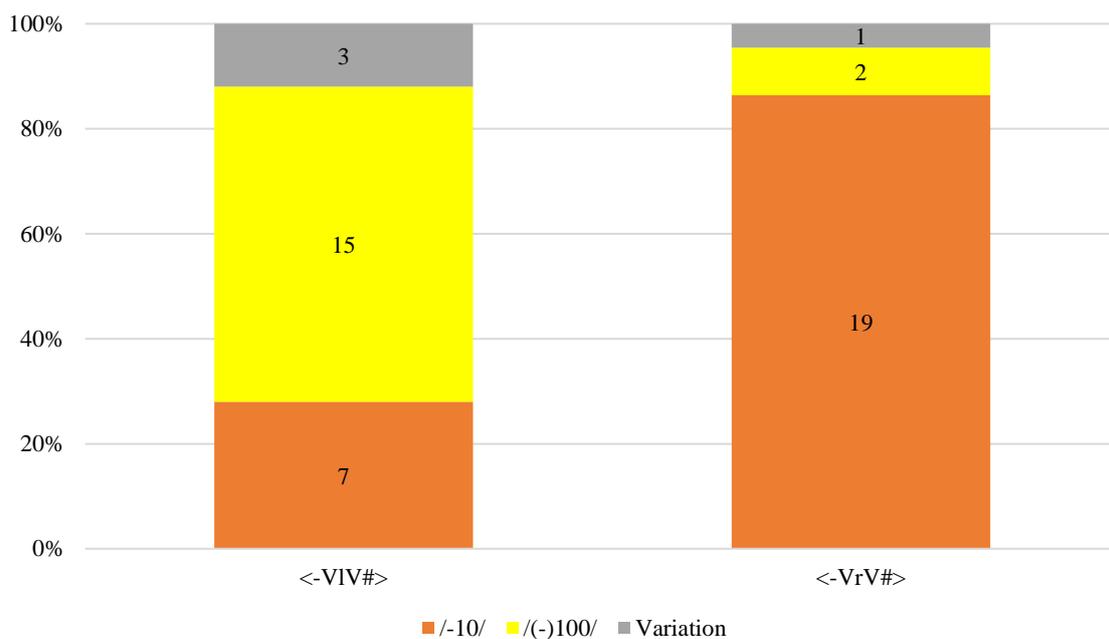
position de la syllabe accentuée en italien et en anglais sont consignés dans le tableau 2 extrait de mon étude.

Correspondance totale / -1/ → / -1/ ³¹	Correspondance totale /(-)10/ → /(-)10/	Correspondance totale /(-)100/ → /(-)100/
1 cas <i>tirami'sù</i> → <i>tirami'su</i>	86 cas <i>can'tata</i> → <i>can'tata</i> <i>macchi'ato</i> → <i>macchi'ato</i>	2 cas <i>'crimine</i> → <i>'crimini</i> <i>'timpano</i> → <i>'timpani</i>
Cas particuliers	Cas particuliers	Cas particuliers
<i>pie'tà</i> → <i>pie'ta</i> + variante en /100/ <i>ca'sino</i> (maison) et <i>casi'no</i> (jeu) → <i>ca'sino</i>	/	<i>in'cognito</i> → <i>inco'gnito</i> + variante en /0100/ <i>'patina</i> → <i>'patina</i> + variante en /010/

Tableau 2 : Comparaison de la position de la syllabe accentuée en italien et en anglais dans les emprunts avec une structure en <VCV#> (C = <t, d, n, s, z>).

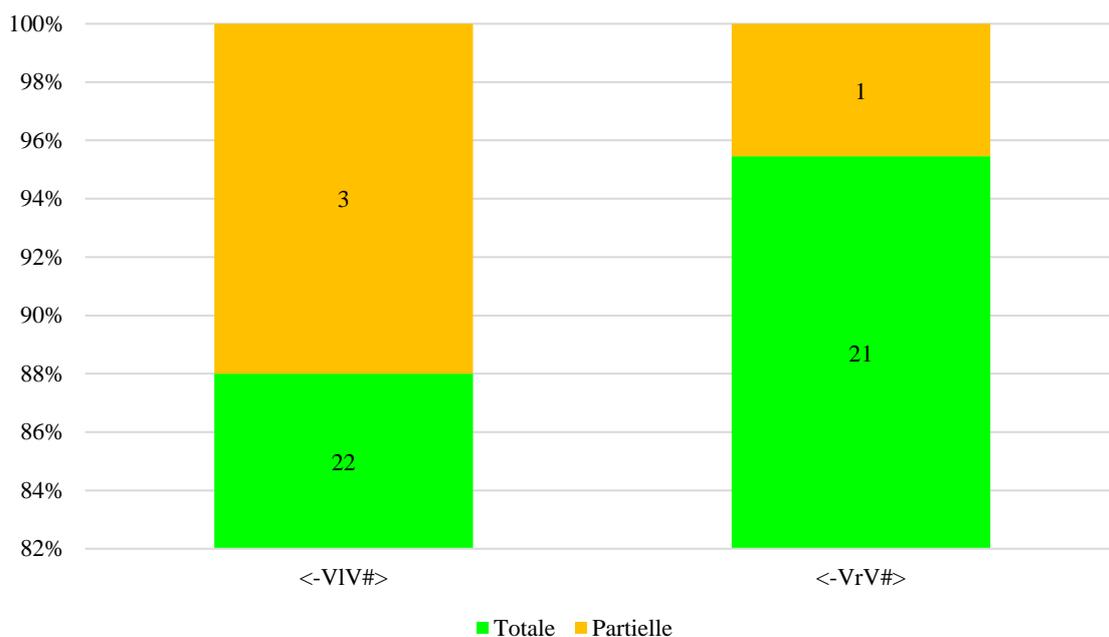
D'une part, les résultats montrent que la plupart des emprunts italiens en anglais sont accentués sur la pénultième, même si quelques rares cas d'accentuation sur la syllabe finale ou sur la syllabe antépénultième sont à noter. D'autre part, la comparaison accentuelle entre l'italien et l'anglais montre que l'emprunt reproduit très souvent l'accentuation italienne d'origine. Les rares cas de différence accentuelle entre les deux langues concernent des accentuations finales ou antépénultièmes, ce qui semble suggérer une instabilité accentuelle en anglais, mais une instabilité toute relative, étant donné le faible nombre d'items concernés (en tout et pour tout 4 cas de variation seulement et aucune réelle exception dans les 93 items testés). J'en conclus dans un premier temps que les emprunts italiens en anglais semblaient massivement reproduire l'accentuation italienne originelle. Les résultats concernant la répartition des schémas accentuels dans les emprunts italiens possédant une structure <-VIV#> et <-VrV#> en anglais sont présentés dans le graphique 5.

³¹ Les termes à gauche de la flèche sont les mots italiens et ceux à droite les emprunts en anglais.



Graphique 5 : Accentuation des emprunts italiens en <-VIV#> et <-VrV#> en anglais.

Les résultats montrent une tout autre réalité accentuelle au sein de ces structures. Les emprunts italiens en <-VIV#> sont davantage accentués sur la syllabe antépénultième (*ta'rantula*, *'broccoli*, *'buffalo*) que pénultième (*ci'cala*, *fi'nale*), ce qui justifie en partie l'exclusion de /l/ du fonctionnement de la règle des mots « italiens ». En revanche, je constatai que les emprunts avec une structure <-VrV#> étaient massivement accentués en /(-)10/ (*bra'vura*) en anglais. Il restait donc à déterminer les raisons pour lesquelles /r/ ne faisait pas partie de la formulation de la règle accentuelle en anglais. Les résultats de l'analyse du taux de correspondance accentuelle entre l'italien et l'anglais figurent au sein du graphique 6.



Graphique 6 : Taux de correspondance accentuelle des emprunts en <-VIV#> et <-VrV#> en italien et en anglais.

L'analyse des résultats du graphique 6 fut très instructive. Il s'avère que les pourcentages de correspondance totale entre la position de la syllabe accentuée en italien et en anglais sont à nouveau très élevés au sein de ces deux structures (88% de correspondance pour <-VIV#> et 95% pour <-VrV#>). Les cas de correspondance partielle, fruits des variations accentuelles observées au sein des sources dictionnairiques, sont marginaux. Cela me permet de démontrer que le taux de reproduction de l'accentuation italienne d'origine était très élevé au sein des emprunts italiens en anglais, et ce, quel que soit le statut de la consonne préfinale. Ainsi, même les mots en <-VIV#> principalement accentués sur la syllabe antépénultième en italien sont accentués sur cette même syllabe en anglais, selon le principe de conservation du schéma accentuel d'origine.

'cupola → *'cupola*
'gondola → *'gondola*
ta'rantola → *ta'rantula*
'broccolo → *'broccoli*
'bufalo → *'buffalo*

La règle des mots « italiens » en anglais est bien fondée sur une réalité du système accentuel italien, car la majorité des mots italiens sont accentués sur la syllabe pénultième. Le principe de conservation du schéma accentuel mis en évidence dans l'étude implique par conséquent

que la quasi-totalité des emprunts italiens accentués en /(-)10/ sont donc accentués de façon identique en anglais et relèvent de cette règle accentuelle. Si l'accentuation massivement antépénultième constatée au sein des structures de type <-VIV#> en italien justifie le fait que la consonne alvéolaire /l/ ne soit pas incluse dans la règle, pourquoi ne pas y avoir inclus les structures en <-VrV#> massivement accentuées en /(-)10/ en italien et en anglais ?

Le paradoxe réside ici dans l'appellation même de la règle des mots « italiens » : les guillemets, on l'a dit, illustrent la proportion importante d'items provenant de l'italien parmi ceux concernés, mais cette règle s'applique à des mots qui ne sont pas tous d'origine italienne. En effet, l'examen des items concernés par la structure <-VrV#> en anglais montre que les mots concernés sont principalement d'origine grecque et latine, comme le montrent les exemples ci-dessous :

'agora, 'amphora, a'naphora, 'camera, 'cholera, 'corpora, 'gerbera, 'viscera, ex'tempore.

Ces emprunts au grec et au latin sont massivement accentués en /(-)100/, ce qui justifie la non-intégration de la structure <-VrV#> au sein de la règle des mots « italiens », alors même que l'étude des mots italiens révèle un pourcentage important de correspondance accentuelle. La règle des mots « italiens » rend compte de l'accentuation des emprunts uniquement sur des critères segmentaux, en dépit du phénomène de reproduction des paramètres phonologiques des langues sources, qui entre en contradiction avec l'aspect pédagogique de cette règle. Le taux de correspondance accentuelle très élevé entre l'italien et l'anglais dans les structures <-VIV#> se traduit par la reproduction de la position de la syllabe accentuée originelle, et justifie la présence en anglais de nombreux mots accentués sur la syllabe antépénultième. Cette accentuation antépénultième explique l'exclusion des structures <-VIV#> de la règle des mots « italiens », alors même qu'elles attestent de l'incroyable régularité accentuelle des mots d'origine italienne en anglais, si le paramètre de reproduction de l'accentuation de la langue source est seul pris en compte.

Ces premiers résultats me confortèrent dans l'idée d'explorer le comportement accentuel des mots italiens dans un corpus plus important et d'étudier aussi le comportement accentuel des emprunts de diverses langues sources afin d'établir une logique globale de l'accentuation des emprunts en anglais.

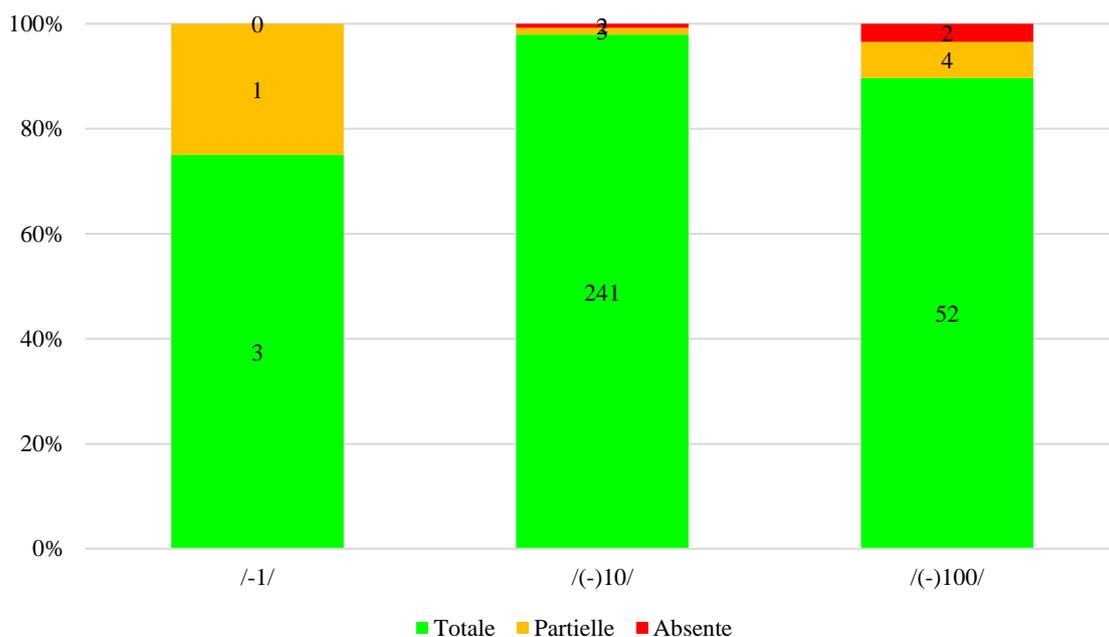
L'étude suivante consacrée aux emprunts italiens fut publiée en 2018³² et marqua un élargissement de mes perspectives de recherche. Je décidai pour cette étude d'élargir le corpus des emprunts italiens en anglais afin de tester les hypothèses formulées lors de la précédente étude et de procéder à un test oral sur des locuteurs anglophones pour observer leur prononciation des mots italiens. L'analyse de données orales constituait une grande nouveauté dans mes recherches. J'utilisai l'ouvrage de L. Pinnavaia (2001) consacré aux emprunts italiens dans l'*OED* pour constituer mon corpus. Son approche résolument lexicographique associée à un travail diachronique sur l'intégration des emprunts italiens en anglais fut d'une aide inestimable pour bâtir mon recueil de données. Je sélectionnai ainsi parmi ses inventaires tous les mots présentant une structure <-VCV#> en excluant les agrégats consonantiques en position préfinale ainsi que les cas de synérèse. Mon étude portant uniquement sur l'accentuation en anglais contemporain, je ne conservai que les emprunts associés à une transcription phonétique dans la base de données et non répertoriés comme obsolètes. Seuls les emprunts directs furent sélectionnés, de même que les structures syllabiques permettant une comparaison accentuelle entre les deux langues (exemple : l'emprunt *teleferic* provenant de l'italien *teleferica* ne fut pas conservé, car la différence segmentale entre les deux termes empêche toute comparaison). Le corpus d'analyse se composait de 308 emprunts provenant de l'*OED* auxquels j'ajoutai les données de *LPD* et *EPD* lorsque les items étaient attestés. En complément de l'étude parue en 2021, j'ajoutai également les schémas accentuels de l'anglais américain, afin d'observer le degré de correspondance entre les deux variétés. Je relevai l'accentuation des mots italiens dans deux dictionnaires de référence : F. Sabatini & V. Coletti (2005) et A. Arrizi & P. Robert (1999).

L'analyse de l'accentuation des mots en italien confirma qu'ils étaient massivement accentués sur la syllabe pénultième :

- 247 mots accentués sur la syllabe pénultième (79.9%).
- 57 mots accentués sur la syllabe antépénultième (18.4%).
- 4 mots accentués sur la syllabe finale (1.4%).

L'analyse de la correspondance accentuelle entre le mot italien et l'emprunt correspondant en anglais allait une nouvelle fois se révéler très instructive (voir graphique 7).

³² Fournier, P. (2018), "Stress Assignment in Italian Loanwords in English and its Impact on the Stressing of Foreign Words by Native English Speakers", Corela HS-24, *Multicultural Spoken English* (<https://journals.openedition.org/corela/5113>).



Graphique 7 : Analyse du taux de correspondance accentuelle entre l’italien et l’anglais britannique et américain dans les 308 mots du corpus.

Les résultats montrent très clairement qu’il y a un très fort taux de correspondance accentuelle entre le mot italien et l’emprunt correspondant en anglais et ce, quel que soit le schéma accentuel concerné. Si les résultats concernant la correspondance accentuelle au niveau de la syllabe pénultième étaient prévisibles, une telle régularité concernant les schémas sur la syllabe finale ou antépénultième l’était nettement moins. Cette étude montra ainsi qu’il n’y avait que quelques différences accentuelles, principalement matérialisées par des variantes.

Exemples : *in'cognito* → /(-)10/ ou /(-)100/ dans les 3 sources.

'pergola → /100/ dans les 3 sources + /010/ dans *LPD* (GB et US).

'rapido → /100/ ou /010/ dans *OED* (GB et US).

Je remarquai que les variantes pouvaient être intra ou inter-dictionnaires mais qu’elles ne relevaient jamais d’une différence dialectale. Les emprunts italiens sont en effet accentués de façon identique en anglais britannique et américain. Les résultats prouvent que l’accentuation italienne d’origine est reproduite lors du phénomène d’emprunt et qu’il s’agit de la principale dynamique accentuelle régissant le système. Le pourcentage très faible de différence accentuelle par l’intermédiaire du jeu des variantes illustre la faible autonomie du système anglais à ce niveau.

L’impact de l’accentuation italienne était tel que je décidai de pousser plus loin le raisonnement en analysant des productions de locuteurs anglophones. Je menai ainsi une

expérience sur 7 locuteurs³³ à qui il était demandé de lire une liste de 150 mots contenant 50 mots italiens (noms et adjectifs). Les mots italiens furent sélectionnés selon plusieurs critères. Il s'agissait de mots non empruntés par l'anglais, d'au moins trois syllabes et présentant une structure segmentale de type <-VCV#> (avec des lieux d'articulation distincts pour les consonnes). Ces mots relevaient de deux schémas accentuels distincts en italien : 26 étaient accentués sur la syllabe pénultième et 24 sur la syllabe antépénultième. Un échantillon représentatif des mots sélectionnés est présenté ci-dessous :

→ mots accentués sur la syllabe pénultième : *sagrestano, pericoloso, naftalina, panchina, premurosa, pedina, intriso, vampata, riposo, morbosio, grattacapo, gattino, lettino, lattuga.*

→ mots accentués sur la syllabe antépénultième : *solita, umido, gomito, prossimo, enfasi, ibrida, ettaro, valvola, ennesimo, gomitolo, edicola, pellicola, ostacolo, ossigeno.*

J'attendais donc des locuteurs anglophones qu'ils accentuent la syllabe pénultième ou antépénultième pour tester l'hypothèse sous-jacente selon laquelle le degré de correspondance entre la syllabe accentuée en italien et la syllabe accentuée dans leurs productions serait lié au niveau de compétence des locuteurs en italien. La corrélation entre reproduction de l'accentuation d'origine et compétences linguistiques des locuteurs me semblait être une piste intéressante à envisager. Cela faisait d'une certaine manière écho à mon hypothèse selon laquelle le niveau d'instruction des locuteurs pouvait avoir un impact sur la prononciation des mots savants en anglais dans mes études relatives à l'accentuation des composés savants. Parmi les participants, si certains n'avaient aucune notion en italien, d'autres l'avaient déjà étudié au lycée. Les 100 mots restants de la liste correspondaient à des items attestés en anglais et relevant de diverses origines afin de ne pas influencer les productions des participants concernant les mots italiens.

Exemples : *hearten, glisten, fumble, forbid, believe, understand, hurdle, listen, answer, kibbutznik, mandapam, mazurka, pastourelle, restaurant, résumé, gendarmerie, t'ai chi ch'uan, kimono, tokonoma, genipapo, guerilla, gemütlich.*

Sur les 350 occurrences réalisées par les participants (7 participants prononçant chacun 50 mots italiens), 340 étaient exploitables. Les schémas accentuels furent consignés par trois personnes (deux doctorants du laboratoire et moi-même) et les résultats mis ensuite en commun afin de

³³ Il s'agissait de lecteurs et de maîtres de langue en poste à l'Université Paris 13 et à l'Université de Tours.

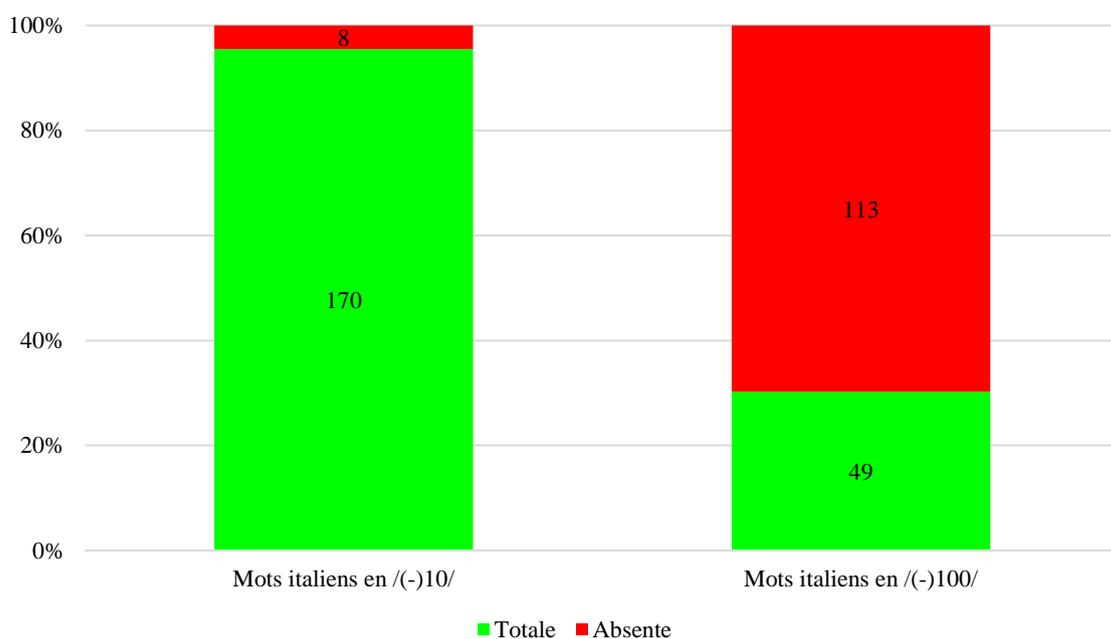
garantir l'objectivité de l'étude. Les données globales montrent que les locuteurs anglophones optent massivement pour l'accentuation pénultième lorsqu'ils sont face à un mot italien.

→ 82.3% des mots accentués en /(-)10/.

→ 16.8% des mots accentués en /(-)100/.

→ 0.9% des mots accentués en /(-)1000/.

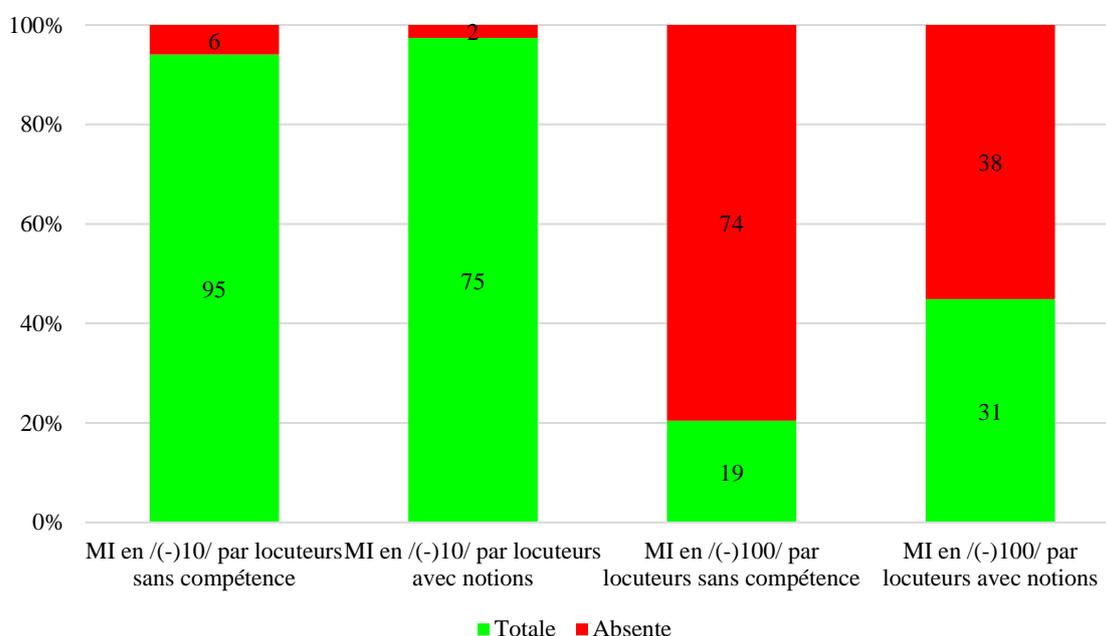
Concernant le degré de correspondance entre l'accentuation italienne et la position de la syllabe accentuée dans les réalisations des locuteurs, 178 occurrences relatives aux mots italiens accentués sur la pénultième furent récoltées contre 162 pour les mots accentués sur la syllabe antépénultième. Les premiers résultats sont édifiants (graphique 8).



Graphique 8 : Taux de correspondance accentuelle entre les mots italiens et les réalisations des locuteurs anglophones.

Les résultats du graphique 8 montrent que les locuteurs reproduisent massivement l'accentuation pénultième des mots italiens, tandis que l'accentuation antépénultième n'est que très imparfaitement reproduite, puisque les locuteurs optent massivement pour une accentuation pénultième. Je notai toutefois des disparités chez les participants, disparités semblant liées au degré d'exposition des locuteurs à la langue italienne, raison pour laquelle ma problématique évolua et prit ensuite en compte le degré de compétence des locuteurs en italien. Je procédai ainsi à la même analyse tout en dissociant les locuteurs possédant des notions même vagues

d'italien des locuteurs n'ayant aucune notion et n'ayant jamais été exposés à la langue italienne. Les résultats de cette seconde analyse sont présentés dans le graphique 9.



Graphique 9 : Taux de correspondance accentuelle entre les mots italiens et les réalisations des locuteurs anglophones selon leur niveau de compétence en italien.

Les résultats concernant les mots accentués sur la pénultième en italien ne sont pas modifiés si l'on intègre la variable du degré de compétence des locuteurs en italien. Les locuteurs anglophones accentuent massivement sur la pénultième les mots en italiens accentués en /(-)10/ quel que soit leur degré de compétence, ce qui génère un taux de correspondance accentuelle très important. En revanche, on constate une augmentation significative du taux de correspondance accentuelle au sein des mots accentués sur la syllabe antépénultième en italien lorsque la variable du degré de compétence est intégrée. Il apparaît en effet que plus un locuteur a été exposé à la langue italienne et plus son taux de reproduction de l'accentuation antépénultième des mots italiens augmente.

Ces résultats me permirent de dégager plusieurs conclusions, plus clairement formulées ici, je pense, que dans mon article de 2018 (datant, je le rappelle, d'une étude de 2012). Les locuteurs exposés à la langue d'origine ont une conscience raisonnée des mécanismes de transmission accentuelle de la langue source à la langue cible, ici l'italien et l'anglais. Plus le niveau de maîtrise d'une langue source augmente et plus le taux de reproduction du schéma accentuel d'origine au sein de la langue cible augmente également. Mais paradoxalement, les

résultats des tests sur les locuteurs anglophones sont tout de même assez éloignés de la correspondance quasi-parfaite démontrée par l'analyse des données dictionnaires, notamment en ce qui concerne la reproduction de l'accent antépénultième. Je formulai alors l'hypothèse selon laquelle les données phonétiques relatives aux emprunts italiens en anglais dans les dictionnaires de prononciation seraient retranscrites par des locuteurs bilingues (ou au minimum très cultivés pour avoir une telle maîtrise de la reproduction accentuelle), hypothèse que mes études suivantes me permirent d'ailleurs d'étayer. Cette étude démontra également que confrontés à des mots d'origine étrangère, les locuteurs anglophones *lambda* tendent à appliquer systématiquement un accent pénultième qui serait ainsi caractéristique et représentatif de l'accentuation des mots étrangers en anglais, créant ainsi un conflit entre influence des langues sources et autonomie du système phonologique anglais. Ce point sera développé dans la suite de ma synthèse et notamment dans la partie proposant une formalisation de la logique accentuelle des emprunts en anglais. Quant à l'hypothèse maladroitement formulée en fin d'analyse, selon laquelle l'afflux massif de mots italiens accentués sur la syllabe pénultième aurait remodelé le système accentuel de la totalité des mots empruntés en anglais en générant un schéma accentuel préférentiel en /(-)10/, je n'y crois plus guère aujourd'hui, car mes recherches montrèrent par la suite que l'accentuation pénultième était également très fréquente dans d'autres langues, et pas seulement en italien.

Ma dernière étude en date de l'accentuation des emprunts italiens remonte à 2016³⁴, mais elle constitue plus une tentative de formalisation du système qu'une étude générant de nouvelles données. Cette étude souligne la correspondance accentuelle quasi-parfaite existant entre l'italien et l'anglais en reprenant les résultats de l'article de 2018. Je remarquai dans cet article que si l'italien constituait un exemple parfait de transmissions des propriétés accentuelles d'une langue source vers l'anglais, les emprunts français et surtout japonais ne montraient pas une telle régularité, ce qui me conduisit à proposer une première tentative de formalisation du système dont je dévoilerai les grandes lignes dans le troisième chapitre.

La sous-partie suivante est consacrée à mes études sur les emprunts au français auxquels j'ai consacré un nombre incalculable d'heures.

³⁴ Fournier, P. (2016), « Nouvelles perspectives sur l'accentuation des emprunts en anglais contemporain », *Itinéraires*, 2015-2, *Stumbling blocks. Entraves et obstacles aux circulations. Approches pluridisciplinaires*, M. Lévêque & C. Parfait (éds.) (<http://itineraires.revues.org/2827>).

3.2 Les emprunts au français

Mon intérêt vis-à-vis du français s'explique par plusieurs raisons qui ne tiennent pas seulement au fait qu'il s'agit de ma langue maternelle. Le français est la langue à laquelle l'anglais a le plus emprunté pour des raisons principalement historiques. C'est pourquoi l'anglais a un stock lexical très important hérité du français, qui permet d'envisager un traitement quantitatif des faits accentuels. De plus, une partie des terminaisons contraignantes de l'anglais provient du français et notamment les terminaisons associées à une accentuation finale (pour des raisons que je développerai dans cette section), ce qui ne manqua pas de m'interpeler lors de mon doctorat. Enfin, les emprunts français constituent une sous-classe atypique d'emprunts car ils possèdent des propriétés phonologiques et notamment accentuelles spécifiques qui les distinguent des autres emprunts en anglais, ce qui m'incita à les étudier en profondeur.

Avant même d'entreprendre le résumé de mes résultats principaux concernant les emprunts français, je présente en quelques lignes les spécificités de l'accentuation française, car la compréhension du système accentuel français est indispensable pour rendre compte des phénomènes se produisant dans les emprunts français en anglais. Bien que le français adopte une logique accentuelle et un fonctionnement intonatif radicalement différents de ceux de l'anglais (le français relève de la catégorie des *syllable-timed languages* tandis que l'anglais appartient à la catégorie des *stress-timed languages*), certaines spécificités du système phonologique du français ont été reproduites en anglais. L. Hyman (1985) classe le français dans les langues fonctionnant avec un accent sur la syllabe finale. Le français n'a pas un accent dit lexical, mais un accent de groupe rythmique qui tombe sur la dernière syllabe de ce groupe (en dehors des contextes emphatiques). L'accent a donc un rôle essentiellement démarcatif en français, dans la mesure où il sert à indiquer les frontières rythmiques / intonatives qui sont généralement modelées sur les structures syntaxiques. Les domaines sémantique et pragmatique sont donc essentiels dans l'identification de tels groupes. Concernant la syllabe accentuée en fin de groupe, elle est systématiquement réalisée avec un ton montant, sauf au niveau du dernier groupe de la phrase, dont la syllabe finale est réalisée avec un ton descendant. Le ton montant sert à segmenter et à indiquer qu'un autre groupe intonatif vient compléter le précédent tandis que le ton descendant traduit la fin d'une phrase.

Exemple : Dans une ferme / il y a / des vaches / des moutons / des cochons.

↑ ↑ ↑ ↑ ↓

Je dois beaucoup aux travaux de A. Di Cristo & D. Hirst (1993) et M. Léon & P. Léon (2009) dans ma compréhension globale du système intonatif du français, dont je ne restitue ici qu'une minuscule part, celle qui fut essentielle dans mes études accentuelles sur les emprunts français en anglais.

Ainsi, même si les deux systèmes accentuels sont très différents, l'anglais utilise une accentuation finale au niveau lexical pour retranscrire la logique de l'accentuation finale des groupes intonatifs en français. Ce schéma accentuel final devient dès lors un marqueur phonologique permettant d'identifier l'origine française de l'emprunt. Toutefois, si les chercheurs s'accordent volontiers pour qualifier l'accentuation des emprunts français en anglais de finale, aucune étude quantitative mêlant dimension diachronique, comparaison inter-dictionnaire et paramètre variétal n'a réellement été menée.

En 2011³⁵, je me rendis à l'Université de Craiova en Roumanie pour présenter une étude en collaboration avec trois doctorantes du laboratoire sur les emprunts français en anglais. Cette étude centrée sur les trois terminaisons contraignantes d'origine française *-eur*, *-euse* et *-aire* visait à observer les régularités accentuelles liées à ces terminaisons en anglais, mais également à évaluer leur potentiel de productivité au sein de la langue anglaise. Si cette première publication sur les emprunts français faisait la part belle aux raisons historiques justifiant la présence d'un tel stock lexical en anglais en détaillant l'impact de l'invasion de l'Angleterre par Guillaume le Conquérant ou celui de la Guerre de Cent Ans, l'intérêt de l'étude résidait surtout dans les conséquences phonologiques de tels faits historiques.

Nous constatâmes que les mots en *-eur*, *-euse* et *-aire* furent principalement empruntés au français entre la Renaissance et le début du 20^e siècle. Les données étymologiques de l'*OED* nous permirent de dater avec précision l'entrée de chaque mot dans la langue anglaise en utilisant le principe de datation présenté dans les sources, et également de déterminer le domaine sémantique dont relevait chacun de ces emprunts. Les principaux domaines sémantiques concernés sont les suivants : financier, religieux, politique, guerrier, administratif et artistique. J'allais désormais adopter une méthode de recueil de données de type lexicographique similaire à celle-ci pour chacune de mes études, dans un souci d'exhaustivité, d'exactitude et de transparence.

Le corpus de l'étude comprend 287 items, dont 168 en *-eur*, 42 en *-euse* et 87 en *-aire*. Au sein de ces données, certains items relèvent d'un processus d'emprunt direct au français tandis

³⁵ Descloux, E., Fournier, P., Martin, M. & S. Vanhoutte (2011), « Les mots français en *-eur* (*-euse*) / *-aire* en anglais contemporain : emprunt et création lexicale », *Actes du Colloque international - Les emprunts lexicaux au français dans les langues européennes*, M. Iliescu et alii (éds.), Craiova : Universitaria, 145-162.

que d'autres sont le fruit d'une création lexicale utilisant les matrices lexicogéniques de l'anglais. Ainsi, il y a 34 créations lexicales (17 en *-aire* et 17 en *-eur* et *-euse*) sur les 287 items que compte le corpus. La première partie de l'étude consista à mettre en évidence les processus lexicologiques à l'œuvre au sein de ces créations lexicales. Lorsqu'un item d'origine étrangère est emprunté, sa structure morphologique est opaque car les locuteurs anglophones sont dans l'incapacité d'identifier les frontières morphologiques et d'attribuer un statut sémantique aux affixes. Cependant, avec l'afflux d'emprunts possédant les mêmes terminaisons, la sensibilité morphologique des locuteurs augmente à tel point qu'ils sont en mesure d'attribuer des caractéristiques sémantiques aux bases et aux affixes qui peuvent alors être ré-utilisés pour générer du lexique. L'analyse des 17 créations lexicales en *-aire* démontra la faible productivité de ce suffixe d'origine française en anglais, dont l'utilisation revêt un caractère principalement humoristique comme le montrent les deux exemples suivants :

thousandaire = « qui a mille livres ».

pillionaire = « personne chevauchant le siège arrière d'une moto ».

L'étude des formations restantes (voir l'inventaire ci-dessous) nous permit d'isoler un sème commun en relation avec les notions de millionnaire et d'argent.

bazillionaire, bimillionaire, billionnaire, demi-millionnaire, gazillionaire, megamillionaire, multibillionaire, narco-billionnaire, oilionaire, quadrillion(n)aire, squillionaire, trimillionaire, trillionaire, zillionaire.

Nous interprêtâmes tout d'abord ces formations comme des suffixations en *-aire* mais le cas de *oilionaire* nous conduisit à réévaluer la structure morphologique de ces formations. Le terme *oilionaire* relève en effet de la catégorie des amalgames.

oil + *-ionaire* (troncation antérieure de *millionnaire*)

Par extension, les formations restantes pouvaient légitimement être considérées comme des amalgames. Cette hypothèse nous sembla la plus adaptée pour décrire le comportement morphosémantique de ces formations véhiculant la notion de millionnaire, car le suffixe *-aire* à lui seul ne le pouvait pas. Cette réévaluation du statut des formations en *-aire* en tant que structures amalgamées (à l'exception de *thousandaire* et *pillionaire*) fut d'autant plus nécessaire que les formations en *-eur* / *-euse* firent apparaître un comportement similaire.

L'analyse lexicologique montra en effet que 8 termes en *-eur* / *-euse* relevaient du processus de composition.

Exemples : *art connoisseur, delusions of grandeur, hunt saboteur, hunt saboteuse, physician-accoucheur.*

La composition germanique (N₂ N₁ : *art connoisseur*) est majoritaire dans l'inventaire, par rapport à la composition romane (N₁ préposition N₂ : *delusions of grandeur*), ce qui se justifie sans doute par le caractère principalement hybride des constructions qui combinent un élément d'origine latine avec un autre élément, d'origine romane. Le processus de composition étant majoritairement germanique en anglais, il est logique que de tels composés formés en anglais relèvent de ce type de composition. Trois cas de suffixation en *-eur* / *-euse* furent mis en évidence : *restauranteur*, *strippeur* et *strippeuse*. Ces créations en anglais restent tout de même sémantiquement transparentes pour les francophones (il s'agit des équivalents de *restaurateur* et de *stripteaseur* et *stripteaseuse* qui sont par ailleurs paradoxalement attestés en anglais) car elles n'ont pas subi de glissement sémantique majeur. Enfin, nous notâmes que, comme pour les mots en *-aire*, il existait quelques structures amalgamées qui relevaient du champ sémantique des nouvelles technologies (à l'exception de *shamateur*) : *infopreneur*, *intrapreneur*, *netpreneur*, *netrepreneur* et *technopreneur*. Le patron morphologique des cinq structures amalgamées relative à la notion d'entrepreneur se résume de la façon suivante avec une troncation antérieure de *entrepreneur* :

X- + (ent)(re)preneur

info-

intra-

net-

techno-

Je mesure à présent l'aspect innovant de cette recherche lexicologique et sémantique du point de vue de la tradition guierrienne, qui n'investissait que trop rarement ces sous-domaines linguistiques. L'originalité de cette étude réside davantage dans son traitement lexicologique des emprunts français que dans son traitement phonologique. Toutefois, sans être révolutionnaire par ses résultats, ce travail nous permet de tester des paramètres fréquentiels.

La problématique qui sous-tendait notre analyse phonologique était la suivante : les mots en *-aire*, *-eur* et *-euse* conservent-ils une accentuation finale héritée du français ou bien sont-ils sujets à un phénomène de rétraction accentuelle imposé par le système anglais ? Avec du recul, je constate à quel point la problématique initiale de mon étude était biaisée, car il était en réalité impossible de déterminer si la proportion importante d'accentuation finale était due à une reproduction de l'accent final du français ou bien à une influence des terminaisons contraignantes déterminantes en anglais. Ces terminaisons étant d'origine française mais également associées à un accent final en anglais au sein du système des terminaisons contraignantes, le problème était insoluble.

Le recours aux données fréquentielles extraites du *COCAE* montra que 51% des 297 termes avaient une fréquence nulle et que 41% avaient une fréquence comprise entre 1 et 100 sur 425 millions. Seuls 8% des items avaient une fréquence pouvant être qualifiée de significative au sein du lexique anglais. Concernant l'analyse phonologique menée à partir des données extraites de quatre dictionnaires (*OED*, *LPD*, *EPD* et *MCQ*), elle englobait 200 items (les monosyllabes et les termes sans transcription phonétique dans l'*OED* étaient écartés). 90% des items ont une accentuation finale en anglais britannique, ce qui était prévisible. Je ne mentionnerai ici que les mots présentant une accentuation différente et dont nous fîmes l'inventaire. Parmi les exceptions notables en *-eur* / *-euse* avec une fréquence significative en anglais, je citerai :

charmeuse /10/ (fréquence 45), *derailleur* /010/ (fréquence 634), *grandeur* /10/ (fréquence 1181), *chauffeur* /10/ ou /01/ (fréquence 684), *amateur* /100/ ou /201/ (fréquence 5221).

Concernant les mots en *-aire*, je citerai :

laissez-faire /100/ (fréquence 446), *luminaire* /100/ (fréquence 14), *solitaire* /100/ ou /201/ (fréquence 280).

