
Pour une phonologie de la Parole

données, outils, modélisation

Mémoire d'habilitation à diriger des recherches
présenté par

Julien EYCHENNE

le 2 décembre 2016

Tome I : volume de synthèse

Jury :

Philippe BOULA DE MAREÛIL, Directeur de recherche au LIMSI-CNRS

Haike JACOBS, Professeur à l'université Radboud de Nimègue

Laurence LABRUNE, Professeur à l'université Bordeaux Montaigne

Bernard LAKS, Professeur à l'université Paris Ouest Nanterre (garant)

Chantal LYCHE, Professeur à l'université d'Oslo

Marc VAN OOSTENDORP, Senior Researcher au Meertens Instituut

“It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts.”

—SHERLOCK HOLMES

Adventures of Sherlock Holmes, Arthur Conan Doyle

Remerciements

Mettre en place un jury d'HDR n'est jamais chose facile, et ça l'est d'autant moins lorsque le candidat, situé à l'autre bout du monde, se montre peu flexible. Je tiens donc tout d'abord à remercier chaleureusement Philippe Boula de Mareüil, Haike Jacobs, Bernard Laks, Laurence Labrune, Chantal Lyche et Marc van Oostendorp d'avoir accepté de faire une place dans leurs agendas bien remplis pour prendre part à ce jury et porter un regard critique sur mes travaux. Je remercie par ailleurs Chantal Lyche et Haike Jacobs d'avoir accepté la tâche additionnelle de rapporteurs.

Je remercie également Bernard Laks de s'être porté garant de ce travail et d'avoir été toujours disponible malgré la distance et le décalage horaire. Ses conseils et encouragements ont été une aide précieuse.

Le métier d'enseignant-chercheur est un métier dans lequel la frontière entre travail et loisir est parfois ténue : je tiens à remercier ici tous les collègues et amis avec qui j'ai collaboré, échangé et partagé des moments inoubliables. Parmi eux, je tiens à citer nommément, en vrac, Jacques Durand, Sylvain Detey, Isabelle Racine, Helene Andreasen, Chantal Lyche (à nouveau), Elissa Pustka, Marie-Hélène Côté et Jeff Tennant. C'est toujours un plaisir de les revoir lorsque j'en ai l'opportunité.

Je tiens également à remercier Jisun Nam pour son amitié, sa bonne humeur contagieuse, et tout ce qu'elle a fait lorsqu'elle était directrice pour améliorer mes conditions de travail ici en Corée. Le quotidien serait bien différent sans elle.

Les deux années que j'ai passées au Canada n'auraient pas été ce qu'elles ont été sans l'aide, le soutien et l'amitié de Doug et Tracy Walker. Ils ont été d'une générosité extraordinaire et je les en remercie infiniment.

Enfin, un **grand merci** (en police 25, comme convenu) à ma femme Sulki. Au-delà de son soutien constant et indéfectible, elle m'a aidé dans la préparation du volume II en scannant méticuleusement tous mes articles, m'économisant ainsi un temps précieux.

Table des matières

Introduction générale	1
Résumé des travaux de second et troisième cycle	1
Parcours depuis l'obtention du doctorat	4
Postes occupés	4
Activités d'enseignement	6
Implication dans des projets de recherche	7
Organisation du mémoire	8
1 Pour une phonologie de la Parole	9
1.1 Phonologie et cognition	9
1.1.1 Système de traitement de l'information	10
1.1.2 Phonologies du lexique et phonologies du vocabulaire	12
1.1.3 Phonologie et Grammaire Universelle	14
1.2 Langue et Parole	16
1.3 Critique du style « galiléen » en linguistique	20
1.4 Phonologie de corpus	24
1.4.1 La phonométrie de Zwirner	24
1.4.2 La phonologie de corpus aujourd'hui	28
1.5 Conclusion	31
2 Description de la variation phonologique	33
2.1 Français	33
2.1.1 Le français méridional	33
2.1.2 Le schwa	36
2.1.3 La liaison	39
2.1.4 Français non natif et enseignement du français	44
2.2 Coréen	46
2.2.1 Système phonologique du coréen	46
2.2.1.1 Consonnes	46
2.2.1.2 Voyelles	47
2.2.2 Fusion de /e/ et /ɛ/	48
2.3 Conclusion	52

3	Ressources et outils	54
3.1	Plateforme PFC	55
3.2	Dolmen	57
3.2.1	Architecture	58
3.2.2	Métadonnées	61
3.2.3	Modèle de données	61
3.2.4	Requêtes	65
3.2.5	Plugins	65
3.3	K-SPAN : corpus de coréen phonétisé	68
3.3.1	Correspondance graphie/phonie en coréen	69
3.3.2	Stratégie de phonétisation	71
3.3.3	Perspectives d'exploitation linguistique	74
3.4	DecoLex : annotation des langues à morphologie riche	75
3.5	Conclusion	77
4	Modélisation symbolique	78
4.1	Liaison	78
4.2	Le pied en français	81
4.2.1	Schwa final en français	81
4.2.2	Le pied en français	82
4.2.3	Approche stratifiée avec grille seule	84
4.3	Schwa opaque en français du midi	88
4.3.1	Contre-blocage et/ou contre-alimentation ?	90
4.3.2	Taxonomie du schwa opaque	93
4.3.3	Analyse d'un idiolecte	95
4.3.3.1	Méthode	95
4.3.3.2	Principaux résultats	96
4.3.3.3	Interprétation	98
4.4	Conclusion	99
5	Modélisation quantitative	100
5.1	Structure suprasegmentale	100
5.1.1	Connexionnisme	101
5.1.2	Réseau calculatoire dynamique	103
5.1.3	Algorithme d'apprentissage	105
5.1.4	Apprenants biaisé et non biaisé	107
5.2	Phonotaxe	111
5.2.1	Le modèle MaxEnt	112
5.2.2	Application au français	113
5.3	Distinctivité, allophonie et rendement fonctionnel	116
5.3.1	Distinctivité gradiente	117
5.3.2	Le modèle	119

5.3.2.1	Rendement fonctionnel	119
5.3.2.2	Distance phonétique	121
5.3.2.3	Indice de distinctivité	122
5.3.3	Les voyelles moyennes en français	123
5.3.3.1	Rendement fonctionnel	124
5.3.3.2	Distance phonétique	127
5.3.3.3	Distinctivité	128
5.3.4	Fusion de /e/ et /ɛ/ en coréen	133
5.3.4.1	Rendement fonctionnel	133
5.3.4.2	Distance phonétique	134
5.3.4.3	Indice de distinctivité	137
5.4	Conclusion	138
6	Apprenabilité des langues naturelles	141
6.1	Complexité formelle des grammaires	141
6.1.1	Hierarchie de Chomsky « classique »	141
6.1.1.1	Langages rationnels	142
6.1.1.2	Langages algébriques	144
6.1.1.3	Langages contextuels	147
6.1.1.4	Langages récursivement énumérables	147
6.1.2	Extensions de la Hiérarchie de Chomsky	149
6.1.3	Où se situe la phonologie ?	149
6.2	Apprentissage formel	151
6.2.1	Le théorème de Gold	151
6.2.2	Identifiabilité vs acquérabilité	152
6.3	Conclusion	154
	Conclusion générale	156
	A Extrait du Plugin PFC	158
	Bibliographie	160

Introduction générale

Ce mémoire est une synthèse de mes travaux sur une période de dix années, depuis l'obtention de mon doctorat en décembre 2006. Il constitue également pour moi une forme d'émancipation intellectuelle. Longtemps partagé sur le plan théorique entre grammaire générative et grammaire d'usage, comme je l'expliquerai au fil de ces pages, j'ai souvent eu le sentiment d'être pour ainsi dire assis entre deux chaises, n'étant pleinement à l'aise ni d'un côté ni de l'autre. Pour pénible qu'elle ait été, cette situation m'a pourtant permis de me nourrir de points de vue différents, souvent contradictoires, et m'a forcé à explorer de nouveaux horizons au-delà du champ étroit de la phonologie (générative). Ceci m'a en retour amené à développer au fur et à mesure une conception assez personnelle de ce champ, et c'est cette conception que je souhaite exposer et défendre dans ce mémoire.

Ce bref prologue vise à poser les jalons de mon parcours professionnel, afin d'en dégager les points saillants et de permettre au lecteur d'en avoir une vue d'ensemble. Je discuterai plus en détail mes travaux dans les différents chapitres de ce mémoire.

Résumé des travaux de second et troisième cycle

Mon intérêt pour la phonologie de corpus a été en grande partie conditionné par ma formation universitaire à l'Université de Toulouse-Le Mirail, où j'ai effectué la totalité de mes études, du DEUG au doctorat. C'est au début de ma maîtrise, réalisée en 2001-2002 sous la direction de Jean-Michel Tarrier et sous le parrainage de Jacques Durand, que j'ai commencé à travailler sur des corpus phonologiques. Jacques Durand, qui avec Bernard Laks et Chantal Lyche venait de lancer le programme de recherche « Phonologie du français contemporain (PFC) : usages, variétés, structure » (Durand, Laks & Lyche 2011b), m'a proposé de travailler sur le schwa en français québécois, à partir d'une pré-enquête de terrain qu'il avait réalisée à l'Université Laval. C'est alors que j'ai commencé à m'initier au travail de transcription, de codage et d'analyse des données.

En Septembre 2003, j'ai commencé, sous la direction de Jacques Durand, un mémoire de DEA pompeusement intitulé « Prolégomènes à une étude comparative du schwa en français : aspects méthodologiques, empiriques et théoriques ». Ce mémoire posait les bases de ce qui devait constituer le socle de mon travail jusqu'à aujourd'hui. Sur le plan empirique, il proposait une ébauche d'analyse contrastive de la variation diatopique, à partir de deux enquêtes de terrain : l'enquête québécoise que j'avais analysée dans mon travail de maîtrise, ainsi qu'une enquête réalisée dans un village viticole du Languedoc.

Sur le plan théorique, il proposait une analyse formelle des deux variétés étudiées dans un formalisme basé sur contraintes, à savoir le modèle développé par Jean-Pierre Angoujard (Angoujard 1997). Ce modèle a deux caractéristiques essentielles. Tout d'abord, il s'agit d'un modèle *déclaratif* : situé dans le prolongement de la PHONOLOGIE DÉCLARATIVE développée par Bird (1995), ce modèle postule que les généralisations phonologiques sont exprimées sous forme de *contraintes inviolables* portant sur les formes phonologiques de surface. Ce modèle est également non dérivationnel et *monostratal*, puisque les opérations phonologiques sont appliquées simultanément, dans un seul niveau. L'autre caractéristique importante de ce modèle est qu'il est fortement *représentationnel* : les contraintes portent sur des structures et primitives symboliques, parcimonieuses et explicitement définies¹. Les opérations permises sur ces structures sont très limitées. Ainsi, toutes les contraintes doivent s'appliquer de manière monotone, ce qui interdit toute opération destructrice.

Après avoir exploré de manière assez approfondie cette variante de la phonologie déclarative, j'avais acquis la conviction que ce modèle n'était pas l'outil dont j'avais besoin pour pouvoir traiter la variation diatopique à plus grande échelle, comme je l'envisageais pour mon travail de thèse. Par un heureux concours de circonstances, Jean-Pierre Montreuil, professeur à l'Université Austin au Texas, avait été invité pour un semestre à l'Equipe de Recherche en Syntaxe et Sémantique (ERSS), laboratoire au sein duquel j'entamais alors mon doctorat. Après lui avoir expliqué mon projet de thèse, il a su me convaincre de ce que la THÉORIE DE L'OPTIMALITÉ (*Optimality Theory*, Prince & Smolensky 1993), qui s'était imposée comme le cadre dominant sur la scène internationale dès le milieu des années 90, était sans doute le cadre théorique le plus adapté pour traiter la variation phonologique. Il a su patiemment dissiper la plupart des critiques, souvent naïves ou mal informées, que je lui adressais. A l'instar de la phonologie déclarative, la théorie de l'optimalité, dans sa version canonique, postule que le savoir phonologique est exprimé sous forme de contraintes, organisées de manière monostratale. Il y a néanmoins deux différences importantes : tout d'abord, les contraintes en OT sont violables et hautement conflictuelles. On en distingue deux types fondamentaux : les contraintes structurales (aussi appelées contraintes de marque), qui pénalisent les structures marquées dans les formes de surface (ex : une coda complexe, une syllabe sans attaque), et les contraintes de fidélité, qui pénalisent les disparités entre la forme sous-jacente et la forme de surface. Les conflits entre contraintes sont résolus en hiérarchisant les contraintes, qui sont supposées universelle, chaque langue correspondant à une hiérarchie, ou à un ensemble de hiérarchies, de contraintes.

C'est donc dans le cadre de la théorie de l'optimalité que je me suis décidé à aborder la variation phonologique dans la réalisation du schwa français, à partir de quatre enquêtes du projet PFC, représentant une quarantaine de locuteurs au total. En plus de l'enquête en Languedoc que j'avais déjà traitée dans le cadre de mon DEA, et qui représentait une

1. Il s'agit essentiellement de la « grille rythmique » au niveau métrique, et des « éléments » au niveau infra-segmental. Ces primitives infra-segmentales sont héritées de la phonologie du gouvernement (Kaye et al. 1985).

variété de français méridional « conservateur », j'ai analysé une enquête au Pays basque, représentative d'un français méridional plus « innovateur »², ainsi qu'une enquête en Vendée qui, du point de vue du schwa, présente les propriétés typiques que l'on rencontre dans les variétés de français septentrionales. En plus de ces variétés de français métropolitaines, j'ai également inclus une enquête réalisée en Alberta, dans l'Ouest canadien, dans un contexte francophone minoritaire (l'Alberta est une province anglophone). Ce travail sur le français canadien m'a donné l'opportunité d'effectuer plusieurs séjours à l'Université de Calgary, à l'invitation de Douglas Walker. Ces séjours m'ont non seulement permis de travailler sur les données récoltées et d'obtenir des sources bibliographiques difficiles d'accès, mais aussi de mieux appréhender la situation sociolinguistique spécifique du français albertain.

Alors que j'avancçais dans le traitement et l'analyse des données d'une part, et dans mes lectures en phonologie théorique d'autre part, j'ai peu à peu commencé peu à peu à entrevoir le fossé qui séparait les travaux théoriques et les données sur lesquelles je travaillais. Les travaux théoriques étaient généralement basés sur quelques exemples isolés, provenant soit de l'intuition du phonologue, soit d'une littérature de deuxième main ; les données à ma disposition étaient au contraire volumineuses, bruitées, et ne semblaient guère se prêter à l'établissement de généralisations à caractère absolu, comme on en voit dans les travaux théoriques. Cette situation m'a amené à explorer quelques travaux en grammaire d'usage (notamment Bybee 2001), mais ceux-ci m'ont laissé profondément insatisfait à cause de leur manque de formalisation.

Bon gré mal gré, j'ai donc soutenu une thèse intitulée « Aspects de la phonologie du schwa en français contemporain : optimalité, visibilité prosodique, gradience » , qui adoptait la théorie de l'optimalité pour cadre théorique. Les points saillants de ce travail sont à mes yeux les suivants :

- une étude détaillée de la phonologie du français méridional qui mêle analyse descriptive et modélisation théorique
- des propositions théoriques sur la nature de l'interface phonétique/phonologie
- des propositions pour traiter certains cas d'opacité phonologique et certains phénomènes gradients (dévoisement partiel, masquage articulatoire) en français

Je reviendrai en détail sur certains de ces aspects de ce travail, pour nuancer certains points ou pour indiquer les endroits où j'avais fait fausse route (le traitement de l'opacité notamment). Il y a néanmoins deux mots-clés qui apparaissent dans le sous-titre de ma thèse et qui constituent le fil d'Ariane de ma démarche : *gradience* et *optimalité*. Mon intérêt pour les phénomènes gradients m'a poussé à m'éloigner des approches purement symboliques et représentationnelles pour me tourner vers les méthodes d'analyse quantitative modernes et les grammaires stochastiques. Par ailleurs, bien que je me sois peu à peu distancié vis-à-vis de la théorie de l'optimalité, je reste convaincu que la notion d'optimalité, et plus précisément la notion d'harmonie³ telle que l'a définie Paul Smolensky dans ses premiers travaux (*i.e.* la mesure du degré de grammaticalité d'une forme), a un

2. Je reviendrai sur les qualificatifs de « conservateur » et « innovateur » à la page 33.

3. L'optimalité est l'équivalent discret de l'harmonie, qui est une notion continue.

rôle fondamental à jouer en phonologie. Comme on le verra au chapitre 5 en particulier, ces deux dimensions ont pris une place importante dans mon travail récent.

Parcours depuis l'obtention du doctorat

Mon parcours universitaire est quelque peu atypique puisque, depuis l'obtention de mon doctorat, j'ai travaillé dans cinq universités différentes, en Amérique du Nord, Europe et Asie. Il me semble donc opportun de présenter ce parcours dans les grandes lignes, dans la mesure où il a façonné ma conception de l'enseignement et de la recherche.

Postes occupés

University of Calgary (Canada, 2007-2008) Peu après l'obtention de mon doctorat, j'ai été engagé, pour une durée d'un an, comme post-doctorant au sein du Language Research Centre (LRC) de l'Université de Calgary, située dans la partie anglophone du Canada (dans la province de l'Alberta). Le LRC a une expertise reconnue dans le domaine de l'acquisition des langues étrangères, et le projet de recherche sur la base duquel j'avais été engagé se situait pleinement dans cette perspective, l'objectif étant de développer des outils et ressources pédagogiques à partir des données récoltées dans le cadre du projet PFC. C'est à cette époque que j'ai commencé à collaborer avec Sylvain Detey, spécialiste de l'acquisition de la phonologie du français langue étrangère, alors maître de conférences à l'Université de Rouen, et maintenant en poste à L'Université de Waseda (Tokyo, Japon). C'est aussi dans le cadre de ce post-doctorat que j'ai développé la « Plateforme PFC », un concordancier permettant de travailler avec les données du projet PFC (voir §3.1).

Simon Fraser University (Canada, 2008-2009) Après ce premier post-doctorat, j'ai été engagé pour une durée d'un an sur un poste d'*Assistant Professor*⁴, en remplacement d'une collègue partie en congé sabbatique, au sein du département de français et d'italien de l'Université Simon Fraser. J'y ai poursuivi mes recherches sur la description des variétés de français.

Centre Valibel (Belgique, 2011) Après une année d'hiatus, j'ai été recruté en tant que post-doctorant du Fonds National de la Recherche Scientifique (FNRS) belge au sein du centre Valibel, à l'Université catholique de Louvain-La Neuve, dans le cadre d'un contrat initialement prévu pour une durée de deux ans. Ce post-doctorat s'inscrivait dans le cadre d'un projet intitulé « Périphérie gauche des unités de discours » et coordonné par Anne-Catherine Simon et Liesbeth Degand, dont l'objectif était d'établir, sur la base d'un corpus multi-genre annoté, une typologie des unités discursives que l'on rencontre sur le bord gauche des unités de discours, en se basant exclusivement sur des critères prosodiques

4. « Professeur adjoint » dans la terminologie canadienne francophone.

et macro-syntaxiques formels, et non perceptifs ou interprétatifs. En collaboration avec un doctorant, Thomas Van Damme, j'ai constitué le corpus dans sa phase initiale et ai implémenté divers scripts pour l'analyse des données. J'ai écourté ce post-doctorat au bout de 7 mois après avoir obtenu un poste de 5 ans aux Pays-Bas.

Rijksuniversiteit Groningen (Pays-Bas, 2011-2012) J'ai rejoint le département de langues romanes de l'Université de Groningen, où j'étais responsable des enseignements de linguistique française, en tant que *Universitair Docent* (= Assistant Professor). La faculté que j'ai intégrée a subi de nombreux changements structurels, dus à de sévères restrictions budgétaires, et d'autres changements plus drastiques encore poignaient à l'horizon. Les conditions imposées par l'université pour la deuxième année n'étant plus celles que l'on m'avait proposées au début du contrat, avec notamment une charge d'enseignement très alourdie, j'ai décidé de quitter ce poste au terme de la première année, après avoir obtenu un nouveau poste en Corée du Sud.

Bien que mon séjour à Groningen ait été bref, il a été fructueux sur le plan intellectuel, puisqu'il m'a permis de me familiariser avec les travaux des membres du *Center for Language and Cognition Groningen* auquel j'étais rattaché, et en particulier les travaux de John Nerbonne et de ses collaborateurs en dialectométrie computationnelle. Ces travaux ont été et continuent d'être une source d'inspiration pour moi.

Hankuk University of Foreign Studies (Corée du Sud, depuis 2012) J'ai intégré l'université Hankuk University of Foreign Studies (HUFS) (Yongin et Séoul⁵, Corée du Sud)⁶ en septembre 2012 sur un poste en phonétique et phonologie, au sein du département de linguistique et de sciences cognitives, d'abord en tant qu'*Assistant Professor* (조교수) puis, à partir de septembre 2014, comme *Associate Professor* (부교수).

Cette prise de poste a eu pour moi plusieurs effets bénéfiques. Elle m'a d'abord permis, en me donnant une certaine stabilité, de devenir plus productif au terme de la première année puisqu'il n'était dès lors plus nécessaire de re-développer de nouveaux cours de A à Z à chaque semestre et de passer la majeure partie de l'été à déménager. Par ailleurs, ayant intégré un département de linguistique, et non de langues, j'ai pu recentrer une partie de mes travaux sur des thématiques qui me tiennent à cœur (modélisation phonologique), tout en élargissant mes perspectives à la phonologie du coréen. Enfin, ce poste étant en phonologie et phonétique⁷, j'ai été amené à interroger de manière plus systématique l'interface entre phonologie et phonétique, et plusieurs de mes travaux récents témoignent de cette évolution.

5. Cette université a la particularité de posséder deux campus, l'un à Séoul, et l'autre à Yongin, à une heure environ de la capitale. Je travaille principalement à Yongin, mais j'enseigne une fois par semaine mes séminaires de maîtrise/doctorat à Séoul.

6. Le choix de la Corée surprendra peut-être, mais il est en partie motivé par des raisons personnelles, mon épouse étant coréenne.

7. J'assure tous les enseignements liés à ces deux disciplines au sein du département.

Activités d'enseignement

Mes activités d'enseignement ont bien sûr été conditionnées par les différents postes que j'ai occupés et les publics visés. Je dégagerai néanmoins trois grands domaines, au-delà des cours d'introduction à la linguistique que j'ai donnés dans le cadre de mon monitorat, lorsque j'étais doctorant :

Linguistique française Mes cours de linguistique française ont tous été conçus pour des apprenants du français, anglophones ou néerlandophones. Ces cours ont couvert l'ensemble des disciplines fondamentales en linguistique (phonétique articulatoire, phonologie, morphologie, syntaxe et sémantique), mais toujours avec une visée applicative puisqu'il s'agissait *in fine* d'aider les étudiants à mieux maîtriser le français plutôt que de les convaincre du bien-fondé de telle ou telle approche théorique. Dans la mesure du possible, je me suis toujours efforcé d'accorder une place importante à la variation, en me basant souvent sur des matériaux issus de mes propres recherches ou de celles de mes collaborateurs, au sein du projet PFC notamment. J'ai notamment pu, à Simon Fraser University, développer et enseigner pendant un semestre un cours entièrement dédié aux variétés de français.

Langue et culture françaises Une partie non négligeable de mes activités d'enseignement au Canada et aux Pays-bas a été consacrée à des cours de langue et de culture « traditionnels ». J'ai ainsi enseigné un cours de grammaire française (réparti sur deux semestres) et plusieurs cours sur la Francophonie, seul ou en co-enseignement. Là encore, mes cours sur la Francophonie ont toujours accordé une place aux variétés de français, en essayant toujours de mettre en relation les grands événements historiques et la structure des communautés francophones et des variétés qu'elles parlent.

Phonétique et phonologie générales Depuis ma prise de poste en Corée du Sud, tous mes cours ont trait à la phonologie et à la phonétique générales, et l'enseignement s'effectue en anglais. Au niveau de la licence (qui s'effectue en quatre ans en Corée), les cours que je dispense sont des cours d'introduction classiques à la phonétique et à la phonologie, ainsi qu'un cours d'introduction à la phonétique expérimentale et un cours d'introduction au traitement automatique de la parole (reconnaissance et synthèse). Au niveau maîtrise/doctorat, j'enseigne en alternance un séminaire de phonétique acoustique et un séminaire thématique de phonologie. Les thèmes que j'ai développés pour ce dernier sont la phonologie générative (allant de la phonologie dérivationnelle standard aux variantes stochastiques de la théorie de l'optimalité), la phonologie suprasegmentale et la phonologie diachronique. Je mets également en place un nouveau séminaire sur la phonologie de corpus, en lien direct avec mes intérêts de recherche centraux.

Implication dans des projets de recherche

La majeure partie de mes activités de recherche a été menée en équipe, au sein de projets collaboratifs. Je les mentionne ici succinctement.

Programme PFC Le programme « Phonologie du français contemporain : usages, variétés, structure » (Durand et al. 2011b)⁸ est un programme de recherche fédératif dont le but initial était de constituer un corpus oral de référence pour le français parlé dans l'espace francophone. Ce corpus contient à l'heure actuelle des échantillons de paroles transcrits et annotés pour plus de 400 locuteurs (environ 1 heure par locuteur) dans toute la francophonie. Le matériel a été recueilli selon un protocole unifié à mi-chemin entre la dialectologie traditionnelle et la sociolinguistique variationniste. Le protocole standard contient la lecture d'une liste de mots et d'un texte, ainsi que deux entretiens (guidé et libre) de 20 à 30 minutes environ. Mon travail au cours des 10 dernières années a porté sur l'analyse et la description de données, mais aussi sur le développement d'outils pour l'analyse des données PFC.

ANR PHONLEX J'ai participé au projet ANR-07-BLAN-0016 « PHONLEX, De la phonologie aux formes lexicales : liaison et cognition en français contemporain », coordonné par Jacques Durand. Ma contribution s'est manifestée à travers plusieurs communications et deux publications : Eychenne (2011)^[2] et Eychenne, Lyche, Durand & Coquillon (2014)^[17].

Projet IPFC Le projet « Interphonologie du français contemporain » coordonné par Sylvain Detey (Université Waseda, Japon), Isabelle Racine (Université de Genève, Suisse) et Yuji Kawaguchi (Tokyo University of Foreign Studies, Japon) est un projet adossé au projet PFC et qui adopte un protocole d'enquête similaire pour l'étude des apprenants du français dans le monde. Je suis impliqué dans ce projet depuis 2009.

Projet PAC Le projet PAC « La Phonologie de l'Anglais Contemporain : usages, variétés et structure / The Phonology of Contemporary English : usage, varieties and structure »⁹ est un projet coordonné par Philip Carr, Jacques Durand, Sophie Herment et Anne Przewozny-Desriaux. Il s'agit d'un projet similaire dans sa méthodologie au projet PFC, mais portant sur l'anglais. Mon rôle a essentiellement porté sur l'adaptation des outils informatiques que j'ai développés pour qu'ils fonctionnent sur les données PAC. Ce travail a été mené en collaboration étroite avec Jacques Durand et Sylvain Navarro (Paris Diderot).

« **Périphérie gauche des unités de discours** » Comme je l'ai mentionné plus, j'ai participé à la phase initiale du projet « Périphérie gauche des unités de discours » piloté

8. <http://www.projet-pfc.net>

9. <http://www.projet-pac.net>

par Anne-Catherine Simon et Liesbeth Degand (Valibel, Louvain-La-Neuve, Belgique), dans le cadre d'une allocation post-doctorale.

Organisation du mémoire

Ce mémoire est organisé en deux volumes. Le premier tome est le volume de synthèse qui vise d'une part, comme il est attendu, à présenter mon profil de recherche et à donner un panorama de mes travaux universitaires depuis le doctorat, et d'autre part à présenter le type d'approche vers lequel je tends maintenant, de manière à ce que le jury puisse se faire une idée plus complète de mon cheminement intellectuel. Le mémoire est organisé comme suit. Le chapitre 1 est une introduction générale qui a pour but de contextualiser ma démarche et de clarifier mon positionnement méthodologique et théorique. Le chapitre 2 donne un aperçu de mes travaux portant sur la description de la variation phonologique. Le chapitre 3 décrit les outils et ressources pour la linguistique de corpus que j'ai construits, seul ou en collaboration. Les chapitres 4 et 5 décrivent mes travaux de modélisation, symbolique et quantitative respectivement. Enfin, le chapitre 6 sert d'épilogue et aborde une question à laquelle je me suis intéressé récemment, à savoir la question de l'apprenabilité des langues naturelles.

Le deuxième volume de cette HDR est le volume de travaux. Je n'y ai inclus que des travaux publiés après l'obtention de mon doctorat, à l'exception d'un article de 2004 co-écrit avec Jacques Durand, qu'il me semblait pertinent d'inclure car il s'agit à ma connaissance de la première étude comparative de la réalisation du schwa qui soit basée sur corpus. En ce sens, ce travail est tout à fait représentatif de la démarche dans laquelle je m'inscris et que j'étayerai tout au long de ce mémoire. Les publications ont été classées par type et ont été numérotées de manière séquentielle, le numéro de la publication apparaissant entre crochets dans la table des matières du second volume. Les références à ces publications dans le premier volume sont toujours accompagnées de ce numéro, entre crochets et en exposant. Par exemple, la publication Eychenne (2014)^[6] fait référence à l'article intitulé « Schwa and the loi de position in Southern French » et publié dans le *Journal of French Language Studies*. On le trouvera à la page 150 du second volume.

Chapitre 1

Pour une phonologie de la Parole

La phonologie étant à l'heure actuelle un champ d'investigation particulièrement éclaté, il m'a semblé opportun, plutôt que de commencer *in medias res* par une présentation de mes recherches, de situer mon travail dans une perspective plus large. Je ne chercherai pas ici à me positionner en tant qu'épistémologue de la discipline : je n'en ai ni l'ambition, ni les compétences. Je voudrais en revanche profiter de l'espace dont je dispose pour tenter d'organiser et de préciser diverses idées et réflexions, souvent fugaces ou nébuleuses, concernant la manière dont j'envisage aujourd'hui le travail de phonologue, notamment quant à la question du rapport aux données. Ce travail de synthèse permettra de contextualiser les chapitres suivants et aidera ainsi le lecteur à mieux comprendre mon cheminement.

1.1 Phonologie et cognition

Le langage tel que nous le connaissons étant le propre de l'Homme, la plupart des chercheurs s'accorde sur le fait qu'une théorie du langage adéquate doit être en mesure de rendre compte de la dimension cognitive de cette faculté si singulière dans le règne animal¹. Mais passée cette entente de principe, la communauté est profondément divisée quant à la manière d'y parvenir. Comme le souligne Taylor (2003 : 6-16), si le courant chomskyen a mis l'accent sur l'analyse formelle du langage, la modularité ou la quête de généralisations profondes, le courant se réclamant de la *Linguistique Cognitive* (Croft & Cruse 2004), qui pris au sens large englobe les approches exemplaristes et les grammaires d'usage (Bybee 2001 ; Tomasello 2003 ; Langacker 2008), met au contraire l'accent sur des mécanismes cognitifs généraux tels que la catégorisation, l'organisation figure-fond ou le partage d'attention pour tenter d'expliquer les propriétés des langues du monde. Se pose donc pour le phonologue la question de savoir quelle est la place de la composante phonologique dans une théorie du langage complète. Répondre à cette question est un programme de recherche à part entière qui dépasse largement le cadre de ce mémoire, mais

1. Voir néanmoins Chater, Clark, Goldsmith & Perfors (2015 : 58-59) pour une position non cognitiviste assumée.

la question ne peut être totalement éludée puisqu'elle affecte nécessairement la manière dont on envisage la phonologie.

1.1.1 Système de traitement de l'information

J'adopterai vis-à-vis de la faculté de langage un point de vue multi-componentiel au sens de Fitch (2010 : 17–18), c'est-à-dire que je considère le langage comme un système complexe composé de sous-systèmes qui peuvent, au moins à un certain niveau, s'analyser de manière (relativement) autonome. Je rejette en revanche la conception modulariste fodorienne (Fodor 1983), selon laquelle les différentes composantes (ou modules) sont étanches et encapsulées. Dans cette perspective multi-componentielle, la phonologie peut être conçue comme une interface qui permet, par un processus d'encodage en production et de décodage en perception, d'actualiser la faculté de langage de l'être humain dans sa modalité par défaut, à savoir la modalité orale. On peut donc la considérer comme un **SYSTÈME DE TRAITEMENT DE L'INFORMATION**, qui prend en entrée un certain type de représentation (un input) et produit en sortie une autre représentation (un output). Là encore, ce schème est bidirectionnel : en production, le système transforme une représentation mentale, qu'elle qu'en soit la forme exacte, en un ensemble de gestes articulatoires plus ou moins coordonnés dans le temps, alors qu'en perception une représentation acoustique est convertie en une représentation mentale abstraite. Un tel système doit nécessairement posséder une certaine complexité interne, et l'une des tâches du phonologue (la seule pour certains) est de comprendre cette complexité, tout en montrant comment cette composante phonologique s'articule avec les autres niveaux structurels de la faculté langagière.

Je prendrai donc pour point de départ la conception maintenant classique de David Marr qui, dans son travail pionnier sur la vision (Marr 1982), pose que tout système complexe de traitement de l'information ne peut être appréhendé de manière holistique :

Almost never can a complex system of any kind be understood as a simple extrapolation from the properties of its elementary components [...] If one hopes to achieve a full understanding of a system as complicated as a nervous system, a developing embryo, a set of metabolic pathways, a bottle of gas, or even a large computer program, then one must be prepared to contemplate different kinds of explanation at different levels of description that are linked, at least in principle, into a cohesive whole, even if linking the levels in complete detail is impractical.

(Marr 1982 : 19-20)

Marr distingue ainsi trois niveaux fondamentaux. Le premier niveau, le plus abstrait, est le **NIVEAU COMPUTATIONNEL** : il vise à caractériser la nature des opérations calculatoires effectuées par le système, ainsi qu'à en déterminer le but : pourquoi le système réalise-t-il telle opération plutôt qu'une autre ? Le second niveau est le **NIVEAU ALGORITHMIQUE** : il vise à expliciter la manière dont les opérations décrites au niveau supérieur sont effectuées. Cela implique de choisir, d'une part, une forme de *représentation* pour l'input

et l'output du système et, d'autre part, un algorithme qui permette de modéliser le(s) *processus* qui transforme(nt) un input donné en un certain output. Le troisième niveau est le NIVEAU IMPLÉMENTATIONNEL (« hardware implementation »), qui décrit la manière dont l'algorithme est réalisé physiquement dans son support matériel, quel qu'il soit (ex : une machine ou un cerveau). Par exemple, un enfant qui effectue une addition utilisera sans doute le même algorithme qu'une caisse enregistreuse de supermarché, bien que la réalisation physique de cet algorithme soit différente dans les deux cas (Marr 1982 : 24).

La division entre le niveau implémentationnel d'une part, et les niveaux algorithmique et computationnel d'autre part, correspond grosso modo à la distinction traditionnelle en sciences cognitives entre cerveau et esprit (Thagard 2005). Marr (1982) insiste néanmoins sur la nécessité de distinguer le *quoi* du *comment*. Au niveau computationnel (le *quoi*), l'objectif est de développer une *théorie computationnelle* du système qui élicite les opérations logiques effectuées par celui-ci. Au niveau algorithmique (le *comment*), la question centrale qui se pose est celle du choix de la représentation, que Marr (1982 : 20) définit comme « a formal system for making explicit certain entities or types of information, together with a specification of how the system does this ». Il donne pour illustrer cette notion l'exemple des différents systèmes numériques (arabe, romain, binaire). Le nombre trente-sept peut ainsi être représenté de différentes manières, soit 37 dans le système arabe, 100101 dans le système binaire (utilisant un seul octet), et XXXVII dans le système romain. Chaque représentation peut être plus ou moins adaptée à un problème donné, mais il n'y a pas nécessairement de représentation qui soit meilleure que les autres dans l'absolu.

Le niveau algorithmique joue un rôle pivot entre les niveaux computationnel et implémentationnel : il faut toutefois se garder de n'envisager la notion d'algorithme que d'un point de vue procédural. S'il est bien entendu tout à fait possible que le niveau algorithmique opère sur des représentations symboliques de manière sérielle, comme c'est le cas dans les modèles d'intelligence artificielle classique, ou encore dans la phonologie dérivationnelle standard de type SPE (*The Sound Pattern of English*, Chomsky & Halle 1968), rien n'interdit d'envisager d'autres architectures, telles que des réseaux de neurones formels traitant l'information en parallèle. Cette architecture pourra à son tour être réalisée de manières bien différentes au niveau implémentationnel, en fonction du support matériel qui la réalise (processeur monocœur ou multicœur, réseau de neurones biologiques, etc.). Si ce dernier niveau joue comme les autres un rôle majeur, puisqu'il détermine *in fine* la réalisation physique du processus étudié, il ne saurait à lui seul expliquer l'ensemble des propriétés du système. Il serait donc illusoire d'espérer comprendre quoi que ce soit au langage en observant des décharges électriques se propager dans un réseau de neurones biologiques. Comme l'a très bien souligné Marr :

[A]n algorithm is likely to be understood more readily by understanding the nature of the problem being solved than by examining the mechanism (and the hardware) in which it is embodied. In a similar vein, trying to understand perception by studying only neurons is like trying to understand bird flight by studying only feathers : It just cannot be done. In order to understand

bird flight, we have to understand aerodynamics ; only then do the structure of feathers and the different shapes of birds' wings make sense.

(Marr 1982 : 27)

En résumé, bien que ces trois niveaux d'analyse ne soient pas complètement autonomes, puisque les choix opérés à un niveau influencent en partie les choix possibles à un autre niveau (par exemple, le choix d'un algorithme impose des contraintes sur ses possibilités de réalisation matérielle, et vice versa), il est souvent possible, voire nécessaire, de ne se concentrer que sur un ou deux de ces niveaux à la fois, selon la nature du problème traité, puisque les questions qui se posent à chaque niveau sont relativement indépendantes les unes des autres. Si l'on espère aboutir à une description globale du système, il est donc crucial que les modélisations qui sont avancées à chacun de ces niveaux soient concordantes les unes avec les autres, de sorte qu'il soit possible, au moins en principe, de les réassembler en un tout cohérent.

1.1.2 Phonologies du lexique et phonologies du vocabulaire

La majeure partie des travaux en phonologie à partir de la « révolution cognitive » initiée par Chomsky se sont consacrés au développement de théories computationnelles au sens donné plus haut². Tenter de développer une théorie computationnelle indépendamment de toute autre considération, en s'assurant que la théorie est intrinsèquement cohérente et suffisamment précise pour pouvoir formuler des hypothèses falsifiables, est une stratégie scientifique tout à fait valable (Carr 2000 : 74-75). La plupart des approches génératives depuis SPE, par leurs succès mais aussi par leurs échecs, nous ont permis d'avancer dans notre compréhension de la phonologie sur des questions aussi diverses que la nature des systèmes tonaux (Goldsmith 1976), la typologie des systèmes accentuels (Hayes 1995), la distance entre forme de surface et forme sous-jacente (McCarthy & Prince 1995) ou encore l'opacité phonologique (Kiparsky 1971, 1973 ; Baković 2007).

Pour autant, il est essentiel que la théorie ne devienne pas autarcique en ne cherchant plus à s'articuler avec les autres niveaux, notamment avec le niveau algorithmique. Les avancées tant en neurosciences qu'en informatique nous permettent maintenant (et depuis déjà plusieurs décennies) d'envisager le développement de modèles plus intégratifs, ce qui implique d'accepter de dialoguer avec les disciplines connexes, en particulier la phonétique et la psycholinguistique. Or les avancées dans ces domaines doivent inviter le phonologue à s'interroger sur la pertinence de ses modèles : de nombreux résultats suggèrent en effet que le lexique n'est pas une simple liste d'entrées dictionnairiques, mais qu'il est au contraire structuré en un réseau richement interconnecté et vastement redondant. Il a ainsi été montré dans plusieurs langues que l'accès à un item lexical est affecté par le nombre de formes phonologiquement proches (sa densité de voisinage), en production,

2. Il est intéressant de noter que Marr (1982 : 28-29) donne explicitement comme exemple d'une théorie computationnelle la grammaire transformationnelle de Chomsky. Il suggère que la distinction entre compétence et performance correspond à peu près (mais pas exactement) à la distinction entre niveaux computationnel et algorithmique.

perception et acquisition (voir la discussion en §3.3.3). On sait par ailleurs que des formes phonologiquement homophones, telles que *thyme* et *time* en anglais, peuvent subir un traitement différencié en fonction de leur fréquence, la forme la plus fréquente ayant tendance à être phonétiquement réduite (Gahl 2008). Ces effets de la fréquence dans le changement et l'accès au lexique sont maintenant établis au-delà de tout doute raisonnable (Bybee 2001 ; Pustka 2007 ; Ernestus 2014) ; pourtant, ils peinent à être intégrés dans les approches génératives.

La pierre d'achoppement du débat est ce que Langacker a appelé la *Rule-List Fallacy* (cf. Langacker 2008) : dans le domaine de la phonologie, la façon la plus utile d'éclairer ce débat est à mon sens de prendre en compte la distinction entre le VOCABULAIRE, qui est l'ensemble des mots existants dans une langue, et le LEXIQUE, qui est l'ensemble des mots *possibles* dans cette langue³. L'exemple canonique est le triplet de formes *brick/blick/*bnick* en anglais : alors que *brick* est un mot du vocabulaire, *blick* n'est pas un mot du vocabulaire mais respecte les contraintes phonotactiques de l'anglais (il fait donc partie du lexique), alors que **bnick* ne fait ni partie du vocabulaire ni du lexique puisque le groupe /bn/ n'est pas un groupe d'attaque possible dans cette langue. Dans la perspective générativiste, le lexique est une propriété de la grammaire et il est donc exprimé sous formes de généralisations structurelles, qui prennent la forme de règles ou de contraintes selon le cadre. Le vocabulaire est indépendant du lexique, et constitue le dépôt des idiosyncrasies lexicales. Il n'a à ce titre qu'un intérêt marginal, sinon nul. Pour les tenants des grammaires d'usage au sens large, lexique et vocabulaire sont indissociables : les régularités structurelles émergent du vocabulaire de manière gradiente, selon leur degré de systématité. Le point crucial dans cette perspective est que le lexique (*i.e.* les régularités grammaticales) n'a pas d'existence autonome : il n'est que la réalisation de schèmes immanents plus ou moins spécifiques dans le vocabulaire. Une difficulté importante à laquelle se heurtent les modèles générativistes est que la productivité des généralisations linguistiques n'est pas catégorique : les approches générativistes traditionnelles, que j'appellerai *phonologies du lexique*, prédisent en effet que tout ce qui est exceptionnel ne peut donner lieu à des généralisations productives puisque le vocabulaire n'encode que l'information non prédictible, qui ne fait donc pas partie de la grammaire. Les phonologies exemplaristes et d'usage, que j'appellerai *phonologies du vocabulaire*, prédisent quant à elles divers degrés de productivité en fonction de la force et du degré d'engramme des schèmes présents dans le vocabulaire⁴. Or, comme l'ont bien montré Bybee & Moder (1983), il existe divers degrés de productivité et les « exceptions » peuvent donner lieu à l'émergence de sous-régularités pouvant être productivement étendues à des formes nouvelles telles que des logatomes (ex : [splɪŋ] → [splʌŋ] par analogie avec [rɪŋ]/[rʌŋ], [dɪŋ]/[dʌŋ], etc.). Conscients de ces difficultés, certains générativistes ont tenté de réconcilier les deux approches en proposant des modèles hybrides (voir Pinker 1991 et plus récemment Bermúdez-Otero 2012). De telles approches ne remettent pas en cause l'existence d'une grammaire autonome, mais

3. J'emprunte cette distinction essentielle à McCarthy (2002).

4. Rappelons que dans ces approches le vocabulaire ne se limite pas à un ensemble de mots et morphèmes mais peut inclure des constructions de taille variable.

elles acceptent, comme la plupart des grammaires d'usage, que le vocabulaire est structuré sous forme de mémoire associative qui permet l'émergence de patrons structurels partiellement productifs. Ce type d'approche hybride me semble toutefois problématique pour deux raisons. D'un point de vue théorique d'abord, un tel modèle semble peu parcimonieux : si l'on admet qu'une organisation de type associative est nécessaire dans le vocabulaire, et puisque ce mode d'organisation est à lui seul capable de rendre compte de toutes les généralisations productives, qu'elles soient absolues ou gradientes, on doit se demander, en vertu du principe du rasoir d'Occam, pourquoi il serait nécessaire d'inclure par ailleurs une grammaire autonome (voir également Taylor 2003 : 318–319). Le second problème est qu'il n'est pas clair dans ces modèles, pour autant que je puisse en juger, de quelle manière et à quel stade une généralisation fait le saut qualitatif de passer de généralisation associative dans le vocabulaire à généralisation grammaticale dans le lexique. Les phonologies du vocabulaire apportent une réponse claire et nette : il n'y a *aucun* saut qualitatif, et toutes les généralisations existent dans le vocabulaire, avec des degrés d'ancrage divers.

1.1.3 Phonologie et Grammaire Universelle

Toute phonologie cognitive, qu'elle soit une phonologie du lexique ou du vocabulaire, se doit d'être en mesure d'expliquer la nature des invariants structuraux que partagent les langues. La tradition générativiste considère que la Grammaire Universelle (GU), qui est innée, est le siège de ces universaux (Chomsky 1986). Ce point de vue a pris une ampleur particulière en phonologie avec l'avènement de la théorie de l'optimalité (dorénavant OT, Prince & Smolensky 1993), puisque cette théorie, dans sa version forte, suppose que toutes les contraintes phonologiques sont universelles (et sont donc fournies par la GU). Cette conception innéiste forte de la GU a été défendue encore assez récemment par Soderstrom, Mathis & Smolensky (2006), qui se sont proposés d'examiner les corrélats biologiques du « gène CV » dans la théorie de la syllabe en OT. Ce substrat génétique expliquerait l'invariant phonologique absolu, l'un des rares d'ailleurs, selon lequel toutes les langues ont des syllabes de type CV. Cette interprétation innéiste de l'universalité n'est cependant pas la seule possible, et un nombre croissant de chercheurs considère que de nombreux universaux phonologiques, peut-être même tous, ont une base phonétique, qu'elle soit articulatoire ou perceptive. Prenons l'exemple donné par Hayes & Steriade (2004 : 6–8), bien attesté d'un point de vue typologique, de la contrainte bannissant les géménées voisées. Il semble bien établi qu'il n'existe pas de langue (ayant par ailleurs une opposition de voisement) qui autorise les géménées voisées à l'exclusion des géménées non voisées. Il s'agit donc d'un universel implicationnel, la présence de géménées voisées impliquant la présence de géménées non voisées, mais pas réciproquement. Alors qu'une approche strictement innéiste consisterait à postuler une base biologique à cette contrainte structurelle, on peut donner une explication phonétique (en l'occurrence, aérodynamique) à cet universel : lors de la production d'obstruantes, la constriction empêche le flux d'air supraglottique de se propager librement vers l'extérieur, et il est donc nécessaire de contrôler

activement l'expansion du conduit vocal pour maintenir le voisement, ce qui demande un effort et une coordination particulièrement importants dans le cas de consonnes longues. La plus grande difficulté articulatoire que pose les géménées voisées permet donc de rendre compte de cet universel phonologique. Cette approche phonétique me semble largement préférable à une approche innéiste car elle permet d'*expliquer* le phénomène au lieu de simplement le constater : l'explication proposée repose sur des lois physiques générales, indépendamment motivées, et non sur une postulation *ad hoc* sans valeur explicative (si les langues évitent les géménées voisées, c'est parce que l'être humain est biologiquement programmé pour les éviter). Si l'on interprète la GU dans son sens fort, on est en droit de se demander quel serait l'avantage évolutif qu'il y aurait à enrichir le patrimoine génétique d'une restriction sur les géménées voisées, plutôt que sur les géménées non voisées par exemple.

Il reste à déterminer si tous les universaux phonologiques ont une base phonétique (ou cognitive), et il n'est bien sûr pas impossible qu'un certain nombre de ces universaux aient une base génétique. Mais plus nous progressons dans notre compréhension des langues, plus la possibilité que cela soit nécessaire semble s'amenuiser. Prenons l'exemple des systèmes vocaliques : il est bien établi d'un point de vue typologique que les langues préfèrent les voyelles antérieures non arrondies et les voyelles postérieures arrondies. Les voyelles antérieures arrondies telles que [y ø] ou postérieures non arrondies telles que [u ʌ] sont donc plus *marquées*. On peut là encore postuler l'existence de contraintes de marque universelles, qui ne font que constater une observation typologique sans expliquer pourquoi les voyelles antérieures arrondies seraient plus complexes/marquées que les voyelles antérieures non arrondies, alors que c'est précisément l'inverse pour les voyelles postérieures. Or, dès lors que l'on prend en compte la phonétique, cet universel reçoit une explication perceptive simple et naturelle : parmi tous les systèmes à 5 voyelles imaginables, le système /i e a o u/ est maximalelement dispersé dans l'espace acoustique, notamment dans le plan F_1/F_2 (Lindblom 1986), et il est donc optimal d'un point de vue perceptif. Divers travaux de simulation ont d'ailleurs montré qu'il était possible de modéliser l'émergence de ce type de système vocaliques avec très peu, voire pas d'information pré-cablée (de Boer 2001 ; Oudeyer 2005, 2006). Le travail de Pierre-Yves Oudeyer mérite ici de retenir l'attention : en utilisant un modèle simplifié du conduit vocal et du système perceptif, il montre comment une population d'agents, qui au départ produisent des vocalisations aléatoires, convergent peu à peu vers un code vocal partagé très similaire d'un point de vue typologique à ce que l'on observe dans les langues humaines : ainsi, sur la base des simulations rapportées, la population d'agents converge dans 25% des cas vers un système de type /i e a o u/, ce qui correspond de très près au 28% observés dans la base UPSID (Maddieson 1984). L'aspect le plus intéressant de ce travail est que le modèle est dépourvu de toute téléologie : il n'y a pas de pression fonctionnelle tendant à maximiser la dispersion dans l'espace vocalique comme c'est par exemple le cas dans le modèle de Flemming (2004, 2005). Ces systèmes *optimaux* émergent simplement des propriétés du système perceptif (artificiel), qui fusionne deux catégories perceptives dès lors que leurs centroïdes sont dans une distance critique l'un de l'autre. Ce type de travail

montre qu'il est possible, au moins en principe, de simuler certains aspects de la phonologie des langues humaines en utilisant des mécanismes d'apprentissage très généraux couplés aux contraintes biologiques imposées par les systèmes sensorimoteurs. Dès lors, encoder ce type d'universel dans la GU semble tout à fait superfétatoire.

Pour résumer, s'il est indéniable que la phonologie générative a permis de faire des avancées réelles dans notre compréhension des patrons et des processus phonologiques, il ne semble plus dorénavant possible d'ignorer les résultats obtenus dans les disciplines connexes : d'une part, ils nous fournissent des données difficilement réconciliables avec des modèles de compétence complètement découplés de la performance et, d'autre part, ils offrent souvent des explications (ou à défaut, des pistes d'explication) pour des phénomènes que la phonologie se satisfait trop souvent de constater de manière arbitraire. Contrairement à la position que j'ai pu défendre (en filigrane) dans ma thèse, je considère aujourd'hui qu'une phonologie qui se veut « cognitive » doit passer de la parole aux actes. Reconnaître, comme je l'ai suggéré au début de cette section, que la phonologie est un système d'interface avec le monde externe, qui nous permet d'interagir avec nos congénères, a plusieurs implications. Cela nous conduit tout d'abord à prendre en compte la dimension sociale et inter-individuelle du langage dans l'élaboration de modèles théoriques. Cela nous amène également à admettre (*pace* Hale & Reiss 2000 ou encore Scheer 2004) que la phonologie a un pied dans le monde cognitif, mais aussi un pied dans le monde physique, et que les propriétés du conduit vocal et du système perceptif imposent des contraintes sur la structure des systèmes phonologiques et la nature des patrons phonotactiques. Cela doit enfin nous amener à développer des modèles psychologiquement plus réalistes, qui permettent de développer un dialogue plus constructif avec les autres sciences cognitives et la biologie notamment.

1.2 Langue et Parole

Si l'on accepte les prémisses posées ci-dessus, l'on est nécessairement amené à s'interroger sur la dichotomie saussurienne sur laquelle s'est fondée la linguistique moderne, à savoir la distinction entre Langue et Parole⁵. L'exercice est quelque peu périlleux, compliqué par le fait que le contenu du *Cours de Linguistique Générale* (CLG, de Saussure 1916⁶), basé comme on le sait sur les notes des étudiants de Saussure, semble ne pas refléter complètement la pensée de Saussure. Le CLG est essentiellement basé sur le troisième des trois cours qu'il a délivrés, mais un examen des sources à partir desquelles il a été rédigé (Godel 1969) et des notes manuscrites de Saussure lui-même, y compris celles qui ont été retrouvées en 1996 dans l'orangerie familiale (de Saussure 2002), suggèrent que sa pensée et sa terminologie ont quelque peu fluctué. Je ne me lancerai pas ici dans une exégèse de la pensée de Saussure, et je renvoie notamment à Bergounioux & Laks (2003) et Laks (2003) pour une analyse de l'héritage saussurien dans le domaine de la

5. J'écrirai Langue et Parole avec une majuscule, comme on le fait généralement, afin de lever toute ambiguïté avec leur sens courant, sauf dans les citations où ces termes sont écrits en minuscules.

6. Je me base ici sur l'édition critique de Tullio de Mauro.

phonologie.

Si l'on s'en tient d'abord à la dichotomie classique telle qu'elle est exposée dans le CLG, la LANGUE est « à la fois un produit social de la faculté du langage et un ensemble de conventions nécessaires, adoptées par le corps social pour permettre l'exercice de cette faculté chez les individus » (CLG : 25). La PAROLE est au contraire un acte individuel, « à la fois l'instrument et le produit » de la Langue (CLG : 37). Comme cela a été souligné (Anderson 1985 : 24), cette dichotomie est très similaire à la distinction entre compétence, qui représente le savoir linguistique que possède un sujet, et performance, qui est l'actualisation de la compétence, pouvant être affectée par des facteurs tels que l'attention, la mémoire, l'état psychologique du sujet, etc⁷. Néanmoins, il faut mettre en exergue l'insistance de Saussure sur le caractère social de la Langue, qui est une convention *partagée* par une communauté donnée. Si le Cours mentionne la dimension *psychique* de la Langue à plusieurs reprises, c'est parce qu'elle est la « somme d'empreintes déposées dans chaque cerveau » (CLG : 38). Les deux dimensions, sociale et individuelle, sont donc indissociables dans la Langue, bien que sa dimension individuelle ait souvent été négligée à la faveur d'un parallélisme simplificateur du type Langue : Parole = social : individuel (voir également Anderson 1985 : 25).

La dimension cognitive (ou « psychique ») de la Langue transparaît de manière encore plus nette dans les sources qui se réfèrent au contenu du premier cours, et la complexité du rapport entre Langue et Parole ne peut sans doute s'envisager pleinement qu'au sein de la trichotomie Langage vs Langue vs Parole : comme le souligne Godel (1969 : 146), « parole et langue sont deux aspects du même fait social (le langage), mais susceptibles l'un et l'autre d'être observés chez l'individu ». A cet égard, il me semble utile d'examiner un passage, assez long mais qui mérite d'être repris presque intégralement, extrait de l'un des cahiers d'étudiants du premier cours :

S'il est vrai que l'on a toujours besoin du trésor de la langue pour parler, réciproquement tout ce qui entre dans la langue a d'abord été essayé dans la parole un nombre de fois suffisant pour qu'il en résulte une impression durable ; la langue n'est que la consécration de ce qui avait été évoqué par la parole.

(p. 24) Cette opposition de langue et de parole [...] est très importante par la clarté qu'elle jette dans l'étude du langage. Un moyen de rendre particulièrement sensible et observable cette opposition, c'est d'opposer langue et parole dans l'individu (le langage est un fait social, il est vrai, mais pour nombre de faits il est plus commode de le concentrer dans l'individu). On pourra alors distinguer presque tangiblement ces deux sphères : langue et parole.

7. Chomsky lui-même a noté cette similitude, mais il insiste sur le caractère universaliste et génératif de la compétence : « The distinction I am noting here [competence vs performance] is related to the *langue-parole* distinction of Saussure ; but it is necessary to reject his concept of *langue* as merely a systematic inventory of items and to return rather to the Humboldtian conception of underlying competence as a system of generative processes ». (Chomsky 1965 : 4)

Tout ce qui est amené sur les lèvres par les besoins du discours, et par une opération particulière, c'est la *parole*.

Tout ce qui est contenu dans le cerveau de l'individu, le dépôt des formes entendues et pratiquées et de leur sens, c'est la *langue*.

De ces deux sphères, la sphère parole est la plus sociale, l'autre est la plus complètement individuelle. La langue est le réservoir individuel ; tout ce qui entre dans la langue, c'est-à-dire dans la tête, est individuel.

(p. 25) Du côté interne (sphère langue), il n'y a jamais préméditation, ni même de méditation, de réflexion sur les formes, en dehors de l'acte, de l'occasion de la parole, sauf une activité inconsciente presque passive, en tout cas non créatrice : l'activité de classement.

Si tout ce qui se produit de nouveau s'est créé à l'occasion du discours, c'est dire en même temps que c'est du côté social du langage que tout se passe. D'autre part, il suffira de prendre la somme des trésors de langue individuels pour avoir *la langue*. Tout ce que l'on considère en effet dans la sphère intérieure de l'individu est toujours social, parce que rien n'y a pénétré qui ne soit d'abord consacré par l'usage de tous dans la sphère extérieure de la parole.

(Godel 1969 : 145–146, notes de R2, italiques dans l'original)

Comme on le voit clairement ici, ce n'est pas la Langue, mais la Parole qui est considérée comme l'aspect le plus social, puisqu'elle est nécessairement le fruit d'une interaction inter-individuelle ; la Langue, ou plutôt le « trésor de langue », est au contraire fondamentalement cognitive puisqu'il s'agit de l'engramme, chez l'individu, des faits de parole (le « dépôt des formes [...] et de leur sens », la « consécration de ce qui avait été évoqué par la parole »). Il apparaît donc que la Langue tout comme la Parole ont chacune une dimension sociale et une dimension individuelle, puisque toute interaction sociale est *in fine* actualisée au niveau de l'individu. Il semble donc raisonnable de considérer, comme le fait Laks (2003) sur la base d'autres textes, que Saussure concevait l'étude de la Langue comme une science cognitive au sens où on l'entend aujourd'hui.

L'introduction de la dichotomie entre Langue et Parole, telle qu'elle est présentée dans le Cours, avait avant tout pour but d'affirmer la primauté de la Langue, essentielle, sur la Parole, secondaire (CLG : 37) : la Langue peut et doit s'étudier de manière indépendante de la Parole. De nombreux chercheurs voient encore dans cette opposition le fondement de la linguistique moderne, et la justification de la nécessité perçue de devoir débarrasser la Langue de tout ce qui relève de son actualisation dans la Parole (voir Scheer 2004 pour une défense radicale de ce point de vue en phonologie).

Tous n'acceptent cependant pas cette interprétation : si l'on prend au sérieux la proposition que la Langue, dans sa manifestation individuelle, est l'engramme des formes manifestées dans la Parole, comme évoqué plus haut, il en découle inexorablement que la Langue ne peut s'appréhender que dans la Parole, qui en est la manifestation concrète. Ce point de vue a notamment été défendu par Gustav Herdan dans son ouvrage intitulé *Quantitative Linguistics* (Herdan 1964), où il défend une interprétation statistique de

la dichotomie saussurienne. Il propose d’abord de ne pas envisager la Langue comme le simple dépôt des formes, mais comme le dépôt des formes et de leur *probabilité d’occurrence* (Herdan 1964 : 32). Comme il le souligne, cette modification de la notion de Langue permet de reformuler la distinction Langue/Parole dans un cadre quantitatif, et plus précisément dans un cadre statistique. On peut ainsi envisager la Langue comme une population statistique, caractérisée par un ensemble de probabilités déterminées, la Parole étant alors un échantillon (aléatoire) de la Langue, auquel est généralement associée une certaine marge d’erreur (voir figure 1.1). Selon cette interprétation, la tâche du linguiste est d’inférer, à partir d’un échantillon fini de la Parole et à l’aide de méthodes quantitatives, les propriétés du système sous-jacent, à savoir la Langue, auquel on n’a pas directement accès.

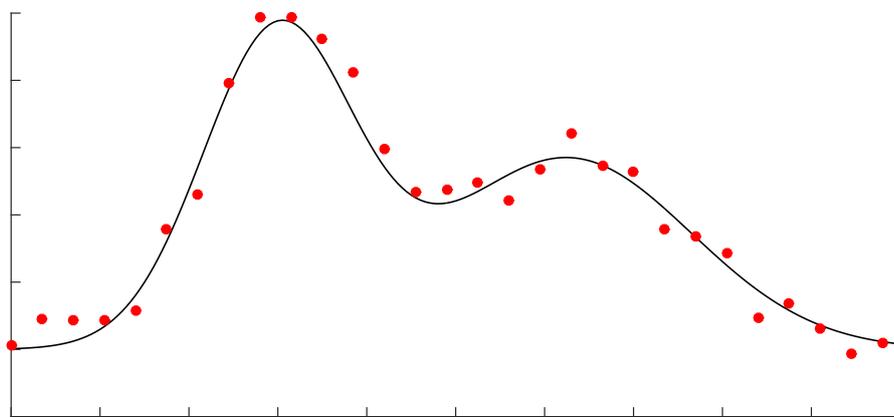


Figure 1.1. Echantillon (points) vs population (courbe)

L’interprétation de Herdan me paraît importante à plusieurs égards. Elle réaffirme tout d’abord le caractère fondamentalement *empirique* de la linguistique, y compris la phonologie à laquelle il consacre d’ailleurs une partie de son ouvrage. Cette interprétation montre par ailleurs que, bien que les concepts de Langue et de Parole soient distincts, il ne peut y avoir d’étude de la Langue sans étude de la Parole. Enfin, en proposant d’associer aux formes linguistiques des probabilités correspondant à leur fréquence d’usage, son approche préfigure les grammaires d’usage contemporaines (Bybee 2001 ; Tomasello 2003 ; Langacker 2008 ; Christiansen & Chater 2016). Au-delà des différences de détail, ces approches ont en commun le fait qu’elles visent à intégrer les dimensions sociale et cognitive dans l’étude du langage. L’article de positionnement de Beckner, Blythe, Bybee, Christiansen, Croft, Ellis, Holland, Ke, Larsen-Freeman & Schoenemann (2009) offre un socle commun à ces approches, en proposant d’envisager le langage comme un **SYSTÈME COMPLEXE ADAPTATIF**, qui a quatre propriétés fondamentales :

1. le système consiste en un réseau d’agents (les locuteurs/auditeurs) en interaction
2. le système est adaptatif (et donc dynamique) puisque le comportement d’un locuteur à un instant $t + 1$ dépend de l’état du système à un instant t , lui-même

dépendant des interactions passées

3. le comportement d'un locuteur est le résultat d'un ensemble de facteurs potentiellement concurrents (perception, motivation sociale)
4. les structures linguistiques émergent de l'effet combiné de l'expérience passée, des interactions sociales et des processus cognitifs mis en œuvre.

Il s'agit bien sûr d'un programme de recherche, destiné à guider et orienter la modélisation, plutôt que d'un cadre théorique qui permet de formuler des hypothèses précises. Ce programme offre néanmoins un changement de perspective profond par rapport au programme générativiste : l'objectif n'est plus de décrire la compétence d'un locuteur/auditeur idéal dans une communauté linguistique homogène, mais de le resituer *dans son interaction* avec le reste de sa communauté. Un tel programme suppose donc de renouer avec une approche empirique des faits de langue, puisque ceux-ci, comme l'a suggéré Herdan (1964), ne peuvent être appréhendés que dans leurs usages. Cela implique également d'abandonner un mode de raisonnement strictement hypothético-déductif dans l'étude du langage.

1.3 Critique du style « galiléen » en linguistique

L'avènement de la grammaire générative a bouleversé la linguistique et les sciences cognitives, en développant des outils formels d'une précision remarquable. Les premiers travaux de Chomsky sur les grammaires formelles ont par ailleurs contribué de façon décisive à la théorie de la computation (voir la discussion de la hiérarchie de Chomsky en §6.1). Malgré ces avancées réelles, un pan non négligeable de la linguistique générative s'est peu à peu enfermé dans un mode argumentatif singulier, jusqu'à développer une conception idiosyncrasique de la méthode scientifique en rupture avec toutes les pratiques en vigueur. C'est ce que Behme (2014) a appelé la linguistique « galiléenne » dans sa critique de *On Nature and Language* (Chomsky 2002), en référence au « style galiléen » souvent invoqué par Chomsky (voir également Botha 1982). Pour comprendre l'origine de cette dérive, il est utile de revenir brièvement sur les fondements de la linguistique générative.

