



Volume 1 : Mémoire

Habilitation à Diriger des Recherches

Discipline : Sciences du Langage
Mémoire présenté et soutenu publiquement le 16 novembre 2017

Corine Astésano
(URI Octogone-Lordat, UT2J)

**Le statut de l'accent initial dans la phonologie prosodique du français
– Enjeux descriptifs et psycholinguistiques –**

JURY

Ellen Gurman BARD
Mireille BESSON
Elisabeth DELAIS-ROUSSARIE
Jacques DURAND
D. Robert LADD
Jean-Luc NESPOULOUS

Professeur, Université d'Edinburgh (GB)
Directrice de Recherche CNRS, Aix-Marseille Université
Directrice de Recherche CNRS, Université de Nantes – *Rapporteur*
Professeur émérite, Université Toulouse Jean Jaurès – *Rapporteur*
Professeur émérite, Université d'Edinburgh (GB) – *Rapporteur*
Professeur émérite, Université Toulouse Jean Jaurès – *Garant*

*LE STATUT DE L'ACCENT INITIAL DANS LA
PHONOLOGIE PROSODIQUE DU FRANÇAIS*

— ENJEUX DESCRIPTIFS ET PSYCHOLINGUISTIQUES —

Corine Astésano
URI Octogone-Lordat
Université de Toulouse 2

A mes parents, qui m'ont soutenue jusque dans mes errances musicales

Table des matières

Introduction	11
<i>PRELUDE</i>	17
Considérations théoriques et méthodologiques	19
1. les fonctions de la prosodie	19
2. Les paramètres de la prosodie	20
2.1. Paramètres phonologiques.....	20
2.2. Paramètres physiques	20
2.3. Mise en relation des niveaux d'analyse.....	21
2.4. Mesures des paramètres prosodiques.....	22
2.5. Pourquoi l'intonation occupe-t-elle une place centrale dans les modèles prosodiques ?.....	23
3. Phonologie intonative et théorie Métrique - Eléments définitoires	24
3.1. Principes de la représentation métrique de la prosodie	24
3.2. Niveaux de constituance métriques et indépendance prosodie/syntaxe	25
3.3. Sur la notion d'« Accent »	27
3.4. Les principes de l'approche Métrique Autosegmentale (AM).....	29
Descriptions prosodiques du français	33
4. Particularités accentuelles du français	33
5. Contexte théorique	34
5.1. Approches intonosyntaxiques vs métriques et AM du français.....	34
5.2. Les niveaux de constituance en français et leurs liens avec IA et FA.....	37
5.3. Fonctions de l'accentuation en français.....	42
5.4. Quel domaine pour rendre compte de l'accentuation en français ?	44
5.5. En finir avec la notion de <i>français, langue sans accent</i> ? Implications pour les statuts de IA et FA.....	50
6. Conclusion	56
Modèles neuro- et psycholinguistiques du traitement du langage	59
7. Bases neurales du langage	59
8. Les modèles pré-lexicaux de reconnaissance des mots et d'accès au sens	61
8.1. Limites des modèles psycholinguistiques pour le français	63
8.2. Notre proposition pour les modèles de traitement de la parole.....	66

FUGUE	69
Pour un ancrage métrique des descriptions prosodiques du français	71
9. Mètre, rythme et paramètre de durée	71
9.1. Notion de 'battement' métrique ('beat')	72
9.2. Perception du mètre et du rythme – prise en compte nécessaire de la durée en français	74
10. La perception comme fenêtre sur la proéminence cognitive	76
11. Quelle approche pour rendre compte de la prosodie du français ?	77
11.1. L'approche métrique de Di Cristo (2000)	78
11.2. Caractéristiques acoustiques de l'accentuation (Astésano, 2001)	79
11.3. Une approche métrique mixte intégrant les apports de AM	84
Statut phonologique de IA et FA – Résultats expérimentaux	87
12. Le rôle de IA dans la structuration prosodique du français – Etudes de production et de perception	87
12.1. Représentation phonologique de IA : une étude exploratoire de perception (Jankowski <i>et al</i> , 1999)	87
12.2. IA marqueur de structure (Astésano <i>et al</i> , 2007 ; 2003 ; 2002)	88
12.3. Réalisations d'IA en plateau (Welby <i>et al</i> , 2016)	92
13. Perception des frontières, des proéminences et du phrasé prosodique	95
13.1. Perception de IA et FA et leurs liens avec les frontières prosodiques (Astésano <i>et al</i> , 2012 ; 2013 ; en cours)	96
13.2. Perception de FA produit comme L ou !H à l'intérieur d'un constituant (Mendez & Astésano, 2017 ; Mendez, en cours)	100
13.3. Perception de IA et alignement tonal de H ₁ ; Perception de DFA en lien avec ses corrélats acoustiques (Astésano & Bertrand, 2016 ; Garnier ; Mendez, en cours)	103
13.4. Conclusion intermédiaire sur le statut phonologique de IA	105
13.5. Perception du phrasé prosodique (Garnier <i>et al</i> , 2016 a&b ; Garnier, en cours)	106
13.6. Analyse des syntagmes dits « d'Edinburgh » dans une tâche de MapTask (Bard <i>et al</i> , 2013 ; Gorish <i>et al</i> , 2014)	110
13.7. Perception de IA et FA, et liens avec les frontières en parole spontanée (Roux <i>et al</i> , 2016 ; en cours)	112
14. Bases neurales de la prosodie	114
14.1. Spécialisation hémisphérique du traitement de la prosodie	115
14.2. Traitement des modalités (Astésano <i>et al</i> , 2004a)	116
14.3. Traitement du focus (Magne <i>et al</i> , 2004 ; 2005a)	119