Cette étude à la frontière entre lexicologie, sémantique et phonologie marqua le point de départ de mes recherches sur les emprunts français, qui se concentrèrent toutefois par la suite sur l'aspect accentuel. Mon étude suivante, publiée en 2013³⁶, marqua une nette évolution tant théorique que méthodologique par rapport à cette première étude de 2011. Je mis davantage en exergue le caractère hybride du système anglais, résultat de la fusion des logiques romane et germanique, comme moteur de la dynamique accentuelle des emprunts français. En effet, les travaux de D. Minkova (2009) ou J.-M. Fournier (2007) montrent qu'avant l'arrivée massive des emprunts français en anglais, la logique germanique consistant à accentuer la première syllabe d'une base était prédominante en anglais. L'arrivée des emprunts français a eu une conséquence importante sur le système phonologique de l'anglais, car si l'influence de la logique romane importée par les mots d'origine française ne fut pas immédiate (dans les premiers siècles suivant la conquête de Guillaume le Conquérant, l'accentuation des emprunts français était gouvernée par la logique germanique), petit à petit, cette logique romane consistant à accentuer un terme à partir de la droite en raison de l'influence décisive des terminaisons sur la position de l'accent, commença à remodeler le système phonologique de l'anglais. Les emprunts français sont donc une parfaite illustration du caractère hybride de

³⁶ Fournier, P. (2013), « L'accentuation des emprunts français en anglais britannique et américain », *Annales de l'Université de Craiova - Série langues et littératures romanes*, AN XVII, Nr.1, Craiova : Universitaria, 105-129.

l'anglais, qui se traduit par un équilibre permanent entre l'accentuation finale héritée de la logique romane et les rétractions accentuelles sur les syllabes pénultièmes ou antépénultièmes caractéristiques de la logique germanique.

Je décidai pour cette étude d'analyser le comportement accentuel des emprunts français dans un corpus plus large. J'utilisai le dictionnaire *EPD* pour faire une recherche par terminaison et je sélectionnai les terminaisons contraignantes d'origine française (-aire, -eur, -euse, -eux, -C'C'e³⁷, -que) auxquelles j'ajoutai des terminaisons typiquement françaises (-ier, -ais, -eau, -aux). Il me paraissait en effet hautement probable que les terminaisons dites contraignantes soient massivement associées à un schéma accentuel sur la syllabe finale, tandis qu'aucun ouvrage ne mentionnait à ma connaissance l'impact des terminaisons typiquement françaises telles que -ier ou -ais sur la position de l'accent. Après avoir croisé les résultats de ma première extraction avec les données étymologiques de l'*OED* pour ne conserver que les emprunts polysyllabiques directement issus du français, je disposais d'un corpus de 403 emprunts. Je notai ensuite les schémas accentuels en anglais britannique et américain avec un double objectif : observer le comportement accentuel de terminaisons typiquement françaises et déterminer le degré de correspondance entre l'anglais britannique et américain.

J'eus toutefois quelques problèmes méthodologiques au moment de comparer la position de la syllabe accentuée en anglais britannique et américain. Plutôt que de passer sous silence ces problèmes liés à la syllabation en anglais, je décidai d'y consacrer une partie de ma réflexion. Ma conception purement phonologique de la syllabe héritée de la logique de l'École de Guierre fut mise à mal par le traitement accentuel de ces emprunts, pour lesquels une conception phonétique de la syllabe semblait plus adaptée. En effet, dans un item tel que *millionaire*, il y a quatre syllabes phonologiques calculées à partir des voyelles graphiques. Toutefois la prononciation attestée dans *LPD* pose problème :

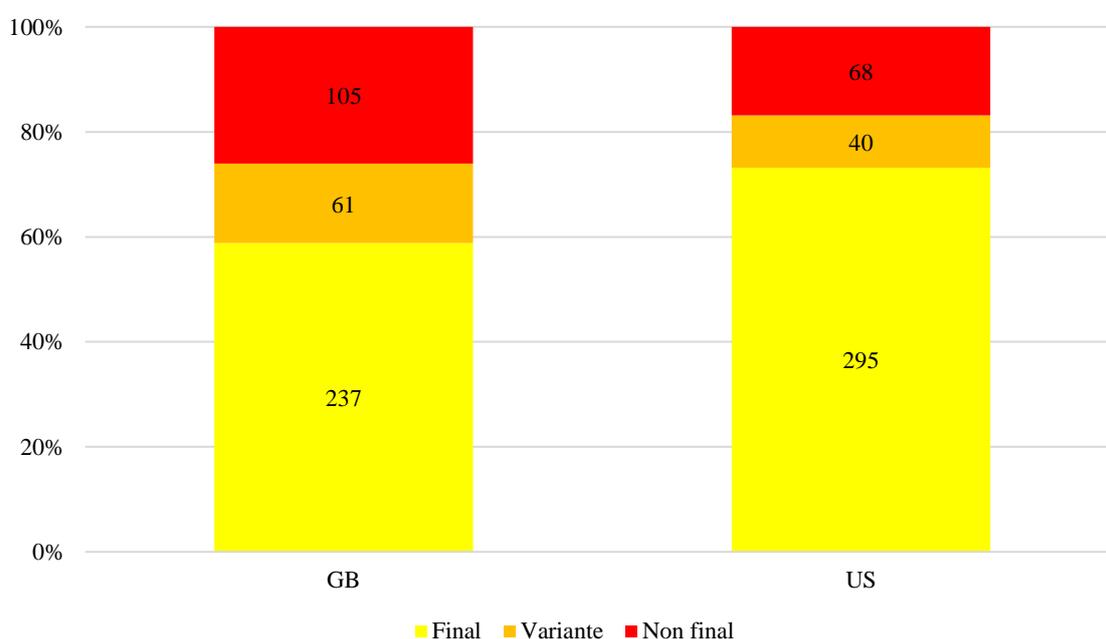
millionaire = /,mɪljə'neə, -iə- || -'neər, -'næər/

En effet, si la prononciation principale en anglais britannique et américain est trisyllabique, la variante attestée en anglais britannique démontre qu'une prononciation en quatre syllabes est également possible, même si la palatalisation de la voyelle <i> en /lj/ reste privilégiée. J'étais donc très embarrassé au moment de noter les schémas accentuels de deux variétés. De la même façon, le statut de la terminaison -ier était problématique car selon les entrées, il pouvait être monosyllabique (/jeɪ/) ou dissyllabique (/iɛɪ/). Par conséquent, conclure que *a'telier* était accentué en /010/ ou en /0100/ en anglais britannique, alors même que l'accent était placé sur

³⁷ Cette notation englobe toutes les terminaisons avec gémées consonantiques suivies d'un <e#> muet (exemples : -lle, -tte).

la même syllabe, n'avait pas vraiment de sens. J'étais face à mon premier problème méthodologique de jeune chercheur et je devais adopter une position claire pour garantir l'intégrité de mon étude. J'optai donc simplement pour un codage binaire du schéma accentuel en « final » et « non final », position justifiée par l'objectif même de mon étude, qui consistait à évaluer la proportion d'accent final au sein des emprunts français. Cette simplification méthodologique attribuable à ma faible expérience en tant que chercheur me conduisit petit à petit à réviser le statut de la syllabe et à la considérer comme une entité plus phonétique que phonologique, lorsque mes travaux s'inscrivirent dans les processus phonétiques et phonologiques à l'interface langue source / langue cible. Par conséquent, *atelier* (/ə'tel.i.ɛr , 'æt.el- , 'æt.ə.li- , -jeɪ/) relevait du codage « non final » en anglais britannique mais du codage « non final / final » (indiquant un phénomène de variation) en anglais américain (/ə'tɹl.jeɪ, æt'el , 'æt.el'jeɪ/). De la même façon, *Charpentier* est codé « non final » en anglais britannique (/ʃɑ:.'pɑ:n.ti.ɛɪ/) mais « final » en anglais américain (/ˌʃɑ:r.pɑ:n.ti'ɛɪ , -'tjeɪ/).

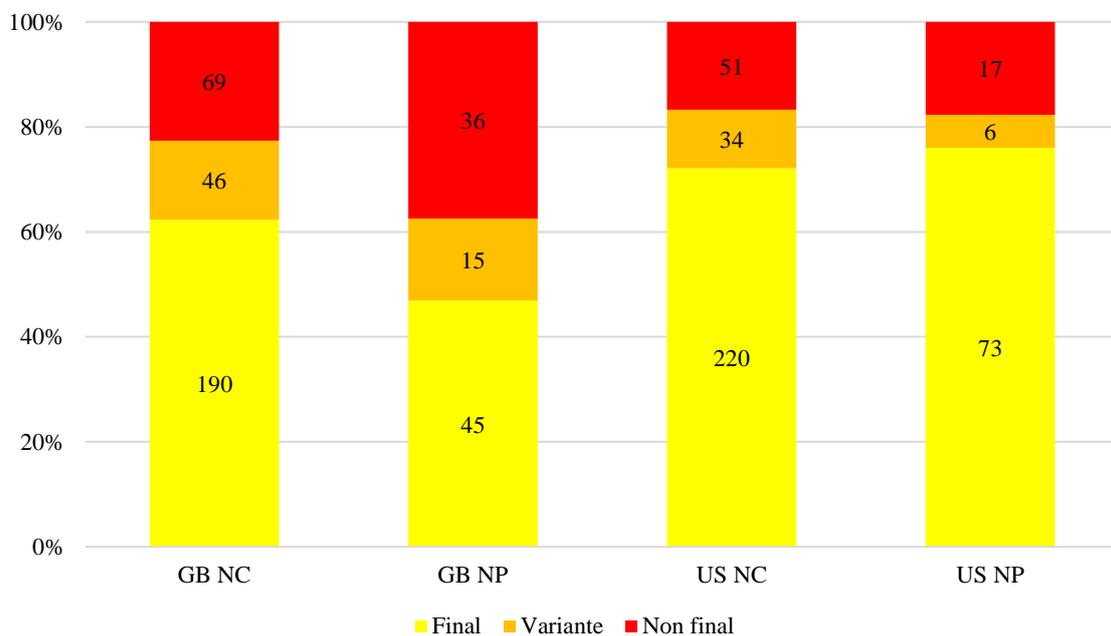
Les résultats concernant la répartition des schémas accentuels en anglais britannique et américain sont présentés dans le graphique 10.



Graphique 10 : Accentuation des emprunts français et anglais britannique et américain dans *EPD*.

Le codage « variante » correspond aux emprunts français possédant à la fois une accentuation finale et non finale au sein des données d'*EPD*. Je remarquai que le pourcentage

d'accentuation finale était nettement plus important en anglais américain qu'en anglais britannique et qu'inversement les accentuations non finales (sur les syllabes pénultièmes ou antépénultièmes) étaient plus fréquentes en anglais britannique. Toutefois, en parcourant les données, je m'aperçus qu'un autre facteur que je n'avais jamais pris en considération auparavant semblait avoir une influence décisive sur la logique accentuelle des emprunts français, à savoir la distinction nom commun / nom propre. Je codai ensuite mes données en utilisant NC pour nom commun et NP pour nom propre et procédai ensuite à la même analyse que dans le graphique 10. J'eus toutefois quelques problèmes de catégorisation de certains emprunts, à la frontière entre nom commun et nom propre. Je pense notamment à des cas tels que *sans-culotte* ou *deux-chevaux* (qui sont classés tous les deux dans les noms propres même si une telle simplification est évidemment discutable) ou encore à des expressions telles que *comme il faut* (qui sont assimilées à la catégorie des noms communs). Je me basai sur la classification proposée par le dictionnaire *Merriam Webster's* pour trancher. J'identifiai ainsi 305 noms communs et 96 noms propres. Les résultats prenant en compte la distinction nom commun / nom propre sont présentés dans le graphique 11.



Graphique 11 : Accentuation des emprunts français en anglais britannique et américain avec la distinction nom commun / nom propre dans *EPD*.

Les résultats montrent clairement que la distinction nom commun / nom propre n'est pas déterminante en anglais américain, car la proportion d'accentuation finale y reste sensiblement

équivalente. En revanche, l'introduction de cette variable semble avoir un effet significatif en anglais britannique, qui montre un pourcentage d'accentuation finale encore plus faible au sein des noms propres (moins de 50%).

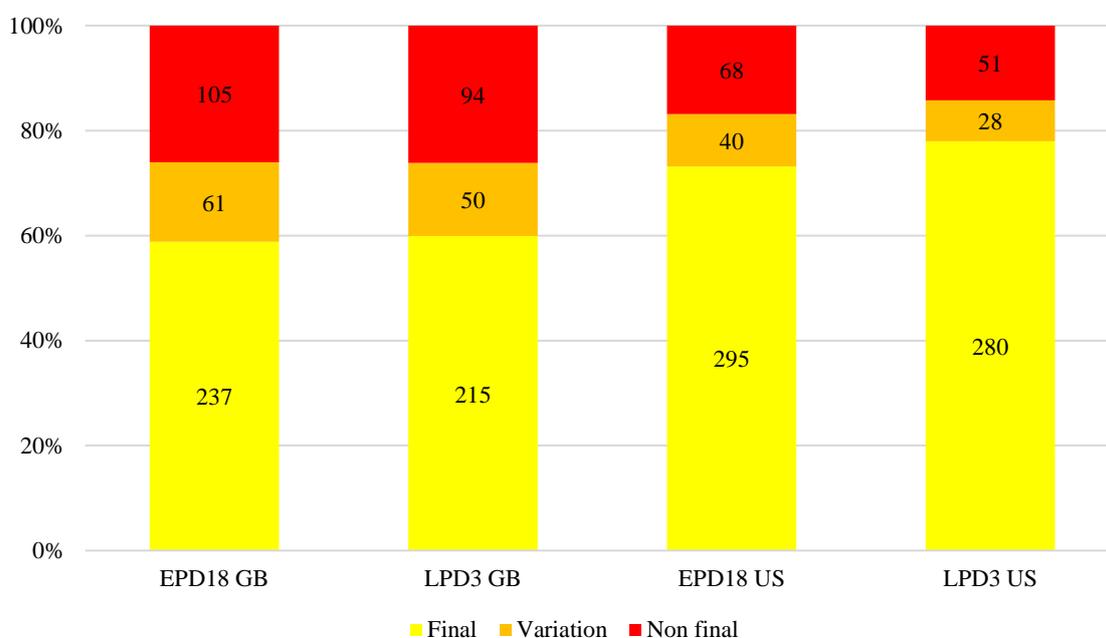
Cette étude de 2013 démontra ainsi que l'anglais américain reproduisait mieux l'accent final démarcatif hérité du français que l'anglais britannique, quelle que soit la catégorie de l'item, et que la distinction nom commun / nom propre avait un effet significatif sur l'accentuation des emprunts français en anglais britannique, avec un taux très faible de noms propres d'origine française accentués sur la finale. Quant aux raisons pour lesquelles l'anglais britannique avait moins recours à l'accentuation finale au sein des emprunts français, je fis dans la conclusion de cette étude une série d'hypothèses pour tenter de justifier le phénomène. J'avoue que j'étais à l'époque très surpris de cette différence accentuelle notable entre les noms communs et les noms propres en anglais britannique. Un nom propre est auto-référentiel et il me semblait dès lors incongru que les noms propres d'origine française ne soient pas accentués sur la syllabe finale en anglais britannique. Je repris donc l'hypothèse sociolinguistique formulée par J.-M. Chadelat (2000), qui prétendait que l'attitude consistant à s'éloigner de la prononciation d'origine pouvait être attribuée à la persistance d'une rivalité socioculturelle tenace entre la France et l'Angleterre. La prononciation britannique, longtemps influencée par le français pour des raisons historiques, marquerait ainsi son émancipation en adoptant plus volontiers le principe de rétraction accentuelle que l'anglais américain. Il était évidemment extrêmement difficile de tester la validité d'une telle hypothèse car mon étude en synchronie basée sur les données de la 18^e édition d'*EPD* m'empêchait de montrer une telle évolution au sein des dictionnaires. Mais cet argument m'incita à inclure une dimension diachronique dans mon étude suivante, en utilisant différentes éditions des mêmes dictionnaires de prononciation afin de travailler sur l'évolution de l'accentuation des emprunts français en anglais britannique et américain.

Le dernier résultat notable de mon étude fut l'identification de structures morphologiques déterminantes différentes en anglais britannique et américain, ce qui n'avait à ma connaissance jamais été démontré. Ainsi, en isolant les terminaisons <ais>, <eau>, <ier> et <é>, il apparaissait très clairement qu'elles étaient principalement associées à une accentuation pénultième en anglais britannique et finale en anglais américain.

Exemples : 'Calais, 'bandeau, fi'nancier, 'piqué en GB

Ca'lais, ban'deau, finan'cier, pi'qué en US

C'est sur la base de ces conclusions que je publiai en 2016³⁸ une nouvelle étude sur l'accentuation des emprunts français me permettant de tester certaines hypothèses formulées précédemment. Je repris le même corpus composé de 403 emprunts français extraits d'*EPD* que lors de l'étude précédente de 2013. Dans un premier temps, je procédai à une comparaison inter-dictionnaire entre la 18^e édition du dictionnaire *EPD* (*EPD18*) et la dernière en date de *LPD*, à savoir la troisième (*LPD3*). Cependant, même si les données au sein des deux dictionnaires sont souvent similaires, tous les emprunts d'*EPD18* ne figuraient pas dans *LPD3*. Je parvins ainsi à identifier 359 emprunts communs aux deux sources sur les 403 du corpus initial et je repris le système de codage utilisé dans l'étude précédente à savoir « final », « non final » et « variation » pour représenter la position de l'accent principal en anglais britannique et américain. Les résultats de cette première comparaison apparaissent au sein du graphique 12.

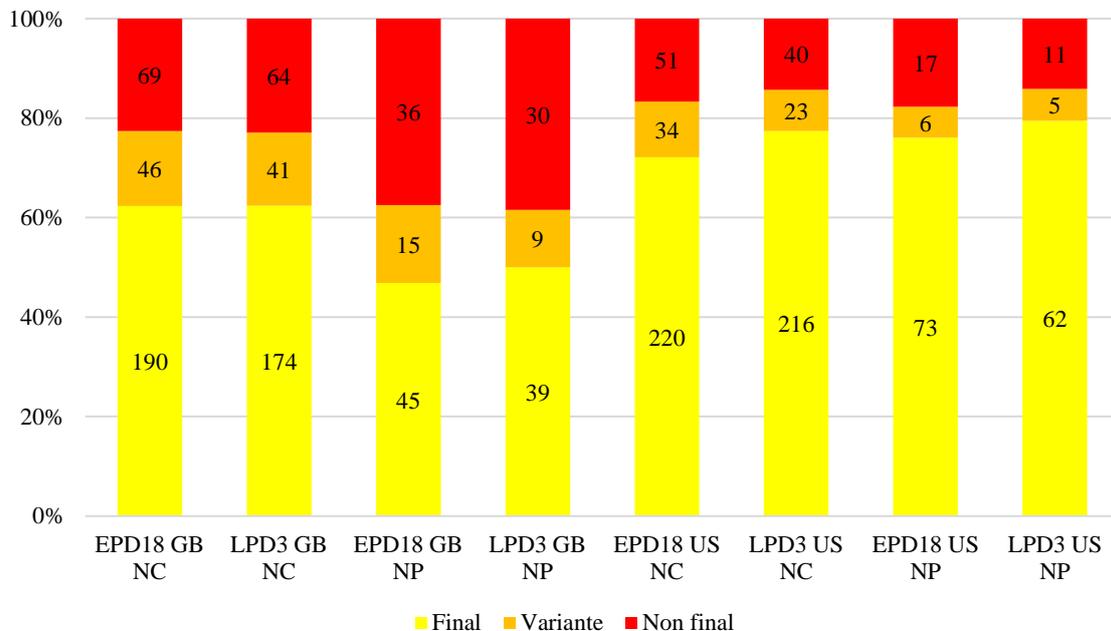


Graphique 12 : Accentuation des emprunts français en anglais britannique et américain dans *EPD18* et *LPD3*.

Les résultats du graphique 12 montrent que la proportion d'accentuation finale et non finale est sensiblement équivalente dans les deux sources dictionnaires, tant en anglais britannique qu'américain. Par conséquent, les déductions de la précédente étude faites à partir des données

³⁸ Fournier, P. (2016), « L'accentuation des emprunts français dans les dictionnaires anglais de prononciation », *Proceedings of the XVII EURALEX International Congress - Lexicography and Linguistic Diversity*, T. Margalidze & G. Meladze (éds.), Tbilisi: Ivane Javakhishvili Tbilisi State University, 424-435 (<http://euralex2016.tsu.ge/publication.html>).

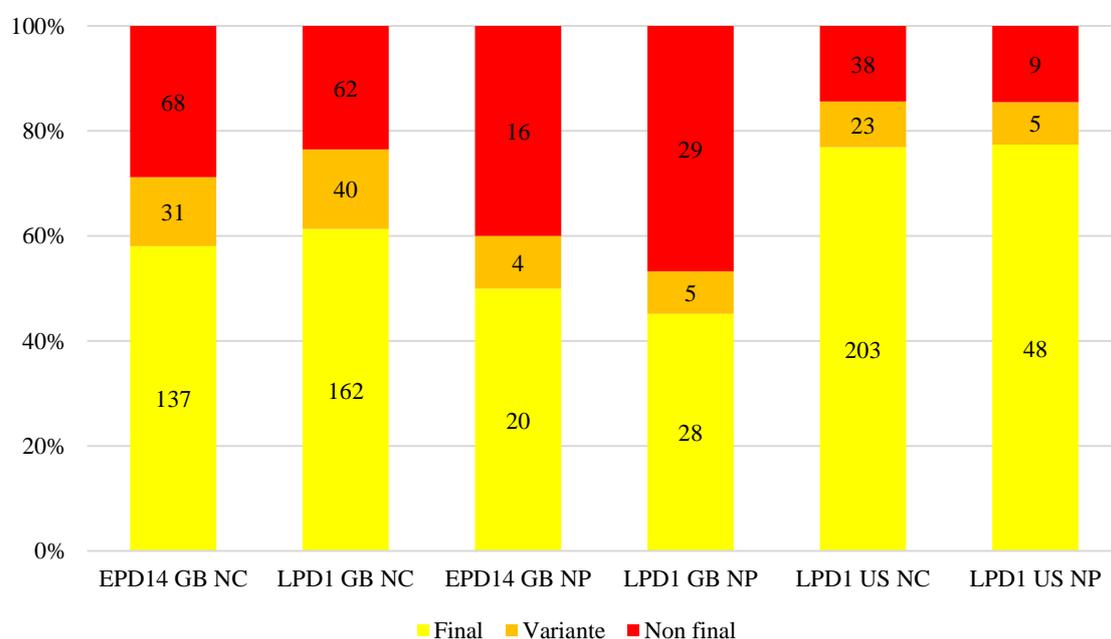
d'*EPD18* sont identiques à celles que l'on peut faire concernant les données de *LPD3*, à savoir que les emprunts français sont davantage accentués sur la syllabe finale en anglais américain qu'en anglais britannique et que la proportion de variation accentuelle est plus importante en anglais britannique, ce qui traduit une forme d'instabilité accentuelle, reflet de l'opposition de deux logiques : la logique de reproduction de l'accentuation finale, en vertu de l'héritage français de l'emprunt, et le principe de rétraction accentuelle, caractéristique de l'influence germanique au sein du système anglais. Ce dernier argument, que je ne formulai réellement qu'en 2016, s'appliquait tout autant aux résultats de ma précédente étude. J'ajoutai à nouveau la variable nom commun / nom propre à l'analyse afin de vérifier si la tendance dégagée à partir des données d'*EPD18* se confirmait avec l'analyse des données de *LPD3*. Les résultats de cette comparaison sont présentés dans le graphique 13.



Graphique 13 : Accentuation des emprunts français en anglais britannique et américain dans *EPD18* et *LPD3* avec la distinction nom commun / nom propre.

Les résultats du graphique 13 confirmèrent que la catégorie de l'emprunt n'exerçait pas d'influence significative en anglais américain, puisque la proportion d'accentuation finale était sensiblement équivalente entre les noms communs et les noms propres dans les deux dictionnaires. L'ajout des données de *LPD3* confirma que la catégorie de l'emprunt avait un impact sur la structure accentuelle des emprunts français en anglais britannique, avec un faible taux d'accentuation finale (50% environ) pour les noms propres.

La dernière partie de cette étude fut consacrée à une analyse diachronique de l'accentuation des emprunts français, effectuée à partir de versions antérieures des mêmes dictionnaires de prononciation. Cette analyse fut réalisée afin de tester l'hypothèse sociolinguistique émise par J.-M. Chadelat (2000) concernant la prononciation de l'anglais britannique. Ainsi, si l'anglais britannique contemporain tendait à se désolidariser de la prononciation française et notamment de l'accentuation finale, marqueur de l'origine française de l'emprunt, il me paraissait possible d'observer la trace d'une telle évolution en analysant les données accentuelles proposées par les versions antérieures des dictionnaires. Les données de *LPD3* (2008) furent ainsi comparées aux données de la toute première édition du dictionnaire (*LPD1* 1990) tandis que celles d'*EPD18* (2011) furent comparées à celles de la 14^e édition du dictionnaire (*EPD14* 1977). *EPD14* ne mentionne pas les transcriptions phonétiques de l'anglais américain qui apparaissent dans les éditions suivantes. Elles sont toutefois présentes dans *LPD1*. Nous fîmes donc les mêmes analyses que celles effectuées dans la première partie de l'étude sur les versions antérieures. *EPD14* et *LPD1* mentionnent respectivement 276 et 326 des emprunts du corpus. Les résultats de l'accentuation des emprunts français dans *EPD14* et *LPD1* sont consignés dans le graphique 14.



Graphique 14 : Accentuation des emprunts français en anglais britannique et américain dans *EPD14* et *LPD1* avec la distinction nom commun / nom propre.

L'accentuation des emprunts français dans *LPDI* présente les mêmes caractéristiques que dans *LPD3* et *EPD18*, avec un taux sensiblement équivalent d'accentuation finale en anglais américain, quelle que soit la catégorie de l'item (77% environ). Cela démontrait que l'accentuation des emprunts français n'avait pas évolué en anglais américain au cours des trois décennies précédentes. Concernant l'anglais britannique, l'introduction du paramètre nom commun / nom propre fit apparaître la même tendance dans *LPDI* et *EPD14*, avec un taux d'accentuation finale inférieur à 50% pour les noms propres et un taux d'accentuation finale se situant aux alentours des 60% pour les noms communs. Par conséquent, la dimension diachronique introduite par l'analyse des données d'*EPD14* et *LPDI* ne permit pas de mettre en évidence une quelconque évolution accentuelle des emprunts français en anglais britannique. L'hypothèse sociolinguistique émise par J.-M. Chadelat était donc mise à mal, au vu de cette étude diachronique menée sur une période certes relativement courte (de 1977 à 2011) mais sans aucun doute représentative d'une tendance plus globale sur la durée. Comment dès lors justifier le phénomène de rétraction accentuelle observé en anglais britannique ?

La présence de micro-paradigmes identifiés à partir de la morphologie (exemples : <-ais>, <-eau>) et possédant une influence distincte en anglais britannique et américain, hypothèse qui me paraissait pertinente en 2013, me sembla en 2016 d'une portée plus limitée, considérant les milliers de mots empruntés au français ne relevant pas de micro-paradigmes morphologiques. Je ne cessai de m'interroger sur mes trois études consacrées aux emprunts français et sur la pertinence des critères de constitution de mes différents corpus. Élaborer un corpus à partir de terminaisons typiques du français orientait nécessairement mon étude vers un traitement morphologique de la question. N'était-il pas plus pertinent de sélectionner les emprunts non pas à partir de paramètres segmentaux ou morphologiques, mais seulement sur la base de leur origine française en dépit de tout autre considération ? En effet, si je considérais l'origine française de l'item comme le critère essentiel régissant la logique globale du système, le phénomène était donc observable et testable pour tous les emprunts français et pas seulement pour une sous-partie soigneusement sélectionnée à partir de critères biaisés. De plus, la présence de terminaisons contraignantes systématiquement associées à un accent final comme paramètre de sélection des emprunts faussait probablement les résultats, en surévaluant la proportion d'accentuation finale au sein des emprunts d'origine française. Je me lançai donc dès 2016 dans la constitution d'un corpus massif d'emprunts français sélectionnés seulement à partir de leur origine.

Je m'efforçai de faire preuve d'objectivité au moment de constituer ce corpus, car je ne souhaitais pas inclure un paramètre qui biaiserait à nouveau mon étude. Je décidai d'extraire

manuellement tous les emprunts répertoriés comme français dans la troisième édition de *LPD*. Ce dictionnaire présente l'avantage d'indiquer la transcription phonétique française d'un emprunt dès lors qu'il est considéré comme tel, et je relevai donc tous les mots accompagnés d'une telle transcription. La distinction entre les différentes variétés de français ne figure pas dans la base de données et ma méthode présente bien évidemment des limites, car elle ne garantit pas l'exhaustivité des données ; mais elle permet toutefois de construire un corpus conséquent avec une intervention quasi-nulle du chercheur, ce qui garantit une certaine objectivité. Après le nettoyage des données issues de la première extraction, j'obtins un corpus de 1043 items, dont la seule langue source attestée dans *LPD3* était le français. Je réalisai un codage similaire à celui des précédentes études, car bien que le codage « final » / « non final » ait été au départ une façon d'éviter de me positionner sur le statut attribué à la syllabe au sein de mes études, je réalisais son bien-fondé, et je testai ensuite la proportion d'accentuation finale en anglais britannique et américain. Cette étude fut soumise à la revue *Journal of Linguistics*, dont les relectures intransigeantes, mais somme toute bienveillantes, m'encouragèrent à l'améliorer.

Les exigences d'une revue comme *Journal of Linguistics* sont telles qu'elles me poussèrent à adopter un positionnement toujours plus rigoureux tant le plan méthodologique que théorique. Je décidai de collaborer par la suite sur ce sujet avec Q. Dabouis³⁹ et je pus ainsi bénéficier de sa maîtrise de l'outil statistique pour tester l'influence des différents paramètres globalement reconnus comme déterminants dans l'assignation des accents par la littérature mondiale et de sa capacité à rendre accessibles les concepts des cadres théoriques majeurs les plus complexes. Les résultats de l'étude que je restitue ci-dessous sont donc le fruit d'un long cheminement intellectuel entre 2016 et 2023 et d'une collaboration fructueuse.

Notre postulat s'appuie sur les idées généralement véhiculées par les ouvrages majeurs de phonologie anglaise au niveau mondial, et notamment celle de M. Hammond (1999), qui considère que les mécanismes régissant la prononciation des emprunts ne sont pas différents de ceux du reste du lexique. Cependant, au sein des inventaires présentant les exceptions aux règles accentuelles majeures de l'anglais, on remarque que de nombreuses exceptions accentuées sur la syllabe finale sont en réalité empruntées au français (exemples : *machine*, *prestige*). Nous mîmes tout d'abord en évidence les spécificités sémantiques, morphologiques, phonologiques et segmentales des emprunts français qui les distinguent du reste du lexique en anglais. Il s'agit

³⁹ Dabouis, Q. & P. Fournier (2024), "Stress in French Loanwords in British and American English", *Journal of Linguistics*, 1-26.

ici de notre postulat concernant l'existence de sous-systèmes en anglais, dont je reparlerai dans le troisième chapitre de cette synthèse.

L'accentuation finale est assez peu attestée en anglais, si l'on excepte les emprunts français, et elle est généralement associée à certaines terminaisons contraignantes telles que *-self* ou *-teen* ou aux items préfixés autres que des substantifs (exemples : *again, forgive, distinct, detect*). La pression du phénomène de rétraction accentuelle, qui traduit l'héritage de la logique germanique, est donc importante au sein du système phonologique anglais. Dans cet article à paraître, nous avons évalué l'impact des paramètres potentiellement déterminants dans l'accentuation des emprunts français. Mes études précédentes avaient mis en évidence des différences accentuelles entre l'anglais britannique et américain au sein des emprunts français, raison pour laquelle nous décidâmes de tester l'efficacité des variables séparément sur l'anglais britannique et américain. Concernant les variables potentiellement déterminantes, nous choisîmes de tester l'efficacité des arguments généralement avancés par les chercheurs pour justifier l'accentuation des emprunts français, à savoir la structure segmentale, la distinction nom commun / nom propre et la morphologie. Nous ajoutâmes deux variables supplémentaires rarement citées, mais qui nous parurent potentiellement déterminantes : le nombre de syllabes et le degré de complexité de l'item (mot simple vs mot composé). Aucune autre étude quantitative de cette dimension n'ayant été entreprise précédemment, l'étude nous paraissait extrêmement ambitieuse et prometteuse.

Le codage des variables et des données fut effectué de la façon suivante. Si une terminaison est attestée dans au moins 10 items au sein du corpus, la variable <terminaison> est activée. Seize terminaisons furent ainsi identifiées au sein des données : *-age, -aire, -ant, -ard, -C'C'e, -é(e), -el, -et, -eur / -euse, -i, -ie, -ier, -in, -on, -que, -y*. Si certaines de ces terminaisons ont un effet sur l'accentuation clairement identifié dans la littérature, l'impact de certaines terminaisons sur l'accent n'a en revanche jamais été testé.

Les catégories syntaxiques NC / NP sont codées à partir de l'édition en ligne du *Merriam Webster's*.

Concernant la variable relative à la complexité de l'emprunt, les entrées sont considérées comme complexes dès lors qu'elles présentent un espace (exemple : *bon appetit*) ou un trait d'union (*passé-partout*). L'hypothèse selon laquelle les composés ou les structures syntaxiques complexes présentaient des spécificités accentuelles distinctes des mots dits simples ne nous semblait pas improbable.

Le poids syllabique de la syllabe finale fut également codé afin d'évaluer la corrélation entre accentuation finale et poids syllabique. Trois codages furent effectués à partir de la

prononciation principale proposée dans *LPD3*. Le premier codage est relatif à la structure de la syllabe (exemples : V, VC, VCC). Le deuxième codage binaire détermine la présence ou non d'une voyelle longue au sein de la syllabe finale, car les voyelles longues ont longtemps été perçues comme des éléments attractifs. Enfin, le troisième codage est basé sur le nombre de mores⁴⁰, calculé de la façon suivante :

- voyelles courtes et réduites = 1 more.
- voyelles longues = 2 mores.

Concernant les consonnes, deux représentations furent testées, l'une pour laquelle les consonnes en position coda correspondent à une seule more et une seconde selon laquelle les séquences de deux consonnes représentent deux mores, à l'exception de certaines séquences finales (exemple : /st/) dont le poids était équivalent à une seule more (selon M. Hammond 1999).

Enfin, le nombre de syllabes de chaque entrée fut déterminé à partir des syllabes phonétiques.

Les cinq variables furent incluses dans un modèle de régression logistique binaire, dont la seule variable dépendante était la position de l'accent (« final » ou « non final ») déterminée à partir de la prononciation principale de *LPD3*. Cet outil statistique permet de calculer précisément l'influence de chaque variable au sein du système lorsque les autres sont égales par ailleurs. Avant de donner les résultats du test statistique, nous observâmes que la distribution des schémas accentuels au sein du corpus, basée sur la prononciation principale, s'opérait de la façon suivante :

- 524/1043 = accent final en anglais britannique (50%).
- 813/1043 = accent final en anglais américain (78%).

Ces premiers résultats confirmaient les résultats de mes études précédentes, avec une proportion d'accentuation finale nettement plus importante en anglais américain qu'en anglais britannique.

Pour des raisons méthodologiques, l'analyse de régression logistique binaire fut effectuée sur 899 entrées. Les résultats de ce test remirent en question certaines de mes conclusions précédentes, car si toutes les variables testées montrèrent un effet déterminant sur la position de l'accent principal, tant en anglais britannique qu'américain, la seule pour laquelle aucune influence déterminante n'émergea fut celle de la catégorie syntaxique (distinction nom commun / nom propre). Ma déception fut de courte durée et laissa place à une forme de

⁴⁰ La more est une unité inférieure à la syllabe et qui permet un traitement plus approprié des phénomènes prosodiques.

satisfaction de voir mes réflexions sur l'accentuation des emprunts français avancer de la sorte. Je donne ci-dessous les résultats les plus significatifs de cette riche étude.

Concernant l'impact des terminaisons sur l'accent, toutes les terminaisons identifiées ne sont pas déterminantes. De plus, les terminaisons influentes ne sont pas strictement identiques en anglais britannique et américain et ne sont pas toutes corrélées à une accentuation finale. En anglais britannique, trois terminaisons sont associées à un schéma final :

-aire : affaire, laisser-faire, Voltaire.

-C'C'e : cayenne, cocotte, demoiselle.

-eur / -euse : agent provocateur, danseuse, entrepreneur.

Quatre terminaisons sont associées à un accent non final :

-é(e) : actualité, démodé, pâté.

-et : beignet, filet, Monet.

-in : au gratin, navarin, Rodin.

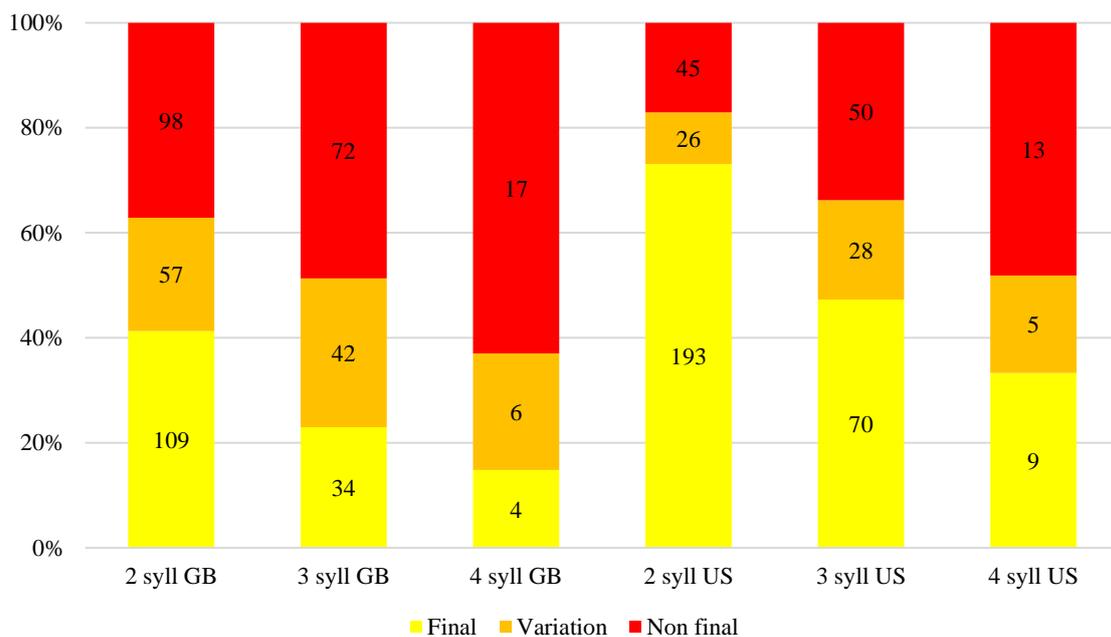
-on : Bourbon, cabochon, liaison.

En anglais américain, les terminaisons *-C'C'e* et *-eur / -euse* sont également associées à un accent final, de même que *-in* et *-on* le sont à un accent non final. Les terminaisons *-aire* et *-et* ne sont pas déterminantes en anglais américain, à l'inverse de l'anglais britannique. Enfin, la terminaison *-é(e)* a une influence distincte en anglais américain, car elle est massivement associée à un accent final (*actualité, Champs Elysées, crudités, passé, repoussé*).

Afin de travailler sur les mêmes données dans les deux variétés, tous les termes présentant l'une des terminaisons mentionnées ci-dessus furent écartés de l'analyse pour ne laisser que 623 items. Sur ces 623 emprunts, 58% ont un accent final en anglais britannique contre 77% en anglais américain. En regardant les résultats relatifs à la complexité des emprunts (simples vs complexes), nous notons une nette différence, car les emprunts dits complexes sont presque toujours associés à une accentuation finale dans les deux variétés, ce qui n'est pas le cas des emprunts dits simples. Les exemples suivants illustrent ce phénomène en anglais britannique et américain : *coup d'état, faux ami*. Il y a toutefois quelques termes présentant une accentuation non finale en anglais britannique et finale en anglais américain :

Arc de Triomphe, hors de combat, par excellence, prêt-à-porter.

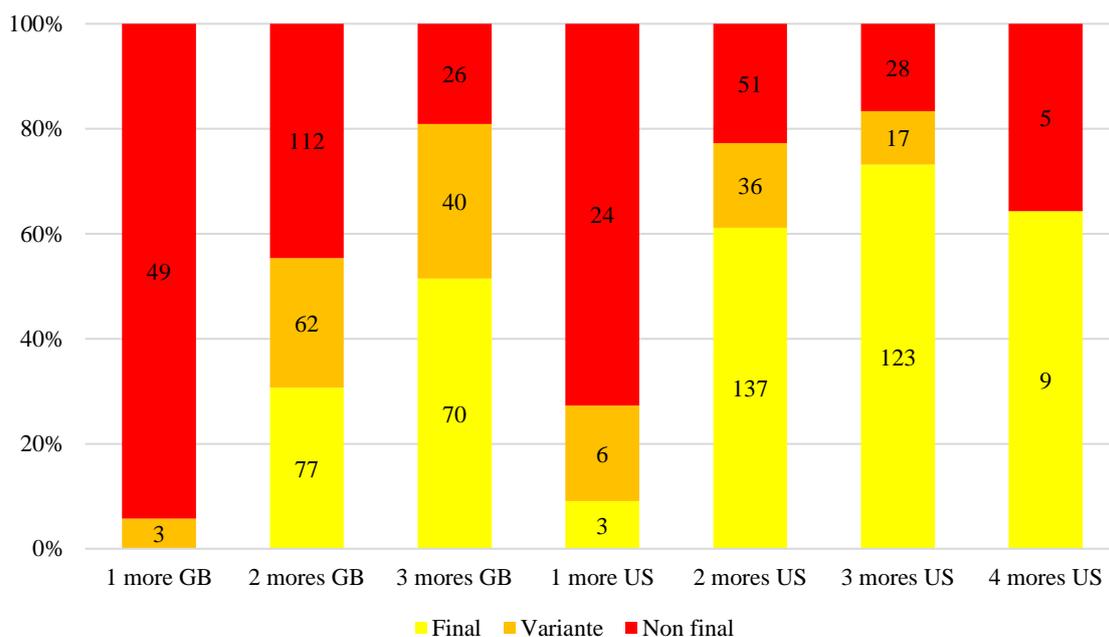
L'influence du nombre de syllabes sur la position de l'accent fut ensuite évaluée sur les mots simples uniquement, car les mots complexes étaient presque exclusivement associés à une accentuation finale. Je reproduis les résultats au sein du graphique 15 ci-dessous, sans lequel la tendance globale serait plus difficile à retranscrire.