Dès ses premiers travaux, Chomsky définit trois niveaux d'adéquation (ou « succès ») d'une description grammaticale (Chomsky 1964 : chap. 2) qui ont profondément marqué le champ. Une théorie satisfait à l'ADÉQUATION OBSERVATIONNELLE si elle est capable d'énumérer toutes les structures possibles de la langue. Une grammaire observationnellement adéquate possède alors une CAPACITÉ GÉNÉRATIVE FAIBLE. Une théorie satisfait à l'ADÉQUATION DESCRIPTIVE si elle est par ailleurs capable d'associer à ces formes des descriptions structurales correctes, qui rendent donc compte (en principe) de l'intuition des locuteurs natifs sur leur langue. Une grammaire descriptivement adéquate possède une CAPACITÉ GÉNÉRATIVE FORTE. A titre d'exemple, considérons les descriptions structurales, présentées à la figure 1.2, associées à une phrase simple par deux grammaires, G_1 (à gauche) et G_2 (à droite) : bien que ces deux grammaires soient observationnellement adéquates (elles sont *faiblement équivalentes*), la plupart des syntacticiens s'accordera

sur le fait que G_2 est une meilleure grammaire, puisque G_1 ne rend pas compte de la structure interne du syntagme nominal.

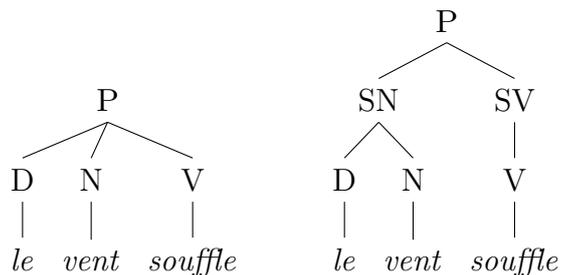


Figure 1.2. Deux descriptions structurales de la même phrase

Le dernier niveau d'adéquation, l'ADÉQUATION EXPLICATIVE, est atteint si, étant donné un ensemble de grammaires observationnellement adéquates, la théorie fournit une mesure d'évaluation qui permette de sélectionner la meilleure grammaire parmi toutes les grammaires observationnellement adéquates.

Dans la conception chomskyenne, ces trois niveaux sont hiérarchisés et le but ultime d'une théorie linguistique est de satisfaire à l'adéquation explicative, considérée comme la plus difficile à atteindre. L'adéquation observationnelle (ou son corrélat, la capacité générative faible), n'ont que peu de valeur aux yeux de Chomsky, qui affirme ainsi que :

The study of weak generative capacity is of rather marginal linguistic interest. It is important only in those cases where some proposed theory fails even in weak generative capacity – that is, where there is some natural language even the *sentences* of which cannot be enumerated by any grammar permitted by this theory.

(Chomsky 1965 : 60)

La quête de l'adéquation explicative s'est ainsi subrepticement opérée au détriment de la satisfaction de la forme d'adéquation la plus élémentaire : l'adéquation observationnelle. L'intuition du locuteur natif, qui devait être l'objet premier de la linguistique, a été rapidement remplacée par l'intuition du linguiste, et il est ainsi devenu tout à fait acceptable de ne travailler à partir que de ses seules intuitions, ou au mieux à partir de données de deuxième main dont la validité n'est pas remise en cause. A l'étude minutieuse et rigoureuse des faits langagiers qui caractérisait la période structuraliste, on a substitué une pratique où les données ont pris une valeur secondaire, et avant tout illustrative : comme l'a montré Laks (2008), alors que les linguistiques des XIX^e et XX^e siècles s'étaient constituées en sciences du *datum*, visant à constituer de larges corpus de données structurés, la linguistique générative se présente comme une science de l'*exemplum*, où les faits linguistiques ne sont plus que des exemples isolés et dont la validité n'est que rarement questionnée. Le constat sévère que dresse Morin (1987) illustre parfaitement cet état de faits dans le domaine de la phonologie du français. En ce qui concerne la syntaxe, Sorace & Keller (2005) ont clairement montré que les jugements de grammaticalité étaient

gradients et fort variables d'un individu à l'autre : sans remettre en cause la valeur de ces intuitions, il apparaît clairement, comme ils le suggèrent, qu'ils doivent être collectés en suivant un protocole expérimental rigoureux et contrôlé pour être scientifiquement exploitable.

Le dédain pour les données a atteint son paroxysme dans l'interprétation qu'a proposée Chomsky du « style galiléen », une expression empruntée au physicien Steven Weinberg, qui dit l'avoir repris de Husserl. Une note de fin (Chomsky 2002 : 189, note 4) nous renvoie à Weinberg (1976 : 28–29), et je reprends ici l'original duquel les bribes citées par Chomsky sont extraites :

We have all been working in what Husserl called the Galilean style; that is, we have all been making abstract mathematical models of the universe to which at least the physicists give a higher degree of reality than they accord to the ordinary world of sensation. As our knowledge increases, the abstract mathematical world becomes farther removed from the world of sensation.
Weinberg (1976 : 28)

L'interprétation qui est faite par Chomsky de ce « style galiléen » est qu'il est tout à fait justifiable de mettre de côté les données qui ne s'accordent pas avec la théorie :

[Galileo] dismissed a lot of data; he was willing to say : “Look, if the data refute the theory, the data are probably wrong.” And the data that he threw out were not minor.
(Chomsky 2002 : 98)

Il est bien entendu presque toujours nécessaire d'idéaliser et d'abstraire le problème ou les données que l'on analyse si l'on souhaite être en mesure de les expliquer, comme le souligne Carr (2000). En outre, il est clair que la postulation d'entités ou de mécanismes abstraits qui n'ont aucun corrélat direct dans le monde sensible peuvent souvent nous amener à une meilleure compréhension des phénomènes, et c'est en ce sens que je comprends la citation de Weinberg. Ainsi en va-t-il de la syllabe en phonologie ou du syntagme en syntaxe : quel que soit leur statut ontologique, ils nous permettent de faire émerger des généralisations profondes sur la structure des langues tout en nous éloignant de la réalité sensible du signal de parole. Mais l'interprétation de Chomsky (via Weinberg) du style galiléen va bien plus loin, et c'est là qu'elle me paraît extrêmement problématique. Il faut tout d'abord noter que Husserl n'a apparemment jamais utilisé l'expression « style galiléen » : la notion de style apparaît dans l'expression « style du monde » (all. *Weltstil*), mais elle y désigne le monde environnant auquel on a accès par l'expérience des sens (Hacking 2004 : 179 ; Dodd 2004 : 89–90). Ce point est certes mineur mais il n'est pas tout à fait anodin puisque certains disciples de Chomsky passent allègrement du « style galiléen » au « mode de pensée galiléen » (Boeckx 2006 : 123), attribuant à Galilée un mode de raisonnement qui lui est manifestement étranger. Par ailleurs, la présentation de Galilée par Husserl, sur laquelle s'appuie Chomsky (via Weinberg), n'est pas exempte de tout reproche : Moran (2012) souligne ainsi que Husserl minimise largement la contri-

bution de Galilée au développement d’une science empirique, et le présente comme un platonicien :

Husserl presents Galileo as a Pythagorean Platonist mathematician, who perfected the ideal of exact measurement (and applied it to the recalcitrant area of motion), but who was aware that nature only imperfectly lives up to that ideal. Husserl’s primary claim is that Galileo, in order to make nature more responsive to mathematics, had to *idealize* nature itself. Husserl does not comment on Galileo’s use of the telescope or his other technological achievements. [...] Husserl completely plays down empirical discovery, the empirical method and induction. For him, the real breakthrough is that nature conforms to purely a priori ideal rules.

(Moran 2012 : 81–82, italiques dans l’original)

De fait, la contribution de Galilée à la science moderne est présentée sous un jour bien différent par un autre physicien, Richard P. Feynman, dans son ouvrage classique d’introduction à la physique :

[P]hysics, as do all the sciences, depends on *observation*. One might also say that the development of the physical sciences to their present form has depended to a large extent on the emphasis which has been placed on the making of *quantitative* observations. Only with quantitative observations can one arrive at quantitative relationships, which are the heart of physics.

Many people would like to place the beginnings of physics with the work done 350 years ago by Galileo, and to call him the first physicist. Until that time, the study of motion had been a philosophical one based on arguments that could be thought up in one’s head. Most of the arguments had been presented by Aristotle and other Greek philosophers, and were taken as “proven”. Galileo was skeptical, and did an experiment on motion which was essentially this: He allowed a ball to roll down an inclined trough and observed the motion. He did not, however, just look; he measured *how far* the ball went in *how long a time*.

(Feynman 2011 : 5-1, italiques dans l’original)

Ce n’est donc pas pour avoir su ignorer les exemples récalcitrants que Galilée est resté dans la mémoire collective, mais bien parce que sa démarche a introduit des méthodes quantitatives basées sur l’*observation* et la *mesure* des phénomènes naturels (sur ce point, voir aussi Chater et al. 2015 : 95–97).

La position exprimée par Chomsky est certes extrême et elle ne fait pas l’unanimité parmi ceux qui adhèrent par ailleurs dans l’ensemble à son programme de recherche. Néanmoins, comme le souligne à juste titre Behme (2014), Chomsky a eu, et continue d’avoir, une influence considérable, directe ou indirecte, dans le champ de la linguistique, et ce message continue à recevoir un fort écho dans la communauté (voir par exemple Boeckx 2006 : §4). Qui pourra blâmer le linguiste en herbe de s’être laissé séduire par

cette démarche, alors qu'il est invité à marcher dans les pas de Newton et de Galilée en se tenant sur les épaules du géant Chomsky? Le problème est que Chomsky a une vision très sélective de la contribution des pionniers de la science, et il y a un risque réel de laisser prospérer en linguistique une méthodologie solipsiste se réclamant scientifique, mais qui est sans pareil dans toutes les autres sciences (Behme 2014).

1.4 Phonologie de corpus

J'ai défendu une conception de la phonologie qui replace les observables linguistiques au centre de sa démarche. Une telle approche s'inscrit naturellement dans le champ de la LINGUISTIQUE DE CORPUS (Biber et al. 1998 ; Kennedy 2014), et plus particulièrement dans le champ émergent de la PHONOLOGIE DE CORPUS. La reconnaissance du besoin de fonder les analyses linguistiques sur des données primaires auquel on assiste depuis une vingtaine d'années n'est cependant pas un fait nouveau : comme on l'a évoqué plus haut, la rupture d'avec la grammaire traditionnelle opérée dès le XIX^e siècle s'est établie sur l'analyse de données systématisées (Laks 2008 : §2.2), et le courant structuraliste américain s'est consacré à mettre en place ce que Chomsky (1957 : 51) appellera des PROCÉDURES DE DÉCOUVERTE afin de dégager la structure phonologique, morphologique et syntaxique des langues à partir de larges corpus de données (Harris 1951). Le courant générativiste, fortement ancré dans une perspective hypothético-déductive des phénomènes langagiers, constitue donc, singulièrement, une parenthèse historique d'un point de vue épistémologique. Je ne tenterai pas ici de retracer l'histoire de la phonologie de corpus. J'ai été amené dans mes travaux récents à revisiter la contribution des pionniers structuralistes à une phonologie quantitative, notamment Charles Hockett : je reviendrai sur ce travail en détail en §5.3 et ne m'y attarderai donc pas ici. J'évoquerai cependant les travaux d'un autre pionnier dont l'importance semble avoir été négligée.

1.4.1 La phonométrie de Zwirner

Si la contribution des structuralistes au développement d'une phonologie empirique est généralement reconnue, quoique pas toujours appréciée à sa juste valeur, il est un autre pionnier important de la phonologie de corpus qui ne semble pas recevoir aujourd'hui la reconnaissance qu'il mérite : Eberhard Zwirner⁸. Neurologue de formation, cet allemand a publié, en collaboration avec le statisticien Kurt Zwirner⁹ un ouvrage intitulé *Grundfragen der Phonometrie* et paru en 1936, qui ne sera porté à l'attention de la communauté internationale, dans sa deuxième édition étendue, qu'en 1966, grâce à une traduction anglaise parue sous le titre *Principles of Phonometrics*¹⁰ (Zwirner & Zwirner 1970).

8. En plus du développement de la phonométrie, Zwirner est reconnu pour avoir fondé la revue qui en 1957 est devenue *Phonetica* (Kohler 2013).

9. Il n'y a pas de lien de parenté entre eux.

10. Ne maîtrisant pas l'allemand, je m'appuie ici sur la traduction anglaise.

La PHONOMÉTRIE de Zwirner est fondée sur deux piliers fondamentaux. Selon lui, l'étude de la structure sonore des langues doit tout d'abord se fonder sur l'analyse d'échantillons de parole naturelle, récoltés et annotés selon une méthodologie précise (voir ci-après). Cette analyse doit par ailleurs, pour pouvoir rendre compte de la variabilité de la parole, être fondée sur la statistique. Le développement de la phonométrie s'est fait principalement en réaction à la phonétique expérimentale. L'essor de la phonétique expérimentale, avec les travaux de l'abbé Rousselot notamment, avait permis de mettre en évidence le caractère continu du flux de parole et avait conduit nombre de phonéticiens, en particulier Edward Wheeler Scripture, à considérer que la segmentation du signal de parole en unités discrètes n'était qu'illusoire. A l'autre bout du spectre, Troubetzkoy (1938) prônait une séparation franche entre la phonologie et la phonétique. La position de Zwirner représente une troisième voie entre ces deux extrêmes : il reconnaît sans ambiguïté la distinction entre langue (*language*) et parole (*speech*), et donne clairement préférence à la linguistique, seule à même de déterminer ce qu'est un objet linguistique par des procédures qui lui sont propres (Zwirner & Zwirner 1970 : 89), mais il considère en même temps que le système de la langue et sa réalisation concrète dans la parole, et plus précisément la phonématique et la phonétique, sont *indissociables* :

Discussion of the relationship of phonemics to phonetics [is not] a matter of two different objects: language structure *and* speech act, which then on their part require two independent sciences [...] Instead, it is a matter of two aspects from which one and the same object can and must be treated.

(Zwirner & Zwirner 1970 : 78–79, italiques dans l'original)

La démarche de la phonométrie est ainsi entièrement fondée sur la reconnaissance d'un ensemble d'unités discrètes, les CLASSES SONORES (*sound classes*) : d'un point de vue linguistique, elles représentent les *normes*, partagées par une communauté, des réalisations auditives, acoustiques et physiologiques des *réalisations* des sons de la parole. Ces classes sonores ne correspondent pas aux phonèmes mais plutôt à ce que l'on appellerait segments ou phones : ainsi, [ç] et [χ], qui ont une distribution allophonique en allemand, correspondent-ils à deux classes sonores différentes, en ce sens qu'ils représentent chacun la norme d'un ensemble d'expériences articulatoires et perceptives (Zwirner & Zwirner 1970 : 93). Là encore, il faut insister sur le fait que les dimensions linguistique et physique sont inextricablement liées dans la perspective phonométrique :

It is not that there are, first, perceptible sounds and, secondly, imperceptible sound-norms. The *same* linguistic structures can be considered from the viewpoint of their non-evident, normative structure—that is, from the viewpoint of linguistics—and also from the viewpoint of their realization in a particular, unrepeatable conversation.

L'approche phonométrique propose une méthodologie explicite pour déterminer le système phonologique d'une langue. La première étape, préliminaire, consiste à enregistrer un échantillon de parole, à l'aide d'un phonographe par exemple. Le travail du

1. Der Blick allein , der Römerart u
2. der blik 2alain , der røme2art u
3. der blikh 2alain , der Røme2arth u

2.	3.	Q ₁	Q ₂	q ₁	q ₂	P	p	M	m ₁	m ₂	m ₃	m ₄	A ₁	A ₂	K/R	α ₁ /a ₁	α ₂ /a ₂	α ₃ /a ₃
d	d			10	10			\\ \\	+2,1°	1,24	77,2	0,4	\\ \\ \\		2	14	87,5	8,75
e	e			4											+11	1,14	0,93	0,85
b	b			11	9 (2)													
l	l			11	10 (1)			///	+6,6°	1,14	86,7	1,0	/ ///		9	16	81,5	7,41
k	k			6											-12	0,89	0,70	0,64
h	h			6														
2	2			10	10			\\ \\	-1,0°	0,54	81,9	0,2	\\ \\ \\		1	18	116,5	11,65
l	l			7	7										+7	1,06	2,18	1,11
ai	a			21	21			/ \ /	+12,4°	1,06	81,4	4,4	/ \ /		6	19	254,0	12,95
i	i			11	9 (2)		36											
n	n			11	9 (2)													
d	d			13	12 (1)			^ \ ^	-6,2°	0,66	75,4	1,2	/ \\ \\		2	14	131,0	10,08
e	e			7	(1) 6										+11	1,21	1,23	1,15
r	r			14	14			///	+23,5°	1,00	83,5	5,7	/ ///		9	17	161,5	11,54
ø	ø			8	8										-12	0,90	1,19	1,19
m	m			14	7 (7)			\\ \\	-31,3°	1,23	83,1	3,6	\\ \\ \\		1	19	136,0	9,71
e	e			14	7 (7)										+11	1,05	1,12	1,31
2	2			12	12			- \\	+31,0°	1,36	80,0	6,6	\\ / \		8	20	152,5	12,71
a	a			5	4 (1)										-11	3,33	7,44	3,10
r	r			6														
t	t			1														
h	h			5	(1) 4			\\ \\	+4,79°	1,30	74,2	3,3	/ \\ \\		1	6	20,5	4,40
u	u			5	(1) 4													

Figure 1.3. Texte-liste phonométrique (Zwirner & Zwirner 1970 : 144)

phonomètre commence alors, qui consiste à retranscrire des *textes* utilisables pour l'analyse phonométrique. Le premier type de texte est une transcription allophonique, réalisée par un transcrip-teur entraîné qui connaît la langue (idéalement un locuteur natif), en utilisant un système symbolique tel que l'alphabet phonétique international (API). Il est en effet essentiel pour Zwirner que le transcrip-teur soit sensible aux oppositions qui existent dans la langue et n'impose pas le filtre interprétatif qu'imposerait sa langue maternelle s'il venait à transcrire une langue qui lui est étrangère. Il rejette donc l'idée d'une transcription objective du signal : toute transcription est un acte interprétatif, et le locuteur natif est le plus à même de mener cette tâche à bien. Zwirner est bien entendu conscient de l'écueil potentiel que présente cette approche : deux transcrip-teurs, même entraînés, ne s'accorderont pas forcément sur le symbole à choisir. Or cet aspect est intégralement pris en compte dans la méthode phonométrique : le texte doit être transcrit par plusieurs transcrip-teurs (au moins trois) et, en cas de désaccord, l'ensemble des *variantes auditives* seront listées dans le texte avec leur degré d'accord inter-juge. Une fois ce *texte phonétique* établi, on procède à la transcription d'un *texte phonémique-allophonique*, qui correspond plus ou moins à une transcription allophonique large¹¹. Enfin, on obtient un *texte-liste phonométrique* combinant les données de la transcription avec des mesures expérimentales. Un exemple en est donné à la figure 1.3. Les couches de transcriptions numérotées 1, 2 et 3 représentent respectivement le texte orthographique, le texte phonémique-allophonique et le texte phonétique. Les deux derniers sont également répétés verticalement, chaque segment correspondant à une ligne du tableau. Je ne discuterai que les 4 premières colonnes (Q_1 , Q_2 , q_1 , q_2) qui représentent l'analyse de la quantité (les autres colonnes servent à noter les pauses, la hauteur et l'accent lexical). Le symbole \smile représente une voyelle brève alors que $-$ représente une voyelle longue. Les symboles peuvent être empilés (\smile) pour représenter une voyelle plus ou moins brève (ou longue selon le cas). La première colonne (Q_1) représente la longueur phonologique. La deuxième colonne (Q_2) représente la transcription donnée par trois transcrip-teurs : on voit que les transcriptions s'accordent presque partout, sauf sur le dernier /a/. Les troisième (q_1) et quatrième (q_2) colonnes représentent respectivement la durée des segments et la durée de la partie voisée des voyelles et des sonantes mesurées à l'aide d'un kymographe.

L'approche phonométrique, bien qu'elle semble avoir été largement oubliée aujourd'hui, préfigure à bien des égards la phonologie de corpus, en se basant sur des corpus de parole construits selon une méthodologie rigoureuse et explicite : de fait, le corpus tel que le concevait Zwirner s'éloigne assez peu de la manière dont l'envisagent Gut & Voormann (2014) près de quatre-vingts ans plus tard (voir *infra*). On y retrouve également une conception voisine de la conception saussurienne, quoique formulée semble-t-il de manière indépendante mais dans tous les cas résolument ancrée dans la Parole : Zwirner avait à cœur de développer une discipline unifiée englobant la phonétique et la phonologie, donnant priorité à l'analyse phonologique mais en l'adossant toujours au signal de parole.

11. Les différences entre le symbole attendu et la réalisation effective d'un locuteur sont appelées *variantes de parole* (*speech variants*) et sont listées au même titre que les variantes auditives produites par les transcrip-teurs.

L'analyse quantitative y est omniprésente mais ne fait sens pour lui que si elle est guidée par l'analyse linguistique et, réciproquement, tout phénomène linguistique se doit d'être quantifié de manière précise afin d'en déterminer les latitudes de variation.

1.4.2 La phonologie de corpus aujourd'hui

Le champ de la phonologie de corpus commence à émerger, depuis une quinzaine d'années environ, comme un courant plus ou moins distinct (Laks 2008 ; Durand et al. 2014 ; Durand 2016) qui se trouve au confluent des travaux en phonologie et en linguistique de corpus (au sens large). La variété d'approches que l'on rencontre en phonologie de corpus fait qu'il est difficile d'en donner une vision unifiée, mais il est clair que tous les chercheurs qui s'inscrivent dans ce paradigme partagent un même souci de (re)placer l'analyse des données au centre de leur démarche. La phonologie de corpus « moderne », telle qu'elle s'est développée depuis le début du XXI^e, puise ses racines dans la sociolinguistique variationniste (Labov 1976, 1994, 2001, 1994), le traitement automatique du langage (Manning & Schutze 1999), et a continuellement bénéficié des progrès importants qu'ont connus l'informatique et l'analyse quantitative des données (Gries 2015). Enfin, le souci d'ancrer la recherche dans l'étude empirique de données quantifiées la rapproche de la phonologie de laboratoire (Pierrehumbert et al. 2000), et il serait sans doute difficile de tracer une frontière nette entre ces deux champs. Plus généralement, on peut considérer que la phonologie de corpus, la sociophonétique variationniste et la phonologie de laboratoire sont trois mouvements plus ou moins indépendants qui ont réinitié un programme de recherche proche de ce qu'était la phonométrie de Zwirner, bien que cette filiation n'ait à ma connaissance jamais été revendiquée, ni même reconnue.

Puisque la notion de CORPUS joue un rôle central dans la phonologie de corpus, il apparaît nécessaire de s'entendre sur l'usage de ce mot. Bien que dans la pratique on utilise souvent BASE DE DONNÉES et CORPUS de manière interchangeable, il est utile de les distinguer. Je suivrai ici Laks (2008), qui définit la BASE ou BANQUE DE DONNÉES comme un « ensemble de données, souvent composites, agrégées sans autre objectif explicite que quantitatif et de documentation empirique ». Le corpus, en revanche, est « un objet explicitement construit, tributaire d'une méthodologie et d'un cadre théorique, fût-il implicite, dont la construction répond à un objectif documentaire très précis et affiché ». Il reste à déterminer ce qui constitue un corpus *phonologique* : Gut & Voormann (2014) en donnent une définition étroite, qui est un bon point de départ et que je reprends en (1).

(1) Définition étroite d'un corpus phonologique

A phonological corpus is [defined] as a representative sample of language that contains

1. *primary data* in the form of audio or video data ;
2. *phonological annotations* that refer to the raw data by time information (time-aligned) ; and
3. *metadata* about the recordings, speakers and corpus as a whole.

(Gut & Voormann 2014 : 16)

Les DONNÉES PRIMAIRES correspondent à un ensemble de fichiers audio ou vidéo qui ont été assemblés en un tout relativement cohérent. Bien que les données puissent être collectées dans un but bien défini, Gut & Voormann (2014) considèrent que ce n'est pas une condition nécessaire, et de fait de nombreux corpus sont utilisés pour examiner des hypothèses pour lesquelles ils n'avaient pas été conçus à l'origine. Une ANNOTATION PHONOLOGIQUE peut prendre plusieurs formes, par exemple une transcription en API ou plus généralement, tout type de codage symbolique qui fournisse des informations utiles sur les données primaires. Notons qu'une condition importante imposée par ces auteurs est que les annotations soient alignées sur le signal, de manière à permettre des allers-retours entre l'annotation et les données primaires. Cela implique que l'annotation soit réalisée à l'aide d'un outil tel que Praat (Boersma & Weenink 2016) ou WaveSurfer (Sjölander & Beskow 2009) par exemple. Enfin, les MÉTADONNÉES représentent toute l'information qui permet de lier les données primaires et les annotations phonologiques, et qui fournissent des relations structurelles sur les données pour en permettre l'organisation et l'exploitation linguistique.

Il est clair que tout corpus qui satisfait à ces exigences peut légitimement prétendre au statut de « corpus phonologique », et Durand (2016) choisit lui aussi d'adopter cette définition dans son article de synthèse sur la phonologie de corpus. Cette définition me paraît toutefois trop restrictive car elle exclut du champ de la phonologie de corpus certains types de corpus qui ont à mon avis droit de cité. Je parlerai donc de CORPUS ORAL pour faire référence au corpus phonologique au sens de Gut & Voormann (2014). Il existe à côté de ces corpus oraux tout en ensemble de jeux de données phonologiques structurés, qui ressortissent à ce que j'appellerai le CORPUS DICTIONNAIRIQUE : ces corpus se présentent sous la forme d'une liste de mots contenant *a minima* des transcriptions phonétiques et/ou phonologiques, mais qui peuvent contenir d'autres informations telles que la fréquence textuelle, la partie du discours, le lemme, etc. on peut citer le CMU Pronouncing Dictionary¹² pour l'anglais, la base LEXIQUE¹³ pour le français (New et al. 2001), ou même le corpus K-SPAN pour le coréen (cf. §3.3 p. 68 suiv.). Ce type de corpus est utilisé dans de nombreuses études : par exemple, Alderete & Bradshaw (2013) docu-

12. <http://www.speech.cs.cmu.edu>

13. <http://www.lexique.org>

mentent, sur la base d'un corpus d'environ 1800 formes phonétisées, les patrons phonotactiques dans les mots radicaux du samoan, et ont permis de mettre au jour ou d'affiner un certain nombre de généralisations sur cette langue, par exemple sur les restrictions de cooccurrence visant les consonnes homorganiques. A ce type de travail très empirique s'ajoutent des travaux d'orientation plus théorique, où le corpus est utilisé pour éprouver des modèles : c'est ce qu'ont fait Hayes & Wilson (2008) en développant leur modèle d'inférence de contraintes phonotactiques, sur lequel je reviendrai (cf. §5.2 p. 111 suiv.). Ce genre de corpus me paraît utile au même titre que les textes phonémiques-allophoniques de Zwirner, et peuvent s'avérer précieux pour examiner des phénomènes (portant sur les mots rares notamment) que l'on pourrait difficilement observer dans d'autres corpus, par exemple des corpus oraux de parole spontanée. On peut en outre imaginer bien d'autres types de données phonologiques qui puissent être constituées en corpus, par exemple des jugements de grammaticalité portant sur des phénomènes phonologiques et mesurés sur une échelle de Likert. L'essentiel à mes yeux est que le corpus contienne des données phonologiques codées sous forme symbolique, et que la méthodologie qui ait guidé la construction du corpus soit explicite. Je propose donc en (2) une définition élargie du corpus phonologique, qui recouvre un éventail plus large et qui est de ce fait beaucoup plus inclusive¹⁴.

(2) Définition élargie d'un corpus phonologique

Un corpus phonologique est un ensemble structuré de données, construit selon une méthodologie explicite, et contenant *a minima* des données codifiées résultant d'un processus interprétatif au niveau phonologique. Il peut également inclure des données primaires (audio ou vidéo), d'autres données non phonologiques utiles à l'exploitation des données phonologiques (annotation aux niveaux syntaxique, pragmatique, etc.) et des métadonnées.

Dans le domaine francophone, un nombre non négligeable de corpus phonologiques ont été construits, bien que peu d'entre eux soient aujourd'hui accessibles. Comme Atanas Tchobanov et moi l'avons noté (Eychenne & Tchobanov 2013)^[15], la France a longtemps été le parent pauvre des études sur corpus et c'est paradoxalement en Amérique du Nord, au Québec en particulier, que la plupart des grands corpus phonologiques ont vu le jour (la majorité d'entre eux sont des corpus *oraux* au sens défini plus haut). Le besoin de répertorier les usages endogènes et de mettre en place des politiques linguistiques prenant en compte la spécificité des parlars nord-américains a permis de faire naître

14. Dans un monde idéal, où les universitaires auraient accès à des ressources humaines et financières illimitées, tous les corpus phonologiques seraient des corpus oraux, et les corpus dictionnaires seraient alors accompagnés d'échantillons sonores sur lesquels ils seraient alignés. Malheureusement nous ne vivons pas dans un monde idéal, et il faut parfois se contenter de ressources qui, sans être optimales, nous permettent néanmoins d'avancer dans notre compréhension des phénomènes langagiers.

d’ambitieux programmes de recherche, soutenus au Canada par les gouvernements fédéral et/ou provincial (au Québec notamment). Parmi ces corpus, construits pour la plupart selon une méthodologie labovienne, on peut mentionner le corpus de Sankoff et Cedergren sur le français parlé à Montréal dans les années 70 (120 entrevues auprès de francophones âgés de 15 à 85 ans), le corpus Ottawa-Hull¹⁵ (120 locuteurs, 3,5 millions de mots), ou encore le corpus de Mougeon et Beniak sur les francophones en situation minoritaire en Ontario (Mougeon & Beniak 1991) ; le corpus de Dubois et ses collaborateurs sur le français cadien de Louisiane (Dubois 2005) est quant à lui composé de 137 entrevues (soit environ 3 millions de mots) et couvre cinq générations, y compris des locutrices nées à la fin du XIX^e siècle¹⁶. En France, le travail d’enquête mené par Martinet (1945) dans les camps de prisonniers et plus tard par Walter (1977) ont assurément joué un rôle précurseur ; Walter (1977) en particulier a documenté la variation diatopique de manière systématique, sur un corpus de taille certes modeste, mais qui fournit des données précieuses pour l’étude des variétés qu’elle a décrites (cf. français au Pays basque p. 34). A l’heure actuelle, les corpus ESLO (enquêtes sociolinguistiques à Orléans)¹⁷ et PFC¹⁸ (Phonologie du français contemporain : usages, variétés, structure, Durand et al. 2011b) représentent les deux grands corpus oraux publiquement accessibles et construits selon des standards modernes¹⁹. Une grande partie du travail que je décrirai dans les chapitres suivants se base sur les données du corpus PFC.

1.5 Conclusion

J’ai tenté de plaider dans ce chapitre pour une phonologie de la Parole. Une telle phonologie reconnaît tout d’abord les dimensions cognitive et sociale du langage, et le fait qu’elles soient inextricablement liées. Cela implique d’accorder une place centrale aux observables, de les mesurer et de les quantifier, et la phonologie de corpus moderne offre de ce point de vue un cadre méthodologique fécond. J’insiste néanmoins sur le fait que mon objectif n’est ni de défendre une phonologie qui conçoit le corpus comme une fin en soi, ni de rejeter les apports de la phonologie formelle. Je pars simplement du constat que nous n’avons pas, à l’heure actuelle, de modèle théorique qui soit pleinement satisfaisant : les phonologies génératives (ce que j’ai appelé les *phonologies du lexique*) nous permettent de modéliser un nombre conséquent de phénomènes de manière précise, mais ont une grande difficulté à prendre en compte les phénomènes gradients, partiellement productifs ou exceptionnels, si tant est qu’elles cherchent à le faire par ailleurs ; les pho-

15. <http://www.sociolinguistics.uottawa.ca/holdings/canadian-fe.html>

16. Ces corpus ne sont pas publiquement accessibles, mais peuvent toutefois être utilisés à des fins de recherche.

17. Voir <http://eslo.huma-num.fr/>. Il existe deux sous-corpus : ESLO1, réalisé au début des années 70, et ESLO2, un corpus plus récent qui a été initié en 2008.

18. <http://www.projet-pfc.net>

19. Il faut également mentionner la base de données VALIBEL en Belgique (Simon et al. 2014), qui n’est pas un corpus à proprement parler, mais plutôt un « méta-corpus » de français belge, *i.e.* un ensemble hétérogène de corpus ayant chacun été construit à des fins spécifiques.

nologies d'usage (ce que j'ai appelé *phonologies du vocabulaire*) restent quant à elles dans l'ensemble formulées de manière vague et imprécise, souvent d'ailleurs par rejet, implicite ou explicite, de la linguistique formelle.

On ne peut à mon avis se satisfaire de la situation actuelle, en continuant à soutenir d'un côté que « grammar is grammar and usage is usage » (Newmeyer 2003), et en continuant de l'autre à présenter la phonologie formelle comme si elle s'était arrêtée à SPE (Bybee 2001), en ignorant les résultats tangibles qu'elle a à son actif. Un programme synthétique ne pourra être l'œuvre d'une seule personne et sera fondamentalement interdisciplinaire, intégrant autant que possible les apports de la phonétique, de la neuropsycholinguistique, de la modélisation informatique, de la description systématique ou encore de l'analyse formelle. A défaut de disposer d'un tel cadre, il me semble important de jeter des ponts entre la phonologie et les disciplines connexes, de manière à favoriser les échanges et établir un dialogue plus fructueux.

Chapitre 2

Description de la variation phonologique

Ayant défendu au chapitre précédent une conception empirique de la phonologie, je commencerai ce travail de synthèse par une présentation d'un certain nombre de mes travaux qui ont un fort ancrage descriptif, visant à documenter la variation phonologique sous diverses formes. Si un large pan de ces travaux aborde la phonologie du français natif, mes intérêts se sont également élargis à la phonologie du français non natif et, plus récemment, à la phonologie du coréen.

Dans la mesure où description et modélisation sont souvent inextricablement liées, il est parfois difficile d'extraire la seule dimension empirique d'un certain nombre de publications. Par conséquent, certains travaux sont discutés soit dans ce chapitre, soit dans un chapitre ultérieur selon que l'accent est mis sur la description ou la modélisation, alors que d'autres le sont de manière transversale.

2.1 Français

2.1.1 Le français méridional

Etant donné qu'une grande partie de mon travail s'est consacrée au français méridional, il me semble d'abord utile de faire une petite mise au point terminologique. J'ai presque toujours utilisé, à l'instar de plusieurs collègues, les qualificatifs de « conservateur » et « innovateur » pour parler des variétés de français méridional, tout en soulignant qu'il ne fallait pas interpréter ces qualificatifs comme des jugements de valeur mais comme des étiquettes commodes. Néanmoins, avec le recul, ce choix me paraît plutôt malheureux, aussi je préfère dorénavant parler de FRANÇAIS MÉRIDIONAL COMMUN, qui correspond *grosso modo* à ce qu'Armstrong & Pooley (2010 : 188) appellent le « Dominant Southern Pattern » (présence d'un schwa final lexical avec taux d'élosion faible ou nul, appendice nasal, loi de position systématique), et de FRANÇAIS MÉRIDIONAL TRANSITIONNEL, qui présente une érosion des traits communs allant généralement dans le sens du français septentrional. Ces étiquettes ne visent pas plus que les anciennes à recouvrir l'éventail des

variétés que l'on rencontre dans le Midi, mais elles me semblent malgré tout préférables dans la mesure où elles ne sont pas connotées.

Depuis mon travail de thèse, mes travaux descriptifs se sont concentrés pour l'essentiel sur les variétés du Sud Ouest, au Pays basque et dans l'aire languedocienne. J'ai notamment contribué deux chapitres¹ (Eychemme 2009a^[11] et Eychemme 2009b^[12]) sur la phonologie de ces deux variétés dans un ouvrage collectif intitulé *Phonologie, variation et accents du français* (Durand, Lyche & Laks 2009), qui offre un panorama de la phonologie de nombreuses variétés de français à partir des enquêtes PFC.

Je me concentrerai ici sur la phonologie du français au Pays basque dans la mesure où il s'agit d'un terrain qui a été peu étudié, et le seul véritable point de référence est la description d'un locuteur d'Hasparren âgé de 72 ans par Walter (1977). Cette variété est pourtant d'un grand intérêt sociolinguistique puisque le français y a été en contact intense avec le basque et qu'une partie non négligeable de la population a une connaissance de cette langue² (Pooley 2003). Le chapitre que j'ai consacré à cette variété (Eychemme 2009a)^[11] vise donc à mieux la documenter, à partir de l'enquête PFC réalisée à Bayonne et Saint-Jean-Pied-de-Port et des premiers travaux auxquels ce corpus a donné lieu (Aurnague & Durand 2003 ; Durand et al. 2004). On trouvera également des éléments de description dans Eychemme (2010)^[14], notamment sur la question du /R/³, mais avec une visée plus pédagogique. Je ne reprendrai ici que quelques aspects saillants qu'il me paraît important de mettre en valeur. Je ne discuterai ni de la liaison, qui présente peu de spécificités, ni du schwa, que j'aborde en §2.1.2 dans une perspective comparative.

Le système consonantique de cette variété, tel qu'il apparaît à travers un examen de la liste de mots du protocole PFC, est particulièrement intéressant puisqu'il présente des traits archaïques du consonantisme français. Ainsi, on trouve tout d'abord chez la locutrice la plus âgée (JM1, 92 ans) une fricative glottale /h/ (ex : *le hasard* [løhazar]). la locutrice JM1 et sa fille (MA1, 66 ans) maintiennent également une opposition entre une latérale palatale /ʎ/, qui historiquement correspond à la graphie *ill* en français, et une latérale alvéolaire /l/ (ex : *étriller* [etriʎe] vs *étrier* [etrije]). Ces deux locutrices possèdent en outre une véritable nasale palatale (ex : *baignoire* [bejwarə]), et la locutrice MA1 oppose par exemple clairement *gnôle* [ɲələ] à *nièce* [nijəs], réalisé avec une diérèse. La nasale palatale a une distribution plus restreinte chez quelques locuteurs (elle est limitée à la position intervocalique comme dans *agneau*) et a disparu chez les autres (y compris la fille de MA1, MA2, 38 ans) au profit de [ɲ]. L'enquête PFC au Pays basque nous offre une perspective intéressante sur le changement diachronique à travers les générations, puisque la locutrice la plus âgée possède les trois consonnes /h ʎ ɲ/, sa fille MA1 a perdu la glottale mais conserve /ʎ/ et /ɲ/, et la petite fille de JM1, MA2, a perdu ces trois consonnes.

1. Ces contributions sont partiellement basées sur les résultats obtenus dans ma thèse.

2. Sur les 12 locuteurs de l'enquête PFC, 6 sont bilingues, et tous ont au moins des notions ou une connaissance passive de la langue.

3. Je suis l'usage qui tend à se répandre et représente la rhotique du français par le symbole /R/ au niveau phonologique.

Une autre caractéristique saillante du consonantisme du français au Pays basque est la réalisation du /R/. Trask (1997) note tout d’abord une différence phonologique importante entre les variétés de basque parlées du côté espagnol et du côté français : alors que les variétés espagnoles opposent un /r/ apical battu à un /r/ apical roulé (ex : *zori* « chance » vs *zorri* « poux »), la vibrante apicale est passée du côté français à une fricative uvulaire, qui présente un grand degré de friction en labourdin et en bas-navarrais (deux des trois provinces bascophones au Pays basque français). Dans la variété de basque des locuteurs du corpus, ce *r* est normalement dévoisé, soit généralement [χ]. En ce qui concerne le français, tous les locuteurs, à l’exception de la doyenne JM1, ont un /R/ d’arrière. JM1 présente une variation libre entre une fricative uvulaire et une apicale, qui peut être battue ou vibrante, chacune se présentant voisée ou non voisée selon le contexte consonantique. A titre d’exemple, cette locutrice alterne dans la lecture de texte entre [ministr̥] et [ministχ] pour *ministre*. Les autres locuteurs ont toujours un /R/ d’arrière, qui a trois réalisations : une fricative voisée [ʁ], une fricative uvulaire dévoisée [ʁ̥] et une fricative uvulaire avec un fort degré de friction que je note [χ̥]. Je reprends ici mes observations sur le locuteur PI1 (31 ans), l’un des plus jeunes locuteurs du corpus. En position initiale de mot, on trouve chez ce locuteur la variante voisée comme dans *rigolo* [ʁigolo] et *résultats* [ʁezylta], mais on rencontre également la variante dévoisée : *redonner* [χ̥ɔdone], *réappris* [χ̥eapʁi]. La paire *réappris* [χ̥eapʁi] vs *réapprendre* [ʁeapʁãd] nous montre qu’il s’agit bien de variation « libre ». En position intervocalique, la variante voisée semble prédominante (*couramment* [kuvamã], *parents* [pavã]), mais on note également des variantes non voisées comme dans *naturellement* [natyχ̥ɛlmã]. En fin de mot devant pause, le /R/ est généralement non voisé, avec un degré de friction plus ou moins important : *mes frères* [mefʁ̥ɛʁ̥], *la culture* [lakyltyχ̥]. La prédominance des /R/ uvulaires dévoisés très fricatifs est un trait très perceptible de cet accent et le distingue d’ailleurs des variétés du Midi plus orientales. Cette réalisation est souvent dépréciée d’un point de vue sociolinguistique : ainsi, au début de l’entretien guidé, PI1 raconte-t-il que l’un de ses professeurs le reprenait systématiquement sur sa prononciation du /R/ dans un mot comme *rugby*, qu’il prononce [χ̥ygbi].

Les voyelles, à l’exception du schwa (cf. §2.1.2), n’appellent comparativement que peu de remarques. Les voyelles orales suivent de manière générale le patron du français méridional commun, avec une seule voyelle basse /a/ et alternance des voyelles mi-fermées [e ø o] et mi-ouvertes [ɛ œ ɔ] régie par la loi de position, les voyelles mi-fermées se trouvant en syllabe ouverte et les voyelles mi-ouvertes se trouvant en syllabe fermée ou en syllabe ouverte suivie d’un schwa (cf. §4.3 p. 88 suiv.). Il faut néanmoins mentionner le cas de la locutrice JM1, chez laquelle la loi de position n’est pas aussi absolue : les voyelles moyennes en syllabe ouverte non finale tendent à s’ouvrir sans toutefois devenir complètement mi-ouvertes, par exemple *des jeunets* [dezɔ̃ne], *épais* [ɛpe], *beauté* [bɔ̃te].

Enfin, du point de vue des voyelles nasales, on distingue 4 voyelles distinctives, qui sont souvent accompagnées d’un appendice nasal, mais pas de manière aussi systématique qu’en Languedoc. Je prendrai ici pour acquis le fait que le français méridional commun ne possède pas de voyelles nasales au niveau sous-jacent, et que les voyelles nasales sont

des voyelles orales suivies d'un appendice nasal, soit /a^N ε^N œ^N ɔ^N/ (Durand 1988 ; Eychenne 2006 : §5.1). L'appendice peut souvent fusionner avec la voyelle pour donner lieu à de véritables voyelles nasales. Une observation intéressante, mais qui mériterait d'être étudiée de manière plus systématique, est le fait que l'appendice nasal est globalement plus saillant en tâche de lecture qu'en conversation : ceci peut aisément s'interpréter comme un phénomène de réduction en parole rapide et non contrôlée, mais cela soulève une question sociolinguistique intéressante. En effet, comme je l'ai fait observé, dans la mesure où l'appendice nasal est un trait stéréotypique, voire stigmatisé des variétés méridionales (Pustka 2007), on s'attendrait à ce que les locuteurs, s'ils avaient de véritables voyelles nasales au niveau sous-jacent, réalisent presque systématiquement de vraies voyelles nasales en tâche de lecture de texte et de mots, puisqu'il s'agit des tâches les plus contrôlées et donc celles où ils sont les plus susceptibles d'éprouver un sentiment fort d'insécurité linguistique. Or c'est bien l'inverse qui se produit et c'est dans la lecture de mots qu'il est le plus saillant. Du point de vue du timbre, on observe une certaine variation intra- et inter-individuelle : la voyelle /a^N/, qui est généralement centrale, est systématiquement une voyelle d'arrière /ɑ^N/ pour plusieurs des locuteurs, certains autres locuteurs pouvant alterner entre les deux réalisations. On note chez l'un des locuteurs en particulier (PI1) une tendance à la neutralisation de l'opposition /ɑ^N/ vs /ɔ^N/, un phénomène par ailleurs bien attesté, notamment en français parisien (Hansen 2012 : 159). L'opposition /ε^N/ vs /œ^N/ (ex : *brin* vs *brun*) n'est quant à elle pas menacée, et la voyelle /ε^N/ n'est jamais réalisée mi-fermée, comme c'est par exemple le cas en Languedoc chez certains des locuteurs les plus âgés.

2.1.2 Le schwa

Le schwa a été le thème central de ma thèse, et il est resté depuis l'un de mes axes de recherche privilégiés. Dans la mesure où la plupart de mes travaux sur ce sujet intègrent une dimension théorique, je serai amené à revisiter cette question dans les chapitres suivants à la lumière de divers cadres théoriques. Je me concentrerai donc ici sur les aspects essentiellement descriptifs de mon travail.

Jacques Durand et moi, dans un article intitulé *Le schwa en français : pourquoi des corpus ?* (Durand & Eychenne 2004)^[1], avons développé ce qui est à ma connaissance la première étude comparative du schwa en position de fin de mot. Ce travail s'est basé sur trois enquêtes du corpus PFC, enquêtes que j'ai par la suite intégrées dans mon corpus de thèse : deux enquêtes dans le sud (Languedoc et Pays basque) et une enquête dans le nord de la France (Vendée). S'il existait des descriptions précises des variétés méridionales (notamment Durand et al. 1987 ; Moreux 1985a), ces variétés n'avaient que très peu porté dans le débat sur la question du statut du schwa, et notamment sur la présence ou non d'un schwa lexical en position finale. La question théorique qui se pose est de savoir s'il existe une différence au niveau phonologique entre les mots du type *net*, sans *e* graphique, et les mots de type *nette*, avec *e* graphique. Je synthétise ici les principaux résultats de cet article : ces données ont depuis été réanalysées de manière plus approfondie dans ma

thèse (Eychenne 2006), mais celle-ci ne remet pas en question les conclusions auxquelles nous avons abouti.

Tout d'abord, la variété languedocienne décrite représente bien une variété méridionale traditionnelle où la corrélation graphie/phonie correspond presque systématiquement à une opposition phonologique : le schwa est presque toujours réalisé dans les mots de type *nette* (1523/1784, soit 86,8% de réalisation), qui a donc la représentation phonologique /nĒtə/⁴, mais presque jamais dans les mots de type *net* (19/758 soit 2,5%), qui a donc la représentation phonologique /nĒt/. Si l'on observe bien une érosion du schwa (lexical) dans les mots à *e* graphique chez les plus jeunes locuteurs, on constate en revanche qu'il n'y a pas de schwa « épenthétique » dans les mots de type *net*. Parmi les 19 cas de mots à *e* graphique qui ont été codés avec une réalisation vocalique claire (sur 758 codages), 18 apparaissent dans *donc*, *avec* ou des composés de *vingt*. Comme nous l'avons rappelé, les formes *donc* et *avec* ont longtemps coexisté avec les formes *doncques* et *avecque*, et il est donc tout à fait possible que certains locuteurs aient enregistré une forme phonologique avec schwa sous-jacent malgré la graphie actuelle. La présence de schwa dans des formes comme *vingt-deux* [vɛⁿtədø] ou *vingt-trois* [vɛⁿtətɔwa]⁵ est un phénomène attesté et qui s'explique aisément par analogie avec les autres numéraux (*trente*, *quarante*, *cinquante*, *soixante*). Il ne reste qu'une seule occurrence de schwa dans *circuits habituels* en contexte prépausal (en tâche de lecture). Cette occurrence isolée a été produite par le locuteur dont le taux d'effacement était le plus élevé (et se rapprochant donc le plus du français méridional). Il est difficile de tirer quelque conclusion que ce soit de cette seule réalisation, mais il est fort probable que ce schwa soit une détente vocalique de la sonante /l/⁶.

Les résultats de l'enquête au Pays basque montrent qu'il s'agit d'une variété transitionnelle avec un taux de maintien bien plus faible qu'en Languedoc dans les mots à *e* graphique (1023/2174 soit 47,1%), mais la corrélation graphie/phonie reste nette puisque le taux de réalisation n'est que de 6,8% (74/1091) lorsqu'il n'y a pas de *e* graphique. Ces résultats globaux doivent être affinés : tout d'abord, les résultats moyens masquent une forte variabilité inter-individuelle, avec des locuteurs ayant un système proche du système des locuteurs languedociens, et d'autres ayant un système où l'opposition entre /Cə/ et /C/ n'existe plus qu'à l'état vestigial. Nous avons tenté de déterminer si le taux de réalisation du schwa pour les mots à *e* graphique était corrélé avec l'âge. Pour ce faire, nous avons examiné les formes à *e* graphique en contexte VC_# {C,#}, soit l'ensemble des mots à *e* graphique dont la consonne finale est précédée d'une voyelle (ex : *petite*) et se trouvant en contexte préconsonantique ou prépausal. Il s'agit des contextes où la voyelle est la plus susceptible de ne pas être réalisée puisque sa réalisation est *a priori* moins susceptible d'être influencée par des contraintes phonotactiques telles que l'évitement de

4. Le symbole /Ē/ représente une voyelle non spécifiée pour l'aperture, et qui se réalise [e] ou [ɛ] en fonction de la loi de position (cf. §4.3).

5. J'utilise le symbole [ʷ] en transcription large pour représenter l'appendice nasal du français du midi.

6. J'ai proposé dans ma thèse d'interpréter certains schwas brefs comme des schwas « non syllabiques », phénomène pour lequel j'ai développé une analyse basée sur la géométrie des traits et la théorie de l'optimalité (Eychenne 2006 : §7.3.2; voir également Durand & Eychenne 2007^[24]).

séquences triconsonantiques.

Je reprends sous forme graphique la corrélation entre l'âge des locuteurs et le taux de réalisation du schwa que nous avons donnée sous forme de tableau (figure 2.1). Un examen de ces données à l'aide du coefficient de corrélation de Pearson montre qu'il existe une corrélation positive significative ($r = 0,63$; $p < 0,05$) entre l'âge et le taux du schwa. L'âge explique donc environ 40% de la variance observée ($r^2 = 0.399$) mais la magnitude de cet effet est plutôt modeste puisque, comme le révèle la ligne de régression, elle est d'environ 0.37% par année. Ceci corrobore néanmoins les observations que nous avons formulées sur une base plus impressionniste.

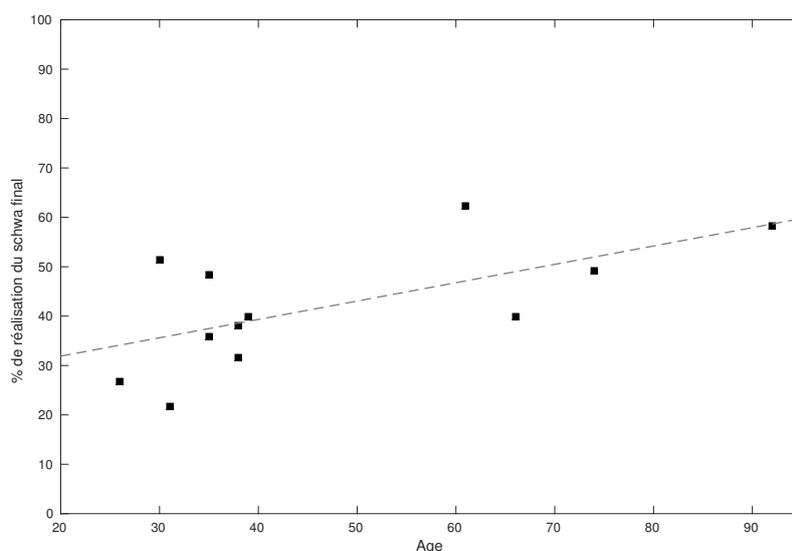


Figure 2.1. Corrélation âge/taux de réalisation du schwa en contexte VC_# {C,#} au Pays basque

Les voyelles épenthétiques que l'on observe pour les mots sans *e* graphique proviennent d'un petit nombre de locuteurs : nous avons relevé quelques épenthèses en contexte CC_#C (ex : *Marc - Blanc*), un contexte où il n'y a pas d'épenthèse vocalique en français méridional commun, et un certain nombre de réalisations vocaliques en contexte prépausal.

Concernant l'enquête en Vendée, nos résultats ont montré qu'il n'y avait pas de corrélation graphie/phonie, avec un taux de réalisation de 15,2% (183/1208) en présence d'un *e* graphique, et de 9,8% (68/684) pour les mots sans *e* graphique⁷. Ces résultats ne permettent pas de trancher la question du statut sous-jacent du schwa de manière définitive, mais ils imposent certaines contraintes sur le type d'analyse compatible avec

7. Les codages sur lesquels cet article est basé provenaient de codeurs septentrionaux. Dans le cadre de ma thèse (Eycheune 2006), afin de m'assurer de l'homogénéité de mon corpus, j'ai été amené à recoder la position finale dans la mesure où la plupart des « schwas » finaux correspondaient à une détente vocalique et non pas à un schwa tel qu'on le trouve dans les variétés méridionales (y compris dans mon propre accent).

les données. On peut distinguer trois classes d’approches en français septentrional : (i) celles qui postulent qu’il n’y a pas de schwa final sous-jacent (Tranel 1981) ; (ii) celles qui postulent qu’il y a un schwa final quelle que soit la forme graphique, ce schwa pouvant être lexical ou la réalisation d’un noyau vide (Scheer 1999) ; (iii) les approches qui postulent une corrélation (approximative) entre graphie et phonie (Schane 1968 ; Dell 1973). Nos résultats s’ajoutent aux arguments avancés par Tranel (1981) et posent de sérieuses difficultés pour les approches qui postulent qu’il y a une asymétrie entre mots à schwa final et mots sans schwa final dans les variétés non méridionales. Ils restent en revanche compatibles dans une certaine mesure avec les deux premiers types d’approche, et la réfutation de telle ou telle approche doit passer par un examen précis de ses postulats théoriques et de ses prédictions. Une telle analyse dépassait néanmoins le cadre de cette étude, dans la mesure où notre objectif était de montrer la nécessité d’utiliser des corpus pour comprendre la complexité du comportement du schwa en allant au-delà de l’empirie appauvrie sur laquelle s’est basée la grande majorité des analyses générativistes (Morin 1987). Je renvoie néanmoins le lecteur à la section §4.2.1 (p. 81) pour des éléments de discussion.

2.1.3 La liaison

Une autre partie importante de mon travail de description s’est attelée à documenter, seul ou en collaboration, la réalisation de la liaison à partir des enquêtes PFC. Mon travail sur le schwa m’a naturellement amené à m’intéresser à la question de la liaison puisque la question du schwa engage nécessairement une réflexion sur la question des consonnes finales et des contextes de sandhi, et donc des consonnes de liaison⁸.

Ma première contribution sur ce thème (Eychenne 2011)^[2] a été un article paru dans un numéro spécial de *Langue française*, coordonné par Bernard Laks et consacré à la phonologie du français contemporain à partir du corpus PFC. Il s’agit pour moi d’une publication charnière dans la mesure où elle représente une étape de transition entre le type de travail que j’ai mené pendant mon doctorat et l’orientation plus explicitement quantitative que j’ai prise par la suite. Je discuterai les analyses théoriques que je développe dans cet article dans un chapitre ultérieur (cf. §4.1 p. 78 suiv.) et me concentrerai sur la dimension empirique de cette contribution. J’ai donc constitué un corpus de 7 enquêtes tirées du corpus PFC, à savoir : Douzens (Languedoc, 11a), Roanne (Loire, 42a), Brécey (Basse-Normandie, 50a), Biarritz (Pays basque, 64a), Lyon (Rhône, 69a), Paris centre (75c), Treize-Vents (Vendée, 85a)⁹. J’ai étiqueté ce corpus, qui était déjà codé pour la liaison, pour les parties du discours avec l’étiqueteur TreeTagger (Schmid 1994, 1995) et ai extrait chaque occurrence du code liaison en prenant en compte un certain nombre de variables : la (non-)réalisation de la liaison, la longueur du mot liaison-

8. Cet intérêt pour les approches empiriques de la liaison m’a poussé à mettre en place et coordonner, en collaboration avec Bernard Laks, un numéro spécial du *Journal of French Language Studies* intitulé *La liaison en français : normes, usages, acquisition*. Le volume est en cours de finalisation et paraîtra en mars 2017 (volume 27 numéro 1).

9. Le code alphanumérique représente le code de l’enquête dans la base PFC.

nant (monosyllabique ou polysyllabique), la catégorie grammaticale du mot liaisonnant, la fréquence du mot liaisonnant (compressée sur une échelle logarithmique), l'enquête, le sexe du locuteur et sa classe d'âge (moins de 30 ans ; entre 30 et 60 ans ; plus de 60 ans). Au total, 10934 codages liaison ont ainsi été extraits. Ces 10934 tokens correspondent à seulement 1264 types. Sans surprise, les données du corpus suivent une loi de Zipf : un petit nombre de types sont très fréquents, et leur fréquence décroît exponentiellement avec leur rang, alors que la majorité des types sont rares et n'apparaissent qu'une fois (voir figure 2.2). Ainsi, sur les 1264 types codés, les trois mots les plus fréquents, à savoir *on* (886 occurrences), *est* (823 occurrences) et *les* (428 occurrences), représentent à eux seuls 19,5% des sites de liaison codés. Ces résultats concordent, là encore sans surprise, avec les résultats de Durand et al. (2011a), et plus récemment de Barreca (2015), qui ont utilisé une base plus large et une approche différente.

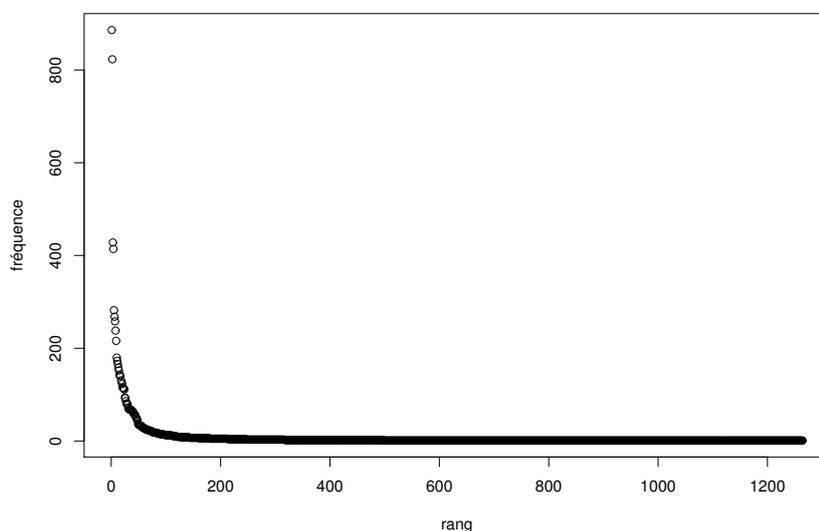


Figure 2.2. Fréquence des 1264 types codés pour la liaison (Eychenne 2011 : 90)^[2]

Afin de déterminer l'importance relative des différentes variables, les données ont été analysées à l'aide d'un modèle de classification et de régression linéaire (*classification and regression trees*, CART)¹⁰. Le prédicteur le plus important est la catégorie du mot grammatical, avec les adjectifs, déterminants (y compris les numéraux), pronoms et prépositions (catégorie I) d'un côté, et les adverbes, conjonctions, noms et verbes (catégorie II) de l'autre. La distinction correspond donc à peu près à la distinction entre mots où la liaison est obligatoire et mots où elle est facultative ou erratique. Les adjectifs sont groupés avec la première catégorie à cause de leur possibilité d'être antéposés, auquel cas la liaison est presque systématique. Le deuxième critère de partition pour les deux cas est la taille du mot, les monosyllabes favorisant là encore la réalisation de la liaison. On

10. Ce type de modèle partitionne les données de manière arborescente (par branchement binaire) en sélectionnant à chaque étape la variable qui offre la meilleure séparation des données.

sait que les mots monosyllabiques ont tendance à perdre leur autonomie accentuelle : ils se présentent donc comme des proclitiques au niveau prosodique, ce qui peut expliquer pourquoi ils montrent des taux de réalisation de la liaison plus élevés. A titre d'exemple, comparons les prépositions *dans* et *après*. Dans le cas de la préposition *dans*, la liaison est réalisée dans 113 cas, mais on ne relève que 4 occurrences de liaison non réalisée : *dans* \cap *un camping* ; *dans* \cap *un mobil-home* ; *dans* \cap *un sens* ; *dans* \cap *un truc chez moi*. Dans le cas d'*après* en revanche, on ne trouve aucune liaison réalisée (0/56). Le troisième paramètre important que l'on peut dégager est la fréquence : elle joue un rôle important dans les monosyllabes, les mots fréquents étant plus sujets à la liaison que les mots rares. La tâche a elle aussi une influence dans la réalisation de la liaison. Les trois niveaux de style qui ont été codés (entretien libre vs entretien guidé vs texte lu) étaient destinés à éliciter trois registres différents. On observe effectivement une différence claire entre les différents registres : pour les mots monosyllabiques de type {ADJ, DET, PRO, PREP} peu fréquents, il y a une différence significative entre tâches, la conversation guidée étant plus proche du texte lu que de la conversation libre. Par ailleurs, dans les noms, verbes et adverbes, on observe une dichotomie nette entre texte lu d'une part et conversation (libre et guidée) d'autre part. L'âge des locuteurs joue un rôle pour les mots de catégorie II, qu'ils soient monosyllabiques ou polysyllabiques, mais seulement en conversation (libre et guidée) : dans les monosyllabes, les seniors ont un taux de réalisation plus élevé que les autres ; dans les polysyllabes, les seniors et les moyens se distinguent des juniors par un taux de réalisation plus élevé. Ceci peut donc s'interpréter comme une tendance générale des locuteurs plus âgés à réaliser la liaison variable plus souvent que les locuteurs plus jeunes. Enfin, le sexe des locuteurs semble jouer un petit rôle dans les prépositions monosyllabiques (cf. *dans* \cap *un camping* mentionné plus haut), les hommes ayant un taux de non-réalisation plus élevé que les femmes. De manière intéressante, le seul paramètre qui ne semble jouer aucun rôle est le point d'enquête, ce qui suggère, comme cela a été noté (voir notamment Côté 2017), que la variation diatopique en France métropolitaine est négligeable eu égard aux autres variables. Je donne à la figure 2.3 le modèle CART dans son ensemble que je n'avais pas pu inclure dans l'article faute de place.

J'ai revisité la question de la liaison plus récemment en collaboration avec plusieurs collègues (Eychenne, Lyche, Durand & Coquillon 2014)^[17], dans le cadre d'une contribution à un volume collectif intitulé *La liaison : approches contemporaines* (Soum-Favaro, Coquillon & Chevrot 2014), résultat du projet ANR PHONLEX sur la liaison en français. Cet article (Eychenne, Lyche, Durand & Coquillon 2014)^[17] fait d'une certaine manière écho au travail (décrit plus haut) que j'ai mené avec Jacques Durand sur le schwa, tant dans sa méthode que dans ses objectifs. Nous y défendons la nécessité de travailler sur corpus pour l'analyse de phénomènes aussi complexes que la liaison. En effet, s'il est un domaine où les phonologues se sont consacrés à tenter de satisfaire à l'adéquation explicative chomskyenne au détriment de l'adéquation observationnelle, c'est bien celui de

11. Le facteur *fw* représente la log-fréquence lexicale, les catégories 1, 2 et 3 pour l'âge représentent respectivement les juniors, moyens et seniors, et les tâches *l*, *g* et *t* représentent respectivement la conversation libre, la conversation guidée et le texte lu du protocole PFC.

la liaison : comme l’ont très bien montré Durand et al. (2011a), la plupart des phonologues générativistes ont, dans le sillage de Schane (1968), considéré la liaison comme un phénomène unitaire et purement phonologique. L’essentiel de ces efforts s’est donc porté sur les aspects représentationnels et/ou computationnels de la liaison, en négligeant de prendre en considération l’ensemble des phénomènes, notamment les dimensions morphologique et syntaxique (Morin & Kaye 1982). La base empirique s’appuie alors le plus souvent sur les travaux normatifs (en particulier Fouché 1956) : on rappellera ainsi que Selkirk (1972) écarte des liaisons données par Fouché telles que *les enfants* [z] *obéissent*, qu’elle juge obsolète, alors qu’elle inclut *long* [k] *hiver*, qui ne l’est pourtant pas moins. On pourrait penser qu’il ne s’agit là que d’un effet collatéral de l’approche hypothético-déductive poussée à l’extrême qui caractérise la période SPE, mais ce type de « données » perdure encore et l’on trouve par exemple *sot* [t] *aigle* chez Féry (2003).