14.4. Traitement du mètre (Astésano <i>et al</i> , 2004c ; Magne <i>et al</i> , 2007).....	121
14.5. Traitement de l'accent Initial IA (Astésano <i>et al</i> , 2013; Aguilera <i>et al</i> , 2014 ; te Rietmolen <i>et al</i> , 2016 a, b, c)	124
STRETTE	135
Quelles implications pour les modèles phonologiques et psycholinguistiques ?	137
15. Principales tendances	137
16. Proposition de pied spondéïque marqueur du <i>pw</i>	140
17. La surdit� accentuelle en question	142
18. Le <i>pw</i> comme base m�trique pour les mod�les psycholinguistiques	143
VARIATIONS	145
19. Cognition compar�e du Langage et de la Musique.....	147
19.1. Rythme musical et linguistique.....	148
19.2. Interactions Langage & Musique	149
20. Prosodie et pathologies de la parole	155
1.1 La prosodie � l'aide de l'acc�s au sens dans les cas d'aphasie d�veloppementale	157
1.2 Sp�cialisation h�misph�rique du traitement de la prosodie chez des patients atteints de tumeurs c�r�brales envahissantes.....	159
20.1. Conservation des fonctions prosodiques apr�s traitement des cancers de la cavit� buccale et du pharynx	163
21. Prosodie et apprentissage d'une Langue Seconde.....	167
21.1. Impact de l'entra�nement prosodique et MVT sur la fluence en langue �trang�re	167
21.2. Mesure de la capacit� d'imitation prosodique en langue maternelle et langue �trang�re	169
22. Autres recherches en phon�tique	172
CODA	177
BIBLIOGRAPHIE	183

Un tel travail de recherche est forcément le fruit de riches interactions avec les collègues, depuis les débuts de ma carrière à Aix-en-Provence jusqu'à aujourd'hui à Toulouse.

Je commencerai par les Aixois, avec lesquels j'ai grandi scientifiquement : Ludovic Jankowski, dit Loud, Comte d'Outre-Temps, pour l'inspiration et l'admiration que suscite son travail, arrêté bien trop tôt. Si tu avais continué, la face de la prosodie en France aurait été changée. Roxane Bertrand sans laquelle ce travail ne serait pas ce qu'il est. Elle a su me pondérer, toujours avec délicatesse, et m'a suivie dans cette folie de révision du système prosodique français avec enthousiasme. C'est elle qui m'a permis de faire le pont entre mon approche très ancrée métrique et l'AM, c'est encore elle qui m'a ouvert les portes sur la parole spontanée. Pour tout ça, pour tous les fous rires (parfois nerveux après de longues heures de brainstorming), ma plus amicale reconnaissance. Robert Espesser pour les très longues séances de quête du meilleur modèle statistique, pour les fous-rires également, et les moments d'incompréhension. Depuis le tout début, je te dois tout, dont mes pires crises d'angoisse. Alain Ghio, l'éternel collaborateur et néanmoins ami, dont l'enthousiasme et la perspicacité m'ont plus d'une fois aidée à dépasser mes craintes et mes doutes : tout semble simple avec toi. Pauline Welby, pour nos discussions passionnantes sur la prosodie et l'AM, pour sa gentillesse et son attention permanente. Cristel Portes, pour les mêmes discussions et aussi pour notre équipée réjouissante à la 'Thelma et Louise' à Dublin 2014. Et puis, il y a les compagnons de route historique : Christine, Yohann, Noël, Piet, Béa, Muriel ... Et les mentors : Albert Di Cristo, Daniel Hirst, Mario Rossi, Bernard Teston...

Il y a les Marseillais : Mireille Besson, avant tout, qui m'a énormément appris après ma thèse et avec laquelle j'ai pris un immense plaisir à collaborer. Merci pour ta rigueur et ton amitié, ta joie de vivre et ta bienveillance. Je te dois d'être là aujourd'hui à Toulouse. Il y a aussi Daniele Schön, Cyrille Magne, et Radouane el Yagoubi que j'ai la chance d'avoir maintenant pour collègue toulousain. Je souhaite que nos recherches durent encore très longtemps.

Il y a les Edinbourgeois : Ellen Gurman Bard et Alice Turk auxquelles je dois une grande partie de ma recherche aujourd'hui. C'est à Edinburgh que j'ai appris que rigueur et exigence scientifiques riment également avec humour et bienveillance. Quel bonheur d'avoir appris à vos côtés ! C'est à Edinburgh aussi que j'ai rencontré Bob Ladd pour lequel j'ai également une grande admiration. Je soumets mon travail à son jugement avec une grande fébrilité.