Graphique 15 : Position de l'accent dans les emprunts français en anglais britannique et américain dans les mots simples et selon le nombre de syllabes.

Le graphique 15 montre la même tendance en anglais britannique et américain, à savoir que plus le nombre de syllabes augmente, plus la proportion d'accentuation finale diminue. Par conséquent, plus un mot simple est long, plus il aura tendance à adopter une accentuation non finale synonyme de rétraction accentuelle sur la syllabe pénultième ou antépénultième.

Les résultats concernant l'influence du poids syllabique de la syllabe finale et sa corrélation avec l'accentuation finale sont présentés dans le graphique 16.



Graphique 16 : Position de l'accent en anglais britannique et américain selon le poids syllabique (calculé en mores) de la syllabe finale.

Les résultats du graphique 16 montrent que plus la syllabe finale est lourde (plus elle contient de mores), plus l'accent tend à être sur la syllabe finale en anglais britannique et en anglais américain. Cependant, nous nous revendiquons, Q. Dabouis et moi-même, de l'École de Guierre, ce qui nous conduit ici à remettre en question l'influence du poids syllabique, en raison du problème de circularité qu'il soulève. En effet, en observant les syllabes finales composées d'une seule more en anglais britannique et américain, nous constatons la présence de nombreuses voyelles réduites (exemples : /ə/ dans *Broca*, *Casablanca*, *Nouméa*, *Zola*). Ces voyelles réduites sont-elles inaccentuées parce que, par définition, les voyelles réduites ne sont jamais accentuées ? Ou, inversement, sont-elles réduites parce qu'elles ne sont pas accentuées ? Je penche évidemment pour la seconde hypothèse, qui privilégie un phénomène accentuel basé sur des critères segmentaux ou morphologiques, les réalisations des voyelles n'étant que la conséquence du caractère accentué ou non de la syllabe.

Notre recherche s'achevait sur quelques analyses visant à rendre compte des accentuations non finales dans les emprunts d'au moins trois syllabes dans les deux variétés, soit 228 emprunts. Les mots complexes sont plus souvent accentués sur la syllabe pénultième que les mots simples en anglais britannique, mais cela tient en partie à la nature même du corpus de mots complexes massivement composés de constituants finaux dissyllabiques.

Exemples : *au gratin*, *aide-mémoire*, *bon voyage*, *folie de grandeur*.

Si la structure du second constituant diffère, le pourcentage d'accentuation sur la syllabe antépénultième augmente.

Exemples : *Cyrano de Bergerac, nom de plume, cul-de-sac, pain au chocolat.*

Les résultats sont identiques en anglais américain. Cette étude démontra donc que le système accentuel des mots empruntés au français, généralement considéré comme stable car associé à une accentuation finale dans les ouvrages majeurs de phonologie, n'était pas aussi régulier et que la proportion d'accentuation non finale était très importante en anglais britannique. L'analyse mit en évidence les paramètres déterminants dans l'assignation de l'accent principal, tout en discréditant l'hypothèse émise lors de mes précédentes études, selon laquelle la distinction nom commun / nom propre était déterminante en anglais britannique. Cette erreur n'était pas due à un problème méthodologique, mais à l'absence d'outils statistiques permettant de déterminer l'influence de ce critère par rapport aux autres paramètres déterminants au sein du système. L'analyse de paramètres tels que le nombre de syllabes ou la complexité de l'item représenta un développement majeur de mes recherches sur le sujet tout comme l'appui statistique. Le fruit de cette collaboration verra le jour en 2024 dans la revue *Journal of Linguistics*.

Même si les résultats relatifs aux emprunts italiens et français occupent une place prépondérante dans cette deuxième partie de ma synthèse, j'eus également l'occasion de mener quelques travaux plus ponctuels sur d'autres langues sources et dont je présente les principaux résultats dans la partie suivante.

3.3 Les emprunts aux autres langues

3.3.1 Les emprunts au japonais

À l'occasion d'une conférence organisée à Tokyo, je présentai⁴¹ avec une doctorante de l'Université de Tours les résultats d'une recherche consacrée à l'accentuation des emprunts japonais en anglais, résultats publiés en 2016 dans une perspective comparative du comportement accentuel des emprunts français, italiens et japonais⁴². Je choisis de travailler à

⁴¹ Fournier, P. & S. Vanhoutte (2013), "Stress in Japanese Loanwords in English: Faithfulness or Adaptation?", poster présenté lors de l'International Conference on Phonetics and Phonology (ICPP2013), National Institute for Japanese Language and Linguistics (NINJAL), Tachikawa (Tokyo) (25-27 janvier).

⁴² Fournier, P. (2016), « Nouvelles perspectives sur l'accentuation des emprunts en anglais contemporain », *Itinéraires*, 2015-2, *Stumbling blocks. Entraves et obstacles aux circulations. Approches pluridisciplinaires*, M. Lévêque & C. Parfait (éds.).

l'époque sur les emprunts au japonais en raison du nombre important de mots empruntés à cette langue en anglais. J'utilisai l'outil de recherche avancée de l'*OED* pour sélectionner tous les emprunts directs provenant du japonais, tout en écartant les monosyllabes, les traductions littérales, les emprunts complexes hybrides et les termes obsolètes. J'obtins un corpus de 464 emprunts dont je notai la position de l'accent principal en croisant les données de l'*OED* avec celles de *LPD* et *EPD*. J'identifiai 117 emprunts communs aux trois sources. La comparaison accentuelle entre les emprunts japonais en anglais et les items en japonais ne fut pas aisée, en raison des différences fondamentales entre les deux systèmes phonologiques. Le japonais est une langue possédant un accent mélodique (*pitch-accent language*) (T. Vance 2008) et la syllabe accentuée précède une chute au niveau mélodique (de *High* à *Low*). La présence d'une syllabe accentuée en japonais me permit ainsi d'effectuer un travail de comparaison. Cependant, la présence d'une chute mélodique n'est pas systématique en japonais, ce qui signifie que certains mots sont considérés comme inaccentués. J'utilisai le *Kenkyusha's New Japanese-English Dictionary* de M. Koh (1974)⁴³ proposant une écriture romanisée, et je notai la position de la syllabe accentuée en japonais. J'identifiai 326 mots japonais sur les 464 attestés dans l'*OED*⁴⁴. Voici quelques exemples de schémas accentuels proposés par le dictionnaire japonais (les syllabes inaccentuées sont notées en italiques) :

shu↑*ri-ken* LHH = pas de syllabe accentuée

a*lwabi* HLL = accentuation en /100/

ha↑*ra-kiri*↓ LHHH(L) = accentuation en /-1/

hi↑*ra-ga* *na* LHHL = accentuation en /-10/

i↑*ke* *l**bana* LHLL = accentuation en /0100/

Les syllabes prononcées sur un ton bas sont notées L (pour *low*) et celles réalisées sur un ton haut sont notées H (pour *high*). Je codai la position de la syllabe accentuée en utilisant le système de représentation numérique.

Mes premières investigations sur le comportement accentuel des emprunts japonais ayant démontré que le taux de variation entre l'anglais britannique et américain était pratiquement nul (comme d'ailleurs pour les emprunts italiens), je décidai de noter l'accent en anglais britannique uniquement. Les résultats concernant l'accentuation des emprunts japonais dissyllabiques en anglais britannique sont assez révélateurs :

- 92% sont accentués en /10/ (exemples : *bonsai*, *haiku*, *kana*, *manga*).

⁴³ Sur les conseils de Florence Labrune, spécialiste de la phonologie du japonais, que je remercie de nouveau.

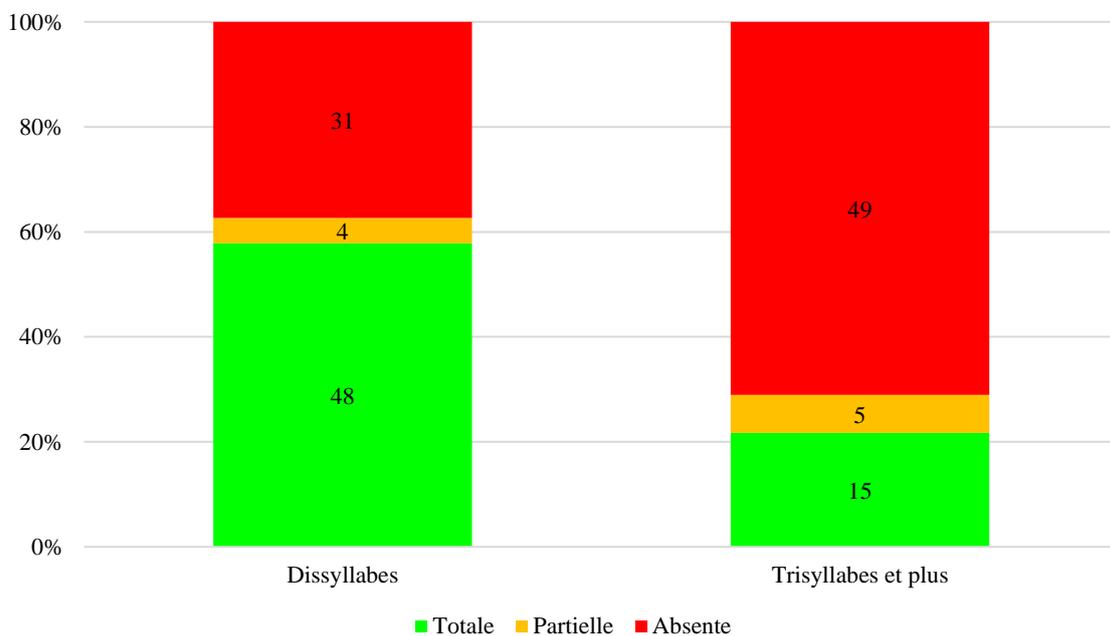
⁴⁴ Mon manque de maîtrise du japonais et des variantes orthographiques explique sans doute que je passai ainsi à côté de nombreux items.

- 4% sont accentués en /01/ (*gobang, sensei, tycoon*).
- 4% montrent une variation accentuelle (*banzai, gaijin, Nikkei, nisei*).

Les résultats concernant l'accentuation des emprunts japonais de trois syllabes et plus confirment la tendance observée dans les dissyllabes :

- 65% sont accentués en /-10/ : *adzuki, hara-kiri, kakemono*.
- 18% sont accentués en /(-)100/ : *anime, netsuke, samurai*.
- 2% sont accentués en /-1/ : *jigotai, shuriken*.
- 1% est accentué en /(-)1000/ : *Kamakura, sewamono*.
- 9% de variation accentuelle (principalement /(-)10/-(-)100/) : *aikido, sashimi, Sudoku*.

Les pourcentages restants correspondent à des erreurs de transcription imputables à l'*OED*⁴⁵. Même si la tendance à l'accentuation pénultième est moins nette dans les mots d'au moins trois syllabes que dans les dissyllabes, elle se démarque nettement des autres schémas accentuels. Je cherchai ensuite à évaluer le taux de correspondance entre la syllabe accentuée dans les mots japonais et la syllabe accentuée dans les emprunts correspondants en anglais. Les résultats de cette analyse sont présentés dans le graphique 17.



Graphique 17 : Étude de la correspondance accentuelle entre les mots japonais et les emprunts correspondants en anglais.

⁴⁵ Il s'agit principalement de termes avec deux accents primaires.

Si les résultats concernant la correspondance accentuelle au sein des dissyllabes pouvaient laisser planer le doute sur une potentielle correspondance avec un taux d'environ 60%, les résultats concernant les mots d'au moins trois syllabes le balayèrent. Le pourcentage de correspondance accentuelle est si faible au sein des mots de trois syllabes et plus (il l'eût d'ailleurs été encore plus si j'avais inclus les mots japonais sans accent dans la catégorie « absence de correspondance ») que je pouvais légitimement conclure que la dynamique accentuelle régissant l'accentuation lexicale des emprunts japonais en anglais n'était pas liée à une quelconque reproduction du schéma accentuel d'origine. Je penchai à l'époque, mais je reviendrai sur ce point dans ma présentation d'une formalisation du système, pour l'hypothèse du schéma accentuel préférentiel appliqué par défaut aux emprunts en anglais, à savoir l'accentuation pénultième qui est majoritaire au sein des données.

3.3.2 Les emprunts au gaélique écossais

J'eus également l'occasion de travailler sur les emprunts au gaélique écossais à l'occasion du Congrès International de la SFEE (Société Française d'Études Écossaises) organisé à l'Université Paris 13, les 16 et 18 novembre 2017, et dont l'invitée d'honneur était J. Stuart-Smith, professeur de phonétique à l'Université de Glasgow. Je ne manquai pas cette opportunité de présenter mes travaux qui s'inscrivent dans la lignée directe de mes autres recherches sur les emprunts, à savoir une partie lexicographique orientée vers le recueil de données et une exploitation phonologique du corpus. Cette recherche se distingua toutefois des autres, car elle me permit également de faire quelques excursions dans le domaine de la graphophonématique, et notamment un travail sur /x/, dont la réalisation en gaélique écossais et sa reproduction dans les dictionnaires anglais de prononciation reposent sur des considérations sociolinguistiques, comme me le fit justement remarquer J. Stuart-Smith lors des échanges suivant la présentation. Cette étude donna lieu à une publication⁴⁶ regroupant l'essentiel des communications présentées lors de ce congrès et s'inscrivant dans la thématique de la « différence ».

La partie lexicographique de l'étude débuta par une extraction de données effectuée à partir de la base de données étymologiques de l'*OED*. Je pris soin de distinguer le gaélique écossais (*Scottish Gaelic*) du gaélique irlandais (*Irish*) dans le moteur de recherche avancée de la base

⁴⁶ Fournier, P. (2021), « Les emprunts au gaélique écossais en anglais », *L'Ecosse, la différence - Scotland, the Difference*, S. Juillet Garzon, P. Fournier & A. Fiasson (éds.), Série Caledonia : Presses Universitaires de Franche-Comté, 215-229.

de données. Cette première extraction me donna une liste de 56 items, qui me permit d'obtenir, après nettoyage des données, 41 emprunts exploitables pour mon analyse. Afin d'évaluer la date d'entrée de ces items dans les dictionnaires anglais, je me servis de la date de premier référencement indiquée dans les sources bibliographiques de l'*OED*. Toutefois, devant les nombreuses interrogations suscitées par certaines sources, je tombai presque par hasard sur un article de P. Gilliver (2015) détaillant le fonctionnement des sources de l'*OED*. Les lecteurs sont employés par le dictionnaire pour parcourir des manuscrits et dénicher les premières occurrences d'un terme en anglais, mais le résultat dépend en partie de la fiabilité du travail effectué par chaque relecteur et de sa capacité à associer les variantes orthographiques à la forme contemporaine. Quelle ne fut pas ma surprise d'apprendre que la référence mondiale en termes de ressources dictionnairiques pour la langue anglaise pouvait donc faire preuve d'approximation dans sa collecte de données ! Cela m'incita ainsi à relativiser l'exactitude de mes propres recherches et à accepter le fait que certaines hypothèses visant à rendre compte d'un fait linguistique ne pouvaient en aucun cas constituer des vérités absolues sur le fonctionnement de la langue.

Les résultats relatifs à la date de première attestation des emprunts au gaélique écossais au sein de la langue anglaise montrent que la distribution des emprunts s'opère du 14^e au 20^e siècle avec un pic au 19^e siècle. À l'inverse de mes précédentes études, pour lesquelles j'avais eu recours aux données fréquentielles proposées par le *COCAE*, j'utilisai ensuite l'outil de fréquence de l'*OED*, qui détermine la fréquence d'un item sur une échelle allant de 1 pour les mots rares à 8 pour les plus attestés. Les résultats montrent que 27 emprunts au gaélique écossais sont classés aux niveaux 2 et 3, contre seulement 8 aux niveaux 4 et 5, qui constituent les items empruntés à cette langue les plus attestés en anglais (exemples : *Gael, loch, oy, plaid, ptarmigan, mod*). Les trois catégories sémantiques couvertes par ces emprunts sont la culture et les traditions locales, la géographie et la zoologie. Le dernier champ d'investigation de la partie lexicographique concernait l'étude des correspondances graphiques entre les mots en gaélique écossais et les emprunts correspondants en anglais. De façon assez surprenante, certains emprunts ne montrent aucune différence orthographique par rapport au terme existant en gaélique écossais, même parfois après plusieurs siècles de présence au sein du lexique anglais. Les exemples suivants illustrent cette correspondance graphique parfaite :

bothan, cranachan, loch, lochan, machair, mod, och, partan, piob mhor, port, port-a-beul, sporan et strath.

À l'inverse, certains emprunts ont subi de profondes modifications orthographiques, provoquées vraisemblablement par une adaptation aux règles en vigueur au sein du système

orthographique anglais. Les deux exemples suivants sont particulièrement représentatifs de ces modifications qui me laissaient dubitatif quant aux résultats à venir concernant les éventuelles transmissions phonologiques du gaélique écossais vers l'anglais :

piobaireachd → *pibroch*

tàrmachan → *ptarmigan*

L'étude phonologique visait avant tout à rendre compte de la prononciation des emprunts au gaélique écossais en anglais, ainsi qu'à étudier les relations graphophonématiques pertinentes potentiellement héritées de la langue source. Bien qu'il soit à tort souvent considéré comme approximatif, concernant les transcriptions phonétiques proposées, l'*OED* propose des transcriptions phonétiques vocaliques pertinentes et parfois dans quatre variétés : l'anglais britannique, l'anglais américain, l'anglais écossais et l'anglais irlandais. Les deux dernières variétés n'étaient toutefois pas systématiquement attestées dans la base de données, ce qui m'empêcha de proposer une réelle analyse de ces deux variétés en question.

L'analyse des 38 transcriptions phonétiques de l'*OED* fit apparaître quelques résultats notables. Au niveau de l'accent lexical, la majorité des emprunts au minimum dissyllabiques sont accentués sur la syllabe pénultième, avec toutefois quelques accentuations finales :

exemples : *mormaor*, *ochone* et *piob mhor*,

et antépénultièmes :

exemples : *cranachan* et *ptarmigan*,

et ce dans toutes les variétés proposées. Le terme *ochanee* constitue le seul cas de variation intra- et inter-dialectale avec une accentuation finale en anglais britannique et américain, une accentuation pénultième en anglais écossais et un phénomène de variation accentuelle /-1/ - /(-)100/ en anglais irlandais.

L'analyse des règles graphophonématiques démontra quelques exceptions par rapport aux règles usuelles (A. Deschamps 1994, J.-M. Fournier 2010). Toutefois, les cas particuliers que je mis en évidence lors cette étude correspondaient aux variétés d'anglais écossais et irlandais, qui n'ont à ma connaissance jamais été incluses au sein du fonctionnement systémique des règles graphophonématiques des ouvrages majeurs. Voici quelques cas particuliers relevés :

mink = /ɪ/ ou /ʌ/ (SC).

mod = /ɔ:/ ou /ɒ/ (GB).

piob mhor = /i:/ avec <o> muet (GB / US).

port-a-beul = réalisation dissyllabique de <eu> (GB / US / SC).

La proportion très élevée de voyelles non réduites en contexte inaccentué en anglais britannique notamment m'interpella.

Exemples : *mormaor* (/mɔ:'mɛ:/), *ochone* (/ə'xəʊn/, /ə'həʊn/, /ɒ'xəʊn/), *piob mhor* (/pi:b 'mɔ:/), *fulmar* (/ 'fʊlmə/, / 'fʊlmɑ:/), *pibroch* (/ 'pi:brɒx/, / 'pi:brɒk/).

Je faisais par la suite le lien entre ce phénomène et le fonctionnement en contexte inaccentué de tous les emprunts en anglais qui ont une proportion plus élevée de voyelles pleines que dans le reste du lexique, où traditionnellement, il y a davantage de voyelles réduites en contexte inaccentué.

Au niveau des règles graphie-phonie des consonnes, la prononciation du graphème <m> de *piob mhor* réalisé /v/ ou /m/ en variante tant en anglais britannique qu'en anglais américain constitue un cas notable. Je ne savais pas à l'époque quel statut accorder à la prononciation /v/ même si, de par mes études passées, j'avais conscience de l'influence potentielle des langues sources non seulement au niveau accentuel mais également au niveau des correspondances graphophonématiques. La réalisation /v/ du graphème <m> était en réalité un héritage direct d'une correspondance graphie-phonie attestée en gaélique écossais, comme cela fut confirmé par plusieurs locuteurs de cette langue présents à la conférence. Cela semblait donc également corroborer l'hypothèse que seuls les locuteurs bilingues ou extrêmement cultivés sur un plan linguistique (mais on imagine aisément que les lexicographes à l'origine de la création des dictionnaires de prononciation majeurs en anglais le sont) soient à l'origine de l'introduction de tels items au sein des dictionnaires. Il est autrement impossible de justifier cette réalisation en /v/ du graphème <m> au sein des données de l'*OED*.

Je consacrai la dernière partie de cette étude à la réalisation phonétique du digraphe consonantique <ch>, attesté à plusieurs reprises dans les données. Le tableau 3 ci-dessous est extrait de ma publication et répertorie toutes les prononciations de cette séquence dans l'*OED* au sein des différentes variétés.

Emprunt	GB	US	SC	IR
<i>quaich</i>	/x/-/k/	/k/-/x/	/x/	////////////////
<i>loch</i>	/x/-/k/	/k/-/x/	/x/	////////////////
<i>och</i>	/x/-/k/	/k/-/x/	/x/	/x/-/k/
<i>cranachan</i>	/k/	/k/	/k/	////////////////
<i>ochanee</i>	/x/	/k/	/x/	/x/
<i>ochone</i>	/x/-/h/	/h/-/x/	/x/	/x/
<i>lochan</i>	/x/-/k/	/k/-/x/	/x/	////////////////
<i>machair</i>	/x/	/x/	/x/	////////////////
<i>pibroch</i>	/x/-/k/	/x/-/k/	////////////////	////////////////

Tableau 3 : Réalisation phonétique du digraphe consonantique <ch> dans l’*OED*.

Les données du tableau 3 montrent que le digraphe <ch> est réalisé /k/ ou /x/ dans les emprunts au gaélique écossais (la variante en /h/ pour *ochonee* en anglais américain constitue un cas marginal). L’anglais écossais opte systématiquement pour /x/ (à l’exception de *cranachan*) tout comme l’anglais irlandais sur la base des quelques cas retranscrits. En revanche, le phénomène de variation est très important en anglais britannique et américain avec /x/ globalement utilisé en variante principale en anglais britannique et en variante secondaire en anglais américain. Ces résultats me paraissaient à première vue surprenants et je croisai les données de l’*OED* avec celles de *LPD* et *EPD*. Quatre emprunts étaient attestés dans au moins l’un des deux dictionnaires : *loch*, *och*, *ochonee*, *pibroch* (voir tableau 4).

Emprunt	LPD3 GB	LPD3 US	EPD18 GB	EPD18 US
<i>loch</i>	/x/-/k/	/k/-/x/	/k/ - as if Scots /x/	/k/ - as if Scots /x/
<i>och</i>	/x/	/k/	/x/	/k/ - as if Scots /x/
<i>ochone</i>	/x/	/x/	////////////////	////////////////
<i>pibroch</i>	/k/-/x/-/f/	/k/	/k/-/x/	/k/

Tableau 4 : Réalisation phonétique du digraphe consonantique <ch> dans *LPD* et *EPD*.

Un commentaire inséré dans les entrées de *loch* et *och* dans *EPD* ne manqua pas de me surprendre : « *as if Scots - /x/* ». L'entrée de *loch* est également complétée par la citation suivante (2011 : 93) :

The ability to pronounce /x/ in 'loch' is often considered to be a test of Scottishness by Scottish speakers.

Je remarquai toutefois que la distribution de /x/ était assez inégale au sein des données du dictionnaire, car si /x/ est mentionné en variante possible de /k/ en anglais américain dans *loch* et *och*, la seule prononciation attestée en anglais britannique dans *och* est /x/, l'absence de la variante /k/ semblant étrange. De même, si /x/ figure en variante de /k/ en anglais britannique dans *pibroch*, /k/ est la seule réalisation attestée en anglais américain. Plusieurs choses me paraissaient étranges dans les données proposées par les dictionnaires. Tout d'abord, les locuteurs américains avaient-ils réellement recours à la réalisation /x/ et étaient-ils en capacité de reproduire /x/, considéré comme un critère de *Scottishness* par *EPD* ? Si cette prononciation semblait attestée en anglais britannique, j'avais un doute concernant la légitimité d'une telle réalisation en anglais américain. De plus, la distribution de /x/ était inégale dans les deux dictionnaires de prononciation, parfois absente, souvent en variante de /k/ mais parfois référencée en tant que seule réalisation possible. J. Stuart-Smith, invitée d'honneur de la conférence, m'apporta quelques éléments de réponse, d'autant plus qu'elle avait auparavant travaillé sur la réalisation de /x/ à Glasgow. Elle me confirma que les données dictionnaires relatives à /x/ dans les dictionnaires de référence ne reflétaient plus une réalité en perpétuelle évolution et ce en raison de facteurs principalement historiques et sociolinguistiques. En d'autres termes, /x/ n'était déjà pratiquement plus utilisé à Glasgow (E. Lawson & J. Stuart-Smith 1999), au profit d'une réalisation en /k/. Quant à imaginer une telle prononciation en anglais britannique standard et en anglais américain standard, elle me dit que les dictionnaires de prononciation véhiculaient une forme de caricature de la *Scottishness* fondée sur une réalité dorénavant partiellement disparue.

3.3.3 Les emprunts à l'arabe

La dernière langue source à laquelle je consacrai une étude et sur laquelle je travaille actuellement en vue d'une seconde publication est l'arabe. La publication de 2020⁴⁷ réalisée en

⁴⁷ Fournier, P. & R. Latrache (2020), "Arabic Loanwords in English: a Lexicographical Approach", *Proceedings of XIX EURALEX Congress: Lexicography for Inclusion*, Vol. I, Z. Gavriilidou, M. Mitsiaki & A. Fliatouras (éds.), Democritus University of Thrace, 343-349.

collaboration avec ma collègue R. Latrache, historienne arabophone spécialiste des États-Unis, a un ancrage purement lexicographique et cherche à inventorier les emprunts arabes attestés dans l'*OED* et à retracer leur origine grâce au travail de datation et de collecte de données effectué par la version en ligne du dictionnaire. L'arabe est perçu comme la langue orientale ayant fourni le plus d'emprunts à l'anglais, même si nous nous sommes aperçus au moment de l'extraction des données que de nombreux emprunts d'origine arabe par la langue anglaise avaient été opérés par l'intermédiaire de l'espagnol. Par conséquent, si le nombre d'emprunts d'origine arabe en anglais est extrêmement important – H. Salloum & J. Peters (1996) estime ce nombre aux alentours de 6500 – la proportion d'emprunts directement empruntés par l'anglais à la langue arabe est en réalité relativement réduite. Le but de cette étude consistait à établir un corpus d'emprunts directs, dans la perspective d'une seconde étude traitant de l'accentuation des emprunts arabes en anglais contemporain. La collaboration avec une arabophone fut extrêmement précieuse au moment d'étudier l'origine des items extraits de la base de données. Ainsi, des items tels que *Islam* (dont l'origine est considérée comme turque dans l'*OED* car cet item a intégré l'anglais par l'intermédiaire du turc) ne furent pas conservés dans le corpus final puisqu'ils ne correspondaient pas à la notion d'emprunt direct. Le corpus sur lequel notre étude de 2020 est basée se compose de 409 emprunts directs à l'arabe.

L'afflux des emprunts arabes en anglais débute aux alentours du 16^e siècle avec un pic au 19^e siècle. Cette analyse démontre que le domaine sémantique le plus représenté au sein des emprunts arabes en anglais est celui de la religion, avec ensuite, par ordre décroissant, les domaines de la science, de la vie quotidienne (vêtements, nourritures et boissons), du commerce et des arts. Lors du 19^e siècle, il y eut un apport massif d'emprunts relatifs aux domaines religieux et scientifiques et nous cherchâmes à en comprendre les raisons à partir d'un travail sur les sources. Cela souleva les mêmes problématiques que celles abordées lors de mon étude sur les emprunts au gaélique écossais et la question de la pertinence et de la représentativité des sources de l'*OED* soulevées par P. Gilliver (2015). Le système même du travail de datation des items et de collecte des sources est biaisé et remet en perspective toute étude lexicographique qui se prétendrait exacte et exhaustive. En effet, en travaillant uniquement sur les deux premières apparitions de chaque emprunt au sein des données, il apparaît que 50 emprunts attestés pour la première fois au 19^e siècle en anglais sont en réalité extraits des deux ouvrages des célèbres orientalistes E.W. Lane et R.F. Burton. À la suite de leurs traductions du célèbre ouvrage *Arabian Nights* (ou *A Thousand and One Nights*), on constate un afflux plus important d'emprunts arabes dans la langue anglaise. Ainsi, huit emprunts proviennent directement de la traduction proposée par E.W. Lane, tandis que deux sont empruntés à celle de R.F. Burton. Le

paradoxe pourrait s'arrêter là, mais nous avons constaté que 40 emprunts arabes étaient également apparus pour la première fois en anglais à partir de travaux des deux mêmes auteurs : *An Account of the Manners and Customs of the Modern Egyptians* (E.W. Lane 1836) et *Personal Narrative of a Pilgrimage to El-Medinah and Meccah* (R.F. Burton 1855). Il semblerait donc que le travail de datation effectué par les relecteurs du dictionnaire soit biaisé dès la sélection même des ouvrages à parcourir.

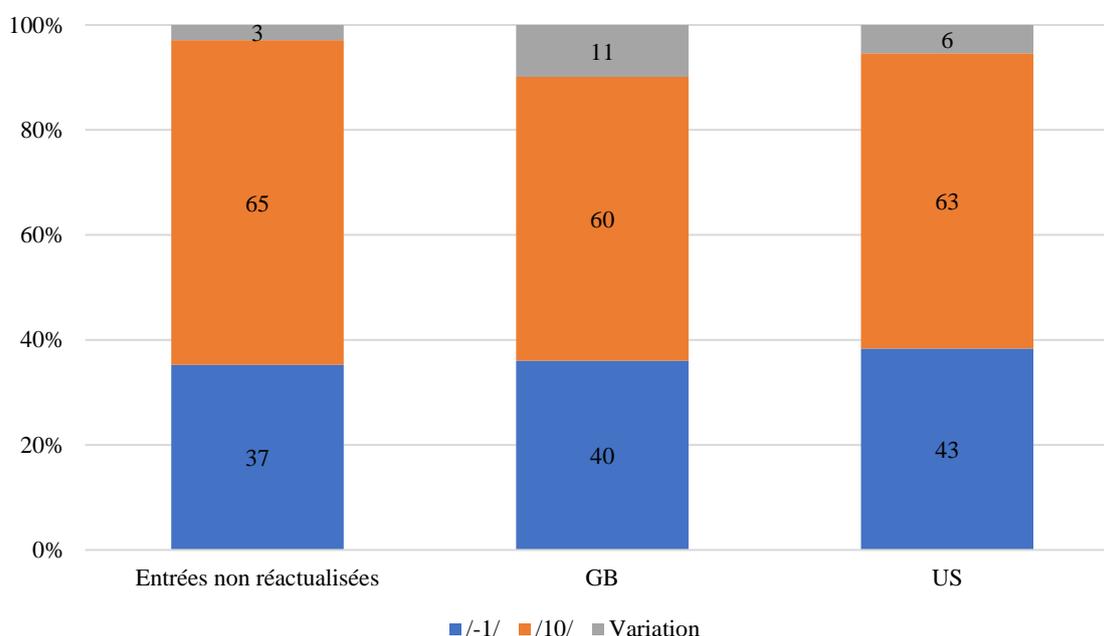
Nous terminâmes cette étude lexicographique par une analyse d'un phénomène sémantique fréquent lors du phénomène d'emprunt, celui de la réduction du sens d'origine. En effet, la langue anglaise associe souvent un emprunt à un seul domaine, alors même qu'il véhicule un sémantisme plus étendu dans sa langue d'origine. Je ne citerai ici que l'exemple de *madrassa*, qui illustre parfaitement ce phénomène. Le terme arabe désigne une école au sein de laquelle tout type d'enseignement est dispensé, tandis que l'*OED* restreint le sens de l'item en lui attribuant uniquement un enseignement de type religieux.

La création d'un corpus d'emprunts arabes en anglais me permit ensuite de travailler sur l'accentuation de ces items dans les dictionnaires anglais pour en aborder les problématiques sous-jacentes : les règles du système phonologique de l'anglais rendent-elles compte de l'accentuation des emprunts arabes ? Un recours aux propriétés phonologiques de la langue arabe permettrait-il de le faire plus efficacement ? L'étude est actuellement terminée et je suis en train d'en achever la mise en forme afin de la soumettre à une revue⁴⁸. J'hésite toutefois à agrémenter cette analyse d'une sous-partie complète sur les principes accentuels de la langue arabe en m'adjoignant les services d'un phonologue spécialisé dans la prononciation de l'arabe, pour des raisons que je vais développer dans les paragraphes suivants.

L'analyse de l'accentuation des emprunts arabes repose sur 335 des 409 items du corpus de 2020. En effet, les items monosyllabiques ne sont pas conservés, tout comme les emprunts n'ayant pas de transcription phonétique associée dans l'*OED*. La première partie de l'étude consiste à rendre compte de la position de l'accent principal dans la version en ligne de l'*OED*, en traitant séparément les transcriptions appartenant aux sections réactualisées de celles relevant des parties non remises à jour datant de 1989, c'est-à-dire l'année de création de la version en ligne du dictionnaire. En effet, si les parties réactualisées mentionnent la transcription britannique et américaine, les parties les plus anciennes ne mentionnent qu'une seule transcription, que l'on peut légitimement considérer comme britannique, mais dont les origines ne sont pas spécifiées. Je décidai donc de les traiter séparément. Le graphique 18

⁴⁸ Fournier, P. (en préparation), "Stress in Arabic Loanwords in English".

présente l'accentuation des emprunts arabes dissyllabiques dans les sections non réactualisées de l'*OED*, ainsi que dans les sections remises à jour, en anglais britannique et américain.



Graphique 18 : Accentuation des emprunts arabes dissyllabiques dans l'*OED*.

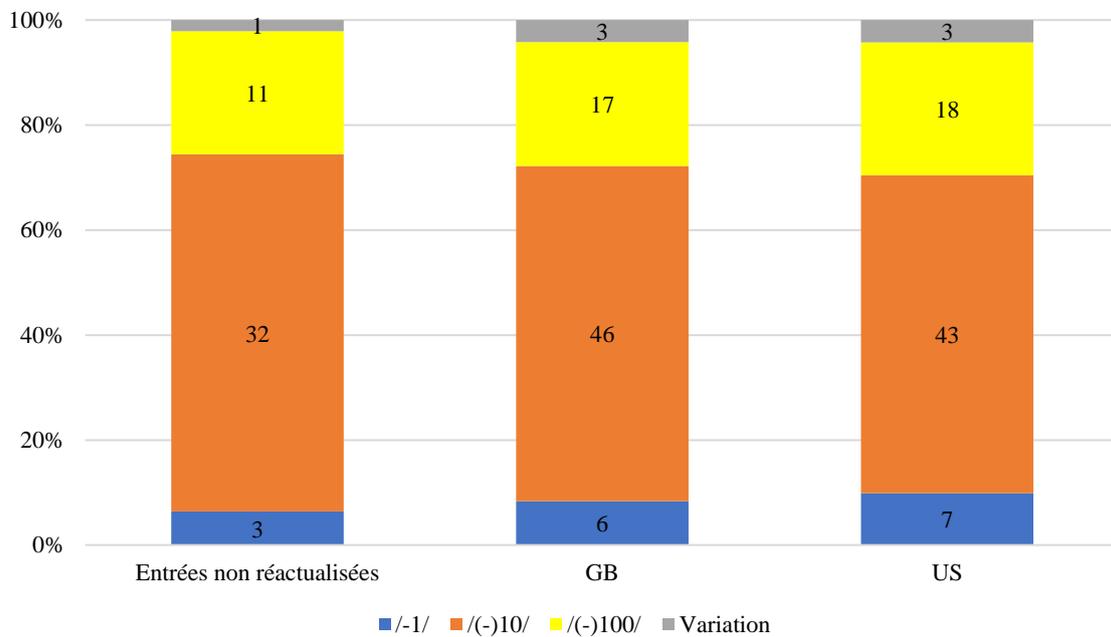
Les résultats du graphique 18 montrent tout d'abord une répartition équivalente des schémas accentuels provenant des zones non réactualisées du dictionnaire et de celles relatives à l'anglais britannique et américain. Les transcriptions provenant des entrées non renouvelées sont donc a priori suffisamment fiables pour être incluses à l'analyse. La différence accentuelle entre l'anglais britannique et américain est quasiment nulle. Ces premiers résultats montrent un taux non négligeable d'emprunts dissyllabiques accentués sur la syllabe finale, schéma accentuel généralement caractéristique des emprunts d'origine française. On constate ainsi qu'environ 40% des emprunts arabes sont accentués sur la syllabe finale.

Exemples provenant des entrées non remises à jour : *Kitab, kitar, Koreish, litham, serab, seyal, shamal, Takbir, talak, tarfa, tumbak, wazir.*

Les exemples ci-dessous sont accentués en /-1/ en anglais britannique et en anglais américain dans les zones réactualisées du dictionnaire :

adhan, agal, Al-Hajj, ansar, emir, feddan, halal, Hamas, haram, imam, khimar, mafeesh.

Les résultats concernant l'accentuation des emprunts arabes d'au moins trois syllabes sont présentés au sein du graphique 19, avec une nouvelle fois la distinction entre les entrées réactualisées et non réactualisées.



Graphique 19 : Accentuation des emprunts arabes au minimum trisyllabiques dans l'OED.

Les tendances sont une nouvelle fois identiques entre les entrées remises à jour et celles non remises à jour, ainsi qu'entre l'anglais britannique et américain, ce qui confirme, d'une part, que les emprunts arabes ne sont pas sujets à la variation dialectale, à l'inverse des emprunts français, et qui rend légitime, d'autre part, l'utilisation des transcriptions des entrées non renouvelées du dictionnaire. Les résultats montrent que les emprunts d'au moins trois syllabes sont principalement accentués sur la syllabe pénultième.

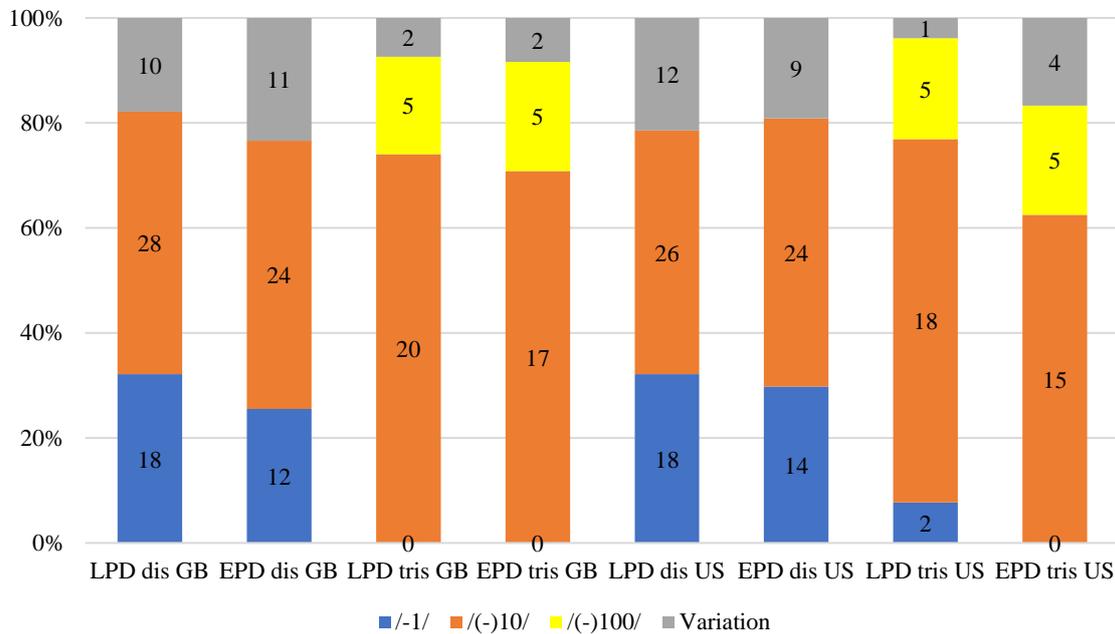
Exemples : *felafel, gandoura, inshallah, tabbouleh, tariqa*.

Toutefois, la proportion d'emprunts accentués sur la syllabe antépénultième n'est pas négligeable, avec 20% des cas recensés.

Exemples : *Cabila, frasilah, Maliki, mandarah, ziraleet*.

Après avoir démontré grâce à une analyse du taux de correspondance entre l'anglais britannique et américain que la différence dialectale n'était pas suffisamment importante (+ 90% de correspondance totale) pour exercer une influence sur les résultats, je croisai ensuite les résultats obtenus à partir des données de l'OED avec les transcriptions proposées par LPD et EPD. Les données de l'OED attestant des mots très rares en anglais, à l'inverse de LPD et

EPD, je ne recensai ainsi que 83 mots dans *LPD* et 71 dans *EPD*. Le graphique 20 présente les résultats de l'accentuation des emprunts arabes dans ces deux dernières sources.



Graphique 20 : Accentuation des emprunts arabes dans *LPD* et *EPD*.

Les résultats concernant les dissyllabes montrent une forte correspondance entre les deux sources dictionnaires, tant pour l'anglais britannique que pour l'anglais américain. La proportion d'accentuation finale dans les dissyllabes n'est pas négligeable, notamment si l'on prend en considération les cas de variation illustrant un conflit entre /01/ et /10/. Les emprunts au minimum trisyllabiques sont principalement accentués sur la syllabe pénultième dans les deux dictionnaires et à nouveau il n'y a pas de différence variétale. La proportion d'accentuations antépénultièmes observée dans l'*OED* est équivalente avec environ 20% des emprunts accentués en /(-)100/.