Nous rappelons donc les contributions majeures que l’on doit aux études sur corpus : Ågren (1973), qui a mis en évidence de manière précise le lien entre degré de formalité et taux de réalisation variable dans le discours radiophonique, de Jong (1994), qui a notamment mis en évidence le rôle de la fréquence lexicale dans un corpus d’entretiens semi-formels à Orléans, ou encore Encrevé (1988), qui a mis au jour l’existence de la liaison non enchaînée dans le discours des hommes politiques, ont chacun contribué à renouveler notre compréhension de la liaison. Je ne reprendrai pas ici l’ensemble des résultats issus du corpus PFC (voir notamment Durand & Lyche 2008 ; Mallet 2008 ; Durand et al. 2011a ; Barreca 2015) mais je me concentrerai sur l’une des questions que nous avons abordées en détail dans cet article : la question du registre. La forme *est* se trouvant le plus souvent en contexte de liaison variable et étant, comme nous l’avons vu plus haut, l’une des plus fréquentes, elle permet d’obtenir des données suffisamment massives pour pouvoir tirer des généralisations fiables sur le paramètre du registre. Nous avons extrait 1498 codages de la forme *est* de la base PFC pour la France métropolitaine. Nous avons distingué le présentatif *c’est* des autres formes de *est*, puisque le présentatif est un trait plutôt caractéristique du français oral. Je reprends au tableau 2.1 la répartition des résultats par tâche (notez que le présentatif n’apparaît pas dans le texte lu).

	<i>c’est</i>		<i>est</i>		total	
	liaison	non-liaison	liaison	non-liaison	liaison	non-liaison
libre	86	317	65	83	151	400
	21,34%		43,92%		27,40%	0
guidé	107	257	93	100	200	357
	29,40%		48,19%		35,91%	0
texte			300	90	300	90
			76,92%		76,92%	0

Tableau 2.1. Réalisation de la liaison pour la forme *est* dans 14 enquêtes en France (Eychemme et al. 2014 : 47)

L'analyse de ces résultats au moyen d'un test χ^2 a permis d'établir que le taux de réalisation dans le présentatif *c'est* est significativement moins élevé que dans les autres formes de *est*. En revanche, il n'y a pas de différence significative pour *c'est* comme pour *est* entre les conversations libre et guidée, mais on observe bien une asymétrie entre le texte et la conversation pour la forme *est*, le texte provoquant comme on s'y attend un taux de réalisation plus élevé (76,92%) que les conversations guidée (35,91%) et libre (27,40%). Ce résultat montre clairement que la liaison est bien sensible au style de parole (lecture vs conversation) : l'absence de différence entre libre et guidé est sans doute due au fait que la différence entre ces deux styles n'est souvent guère marquée dans de nombreux entretiens PFC. Ce résultat montre par ailleurs qu'il faut traiter la liaison non pas comme un phénomène global, mais qu'il est au contraire nécessaire, comme l'ont mis en évidence divers travaux (Bybee 2001 ; Laks 2005b ; Chevrot et al. 2007 ; Durand et al. 2011a ; Barreca 2015), de prendre en compte toutes les constructions dans lesquelles la liaison apparaît ou peut apparaître.

2.1.4 Français non natif et enseignement du français

Comme je l'ai indiqué dans l'introduction de ce volume, le premier poste que j'ai occupé après l'obtention de ma thèse était en enseignement et acquisition des langues secondes : ceci m'a naturellement amené à m'interroger sur les applications concrètes possibles du travail de description et de modélisation que je menais. Ceci m'a conduit à collaborer avec Sylvain Detey (Université de Waseda, Japon) et plus tard également avec Isabelle Racine (Université de Genève, Suisse) et Yuji Kawaguchi (Tokyo University of Foreign Studies, Japon), d'abord de manière épisodique puis, à partir de 2011, de manière plus soutenue. Ce travail collaboratif a notamment abouti sur une étude de l'acquisition des voyelles nasales par les apprenants japonophones (Detey et al. 2014)^[27] et la coordination d'un ouvrage actuellement sous presse (voir ci-dessous).

Comme pour la phonologie du français natif, ma démarche s'inscrit dans une approche de corpus. J'ai donc rejoint peu après son lancement le projet IPFC, coordonné par Sylvain Detey, Isabelle Racine et Yuji Kawaguchi qui est un projet adossé au projet PFC et dont le but est d'étendre et d'adapter le protocole méthodologique de PFC à la phonologie des locuteurs non natifs du français (Detey & Kawaguchi 2008 ; Racine, Detey, Zay & Kawaguchi 2012). Il fédère aujourd'hui plus d'une quinzaine d'équipes couvrant autant de langues. Comme le projet PFC, le projet IPFC est basé sur un ensemble de tâches diversifiées, mais il est organisé de manière progressive de sorte qu'il puisse être adapté en fonction du niveau des apprenants. Les tâches sont, par ordre croissant de difficulté : répétition de mots isolés, lecture de mots, lecture de texte, conversation guidée sous forme de questions/réponses entre un étudiant et un instructeur, et une conversation entre étudiants sur un sujet prédéterminé.

Dans le cadre de ce projet, j'ai constitué, pendant l'année que j'ai passée à l'Université de Groningen aux Pays-Bas, un corpus d'apprenants néerlandophones de cette université. Ce corpus est venu compléter le corpus réalisé à l'université de Nijmegen par Dominique

Nouveau et Janine Berns. L'exploitation du corpus a donné lieu à deux mémoires de mastère sous ma direction : l'un sur l'acquisition de la liaison (Jansma 2011), l'autre sur l'acquisition des voyelles nasales (Oosterhof 2012). Ces deux études se sont basées sur les codages perceptifs pour la liaison et les voyelles nasales que S. Detey, I. Racine et moi-même avons mis en place au sein du projet IPFC, et les résultats ont été analysés à l'aide de l'outil Dolmen, présenté au chapitre suivant (§3.2 p. 57 suiv.), et des extensions que j'ai réalisées pour IPFC. Ces deux études nous ont permis d'éprouver ces deux codages et de clarifier certains points qui se sont révélés problématiques à l'usage.

Peu après mon arrivée à Hankuk University of Foreign Studies (HUFS) en Corée, j'ai contribué à mettre en place le sous-projet IPFC-coréen, en collaboration avec Mun Hi Han de l'université nationale de Séoul. Nous avons adapté le protocole IPFC pour tenir compte des difficultés phonologiques propres aux apprenants coréens, et avons constitué un corpus pilote dont le matériau a servi de base à un chapitre d'ouvrage (Han & Eychenne 2017^[23], voir ci-après).

Au-delà du travail de constitution de corpus, j'ai contribué à plusieurs publications collectives visant soit à décrire la variation phonologique du français natif pour la rendre accessible à des locuteurs non natifs, soit à documenter la prononciation des locuteurs non natifs. Ces publications collectives représentent un effort de pallier deux carences dans l'enseignement de la prononciation du français langue étrangère (FLE) : le fait tout d'abord que la prononciation soit le parent pauvre de l'enseignement en FLE, étant le plus souvent reléguée à un bref chapitre en début d'ouvrage ; le fait ensuite que les modèles de prononciation, par exemple dans le cas de la liaison, soient souvent vagues ou peu clairs et se fondent sur des modèles normatifs qui ne reflètent pas la réalité des usages (Detey & Racine 2012). Or, divers travaux tant en L1 qu'en L2 ont mis en évidence le fait qu'une exposition à la variation permet l'émergence de catégories perceptives plus robustes (Detey 2009, voir également Eychenne 2016b^[10]).

Je ne reprendrai pas ici ces contributions dans le détail et je me contente donc d'en faire mention. Dans le cadre du volume intitulé *Les variétés du français parlé dans l'espace francophone : ressources pour l'enseignement* (Detey et al. 2010), j'ai contribué un panorama du français parlé en Amérique du Nord (Eychenne & Walker 2010)^[13] et au Pays basque (Eychenne 2010)^[14]. La plupart des chapitres de ce volume, dont le chapitre sur le Pays basque, sont accompagnés d'une fiche pédagogique sous forme d'analyse d'un extrait de conversation avec un locuteur de la variété décrite (la fiche et l'échantillon sonore sont présents sur le DVD accompagnant l'ouvrage). Par ailleurs, en collaboration avec S. Detey, I. Racine et Y. Kawaguchi, j'ai assuré la coordination d'un ouvrage à paraître chez CLE International intitulé *La prononciation du français : du natif à l'apprenant* (Detey, Racine, Kawaguchi & Eychenne 2017), une publication qui est à ma connaissance unique puisqu'elle propose dans un même volume une description de la variation chez les locuteurs natifs et non natifs. J'ai co-écrit trois chapitres de cet ouvrage : une présentation générale des notions de phonétique et de phonologie pour les non spécialistes (Durand & Eychenne 2017a)^[21], une description du français méridional (Durand & Eychenne 2017b)^[22] et une description des apprenants coréanophones (Han & Eychenne 2017)^[23]. Les chapitres sur le

français méridional et sur les apprenants coréens suivent le format général de l’ouvrage et sont accompagnés d’échantillons sonores, extraits pour la plupart des corpus PFC (natif) et IPFC-coréen (non natif).

Comme je l’ai souligné dans une contribution récente sur l’utilité des corpus en enseignement des langues (Eychenne 2016b)¹², ce type d’ouvrage, qui s’appuie sur des corpus authentiques, donne aux enseignants des outils « clé en main » qu’ils peuvent intégrer directement dans leur cursus, et contribue de ce fait à renouveler les matériaux pédagogiques dans l’enseignement du FLE.

2.2 Coréen

Dès mon arrivée à HUFES, j’ai commencé à m’intéresser de près à la phonologie du coréen, d’abord pour les besoins de mon enseignement, après quoi j’ai développé mes propres thématiques de recherche à travers plusieurs collaborations. Je n’aborderai ici que la dimension descriptive de ce volet de mon travail : je renvoie le lecteur à la section §5.3.4 (p. 133 suiv.) pour une présentation du travail de modélisation qui prolonge les analyses présentées dans cette section ; j’ai par ailleurs développé, en collaboration avec deux collègues, le corpus K-SPAN, un corpus de coréen phonétisé, qui est lui présenté au chapitre suivant (§3.3 p. 68 suiv.).

Avant de présenter le travail de documentation de la variation que j’ai entrepris avec mon collègue Tae Yeoub Jang, je donnerai un bref aperçu du système phonologique du coréen pour les lecteurs qui ne sont pas familiers avec cette langue.

2.2.1 Système phonologique du coréen

Je me contenterai de présenter le système phonologique et n’aborderai pas le système graphique du coréen (*Hangeul*) dans la mesure où ce sujet est traité de manière détaillée dans la discussion de la construction du corpus K-SPAN.

2.2.1.1 Consonnes

Le système consonantique possède un riche sous-système de plosives à trois séries, toutes non voisées : lâches (lenis), aspirées (fortis) et tendues (fortis). Les consonnes dites « fortis », se caractérisent par plusieurs traits par rapport aux consonnes lâches : elles provoquent une élévation de la F_0 sur la phase initiale de la voyelle suivante ; elles sont caractérisées par une tenue plus longue ; elles sont articulées avec un contact linguo-palatal plus important (Cho et al. 2002 : 196–197). Les consonnes tendues se caractérisent en outre par une tension glottale plus élevée et un VOT très bref. Les consonnes lâches ont un VOT

12. Cet article est une version écrite de la communication que j’ai donnée au colloque *Construction and Use of Spoken Corpora of learners of Korean* (Yonsei University, Séoul, 7 mai 2016), à l’invitation de la Société d’Education de Langue et Culture Coréenne.

intermédiaire entre les tendues et les aspirées¹³. Le sous-système des fricatives ne contient quant à lui que trois segments : une glottale et une paire de sibilantes lâche/tendue. Le voisement n'est pas distinctif, mais les plosives lâches (et de manière facultative, la fricative lâche) deviennent voisées entre deux sonantes (ex : /pata/ → [pada] « mer »).

Il n'y a malheureusement pas de diacritique API officiel pour noter les consonnes tendues du coréen. On rencontre parfois le symbole [ᵸ] qui fait partie de l'API étendu, mais il reste peu utilisé. Je suivrai ici l'usage majoritaire et j'adopterai l'étoile (ex : [k*]), bien qu'il ne s'agisse pas d'un symbole API officiel (cf Shin et al. 2013 : chap. 4). Le tableau 2.2 offre un aperçu du système consonantique dans son ensemble :

	bilabial	alvéolaire	alvéolo-palatal	vélaire	glottal
plosives lâches	p	t	t͡ɕ	k	
plosives aspirées	p ^h	t ^h	t͡ɕ ^h	k ^h	
plosives tendues	p*	t*	t͡ɕ*	k*	
fricatives lâches		s			h
fricative tendue		s*			
nasales	m	n		ŋ	
liquide		l			

Tableau 2.2. Consonnes du coréen standard

Le coréen possède un riche éventail de processus phonologiques. Je reviendrai plus en détail sur ce thème à la section §3.3 (p. 68 suiv.) et renvoie également le lecteur intéressé à Shin et al. (2013) pour un bon aperçu.

2.2.1.2 Voyelles

La structure du système vocalique du coréen fait débat parmi les spécialistes, et la discussion est compliquée par le fait que certains chercheurs, essentiellement formés en Europe, suivent la tradition de l'API, alors que d'autres, formés en Amérique du Nord, utilisent le plus souvent un système de transcription américaniste. Je m'appuierai ici sur Shin et al. (2013), qui suivent la tradition API. Dans son extension maximale, la « prononciation du coréen standard » (qui correspond plus ou moins à la prononciation des séoulites) contient 10 monophthongues, à savoir /i e ε y ø u ʌ a u o/. Les descriptions phonologiques reconnaissent traditionnellement trois degrés de hauteur (haut vs moyen vs bas), deux lieux d'articulation (antérieur vs postérieur) et un arrondissement distinctif (Ahn 1998 : 34). Le tableau 2.3 donne un aperçu du système.

Ce système appelle néanmoins plusieurs remarques. Premièrement, les voyelles /y ø/ sont maintenant presque toujours prononcées comme des diphtongues, transcrites /wi/ et

13. D'un point de vue phonétique, les lâches sont légèrement aspirées, alors que les tendues sont fortement aspirées. Notons que le /s/ lenis est lui aussi phonétiquement légèrement aspiré, ce qui est parfois noté [s^h] en transcription étroite (Cho et al. 2002 ; Holliday 2012).

	V antérieures		V postérieures	
	non arrondies	arrondies	non arrondies	arrondies
V hautes	i	(y)	ɯ	u
V moyennes	e	(ø)	ʌ	o
V basses	ɛ		ɑ	

Tableau 2.3. Monophtongues du coréen standard

/we/ respectivement. La glissante est généralement antériorisée, surtout pour /wi/ (H.-Y. Lee 1996 : 122), et elle est le plus souvent réalisée comme une articulation secondaire de la consonne précédente (Shin et al. 2013 : 96-97). Deuxièmement, le coréen (standard) connaît une opposition de longueur (ex : /nun/ « œil » vs /nu:n/ « neige »), mais ce trait n'est distinctif qu'en syllabe initiale (pour les mots polysyllabiques) et joue maintenant au mieux un rôle marginal : il semble avoir disparu pour la plupart des jeunes locuteurs (Shin et al. 2013 : 152-153; Sohn 1999 : 14). Troisièmement, l'opposition /e/ vs /ɛ/, sur laquelle j'ai travaillé, est la cible d'une fusion en cours (voir notamment Hong 1988 : 36-89; Shin et al. 2013 : 99-101), plus ou moins achevée : selon Lee & Ramsey (2011 : 295), aucun locuteur de moins de 50 ans ne fait plus d'opposition entre ces deux voyelles.

Il existe par ailleurs un ensemble de diphtongues dont le statut phonologique fait débat. Dans la mesure où ce débat n'est pas central pour les travaux que je présenterai, je suivrai simplement Shin et al. (2013 : 109) et transcrirai ces diphtongues comme des séquences glissante + voyelle au niveau sous jacent, soit /wi we jɛ wɛ ɯj jʌ wʌ ja wa ju jo/. Notons que les paires /je ~ jɛ/ d'une part, et /we ~ wɛ/ d'autre part, fusionnent en coréen contemporain au même titre que /e ~ ɛ/.

2.2.2 Fusion de /e/ et /ɛ/

Peu après mon arrivée à HUFS, j'ai commencé à collaborer avec Tae Yeoub Jang, phonéticien au département de linguistique anglaise de cette université. Nous avons donc commencé à élaborer un projet, toujours en cours, sur la variation dialectale dans la production et la perception des voyelles en coréen. Nous avons sélectionné trois localités (voir carte 2.4) représentant trois aires dialectales différenciées : Séoul (aire centrale) pour le standard, Gwangju au sud-ouest (aire dialectale Jeolla) et Daegu au sud-est (aire dialectale Gyeongsang)^{14 15}. Nous avons enregistré 10 sujets par localité (5 hommes et 5 femmes) : tous étaient étudiants à Séoul au moment de l'enquête, mais ils avaient passé

14. Voir Sohn (1999) pour une présentation des aires dialectales du coréen.

15. La littérature sur le coréen suit la tradition anglo-saxonne et appelle *dialect* ce que l'on appelle généralement *variétés* en français. Il existe une autre langue coréanique parlée sur l'île de Jeju (le *jejumul*) qui est parfois considérée comme un « dialecte » du coréen, mais les deux langues sont mutuellement intelligibles.

la majeure partie de leur vie dans la localité ciblée ou, pour quelques sujets, à proximité.

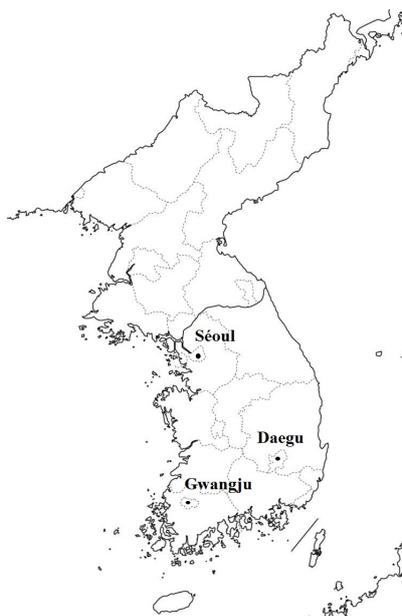


Figure 2.4. Points d'enquête du corpus d'Eychenne & Jang (2015)

Le test de perception est inspiré du travail de Chladkova & Escudero (2012) : nous avons créé un ensemble de stimuli artificiels (à l'aide de Praat) en quadrillant le plan acoustique F1/F2 de points équidistants sur une échelle perceptive (l'échelle ERB¹⁶). Le test de production visait à enregistrer les monophthongues du coréen en contexte /h_d/ (Yang 1996). Nous avons donc utilisé des logatomes de type /h_da/ (ex : /hoda/) enchâssés dans une phrase porteuse de type /igastun hoda jejo/ « ceci est un hoda ».

Nous avons d'emblée décidé d'exclure les voyelles orthographiées ㅟ et ㅠ qui, bien qu'elles correspondent respectivement aux monophthongues /y/ et /ø/ dans la prononciation normative, sont en pratique réalisées comme des diphtongues (/wi/ et /we/ respectivement), comme nous l'avons vu plus haut, et il aurait été complexe d'inclure des diphtongues dans notre protocole (pour le test de perception notamment). Nous avons donc décidé de les exclure. Les voyelles antérieures moyennes non arrondies /e/ (orthographiée ㅕ) et /ɛ/ (orthographiée ㅖ) étant toujours des monophthongues, nous les avons incluses dans notre protocole. Pour le test de production, nous avons donc inclus 8 voyelles, à savoir /i e ɯ ʌ ɑ u o/. Toutes les phrases servant de stimulus ont été écrites en utilisant l'orthographe standard. Le test de perception était un test à choix forcé pour lequel le sujet devait choisir la voyelle, parmi les 8 voyelles présentées sous forme orthographique, qui se rapprochait le plus du stimulus qu'il avait entendu. Les voyelles étaient disposées sur l'écran en ordre pseudo-aléatoire, de telle sorte que les voyelles phonologiquement ou graphiquement proches (notamment /e/ et /ɛ/) ne soient pas côte à côte.

Alors que notre objectif d'ensemble était l'étude des monophthongues, nous avons décidé de nous focaliser dans un premier temps sur les voyelles /e/ et /ɛ/, et ce afin

16. Chladkova & Escudero (2012) ont utilisé l'échelle mel, mais l'échelle ERB nous a paru plus pertinente puisqu'elle est basée sur des travaux psycho-acoustiques plus récents (Moore 2013).

de déterminer s'il était préférable de les traiter comme deux voyelles distinctes ou comme une seule voyelle, surtout du point de vue du test de perception. Les sujets n'ayant reçu aucune instruction spécifique quant à ces deux voyelles, nous nous attendions à ce que, s'ils avaient perdu l'opposition en perception, leur choix soit aléatoire entre les deux voyelles orthographiques. Nous nous sommes donc spécifiquement concentrés sur ces deux voyelles¹⁷, et les résultats de ce travail ont été publiés dans Eychenne & Jang (2015)^[7]. L'analyse du reste du corpus est un travail en cours, qui fera l'objet d'une publication ultérieure.

Pour le sous-corpus restreint aux voyelles /e/ et /ɛ/, nous avons obtenu 300 tokens en production (2 voyelles × 10 locuteurs × 3 variétés × 5 répétitions par voyelle), qui ont été mesurés à l'aide de Praat, et 10245 tokens au total en perception. Pour le test de perception, il était inévitable que certains stimuli soient identifiés de manière erronée, ce qui pouvait se produire par exemple si le sujet appuyait par erreur sur la mauvaise voyelle. Afin d'éliminer les points les plus aberrants, nous avons adopté le filtre suivant : toute voyelle dont le F1 ou le F2 (en unités ERB) était au-delà de 3 écarts-types de la moyenne¹⁸ ont été écartées. La figure 2.5 donne un exemple de la carte perceptuelle représentant les résultats obtenus pour un locuteur : on y voit clairement une voyelle aberrante reconnue comme /u/ dans la région la plus antérieure de /i/.

Les résultats du test de production sont synthétisés au tableau 2.4. Comme on peut le voir, pour chaque sexe les valeurs des deux voyelles pour F1 et F2 sont très proches, et elles sont également voisines dans les trois variétés : les différences sont presque toujours inférieures à l'écart-type. Les données ont été analysées à l'aide de MODÈLES LINÉAIRES À EFFETS MIXTES avec intercept aléatoire (Pinheiro & Bates 2000 ; Gries 2015) pour lesquels les facteurs VARIÉTÉ, SEXE et VOYELLE ont été traités comme des effets fixes, et les facteurs SUJET et RÉPÉTITION de la voyelle comme des effets aléatoires. Un modèle a été construit séparément pour chaque formant, avec le formant (F1 ou F2) comme variable de réponse. L'analyse a montré que seul l'effet SEXE était (sans surprise) significatif pour F1 et F2 ($p < 0,001$ dans les deux cas), et que ni VARIÉTÉ ni VOYELLE ne l'étaient : aucun des trois groupes ne fait donc de distinction entre /e/ et /ɛ/ en production, et la voyelle est la même dans les trois variétés.

Si l'on tourne maintenant notre attention du côté de la perception, les résultats sont quelque peu différents. Un examen visuel des résultats (cf. figure 2.6) suggère qu'il y a une possible légère différence entre les deux voyelles pour les hommes à Gwangju, et une différence plus marquée pour les femmes à Séoul. Nous avons construit des modèles linéaires à effets mixtes similaires aux modèles de production, en analysant là encore F1 et F2 séparément. L'analyse des résultats montre que l'interaction VARIÉTÉ × SEXE × VOYELLE était clairement significative ($p < 0,0001$) pour F1, ce qui rend difficile

17. En plus de l'analyse phonétique de notre corpus, Eychenne & Jang (2015) contient une discussion du rendement fonctionnel de cette opposition. Ce point sera abordé en détail en §5.3.4 et je me concentre ici sur la partie phonétique de ce travail.

18. Dans une population distribuée selon une loi gaussienne, environ 99,7% des valeurs se situent dans un intervalle de ± 3 écarts-types autour de la moyenne.

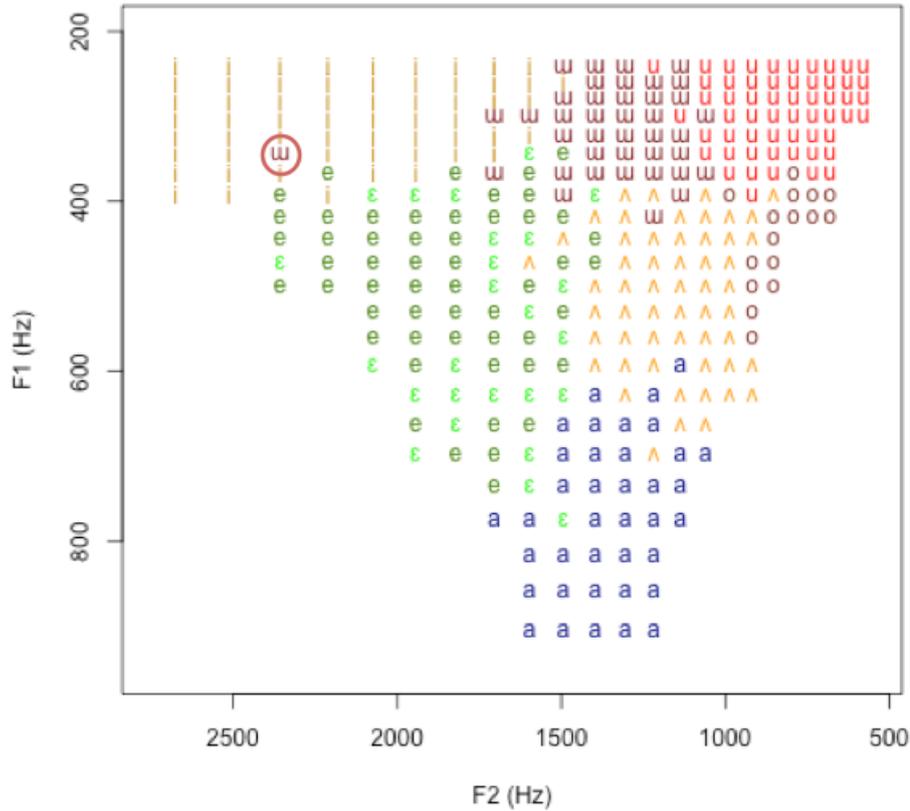


Figure 2.5. Carte perceptuelle pour un sujet coréen (Eychenne & Jang 2015)

l'interprétation des effets principaux (ainsi que des interactions à deux termes) puisqu'on considère en général qu'ils ne peuvent être interprétés de manière fiable en présence d'une interaction significative. Il apparaît toutefois clairement que les femmes de Séoul ont perçu deux voyelles distinctes. Aucun des effets fixes n'était significatif pour F2, mais il faut noter que l'effet aléatoire SUJET était significatif pour F1 et pour F2, ce qui suggère une forte variation inter-individuelle.

En résumé, en ce qui concerne la production, mis à part la différence attendue entre hommes et femmes attribuable au moins en partie à des différences physiologiques, nous n'avons pas observé de différence entre les deux voyelles à travers les trois variétés étudiées ; en perception, seules les femmes de Séoul ont perçu deux voyelles distinctes, quoique celles-ci soient proches et se chevauchent largement dans l'espace acoustique. On peut donc conclure que, bien que la fusion des voyelles soit achevée en production, il demeure une légère distinction en perception, au moins dans l'un des groupes. Une piste d'explication envisageable est qu'il est possible que les femmes de la capitale soient plus sensibles à la norme orthoépique que les autres groupes : le système scolaire enseigne encore la tradition normative selon laquelle il faut distinguer les deux voyelles, et les ma-

		Séoul		Daegu		Gwangju	
		H	F	H	F	H	F
e	F1	531	573	525	557	482	566
		(73)	(91)	(58)	(118)	(46)	(101)
	F2	1868	2261	1896	2246	1981	2210
		(115)	(148)	(243)	(176)	(271)	(227)
ε	F1	539	573	534	578	494	571
		(63)	(87)	(49)	(142)	(51)	(90)
	F2	1910	2268	1856	2209	1917	2159
		(200)	(172)	(223)	(141)	(231)	(238)

Tableau 2.4. Valeurs F1 et F2 (Hz) pour les voyelles /e/ et /ε/ (Eychenne & Jang 2015)

nuels scolaires présentent encore des exercices contenant des paires minimales (ex : /ke/ « crabe » vs /kε/ « chien ») que les élèves sont enjoins à distinguer. La sensibilité des femmes aux formes de prestige a été mise en lumière par Labov (2001 : 266–267) pour l’anglais, mais il faudrait un échantillon bien plus important que celui qui a été analysé pour pouvoir l’établir de manière plus catégorique dans ce cas-ci. A ce stade, cette piste reste une conjecture quoiqu’elle mérite à mes yeux d’être explorée plus avant. La sociolinguistique variationniste étant bien moins développée en Corée qu’elle ne l’est en Europe ou en Amérique du Nord¹⁹, il y a un large travail de documentation des usages à réaliser sur cette langue.

2.3 Conclusion

J’ai tenté de donner dans ce chapitre un aperçu du premier volet de mon travail, à savoir sa dimension descriptive. Ce travail constitue le soubassement de mon travail théorique et il est en ce sens essentiel. Du point de vue du français, ce travail s’inscrit dans une dynamique plus large, issue principalement du projet PFC, visant à renouveler la base empirique sur laquelle les analyses phonologiques du français se fondent, en donnant une place centrale à la variation diatopique. Ces travaux ont donné lieu à plusieurs publications à visée pédagogique, qui permettent de valoriser ce travail descriptif en le rendant accessible aux enseignants et apprenants du français qui souhaitent s’imprégner de la réalité des usages dans leur pluralité. Ce travail descriptif se poursuit bien entendu : le principal chantier en cours de ce point de vue est une analyse des voyelles nasales en français du Midi, travail que je mène en collaboration avec Léa Courdès-Murphy, doctorante à l’Université Toulouse-Jean Jaurès. Nous avons mis en place un codage per-

19. L’étude sociolinguistique la plus conséquente dont j’aie connaissance a été réalisée par Yunsook Hong, une étudiante de Labov, dans le cadre de sa thèse (Hong 1988).

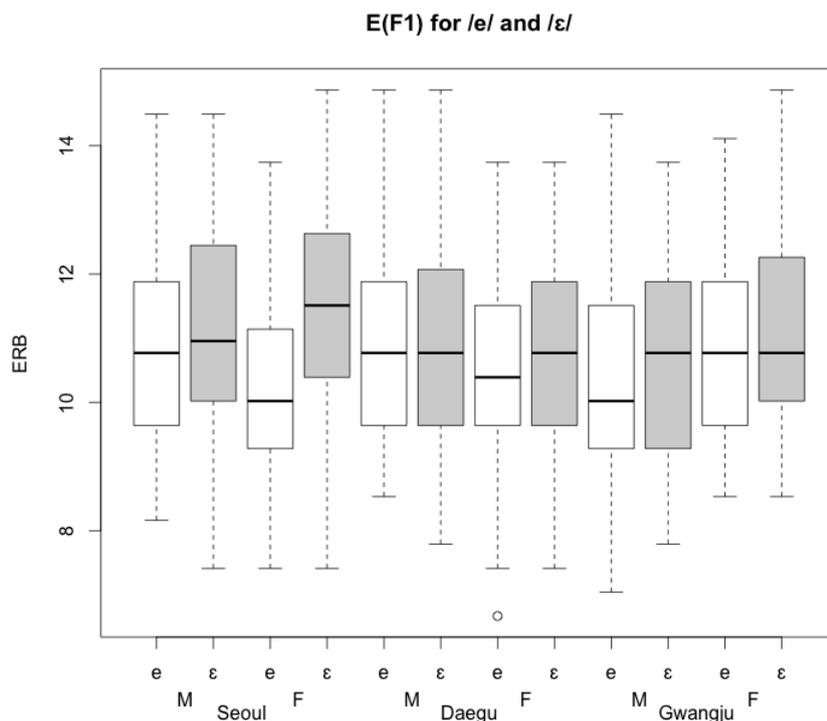


Figure 2.6. F1 (en unités ERB) pour /e/ et /ɛ/ (Eychenne & Jang 2015)

ceptif, inspiré du codage schwa du protocole PFC, pour les voyelles nasales en français méridional. Ce codage sera couplé à une analyse acoustique de la nasalité des voyelles et appliqué à un corpus de locuteurs marseillais (il est également utilisé par Léa Courdès-Murphy dans le cadre de sa thèse, pour le codage d'un corpus plus large).

Mes travaux descriptifs sur le coréen sont pour l'heure de portée plus modeste que mes travaux sur le français, mais la démarche que j'ai développée avec mon collègue Tae Yeoub Jang s'inscrit dans la continuité de mon travail sur le français, en accordant une place centrale à la variation. Nous sommes maintenant en train d'étendre notre analyse du lien entre perception et production à l'ensemble du corpus que nous avons constitué. J'ai par ailleurs entrepris avec Young-Eun Kim, doctorante à Hankuk University of Foreign Studies, une étude phonético-phonologique sur la réalisation des voyelles postérieures non arrondies dans le parler séoulite. Notre objectif est de mieux circonscrire la variation dans la réalisation des voyelles et de déterminer l'influence éventuelle des consonnes environnantes, un phénomène qui est très peu documenté. Nous avons établi le protocole expérimental et le corpus est en cours de constitution.

Chapitre 3

Ressources et outils

Tout travail empirique suppose que l'on dispose d'instruments adéquats. S'il existe des outils génériques permettant d'analyser statistiquement des données, le plus utilisé à l'heure actuelle étant sans doute le langage R, l'utilisation de tels outils présuppose que les données à analyser aient été pré-traitées et soient disponibles dans un format directement exploitable, typiquement un fichier tabulé. En linguistique de corpus, l'obtention d'un tel « produit fini » requiert néanmoins un important travail en amont, allant de la collecte à l'extraction des données, en passant par la transcription, l'annotation et la validation du corpus. Pour tout corpus non trivial, chacune de ces étapes requiert des outils adaptés, si bien qu'il est nécessaire de se doter d'une politique d'outillage appropriée.

Une partie non négligeable de mon travail a été consacrée à développer des outils pour répondre à ces besoins. Ce travail d'ingénierie n'est pas toujours très valorisé, mais je le considère comme un aspect important de ma démarche puisque l'outil informatique joue un rôle pivot entre les données brutes et l'exploitation linguistique qui en est faite. J'ai de ce point de vue, à de multiples reprises et avec divers collègues, contribué à promouvoir l'utilisation d'outils et de méthodes issues de la linguistique de corpus auprès de publics encore peu familiers des développements technologiques récents (Eychemme & Tchobanov 2013^[15] ; Eychemme et al. 2016^[19] ; Eychemme 2016b^[10]).

L'essentiel de mes efforts dans le domaine de la linguistique outillée se sont portés sur le développement de Dolmen (cf. §3.2), un logiciel pour l'analyse de corpus oraux, au sens défini au premier chapitre (cf. §1.4.2). Je commencerai néanmoins par contextualiser cette réalisation en discutant le « prédécesseur » de Dolmen, à savoir la Plateforme PFC (cf. §3.1). Je discuterai ensuite deux autres réalisations, à savoir un corpus de coréen phonétisé que j'ai réalisé en collaboration avec deux collègues (§3.3), et DecoLex, une plateforme d'annotation pour les langues à morphologie riche (notamment le coréen) qui est en cours de développement à HUFs dans le cadre d'un projet de recherche portant sur l'analyse multilingue d'opinions et de sentiments (cf. §3.4).

3.1 Plateforme PFC

Comme je l’ai brièvement mentionné dans le prologue de ce mémoire, j’ai réalisé à partir de ma maîtrise divers outils plus ou moins *ad hoc* pour le codage et l’analyse des données PFC. Ces outils m’ont permis, de la maîtrise au doctorat, d’extraire et analyser mes données, et ont été mis à la disposition des collègues du projet PFC¹. A l’heure où le site PFC n’était pas encore pleinement opérationnel, ils constituaient le moyen le plus commode d’extraire des résultats quantitatifs des codages PFC. Ces outils ont été documentés dans divers bulletins du projet PFC, mais je ne m’y attarderai pas ici.

Ces outils de « première génération » présentaient un inconvénient majeur pour la plupart des collègues : ils étaient en ligne de commande et donc très peu ergonomiques, ce qui était une barrière insurmontable pour beaucoup. Si ma réaction initiale a été de blâmer les utilisateurs potentiels, j’ai fini par réaliser que ce n’était pas à l’utilisateur de s’adapter à l’outil, mais bien à l’outil de s’adapter à l’utilisateur. Par ailleurs, ayant travaillé de manière extensive avec les données PFC, j’avais une idée assez précise du type de programme que j’aurais moi-même voulu avoir à ma disposition lors de la préparation de mon doctorat. Comme à la même époque je venais d’être engagé en tant que post-doctorant à l’université de Calgary, et que le projet sur lequel j’avais été sélectionné consistait précisément à développer une plateforme permettant de faciliter l’utilisation et l’analyse des données PFC, je me suis investi dans le développement d’un nouvel outil, à savoir la « Plateforme PFC ».

Je ne présenterai cet outil que dans les grandes lignes, puisqu’il est maintenant supplanté par Dolmen (cf. §3.2), mais on en trouvera une description détaillée dans l’un des bulletins du projet PFC (Eychenne 2007)^[29]. Il est également décrit dans un chapitre que j’ai écrit en collaboration avec Atanas Tchobanov, ingénieur de recherche à l’Université Paris Ouest Nanterre La Défense (Eychenne & Tchobanov 2013)^[15], paru dans un volume intitulé *Méthodes et outils pour l’analyse phonétique des grands corpus oraux* (Nguyen & Adda-Decker 2013).

Ce logiciel répondait à plusieurs exigences :

- *libre (open source)* : comme tous les outils que j’ai réalisés, cet outil est un logiciel libre, diffusé sous la General Public License (GPL) publiée par la Free Software Foundation². Cela signifie non seulement qu’il est disponible gratuitement, mais aussi et surtout que le code source peut être inspecté et modifié par quiconque a les connaissances techniques requises.
- *multiplateforme* : travaillant moi-même exclusivement sur systèmes UNIX (Linux et macOS), il était essentiel que l’outil fonctionne sur ces plateformes. Bien entendu, l’outil devait également pouvoir fonctionner sur Windows de manière à être facilement utilisable par la majorité des utilisateurs.
- *interopérabilité avec Praat* : les données PFC étant annotées et codées dans le lo-

1. On trouvera une brève description du « classeur-codeur » pour le schwa (et la liaison) dans Durand & Eychenne (2004 : §3.3)^[1].

2. <https://www.gnu.org/licenses/old-licenses/gpl-2.0.html>

giciel Praat, il était important que l'outil puisse communiquer avec ce programme. Cette tâche a été possible du fait que les auteurs de Praat aient mis à la disposition de la communauté un module appelé `sendpraat` qui permet précisément d'envoyer et de recevoir des données vers/depuis Praat.

La plateforme PFC, qui est écrite en python, se présente sous la forme d'une extension (*plugin*) Praat, et peut donc être invoquée directement depuis Praat. Elle offre plusieurs modules, présentés sous la forme d'onglets, qui permettent d'accomplir différentes tâches telles que la gestion des enquêtes et l'analyse des différentes couches d'annotation du projet PFC (orthographe, codage schwa et codage liaison). La figure 3.1 présente l'onglet liaison, avec une partie des résultats correspondant aux liaisons réalisées pour une enquête donnée.

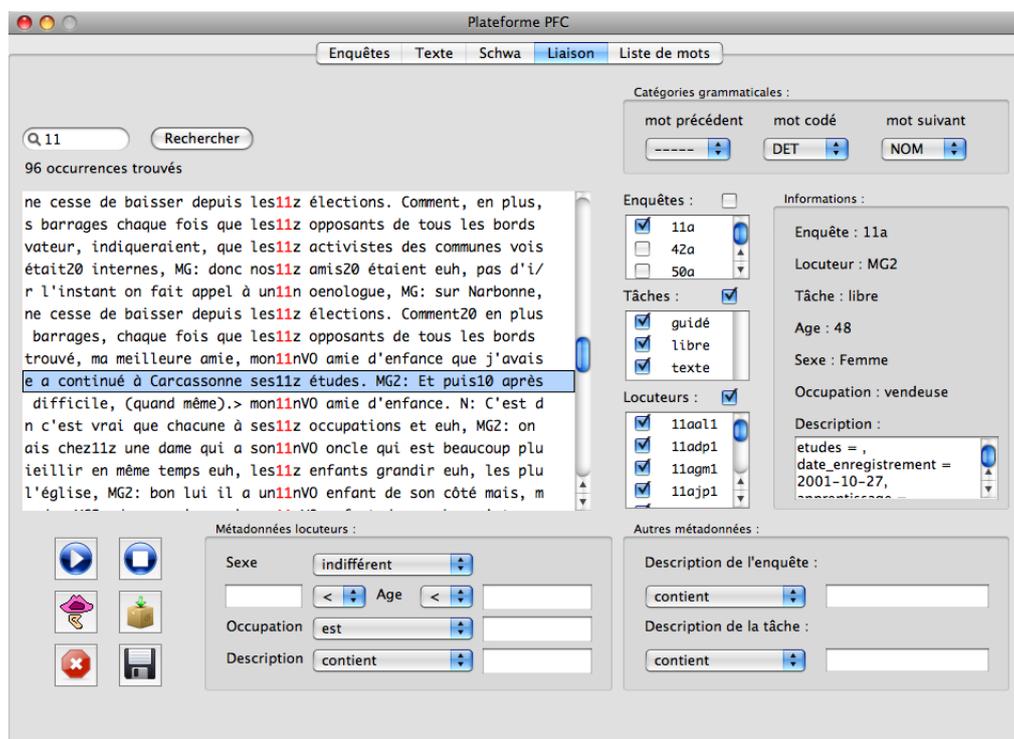


Figure 3.1. Onglet liaison de la plateforme PFC

Les requêtes utilisateur peuvent être filtrées sur la base des métadonnées (âge ou sexe du locuteur, point d'enquête, etc.) et des données elles-mêmes. Les données peuvent quant à elles être filtrées à l'aide d'une simple chaîne de caractères, mais l'outil accepte également tout motif correspondant à une expression rationnelle³ valide, ce qui permet de construire des motifs plus complexes. A titre d'exemple, le motif `"11[^n]"` extraira la chaîne de caractères `"11"` suivie d'un caractère autre que `"n"`, ce qui dans les conventions PFC

3. Une expression rationnelle (ou expression régulière) est une chaîne qui décrit un langage rationnel (cf. §6.1.1.1 p. 142 suiv.). La syntaxe exacte varie selon l'outil utilisé, et la plupart des modules d'expression rationnelles, dont ceux utilisés par la plateforme PFC et Dolmen (décrit à la section §3.2), permettent également de décrire des langages plus puissants que les langages rationnels. On en trouvera une bonne présentation dans Karttunen et al. (1996).

correspond à toute liaison réalisée avec une consonne autre que /n/. Une fois extraits, les résultats sont affichés sous la forme de concordances présentées avec leur cotexte gauche et droit, et dont l’empan peut être ajusté. L’alignement temporel des occurrences étant préservé, chaque concordance peut-être écoutée, directement modifiée ou même ouverte dans Praat si besoin est.

La plateforme PFC, malgré ses avantages, avait plusieurs inconvénients. Le fait qu’elle soit écrite en Python en rendait l’installation complexe pour de nombreux utilisateurs (en particulier les étudiants), problème qui a été exacerbé à l’époque par la transition des architectures 32 bits aux architectures 64 bits, ce qui a provoqué diverses incompatibilités, notamment sur Mac. Par ailleurs, la plateforme était entièrement dépendante de Praat, ce qui complexifiait davantage le déploiement de l’outil. Mais le problème le plus fondamental est que j’avais au départ conçu ce logiciel pour traiter les données PFC, mais de nombreux collègues ont exprimé des besoins allant bien au-delà de ceux que j’avais pu anticiper. Ainsi, les collègues du Projet PAC m’ont tout d’abord demandé d’adapter l’outil pour leurs propres besoins, ce que j’ai fait sans trop de difficulté puisque les deux projets ont des conventions très proches. Cependant, le logiciel était ajusté pour les besoins standard de ces deux projets : il n’était d’une part pas vraiment possible de l’utiliser pour des tâches autres que celles qui avaient été prévus (par exemple, l’ajout d’une tire de codage pour l’analyse du comportement des glissantes) ; d’autre part, plusieurs chercheurs qui avaient pris connaissance de l’existence de cet outil sont entrés en contact avec moi car ils souhaitaient l’utiliser dans leurs propres projets de recherche. L’architecture de l’outil étant, comme je l’ai évoqué, trop rigide, elle ne le permettait tout simplement pas sans une réorganisations structurelle profonde. Ces différentes limitations m’ont donc poussé à repenser cet outil de manière à le rendre plus générique et plus adaptable. C’est alors, en 2010, que j’ai commencé le développement de Dolmen.

3.2 Dolmen

Dolmen est une plateforme pour l’analyse de corpus oraux⁴. La version du logiciel que je discuterai ici est la version 2.0, qui est actuellement en phase de test et de mise au point. Les fonctionnalités restent globalement les mêmes que la version 1.3, telle qu’elle est décrite dans Eychenne, Navarro, Tchobanov & van Leussen (2016)^[19] et Eychenne & Paternostro (2016)^[20], et qui est par ailleurs disponible sur le site web accompagnant le volume *Varieties of French* (Detey, Durand, Laks & Lyche 2016), dans lequel les deux chapitres sus-mentionnés ont été publiés. La version 2.0 apporte néanmoins diverses améliorations, notamment d’un point de vue ergonomique, et elle est construite autour d’une nouvelle architecture et d’un système d’extension plus robuste, ce qui permettra de faciliter la maintenance et l’ajout de nouvelles fonctionnalités sur le long terme.

Mon but n’est pas ici de documenter le fonctionnement de l’outil dans le détail (je renvoie à Eychenne & Paternostro 2016^[20] ainsi qu’au manuel du programme), mais plutôt

4. L’outil est téléchargeable sur ma page web : <http://www.julieneychenne.info/dolmen>

de présenter Dolmen d'un point de vue ingénierique.

3.2.1 Architecture

Si l'outil Dolmen reprend les grands principes (ouverture, portabilité, interaction avec Praat) et les principales fonctionnalités de la Plateforme PFC, il a dès le départ été conçu comme une plateforme modulaire et extensible. Pour ce faire, Dolmen est construit autour d'un système d'extensions (ou *plugins*). De nombreux logiciels, dont Praat, offrent un tel mécanisme : chaque plugin est un module (optionnel) qui peut être ajouté au programme de base, lui conférant de nouvelles fonctionnalités. Ceci offre donc aux utilisateurs la possibilité de configurer et de personnaliser l'outil selon leurs besoins. Ce mécanisme me permet quant à moi d'adapter l'outil pour chacun des projets auxquels je collabore, indépendamment les uns des autres : ainsi, ajouter des fonctionnalités pour le codage d'un nouveau phénomène pour le projet IPFC (cf. §3.2.5) ne requiert qu'une mise à jour du plugin IPFC, sans modification du programme de base.

L'architecture d'ensemble de Dolmen s'articule donc autour de ce système d'extensions. Par ailleurs, fort de l'expérience acquise à travers le développement de la Plateforme PFC, il m'a semblé essentiel que Dolmen soit un outil autonome, installable comme n'importe quel autre programme sur chacune des plateformes supportées (Mac, Windows, Linux), et qu'il soit basé sur des technologies aussi pérennes que possible. Le langage C++ s'est donc imposé comme le meilleur choix pour le développement de ce logiciel : c'est un langage mûr, standardisé, et qui permet de construire des programmes robustes, performants et facilement déployables (toutes choses égales par ailleurs...). Dolmen est écrit dans la version du langage C++ communément appelée « C++11 », qui correspond à la norme ISO/CEI 14882:2011⁵. Le fait de bâtir un logiciel sur un tel standard offre deux garanties importantes : d'une part, l'application sera portable puisqu'elle pourra être compilée sur toute plateforme pour laquelle il existe un compilateur respectant le standard ; d'autre part, le code source de l'application sera disponible sur le long terme puisque les standards C et C++⁶ offrent une excellente rétro-compatibilité. Toutes les bibliothèques utilisées par Dolmen sont également des technologies libres, portables, et écrites en C ou C++ standard, ce qui réduit grandement les risques d'obsolescence.

Quoique le langage C++ offre une base solide pour construire un programme portable et facilement déployable, son inconvénient majeur est qu'il demande un temps de développement plus conséquent que des langages de script tels que Python. Par ailleurs, bien qu'il soit techniquement possible d'écrire des extensions en C++, ce n'est en pratique pas une solution envisageable pour des raisons de temps de développement, mais

5. Cette version offre de nombreuses améliorations en matière de robustesse et de performances, tout en offrant un plus large éventail d'algorithmes et de structures de données standard. Il existe une version plus récente du standard (C++14) et une autre version devrait voir le jour l'an prochain (C++17) ; bien qu'elles apportent elles aussi diverses améliorations, ces versions ne sont pas encore pleinement implémentées dans les compilateurs C++ actuels.

6. Le langage C est l'ancêtre du langage C++ et est considérée comme la *lingua franca* des langages informatiques.

aussi et surtout de difficulté technique pour les utilisateurs et de portabilité⁷. La stratégie que j'ai adoptée, et qui est la plus répandue pour implémenter un système d'extensions, a consisté à intégrer un moteur de script à la plateforme. Après avoir évalué plusieurs options, j'ai décidé d'intégrer le langage de script Lua (Ierusalimschy et al. 1995)⁸, un langage développé au sein du département d'informatique de l'Université pontificale catholique de Rio de Janeiro (Brésil). Contrairement à la plupart des langages de script (par exemple, Python), Lua a été conçu pour être embarqué dans des applications écrites en C ou C++ : il est donc très léger et facile à intégrer dans une application hôte. Ce langage de script est très utilisé dans le monde industriel, en particulier dans l'industrie du jeu vidéo où il est reconnu pour sa facilité d'intégration et ses performances. Il a par ailleurs l'avantage d'être facile d'utilisation et bien documenté (Ierusalimschy 2006).

L'architecture générale de Dolmen, qui est présentée à la figure 3.2, s'articule donc autour de ce moteur de script. L'outil est bâti sur une couche d'abstraction qui permet de s'affranchir des spécificités du système d'exploitation et de l'architecture matérielle hôtes (ex : structure du système de fichiers, ordre des octets en mémoire, etc.). Cette couche intègre diverses bibliothèques libres et portables pour les tâches de base telles que la lecture des fichiers audio ou la gestion de l'Unicode (voir ci-après) ; elle inclut notamment un moteur de base de données relationnelle léger implémentant le standard SQL (*Structured Query Language*) pour la gestion des métadonnées. L'outil intègre ensuite une couche de bas-niveau⁹, écrite en C++, qui est responsable de la gestion des tâches computationnellement lourdes, notamment la gestion du son, la lecture/écriture de fichiers et le traitement des requêtes utilisateur. Enfin, la couche de haut-niveau est presque entièrement écrite dans le langage de script Lua. Cette couche est composée, outre le moteur de script lui-même, d'un ensemble de modules standard pilotant le programme, ainsi que des plugins utilisateurs qui permettent de configurer et d'étendre les fonctionnalités du logiciel. L'intégration d'un moteur de script dans la plateforme permet donc d'écrire une grande partie de la logique de haut-niveau dans un langage flexible et facile d'utilisation, tout en déléguant les tâches computationnellement lourdes au code écrit en C++ lorsque cela s'avère nécessaire. La communication entre le code Lua et le code C++ s'effectue au moyen d'une interface de programmation qui permet d'une part d'effectuer des tâches de programmation courantes telles que la manipulation de chaînes de caractères et l'accès au système de fichiers¹⁰, et d'autre part de contrôler certains aspects de la plateforme (cf. §3.2.5).

7. Chaque extension devrait alors être compilée pour chaque plateforme.

8. <http://www.lua.org>

9. En informatique, on désigne généralement par « bas-niveau » tous les aspects qui dépendent plus ou moins directement de l'architecture matérielle ; les aspects de « haut-niveau » correspondent aux structures de données et opérations qui font abstraction des aspects de bas-niveau. Ainsi, pour un utilisateur (haut-niveau), la chaîne "été" comprend trois « caractères ». Pour la machine (bas-niveau), sa représentation varie selon l'encodage utilisé : elle est par exemple composée de cinq octets lorsqu'elle est encodée en UTF-8, le "é" étant encodé sur deux octets, contrairement au "т" qui n'occupe qu'un seul octet.

10. Lua étant un langage de script embarqué, il n'offre que très peu de fonctionnalités par défaut : c'est donc l'application hôte qui est en charge de fournir les fonctionnalités additionnelles requises.

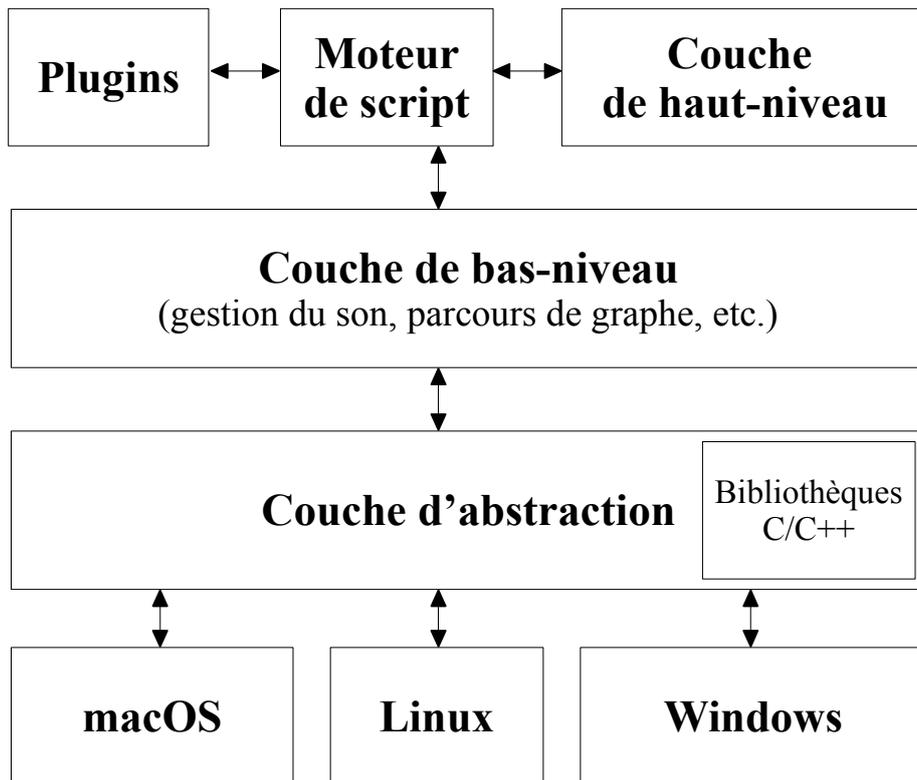


Figure 3.2. Architecture de Dolmen

Enfin, afin de garantir la pérennité et la portabilité de l’outil, l’architecture de Dolmen s’appuie sur trois standards informatiques importants :

- **UNICODE** : il s’agit d’un ensemble de standards permettant de représenter les caractères de manière non ambiguë, quelle que soit la plateforme sous-jacente (The Unicode Consortium 2016). Dolmen 2.0 utilise en interne le standard UTF-8 (le plus courant, en particulier sur le Web), mais il supporte également le standard UTF-16 qui est utilisé par Praat pour l’encodage des fichiers TextGrid contenant des caractères « spéciaux » (c’est-à-dire tous les caractères autres que les lettres de base de l’alphabet latin, les chiffres, et quelques signes de ponctuation).
- **XML** (*eXtensible Markup Language*) : le standard XML est un standard qui s’est imposé avec l’essor du Web pour la représentation et l’échange de données structurées.
- **SQL** (*Structured Query Language*) : il s’agit d’un langage de requête standard pour les bases de données relationnelles. Dolmen utilise un moteur de base de données relationnelle embarqué pour la gestion des métadonnées (voir §3.2.2 ci-après).

Comme Atanas Tchobanov et moi l'avons souligné à plusieurs reprises (cf. Eychenne & Tchobanov 2013^[15] et Eychenne et al. 2016^[19]), l'utilisation de tels standards permet de garantir l'interopérabilité et la bonne conservation des données à long terme, puisque celles-ci peuvent être stockées et lues dans un format ouvert, c'est-à-dire auquel quiconque peut avoir accès, et non ambigu. Par exemple, l'adoption d'Unicode par Praat permet maintenant de s'assurer que les fichiers sont lisibles sur toutes les plateformes, mais il n'en a pas toujours été ainsi : l'échange de fichiers au sein du projet PFC a longtemps été compliqué par le fait que Praat n'implémentait pas ce standard, et plusieurs collègues ont fait la malheureuse expérience de perdre des données en échangeant des fichiers entre Mac et Windows. Si de tels problèmes d'incompatibilité deviennent de moins en moins fréquents, ils n'ont toutefois pas encore totalement disparu, comme mes collègues et moi en avons fait l'expérience dans la construction du corpus K-SPAN pour le coréen (cf. §3.3).

3.2.2 Métadonnées

Comme je l'ai évoqué plus haut, l'un des objectifs majeurs était d'être suffisamment flexible pour pouvoir être adapté aux besoins de projets divers. Il était crucial d'implémenter un système de métadonnées¹¹ qui n'impose que très peu de contraintes aux utilisateurs. Dolmen offre un mécanisme simple, mais qui s'est avéré efficace à l'usage, appelé *propriétés* : une propriété est une paire ⟨attribut, valeur⟩ qui est associée à un fichier. L'attribut, qui est appelé *catégorie*, représente une variable (socio-)linguistique (ex : LOCUTEUR, ÂGE, SEXE) ; la *valeur* permet quant à elle de mettre en relation et de comparer les propriétés des fichiers pour une catégorie donnée. À partir de la version 2.0, les propriétés sont explicitement typées : le type de la valeur de la propriété peut ainsi être TEXTE (une chaîne de caractères, c'est le choix par défaut), NOMBRE (un nombre réel) ou BOOLÉEN (vrai ou faux). Les propriétés d'un fichier peuvent être gérées à partir d'un éditeur intégré à l'interface utilisateur (cf. figure 3.3). Elles peuvent également être importées via un fichier au format CSV ou éditées via l'interface de programmation (cf. §3.2.5).

3.2.3 Modèle de données

L'un des prérequis fondamentaux pour Dolmen était d'être compatible avec le format d'annotation de Praat, à savoir le format TextGrid. Il s'agit d'un format organisé en niveaux d'annotation autonomes appelés *tires*, chaque tire étant une série d'intervalles ou d'instantanés annotés et alignés sur le signal de parole. Malgré son utilité et ses indéniables qualités, le modèle de données de Praat, bien qu'il soit très utilisé, souffre de certaines limitations. Il n'est ainsi pas possible de représenter directement des annotations pa-

11. On entend par « métadonnées » toutes les informations sémantiques qui permettent de structurer ou de mettre en relation les données entre elles.

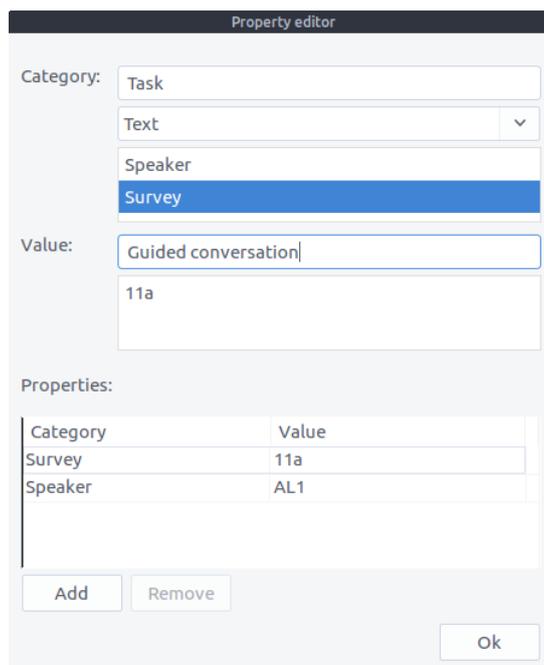


Figure 3.3. Editeur de propriétés de Dolmen

rallèles, correspondant par exemple à des mots alignés avec des parties du discours, ou encore hiérarchiques, c'est-à-dire différents niveaux structurels imbriqués. Afin d'illustrer le problème, la figure 3.4 fournit un exemple d'une segmentation hiérarchique dans Praat telle qu'elle est obtenue par alignement semi-automatique, à l'aide du plugin Praat EasyAlign (Goldman 2011). On y voit trois tires représentant des niveaux hiérarchiques successifs : segments, syllabes et mots. Or si la structure hiérarchique peut être déduite à partir de l'organisation visuelle des tires et intervalles (cf. figure 3.5), elle n'est encodée nulle part : chaque tire étant une couche d'annotation indépendante des autres, il n'y a pas de moyen direct de mettre en relation les différents niveaux. Rien dans le format TextGrid n'indique que la structure est effectivement hiérarchique (ou parallèle, selon le cas), et l'on ne peut donc s'assurer de ce qu'elle est effectivement valide. Ceci peut parfois occasionner des problèmes puisqu'il peut arriver qu'un annotateur déplace accidentellement une frontière d'intervalle, ce qui a pour effet d'invalider la représentation structurelle de l'annotation et peut occasionner des bugs subtils.

Il est donc souhaitable de pouvoir encoder l'information structurelle inter-niveau de manière explicite. Par ailleurs, Praat n'est pas le seul logiciel permettant d'annoter le signal de parole. On citera notamment WaveSurfer (Sjölander & Beskow 2009), EMU (John & Bombien 2014) ou encore EXMARaLDA (Schmidt & Wörner 2014), chacun ayant son propre format de données, et chacun ayant ses propres atouts et faiblesses et/ou servant une communauté d'utilisateurs particulière. Par exemple, le logiciel EMU propose de nombreuses fonctionnalités pour l'analyse phonétique de bases de données parole,

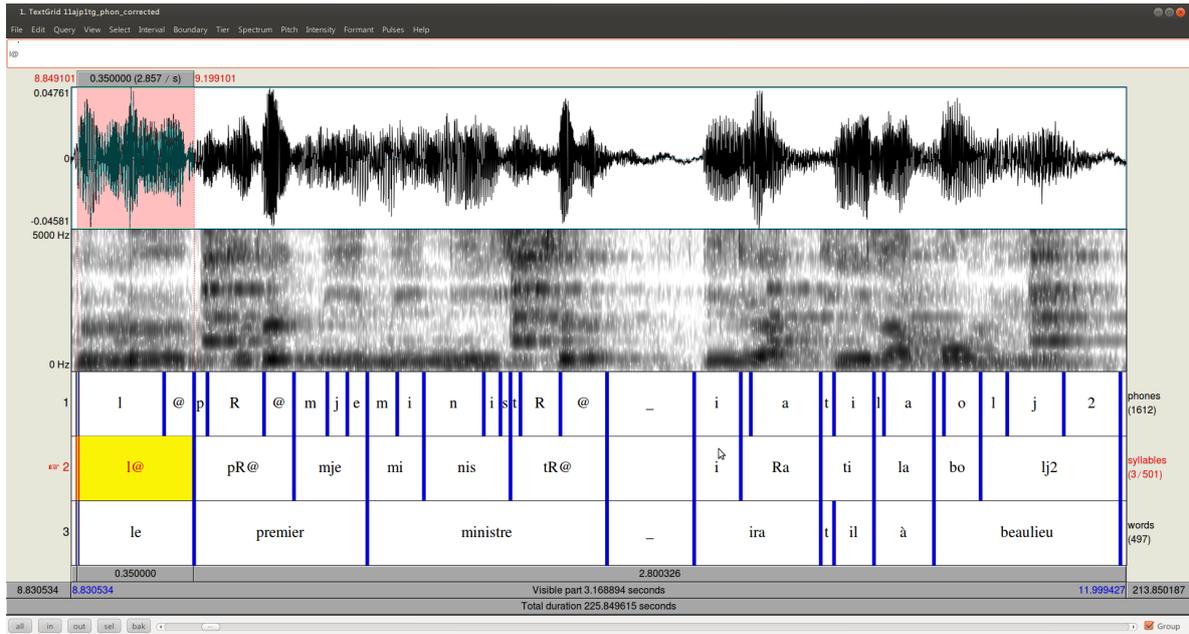


Figure 3.4. Segmentation hiérarchique d'un extrait du texte PFC dans Praat

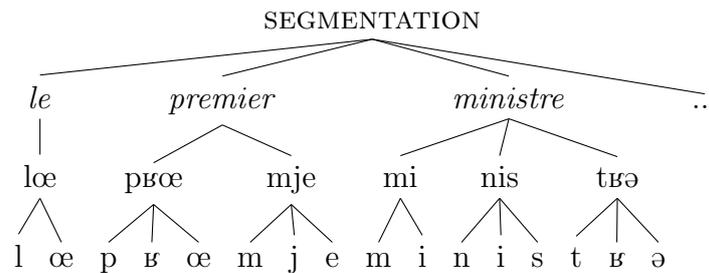


Figure 3.5. Structure hiérarchique partielle correspondant à la fig. 3.4

tout en permettant (à l'inverse de Praat) de représenter explicitement des structures hiérarchiques; EXMARaLDA est quant à lui davantage utilisé dans le domaine de la pragmatique et de l'analyse conversationnelle, domaines dans lesquels il est apprécié pour ses capacités d'annotation de fichiers vidéo.