Il y a enfin les Toulousains : Jean-Luc Nespoulous, avant tout, qui m'a permis de remonter en selle après quelques années de découragement. A l'époque, il n'y avait bien que toi en France que mon parcours interdisciplinaire n'effrayait pas. Je t'en serai éternellement reconnaissante. J'espère faire honneur à la confiance que tu as su m'accorder. Lorraine Baqué, assimilée Toulousaine, ma sœur scientifique et spirituelle (à tous les sens du terme), ma très chère amie (tu peux mettre sur cette dernière phrase tous les accents initiaux à valeur emphatique que tu veux). Michel Billières, notre frère de bonne chère et de bons mots. Mes collègues et amies de laboratoire/département : Barbara Köpke, pour sa pugnacité et ténacité qui nous ont permis, contre toute attente, de passer les pires tempêtes institutionnelles. Merci pour ton soutien et ton amitié. Christiane Soum, dont la zenitude et la sagesse sont un modèle pour moi. Mélanie Jucla, parce que c'est un vrai bonheur de partager mon bureau et mon travail avec quelqu'un d'aussi pétillant ! Nathalie Rossi-Gensane et ses anacoluthes, avec laquelle j'ai beaucoup ri et dont je regrette le départ de Toulouse : quel humour décalé réjouissant ! Comment pourrais-je oublier, enfin, la 'fine équipe' de doctorants constituant depuis quelques années le 'groupe parole' du laboratoire : Laury Garnier, Olivier Nocaudie, Noémie te Rietmolen et Rogelio Mendez. Ils partagent non seulement ma passion pour la prosodie mais aussi une certaine vision de la recherche, de la camaraderie, de la bonne chère et du travail d'équipe. J'ai énormément de chance de vous avoir à mes côtés. Merci aussi à Charlotte, une de mes premières doctorantes que j'ai le plaisir d'avoir pour collègue maintenant.

Merci à Bach, Satie, Chopin, Pixies, Radiohead, Noir Désir, Chet Baker, Bill Evans, Al Jarreau, Michael 'the king of Pop', et tant d'autres, qui m'ont accompagnée pendant la rédaction de ce document au gré de mes humeurs.

Le français tient une place particulière parmi les langues du monde. Elle ferait partie des rares langues ‘exotiques’ sans tons ni accentuation lexicale. L’essence même du caractère métrique des langues, selon lequel la parole est organisée par l’alternance régulière de proéminences hiérarchiquement structurées, lui serait donc refusée. Certes, les Français n’en sont pas à un ‘paradoxe’ ou une ‘exception’ près, mais cela concerne en général leur façon de cuisiner et de consommer du vin, ou leur vision fantasmée d’une hégémonie culturelle mondiale. Dès le début de ma carrière scientifique, cette ‘exception’ prosodique m’a dérangée. Ne s’agissait-il pas encore d’une illusion, linguistique cette fois, nous maintenant en marge du reste du monde ?

Musicienne à mes heures, j’avais la sensation en entendant ma langue d’une récurrence régulière de battements donnée par les proéminences. Certes, le rythme syllabique est présent, certes l’allongement final aux frontières intonatives majeures est prégnant perceptivement, mais il me semblait qu’il y avait une vie accentuelle à part entière dans cette langue. Fallait-il être musicien ou encore de langue maternelle étrangère comme l’un des plus éminents prosodistes francophones, Iván Fónagy, pour entrevoir la rémanence métrique du français en surface ?

J’ai donc entrepris ma thèse dans les pas d’Albert Di Cristo à l’institut de Phonétique d’Aix-en-Provence, avec l’idée de tester expérimentalement les propositions métriques de son modèle. Nous étions en quête d’invariants rythmiques et accentuels constituant le système noyau du français, à travers l’étude de différents styles de parole allant de la lecture d’un conte au journal radiophonique en passant par l’entretien libre radiophonique (interviews). D’approche métrique, ce travail ne se basait cependant pas sur la transcription métrique des ‘textes’, mais sur un repérage perceptif des phénomènes accentuels et de frontières. Le prisme de la perception nous permettait par ailleurs de dépasser le cadre d’analyse habituel du syntagme et de la phrase dans ce type d’approches, et d’envisager de manière plus libre le lien entre proéminences et constituance. Le modèle de Di Cristo rompait également avec les traditions de descriptions intonosyntaxiques du français, et permettait de questionner le caractère lexical de l’accent en français, à un certain niveau de représentation. Les résultats de ce travail m’ont permis d’envisager un rôle plus important de l’accent initial (IA), sans doute moins ‘secondaire’ que ce que l’ensemble des modèles prosodiques lui conférait. Il participait de plein droit avec l’accent final (FA) à la structuration rythmique et métrique de la langue, et marquait en surface une unité prosodique proche du mot lexical, particulièrement en parole

spontanée. Je ne parle pas ici de la version emphatique de IA à laquelle on fait le plus souvent référence, mais bien de sa version ‘rythmique’ et métrique.

La conclusion de mon travail de thèse posait donc la question du rôle que devait jouer IA dans le découpage du flot de parole en unités linguistiques interprétables par l’auditeur. A la fin des années 90, très peu d’études s’intéressaient au rôle de l’accentuation dans l’accès au sens en français, de par son caractère post-lexical supposé qui ne cadrerait pas avec les modèles psycholinguistiques. Plus grave, le bruit commençait à courir que les francophones seraient ‘sourds’ à l’accent métrique (*stress*) en langue étrangère. Certes, cela permettait d’expliquer pourquoi les français sont si mauvais en langue étrangère : leur système phonologique ne leur permettrait pas de se représenter l’accent lexical, et donc de le percevoir et de l’implémenter en langue étrangère¹. Malheureusement, cette proposition a été surinterprétée, et généralisée à la capacité des francophones à percevoir *leur propre système accentuel*. Cette étude (et les suivantes) devenait la preuve de l’absence même de représentation phonologique de l’accent en français, alors que ce n’est pas ce qu’elle testait ! Ces recherches ont donc eu un impact² délétère sur le développement des recherches en français, aussi bien en psycholinguistique qu’en phonologie, d’ailleurs. Le défi est en effet, dans ces conditions, très difficile à relever dans les deux domaines.