Comment expliquer la proportion importante d'accentuation finale au sein des dissyllabes d'origine arabe et la proportion non négligeable d'accentuation antépénultième au sein des emprunts au minimum trisyllabiques ? Dans la partie suivante de mon étude, je décidai de tester l'efficacité des principes accentuels régissant le système anglais sur les emprunts arabes, afin de déterminer si ces phénomènes accentuels étaient déclenchés par l'application des paramètres phonologiques de l'anglais. J'utilisai pour ce faire le système accentuel algorithmique élaboré par J.-M. Fournier (2010) et qui consiste en une combinaison de paramètres syntaxiques, morphologiques et segmentaux.

Je ne travaillai pour cette analyse que sur les emprunts présentant un schéma accentuel unique, ce qui excluait les cas de variation. Je remarquai ainsi que l'accentuation finale de certains dissyllabes était justifiée par la présence d'une terminaison contraignante.

-C'C'e# : *gi'raffe*.

-V'V'(C₀(e))# : *ca'beer, ha'boob, hu'zoor, jin'nee, ma'feesh, sa'laam*.

Un cas seulement relève de la règle des composés consistant à accentuer le premier élément lexical : '*sim-sim*. La règle segmentale des dissyllabes substantifs consistant à accentuer la première syllabe de l'item justifie l'accentuation en /10/ des emprunts suivants :

'mukhtar, 'mulai, 'murshid, 'Muslim, 'naib, 'naskhi, 'omda, 'rayah, 'riba, 'rottol, 'sabkha, 'sarsar, 'semsem, 'shadda, 'Shammar, 'Sheikha, 'sherwal, 'Shiah, 'shisha, 'shura, 'Siddi, 'Sufi, 'sulham.

Cependant, je remarquai un nombre important d'exceptions à la règle des dissyllabes accentués en /01/ (tout comme de nombreux dissyllabes d'origine française d'ailleurs). Le pourcentage d'exceptions est équivalent à 31% dans les dissyllabes (57 exceptions sur 183). Concernant les 101 emprunts au moins trisyllabiques présentant un schéma accentuel unique, sept cas relèvent de l'application la loi de la dérivation neutre :

I'raqi > I'raq, Ku'waiti > Ku'wait, 'Maghribi > 'Maghrib, 'Muslimah et 'Muslimin > 'Muslim, Qa'tari > Qa'tar⁴⁹ et 'Yemeni > 'Yemen.

Dans tous ces cas, les terminaisons sont neutres et par conséquent l'accentuation des formes dérivées est déduite de la structure accentuelle des dérivants. Je décidai donc d'écarter ces formes dérivées de l'analyse car l'accentuation se calcule en vertu de la relation isomorphique à partir des dérivants dissyllabiques qui relèvent de la catégorie précédente. La terminaison contraignante - {i,e,u}+(VC₀(e)) justifie l'accentuation des trois emprunts

'ariel, 'Roumia, 'sakia,

mais ne permet pas de rendre compte de l'accentuation des cinq emprunts suivants, qui sont donc des exceptions :

mudi'rieh, Mur'ji'ah, Senus'sia, Sha'ria, ta'zia.

La règle des mots composés consistant à placer l'accent sur le premier élément lexical ne s'applique pas dans les emprunts *baba ga'noush* et *quai'ss ki'tir*, qui sont accentués sur le second élément de composition. Il est fort à parier que les locuteurs anglophones interpréteraient l'espace entre les deux éléments lexicaux comme un signe propre à la catégorie des composés.

⁴⁹ L'accentuation des dérivants est extraite de *LPD*. Le terme *Qatar* est accentué en /10/ ou en /01/ en anglais britannique et américain. L'accentuation de *Qatari* est par conséquent déduite de *Qa'tar*.

La règle des mots « italiens » détaillée dans les sections précédentes justifie l'accentuation sur la syllabe pénultième des 12 emprunts suivants :

Gar'shuni, ham'mada, inti'fada, mau'lana, me'dina, Nas'rani, qa'sida, sha'hada, ta'hina, ta'rada, Ya'hudi, Yu'nani.

Une seule exception à la règle des mots « italiens » est à noter : *'Adeni*. La présence d'un agrégat consonantique en position préfinale permet de justifier l'accentuation pénultième dans les 15 emprunts suivants :

Allahu 'Akbar, bis'millah, gai'assa, ha'rissa, ma'halla, Me'rissa, Mu'hammad / Mu'hammed, Mu'harram, mu'tawwa, pa'stilla, Sala'fiyya, se'besten, Se'nussi, sha'warma, ta'halli.

Je n'ai observé que deux exceptions à cette règle : *Alhamdulillah* et *Subbhanal'lah*. Enfin, la Règle Normale d'accentuation qui est associée à un accent antépénultième dans les mots longs ne relevant d'aucune des règles mentionnées précédemment justifie l'accentuation de 20 emprunts.

Exemples : *'Alawi, Al'debaran, 'athanor, 'cafila, 'Fatiha / 'Fatihah, 'Fomalhaut, 'frasilah, 'habara, 'Maliki.*

Les accentuations sur la syllabe finale ou pénultième attestées dans les mots suivants sont donc hors-système. Aucune règle formulée à partir de l'algorithme accentuel de J.-M. Fournier ne permet de rendre compte des schémas accentuels des emprunts de la liste suivante, qui sont donc considérés comme des exceptions :

a'baya, An'sayri, ar'gileh, dar'buka, fe'lafel, fel'lagha, gan'doura, gha'wazee, gha'zeeyeh, ha'lala, ha'lawi, ha'wala, Howei'tat, ijtihad, Ka'babish, keffiyeh, Khl'lafat, ma'haila, mu'hajir, Na'sara, Nu'sayri, Sa'hara, Sah'rawi, Sa'lafi, sa'luki, Shera'rat, She'rari, she'reefa, she'rifi, Swa'hili, ta'riqa, tum'baki, za'riba / za'reba.

Sur les 94 emprunts arabes d'au moins trisyllabes, 50 d'entre eux (c'est-à-dire 53%) sont accentués conformément aux règles en vigueur au sein du système accentuel proposé par J.-M. Fournier, mais 44 (c'est-à-dire 47%) constituent en revanche des exceptions, dans le sens où ils sont accentués différemment de ce que l'on pourrait attendre de l'application des règles en vigueur. Le pourcentage d'exceptions augmenterait de façon significative si les cas de variation accentuelle étaient inclus dans l'analyse. Nul système ne peut prétendre rendre compte de l'intégralité des faits ; toutefois, la proportion importante d'exceptions et la présence de schémas accentuels spécifiques au sein des dissyllabes et des trisyllabes (et plus) me conduisirent à envisager d'autres pistes permettant de rendre compte de l'accentuation de tous les emprunts arabes.

Je formulai tout d'abord l'hypothèse d'une application de l'accentuation pénultième qui constitue le schéma préférentiel au sein des emprunts en anglais (à l'exception des emprunts au français). Les recherches précédentes mentionnées dans cette synthèse et notamment celles relatives aux emprunts italiens ou japonais ont démontré que les emprunts sont massivement accentués sur la syllabe pénultième en anglais et que ce schéma accentuel pourrait même être un marqueur phonologique permettant d'identifier l'origine étrangère d'un item. Si cette piste est envisageable pour les emprunts arabes d'au moins trois syllabes, elle ne permet pas de rendre compte de l'accentuation finale fréquemment attestée au sein des dissyllabes.

La deuxième piste que j'explorai était de nature segmentale, car je constatai que de nombreux emprunts arabes accentués sur la syllabe finale relevaient d'une structure de type <- VC#> notamment au sein des dissyllabes.

Exemples : *adhan, Hamas, imam, salat*.

L'argument d'une structure segmentale spécifique exerçant une influence déterminante sur la position de l'accent ne me parut toutefois pas pertinente, car des emprunts relevant de la même structure étaient à l'inverse accentués en /10/.

Exemples : *Deneb, loofah, marid*.

La dernière hypothèse concernait la reproduction de l'accentuation arabe d'origine. Les transcriptions phonétiques de l'arabe extraites de *LPD* me furent précieuses dans cette optique, car elles mentionnent la position de la syllabe accentuée en arabe. Je parvins ainsi à extraire du dictionnaire 21 entrées mentionnant la transcription arabe. Le tableau 5 répertorie les transcriptions arabes extraites de *LPD*.

Emprunt	GB	US	Transcription arabe
<i>halal</i>	/01/-/10/	/01/-/10/	[ħa 'la:l]
<i>Hadith</i>	/01/	/01/	['ħa 'di:θ]
<i>inshallah (insha'allah)</i>	/010/	/201/	[in ʃa: ʔaħ 'ħa:h, in ʃa: 'ħa:h]
<i>Hamas</i>	/10/-/01/	/10/-/01/	[ħa 'ma:s]
<i>Allah</i>	/10/-/01/	/10/-/01/	[aħ 'ħa:h]
<i>al-Qaeda</i>	/01(0)0/	/01(0)0/	[al 'qa: ʃi da]
<i>Saudi</i>	/10/-/010/	/10/	[sa 'ʃu: di:]
<i>jihad</i>	/01/	/01/	[dʒi 'ha:d]
<i>Koran</i>	/01/	/01/	[qur 'ʔa:n]
<i>Fatah</i>	/10/	/10/	['fa taħ]
<i>umma (ummah)</i>	/10/	/10/	['ʔum mah]
<i>haram</i>	/01/	/01/	[ħa 'ra:m]
<i>niqab</i>	/01/	/01/	[ni 'qa:b]
<i>Qur'an (Quran)</i>	/01/	/01/	[qur 'ʔa:n]
<i>kibla (kiblah)</i>	/10/	/10/	['qib lah]
<i>Eid-al-Fitr (Eid-ul-Fitr)</i>	/2010/	/2010/	[ʃi:d al 'fitʃr]
<i>intifada</i>	/2010/	/2010/	[in ti 'fa: dʃa]
<i>Medina (medina)</i>	/010/	/010/	[me 'di: na]
<i>Muhammed</i>	/010/	/010/	[mu 'ħam mad]
<i>sharia (shari'ah)</i>	/010/	/010/	[ʃa 'ri: ʃa]
<i>hijab</i>	/01/	/01/	[ħi dʒa:b]

Tableau 5 : Transcriptions phonétiques arabes extraites de *LPD*.

Mis à part les cas de *hijab* qui n'est pas accentué dans la transcription arabe et de *Hadith* pour lequel deux accents principaux sont notés en arabe, les résultats de la comparaison accentuelle sont saisissants. Il y a en effet une correspondance quasi-parfaite entre la position de la syllabe accentuée en arabe et en anglais.

Exemple : *ni'qab* GB / US et [ni 'qa:b] en arabe.

Les quelques cas de variation accentuelle présents en anglais traduisent un conflit latent entre reproduction de l'accentuation arabe et application des règles relatives au système phonologique de l'anglais. Le pourcentage de correspondance est tel que l'hypothèse d'une conservation ou reproduction de l'accentuation arabe rend bien mieux compte du fonctionnement accentuel des emprunts arabes en anglais que les principes testés précédemment. Mon étude se conclut sur cette hypothèse prometteuse, en raison de mon incapacité momentanée à traiter les phénomènes accentuels en arabe et à procéder ainsi à la comparaison systématique de l'accentuation des emprunts de mon corpus avec les items arabes correspondants. J'envisage dans un premier temps de publier ces résultats qui me paraissent

révélateurs de l'influence des langues sources sur l'accentuation des emprunts en anglais et, dans un second temps, de m'adjoindre les services d'un collègue phonologue spécialisé en arabe afin de poursuivre le travail entrepris dans cette étude.

Cette partie a présenté de manière détaillée les résultats de mes travaux sur l'accentuation des emprunts au japonais, au gaélique écossais et à l'arabe. Mes réflexions se sont également nourries de mon travail de supervision des mémoires de recherche des étudiants de Master, certains d'entre eux ayant travaillé sur l'accentuation des emprunts provenant d'autres langues sources. Je pense notamment au mémoire de R. Ramdane Chérif⁵⁰ (2018) consacré à l'évolution de l'accentuation des emprunts français dans les dictionnaires anglais de prononciation. Ce travail permet de tester la validité de mon étude diachronique de 2016 effectuée à partir de deux versions antérieures des dictionnaires *LPD* et *EPD* (voir les résultats en Annexe 1). Le mémoire d'A. Chaou (2021) sur l'accentuation des emprunts portugais en anglais⁵¹ permet de recueillir des données relatives à une nouvelle langue source (voir les résultats en Annexe 2). Les résultats significatifs du mémoire de K. Si Amer⁵² (2021) consacré à l'accentuation des emprunts espagnols en anglais figurent en Annexe 3 et les résultats des travaux menés par A. Hussein⁵³ (2020) sur la prononciation des emprunts arabes en anglais, et plus spécifiquement sur les règles graphophonématiques sont en Annexe 4. J'ai supervisé d'autres mémoires qui, s'ils ne sont pas directement centrés sur mes propres perspectives, investissent des domaines intrinsèquement liés à mes problématiques de recherche ; comme par exemple celui d'A. Mahtout (2022) sur les mutations morphologiques des emprunts allemands en anglais ou encore celui de G. Zhao (2023) sur la lexicalisation des emprunts chinois en anglais. La supervision de ces travaux liés au domaine des emprunts en anglais suggère autant de pistes de réflexion permettant d'avancer sur mes propres problématiques.

Cette troisième partie consacrée à la description des résultats est certes longue en comparaison des autres parties, mais elle est nécessaire pour mieux comprendre les étapes du cheminement intellectuel m'ayant conduit à proposer diverses tentatives de formalisation du système des emprunts en anglais.

⁵⁰ Ramdane Chérif, R. (2018), *L'évolution de l'accentuation des emprunts français dans les dictionnaires anglais de prononciation*, Mémoire de Master 2 recherche, Université Sorbonne Paris Nord.

⁵¹ Chaou, A. (2021), *Stress assignment in Portuguese loanwords in contemporary English*, Mémoire de Master 2 recherche, Université Sorbonne Paris Nord.

⁵² Si Amer, K. (2021), *Stress assignment in Spanish loanwords in contemporary English*, Mémoire de Master 2 recherche, Université Sorbonne Paris Nord.

⁵³ Hussein, A. (2020), *Vowel pronunciation of Arabic loanwords in English*, Mémoire de Master 2 recherche, Université Sorbonne Paris Nord.

Chapitre 4 : Tentatives de formalisation du sous-système des emprunts en anglais

En préambule à cette tentative de formalisation des résultats de mes recherches relatives à la prononciation des emprunts en anglais, domaine auquel je souhaite pleinement me consacrer à l'avenir, je mentionnerai les ouvrages, citations et approches qui m'ont profondément inspiré dans ce cheminement.

La première source d'inspiration est en réalité une simple citation de M. Hammond, extraite de son célèbre ouvrage *The Phonology of English : A Prosodic Optimality-Theoretic Approach* (1999). Ces quelques lignes que je restitue ci-dessous ne manquèrent pas de provoquer chez moi un vif étonnement (1999 : 284) :

In some languages, borrowed words exhibit rather different phonological patterns from those of native words, but this does not appear to be the case in the stress system of English.

Je tombai tardivement sur cette citation, dans le sens où mes recherches sur la prononciation des emprunts avaient déjà débouché sur des pistes et des hypothèses qui me paraissaient satisfaisantes pour justifier l'existence d'un comportement spécifique des emprunts en anglais. Quelle ne fut donc pas ma surprise de constater qu'une telle citation figurait dans l'un des ouvrages majeurs de la phonologie moderne. Les échanges que j'ai pu avoir sur le sujet avec Q. Dabouis par la suite ne firent que confirmer le fait qu'au sein du paysage actuel de la phonologie, la croyance selon laquelle les paramètres phonologiques régissant les emprunts en anglais ne sont pas différents de ceux utilisés pour rendre compte de la prononciation du reste du lexique reste profondément ancrée au sein des théories phonologiques majeures. Il faut sans doute y voir l'influence déterminante de *The Sound Pattern of English* (1968) de N. Chomsky et M. Halle, qui utilise de nombreux emprunts pour servir d'exemples aux règles énoncées au sein des inventaires. La distinction entre vocabulaire natif et vocabulaire étranger fut donc abolie dès 1968 concernant les règles permettant de déterminer la prononciation du lexique, ce qui a sans doute grandement contribué par la suite à considérer que des positions telles que celle adoptée par M. Hammond allaient de soi. Je n'eus de cesse depuis lors de démontrer que les emprunts constituaient un sous-système unique au sein du système anglais.

L'ouvrage *Loan Phonology* édité par A. Calabrese et L. Wetzels en 2009 fut très inspirant, car il m'ouvrit à un monde de travaux quasiment inconnus pour moi. Cet ouvrage qui propose un état de l'art des recherches traitant des phénomènes phonologiques à l'interface langue source - langue cible est d'une immense richesse et les travaux proposés sont très aboutis. L'introduction scientifique de l'ouvrage rédigée par les éditeurs est à elle seule une source

exceptionnelle d'informations. Je citerai dans ce préambule consacré aux sources fondatrices de mon propre travail de structuration du système les deux pensées qui s'affrontent au sein de ce courant théorique concernant le mode d'intégration et la prononciation des emprunts dans une langue cible (1999 : 1-3). Les chercheurs défendant le **modèle phonologique** (*phonological stance model*) (L. Hyman 1970, H. Jacobs & C. Gussenhoven 2005, C. Paradis & D. LaCharité 2005) considèrent que les emprunts sont intégrés aux langues cibles par l'intermédiaire de locuteurs bilingues. De ce fait, les emprunts sont prononcés grâce aux mécanismes phonologiques régissant la langue cible. Il s'agit là de la position dominante au sein de ce courant, même si d'autres chercheurs pensent que, puisque les locuteurs sont bilingues, les emprunts peuvent également légitimement être réalisés par l'intermédiaire des paramètres phonologiques de la langue source. Le second modèle correspond à la **position perceptuelle** (*perceptual stance model*) (D. Silverman 1992, S. Peperkamp & E. Dupoux 2008), et défend la position selon laquelle les emprunts sont intégrés à des langues cibles par des locuteurs ne maîtrisant pas les mécanismes de la langue source au sein de laquelle ils ont puisé le lexique. La prononciation de l'emprunt dans la langue cible est donc basée sur la perception qu'ont les locuteurs de la prononciation de l'item dans la langue source et sur leur capacité à reproduire cette prononciation avec les outils dont ils disposent dans la langue cible, ce qui implique une restitution plus ou moins fidèle, en fonction des qualités de perception dudit locuteur. Le processus comprend plusieurs étapes qui sont autant de phases d'ajustement et d'adaptation aux mécanismes de la langue cible, tout en conservant l'idéal d'une reproduction aussi fidèle que possible de la langue d'origine.

Les résultats de mes travaux me font pencher vers le modèle phonologique, car le taux de reproduction de l'accentuation d'origine constaté dans certaines de mes études est trop parfait pour que le processus ne soit pas initié par des locuteurs bilingues (ou extrêmement cultivés, comme je le disais au sujet des lexicographes intégrant les composés savants dans les bases de données dictionnaires). Quant à savoir si la prononciation des emprunts est basée sur les propriétés phonologiques de l'anglais ou celles des différentes langues sources, le conflit me semble à l'heure actuelle insoluble, mais la vérité se situe peut-être même au croisement des deux hypothèses, ce qui justifierait d'un côté la présence d'un sous-système spécifique aux emprunts au sein du système phonologique de l'anglais et de l'autre l'influence indéniable des langues sources sur la prononciation des emprunts correspondants en anglais. Quoi qu'il en soit, l'intérêt de l'approche développée par la *Loan Phonology* réside dans le fait que tous les chercheurs inscrivant leurs travaux au sein de ce courant considèrent comme possible, si ce

n'est probable, le fait que les langues sources exercent une influence par l'intermédiaire des locuteurs sur la façon dont un emprunt sera réalisé dans une langue cible, quelle qu'elle soit.

La citation de K. Church (1985) dont j'ai déjà mentionné le caractère décisif lors de mes études sur les emprunts italiens eut en réalité une portée théorique encore bien plus importante. Mes tout premiers résultats sur les emprunts italiens en anglais et la correspondance accentuelle quasi parfaite entre l'italien et l'anglais me laissèrent penser – et il s'agissait là d'une erreur de « jeune chercheur » – que l'accentuation de tous les emprunts et pas seulement celle des emprunts italiens pouvait se justifier en ayant recours à l'accentuation des langues d'origine. J'avais ainsi tendance à interpréter tous mes résultats au prisme de cette hypothèse. Mon erreur fut de courte durée car, d'une part, les résultats concernant l'étude de deux autres langues sources (à savoir le français et le japonais) ne reflétèrent absolument pas une telle correspondance accentuelle, et, d'autre part, l'hypothèse d'un « pseudo-accent spécifique aux mots étrangers en anglais » mentionnée par K. Church me fit reconsidérer les choses. En effet, si l'influence des langues sources sur l'anglais me semblait indéniable, pour autant, des locuteurs anglophones non bilingues étaient tout de même en mesure de prononcer les emprunts, en utilisant des paramètres internes au système anglais. Plusieurs facteurs, dont le niveau d'instruction, jouaient indéniablement un rôle dans la maîtrise de ces paramètres. Il existait donc des paramètres internes au système anglais permettant de rendre compte de la prononciation des mots d'origine étrangère. Ce constat me permit d'identifier la présence d'un conflit perpétuel entre influence des langues sources et autonomie du système anglais.

Les travaux de Y. Kang (2010) sur l'adaptation prosodique des emprunts m'ont également inspiré. Il ressort de son analyse que le bilinguisme des locuteurs introduisant des mots dans une nouvelle langue leur permet de percevoir la position des syllabes accentuées dans la langue source et d'être ainsi en mesure de reproduire ces schémas prosodiques dans la langue cible (2010 : 2308) :

the closer the contact and the higher the level and rate of bilingualism, the more likely the adapters are to perceive stress contrasts of the input language correctly, in turn leading to the stronger possibility of the preservation of input stress.

Le modèle d'adaptation prosodique développé dans S. Davis, N. Tsujimura & J. Tu (2012) me donna des arguments en faveur d'une formalisation du système. Leur modèle s'organise autour de trois processus interdépendants : la comparaison systématique de phénomènes prosodiques dans la langue source et dans la langue cible (certains phénomènes résistent en effet davantage à la transmission que d'autres), les mécanismes permettant de rendre compte de la prononciation des emprunts dans une langue cible (certaines langues possèdent des

paramètres spécifiques permettant de rendre compte de la prononciation des emprunts tandis que d'autres langues utilisent les mêmes paramètres pour tout le lexique) et la détermination de la structure syllabique et segmentale des deux langues, paramètre potentiellement déterminant dans le processus d'adaptation ou de reproduction.

Enfin, la collaboration avec Q. Dabouis a également été décisive dans la structuration et la formalisation du système. Il ressort de la plupart de ses travaux une croyance profonde en l'autonomie du système anglais et en l'existence de strates distinctes mais interdépendantes, ce qui, associé à mes travaux sur l'influence des langues sources et sur les spécificités des structures empruntées au sein du système anglais, nous permet de proposer une formalisation du système organisé en réalité en sous-systèmes (section 4.3) possédant chacun des propriétés linguistiques distinctes des autres sous-systèmes.

La dernière partie de ma synthèse est structurée en trois sections qui traduisent l'évolution de ma pensée au fur et à mesure des publications et des collaborations. La première d'entre elles est consacrée à ma toute première tentative de formalisation de l'influence des langues sources sur l'anglais.

4.1 Les différents degrés d'influence des langues sources

Cette première tentative de synthétisation de l'influence des langues sources correspond à une publication de 2016⁵⁴ dans la revue *Itinéraires*, faisant suite à une journée d'études consacrée à la thématique « Entraves et obstacles aux circulations » et rassemblant une partie des chercheurs du laboratoire Pléiade. À cette occasion, j'essayai de synthétiser mes premiers résultats à la lumière de ce thème qui paraissait propice à une réflexion d'ensemble sur les travaux déjà menés.

Mes travaux sur l'accentuation des trois langues sources que sont l'italien, le japonais et le français illustrent trois formes d'influences distinctes sur l'accentuation des emprunts correspondants en anglais. Les résultats relatifs à l'accentuation des emprunts italiens montrent une correspondance quasi parfaite entre la position de la syllabe accentuée en italien et en anglais. Ce taux de correspondance exceptionnel n'est pas uniquement attesté pour les accentuations pénultièmes, auquel cas l'argument selon lequel les emprunts sont massivement

⁵⁴ Fournier, P. (2016), « Nouvelles perspectives sur l'accentuation des emprunts en anglais contemporain », *Itinéraires*, 2015-2, *Stumbling blocks. Entraves et obstacles aux circulations. Approches pluridisciplinaires*, M. Lévêque & C. Parfait (éds.) (<http://itineraires.revues.org/2827>).

accentués sur la syllabe pénultième en anglais et ce, quelle que soit leur origine, pourrait entrer en contradiction avec mon hypothèse. Le taux de correspondance accentuelle est également très élevé pour les accentuations finales ou antépénultièmes. Ces résultats me permirent de conclure que les emprunts italiens en anglais reproduisaient systématiquement la structure prosodique d'origine. Aucun obstacle ne semblait interférer dans le processus.

L'accentuation des emprunts japonais montre à l'inverse une absence totale de correspondance accentuelle entre la syllabe accentuée en japonais et la syllabe accentuée dans l'emprunt en anglais. Le taux de correspondance est si peu significatif que la reproduction de la structure prosodique de la langue source ne peut en aucun cas constituer une piste permettant de rendre compte de l'accentuation des emprunts japonais en anglais. Par conséquent, si la reproduction des paramètres prosodiques de la langue source était entravée, j'en conclus que le système phonologique de l'anglais attribuait un accent pénultième par défaut pour ces emprunts. Le schéma accentuel pénultième était en effet le schéma le plus attesté dans les emprunts en anglais et ce, quelle que soit leur origine. Ce schéma accentuel pénultième préférentiel au sein du système des emprunts serait un marqueur phonologique d'identification de l'origine étrangère d'un item et d'appartenance à ce sous-système.

Enfin, les résultats relatifs aux emprunts français sont plus complexes à interpréter, car la variation dialectale entre l'anglais britannique et l'anglais américain brouille les pistes. Mes données montrent que l'accent final démarcatif du français est mieux reproduit en anglais américain qu'en anglais britannique. Il y a donc, certes, une reproduction de la structure prosodique du français mais elle est toute relative en anglais britannique où le pourcentage d'emprunts concernés par le phénomène de rétraction accentuelle est important. J'en conclus donc que les emprunts français constituaient une sous-catégorie spécifique d'emprunts car, d'une part, le schéma accentuel préférentiel était final et, d'autre part, seuls les emprunts français semblaient sujets à la variation dialectale anglais britannique / anglais américain au sein du système. De plus, les propriétés morphologiques, segmentales et sémantiques entre autres constatées au sein des emprunts français se distinguaient de celles des emprunts provenant d'autres langues. Toutes ces raisons, associées à l'afflux historique massif d'emprunts d'origine française et au remodelage du système phonologique de l'anglais en raison de l'influence de la logique romane véhiculée par ces milliers d'emprunts, me laissèrent penser que les emprunts français constituaient un sous-système unique et hybride au sein du système anglais.

Quant à déterminer les raisons pour lesquelles la structure prosodique italienne est parfaitement reproduite en anglais alors que les paramètres prosodiques du japonais sont

bloqués, le recours à la nature même des systèmes prosodiques en question me semblait une piste prometteuse. L'italien étant une langue avec une accentuation de type lexical, tout comme l'anglais, cela me paraissait favoriser la transmission des propriétés prosodiques de la langue source vers la langue cible. À l'inverse, en raison des différences systémiques profondes entre les paramètres prosodiques du japonais et de l'anglais, notamment au niveau de la matérialisation acoustique de la syllabe accentuée qui se traduit en japonais par une chute mélodique, la reproduction de l'accentuation d'origine semblait plus difficile à réaliser. J'en conclus qu'il était possible de mettre en place une typologie des langues sources basée sur la capacité de reproduction de la structure prosodique d'origine en anglais, avec comme variable principale pour organiser la typologie, la proximité des structures prosodiques. En d'autres termes, il y a une forte correspondance accentuelle entre l'italien et l'anglais car la similitude des structures prosodiques le permet tandis que la profonde différence entre les deux systèmes prosodiques du japonais et de l'anglais ne permet pas de reproduire l'accentuation d'origine, en conséquence de quoi un accent pénultième est automatiquement généré en anglais.

Plusieurs résultats présentés dans cette synthèse suggèrent que les locuteurs bilingues sont à l'origine de l'introduction des emprunts en anglais, en raison du fort taux de correspondance accentuelle entre l'anglais et certaines langues sources (exemples : mes travaux sur les emprunts italiens, les conclusions des mémoires de recherche supervisés, portant sur l'accentuation des emprunts espagnols et portugais) ou de la reproduction de certaines propriétés phonétiques ou graphophonématiques de la langue source (exemples : les voyelles libres étrangères dans les emprunts arabes ou les correspondances graphie-phonie spécifiques au gaélique écossais en anglais). Même si les résultats me conduisent à privilégier cette hypothèse, toujours est-il que lorsque les structures prosodiques sont trop différentes, comme dans le cas du japonais et de l'anglais, les locuteurs bilingues vont opter pour les paramètres phonologiques de la langue cible (ici l'anglais) pour matérialiser la position de la syllabe accentuée car le matériel prosodique à disposition dans la langue source ne permettrait qu'imparfaitement de le faire⁵⁵.

La prononciation des emprunts en anglais est donc fortement influencée par les langues sources lorsque la reproduction de l'accentuation d'origine n'est pas entravée par des systèmes phonologiques incompatibles. Lorsque l'interaction entre les deux systèmes phonologiques ne

⁵⁵ Mes résultats préliminaires concernant l'accentuation des emprunts arabes en anglais qui montrent une correspondance quasi parfaite entre la position de la syllabe accentuée en arabe et en anglais me laissent perplexes. Les systèmes prosodiques des deux langues semblent à première vue très différents, pour autant on constate une remarquable reproduction de l'accentuation d'origine dans les données préliminaires. Cette piste mérite d'être creusée dans le futur.

permet pas une telle reproduction, l'anglais attribue un accent pénultième « par défaut » aux emprunts, qui constitue donc le schéma accentuel préférentiel de cette catégorie. Les résultats du test oral⁵⁶ concernant les emprunts à l'italien ont confirmé que lorsque les locuteurs anglophones se retrouvent face à un mot d'origine étrangère, ils ont tendance à privilégier une accentuation pénultième (même si cet item n'est pas accentué sur la syllabe pénultième dans sa langue d'origine) et ce schéma traduit l'origine étrangère de l'item. Dans cette étude datant de 2018, j'émis l'hypothèse selon laquelle l'accentuation pénultième, marqueur phonologique propre au système anglais, pourrait être héritée de l'afflux massif en anglais de mots d'origine italienne accentués sur la syllabe pénultième. La reproduction quasi parfaite du schéma accentuel pénultième italien (en effet, 80% des mots italiens sont accentués sur la pénultième), aurait ainsi généré cet accent préférentiel. Cette hypothèse me paraît toutefois peu probable aujourd'hui car d'une part, elle est extrêmement difficile à prouver et d'autre part, une étude à paraître en 2024⁵⁷ sur l'influence des langues sources sur l'accentuation des emprunts en anglais montre que de nombreuses langues sources accentuent massivement la syllabe pénultième à l'instar de l'italien. De plus, le taux de reproduction élevé de la position de la syllabe accentuée au sein de ces langues sources optant massivement pour une accentuation pénultième vient nuancer mon hypothèse initiale.

Le schéma accentuel préférentiel sur la syllabe pénultième n'est en réalité qu'un des paramètres parmi d'autres permettant aux locuteurs de notifier l'origine étrangère d'un mot. Il existe d'autres propriétés morphologiques, segmentales, sémantiques ou graphophonématiques spécifiques aux emprunts que mes nombreuses études basées sur des corpus d'emprunts m'ont permis au fil des années d'identifier au sein du système anglais.

4.2 Les propriétés du sous-système des emprunts en anglais

Je distingue dans cette section les emprunts français du reste des emprunts, car comme je vais le mettre en évidence, ils possèdent des spécificités propres, qui les distinguent des autres emprunts.

Concernant tout d'abord les propriétés sémantiques des emprunts, elles sont dites universelles. M. Haspelmath (2009) considère que le phénomène d'emprunt est lié à une

⁵⁶ Fournier, P. (2018), "Stress Assignment in Italian Loanwords in English and its Impact on the Stressing of Foreign Words by Native English Speakers", *Corela* HS-24, *Multicultural Spoken English* (<https://journals.openedition.org/corela/5113>).

⁵⁷ Fournier, P. (2024), "The impact of source languages on the stressing of loanwords in English", *Lexis* 23.

motivation culturelle et qu'il est régi par deux facteurs : un critère social (l'emprunt est souvent associé à une notion de prestige) et un critère grammatical (les noms sont plus souvent empruntés que les verbes, qui doivent subir plusieurs phases d'adaptation morphologique avant de pouvoir être pleinement utilisés). M. Haspelmath fait une distinction dans son chapitre entre les différents types d'emprunts et observe des restrictions sémantiques inhérentes à la catégorie. Je me contenterai simplement d'ajouter que l'emprunt doit impérativement être distingué du processus de *code-switching* (C. Myers-Scotton 1993) qui se définit globalement comme l'introduction d'un item d'une langue étrangère L2 dans le discours d'un individu relevant d'une langue L1. Ce processus est ponctuel, propre à chaque locuteur et intervient en **discours** et non pas en **langue**. Le recours aux dictionnaires pour construire mes corpus d'emprunts me garantit toutefois un travail sur des formes institutionnalisées, ne pouvant donc par définition relever du phénomène de *code-switching*.

L'emprunt vise à combler un vide lexical dans une langue et l'on parle dans ce cas d'emprunts « culturels ». Il sert souvent à introduire une notion étrangère, quelle qu'elle soit, dans une langue cible. J'ai constaté que les principaux domaines sémantiques couverts par les emprunts au sein de mes corpus étaient les suivants :

- habitudes culturelles et vie quotidienne : *bodega, matador, anime, geisha, halal*.
- nourriture et boissons : *ravioli, risotto, spaghetti, tortellini, cranachan, tortilla*.
- sports et loisirs : *bonsai, judo, karate*.

Certains sous-domaines sémantiques de l'anglais sont constitués d'items exclusivement empruntés à une seule langue source. On peut citer par exemple le domaine de la musique qui a énormément emprunté à la langue italienne au fil des siècles (exemples : *sonata, sotto voce, cello*). Je pense toutefois qu'il est important de faire une distinction entre les emprunts français et le reste des emprunts au niveau sémantique. En effet, si les emprunts aux autres langues sont, comme je viens de le montrer, des emprunts culturels qui ont globalement une connotation exotique, les emprunts français sont à la croisée entre emprunts culturels et emprunts relevant du cœur du système lexical de l'anglais (*core borrowings* chez M. Haspelmath (2009)). En effet, si pour les emprunts culturels, il n'existe pas d'équivalent lexical dans la langue cible, ce qui constitue d'ailleurs la raison principale du phénomène d'emprunt, les *core borrowings* sont en concurrence avec d'autres items. C'est le cas des emprunts français associés à une notion de prestige et attestés au sein du lexique anglais, malgré la présence d'un concept équivalent au sein du stock lexical d'origine germanique. Il me semble donc, pour conclure sur les propriétés sémantiques des emprunts qu'il est indispensable de dissocier les emprunts français qui, pour des raisons historiques, se situent au cœur du stock lexical de l'anglais et véhiculent une notion

de prestige, des autres emprunts culturels à propos desquels la notion de prestige n'est pas aussi développée. Je notai également la proportion importante de noms propres étrangers figurant au sein des inventaires. Ces noms propres sont généralement des marques, des noms de lieux célèbres ou des figures majeures du monde politique, artistique ou sportif. Il est donc tout à fait possible de les rattacher aux grands domaines sémantiques évoqués ci-dessus.

La deuxième spécificité qui distingue les emprunts du reste du lexique est de nature phonologique et concerne la position de l'accent lexical. Là encore, il importe de dissocier les emprunts français des autres, car si le schéma préférentiel des emprunts français est l'accentuation finale (même si les données concernant l'accentuation en anglais britannique remettent quelque peu en question cette affirmation), les emprunts aux autres langues sont massivement accentués sur la syllabe pénultième. L'accent final est le schéma préférentiel des emprunts français, qui reproduisent ainsi l'accentuation démarcative finale des groupes intonatifs du français. De plus, la présence de terminaisons contraignantes d'origine française associées à une accentuation finale renforce ce phénomène. J'ai déjà longuement évoqué les propriétés accentuelles des emprunts tout au long de ma synthèse et je ne m'étendrai pas davantage ici.

Une autre spécificité d'ordre phonologique concerne la réalisation des voyelles en contexte inaccentué. Au sein du système anglais, les voyelles en contexte inaccentué ont généralement tendance à être réduites. Or, j'ai constaté au sein de mes différents corpus une proportion importante d'emprunts provenant tant du français que des autres langues qui adoptaient une voyelle pleine en contexte inaccentué, ou présentaient un cas de variation de type voyelle pleine / voyelle réduite. Je remarquai que cette particularité constituait une propriété phonologique spécifique des emprunts. Les exemples ci-dessous montrent cette spécificité phonologique des emprunts en anglais britannique (mais le phénomène est également attesté en anglais américain).

Aalborg /'ɔ:l bɔ:g, 'ɑ:l-/ , *anime* /'ænimɛɪ, -ə-/ , *halal* /hə'la:l, 'hælæl, hæ'læl/ , *imam* /ɪ'mɑ:m, 'i:mɑ:m/ , *jihad* /dʒɪ'hæd, dʒə-, -'hɑ:d/ , *kakemono* /,kæki'məʊnəʊ/ , *pizzicato* /,pɪtsɪ'ka:təʊ, -ə-/.

Ces exemples montrent que la position de l'accent est déterminante dans la réalisation des voyelles inaccentuées et qu'un contexte de variation accentuelle, présent ci-dessus dans *halal* et *imam*, pourrait augmenter la proportion de variation voyelle pleine / voyelle réduite en contexte inaccentué. Les contextes typiquement inaccentués dans *anime*, *jihad*, *kakemono* ou *pizzicato* montrent toutefois un phénomène de variation presque systématique. Concernant les emprunts français, le phénomène est également attesté dans les exemples suivants :

Gare du Nord /,ga:dju:'nɔ:/, *gateau* /'gætəʊ/, *laisser-faire* /,leisei'feə, ,les-/.

Les emprunts se caractérisent également par des correspondances graphie-phonie spécifiques tant au niveau vocalique que consonantique. Le tableau 6 répertorie les correspondances graphie-phonie spécifiques des voyelles libres étrangères en anglais restituées à partir de l'ouvrage de J.-M. Fournier (2010 : 113).

Voyelle libre	Graphie	Voyelle libre étrangère
/eɪ/	<a>	/ɑ:/
/i:/	<e>	/eɪ/
/aɪ/	<i>	/i:/

Tableau 6 : Les trois voyelles libres étrangères principales de l'anglais.

Les voyelles libres étrangères sont essentiellement attestées dans les syllabes accentuées des mots d'origine étrangère en anglais. De telles correspondances graphie-phonie sont un indicateur phonétique de l'origine étrangère de l'item et il est même possible de les rencontrer, comme je l'ai montré dans la partie sur les emprunts arabes, dans des contextes généralement entravants (exemples : *halal*, *imam* /ɑ:/). On peut également observer un phénomène de variation voyelle libre / voyelle libre étrangère lié à une différence dialectale :

Exemple : *tomato* = /tə'mɑ:təʊ/ (GB) vs /tə'meɪtəʊ/ (US)

L'exemple de *tomato* montre que l'anglais britannique opte pour une correspondance graphie-phonie de type étranger quand l'anglais américain adopte une correspondance classique au sein du lexique anglais. Voici quelques exemples de voyelles libres étrangères extraites de mes différents corpus :

<a> = /ɑ:/ : *banana*, *makiwara*, *marumage*.

<e> = /eɪ/ : *cliché*, *mokume*.

<i> = /i:/ : *aikido*, *bikini*, *bushido*, *mosquito*.

Il existe également une série de correspondances graphie-phonie spécifiques aux emprunts concernant les voyelles libres étrangères colorées par <r> :

<a> = /ɑ:/⁵⁸ : *safari*.

<e> = /eə/ (GB⁵⁹) : *sombrero*.

⁵⁸ La correspondance phonétique n'est pas différente pour la voyelle monographe <a>, que la voyelle libre étrangère soit colorée ou non par <r>.

⁵⁹ Les diphtongues centralisantes /eə/ et /ɪə/ n'existent pas en anglais américain.

<i> = /iə/ (GB) : *fakir*.

Les voyelles libres étrangères sont également attestées dans les emprunts français :

coup d'état, faux pas = /ɑ:/

corvée, cortège = /ɛ/

coquille, frisson = /i:/.

Les emprunts français possèdent également toute une série de correspondances graphie-phonie qui leur est propre. En voici quelques exemples :

<i> = /æ/ : *coq au vin, meringue*.

<eu> = /ɜ:/ : *danseuse*.

<ou> = /u:/ : *route*.

<oi> = /wa:/ : *memoir*.

Concernant les correspondances graphie-phonie au niveau consonantique, je notai dans l'ouvrage co-écrit avec I. Girard (2022) que certaines correspondances consonantiques étaient spécifiques aux emprunts et parfois même à une seule langue source. Voici quelques exemples extraits de l'ouvrage concernant les emprunts :

<c(c)> = /tʃ/ : *carpaccio, cello*.

<gg> = /dʒ/ : *arpeggio, solfeggio*.

<sch> = /ʃ/ : *putsch, schwa*.

<tz#> = /ts/ : *blitz, quartz*.

<z(z)> = /ts/ : *panzer, pizza*.