Face à cette diversité d'outils et de formats, il semble souhaitable de converger vers un modèle de représentation des données standardisé. L'un des standards les plus reconnus en linguistique de corpus est sans doute la *Text Encoding Initiative* (TEI, TEI Consortium 2014)¹², dont l'objectif est de permettre de représenter tout type de données textuelles au format numérique. Il s'agit toutefois d'un format lourd, très complexe (la documentation contient plus de 1600 pages!) et peu adapté à la représentation d'annotations alignées sur le signal de parole car il se concentre sur la structuration et l'annotation de données

12. <http://www.tei-c.org/Guidelines/P5/>

textuelles¹³.

Un pas important dans l'émergence d'un standard pour l'annotation de la parole a été accompli par Bird & Liberman (2001). Leur approche est basée sur deux prémisses fondamentales. La première est qu'il est illusoire d'espérer créer un *format de données* universel, qui puisse être utilisé par l'ensemble de la communauté. La seconde est que, malgré la diversité et la multiplicité des outils et formats existants, il est possible de construire un *modèle de données* qui recouvre la (quasi-)totalité des usages, et qui permette de faciliter l'*inter-opérabilité* des formats existants. Un tel modèle peut être construit à partir de ce que les auteurs appellent un GRAPHE D'ANNOTATION (*Annotation Graph*). Comme son nom l'indique, ce modèle propose de représenter une annotation linguistique sous la forme d'un graphe : les sommets du graphe représentent des bornes temporelles, ancrées dans la ressource audio ou vidéo, alors que les arcs, qui sont étiquetés, représentent les éléments d'annotation. La spécification du contenu des étiquettes des arcs ne fait pas partie du formalisme : ce choix dépend de la nature des données. L'étiquette peut néanmoins être représentée, d'un point de vue formel, comme un n -uplet de paires attribut/valeur. J'illustrerai le fonctionnement de ce formalisme à l'aide d'un exemple concret. Reprenons l'annotation hiérarchique présentée à la figure 3.4 et correspondant à l'arbre de la figure 3.5. Un graphe (partiel) simplifié en est donné à la figure 3.6 : chaque arc est étiqueté avec le niveau structurel auquel il appartient (segment, syllabe, mot) et le label correspondant à l'intervalle qu'il représente dans l'annotation TextGrid. On peut enrichir cette représentation selon les besoins, par exemple un ajoutant un indice positionnel représentant l'ordre dans lequel les tires sont ordonnées, une relation de dominance pour indiquer l'ordre hiérarchique (ex : mot > syllabe), ou toute autre information qui peut être utile à la représentation des données.

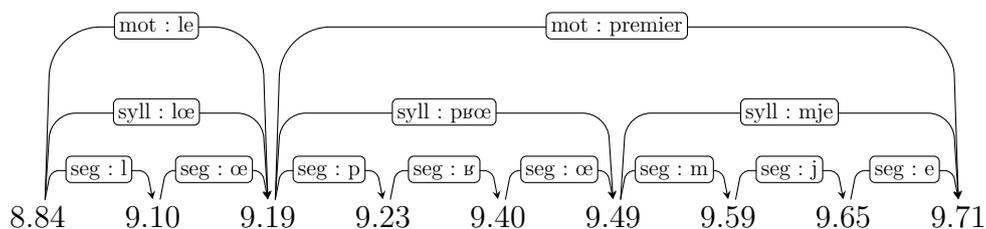


Figure 3.6. Graphe d'annotation partiel pour la fig. 3.4

Ce modèle de données est le modèle utilisé en interne par Dolmen pour représenter les annotations de fichiers sonores. A l'heure actuelle, Dolmen accepte donc trois formats de fichiers pour les annotations : le format TextGrid de Praat, le format LAB de WaveSurfer, et son propre format natif basé sur XML. Quel que soit le format d'entrée, l'annotation est convertie en un graphe d'annotation, qui peut ensuite être exporté vers chacun de ces trois formats selon les besoins.

13. La section 8 de la TEI (TEI Consortium 2014 : 233–255), intitulée « Transcriptions of Speech », offre un ensemble de recommandations pour la transcription de la parole, mais celles-ci restent très génériques et ne semblent pas avoir été adoptées par la communauté (cf. Bird & Liberman 2001 : 26).

Le modèle développé par Bird & Liberman (2001) a donné lieu à des développements ultérieurs, et un standard ISO en a même récemment émergé : le LANGUAGE ANNOTATION FRAMEWORK (LAF, Ide & Suderman 2014). Il s'agit d'un modèle de données plutôt que d'un format, mais les concepteurs du LAF ont proposé un format de données XML basé sur le LAF : le GRAPH ANNOTATION FRAMEWORK (GrAF, Ide & Suderman 2007). Ce format n'a à ma connaissance été implémenté par aucun des principaux outils utilisés en phonétique et en phonologie de corpus. Dolmen ne permet donc pas à l'heure actuelle d'exporter les fichiers dans le format GrAF ; il s'agit d'ailleurs d'un format particulièrement verbeux et qui génère par conséquent des fichiers très volumineux qu'il serait peu efficace de traiter. Néanmoins, si ce format devient largement adopté par la communauté et s'impose effectivement comme un format d'échange, il sera très facile d'implémenter un filtre de sortie puisque la traduction d'un graphe d'annotation à la Bird & Liberman (2001) en un graphe de type GrAF est un tâche très simple.

3.2.4 Requêtes

La fonctionnalité centrale de Dolmen est la fonction de concordancier pour les annotations de parole (figure 3.7). Elle est similaire dans son principe à la fonction offerte par la Plateforme PFC, et permet donc d'effectuer des requêtes en texte brut ou à l'aide d'expressions rationnelles, d'écouter la portion du signal correspondant à la transcription, d'éditer le texte de l'annotation ou de l'ouvrir dans Praat (s'il s'agit d'un TextGrid), ou encore d'exporter les résultats dans un fichier tabulé. Parmi les fonctions supplémentaires, j'en signalerai trois qui n'existent pas dans la plateforme : (i) la possibilité de marquer des signets (*bookmarks*) dans le corpus, de manière à pouvoir retrouver facilement certains éléments d'annotation ; (ii) la possibilité d'effectuer des requêtes sur plusieurs tires simultanément, ce qui permet de faire des recherches dans des annotations organisées en tires hiérarchisées (arbres prosodiques ou syntaxiques) ou parallèles (ex : mots alignés avec leur catégorie grammaticale) ; (iii) la possibilité de comparer plusieurs annotateurs, de manière à pouvoir calculer leur taux d'accord et à déterminer les occurrences sur lesquelles ils divergent.

3.2.5 Plugins

Comme je l'ai mentionné plus haut, il est possible d'écrire des extensions (implémentées dans le langage Lua) permettant d'étendre les possibilités de l'outil. La communication entre Dolmen et les extensions s'effectue au moyen d'un mécanisme appelé « signal/slot » : lors de certaines tâches (par exemple, le chargement d'un fichier), la plateforme émet un signal spécifique ; les extensions peuvent connecter ce signal à une ou plusieurs fonction(s) appelées « slot(s) ». Chaque fois qu'un signal est émis, les slots auxquels il est associé sont exécutés. Les extensions peuvent par ailleurs créer de nouveaux signaux, utiliser des slots standard fournis par le programme et étendre l'application à l'aide de menus personnalisés. Ce principe de conception logicielle permet d'accroître la modularité

du logiciel en découplant la plateforme logicielle de ses extensions.

Afin d'illustrer les bénéfices de cette implémentation, je prendrai l'exemple des métadonnées PFC de base. Le nommage des fichiers dans le projet PFC suit des conventions strictes, qui permettent d'identifier chaque fichier de manière unique. Le nom de base du fichier (dépourvu d'extension) est composé de trois caractères correspondant au code de l'enquête, suivis de trois caractères correspondant au code du locuteur, suivis d'un caractère indiquant la tâche (lecture de mots, lecture de texte, conversation libre ou guidée) et d'un caractère représentant le type de fichier (fichier son vs annotation). Par exemple, le nom de fichier 85a_{gm1}mg correspond (de droite à gauche) à l'annotation (g) de la liste de mots (m) du locuteur gm1 de l'enquête 85a (Treize-Vents, Vendée). Puisque ces métadonnées sont entièrement prédictibles, il est inutile de demander à l'utilisateur de les entrer manuellement dans le logiciel. La version 1 de Dolmen incluait du code C++ spécifique au projet PFC, qui pouvait être activé via une option dans les préférences du programme. Cette stratégie n'est évidemment pas généralisable¹⁴ et elle a l'inconvénient de complexifier la structure du logiciel. *A contrario*, la nouvelle architecture découple complètement l'ajout des métadonnées PFC de l'outil lui-même. L'ajout des métadonnées s'effectue au moyen d'un petit module écrit en Lua, qui est fourni au sein du plugin PFC. Ce module ajoute une entrée dans un menu intitulé PFC (lui-même créé par le plugin), ce qui permet à l'utilisateur d'ajouter ces métadonnées au moyen d'un simple clic (le code de ce module est inclus dans l'annexe A.1 à des fins d'illustration).

En plus de l'interface de programmation qui permet de tirer profit des fonctionnalités de la plateforme, Dolmen offre un mécanisme appelé GRAMMAIRES DE REQUÊTES (*query grammars*)¹⁵ qui permet de faciliter l'exécution de requêtes pour les besoins d'un projet précis. Le problème que cette fonctionnalité vise à résoudre est le suivant : de nombreux chercheurs, en phonologie de corpus et en linguistique variationniste notamment, mettent en place des systèmes de codage (le plus souvent perceptif) pour l'analyse des variables linguistiques qu'ils étudient. C'est par exemple le cas du schwa et de la liaison dans le projet PFC, et Sylvain Detey, Isabelle Racine et moi avons besoin de mettre en place de nouveaux codages dans le cadre du projet IPFC sur la phonologie des locuteurs non natifs. S'il est bien sûr possible d'effectuer des requêtes à l'aide d'expressions rationnelles pour tout type de codage, il nous a paru nécessaire de mettre en place une interface graphique conviviale pour chaque codage, de sorte que les collaborateurs peu rompus à l'utilisation de l'outil informatique puisse exploiter les données sans difficulté. Les grammaires de requêtes permettent donc de spécifier la structure du codage dans un format déclaratif stocké dans un fichier, qui est ensuite interprété par Dolmen pour créer dynamiquement une interface graphique correspondant au codage : l'utilisateur n'a plus qu'à cliquer sur des boutons pour construire et exécuter sa requête. Les expressions rationnelles corres-

14. De fait, la version 1 contient également du code spécifique au projet PAC sur l'anglais.

15. Cette fonctionnalité a été implémentée grâce au soutien de Sylvain Detey (Université de Waseda, Tokyo), dans le cadre du projet *A corpus-based longitudinal study of the interphonological features of Japanese learners of French* (Japanese Society for the Promotion of Science, Grant-in-Aid for Scientific Research (B) n°23320121 (2011-2014). Qu'il en soit à nouveau remercié !

pondantes sont construites par l’outil, en fonction des choix de l’utilisateur, et celui-ci n’a donc pas besoin d’en connaître la syntaxe.

J’illustrerai cette fonctionnalité à l’aide du codage des voyelles nasales dans IPFC. Ce codage comprend 6 champs répartis sur 12 caractères alphanumériques. Ils correspondent à (i) la qualité perçue de la voyelle cible, (ii) le contexte phonologique gauche, (ii) contexte phonologique droit, (iv) l’évaluation perceptive du degré de nasalité, (v) l’évaluation perceptive du timbre de la voyelle et (vi) l’évaluation perceptive de la présence d’un appendice consonantique (ou d’une consonne de liaison). En voici un exemple pour la voyelle /ã/ réalisée [am] dans *campagne* par un apprenant japonophone : 11_3P_40P_2_10_2M. Comme on peut le voir, ce codage est assez complexe. L’interface correspondant à la grammaire de requête est présentée à la figure 3.8. Bien qu’il soit nécessaire aux annotateurs de connaître le codage, les données codées peuvent ensuite être exploitées via cette interface par n’importe quel utilisateur, sans qu’il y ait besoin de connaître le codage. Il est par ailleurs aisé de modifier cette interface si le codage évolue (il suffit de modifier le fichier texte correspondant) et l’on peut créer autant de grammaires que nécessaire (le projet IPFC en compte cinq, et d’autres sont en cours de réalisation).

The interface consists of six panels arranged in a 2x3 grid:

- Cible:** A list of checkboxes for target vowels: [ã], [ô], [è], [ôè], [an/am], [on/om], [en/em], [un/um], [in/im].
- Contexte gauche:** Options for context: (longue) pause à gauche, frontière de mot à gauche, and élément consonantique à gauche (with a dropdown menu).
- Contexte droit:** Options for context: pause + consonne, frontière de mot + consonne, élément consonantique + frontière/pause (with a dropdown menu), milieu de mot (with a dropdown menu), liaison potentielle non réalisée (with a dropdown menu), and liaison potentielle réalisée (with a dropdown menu).
- Nasalité:** Options for nasality: voyelle nasale, voyelle orale + nasalisation, voyelle orale, and voyelle nasale non réalisée.
- Qualité spectrale:** Options for spectral quality: timbre correct (with a dropdown menu) and timbre incorrect (with a dropdown menu).
- Appendice ou consonne de liaison:** Options for the appendix: aucun élément postvocalique dans le contexte _C, élément consonantique postvocalique (with a dropdown menu), and site de liaison (_V).

At the bottom, there are three buttons: 'Compare annotators', 'Reference annotator:' (with a text input field), and 'Select fields'.

Figure 3.8. Interface pour l’analyse des voyelles nasales dans IPFC

3.3 K-SPAN : corpus de coréen phonétisé

L’un des projets auquel je me suis attelé au cours des douze derniers mois a été, en collaboration avec Jeffrey Holliday (spécialiste du coréen) et Rory Turnbull (psycholinguiste), la construction d’un corpus phonétisé de coréen. Le point de départ de ce travail a été la constatation qu’il n’existait pas de telle ressource librement disponible. Jeffrey

Holliday, qui venait d’obtenir un poste à l’université de Corée (고려 대학교) à Séoul en septembre 2015, m’avait fait part de ses travaux en cours, en collaboration avec Rory Turnbull (actuellement post-doctorant à l’Ecole Normale Supérieure à Paris), sur la densité de voisinage (*neighborhood density*) en coréen (voir §3.3.3 ci-après). Dans une étude préliminaire (Holliday & Turnbull 2015), ils avaient utilisé un corpus orthographique qu’ils avaient tenté de phonétiser sur la base de règles de conversion graphème-phonème : cette stratégie est adoptée par la quasi-totalité des travaux en phonologie du coréen (voir notamment Wedel et al. 2013a ; Silverman 2010 ; Oh et al. 2015), mais elle produit un résultat imparfait puisque la forme phonétique de surface¹⁶ n’est souvent pas prédictible à partir de la forme orthographique. Or, j’avais de mon côté, dans un travail réalisé avec Tae-Youb Jang et que je décrirai au chapitre 5 (voir §5.3.4 p. 133), adopté une stratégie similaire à la leur mais en utilisant le corpus Sejong, qui est le corpus écrit de référence pour le coréen (Kim 2006). J’avais rencontré des problèmes identiques aux leurs, ce qui m’a amené à explorer la possibilité d’utiliser un dictionnaire en ligne en libre accès pour obtenir des phonétisations fiables. Nous avons donc décidé de mutualiser nos efforts de manière à créer un corpus de coréen phonétisé, qui serait librement accessible à la communauté.

Le corpus résultant de cette collaboration est disponible en libre accès sur l’archive TROLLing (voir Holliday, Turnbull & Eychenne 2016a). Le fruit de ce travail est synthétisé dans une publication (Holliday, Turnbull & Eychenne 2016b)^[32], dont la version révisée est en cours d’évaluation. Le reste de cette section décrit la construction du corpus dans les grandes lignes.

3.3.1 Correspondance graphie/phonie en coréen

Afin de comprendre les difficultés spécifiques que pose la phonétisation du coréen, il est nécessaire de comprendre la structure du système d’écriture de cette langue. Le système graphique du coréen, appelé *Hangeul*, est un système alpha-syllabique qui a été conçu au XV^e siècle pour transcrire le coréen, jusque là écrit en caractères chinois (*Hanja*), ce afin de démocratiser l’accès à l’écriture (Kim-Renaud 1997)¹⁷.

Le Hangeul est composé de 40 lettres (*jamo*), qui ont pour la plupart une valeur phonémique, et qui s’organisent en syllabes graphiques appelées *eumjeol*. A titre d’exemple, le mot 사진 « photographie », qui correspond à la forme phonologique /satɕin/, est composé de deux eumjeol. Le premier, 사, est composé des jamo ㅅ /s/ et ㅏ /a/. Le second, 진, est composé des jamo ㅈ /tɕ/, ㅣ /i/, et ㄴ /n/. L’ordre de lecture des jamo dans un

16. Dans le cadre de ce travail, ce qui est appelé « forme phonétique » ou « forme de surface » correspond à une transcription phonétique large qui ignore les processus strictement allophoniques. Par exemple, les consonnes lâches deviennent voisées entre sonantes (ex : 바다 /pata/ « mer » → [pada]), et la fricative sibilante /s/ est palatalisée devant /i/ ou /j/ (ex : 신 /sin/ « dieu » → [ɕin]). Ces variations allophoniques ne sont pas pertinentes pour les mesures de similarité lexicale telles que la densité de voisinage.

17. Les caractères Hanja sont de nos jours essentiellement utilisés afin de désambigüiser les homonymes et de préciser l’étymologie ou le sens des mots.

eumjeol est gauche/droite s'il n'y a pas de coda (ex : 사), ou gauche/droite/bas s'il y en a une (ex : 진).

Le caractère syllabique de cet alphabet permet d'encoder certaines différences morphologiques. Ainsi, la forme phonologique /tali/ peut-elle être écrite de deux manières différentes. La première, 다리 /ta.li/, est un mot monomorphémique qui signifie « jambe », alors que la seconde, 달이 /tal.i/, est un mot morphologiquement complexe composé du nom /tal/ « lune » et du suffixe nominatif /i/. On remarquera que dans cette deuxième forme, le second eumjeol commence par le jamo ㅇ, qui représente ici une attaque vide. Le processus de resyllabification général que connaît le coréen, et qui est d'ailleurs similaire au phénomène d'enchaînement en français (Encrevé 1988), provoque la resyllabation du /l/ de coda en attaque de la syllabe suivante. Ainsi, bien que les deux mots soient prononcés de manière identique, la différence morphologique est préservée par l'orthographe.

Comme l'exemple précédent l'illustre, la relation entre graphie et phonie n'est pas bi-univoque. Il existe par ailleurs tout un ensemble de processus phonologiques qui rendent cette mise en correspondance encore moins directe. Tout d'abord, comme nous l'avons vu en §2.2.2, certaines distinctions vocaliques comme l'opposition /e ~ ε/ sont maintenant perdues dans le coréen séoulite contemporain. De plus, le coréen connaît de nombreux processus qui neutralisent des oppositions phonologiques dans divers contextes (Ahn 1998 ; Shin et al. 2013 : chap. 8). A titre d'exemple, l'opposition entre /p/ (lenis), /p^h/ (fortis aspirée) et /p*/ (fortis tendue) est neutralisée en coda, et ces trois consonnes sont réalisées comme une plosive bilabiale non relâchée [p̚]. Si ces phénomènes de neutralisation sont nombreux et interagissent parfois, ils ne sont en général pas problématiques pour un système de phonétisation automatique puisqu'ils sont entièrement prédictibles.

Il existe cependant divers processus, parfois idiosyncrasiques, qui sont sensibles à la structure morphologique du mot, et qui ne sont pas reflétés dans la forme orthographique. Pour ne prendre qu'un seul exemple, dans certains mots dont le second morphème commence par /i/ ou /j/, la consonne /n/ est « insérée » entre les deux morphèmes. Ainsi, 담요 /tam#jo/ « couverture » est construit à partir du morphème sino-coréen 담 /tam/ « couverture » et du morphème natif 요 /jo/ « matelas coréen ». Ce mot, dont la structure morphologique est /tam#jo/, subit ce processus d'épenthèse¹⁸ et est donc réalisé [tamnjo]. Toutefois, ce processus ne s'applique pas à tous les morphèmes commençant par /i/ ou /j/. Ainsi, le mot 금요일 /kum#jo/ « vendredi », qui contient les morphèmes 금 /kum/ « or » et 요 /jo/ « brillant », est réalisé [kumjo] de manière transparente. Je ne rentrerai pas ici dans le détail de l'ensemble des processus morpho-phonologiques du coréen (voir Shin et al. 2013 : ch. 9 pour un aperçu des principaux phénomènes) ; ce

18. D'un point de vue historique, le /n/ faisait partie du second morphème et a subi un processus d'effacement en position initiale suivie de /i/ ou /j/. Comme pour la liaison en français, les spécialistes sont partagés sur le statut sous-jacent vs épenthétique de ces consonnes en synchronie. On notera que dans certaines variétés, notamment dans la province de Hamgyeong (Corée du nord), le /n/ peut-être réalisé dans des contextes où il n'est pas attendu d'un point de vue étymologique. Ainsi, le mot /mokjoil/ « jeudi », qui est réalisé [mogjoil] en coréen standard, est réalisé [moŋnjoil] dans cette variété, avec nasalisation régressive de l'obstruante précédant /n/ (Martin 2006 : 17). Ces faits semblent soutenir l'approche épenthétique, du moins dans ces variétés.

qui importe ici est que dans de nombreux cas, il n'est pas possible de prédire la forme phonétique à partir de la forme orthographique, et celle-là ne peut donc être dérivée automatiquement par règle à partir de celle-ci.

3.3.2 Stratégie de phonétisation

Afin d'obtenir un corpus de coréen phonétisé fiable, nous avons utilisé conjointement deux ressources librement accessibles. La première est un corpus de coréen édité par le National Institute of the Korean Language (NIKL). Ce corpus, qui est disponible en libre accès¹⁹, est un corpus dictionnaire lemmatisé (non phonétisé). Nous avons ignoré les morphèmes grammaticaux pour nous concentrer sur les 82501 lemmes présents dans ce corpus. Chaque lemme est accompagné de diverses informations, notamment sa fréquence lexicale dérivée d'un corpus de textes, sa catégorie grammaticale, ainsi qu'une éventuelle désambiguïsation, en Hanja pour les mots d'origine sino-coréenne, en caractères latins pour les mots d'emprunt récents (à l'anglais notamment). Pour obtenir les formes phonétiques, nous avons utilisé le dictionnaire en ligne de l'entreprise Naver^{20 21} : celui-ci fournit, pour chaque forme graphique, une liste des différentes entrées disponibles (en cas de polysémie) et, pour les mots dont la prononciation ne correspond pas directement à l'orthographe, Naver fournit également une pseudo-transcription phonétique en Hangeul.

Reprenons les exemples de **금요**, dont la réalisation est transparente (à la resyllabation près), et de **담요**, qui subit le processus d'épenthèse de /n/. Pour le premier, Naver ne fournit aucune indication phonétique, alors qu'il fournit la phonétisation [담:뇨] pour le second. Ces représentations pseudo-phonétiques n'utilisent que les caractères du Hangeul, augmentés du symbole de longueur pour indiquer les voyelles longues, qui ne sont pas notées en Hangeul. Pour obtenir la forme phonétique de **담요**, il suffit donc d'appliquer les règles de conversion graphème/phonème à la représentation pseudo-phonétique [담:뇨], ce qui donne [ta:mnjo]. La phonétisation de **금요** est quant à elle obtenue directement à partir de la forme orthographique. Les formes homographes étaient en général, mais pas de manière systématique, distinguées par un code numérique identique dans les deux ressources (rappelons qu'elles sont toutes les deux éditées par le NIKL). Par exemple, aucune phonétisation n'est fournie pour **가마**₄ (une unité de volume), ce qui signifie que ce mot est réalisé de manière transparente [가마] (= [kama]), alors que **가마**₅ « chaise à porteurs » est phonétisé [가:마] (= [ka:ma]), avec une voyelle longue dans la syllabe initiale.

Afin de phonétiser le corpus NIKL, nous avons adopté la stratégie suivante. Nous avons converti le fichier du corpus NIKL en Unicode (UTF-8) car le corpus était fourni dans un

19. http://korean.go.kr/front/reportData/reportDataView.do?mn_id=45&report_seq=1&pageIndex=1

20. <http://krdic.naver.com/>

21. Le dictionnaire Naver est lui-même basé sur un dictionnaire en ligne édité par le NIKL ; en revanche, la version de Naver est accessible via le protocole standard Common Gateway Interface et produit du code HTML facile à traiter, ce qui permet d'obtenir des résultats de manière automatique à l'aide d'un script dédié.

encodage non standard (Windows code page 949). Dans ce processus, nous avons identifié 343 formes qui contenaient des caractères Hanja invalides dans leur désambiguïsation ; 314 de ces formes apparaissent dans le corpus final. Une fois le corpus converti, nous avons tenté d'obtenir une forme phonétique dans le dictionnaire de Naver. Pour les mots non ambigus, nous avons extrait la seule forme fournie par Naver. Pour les mots polysémiques, nous avons utilisé la désambiguïsation en Hanja et/ou le code numérique, selon l'information disponible. Les mots qui n'ont pas pu être identifiés dans le dictionnaire en ligne, soit parce qu'ils étaient absents, soient parce qu'ils renvoyaient à une forme similaire mais différente (ex : 가가소소하다, un mot très rare, renvoie à 소소하다), ont été éliminés. Sur les 82501 lemmes du corpus NIKL, 18665 ont ainsi été éliminés. Nous avons donc obtenu 63836 formes phonétiques, soit 77,4% du corpus original.

Pour un nombre non négligeable de formes (5018 items, soit 7,9% du corpus), Naver nous a fourni deux variantes de prononciation, l'une avec l'application de divers processus de neutralisation optionnels, et l'autre sans. L'application de ces processus, tels que l'effacement de /j/ après une obstruante vélaire (ex : /sikje/ « montre » → [sike]) est censée rendre compte de la prononciation actuelle du coréen. Toutefois, si chaque processus, pris de manière isolée, était appliqué systématiquement à l'une ou l'autre variante, l'ensemble des processus ne l'était pas. Ainsi, bien que la plupart des processus typiques de la prononciation actuelle à Séoul aient été appliqués à la seconde variante, certains, comme l'insertion de la glissante /j/ entre /i/ et /ʌ/, l'ont été à la première. En outre, plusieurs traits importants de la prononciation contemporaine, notamment la perte d'opposition de longueur et la fusion de /e/ et /ɛ/, n'étaient tout simplement pas indiqués.

Nous avons donc décidé de créer, à partir de l'information obtenue sur le dictionnaire en ligne, deux variantes de prononciation. La première, que nous avons appelée « conservatrice », représente un système plus archaïque, et quelque peu artificiel, dans lequel toutes les oppositions ont été préservées et aucun des processus optionnels n'a été appliqué. La seconde variante, que nous avons qualifiée de « moderne », représente une forme à laquelle a été appliqué l'ensemble des processus de neutralisation et de fusion que connaît le coréen séoulite. Par exemple, les voyelles ㅟ et ㅠ ont été rendues comme les monophthongues /y/ et /ø/ respectivement dans la variante conservatrice, alors qu'elles ont été transcrites comme /wi/ et /wɛ/ dans la variante moderne. Pour obtenir les deux variantes, nous avons tout d'abord linéarisé les deux transcriptions pseudo-phonétiques fournies par Naver à l'aide d'un algorithme de décomposition standard (The Unicode Consortium 2016 : §3.12) qui traduit chaque eumjeol en la séquence de jamo correspondante. Ainsi, la chaîne "가공" est elle linéarisée en "ㄱ ㅏ ㅓ ㅓ ㅓ ㅓ" (= /kakoŋ/). Nous avons ensuite aligné les deux chaînes ainsi linéarisées en utilisant l'algorithme *Minimum Edit Distance*, tel qu'il est implémenté dans Cock et al. (2009) et avons ensuite créé chacune des variantes en assignant chaque non-correspondance entre les deux chaînes à la variante appropriée. Nous avons par ailleurs appliqué à la variante moderne les processus dont Naver ne rend pas compte. Premièrement, nous avons appliqué la neutralisation de l'opposition /e ~ ɛ/ en /ɛ/ mentionnée plus haut. Notons que cette neutralisation a pour effet de neutraliser l'opposition entre ㅟ /we/ et ㅠ /wɛ/ en /wɛ/ d'une part, et ㅟ /je/ et

ㅈ /jɛ/ en /jɛ/ d'autre part. Etant donné que la diphtongue ㅈ /ø/ est par ailleurs neutralisée en /wɛ/, phénomène qui est lui pris en compte par Naver, cette étape a pour effet de neutraliser l'opposition tripartite ㅈ /ø/ vs ㅉ /we/ vs ㅊ /wɛ/ en /wɛ/. Deuxièmement, nous avons neutralisé l'opposition de longueur puisque celle-ci ne semble jouer au mieux qu'un rôle marginal dans le coréen moderne, et la plupart des spécialistes s'accordent sur le fait que l'opposition a maintenant disparu dans la prononciation séoulite (Lee & Ramsey 2011 : 296-297, Shin et al. 2013 : 153, Sohn 1999 : 14).

J'illustrerai ce processus d'alignement à l'aide d'un exemple concret. Prenons le cas du mot 가갯방 /kakes#paŋ/ « petite boutique ». Naver propose la phonétisation suivante, avec deux variantes : [가:게뻥/가:겜뻥], soit [ka:kep*ɑŋ] et [ka:kɛp*ɑŋ]. Nous voyons ici que Naver a appliqué le processus (obligatoire) de tensification des obstruantes après obstruante aux deux formes (/p/ → [p*]), et a appliqué le processus (facultatif) de simplification des séquences homorganiques (/t/ → /p/ → ∅) à la première. Aucune des formes n'a en revanche subi les processus de neutralisation de longueur ou de neutralisation de l'opposition /e ~ ɛ/, caractéristiques de la prononciation moderne. Pour obtenir les formes attendues, soit [ka:kɛp*ɑŋ] pour la forme conservatrice et [kakep*ɑŋ] pour la forme moderne, nous procédons en deux étapes. Tout d'abord, les formes sont linéarisées et alignées, comme on le voit à la figure 3.9a, après quoi on applique à la première variante (la forme moderne) les processus qui ne sont pas pris en compte par Naver (3.9b).

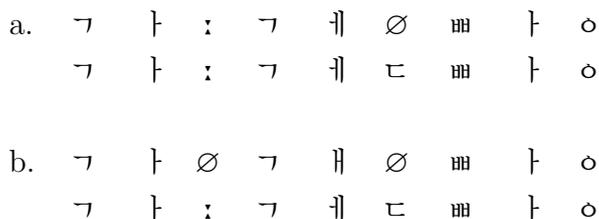


Figure 3.9. Alignement des variantes pseudo-phonétisées pour 가갯방

Ce processus de phonétisation s'est avéré très efficace dans l'ensemble. Un examen attentif des données a cependant révélé quelques problèmes qu'il nous a fallu corriger. Ces problèmes sont dus au fait que la phonétisation proposée par Naver n'est pas parfaite, et qu'elle introduit donc parfois certaines erreurs. Par exemple, le mot d'emprunt 로봇 /lopos/ « robot », qui est régulièrement réalisé [lopot], était phonétisé par Naver en [로뿏], ce qui aboutit à la forme de surface erronée *[lobos]). Nous avons donc appliqué un ensemble de règles auxiliaires aux formes phonétiques obtenues de manière à nous assurer qu'aucune forme de surface ne viole la phonotaxe du coréen (en l'occurrence, [s] n'apparaît jamais en coda). Nous avons ainsi corrigé 98 formes erronées. Enfin, afin de s'assurer de la qualité du processus de phonétisation global, nous avons extrait un échantillon de 1000 formes sélectionnées de manière aléatoire, que nous avons vérifiées manuellement. Aucune erreur n'a été identifiée lors de cette étape.

Le tableau 3.1 illustre quelques exemples de phonétisation typiques résultant de la chaîne de traitement dans son ensemble. Les représentations phonétiques sont proposées

dans le corpus final sous deux formes différentes. La première utilise le système Worldbet (Hieronymus 1994), qui est une forme étendue l’alphabet SAMPA, lui-même une représentation de l’API n’utilisant que des caractères ASCII ; l’autre forme repose sur une correspondance bi-univoque entre jamo et caractères ASCII. La première forme est destinée à être lue par les linguistes, alors que la seconde est destinée à être traitée par des algorithmes informatiques dédiés.

Orthographe	Forme conservatrice	Forme moderne
가갯방	ka:kɛtp*aŋ	kakɛp*aŋ
간행되다	kanhɛŋtøta	kanhɛŋtwɛta
시계	si:kje	sikɛ
전의	tɕɛɹni	tɕɛɹni
최고가	tɕʰø:kok*a	tɕʰwɛkok*a
밤일	pamnil	pamnil

Tableau 3.1. Illustration du processus de phonétisation

3.3.3 Perspectives d’exploitation linguistique

L’objectif premier de ce travail était la création d’un lexique phonétisé librement disponible pour la communauté des linguistes. Ce corpus a deux atouts majeurs : il offre d’une part des représentations phonétiques fiables, qui permettent de pallier les limitations d’une approche de conversion graphème/phonème naïve ; d’autre part, il permet aux chercheurs qui ne sont pas spécialistes du coréen de pouvoir tester des hypothèses sur la structure du lexique de cette langue.

L’une des lignes de recherche que mes collègues et moi poursuivons, et qui sera entreprise dans une étape ultérieure de ce travail, est de tester la réalité psychologique de la notion de densité de voisinage en coréen. Dans son acception la plus simple, et la plus répandue, la densité de voisinage d’un mot est le nombre d’items dans le lexique qui diffèrent de ce mot par exactement une opposition phonémique²². Comme nous le soulignons dans Holliday, Turnbull & Eychenne (2016b), de nombreux travaux ont mis en évidence le rôle de la densité de voisinage dans l’accès au lexique. En production, il a ainsi été découvert que l’accès à un item lexical est plus lent lorsque celui-ci a une densité de voisinage élevée (Luce & Pisoni 1998). En production, plusieurs études ont mis au jour une tendance à l’hyperarticulation dans les mots ayant une densité de voisinage élevée (cf. Munson & Solomon 2004 ; Wright 2004), ce qui a été interprété, dans une perspective fonctionnaliste, comme une manifestation du principe de maximisation de la distinctivité (Grammont 1933 : 176). D’autres études enfin ont également suggéré que cette notion joue un rôle important dans la phase d’acquisition : le lexique des enfants tend à contenir

22. Cette notion est proche, mais distincte, de la notion de rendement fonctionnel que je discuterai en détail ultérieurement (cf. §5.3 p. 116 suiv.).

davantage de mots ayant une densité de voisinage élevée au début, puis à devenir moins dense par la suite (Coady & Aslin 2003 ; Stokes 2010).

La grande majorité de ces études porte cependant sur l'anglais, et le rôle éventuel que cette notion pourrait jouer dans d'autres langues, *a fortiori* dans des langues typologiquement très différentes telles que le coréen, reste largement à élucider. En plus des formes phonétiques de surface, notre corpus inclut diverses instantiations de la mesure de densité de voisinage, pour les deux formes phonétiques (conservatrice et moderne) que nous avons dérivées mais aussi pour la forme orthographique, et ce au niveau segmental et au niveau syllabique. L'étape ultérieure de notre travail est donc de tester la pertinence de la notion de densité de voisinage en coréen. La question de fond à laquelle nous espérons contribuer est de savoir si l'effet de la densité de voisinage est universel et constant de langue à langue, et si ce n'est pas le cas de tenter de déterminer pourquoi, et dans quelle mesure, cette notion est active dans certaines langues mais pas dans d'autres. Nous n'envisageons bien entendu pas de trancher cette question, mais mes collègues et moi espérons que notre travail sur le coréen apportera des éléments de réponse.

3.4 DecoLex : annotation des langues à morphologie riche

Je terminerai cette discussion des principaux outils et ressources que j'ai réalisés par une brève présentation de DecoLex, un outil en cours d'élaboration, que j'implémente en collaboration étroite avec Jisun Nam (HUFS) dans le cadre d'un projet interdisciplinaire subventionné par la *National Research Foundation of Korea* (NRF-2015R1A5A7036384), dont J. Nam coordonne la partie linguistique. Je ne décrirai pas le fonctionnement de l'outil en détail ici (je renvoie le lecteur intéressé à Eychemme 2016a^[30]) et j'en présenterai simplement les grands principes.

Cet outil, écrit en C++ et construit sur les mêmes technologies que Dolmen, est destiné à l'annotation morpho-syntaxique et sémantique de corpus écrits : l'objectif du projet pour la partie linguistique est le développement d'un large corpus multilingue²³ annoté pour l'analyse d'opinions et de sentiments. Les modèles les plus couramment utilisés pour ces tâches sont des variantes de modèles statistiques dits BAG-OF-WORDS, qui traitent un document comme un espace à n dimensions, où chaque mot représente une dimension et sa fréquence représente un point le long de cette dimension. Ce type de modèles, extrêmement naïfs d'un point de vue linguistique, fonctionnent assez bien pour des langues à morphologie pauvre comme l'anglais mais sont peu adaptés à des langues agglutinantes telles que le coréen.

Notre objectif est donc de développer un corpus multilingue annoté qui, dans une deuxième phase du projet, sera utilisé par des informaticiens et des statisticiens pour tenter de construire des modèles plus adaptés et qui puissent dans la mesure du possible

23. Les langues prises en compte incluent notamment le coréen, le vietnamien, le russe, le mongol et le tagalog.

fonctionner sur plusieurs langues. Mon travail de développement se situe donc en amont, et vise à faciliter la tâche des annotateurs (une dizaine d'étudiants et doctorants) puisqu'il n'est pas envisageable d'étiqueter le corpus manuellement.

L'outil accepte en entrée des fichiers au format texte brut (une phrase par ligne) ou bien des graphes produits par le programme Unitex²⁴ où chaque phrase est représentée sous forme d'un AUTOMATE FINI (voir §6.1.1.1 p. 142 suiv. pour une discussion détaillée des automates finis). Ces graphes permettent d'une part de représenter la décomposition morphologique de chaque mot, mais ils permettent également de prendre en compte les cas d'ambiguïté morphologique, fréquents en coréen. Cette méthode d'annotation s'appuie sur l'approche lexicale/grammaire développée par Maurice Gross et ses collaborateurs (cf. Gross 1987).

J'en donnerai un exemple avec la phrase 기린 유명하다고 하던데... « Il paraît que la (bière) Kirin est célèbre ». Le mot 기린 *kirin* a trois analyses morphologiques possibles :

- 기리 (NOM) + ㄴ (TOPIQUE) → « aubaine »
- 기리 (VERBE) + ㄴ (PASSÉ) → « commémoré »
- 기린 (NOM) → marque de bière ou « girafe »

L'interface de l'outil offre donc un premier module qui permet à l'annotateur de sélectionner l'analyse morphologique correcte pour chaque mot et, si nécessaire, de la modifier manuellement (dans l'exemple ci-dessus, c'est la troisième analyse qui est correcte). L'outil permet ensuite d'annoter le corpus à l'aide d'étiquettes sémantiques, qui sont définies dans un fichier texte et peuvent donc être modifiées ou étendues au besoin. Le modèle d'annotation est basé sur ce que nous avons appelé des CHAÎNE D'ÉTIQUETAGE (*tag chain*), qui est une séquence ordonnée (mais pas nécessairement contiguë) de morphèmes, chaque morphème étant représenté de manière unique par sa position dans le mot et la position du mot auquel il appartient dans la phrase. Par exemple, dans le syntagme 삼성 갤럭시가 /samsʌŋ#kɛllʌksi+kɑ/ « le Galaxy de Samsung + SUJET », le marqueur du sujet /kɑ/ est orthographiquement rattaché au mot /kɛllʌksi/ bien qu'il dépende de l'entité /samsʌŋ kɛllʌksi/ que l'on souhaite ici étiqueter. On annotera donc dans ce cas la séquence de morphèmes /samsʌŋ kɛllʌksi/ comme une seule chaîne d'étiquetage. Cette approche permet par ailleurs d'annoter des chaînes de morphèmes non contiguës telles que la négation *ne... pas* en français par exemple.

Une fois le fichier d'entrée étiqueté, il peut être sauvegardé dans un format XML conçu pour les besoins du projet. C'est dans ce format XML que le corpus sera fourni aux informaticiens pour qu'ils entraînent leurs modèles statistiques. Il est également prévu que le corpus et l'outil d'annotation soient mis à la disposition de la communauté universitaire une fois le projet achevé.

24. <http://www-igm.univ-mlv.fr/~unitex/>

3.5 Conclusion

J'ai présenté un ensemble de réalisations technologiques pour l'annotation et l'analyse de corpus. Le plus significatif de ces outils est Dolmen, un outil auquel j'ai abouti de manière itérative, en tenant compte de mes propres besoins et de ceux exprimés par les utilisateurs. On pourra légitimement se demander si tous ces efforts de développements sont bien nécessaires. Je répondrai à cela que l'écriture du code source d'un programme n'est que la partie émergée de l'iceberg : il y a en amont tout un travail de conception, qui doit prendre en compte de multiples contraintes et accepter divers compromis (temps de développement initial, facilité de maintenance à long terme, facilité d'utilisation, performance, etc.) et un aval un important travail de maintenance qui vise à corriger les problèmes existants, à s'assurer que le programme reste opérationnel malgré les évolutions technologiques, ou encore à étendre les fonctionnalités de l'outil pour répondre aux besoins des utilisateurs. Par ailleurs, il en va du développement de programmes informatiques comme il en va du développement de modèles théoriques : il est presque impossible d'obtenir un résultat satisfaisant du premier coup et il s'agit donc d'un travail cumulatif et itératif. De même que personne n'utilise aujourd'hui la théorie de l'optimalité dans la version qui a été introduite par Prince & Smolensky (1993), le modèle ayant été étendu et raffiné d'innombrables manières en fonction des besoins des analystes, de même l'outil informatique évolue en permanence afin de corriger les failles techniques, malheureusement inévitables, et de prendre en compte les nouveaux besoins.

Parmi les développements en cours et à venir, je travaille actuellement en collaboration étroite avec Sylvain Detey de manière à automatiser davantage l'extraction de résultats dans Dolmen. L'objectif est d'extraire automatiquement des statistiques descriptives par groupe et par sujet dans le cadre du projet IPFC. Ici comme ailleurs, cette fonctionnalité sera implémentée de telle sorte qu'elle puisse bénéficier à l'ensemble des utilisateurs. Nous envisageons par ailleurs le développement d'une fonctionnalité permettant d'accélérer le précodage des données. Le codage est une étape extrêmement coûteuse, si bien qu'il est difficile de recourir à plus de deux annotateurs par fichier²⁵. Notre objectif est de mettre en place une procédure qui puisse pré-coder, dans la mesure du possible, les champs prédictibles, de sorte que l'annotateur n'ait qu'à se focaliser sur les champs perceptifs à l'aide d'une interface ergonomique. Ceci nous permettra d'accroître le nombre d'annotations par fichier et d'obtenir ainsi des données plus fiables.

25. Rappelons que Zwirner recommandait au moins trois annotateurs (cf. §1.4.1).

Chapitre 4

Modélisation symbolique

Les deux précédents chapitres ont fait état de la dimension empirique et appliquée de mon travail. Le présent chapitre, ainsi que le chapitre suivant, présentent mon travail de modélisation théorique. Je me concentrerai ici sur le travail de modélisation symbolique. S'il est vrai comme je l'ai déjà souligné que mes recherches ont pris une orientation plus explicitement quantitative ces dernières années, ceci ne doit pas être interprété comme un reniement du travail symbolique que j'ai accompli, même si je serai amené à nuancer sur certains points les positions que j'ai défendues et les analyses que j'ai proposées. La modélisation symbolique est un instrument utile pour progresser dans notre compréhension des phénomènes langagiers, et modélisation symbolique et quantitative sont selon moi complémentaires, et non antagoniques. Je donnerai donc dans ce chapitre un aperçu des travaux de modélisation que j'ai menés et qui ont porté sur la liaison et le schwa en français.

4.1 Liaison

Je commencerai par synthétiser un travail sur la liaison qui est partiellement basé sur le dernier chapitre de ma thèse (Eychenne 2011)^[2]. Comme je l'ai mentionné au chapitre 2, cet article représente une étape charnière dans mon parcours puisqu'il amorce la transition vers un travail plus quantitatif tout en gardant un pied en phonologie générative classique. Je considère avec le recul que je n'ai pas réussi à intégrer dans ce travail les dimensions empirique et théorique de manière pleinement satisfaisante, c'est d'ailleurs la raison pour laquelle j'ai choisi de discuter la partie empirique séparément (cf. §2.1.3 p. 39 suiv.). Je considère néanmoins que l'analyse que j'ai proposée pour la liaison en théorie de l'optimalité va plus loin que les analyses précédemment formulées dans ce cadre (Tranel 2000), aussi me paraît-il utile d'en présenter les grandes lignes.

Deux questions se posent pour toute analyse de la liaison en phonologie générative : quelle est la représentation des consonnes de liaison (CL) et quelles sont les opérations calculatoires qui rendent compte de la (non-)réalisation de ces consonnes ? La solution que j'ai proposée se fonde sur deux innovations en théorie de l'optimalité (OT) qui, prises

conjointement, permettent de modéliser la liaison d’une manière digne d’intérêt. Je rappelle que la théorie de l’optimalité, dans sa forme canonique, considère qu’une forme de surface est le résultat de la satisfaction d’un ensemble de contraintes hiérarchiquement ordonnées : des contraintes structurales (aussi appelées contraintes de marque), qui pénalisent certaines configurations dans les formes de surface (ex : une syllabe sans attaque) et des contraintes de fidélité qui pénalisent toute disparité entre la forme sous-jacente et la forme de surface. Les contraintes de fidélité sont une réponse salutaire à l’abstraction débridée qui caractérisait la phonologie SPE et qui permettait à Schane (1968) de dériver en synchronie *père* [pɛr] de /patɛrn/.

L’analyse que j’ai développée se base sur une critique de l’analyse proposée par Tranel (2000), qui considère que la liaison relève d’une tendance générale à l’évitement de l’hiatus. J’ai montré que cette analyse posait au moins deux problèmes : le premier problème, d’ordre théorique, est la manière dont les contraintes de fidélité fonctionnent dans son approche. Selon Tranel, l’apparition de la liaison est crucialement due à une contrainte anti-hiatus, comme on peut le voir en 4.1 (la contrainte DEP(L) interdit l’épenthèse d’une consonne de liaison).

/pɛti ⟨t⟩ ano/	*VV	DEP(L)
☞ [pɛeti t ano]		*
[pɛeti ano]	*!	

Tableau 4.1. Réalisation de la liaison (Tranel 2000 : 52)

Or selon Tranel (2000 : 51) les CL ont une « détermination lexicale cruciale », et il est intéressant de noter au passage que Tranel rejoint ici l’analyse qu’il avait proposée en 1981 dans une perspective linéaire. Si les CL sont effectivement présentes lexicalement, il est alors impossible qu’elles violent une contrainte de la famille DEP comme en 4.1, puisque ces contraintes ne sont violées que lorsqu’un élément en output ne correspond à aucun élément en input. Le deuxième problème soulevé par cette approche anti-hiatus est qu’elle ne permet pas d’expliquer la réalisation de la liaison lorsque le mot de gauche se termine par une consonne, comme dans *belles* [z] *années*.

J’ai donc proposé une analyse alternative qui s’appuie sur la dynamique diachronique de la liaison, à savoir l’effacement des consonnes finales en position de coda (Laks 2005a). Cette approche se base par ailleurs sur un principe fondamental de la théorie de l’optimalité, relatif à la correspondance input/output, et qui a été formulé par McCarthy & Prince (1993) mais qui a été assez vite délaissé, pour n’être « redécouvert » que plus tard (van Oostendorp 2005). Il s’agit du principe de COHÉRENCE DE L’EXPONENCE, énoncé ci-dessous :

(1) Cohérence de l'exponence

L'exponence morphologique d'un morphème phonologiquement spécifié ne peut être modifiée.

(McCarthy & Prince 1993 : 21)

Ce principe stipule que l'affiliation morphologique (ou exponence) d'un élément ne peut être modifiée de l'input à l'output. Par exemple, un schwa épenthétique dans le néerlandais /arm/ → [arəm] « bras » ne sera jamais reconnu comme faisant partie du morphème en output puisqu'il a une exponence nulle, et ce bien qu'il soit intégré prosodiquement dans celui-ci. Ceci peut être représenté de la manière suivante :

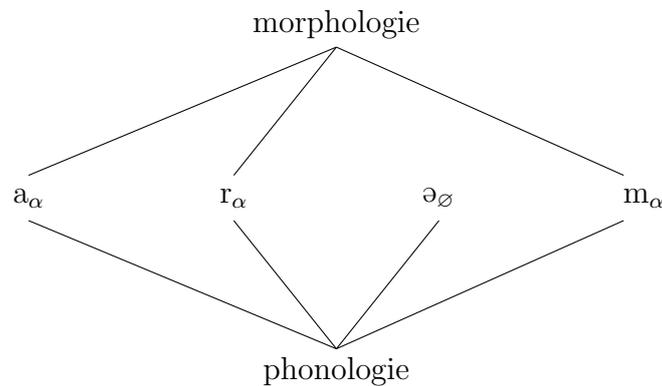


Figure 4.1. Exponence morphologique (d'après van Oostendorp 2005)

J'ai donc proposé d'appliquer ce principe à l'analyse de la liaison en assignant aux consonnes latentes une exponence morphologique spécifique (notée λ). Ainsi, les formes masculine /pəti<t>/ et féminine /pətit/ de *petit(e)* ont respectivement la forme $p_\alpha \emptyset_\alpha t_\alpha i_\alpha t_\emptyset$ et $p_\alpha \emptyset_\alpha t_\alpha i_\alpha t_\alpha$.

Du point de vue computationnel, cette analyse exploite un développement théorique proposé par Pater (2007) qui permet de dédoubler certaines contraintes en une version générique et une version spécifique, la version spécifique étant co-indexée avec certains morphèmes et ne s'appliquant qu'à ceux-ci. La contrainte spécifique doit toujours dominer la contrainte générique, sans quoi elle serait toujours sans effet dans la grammaire. J'ai donc proposé de découpler la contrainte structurale PASDECODA de telle sorte que la variante spécifique PASDECODA $_\lambda$ cible spécifiquement les CL. Le tableau 4.2 donne une illustration de la réalisation (en contexte prévocalique) et non-réalisation (en contexte préconsonantique) de la liaison. Pour rendre compte de la réalisation du /z/ du pluriel dans des contextes comme *belles* [z] *années*, j'ai proposé, sur la base de travaux antérieurs (Rice 2006), d'intégrer une contrainte de fidélité portant sur l'expression des catégories grammaticales, plus spécifiquement la contrainte MAX(Plur) qui demande à ce que tout morphème du pluriel soit réalisé en output. J'ai montré comment cette contrainte, en inter-agissant avec les contraintes précédentes, permettait de rendre compte de manière adéquate de la réalisation du morphème du pluriel /z/ (Eychenne 2011 : 96–97).

/grãt _λ emwa/	ATTAQUE	PASDECODA _λ	Max(C)	PASDECODA	*λ
☞ [grã.t _λ e.mwa]					*
[grã.e.mwa]	*!		*		
/grãt _λ denivele/					
[grãt _λ .de.ni.və.le]		*!		*	*
☞ [grã.de.ni.və.le]			*		

Tableau 4.2. *grand émoi* vs *grand dénivelé* (Eychenne 2011 : 95)

L’approche que j’ai proposée me semble avoir plusieurs atouts : (i) elle s’accorde avec la dynamique diachronique de la liaison ; (ii) elle rend compte du statut lexical spécifique de la liaison, sans recourir à la notion par ailleurs problématique de « flottance » ; (iii) la représentation spécifique attribuée aux CL s’accorde avec les travaux qui ont montré que les CL diffèrent des consonnes d’attaque (voir Tremblay & Spinelli 2014 : 95–96 et les références citées). Pour autant, elle pose (ou du moins ne résout pas) certains problèmes : (a) elle ne prend pas en compte l’effet de la fréquence lexicale ; (b) elle ne prend pas en compte les constructions spécifiques dans lesquelles la liaison peut apparaître et qui conditionnent pourtant sa réalisation (voir la discussion de *c’est* vs *est* p. 43) ; (c) elle ne propose aucun mécanisme pour rendre compte des liaisons variables. De manière plus générale, la difficulté à laquelle se heurte cette approche est qu’elle tente de traiter dans une phonologie du lexique, au sens défini au chapitre 1, ce qui est un problème relevant pour l’essentiel du vocabulaire.

4.2 Le pied en français

Une partie non négligeable de mon travail a été consacrée au pied en français et, de manière connexe, au mot prosodique. A l’occasion d’un festschrift en l’honneur de Jacques Durand, ma collègue Helene Andreassen et moi avons décidé de revisiter la question du pied en français puisque Jacques Durand y avait consacré de nombreux travaux (notamment Durand 1976, 1986, 1995). Différentes analyses ont été proposées pour analyser la structure (micro-)prosodique du français : certains auteurs ont soutenu l’existence d’un pied trochaïque, alors que d’autres ont soutenu l’existence d’un pied iambique. Il nous a donc paru utile de tenter de faire le point sur cette question.

4.2.1 Schwa final en français

Pour saisir tout l’enjeu du débat, il est nécessaire de rappeler qu’il n’y a pas de consensus sur le statut du schwa en français (cf. §2.1.2 p. 36 suiv.), notamment en position finale de mot. Dans la lignée de Schane (1968) et Dell (1973), certains postulent l’existence

d'un schwa dans des formes telles que *verte* /vert+ə/, le schwa pouvant correspondre à un marqueur du féminin ou à une marque flexionnelle, alors que d'autres (ex : Tranel 1981) rejettent l'existence de tels schwas « protecteurs » dont le rôle est pour l'essentiel de permettre aux consonnes latentes d'être réalisées : par exemple, le singulier /vert/ verrait sa consonne finale effacée alors qu'elle est préservée dans le féminin, sous l'action du schwa qui la déroberait à l'effacement. Dans la perspective linéaire et dérivationnelle adoptée par ces auteurs, ces schwas abstraits sont effacés par une règle tardive. Or il se trouve que le français méridional commun offre une situation très similaire, à ceci près que le schwa final n'est effacé que devant voyelle (bien qu'il tende maintenant à être effacé de plus en plus par les jeunes locuteurs). Ceci a amené des auteurs tels que Schane à suggérer l'existence de représentations supradialectales communes, le français non méridional ayant un processus catégorique additionnel qui efface le schwa une fois qu'il a rempli son office (s'assurer de la réalisation des consonnes finales).

L'un des problèmes majeurs rencontrés par les analyses linéaires abstraites qui supposent une règle d'effacement des consonnes finales (*petit* /pətit/ → [pəti]) est que certaines consonnes ne sont jamais effacées : il faut donc soit postuler un schwa qui n'a pas de motivation indépendante (ex : *chef* /ʃɛfə/, Dell 1970), soit traiter ces consonnes finales comme exceptionnelles (Schane 1968). Les données du français méridional commun permettent de départager ces approches sans équivoque, puisque le *e* graphique correspond à un schwa phonologique (comme dans les approches abstraites) mais qu'il n'y a pas de schwa dans des formes du type *chef* : il faudrait donc, dans une perspective linéaire, traiter ces consonnes comme exceptionnelles au moins pour ces variétés. Les analyses non linéaires subséquentes ont affiné les analyses (ex : Encrevé 1988) en proposant que la consonne latente soit une consonne FLOTTANTE qui puisse être ancrée prosodiquement par l'action d'une voyelle (et donc possiblement, d'un schwa), mais ces analyses ne remettent pas fondamentalement en cause l'idée qu'il existe des schwas abstraits en finale de mot qui ne sont jamais réalisés phonétiquement.

4.2.2 Le pied en français

Pour rentre compte entre autres choses du caractère inaccentué du schwa dans le Midi (cf. *correct* [ko.'rɛkt] vs *correcte* [ko.'rɛk.tə]), Durand (1976) a proposé une analyse en phonologie de dépendance selon laquelle toute syllabe contenant une voyelle pleine (autre que schwa) projette un pied, alors que les syllabes dont le noyau est un schwa sont rattachées à la syllabe précédente et forment avec elle un trochée. Une analyse similaire a été indépendamment proposée par Selkirk (1978) pour le français de référence (en supposant l'existence de schwas abstraits). La figure 4.2 donne une représentation dans l'esprit des deux approches (mais plus proche de l'analyse de Selkirk), en utilisant la hiérarchie prosodique standard (Nespor & Vogel 1986) où, comme il est d'usage, les symboles σ , Σ , ω représentent respectivement la syllabe, le pied et le mot prosodique.

Ce type d'analyse n'a, à ma connaissance, pas été remise en cause pour le français méridional. En revanche, bien que le pied trochaïque soit encore utilisé par certains au-

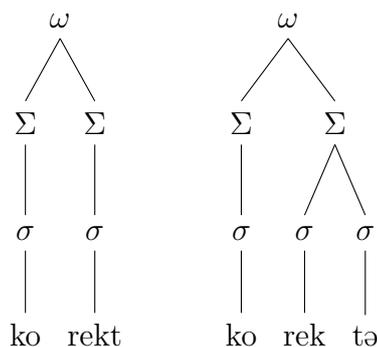


Figure 4.2. Pédification de *correct* et *correcte*

teurs qui défendent une analyse abstraite de la (morpho-)phonologie du français septentrional (Montreuil 2002), d'autres auteurs ont défendu l'existence de pieds iambiques pour rendre compte de l'accent lexical final du français (Hammond 1995), de la morphologie non concaténative (*faculté* → *fac*, *cinéma* → *ciné*) (Scullen 1997) et des données acquisitionnelles (Rose 2000).

Du point de vue de la morphologie non concaténative d'abord, il faut noter que, bien qu'il existe un grand nombre de mots qui suivent le patron iambique, il existe aussi, comme Scullen (1997 : §3.2.3.2) l'avait d'ailleurs noté elle-même, un nombre non négligeable de formes déviantes qui ne se conforment pas aux patrons iambiques canoniques (à savoir H, LL ou LH).

(2) Formes apocopées déviantes¹ :

<i>professionnel</i>	→	<i>pro</i>	L
<i>distribution</i>	→	<i>distrib</i>	HH
<i>hebdomadaire</i>	→	<i>hebdo</i>	HL
<i>écologiste</i>	→	<i>écolo</i>	LLL

Il semble tout au plus que les formes apocopées montrent une préférence pour les mots prosodiques binaires, sans nécessairement faire référence au poids syllabique.

Par ailleurs, un examen détaillé de la distribution de l'accent lexical (primaire et secondaire) soulève d'importantes difficultés dès lors que l'on prend en considération et les données septentrionales et les données méridionales. Je ne reprendrai pas le détail de notre argumentation ici (Andreassen & Eychenne 2013 : 129–131), mais j'en illustrerai tout de même la teneur à l'aide de quelques exemples :

(3) Accent lexical primaire et secondaire :

1. L et H représentent respectivement une syllabe légère (c'est-à-dire ouverte en français) et une syllabe lourde (fermée).

- | | | |
|----|-----------------------|-----------------------|
| a. | <i>expérimental</i> | [ɛks.pe.ri.mã.'tal] |
| | <i>caricatural</i> | [ka.ri.ka.tu.'ral] |
| b. | <i>gouvernemental</i> | [gu.vɛr.nə.mã.tal] |
| | <i>événementiel</i> | [e.vɛ.nə.mã.sjɛl] |
| c. | <i>expérimentale</i> | [ɛks.pe.ri.mã.'ta.lə] |
| | <i>caricaturale</i> | [ka.ri.ka.tu.'ra.lə] |

Nous voyons dans les pentasyllabes en (3a) ce qui s'apparente visiblement à un patron iambique dont le pied principal est aligné à droite (Hammond 1995). Ce patron est perturbé en (3b), où la troisième syllabe, dont la tête est un schwa, ne reçoit pas d'accent secondaire : c'est la syllabe précédente qui le porte, ce qui apparaît compatible avec l'hypothèse d'un pied trochaïque, la syllabe avec schwa étant prosodifiée comme le membre faible du trochée (ex : [gu.⟨vɛr.nə⟩_Σ.mã.tal]). Les formes (3c), que l'on rencontre en français méridional, posent alors un véritable problème : la présence du schwa final suggère que les deux dernières syllabes forment un trochée, et ceci est d'ailleurs nécessaire pour rendre compte de l'ajustement des voyelles moyennes selon la loi de position (voir §4.3 *infra* pour une discussion détaillée). Mais il n'est alors plus possible d'expliquer le parallélisme accentuel entre (3a) et (3c) puisqu'en (3c) la syllabe pénultième ne peut être à la fois la tête du trochée final (ex : [ɛks.pe.ri.mã.⟨'ta.lə⟩_Σ]) et du iambe formé avec la syllabe précédente (ex : [ɛks.pe.ri.⟨mã.'ta⟩_Σ.lə]). La seule solution à ce casse-tête serait de traiter le schwa final comme extramétrique, comme l'ont par exemple suggéré Delais-Roussarie & Rialland (2007) : cette approche prédit cependant que l'on ne devrait pas rencontrer de schwa interne de morphème, puisque les constituants extramétriques se situent normalement au bord gauche ou droit de leur domaine (Hayes 1995 : 57–58). L'existence de formes telles que *céleri* [sɛ.lə.ri] en français du midi montre que cette hypothèse n'est pas tenable (je reviendrai en détail sur la question du schwa dans ce type de formes en §4.3.1).

4.2.3 Approche stratifiée avec grille seule

Pour pallier les difficultés rencontrées avec le pied, nous avons exploré la possibilité de rendre compte de la structure métrique du français dans le modèle GRILLE SEULE proposé par Prince (1983)². Nous avons par ailleurs intégré cette approche dans une version stratifiée de la théorie de l'optimalité (Kiparsky 2000 ; Bermúdez-Otero 2007).

Notre approche reconnaît l'existence de trois lignes métriques au niveau lexical, notées L_0 , L_1 , L_2 qui correspondent aux niveaux de la more, du pied et du mot prosodique. Chaque ligne métrique est calculée dans une strate différente, et les contraintes actives dans la strate n sont actives dans la strate $n + 1$, où elles peuvent être réordonnées. Je donne ici l'exemple de la structure métrique de *hebdomadaire* dans le français méridional. L'ajustement des voyelles moyennes se fait ici de manière très similaire à ce qu'avait

2. Je renvoie à Laks (1997) pour une critique détaillée de la constituance et une défense du modèle grille seule.

proposé Durand (1976), puisque les voyelles moyennes suivies d'un creux rythmique sont réalisées mi-ouvertes, alors que les autres sont réalisées mi-fermées. L'apparition d'un creux rythmique peut être due à une coda, comme dans la première syllabe, ou à la présence d'un schwa, ni l'un ni l'autre ne projetant de pic au niveau L_1 (voir Andreassen & Eychenne 2013 : 133–134). Ces généralisations sur le mot peuvent cependant être obscurcies lorsque le mot est intégré dans un domaine prosodique plus large (*i.e.* au niveau « post-lexical »).

L_2						×				
L_1	×		×	×	×	×				
L_0	×	×	×	×	×	×				
	ε	b	d	o	m	a	d	ε	r	\emptyset

Figure 4.3. Grille métrique au niveau lexical pour *hebdomadaire*

Considérons une forme telle que *des recherches généalogiques humoristiques* [de $\text{v}\emptyset.\text{'}\varepsilon\text{b}.\text{'}\emptyset // \text{'}\varepsilon\text{e}.\text{ne}.\text{'a}.\text{lo}.\text{'}\varepsilon\text{i}.\text{k}\emptyset \text{'y}:\text{m}o.\text{ris}.\text{'t}i.\text{k}\emptyset$] avec un accent focal initial dans *humoristiques* (11ald1, entretien guidé, corpus PFC) : l'accent est marqué par un ton haut (H), comme on peut le voir à la figure 4.4, et la voyelle est environ deux fois plus longue que la voyelle suivante (117 vs 56 ms). Commençons par examiner le patron accentuel du mot *généalogiques*. Nous posons l'existence de contraintes militent contre les collisions accentuelles paramétrées par ligne rythmique (schéma *CLASH(L_n)) et admettons que les contraintes de fidélité peuvent faire référence aux proéminences métriques ; les schémas MAX(L_n) et DEP(L_n) militent respectivement contre l'effacement et l'insertion d'un sommet au niveau n . Par ailleurs, notre approche suppose que chaque strate contient des contraintes de mise en relief (*enhancement*) des bords du domaine construit dans cette strate. Le schéma de contraintes est le suivant :

- (4) EDGE (L_n , E) : la position métrique gauche/droite sur la ligne L_n est un sommet rythmique.
 où $E \in \{L(\text{eft}), R(\text{ight})\}$

Etant donné cet ensemble de contraintes, la structure métrique de *généalogiques* peut être modélisée à la strate 2 comme il est montré au tableau 4.3, en supposant un input depuis la strate 1 où toutes les voyelles pleines projettent une proéminence en L_1 . La contrainte *CLASH(L_1) s'assure de ce que cette forme présente une alternance pic/creux.