Le statut ambigu de l’accentuation en français émane également d’une surinterprétation de l’article de Rossi ‘*le français, langue sans accent*’, nous le verrons. Cette ambiguïté constitue une ‘épine dans le pied’ des linguistes et psycholinguistes, sur laquelle on met le plus souvent un pansement pour avancer tout de même dans la description du fonctionnement de cette langue. Les modèles prosodiques et psycholinguistiques reposent le plus souvent sur des préceptes élaborés à partir de l’anglais et d’autres langues à accentuation lexicale. Ces modèles semblent correspondre à la description d’un nombre non négligeable de langues, au moins les langues à accentuation les mieux décrites. On s’accommode des contradictions générées par les tentatives de modélisation du français dans ces modèles, comme lorsque l’on manipule la notion d’« accent » alors même que son existence est, à bien des égards et à bien des niveaux, remise en question. On s’accommode de cette même notion lorsqu’on montre que l’accentuation est utilisée pour segmenter et accéder au sens, alors même qu’on ne reconnaît pas l’existence de l’accent au niveau des mots que l’on manipule expérimentalement.

¹ Une autre interprétation est cependant tout à fait possible, relative encore une fois à « l’exception française » : notre hégémonie linguistique et culturelle passée est inscrite dans nos gènes, dans notre système éducatif, nous dispensant moralement de faire l’effort de parler d’autres langues. L’hégémonie linguistique et culturelle actuelle du monde anglophone nous livre d’ailleurs le même constat : à quelques exceptions près, les locuteurs anglophones ne brillent pas non plus par leurs capacités en langue étrangère... Tout ceci relève donc sans doute plus de la politique que de la linguistique.

² Factor ? ...

Finally, the question that has been bothering me since then is the following : is French not fitting into these models, is it really because it is an exotic language ? Or is it because the syncretism between accentuation and intonation, a phenomenon so particular to French, obscures the analysis of the domains governed by accentuation in this language, domains on which the descriptive models are based ? The question may seem idiotic, and somewhat provocative, but how can one suppose that a tiny part of humanity would have a different cognitive functioning from its congeners ? Even the cultural exception cannot explain this !

It is therefore to this epistemological contradiction, on which one throws a veil of pudicity when one works on French, that I wanted to focus after my thesis. To determine the contours, I had to adopt an interdisciplinary approach, at the crossroads of linguistics and of psycho- and neurolinguistics.

My interdisciplinary formation was made around two major collaborations : the first with Mireille Besson at CNRS in Marseille in cognitive neuroscience, and the second with Ellen Gurman Bard and Alice Turk at the University of Edinburgh. The two post-doctoral fellowships with Mireille Besson led me to become interested in the neural bases of prosody and in the links between prosodic system and music. It was at this occasion that I learned to familiarize myself with theoretical and methodological notions of brain imaging, particularly Electroencephalography (EEG) which is still today a methodological tool that I value for the evidence of phonological processing of accentuation. These two post-doctoral fellowships with Mireille Besson were interrupted by a two-year post-doctoral stay at the University of Edinburgh, where Ellen Bard, Alice Turk and I ourselves searched to precisely bring to light the rules of IA occurrences as a function of structural and prosodic constraints. It was first to judge of the metric reality of IA in the prosodic structuring of French before its potential role in access to meaning. Before my recruitment in the laboratory Lordat directed by Jean-Luc Nespoulous in 2006, I had thus been able to advance on the evidence of the metric reality of accentuation in French : on the one hand, a study with Mireille Besson had allowed us to show that the meter is processed in French and that its manipulation influences the semantic processing efficiency ; on the other hand, the study with Ellen Bard and Alice Turk on the rules of IA occurrence showed that IA is systematically used as a structure marker, from the level of the *prosodic word*, i.e. a prosodic unit close to the lexical word.

L'intégration dans l'équipe de Jean-Luc Nespoulous m'a permis de m'ouvrir encore à d'autres pans de recherche en prosodie, notamment sur le versant applicatif, en apprentissage d'une langue seconde et en pathologies de la parole. C'est là que je me suis rendu compte (avec un certain désarroi) de l'impact des présupposés théoriques sur le français dans ces deux domaines applicatifs. Comment enseigner la prosodie d'une langue lorsqu'on ne lui reconnaît pas de système accentuel? Comment monter des protocoles de tests et de remédiation efficaces pour les diverses pathologies de la parole, quand les praticiens pensent qu'« il n'y a pas d'accent et de prosodie en français »? Voici quelques questions, variations autour de mon thème de recherche central, qui contribuent néanmoins à lui apporter un éclairage très informatif.

Mon enquête sur le système accentuel du français a continué en parallèle et a pu considérablement se développer lorsque j'ai pu me recentrer sur ma recherche lors d'une délégation CNRS de deux ans à Aix-en-Provence. Cela a été l'occasion de déposer et d'obtenir (en collaboration avec Roxane Bertrand au LPL) un financement ANR sur le 'rôle de IA dans la structuration prosodique du français : de la phonologie au traitement de la parole' (*PhonIACog*, 2013-16). Ces conditions de recherche m'ont permis de construire une dynamique autour d'une petite équipe de doctorants et post-doctorants avec lesquels nous avons bien progressé dans notre compréhension du système prosodique du français. Nous travaillons aux confins de l'interdisciplinarité si chère à Jean-Luc Nespoulous³ et moi-même, aussi bien sur les versants linguistiques (description du système sur de gros corpus variés), que psycho- et neurolinguistiques (études de perception et de neuroimagerie).