Là encore, nous remarquâmes au moment de constituer les inventaires d'exemples de l'ouvrage de transcription phonétique, que les emprunts français présentaient des correspondances distinctes des autres emprunts :

<ch> = /ʃ/ : *charade, machine*.

<g> = /ʒ/ : *genre, prestige*.

<j> = /ʒ/ : *jalousie, jongleur*.

<qu> = /k/ : *liqueur, unique*.

J'ai également remarqué au sein des emprunts français que de nombreuses correspondances graphie-phonie au niveau consonantique étaient sujettes à des différences variétales. Par exemple, la correspondance en /ʒ/ du graphème <j> dans *jalousie* est attestée en anglais britannique uniquement, l'anglais américain attestant une correspondance héritée du système anglais avec /dʒ/. Enfin, pour en terminer sur les correspondances graphie-phonie, de très nombreux emprunts français possèdent une consonne finale muette, particularité absente des mots d'origine germanique.

Exemples : *bourgeois, buffet, confit, coup, rendez-vous*.

Les emprunts possèdent également des propriétés segmentales qui les distinguent du reste du lexique. Concernant le nombre de syllabes, je remarquai dans mes différents corpus la très faible quantité de monosyllabes, et la proportion moindre de dissyllabes par rapport aux emprunts trisyllabiques (et plus). Les termes empruntés tendent donc à être plus longs que les mots d'origine germanique en anglais. Le critère du nombre de syllabes est moins marqué dans les emprunts français, au sein desquels je notai une proportion importante de termes également dissyllabiques.

Les emprunts présentent généralement une structure de type <-VCV#>⁶⁰, héritage de la structure syllabique romane.

<-VCV#> : *bikini, Bratislava, guanaco, legato, maestoso, ravioli, volcano*.

Une restriction segmentale notable concerne le statut de la voyelle finale qui est nécessairement prononcée (ce qui exclut les <e#> muets) et différente de <y>. La structure segmentale des emprunts français diffère radicalement de celle des autres, car j'ai constaté qu'aucune formalisation de la structure segmentale ne permettait de rendre compte de toutes les séquences attestées au sein des corpus. Les principales structures segmentales rencontrées sont les suivantes :

<-VCe#> : *gite, confrere, de luxe*.

<-VCCe#> : *coquille, demoiselle, en masse*.

<-VC(C)é(e)#> : *conté, Doré, écarté*.

<-V(V)C#> : *cordon, Delacroix, filet*.

La structure morphologique des emprunts est traditionnellement qualifiée d'opaque, car les locuteurs anglophones dits *lambda* sont dans l'incapacité de repérer les frontières morphologiques présentes au sein de certains items. De plus, les emprunts étant principalement des substantifs, les structures morphologiques attestées sont moins complexes que pour les verbes notamment. Les locuteurs anglais les plus cultivés repèreront sans doute les suffixes italiens tels que *-issimo* attestés dans *fortissimo* et *prestissimo* ou *-ino* et *-ina* présents dans *andantino, concertino* et *Signorina*, de même que les suffixes d'origine espagnole tel que *-ita* dans *Senorita*. Mais de manière générale, les processus morphologiques présents au sein des emprunts et reconnus comme tels par les locuteurs sont considérés comme inexistantes. En

⁶⁰ Il s'agit là bien évidemment d'un squelette segmental prototypique permettant de représenter la structure du plus grand nombre d'emprunts. Ce modèle ne rend évidemment pas compte de toutes les structures rencontrées, d'autant plus que j'ai remarqué lors de ma dernière étude sur l'influence des langues sources que les emprunts aux langues germaniques (exemples : allemand, danois, suédois) présentent des structures segmentales très éloignées de la structure héritée des langues romanes.

revanche, au sein des emprunts français, la présence de terminaisons contraignantes associées à un schéma accentuel final, terminaisons correspondant souvent à des suffixes français (exemples : *-ette*, *-esque*, *-aire*, *-eur*, *-euse*), démontre une structure morphologique complexe et dont les locuteurs anglophones sont conscients. En effet, mon étude de 2011 consacrée à la productivité des suffixes d'origine française *-aire*, *-eur* et *-euse* en anglais a démontré que même si la productivité de tels suffixes est limitée, elle est néanmoins attestée au sein des bases de données dictionnairiques, preuve d'une réelle autonomie morphosémantique au sein du système anglais.

La fréquence constitue le dernier paramètre notable montrant une différence entre les emprunts et le reste du lexique. Mes différentes études, à l'occasion desquelles les données du *COCAE* ont été utilisées, ont montré une fréquence des emprunts très en-deçà du reste du lexique. Ces données fréquentielles concernaient les emprunts français mais il est fort à parier que les données sont sensiblement équivalentes pour le reste des emprunts, même si toutefois, ces arguments sont à prendre avec du recul, si l'on considère la fréquence importante d'emprunts comme *coffee* ou *tomato* en anglais.

Le tableau 7 regroupe d'un côté les spécificités des emprunts français et de l'autre les caractéristiques du reste des emprunts. L'emprunt prototypique en anglais serait donc un mot possédant toutes ces caractéristiques.

Propriétés	Emprunts	Emprunts français
Sémantiques	- Emprunts culturels	- Emprunts culturels - Notion de prestige
Phonologiques	- Accent /(-)10/ préférentiel - Moins de réduction en contexte inaccentué.	- Accent /-1/ préférentiel - Moins de réduction en contexte inaccentué.
Graphie-phonie	- Voyelles libres étrangères - Correspondances consonantiques spécifiques	- Voyelles libres étrangères - Correspondances consonantiques spécifiques - Consonnes finales muettes
Segmentales	- Polysyllabiques - <-VCV#>	- Structures segmentales spécifiques
Morphologiques	- Opaques	- Terminaisons contraignantes - Productivité suffixale
Fréquentielles	- Faibles	- Faibles

Tableau 7 : Propriétés des emprunts français et des autres emprunts observables au sein du système anglais.

Pour en revenir à la définition prototypique de l'emprunt, et si l'on se concentre uniquement sur les emprunts distincts du français, je considère donc que les deux exemples suivants sont les archétypes parfaits réunissant toutes les propriétés définitoires d'un emprunt en anglais :

amigo /ə'mi:ɡəʊ, æ-/ (GB) - /ə'mi:gou, ɑ:-/ (US)

linguini /lɪŋ'ɡwi:ni/ (GB / US)

Ces deux termes sont des emprunts culturels, *amigo* faisant référence à la vie quotidienne espagnole, tandis que *linguini* provient de la tradition culinaire italienne et ils ont une fréquence relativement faible au sein de la langue anglaise. Sur un plan phonologique, ils sont tous les deux accentués sur la syllabe pénultième, schéma accentuel préférentiel des emprunts ; et ils ont des voyelles pleines en contexte inaccentué non final. Si dans *amigo*, il y a variation entre voyelle réduite (/ə/) et voyelle non réduite (/æ/ en anglais britannique et /ɑ:/ en anglais américain), le terme *linguini* atteste seulement une forme pleine (/ɪ/) de la voyelle en contexte inaccentué. Concernant les correspondances graphie-phonie, on constate la présence d'une

voyelle libre étrangère (/i:/) au niveau de la voyelle accentuée. Enfin, au niveau segmental, les deux emprunts sont trisyllabiques et présentent tous deux une structure de type <-VCV#>.

Il est évident que tous les emprunts en anglais ne combinent pas tous les paramètres définitoires de la catégorie. Il faut concevoir le système des emprunts comme un gradient dont le centre du domaine constituerait l'essence même d'un emprunt, concentrant la totalité ou la quasi-totalité des caractéristiques présentées dans le tableau 7. Les autres emprunts adopteraient seulement certaines des caractéristiques, en fonction de leur éloignement par rapport à la représentation prototypique figurant au centre et symbolisée par la croix dans la figure 2 ci-dessous. Cette figure illustre le concept de micro-système et de gradient avec au centre les emprunts les plus représentatifs de la notion et un éloignement progressif par rapport au centre en fonction des caractéristiques adoptées ou non par les items.

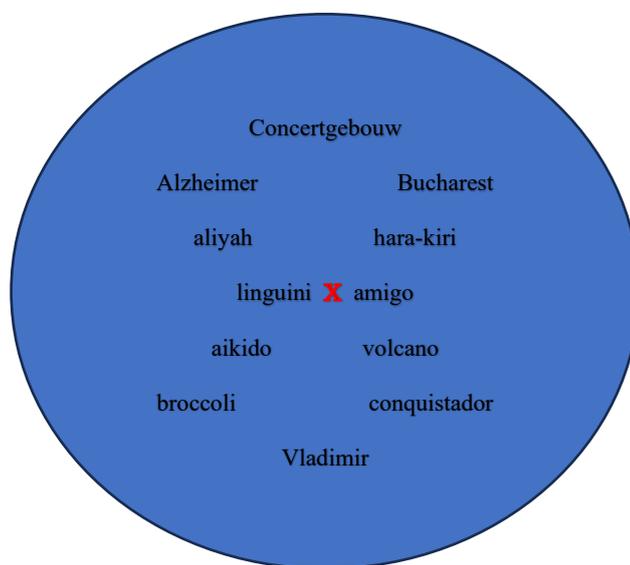


Figure 2 : Conceptualisation du sous-système des emprunts avec des exemples.

Les emprunts *aliyah*, *hara-kiri*, *aikido*, *volcano* et *bodega* s'éloignent légèrement du centre prototypique de la notion, bien qu'ils en respectent la plupart des caractéristiques, et ce pour des raisons différentes. Le terme *aliyah* présente une structure segmentale de type <-CVC#>, distincte de la représentation prototypique en <-VCV#>. Le terme *hara-kiri* présente une structure morphologique complexe probablement héritée du japonais, très différente de la structure morphologique opaque traditionnellement adoptée par les emprunts. L'emprunt *aikido* présente une variation accentuelle /010/-/100/ en anglais britannique et /010/-/201/ en anglais américain. La présence d'une variation accentuelle éloigne ainsi l'item du centre de la notion près duquel figurent les emprunts uniquement accentués selon le schéma préférentiel sur la

pénultième. Le terme *volcano*, quant à lui, ne présente pas de voyelle libre étrangère, ce qui constitue l'une des propriétés phonologiques essentielles de la catégorie des emprunts (/vɒl'keɪnəʊ/ (GB) et /vɑ:l'keɪnəʊ/ (US)).

Les emprunts *broccoli* et *conquistador* me paraissent encore plus éloignés du centre définitoire de l'emprunt. Tout d'abord, ces deux items ont une accentuation antépénultième et de plus *broccoli* (/ˈbrɒkəli/ (GB) et /ˈbrɑ:kəli/ (US)) montre un tel degré de réduction en contexte inaccentué qu'un phénomène de syncope est suggéré par *LPD*. L'emprunt *conquistador* présente également une variante accentuelle sur la syllabe finale, relève d'une structure segmentale de type <-CVC#> et n'atteste la présence d'une voyelle libre étrangère (<i> = /i:/) qu'en variante en anglais américain (et pas du tout en anglais britannique). Lorsque plusieurs spécificités propres aux emprunts ne sont pas respectées, l'emprunt tend à s'éloigner du centre.

Enfin, j'ai mentionné à titre indicatif des emprunts situés en marge de la catégorie, à savoir *Alzheimer*, *Bucharest*, *Concertgebouw* et *Vladimir*. Ces items font référence à une notion culturelle étrangère au sens large et la connotation sémantique étrangère qu'ils véhiculent, associée à une faible fréquence, leur permet d'être considérés comme des emprunts. Ils ne respectent toutefois aucun des autres paramètres définitoires de la catégorie tant au niveau phonologique que graphophonématique ou segmental. Ils sont toutefois perçus comme des emprunts par les locuteurs anglophones en raison de spécificités phonotactiques sur lesquelles je ne me suis pas encore penché. Par exemple, les séquences <lz> dans *Alzheimer* ou <vl> dans *Vladimir* ne correspondent pas à des séquences attestées au sein du lexique anglais. Les locuteurs accorderont donc vraisemblablement un statut d'emprunt à ces items en raison de la transgression des règles phonotactiques du système anglais.

La modélisation du sous-système des emprunts doit être conçue comme un ensemble dynamique à l'intérieur duquel les emprunts se rapprochent ou s'éloignent du centre de l'essence même de la notion. Les raisons pour lesquelles les emprunts respectent plus ou moins les propriétés définitoires de la catégorie tiennent selon moi à leur degré d'assimilation à la langue anglaise, qui peut être corrélé au facteur temporel. Plus un emprunt aura passé de temps au sein du lexique anglais et plus il aura tendance à adopter les propriétés définitoires du système anglais, ce qui se traduit par une assimilation plus ou moins marquée et en perpétuelle évolution. Je ne peux évidemment m'empêcher d'ajouter dès à présent l'hypothèse selon laquelle les langues sources exercent également une influence essentielle sur le sous-système des emprunts. Cela se matérialise par un conflit permanent au sein de ce sous-système entre

assimilation aux règles en vigueur au sein du système anglais et préservation des spécificités linguistiques inhérentes à certaines langues sources.

Il serait tout à fait possible de concevoir de la même façon le sous-système des emprunts français, avec un centre du système correspondant à la somme des paramètres du tableau 7 et qui illustrerait la notion prototypique de « français » au sein du système anglais, tant sur un plan phonologique que graphophonématique, morphologique, sémantique, segmental ou fréquentiel. Je me contenterai simplement de rappeler à ce stade que les propriétés définitoires des emprunts français sont en partie distinctes de celles des emprunts aux autres langues et qu'il convient de les faire relever de deux sous-systèmes spécifiques.

La section suivante présente la dernière étape de mon travail de formalisation de l'ensemble du système anglais.

4.3 Les sous-systèmes de l'anglais

La dernière étape de la formalisation des sous-systèmes de l'anglais est le fruit de ma collaboration avec Q. Dabouis, lequel a également travaillé sur des sujets connexes. La somme de nos travaux nous a permis de présenter une formalisation du système anglais basée sur le principe de sous-systèmes dans une publication datant de 2022⁶¹. Cette publication représente l'aboutissement de nombreuses années de recherche et de réflexions sur les emprunts, mais aussi sur le fonctionnement global du système anglais. Ce travail ne constitue toutefois pas une fin en soi, mais un point de départ vers des projets de recherche illimités.

De la même façon que J. Ito & A. Mester (1995) postulent l'existence de sous-systèmes imbriqués au sein de la langue japonaise, il nous semblait fondamental de distinguer les mécanismes régissant le comportement des emprunts en anglais de ceux permettant de rendre compte de l'organisation du reste du lexique, sans toutefois déconnecter les microsystèmes les uns des autres. Les emprunts fonctionnent en effet comme un sous-système au sein du système anglais, et nous avons réalisé que ce système était en réalité composé de plusieurs sous-systèmes possédant des caractéristiques segmentales, phonologiques, morphologiques, sémantiques, fréquentielles et graphophonématiques propres à chacun. Nous concevons ce modèle comme un système hybride et dynamique, ce qui signifie que même si chaque sous-système possède ses propres caractéristiques, les frontières entre les sous-systèmes sont fluctuantes. Cela montre ainsi que les sous-systèmes partagent certaines propriétés. Les

⁶¹ Dabouis, Q. & P. Fournier (2022), "English PhonologieS?", *Modèles et modélisation en linguistique / Models and Modelisation in Linguistics*, V. Arigne & C. Rocq-Migette (éds.), Peter Lang, 215-258.

sous- systèmes dont nous postulons l'existence sont les suivants (la formalisation du système repose sur une nomenclature qui doit beaucoup aux travaux de E. Carney sur le système graphique de l'anglais (1994)) :

- §*Core*. Il est cependant essentiel de distinguer deux ensembles distincts au sein de ce sous-système :
 - §*Core-Native*.
 - §*Core-Latinate*.
- §*French* ;
- §*Learned* ;
- §*Foreign*.

Le sous-système §*Core* correspond au cœur du stock lexical de l'anglais, et pour des raisons historiques, il se subdivise en deux ensembles : l'ensemble §*Core-Native* constitué du vocabulaire d'origine germanique et l'ensemble §*Core-Latinate* composé du vocabulaire d'origine romane. Le sous-système §*French* regroupe tous les mots d'origine française qui ont conservé des propriétés linguistiques spécifiques aux mots d'origine française. Ces propriétés peuvent être de nature diverse (morphologique, phonologique, lexicologique, sémantique, etc.) et elles peuvent se combiner. Toutefois, les items présents dans ce sous-système sont différents des mots d'origine française totalement assimilés par le système anglais et qui figurent donc au sein du sous-système §*Core*. Les emprunts français peuvent transiter du sous-système français vers les sous-systèmes natifs, ce qui démontre le caractère dynamique du modèle. Le sous- système §*Learned* regroupe le lexique savant. Il faut concevoir ce sous-système comme le principal pourvoyeur de vocabulaire scientifique et technique, avec des processus morphologiques plus ou moins complexes permettant de générer du lexique. Des items comme *horizon* ou *papyrus* appartiennent à ce sous-système savant, de la même façon que les composés savants dont les éléments constitutifs sont d'origine classique. Les composés savants sont soit empruntés directement aux langues classiques, soit construits au sein du système par l'intermédiaire des matrices lexicogéniques à l'œuvre en anglais. Enfin, le sous-système §*Foreign* regroupe les mots d'origine étrangère (autre que classique), c'est-à-dire les emprunts. Si la formulation même de la règle des mots « italiens » montre une représentation prototypique de l'emprunt avec une structure segmentale en <-VCV#> et une origine essentiellement romane, ce sous-système regroupe tous les mots étrangers et perçus comme tels par les locuteurs anglophones qui leur attribuent des caractéristiques linguistiques différentes des autres sous-systèmes. Ainsi, le sous-système §*Foreign* est composé de mots provenant de toutes les langues auxquelles l'anglais a emprunté.

Il n'est pas utile de refaire ici l'inventaire de tous les paramètres spécifiques aux sous-systèmes §*Foreign* et §*French*, car la section 4.2 a déjà mis en évidence les prototypes même des emprunts d'origine française et des emprunts aux autres langues, tout en montrant l'interaction entre les différents niveaux linguistiques. Je mentionnerai à titre d'illustration les propriétés définitoires du sous-système §*Learned* pour montrer que, si certains paramètres sont partagés par plusieurs sous-systèmes, la combinaison de tous les paramètres fait de chaque sous-système un ensemble unique.

Le sous-système §*Learned* possède les caractéristiques suivantes :

- niveau sémantique : vocabulaire scientifique ou technique ;
- niveau morphologique : pluralité de processus morphologiques dont la composition savante (interaction entre bases et affixes) et nombreuses bases savantes attestées au sein de ce sous-système ;
- niveau phonologique : accentuation antépénultième majoritaire, même si on constate une forte proportion d'accentuation pénultième (héritage des processus phonologiques des langues classiques) et une faible proportion de voyelles réduites en contexte inaccentué, surtout au sein de la composition savante (les bases non accentuées tendent à conserver une voyelle pleine) ;
- correspondances spécifiques au niveau graphophonématique :
 - au niveau vocalique : présence de voyelles libres étrangères tant pour les voyelles monographe (voir tableau 6) que pour les voyelles digraphes (exemples : <ae> et <oe> = /i:/ dans *Caesar* et *foetus*).
 - au niveau consonantique : <ph> = /f/ dans *phanton* ou *sphere*, <ch> = /k/ dans *chaos* ou *stomach*, <rh> = /r/ dans *rhetoric* ou *rhythm*.
- niveau fréquentiel : très faible productivité.
- niveau segmental : mots en général polysyllabiques (composition savante) et structure <-CVC-> fréquemment attestée notamment au sein des bases savantes (exemples : *-log-*, *-taph-*, *-vor-*, *-phil-*).

La liste des paramètres linguistiques utilisés pour caractériser chaque sous-système ne se veut en aucun cas exhaustive. Il nous a toutefois semblé que ces paramètres étaient les plus pertinents pour permettre une telle formalisation en sous-systèmes. Je ne ferai pas ici une énumération des paramètres spécifiques au sous-système §*Core*, dont les données proviennent essentiellement des travaux de Q. Dabouis. Le tableau regroupant la liste de tous les paramètres pertinents pour caractériser chaque sous-système avec pour chaque paramètre une série d'exemples est disponible dans notre publication de 2022 (2022 : 236).

La mise en évidence des paramètres jouant un rôle essentiel dans la définition des sous-systèmes s'appuie essentiellement sur des données dictionnaires. Les dictionnaires représentent l'état de la langue à un moment donné et ne sont pas représentatifs de tous les locuteurs anglophones. De la même façon, les études basées sur les productions de locuteurs anglophones ne peuvent prétendre être représentatives du fonctionnement global de la langue anglaise. Nous avons beaucoup réfléchi à la dichotomie langue - parole et à la distinction dictionnaire - locuteurs au moment de procéder à la représentation dynamique des sous-systèmes. Le modèle que je vais présenter ci-dessous est représentatif du fonctionnement du système. Il faut concevoir chaque sous-système comme la somme de paramètres provenant de différents niveaux linguistiques (phonologiques, morphologiques, sémantiques, etc.) qui sont autant d'outils dont les locuteurs de l'anglais disposent. La maîtrise des paramètres relatifs à chaque sous-système est évidemment corrélée au niveau d'instruction de chaque locuteur. Plus un locuteur anglophone est instruit et plus il sera en mesure de maîtriser les différents niveaux linguistiques de chaque sous-système. Il n'y a toutefois pas de bonne ou de mauvaise façon d'utiliser les outils de la langue et je reviens vers la strate phonologique, puisque ma synthèse repose en grande partie sur l'accentuation lexicale. Il y a la norme et les différentes réalisations des locuteurs qui s'écartent plus ou moins de cette norme. Par conséquent, si un locuteur peu instruit ne maîtrise pas les spécificités phonologiques du sous-système des emprunts qui lui permettraient de réaliser ces items de la façon préconisée par les dictionnaires de prononciation, il est fort probable qu'il utilise les outils d'un sous-système qu'il maîtrise (vraisemblablement ceux du sous-système § *Core*), pour prononcer les mots d'origine étrangère. Cependant, les résultats de mes tests oraux menés sur les mots italiens ont montré que les locuteurs anglophones, quel que soit leur niveau d'instruction, semblent avoir une représentation phonologique stéréotypée des emprunts.

Enfin, il me semblait essentiel d'intégrer les langues sources au sein de ce modèle dynamique, car les résultats de mes travaux ont démontré que les langues sources exerçaient une influence non négligeable sur les structures phonologiques et graphophonématiques⁶² (position de l'accent lexical et réalisation des voyelles accentuées). La maîtrise des paramètres des sous-systèmes repose, comme je l'ai mentionné précédemment, sur le niveau d'instruction des locuteurs. Il est fort probable que les emprunts intègrent la langue anglaise par l'intermédiaire de locuteurs bilingues ou extrêmement cultivés, comme tout lexicographe serait l'être. On imagine aisément les lexicographes des maisons d'édition Cambridge, Longman ou

⁶² Je n'ai pour l'instant pas testé l'influence des langues sources sur les autres niveaux linguistiques constitutifs du sous-système des emprunts.

Oxford maîtriser les rudiments de plusieurs langues à défaut d’être polyglottes. L’influence des langues sources sur le paramétrage des sous-systèmes se manifeste par le biais des locuteurs bilingues. Le modèle proposé ci-dessous matérialise donc le dynamisme interactionnel entre les sous-systèmes et l’influence que peuvent exercer les langues sources sur les sous-systèmes suivants : §*Learned*, §*French* et §*Foreign*.

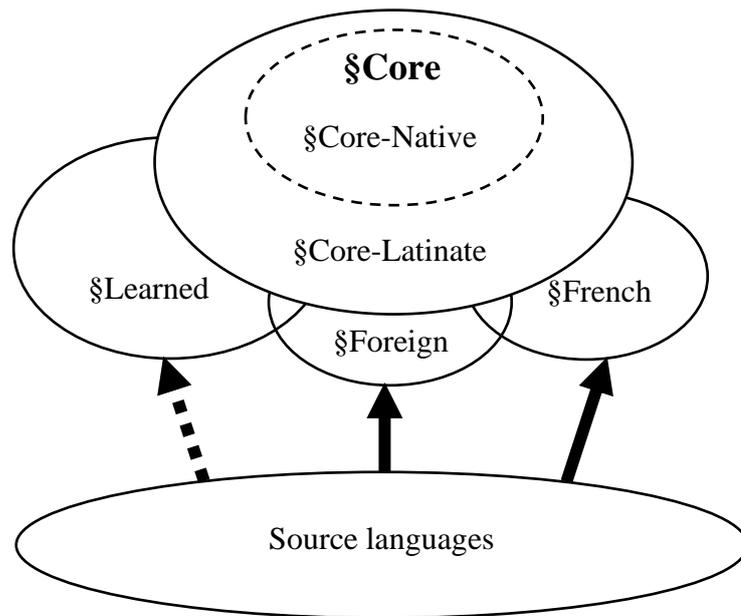


Figure 3 : Représentation des sous-systèmes de l’anglais (2022 : 248).

L’influence des langues sources sur les sous-systèmes §*Foreign* et §*French* est clairement établie et les résultats présentés dans cette synthèse le prouvent. Les langues sources ont un impact déterminant sur la prononciation des mots d’origine étrangère en anglais. En revanche, l’influence des langues sources sur le sous-système §*Learned* est certes réelle, mais moins directe car en fait de langues sources, il s’agit essentiellement du grec et du latin. De plus, de nombreux termes appartenant au sous-système savant sont en réalité construits en anglais à partir de bases d’origine classique. C’est la raison pour laquelle nous avons opté pour une flèche en pointillés, qui traduit une influence moins directe.

La conception des sous-systèmes matérialise le dynamisme et l’interaction au sein du modèle. Les frontières entre les différents sous-systèmes sont fluctuantes et les chevauchements montrent que même s’ils sont distincts les uns des autres, ils partagent toutefois certaines caractéristiques. Ces chevauchements signifient donc que certains items sont à la frontière entre

deux sous-systèmes car ils ont des caractéristiques hybrides communes aux deux. Les sous-systèmes §*Foreign* et §*French* sont connectés, de même que les sous-systèmes §*Foreign* et §*Learned*. En revanche, rien ne laisse supposer dans les données recueillies que les sous-systèmes §*Learned* et §*French* le soient.

Le point central de chaque sous-système concentre toutes les propriétés définitoires. Plus un item s'éloigne du centre du sous-système, parce qu'il concentre moins de paramètres définitoires, plus il arrive en marge du sous-système et se prête à une arrivée possible dans la zone de chevauchement avec un autre sous-système, et même, à terme, à un passage définitif dans un autre sous-système. Les cas de transition des sous-systèmes §*Learned*, §*Foreign* et §*French* vers le sous-système §*Core* sont les plus fréquents et reposent sur le principe d'assimilation. Plus un item est assimilé au sein du lexique anglais, plus il combine de propriétés linguistiques propres au cœur du stock lexical. En effet, les cas de transition du sous-système §*Core* vers l'un des autres sous-systèmes semblent peu probables.

Les transitions et les chevauchements soulignent l'importance du facteur temporel au sein du système. Mais pour autant, réduire le dynamisme du système à ce facteur temporel me semble problématique, car certains emprunts français attestés dans les dictionnaires depuis plusieurs siècles présentent toujours les spécificités du sous-système §*French*, et ne semblent pas en voie d'intégrer le sous-système §*Core*. De plus, j'ai montré lors de cette synthèse que, si l'accentuation des emprunts français était sensible à la variation dialectale, ce n'était absolument pas le cas des emprunts aux autres langues. Il y a donc des facteurs plus ou moins déterminants selon le sous-système donné qui ont un impact direct sur la matérialisation des zones de chevauchement. Enfin, ce modèle représente un fonctionnement du système, mais cette représentation en sous-systèmes fluctue selon le niveau d'instruction des locuteurs. Comme nous l'avons mentionné précédemment, les locuteurs n'ont pas tous le même niveau d'instruction, et par conséquent, la même maîtrise des spécificités propres à chaque sous-système. Il existe donc également des phénomènes de variation et de fluctuation imputables aux locuteurs.

L'architecture du modèle illustre le conflit permanent entre deux logiques. La structuration du modèle en sous-systèmes démontre l'autonomie de chaque sous-système, mais aussi le potentiel d'interaction entre les sous-systèmes. Toutefois, les langues sources exercent une influence fondamentale sur les éléments d'origine étrangère au sein du système. La terminologie utilisée pour qualifier les sous-systèmes illustre l'importance et l'influence des éléments d'origine étrangère au sein du système.

Bilan et perspectives

De mon initiation à la recherche au sein de l'École de Guierre à mes travaux actuels sur les emprunts, je n'ai eu de cesse, lors de mon parcours, de mettre en évidence les mécanismes qui organisent la logique du système accentuel de l'anglais. Après avoir travaillé sur le processus de suffixation par l'intermédiaire des adjectifs suffixés en *-ous*, j'ai ensuite orienté mes recherches vers le domaine de la composition savante, et dorénavant l'essentiel de mes recherches porte sur la prononciation des emprunts. La structure de la synthèse a montré mon évolution tant au niveau théorique qu'au niveau méthodologique. Ma compréhension globale des faits accentuels s'est développée pour intégrer une dimension interlangue et élaborer notamment des réflexions théoriques portant sur les contacts entre langues et les phénomènes linguistiques se produisant à l'interface entre langue source et langue cible. La méthodologie adoptée lors de ma thèse s'est enrichie et certains de mes travaux prennent désormais en compte la dimension diachronique des faits accentuels, proposent des comparaisons inter-dictionnairiques et sont également orientés vers l'analyse de données orales par l'intermédiaire de tests sur des locuteurs anglophones ou sur les quotatifs en interaction, comme lors de ma collaboration avec Y. Fuchs.

Mes perspectives de recherche sont donc plurielles et je synthétise mes futurs projets dans les paragraphes suivants.

Je souhaite en premier lieu poursuivre mes travaux relatifs à l'accentuation des composés savants. L'accentuation de telles structures repose sur des critères morphosémantiques et les éléments d'origine classique y jouent un rôle essentiel, car certaines propriétés phonologiques des langues classiques relatives à l'accentuation lexicale sont reproduites en anglais. Ces composés appartiennent au sous-système savant et j'ai formulé dans la synthèse l'hypothèse selon laquelle leur prononciation serait corrélée au degré d'instruction des locuteurs. Plus un locuteur serait instruit et plus il prononcerait ces mots savants de la manière préconisée par les dictionnaires anglais de prononciation. La mise en place d'un protocole expérimental évaluant la corrélation entre degré d'instruction et prononciation des mots savants relevant de discours de spécialité dans des productions de locuteurs anglophones serait particulièrement révélatrice.

La formalisation en sous-systèmes proposée dans la synthèse constitue un nouveau modèle théorique permettant de mieux rendre compte des spécificités des différentes catégories d'items rencontrés en anglais. Cette modélisation constitue plus une ouverture sur d'autres projets qu'un réel aboutissement. Plusieurs de mes travaux soumis à des revues spécialisées s'inscrivent directement dans ce cadre théorique et seront publiés en 2024. Après plusieurs

allers-retours très formateurs tant sur un plan théorique que méthodologique, l'article « Stress in French loanwords in British and American English », écrit en collaboration avec Q. Dabouis, sera publié en 2024 dans la revue *Journal of Linguistics*. Il montre que les mots appartenant au sous-système §*French* possèdent des caractéristiques spécifiques qui les distinguent des autres, et nous mettons en évidence dans cette étude toutes les propriétés linguistiques des mots d'origine française. Nous démontrons ensuite que l'accentuation des emprunts français est régie par des paramètres segmentaux (nombre de syllabes et poids syllabique) et morphologiques (terminaisons contraignantes et mots simples vs mots complexes) au moyen d'une analyse statistique. Nos résultats prouvent également que les paramètres déterminants dans la dynamique accentuelle des emprunts français sont différents en anglais britannique et américain, ce qui avait été suggéré par le passé, mais n'avait jamais été prouvé par un modèle de régression statistique.

Une deuxième publication inscrite dans les mêmes perspectives de recherche verra le jour en 2024 dans la revue *Lexis*. Elle s'intitule « The impact of source languages on the stressing of loanwords in English » et j'ai décidé de l'insérer en tant qu'inédit en annexe de cette synthèse, car les résultats en sont significatifs. J'y démontre l'impact des langues sources sur l'accentuation des emprunts correspondants en anglais, en comparant de manière systématique la position de la syllabe accentuée dans les mots provenant de 19 langues sources et dans les emprunts en anglais. Mon analyse est basée sur les données phonétiques extraites du dictionnaire *Routledge* (2017). Le taux de correspondance accentuelle très élevé démontre de manière irréfutable que la préservation de l'accentuation des langues sources réalisée par l'intermédiaire de locuteurs bilingues constitue un paramètre fondamental permettant de rendre compte de la logique accentuelle des emprunts en anglais. Cela prouve également par extension que les langues sources ont une influence déterminante sur la modélisation des sous-systèmes de l'anglais, ce que la formalisation du modèle présenté dans la quatrième partie laissait déjà suggérer.

Mon étude sur l'accentuation des emprunts arabes en anglais est en cours de finalisation et les résultats de la concordance accentuelle entre l'arabe et l'anglais sont prometteurs. Une telle correspondance n'était pas anticipée, car les deux systèmes phonologiques me semblaient trop différents pour permettre une telle concordance. En conséquence, je vais collaborer très prochainement avec des phonologues spécialisés dans la phonétique et la phonologie de l'arabe pour pousser plus loin mes investigations.

De manière plus générale, je compte poursuivre mes travaux dans le domaine de l'accentuation lexicale des emprunts en anglais. Je finalise actuellement une base de données

d'emprunts extraits des données de *LPD3* (3000 items environ) en sélectionnant tous les mots d'origine étrangère répertoriés comme tels par le dictionnaire (à l'exception des emprunts français). Ce travail titanesque d'extraction inclut également une phase de codage des facteurs déterminants au niveau accentuel, ce qui me permettra très prochainement de tester, à l'aide d'un modèle de régression statistique, l'impact de divers paramètres potentiellement déterminants sur l'accentuation des emprunts, dont celui de la conservation du schéma accentuel de la langue d'origine. Cette base de données facilitera également mes travaux sur le comportement accentuel spécifique de certaines familles de langues. En effet, j'ai constaté lors de mes différents travaux que les emprunts d'origine russe avaient un comportement accentuel totalement atypique en anglais, que l'hypothèse du principe de conservation de la structure prosodique du russe ne permettait pas de justifier. De la même façon, les emprunts au chinois constituent une classe d'emprunts présentant des spécificités accentuelles particulières en anglais, avec des séquences de syllabes accentuées qui sont traditionnellement présentées comme des violations manifestes des principes organisant le système phonologique de l'anglais. Ces travaux sur l'accentuation des emprunts provenant de différentes langues sources ne seront toutefois pas déconnectés les uns des autres sur un plan théorique. J'aspire en effet, après avoir mis en évidence les caractéristiques accentuelles de chaque famille de langues, à établir une typologie des langues dont le critère de classification serait le taux de reproduction de l'accentuation d'origine. J'ai démontré dans ma synthèse que si les paramètres prosodiques de certaines langues sources ont une influence sur l'accentuation des emprunts en anglais en raison d'une compatibilité des systèmes prosodiques (exemple : l'italien), certaines langues ne permettent pas un tel transfert de propriétés phonologiques (exemple : le japonais). Ce projet de recherche très ambitieux nécessiterait la mise en place de collaborations avec des linguistes appartenant à d'autres structures de recherche européennes et des locuteurs natifs des différentes langues sources et pourrait faire l'objet de projets financés.

La rédaction de cette synthèse m'a également permis de réactiver certaines pistes de réflexion et notamment celle relative à l'interaction entre les différents niveaux linguistiques constitutifs des sous-systèmes. Un sous-système est conçu comme une somme de propriétés morphologiques, phonétiques, phonologiques, lexicologiques, morphologiques, segmentales, sémantiques, graphophonématiques, fréquentielles, syntaxiques et pragmatiques caractéristiques de cet ensemble. Ces différents niveaux linguistiques en interaction au sein des sous-systèmes sont autant de perspectives de collaborations futures avec des chercheurs travaillant dans ces domaines. Un travail sur le sous-système des emprunts en collaboration avec des chercheurs représentant chacune des disciplines mentionnées ci-dessus serait

bénéfique pour la compréhension globale du système et sur les interactions entre ces domaines. Je souhaite travailler notamment sur la représentation prototypique de l'emprunt en anglais. La façon dont j'ai présenté le sous-ensemble des emprunts dans la section 4.2 de la synthèse est un premier pas dans cette voie. Le cœur du sous-système des emprunts concentre les paramètres linguistiques définitoires de la catégorie. Je souhaite travailler sur la perception qu'ont les locuteurs anglophones de la notion d'emprunt et sur la représentation prototypique de l'emprunt en anglais. J'ai formulé dans cette synthèse de nombreuses pistes de réflexion concernant les caractéristiques prototypiques des emprunts en anglais, mais n'étant pas spécialiste de tous les domaines évoqués, je souhaiterais travailler sur la réalisation d'un squelette prototypique en collaboration avec des spécialistes des autres domaines. La création d'une forme prototypique de l'emprunt en anglais repose inmanquablement sur la notion de gradient. Jusqu'à quel point un emprunt est-il considéré comme tel par les locuteurs anglophones et sur quels critères ? Ce travail fera ainsi émerger les formes prototypiques, de même que celles plus éloignées du centre et concentrant ainsi moins de paramètres définitoires. Les formes en marge du système correspondront donc aux emprunts regroupant le moins de critères constitutifs de cette classe d'items.

Le modèle théorique des sous-systèmes de l'anglais tel que je l'ai ébauché est basé sur des données relatives à l'anglais contemporain. J'ai mentionné précédemment que les mots relevant d'un sous-système pouvaient au fil du temps atteindre les frontières de ce sous-système, voire transiter vers un sous-système voisin. Il serait intéressant de déterminer les parties du lexique concernées par ce phénomène de transition au moyen d'une étude diachronique sur l'accentuation dans les dictionnaires de prononciation du 17^e siècle jusqu'à aujourd'hui. De plus, si le temps est un facteur déterminant dans la délimitation des frontières et des zones de chevauchement entre sous-systèmes, d'autres facteurs exercent-ils également une influence ?

Enfin, j'ai eu de nombreux débats avec des collègues phonologues concernant le mode d'intégration des emprunts dans le lexique anglais. Si je suis fondamentalement persuadé, car mes données confirment cette hypothèse, que les emprunts intègrent le lexique anglais et deviennent ainsi des formes institutionnalisées au sein des dictionnaires de prononciation par l'intermédiaire de locuteurs bilingues, d'autres collègues penchent vers l'hypothèse de locuteurs cultivés, mais pas nécessairement bilingues. Ce point est d'une importance capitale dans le cadre théorique de la *Loan Phonology* et les deux positions s'affrontent. Des protocoles expérimentaux mis en place sur des locuteurs bilingues et non bilingues testant la corrélation entre le degré de bilinguisme et la capacité des locuteurs à reproduire en anglais les phénomènes prosodiques attestés dans les langues sources permettraient de trancher.

Mes projets de recherche dans le domaine de l'accentuation lexicale des emprunts et des sous-systèmes de l'anglais permettent d'envisager de nombreuses collaborations. De plus, il s'agit d'un champ disciplinaire encore peu exploré en phonétique et phonologie de l'anglais dans les universités françaises, et l'encadrement de travaux de recherche dans ce domaine permettrait de le développer.

J'ai conscience que les fonctions auxquelles j'aspire, celles de Professeur des Universités, impliquent de nouvelles tâches en termes d'animation et de direction de la recherche. Je m'y suis préparé en devenant directeur-adjoint de l'axe 1 « Marges, Inégalités, Vulnérabilités » du laboratoire Pléiade en septembre 2023. Ce travail au sein d'une équipe pluridisciplinaire va me permettre de développer cet aspect essentiel de pilotage de la recherche et d'élargir encore mes perspectives théoriques au contact de chercheurs provenant d'autres domaines des sciences humaines. L'organisation de manifestations scientifiques de dimension internationale constitue également l'une des tâches essentielles de la fonction de Professeur des Universités, car elle permet de mettre sur le devant de la scène des thématiques pouvant générer des vocations chez de jeunes chercheurs. Il est donc essentiel de cultiver cet aspect de la recherche en organisant des colloques et des journées d'études en collaboration avec d'autres unités de recherche. J'ai déjà acquis une solide expérience dans l'organisation de manifestations scientifiques, notamment grâce à la co-organisation du Colloque de Villeteuse sur l'anglais oral. Enfin, l'un des aspects fondamentaux de la fonction de Professeur des Universités consiste à encadrer des thèses. Je me suis préparé à cette tâche en encadrant de nombreux mémoires de recherche au sein du Master Ledilangt. Le fait d'exercer dans une université telle que celle de Sorbonne Paris Nord, qui autorise les Maîtres de Conférences à encadrer des mémoires de recherche, même sans HDR, m'a sensibilisé aux problématiques d'encadrement et d'accompagnement intellectuel des jeunes chercheurs. De plus, mes activités d'expertise auprès de revues scientifiques telles que *Anglophonia* ou *Lexis* et ma participation à plusieurs comités de sélection en tant que membre interne ou externe m'ont permis de cultiver une rigueur scientifique indispensable pour encadrer les travaux de jeunes chercheurs.

Les tâches administratives occupent également une place prépondérante dans l'activité des Professeurs des Universités. Je me suis beaucoup investi dans l'aspect administratif depuis mes débuts en tant que Maître de Conférences à l'Université Sorbonne Paris Nord. J'ai exercé de nombreuses fonctions telles que celles de directeur de département, co-responsable de la double licence anglais-espagnol, co-responsable de la confection des emplois du temps ou encore responsable de la coordination des enseignements en phonétique et en anglais oral. Le fait d'exercer dans une UFR pluridisciplinaire et d'assurer très tôt de nombreuses tâches

administratives m'a permis de prendre rapidement conscience du fonctionnement de l'université et de la réalité du travail d'enseignant-chercheur.