L'accent focal dans *humoristiques* est quant à lui attribué au niveau phrastique (strate 3), comme illustré au tableau 4.4. En plus des contraintes sus-mentionnées, nous voyons à l'œuvre la contrainte de poids WEIGHTTOSTRESS, qui demande à ce qu'une voyelle accentuée soit lourde (bimorique), ainsi que la contrainte H/STRESS(L_3), qui demande à ce que la voyelle accentuée soit associée à un ton haut.

Nous avons également appliqué ce modèle à l'accentuation pénultième que l'on rencontre dans les variétés suisses : cet accent est lui aussi caractérisé par un ton haut, au

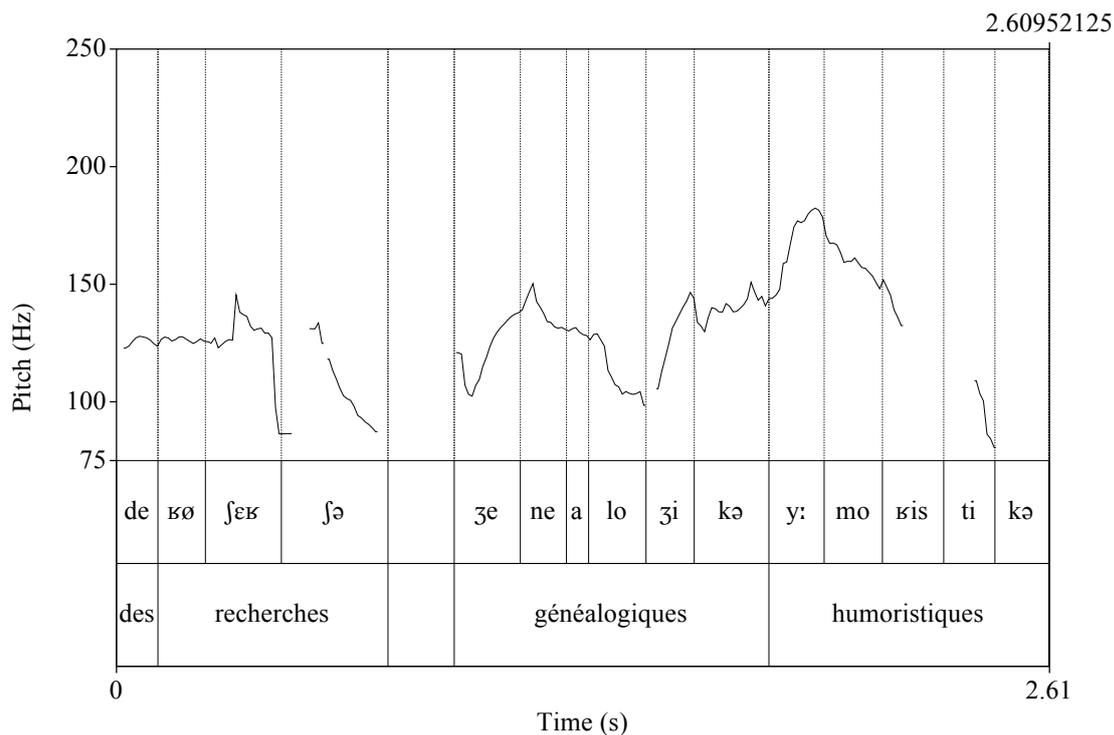


Figure 4.4. Contour de la F_0 pour *des recherches généalogiques humoristiques* (11ald1, entretien guidé)

niveau du syntagme intonatif, comme dans *Venir sonner **chez** lui ou l'arrêter dans **la** rue* (cf. figure 4.5).

Bien que l'accent puisse être réalisé sur un mot lexical comme dans *connu*, les deux exemples précédents montrent qu'il est tout à fait possible qu'il soit réalisé sur un clitique. Pour rendre compte de ce phénomène, nous avons intégré un schéma de contraintes (flottante, puisque le phénomène est optionnel) appelée $PENULTIMATESTRESS(L_n)$, qui comme son nom l'indique demande à ce que la pénultième porte une proéminence, ainsi qu'un schéma interdisant d'accentuer les clitiques ($*STRESSEDCLITIC(L_n)$). Cette contrainte peut devenir dominante dans la strate 3, comme illustré au tableau 4.5. Lorsqu'elle est dominée par $*CLASH(L_2)$ et/ou $*STRESSEDCLITIC(L_2)$, c'est le candidat (c) qui émerge comme candidat optimal. La présence de $*STRESSEDCLITIC(L_3)$ assure que le clitique ne peut pas recevoir d'accent principal au niveau du syntagme intonatif. Le lecteur se demandera peut-être pourquoi nous avons inclus le schéma $PENULTIMATESTRESS$ au lieu de $NONFINALITY$ qui est une contrainte bien établie dans la littérature. La raison en est que la proéminence antépénultième *coexiste* avec l'accent final, notamment avec des voyelles finales de syntagme intonatif allongées au niveau post-lexical, par exemple dans *décevante* [dəsvã:rt].

Le modèle stratifié présenté ci-dessus permet de résoudre l'une des difficultés majeures

	*CLASH(L ₁)	MAX(L ₁)	EDGE(L ₂ ,R)	EDGE(L ₂ ,L)
[☞] a. L ₂ : x L ₁ : x x x L ₀ : x x x x x zɛnealozikə		**	*	****
b. L ₂ : x L ₁ : x x x x L ₀ : x x x x x zɛnealozikə	*!***			****
c. L ₂ : x L ₁ : x x L ₀ : x x x x x zɛnealozikə		***!	**	***

Tableau 4.3. Calcul de l’accent pour *généalogiques* (Strate 2)(Andreassen & Eychenne 2013 : 135)

	EDGE(L ₃ ,R)	WEIGHT ToSTRESS	H/STRESS(L ₃)	DEP(L ₂)
[☞] a. L ₃ : x L ₂ : x x L ₁ : x x x x L ₀ : xx x x x x y:moristikə H				*
b. L ₃ : x L ₂ : x x L ₁ : x x x x L ₀ : x x x x x ymoristikə		*!	*	*
c. L ₃ : x L ₂ : x L ₁ : x x x x L ₀ : x x x x x ymoristikə	*!		*	

Tableau 4.4. Accent phrastique dans *humoristiques* (Strate 3)(Andreassen & Eychenne 2013 : 136)

auxquelles je m’étais heurté dans ma thèse, à savoir le fait que les généralisations prosodiques à différents niveaux ne sont pas toujours congruentes, et qu’un niveau supérieur peut procéder à des réajustements rythmiques aux niveaux inférieurs. Il est alors nécessaire de recourir à des dispositifs théoriques tels que la non-prosodification de syllabes ou le super-pied, une structure intermédiaire entre le pied et le mot prosodique (Selkirk 1980). Le rejet de la constituance au profit d’une expression plus directe des relations de proéminence permet donc d’éliminer ce type de problèmes. Il offre du même coup une solution simple au problème de la pédification de formes telles que *expérimentale* [ɛks.pe.ri.mã.'ta.lə] (cf 3 ci-dessus) : les formes méridionales et septentrionales sont alors rigoureusement identiques, à ceci près que le français méridional présente une syllabe finale additionnelle non accentuée.

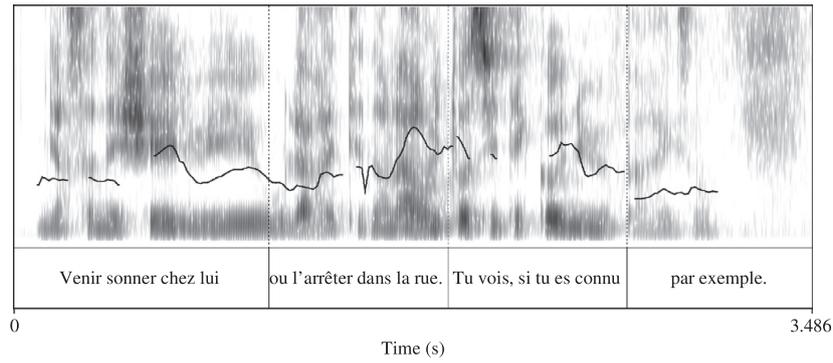


Figure 4.5. Spectrogrammes de montées pénultièmes dans *chez lui – la rue – connu* (svaje1g, Nyon)(Andreassen & Eychenne 2013 : 137)

	*STRESSED CLITIC(L ₃)	PENULTIMATE STRESS(L ₂)	*CLASH(L ₂)	*STRESSED CLITIC(L ₂)	H/STRESS ^a
a. L ₃ : × L ₂ : × × L ₁ : × × L ₀ : × × fe lqi	*!		*	*	**
^{or} b. L ₃ : × L ₂ : × × L ₁ : × × L ₀ : × × fe lqi H			*	*	*
c. L ₃ : × L ₂ : × × L ₁ : × × L ₀ : × × fe lqi		*!			**

Tableau 4.5. Accent pénultième dans *chez lui* (Strate 3)(Andreassen & Eychenne 2013 : 137)

4.3 Schwa opaque en français du midi

La loi de position en français méridional, dont la définition est rappelée en (5), est l'un des problèmes phonologiques auxquels j'ai consacré une attention toute particulière.

(5) Loi de position

Une voyelle moyenne est :

- a. mi-fermée en syllabe ouverte ;
- b. mi-ouverte en syllabe fermée ou en syllabe fermée suivie d'une syllabe dont le noyau est un schwa.

En français méridional commun, l'application de la loi de position est (presque) sans exception : comme je l'ai montré de manière détaillée dans ma thèse, ce patron est pleinement productif et se retrouve notamment dans les emprunts (ex : *jogging* [dʒogɪŋ]) et

dans la morphologie non concaténative (*fofolle* [fɔfɔlə]). Il en résulte que les paires de voyelles [e ε], [ø œ] et [o ɔ] ont une distribution allophonique :

(6) Illustration de la loi de position :

- | | | | | | | |
|----|--------------|----------|--------------|----------|-------------|----------|
| a. | <i>sait</i> | [se] | <i>ceux</i> | [sø] | <i>seau</i> | [so] |
| b. | <i>sel</i> | [sɛl] | <i>seul</i> | [sœl] | <i>sol</i> | [sɔl] |
| c. | <i>selle</i> | ['sɛ.lə] | <i>seule</i> | ['sœ.lə] | <i>sole</i> | ['sɔ.lə] |

Ces paires représentent trois phonèmes que l'on notera /E œ O/. Afin de guider le reste de la discussion, je donne en (7) les critères définitoires du schwa en français méridional commun. On voit que, contrairement au schwa septentrional, l'alternance voyelle/zéro est limitée au contexte prévoicalique ; par ailleurs, schwa interagit de manière systématique avec la loi de position, comme on l'a noté dans les exemples ci-dessus. Ces critères sont des conditions suffisantes, mais non nécessaires, pour identifier un schwa. Par exemple, la condition d'abaissement ne s'applique pas dans les mots où la voyelle précédente n'est pas moyenne, mais d'autres critères, notamment la condition d'accentuation, peuvent s'appliquer par ailleurs (ex : *patte* ['patə]).

(7) Propriétés du schwa en français méridional

- a. *condition d'effacement* : schwa est effacé devant voyelle (ex : *bulle éclatée* [byleklate]).
- b. *condition d'accentuation* : schwa n'est pas accentué (ex : *mate* ['matə]).
- c. *condition réalisation* : la qualité phonétique du schwa est variable (ex : *patte* [patə], [patœ], [patɐ] selon la variété et/ou le locuteur).
- d. *condition d'abaissement* : un schwa abaisse la voyelle précédente s'il s'agit d'une voyelle moyenne (ex : *nette* [netə]).

(adapté d'après Durand 1995 : 40–42)

La loi de position soulève de nombreuses questions empiriques et théoriques. La première, et celle qui a le plus retenu l'attention des chercheurs, est la caractérisation du contexte disjonctif (2b). J'ai déjà abordé cette question à la section précédente, à travers l'approche grille seule que j'ai développée avec Helene Andreassen. Parmi les autres problèmes importants que soulève la loi de position, deux ont retenu mon attention : la question des exceptions, que j'ai abordée en détail dans Eychenne (2014)^[6], et la question de l'opacité (Kiparsky 1971, 1973), que j'ai également abordée dans Eychenne (2014)^[6] mais dont le résultat le plus abouti se trouve dans Eychenne (2015b)^[9]. Je serai amené à rediscuter la question des exceptions dans le chapitre ?? ; aussi, je me concentrerai ici sur la question de l'opacité, en me basant pour l'essentiel sur Eychenne (2015b)^[9].

4.3.1 Contre-blocage et/ou contre-alimentation ?

On admettra tout d’abord qu’il est correct de considérer le schwa comme une voyelle dégénérée au niveau sous-jacent en français méridional, comme le font notamment Durand (1976, 1995) et Moreux (1985a) (voir Watbled 1995 pour une position alternative et ma critique détaillée dans Eychenne 2014 : §3.2^[6]). Ceci étant posé, le problème que soulève le schwa est le suivant : bien qu’il puisse être réalisé comme une voyelle centrale en finale de mot (*bête* [betə]), il est très souvent réalisé comme une voyelle moyenne antérieure arrondie en position non finale (*bêtement* [betøma^N]) (voir Rochet 1980 ; Durand 1995). Il existe donc trois processus qui interagissent : un processus de pédification, un processus d’ajustement des voyelles moyennes qui dépend du premier, et un processus de « coloration » du schwa. Comme on va le voir, ces processus interagissent de manière OPAQUE. Dans ma thèse, j’avais analysé ce phénomène comme un cas de CONTRE-BLOQUAGE (*counterbleeding*). Pour le comprendre, partons d’une approche dérivationnelle à la Selkirk (1978) : cette auteure postule deux règles essentielles pour la construction des pieds : une règle appelée PIED-SIMPLE, selon laquelle toutes les voyelles (y compris le schwa) projettent un pied, et une règle subséquente appelée PIED-DÉRIVÉ, qui fusionne un pied dont le noyau est un schwa avec le pied précédent pour former un trochée (voir figure 4.6). Nous admettrons qu’il existe par ailleurs pour le français méridional commun un processus d’ajustement des voyelles moyennes, que l’on notera informellement AJUST-VM et qui regroupe les conditions a et b de (2), et une règle de coloration du schwa (COLOR-SCHWA).

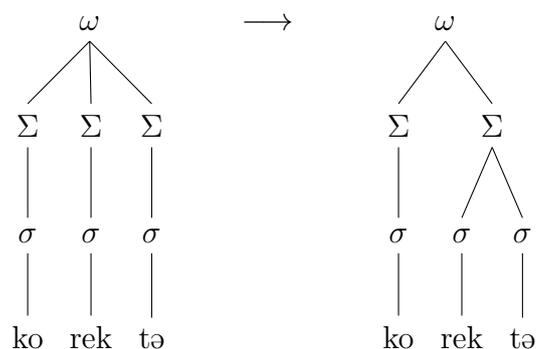


Figure 4.6. Construction du pied dérivé selon Selkirk (1978)

Observons la dérivation suivante (de type SPE), où RSJ et RS représentent respectivement la représentation sous-jacente et la représentation de surface :

(8) Dérivation de *bêtement* avec schwa opaque :

RSJ	/##bEtə+ma ^N ##/
SYLLABIFICATION	⟨bE⟩ _σ ⟨tə⟩ _σ ⟨ma ^N ⟩ _σ
PIED-SIMPLE	⟨⟨bE⟩ _σ ⟩ _Σ ⟨⟨tə⟩ _σ ⟩ _Σ ⟨⟨ma ^N ⟩ _σ ⟩ _Σ
PIED-DÉRIVÉ	⟨⟨bE⟩ _σ ⟨tə⟩ _σ ⟩ _Σ ⟨⟨ma ^N ⟩ _σ ⟩ _Σ
COLOR-SCHWA	⟨⟨bE⟩ _σ ⟨tɛ⟩ _σ ⟩ _Σ ⟨⟨ma ^N ⟩ _σ ⟩ _Σ
AJUST-VM	⟨⟨bɛ⟩ _σ ⟨tə⟩ _σ ⟩ _Σ ⟨⟨ma ^N ⟩ _σ ⟩ _Σ
RS	[betəma ^N]

L'ordre des règles est ici crucial : observons donc ce qui se passe dans la dérivation contre-factuelle dans laquelle la coloration du schwa précède la construction de pieds dérivés :

(9) Dérivation contre-factuelle *bêtement* :

RSJ	/##bEtə+ma ^N ##/
SYLLABIFICATION	⟨bE⟩ _σ ⟨tə⟩ _σ ⟨ma ^N ⟩ _σ
PIED-SIMPLE	⟨⟨bE⟩ _σ ⟩ _Σ ⟨⟨tə⟩ _σ ⟩ _Σ ⟨⟨ma ^N ⟩ _σ ⟩ _Σ
COLOR-SCHWA	⟨⟨bE⟩ _σ ⟩ _Σ ⟨⟨tɛ⟩ _σ ⟩ _Σ ⟨⟨ma ^N ⟩ _σ ⟩ _Σ
PIED-DÉRIVÉ	_____
AJUST-VM	⟨⟨be⟩ _σ ⟩ _Σ ⟨⟨tə⟩ _σ ⟩ _Σ ⟨⟨ma ^N ⟩ _σ ⟩ _Σ
RS	*[betəma ^N]

Nous voyons que l'application de COLORATION *avant* PIED-DÉRIVÉ empêche cette dernière de s'appliquer puisque son contexte d'application, qui dépend crucialement de la présence d'un schwa, a été détruit. Dans cette dérivation contre-factuelle, COLOR-SCHWA bloque PIED-DÉRIVÉ : il en résulte que COLOR-SCHWA contre-bloque PIED-DÉRIVÉ en (8). Dans la terminologie introduite par McCarthy (1999), on dira qu'il y a eu SUR-APPLICATION de PIED-DÉRIVÉ puisqu'un trochée ne peut être construit que s'il y a un schwa, et ce schwa n'est pas manifeste en surface. Ce type d'opacité pose un problème majeur pour une théorie non dérivationnelle comme la théorie de l'optimalité standard, puisqu'il est nécessaire de faire référence à une représentation intermédiaire entre l'input et l'output. Ce problème était l'une des raisons principales qui m'avaient poussé à adopter des représentations enrichies dans ma thèse (Eychenne 2006 : §6.3.1.3), un point sur lequel je reviendrai à la fin de cette section.

On peut néanmoins envisager le problème de manière légèrement différente. En effet, l'analyse à la Selkirk présentée ci-dessus suppose la présence d'opérations destructrices : on construit d'abord un pied dont la tête est un schwa, que l'on détruit ensuite pour rattacher sa syllabe au pied précédent, comme on l'a vu à la figure 4.6. Une analyse alternative, dans l'esprit de ce qu'a proposé Durand (1986) en phonologie de dépendance, serait de construire le pied trochaïque au moyen d'opérations non destructrices. Modifions donc légèrement la règle PIED-SIMPLE (et appelons-la PIED-SIMPLE'), de sorte qu'elle ne projette un pied que si la voyelle est une voyelle pleine ; modifions de même PIED-DÉRIVÉ (que l'on appellera PIED-DÉRIVÉ'), afin qu'elle intègre toute syllabe non intégrée dans le

ped précédent. Dans ce cas, la structure prosodique est construite de manière monotone. La dérivation correspondante est donnée en (10).

(10) Dérivation monotone de *bêtement* :

RSJ	/##bEtə+ma ^N ##/
SYLLABIFICATION	⟨bE⟩ _σ ⟨tə⟩ _σ ⟨ma ^N ⟩ _σ
PIED-SIMPLE'	⟨⟨bE⟩ _σ ⟩ _Σ ⟨tə⟩ _σ ⟨⟨ma ^N ⟩ _σ ⟩ _Σ
PIED-DÉRIVÉ'	⟨⟨bE⟩ _σ ⟨tə⟩ _σ ⟩ _Σ ⟨⟨ma ^N ⟩ _σ ⟩ _Σ
COLOR-SCHWA	⟨⟨bɛ⟩ _σ ⟨tɛ⟩ _σ ⟩ _Σ ⟨⟨ma ^N ⟩ _σ ⟩ _Σ
AJUST-VM	⟨⟨bɛ⟩ _σ ⟨tø⟩ _σ ⟩ _Σ ⟨⟨ma ^N ⟩ _σ ⟩ _Σ
RS	[betøma ^N]

Observons maintenant la dérivation contre-factuelle dans laquelle COLOR-SCHWA précède PIED-SIMPLE :

(11) Dérivation monotone contre-factuelle de *bêtement* :

RSJ	/##bEtə+ma ^N ##/
SYLLABIFICATION	⟨bE⟩ _σ ⟨tə⟩ _σ ⟨ma ^N ⟩ _σ
COLOR-SCHWA	⟨bE⟩ _σ ⟨tɛ⟩ _σ ⟨ma ^N ⟩ _σ
PIED-SIMPLE'	⟨⟨bE⟩ _σ ⟩ _Σ ⟨⟨tɛ⟩ _σ ⟩ _Σ ⟨⟨ma ^N ⟩ _σ ⟩ _Σ
PIED-DÉRIVÉ'	_____
AJUST-VM	⟨⟨be⟩ _σ ⟩ _Σ ⟨⟨tø⟩ _σ ⟩ _Σ ⟨⟨ma ^N ⟩ _σ ⟩ _Σ
RS	*[betøma ^N]

Nous voyons que, comme en (9) ci-dessus, PIED-DÉRIVÉ' est bloquée et nous avons à nouveau une situation de contre-bloquage. Mais il y a dans cette nouvelle dérivation une interaction supplémentaire : en effet, COLOR-SCHWA alimente PIED-SIMPLE' puisqu'elle lui fournit un contexte d'application supplémentaire (le pied dont la tête était un schwa avant l'application de COLOR-SCHWA). Il en résulte que dans la dérivation (10), COLOR-SCHWA contre-alimente (*counterfeeds*) PIED-SIMPLE'. Dans la terminologie de McCarthy (1999), il y a SOUS-APPLICATION de PIED-SIMPLE', puisque l'une des voyelles pleines (le schwa coloré) ne projette pas son propre pied. Dans le même temps, il y a sur-application de PIED-DÉRIVÉ' puisque le noyau de la syllabe dépendante n'est pas un schwa.

Il ressort de cette discussion deux points importants : le premier, qui n'est pas une surprise, est que la nature de l'opacité en jeu dépend de la manière dont le phénomène est analysé (Turcsan 2005). Si l'on considère qu'une analyse monotone du phénomène est préférable, on est par ailleurs amené à considérer que la loi de position est à la fois un cas de contre-blocage et de contre-alimentation. La raison en est que l'une des généralisations en jeu implique la construction de structure prosodique, ce qui dans une théorie dérivationnelle nécessite plusieurs règles (ici, deux règles). Comme l'a souligné à juste titre Baković (2011 : 42), « most if not all definitions of pairwise ordered rule relations provided in textbooks and in the scholarly literature are insufficiently precise

about situations involving more than two rules ». La loi de position présente un tel cas où la même règle peut interagir avec deux autres règles. Ces résultats viennent s'ajouter aux travaux récents (notamment Baković 2007, 2011) qui montrent que la conception héritée des travaux de Kiparsky dans les années 70 ne couvre pas l'ensemble des interactions possibles.

4.3.2 Taxonomie du schwa opaque

Ayant discuté le caractère opaque du schwa, il est intéressant d'analyser le phénomène pour un apprenant qui acquiert ce système phonologique. J'ai donc proposé une taxonomie destinée à identifier les cas de figure auxquels l'apprenant est confronté. Cette taxonomie distingue cinq catégories.

schwa morphologiquement récupérable Ces schwas apparaissent entre une base et un suffixe, et leur existence peut être déduite sans ambiguïté de la structure morphologique du mot. Prenons par exemple le mot *utilement* réalisé [ytiløma^N]. La voyelle [ø] peut être mise en correspondance systématique avec le schwa final de la base *utile* [y'tilø] (cf Rochet 1980 : 92 ; Durand et al. 1987 : 992), qui satisfait la condition d'accentuation puisqu'il est inaccentué.

Ce type de schwa est présent dans la morphologie flexionnelle, et il apparaît notamment dans la conjugaison (*chante* [ʃa^Ntə] vs *chanterai* [ʃa^Ntøre]) et comme marqueur du féminin (*net* [nɛt] vs *nette* [nɛtə]).

schwa phonologiquement récupérable Un schwa peut être récupéré sur des bases purement phonologiques s'il est précédé d'une voyelle moyenne. Prenons l'exemple du mot monomorphémique *céleri*³. Durand, Slater & Wise (1987) soulignent qu'il existe deux prononciations, l'une avec un [e] mi-fermé ([seløri]), et l'autre avec un [ɛ] mi-ouvert ([sɛløri]). En vertu de la condition d'abaissement (7d), ces auteurs soutiennent que la voyelle médiale est un schwa lorsque la voyelle précédente est mi-ouverte (soit la forme sous-jacente /sEløri/) alors qu'il s'agit d'un /ɛ/ stable si la voyelle est mi-fermée (soit /sElɛri/).

Il faut noter qu'un schwa est souvent récupérable par la morphologie et par la phonologie. C'est par exemple le cas du [ø] de *netteté* [nɛtøte] : la morphologie permet d'identifier un schwa marqueur du féminin dans *nette*, et l'abaissement de la voyelle moyenne précédente indique là encore, sur la base de critères strictement phonologiques, qu'il s'agit d'un schwa.

schwa putatif non récupérable La situation se complique dès lors que le schwa n'est récupérable ni phonologiquement, ni morphologiquement. Considérons la voyelle orthographiée *e* dans les mots monomorphémiques *hameçon* [amøso^N] et *allemand* [aløma^N]. La voyelle est *a priori* réalisée de manière identique au /ɛ/ stable orthographié *eu* dans

3. Il existe d'autres mots qui présentent les mêmes propriétés, par exemple *écrevisse*.

pharmaceutique [farmasøtikə]. S'il n'y a pas de différence entre les deux types de voyelles, on peut supposer que la voyelle de *hameçon* et *allemand* a été réanalysée comme un / E / stable, comme l'on suggéré Durand et al. (1987). Or tout l'enjeu est ici est de déterminer que les deux voyelles sont bel et bien identiques, autrement dit de vérifier si la réalisation du schwa est bel et bien opaque : il se peut en effet qu'il subsiste des différences phonétiques fines entre les deux voyelles, puisqu'on attend d'un schwa qu'il forme un trochée avec la voyelle précédente alors qu'une voyelle pleine devrait projeter son propre pied. Autrement dit, il est possible qu'un schwa *puisse* être réalisé comme un [ø] (mais il pourrait aussi être réalisé comme une voyelle centrale), alors qu'un / E / stable *doit* être réalisé [ø] en syllabe ouverte. C'est ce que j'ai appelé l'HYPOTHÈSE D'ASYMÉTRIE dans Eychenne (2014 : 230) :

(12) Hypothèse d'asymétrie

/ə/ peut être réalisé comme une voyelle moyenne antérieure arrondie
mais / E / ne peut pas être réalisé comme une voyelle centrale [ə].

Si cette hypothèse est vérifiée en position non finale, il y aurait là un argument de poids à l'appui de la présence d'un schwa dans *hameçon* ou *allemand* ; dans le cas contraire, l'hypothèse la plus probable est que cette voyelle a été réanalysée comme une voyelle stable. Bien entendu, il est possible (et même très probable) que les deux systèmes existent, et que certains locuteurs aient une voyelle stable alors que d'autres auraient un schwa. Dans tous les cas, cette question ne peut être tranchée sans recourir à une analyse précise de ces contextes, et l'étude de cas que je rapporte plus bas est à ma connaissance la seule qui ait étudié cette question d'un point de vue phonétique.

schwa putatif alternant Une autre classe de schwas potentiels provient des alternances avec [ɛ] du type *amener* [amøne] vs *amène* [amɛnə]. C'est ce qui correspond à l'AJUSTEMENT EN SYLLABE FERMÉE du français septentrional (Dell 1973 ; Montreuil 2002). Il est difficile de trancher cette question de manière catégorique puisqu'elle engage l'interface phonologie/morphologie et qu'elle recoupe en partie les autres catégories mentionnées précédemment.

Il y a toutefois deux éléments importants qui montrent que cette question ne peut être tranchée de manière globale, et qu'elle doit être traitée au cas par cas. Tout d'abord, comme je l'ai observé (Eychenne 2006 : 219), certains locuteurs qui ont une application très générale de la loi de position peuvent présenter dans ce contexte des patrons qui ne sont pas systématiques, par exemple *hôtel* ~ *hôtelier* [ɔtølje], qui suggère la présence d'un schwa, mais *vaisselle* ~ *vaisselier* [vesølje], auquel cas la voyelle n'abaisse pas le / E / précédent et est donc une voyelle stable. Le deuxième élément est le fait que, comme l'a très bien montré Morin (1988), les alternances schwa/ɛ ne peuvent pas être considérées indépendamment des autres alternances du type *venir* ~ *viennne*, *devoir* ~ *doive*, etc. L'alternance schwa/ɛ (ou $\text{E}/\text{ɛ}$) est donc un problème morphologique et non pas phonologique.

/œ/ **stable** Les voyelles stables, enfin, sont les voyelles telles que le *eu* de *pharmaceutique* discuté plus haut et pour lesquelles il n'existe aucune raison de postuler l'existence d'un schwa.

4.3.3 Analyse d'un idiolecte

Afin de tenter de mieux comprendre la nature du schwa opaque, j'ai décidé d'entreprendre une étude phonétique détaillée de cette question. Bien que j'aie initialement envisagé de me baser sur les données disponibles dans la base PFC, j'ai dû y renoncer pour deux raisons : tout d'abord, la qualité acoustique des fichiers permettait difficilement de mesurer les schwas de manière fiable, surtout en position finale où ils atones, sont parfois dévoisés, et interagissent de manière complexe avec le *eah* d'hésitation ; en outre, de par la nature même du corpus PFC, certains types de schwas tels que les schwas putatifs non récupérables étaient sous-représentés. J'ai donc décidé d'étudier un seul locuteur de manière détaillée.

4.3.3.1 Méthode

J'ai choisi pour cette étude un locuteur typique du français méridional commun vivant à Figeac (dans le Lot), à deux heures environ de Toulouse. Ce locuteur, âgé de 64 ans au moment de l'enregistrement, est né et a vécu jusqu'à environ 20 ans en Midi-Pyrénées, a ensuite vécu à près de Marseille (à Martigues, puis Istres) pendant 25 ans, après quoi il est retourné en Midi-Pyrénées à la frontière Lot/Aveyron où il vit depuis 1995. Les données ont été enregistrées chez le locuteur avec un enregistreur PCM Olympus LS-5 maintenu à distance constante de la bouche (environ 40 cm). Le matériel expérimental était une liste de 260 mots isolés (contenant des distracteurs), qui a été répétée trois fois par le sujet (avec une courte pause entre chaque répétition).

Les données ont été segmentées et analysées à l'aide de Praat. J'ai procédé à un premier alignement grossier à l'aide d'EasyAlign (Goldman 2011), qui a été corrigé manuellement en tenant compte des indices acoustiques présents dans l'onde sonore et dans un spectrogramme à bande large. J'ai ensuite extrait 4 jeux de données :

- I Toutes les voyelles accentuées
- II Tous les schwas finaux avec la voyelle tonique précédente
- III Les voyelles correspondant à *e* ou *eu* en position initiale
- IV Les voyelles correspondant à *e* ou *eu* en position médiale

J'ai d'abord extrait les 4 premiers formants, échantionnés à 5 ms d'intervalle, à l'aide d'une analyse LPC (de 0 à 5000 Hz, 12 pôles). Afin de limiter le nombre d'erreurs de détection, j'ai appliqué conjointement deux méthodes : les trois premiers formants ont tout d'abord été lissés en utilisant une moyenne mobile à 3 points ; ces trois formants ont ensuite été mesurés au milieu de la voyelle en utilisant une interpolation linéaire basée sur les formants lissés. J'ai par la suite appliqué un filtre, basé sur le travail de Gendrot & Adda-Decker (2005), visant à exclure toute voyelle dont l'un des formants aurait une

valeur aberrante⁴. Le corpus final comprend, pour chaque jeu de données :

I 702 voyelles orales accentuées (9 ont été rejetées par le filtre)

II 330 paires voyelle pleine/schwa : 13 paires ont été inspectées manuellement, 9 ont été corrigées, les autres ont été rejetées

III 66 voyelles (aucune n'a été rejetée)

IV 228 paires voyelle pleine/schwa (ou voyelle pleine/*eu*) : une seule paire a été écartée

Les 3 premiers formants, mesurés en Hertz (Hz), ont été convertis sur l'échelle perceptive ERB. J'ai par ailleurs extrait la F0 (qui a été convertie en demi-tons, DT), l'intensité en décibels (dB) et la durée en millisecondes (ms) de la voyelle. L'analyse statistique a été effectuée dans l'environnement MATLAB 2014b, édité par MathWorks.

4.3.3.2 Principaux résultats

Je ne discuterai ici que les résultats les plus saillants, et renvoie le lecteur à Eychenne (2015b : §3.2) pour une présentation détaillée. Il ressort tout d'abord de cette analyse que le schwa final non accentué est le plus proche de [ø] accentué chez ce locuteur.

Les modèles statistiques que j'ai construits visaient à déterminer si les différents paramètres acoustiques mesurés (convertis sur une échelle perceptive) permettaient de prédire une différence entre deux types de voyelles : les /E/ stables (orthographiés *eu*) et les schwas (orthographiés *e*). L'analyse statistique des résultats a donc été menée à l'aide de MODÈLES LOGISTIQUES À EFFETS MIXTES, qui sont à la fois une extension de la méthode de régression classique qui prend en compte, en plus des effets fixes, des effets aléatoires, et à la fois une extension de la régression logistique qui permet de modéliser une variable de réponse binaire.

Le premier modèle, basé sur le jeu de données I, visait à déterminer s'il existait une différence entre les voyelles accentuées et le schwa. Les moyennes et écarts-types (entre parenthèses) sont donnés au tableau 4.6⁵. J'ai donc modélisé l'ACCENT comme variable de réponse, la DURÉE, l'INTENSITÉ et la HAUTEUR comme effets fixes, et le MOT comme effet aléatoire. Durée ($p < 0,001$), intensité ($p < 0,01$) et hauteur ($p < 0,001$) étaient tous les trois significatifs, la voyelle accentuée étant plus longue, plus intense et plus haute que le schwa.

durée		intensité		hauteur	
V	ə	V	ə	V	ə
145 ms	77 ms	75,5 dB	68,2 dB	141 Hz	110 Hz
(35)	(18)	(2,2)	(3,9)	(6)	(8)

Tableau 4.6. Voyelle accentuée vs schwa en trochée final

4. Notons que ce filtre est assez permissif, et il n'est donc pas exclu qu'il ait subsisté quelques erreurs de mesure, comme dans tout travail de ce type.

5. Bien que les modèles statistiques aient été bâtis en utilisant des valeurs converties sur une échelle perceptive (demi-tons ou ERB), je donne les valeurs en Hertz dans les tableaux pour plus de lisibilité.

Le deuxième modèle visait à vérifier s'il existait une différence phonétique entre les *e* (ex : *brevet*) et les *eu* (ex : *breuvage*) en position initiale de monosyllabe, puisqu'il est généralement admis que ces voyelles sont phonétiquement identiques (Durand et al. 1987). Le modèle tentait cette fois de prédire la CATÉGORIE de la voyelle (/ɛ/ vs /ə/) comme variable de réponse, en utilisant la DURÉE, l'INTENSITÉ, la HAUTEUR et les trois premiers formants comme effets fixes, avec toujours le MOT comme effet aléatoire. Les résultats sont présentés au tableau 4.7. Comme on peut le voir, la différence entre les moyennes est toujours inférieure à l'écart type de l'une ou l'autre voyelle, ce qui suggère qu'il n'y a pas de différence. C'est ce que confirme l'analyse statistique, puisqu'aucun des paramètres acoustiques n'était significatif.

	durée	intensité	F0	F1	F2	F3
<i>eu</i>	93 ms (21)	80,1 dB (1,0)	138 Hz (4)	416 Hz (17)	1359 Hz (98)	2468 Hz (111)
<i>e</i>	82 ms (21)	79,7 dB (2,1)	140 Hz (5)	424 Hz (23)	1339 Hz (117)	2450 Hz (140)

Tableau 4.7. *e* vs *eu* en position initiale

Le troisième et dernier modèle que je discuterai est le modèle pour la position médiale. J'ai regroupé les 5 catégories de la taxonomie présentée en §4.3.2 en 3 catégories : les schwas récupérables, les schwas putatifs et les /ɛ/ stables. Les paramètres acoustiques testés étaient les mêmes que pour la position initiale, et les résultats sont consignés au tableau 4.8. Puisqu'il y avait cette fois 3 catégories, j'ai construit un modèle multinomial à effets mixtes, avec la CATÉGORIE comme variable de réponse, les mêmes effets fixes que pour la position initiale et le MOT comme effet aléatoire. Comme pour la position initiale, aucun des paramètres acoustiques ne s'est avéré significatif.

	durée	intensité	F0	F1	F2	F3
/ə/	68 ms (15)	79,8 dB (1,4)	141 Hz (5)	427 Hz (21)	1438 Hz (82)	2470 Hz (175)
/ə/ putatif	69 ms (13)	79,9 dB (1,4)	142 Hz (3)	430 Hz (22)	1443 Hz (91)	2535 Hz (163)
/ɛ/ stable	74 ms (19)	80,3 dB (1,4)	141 Hz (5)	411 Hz (17)	1483 Hz (61)	2396 Hz (226)

Tableau 4.8. *e* vs *eu* en position médiale

Je résume brièvement les résultats obtenus pour ce locuteur :

- le schwa final est une voyelle atone phonétiquement proche de [ø] accentué
- il n'y a pas de différence phonétique entre *e* et *eu* en position initiale de polysyllabe
- il n'y a pas de différence phonétique entre schwas récupérables, schwas putatifs et /ɛ/ stables en position médiale

4.3.3.3 Interprétation

Eu égard à ces résultats, il apparaît que, pour ce locuteur, l’hypothèse d’asymétrie est falsifiée : la réalisation du schwa en position non finale est bel et bien opaque. Ceci a donc des répercussions pour l’apprenabilité de l’opposition schwa vs / œ /, et il est donc légitime de considérer que, dans les contextes où aucun indice phonologique ou morphologique ne permet d’identifier un schwa, les locuteurs ont réanalysé les voyelles de façon transparente et ont maintenant un / œ / stable. Ainsi, les mots *brevet* ou *hameçon* auraient respectivement les formes sous-jacentes /bR œ vE/ et /am œ sO^N/, en accord avec l’analyse de Durand et al. (1987).

Le fait que le schwa soit réalisé de manière opaque a par ailleurs des conséquences importantes pour la théorie phonologique. Comme je l’ai mentionné plus haut, la théorie de l’optimalité standard n’est pas capable de traiter ce type de phénomènes. Pour sortir de cette impasse sans renoncer à un cadre non dérivationnel, j’avais proposé dans ma thèse d’enrichir les représentations phonologiques à l’aide de la théorie de la TURBIDITÉ (Goldrick 2001 ; van Oostendorp 2008), qui propose de découpler les relations d’association en une relation de *projection* \uparrow et une relation de *licencement* \downarrow (aussi appelée *prononciation*). Un mot comme *bêtement* [b œ t œ ma^N] reçoit dans cette approche la représentation donnée à la figure 4.7 : les traits associés au schwa sont projetés mais ne sont pas pas licenciés.

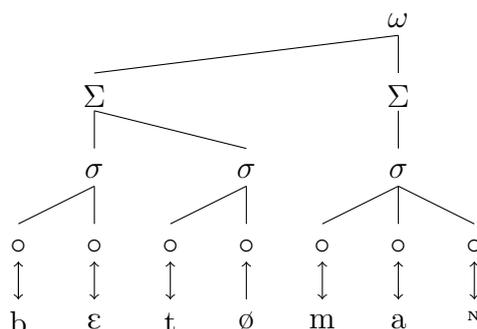


Figure 4.7. Représentation turbide de *bêtement*

J’avais utilisé ces relations turbides pour deux classes de phénomènes différents : la première était précisément ce phénomène d’opacité, et s’appuyait plus directement sur la proposition de Goldrick (2001) ; la seconde, qui s’inspirait de van Oostendorp (2008), était l’ensemble des phénomènes gradients tels que le dévoisement partiel ou le marquage articulaire. Puisque ces représentations turbides prédisaient une différence de réalisation dans le cas du dévoisement partiel ([d] \neq [d̥] \neq [t]), l’analyse prédisait également, en toute rigueur, une différence phonétique entre un / œ / stable et un schwa coloré en [ø]. Or nous voyons que cette approche n’est pas tenable pour l’idiolecte que j’ai décrit ici. L’alternative la plus viable à mes yeux est l’adoption d’un modèle stratifié, comme je l’ai défendu dans Andreassen & Eychenne (2013) et dans Eychenne (2014).

4.4 Conclusion

J'ai discuté dans ce chapitre un sous-ensemble de mes recherches qui sont représentatives du travail de modélisation symbolique que j'ai entrepris au cours des dix dernières années. Ce travail, tout en continuant à se baser sur des corpus, s'est efforcé d'intégrer des méthodes statistiques modernes afin de pouvoir tester des hypothèses précises sur les données que j'ai analysées. Si cette intégration s'est d'abord faite de façon balbutiante dans mon travail sur la liaison (Eychemme 2011), elle me semble avoir atteint aujourd'hui une certaine maturité dans l'analyse que j'ai proposée du schwa opaque : je n'ai connaissance d'aucune autre étude de ce phénomène qui ait intégré analyse formelle et analyse phonétique de telle manière⁶. Ce travail sur l'opacité et le travail que j'ai mené en collaboration avec Helene Andreassen me semblent représentatifs du type de phonologie auquel j'aspire dorénavant, à savoir une phonologie solidement ancrée dans l'analyse phonétique et quantitative des données, mais qui reste informée par, et tente de contribuer à, la théorie phonologique.

Ceci étant dit, comme on le verra au prochain chapitre, mon travail a maintenant pris un tournant plus quantitatif, y compris du point de vue de la modélisation. L'approche purement symbolique que j'avais adoptée pour des problèmes tels que la gradience, notamment l'utilisation de représentations « turbides » que j'ai évoquée plus haut, me paraît aujourd'hui, avec le recul, inadéquate. Je reste néanmoins convaincu que la modélisation symbolique est un outil utile car elle permet de formuler des hypothèses précises, et idéalement falsifiables.

6. Il faut néanmoins mentionner l'excellent travail de Moreux (1985a,b), mais il n'avait pas à sa disposition les outils statistiques et les instruments d'analyse phonétique dont on dispose aujourd'hui.

Chapitre 5

Modélisation quantitative

Ce chapitre vise à donner un aperçu du travail de modélisation quantitative que j'ai entrepris au cours des cinq dernières années. Bien que j'aie commencé à intégrer des méthodes quantitatives du point de vue de l'analyse des données depuis un certain temps, ce n'est que plus récemment que j'ai cherché à explorer des modèles quantitatifs de la *grammaire*. Cette transition a été dans un premier temps difficile, puisqu'elle a radicalement changé ma perspective sur les phénomènes langagiers. Ce changement de perspective était néanmoins inéluctable, et il s'est avéré salutaire : tout linguiste qui prend au sérieux le constat que les phénomènes grammaticaux sont gradients, et qui en tire pleinement les conséquences, sera inévitablement amené à considérer qu'il est nécessaire de recourir aux mathématiques du continu pour traiter de tels phénomènes de manière adéquate.

Je présenterai dans cette section trois modèles avec lesquels j'ai travaillé, chacun visant à traiter un domaine particulier de la phonologie, à savoir la structure suprasegmentale, la phonotaxe et la distinctivité. Le dernier modèle est un modèle original que j'ai développé en collaboration avec Tae Yeoub Jang à HUFS, et je le considère comme mon travail le plus abouti sur le plan théorique.

5.1 Structure suprasegmentale

La porte d'entrée vers le monde de la modélisation quantitative a été pour moi la lecture de l'article intitulé « Harmonic Phonology » (Goldsmith 1993a) paru dans le volume *The Last Phonological Rule* (Goldsmith 1993b). Cet article pose les bases d'un programme de recherche pour une phonologie stratifiée à trois niveaux et suggère, dans sa dernière section, un lien avec un modèle computationnel, appelé RÉSEAU DYNAMIQUE CALCULATOIRE (RCD, ang. *dynamic computational network*)¹, sur lequel John Goldsmith travaillait en collaboration avec l'un de ses doctorants, Gary Larson. Bien que j'aie lu cet article lors de mon doctorat, je n'en avais survolé la dernière partie que très superficiellement et les séries de nombres qu'elle présentait m'avaient alors laissé perplexe.

1. Prince (1993) montre que ce modèle est une approximation discrète d'un *modèle dynamique linéaire*, et cette terminologie est parfois employée.

J'ai donc décidé en le relisant de tenter de véritablement comprendre cette démarche en l'abordant sans *a priori*. Je peux maintenant dire avec le recul qu'il a profondément changé ma manière d'envisager la phonologie.

Il me semble ici utile de replacer l'approche RCD dans une perspective plus large dans la mesure où il existe des liens profonds entre cette approche et les approches connexionnistes qui ont émergé à la fin des années 80, avec la parution notamment des deux volumes du groupe *Parallel Distributed Processing* (PDP, Rumelhart et al. 1986a,b). Ces travaux partent du postulat que la cognition peut être modélisée, et *in fine* expliquée, sous la forme d'architectures de neurones formels inspirés des neurones biologiques. L'esprit humain n'est alors plus envisagé comme une machine manipulant des symboles discrets : l'activité mentale est modélisée comme le résultat de la dynamique d'architectures distribuées et le savoir est encodé directement dans les connexions synaptiques de ces réseaux.

5.1.1 Connexionnisme

Le connexionnisme prend ses racines dans les années 40 avec les travaux de Warren McCulloch et Walter Pitts, qui ont développé le premier modèle de NEURONE FORMEL. Il s'agit d'un modèle mathématique simplifié inspiré des neurones biologiques : le neurone formel est une unité calculatoire qui reçoit des valeurs en entrée provenant d'autres neurones (grâce à l'équivalent des dendrites d'un neurone biologique), qu'il intègre pour produire une valeur de sortie (envoyée dans l'équivalent de l'axone). Une illustration en est donnée à la figure 5.1. Le modèle de McCulloch et Pitts était relativement simple : les neurones formels ne pouvaient produire en sortie qu'une réponse binaire (0 ou 1) ; la valeur reçue en entrée était la somme des valeurs de sortie des neurones activés auxquels le neurone était connecté, chacun pouvant avoir un effet excitatoire ou inhibitoire ; enfin, l'activation du neurone était déterminée par une fonction pallier avec un seuil au-delà duquel le neurone était activé (valeur 1) et en dessous duquel il restait inactivé (valeur 0). McCulloch et Pitts ont montré que cette architecture simple permettait de simuler la logique booléenne de base.

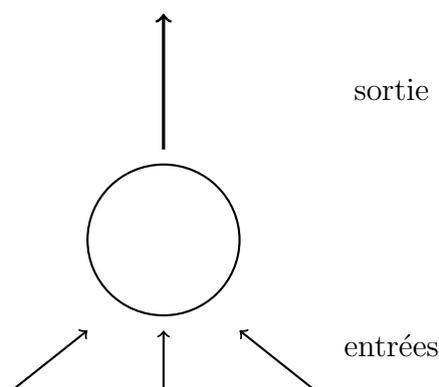


Figure 5.1. Neurone formel

Ce modèle simple a été ultérieurement raffiné par Rosenblatt (1958) qui a développé

le premier véritable réseau neuronal à travers le modèle qu'il a appelé PERCEPTRON. Comme le soulignent McLeod et al. (1998), la perspective de Rosenblatt diffère de manière importante de celle de McCulloch et Pitts : il a développé une véritable architecture *distribuée* de nature probabiliste, de sorte que si l'une des unités du système devient défaillante, la perte est compensée par l'activité des autres unités du réseau. Dans sa forme canonique, un perceptron est une couche de neurones dont toutes les unités sont connectées à une autre couche, composée d'un ou plusieurs neurones (cf. figure 5.2). A chaque connexion d'entrée entre deux neurones j (entrée) et i (sortie)² est associée un poids, qui indique la force de la connexion synaptique entre ces deux neurones. Dans le cas le plus simple, les valeurs des connexions d'entrée sont intégrées à l'aide d'une simple somme pondérée ; autrement dit, le signal d'entrée sur chaque synapse est multiplié par le poids de la synapse correspondante et les signaux sont additionnés. Lorsque le poids est positif, j aura un effet excitatoire sur i , alors qu'il aura un effet inhibitoire si ce poids est négatif. Formellement, pour un neurone i connecté à n neurones en entrée, $\mathbf{a} = (a_1, a_2, \dots, a_n)$ représente le vecteur d'activation de i et $\mathbf{w} = (w_{i1}, w_{i2}, \dots, w_{in})$ représente le vecteur des poids synaptiques d'entrée de i ; on définit alors le NETINPUT de i comme la somme des valeurs d'activation multipliées par les poids synaptiques correspondants (McLeod et al. 1998 : 17) :

$$(1) \quad \text{netinput}_i = \sum_{j=1}^n w_{ij} \cdot a_j$$

L'activation d'un neurone formel est déterminée par le netinput mais elle ne lui correspond en général pas de manière linéaire. L'activation de i est déterminée par sa FONCTION D'ACTIVATION qui prend en entrée le netinput de i et produit en sortie une valeur, qui peut être binaire (0 ou 1) comme dans le cas du modèle de McCulloch et Pitts, ou continue (une valeur sur l'intervalle $[0,1]$). Ceci permet de modéliser les effets de seuil que l'on observe dans les neurones biologiques.

Il a été montré dans les années 70 que le perceptron avait d'importantes limitations et qu'il existait certaines fonctions, telles que le OU exclusif³, qu'il n'était pas capable de calculer. Ces résultats négatifs ont momentanément freiné les recherches sur le perceptron et les réseaux de neurones, mais une série de travaux à la fin des années 70 et au début des années 80 ont relancé l'intérêt pour ce type de modèles en développant de nouvelles architectures neuronales et de nouvelles méthodes d'apprentissage palliant les limitations du perceptron. Rumelhart et al. (1988) ont ainsi montré comment une architecture neuronale dotée d'une couche de neurones interne (ou « cachée »), disposée entre la couche d'entrée et la couche de sortie, était capable d'apprendre des fonctions telles que le OU exclusif.

2. Selon la convention acceptée mais dont la logique m'échappe complètement, on désigne le neurone d'entrée par la lettre j et le neurone connecté en sortie par la lettre i . Je suis donc ici la pratique consacrée par l'usage.

3. A ou B mais pas A et B.

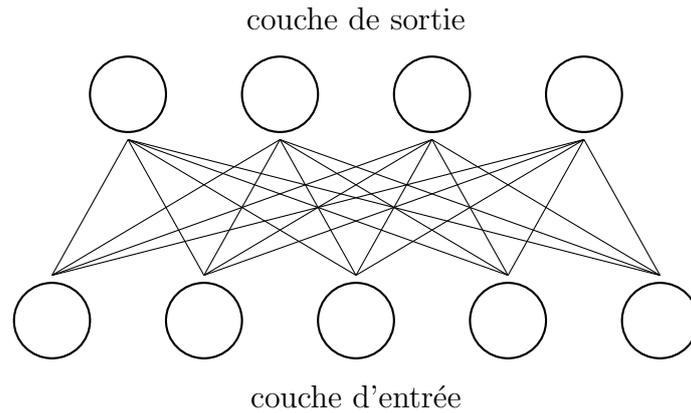


Figure 5.2. Perceptron

5.1.2 Réseau calculatoire dynamique

Les réseaux calculatoires dynamiques (RCD) ont été développés et étudiés dans une série de travaux, notamment Larson (1990) ; Goldsmith & Larson (1990, 1993) ; Goldsmith (1992, 1994) ; Laks (1995, 1997) et dont la présentation la plus détaillée est la thèse de doctorat de Gary Larson (1993). La version du modèle présentée par Larson (1993) comporte trois niveaux, à l'instar de la Phonologie Harmonique de Goldsmith (1993a) : un niveau autosegmental, qui spécifie l'organisation des segments en termes de traits et détermine la sonorité des segments (niveau A) ; un niveau syllabique, qui détermine la courbe de sonorité de la chaîne (niveau S) ; un niveau métrique, qui détermine le patron accentuel de la chaîne (niveau M). Ces trois niveaux s'alimentent mutuellement ($A \rightarrow S \rightarrow M$) : la dynamique du niveau A produit en sortie une valeur correspondant à la sonorité de chaque segment ; le niveau S identifie des pics de sonorité locaux et globaux, lesquels servent d'entrée au niveau M pour l'identification des proéminences accentuelles. Les niveaux syllabique et métrique sont modélisés à l'aide de RCD, alors que le niveau A peut être modélisé de diverses manières, mais est le plus souvent modélisé à l'aide de valeurs numériques *ad hoc* (voir néanmoins Larson 1993 : 22–28).

Ce modèle propose une architecture connexionniste originale pour l'analyse des phénomènes de syllabation et d'accentuation. L'hypothèse générale qui sous-tend cette approche est que le profil rythmique de la chaîne parlée est le résultat de l'interaction entre une valeur de sonorité (ou d'activation) inhérente et des dynamiques latérales de coopération et de compétition (c'est-à-dire d'excitation et d'inhibition) entre unités voisines (segments, syllabes). Un RCD est un perceptron unicouche à inhibition bilatérale, dans lequel un input de longueur n est représenté comme un réseau de n nœuds, chaque nœud possédant deux connexions synaptiques avec ses voisins adjacents. Les connexions de droite à gauche sont contrôlées par un paramètre (poids synaptique) α alors que les connexions de gauche à droite sont contrôlées par un paramètre β (cf. figure 5.3). Le modèle permet par ailleurs de prendre en compte un BIAIS initial ou final (voire pénultième) pour modéliser les

systèmes accentuels non quantitatifs.

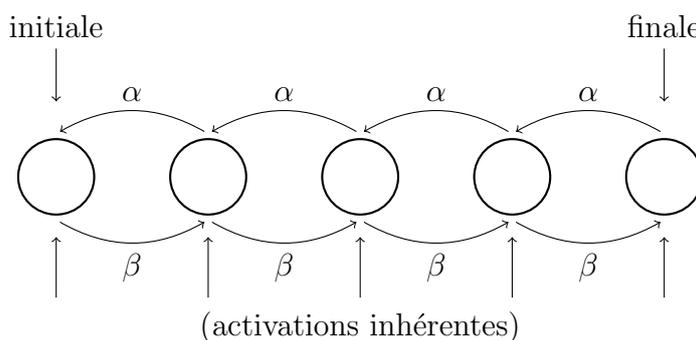


Figure 5.3. Réseau calculatoire dynamique

Dans un réseau syllabique, chaque unité représente un segment et reçoit en entrée une valeur correspondant à sa SONORITÉ INHÉRENTE (qui correspond à la sortie du niveau autosegmental pour cette unité), ainsi que la sortie de ses voisins adjacents pondérés par les poids synaptiques α et β . Elle transmet en sortie la somme des valeurs reçues en entrée, ce qui correspond à la SONORITÉ DÉRIVÉE de l'unité. En termes traditionnels, la sonorité inhérente correspond *grosso modo* à la forme sous-jacente alors que la sonorité dérivée correspond à la forme de surface. Le réseau évolue dynamiquement dans le temps jusqu'à ce qu'il se stabilise dans un état d'équilibre où de nouvelles mises à jour de l'activation des unités ne modifient plus (ou très peu) l'état global du système. C'est ce qui caractérise les propriétés *dynamiques* de cette architecture. La mise à jour du système est décrite par l'équation suivante (Goldsmith 1992 : 223) :

$$(2) \quad d_i^{t+1} = u_i + \alpha \cdot d_{i+1}^t + \beta \cdot d_{i-1}^t$$

où u est la sonorité inhérente et d_i^t représente la sonorité dérivée de l'unité i à l'instant t . En pratique, on détermine souvent un seuil θ en dessous duquel on considère que le système s'est stabilisé (mais voir le paragraphe suivant).

Prince (1993) a contribué de manière décisive à notre compréhension des RCD puisqu'il a prouvé les conditions dans lesquelles un RCD pouvait se stabiliser (ou *converger*). Je le donne en (3). Il a par ailleurs proposé une équation matricielle équivalente à (2) qui permet d'obtenir une solution exacte pour tout réseau convergent ⁴.

4. On trouvera sur ma page web (<http://julieneychenne.info/tools>) un module Python qui implémente cette équation ainsi que le modèle d'apprentissage décrit en §5.1.3. C'est le code que j'ai utilisé dans les travaux décrits ci-après.

(3) Théorème de Prince

Tout RCD satisfaisant à la condition $|\alpha\beta| \leq \frac{1}{4}$ est convergent pour tout n , n étant le nombre d'unités du réseau.

(adapté de Prince 1993 : 53)

Il y a donc en théorie une infinité de modèles convergents possibles. En pratique, on limite généralement les valeurs possibles de α et β à l'intervalle $[-0,5, 0,5]$.

Pour comprendre le fonctionnement de la propagation de l'onde de sonorité, examinons un exemple simple. Soit l'input $(0, 0, 0, 0, 0, 0, 0, 0, 1)$, qui représente la séquence de syllabes $\sigma\sigma\sigma\sigma\sigma\sigma\sigma\sigma$ dans un réseau métrique, et un modèle dont les paramètres sont $\alpha = -0,5$ et $\beta = 0$. La figure 5.4 montre la sonorité dérivée des unités du réseau : comme on peut le voir, l'unité la plus à droite propage sa sonorité vers la gauche d'un facteur α^n , où n est la distance entre deux unités. Ce poids étant négatif, le patron résulte en une succession de pics et de creux d'amplitude décroissante. Ce patron représente une alternance rythmique de type $\delta\sigma\delta\sigma\delta\sigma\delta\sigma\delta$ ⁵ : les unités impaires sont inhibées alors que les unités paires sont excitées. Au niveau syllabique, cette alternance peut être interprétée comme l'alternance entre syllabe accentuée et inaccentuée. L'aspect le plus intéressant de ce modèle à mes yeux est que cette propriété est une *conséquence de l'architecture du réseau*. Il n'est donc pas nécessaire de postuler des contraintes de type *CLASH ou *LAPSE pour éviter les séquences de pics ou de creux.

5.1.3 Algorithme d'apprentissage

S'il est possible de calibrer manuellement un réseau RCD pour modéliser des systèmes relativement simples, on souhaite évidemment pouvoir apprendre les paramètres du modèle à partir de données. L'algorithme d'apprentissage associé aux RCD est décrit en détail dans Larson (1993 : chap. 6) : il s'agit d'une variante de l'algorithme de RECUIT SIMULÉ (*simulated annealing*), qui permet d'inférer les paramètres α et β mais aussi les sonorités inhérentes des unités du modèle. Le modèle est un modèle d'APPRENTISSAGE SUPERVISÉ⁶ probabiliste qui intègre un paramètre τ appelé « température », initialement très élevée ($\tau = 1,0$). À chaque présentation d'une nouvelle forme, la température décroît en fonction d'un taux de refroidissement $\Delta\tau$, indépendamment du succès ou de l'échec. Lorsque le système prédit une forme erronée, les paramètres α et β sont modifiés

5. Dans cet exemple, l'écho d'accent s'atténue rapidement, mais un α dont la magnitude serait plus élevée s'atténuerait plus lentement. Voir Laks (1997 : chap. 5) pour de nombreux exemples appliqués à différents systèmes accentuels.

6. Dans l'apprentissage supervisé, les données sont classifiées par avance et le modèle exploite cette information lors de l'apprentissage. Dans l'APPRENTISSAGE NON SUPERVISÉ, les données ne sont pas étiquetées et le modèle classe automatiquement les données. Notons que les modèles supervisés ne sont pas forcément plus réalistes : par exemple, l'algorithme des k -moyennes, qui classe un ensemble de données en k catégories, dépend du paramètre k qui doit être spécifié à l'avance. Ses performances sont également en partie conditionnées par la phase d'initialisation du modèle, qui est en général aléatoire.

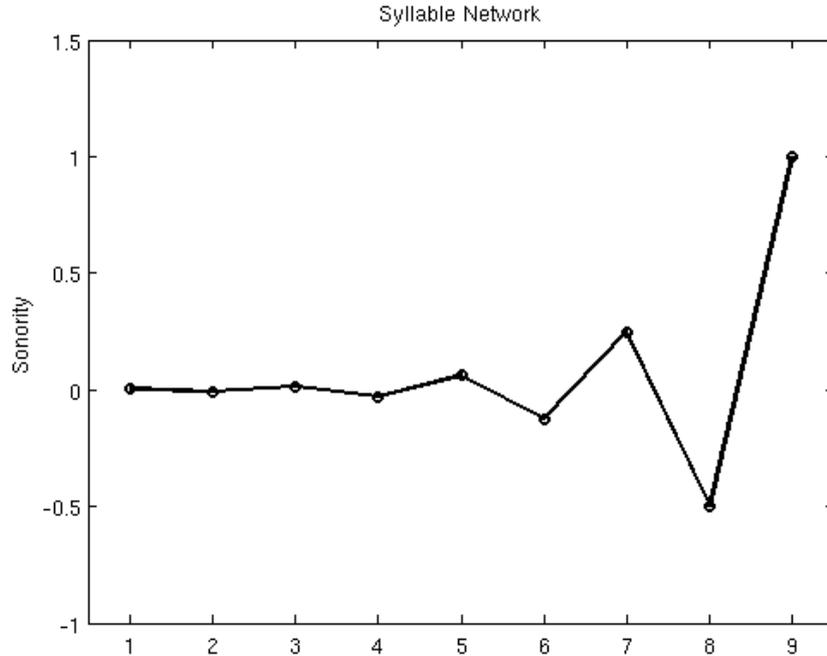


Figure 5.4. Propagation à gauche de l'onde rythmique

aléatoirement d'un facteur tiré d'une loi normale de moyenne zéro et de variance τ , normalisé par une constante κ . La température du système est également augmentée de la distance euclidienne du degré de modification de α et β . Enfin, la sonorité inhérente des segments dont la sonorité dérivée est incorrectement prédite est elle aussi modifiée, d'un facteur qui dépend de la température et d'une constante de normalisation que l'on notera λ . Lorsque la température du système descend sous un seuil θ , on considère que le système est « gelé » dans un état stable et la procédure d'apprentissage s'interrompt. L'intuition qui sous-tend cet algorithme est que plus le système fait d'erreurs, plus les configurations qu'il explore devraient être différentes de la configuration actuelle ; réciproquement, lorsque le système fait peu d'erreurs, sa grammaire devrait peu évoluer.

Il est important de distinguer ces *hyperparamètres*, qui appartiennent au modèle d'apprentissage, des *paramètres* α et β propres au RCD. Je résume ici les 5 hyperparamètres de cet algorithme d'apprentissage :

- τ : la température du système, qui commence toujours à 1,0.
- Δ_τ : le facteur de refroidissement du système
- θ : le seuil de gel du système
- κ : contrôle l'amplitude de la modification de α et β en cas d'erreur
- λ : contrôle l'amplitude de la modification de la sonorité inhérente en cas d'erreur

Le modèle rythmique du RCD, que ce soit au niveau syllabique ou métrique, repose fondamentalement sur la notion d'onde de sonorité, et abandonne de ce fait la notion de

constituance (Laks 1997 : §5.6). La « syllabation » d'un RCD est donc déterminée par l'identification de pics et creux locaux, en suivant un algorithme similaire au SONORITY SEQUENCING PRINCIPLE (SSP, cf. Clements 1990). L'algorithme se contente simplement d'indiquer les pics par H (*high*), les creux par L (*low*) et les autres segments par O (*other*) (Larson 1993). A titre d'exemple, un mot tel que *docteur* [dɔk.tœʁ] se verra attribuer la « syllabation » LHOLHL.

5.1.4 Apprenants biaisé et non biaisé

J'ai exploré les RCD dans deux études : la première, qui a été présentée au *Congrès Mondial de Linguistique Française* (CMLF) de 2012 (Eychenne 2012)^[26], est une étude de la structure syllabique en français méridional : la contribution principale de ce travail a été de montrer qu'il n'était pas nécessaire de stipuler a priori la sonorité inhérente des unités du réseau, et que l'architecture d'un RCD était suffisante pour en expliquer l'acquisition. La seconde étude (Eychenne 2013)^[4] approfondit la première et propose une modélisation du système métrique du français méridional dans ce modèle. Je me concentrerai ici sur la première étude.