Ce document va donc présenter ces différentes étapes de ma recherche. J'ai emprunté au vocabulaire musical contrapuntique pour la construction de ce document. J'aborderai en *Prélude* quelques notions relatives au contexte théorique et méthodologique de mon travail. J'y discuterai notamment du statut ambigu de l'accentuation du français (IA et FA) dans les différents modèles (intonosyntaxiques, métriques ou issus de la Métrique Autosegmentale AM). La *Fugue*, au centre de ce document, présente mon approche théorique, mêlant l'ancrage métrique fort de Di Cristo et les outils conceptuels et descriptifs puissants de AM. Les principaux résultats y sont présentés sur les trois versants d'investigation (production, perception et neuroimagerie) : IA y occupe une grande place mais certains résultats éclairent également le comportement métrique de FA à tous les niveaux de constituance. Cette *Fugue* est suivie d'une courte *Strette*, (*i.e.* ici d'une conclusion intermédiaire), permettant

³ Je fais référence ici au fameux triptyque des niveaux d'analyse en neuropsycholinguistique de Jean-Luc Nespoulous, où la linguistique siège au cœur du triptyque comme base essentielle à toute connaissance psycho-et neurolinguistique.

d'envisager les implications de nos propositions sur les modèles prosodiques et psycholinguistiques pour le français. Les *Variations* présentent les travaux que j'ai pu mener en neurosciences cognitives sur la comparaison 'langage et musique', où je discute de l'utilité de l'interface prosodique, ainsi que ceux menés en L2 et pathologies du langage. A chaque étape de présentation des travaux, nous renvoyons le lecteur aux articles les illustrant, présentés en annexe. Enfin, une *Coda* tentera d'envisager les pistes de recherche à venir.

PRELUDE

Considérations théoriques et méthodologiques

Ce chapitre *Prelude* n'a certainement pas pour vocation de présenter ce qu'est la prosodie: un ouvrage entier n'y suffirait d'ailleurs pas et je serais bien en peine d'en esquisser les bases. Il s'agit pour moi de présenter le cadre dans lequel j'envisage la prosodie du français, et l'accentuation en particulier. Pour cela, je commence par un rapide aperçu des fonctions de la prosodie, et de ses principaux niveaux d'analyse et paramètres, afin de situer son étude dans le champ de la linguistique. J'en arrive ensuite à la question de la représentation de la prosodie, au travers des approches métrique et Métrique Autosegmentale (AM) dans lesquelles mes travaux s'inscrivent. Vient enfin une présentation de l'accentuation du français à travers les principales approches descriptives (intonosyntaxiques, métriques et AM) dont émerge une pluralité de conceptions de l'accentuation participant de son ambiguïté. Le *Prelude* s'achève par la présentation des modèles psycho- et neurolinguistiques de traitement du langage et des limites auxquelles ils sont confrontés pour rendre compte du fonctionnement en français.

1. LES FONCTIONS DE LA PROSODIE

La prosodie entre en jeu dès lors que le langage est oralisé⁴. La substance même de la parole se définit par la prosodie : elle est extralinguistique, en ce sens qu'elle véhicule les caractéristiques idiosyncrasiques du locuteur (régionales, sexuelles etc...); elle est paralinguistique quand elle véhicule l'affect, les émotions du locuteur et son état d'esprit dans la situation de communication (ironie, sarcasme, neutralité...); elle est enfin linguistique lorsqu'elle participe de l'actualisation et la structuration des différents niveaux linguistiques (lexique, morpho-syntaxe, sémantique, pragmatique etc...). Si je rappelle ici brièvement les fonctions de la prosodie, c'est d'une part parce que sa multifonctionnalité la rend difficile à appréhender, mais aussi parce que, pour des raisons historiques et épistémologiques que je ne vais pas développer, elle a paradoxalement longtemps été considérée comme la « *Cendrillon de la linguistique* » (voir Bolinger, 1986, Fónagy, 1989, et Di Cristo, 2012 pour une discussion). Longtemps son analyse a été considérée comme redondante par rapport à l'information syntaxique, bien qu'il soit maintenant largement admis une certaine indépendance entre prosodie et syntaxe (Shattuck-Huffnagel & Turk, 1996), notamment lorsque la prosodie a un rôle de mise en valeur sémantique et pragmatique (« *C'est sans doute dans la mesure où elle assume ce rôle qu'elle échappe à la syntaxe.* » Rossi, 1985a :144).

⁴ On parle cependant aussi de prosodie en lecture silencieuse (Fodor, 2002a) et en langue des signes.

Longtemps, également, a-t-on réduit la prosodie à ses fonctions para- et extralinguistiques. La domination de la syntaxe en linguistique et la complexité fonctionnelle inhérente de la prosodie font que la prosodie trouve une place encore très marginale dans les modèles de traitement du langage.

2. LES PARAMETRES DE LA PROSODIE

Il faut distinguer les paramètres phonologiques de la prosodie (accentuation, intonation, rythme) des paramètres acoustico-phonétiques qui permettent de les actualiser (fréquence fondamentale ou f_0 , durée, intensité).