Je pense avoir pu démontrer tout au long de cette synthèse que j'avais acquis un recul critique par rapport à la recherche en général, mais surtout par rapport à la mienne, en n'hésitant pas par exemple à infirmer les résultats de mes études précédentes. J'ai évoqué les dilemmes langue / parole ou dictionnaires de prononciation / locuteurs, ce qui montre que je suis tout à fait conscient des limites de mon propre positionnement théorique préférentiel. Il est en effet pour moi essentiel de croiser les différentes approches relatives aux faits phonétiques et phonologiques, ce qui, je l'espère, conduira à nouveau à de nombreuses collaborations fructueuses au niveau de la recherche et de l'encadrement. Ce recul me permet légitimement d'envisager un accompagnement approprié de doctorants, et même de ceux qui émargeraient à des domaines un peu plus éloignés de mes sujets de prédilection, ou s'inscriraient dans une approche théorique différente. Je compte ainsi mettre à profit l'expérience acquise en tant que Maître de Conférences au sein de l'Université Sorbonne Paris Nord pour développer les thématiques de recherche abordées dans cette synthèse, mettre en valeur la recherche notamment par l'intermédiaire d'organisation de manifestations scientifiques, et encadrer de jeunes chercheurs.

Références

- Abasq, V., Fournier, J.-M., Fournier, P. & M. Martin (2009), “A Dictionary Database of Contemporary English”, poster présenté au PAC Workshop : “Models, Variation & Phonological Corpora », Université d’Aix-en-Provence (10-11 septembre).
- Abasq, V., Dabouis, Q., Descloux, E., Fournier, J.-M., Fournier, P., Girard, I., Martin, M. & S. Vanhoutte (2012), “Stress in Prefixed Disyllabic Verb / Noun Pairs”, poster présenté lors du 20^e Manchester Phonology Meeting (MFM) (24-26 mai).
- Abasq, V., Dabouis, Q., Descloux, E., Fournier, J.-M., Fournier, P., Girard, I., Martin, M. & S. Vanhoutte (2012), “Multicategorical Prefixed Words Stress Behaviour: Variation and Frequency”, communication présentée lors du Colloque International PAC 2012, “The Phonology of Contemporary English: Variation and Change”, Université Toulouse II-Le Mirail (29 février-2 mars).
- Andersen, G. (2000), “The role of the pragmatic marker like in utterance interpretation”, in Andersen, Gisle & Fretheim, Thorstein (éds.), *Pragmatic Markers and Sociolinguistic Variation: A Relevance-Theoretic Approach to the Language of Adolescents*, Amsterdam & Philadelphia: John Benjamins, 17-38.
- Arrizi, A. & P. Robert (1999), *Le Robert & Signorelli. Dictionnaire français-italien / italien-français*, Le Robert.
- Ballier, N. (2013), « Approche prosodique de la cohésion en anglais : le cas de AND en texte lu », Deschamps, A., Trevian, I. & Zumstein, F. (éds.), Actes du 11^e Colloque d’avril sur l’anglais oral (2002), Paris : Les Langues Modernes, Diffusion APLV, 9-22.
- Bauer, L. (1983), *English word-formation*, Cambridge University Press.
- Buchstaller, I. & I. Van Alphen (éds.) (2012). *Quotatives*, Amsterdam & Philadelphia: John Benjamins.
- Burzio, L. (1994), *Principles of English Stress*, Cambridge University Press.
- Calabrese, A. & L. Wetzels (éds.) (2009), *Loan phonology*, John Benjamins Publishing Company.
- Carney, E. (1994), *A Survey of English Spelling*, London: Routledge.
- Carr, P. (1999), *English Phonetics and Phonology: An Introduction*, London: Blackwell.
- Castanier, J. (2016), *L’évolution accentuelle du lexique anglais contemporain appréhendée à travers les dictionnaires de prononciation (XVIII^e - XXI^e siècles)*, Thèse de doctorat, Université de Poitiers.

- Chadelat, J.-M. (2000), *Valeurs et fonctions des mots français en anglais à l'époque contemporaine*, Paris : L'Harmattan.
- Chaou, A. (2021), *Stress assignment in Portuguese loanwords in contemporary English*, Mémoire de Master 2 recherche, Université Sorbonne Paris Nord.
- Chomsky, N. & M. Halle (1968), *The Sound Pattern of English*, New-York: Harper & Row.
- Church, K. (1985), "Stress Assignment in Letter-to-Sound Rules for Speech Synthesis", *The Journal of the Acoustical Society of America*, 78 (S1).
- Dabouis, Q. & P. Fournier (2022), "English PhonologieS?", *Modèles et modélisation en linguistique / Models and Modelisation in Linguistics*, V. Arigne & C. Rocq-Migette (éds.), Peter Lang, 215-258.
- Dabouis, Q., Fournier, J.-M., Fournier, P. & M. Marjolaine (2023), "English Word Stress and the Guierrian School", *New Perspectives on English Word Stress*, E. Yamada, A. Przewozny, J.-M. Fournier & N. Ballier (éds.), Edinburgh University Press, 53-82.
- Dabouis, Q. & P. Fournier (2024), "Stress in French Loanwords in British and American English", *Journal of Linguistics*, 1-26.
- Dailey O'Cain, J. (2000), "The sociolinguistic distribution and attitudes towards focuser like and quotative like", *Journal of Sociolinguistics* 4, n°1, 60-80.
- Davis, S., Tsujimara, N. & J. Tu (2012), "Toward a taxonomy of loanword prosody", *Catalan Journal of Linguistics* 11, 13-39.
- Deschamps, A. (1994), *De l'écrit à l'oral et de l'oral à l'écrit. (Phonétique et orthographe de l'anglais)*, Paris : Ophrys.
- Deschamps, A., Duchet, J.-L., Fournier, J.-M. & M. O'Neil (2004), *English Phonology and Graphophonemics*, Paris : Ophrys.
- Descloux, E., Fournier, P., Martin, M. & S. Vanhoutte (2011), "A Dictionary Database of Contemporary English: When the Tool Meets the Needs. The Particular Case of Word Stress Assignment in Disyllabic Verbs", poster présenté lors du 19^e Manchester Phonology Meeting (MFM) (19-21 mai).
- Descloux, E., Fournier, P., Martin, M. & S. Vanhoutte (2011), « Les mots français en *-eur* (*-euse*) / *-aire* en anglais contemporain : emprunt et création lexicale », *Actes du Colloque international - Les emprunts lexicaux au français dans les langues européennes*, M. Iliescu et alii (éds.), Craïova : Universitaria, 145-162.
- Di Cristo, A. & D. Hirst (1993), « Rythme syllabique, rythme mélodique et représentation hiérarchique de la prosodie du français », *Travaux de l'Institut de Phonétique d'Aix* 15, 9- 24.

- Duchet, J.-L. (1994), *Code de l'anglais oral*, Paris : Ophrys.
- Fournier, J.-M. (1990), *Etude linguistique des adjectifs anglais en -ic*, Thèse de doctorat, Université de Paris VII.
- Fournier, J.-M. (1998), « Que contraignent les terminaisons contraignantes ? », *Topiques, Nouvelles recherches en linguistique anglaise*, Travaux XCIII du CIEREC, Publications de l'Université de Saint-Etienne.
- Fournier, J.-M. (2007), "From a Latin Syllable-Driven Stress System to a Romance Versus Germanic Morphology-Driven Dynamics: in Honour of Lionel Guierre", in P. Carr & P. Honeybone (éds.), *English Phonology, Language Sciences* 29, 218-236.
- Fournier J.-M., Fournier, P., Girard, I. & M. Martin (2010), "Structure, Variation, Usage and Corpora: The Case of Word Stress Assignment in Disyllabic Verbs", communication présentée lors du PAC Workshop: "The Phonology of English: Usage, Varieties and Structure", Université Montpellier III (13-14 septembre).
- Fournier, P. (2010), « Le suffixe *-ous* en anglais contemporain : à la croisée des principes accentuels », communication effectuée lors des Journées de l'École Doctorale SHS de l'Université de Tours (3 juin).
- Fournier, P. (2011), *Accentuation et prononciation des suffixés en -ous en anglais contemporain*, Thèse de doctorat, Université de Tours.
- Fournier, P. & S. Vanhoutte (2013), "Stress in Japanese Loanwords in English: Faithfulness or Adaptation?", poster présenté lors de l'International Conference on Phonetics and Phonology (ICPP2013), National Institute for Japanese Language and Linguistics (NINJAL), Tachikawa (Tokyo) (25-27 janvier).
- Fournier, P. (2013), « Accentuation des composés savants en anglais contemporain : identification morphosémantique et enjeu phonologique », *CORELA* - Numéro 2, Volume 11 (<https://journals.openedition.org/corela/2971>).
- Fournier, P. (2013), « L'accentuation des emprunts français en anglais britannique et américain », *Annales de l'Université de Craïova - Série langues et littératures romanes*, AN XVII, Nr.1, Craïova : Universitaria, 105-129.
- Fournier, P. (2016a), « L'accentuation des emprunts français dans les dictionnaires anglais de prononciation », *Proceedings of the XVII EURALEX International Congress - Lexicography and Linguistic Diversity*, T. Margalitadze & G. Meladze (éds.), Tbilisi: Ivane Javakhishvili Tbilisi State University, 424-435 (<http://euralex2016.tsu.ge/publication.html>).
- Fournier, P. (2016b), « Nouvelles perspectives sur l'accentuation des emprunts en anglais contemporain », *Itinéraires*, 2015-2, *Stumbling blocks. Entraves et obstacles aux*

- circulations. Approches pluridisciplinaires*, M. Lévêque & C. Parfait (éds.) (<http://itineraires.revues.org/2827>).
- Fournier, P. (2017), « Vers une tentative d'uniformisation du traitement accentuel des composés savants suffixés en anglais contemporain », *Anglophonia* 23, Nouvelles frontières de la grammaticalité (<http://journals.openedition.org/anglophonia/1210>).
- Fournier, P. (2018), “Stress Assignment in Italian Loanwords in English and its Impact on the Stressing of Foreign Words by Native English Speakers”, *Corela* HS-24, *Multicultural Spoken English* (<https://journals.openedition.org/corela/5113>).
- Fournier, P. & I. Girard (éds.) (2018), *Multicultural Spoken English*, *Corela* HS-24 (<https://journals.openedition.org/corela/5084>).
- Fournier, P., Herment, S. & H. Le Priault (éds.) (2019), *Voice Quality in English*, *Anglophonia* 27 (<https://journals.openedition.org/anglophonia/1869>).
- Fournier, P. & R. Latrache (2020), “Arabic Loanwords in English: A Lexicographical Approach”, *Proceedings of XIX EURALEX Congress: Lexicography for Inclusion*, Vol. I, Z. Gavriilidou, M. Mitsiaki & A. Fliatouras (éds.), Democritus University of Thrace, 343- 349.
- Fournier, P. (2021), « Transcription et formalisation de l'accentuation des mots « italiens » en anglais contemporain », *Transcrire, écrire, formaliser 2*, F. Toupin & C. Collin (éds.), Collection des Travaux Linguistiques du CerLiCO : Presses Universitaires de Rennes, 53- 68.
- Fournier, P. (2021), « Les emprunts au gaélique écossais en anglais », *L'Ecosse, la différence - Scotland, the Difference*, S. Juillet Garzon, P. Fournier & A. Fiasson (éds.), Série Caledonia : Presses Universitaires de Franche-Comté, 215-229.
- Fournier, P. & I. Girard (2022), *Manuel de transcription phonétique de l'anglais*, Paris : Ellipses.
- Fournier, P. (2024), “The impact of source languages on the stressing of loanwords in English”, *Lexis* 23.
- Fournier, P. (en préparation), “Stress in Arabic Loanwords in English”.
- Fuchs, Y. (2012), *Les quotatifs en interaction. Approche synchronique d'un paradigme en mouvement, dans un corpus d'anglais oral britannique et irlandais*, Thèse de doctorat : Université Paris 3.
- Fuchs, Y. (2014), *Les quotatifs en interaction en anglais contemporain*, Paris : Presses de la Sorbonne Nouvelle.
- Giegerich, H.J. (1992), *English Phonology: An Introduction*, Cambridge University Press.

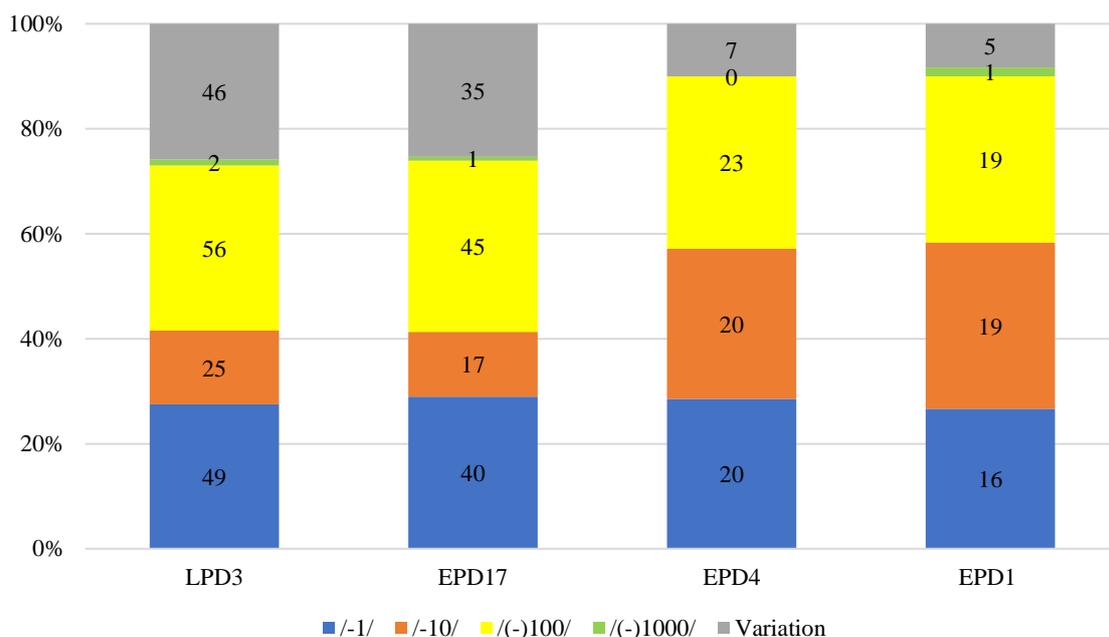
- Gilliver, P. (2015), "The Quotation Collectors: A Conspectus of Readers for the *Oxford English Dictionary*", *Dictionaries. Journal of the Dictionary Society of North America* 36, 47-71.
- Ginésy, M. (2000), *Phonétique et phonologie de l'anglais*, Paris : Ellipses.
- Guierre, L. (1975), *Drills in English Stress Patterns*, Paris: Armand Colin - Longman.
- Guierre, L. (1979), *Essai sur l'accentuation en anglais contemporain - Eléments pour une synthèse*, Thèse de Doctorat : Université Paris VII.
- Guierre, L. (1984), *Drills in English Stress-Patterns* (4th ed.), London: A. Colin - Longman.
- Guierre, L. (1987), *Règles et exercices de prononciation anglaise*, Paris : Armand Colin-Longman.
- Halle, M. & J.R. Vergnaud (1987), "Stress and the cycle", *Linguistic Inquiry* 18(1), 45-84.
- Hammond, M. (1999), *The Phonology of English: A Prosodic Optimality-Theoretic Approach*, Oxford: Oxford University Press.
- Haspelmath, M (2009), "Lexical Borrowing: Concepts and Issues", *Loanwords in the World's Language: A Comparative Handbook*, M. Haspelmath et U. Tadmor (éds.), New York et Berlin: de Gruyter Mouton, 35-54.
- Hayes, B. (1980), *A Metrical Theory of Stress Rules*, Doctoral dissertation, Massachusetts Institute of Technology.
- Hayes, B. (1982), "Extrametricity and English stress", *Linguistic Inquiry* 13 (2), 227-276.
- Hussein, A. (2020), *Vowel pronunciation of Arabic loanwords in English*, Mémoire de Master 2 recherche, Université Sorbonne Paris Nord.
- Hyman, L. (1970), "The role of borrowing in the justification of phonological grammars", *Studies in African Linguistics* 1 (1), 1-48.
- Hyman, L. (1985), *A Theory of Phonological Weight*, Dordrecht: Foris.
- Itô, J. & A. Mester (1995), "Japanese Phonology", *The Handbook of Phonological Theory*, J. Goldsmith (eds.), Oxford: Blackwell, 816-838.
- Jacobs, H. & C. Gussenhoven (2005), "Loan phonology: salience, the lexicon and optimality theory", *Linguistic Inquiry* 36 (2), 223-258.
- Jones, D. (2011), *Cambridge English Pronouncing Dictionary* (18th edition), Cambridge: Cambridge University Press.
- Kang, Y. (2010), "Tutorial overview: suprasegmental adaptation in loanwords", *Lingua*, 120 (9), 2295-2310.
- Kingdon, R. (1958), *The Groundwork of English Stress*, London: Longman, Green.
- Kiparsky, P. & I.S. Yang (1982), "Lexical morphology and phonology", 1982, 3-9.

- Lawson, E. & J. Stuart-Smith (1999), “A sociophonetic investigation of the ‘Scottish’ consonants (/x/ and /hw/) in the speech of Glaswegian children”, *Proceedings of the 14th International Congress of Phonetic Sciences*, Berkeley, University of California, 2541-2544.
- Le Priault, H. & P. Fournier (éds.) (2016), *Approches pluridisciplinaires en phonologie anglaise*, *Anglophonia* 21 (<http://anglophonia.revues.org/694>).
- Léon, M. & P. Léon (2009), *La prononciation du français* (2nd edition), Paris : Armand Colin.
- Lieberman, M. & A. Prince (1977), “On stress and linguistic rhythm”, *Linguistic Inquiry* 8(2), 249-336.
- Macquarie Dictionary* (2005), 4th edition, Sydney: Macquarie Dictionary Publishers Pty Ltd.
- Mahtout, A. (2022), *Les mutations morphologiques des emprunts allemands en anglais*, Mémoire de Master 2 recherche, Université Sorbonne Paris Nord.
- Marchand, H. (1969), *The Categories and Types of Present-Day English Word-Formation*, München, CH: Beck.
- Martin, M. (2011), *De l’accentuation lexicale en anglais australien standard contemporain*, Thèse de doctorat, Université de Tours.
- Masuda, K. (1974), *Kenkyusha’s New Japanese-English Dictionary* (4th ed.), Tokyo: Kenkyusha.
- McMahon, A. M. (2002), *An Introduction to English Phonology*, Edinburgh: Edinburgh University Press.
- Minkova, D. (2009), “Old and Middle English Prosody”, Ans van Kemenade & Bettelou Los (eds.), *The Handbook of the History of English*, Wiley-Blackwell, 95-124.
- Mohanan, K. P. (1982), *Lexical phonology*, Doctoral dissertation, Massachusetts Institute of Technology.
- Myers-Scotton, C. (1993), “Common and uncommon ground: social and structural factors in codeswitching”, *Language in Society* 22 (4), 475-503.
- Paradis, C. & D. LaCharité (2005), “Category preservation and proximity versus phonetic approximation in loanword adaptation”, *Linguistic Inquiry* 36 (2), 223-258.
- Peperkamp, S. & E. Dupoux (2008), “A typological study of stress ‘deafness’”, *Laboratory Phonology* 7, 203-240.
- Pinnavaia, L. (2001), *The Italian Borrowings in the Oxford English Dictionary*, Bulzoni Editore Edition.
- Plag, I. (2003), *Word-Formation in English*, New York: Cambridge University Press.

- Ramdane Chérif, R. (2018), *L'évolution de l'accentuation des emprunts français dans les dictionnaires anglais de prononciation*, Mémoire de Master 2 recherche, Université Sorbonne Paris Nord.
- Roach, P. (2000), *English Phonetics and Phonology*, Cambridge: Cambridge University Press.
- Romaine, S. & D. Lange (1991), "The use of like as a marker of reported speech and thought: a case of grammaticalization in progress", *American Speech* 66, 227-279.
- Roach, P. (1983), *English Phonetics and Phonology, a Practical Course*, Cambridge University Press.
- Sabatini, F. & V. Coletti (2005), *Dizionario della Lingua Italiana*, Milano : Rizzoli-Larousse.
- Salloum, H. & J. Peters (1996), *Arabic Contributions to the English Vocabulary: English Words of Arabic Origin; Etymology and History*, Librairie du Liban Publ.
- Si Amer, K. (2021), *Stress assignment in Spanish loanwords in contemporary English*, Mémoire de Master 2 recherche, Université Sorbonne Paris Nord.
- Silverman, D. (1992.), "Multiple scansion in loanword phonology: evidence from Cantonese", *Phonology* 9 (2), 289-328.
- Simpson, J. & E. Weiner (1989), *Oxford English Dictionary. Online Edition*, Oxford: Clarendon Press.
- Tagliamonte, S. & R. Hudson (1999), "Be like et al. beyond America: the quotative system in British and Canadian youth", *Journal of Sociolinguistics* 3, n°2, 47-172.
- Tannen, D. ([1989] 2007), *Talking Voices: Repetition, Dialogue and Imagery in Conversational Discourse* (2nd edition), Cambridge: Cambridge University Press.
- Trevian, I. (2003), *Morphoaccentologie et processus d'affixation de l'anglais*, Bern : Peter Lang.
- Vance, T. J. (2008), *The Sounds of Japanese (with audio CD)*, Cambridge: Cambridge University Press.
- Wells, J.C. (2008), *Longman Pronunciation Dictionary* (3rd edition), London: Longman.
- Wijk, A. (1966), *Rules of Pronunciation for the English Language. An Account of the Relationship Between English Spelling and Pronunciation*, Oxford: Oxford University Press.
- Zhao, G. (2023), *Etude de la lexicalisation des emprunts chinois en anglais*, Mémoire de Master 2 recherche, Université Sorbonne Paris Nord.

Annexe 1 - Mémoire de R. Ramdane Chérif sur l'évolution de l'accentuation des emprunts français dans les dictionnaires anglais de prononciation

Ce mémoire de recherche teste l'hypothèse formulée dans mon étude diachronique de 2016 selon laquelle les emprunts français n'ont pas subi de modification majeure de leur schéma accentuel au cours des dernières décennies. Les résultats de ce mémoire sont basés sur les emprunts français extraits de *LPD3*, dès lors qu'une transcription française est attestée dans la base de données. L'origine française des items est vérifiée dans l'*OED* et l'analyse se focalise uniquement sur les noms communs d'au moins trois syllabes en anglais britannique, car l'évolution de la variation accentuelle est plus visible au sein des mots dits longs. La première partie de l'étude est synchronique et compare l'accentuation des emprunts français sélectionnés dans *LPD3* et *EPD17*, puis la seconde partie procède à une analyse diachronique de l'accentuation en comparant les données d'*EPD17* avec celles des dictionnaires *EPD1* (1917) et *EPD4* (1937). Le choix d'*EPD* pour une étude diachronique portant sur l'accentuation se justifie par le fait qu'il a subi davantage de réactualisations que *LPD* et que l'évolution de l'accentuation des emprunts français y est donc peut-être davantage perceptible. L'étude concerne uniquement l'anglais britannique car l'anglais américain n'est pas recensé dans les premières éditions d'*EPD*. Le corpus se compose de 178 emprunts identifiés à partir de *LPD3*, dont 138 figurent dans *EPD17*, 70 dans *EPD4* et 60 dans *EPD1*. Le graphique 21 synthétise la répartition des schémas accentuels mis en évidence dans les quatre dictionnaires de prononciation.



Graphique 21 : Accentuation des emprunts français en anglais britannique dans *LPD3*, *EPD17*, *EPD4* et *EPD1*.

Les résultats montrent que la répartition des schémas accentuels est quasiment identique dans *LPD3* et *EPD17*, avec une proportion étonnamment faible d'accentuation finale (30% environ) au profit notamment d'une accentuation antépénultième qui constitue une proportion sensiblement équivalente. Je fus à l'époque assez surpris du faible taux d'accentuation finale en anglais britannique, taux bien inférieur à celui de 50% mis en évidence pour la catégorie des noms propres dans mes précédentes études. Je l'interprète différemment aujourd'hui, car ma toute dernière étude en collaboration avec Q. Dabouis a montré que plus le nombre de syllabes d'un emprunt français est important et plus la tendance à la rétraction accentuelle sur la syllabe pénultième ou antépénultième est avérée. Ainsi, le critère même de sélection du corpus consistant à ne sélectionner que des emprunts au minimum trisyllabiques aurait quelque peu faussé l'étude en surreprésentant les schémas pénultièmes et antépénultièmes au détriment de l'accentuation finale.

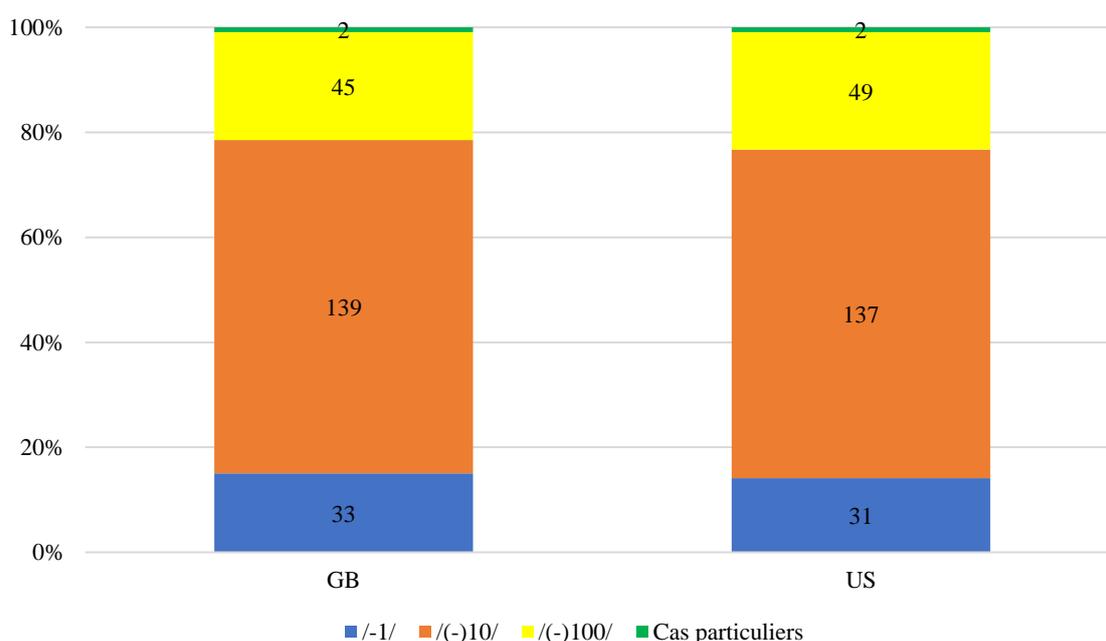
La comparaison diachronique avec les versions antérieures d'*EPD* ne montre pas d'évolution accentuelle majeure concernant les emprunts français en anglais britannique. On constate une répartition égale entre les trois schémas accentuels finaux, pénultièmes et antépénultièmes dans *EPD1* et *EPD4*. On note toutefois une légère diminution de l'accentuation pénultième dans les dictionnaires contemporains, au profit d'une augmentation des cas de variation (principalement accentuation finale vs non finale). Là encore, cela ne va pas dans le

sens de l'hypothèse sociolinguistique de J.-M. Chadelat (2000) selon laquelle l'anglais britannique tendrait à se démarquer de la prononciation française, le phénomène de rétraction accentuelle sur la syllabe pénultième pouvant constituer un exemple concret au niveau accentuel. Les résultats ne montrent pas une telle tendance à la rétraction accentuelle en anglais britannique : il y a même une baisse de l'accentuation pénultième notée dans la dernière édition d'*EPD*.

Le reste de l'étude consiste à évaluer le taux de correspondance accentuelle entre les emprunts attestés dans *LPD3* et *EPD17* (80% de correspondance totale), dans *EPD1* et *EPD4* (93% de correspondance totale) et entre *EPD1*, *EPD4* et *EPD17* (60% de correspondance totale et 22% de correspondance partielle) afin de faire émerger les microsystèmes ayant subi une évolution accentuelle au fil des éditions.

Annexe 2 - Mémoire d'A. Chaou sur l'accentuation des emprunts portugais en anglais

À partir d'une recherche avancée basée sur l'origine étymologique portugaise des items au sein de l'*OED*, 219 emprunts furent sélectionnés pour l'analyse, dont 125 étaient également attestés dans *LPD3*. Le graphique 22 montre la répartition des schémas accentuels des emprunts portugais dans l'*OED* en anglais britannique et américain.



Graphique 22 : Accentuation des emprunts portugais en anglais britannique et américain dans l'*OED*.

L'étude présente quelques lacunes méthodologiques dans l'extraction des emprunts, dans la détermination de la variété de portugais sélectionnée (portugais lusitanien ou brésilien), ainsi que dans la présentation des résultats accentuels. En effet, sont regroupés dans le graphique 22 les emprunts dissyllabiques et trisyllabiques (et plus), ce qui a pour fâcheuse conséquence de minimiser la proportion d'accentuation antépénultième au sein des résultats. En effet, les dissyllabes ne pouvant pas, par définition, relever d'une accentuation antépénultième, les résultats sont quelque peu faussés. J'ai tout de même choisi de présenter les résultats de ce mémoire, car j'ai remarqué que la proportion d'emprunts dissyllabiques était relativement faible au sein du corpus, ce qui permettait tout de même de conclure à la représentativité des données du graphique 22.

On observe une très faible variation dialectale entre l'anglais britannique et américain. Environ 15% des emprunts portugais sont accentués sur la syllabe finale :

exemples : *Babassu, candiru, igarape* ;

20% sont accentués sur la syllabe antépénultième :

exemples : *adigar, agilar, ferrinho* ;

et 65 % environ sont accentués sur la syllabe pénultième :

exemples : *Albino, macumba*.

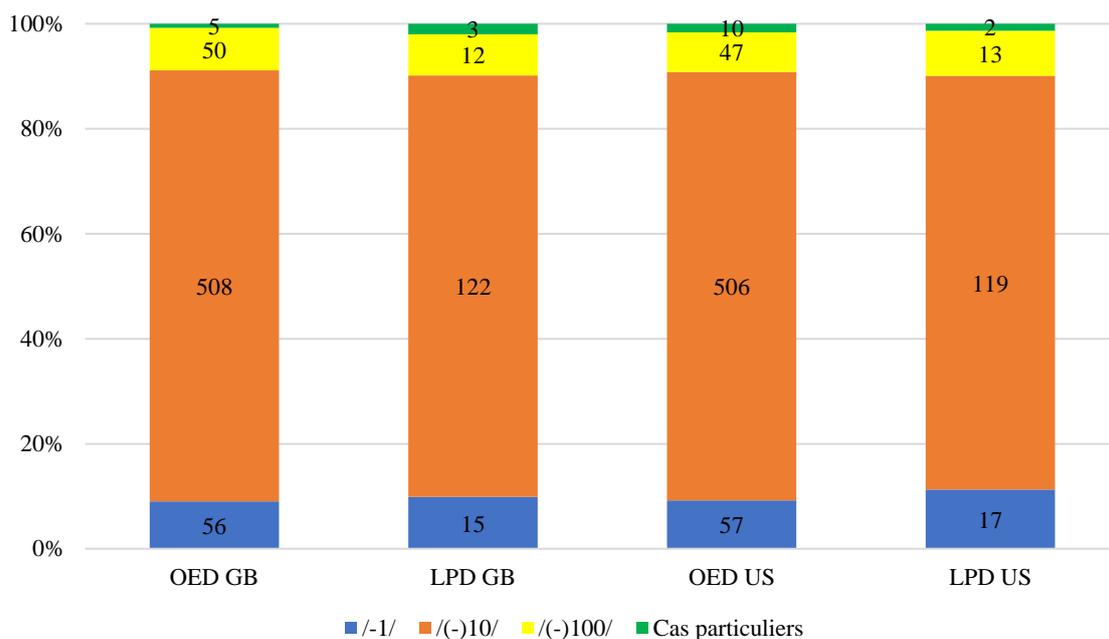
Même si l'accentuation pénultième reste le schéma le plus attesté dans les emprunts portugais, on constate que la proportion d'accentuation finale et antépénultième est tout de même significative. Les données concernant la répartition des schémas accentuels sont globalement identiques dans les 125 emprunts attestés dans *LPD*. Les taux significatifs d'accentuation finale et antépénultième s'expliqueraient-ils par une reproduction de l'accentuation portugaise d'origine ? C'est ce que semblent suggérer les résultats obtenus par A. Chaou dans la partie finale de son mémoire, avec un taux de correspondance supérieur à 80% entre la position de la syllabe accentuée en anglais britannique et américain dans *LPD* et la position de la syllabe accentuée en portugais déterminée à partir des données du dictionnaire en ligne *Portuguese Online Pronunciation Dictionary Portuguesepod101.com*.

Annexe 3 : Mémoire de K. Si Amer sur l'accentuation des emprunts espagnols en anglais

Ce mémoire est organisé de la même façon que celui d'A. Chaou en Annexe 2. L'étude est toutefois beaucoup plus riche, en raison de l'influence majeure qu'exerce la langue espagnole sur l'anglais aux Etats-Unis notamment. La sélection des emprunts directs à la langue espagnole par l'intermédiaire d'une recherche avancée effectuée dans l'*OED* permet de recueillir 619 items dont 151 figuraient également dans *LPD*. Voici quelques exemples représentatifs de ce corpus :

adobe, aficionado, aguardiente, alameda, Alcazar, alfalfa, alpaca, altiplano, armadillo, arroyo, babaco, balata, barracuda, alcalde, barrio, bodega, bolero, boliviano, bonanza, bonito, bracero, burrito, caballero, cabana, camarilla, cascarilla, caudillo, cigarillo, conquistador, ciguatera, cojones, condor, coquina, chilli, chipotle, chorizo, cha-cha / cha-cha-cha.

La première partie de l'étude consiste à rendre compte de l'accentuation des emprunts espagnols en anglais britannique et américain à partir des transcriptions de l'*OED* et de *LPD*. Le graphique 23 présente les résultats de la répartition des schémas accentuels observés. Le taux de correspondance accentuelle entre l'anglais britannique et américain fut évalué en amont de ce travail afin de déterminer si l'accentuation des emprunts espagnols était sujette à une variation de type dialectal, ce qui s'avéra ne pas être le cas. Le taux de correspondance accentuelle entre les deux variétés est de 97% dans l'*OED* et de 98% dans *LPD*.



Graphique 23 : Accentuation des emprunts espagnols en anglais britannique et américain dans l'OED et LPD.

Même si, une nouvelle fois, un problème méthodologique complique l'interprétation de ces résultats, en raison de la non prise en compte du nombre de syllabes, ce qui conduit à une sous-représentation de la proportion d'accentuation antépénultième au sein des emprunts de trois syllabes et plus, des résultats intéressants sont à noter. Les résultats sont sensiblement équivalents entre les deux dictionnaires : les emprunts espagnols sont massivement accentués sur la syllabe pénultième (environ 80%) et les accentuations sur la syllabe finale (environ 10%) et sur la syllabe antépénultième (environ 10%) restent minoritaires.

La seconde partie de l'étude, centrée sur l'analyse de la correspondance accentuelle entre l'anglais et l'espagnol, conduit à des conclusions sensiblement équivalentes à celles formulées pour les emprunts portugais dans le mémoire précédent, à savoir que le taux de correspondance s'élève à environ 80%, ce qui semble suggérer que la dynamique accentuelle des emprunts espagnols en anglais est basée sur une reproduction du schéma accentuel d'origine.

Annexe 4 - Mémoire d'A. Hussein sur l'application des règles graphophonématiques dans les emprunts arabes en anglais

Ce travail souligne la présence massive au sein des mots empruntés à l'arabe de voyelles libres étrangères (<a> = /ɑ:/, <e> = /eɪ/, <i> = /i:/) au niveau des syllabes accentuées.

Exemples : *Baath*, *baba ganoush*, *Salaam*, *Sharia*, *ulema*.

A. Hussein utilise l'algorithme de J.-M. Fournier (2010) permettant de déterminer la valeur d'une voyelle accentuée en fonction de son environnement graphique. Le corpus se compose de 100 emprunts directs à l'arabe attestés dans l'*OED* et *LPD*. L'influence des voyelles libres étrangères est telle sur la prononciation des emprunts arabes que l'on en rencontre même dans des contextes généralement entravants.

Exemples : C₂ → voyelle entravée mais *hadj*, *inchallah* = /ɑ:/.

C# → voyelle entravée mais *halal*, *imam* = /ɑ:/ et *Hanif* = /i:/.

Ce mémoire contient également dans la partie théorique un petit précis de phonologie de l'arabe permettant d'observer d'éventuelles correspondances phonétiques entre l'arabe et les emprunts correspondants en anglais. Les consonnes et les voyelles de l'arabe y sont présentées, de même que les différentes structures syllabiques rencontrées. La seconde partie du mémoire tente de justifier les exceptions aux règles de prononciation (et notamment la présence massive de voyelles libres étrangères) en mettant en parallèle la prononciation des emprunts arabes en anglais avec la prononciation arabe. Si le format du mémoire limite d'emblée la portée d'une telle étude, les résultats sont toutefois significatifs au niveau de la réalisation des voyelles. La présence de voyelles libres étrangères dans les dictionnaires anglais de prononciation est le reflet direct de la prononciation arabe avec des prononciations vocaliques sensiblement équivalentes. L'anglais utilise les phonèmes dont il dispose et qui sont suffisamment proches de ceux de l'arabe pour reproduire le plus fidèlement possible la prononciation d'origine. Les systèmes phonologiques étant radicalement différents, le processus d'adaptation inhérent à l'intégration d'un emprunt arabe en anglais reproduit une prononciation certes différente, mais la plus proche possible en vertu des paramètres phonétiques disponibles en anglais. Ainsi, ces travaux montrent que l'influence des langues sources sur les langues cibles ne se limite pas au seul domaine de l'accentuation lexicale, mais que la sélection des phonèmes utilisés pour prononcer un emprunt est aussi dictée par le souhait de reproduire le plus fidèlement possible la prononciation d'origine. Ce phénomène n'est toutefois pas systématique et la présence dans les emprunts arabes de cas de variation de type voyelle libre / voyelle libre étrangère ou voyelle entravée / voyelle libre étrangère traduit le conflit au sein du système qui oscille entre

reproduction de la prononciation d'origine et autonomisation et adaptation aux paramètres du système phonologique de l'anglais. Les résultats de ce mémoire montrent une fois encore que l'hypothèse selon laquelle les emprunts intègrent une autre langue par l'intermédiaire de locuteurs bilingues paraît hautement probable.

The impact of source languages on the stressing of loanwords in English

This study suggests that the parameters accounting for the stressing of loanwords in English should be distinguished from those governing the rest of the English lexicon. This dictionary-based analysis proves that the stress characteristics of source languages are determining factors in the stressing of words of foreign origin in English. Throughout history, the massive influx of words from source languages stressed on the penultimate syllable might have led to a default stress pattern in English being used as a phonological marker to indicate the foreign origin of items. Even though stress preservation is the main parameter accounting for the stress patterns of English loanwords, there are proofs of an ongoing conflict in English between the preservation of the original stress pattern derived from source languages and the application of the default stress pattern instead.

Key words: English loanwords, stress patterns, source languages, dictionary data.

0. Introduction.

Hammond (1999: 284) states that the rules governing the pronunciation of loanwords in English are not different from those applying to the rest of the lexicon:

In some languages, borrowed words exhibit rather different phonological patterns from those of native words, but this does not appear to be the case in the stress system of English.

The assumption according to which loanwords exhibit no specific prosodic features (especially in terms of stress assignment) is shared by many researchers. As a consequence, loanwords are often cited in reference works (Burzio 1994; Chomsky & Halle 1968) to illustrate the regularity of English stress rules. If segmental parameters (among which syllable weight) are often used to account for the stressing of loanwords in English, little has been said about the impact of source languages on the pronunciation of loanwords in English.

Itô & Mester (1999) insist on the need to distinguish the mechanisms applying to the pronunciation of loanwords from those accounting for that of the rest of the lexicon. Although their study deals with Japanese, Dabouis & P. Fournier (2022) show that this hypothesis also stands for English. They posit that English is actually composed of subsystems and that words which belong to one specific subsystem share similar semantic, morphological, phonological,

segmental and letter-to-sound characteristics which differ from the words from the other subsystems. Their model highlights the interaction between the different subsystems but also shows that the pronunciation of loanwords might also be influenced by the preservation of linguistic properties from source languages. These assumptions echo the positions developed in major works in Loan Phonology and according to which loanwords are introduced in English by bilingual speakers. This position is not shared by all researchers because some think that the speakers who introduce words from foreign languages are not bilingual. However, the probability that there is some interaction between the phonological parameters of source languages and the rules accounting for the pronunciation of loanwords in English is very high. Although there have been many studies concerning the pronunciation of words of English origin in other languages (Paradis & La Charité (2008) for French, Kubozono, Ito & Mester (2008) for Japanese or Kenstowicz (2007) for Fijian for example), few studies investigating the pronunciation of loanwords in English in relation with the phonological properties of source languages are attested in the literature.

In English, loanwords are usually stressed on the penultimate syllable (except French loanwords which are generally stressed on the final one) and this stress pattern is considered to be the default pattern in loanwords and a phonological indicator of the foreign origin of items (P. Fournier 2018). However, no quantitative analysis has been carried out to determine if this assumption is true. So, this study first investigates the distribution of stress patterns in English loanwords on a representative corpus compiled from the *Routledge* dictionary (Upton & Kretschmar 2017). The overall results show that even if loanwords are usually stressed on the penultimate syllable, the proportion of words stressed on the antepenultimate stress is significant. The results need to be refined as it is necessary to distinguish between the different source languages attested in the dictionary because words from specific source languages show regular stress patterns.

In order to determine the impact of the phonological properties of source languages on the stressing of loanwords in English, the location of the stressed syllable found in source languages has systematically been compared (whenever such information was available in the pronouncing dictionary) with the stress patterns attested in English loanwords. Even though the transmission of phonological properties from source languages to English proves out to be a determining factor in the stressing of loanwords (the efficiency of this parameter has yet to be tested among all the other determining parameters, however), the reproduction of the original stress pattern is not systematic and variation is strongly linked to the source languages. The examples below illustrate the different scenarios.