Cette étude utilise un corpus de français méridional basé sur l'enquête PFC en Languedoc, à partir de laquelle j'ai constitué un corpus dictionnaire phonétisé, qui comprend 3314 types correspondant à 46 682 tokens. J'en donne quelques exemples en (4). Ce corpus a été divisé, après randomisation, en un corpus d'entraînement contenant 60% des données (1492 mots), un corpus de validation et un corpus de test contenant chacun 20% (497 mots). Le corpus d'entraînement, comme son nom l'indique, a été utilisé pour entraîner les dispositifs d'apprentissage ; le corpus de validation a été utilisé pour ajuster les hyperparamètres du dispositif d'apprentissage ; enfin, le corpus de test a été utilisé pour mesurer la performance du modèle sur des données nouvelles, qui n'ont pas été rencontrées pendant la phase d'apprentissage et de mise au point. Cette étape permet de vérifier les capacités de généralisation du modèle au-delà des données qu'il a rencontrées.

(4) Exemples du corpus dictionnaire de français méridional :

<i>génial</i>	ʒe.njal
<i>généalogiques</i>	ʒe.ne.a.lo.ʒi.kə
<i>général</i>	ʒe.ne.ral
<i>générale</i>	ʒe.ne.ra.lə
<i>génération</i>	ʒe.ne.ra.sjɔ ^N
<i>génétique</i>	ʒe.ne.ti.kə

La question centrale que j'ai abordée dans ce travail est la suivante : la notion de sonorité doit-elle être pré-câblée dans un dispositif d'apprentissage ou peut-elle être induite à partir des données ? Les RCD et l'algorithme d'apprentissage associé, en permettant d'ajuster la sonorité inhérente des segments au cours de l'apprentissage, offrent un cadre particulièrement fécond pour tenter de répondre à cette question. J'ai donc construit deux

modèles, dont l'un est un dispositif d'apprentissage BIAISÉ, qui dispose d'une échelle de sonorité donnée *a priori* (possiblement par une Grammaire Universelle innée), et l'autre un dispositif NON BIAISÉ qui commence *tabula rasa* avec une initialisation uniforme de la sonorité inhérente des segments. L'échelle de sonorité que j'ai adoptée pour le dispositif biaisé, et qui est adaptée d'après Larson (1993 : 58), est donnée en (5). En ce qui concerne les hyperparamètres du modèle, j'ai utilisé les valeurs $\kappa = 0,2$ et $\lambda = 0,125$ données par Larson (1993). Les meilleures performances pour les deux dispositifs ont été obtenues avec un taux de déclin très lent ($\Delta_\tau = 0,995$) et un seuil de refroidissement très bas ($\theta = 0,001$). Le dispositif d'apprentissage étant probabiliste, les résultats varient d'une tentative à l'autre. Pour cette raison, j'ai mesuré deux séries de 1000 tentatives sur le corpus d'entraînement et de validation avec les paramètres que je viens d'évoquer, l'une pour l'apprentissage biaisé et l'autre pour l'apprentissage non biaisé⁷. Dans les deux cas, pour la phase de test, j'ai préféré construire un « système moyen » à partir des valeurs obtenues pour α , β et le vecteur de sonorité inhérente sur les 1000 tentatives plutôt que de sélectionner le modèle dont la performance était la meilleure. Cette approche me semble refléter la performance globale du système de manière plus réaliste que ne le ferait le meilleur modèle. A titre de comparaison, j'ai également créé un modèle étalon qui tente de prédire la syllabation sur la seule base de l'échelle de sonorité (5) et du SSP, mais sans RCD : ce modèle obtient un taux de reconnaissance de 93,56% sur le corpus de test.

(5) Echelle de sonorité pour le dispositif biaisé :

occlusives non voisées	1
occlusives voisées	2
fricatives non voisées	3
fricatives voisées	4
occlusives nasales	5
liquides	6
glissantes/appendice nasal	7
schwa	8
voyelles hautes	9
voyelles moyennes	11
voyelle basse	13

Observons maintenant les résultats du dispositif biaisé. La performance du système est globalement très bonne : la moyenne sur le corpus de validation pour 1000 tentatives se situe à 98,98%, avec un minimum de 96,78% et un maximum de 100%. Il converge en moyenne au bout de 4,8 présentations du lexique (ou époques). Le modèle se stabilise en moyenne avec $\alpha = 0,210659$ et $\beta = -0,015455$. Sur le corpus de test, le modèle moyen obtient une performance de 98,59%, soit 7 erreurs⁸. Cette performance est supérieure à

7. A chaque simulation, les valeurs initiales de α et β étaient choisis aléatoirement dans l'intervalle $[-0,3 \ 0,3]$ selon une distribution uniforme.

8. Dans les deux modèles (biaisé et non biaisé), les erreurs sont presque toutes dues aux groupes /sC/

celle du modèle étalon sur le corpus de test (93,56%), ce qui montre que cette architecture est supérieure à une approche basée sur le seul SSP. Les valeurs de sonorité inhérente moyennes après apprentissage (arrondies au centième) sont les suivantes (comparer à (5)) :

(6) Sonorités inhérentes du dispositif biaisé après apprentissage :

a : 14,33	ɔ : 12,48	ɛ : 12,26	e : 12,15	o : 11,33
œ : 11,20	ø : 11,11	ə : 10,41	i : 10,33	u : 9,59
y : 9,41	w : 7,75	ɥ : 7,06	j : 7,05	ⁿ : 6,36
r : 5,48	l : 5,24	ʁ : 4,90	n : 3,99	ʒ : 3,69
z : 3,47	m : 3,46	v : 3,06	ʃ : 2,79	f : 2,42
g : 1,92	k : 1,73	b : 1,50	s : 1,27	d : 1,24
p : 0,54	t : -0,40			

Ces valeurs sont dans l'ensemble très proches des valeurs de départ, mais il faut noter en particulier la sonorité du /s/ qui, avec une valeur de 1,27, fait partie des segments les moins sonores avec /p/, /t/ et /d/, et se démarque nettement des autres fricatives.

Le dispositif non biaisé a lui aussi été évalué sur une série de 1000 présentations. Au départ de chaque simulation, la sonorité inhérente des segments était uniformément initialisée à 1. Le modèle obtient une performance moyenne sur le corpus de validation de 99,13%, avec un minimum de 97,38% et un maximum de 100%. Ce modèle converge plus lentement que le modèle biaisé (7,8 époques en moyenne), ce qui n'est pas surprenant puisqu'il doit déterminer la sonorité des segments sans aucune aide extérieure.

Les paramètres moyens qui émergent sont $\alpha = 0,204843$ et $\beta = -0,020909$, ce qui est similaire aux valeurs obtenues avec le modèle biaisé. Les valeurs de sonorité moyennes après apprentissage sont les suivantes :

(7) Sonorités inhérentes du dispositif non biaisé après apprentissage :

a : 4,58	ɔ : 3,99	ɛ : 3,84	ə : 3,76	e : 3,62
i : 3,33	o : 2,71	u : 2,67	y : 2,45	œ : 2,44
ø : 2,05	w : 1,98	ɥ : 1,51	j : 1,39	ⁿ : 1,12
r : 0,72	ʁ : 0,70	l : 0,60	ʃ : -0,01	ʒ : -0,06
z : -0,31	n : -0,57	g : -0,59	f : -0,81	v : -0,86
k : -0,87	m : -1,02	b : -1,15	s : -1,28	d : -1,37
p : -1,50	t : -2,13			

Bien que les valeurs numériques soient différentes, la hiérarchie de sonorité apprise est très semblable à celle obtenue en (6), mais sans aucun présupposé quant à la sonorité relative des unités. La corrélation entre les sonorités apprises par les deux systèmes est donnée à la figure 5.5 : le taux de corrélation est de plus de 95% ($r^2 = 0,9561$). Cela

à l'initiale de mot (ex : *spécialisé, spécialité, station, stage*).

suggère que quelle que soit la configuration de sonorité inhérente de départ, le modèle aboutit à des solutions très similaires au terme de l'apprentissage.

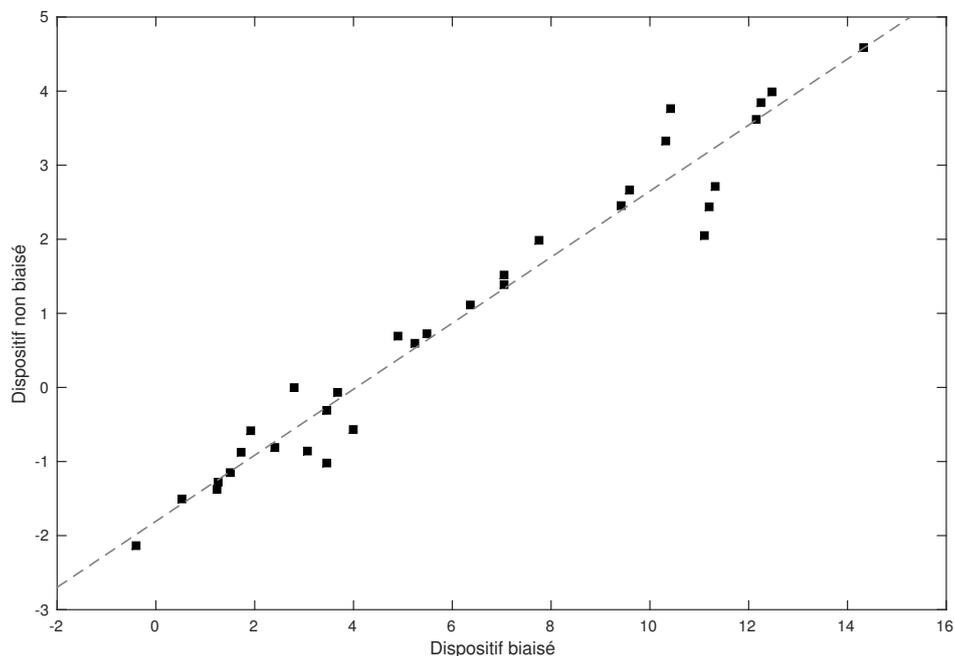


Figure 5.5. Corrélation entre les sonorités inférées par les modèles biaisé et non biaisé

La capacité de généralisation du modèle non biaisé moyen sur le corpus de test est par ailleurs très bonne et est en fait légèrement supérieure à celle du modèle biaisé moyen, à une forme près (98,79%, soit 6 erreurs). Là encore, la performance est supérieure au modèle étalon sur le corpus de test, ce qui suggère non seulement que le RCD est supérieur au seul principe de sonorité comme précédemment, mais aussi qu'il n'est pas nécessaire de disposer d'une échelle de sonorité donnée *a priori* : le RCD est à même d'inférer la sonorité des segments, et il semble le faire de manière plus efficace en partant d'une distribution uniforme puisque sa performance est légèrement supérieure, à la fois sur le corpus de validation et sur le corpus de test. Pour vérifier si la différence de performance sur les 1000 simulations du corpus de validation est significative, j'ai procédé à un test de Wilcoxon pour la médiane⁹ : la différence est significative ($p < 0,001$) et le modèle non biaisé s'avère plus performant que le modèle biaisé. Ces résultats montrent donc qu'il n'est pas nécessaire de disposer d'une échelle de sonorité pré-câblée : l'architecture du RCD, couplée à son algorithme d'apprentissage, permet d'inférer la sonorité relative des unités à partir de données auxquelles l'apprenant a été exposé (voir également Laks 1995, qui obtient des résultats allant dans le même sens, avec une architecture et des hypothèses de départ sensiblement différentes).

9. Les données n'étant pas normalement distribuées, un t-test pour la moyenne n'aurait pas été approprié.

Lorsque j'ai présenté ce travail au CMLF, on m'a adressé la critique suivante : ce modèle repose sur l'existence de données étiquetées données à l'avance (en d'autres termes, l'apprenant a accès aux descriptions structurales correctes des formes qu'il reçoit) ; il existe par ailleurs un « instructeur » qui indique à l'apprenant qu'il a fait une erreur, après quoi celui-ci modifie son comportement jusqu'à ce que ses prédictions correspondent aux données étiquetées. Or on sait que le cerveau ne fonctionne pas comme cela : les données à partir desquelles l'apprenant infère sa grammaire ne sont pas pré-étiquetées et il n'y a d'agent externe qui guide l'apprenant. Ce modèle n'est donc pas du tout réaliste... Je crois que cette critique repose sur deux erreurs d'interprétation, et je voudrais les dissiper au cas où elles se présenteraient à l'esprit du lecteur. Il ne faut tout d'abord pas interpréter trop littéralement ce que j'appellerai la MÉTAPHORE DE L'INSTITUTEUR PUNITIF, selon laquelle l'apprenant, dès qu'il produit une forme erronée, reçoit un coup de règle sur les doigts, puis adapte en conséquence son comportement de manière à éviter la même sanction dans le futur. S'il s'agit bien d'une manière imagée d'envisager l'apprentissage supervisé, il y a deux aspects importants qu'il faut garder à l'esprit. En premier lieu, le changement induit par la boucle de rétro-action est *aléatoire* : l'apprenant a détecté que la solution actuelle était sub-optimale, et va explorer une solution voisine afin de déterminer si elle lui est supérieure. Deuxièmement, et ce point me semble capital, rien n'oblige *a priori* dans les modèles supervisés à ce que la source de rétro-action soit un *agent externe* (l'instituteur punitif) : il est tout à fait plausible que la boucle de retour soit un module interne au système cognitif de l'apprenant. La deuxième erreur d'interprétation à mon sens est que le problème de l'accès aux descriptions structurales correctes ne se pose que si ces descriptions structurales ont une structure complexe et non manifeste, comme c'est par exemple le cas des constituants de la hiérarchie prosodique. La théorie syllabique des RCD repose sur la notion d'onde sonore, et les descriptions se résument donc à une chaîne de pics (H), creux (L) et positions intermédiaires (O). De telles descriptions pourraient tout à fait être *construites* par le système perceptif de l'apprenant. On peut ainsi imaginer que le système perceptif de l'apprenant soit un module d'apprentissage non supervisé qui attribue des descriptions structurales aux chaînes reçues en input (il s'agit d'une tâche de classification assez simple), et que ce module est connecté en série à un système d'apprentissage supervisé tel que celui utilisé ici. Dans cette perspective, apprentissage supervisé et non supervisé ne sont pas des paradigmes antinomiques, mais peuvent au contraire coopérer au sein d'un système d'apprentissage plus large et composite.

5.2 Phonotaxe

Les résultats encourageants obtenus avec les RCD m'ont poussé à explorer de manière plus actives d'autres approches, et je présenterai brièvement dans cette section une étude que j'ai réalisée sur la phonotaxe du français (Eychenne 2015a)^[8]. Ce travail a été publié dans un numéro spécial de *Langages*, dirigé par Gabriel Bergounioux et Tobias Scheer et intitulé *Phonologies contemporaines*. Il s'agit à ma connaissance de la première application

à la phonotaxe du français du modèle phonotactique développé par Hayes & Wilson (2008).

5.2.1 Le modèle MaxEnt

Le modèle de Hayes & Wilson (2008) est basé sur le modèle inférentiel appelé MAXIMUM ENTROPY (MaxEnt), un modèle très utilisé en traitement automatique du langage pour des tâches de classification, et parfois également pour le développement d'étiqueteurs grammaticaux (Manning & Schutze 1999). Goldwater & Johnson (2003) ont les premiers identifié les liens étroits qui existent entre la Grammaire Harmonique (GH) développée par Paul Smolensky et ses collaborateurs (Legendre et al. 1990 ; Pater 2009 ; Potts et al. 2010 ; Smolensky & Legendre 2006a,b), OT et le modèle MaxEnt. La différence la plus fondamentale est que dans la GH et comme dans MaxEnt, les contraintes sont pondérées et additionnées linéairement, alors qu'elles sont ordonnées en OT.

Le modèle MaxEnt offre un cadre théorique qui permet de modéliser des grammaires gradientes, sous forme probabiliste. Il repose de manière cruciale sur la notion d'ENTROPIE issue des travaux de Claude Shannon en théorie de l'information (Shannon 1948). L'entropie est une mesure du taux moyen d'information (ou, de manière équivalente, d'incertitude) contenu dans un signal. La relation entre information et incertitude n'est pas intuitive au premier abord, mais elle peut être illustrée à l'aide d'un exemple simple. Imaginons deux systèmes de communication rudimentaires qui envoient des mots (lexicaux) isolés entre un émetteur et un récepteur, l'un en finnois, et l'autre en espagnol : afin de minimiser le coût de transfert, on cherche à réduire la taille du message. En finnois, l'accent lexical primaire est entièrement prédictible puisqu'il tombe toujours sur la première syllabe, et il est donc entièrement redondant de le transmettre (en termes classiques, il n'est nécessaire de le spécifier dans la représentation sous-jacente). Il ne contient donc aucune information utile pour le récepteur puisque celui-ci est toujours certain de la position de l'accent. En espagnol, l'accent est beaucoup moins prédictible et il faudra dans de nombreux cas le transmettre avec le message (autrement dit, il doit être indiqué dans la représentation sous-jacente de certains mots, par exemple *término* « fin »). Le récepteur espagnol est beaucoup moins certain qu'en finnois quant à la position de l'accent dans un mot, et l'accent communique de ce fait un certain contenu informationnel. La notion d'entropie formalise précisément cette intuition. Il s'agit d'une mesure du nombre de choix binaires (ou bits) moyen nécessaires pour l'encodage d'un message dans un système de communication.

Le principe de maximisation de l'entropie est lui aussi assez simple, et je prendrai à nouveau un exemple pour l'illustrer. Imaginons un concours qui récompense les 5 premiers gagnants. Sachant que l'on dispose d'une somme de 1400 € et que l'on veut attribuer 500 € au premier et 300 € au second : combien doit-on attribuer aux trois suivants ? Il existe une infinité de manières de diviser les 600 € restants en trois, mais le principe de maximisation de l'entropie stipule qu'en l'absence de contraintes additionnelles (ex : le troisième doit recevoir plus que le quatrième), la meilleure distribution est la plus uniforme

(en l'occurrence, 200 € chacun). Le principe de maximisation de l'entropie formalise cette intuition simple : la distribution $\{500, 300, 200, 200, 200\}$ est celle qui a l'entropie la plus élevée eu égard aux contraintes que l'on s'est imposées au départ.

Hayes & Wilson (2008) ont appliqué le modèle MaxEnt à l'apprentissage de divers patrons phonotactiques, en l'adaptant de sorte qu'il soit capable d'inférer des contraintes négatives à partir d'un ensemble de données fournies en entrée, et ils ont implémenté leur modèle dans un programme informatique librement disponible¹⁰. Ce modèle abandonne la notion de forme d'entrée puisque l'objectif est ici d'induire une grammaire phonotactique à partir de données observées. Toutes les contraintes sont donc des contraintes structurales et il n'y a pas de contraintes de fidélité (la notion de fidélité n'a plus aucun sens dès lors qu'il n'y a plus d'input).

Formellement, le modèle attribue à chaque forme une probabilité calculée à partir de son profil de violation des contraintes de la grammaire. Cette probabilité est obtenue à partir d'un score défini par l'équation suivante (Hayes & Wilson 2008 : 383) :

$$(8) \quad h(x) = \sum_{i=1}^N w_i C_i(x)$$

où w_i est le poids de la i ème contrainte (avec $w_i \geq 0$), $C_i(x)$ le nombre de violations de la contrainte i par le candidat x , et Σ la somme des violations des contraintes pondérées par leur poids respectif. Le score $h(x)$ étant positif, il peut être directement interprété comme une mesure de la dysharmonie d'un candidat. On peut à partir de ce score déterminer pour chaque forme une valeur MaxEnt P^* telle que :

$$(9) \quad P^*(x) = e^{-h(x)}$$

où e est la base du logarithme naturel, avec en exposant l'opposé du score de x . Le signe négatif garantit que les formes les plus harmoniques (i.e. celles qui ont un score $h(x)$ relativement faible) recevront une valeur MaxEnt plus élevée que les candidats moins harmoniques. La probabilité $P(x)$ d'un candidat est alors obtenue en normalisant la valeur Maxent :

$$(10) \quad P(x) = \frac{P^*(x)}{\sum_y P^*(y)}$$

5.2.2 Application au français

Pour appliquer ce modèle au français, j'ai dans un premier temps envisagé d'utiliser des données issues du corpus PFC, mais j'y ai assez rapidement renoncé pour plusieurs

10. <http://www.linguistics.ucla.edu/people/hayes/Phonotactics/Index.htm>

raisons. Tout d'abord, l'outil fourni par Hayes et Wilson ne peut traiter que des corpus dictionnaires : on ne peut donc travailler qu'au niveau du mot, et il n'est pas possible à ma connaissance de travailler au niveau des énoncés, ce qui limitait déjà fortement l'intérêt d'utiliser des corpus de parole authentiques. De plus, le seul module de conversion graphème/phonème librement accessible dont j'aie connaissance (Goldman 2011) produit des formes phonétiques très standardisées : il était donc bien plus simple d'utiliser le corpus dictionnaire LEXIQUE, qui contient lui aussi des formes standard mais a l'avantage d'être « prêt à l'emploi ». Enfin, LEXIQUE contient un nombre de types et de tokens bien plus important que ce que j'aurais pu obtenir à partir d'enquêtes PFC. Néanmoins, le modèle étant computationnellement très lourd, je n'ai pas pu utiliser la base LEXIQUE en entier, et j'ai donc extrait (de manière aléatoire) un corpus d'entraînement de 10000 formes et un corpus de test de 2000 formes.

Les contraintes inférées par le modèle sont des bigrammes et trigrammes de structures de traits. Il a donc été nécessaire de coder l'inventaire du français à l'aide de traits distinctifs binaires. Notons également que par commodité, les frontières de mots sont traitées dans ce cadre comme ayant le trait [+mot], dans le style SPE : cela permet de construire des contraintes portant sur les bords gauche et droit du mot. Par ailleurs, ce modèle est un modèle probabiliste, ce qui implique les résultats varient d'une simulation à l'autre : dans la mesure où ils étaient très similaires dans les quatre simulations que j'ai réalisées, je n'ai discuté que les résultats obtenus pour la dernière simulation. J'en reprends ici en (11) les 10 contraintes les plus fortes inférées par la grammaire.

(11) Contraintes les plus fortes dans la grammaire du français¹¹ :

contrainte	exemples	score
*[-syll][-syll,+son][+cons]	*snt *klt	5,48
*[+syll,+haut,-arron][+dors,+nas]	*uŋ *aŋ	5,09
*[-tendu,+arron][+mot]	*œ# *ɔ#	5,08
*[+mot][-syll,+son][+cons]	*#ɛt *#lm	4,78
*[-syll][-syll,-cons][+mot]	*mj# *tw#	4,60
*[+mot][-bas,-arron,-avant]	*#ə	4,57
*[+syll,+nas][-cons]	*ãi ěj	4,39
*[-bas,-arron,-avant][+mot]	*ə#	4,38
*[+mot][+dors,+nas]	*#ŋ	4,37
*[-bas,-arron,-avant][+voisé,+dors][+mot]	*ək# *ən#	4,31

Comme on le voit à travers les exemples donnés, ces contraintes sont aisément interprétables et correspondent à de véritables généralisations sur la phonotaxe du français. La première pénalise les séquences triconsonantiques dont le second élément est un pic de sonorité local. La seconde exprime la généralisation selon laquelle /ŋ/ est presque toujours précédé de /i/, puisqu'il s'agit d'un segment emprunté à l'anglais (*parking*, etc.).

11. L'opérateur $\hat{}$ représente le complément (tout segment qui ne partage pas les traits indiqués).

La troisième contrainte pénalise les voyelles moyennes arrondies mi-ouvertes en position finale, ce qui est effectivement une contrainte active de la phonologie du français, que l'on peut par exemple voir à l'œuvre dans la morphologie non concaténative (*professionnel* [pʁɔfɛsjɔnel] → [pʁɔ], et non *[pʁɔ]). Nous voyons également émerger trois contraintes sur la distribution de schwa, qui pénalisent schwa en début et fin de mot, ainsi qu'en syllabe fermée en fin de mot. Cette dernière contrainte est intéressante à double titre : d'abord, le fait qu'elle cible un groupe *ad hoc* de consonnes (les segments qui ne sont pas voisés et dorsaux, *i.e.* toute consonne autre que [g]) tient au fait qu'il y avait dans le corpus d'apprentissage le mot d'emprunt *atabeg*, transcrit [atabɛg] ; ensuite, le fait que la grammaire n'ait pas pu induire la généralisation selon laquelle schwa est interdit en syllabe fermée est attribuable au fait que les séquences inférées sont toutes des *n*-grammes de structures de traits, ce qui fait que la notion de syllabe n'est pas accessible au dispositif inférentiel. Il faudrait ici enrichir le modèle, comme le font Hayes & Wilson (2008) dans certaines de leurs simulations. Néanmoins, l'analyse détaillée des contraintes inférées par le modèle, que je ne reprendrai pas ici, montre que la plupart de ces contraintes font sens, bien que certaines soient formulées de manière moins intuitive que ce à quoi pourrait s'attendre un phonologue (voir à ce titre la spécification de schwa en tant que [-bas,-arron,-avant]). J'ai par ailleurs montré que les prédictions du modèle sur le corpus de test s'accordaient avec ce à quoi l'on pouvait s'attendre, le modèle attribuant notamment un score de dysharmonie élevé aux mots d'emprunts.

Comme je l'ai mentionné plus haut, le score $h(x)$ (et donc, P^*) offre une mesure directe de l'harmonie d'une forme. Alors qu'en théorie de l'optimalité l'harmonie est *discrète* (elle compte le nombre de violations) et *locale* (elle considère les contraintes une par une, sans se soucier des contraintes subordonnées), l'harmonie dans le cadre MaxEnt (et dans la Grammaire Harmonique) est *gradiente*, puisqu'elle prend des valeurs numériques continues, et *globale*, puisqu'elle correspond à la somme pondérée de l'harmonie de chaque contrainte. C'est ce qui permet l'apparition de ce qui est connu dans la littérature anglophone sous le nom de GANG EFFECT, où plusieurs contraintes de poids faible peuvent « s'entraider » pour annuler l'effet d'une contrainte dont le poids est plus élevé (Pater 2009). En théorie de l'optimalité, une contrainte super-ordonnée a toujours précedence sur une contrainte subordonnée, sauf à utiliser des mécanismes additionnels tels que la conjonction de contraintes (Moreton & Smolensky 2002).

Le caractère gradient de l'harmonie dans ce modèle permet d'explorer de manière précise le lien potentiel entre harmonie et fréquence, sous l'hypothèse que les formes fréquentes seraient comparativement plus harmoniques (moins « marquées ») que les formes plus rares¹². Leur corrélation est illustrée à la figure 5.6 : comme on peut le voir, il n'y a pas de corrélation linéaire entre ces deux variables, ce qui est confirmé par un test de corrélation pour le rang de Kendall ($p = 0,311$)¹³. Bien que les formes les plus fréquentes aient toutes une harmonie élevée, il existe des formes dysharmoniques relati-

12. Je précise que le modèle a été entraîné sur les types et non sur les tokens.

13. Une investigation plus avancée des résultats dans Matlab montre qu'il n'y a pas non plus de corrélation non linéaire.

vement fréquentes et moins fréquentes, et de nombreuses formes rares ont également une harmonie élevée.

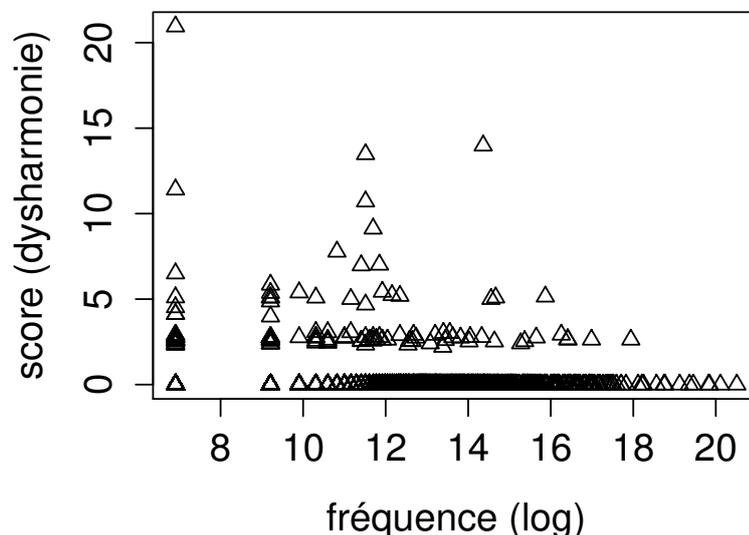


Figure 5.6. Corrélation entre fréquence et dysharmonie (Eychenne 2015a : 86)

Les résultats obtenus avec ce modèle sont là encore encourageants, car ils montrent qu’il est possible d’induire un pan non trivial de la phonotaxe du français à partir d’un mécanisme d’apprentissage très général et d’heuristiques assez simples. Ce modèle n’est pas un modèle *tabula rasa* et Hayes et Wilson montrent clairement que leur modèle étalon échoue sur des problèmes d’apprentissage de relations non locales (telles que l’harmonie vocalique) s’il le modèle n’est pas équipé d’heuristiques additionnelles (en l’occurrence, une tire où seules les voyelles sont projetées, de sorte que les restrictions de co-occurrence soient locales sur cette tire). Ces résultats montrent néanmoins qu’il est possible d’aller assez loin avec assez peu, et ils me confortent dans l’idée qu’il est nécessaire de continuer à explorer des modèles inductifs dans la modélisation grammaticale.

5.3 Distinctivité, allophonie et rendement fonctionnel

Mes travaux de modélisation les plus récents ont porté sur la notion de distinctivité, que j’ai cherché à revisiter dans une perspective quantitative. Je tenterai ici de retracer la genèse de cette réflexion, qui a commencé par l’étude de l’allophonie des voyelles moyennes en français et qui, dans sa forme la plus aboutie, s’est concrétisée dans une étude de la fusion de phonèmes en coréen, dans un travail en cours d’évaluation (Eychenne & Jang

2016)^[31].

5.3.1 Distinctivité gradiente

Le point de départ de ma réflexion a été une interrogation simple. Le français méridional commun et le français parisien qui ont tous les deux les six voyelles phonétiques [e ε ø œ o ɔ]; or ces voyelles représentent respectivement trois phonèmes dans le Sud (à savoir /E (E O/) et six dans le Nord. On sait par ailleurs que dans le Nord ces oppositions perdent du terrain, notamment l'opposition /ø/ vs /œ/ et l'opposition /e/ vs /ε/ en syllabe ouverte (Hansen & Juillard 2011). La question que je me suis posée était la suivante : comment peut-on quantifier précisément le degré d'allophonie (ou inversement de distinctivité) de ces voyelles à partir de corpus ?

Je me suis donc tourné dans un premier temps vers l'approche la plus susceptible de m'orienter dans la bonne direction : le distributionnalisme, et plus particulièrement l'ouvrage *Structural Linguistics* de Zellig Harris (1951). La méthode de Harris pour identifier les oppositions et relations allophoniques est simple et bien connue, mais il n'est pas inutile de la rappeler. Pour identifier les phonèmes et leur distribution, la méthode distributionnelle de Harris utilise un tableau, dont les colonnes représentent l'ensemble des environnements linguistiquement pertinents et dont chaque ligne représente un segment. Admettons que l'on ait identifié l'ensemble des contextes pertinents (ce qui est une tâche difficile à laquelle Harris ne s'attelle d'ailleurs pas) : pour identifier les relations allophoniques, il suffit de cocher toutes les cases (i, j) pour lesquelles le segment i apparaît dans le contexte j . Un exemple en est donné à la figure 5.7.

	W-X	Y-Z
a	✓	
b	✓	
c		✓
d		✓

Figure 5.7. Arrangement de segments d'après leur distribution (Harris 1951 : 66)

Pour toute paire de segments x et y , on peut identifier deux cas prototypiques : x et y apparaissent toujours dans le(s) même(s) contexte(s), auquel cas leur distribution est strictement ÉQUIVALENTE, ou bien ils n'apparaissent jamais dans le même contexte, auquel cas leur distribution est NON SÉCANTE. Le terme *non sécant* est ma traduction de *non-intersecting* que j'emprunte à Hockett (1966), et je réserverai le terme COMPLÉMENTAIRE pour une relation qui s'établit entre deux allophones. La question est donc précisément : quant peut-on dire qu'une distribution non sécante est complémentaire ? Prenons l'exemple de l'espagnol, où les plosives voisées alternent avec les fricatives voisées¹⁴

14. Elles sont généralement réalisées comme des approximantes, mais ce n'est pas essentiel ici.

correspondantes. Supposons que a , b , c et d à la figure 5.7 représentent respectivement les segments $[b \ d \ \beta \ \delta]$. Selon Harris (1951 : 66), on peut dire que « $[a] + [c] = /A/$, and $[b] + [c] = /B/$, if, in respect to some criterion, $a : c = b : d$ ». Le critère en question, tellement évident qu'il est laissé implicite, est la SIMILARITÉ PHONÉTIQUE (ou de manière équivalente, la DISTANCE PHONÉTIQUE). On peut établir que $[b \sim \beta]$ et $[d \sim \delta]$ car $[b]$ et $[\beta]$ sont plus proches l'un de l'autre que $[b]$ et $[\delta]$, et de même $[d]$ et $[\delta]$ sont plus proches que ne le sont $[d]$ et $[\beta]$. C'est également la similarité phonétique qui justifie qu'en anglais $[t]$ et $[t^h]$ soient considérés comme des allophones alors que $[h]$ et $[\eta]$ représentent traditionnellement deux phonèmes distincts bien qu'ils aient une distribution (quasi) non sécante. La distribution, ou plus précisément ce que j'appellerai ici l'ÉQUIVALENCE DISTRIBUTIONNELLE, et la DISTANCE PHONÉTIQUE co-définissent donc la notion de RELATION ALLOPHONIQUE et, partant de là, de PHONÈME. Il existe tout un débat sur la nature du phonème sur lequel je ne reviendrai pas ici (Twaddell 1935), mais comme on le verra au fil de la discussion, le phonème n'a pas, dans l'approche que je défends, de statut autonome : il s'agit simplement d'un ensemble de segments entretenant une relation allophonique. Autrement dit, dans l'exemple espagnol, $/d/$ n'est rien de plus qu'un raccourci commode pour $\{d, \delta\}$, et l'on aurait pu choisir tout autre symbole.

Si la méthode de Harris permet d'identifier les relations allophoniques et les phonèmes d'une langue, elle soulève un problème majeur que n'ont pas manqué de relever les linguistes : toutes les oppositions n'ont pas la même importance dans une langue (Hockett 1955 : 213; Hall 2013; Scobbie & Stuart-Smith 2008). A titre d'exemple, la fricative vélaire $[x]$ de l'anglais écossais que l'on trouve dans *loch* $[lɔx]$, bien qu'elle soit toujours en usage et qu'elle constitue un trait saillant d'un point de vue sociolinguistique, n'apparaît que dans une seule paire minimale (*loch* vs *lock*) et a une fréquence (à la fois de type et de token) très faible (Scobbie & Stuart-Smith 2008 : 95–96). J'ai donc cherché à développer une approche qui permette de quantifier le degré de distinctivité d'une paire de segments, en prenant en compte son équivalence distributionnelle (dans une perspective de phonologie de la Parole, c'est-à-dire en prenant en compte la fréquence d'usage), et sa distance phonétique. J'ai présenté une première version de ce travail aux *Journées FLOORAL* qui se sont tenues à Paris en décembre 2014 dans une étude des voyelles moyennes : ce travail était exploratoire et les résultats étaient encourageants, mais mon approche avait deux limitations : elle se cantonnait aux voyelles moyennes sans prendre en compte les autres voyelles (orales) et la mesure de l'équivalence distributionnelle était plutôt *ad hoc* puisqu'elle avait été « bricolée » par mes soins.

Par un heureux concours de circonstance, j'ai donné au semestre de printemps 2015 un séminaire de phonologie diachronique, au cours duquel j'ai donc été amené à revisiter la notion de RENDEMENT FONCTIONNEL. Je ne connaissais cette notion qu'à travers le travail de Martinet (en particulier son traité de phonologie diachronique, Martinet 1955) : dans son acception la plus simple, mais aussi la plus courante, le rendement fonctionnel d'une opposition correspond au nombre de paires minimales qu'elle permet de distinguer dans le vocabulaire. Mes recherches sur ce thème m'ont permis de prendre connaissance de la définition du rendement fonctionnel qu'a proposée Hockett (1955), et qui est exprimée

dans le cadre de la théorie de l'information (cf. §5.2 ci-dessus). L'approche de Hockett prend non seulement en compte les propriétés distributionnelles, mais aussi la fréquence d'usage puisque, comme il le note lui-même :

Assuming that two phonemes, x and y , can contrast at all, then the functional load carried by the contrast will be greater if both x and y have relatively high text frequencies than if one has a high frequency and the other a low frequency, and greater under those second conditions than if both x and y have low frequencies.

(Hockett 1955 : 215)

Cette approche, qui a été redécouverte il y a une dizaine d'années par Partha Niyogi et l'un de ses étudiants, Dinoj Surendran (Surendran & Niyogi 2003, 2006), m'a semblé très intéressante car elle offrait une mesure inductive sur corpus du rendement fonctionnel. J'ai dans un premier temps appliqué cette méthode à l'analyse du rendement fonctionnel à la fusion de /e/ et /ɛ/ en coréen. Tae Yeoub Jang et moi avons donc, en complément de notre étude phonétique de cette opposition (discutée en §2.2.2, p. 48 suiv.), entrepris de mesurer le rendement de cette opposition (Eychenne & Jang 2015 : §4)^[7]. Nos résultats suggéraient que cette opposition était parmi les trois oppositions qui avaient le rendement le plus bas. Nous avons donc observé que « the low functional load of the /e/ vs /ɛ/ contrast appears to be a factor that may have facilitated, if not triggered the merger currently observed in contemporary Korean ». Nous n'avons cependant fait aucune proposition quant à la manière dont le rendement fonctionnel aurait pu contribuer à la fusion de ces voyelles.

Peu après la publication de cet article, l'idée nous est venue de mettre en relation le travail que j'avais mené sur les voyelles moyennes d'un côté et notre travail sur le coréen de l'autre. Nous avons alors décidé de réfléchir ensemble à la manière d'intégrer mon approche de la distinctivité avec la mesure du rendement fonctionnel de Hockett. Le fruit de ce travail est synthétisé dans Eychenne & Jang (2016)^[31]. Je présente les aspects formels du modèle dans sa version la plus récente (et la plus aboutie) à la sous-section suivante.

5.3.2 Le modèle

Le modèle que nous avons développé offre une mesure gradiente de la distinctivité : pour toute paire de segment, nous définissons un INDICE DE DISTINCTIVITÉ qui prend en compte le rendement fonctionnel, mesuré comme la perte d'information selon l'approche de Hockett, et la distance phonétique entre les segments.

5.3.2.1 Rendement fonctionnel

Comme on l'a vu plus haut, toutes les oppositions phonologiques n'ont pas la même importance dans une langue : elles peuvent apparaître plus ou moins souvent dans le vocabulaire, et les items dans lesquels elles apparaissent peuvent être plus ou moins fréquents. Comme l'avaient noté aussi bien Hockett (1955 : 215) que Martinet (1955) on s'attend

à ce que, toutes choses égales par ailleurs, une opposition joue un rôle plus important si elle apparaît dans un grand nombre de mots qui sont par ailleurs très fréquents.

Hockett (1955) est le premier auteur à ma connaissance à avoir formalisé cette notion de manière précise, en se basant sur la théorie de l'information de Shannon. Hockett définit le rendement fonctionnel d'une paire d'unités x et y de la manière suivante :

$$(12) \quad R = \frac{H - H^*}{H}$$

où H est l'entropie du système de la langue et H^* représente l'entropie du pseudo-système équivalent dans lequel x et y ont été fusionnés (par exemple en les remplaçant par un autre symbole tel que X). Cette équation représente donc la proportion d'information perdue lorsque l'opposition entre x et y est neutralisée (de manière absolue). Plus l'opposition est importante dans la langue, plus la perte d'information sera grande lorsque les symboles sont fusionnés. Si les unités ont une distribution strictement non sécante, la perte d'information sera nulle puisqu'ils n'apparaissent jamais dans le même contexte et que leur fusion n'affecte donc pas le système (ils sont interchangeable).

Pour pouvoir utiliser l'équation (12), Il est nécessaire de formaliser le problème de manière plus précise (voir Surendran & Niyogi 2003, 2006). Soit \mathcal{L} un échantillon représentatif (*i.e.* un corpus) de la langue. On notera $H(\mathcal{L})$ l'entropie du corpus. On peut approximer $H(\mathcal{L})$ en faisant l'hypothèse que ce corpus a été généré par un processus de Markov de degré k , ce qui signifie que la probabilité d'occurrence d'une unité ne dépend que des k unités précédentes. La précision de H croît avec k , mais dans les tâches de ce type on utilise le plus souvent des unigrammes ($k = 0$), des bigrammes ($k = 1$) ou des trigrammes ($k = 2$); toutes les mesures rapportées dans cette section ont été calculées à partir de trigrammes. L'entropie de \mathcal{L} est calculée comme suit (Hockett 1955 : 217) :

$$(13) \quad H(\mathcal{L}) = -\frac{1}{k+1} \sum_{x \in X} p(x) \log_2 p(x)$$

où k est l'ordre du processus de Markov, X représente l'ensemble des n -grammes du corpus, et $p(x)$ est la probabilité du n -gramme x estimée à partir du corpus. La probabilité de chaque n -gramme correspond à :

$$(14) \quad p(x) = \frac{c(x)}{\sum_{i=1}^n c(y_i)}$$

où $c(\cdot)$ est le nombre d'occurrences du n -gramme¹⁵ dans le corpus. La formulation finale du rendement fonctionnel (λ) diffère légèrement de celle de Hockett : nous l'avons

15. Un mot tel que /keguli/ « grenouille » contient par exemple les trigrammes suivants : keg, egu, gul, uli.

convertie en pourcentage pour plus de lisibilité, soit :

$$(15) \quad \lambda(x,y) = 100 \times \frac{H(\mathcal{L}) - H(\mathcal{L}_{xy})}{H(\mathcal{L})}$$

Ainsi, si le rendement fonctionne de $/x \sim y/$ est 5,7, cela signifie que 5,7% d'information ont été perdus par la fusion de x et y .

La mesure de Hockett a selon moi trois atouts importants : le fait qu'elle soit basée sur la théorie de l'information lui confère tout d'abord une solide assise formelle puisqu'il s'agit d'une théorie mathématique établie et qui a été notamment appliquée à l'étude et à la réalisation de nombreux systèmes de communication. Elle permet par ailleurs de prendre en compte la distribution des unités sous forme de n -grammes¹⁶, et représente en ce sens une mesure de l'équivalence distributionnelle. Enfin, en permettant de prendre en compte la fréquence, cette mesure se situe dans le sillage des grammaires d'usage et permet d'appréhender (en partie) le système de la Langue à partir d'un échantillon de Parole, en accord avec la position que j'ai défendue dans le premier chapitre de ce mémoire.

5.3.2.2 Distance phonétique

Dans l'approche que nous avons adoptée, la distance phonétique est mesurée au plan perceptif, et nous suivons en cela les nombreux travaux qui considèrent que la perception joue un rôle majeur dans la formation des systèmes phonologiques et dans le changement (cf. Ohala 1981 ; Lindblom 1986 ; Hume & Johnson 2001 ; Hume et al. 2013 *inter alia*).

Notre approche initiale utilisait une simple distance euclidienne¹⁷ mesurée à partir des valeurs F_3-F_2 , F_2-F_1 et F_1-F_0 avec chaque valeur F_i convertie sur une échelle auditive. C'est l'approche qui a été adoptée dans l'étude sur les voyelles moyennes que je rapporte en §5.3.3 ci-après. Nous avons depuis affiné le modèle, en nous basant sur la notion de second formant effectif (F'_2) (voir Bladon & Fant 1978 ; Fant 1973) qui s'appuie sur la théorie de l'intégration spectrale selon laquelle deux formants sont intégrés s'ils sont situés dans une certaine distance critique l'un de l'autre (3 à 3,5 bark¹⁸, voir Beddor 1991). F'_2 est souvent utilisé en modélisation computationnelle des systèmes vocaliques pour représenter la contribution de F_2 , F_3 , et éventuellement F_4 à la perception des voyelles (de Boer 2001 ; Oudeyer 2005 ; Schwartz et al. 1997). Hermansky & Broad (1989) ont par ailleurs suggéré que F'_2 était corrélé avec la taille et la forme de la cavité antérieure du conduit vocal.

Pour estimer F'_2 , nous avons utilisé une formule donnée par Fant (1973 : 52) qui ne repose que sur les trois premiers formants (F_1 , F_2 , F_3) :

16. Il est tout à fait possible d'utiliser des n -grammes d'unités autres que des segments, par exemple des syllabe ou des mots (cf. Surendran & Niyogi 2003, 2006).

17. La distance euclidienne correspond à la plus courte distance en ligne droite entre deux points.

18. L'échelle bark est similaire à l'échelle ERB : elle est moins récente mais beaucoup plus utilisée.

$$(16) \quad F'_2 = f_2 + \frac{0,5 \times (f_3 - f_2) \times (f_2 - f_1)}{f_3 - f_1}$$

où f est la fréquence, exprimée en Hertz ou en mel. Notre approche se démarque légèrement de la formulation de Fant puisque nous avons adopté l'échelle bark, dans la mesure où il s'agit de l'échelle la plus utilisée à l'heure actuelle pour modéliser F'_2 (de Boer 2001 ; Schwartz et al. 1997).

Pour obtenir une représentation auditive de chaque voyelles, les trois premiers formants en Hertz ont donc d'abord été convertis en bark à l'aide de l'équation suivante (Traunmüller 1990) :

$$(17) \quad z = \frac{26.81}{1 + 1960/f} - 0.53$$

où f est la fréquence en Hertz et z est la transformée auditive en bark. Nous avons par la suite calculé le second formant effectif (z'_2) de chaque voyelle à l'aide de l'équation (16) ci-dessus. Par conséquent, chaque voyelle est représentée par une paire de valeur (z_1, z'_2).

Pour calculer la distance entre deux voyelles i et j , nous avons choisi de remplacer la distance euclidienne standard par une distance euclidienne pondérée (d'après de Boer 2001 : 49) :

$$(18) \quad \delta(i,j) = \sqrt{(z_1^i - z_1^j)^2 + \lambda(z_2^i - z_2^j)^2} \quad \text{avec } \lambda = 0.3$$

Le paramètre λ a pour effet d'atténuer l'influence du second formant effectif. Cette valeur est basée sur diverses observations typologiques et études perceptives (de Boer 2001 : 49-50 ; Schwartz et al. 1997) qui suggèrent que F'_2 contribue moins que F_1 à l'identification des voyelles.

5.3.2.3 Indice de distinctivité

La mesure que nous avons développée formalise l'intuition suivante : si deux segments ont un rendement fonctionnel élevé (au sens de Hockett) ou s'ils sont phonétiquement très différents (cf. [h] et [ŋ] en anglais), leur indice de distinctivité devrait être élevé, et il devrait l'être d'autant plus si les deux sont élevés ; si, au contraire, leur rendement fonctionnel est faible, parce qu'ils sont peu fréquents et/ou parce que leur distribution est proche d'une distribution non sécante, et s'ils sont phonétiquement très proches, alors leur distribution sera proche d'une distribution « allophonique » au sens classique et leur indice de distinctivité devrait être proche de zéro.

Pour ce faire, nous avons utilisé la moyenne géométrique, qui est une mesure de la tendance centrale, comme la moyenne arithmétique, mais qui contrairement à elle permet de moyennner des valeurs qui ne sont *pas mesurées sur la même échelle*. Formellement,

la moyenne géométrique correspond à la n ième racine du produit de n nombres. Dans notre cas, l'indice de distinctivité D de deux unités x et y correspond à la racine carrée du produit du rendement fonctionnel (λ) et de la distance phonétique (δ), soit :

$$(19) \quad D(x,y) = \sqrt{\lambda(x,y) \times \delta(x,y)}$$

La figure 5.8 illustre l'effet combiné du rendement fonctionnel et de la distance phonétique dans la détermination de l'indice de distinctivité.

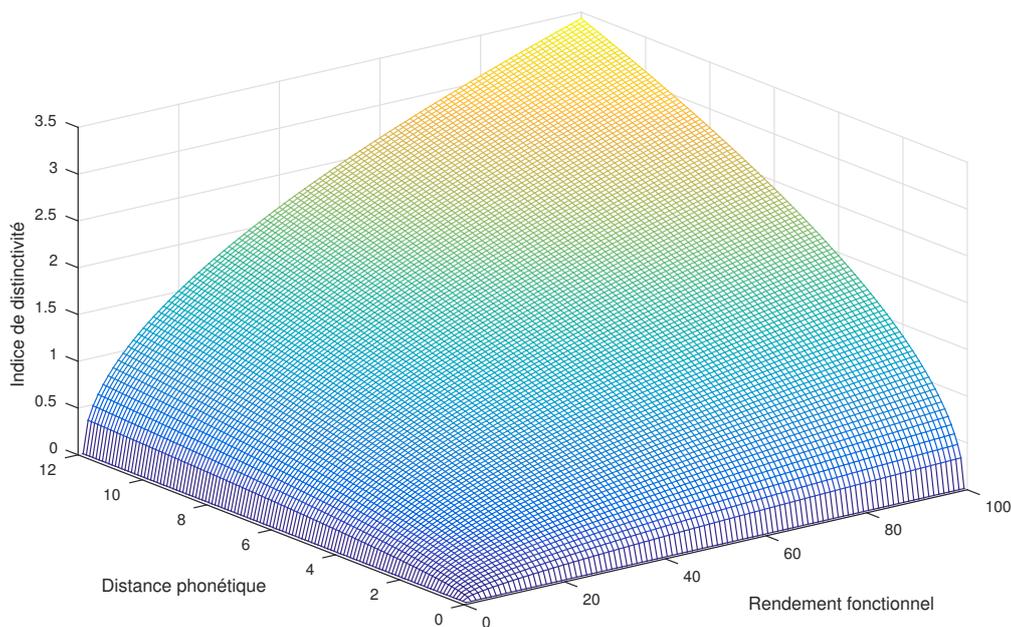


Figure 5.8. Indice de distinctivité

5.3.3 Les voyelles moyennes en français

Une version de ce modèle appliqué aux voyelles moyennes a été présentée à la *International Conference on Speech Sciences* qui s'est tenue à Séoul en 2015. Les principales conclusions sont consignées dans les actes du colloque (Eychenne 2015c)^[28], mais leur taille était malheureusement limitée à deux pages. Je présente donc ici ces résultats de manière plus détaillée.

Ce travail s'appuie sur l'étude préliminaire que j'avais présentée aux Journées FLO-RAL 2014, mais il prend en compte l'ensemble des voyelles orales (pas seulement les voyelles moyennes) et adopte la mesure de Hockett comme mesure de l'équivalence distributionnelle. L'étude vise donc à comparer l'indice de distinctivité relatif des voyelles moyennes en français méridional et septentrional. J'ai utilisé pour les données acoustiques

deux enquêtes PFC représentatives de ces variétés : l'une à Brunoy (à 20 km de Paris) et l'autre à Marseille.

5.3.3.1 Rendement fonctionnel

Il est nécessaire, pour calculer le rendement fonctionnel, de disposer d'un échantillon de parole conséquent. Pour ce faire, j'ai donc décidé d'utiliser la base LEXIQUE, que j'avais précédemment utilisée dans mon travail sur la phonotaxe (cf. §5.2). Toutefois, un examen attentif de la base a fait apparaître que la phonétisation du graphème *o* était problématique, et il était phonétisé comme [o] bien plus souvent qu'attendu. J'ai donc travaillé à partir d'un sous-ensemble de 32324 formes, ce qui correspond à l'ensemble des mots dont la fréquence était supérieure à 500 000. La phonétisation de *o* a été corrigée manuellement en prenant pour référence la prononciation donnée par le Petit Robert 2011. Bien que normatives, les phonétisations du Robert représentent une prononciation septentrionale standard et étaient à ce titre suffisantes pour les besoins de cette étude. Pour obtenir une prononciation correspondant au français méridional (commun), j'ai d'abord ajouté les schwas sur la base de la graphie et ai recodé les voyelles moyennes en leur appliquant la distribution attendue selon la loi de position. Dans les deux jeux de données (Nord et Sud), j'ai par ailleurs introduit 5% de bruit dans la distribution des voyelles moyennes : chaque voyelle mi-fermée avait 5% de chances d'être recodée en la voyelle mi-ouverte correspondante, et *vice versa*. Ceci avait pour but de simuler (dans une certaine mesure) la variation, en particulier dans le Sud où la loi de position, bien qu'elle soit d'application très générale, connaît des exceptions (Durand 2009 ; Eychemme 2014).

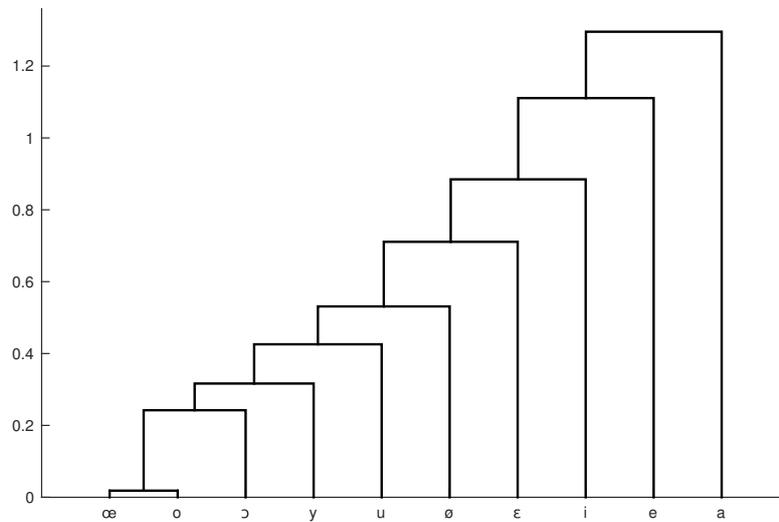
Les résultats numériques sont rapportés sous forme de matrice de distance et sont représentés visuellement sous forme de dendrogramme¹⁹. Comme on peut le voir pour le Nord à la figure 5.9 la paire de voyelle dont le rendement est le plus faible est [œ] / [o] ($\delta = 1,84$) suivie de [ø] et [œ] ($\delta = 2,90$), puis [e] et [œ] ($\delta = 2,47$). Le fait que la voyelle [œ] soit présente dans les paires qui ont le rendement les plus faibles s'explique sans doute par le fait qu'elle soit la voyelle la moins fréquente du français (Wioland 1985).

En ce qui concerne le Sud (figure 5.10), on voit que toutes les paires de voyelles mi-ouverte/mi-fermée ont un rendement très faible et que leurs distributions sont virtuellement non sécantes. La classification hiérarchique fait bien émerger trois paires de voyelles moyennes, mais le regroupement ([o ~ œ], [ø ~ ε], [e ~ O]) n'est pas celui qu'on attendrait sur la base de la loi de position. Nous nous trouvons donc dans la configuration décrite à la figure 5.7 (p. 117) : les propriétés distributionnelles ne sont pas suffisantes pour prédire le regroupement segments.

19. Les dendrogrammes ont été obtenus par classification hiérarchique de la matrice de distance. Il existe trois méthodes courantes pour ce type de classification : le saut minimum, qui se base sur la plus petite distance entre les membres de deux groupes ; le saut maximum, qui se base sur la plus grande distance (ce qui peut donner des résultats aberrants) ; le couplage des moyennes, qui se base sur la moyenne des distances. J'ai utilisé cette dernière méthode en partant de l'hypothèse que lorsque deux voyelles ont été regroupées, elles forment une catégorie dont aucun des membres n'a de statut privilégié.

	i	y	u	e	ε	ø	œ	o	ɔ	a
i		87.96	87.44	170.53	160.74	124.92	13.22	71.27	73.77	173.27
y			61.43	107.05	80.41	84.39	9.24	40.51	45.21	84.78
u				102.59	92.91	77.87	7.64	52.10	49.12	118.33
e					147.15	187.05	2.47	108.21	63.69	251.04
ε						84.49	17.24	75.70	75.80	181.75
ø							2.90	46.23	54.16	152.79
œ								1.84	14.19	13.65
o									34.26	105.37
ɔ										85.04
a										

(a) Matrice de distance

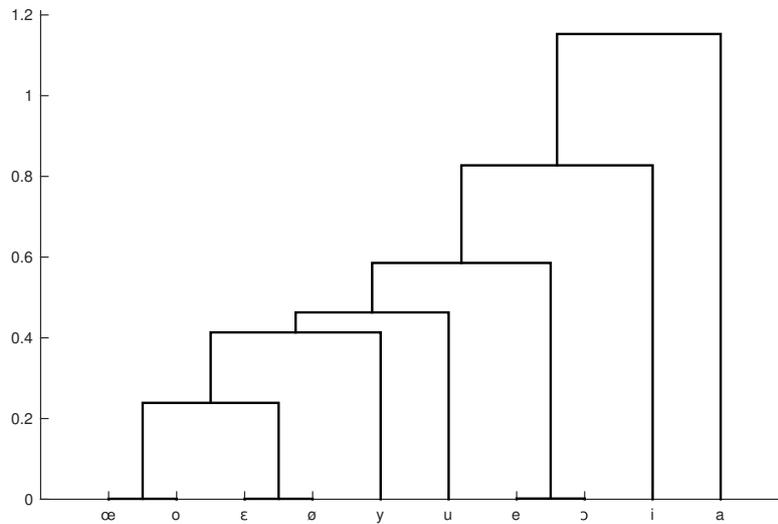


(b) Dendrogramme

Figure 5.9. Rendement fonctionnel dans le Nord ($\times 100$)

	i	y	u	e	ε	ø	œ	o	ɔ	a
i		85.97	85.53	204.75	32.92	126.04	9.56	102.35	14.80	170.24
y			59.33	122.55	19.28	85.69	4.77	55.67	11.47	83.23
u				122.13	19.61	76.27	8.51	67.80	14.03	118.72
e					0.12	215.46	0.11	167.94	0.17	298.95
ε						0.11	14.03	0.11	36.31	40.89
ø							0.11	81.28	0.18	152.77
œ								0.11	12.21	11.50
o									0.18	136.02
ɔ										25.23
a										

(a) Matrice de distance



(b) Dendrogramme

Figure 5.10. Rendement fonctionnel dans le Sud ($\times 100$)

5.3.3.2 Distance phonétique

Pour chacune des deux enquêtes, j'ai sélectionné deux hommes dont les fichiers sonores étaient de bonne qualité et qui me paraissaient représentatifs de la variété. Les conversations guidées et libres ont été alignées à l'aide d'EasyAlign (Goldman 2011) et l'alignement des voyelles orales a été corrigé manuellement. Pour chaque voyelle (à l'exception du schwa), j'ai calculé une valeur moyenne à partir de trois points (à 33, 50 et 67% de la voyelle) pour la F_0 , et les trois premiers formants. J'ai ainsi obtenu un total de 2584 voyelles (954 dans le nord et 1630 dans le sud). Les valeurs moyennes des formants (en Hertz) sont présentées au tableau 5.1.

		F1	F2	F3
i	N	344	2103	2913
	S	301	2014	3135
y	N	383	1903	2731
	S	293	1694	2529
u	N	390	1215	2732
	S	320	963	2673
e	N	418	1927	2807
	S	346	1845	2832
ɛ	N	424	1825	2745
	S	402	1715	2766
ø	N	385	1543	2645
	S	360	1351	2888
œ	N	478	1456	2584
	S	470	1283	2755
o	N	389	1308	2663
	S	387	1038	2732
ɔ	N	456	1270	2618
	S	409	965	2680
a	N	493	1630	2651
	S	524	1292	2645

Tableau 5.1. Valeurs formantiques moyennes (N = Nord, S = Sud)

Comme je l'ai brièvement mentionné en §5.3.2.2, la distance phonétique correspond pour chaque voyelle à la distance euclidienne entre les valeurs E_3-E_2 , E_2-E_1 et E_1-E_0 , chaque valeur E_i représentant la transformé du F_i correspondant sur l'échelle ERB. Cette approche était inspirée du modèle proposé par Syrdal & Gopal (1986).

La matrice de distance et la classification hiérarchique correspondante pour le Nord sont données à la figure 5.11 : comme on peut le voir, les paires de voyelles les plus proches

sont [i] et [y] d'une part, et [ø] et [œ] d'autre part. On voit également que les voyelles arrondies (à l'exception de [y]) forment un groupe, qui est également aggloméré avec [a]. Les voyelles antérieures hautes et antérieures moyennes non arrondies forment un autre groupe.

La figure 5.12 présente les mêmes données pour le Sud. Bien que [i] et [y] soient là aussi les voyelles les plus proches, et qu'elles soient là encore proches de [e], le regroupement des autres voyelles est assez différent de la situation observée à Brunoy. Le point le plus saillant est la proximité de [a] et [œ] : ceci s'explique sans doute par le fait que, d'une part, les voyelles mi-ouvertes dans le Sud sont généralement plus ouvertes que dans le Nord (Nguyen & Fagyal 2008 : 8) et, d'autre part, par le fait que [œ] se trouve fréquemment avant un [ʁ] de coda qui a un effet ouvrant : sa réalisation est donc souvent [œ]. On voit par ailleurs que le regroupement ne correspond en rien à ce que l'on pourrait attendre de la distribution due à la loi de position.

5.3.3.3 Distinctivité

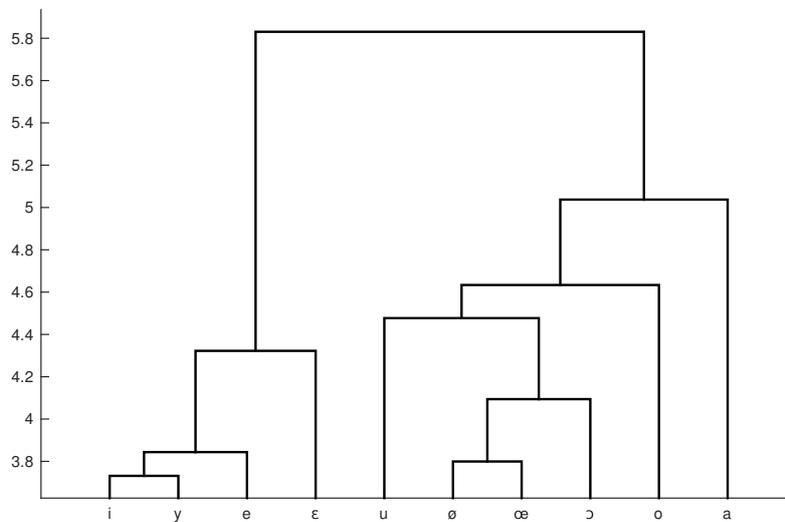
Nous en venons maintenant à la discussion de l'indice de distinctivité, qui intègre le rendement fonctionnel (mesuré comme le taux de perte d'information) et la distance phonétique. Nous voyons à la figure 5.13 que la classification des voyelles est assez similaire au classement obtenu pour le rendement fonctionnel (figure 5.9), à ceci près que [ø] est plus près de [œ] et [o] sous l'effet de la similarité phonétique. De manière intéressante, aucun regroupement n'émerge, mis à part peut-être [œ] / [o], et la classification semble comparable à ce que l'on attendrait d'un système à 10 voyelles (hormis schwa) où tous les segments sont distinctifs, soit /i y u e ø o ε œ ɔ a/.

Si l'on porte notre attention vers le Sud (figure 5.14), on voit que la situation est toute autre : les trois paires [e] ~ [ɛ], [ø] ~ [œ] et [o] ~ [ɔ] ont été regroupées en paires conformément à la loi de position. Aucun autre regroupement n'apparaît, ce qui semble compatible avec l'hypothèse traditionnellement admise selon laquelle les paires de voyelles moyennes représentent trois « phonèmes » distincts et toutes les autres voyelles sont également phonémiques, soit le système à 7 voyelles /i y u E Œ O a/ (hormis schwa).