2.1. Paramètres phonologiques

Les principaux modèles prosodiques s'articulent autour des représentations de l'accentuation et de l'intonation, et de leurs relations avec la structuration du texte ou du message. L'accentuation rend compte de la structuration métrique de la langue, *i.e.* de la distribution des syllabes métriquement fortes (proéminences) dans un domaine grammatical, en général le mot lexical. L'intonation rend compte de la structuration mélodique des énoncés, de leur organisation tonale. Le rythme peut être considéré comme '*l'organisation temporelle des proéminences*' et constituerait donc le socle métrique prééminent à l'intonation (Di Cristo, 2000 ; Astésano, 2001). Nous verrons que ce paramètre de rythme est généralement assez peu pris en compte dans les modèles de la prosodie, davantage centrés sur les paramètres d'accentuation et d'intonation dans leurs liens avec les niveaux structuraux de constituance. Nous verrons également que le modèle AM mêle habilement la représentation des événements métriques locaux d'accentuation et la représentation des mouvements intonatifs plus larges, en une seule et même représentation tonale (§ 3.4, p. 29).

2.2. Paramètres physiques

Les paramètres phonologiques abstraits (accentuation, intonation, rythme) se matérialisent dans la substance par le jeu des variations de f_0 , de durée et d'intensité. Ces trois paramètres physiques vont instancier par exemple les syllabes accentuées ou métriquement fortes : généralement, une syllabe accentuée va se manifester par une variation positive de f_0 (augmentation locale de hauteur), par une augmentation de sa durée (allongement) et le plus souvent par un renforcement articulatoire résultant en une intensité plus forte de la syllabe et de ses composantes spectrales. L'intonation va se manifester par une modulation de la f_0 se développant dans le temps. Le rythme, enfin, serait en quelque sorte le principe organisateur, sommant sur un plan cognitif les effets de ses modulations des paramètres physiques. Il est de

fait moins directement appréhendable d'un point de vue physique et métrique (ici, au sens 'mesure').

Les règles de combinaison des paramètres physiques pour instancier l'accentuation, l'intonation et le rythme vont varier d'une langue à l'autre, constituant ainsi la *grammaire prosodique* de chaque langue.

2.3. Mise en relation des niveaux d'analyse

Dans la description des grammaires prosodiques des langues, toute la difficulté réside dans la mise en relation entre les paramètres physiques observables et quantifiables et les paramètres phonologiques abstraits. Il pourrait être tentant d'opérer une mise en relation directe entre niveau acoustique et unités prosodiques, mais cela serait une grave erreur : on pourrait par exemple interpréter des perturbations micro-prosodiques liées au contenu segmental comme des indices accentuels motivés phonologiquement (décrochages abrupt de f_0 sous l'influence de l'explosion d'une consonne occlusive sourde ; allongement intrinsèque des voyelles nasales ou raccourcissement intrinsèque des voyelles hautes etc... ; Di Cristo, 1985). On met donc un point d'honneur, lorsqu'on étudie la prosodie, à contrôler autant que possible les niveaux d'analyse. Cela passe le plus souvent par la prise en compte d'un niveau intermédiaire entre niveau acoustique et phonologique : le niveau phonétique, où seules les variations acoustiques motivées sont observées (Hirst *et al*, 2000). Ainsi, les propositions de modélisation de la courbe de f_0 (INTSINT, par exemple pour le français ; Hirst & Espesser, 1993) sont-elles le fruit de cette tentative de médiation interprétative. Notons, et on y reviendra, que les propositions de modélisations phonético-phonologiques de la durée sont beaucoup moins développées.

Nous reviendrons au fil de ce document sur la tentation (parfois, et malgré une approche ancrée dans la phonologie prosodique) de mélanger les niveaux d'analyse et de faire un lien un peu trop direct entre substance phonétique et représentations phonologiques. Les transcriptions de la parole basées sur la métrique sont parfois détachées de la substance phonétique (approches purement phonologiques de la description prosodique). Même si elles ne sont pas basées sur l'analyse phonétique des informations prosodiques, certaines approches intonatives peuvent à l'inverse proposer une surspécification des catégories phonologiques induites par le signal de parole (une critique qui a pu être faite au modèle de Mertens, 1987, dans sa proposition de modélisation de l'intonation pour le traitement automatique de la parole) ou au contraire ignorer certains événements (tonals ou non) ne correspondant pas à leurs catégories phonologiques (voir § 11.3, p. 84).

L'approche que j'ai été amenée à choisir au fil des années, passant par le prisme de la perception, permet de se soustraire dans une certaine mesure à cette difficulté d'analyse : la mise en relation des paramètres physiques de la prosodie avec les catégories phonologiques se fait en effet en grande partie sur la base de ce qui est effectivement perçu par les auditeurs (voir § 10, p. 76).

2.4. Mesures des paramètres prosodiques

Ces considérations nous amènent à la manière de mesurer concrètement les paramètres phonétiques de la prosodie. Evacuons d'emblée le paramètre d'intensité, très difficile à mesurer et donc très peu étudié. La mesure de l'intensité nécessite en effet un contrôle strict de la distance de la source (le locuteur) au micro car la moindre variation de distance aura un impact sur la capture des variations d'intensité. Or, quand on parle, on bouge naturellement, particulièrement dans des situations de communication naturelles. Comment alors évaluer les variations d'intensité relatives à la force articulatoire mise pour la réalisation d'une syllabe accentuée, quand l'intensité globale de la parole n'est pas contrôlée ? Quelques corpus de parole 'naturelle' ont été conçus pour contrôler la qualité d'enregistrement (chambre sourde, micro-casque permettant de contrôler la distance à la source...) comme le corpus CID par exemple (Bertrand *et al*, 2008), mais la plupart préfère conserver le naturel de l'interaction en situation de communication 'écologique', *in situ* (par exemple Corpus PFC : Durand *et al*, 2009 ; ou CLAPI : Traverso, 2012). Ces corpus sont malheureusement très difficilement exploitables pour des études de phonétique et de prosodie.