- (1) (a) '*Ostwald* from German [ˈɔstvalt], '*Autobahn* from German [ˈautoba:n];
 (b) '*semplice* from Italian [ˈsemplitʃe], '*sgraffiti* from Italian [zɡrafˈfiti], '*pode'stà*
 from Italian [podeˈsta];
- (2) (a) '*Borodin* from Russian [bərəˈdʲin], '*Chelya'binsk* from Russian [tʃɨˈlʲabʲinsk],
Che'renkov from Russian [tʃɨrʲinˈkof];
 (b) '*Guzmán* from Spanish [ɡuθˈman, ɡusˈman], '*Calderón* from Spanish [ˌkaldeˈron],
 '*matador* from Spanish [mataˈðor].

If the original stress patterns of words of German and Italian origin are perfectly reproduced in English in (1a) and (1b) (even when the primary stress is located on the final or antepenultimate syllable), this does not apply, for instance, to words of Russian and Spanish origin, for which stress differences are attested in (2a) and (2b). That is why it is necessary to distinguish between the different source languages so that distinct behaviours can emerge.

Section 1 presents a review of the literature on the stressing of loanwords in English with an emphasis on the works related to Loan Phonology and the transmission of phonological properties from source languages to target languages. In Section 2, the corpus building process is detailed and the results concerning the distribution of stress patterns in British English are presented in Section 3. The results concerning the comparison between the stress patterns of source languages and those found in English are dealt with in Section 4 and are interpreted in Section 5 in the light of the determining theoretical parameters discussed in Section 1. A series of hypotheses are then formulated to account for whether the original stress pattern is preserved in English.

1. Theoretical background.

1.1 The stressing of loanwords in English.

Syllable weight has long been considered a major parameter to account for the stressing of the English lexicon, including loanwords, and is still found in reference works (cf. introduction), but its influence has notably been questioned in Church (1985) because it appears not to be adapted to account for the stressing of loanwords in English. Church notices that loanwords exhibit different stress patterns from the rest of the English lexicon in comparable weight conditions. Conceiving of stress as based on syllable weight also generates a circularity problem. Stress is said to be related to the weight of syllables which is mainly calculated thanks

to vowel realization. According to this view, the pronunciation of vowels is a key factor in the determination of stress placement, which is highly controversial as the Guierrian School considers that stress assignment is better accounted for by an interaction between morphological, syntactic and segmental parameters and that subsequently the pronunciation of vowels is determined by stress placement (Dabouis *et al.* 2023).

Other studies suggest that the English phonological system is a hybrid and dynamic system made up of two conflicting principles: the Romance principle inherited mainly from Latin and French through the process of borrowing. According to this principle, stress location is reckoned from the rightmost part of the word. However, the English phonological system also relies on the Germanic principle and its root-initial stress conception. (J.-M. Fournier 2007; Minkova 2006). According to J.-M. Fournier (2007), the English stress system relies on syntactic, morphological and segmental parameters and syllable weight alone cannot account for the location of primary stress. The views concerning the interaction between these parameters and the necessity to distinguish the pronunciation of loanwords from the rest of the lexicon are developed in Dabouis *et al.* (2023).

Relying on the arguments put forward by Lionel Guierre in response to the hypotheses presented in the *Sound Pattern of English* (Chomsky & Halle 1968), the researchers from the Guierrian School have identified regular orthographical and segmental patterns in loanwords which are consistent with penultimate stress patterns. This stress rule is known as the “Italian rule” (Guierre 1979: 544-550), because of the considerable proportion of words of Italian origin whose stress pattern can be accounted for according to these parameters. Words of foreign origin with <-VCV#> patterns and whose segmental structure is as follows: a dental obstruent (i.e. /t, d, n, s, z/ except /l/ and /r/) with a pronounced final monographic vowel (i.e. <a, e, i, o, u> except <y>), are generally stressed on the penultimate. The examples in (3) illustrate this stress rule:

(3) *ar* 'mada, *bi* 'kini, *ki* 'mono, *mar* 'tini, *mo* 'squito, *nir* 'vana, *po* 'tato, *tor* 'pedo, *vol* 'cano.

This stress rule is not based on etymology only and it is actually a combination of etymological factors and segmental characteristics. It is a restrictive pedagogical rule as it cannot account for the stressing of words with /l/ and /r/ consonants which are mainly stressed on the antepenultimate (examples: 'broccoli, 'buffalo, 'copula, 'camera, 'Figaro). Furthermore, there are exceptions to this “Italian rule” in English loanwords (examples: 'domino, 'stamina, 'retina, 'wapiti).

1.2 The English subsystems.

Dabouis & P. Fournier (2022) assume that the English phonological system is actually composed of several subsystems characterized by distinct graphophonological, semantic, morphological, phonological and segmental parameters: §Core (which is divided into §Core-Native and §Core-Latinate), §French, §Learned and §Foreign (based on Carney's representation of spelling-to-sound subsystems (1994)). It is necessary to distinguish the pronunciation of loanwords which belong to the §Foreign subsystem from that of the words from the other subsystems because they share linguistic properties. First, there are specific letter-to-sound correspondences for vowels and consonants. In Figure 1, the letter-to-sound vocalic correspondences referred to as foreign free vowels are only attested in loanwords.

Free vowel	Spelling	Foreign free vowel	Examples
/eɪ/	<a>	/ɑ:/	<i>armada, banana, tomato</i>
/i:/	<e>	/eɪ/	<i>cliché, peso, sake</i>
/aɪ/	<i>	/i:/	<i>ballerina, bikini, mosquito</i>

Figure 1: The main foreign free vowels (taken from J.-M. Fournier 2010: 113).

There are also specific letter-to-sound correspondences for consonants. Some of these are illustrated in (4):

- (4) <c(c)> = /tʃ/ in *ciabatta, cello, concerto*;
 <sch> = /ʃ/ in *schnaps, schwa*;
 <z(z)> = /ts/ in *panzer, pizza*.

From the semantic point of view, loanwords are direct references to foreign notions (examples: food, habits, tradition, leisure, clothing). These words usually exhibit <-VCV#> segmental structures (see examples in Figure 1), tend to be polysyllabic and are generally stressed on the penultimate syllable. It is necessary to distinguish the §Foreign subsystem from the §French one especially concerning stress assignment, as words of French origin generally appear to be stressed on the final syllable, especially in American English (Dabouis & P. Fournier 2024). This assumption concerning the existence of a penultimate preferential stress in loanwords echoes what Church said as early as in 1985 concerning a “pseudo-foreign accent” typical of loanwords and used by English speakers to signal their foreign origin:

It is clear (...) that the stress of Italian loans is not dependent on the weight of the penultimate syllable (...) I believe that speakers of English adopt what I like to call a pseudo-foreign accent. That is, when speakers want to communicate that a word is non-native, they modify certain parameters of the English stress rules in simple ways that produce bizarre “foreign sounding” outputs.

The English subsystems mentioned in Dabouis & P. Fournier (2022) interact constantly and their boundaries fluctuate, as is suggested by the overlapping representation of their theoretical models in Figure 2.

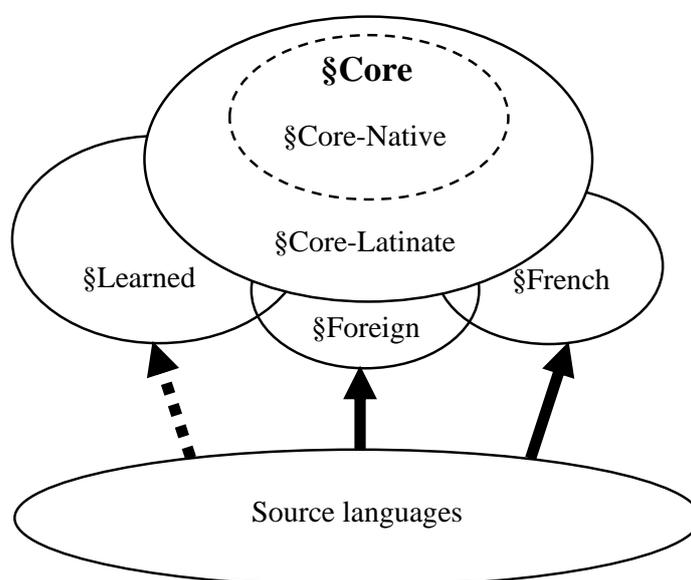


Figure 2: Representation of the English subsystems (Dabouis & P. Fournier 2022).

Even though they consider that English is an autonomous system, source languages have a direct influence on the three subsystems composed of items or elements of foreign origin. This representation testifies to the existence of a conflict between the principles of the English system which illustrate its autonomy and the influence of source languages on items of foreign origin. Therefore, the preferential penultimate stress pattern found in loanwords might be inherited from the prosodic characteristics of source languages because there is evidence in the literature of the direct influence of source languages on English.

1.3 Loanword adaptation.

The investigation of the phonological phenomena observed at the source language / target language interface during the borrowing process has led to the creation of the Loan Phonology theoretical framework. There have been more and more studies investigating the pronunciation of words of foreign origin over the last two decades. There are two scenarios which are put forward to account for the pronunciation of words of foreign origin. They are greatly summarized in the introductory section of Calabrese & Wetzels' book (2009: 1-3). The 'nativization-through-production process' (or the phonological stance model) (Hyman 1970; Jacobs & Gussenhoven 2005; Paradis & LaCharité 2005) posits that loanwords are introduced in target languages by bilingual speakers. The loanwords are mainly pronounced using the phonological principles of target languages. However, even if this view is massively shared, others researchers think that loanwords can also be realized with the oral mechanisms ruling the source languages since speakers are bilingual. The second position, called the 'nativization-through-perception process' (or the perceptual stance model) (Silverman 1992; Peperkamp & Dupoux 2008), states that loanwords are borrowed from source languages by speakers who are not fluent in these languages. The pronunciation of loanwords is based on the perception speakers have of the source language prosodic characteristics and is then reproduced with a variable degree of accuracy in the target language with a necessary phase of adjustments. Peperkamp (2004) shows that the pronunciation of words of foreign origin does not necessarily respect the phonological principles of target languages by using psycholinguistic evidence concerning the perception of non-native sound structures and the mechanisms which are "beyond the listener's awareness".

The literature on the pronunciation of loanwords is productive but the prosodic adaptation of loanwords has received less attention. Studies on prosodic adaptation in loanwords are massively based on the assumption that loanwords infiltrate target languages through bilingual speakers. Bilingualism makes it possible to perceive stress in both languages and therefore to reproduce the original stress features in target languages (Kang 2010):

the closer the contact and the higher the level and rate of bilingualism, the more likely the adapters are to perceive stress contrasts of the input language correctly, in turn leading to the stronger possibility of the preservation of input stress.

An experiment on English native speakers conducted by P. Fournier (2018) has shown that speakers have intuitions about the way words of Italian origin should be stressed. Whenever faced with words they feel to be of Italian origin, they tend to stress the penultimate syllable

whatever the stress pattern of the Italian item may be. The stress match rate between the Italian items and the productions concerning penultimate or antepenultimate stress patterns rises with the proficiency of speakers in Italian. It confirms there is a direct link between the original stress patterns and the way English speakers pronounce words of foreign origin in English. These results echo the conclusions found in Fitt (1996) concerning the pronunciation of trisyllabic foreign city names by English adolescents in which the penultimate stress pattern is used predominantly.

Davis, Tsujimura & Tu (2012) have developed a three-parameter model of prosodic adaptation. The first parameter is the systematic comparison of prosodic features both in source and target languages as it appears that the preservation of the prosodic features from source languages is not systematic in loanwords. Some features may be better reproduced than others. The second parameter is based on the prosodic characteristics some target languages apply to loanwords. Some languages tend to use the same prosodic rules to account for the pronunciation of the whole lexicon while others use parameters which are specific to loanwords. Finally, syllable structure and segmental features are considered as potential determining parameters in the adaptation process.

The model created by Dabouis & P. Fournier (2022) shows that the §Foreign subsystem is a dynamic and hybrid one which relies both on the application of rules which are specific to loanwords in the English system and on the influence of source languages. Loanwords from Italian are a perfect example of stress preservation between source and target languages (P. Fournier 2018) because there are hundreds of Italian loanwords in English and the study shows that there is an outstanding stress match between loanwords of Italian origin and Italian items (more than 95% concerning final, penultimate and antepenultimate Italian stress patterns). This preservation of the Italian stress pattern is assumed to be facilitated by the fact that both languages share similar prosodic features concerning stress representation. P. Fournier (2016) shows that the suprasegmental features of English (stress-language) and Japanese (pitch-accent language) are so different that Japanese stressed syllables are not reproduced in English either in disyllabic or trisyllabic (and more) items. The stressing of Japanese loanwords in English is better accounted for by the application of the preferential stress pattern for loanwords (i.e. penultimate). Nonetheless, Broselow (2009) states that prosodic adaptation is generally required even though the two languages are stress-timed languages.

Therefore, the assumption that the preferential penultimate stress pattern which is applied whenever the reproduction of the original stress pattern is impossible needs to be investigated and tested on a large corpus of loanwords coming from different source languages. Are

loanwords massively stressed on the penultimate syllable in English? If it turns out to be true, is this penultimate stress pattern inherited from source languages and is there a tendency in English to systematically preserve the stress pattern from source languages? Results have shown that stress preservation is attested in loanwords of Italian origin but impossible in loanwords of Japanese origin because of the prosodic differences between English and Japanese. What about other source languages?

2. Corpus building.

In order to create a representative dataset of loanwords in English and to investigate the impact of source languages on the pronunciation of such items, the corpus has been compiled from the *Routledge Dictionary of Pronunciation for Current English* 2nd edition (Upton & Kretschmar 2017). This source is the most updated dictionary for current pronunciation of British and American English and it sometimes mentions the phonetic transcriptions of source languages. These transcriptions give us access to the standard pronunciation of source languages and it is then possible to determine whether phonological properties of source languages are reproduced in English. When only one source language is mentioned, it can legitimately be concluded that the borrowing transfer is direct, that is to say with no intermediate languages which might have influenced the realization of loanwords in English. These loanwords taken from English dictionaries are institutionalized entries, which means that it is necessary to distinguish those institutionalized entries from the foreign words used in code-switching (Myers-Scotton 1993). Indeed, the code-switching process relies on individuals and has a stylistic dimension.

It is clear that the principles underlying the stress systems of source languages throughout the dictionary radically differ. However, whatever the phonetic materialization of stress across languages may be (i.e. intensity, duration, pitch or a combination of several parameters), the *Routledge* editors indicate the location of primary stress in source languages with the API symbol found in English as well (i.e. /'/). It is therefore possible to compare the location of stressed syllables in English as well as in source languages. French loanwords are not included in the dataset because French is categorized among the ‘languages with dominant final stress’ category (Hyman 1985). French has a demarcative stress which falls on the last syllable of intonation units (Di Cristo & Hirst 1993). The stressing of French loanwords has been extensively studied and it has been shown that these constitute a distinct sub-group within the English phonological system (J.-M. Fournier 2010) with morphological, semantic, segmental

but especially phonological characteristics (Dabouis & P. Fournier 2022) which differ from loanwords coming from other languages.

The corpus has been compiled from the *Routledge* database and some loanwords have not been integrated into the final dataset. Monosyllabic loanwords have been excluded and loanwords with two languages of origin are not kept either as it is difficult to determine which of the two source languages may have influenced the pronunciation in English. Entries which are doublets with spelling variations are not kept as it is clear that these minor differences do not affect the pronunciation in English (example: *Tadzhik* is not kept because *Tadjik* is already attested⁶³), as well as masculine / feminine or singular / plural doublets (example: *sgraffito* is not kept because *sgraffiti* is already attested). Finally, certain entries may have different syllable counts in English and in source languages. The realization of vocalic sequences such as <i+V> is generally disyllabic in English but monosyllabic in source languages and especially in Italian with the synaeresis principle. Therefore, variability in the syllable count makes it impossible to compare the stress patterns in English and in source languages. That is why such entries as *siciliana* have been left out.

The final dataset is composed of 768 loanwords which come from 19 source languages. Table 1 compiles the different source languages mentioned in the *Routledge* dictionary as well as the number of items for each language.

Language	Items	Language	Items
Italian	214	Hungarian	12
Russian	166	Norwegian	10
Spanish	116	Polish	10
Welsh	73	Czech	9
German	44	Afrikaans	7
Danish	28	Portuguese	5
Swedish	21	Turkish	5
Dutch	18	Flemish	4
Irish	13	Albanian	1
Greek	12		

Table 1: Source languages mentioned in the *Routledge* dictionary along with the number of loanwords in the final dataset.

The results show that there are major discrepancies between donor languages concerning the number of borrowings transmitted to English. Thus, Italian, Russian and Spanish loanwords

⁶³ When doublets are attested, the first of the two entries has been kept.

respectively represent 27.9%, 21.6% and 15.1% of the dataset (i.e. 64.6%). This massive proportion of loanwords of Italian, Russian and Spanish origin means that some donor languages have only transmitted a few items (examples: Albanian, Flemish or Turkish). Determining the influence of source languages on the basis of such scant evidence is far from easy. That is why the analysis and the conclusions are to be dealt with carefully. Surprisingly, languages such as Arabic or Japanese which are good purveyors of loanwords in English are not present in the data. It does not mean that the *Routledge* editors deny the existence of English words coming from Arabic or Japanese, but simply that the foreign phonetic transcriptions are not mentioned for such words as *halal* or *karate*. The list of loanwords from each source language is available in the appendix along with the disyllabic and trisyllabic (and more) distinction.

The next section deals with the stressing of loanwords of the dataset in British English.

3. The stressing of loanwords in British English.

This section investigates the location of primary stress in loanwords in British English. According to the literature, loanwords coming from languages other than French are mainly stressed on the penultimate syllable (i.e. /(-)10/⁶⁴). The distribution of the stress patterns found in British English is dealt with in 3.1 and it is necessary to distinguish the stress features attested in disyllabic loanwords from those attested in loanwords of three syllables or more.

3.1 The stressing of loanwords in British English.

⁶⁴ /1/ stands for the syllable which carries primary stress and /0/ for an unstressed syllable. The (-) symbol stands for an undetermined number of syllables and the brackets signify that this is an optional feature. Therefore, this numerical representation brings together disyllabic loanwords and loanwords of three syllables or more as well.

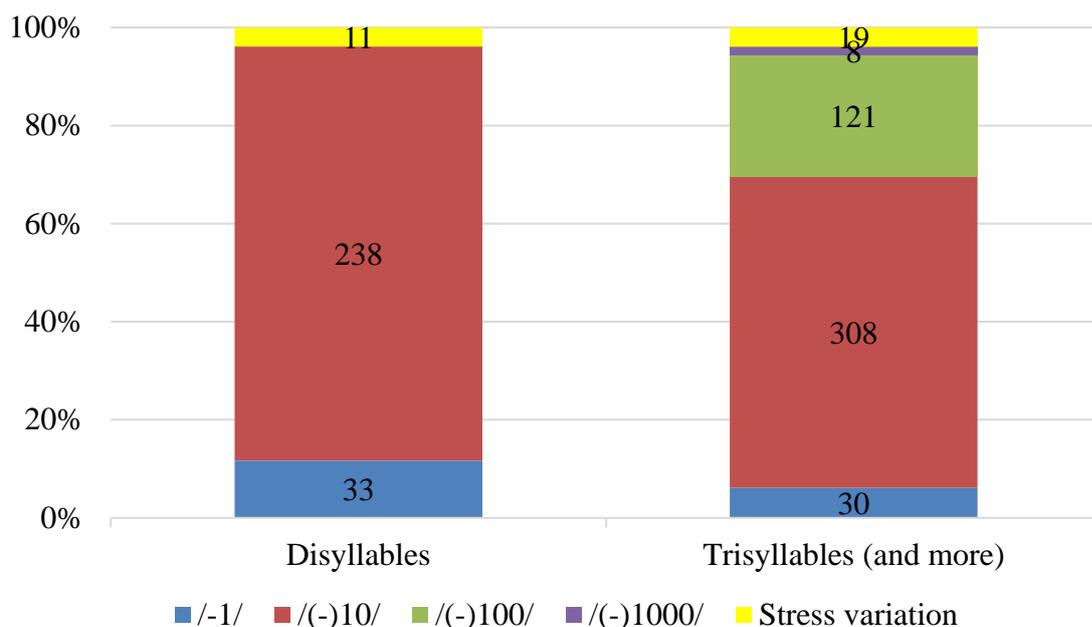


Figure 3: The stressing of loanwords in British English.

The results from Figure 3 show that disyllabic loanwords are generally stressed in /10/ in British English (84.4%) and that only 11.7% of them are stressed in /01/. The proportion of /01/-/10/ variation is very low with only 3.9%. This is not surprising if we consider that final stress patterns in loanwords are mainly found in words of French origin which are excluded from the dataset. Some examples of final stress patterns (5a) and stress variation⁶⁵ (5b) are listed below:

- (5) (a) *Vermeer, Gábor, Sinn Fein, Cavour, Abkhaz, artel, Baikal, kolkhoz, pirog, Bilbao, La Paz, Ortiz;*
 (b) *Loren, signor, Kazakh, Penrhyn, Baku, Kyrgyz, Sergei* = /10/ or /01/.

Concerning loanwords of three syllables (or more), penultimate stress is also dominant (63.4%) but the proportion of antepenultimate stress is significant (24.9%). Other stress patterns are less represented with only 6.2% of loanwords stressed in /-1/, 1.6% in /(-)1000/. Stress variation is not frequent because only 3.9% of loanwords are subject to it. Therefore, even if the penultimate stress pattern is dominant, almost 1 loanword out of 4 of three syllables (and more) is stressed on the antepenultimate syllable in English, and this feature cannot be ignored. The examples in (6) illustrate the final (6a) and the antepenultimate (6b) stress patterns found

⁶⁵ In the *Routledge* introduction (2017: x), it is said that: “The ordering of variant pronunciations does not imply that one form is more desirable or ‘correct’ than another”. Therefore, in this study, the ordering of stress variants is not mentioned as it is not meaningful.

in English loanwords:

- (6) (a) *Kap Farvel, Budapest, Fianna Fáil, pietà, podesta, Astrakhan, Azerbaijan, Chelyabinsk, Tadjikistan, Turkestan, Aguilar, Tenerife, Nant-y-glo, Penybont;*
(b) *Krugerrand, Pietermaritzburg, Jespersen, Kattegat, Hannover, Königsberg, papabile, pianissimo, Amundsen, Aeroflot, Taganrog, tsarevich, Zapotec, Ericsson.*

However, it appears that words from specific source languages have stress characteristics that differ radically from those of loanwords from other source languages and this is particularly attested in words of three syllables (and more). For example, loanwords from Swedish in (7a) are stressed differently from those of Greek origin in (7b):

- (7) (a) 'Ericsson, 'Helsingfors, 'Lagerlöf, 'Norrköping, 'Swedenborg, 'Valdemar, 'Wallenberg, 'Uppsala⁶⁶;
(b) bou 'zouki, Kó 'rinthos, sou 'vlaki, ta 'verna, tzat 'ziki.

Loanwords from Swedish are stressed on the antepenultimate while loanwords from Greek are stressed on the penultimate. It means that source languages might have an influence on the stressing of loanwords in English and that the overall results from Figure 3 need to be refined as it seems essential to isolate the different languages of origin. Therefore, the penultimate stress pattern is usually attested in loanwords in English and it seems to confirm the hypothesis according to which this is the preferential stress pattern in loanwords, but words from specific source languages adopt distinct stress features. It is therefore necessary to distinguish between source languages.

3.2 The stressing of loanwords from major source languages.

In order to determine whether loanwords from some source languages are associated to specific stress patterns which can even be distinct from penultimate stress, it is not necessary to study source languages with few items only as results on such restricted sub-corpora would not be significant. Only source languages with at least 20 items are analysed here, which correspond to seven source languages in the dataset (see Table 1): Italian, Russian, Spanish, Welsh, German, Danish and Swedish. First, the distribution of stress patterns in disyllabic loanwords will be considered in 3.2.1.

⁶⁶ This item can also be stressed on the penultimate syllable.

3.2.1 The stressing of disyllabic loanwords.

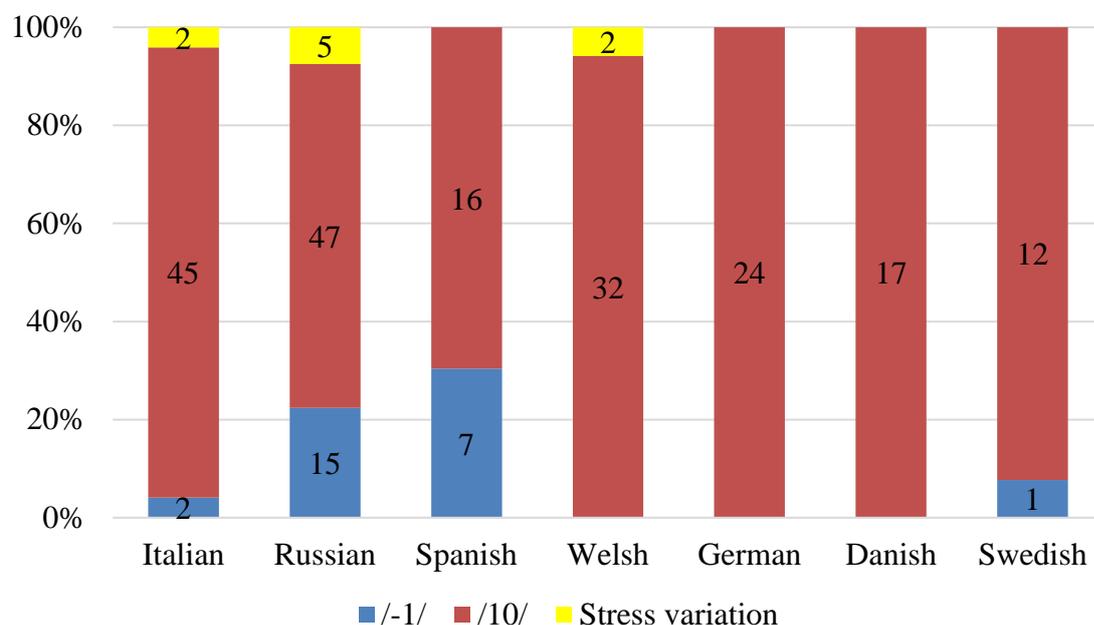


Figure 4: Distribution of stress patterns in disyllabic loanwords from major source languages.

The results from Figure 4 show that disyllabic loanwords coming from Italian, Welsh, German, Danish and Swedish are massively stressed in /10/. This is also the dominant stress pattern in loanwords coming from Russian and Spanish, but the proportion of final stress patterns is nonetheless significant with respectively 22.4% and 30.5%. Examples in (8) illustrate the regular behaviour of disyllabic loanwords coming from Italian (8a), Welsh (8b), German (8c), Danish (8d) and Swedish (8e), which are massively if not all of them stressed in /10/:

- (8) (a) *pasta, pesto, poco, scena, sempre*;
 (b) *Arfon, Dyfed, Elan, Ifor, Mynwy*;
 (c) *Augsburg, Holstein, Leipzig, Ostwald, Steichen*;
 (d) *Aarhus, Bering, Erlang, Nielsen, Verner*;
 (e) *Bergman, gyttja, krona, Nilsson, Rehnquist*.

Examples of disyllabic loanwords coming from Russian (9a) and Spanish (9b) and which are stressed on the final syllable whose rate is quite significant in Figure 4 are listed below:

- (9) (a) *Abkhaz, Baikal, kolkhoz, pirog, Tashkent*;
 (b) *Bilbao, Falange, Jerez, La Paz, Ortiz*.

These results show that the overall results from Figure 3 have to be dealt with carefully because final stress patterns are not attested in all disyllabic loanwords in English but rather restricted to words coming from specific source languages. This stress pattern is attested in disyllabic loanwords coming from Russian and Spanish but is rare in loanwords from other languages. The distribution of stress patterns in loanwords of three syllables (or more) will now be considered in 3.2.2.

3.2.2 The stressing of loanwords of three syllables (or more).

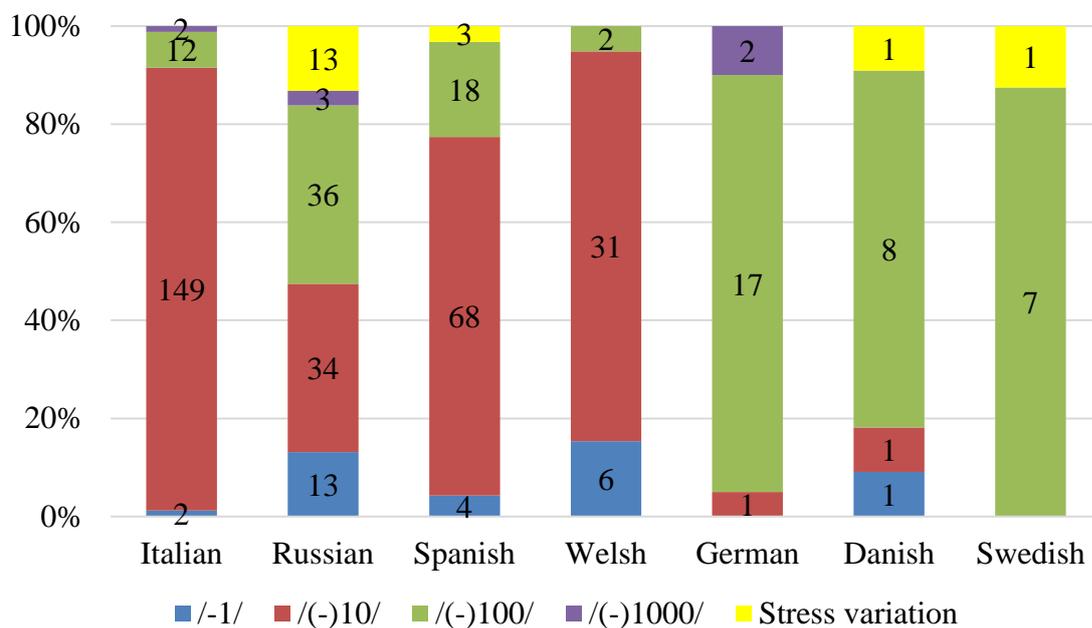


Figure 5: Distribution of stress patterns in trisyllabic (or more) loanwords from major source languages.

The results from Figure 5 show that loanwords of three syllables (or more) coming from Italian and Welsh are generally stressed on the penultimate syllable (90.3% in loanwords from Italian and 79.5% in loanwords from Welsh). This stress pattern is illustrated in (10a) with loanwords from Italian and in (10b) with loanwords from Welsh:

- (10) (a) *Fellini, lasagne, Lugano, maestoso, panettone*;
 (b) *Abersychan, Aneurin, eisteddfod, Llewelyn, Pontardulais*.

The penultimate stress pattern is less dominant in loanwords coming from Spanish with 73.1% of words which are stressed on the penultimate. Consequently, the proportion of antepenultimate stress is significant (19.4%). The results concerning final stress show that this

pattern is rare in loanwords of three syllables or more. There is a small proportion of loanwords from Russian (13.1%) and Welsh (15.4%) which are stressed on the final syllable, which is illustrated with loanwords from Russian (11a) and Welsh (11b):

- (11) (a) *Azerbaijan, Kazakhstan, Novosibirsk, Tadjikistan, Turkmenistan*;
(b) *Aberfan, Penybont, Pontypridd, Talybont, Tal-y-llyn*.

The most striking feature from Figure 5 is the distribution of antepenultimate stress in loanwords. Indeed, it appears that loanwords coming from German, Danish and Swedish are massively stressed on the antepenultimate syllable. Examples in (12) illustrate this phenomenon with words from German in (12a), Danish in (12b) and Swedish in (12c):

- (12) (a) *Autobahn, Fassbinder, Hannover, Rosenwald, Saarbrücken*;
(b) *Andersen, Jespersen, Kattegat, Odense, Rasmussen*;
(c) *Ericsson, Lagerlöf, Swedenborg, Valdemar, Wallenberg*.

The distribution of stress patterns in loanwords from Russian is difficult to interpret as there is no real preferential pattern. Nonetheless, considering the huge proportion of loanwords coming from Russian in the dataset, the proportion of antepenultimate stress is fairly significant (36.3%) as it corresponds to 36 items. The loanwords coming from Russian in (13) are all stressed in /(-)100/:

- (13) *Aeroflot, Bolchevik, Gorbachev, Kaliningrad, Stalingrad*.

Therefore, even if penultimate stress is dominant among loanwords in English, a closer look at the data concerning the major source languages attested in the dictionary shows that the distribution of penultimate stress is not uniform. Loanwords coming from Italian and Welsh (this applies to loanwords from Spanish as well) are usually stressed on the penultimate syllable and the proportion of items from these languages is so important in the dataset that the tendency for loanwords from all languages to be stressed on the penultimate syllable has been in a way ‘overestimated’ in the overall results. It appears that words from German, Danish, Swedish (and also from Russian to a lesser extent) are more likely to be stressed on the antepenultimate syllable. Besides, a massive proportion of antepenultimate stress is also found in loanwords coming from less attested source languages. Antepenultimate stress is the dominant pattern in loanwords coming from Afrikaans (14a), Dutch (14b) or Turkish (14c):

- (14) (a) *'Krugerrand, Pieter 'maritzburg, Ve 'reeniging, Wit 'watersrand*;
(b) *Con 'certgebouw, 'Gelderland, 'Groningen, 'Limburger, 'Nijmegen, 'Suriname*;
(c) *'Atatürk, 'baklava*.

There is thus a clear disparity in the distribution of stress patterns in English loanwords and that this disparity relies on the origin of loanwords. Words coming from Italian are not stressed the same way that those coming from Welsh or German. Antepenultimate stress is not that marginal and is actually the preferential pattern attested in words coming from specific source languages. Therefore, if the language of origin is a significant parameter in the stressing of loanwords in English, it is essential to compare the location of stressed syllables in source languages and in loanwords in English because the distribution of stress patterns attested in English may result from the reproduction of the stress patterns from source languages. Examples in (15) definitely show that there are correspondences between the location of stressed syllables in English loanwords and the location of stressed syllables in source languages. Final stress matches are attested (15a), as well as penultimate stress matches (15b) or antepenultimate ones (15c) suggesting that this phenomenon is not only restricted to penultimate stress:

- (15) (a) *van Gogh* /van'gɒf/ from Dutch /van'xɔx/;
Sinn Feinn /ʃɪn'feɪn/ from Irish /,sɪnɪ'fʲe:nɪ/;
Bakhtin /bak'ti:n/ from Russian /bax'tʲin/;
Kap Farvel /,kɑpfa:'vɛl/ from Danish /,kɑbfa'vɛl/;
Nant-y-glo /,nantə'gləʊ/ from Welsh /,nantʌ'glɔ/;
- (b) *Botha* /'bɒtə(r)/ from Afrikaans /'bʊəta/;
Hašek /'hɑʃɛk/ from Czech /'hʌʃɛk/;
bodega /bə(ʊ)'di:gə(r)/ from Spanish /bo'deyɑ/;
Stavanger /stə'vɑŋə(r)/ from Norwegian /stav'aŋer/;
zucchini /zʊ'ki:ni/ from Italian /tsuk'kini/;
- (c) *Löwenbräu* /'ləʊənbraʊ/ from German /'lœ:vŋbrɔy/;
Ericsson /'ɛrɪksn/ from Swedish /'e:riksɔn/;
Málaga /'maləgə(r)/ from Spanish /'malayɑ/;
Rachmaninov /rak'manɪnɒf/ from Russian /rax'man'inɒf/;
piccolo /'pɪkələʊ/ from Italian /'pikkolo/.

The influential role of the preservation of the stress patterns observed in source languages on the stressing of loanwords in English might explain why loanwords coming from same source languages share the same stress properties. The next section investigates this assumption with a stress match analysis.

4. Stress match analysis between source languages and English.

Before giving the results of the stress match investigation, the terminology used in the figures below primarily needs to be described. The stress match is ‘total’ when the correspondence between the location of the stressed syllable in the source language and in English is identical. The ‘total match’ category brings together stress correspondences attested on last syllables (16a), penultimate syllables (16b), antepenultimate syllables (16c) or even pre-antepenultimate ones (16d):

- (16) (a) *Nuevo León* /,nwɛɪvəʊli:ˈn/ from Spanish /,nwɛβoleˈon/;
(b) *Mato Grosso* /,matəʊˈgrɒsəʊ/ from Portuguese /,matuˈgrɒsu/;
(c) *Swedenborg* /ˈswi:dnɒ:ɡ/ from Swedish /ˈsve:dnɒrj/;
(d) *Peenemunde* /ˈpi:nə,mʊndə(r)/ from German /ˈpe:nəmʏndə/.

There is ‘no match’ when the location of the stressed syllable in the source language and in English is different as in (17):

- (17) *Bratislava* /,bratɪˈslɑ:və(r)/ from Czech /ˈbrʌtɪslʌvʌ/;
Osnabrück /ˈɒznəbrʊk/ from German /ɔsnəˈbrʏk/;
matador /ˈmatədɔ:(r)/ from Spanish /mataˈðor/.

Finally, the match is ‘partial’ when there are stress variants but with at least one stress pattern shared by source languages and English like in (18):

- (18) *Kasparov* /ˈkaspərɒv/ or /kaˈspɑ:rɒv/ from Russian /kaˈsparəf/.

When there are two primary stresses in the transcriptions of source languages as sometimes appears in Italian compounds (example: *sotto voce* /,sɒtəʊˈvəʊtʃi/ from Italian /ˈsottoˈvotʃe/), only the rightmost Italian stress has been taken into account and compared with the primary stress in English. Therefore, in *sotto voce*, the stress match is total. The global results of the stress match investigation are presented in 4.1.

4.1 Overall results of the stress match investigation.

The stress match investigation is carried out on 278 disyllabic loanwords and 477 trisyllabic (or more) loanwords. The overall results are presented in Figure 6.

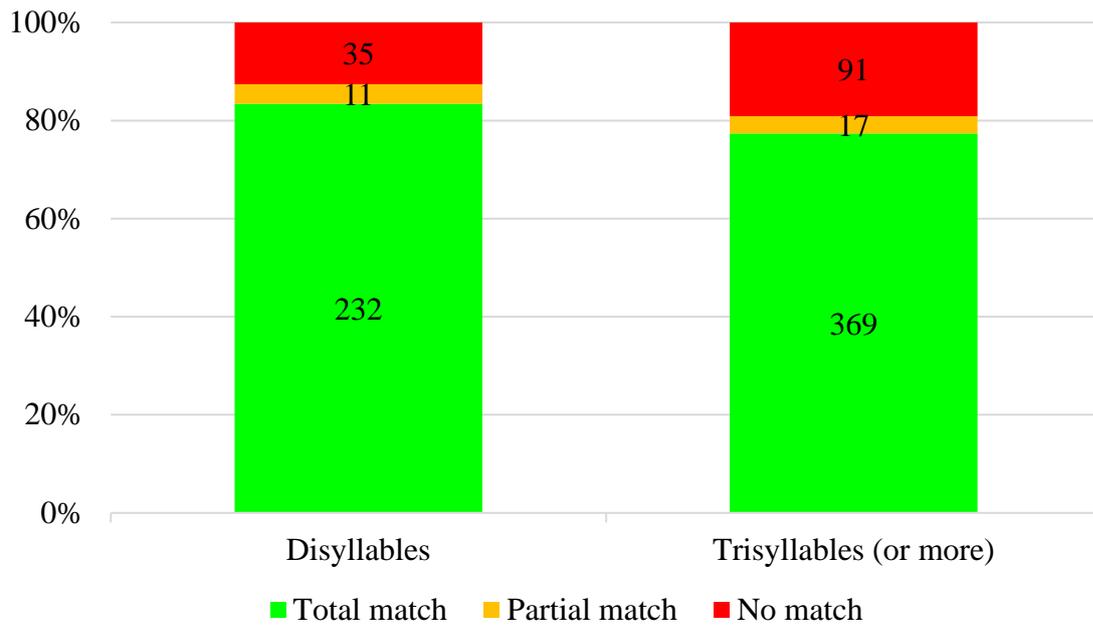


Figure 6: Stress match rates between source languages and English loanwords.

The results show that 83.5% (232 out of 278) of disyllabic loanwords in English are stressed the same way as the items in source languages. Only 12.6% (35 out of 278) of them have distinct stress patterns from those of source languages. This tendency is confirmed by the results found in loanwords of three syllables (or more). Indeed, the location of primary stress is identical in the source language and in English in 369 out of 477 loanwords (77.4%). Almost 4 loanwords out of 5 reproduce the original stress pattern during the borrowing process. Therefore, one cannot deny that there is a direct phonological influence of items from source languages on to the primary stress location in corresponding English loanwords. However, the previous section has shown that the global results needed to be refined because of the distinct stress characteristics found in source languages. One may wonder if the preservation of the original stress pattern is effective whatever the source language may be. In order to prove this, these preliminary results are refined by separating the stress match rate results calculated for each source language, dealing first with disyllabic items in 4.2.1 and with trisyllabic (and more) words in 4.2.2.