Les résultats obtenus sur les voyelles moyennes me semblent encourageants puisqu'ils sont cohérents avec les descriptions traditionnelles des systèmes archétypiques pour le Nord et le Sud. La méthode devra bien sûr être affinée et appliquée sur des données plus réalistes, idéalement sur de très larges corpus alignés au niveau du phone, en calculant le rendement fonctionnel sur le corpus *oral* plutôt que sur un corpus *dictionnaire* comme je l'ai fait ici. D'un point de vue théorique, cette approche invite à repenser les notions d'allophonie et de distinctivité comme un continuum plutôt que comme des notions discrètes. D'un point de vue pratique, une approche inspirée de cette méthode pourrait sans doute permettre de dégager les inventaires vocaliques de manière inductive, ce qui pourrait avoir des applications utiles dans l'étude de la variation (diatopique en particulier).

	i	y	u	e	ɛ	ø	œ	o	ɔ	a
i		3.73	7.26	3.87	4.55	5.04	6.61	6.94	7.32	6.52
y			6.33	3.81	4.36	4.35	5.73	6.39	6.53	6.04
u				6.24	5.82	4.28	4.57	4.70	4.58	5.83
e					4.06	4.21	5.53	6.20	6.29	5.70
ɛ						4.09	4.88	5.45	5.53	4.93
ø							3.80	4.69	4.31	4.75
œ								4.63	3.88	4.56
o									4.51	5.28
ɔ										4.77
a										

(a) Matrice de distance

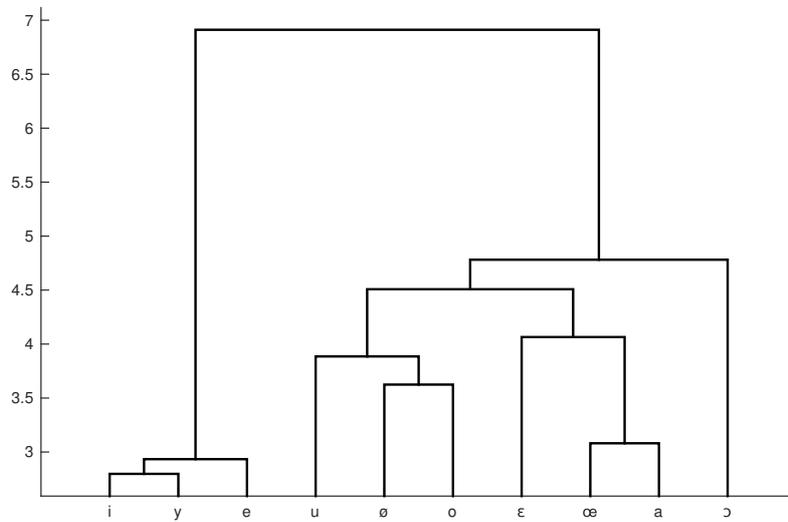


(b) Dendrogramme

Figure 5.11. Distance phonétique dans le Nord

	i	y	u	e	ɛ	ø	œ	o	ɔ	a
i		2.8	8.21	3.03	5.22	5.58	7.32	8.14	10.23	8.22
y			7.55	2.84	4.55	4.95	6.48	7.41	9.49	7.39
u				7.30	6.03	4.14	4.45	3.63	4.55	5.47
e					4.05	4.62	5.85	6.99	8.89	6.70
ɛ						3.79	3.70	5.25	6.53	4.43
ø							3.22	3.62	5.30	4.54
œ								3.40	4.08	3.08
o									3.79	4.42
ɔ										4.43
a										

(a) Matrice de distance

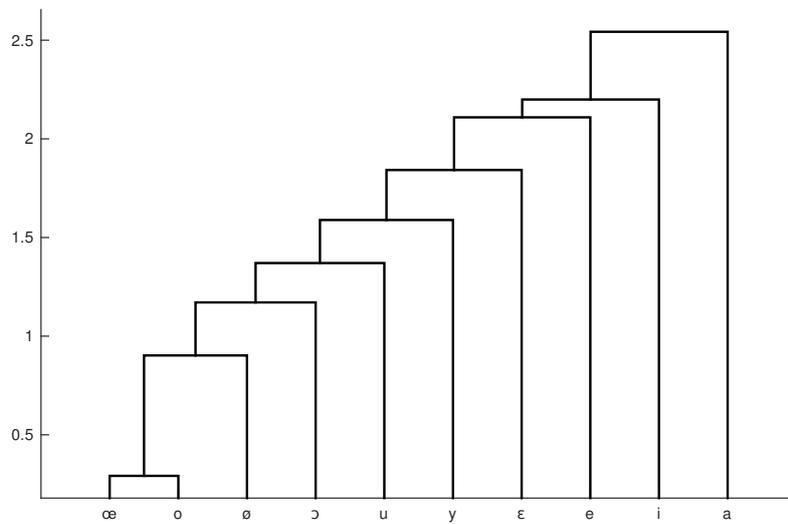


(b) Dendrogramme

Figure 5.12. Distance phonétique dans le Sud

	i	y	u	e	ɛ	ø	œ	o	ɔ	a
i		181.17	252.02	257.04	270.32	250.81	93.47	222.36	232.37	336.05
y			197.25	202.05	187.29	191.59	72.74	160.90	171.86	226.31
u				252.99	232.57	182.61	59.06	156.51	150.04	262.72
e					244.40	280.77	36.99	259.00	200.15	378.53
ɛ						185.79	91.69	203.19	204.70	299.39
ø							33.22	147.23	152.74	269.41
œ								29.19	74.22	78.88
o									124.34	235.76
ɔ										201.41
a										

(a) Matrice de distance ($\times 100$)

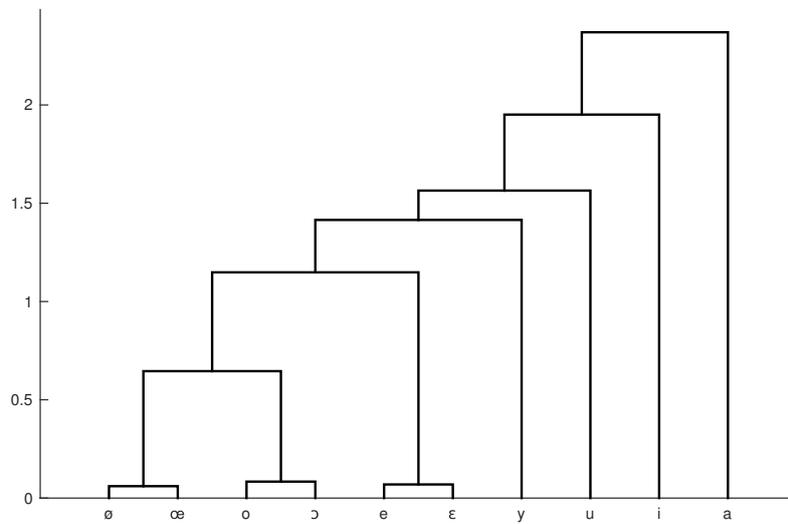


(b) Dendrogramme

Figure 5.13. Indice de distinctivité dans le Nord

	i	y	u	e	ɛ	ø	œ	o	ɔ	a
i		155.08	265.03	249.08	131.12	265.13	83.63	288.66	123.07	374.13
y			211.69	186.47	93.64	206.01	55.58	203.04	104.32	248.01
u				298.67	108.71	177.65	61.57	156.94	79.86	254.88
e					6.93	315.39	8.15	342.52	12.38	447.58
ɛ						6.56	72.08	7.72	154.02	134.51
ø							6.05	171.64	9.83	263.27
œ								6.21	70.61	59.53
o									8.37	245.25
ɔ										105.78
a										

(a) Matrice de distance ($\times 100$)



(b) Dendrogramme

Figure 5.14. Indice de distinctivité dans le Sud

5.3.4 Fusion de /e/ et /ɛ/ en coréen

La méthode telle que je l'ai illustrée sur les voyelles orales du français permet de rendre compte de l'allophonie des voyelles moyennes en français méridional : dans ce modèle, les voyelles « allophoniques » sont des voyelles dont l'indice de distinctivité est le plus faible. On se souvient par ailleurs que la notion de rendement fonctionnel a été souvent évoquée comme étant un facteur possible de changement diachronique (Martinet 1955) : dans une perspective fonctionnaliste, une paire de segments dont le rendement fonctionnel est faible serait plus susceptible de fusionner puisqu'elle serait peu *utile* dans le système. L'exemple souvent donné est la fusion de $\tilde{e} \sim \tilde{\epsilon}$ ²⁰ en français, qui ne distinguent que très peu d'items (*brun* vs *brin*, *emprunte* vs *empreinte*). Or, comme je l'ai évoqué plus haut, nous avons analysé le rendement fonctionnel de $e \sim \epsilon$ et avons observé que, bien qu'il fût parmi les plus bas des paires de voyelles du coréen, son rendement fonctionnel ne permettait pas de prédire à lui seul la fusion de cette paire. Nous avons donc tenté de vérifier l'hypothèse selon laquelle ce n'est pas le rendement fonctionnel, mais la *distinctivité* au sens où je l'ai définie qui permet de prédire la directionnalité d'une fusion de segments. Dans cette perspective, une opposition est d'autant plus susceptible d'être perdue qu'elle contribue peu à l'entropie du système et qu'elle est peu saillante d'un point de vue perceptif.

Je ne reprendrai pas l'analyse dans le détail car elle est assez longue, et je renvoie le lecteur à Eychenne & Jang (2016)^[31], disponible à la fin du volume de travaux. En plus de tester notre méthode, nous comparons la mesure du rendement fonctionnel adoptée ici (perte d'entropie calculée sur des trigrammes de phones) à deux autres mesures, à savoir la perte d'entropie sur des unigrammes de mots et le nombre de paires minimales au sein d'une catégorie grammaticale. En effet, dans de récents travaux typologiques basés sur 8 langues, Wedel et al. (2013a,b) ont suggéré qu'il y avait une corrélation statistique entre rendement fonctionnel et tendance à la fusion, et que cette corrélation était la plus manifeste lorsque le rendement fonctionnel était calculé comme le nombre de paires minimales qu'une opposition permet de distinguer parmi des lemmes de même catégorie. Le fait que la catégorie syntaxique soit un bon prédicteur n'est pas surprenant puisqu'on s'attend à ce qu'une paire minimale soit plus importante si les deux mots peuvent commuter dans la chaîne parlée. Le fait que le nombre de paires minimales soit un meilleur prédicteur que l'entropie nous a paru plus surprenant. Nous avons montré qu'aucune des trois mesures du rendement fonctionnel ne prédisait à elle seule la fusion de $e \sim \epsilon$ et que la notion de distinctivité telle que définie était la seule, parmi celles que nous avons testées, qui était à même de rendre compte de cette fusion.

Je me contenterai ici de donner les principaux résultats de notre travail et renvoie à nouveau le lecteur à Eychenne & Jang (2016)^[31] pour plus de détails.

5.3.4.1 Rendement fonctionnel

Les données utilisées pour calculer le rendement fonctionnel, mesuré comme la perte d'entropie sur des trigrammes de phones, sont le corpus Sejong que nous avons phonétisé

20. J'utilise ici la notation $a \sim b$ pour noter la fusion des « phonèmes » a et b .

par application de règles. Cette méthode n’est pas parfaite, mais le corpus K-SPAN (cf. §3.3) n’était pas encore disponible lorsque nous avons réalisé ce travail. Nous avons utilisé un ensemble de textes du XX^e siècle contenant au total 2 734 643 types, qui correspondent à 33 912 209 tokens.

La matrice de distance est donnée au tableau 5.2. Je donne également à la figure 5.15 une représentation en deux dimensions de la matrice de distance, obtenue par positionnement multidimensionnel (*multidimensional scaling*, MDS), une technique souvent utilisée pour réduire le nombre de dimensions d’un ensemble de données. En l’occurrence, les axes horizontal et vertical représentent les deux dimensions qui représentent au mieux (mais de manière imparfaite) la distance entre chaque point (l’orientation est arbitraire).

Comme l’ont fait observer Oh et al. (2015), la distribution du rendement fonctionnel ne suit pas une loi gaussienne, et un petit nombre d’oppositions ont rendement particulièrement élevé. La médiane ($\tilde{x} = 0,247\%$) est donc un meilleur indicateur de la tendance centrale que la moyenne. La paire dont le rendement est le plus élevé est $\Lambda \sim \alpha$ (0,708%), alors que celle qui a le rendement le plus faible est $\text{u} \sim \varepsilon$ (0,071%). L’opposition qui nous intéresse, à savoir $e \sim \varepsilon$, est l’avant-dernière paire de voyelle (0,089%). Bien que le rendement soit effectivement très faible, comme on s’y attend, ce facteur prédirait à lui seul que c’est l’opposition $\text{u} \sim \varepsilon$, et non $e \sim \varepsilon$, qui aurait dû fusionner.

	i	ɯ	u	e	Λ	o	ε	α
i		0.392	0.351	0.371	0.366	0.354	0.170	0.597
ɯ			0.108	0.127	0.172	0.167	0.071	0.302
u				0.116	0.298	0.279	0.128	0.442
e					0.123	0.192	0.089	0.224
Λ						0.314	0.153	0.708
o							0.176	0.515
ε								0.355
α								

Tableau 5.2. Rendement fonctionnel des monophthongues du coréen

5.3.4.2 Distance phonétique

La fusion de $e \sim \varepsilon$ a commencé vers la fin du XIX^e/début du XX^e siècle. Il est donc malheureusement impossible d’obtenir des données acoustiques représentant un état de langue antérieur à la fusion. Nous nous sommes donc résolus à utiliser des données acoustiques de locuteurs qui présentent toujours l’opposition. Heureusement, les données de Yang (1992, 1996) fournissent les valeurs de F_1 , F_2 et F_3 de 10 hommes et 10 femmes, tous locuteurs de la variété séoulite. Les valeurs rapportées par Yang pour F_1 (l’écart-type est comme d’habitude donné entre parenthèses) sont les suivantes : pour les hommes, $/e/ = 490$ (105) Hz et $/\varepsilon/ = 591$ (75) Hz ; pour les femmes, $/e/ = 650$ (113) Hz and

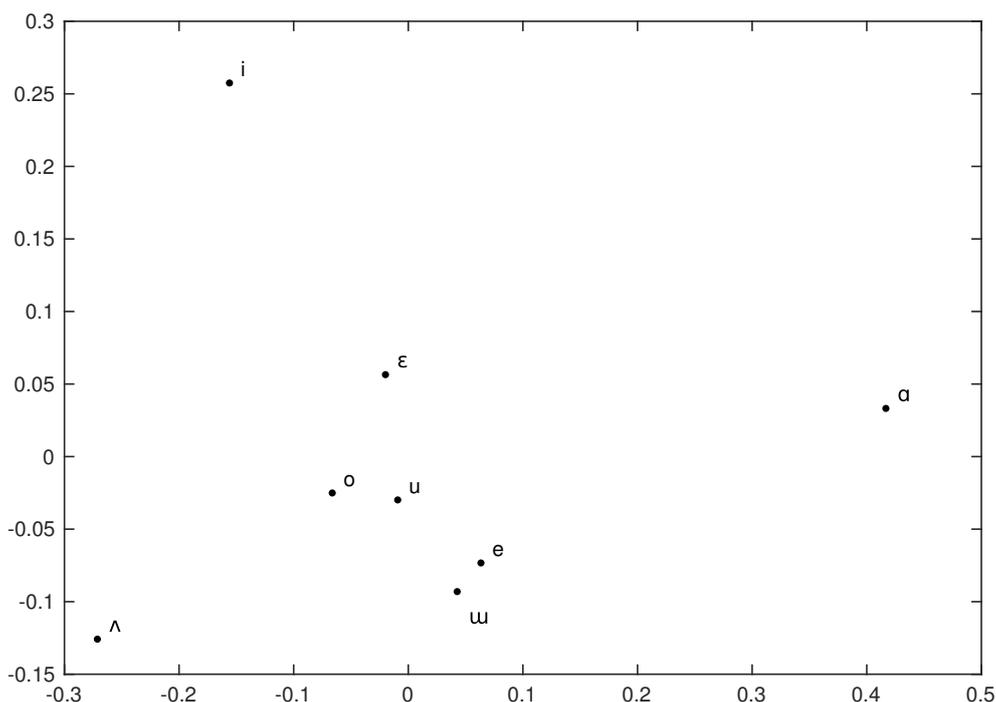


Figure 5.15. Carte MDS pour le rendement fonctionnel des monophthongues du coréen

/ε/ = 677 (108) Hz. Nous n'avons utilisé que les données des hommes car comme on le voit l'opposition est beaucoup moins nette dans les données des femmes, ce qui suggère que la fusion était dans un état plus avancé pour elles. Les données pour l'ensemble des monophthongues des hommes sont reprises au tableau 5.3.

La distance phonétique a été calculée à l'aide du modèle présenté en §5.3.2.2. La valeur médiane est $\tilde{x} = 2,12$ bark, avec des valeurs allant de 0,81 bark pour $u \sim o$ à 3,76 bark pour $i \sim a$. Les paires $u \sim a$ et $i \sim u$ font également partie des paires les plus distantes d'après le modèle. La figure 5.16 fournit le diagramme MDS correspondant à la matrice de distance (les axes ont été pivotés pour plus de lisibilité).

La paire $e \sim \varepsilon$ est l'avant dernière paire (0,87 bark) après $o \sim u$. Le fait que /o/ et /u/ soient les plus proches est intéressant puisque cela a été noté dans la littérature : Hong (1988 : 61-67) identifie ainsi $o \sim u$ comme une « fausse fusion » (*false merger*). Les résultats de son étude sociolinguistique à Séoul montrent que ces voyelles sont acoustiquement très proches, mais que ses sujets étaient néanmoins capables de les distinguer dans des tests de paires minimales (ex : [ori] « canard » vs [uri] « nous »). D'autres travaux plus récents montrent que la distance entre ces voyelles s'est réduite, particulièrement chez les femmes (Han & Kang 2013). En outre, il faut noter qu'un nombre conséquent de mots qui avaient originellement un /o/ ont maintenant une variante libre avec /u/,

	F ₁ (SD)	F ₂ (SD)	F ₃ (SD)
ɑ	738 (87)	1372 (124)	2573 (127)
ε	591 (75)	1849 (106)	2597 (110)
e	490 (105)	1968 (150)	2644 (94)
i	341 (29)	2219 (176)	3047 (146)
o	453 (47)	945 (134)	2674 (156)
(ø)	459 (69)	1817 (163)	2468 (134)
u	369 (43)	981 (141)	2565 (173)
(y)	338 (30)	2114 (140)	2729 (213)
ʌ	608 (76)	1121 (110)	2683 (145)
ɯ	405 (37)	1488 (176)	2497 (80)

Tableau 5.3. F₁, F₂, F₃ pour les locuteurs hommes de Séoul (Yang 1996 : 251)

	i	ɯ	u	e	ʌ	o	ε	ɑ
i		1.44	2.65	1.47	3.34	3.03	2.34	3.76
ɯ			1.38	1.13	2.04	1.6	1.74	2.77
u				2.43	2.12	0.81	2.77	3.22
e					2.11	2.38	0.87	2.31
ʌ						1.41	1.67	1.18
o							2.46	2.57
ε								1.51
ɑ								

Tableau 5.4. Distance perceptive des monophthongues du coréen (en bark)

particulièrement en syllabe finale : [samt^hon] ~ [samt^hun] « oncle » ; [puɕzo] ~ [puɕzu] « argent reçu en cadeau ». Cette situation s'apparente à ce que Labov (1994 : 321) appelle FUSION PAR TRANSFERT, un cas de figure où le remplacement d'une voyelle par une autre se propage dans le vocabulaire.

La faible différence entre les paires o ~ u et e ~ ε appelle deux remarques. Tout d'abord, les données acoustiques que nous avons utilisées représentent un état de langue où la fusion est déjà à un stade avancé : puisque la plupart des auteurs considèrent que le résultat de la fusion est une voyelle intermédiaire [ɛ̞] (Shin et al. 2013 : 99-101), ce que Labov (1994 : 321) appelle une FUSION PAR APPROXIMATION, il est fort probable que la distance entre ces deux voyelles ait été plus grande dans le passé. Par ailleurs, même si ce n'était pas le cas, un modèle qui ne repose que sur la confusion perceptive prédirait que la fusion de o ~ u devrait être au moins aussi avancée que celle de e ~ ε. Or ce n'est pas le cas et comme nous venons de le voir, o ~ u suit une trajectoire évolutive différente

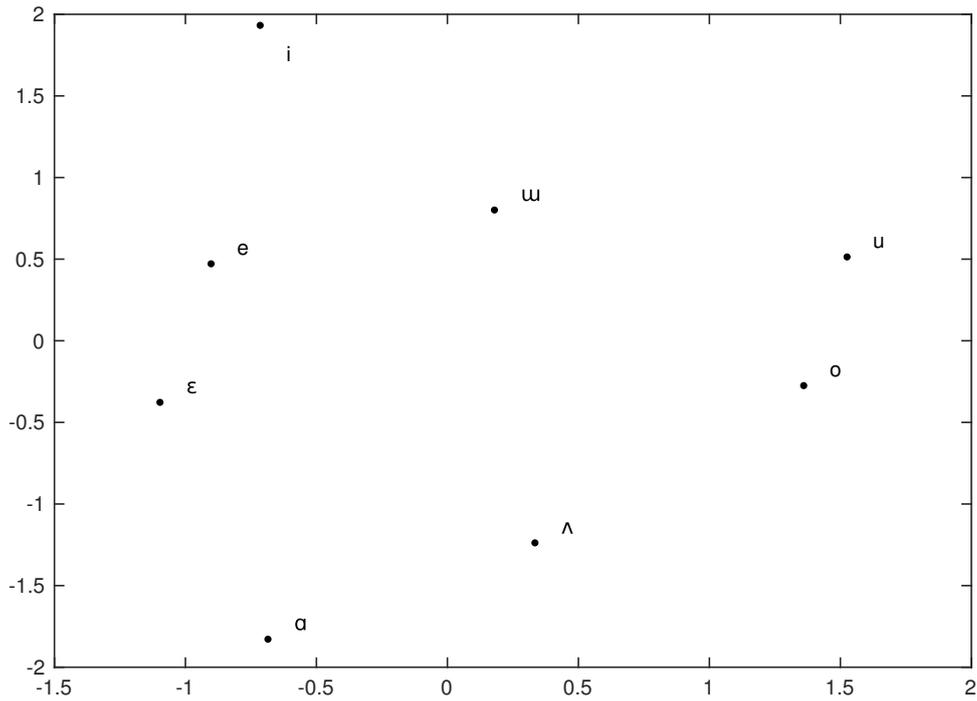


Figure 5.16. Carte MDS perceptive des monophthongues du coréen

puisque'il s'agit vraisemblablement d'une fusion par transfert et non pas d'une fusion par approximation.

5.3.4.3 Indice de distinctivité

Nous avons vu que ni le rendement fonctionnel, ni la distance phonétique ne permettent de prédire à eux seuls la fusion de $e \sim \varepsilon$. Tous deux suggèrent toutefois que cette opposition est relativement faible, mais aucun des deux ne semble déterminant. Il est donc opportun d'examiner leur effet combiné tel qu'il est prédit par l'indice de distinctivité. Les résultats sont présentés au tableau 5.5 et le diagramme correspondant est donné à la figure 5.17.

Examinons tout d'abord les oppositions les plus fortes selon cette mesure. Les paires $i \sim a$ (1,498), $u \sim a$ (1,192) et $o \sim a$ (1,151) sont les oppositions les plus saillantes en ce sens qu'elles ont à la fois un rendement élevé et sont perceptivement distinctes. À l'autre bout du spectre, nous voyons maintenant que $e \sim \varepsilon$ est la paire dont l'indice de distinctivité est le plus faible, puisque comme on l'a vu son rendement est faible et que les deux voyelles sont perceptivement très proches. Les deux autres paires dont le rendement est faible sont $u \sim \varepsilon$ (0,352) et $u \sim e$ (0,379). La paire $o \sim u$ a elle aussi un rendement assez faible (0,467), ce qui est largement dû à la proximité phonétique de /o/ et /u/. Son

	i	ɯ	u	e	ʌ	o	ɛ	ɑ
i		0.751	0.964	0.739	1.106	1.036	0.631	1.498
ɯ			0.385	0.379	0.592	0.517	0.352	0.914
u				0.532	0.795	0.467	0.596	1.192
e					0.510	0.676	0.279	0.719
ʌ						0.665	0.506	0.913
o							0.657	1.151
ɛ								0.732
ɑ								

Tableau 5.5. Distinctivité des monophthongues du coréen

indice de distinctivité est toutefois supérieur à celui de $e \sim \varepsilon$.

Nous voyons donc que la prédiction de l'indice de distinctivité concorde avec la situation observée en coréen : c'est la paire de voyelles dont l'indice est le plus faible qui a fusionné. L'approche devra bien sûr être testée sur d'autres langues pour pouvoir être éprouvée, et elle devra sans doute être affinée : nous avons par exemple montré dans Eychenne & Jang (2016)^[31] comment cette méthode pouvait être généralisée afin d'inclure d'autres dimensions, ou même pour accorder un poids différent aux différentes dimensions si cela s'avérait justifié d'un point de vue théorique ou empirique. Néanmoins, les résultats obtenus me semblent prometteurs. Toutes les fusions dont j'ai connaissance mettent en jeu des unités phonétiquement très proches, que ce soit $w \sim \mathfrak{w}$ et $k \sim x$ en anglais écossais (Scobbie & Stuart-Smith 2008), $\tilde{\varepsilon} \sim \tilde{\alpha}$ en français (Martinet 1955 : 35) ou encore $j \sim \mathfrak{k}$ en espagnol (Navarro 1967), pour ne prendre que quelques exemples parmi tant d'autres. Les auteurs tels que King (1967) qui ont rejeté le rôle du rendement fonctionnel dans les fusions sur la base de résultats négatifs n'ont sans doute considéré que la moitié du problème, puisqu'ils n'ont pas pris en compte la distance phonétique comme cela a été fait dans ce travail.

5.4 Conclusion

J'ai présenté les trois types de modèles quantitatifs que j'ai explorés au cours des cinq dernières années : les réseaux calculatoires dynamiques, qui permettent de modéliser la structure syllabique et métrique, le modèle phonotactique de Hayes et Wilson basé sur le principe de maximisation de l'entropie, qui permet d'induire des contraintes phonotactiques à partir d'exemples *positifs*, et un modèle de la distinctivité que j'ai développé en collaboration avec Tae Yeoub Jang. Bien qu'ils partagent certains traits (les deux derniers modèles reposent par exemple sur la théorie de l'information), ces trois modèles ne forment pas un tout cohérent et ne s'inscrivent donc pas dans une vision globale de la

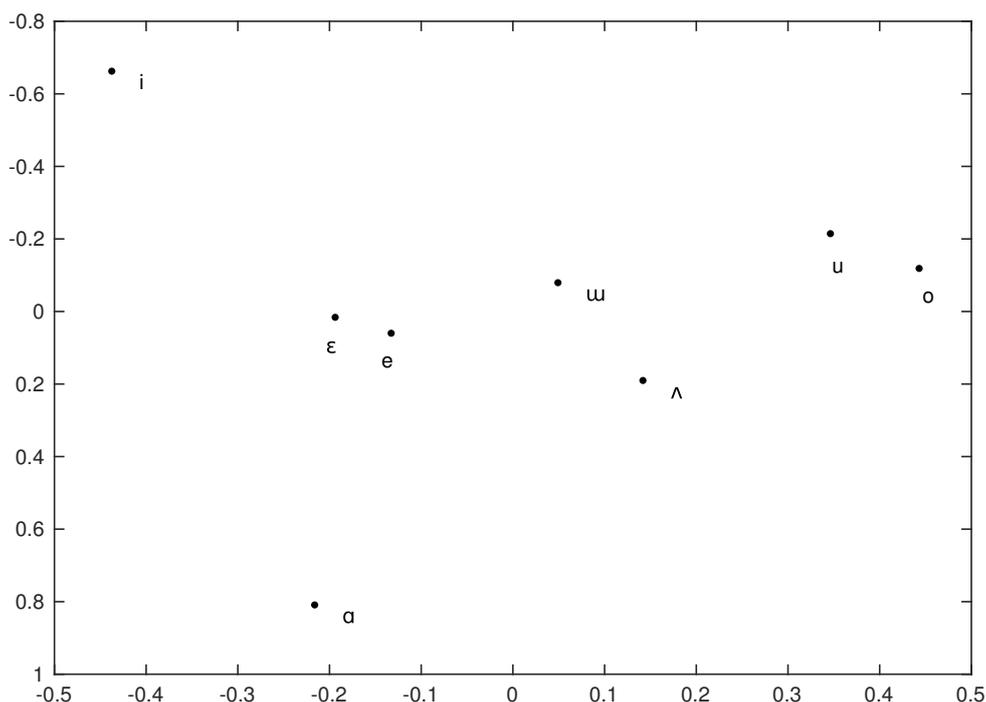


Figure 5.17. Carte MDS de la distinctivité des monophthongues du coréen

grammaire ; ils représentent plutôt trois tentatives, exploratoires, d'appliquer des modèles quantitatifs à l'analyse de tel ou tel problème. Les résultats obtenus, pris ensemble, ont renforcé en moi la conviction que les méthodes quantitatives étaient la manière la plus adéquate de modéliser la gradience grammaticale.

Leur potentiel commence à peine à être reconnu en phonologie, comme on le voit avec le succès du modèle de Hayes et Wilson, la ré-émergence de modèles phonologiques connexionnistes (par exemple Alderete et al. 2013) ou le regain d'intérêt pour la Grammaire Harmonique, restée en dormance pendant plus d'une quinzaine d'années, et qui fait maintenant l'objet d'intenses recherches sous l'impulsion de Paul Smolensky, qui en est à l'origine, mais aussi et surtout de Joe Pater et de ses collaborateurs (*inter alia* Potts et al. 2010 ; Pater 2009, 2014). L'histoire récente de la phonologie générative m'a fait prendre pleinement conscience de ce que la qualité d'une théorie phonologique ne se mesure pas à sa popularité. Le plein potentiel des RCD n'a ainsi à mon sens pas été exploité, et le fait qu'ils soient tombés en « désuétude » doit sans doute beaucoup plus au fait que Gary Larson n'ait pas poursuivi une carrière de phonologue et que John Goldsmith ait tourné son attention vers d'autres types de problèmes (en particulier l'apprentissage non supervisé de la morphologie), plutôt qu'à de véritables critères scientifiques objectifs. Le seul examen critique des RCD dont j'ai connaissance est Prince (1993), qui a notamment

montré que leur capacité générative était trop forte et qu'ils prédisaient des systèmes métriques apparemment non attestés, mais je doute que cela soit suffisant pour expliquer ce désintérêt collectif (le problème de surgénération affecte aussi bien OT que la GH, et ces cadres font malgré tout l'objet de recherches très actives).

Chapitre 6

Apprenabilité des langues naturelles

Les quatre précédents chapitres ont offert une synthèse de mes travaux selon les trois axes (empirique, ingénierique et théorique) structurants de mes recherches. Je voudrais offrir ici, en guise d'épilogue, quelques remarques sur l'apprenabilité des langues naturelles, et j'aborderai pour ce faire la complexité formelle des grammaires et de la phonologie.

6.1 Complexité formelle des grammaires

La contribution scientifique la plus fondamentale de Noam Chomsky est peut-être son travail sur la complexité des grammaires (Chomsky 1956, 1957), qui a jeté les bases de ce qu'il est maintenant convenu d'appeler la HIÉRARCHIE DE CHOMSKY. Il me semble utile d'en rappeler ici les idées essentielles car cela permet de mieux mettre en perspective la complexité formelle de la composante phonologique. Par ailleurs, bien que l'importance de ces notions soit bien établie en traitement automatique du langage (Heinz et al. 2015 ; Jurafsky & Martin 2008) et en informatique, elles n'ont curieusement qu'un rôle marginal en théorie phonologique (mais voir ci-après).

6.1.1 Hiérarchie de Chomsky « classique »

La hiérarchie de Chomsky est généralement présentée sous la forme de quatre niveaux imbriqués, numérotés de 0 à 3 (voir figure 6.1). Le niveau 0 de la hiérarchie englobe les langages dits RÉCURSIVEMENT ÉNUMÉRABLES (*recursively enumerable* en anglais), soit l'ensemble des langages calculables, sans aucune restriction. Les niveaux suivants correspondent à des ensembles de grammaires/langages de plus en plus contraints, appelés respectivement langages CONTEXTUELS (*context sensitive*), langages ALGÈBRIQUES (ou parfois « non contextuels », *context free* en anglais) et les langages RATIONNELS (ou encore « réguliers », par calque de l'anglais *regular*). Dans ce contexte, un LANGAGE est un ensemble (le plus souvent infini) de CHAÎNES, chaque chaîne étant une séquence de symboles provenant de l'ensemble de symboles, ou ALPHABET, qui caractérise le langage. A chaque niveau de la hiérarchie, un langage peut être décrit par un certain type de GRAMMAIRE, telle que définie en (1) :

(1) Grammaire

Une grammaire est un quadruplet $G = (V, \Sigma, R, S)$ où :

1. V est un ensemble fini de symboles non terminaux,
2. Σ est un ensemble fini de symboles terminaux,
3. R est un ensemble fini de règles de réécriture,
4. $S \in V$ est le symbole de départ.

Le type de la grammaire est déterminé par la structure des règles de réécriture qu'elle autorise¹. En outre, chaque langage peut être identifié (on dit aussi « accepté » ou « reconnu ») par un type d'AUTOMATE particulier, c'est-à-dire par une machine abstraite capable d'effectuer un nombre défini d'opérations. Examinons d'abord le type de grammaire le plus restrictif, à savoir les langages rationnels.

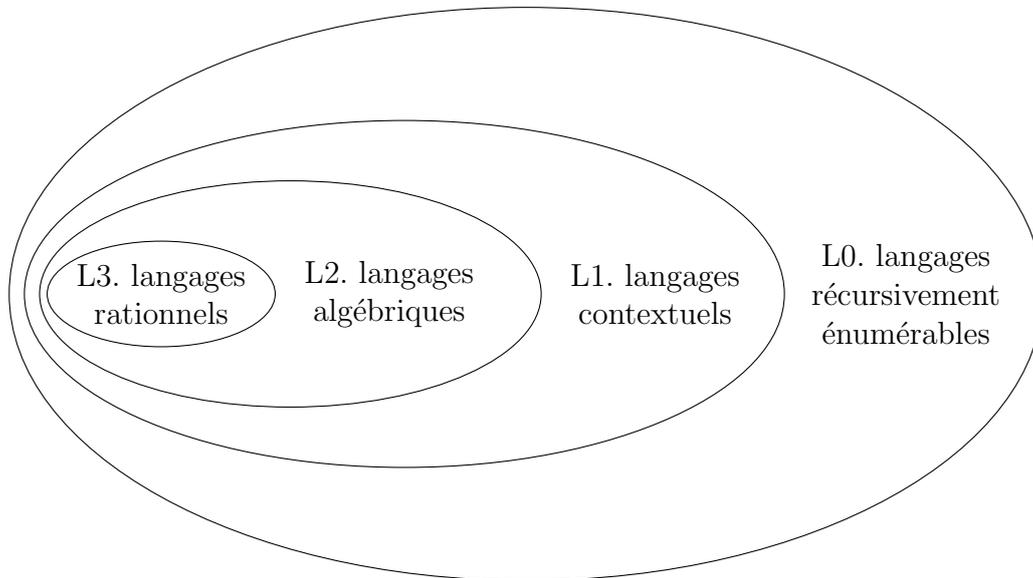


Figure 6.1. Hiérarchie de Chomsky

6.1.1.1 Langages rationnels (L3)

Un langage rationnel peut être décrit par une grammaire rationnelle, c'est-à-dire une grammaire dont les règles de réécriture peuvent prendre les formes suivantes.

1. Dans la suite de cette section, dans les règles de réécriture, une lettre minuscule (ex : a) désigne un symbole terminal, une lettre majuscule (ex : A) représente un symbole non terminal, et une lettre grecque (ex : α) représente une chaîne de symboles terminaux ou non terminaux.

(2) Règles d'une grammaire rationnelle

Une grammaire est rationnelle si et seulement si toutes ses règles de réécriture ont l'une des formes suivantes :

1. $Y \rightarrow y$ où Y est un symbole non terminal et y est un symbole terminal.
2. $Y \rightarrow \epsilon$ où Y est un symbole non terminal et ϵ la chaîne vide.
3. a. $Y \rightarrow Xy$
ou
b. $Y \rightarrow yZ$
où X, Y, Z sont des symboles non terminaux et y est terminal.

Le point important ici est que ce type de grammaire ne peut réécrire qu'un seul symbole non terminal en un symbole terminal (éventuellement vide), précédé ou suivi d'un seul symbole non terminal. Notons par ailleurs que toutes les règles d'une grammaire rationnelle doivent s'accorder sur le fait que le symbole non terminal précède (3a) ou suit (3b) le symbole terminal.

Comme cela a été mentionné plus haut, il est possible de construire un certain type d'automate capable de reconnaître les langages d'un niveau de la hiérarchie. On peut donc donner une définition algorithmique d'un langage rationnel : en l'occurrence, un langage est rationnel si toute chaîne appartenant à ce langage est acceptée par un AUTOMATE FINI (*finite state automaton*). La définition formelle d'un automate fini est donnée en (3) ci-dessous :

(3) Définition d'un automate fini

Un automate fini est un quintuplet $(Q, \Sigma, \delta, q_0, F)$ où :

1. Q est un ensemble fini d'états
2. Σ est un ensemble fini représentant un *alphabet*
3. $\delta : Q \times \Sigma \rightarrow Q$ une *fonction de transition*
4. $q_0 \in Q$ est l'état *initial*
5. $F \subseteq Q$ est un ensemble d'états *terminaux*

(d'après Sipser 2005 : 35)

Comme on le voit, cet automate est composé d'un ensemble d'états et de transitions entre ces états, la fonction de transition δ permettant de passer d'un état à un autre : elle prend en entrée un état et un symbole et produit en sortie un état. Parmi tous les états de l'automate, certains ont un statut particulier : il s'agit de l'état initial q_0 , qui est le point de départ de l'automate, et de l'ensemble d'états terminaux F , qui représente

l'ensemble de points d'arrivée *valides*. L'alphabet est l'ensemble des symboles qui peuvent être manipulés par l'automate.

Pour des raisons qui deviendront plus claires ci-après, il est utile d'imaginer que l'automate dispose d'une bande composée de cellules pouvant accueillir chacune un seul symbole. Toute chaîne peut donc être écrite sur cette bande; la chaîne $xzxz$ sera ainsi représentée sur la bande comme suit : $\boxed{x} \boxed{z} \boxed{x} \boxed{z}$. Pour tenter de reconnaître une chaîne, celle-ci est disposée sur la bande, et l'automate, en partant de l'état initial, lit un symbole sur la bande et essaie de se déplacer en suivant l'une des transitions partant de l'état actuel et correspondant au symbole lu. L'opération est répétée tant qu'il reste des symboles non lus et que l'automate peut se déplacer. La chaîne est reconnue si et seulement elle a été entièrement lue et l'automate s'est arrêté dans l'un des états terminaux au terme de la lecture.

Pour illustrer ce type d'automate, prenons l'exemple, inspiré de Heinz (2009), du système accentuel de l'asmat (langue papoue de Nouvelle-Guinée). Cette langue possède un système accentuel très simple : il s'agit d'une langue à accent fixe sur la finale, avec propagation vers la gauche d'une alternance binaire de syllabes accentuées et inaccentuées (autrement dit, un système de pieds iambiques alignés à droite). Ce système possède trois types de syllabes : syllabe portant un accent primaire, syllabe portant un accent secondaire et syllabe inaccentuée : on peut donc le représenter à l'aide de l'alphabet $\Sigma = \{\acute{\sigma}, \grave{\sigma}, \sigma\}$ où, comme il est d'usage, σ représente une syllabe non accentuée, $\acute{\sigma}$ une syllabe portant un accent primaire, et $\grave{\sigma}$ une syllabe portant un accent secondaire. A l'aide de cet alphabet, il est possible de décrire ce système comme un langage, chaque patron accentuel correspondant à une chaîne du langage :

$$(4) \quad \{\acute{\sigma}, \sigma\acute{\sigma}, \grave{\sigma}\sigma\acute{\sigma}, \sigma\grave{\sigma}\sigma\acute{\sigma}, \grave{\sigma}\sigma\grave{\sigma}\sigma\acute{\sigma}, \sigma\grave{\sigma}\sigma\grave{\sigma}\sigma\acute{\sigma}, \dots\}$$

La figure 6.2 présente un automate fini qui modélise la grammaire accentuelle correspondant à ce langage. On pourra aisément se convaincre de ce que cet automate reconnaît effectivement le système accentuel de l'asmat : tout chemin partant de q_0 et aboutissant à q_1 correspond bien à une chaîne du langage. Par exemple, la séquence d'états $(q_0, q_2, q_3, q_2, q_3, q_1)$ correspond à la chaîne $\grave{\sigma}\sigma\grave{\sigma}\sigma\acute{\sigma}$. Cette chaîne pourrait par exemple être formalisée, de manière équivalente, par une grille rythmique (Prince 1983), comme le montre la figure 6.3 (voir également la discussion en §4.2.3 p. 84). Il est possible d'argumenter en faveur de l'une ou l'autre représentation, mais ces deux formalismes sont *a minima* faiblement équivalents : il est possible de traduire toute grille rythmique du type 6.3 en un automate fini du type 6.2, et *vice versa*.

6.1.1.2 Langages algébriques (L2)

Un langage algébrique peut être décrit par une grammaire dont les règles de réécriture ont la forme $A \rightarrow \alpha$, et permettent donc de réécrire tout symbole non-terminal en une chaîne (éventuellement vide) de symboles terminaux et/ou non terminaux. Ces règles sont

2. Par convention, les états terminaux d'un automate fini sont représentés par un double cercle (ici, q_1).

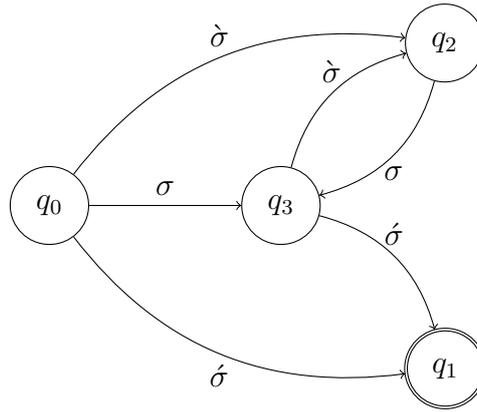


Figure 6.2. Automate fini (d'après Heinz 2009 : 313)²

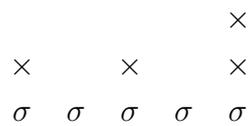


Figure 6.3. Grille rythmique pour $\hat{\sigma}\sigma\hat{\sigma}\sigma\acute{\sigma}$

moins contraintes que les règles des grammaires rationnelles, et l'on peut aisément vérifier que toute règle d'une grammaire rationnelle est également une règle d'une grammaire algébrique : par conséquent, toute grammaire rationnelle est également une grammaire algébrique.

D'un point de vue algorithmique, tout langage algébrique peut être identifié par un AUTOMATE À PILE (*pushdown automaton*), dont la définition formelle est donnée en (4).

(5) Définition d'un automate à pile

Un automate à pile est un sextuplet $(Q, \Sigma, \Gamma, \delta, q_0, F)$, où Q, Σ, Γ, F sont des ensembles finis et :

1. Q est l'ensemble d'états
2. Σ représente l'alphabet d'entrée
3. Γ est l'alphabet de la pile
4. $\delta : Q \times \Sigma_\epsilon \times \Gamma_\epsilon \rightarrow P(Q \times \Gamma_\epsilon)$ est la fonction de transition
5. $q_0 \in Q$ est l'état initial
6. $F \subseteq Q$ est l'ensemble d'états terminaux

(d'après Sipser 2005 : 113)

Comme on peut le voir, cette définition est presque identique à celle donnée en (3) pour l'automate fini, à deux points près : l'ajout d'une *pile* de symboles, qui est une structure de données dont on ne peut manipuler que le sommet (*i.e.* on peut ajouter, lire, et supprimer un symbole au sommet de la pile), et une fonction de transition plus complexe : en plus de l'état courant et du symbole suivant (éventuellement vide)³ sur la bande, elle considère en entrée le symbole au sommet de la pile (éventuellement vide), et produit en sortie un état et un symbole sur la pile⁴.

On voit donc, et c'est ce qu'il faut retenir, qu'un automate à pile n'est rien de plus qu'un automate fini augmenté d'une pile. Les transitions d'un état à un autre sont simplement déterminées, en plus de l'état actuel et du symbole lu sur la bande, par le symbole situé au sommet de la pile. Cet automate est illustré à la figure 6.4 : le contrôle est une représentation schématique de l'ensemble d'états et de la fonction de transition, et la flèche continue représente la tête de lecture. La mémoire additionnelle que représente la pile donne à cet automate une plus grande capacité générative. L'exemple le plus simple qui illustre cette propriété est le langage (algébrique) $\{a^n b^n\}$, soit toutes les chaînes composées de n occurrences de a suivies de n occurrences de b : $\{ab, aabb, aaabbb, \dots\}$. Pour reconnaître une chaîne de ce langage, on peut imaginer que l'automate empile d'abord toutes les occurrences de a , puis vérifie qu'il lit bien un symbole b pour chaque a qu'il désempile. Sans cette mémoire additionnelle, il serait impossible à un automate fini de s'assurer que le nombre de a et de b correspondent, puisque cet automate-là n'a aucune mémoire du « passé »⁵.

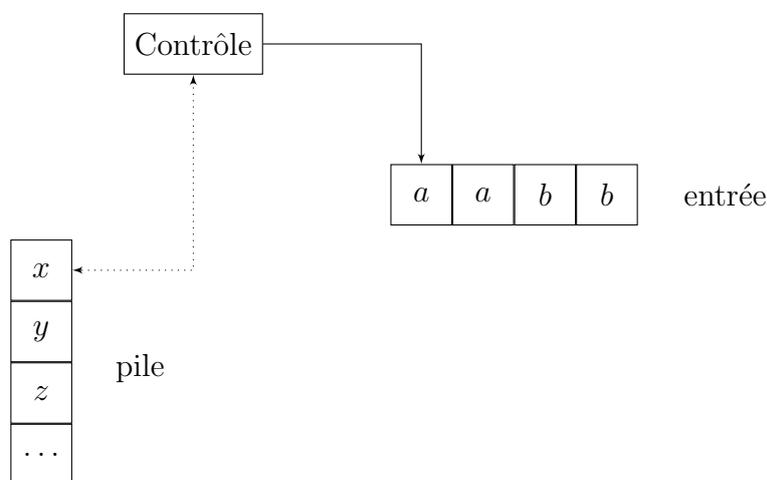


Figure 6.4. Automate à pile (d'après Sipser 2005 : 112)

3. Par convention, Σ_ϵ représente l'union de l'alphabet avec l'ensemble composé de l'élément vide $\{\epsilon\}$. Il en va de même pour Γ_ϵ et la pile de symboles.

4. Plus précisément, cet automate étant non déterministe, la fonction de transition peut produire plusieurs paires (état, symbole sur la pile) valides à partir d'une configuration d'entrée donnée. C'est ce qui est noté $P(Q \times \Gamma_\epsilon)$ dans la définition, soit un ensemble de paires (état, symbole sur la pile).

5. Notons qu'il est possible de construire un automate fini qui *accepte* toutes les chaînes de ce langage (ex : $a^n b^n$), mais pas un automate fini qui n'*accepte que* ce langage.

6.1.1.3 Langages contextuels (L1)

Un langage contextuel peut être décrit par une grammaire contextuelle, dont les règles ont la forme $\alpha B\gamma \rightarrow \alpha\beta\gamma$, où B est un symbole non terminal et α, β, γ sont des chaînes de symboles terminaux ou non terminaux, avec $\beta \neq \epsilon$. En l'occurrence, la transformation opérée par la règle est sensible au contexte $\alpha_ \gamma$. L'exemple canonique de langage contextuel est le langage $\{a^n b^n c^n\}$, soit $\{abc, aabbcc, aaabbbccc, \dots\}$. Je décrirai l'automate qui accepte de tels langages au paragraphe suivant, puisqu'il s'agit d'un cas particulier de l'automate reconnaissant les langages récursivement énumérables. Il est néanmoins possible de vérifier qu'un automate à pile n'est pas assez puissant pour reconnaître un tel langage : l'automate n'ayant accès qu'au sommet de la pile et ne pouvant lire la chaîne d'entrée que de gauche à droite, il n'est pas possible de comparer le nombre de c au nombre de a et de b .

6.1.1.4 Langages récursivement énumérables (L0)

Les langages récursivement énumérables (LRE) peuvent être générés par une grammaire dont les règles ont la forme $\alpha \rightarrow \beta$, où α est une chaîne *non vide* de symboles terminaux et/ou non terminaux, et β est une chaîne (éventuellement vide) de symboles terminaux et/ou non terminaux.

Pour tout LRE, il est possible d'écrire un programme qui énumère récursivement toutes les chaînes du langage. De manière équivalente, tout LRE peut être reconnu par une MACHINE DE TURING, nommée en hommage au mathématicien Alan Turing qui l'a conceptualisée, confirmant par là même les travaux d'Alonzo Church sur la notion de *calculabilité*. La définition formelle de cet automate est donnée ci-dessous :

(6) Définition d'une machine de Turing

Une machine de Turing est un septuplet $(Q, \Sigma, \Gamma, \delta, q_0, q_{accept}, q_{rejet})$, où Q, Σ, Γ sont des ensembles finis et :

1. Q est l'ensemble d'états
2. Σ représente l'alphabet d'entrée qui ne contient pas le **symbole blanc** $_$
3. Γ est l'alphabet de la bande, où $_ \in \Gamma$ et $\Sigma \subseteq \Gamma$
4. $\delta : Q \times \Gamma \rightarrow Q \times \Gamma \times \{G, D\}$ est la fonction de transition
5. $q_0 \in Q$ est l'état initial
6. $q_{accept} \in Q$ est l'état d'acceptation
7. $q_{rejet} \in Q$ est l'état de rejet, avec $q_{accept} \neq q_{rejet}$

(d'après Sipser 2005 : 142)

Cet automate, contrairement aux précédents, dispose d'une bande de longueur infinie (mais bornée à gauche), qu'il peut lire et sur laquelle il peut écrire, et à laquelle est associée un alphabet Γ : celui-ci contient un symbole spécial, le symbole blanc \sqcup , avec lequel toutes les cellules de la bande sont initialisées. La chaîne d'entrée est écrite au début de la bande (*i.e.* sur les n premières cellules), et peut utiliser tout symbole de l'alphabet de la bande sauf \sqcup . Contrairement aux automates précédents, la tête de lecture/écriture peut se déplacer à gauche ou à droite, mais toujours d'une seule cellule à la fois.

La fonction de transition δ prend en entrée l'état actuel et le symbole sous la tête de lecture/écriture, et donne en sortie un nouvel état (qui peut, comme pour un automate fini, être identique au précédent), un nouveau symbole qui remplace le symbole courant sur la bande, et un déplacement d'un pas (= une cellule) vers la gauche (G) ou vers la droite (D), au-dessus du symbole suivant⁶. Au lieu de disposer d'une mémoire annexe telle que la pile dans l'automate à pile, l'automate est libre de lire et écrire sur toute la bande. De ce fait, il n'est pas limité à la lecture du dernier symbole mémorisé, et peut donc reconnaître un ensemble de langages très largement supérieur (tous les langages calculables).

Il est bien entendu possible de limiter les possibilités calculatoires d'une machine de Turing. En particulier, lorsqu'elle ne peut utiliser, en lecture *et* en écriture, que la partie de la bande sur laquelle est écrite la chaîne d'entrée⁷, la machine est appelée AUTOMATE LINÉAIREMENT BORNÉ (*linear bounded automaton*) : il s'agit du type d'automate qui peut reconnaître les grammaires contextuelles discutées précédemment.

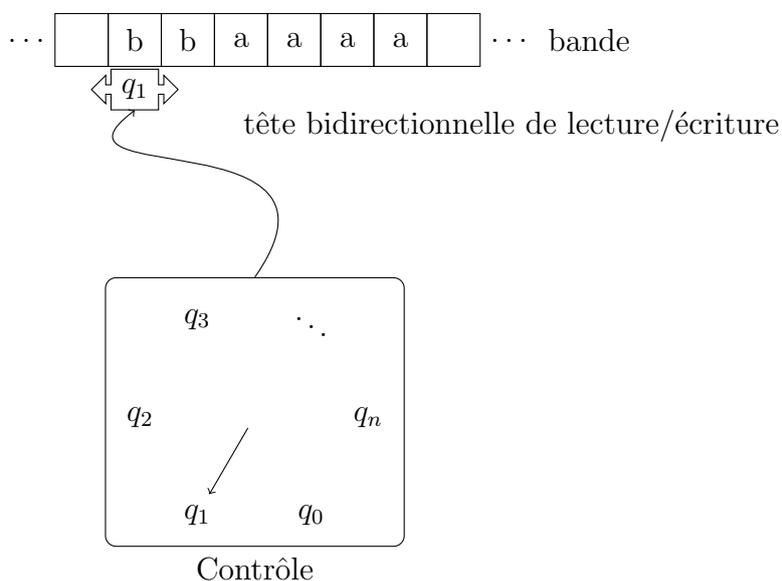


Figure 6.5. Machine de Turing⁸

6. Lorsque la tête de lecture/écriture a atteint le début de la bande et tente de se déplacer à gauche, reste dans cette même position (*i.e.* au-dessus du premier symbole).
 7. Lorsque la machine atteint le dernier caractère de la chaîne et tente de se déplacer à droite, elle reste dans la même position. Voir note 6.
 8. D'après <http://www.texample.net/tikz/examples/turing-machine-2/>.

6.1.2 Extensions de la Hiérarchie de Chomsky

La hiérarchie de Chomsky « standard » a été raffinée de différentes façons, aussi je me contenterai d'évoquer brièvement ici les sous-types de langages qui importent vis-à-vis de la question de la complexité formelles des langues naturelles (voir Jäger & Rogers 2012 pour un aperçu récent).

La proposition originelle de Chomsky était que les grammaires rationnelles n'étaient pas suffisantes pour formaliser la syntaxe de l'anglais (Chomsky 1957). Ce sujet a été âprement débattu et la démonstration de Chomsky a été critiquée (voir Pullum 2011 : pour une synthèse), mais il y a néanmoins aujourd'hui, en traitement automatique du langage du moins, un point de vue assez largement partagé selon lequel la majorité des faits grammaticaux nécessitent au moins des grammaires algébriques, et il existe certains phénomènes, tels que la subordination en suisse allemand (Shieber 1985), qui nécessitent des grammaires légèrement plus puissantes que les grammaires algébriques : les GRAMMAIRES LÉGÈREMENT CONTEXTUELLES (*mildly context-sensitive*, Joshi 1985). Ce point de vue ne fait toutefois pas l'unanimité, et certains auteurs doutent de ce que les grammaires légèrement contextuelles soient suffisantes pour rendre compte de la syntaxe des langues naturelles, *a fortiori* si l'on souhaite prendre en compte la capacité générative forte des grammaires (adéquation descriptive), et pas seulement leur capacité à générer les chaînes du langage (adéquation observationnelle) (Francez & Wintner 2011).

La classe de langages la plus restrictive, *i.e.* L3, a elle aussi connu de nombreux raffinements. On peut ainsi identifier différents sous-types de langages rationnels, qui ne constituent pas nécessairement des classes imbriquées, mais dont les langages partagent certaines propriétés les uns avec les autres. On parle alors de langages SOUS-RATIONNELS (*subregular*). Je n'en donnerai ici que quelques exemples (voir Rogers & Pullum 2011 pour une présentation détaillée, et §6.1.3 ci-après). La première classe que l'on peut identifier, et sans doute la plus triviale, est la classe des langages FINIS, c'est-à-dire la classe des langages contenant un ensemble fini de chaînes. Tous ces langages peuvent être représentés par un automate fini acyclique, et ils sont donc effectivement tous rationnels.

Un autre exemple est la classe des langages STRICTEMENT LOCAUX (*strictly local*). Un langage est strictement k -local (SL_k) si pour toute chaîne du langage, toutes les sous-chaînes de longueur k appartiennent à un ensemble appelé k -facteurs du langage, c'est-à-dire toutes les séquences de k symboles (= k -grammes), pris dans l'alphabet du langage, généralement augmenté de deux symboles représentant le début et la fin de la chaîne (\bowtie et \bowtie respectivement). Par exemple, si la liste de 2-facteurs autorisés est $\{\bowtie C, CV, VC, V \bowtie\}$, le langage résultant sera $\{CV^n\}$ soit toute suite de séquence CV , *i.e.* $\{CV, CVCV, CVCVCV, \dots\}$. La chaîne CVC n'est alors pas une chaîne du langage puisqu'elle contient la sous-chaîne $C\bowtie$, qui n'est pas un 2-facteur du langage.

6.1.3 Où se situe la phonologie ?

S'il est généralement admis que la phonologie est formellement moins complexe que la syntaxe, certains linguistes ont tenté de formaliser les phénomènes phonologiques à l'aide

de formalismes syntaxiques⁹, par exemple en utilisant la théorie X-barre, originellement développée pour représenter la constituance syntaxique, pour représenter la syllabe (ex : van Oostendorp 2000). Certains auteurs ont même défendu une forme de récursivité restreinte pour le pied (voir notamment Sauzet 2004 pour l’analyse du schwa français). Il est toutefois utile de s’interroger sur la complexité formelle des patrons phonologiques puisque, comme nous l’avons vu plus haut (voir p. 142), il est possible de décrire certains systèmes accentuels à l’aide d’automates finis, ce qui signifie que de tels systèmes sont des langages rationnels.

On doit tout d’abord à Johnson (1972) (cité par Karttunen 1993) d’avoir observé que les règles de réécriture du formalisme SPE, bien qu’elles s’apparentent à des règles contextuelles, avaient une puissance équivalente à des automates finis dès lors qu’elles ne s’appliquaient pas à leur propre output. Divers travaux, notamment Koskeniemi (1983), Kaplan & Kay (1994) et Karttunen (1993), ont consolidé les résultats de Johnson et ont montré que la plupart des phénomènes phonologiques peuvent être décrits à l’aide de TRANSDUCTEURS FINIS (*finite-state transducers*), des automates finis qui peuvent écrire un symbole en sortie chaque fois qu’ils lisent un symbole en entrée, ce qui leur permet effectivement d’apparier une forme d’entrée et une forme de sortie.

Plus récemment, dans une étude typologique des patrons accentuels, Heinz (2009) observe que sur les 109 patrons qu’il a étudiés, tous ont une propriété commune : lorsque le patron est représenté sous forme d’automate fini (cf. figure 6.2 p. 145), et que l’on considère le voisinage gauche et droit de longueur $j - k$, où j est le nombre de voisins à gauche et k le nombre de voisins à droite, il n’existe pas deux états qui aient le même voisinage. Plus précisément, 107 des 109 patrons ont un voisinage 1 – 1 (un voisin à gauche et un à droite) ; les deux patrons restants proviennent de descriptions dont la fiabilité n’est pas assurée, mais leur complexité n’excède pas 2 – 2 (deux voisins à gauche et deux à droite). Tous ces patrons appartiennent donc à une sous-classe des langages rationnels, et sont donc sous-rationnel. Heinz nomme cette classe les *neighborhood-distinct languages*, expression que je traduirai sans effort stylistique par langages À VOISINAGE DISTINCT. Heinz (2009) va plus loin et émet l’hypothèse, qui reste à démontrer, que tous les patrons phonotactiques sont à voisinage distinct. Dans un autre article (Heinz 2010), il explore différents autres patrons, y compris l’harmonie consonantique à distance, et montre qu’ils sont également sous-rationnels, mais il n’explore pas, pour autant que je puisse en juger, l’hypothèse selon laquelle ces patrons soient à voisinage distinct.

Ces travaux, pris ensemble, montrent que la plupart des phénomènes sont dans la limite des langages rationnels. Il est toutefois possible que certains phénomènes de reduplication soient au-delà de cette classe de langages (voir Heinz 2009 : 315 et les références citées). Néanmoins, cela signifie que la plupart des phénomènes peuvent être appréhendés à l’aide d’automates finis, éventuellement augmentés de certaines extensions si nécessaire,

9. Ce point de vue a été poussé jusqu’à sa conséquence logique par Anderson (2006, 2013), qui défend un principe d’*analogie structurale* selon lequel les composantes de la grammaire utilisent les mêmes principes structurels, les différences étant dues à la nature de l’alphabet manipulé (ex : primitives articulatoires/acoustiques pour la phonologie *vs* primitives sémantiques pour la syntaxe).

comme dans le cas de la reduplication. Comme le soulignent Karttunen (1993) et Heinz (2009), une telle stratégie peut être utile dans la mesure où cela permet de mesurer précisément la complexité formelle de descriptions ou de vérifier qu'elles sont équivalentes. Néanmoins, il faut garder à l'esprit que de telles comparaisons ne nous renseignent que sur la capacité générative faible des grammaires, et différentes descriptions peuvent correspondre à différentes hypothèses linguistiques. Par exemple, les grammaires correspondant aux structures syllabiques de la figure 6.6 sont faiblement équivalentes et généreraient bien les mêmes chaînes (*i.e.* les mêmes syllabes), mais elles encodent deux hypothèses distinctes quant à la relation qui s'établit entre les segments à l'intérieur de la syllabe.

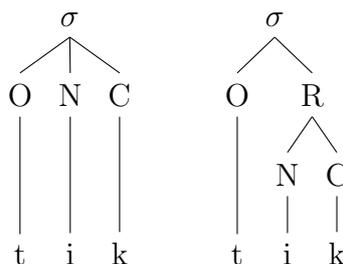


Figure 6.6. Deux descriptions structurales de la même syllabe

6.2 Apprentissage formel

La question de l'apprenabilité des langues naturelles est très souvent discutée dans le contexte d'un résultat obtenu par Gold (1967), et qui a souvent été interprété à l'appui de l'hypothèse que les langues ne peuvent être apprises sur la seule base des exemples positifs auxquels l'apprenant est exposé. Ce résultat a naturellement reçu un écho favorable chez nombre de ceux qui souscrivent à l'hypothèse de la pauvreté du stimulus.

6.2.1 Le théorème de Gold

Le cadre d'apprentissage développé par Gold (1967), comme il le note d'ailleurs lui-même, est un cadre extrêmement idéalisé mais qui permet d'étudier l'apprenabilité des grammaires formelles de manière précise. Dans ce cadre, l'apprenant a pour input un TEXTE, c'est-à-dire un ensemble infini de chaînes provenant du langage cible L . On suppose que le texte est non bruité (il n'y a aucune erreur dans l'input) et que l'apprenant aura été exposé, à un moment ou à un autre, à toutes les chaînes du langage.

L'apprenant est représenté par une fonction G qui reçoit en entrée une nouvelle chaîne i_t provenant du langage à chaque instant t . A chaque fois qu'il reçoit un exemple du langage, l'apprenant émet une hypothèse \hat{g}_t sur la grammaire cible à partir de l'ensemble des chaînes auquel il a été exposé, soit $\hat{g}_t = G(i_1, i_2, \dots, i_t)$. Le langage L est IDENTIFIÉ À LA LIMITE si, après un temps fini, l'apprenant a convergé vers la grammaire cible de

L . Une famille (ou classe) de langages \mathcal{L} est IDENTIFIABLE À LA LIMITE ssi pour tout $L \in \mathcal{L}$, il existe un apprenant qui identifie L à la limite en un temps fini, quel que soit l'ordre de présentation de l'information. Pour être clair, la notion d'identifiabilité (ou d'apprenabilité) est une propriété des familles de langages, pas des langages eux-mêmes. Un langage donné est simplement identifié, ou ne l'est pas. Le résultat le plus important du travail de Gold, en ce qui concerne l'apprentissage des langues humaines, est le résultat négatif suivant :

(7) Théorème de Gold

Si la famille \mathcal{L} correspond à tous les langages finis et au moins un langage infini, alors elle n'est pas apprenable (identifiable) à la limite.

(adapté de Niyogi 2006 : 58 ; Gold 1967)

Ce résultat signifie que même la famille des langages rationnels, pourtant la plus restrictive dans la hiérarchie de Chomsky standard, n'est pas identifiable à la limite puisque, comme nous l'avons vu plus haut, elle contient de nombreux langages infinis. Cela signifie *a fortiori* que la classe des grammaires algébriques et la classe des grammaires (légèrement) contextuelles sont elles aussi inapprenables (Niyogi 2006 : 59). Une interprétation possible, et souvent avancée, de ce résultat est qu'il pose de sérieuses difficultés pour un modèle empiriciste qui suppose que l'apprenant induit la grammaire d'une langue à partir de la seule expérience, puisque les langues appartiennent à la famille des grammaires (légèrement) contextuelles. Autrement dit, l'apprenant doit disposer de biais d'apprentissage innés (ou d'une Grammaire Universelle) pour contraindre l'espace des hypothèses (Niyogi 2006).

6.2.2 Identifiabilité vs acquérabilité

Le résultat de Gold a été critiqué à divers égards, et l'on trouvera une excellente synthèse du débat dans Johnson (2004), sur lequel je m'appuie ici. Une critique souvent avancée est que les apprenants de Gold sont particulièrement abstraits, et qu'ils ne font qu'apprendre des chaînes syntaxiques décontextualisées. Un apprenant plus réaliste et qui aurait accès à plus d'information (phonologique, morphologique, sémantique, pragmatique) pourrait donc exploiter ces différentes sources d'information pour induire la grammaire de la langue. Johnson (2004 : 580) observe cependant que le cadre de Gold n'impose strictement aucune contrainte sur la structure des langages : chaque « chaîne » ou phrase du langage pourrait donc avoir une structure arbitrairement complexe et l'alphabet du langage pourrait contenir tout type d'information utile à l'apprenant sans que cela n'invalide le théorème de Gold.