Il est en revanche plus facile de détecter, visualiser et mesurer les paramètres de f_0 et de durée. Il s'agit cependant d'en interpréter les variations pertinentes en prosodie. Pour cela, des méthodes de normalisation existent, permettant de contrôler la variabilité inter-individuelle des locuteurs ou encore la variabilité intrinsèque des segments (pour une revue sur la variabilité, voir Rossi, 1985b). J'ai moi-même utilisé deux méthodes de normalisation de durée et de f_0 dans mes travaux de thèse sur la caractérisation du système accentuel du français (Astésano, 2001) : 1) la normalisation des durées segmentales, utilisée pour pouvoir inférer dans le cadre de la syllabe les seuls phénomènes d'allongements segmentaux motivés par l'accentuation (Campbell, 1992) ; 2) la normalisation des variations de f_0 permettant la comparaison inter-locuteurs des phénomènes de proéminences (l'échelle ERB de modélisation perceptive de Hermes & Van Gestel, 1991).

La f_0 est cependant le paramètre le plus ‘graphique’, le plus ‘facilement mesurable’, et il occupe donc (avec le paramètre phonologique qu’il permet d’actualiser, l’intonation), une très grande place dans les modèles prosodiques.

2.5. Pourquoi l’intonation occupe-t-elle une place centrale dans les modèles prosodiques ?

L’intonation est le paramètre le plus étudié, le plus modélisé à travers les différentes approches, quelle que soit la langue. Envisagée dans le cadre AM (§ 3.4, p. 29), l’intonation saisit l’essentiel des événements prosodiques (force métrique et contours) ce qui conduit à l’équation « intonation = prosodie ». Le terme ‘intonation’ est donc souvent utilisé en lieu et place du terme ‘prosodie’ (Rossi, 1977 ; Di Cristo, 1978 ; Ladd, 2008). Je préfère conserver le terme ‘prosodie’ car il englobe la notion de rythme, qui, adjointe aux faits intonatifs, permet de saisir le principe de récurrence accentuelle, principe impérieux dans mon approche de la prosodie du français (§ 9, p. 71).

Outre le caractère facilement appréhendable du paramètre de f_0 sur le plan visuel et acoustique, on dit que la f_0 est également le paramètre le plus important (avant la durée et l’intensité) dans la perception des proéminences (Fry, 1958). Cette proposition, tout aussi valide soit-elle, est fondée sur des tests psychoacoustiques de la perception de faits linguistiques où les paramètres acoustiques sont manipulés artificiellement et dans une relation simplifiée avec les autres niveaux linguistiques. Elle doit être tempérée par la prise en compte de phénomènes cognitifs plus complexes présidant à la perception des proéminences (voir § 9.2, p. 74 et 10, p. 76, pour une discussion).

Un argument plus fondamental concerne l’implémentation phonétique du sens véhiculé par l’intonation (Gussenhoven, 2002). Sur le plan de l’expression, les variations de f_0 seraient une grammaticalisation des 3 codes biologiques de production de la parole : 1) le code de Fréquence ou Registre (lié à la taille du larynx et des cordes vocales) permet de véhiculer les informations extralinguistiques d’âge et de sexe du locuteur, les informations de posture des locuteurs par rapport à l’interaction et certaines formes illocutoires de l’intonation (assertion, question...) ; 2) le code d’Effort (lié à l’énergie de production de la parole) est corrélé aux variations de f_0 et est notamment grammaticalisé dans l’expression du focus ; 3) le code de Production (lié à l’énergie globale de la parole entre deux prises de souffle) est grammaticalisé par la hauteur de f_0 décroissante à l’intérieur d’un groupe de souffle, appelée *déclinaison* et signalant les frontières de groupes intonatifs. Ces trois codes permettent donc de véhiculer l’essentiel du sens informationnel (les deux premiers permettant également de

véhiculer des aspects paralinguistiques liés aux affects), ce qui explique la prédominance du paramètre phonologique de l'intonation dans les modèles descriptifs. Dans les approches intonosyntaxiques traditionnelles du français, à la différence des modèles métriques et de phonologie intonative (AM), l'intonation supplée totalement à l'accentuation, reléguée aux phénomènes d'allongement final de frontières ou à la marque d'emphase.

3. PHONOLOGIE INTONATIVE ET THEORIE METRIQUE – ELEMENTS DEFINITOIRES

Une fois posés les différents niveaux d'analyse des faits prosodiques se pose la question de la représentation de la prosodie. Comment articuler les primitives prosodiques rendant compte de la linéarisation de la structure linguistique ? Mon approche de la prosodie du français se situe dans le cadre de l'approche métrique 'fonctionnaliste' développée par Di Cristo, mais j'ai été amenée ces dernières années à adopter les outils descriptifs issus de l'approche Métrique Autosegmentale (AM). Afin de contextualiser l'approche métrique que j'adopte, je décris rapidement quelques notions de base permettant de comprendre l'approche AM et la théorie métrique sur laquelle elle est fondée.