4.2 Stress match rate results for each source language.

4.2.1 Disyllabic words.

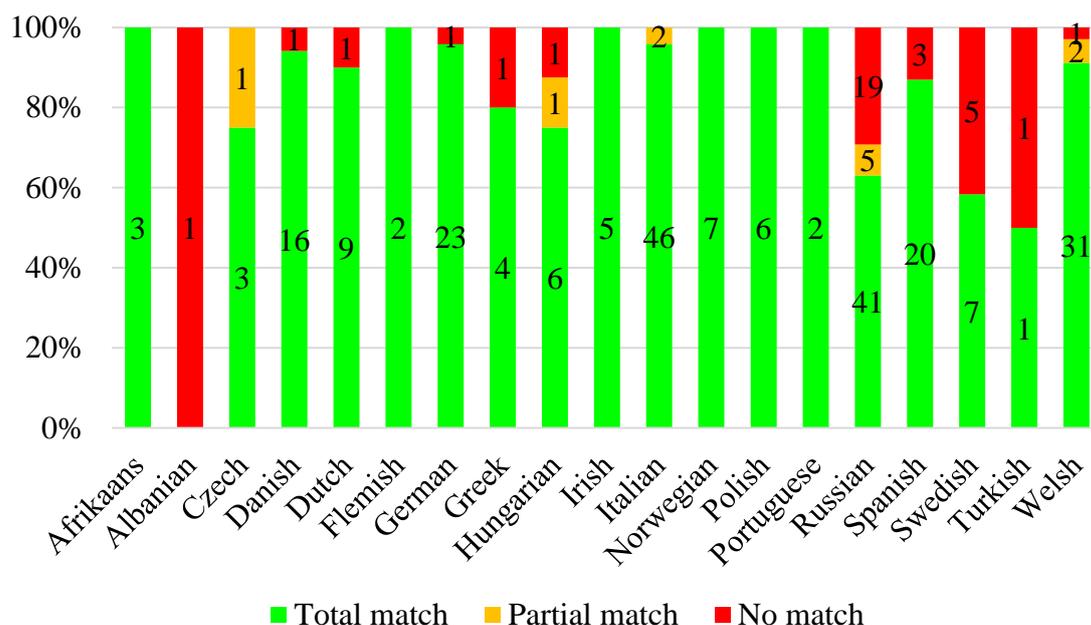


Figure 7: Stress match rate results in disyllabic items for each source language.

Concerning disyllabic words, the results from Figure 7 show that the match between the stress patterns of source languages and the stress patterns attested in English is perfect or close-to-perfect for words coming from Afrikaans, Czech, Danish, Dutch, Flemish, German, Greek, Hungarian, Irish, Italian, Norwegian, Polish, Portuguese, Spanish and Welsh. The results concerning words coming from Albanian, Russian, Swedish and Turkish are less convincing as it appears that there is no systematic correspondence between the location of the stressed syllable in the source language and in English. However, there are only few items coming from Albanian (*Tomor* = /10/ in English and /01/ in Albanian) and Turkish (*Iznik* = /10/ in English and /01/ in Turkish), so these preliminary findings have to be dealt with carefully. The results concerning loanwords coming from Swedish show that the English loanwords can sometimes (19a), but not always (19b) be stressed in the same way as in Swedish:

(19) (a) *gyttja, krona, Malmö, Nilsson, Öland, öre, Vättern* = /10/ in Swedish and in English;

(b) *Alfvén, Bergman, Bofors, Rehnquist, Strindberg* = /01/ in Swedish and /10/ in English.

The most interesting results are found in words from Russian because there are 41 items with total stress matches (examples in 20a) but also 19 items with no matches (examples in 20b):

(20) (a) *Baikal, Bakhtin, kolkhoz, Kurgan, pirog* = /01/ in English and Russian;
dacha, droshky, Duma, Gamow, glasnost = /10/ in English and Russian;

(b) *pogrom, shashlik, Soyuz, taiga, Tiflis, Tuva* = /10/ in English and /01/ in Russian; *Vitebsk* = /01/ in English and /10/ in Russian.

The results found in words of three syllables (or more) will be described in 4.2.2.

4.2.2 Words of three syllables (or more).

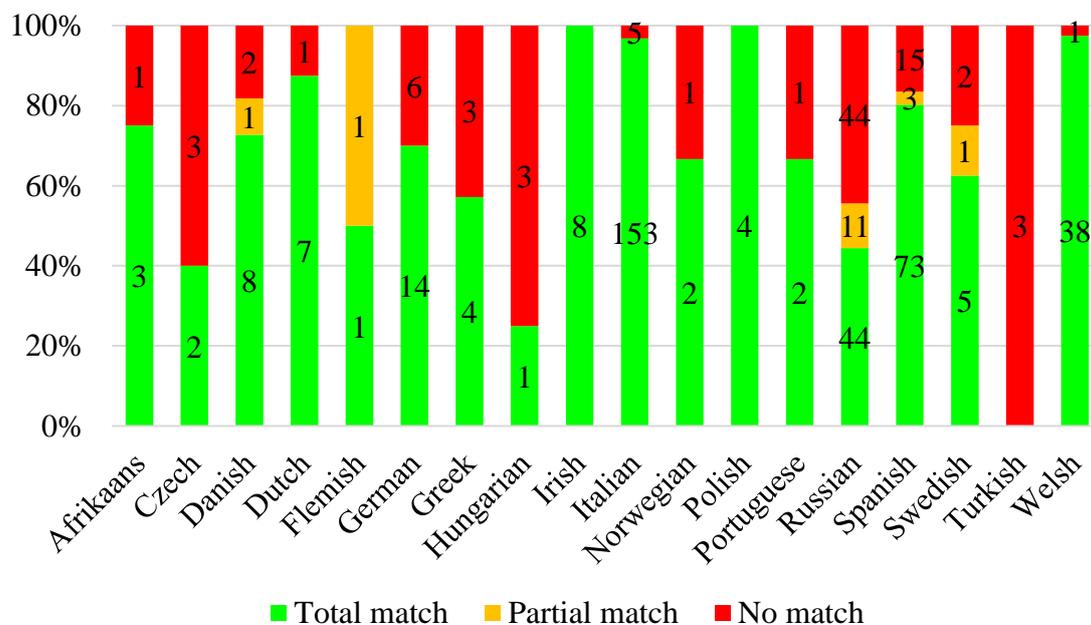


Figure 8: Stress match rate results in trisyllabic (or more) items for each source language.

The stress match rate is high and therefore significant between English and words coming from Afrikaans (21a), Danish (21b), Dutch (21c), Flemish (21d), Irish (21e), Italian (21f), Polish (21g), Swedish (21h) and Welsh (21i).

- (21) (a) *'Krugerrand, Wit'watersrand, Ve'reeniging;*
 (b) *ˌKap Far'vel, 'Andersen, 'Jespersen, 'Kattegat, 'Odense;*
 (c) *ˌvan de 'Velde, 'Groningen, 'Rijksmuseum;*
 (d) *'Passchendale;*
 (e) *ˌFine 'Gael, Cú'chulainn, ˌDáil 'Eireann, Dun 'Laoghaire, 'feiseanna;*
 (f) *im'pasto, Lo'renzo, pa'pabile;*
 (g) *ˌKato'wice, ˌPade'rewski, ˌPende'recki, ˌZako'pane;*
 (h) *'Ericsson, 'Lagerlöf, 'Swedenborg, 'Valdemar;*
 (i) *ˌTaly'bont, ˌTal-y-'llyn, ˌAber'aeron, ˌAber'sychan.*

There are only 12 loanwords coming from those source languages whose stress patterns

radically differ from the original ones (variation is not taken into account). The 12 cases are listed in (22):

- (22) ,*Pieter*'*maritzburg* (English) vs ,*Pieterma*'*ritzburg* (Afrikaans);
,*Erte*'*bölle* (English) vs '*Erte*,*bölle* (Danish);
'*Helsingör* (English) vs ,*Helsin*'*gör* (Danish);
'*Suriname* (English) vs ,*Suri*'*name* (Dutch);
'*auto*,*strada* (English) vs *auto*'*strada* (Italian);
'*auto*,*strade* (English) vs *auto*'*strade* (Italian);
,*mode*'*rato* (English) vs *modera*'*to* (Italian);
,*Quasi*'*modo* (English) vs *Qua*'*simodo* (Italian);
'*Romeo* (English) vs *Ro*'*meo* (Italian);
'*Helsingfors* (English) vs ,*Helsing*'*fors* (Swedish);
'*Norrköping* (English) vs *Norr*'*köping* (Swedish);
,*eistedd*'*fodau* (English) vs *ei*'*steddfodau* (Welsh).

Loanwords coming from Spanish tend to retain their original stress patterns (there are 73 stress matches out of 91 loanwords, i.e. 80%). The results are more difficult to interpret in loanwords coming from German, Greek, Norwegian and Portuguese, as there are just a few more cases for which stress matches are attested but the results are not clear-cut. Finally, there are very few stress matches concerning words coming from Czech, Hungarian, Russian and Turkish.

The results from Figures 7 and 8 make it possible to rank source languages from those for which the reproduction of the location of stressed syllables is clearly attested in English to those for which the original stress patterns is scarcely preserved in English, which means there is no correlation between the stressing of loanwords in English and the conservation of the original stress pattern. Moreover, there are source languages which are difficult to categorize because the results are not conclusive.

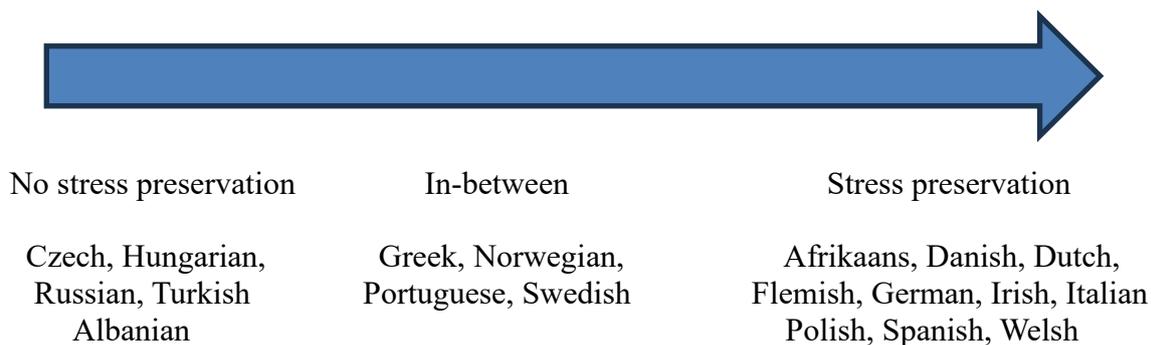


Figure 9: Classification of source languages according to the stress preservation principle with English loanwords.

It is impossible to determine from the results presented in Figures 7 and 8 if the stress match is effective whatever the stress patterns of source languages may be, because there are no details as to whether the distinct stress patterns attested in source languages are reproduced. One might object that the massive proportion of penultimate stress found in English loanwords is not the consequence of the preservation of the stress pattern of source languages but rather the application of the default stress pattern found in loanwords in English, i.e. the penultimate one. The hypothesis according to which there could be direct links between the stressing of loanwords in English and the stress patterns of source languages can only be retained if there are reproductions of final or antepenultimate stress patterns as well. In Section 4.3, the results from the preliminary stress match investigation are refined in order to observe the stress match rates among the different stress patterns attested in source languages.

4.3 Stress match investigation along with stress patterns from major source languages.

This detailed stress match investigation only focuses on source languages with at least 7 disyllabic items and stress variation is not included in order to obtain representative results. The corpus is thus composed of 232 disyllables and the results of the stress match analysis along with the distinct stress patterns found in source languages are presented in Figure 10.

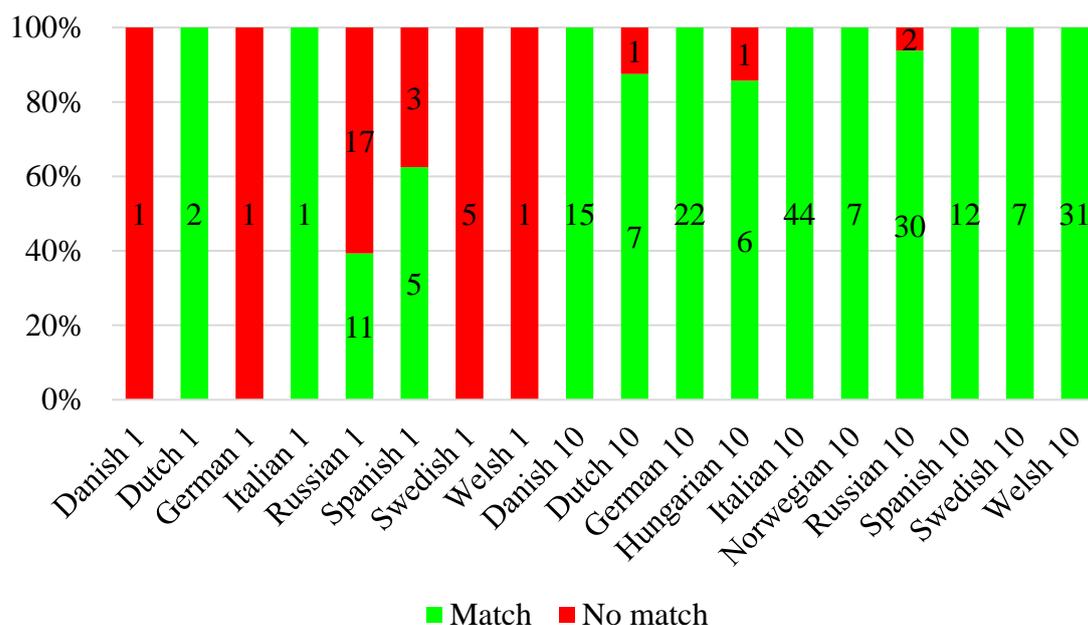


Figure 10: Stress match results in disyllabic items along with the stress patterns attested in source languages.

The stress patterns attested in corresponding source languages are indicated on the horizontal bar and positive matches mean that the original stress patterns are reproduced in English as well. If there are no matches, it means that English stress patterns differ from the ones found in source languages and which are mentioned on the horizontal bar.

In disyllabic words, the results show that there is usually a match concerning penultimate stress, which seems to confirm that the original penultimate stress pattern is almost systematically reproduced in English. There are only 4 items presented in (23) which contradict this phenomenon:

- (23) *Schie'dam* from Dutch ['sxi:dam];
Gá'bor from Hungarian ['gɑ:bɔr];
Mur'mansk from Russian ['murmənsk];
Vi'tebsk from Russian ['vʲitʲipsk].

The stress match rate on the penultimate syllable in disyllabic words is therefore very high (98%, i.e. 181/185 cases), but on the contrary the match rate concerning final stress is very low (40%, i.e. 19/47 cases). Final stress patterns attested in disyllabic items in source languages tend not to be reproduced in loanwords in English, which are massively stressed in /10/. There is a preferential /10/ stress pattern in English disyllabic loanwords and it seems to be inherited from the stress patterns of source languages which are massively stressed in /10/. Indeed, there

is a perfect or close-to-perfect /10/ stress match in disyllables coming from Danish, Dutch, German, Hungarian, Italian, Norwegian, Russian, Spanish, Swedish and Welsh.

As far as words of three syllables or more are concerned, only source languages with at least 7 items in English are conserved and stress variation as well as pre-antepenultimate stress patterns are excluded. The results are presented in Figure 11.

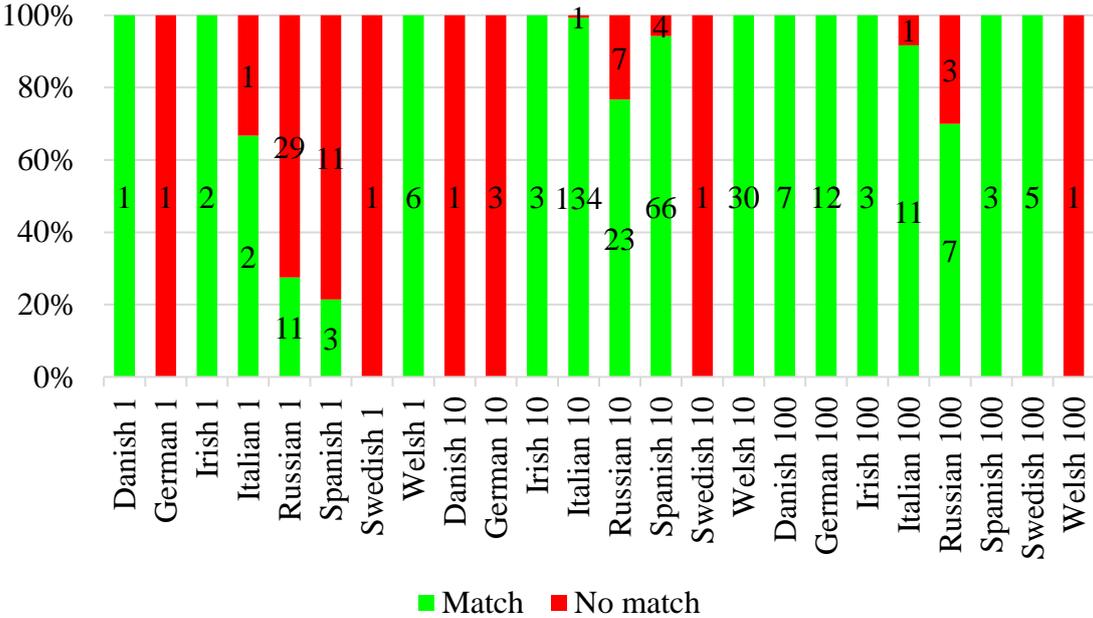


Figure 11: Stress match results in items of three syllables or more along with the stress patterns attested in source languages.

The results concerning the reproduction of final stress patterns from source languages in English show that the stress match rate is actually very low, which echoes the conclusions drawn from the results concerning disyllabic items. There is a perfect match for words coming from Danish, Irish and Welsh and this is also the tendency attested in words coming from Italian, but the number of items is very low. There is a very small proportion of Russian and Spanish words whose final stress patterns are reproduced in English loanwords. The proportion of loanwords which is stressed differently from the original final stress patterns is actually more important. The final stress preservation is only effective in 37% of cases (25 matches out of 68).

Once again, the penultimate stress patterns attested in source languages are generally reproduced in English loanwords. The match rate is perfect or close-to-perfect in words coming from Irish, Italian, Russian, Spanish and Welsh. Surprisingly, this tendency is not attested in the few words coming from Danish, German and Swedish which are all stressed on the

antepenultimate in English (examples: *Hannover*, *Helsingor*, *Kristallnacht*, *Norköpping*, *Saarbrücken*). The proportion of penultimate stress preservation is substantial (94%, i.e. 256 matches out of 273). Therefore, the penultimate stress patterns attested in English loanwords are generally inherited from source languages.

As far as antepenultimate stress patterns attested in English loanwords is concerned, it appears that they are also massively inherited from source languages because there is a 91% match (i.e. 48 matches out of 53). This striking result was not anticipated and definitely proves that source languages have an impact on the stressing of loanwords in English. The stress match is perfect or so in words coming from Danish, German, Irish, Italian, Spanish and Swedish. Once again, the stress preservation principle is not effective concerning words of Russian origin and the only Welsh item with an antepenultimate stress is stressed on the penultimate syllable in English (*eisteddfodau*).

5. Discussion and conclusions.

The results of this study clearly demonstrate that loanwords are massively stressed on the penultimate syllable in English. This stress pattern is attested in disyllabic loanwords but also in loanwords of three syllables or more. This confirms that the penultimate stress pattern is the preferential one found in loanwords in English and echoes the assumption found in Dabouis & P. Fournier (2022) concerning the existence of a preferential penultimate stress which is massively shared by loanwords within the §Foreign subsystem.

However, the study reveals distinct stress features among loanwords in English, which are directly linked to source languages. Final stress is marginal in English loanwords and mainly attested in disyllabic and trisyllabic (and more) loanwords coming from Russian and Spanish. This echoes the argument according to which final stress is mainly found in words of French origin. By contrast, the proportion of antepenultimate stress in English loanwords is fairly significant (around 20%) and this stress pattern is extensively found in loanwords coming from German, Swedish, Danish and Russian. Therefore, the origins of loanwords might be an indicator of the way such words should be stressed in English. Indeed, if antepenultimate stress is mainly found in specific source languages, this is also the case for penultimate stress. Loanwords coming from Italian, Welsh and Spanish are massively stressed on the penultimate. The distribution of stress patterns in English loanwords seems to be directly linked to the stress characteristics of source languages.

The results of the comparison between the location of stressed syllables in source languages

and in English loanwords are striking. The preservation of the prosodic characteristics of source languages is definitely a determining parameter accounting for the stressing of loanwords in English. However, the results need to be refined because there are huge disparities among source languages. There is a perfect or close-to-perfect stress match concerning words coming from Afrikaans, Danish, Dutch, Flemish, German, Irish, Italian, Polish, Spanish and Welsh. On the contrary, the assumption that the location of the stressed syllable in English loanwords is derived from the original stress patterns does not apply to words coming from Czech, Hungarian, Turkish, Albanian and Russian. The stress match rate is so low that the preservation principle cannot account for the stressing of such loanwords. The results are not convincing either concerning loanwords coming from Greek, Norwegian, Portuguese and Swedish.

Considering that the reproduction principle is particularly attested in words coming from source languages which are great purveyors of words of foreign origin in English, like Italian or Spanish, this explains why the overall results of the stress match rate are so high. One cannot deny with such an extensive stress match rate that source languages are not influential on the stressing of loanwords in English. The proportion of stress match would even be higher if words of Russian origin were excluded from the dataset. Indeed, many loanwords of Russian origin in English are stressed differently from the original stress patterns found in Russian. As many loanwords of Russian origin are present in the dataset, this greatly influences the overall results by overrepresenting the “no match” category, when this is obviously mostly restricted to words of Russian origin.

The results show that final stress patterns found in source languages are generally not preserved in English in favour of penultimate stress patterns which are generally attested in disyllabic and trisyllabic (or more) loanwords. Unsurprisingly, penultimate stress from source languages is massively reproduced in English, but this feature was highly anticipated as penultimate stress is the preferential one in English loanwords. On the contrary, such a massive correspondence concerning antepenultimate stress was not expected. If the massive reproduction of stress patterns inherited from source languages was only attested for penultimate stress, one might object that this could actually be the result of the application of the preferential stress pattern on the penultimate and absolutely not linked to reproduction of the stress properties of source languages. But these findings concerning the high proportion of antepenultimate stress directly inherited from source languages are valuable as we can now legitimately conclude that source languages have a determining impact on the stressing of loanwords in English.

The results might also explain why penultimate stress is the preferential one in loanwords

in English and why English speakers, when facing words of foreign origin tend to stress the penultimate as the experiment conducted on Italian words has shown (P. Fournier 2018). This preferential pattern might have been derived from the massive proportion of words of foreign origin stressed on the penultimate which integrated the English lexicon through time. This penultimate stress has been massively preserved in English loanwords because, as this study shows, there is a strong tendency to reproduce the original stress patterns in English, and in turn, penultimate stress has been used to stress words of foreign origin in English. It is now used as a phonological marker to signal the foreign origin of items. It might also explain why words of Russian origin are massively stressed on the penultimate in English and do not reproduce the original stress patterns attested in Russian.

The impact of source languages on the stressing of loanwords in English is undeniable but when this preservation principle is not respected, the preferential penultimate stress is used by default in English loanwords. There is actually an ongoing conflict in English between the preservation of the stress patterns inherited from source languages and the application of the preferential penultimate stress. The application of this preferential stress sometimes generates stress differences between source languages and English.

However, it is hard to explain why there is a perfect or close-to-perfect reproduction of the original stress patterns from some source languages but not from others. In the introduction, it is suggested that loanwords are introduced in English by bilingual speakers who can use either the mechanisms ruling the source language or the target language to pronounce them. This position definitely echoes the two scenarios highlighted in this analysis: the preservation of the stress pattern from source languages or the application of the preferential penultimate stress characteristic of the English system. The usual preservation of original stress patterns attested in this study confirms that loanwords infiltrate English through bilingual speakers who can therefore reproduce the stress patterns of source languages. This is highly probable considering the principle of stress deafness which is shared by non-bilingual speakers. Therefore, if loanwords were introduced in English by non-bilingual speakers, such stress match rates would not exist. Could this mean that loanwords of Russian origin are introduced in English by non-bilingual speakers considering the low stress match rate attested? Probably not, as recent research suggests (Kang 2010; Davis, Tsujimura & Tu 2012) suggest that the hypothesis of different prosodic systems can account for such absence of stress preservation. In other words, the prosodic characteristics of the English and Russian systems would be so different that it might be impossible for speakers to reproduce the Russian stress pattern in English. This argument echoes the results found in P. Fournier (2016) which show that the application of the

penultimate preferential stress pattern in loanwords of Japanese origin is the consequence of the incompatibilities between the English and the Japanese stress systems. Their fundamental characteristics are so different that stress preservation turns out to be impossible. A study of the prosodic characteristics of source languages would therefore be valuable in order to determine if the preservation of the original stress patterns in English can be accounted for with this parameter.

To conclude, this study shows that the phonological properties of source languages and more precisely their stress characteristics have a fundamental impact on the stressing of loanwords in English. The results confirm that the stress patterns of source languages are generally reproduced in English, which accounts for the significant proportion of antepenultimate stress in the dataset. However, stress variation shows that there is a conflict between stress preservation and the application of the preferential stress pattern in English. English loanwords are generally stressed on the penultimate and this is a phonological indicator used in English to illustrate the foreign origin of items. These preliminary results are based on British English only, and American English should be investigated as well. It is also necessary to work on a larger dataset because quite surprisingly, the phonetic transcriptions of major source languages are not attested in the *Routledge* dictionary (examples: Arabic, Chinese, Japanese). Finally, it is essential in future studies to test the influence of the preservation stress hypothesis along with other determining parameters of the English system (examples: segmental structure, grammatical category, syllable weight) in order to determine its efficiency.

References:

- Broselow, Ellen (2009), "Stress adaptation in loanword phonology: Perception and learnability", *Phonology in Perception* (15), 191-234.
- Burzio, Luigi (1994), *Principles of English Stress*, New York: Cambridge University Press.
- Calabrese, Andrea & W. Leo Wetzels (eds.) (2009), *Loan Phonology*, Amsterdam and Philadelphia: John Benjamins Publishing.
- Carney, Edward (1994), *A Survey of English Spelling*, London: Routledge.
- Chomsky, Noam & Morris Halle (1968), *The Sound Pattern of English*, New York: Harper and Row.
- Church, Kenneth (1985), "Stress assignment in letter-to-sound rules for speech synthesis", *The Journal of the Acoustical Society of America* 78 (S1).
- Dabouis, Quentin & Pierre Fournier (2022), "English phonologieS?", *Modèles et modélisation en linguistique / Models and Modelisation in Linguistics*, V. Arigne & C. Rocq-Migette. (eds.), Brussels, Belgium: Peter Lang Verlag, 215-258.
- Dabouis, Quentin, Fournier, Jean-Michel, Fournier, Pierre & Marjolaine Martin (2023), "English word stress and the Guierrian School", *New Perspectives on English Word Stress*, E. Yamada, A. Przewozny, J.-M. Fournier & N. Ballier (eds.), Edinburgh University Press, 53-82.
- Dabouis, Quentin & Pierre Fournier (2024), "Stress in French loanwords in British and

- American English”, *Journal of Linguistics*, 1-26.
- Davis, Stuart, Tsujimura, Natsuko & Jung Yueh Tu (2012), “Toward a taxonomy of loanword prosody”, *Catalan Journal of Linguistics*.
- Di Cristo, Albert & Daniel Hirst (1993), “Rythme syllabique, rythme mélodique et représentation hiérarchique de la prosodie du français”, *Travaux de l’Institut de Phonétique d’Aix 15*, 9-24.
- Fitt, Susan (1996), “Syllable weight, stress patterns and nativisation”, *Proceedings of the Edinburgh Linguistics Department Conference*, University of Edinburgh, 40-51.
- Fournier, Jean-Michel (2007), “From a Latin syllable-driven stress system to a Romance versus Germanic morphology-driven dynamics: in honour of Lionel Guierre”, *Language Sciences* 29 (2-3): 218-236.
- Fournier, Jean-Michel (2010), *Manuel d’anglais oral*, Paris: Ophrys.
- Fournier, Pierre (2016), “Nouvelles perspectives sur l’accentuation des emprunts en anglais contemporain”, *Itinéraires 2015*(2).
- Fournier, Pierre (2018), “Stress assignment in Italian loanwords in English and its impact on the stressing of foreign words by native English speakers”, *Corela* 24.
- Guierre, Lionel (1979), *Essai sur l’accentuation en anglais contemporain: éléments pour une synthèse*, Doctoral dissertation, Université de Paris-VII.
- Hammond, Michael (1999), *The Phonology of English: A Prosodic Optimality-Theoretic Approach*, Oxford: Oxford University Press.
- Hyman, Larry (1970), “The role of borrowing in the justification of phonological grammars”, *Studies in African Linguistics* 1 (1): 1-48.
- Hyman, Larry (1985), *A Theory of Phonological Weight*, Foris: Dordrecht.
- Itô, Junko & Armin Mester (1999), “The phonological lexicon”, *The Handbook of Japanese Linguistics*, N. Tsujimura (ed.), Malden, MA and Oxford, UK: Blackwell Publishers, 62- 100.
- Jacobs, Haike & Carlos Gussenhoven (2005), “Loan Phonology: salience, the lexicon and Optimality Theory”, *Linguistic Inquiry* 36 (2): 223-258.
- Kang, Yoonjung (2010), “Tutorial overview: suprasegmental adaptation in loanwords”, *Lingua* 120 (9): 2295-2310.
- Kenstowicz, Michael (2007), “Salience and similarity in loanword adaptation: a case study from Fijian”, *Language Sciences* 29 (2-3), 316-340.
- Kubozono, Haruo, Itô, Junko & Armin Mester (2008), “Consonant gemination in Japanese loanword phonology”, *18th International Congress of Linguistics*, Seoul, 953-973.
- Minkova, Donka (2006), “Old and middle English prosody”, *The Handbook of the History of English*, A. van Kemenade & B. Los (eds.), Oxford: Blackwell Publishing, 95-125.
- Myers-Scotton, Carol (1993), “Common and uncommon ground: social and structural factors in codeswitching”, *Language in Society* 22 (4): 475-503.
- Paradis, Carole & Darlene LaCharité (2005), “Category preservation and proximity versus phonetic approximation in loanword adaptation”, *Linguistic Inquiry* 36 (2): 223-258.
- Paradis, Carole & Darlene LaCharité (2008), “Apparent phonetic approximation: English loanwords in Old Quebec French”, *Journal of Linguistics* 44 (1), 87-128.
- Peperkamp, Sharon (2004), “A psycholinguistic theory of loanword adaptations”, *Annual Meeting of the Berkeley Linguistics Society*, 341-352.
- Peperkamp, Sharon & Emmanuel Dupoux (2008), “A typological study of stress ‘deafness’”, *Laboratory Phonology* 7, 203-240.
- Silverman, Daniel (1992), “Multiple scansion in loanword phonology: evidence from Cantonese”, *Phonology* 9 (2), 289-328.
- Upton, Clive & William A. Kretzchmar Jr (2017), *The Routledge dictionary of pronunciation for current English*, Taylor & Francis.

Appendix: Final dataset (768 loanwords) extracted from *Routledge 2nd edition*

Afrikaans (7 items): disyllables = *Botha, duiker, Kruger*; trisyllables and more = *Krugerrand, Pietermaritzburg, Vereeniging, Witwatersrand*.

Albanian (1 item): disyllable = *Tomor*.

Czech (9 items): disyllables = *Brno, háleř, Hašek, Praha*; trisyllables and more = *Bratislava, Janáček, koruna, Navrátilova, Novotný*.

Danish (28 items): disyllables = *Aalborg, Aarhus, Ålborg, Århus, Bering, Blixen, Bornholm, Brahe, Dagmar, Erlang, Esbjerg, Falster, Herning, Nielsen, Roskilde, Tórshavn, Verner*; trisyllables and more = *Andersen, Buxtehude, Ertebølle, Helsingør, Jespersen, Kap Farvel, Kattegat, Kierkegaard, Odense, Petersen, Rasmussen*.

Dutch (18 items): disyllables = *Gouda, gulden, Huygens, Limburg, Rembrandt, Schiedam, van Gogh, Vermeer, Zeeland, Zwolle*; trisyllables and more = *Concertgebouw, Gelderland, Groningen, Limburger, Nijmegen, Rijksmuseum, Suriname, van de Velde*.

Flemish (4 items): disyllables = *Leuven, Oostende*; trisyllables and more = *Passchendaele, Zeebrugge*.

German (44 items): disyllables = *Aachen, Adler, Augsburg, Blücher, Diekirch, Duisburg, Gasthof, Hänsel, Heilbronn, Holstein, Karlsbad, Kirchhoff, Kroeber, Leichhardt, Leipzig, Müller, München, Münster, Ostwald, Rügen, Seifert, Steichen, Steuben, Würzburg*; trisyllables and more = *Ansbacher, Autobahn, Baden(-)Württemberg, Baedeker, Erzgebirge, Fassbinder, Gasthöfe, Götterdämmerung, Hannover, Heldentenor, Königsberg, Kristallnacht, Löwenbräu, Niemöller, Osnabrück, Peenemunde, Rosenwald, Saarbrücken, Seligman, Württemberg*.

Greek (12 items): disyllables = *boule, dromoi, Khíos, Pátmos, Vólos*; trisyllables and more = *Áyios Nikólaos, bouzouki, Kórinthos, Seféris, souvlaki, taverna, tztziki*.

Hungarian (12 items): disyllables = *Bartók, csárdás, forint, Gábor, Kádár, Kodály, Korda, Lehár*; trisyllables and more = *Balaton, Budapest, Fertő Tó, Ilona*.

Irish (13 items): disyllables = *Éire, Gaeltacht, Samhain, Sinn Fein, slàinte*; trisyllables and more = *Cúchulainn, Dáil Eireann, Dun Laoghaire, feiseanna, Fianna Fáil, Fine Gael, Oireachtas, Tánaiste*.

Italian (214 items): disyllables = *Arno, Cavour, Croce, Fara, Gigli, Gobbi, Guardi, Lana, lento, lira, Loren, maestro, molto, Parma, pasta, penne, pesto, Pisa, Pitti, pizza, poco, presto, primo, putti, putto, Quattro, Rocco, Roma, Rossi, scampi, scena, scherzo, secco, sempre, signor, solo, squacco, stretto, stucco, tempi, tempo, tondi, tondo, Trieste, troppo, uno, Vespa, Vico, Zante*; trisyllables and more = *Abruzzi, Alberti, Albinoni, Amalfi, Amati, Annigoni, Antonioni, Armani, arrivederci, Ascona, autostrada, autostrade, Avogadro, Bugatti, ciabatta, condottiere, Coppola, Così Fan Tutte, da Vinci, Donizetti, Farnese, Fellini, figurante, Firenze, Galvani, Griselda, impasto, Lambretta, Lambrusco, larghetto, lasagne, legato, libretto, Livorno, Locarno, Lorenzo, Lugano, ma non troppo, Machiavelli, maestoso, Maggiore, manicotti, Manzoni, minestrone, moderato, obbligato, Pacino, Paganini, Pagliacci, Palermo, panettone, panforte, Panini, Pantelleria, papabile, paparazzi, Parmigianino, partita, Pavarotti, pecorino, pentimenti, peperino, peperoni, Pestalozzi, pianissimo, piano nobile, pianoforte, piccolo, pietà, Pirandello, Piranesi, Pisano, pizzeria, pizzicato, podestà, polacca, polenta, Pontormo, portamenti, portolano, pozzolana, predella, prestissimo, prima ballerina, prima donna, primavera, Puccini, Quasimodo, quattrocento, Radice, rallentando, Ravenna, ravioli, Respighi, ricotta, Rimini, ripieno, Risorgimento, risotto, ritardando, ritenuti, ritornelli, Riviera, Romeo, Rossellini, Rossini, rubato, saltarelli, scaloppine, scherzando, Schiaparelli, sciocco, secondi, seicento, semplice, serenata, sforzando, sforzato, sfumati, sgraffiti, signora, signorina, simpatico, smorzando, solfatara, solfeggio, sonata, sonatina, soprano, sordini, sostenuto, sotto voce, spaghetti, spiccato, spumante, spumoni, staccato, stellini, stringendo*.

tagliatelle, tempera, tenuto, terramara, terrazzo, terza rima, terzetti, tessitura, Tetrizzini, toccata, tortellini, tramontana, trattoria, trecento, tutti-frutti, Uffizi, vaporetta, verismo, vermicelli, villeggiatura, Vivaldi, Zeffirelli, zingari, zucchetto, zucchini.

Norwegian (10 items): disyllables = *Bokmål, Hansen, Ibsen, Narvik, Nordkinn, Oslo, Trondheim*; trisyllables and more = *Amundsen, Lofoten, Stavanger*.

Polish (10 items): disyllables = *Gdańsk, Kielce, Kraków, Poznań, Wrocław, złoty*; trisyllables and more = *Katowice, Paderewski, Penderecki, Zakopane*.

Portuguese (5 items): disyllables = *Douro, Faro*; trisyllables and more = *Benguela, Carajás, Mato Grosso*.

Russian (166 items): disyllables = *Abkhaz, Altai, artel, Baikal, Bakhtin, Baku, Balkhash, Bokhara, Cheka, dacha, droshky, Duma, Gamow, glasnost, Gogol, gulag, Kalmuck, Karpov, Kaunas, Kazakh, Kharkov, Khrushchev, Kirov, kolkhoz, koumis, kulak, Kurgan, Kyrgyz, Lada, Laptev, Lena (forename), Lena (river), Lenin, Mikhail, Moskva, Murmansk, muzhik, Neva, Nevsky, Pavlov, Penza, pirog, podsol, pogrom, Pravda, Pushkin, Riga, Rurik, Sergei, shashlik, souslik, Soyuz, sputnik, suslik, Tadjik, taiga, Tashkent, Tiflis, troika, Tula, Tuva, ukase, Ukraine, Verny, Vitebsk, vodka, Yakut, zemstvo*; trisyllables and more = *Aeroflot, Aleksandrovsk, Alma-Ata, apparatchik, Ashkenazy, Ashkhabad, Astrakhan, Azerbaijan, babushka, balalaika, Birobidzhan, Bolshevik, Borodin, Borodino, Bukhara, Chelyabinsk, Cherenkov, Comintern, Ekaterinburg, Eugene Onegin, Gagarin, Godunov, Gorbachev, Gromyko, Horowitz, Kalinin, Kaliningrad, Kaluga, Kamenskoye, Kandinsky, Kasparov, Kazakhstan, Kerensky, Khabarovsk, Koluma, Komsomol, Korsakoff, Kostroma, Kosygin, Leninakan, Leningrad, Leonid, Lysenko, Maksutov, Mandelstam, Menshevik, Minkowski, Mussorgsky, Nabokov, nomenklatura, Novgorod, Novokuznetsk, Novosibirsk, Nureyev, Odessa, Orenburg, Pasternak, Pavlova, Petersburg, Petrograd, pierogi, Politburo, Potemkin, Przewalski, Rachmaninov, Raisa, Rasputin, Rimsky-Korsakov, Romanov, Rostropovich, Sakhalin, Samarkand, samizdat, samovar, samoyed, sastrugi, Sevastopol, Shevardnadze, Sholokhov, Stalingrad, Tadjikistan, Taganrog, tarantass, Tbilisi, tovarich, Trubetskoy, tsarevich, tsarevna, Turkestan, Turkmenistan, Ulanova, Ustinov, Vladimir, Vladivostok, Volgograd, Yekaterinburg, Yerevan, Yevtushenko*.

Spanish (116 items): disyllables = *Bilbao, Chavez, El Cid, Falange, Gaudí, Guzmán, Ifni, Jerez, Juárez, junta, La Paz, León (place in Spain), llano, López, Lorca, Méndez, mesa, Mixtec, Muñoz, Ortiz, Pérez, Ruiz, Vásquez*; trisyllables and more = *Acapulco, Aconcagua, Aguilar, Albacete, Alfonso, Algeciras, Alhambra, Alicante, Allende, Almería, alpargata, Altiplano, Alvarez, Ambato, anchoveta, Andorra, Antofagasta, Arequipa, Argentina, arroyo, Ávila, Ayacucho, Aymara, Ballesteros, Barcelona, Benidorm, bodega, Calderón, Cervantes, chaparajos, Chimborazo, Cochabamba, copaiba, coquina, cordillera, corozo, Cozumel, cuadrilla, Ecuador, García, González, guacharo, guanaco, guerrilla, Ibiza, jalapeño, Jiménez, jipijapa, jojoba, Juanita, La Coruña, La Jolla, Ladino, Lanzarote, Laredo, Las Palmas, llanero, Machu Picchu, Málaga, malaguena, Malvinas, mantilla, matador, Mendoza, Mercedes, Mexicali, Mudéjar, Mudéjares, Murillo, Nuevo León, Orinoco, Ortega, parador, Paraguay, Pinochet, Ramírez, Sajama, San Jacinto, San Salvador, Soledad, Tenerife, Tijuana, Torquemada, Uruguay, Valparaiso, Velásquez, Veracruz, vizcacha, Ximenes, Zacatecas, Zapata, Zapotec, Zaragoza*.

Swedish (21 items): disyllables = *Alfvén, Bergman, Bjorn, Bofors, gyttja, krona, Malmö, Nilsson, Öland, öre, Rehnquist, Strindberg, Vättern*; trisyllables and more = *Ericsson, Helsingfors, Lagerlöf, Norrköping, Swedenborg, Uppsala, Valdemar, Wallenberg*.

Turkish (5 items): disyllables = *Izmir, Iznik*; trisyllables and more = *Atatürk, baklava, vilayet*.

Welsh (73 items): disyllables = *afon, Amlwch, Arfon, Ceinwen, Cleddau, Clydach, Dewi, Dyfed, Dyffryn, Dylan, Elan, Eynon, Gwendraeth, Hywel, Iestyn, Ieuan, Ifor, Iolo, Iorwerth, Morfa, Mynwy, Ogwen, Penrhos, Penrhyn, Powys (surname), Powys (Welsh county), Rhodri, Rhondda, Rhuddlan, Rhydderch, Rhymney, torgoch, Tryfan, Yrnwy*; trisyllables and more = *Aberaeron,*

Aberfan, Abersychan, Aberystwyth, Aneurin, Angharad, Caernarfon, cynghanedd, Dinorwic, Dolgellau, eisteddfod, eisteddfodau, Eluned, Felindre, Hirwaun, Islwyn, Llanwrtyd, Llewelyn, Meredydd, Merthyr Tydfil, Myfanwy, Nantgaredig, Nant-y-glo, Penybont, Plaid Cymru, Pontardawe, Pontardulais, Pontypridd, Preseli, Prestatyn, Pwllheli, Rhiannon, Rhossili, Ruabon, Talybont, Tal-y-llyn, teledu, Tonyrefail, Ystalyfera.