Johnson démontre par ailleurs que le fait de disposer de biais d'apprentissage (qu'ils soient fournis par la GU ou non) n'améliore pas la situation puisque le résultat de Gold

s'applique pour *tous* les apprenants. Ce résultat repose sur le fait qu'il est possible à un instructeur « malicieux » de présenter les chaînes du langage cible de telle sorte que l'apprenant ne puisse jamais converger vers l'hypothèse correcte¹⁰, quels que soient les biais qu'il puisse avoir par ailleurs. Le théorème de Gold est donc aussi problématique pour les approches rationalistes que pour les approches empiricistes (Johnson 2004 : 483–484)

Johnson montre néanmoins qu'il reste une difficulté importante lorsque le théorème de Gold est transposé à l'apprentissage des langues naturelles par l'être humain. Il n'y a en effet aucune garantie que l'apprenant humain convergera *exactement* vers la grammaire cible. Johnson propose donc de distinguer l'identifiabilité au sens de Gold de l'ACQUÉRABILITÉ (*acquirability*), qui consiste à apprendre une grammaire suffisamment proche de la grammaire à laquelle l'apprenant a été exposé.

So far, we've used the term "learnability" as Gold used it, so that a class of languages C is learnable iff there exists a function ϕ such that for any environment E for any language L in C , ϕ permanently converges onto the hypothesis of L as the target language after some finite time. But there is another option. We could also take learnability to have a psychologically more natural meaning, perhaps along the lines of : a class C of natural languages is learnable iff given almost any normal human child and almost any normal linguistic environment for any language L in C , *the child will acquire L (or something sufficiently similar to L) as a native language between the ages of one and five years.*

(Johnson 2004 : 585, c'est moi qui souligne)

De fait, l'existence même de la variation, et de la variation diachronique en particulier, démontre *de facto* que les humains ne sont pas des apprenants de Gold : si l'apprenant convergerait toujours vers la grammaire « cible » (si tant est qu'il soit par ailleurs possible d'identifier *une* grammaire cible), les langues ne devraient jamais changer (Niyogi & Berwick 1997 : 162). Par ailleurs, il est bien établi dans le domaine de l'apprentissage artificiel que certains algorithmes sont capables de converger vers un optimum local, mais pas nécessairement global. C'est le cas de l'algorithme de gradient¹¹ (Manning & Schutze 1999 : 597 suiv.) qui, en partant d'un point quelconque, vise à maximiser localement la valeur de la fonction d'apprentissage : il peut donc « monter » et s'arrêter au sommet du pic le plus proche, sans que celui-ci soit nécessairement le sommet global de la fonction (cf. figure 6.7). Dans ce cas, l'algorithme a atteint une solution qui peut être suffisante mais qui n'est pas nécessairement optimale. Il est donc tout à fait plausible qu'il en aille de même pour les apprenants humains. On sait par ailleurs que les systèmes linguistiques qui se forment en situation de contact sont des systèmes qui ont une grammaire différente des langues sources. C'est le cas des créoles ou encore du mitchif, une langue mixte

10. Sur ce point, voir également Chater, Clark, Goldsmith & Perfors (2015 : 54–55), qui soulignent aussi le rôle de la fréquence dans l'input, qui est exploitée par l'être humain mais qui ne joue aucun rôle chez les apprenants de Gold.

11. Une variante de cet algorithme est utilisée dans l'algorithme de recuit simulé discuté en §5.1.3. Le rôle de la température est notamment de minimiser le risque de converger vers un optimum local.

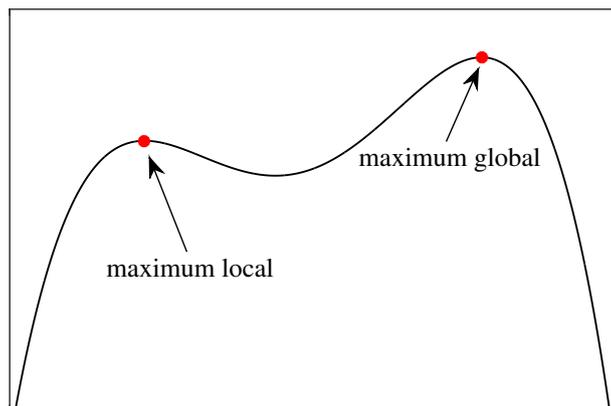


Figure 6.7. Maxima local et global d'une fonction quelconque

parlée dans les prairies canadiennes, et dont le système nominal est à près de 90% à base française, alors que le système verbal est à base crie, une langue algonquienne (Papen 2005). L'émergence de telles langues montre clairement que le rôle de l'apprenant dans l'apprentissage n'est pas d'*identifier*, mais de *construire* une grammaire, à partir de l'input reçu, pour pouvoir communiquer (Tomasello 2003).

6.3 Conclusion

J'ai examiné dans ce chapitre la question de l'apprenabilité des langues naturelles d'un point de vue formel. Nous avons vu que la plupart des patrons phonologiques, au moins du point de vue de la capacité générative faible des grammaires, sont dans la limite des langages rationnels, et sont souvent sous-rationnels (ex : systèmes accentuels). Toutefois, dans la mesure où ils sont infinis (ex : une grammaire accentuelle peut théoriquement être appliquée à une chaîne de longueur arbitraire), ils sont inapprenables au sens de Gold puisque toute classe de langages qui contient tous les langages finis et au moins un langage superfini n'est pas identifiable à la limite. Ce résultat a été interprété comme un argument à l'appui de l'hypothèse selon laquelle l'apprentissage doit être guidé par une Grammaire Universelle qui contraint l'espace des possibles mais, comme nous l'avons vu, le théorème de Gold reste valide quels que soient les biais d'apprentissage, et quelle que soit leur origine (Grammaire Universelle innée, mécanismes cognitifs généraux). Néanmoins, le problème de Gold ne se pose que dans le cas précis où l'on cherche à identifier *exactement* une grammaire dans un ensemble de grammaires dont au moins une est infinie : ceci ne correspond pas à la situation d'apprentissage normale de l'être humain, qui construit simplement *une* grammaire à partir de l'input qu'il reçoit. Bien que les travaux qui se situent dans la lignée du travail de Gold (1967) soient utiles d'un point de vue théorique

puisque'ils nous renseignent sur ce qui est théoriquement possible dans certains cas de figure bien définis, ils deviennent problématiques s'ils sont interprétés tels quels comme des modèles de l'apprentissage humain.

Conclusion générale

J'ai retracé dans ce mémoire les grandes lignes de mon parcours universitaire. Le travail que j'ai entrepris depuis mon doctorat se situe au carrefour de la phonologie de corpus, de l'ingénierie linguistique et de la phonologie théorique. Comme je l'ai souligné au premier chapitre, ma démarche s'inscrit clairement dans une vision empirique de la phonologie : de ce point de vue, le travail de description systématique de la variation que j'ai présenté au premier chapitre fait partie intégrante de cette démarche, puisqu'il en constitue l'ossature. De la même manière, le travail de développement continu que j'ai entrepris tout au long des dix dernières années est un élément essentiel puisqu'il vient soutenir le travail empirique que mes collègues et moi menons sur le français, aussi bien en L1 qu'en L2. De ce point de vue, je continue à collaborer activement avec Sylvain Detey et Isabelle Racine de manière à rendre l'outil Dolmen, qui est ma principale réalisation, plus utile et plus ergonomique au fur et à mesure que la masse de données à traiter augmente et que de nouveaux besoins se font sentir. La frontière entre travail empirique et travail de développement est plutôt ténue dans la mesure où le développement de scripts (Praat ou Python) intervient presque inévitablement dans la plupart de mes travaux. Je considère donc avant tout la programmation comme un outil parmi d'autres dans l'arsenal du linguiste de corpus.

Mes travaux théoriques, on l'a vu, ont exploré un certain nombre de pistes, tant en modélisation symbolique que, plus récemment, en modélisation quantitative. Cet éclectisme s'explique par le fait que le courant dominant en phonologie, à savoir la théorie de l'optimalité dans ses diverses manifestations, me laisse profondément insatisfait. Bien que je l'aie adoptée dans ma thèse et dans quelques publications depuis, cela a toujours été dans des variantes « non standard » dans la mesure où les phénomènes auxquels je me suis intéressé (liaison, schwa, pied en français) deviennent très rapidement difficiles à appréhender dès que l'on essaie de rendre compte (d'une partie) de leur complexité. Les différentes approches et méthodes quantitatives que j'ai explorées récemment me paraissent maintenant des outils bien plus adaptés que les approches strictement logico-symboliques pour modéliser la gradience et la variation linguistique.

En guise d'ouverture, j'évoquerai brièvement les principaux chantiers en cours sur lesquels je travaille. A court terme, je développe actuellement une analyse computationnelle du contact de variétés dans une perspective de transmission culturelle (Cavalli-Sforza & Feldman 1981). Ce travail se focalise plus précisément sur l'influence du français de référence sur le français méridional dans la diffusion du schwa : mon objectif est d'explo-

rer le potentiel des simulations basées sur agents dans la modélisation de la diffusion du changement linguistique, et mon intérêt pour ces approches a été grandement stimulé par la lecture des travaux d'Oudeyer (2005, 2006) et de Kaplan (2001). Ce travail fera l'objet d'une publication dans un volume co-édité par Helene Andreassen et Elissa Pustka sur la phonologie du schwa en français.

Mon principal projet sur l'année à venir, en collaboration avec Marie-Hélène Côté (Université de Lausanne, Suisse), est un ouvrage intitulé *The Phonology of French* qui est en cours d'écriture et qui paraîtra chez Oxford University Press dans la série *The Phonology of the World's Languages* (le contrat a été signé avec l'éditeur et la publication est prévue pour 2017/2018). Cet ouvrage vise d'une part à donner un panorama de la phonologie du français, en accordant une place centrale à la variation, diatopique en particulier, et en nous appuyant largement sur les données et résultats issus du projet PFC, et d'autre part d'offrir une synthèse critique des analyses qui ont été proposées dans la littérature, notamment sur les phénomènes tels que le schwa, la liaison et le *h*-aspiré. Cet ouvrage devrait permettre aux étudiants et phonologues, francophones ou non, de se familiariser avec la phonologie des variétés du français dans leur diversité, au-delà des descriptions traditionnelles basées sur le français de référence, et permettra aussi de porter à la connaissance des lecteurs non francophones un ensemble de travaux et analyses qui n'ont pas reçu l'attention qu'ils auraient pu recevoir s'ils avaient été écrits en anglais. J'ai joint le descriptif détaillé de ce travail en Annexe A du second volume.

A plus long terme, je poursuis l'objectif de tenter d'intégrer les apports des grammaires d'usage et des grammaires formelles. De ce point de vue, les grammaires probabilistes en général, et les approches telles que le modèle phonotactique MaxEnt de Hayes & Wilson (2008) en particulier, que j'ai appliqué à la phonotaxe du français, me semblent ouvrir des perspectives prometteuses pour la modélisation phonologique et constituent une voie médiane intéressante entre grammaires d'usage et grammaires formelles, dans la mesure où ils permettent de prendre en compte les effets de la fréquence de l'input. Bien que mes positions théoriques aient sensiblement évolué depuis la fin de mon doctorat il y a dix ans, je reste convaincu que la notion d'harmonie, au sens de « mesure de la cohérence d'un système » telle qu'elle a été définie par Paul Smolensky dans ses premiers travaux (Smolensky 1983, 1986), continuera à jouer un rôle important, sous une forme ou sous une autre, dans la prochaine génération de modèles théoriques qui émergera.

Annexe A

Extrait du Plugin PFC

Le listing A.1 correspond au code Lua du module `metadata`, dont la tâche est d'ajouter les métadonnées standard pour les annotations issues du projet PFC. Ce module fait partie du « plugin PFC », qui ajoute à Dolmen un certain nombre de fonctionnalités répondant aux besoins spécifiques des collaborateurs de ce projet.

Ce module ajoute une entrée intitulée `Add PFC properties` dans le menu PFC. Lorsque l'utilisateur clique sur ce menu, le module récupère toutes les annotations du projet courant et, pour chaque fichier, il vérifie s'il s'agit d'un fichier TextGrid (le format d'annotation de Praat utilisé au sein du projet PFC) formaté selon les conventions PFC. Si tel est le cas, il extrait les métadonnées à partir du nom du fichier et les ajoute au fichier d'annotation. Il tente également de trouver un fichier son correspondant à cette annotation ; en cas de succès, il lie l'annotation au fichier son et charge automatiquement le fichier son si besoin est.

Listing A.1. Code Lua permettant d'extraire les métadonnées PFC

```
local re_meta = regex.new("^(...) (...) (.)g\\. [tT]ext [gG]rid$")
local fs = filesystem

local tasks = {
  g = "Guided conversation",
  l = "Free conversation",
  t = "Text",
  m = "Word list"
}

local function add_metadata(annot)
  -- set properties
  local file = fs.base_name(annot.path)

  if re_meta:match(file) then
    local survey = string.upper(re_meta:capture(1))
    local speaker = string.upper(re_meta:capture(2))
```

```

    local task = tasks[re_meta:capture(3)]

    annot:add_property("Survey", survey)
    annot:add_property("Task", task)
    annot:add_property("Speaker", speaker)
end

-- set sound file
local extensions = dolmen.supported_audio_formats()
local base = fs.split_ext(annot.path)[1]
base = base:remove_last("g") .. "w"

for i, ext in ipairs(extensions) do
    local sound = base .. ext

    if fs.exists(sound) then
        annot:bind_to_sound(sound)
        break
    end
end
end

local function tag_annotations()
    annots = dolmen.annotations()

    for i, annot in ipairs(annots) do
        add_metadata(annot)
    end

    dolmen.refresh_project()
end

dolmen.add_menu_entry("PFC", "Add PFC properties", tag_annotations)

```

Bibliographie

- Ågren, John. 1973. *Étude sur quelques liaisons facultatives dans le français de conversation radiophonique : fréquences et facteurs*. Uppsala : Acta Universitatis Upsaliensis.
- Ahn, Sang-Cheol. 1998. *An Introduction to Korean Phonology*. Seoul, Korea : Hansin Munhwasa.
- Alderete, John & Mark Bradshaw. 2013. Samoan root phonotactics : Digging deeper into the data. *Linguistic Discovery* 11(1). 1–22.
- Alderete, John, Paul Tupper & Stefan A. Frisch. 2013. Phonological constraint induction in a connectionist network : Learning OCP-Place constraints from data. *Language Sciences* 37. 52–69.
- Anderson, John. 2006. Structural analogy and universal grammar. *Lingua* 116(5). 601–633.
- Anderson, John. 2013. Substance, structural analogy, and universals. *Language Sciences* 39. 15–30.
- Anderson, Stephen R. 1985. *Phonology in the Twentieth Century : Theories of Rules and Theories of Representations*. Chicago : University of Chicago Press.
- Andreassen, Helene N. & Julien Eychenne. 2013. The French foot revisited. *Language Sciences* 39. 126–140.
- Angoujard, Jean-Pierre. 1997. *Théorie de la syllabe. Rythme et qualité*. Paris : CNRS Editions.
- Armstrong, Nigel & Tim Pooley. 2010. *Social and Linguistic Change in European French*. New York : Palgrave Macmillan.
- Aurnague, Michel & Jacques Durand. 2003. Quelques aspects de la phonologie du français au Pays basque. *La Tribune Internationale des Langues Vivantes* 33. 110–6.
- Baković, Eric. 2007. A revised typology of opaque generalisations. *Phonology* 24(2). 217–259.

- Baković, Eric. 2011. Opacity and Ordering. In John A. Goldsmith, Jason Riggle & Alan C. L. Yu (coord.), *The Handbook of Phonological Theory*, 40–67. Wiley-Blackwell 2^e édition.
- Barreca, Giulia. 2015. *L'acquisition de la liaison chez des apprenants italophones. Des atouts d'un corpus de natifs pour l'étude de la liaison en français langue étrangère (FLE)*. Université Paris Ouest Nanterre La Défense & Università Cattolica del Sacro Cuore di Milano, thèse de doctorat.
- Beckner, Clay, Richard Blythe, Joan Bybee, Morten H. Christiansen, William Croft, Nick C. Ellis, John Holland, Jinyun Ke, Diane Larsen-Freeman & Tom Schoenemann. 2009. Language Is a Complex Adaptive System : Position Paper. *Language Learning* 59. 1–26.
- Beddor, Patrice Speeter. 1991. Predicting the Structure of Phonological Systems. *Phonetica* 48(2-4). 83–107.
- Behme, Christina. 2014. A 'Galilean' science of language. *Journal of Linguistics* 50(03). 671–704.
- Bergounioux, Gabriel & Bernard Laks. 2003. Portrait de Saussure en phonologue contemporain. In Simon Bouquet (coord.), *Ferdinand de Saussure*, Paris : Editions de L'Herne.
- Bermúdez-Otero, Ricardo. 2007. Marked phonemes vs marked allophones : Segment evaluation in Stratal OT'. Communication au workshop 'Segment inventories', 30^e GLOW Colloquium.
- Bermúdez-Otero, Ricardo. 2012. The architecture of grammar and the division of labour in exponence. In Jochen Trommer (coord.), *The morphology and phonology of exponence*, vol. 41, 8–83. Oxford : Oxford University Press Oxford.
- Biber, Douglas, Susan Conrad & Randi Reppen. 1998. *Corpus Linguistics : Investigating Language Structure and Use*. Cambridge : Cambridge University Press.
- Bird, Stephen. 1995. *Computational Phonology. A constraint-based approach*. Cambridge : Cambridge University Press.
- Bird, Steven & Mark Liberman. 2001. A Formal Framework for Linguistic Annotation. *Speech Commun.* 33(1-2). 23–60.
- Bladon, Anthony & Gunnar Fant. 1978. A two-formant model and the cardinal vowels. *Speech Transmission Laboratory, Quarterly Progress and Status Report* 19(1). 1–8.
- Boeckx, Cedric. 2006. *Linguistic Minimalism : Origins, Concepts, Methods, and Aims*. OUP Oxford.

- Boersma, Paul & David Weenink. 2016. Praat : Doing phonetics by computer [Computer program]. Version 1.8.5. <http://www.praat.org>.
- Botha, Rudolf P. 1982. On ‘the Galilean style’ of linguistic inquiry. *Lingua* 58(1). 1–50.
- Burton-Roberts, Noel, Philip Carr & Gerard Docherty. 2000. *Phonological Knowledge : Conceptual and Empirical Issues*. Oxford University Press.
- Bybee, Joan. 2001. *Phonology and Language Use*. Cambridge : Cambridge University Press.
- Bybee, Joan L. & Carol Lynn Moder. 1983. Morphological Classes as Natural Categories. *Language* 59(2). 251–270.
- Carr, Philip. 2000. Scientific realism, sociophonetic variation, and innate endowments in phonology. In Burton-Roberts et al. (2000), 67–104.
- Cavalli-Sforza, Luigi Luca & Marcus W. Feldman. 1981. *Cultural Transmission and Evolution : A Quantitative Approach*. Princeton, N.J : Princeton University Press.
- Chater, Nick, Alexander Clark, John A. Goldsmith & Amy Perfors. 2015. *Empiricism and Language Learnability*. Oxford : Oxford University Press.
- Chevrot, Jean-Pierre, Damien Chabanal & Céline Dugua. 2007. Pour un modèle de l’acquisition des liaisons basé sur l’usage : Trois études de cas. *Journal of French Language Studies* 17(01). 103–128.
- Chladkova, Katerina & Paola Escudero. 2012. Comparing vowel perception and production in Spanish and Portuguese : European versus Latin American dialects. *The Journal of the Acoustical Society of America* 131(2). EL119–EL125.
- Cho, Taehong, Sun-Ah Jun & Peter Ladefoged. 2002. Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics* 30(2). 193–228.
- Chomsky, N. 1956. Three models for the description of language. *IRE Transactions on Information Theory* 2(3). 113–124.
- Chomsky, Noam. 1957. *Syntactic structures*. La Haye : Mouton.
- Chomsky, Noam. 1964. *Current issues in linguistic theory*, vol. 38. La Haye & Paris : Walter de Gruyter.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge : MIT Press.
- Chomsky, Noam. 1986. *Knowledge of Language : Its Nature, Origins, and Use*. New York : Praeger.
- Chomsky, Noam. 2002. *On Nature and Language*. Cambridge : Cambridge University Press.

- Chomsky, Noam & Morris Halle. 1968. *The Sound Pattern of English*. Cambridge, Mass : MIT Press.
- Christiansen, Morten H. & Nick Chater. 2016. *Creating Language : Integrating Evolution, Acquisition, and Processing*. Cambridge, MA : The MIT Press.
- Clements, George Nick. 1990. The role of the sonority cycle in core syllabification. In John Kingston & Mary E Beckman (coord.), *Papers in Laboratory Phonology I. Between the Grammar and Physics of Speech*, 283–333. Cambridge University Press.
- Coady, Jeffrey A & Richard N Aslin. 2003. Phonological neighbourhoods in the developing lexicon. *Journal of Child Language* 30. 441–469.
- Cock, Peter J. A., Tiago Antao, Jeffrey T. Chang, Brad A. Chapman, Cymon J. Cox, Andrew Dalke, Iddo Friedberg, Thomas Hamelryck, Frank Kauff, Bartek Wilczynski & Michiel Hoon. 2009. Biopython : freely available python tools for computational molecular biology and bioinformatics. *Bioinformatics* 25(11). 1422–1423.
- Côté, Marie-Hélène. 2017. La liaison en diatopie : Esquisse d'un modèle. *Journal of French Language Studies* 27(1). A paraître.
- Croft, William & D. Alan Cruse. 2004. *Cognitive Linguistics*. Cambridge University Press 1^e édition.
- de Boer, Bart. 2001. *The Origins of Vowel Systems*. Oxford & New York : Oxford University Press.
- de Jong, Daan. 1994. La sociophonologie de la liaison orléanaise. In Chantal Lyche (coord.), *French Generative Phonology : Restrospective and Perspectives*, 95–130. AFLS/ESRI.
- de Saussure, Ferdinand. 1916. *Cours de Linguistique Générale*. Paris : Payot. Edition critique préparée par Tullio de Mauro, 1984.
- de Saussure, Ferdinand. 2002. *Ecrits de linguistique générale*. Paris : Gallimard.
- Delais-Roussarie, Elisabeth & Annie Rialland. 2007. Metrical structure, tonal association and focus in French. In *Romance Languages and Linguistic Theory 2005 : Selected Papers from 'Going Romance', Utrecht, 8-10 December 2005*, 73–98. John Benjamins Publishing Company.
- Dell, François. 1970. *Les règles phonologiques tardives et la morphologie dérivationnelle du français* : thèse de doctorat MIT.
- Dell, François. 1973. *Les règles et les sons. Introduction à la phonologie générative*. Paris : Hermann.

- Detey, Sylvain. 2009. Phonetic input, phonological categories and orthographic representations : A psycholinguistic perspective on why language education needs oral corpora. The case of French-Japanese interphonology development. In Yuji Kawaguchi, Makoto Minegishi & Jacques Durand (coord.), *Corpus Analysis and Variation in Linguistics*, 179–200. Tokyo : John Benjamins.
- Detey, Sylvain, Jacques Durand, Bernard Laks & Chantal Lyche. 2010. *Les Variétés du Français Parlé dans l'Espace Francophone - Ressources pour l'enseignement*. Paris : Ophrys.
- Detey, Sylvain, Jacques Durand, Bernard Laks & Chantal Lyche (coord.). 2016. *Varieties of Spoken French*. Oxford : Oxford University Press.
- Detey, Sylvain & Yuji Kawaguchi. 2008. Interphonologie du Français Contemporain (IPFC) : Récolte automatisée des données et apprenants japonais. Communication aux Journées PFC *Phonologie du français contemporain : variation, interfaces, cognition*.
- Detey, Sylvain & Isabelle Racine. 2012. Les apprenants de français face aux normes de prononciation : quelle(s) entrée(s) pour quelle(s) sortie(s)? *Revue française de linguistique appliquée* XVII(1). 81–96.
- Detey, Sylvain, Isabelle Racine, Julien Eychenne & Yuji Kawaguchi. 2014. Corpus-based L2 phonological data and semi-automatic perceptual analysis, 539–543.
- Detey, Sylvain, Isabelle Racine, Yuji Kawaguchi & Julien Eychenne (coord.). 2017. *La prononciation du français dans le monde : Du natif à l'apprenant*. Paris : CLE International. A paraître.
- Dodd, James. 2004. *Crisis and Reflection : An Essay on Husserl's Crisis of the European Sciences*. New York : Kluwer Academic Publishing.
- Dubois, Sylvie. 2005. Un siècle de français cadien parlé en louisiane : persistance linguistique, hétérogénéité géographique et évolution. In Valdman et al. (2005) 287–305.
- Durand, Jacques. 1976. Generative phonology, dependency phonology and southern French. *Lingua e Stile* 11(1). 3–23.
- Durand, Jacques. 1986. French liaison, floating segments and other matters in a dependency framework. In Jacques Durand (coord.), *Dependency and Non-Linear Phonology*, 161–201. Londres : Croom Helm.
- Durand, Jacques. 1988. Les phénomènes de nasalité en français du Midi : Phonologie de dépendance et sous-spécification. *Recherches Linguistiques* 17. 29–54.
- Durand, Jacques. 1995. Alternances vocaliques en français du Midi et phonologie du gouvernement. *Lingua* 95. 27–50.

- Durand, Jacques. 2009. Essai de panorama phonologique : les accents du Midi. In Luc Baronian & Françoise Martineau (coord.), *Le français, d'un continent à l'autre. Mélanges offerts à Yves Charles Morin*, 123–170. Québec : Presses de l'Université Laval.
- Durand, Jacques. 2016. Corpus Phonology. In Mark Aronoff (coord.), *The Oxford Research Encyclopedia of Linguistics*, Oxford : Oxford University Press. A paraître.
- Durand, Jacques, Michel Aurnague & Julien Eychenne. 2004. La phonologie du français contemporain au Pays basque. In Maria Pilar Perea (coord.), *Dialectologia i recursos informàtics*, 155–198. Barcelona : Promociones y Publicaciones Universitarias.
- Durand, Jacques & Julien Eychenne. 2004. Le schwa en français : pourquoi des corpus? *Corpus* 3. 311–356.
- Durand, Jacques & Julien Eychenne. 2007. Remarks on schwa deletion and epenthesis in French. In *Proceedings of the 5th Journées d'Etudes Linguistiques*, 89–94.
- Durand, Jacques & Julien Eychenne. 2017a. Eléments de linguistique générale. In Detey et al. (2017), 12–22. A paraître.
- Durand, Jacques & Julien Eychenne. 2017b. Le français méridional. In Detey et al. (2017), 34–38. A paraître.
- Durand, Jacques, Ulrike Gut & Gjert Kristoffersen (coord.). 2014. *The Oxford Handbook of Corpus Phonology*. Oxford : Oxford University Press.
- Durand, Jacques, Bernard Laks, Basilio Calderone & Atanas Tchobanov. 2011a. Que savons-nous de la liaison aujourd'hui? *Langue française* (169). 103–135.
- Durand, Jacques & Chantal Lyche. 2008. French Liaison in the Light of Corpus Data. *Journal of French Language Studies* 18(01). 33–66.
- Durand, Jacques, Chantal Lyche & Bernard Laks. 2009. *Phonologie, variation et accents du français*. Paris : Hermes Science Publications.
- Durand, Jacques, Catherine Slater & Hilary Wise. 1987. Observations on schwa in Southern French. *Linguistics* 25(2). 983–1004.
- Durand, Jacques, Bernard Laks & Chantal Lyche. 2011b. Le projet PFC : une source de données primaires structurées. In Durand et al. (2009), 19–61.
- Encrevé, Pierre. 1988. *La liaison avec et sans enchaînement. Phonologie tridimensionnelle et usages du français*. Paris : Editions du Seuil.
- Ernestus, Mirjam. 2014. Acoustic reduction and the roles of abstractions and exemplars in speech processing. *Lingua* 142. 27–41.

- Eychenne, Julien. 2006. *Aspects de la phonologie du schwa dans le français contemporain : Optimalité, visibilité prosodique, gradience*. Université de Toulouse-Le Mirail, thèse de doctorat.
- Eychenne, Julien. 2007. Documentation de la plateforme PFC (version 1.0). In Sylvain Detey & Dominique Nouveau (coord.), *Bulletin PFC n° 7*, 41–54. Toulouse-Le Mirail : CLLE-ERSS.
- Eychenne, Julien. 2009a. La prononciation du français au Pays basque. In Durand et al. (2009), 259–284.
- Eychenne, Julien. 2009b. Une variété de français conservatrice en Languedoc. In Durand et al. (2009), 259–284.
- Eychenne, Julien. 2010. Conversation à bayonne : langue et identité au pays basque. In Detey et al. (2010), 131–143.
- Eychenne, Julien. 2011. La liaison en français et la théorie de l’optimalité. *Langue française* 169. 79–101.
- Eychenne, Julien. 2012. Apprenabilité de la structure syllabique du français méridional dans un réseau calculatoire dynamique standard. In *Actes du congrès mondial de linguistique française 2012*, vol. 1, 1487–1503. EDP Sciences.
- Eychenne, Julien. 2013. Inducing Suprasegmental Structure without Constituency : A Case Study on Southern French. *The Journal of Studies in Language* 29(1). 97–128.
- Eychenne, Julien. 2014. Schwa and the loi de position in Southern French. *Journal of French Language Studies* 24(02). 223–253.
- Eychenne, Julien. 2015a. De l’émergence des contraintes phonotactiques en français. *Langages* (2). 73–90.
- Eychenne, Julien. 2015b. Observations on the phonetic realization of opaque schwa in southern french. *Studies in Phonetics, Phonology and Morphology* 21. 457–494.
- Eychenne, Julien. 2015c. Towards a quantitative metric for allophonic complementarity. In *Proceedings of the 2015 International Conference on Speech Science*, 115–116. Séoul.
- Eychenne, Julien. 2016a. DecoLex : A tool for opinion annotation and classification. Dicora-tr-2016-06 Digital Language and Knowledge Contents Research Association Yongin, Corée du Sud.
- Eychenne, Julien. 2016b. Speech Corpora for Foreign Language Education : Methods and Tools. *Eoneowa Munhwa* 12(3). 1–26.

- Eychenne, Julien & Tae-Yeoub Jang. 2015. On the merger of Korean mid front vowels. *Phonetics and Speech Sciences (Journal of the Korean Society of Speech Sciences)* 7(2). 119–129.
- Eychenne, Julien & Tae-Yeoub Jang. 2016. On the role of contrastivity in the development of the /e ~ E/ merger in Korean.
- Eychenne, Julien, Chantal Lyche, Jacques Durand & Annelise Coquillon. 2014. Quelles données pour la liaison en français : la question des corpus. In Soum-Favaro et al. (2014), 33–60.
- Eychenne, Julien, Sylvain Navarro, Atanas Tchobanov & Jan Willem van Leussen. 2016. Approaching variation in PFC : the tools. In Detey et al. (2016), 387–398.
- Eychenne, Julien & Roberto Paternostro. 2016. Analyzing transcribed speech with Dolmen. In Detey et al. (2016), D35–D52.
- Eychenne, Julien & Atanas Tchobanov. 2013. Les bases de données parole. In Nguyen & Adda-Decker (2013), 121–158.
- Eychenne, Julien & Douglas Walker. 2010. La prononciation du français en Amérique du nord : éléments de synthèse. In Detey et al. (2010), 249–264.
- Fant, C. Gunnar M. 1973. *Speech Sounds and Features*. Cambridge : The MIT Press.
- Féry, Caroline. 2003. Liaison and syllable structure in French. Manuscrit.
- Feynman, Richard. 2011. *The Feynman Lectures on Physics, Vol. I : The New Millennium Edition : Mainly Mechanics, Radiation, and Heat*. New York : Basic Books first trade paper edition édition édition.
- Fitch, W. Tecumseh. 2010. *The Evolution of Language*. Cambridge & New York : Cambridge University Press.
- Flemming, Edward. 2004. Contrast and perceptual distinctiveness. In Hayes et al. (2004) 232–276.
- Flemming, Edward. 2005. Speech Perception and Phonological Contrast. In David B. Pisoni & Robert E. Remez (coord.), *The Handbook of Speech Perception*, 156–181. Malden : Blackwell.
- Fodor, Jerry A. 1983. *The modularity of mind : An essay on faculty psychology*. Cambridge : MIT press.
- Fouché, Pierre. 1956. *Traité de prononciation française*. Paris : Klincksieck. 2ième édition 1959.

- Francez, Dr Nissim & Dr Shuly Wintner. 2011. *Unification Grammars*. Cambridge & New York : Cambridge University Press.
- Gahl, Susanne. 2008. “Time” and “Thyme” Are not Homophones : The Effect of Lemma Frequency on Word Durations in Spontaneous Speech. *Language* 84(3). 474–496.
- Gendrot, Cédric & Martine Adda-Decker. 2005. Impact of duration on F1/F2 formant values of oral vowels : An automatic analysis of large broadcast news corpora in French and German. In *Proceedings of Interspeech 2005*, 2453–2456.
- Godel, Robert. 1969. *Les sources manuscrites du cours de linguistique générale de F. de Saussure*. Genève : Droz.
- Gold, E Mark. 1967. Language identification in the limit. *Information and control* 10(5). 447–474.
- Goldman, Jean-Philippe. 2011. EasyAlign : an automatic phonetic alignment tool under Praat. Proceedings of Interspeech 2011, Firenze, Italie.
- Goldrick, Matt. 2001. Turbid output representations and the unity of opacity. In M. Hiro-tani, A. Coetzee, N. Hall & J-Y Kim (coord.), *Proceedings of the Northeast Linguistics Society*, vol. 30, 231–245. Amherst : GLSA.
- Goldsmith, John. 1976. *Autosegmental phonology* : thèse de doctorat Indiana University (Bloomington) Linguistics Club.
- Goldsmith, John. 1992. Local Modeling in Phonology. In S. Davis (coord.), *Connectionism : Theory and Practice*, 229–246. Oxford : Oxford University Press..
- Goldsmith, John. 1993a. Harmonic Phonology. In Goldsmith (1993b), 21–60.
- Goldsmith, John (coord.). 1993b. *The Last Phonological Rule : Reflections on Constraints and Derivations*. Chicago : University of Chicago Press.
- Goldsmith, John. 1994. A Dynamic Computational Theory of Accent Systems. In *Perspectives in Phonology*, 1–28. Stanford : Center for the Study of Language and Information.
- Goldsmith, John & Gary Larson. 1990. Local modeling and syllabification. In *Papers from the 26th Annual Meeting of the Chicago Linguistic Society : Parasession on the syllable in phonetics and phonology*, 129–141.
- Goldsmith, John & Gary Larson. 1993. Using Networks in a Harmonic Phonology. In *Papers from the 28th Annual Meeting of the Chicago Linguistic Society*, vol. 2, 94–125. Chicago : Chicago Linguistics Society.
- Goldwater, Sharon & Mark Johnson. 2003. Learning OT Constraint Rankings Using a Maximum Entropy Model. In *Proceedings of the Workshop on Variation within Optimality Theory*, 113–122. Stockholm University.

- Grammont, Maurice. 1933. *Traité de phonétique*. Paris : Delagrave.
- Gries, Stephan Th. 2015. The most underused statistical method in corpus linguistics : Multi-level (and mixed-effects) models. *Corpora* 10. 95–125.
- Gross, Maurice. 1987. The use of finite automata in the lexical representation of natural language. In *LITP Spring School on Theoretical Computer Science*, 34–50. Springer.
- Gut, Ulrike & Holger Voormann. 2014. Corpus Design. In Durand et al. (2014) 13–26.
- Hacking, Ian. 2004. *Historical Ontology*. Harvard University Press.
- Hale, Mark & Charles Reiss. 2000. Phonology as Cognition. In Burton-Roberts et al. (2000), 161–184.
- Hall, Kathleen Currie. 2013. A typology of intermediate phonological relationships. *The Linguistic Review* 30(2). 215–275.
- Hammond, Michael. 1995. Syllable parsing in English and French. Rutgers Optimality Archive, 58.
- Han, Jeong-Im & Hyunsook Kang. 2013. Cross-generational Change of /o/ and /u/ in Seoul Korean. *Journal of the Korean Society of Speech Sciences* 5(2). 25–31.
- Han, Mun Hi & Julien Eychenne. 2017. Les coréanophones. In Detey et al. (2017), 117–123. A paraître.
- Hansen, Anita Berit. 2012. A study of young parisian speech : Some trends in pronunciation. In Randall Gess, Chantal Lyche & Trudel Meisenburg (coord.), *Phonological Variation in French : Illustrations from three continents*, 151–172. Amsterdam ; Philadelphia : John Benjamins Publishing Company.
- Hansen, Anita Berit & Caroline Juillard. 2011. La phonologie parisienne à trente ans d'intervalle –Les voyelles à double timbre. *Journal of French Language Studies* 21(03). 313–359.
- Harris, Zellig. 1951. *Structural Linguistics*. Chicago & Londres : The University of Chicago Press.
- Hayes, Bruce. 1995. *Metrical Stress Theory. Principles and case studies*. Chicago : The University of Chicago Press.
- Hayes, Bruce, Robert Kirchner & Donca Steriade (coord.). 2004. *Phonetically Based Phonology*. Cambridge & New York : Cambridge University Press.
- Hayes, Bruce & Donca Steriade. 2004. Introduction : the phonetics bases of phonological markedness. In Hayes et al. (2004), 1–33.

- Hayes, Bruce & Colin Wilson. 2008. A Maximum Entropy Model of Phonotactics and Phonotactic Learning. *Linguistic Inquiry* 39(3). 379–440.
- Heinz, Jeffrey. 2009. On the role of locality in learning stress patterns. *Phonology* 26(02). 303–351.
- Heinz, Jeffrey. 2010. Learning Long-Distance Phonotactics. *Linguistic Inquiry* 41(4). 623–661.
- Heinz, Jeffrey, Colin de la Higuera & Menno van Zaanen. 2015. Grammatical Inference for Computational Linguistics. *Synthesis Lectures on Human Language Technologies* 8(4). 1–139.
- Herdan, Gustav. 1964. *Quantitative linguistics*. Londres : Butterworths.
- Hermansky, H. & D.J. Broad. 1989. The effective second formant F2' and the vocal tract front-cavity, 480–483. IEEE. 1989 International Conference on Acoustics, Speech, and Signal Processing, 1989. ICASSP-89.
- Hieronymus, J. L. 1994. ASCII phonetic symbols for the world's languages : Worldbet. Rapport technique AT&T Bell Laboratories.
- Hockett, Charles F. 1966. The Quantification of Functional Load : A Linguistic Problem. Rapport technique DTIC Document.
- Hockett, Charles Francis. 1955. *A manual of phonology*. Baltimore : Waverly Press.
- Holliday, Jeffrey J. 2012. The acoustic realization of the Korean sibilant fricative contrast in Seoul and Daegu. *Journal of the Korean Society of Speech Sciences* 4(1). 67–74.
- Holliday, Jeffrey J. & Rory Turnbull. 2015. Effects of phonological neighborhood density on word production in Korean. Proceedings of the Eighteenth International Congress of the Phonetic Sciences.
- Holliday, Jeffrey J., Rory Turnbull & Julien Eychenne. 2016a. K-SPAN (Korean Surface Phones and Neighborhoods). <https://opendata.uit.no/dataset.xhtml?persistentId=doi%3A10.18710%2FTWM79F>.
- Holliday, Jeffrey J., Rory Turnbull & Julien Eychenne. 2016b. Surface phonetic forms and phonological neighborhood density statistics from an orthographic Korean corpus. Manuscript soumis à *Behavior Research Methods*.
- Hong, Yunsook. 1988. *A Sociolinguistic Study of Seoul Korean*. Seoul, Korea : Hanshin Publishing Co.

- Hume, Elizabeth, Kathleen Currie Hall, Andrew Wedel, Andrew Ussishkin, Martine Adda-Decker & Cédric Gendrot. 2013. Anti-markedness patterns in French epenthesis : An information-theoretic approach. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 37(1). 104–123.
- Hume, Elizabeth & Keith Johnson. 2001. A model for the interplay of speech perception and phonology. In Elizabeth Hume & Keith Johnson (coord.), *The Role of Speech Perception in Phonology*, 3–26. New York : Academic Press.
- Ide, Nancy & Keith Suderman. 2007. GrAF : A graph-based format for linguistic annotations. In *proceedings of the Linguistic Annotation Workshop*, 1–8. Association for Computational Linguistics.
- Ide, Nancy & Keith Suderman. 2014. The Linguistic Annotation Framework : a standard for annotation interchange and merging. *Language Resources and Evaluation* 10(3-4). 1–24.
- Ierusalimschy, Roberto. 2006. *Programming in Lua*. Rio de Janeiro : Lua.org.
- Ierusalimschy, Roberto, Luiz Henrique de Figueiredo & Waldemar Celes Filho. 1995. Lua – an Extensible Extension Language. *Software : Practice & Experience* 26. 635–652.
- Jäger, Gerhard & James Rogers. 2012. Formal language theory : refining the Chomsky hierarchy. *Philosophical Transactions of the Royal Society B : Biological Sciences* 367(1598). 1956–1970.
- Jansma, Judith. 2011. *Acquisition de la liaison par des apprenants néerlandophones : une étude de cas*. Université de Groningen, mémoire de mastère.
- John, Tina & Lasse Bombien. 2014. EMU. In Durand et al. (2014), 321–341.
- Johnson, C. Douglas. 1972. *Formal aspects of phonological description*. La Haye & Paris : Mouton.
- Johnson, Kent. 2004. Gold’s theorem and cognitive science. *Philosophy of Science* 71. 571–592.
- Joshi, Aravind K. 1985. Tree adjoining grammars : How much context-sensitivity is required to provide reasonable structural descriptions? In D Dowty, L Karttunen & A Zwicky (coord.), *Natural language parsing Studies in Natural Language Processing*, 206–250. Cambridge University Press.
- Jurafsky, Daniel & James H. Martin. 2008. *Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition*. Upper Saddle River, N.J : Prentice Hall.

- Kaplan, Frédéric. 2001. *La naissance d'une langue chez les robots*. Paris : Hermès Science Publications.
- Kaplan, Ronald M. & Martin Kay. 1994. Regular Models of Phonological Rule Systems. *Computational Linguistics* 20(3). 331–378.
- Karttunen, L., J-P. Chanod, G. Grefenstette & A. Schille. 1996. Regular expressions for language engineering. *Natural Language Engineering* 2(04). 305–328.
- Karttunen, Lauri. 1993. Finite-state constraints. In Goldsmith (1993b), 173–194.
- Kaye, Jonathan, Jean Lowenstamm & Jean-Roger Vergnaud. 1985. The internal structure of phonological elements : a theory of charm and government. *Phonology Yearbook* 2. 305–328.
- Kennedy, Graeme. 2014. *An Introduction to Corpus Linguistics*. Londres : Routledge.
- Kim, Hansaem. 2006. Korean national corpus in the 21st century Sejong project. In *Proceedings of the 13th National Institute of Japanese Literature (NIJL) International Symposium*, 49–54.
- Kim-Renaud, Young-Key (coord.). 1997. *The korean alphabet : Its history and structure*. University of Hawaii Press.
- King, Robert D. 1967. Functional Load and Sound Change. *Language* 43(4). 831–852.
- Kiparsky, Paul. 1971. Historical linguistics. In W. O. Dingwall (coord.), *A Survey of Linguistic Science*, 576–642. College Park : University of Maryland Linguistics Program.
- Kiparsky, Paul. 1973. Abstractness, opacity, and global rules. In Osamu Fujimura (coord.), *Three dimensions in linguistic theory*, 57–86. Tokyo : TEC.
- Kiparsky, Paul. 2000. Opacity and cyclicity. *The Linguistic Review* 17. 351–367.
- Kohler, Klaus J. 2013. Farewell. *Phonetica* 70(4). 242–246.
- Koskeniemi, Kimmo. 1983. *Two-level morphology : A general computational model for word-form recognition and production*. Université de Helsinki, thèse de doctorat.
- Labov, William. 1976. *Sociolinguistique*. Paris : Les Editions de Minuit. Traduction d'Alain Kihm.
- Labov, William. 1994. *Principles of Linguistic Change, Vol. 1 : Internal Factors*. Oxford, UK ; Cambridge Mass. : Blackwell Publishers.
- Labov, William. 2001. *Principles of Linguistic Change, Vol. 2 : Social Factors*. Blackwell Publishers.
- Laks, Bernard. 1995. A connectionist account of French syllabification. *Lingua* 95. 51–76.

- Laks, Bernard. 1997. *Phonologie accentuelle. Métrique, autosegmentalité, constituance*. Paris : CNRS Éditions.
- Laks, Bernard. 2003. Saussure's Phonology. In Henrik Galberg Jacobsen, Dorthe Bleses, Thomas O. Madsen & Pia Thomse (coord.), *Take Danish for instance. Linguistic studies in honour of Hans Basbøll*, 199–211. Odense : University Press of Southern Denmark.
- Laks, Bernard. 2005a. La liaison et l'illusion. *Langages* 39(158). 101–125.
- Laks, Bernard. 2005b. Phonologie et construction syntaxique : la liaison, un test de cohésion et de figement syntaxique. *Linx. Revue des linguistes de l'université Paris X Nanterre* (53). 155–171.
- Laks, Bernard. 2008. Pour une phonologie de corpus. *Journal of French Language Studies* 18(1). 3–32.
- Langacker, Ronald W. 2008. *Cognitive Grammar : A Basic Introduction*. Oxford & New York : Oxford University Press, USA.
- Larson, Gary. 1990. Local computational networks and the distribution of segments in the Spanish syllable. In *Papers from the 26th Annual Meeting of the Chicago Linguistic Society : Parasession on the syllable in phonetics and phonology*, .
- Larson, Gary. 1993. *Dynamic Computational Networks and the Representation of Phonological Information* : thèse de doctorat University of Chicago.
- Lee, Ho-Young. 1996. *Gugeo Eumseonghak [Korean Phonetics]*. Seoul, Korea : Taehaksa.
- Lee, Ki-Moon & S. Robert Ramsey. 2011. *A history of the korean language*. Cambridge University Press.
- Legendre, Géraldine, Yoshiro Miyata & Paul Smolensky. 1990. Harmonic grammar : A formal multi-level connectionist theory of linguistic well-formedness : Theoretical foundations. Rapport technique ICS Technical Report 90-5 University of Colorado, Boulder.
- Lindblom, Björn. 1986. Phonetic universals in vowel systems. In John J. Ohala & Jeri J. Jaeger (coord.), *Experimental phonology*, 13–44. Orlando : Academic Press.
- Luce, Paul A & David B Pisoni. 1998. Recognizing spoken words : The neighborhood activation model. *Ear & Hearing* 19(1). 1–36.
- Maddieson, Ian. 1984. *Patterns of sounds*. Cambridge university press.
- Mallet, Géraldine. 2008. *La liaison en français : Descriptions et analyses dans le corpus PFC*. Université Paris X-Nanterre, thèse de doctorat.

- Manning, Christopher D. & Hinrich Schütze. 1999. *Foundations of Statistical Natural Language Processing*. MIT Press.
- Marr, David. 1982. *Vision : a computational investigation into the human representation and processing of visual information*. Cambridge, Mass. : MIT Press.
- Martin, Samuel E. 2006. *Reference grammar of Korean : A complete guide to the grammar and history of the Korean language*. Tuttle Publishing.
- Martinet, André. 1945. *La prononciation du français contemporain. Témoignages recueillis en 1941 dans un camp d'officiers*. Genève : Droz.
- Martinet, André. 1955. *Economie des changements phonétiques. Traité de phonologie diachronique*. Paris : Maisonneuve & Larose. Réédition de 2005.
- McCarthy, John. 1999. Sympathy and Phonological Opacity. *Phonology* 16. 331–99.
- McCarthy, John. 2002. *A Thematic Guide to Optimality Theory*. Cambridge : Cambridge University Press.
- McCarthy, John & Alan Prince. 1993. Prosodic Morphology : constraint interaction and satisfaction. Rapport technique RuCSS-TR-3 Rutgers University Center for Cognitive Science New Brunswick.
- McCarthy, John & Alan Prince. 1995. Faithfulness and reduplicative identity. In Jill Beckman, Laura Walsh Dickey & Suzanne Urbanczyk (coord.), *University of Massachusetts Occasional Papers in Linguistics*, 249–384. Amherst : GLSA.
- McLeod, Peter, Kim Plunkett & Edmund T. Rolls. 1998. *Introduction to Connectionist Modelling of Cognitive Processes*. Oxford & New York : Oxford University Press.
- Montreuil, Jean-Pierre. 2002. Vestigial Feet in French. In *Proceedings of the 2002 Texas Linguistic Society Conference on Stress in Optimality Theory*, Austin : University of Texas at Austin.
- Moore, Brian C. J. 2013. Aspects of auditory processing related to speech perception. In William Hardcastle, John Laver & Fionna E. Gibbon (coord.), *The Handbook of Phonetic Sciences*, 454–488. Malden & Oxford : Wiley-Blackwell 2^e édition.
- Moran, Dermot. 2012. *Husserl's Crisis of the European Sciences and Transcendental Phenomenology : An Introduction*. Cambridge : Cambridge University Press.
- Moreton, Elliott & Paul Smolensky. 2002. Typological Consequences of Local Constraint Conjunction. In *Proceedings of the 21st West Coast Conference on Formal Linguistics*, 306–319. UC Santa Cruz : Cascadilla Press.
- Moreux, Bernard. 1985a. La “loi de position” en français du Midi. 1. Synchronie (Béarn). *Cahiers de grammaire* 9. 45–138.

- Moreux, Bernard. 1985b. La “loi de position” en français du Midi. 2. Diachronie (Béarn). *Cahiers de grammaire* 10. 95–179.
- Morin, Yves-Charles. 1987. French data and phonological theory. *Linguistics* 25. 815–843.
- Morin, Yves-Charles. 1988. De l’ajustement du schwa en syllabe fermée dans la phonologie du français. In S Paul Verluyten (coord.), *La phonologie du schwa français*, 133–90. Amsterdam : John Benjamins.
- Morin, Yves-Charles & Jonathan D. Kaye. 1982. The syntactic bases for French liaison. *Journal of Linguistics* 18(2). 291–330.
- Mougeon, Raymond & Edouard Beniak. 1991. *Linguistic Consequences of Language Contact and Restriction*. Oxford : Oxford University Press.
- Munson, B. & N. P. Solomon. 2004. The effect of phonological neighborhood density on vowel articulation. *Journal of Speech, Language, and Hearing Research* 47. 1048–1058.
- Navarro, Tomás. 1967. Nuevos datos sobre el yeísmo en España. *Thesaurus : Boletín del Instituto Caro y Cuervo* 19(1). 1–17.
- Nespor, Marina & Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht : Foris.
- New, Boris, Christophe Pallier, Ludovic Ferrand & Rafael Matos. 2001. Une base de données lexicales du français contemporain sur internet : LEXIQUE. *L’année psychologique* 101(3). 447–462.
- Newmeyer, Frederick J. 2003. Grammar is grammar and usage is usage. *Language* 79(4). 682–707.
- Nguyen, Noël & Martine Adda-Decker (coord.). 2013. *Méthodes et outils pour l’analyse phonétique des grands corpus oraux*. Paris : Hermes Science Publications.
- Nguyen, Noël & Zsuzsanna Fagyal. 2008. Acoustic aspects of vowel harmony in French. *Journal of Phonetics* 36(1). 1–27.
- Niyogi, Partha. 2006. *The Computational Nature of Language Learning and Evolution*. Cambridge, Mass. : The MIT Press.
- Niyogi, Partha & Robert C. Berwick. 1997. A dynamical systems model for language change. *Complex Systems* 11(3). 161–204.
- Oh, Yoon Mi, Christophe Coupé, Egidio Marsico & François Pellegrino. 2015. Bridging phonological system and lexicon : Insights from a corpus study of functional load. *Journal of Phonetics* 53. 153–176.
- Ohala, John J. 1981. The listener as a source of sound change. In *Papers from the parasession on language and behavior*, 178–203. Chicago : Chicago Linguistics Society.

- Oosterhof, Jantine. 2012. *Acquisition des voyelles nasales chez des apprenants néerlandophones du français*. Université de Groningen, mémoire de mastère.
- Oudeyer, Pierre-Yves. 2005. The self-organization of speech sounds. *Journal of Theoretical Biology* 233. 435–449.
- Oudeyer, Pierre-Yves. 2006. *Self-Organization in the Evolution of Speech*. Oxford ; New York : Oxford University Press.
- Papen, Robert A. 2005. Le mitchif : langue franco-crie des Plaines. In Valdman et al. (2005), 327–347.
- Pater, Joe. 2007. The locus of exceptionality : Morpheme-specific phonology as constraint indexation. *University of Massachusetts Occasional Papers* 32. 259–296.
- Pater, Joe. 2009. Weighted constraints in generative linguistics. *Cognitive Science* 33(6). 999–1035.
- Pater, Joe. 2014. Canadian raising with language-specific weighted constraints. *Language* 90(1). 230–240.
- Pierrehumbert, Janet, Mary E Beckman & D Robert Ladd. 2000. Conceptual foundations of phonology as a laboratory science. In Burton-Roberts et al. (2000), 273–304.
- Pinheiro, José & Douglas Bates. 2000. *Mixed-Effects Models in S and S-PLUS*. New York : Springer.
- Pinker, Steven. 1991. Rules of language. *Science* 253(5019). 530–535.
- Pooley, Tim. 2003. Le basque en France et en Espagne : Perspectives socio-linguistiques. *Cahiers of the Association for French Language Studies* 9(2). 19–50.
- Potts, Christopher, Joe Pater, Karen Jesney, Rajesh Bhatt & Michael Becker. 2010. Harmonic Grammar with linear programming : from linear systems to linguistic typology. *Phonology* 27(01). 77–117.
- Prince, Alan. 1983. Relating to the grid. *Linguistic Inquiry* 14(1). 19–100.
- Prince, Alan. 1993. In defense of the number i : Anatomy of a linear dynamical model of linguistic generalizations. Rapport technique RuCCS TR-1, Rutgers University Center for Cognitive Science.
- Prince, Alan & Paul Smolensky. 1993. Optimality Theory : constraint interaction in Generative Grammar. Rapport technique Ru-CCS-TR2, Rutgers University Center for Cognitive Science New Brunswick.
- Pullum, Geoffrey K. 2011. On the Mathematical Foundations of Syntactic Structures. *Journal of Logic, Language and Information* 20(3). 277–296.

- Pustka, Elissa. 2007. *Phonologie et variétés en contact. Aveyronnais et Guadeloupéens à Paris*. Tübingen : Narr.
- Racine, Isabelle, Sylvain Detey, Françoise Zay & Yuji Kawaguchi. 2012. Des atouts d'un corpus multitâches pour l'étude de la phonologie en L2 : L'exemple du projet « Interphonologie du français contemporain » (IPFC). In A Kamber & C Skupiens (coord.), *Recherches récentes en FLE*, Berne : Peter Lang.
- Rice, Curt. 2006. Nothing is a phonological fact : Gaps and repairs at the phonology-morphology interface. *North East Linguistic Society* 36.
- Rochet, Bernard. 1980. The mid-vowels in Bordeaux French. *Orbis* 29. 76–104.
- Rogers, James & Geoffrey K. Pullum. 2011. Aural Pattern Recognition Experiments and the Subregular Hierarchy. *Journal of Logic, Language and Information* 20(3). 329–342.
- Rose, Yvan. 2000. *Headedness and Prosodic Licensing in the L1 Acquisition of Phonology*. Montréal : thèse de doctorat McGill University.
- Rosenblatt, F. 1958. The perceptron : A probabilistic model for information storage and organization in the brain. *Psychological Review* 65(6). 386–408.
- Rumelhart, David E., George E. Hinton & James L. McClelland. 1988. A general framework for Parallel Distributed Processing. In Rumelhart et al. (1986a), 45–76.
- Rumelhart, David E., James L. McClelland & PDP Research Group. 1986a. *Parallel Distributed Processing : Explorations in the Microstructure of Cognition : Foundations*. Cambridge, Mass : MIT Press.
- Rumelhart, David E., James L. McClelland & PDP Research Group. 1986b. *Parallel Distributed Processing : Explorations in the Microstructure of Cognition : Psychological and Biological Models*. Cambridge, Mass : MIT Press.
- Sauzet, Patrick. 2004. La singularité phonologique du français. *Langue française* 141. 14–35.
- Schane, Sanford. 1968. *French Phonology and Morphology*. Cambridge, Massachusetts : MIT Press.
- Scheer, Tobias. 1999. A theory of consonantal interaction. *Folia Linguistica* 32. 201–237.
- Scheer, Tobias. 2004. Présentation du volume. En quoi la phonologie est vraiment différente. *Corpus* 3. 5–84.
- Schmid, Helmut. 1994. Probabilistic Part-of-Speech Tagging Using Decision Trees. In *Proceedings of International Conference on New Methods in Language Processing*, Manchester, UK.

- Schmid, Helmut. 1995. Improvements in Part-of-Speech Tagging with an Application to German. In *Proceedings of the ACL SIGDAT-Workshop*, Dublin, Ireland.
- Schmidt, Thomas & Kai Wörner. 2014. EXMARaLDA. In Durand et al. (2014), 402–419.
- Schwartz, Jean-Luc, Louis-Jean Boë, Nathalie Vallée & Christian Abry. 1997. The Dispersion-Focalization Theory of vowel systems. *Journal of Phonetics* 25(3). 255–286.
- Scobbie, James M & Jane Stuart-Smith. 2008. Quasi-phonemic contrast and the fuzzy inventory : Examples from Scottish English. In Peter Avery, B. Elan Dresher & Keren Rice (coord.), *Contrast in phonology : Theory, perception, acquisition*, 87–114. Berlin and New York : Mouton de Gruyter.
- Scullen, Mary Ellen. 1997. *French Prosodic Morphology. A Unified Account*. Bloomington : Indiana University Linguistics Club.
- Selkirk, Elisabeth O. 1980. The Role of Prosodic Categories in English Word Stress. *Linguistic Inquiry* 11(3). 563–605.
- Selkirk, Elizabeth. 1972. *The Phrase Phonology of English and French* : thèse de doctorat MIT.
- Selkirk, Elizabeth. 1978. The French foot : On the status of “mute” e. *Studies in French Linguistics* 1(2). 141–150.
- Shannon, Claude Elwood. 1948. A mathematical theory of communication. *The Bell System Technical Journal* 27. 379–423.
- Shieber, Stuart M. 1985. Evidence against the context-freeness of natural language. *Linguistics and Philosophy* 8(3). 333–343.
- Shin, Jiyoung, Jieun Kiaer & Jaeeun Cha. 2013. *The sounds of korean*. Cambridge : Cambridge University Press.
- Silverman, Daniel. 2010. Neutralization and anti-homophony in Korean. *Journal of Linguistics* 46(02). 453–482.
- Simon, Anne-Catherine, Philippe Francard & Philippe Hambye. 2014. The VALIBEL Speech Database. In Durand et al. (2014), 552–561.
- Sipser, Michael. 2005. *Introduction to the Theory of Computation*. Boston : Cengage Learning.
- Sjölander, K & J Beskow. 2009. Wavesurfer [Computer program]. Version 6.0.20. <http://www.speech.kth.se/wavesurfer>.

- Smolensky, Paul. 1983. Schema Selection and Stochastic Inference in Modular Environments. In *Proceedings of AAAI-83*, 378–382.
- Smolensky, Paul. 1986. Information processing in dynamical systems : foundations of harmony theory. In Rumelhart et al. (1986a), 194–281.
- Smolensky, Paul & Géraldine Legendre (coord.). 2006a. *The Harmonic Mind. From neural computation to optimality-theoretic grammar*, vol. 1 : cognitive architecture. Cambridge : MIT Press.
- Smolensky, Paul & Géraldine Legendre (coord.). 2006b. *The Harmonic Mind. From neural computation to optimality-theoretic grammar*, vol. 2 : linguistic and philosophical implications. Cambridge : MIT Press.
- Soderstrom, Melanie, Donald W. Mathis & Paul Smolensky. 2006. Abstract genomic encoding of Universal Grammar in Optimality Theory. In Smolensky & Legendre (2006b), 403–471.
- Sohn, Ho-Min. 1999. *The Korean Language*. Cambridge : Cambridge University Press.
- Sorace, Antonella & Frank Keller. 2005. Gradience in linguistic data. *Lingua* 115(11). 1497–1524.
- Soum-Favaro, Christiane, Annelise Coquillon & Jean-Pierre Chevrot. 2014. *La Liaison : Approches Contemporaines*. Berne : Peter Lang.
- Stokes, Stephanie F. 2010. Neighborhood density and word frequency predict vocabulary size in toddlers. *Journal of Speech, Language, and Hearing Research* 53. 670–683.
- Surendran, Dinoj & Partha Niyogi. 2003. Measuring the usefulness (functional load) of phonological contrasts. Rapport technique Technical Report TR-2003.
- Surendran, Dinoj & Partha Niyogi. 2006. Quantifying the functional load of phonemic oppositions, distinctive features, and suprasegmentals. *Amsterdam studies in the theory and history of linguistic science. Series 4* 279. 43.
- Syrdal, Ann K & Hundrai S Gopal. 1986. A perceptual model of vowel recognition based on the auditory representation of American English vowels. *The Journal of the Acoustical Society of America* 79(4). 1086–1100.
- Taylor, John R. 2003. *Cognitive Grammar*. Oxford & New York : Oxford University Press.
- TEI Consortium. 2014. TEI P5 : Guidelines for Electronic Text Encoding and Interchange. Version 2.6.0. Rapport technique.
- Thagard, Paul. 2005. *Mind : Introduction to Cognitive Science*. Cambridge, Mass : Bradford 2^e édition.

- The Unicode Consortium. 2016. *The unicode standard, version 9.0.0*. The Unicode Consortium. <http://www.unicode.org/versions/Unicode9.0.0/>.
- Tomasello, Michael. 2003. *Constructing a Language : A Usage-Based Theory of Language Acquisition*. Cambridge, Mass. : Harvard University Press.
- Tranel, Bernard. 1981. *Concreteness in Generative Phonology. Evidence from French*. Berkeley : University of California Press.
- Tranel, Bernard. 2000. Aspects de la phonologie du français et la théorie de l'optimalité. *Langue française* 126. 39–72.
- Trask, Robert Lawrence. 1997. *The history of basque*. Londres : Routledge.
- Trautmüller, Hartmut. 1990. Analytical expressions for the tonotopic sensory scale. *The Journal of the Acoustical Society of America* 88(1). 97–100.
- Tremblay, Annie & Elsa Spinelli. 2014. Utilisation d'indices acoustico-phonétiques dans la reconnaissance des mots en contexte de liaison. In Soum-Favaro et al. (2014), 93–116.
- Troubetzkoy, Nicolas S. 1938. *Principes de phonologie*. Klincksieck. Traduction de Jean Cantineau (révisée par Jorge Prieto).
- Turcsan, Gabor. 2005. *Le mot phonologique en français du Midi. Domaines, contraintes, opacité* : thèse de doctorat Université de Toulouse-Le Mirail.
- Twaddell, W. Freeman. 1935. On Defining the Phoneme. *Language* 11(1). 5–62.
- Valdman, Albert, Julie Auger & Deborah Piston-Hatlen (coord.). 2005. *Le Français en Amérique du Nord*. Sainte-Foy, Québec : Presses de l'Université Laval.
- van Oostendorp, Marc. 2000. *Phonological Projection : A Theory of Feature Content and Prosodic Structure*. Berlin & New York : Mouton De Gruyter.
- van Oostendorp, Marc. 2008. Incomplete devoicing in formal phonology. *Lingua* 118(9). 1362–1374.
- van Oostendorp, Mark. 2005. The Theory of Faithfulness. Ms, Meertens Institute.
- Walter, Henriette. 1977. *La phonologie du français*. Paris : Presses Universitaires de France.
- Watbled, Jean-Philippe. 1995. Segmental and suprasegmental structure in Southern French. In John Charles Smith & Martin Maiden (coord.), *Linguistic Theory and the Romance Languages*, vol. 122, 181–200. Amsterdam : John Benjamins.

- Wedel, Andrew, Scott Jackson & Abby Kaplan. 2013a. Functional Load and the Lexicon : Evidence that Syntactic Category and Frequency Relationships in Minimal Lemma Pairs Predict the Loss of Phoneme contrasts in Language Change. *Language and Speech* 56(3). 395–417.
- Wedel, Andrew, Abby Kaplan & Scott Jackson. 2013b. High functional load inhibits phonological contrast loss : A corpus study. *Cognition* 128(2). 179–186.
- Weinberg, Steven. 1976. The forces of nature. *Bulletin of the American Academy of Arts and Sciences* 24(4). 13–29.
- Wioland, François. 1985. *Les structures syllabiques du français*. Genève & Paris : Editions Slatkine.
- Wright, R. 2004. Factors of lexical competition in vowel articulation. In J. Local & R. Ogden (coord.), *Papers in Laboratory Phonology 6*, 26–50. Cambridge : Cambridge University Press.
- Yang, Byunggon. 1992. An acoustical study of Korean monophthongs produced by male and female speakers. *The Journal of the Acoustical Society of America* 91(4). 2280–2283.
- Yang, Byunggon. 1996. A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics* 24(2). 245–261.
- Zwirner, Eberhard & Kurt Zwirner. 1970. *Principles of Phonometrics*. Alabama : University Press Of Alabama. Traduction de Hermann Bluhme.