3.1. Principes de la représentation métrique de la prosodie

L'idée centrale de la théorie métrique (Lieberman, 1975 ; Lieberman & Prince, 1977) réside dans la représentation et la formalisation des relations de hiérarchie accentuelle. La représentation en grille métrique permet, outre la représentation des proéminences relatives (force métrique faible (w) et forte (s) des syllabes), de représenter l'alternance rythmique entre les syllabes fortes et les syllabes faibles, ainsi que les cas de *clashes* accentuels et de désaccentuation (Kager, 1994). La grille métrique suppose donc une représentation hiérarchique des proéminences. L'inclusion du niveau du *ped accentuel* par Halle & Vergnaud (1978) et Selkirk (1980) introduit la notion de *tête* métrique. Chaque pied métrique suppose la présence d'une – et une seule – syllabe forte (la *tête*) coïncidant avec une syllabe accentuée, et il comprend les syllabes faibles associées à la tête. Le niveau du pied induit un principe de *dominance* dans la construction des grilles, qui détermine le côté du pied où la *tête* va être placée. La dominance droite ou gauche du pied métrique va déterminer la *direction* de construction du pied : typiquement, le pied en anglais est construit avec une tête à gauche, et contient les syllabes inaccentuées suivant la tête, jusqu'au pied suivant ; l'inverse est vrai pour la construction du pied en français.

3.2. Niveaux de constituance métriques et indépendance prosodie/syntaxe

Les relations hiérarchiques entre les proéminences se formalisent par la superposition des forces métriques à travers les niveaux de constituance. Nous ne discuterons pas en détail de la pertinence de chaque niveau hiérarchique, abondamment discutés dans la littérature (pour une revue, voir entre autres Shattuck-Hufnagel & Turk, 1996⁵). Le premier niveau de construction de la grille métrique est la Syllabe (ou la more pour le japonais, par exemple). Au niveau immédiatement supérieur se situe le Pied métrique, constituant comportant la proéminence (*stress*, voir ci-dessous pour des précisions terminologiques) et les syllabes inaccentuées associées. L'accent associé au niveau du pied est de nature secondaire. C'est le premier niveau métrique à proprement parler, déterminant notamment la dominance et la direction de construction du pied. Le Mot Prosodique *pw* (ou Mot Phonologique) est en revanche le domaine de l'accent primaire du mot. Au niveau supérieur, le Syntagme Phonologique est le domaine de l'accent de syntagme (et des tons de frontières de syntagme, dans certains modèles). Selon les modèles, on peut le décomposer en Syntagme Majeur et Syntagme Mineur, reflétant ainsi deux niveaux différents de frontières; ils sont également appelés Syntagme Intermédiaire (*ip*) et Syntagme Accentuel (*ap*) dans la phonologie AM. Il est le premier niveau d'interfaçage avec la syntaxe. Enfin, le plus haut niveau de constituance métrique est le Syntagme Intonatif, défini prosodiquement par les tons de frontières intonatives. Il constitue également, outre son interface avec le discours⁶⁷, une interface phonologie/syntaxe.

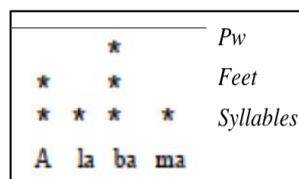


Figure 1 – Représentation en grille métrique du mot anglais 'Alabama' comportant deux pieds métriques. Au niveau du pied sont définis les accents secondaires, moins forts métriquement (ici sur 'A'); le mot prosodique (pw) est le domaine de l'accent primaire, plus fort métriquement ('ba').

Selkirk (1984a) pose une condition de *Strict Layer Hypothesis* contraignant les relations entre les niveaux de constituance. Elle questionne cependant plus tard cette condition (Selkirk, 2011) puisque, si les deux principes de *Layeredness* et *Headedness* (organisation en niveaux et principe de dominance des pieds métriques) semblent universaux, le principe de *Exhaustivity* selon lequel une unité d'un niveau doit être contenu dans le niveau strictement

⁵ Voir également la Figure 6 p. 34 illustrant les différents niveaux de constituance en fonction des modèles.

⁶ Voir pour le français, les approches de Delattre, 1966a ; Rossi, 1999 ; Delais-Roussarie, 2016 où les contours nucléaires sont porteurs d'informations discursives contextuelles.

⁷ Dans le même ordre d'idées, Di Cristo & Hirst (1996), Di Cristo (2011) proposent la 'macro-unité intonative' pour rendre compte de la profondeur de structures notamment en parole spontanée.

supérieur, et le principe de *NonRecursivity*, selon lequel un niveau ne peut pas être contenir un niveau de même nature, sont eux remis en question. Ladd & Campbell (1991) montrent que les domaines prosodiques ne sont pas forcément en nombre limité⁸. Ils proposent par contre l'existence de 'super domaines', ou 'domaines prosodiques composés', du même type que les domaines qu'ils dominent. Ils se fondent sur des études phonétiques sur les corrélats acoustiques des frontières prosodiques qui indiquent une très grande profondeur des structures prosodiques, reflétée par la durée segmentale et les propriétés de f0 dans le voisinage des frontières. Ladd & Campbell (1991) essaient de montrer qu'un approfondissement de la structure prosodique parvient à mieux rendre compte de la variabilité des durées syllabiques. Ainsi, le modèle à quatre niveaux de constituance (opposé aux modèles traditionnels à deux niveaux) rend compte des différences de durées des syllabes finales de syntagmes. Par contre, le modèle à deux niveaux suffit pour rendre compte des différences de durée des syllabes initiales et médianes.

De fait, comme rapporté dans Shattuck-Hufnagel & Turk (1996), Ladd montre que des énoncés contenant deux ou plus Syntagme Intonatif (IP) ont un contour propre qui n'est pas simplement la somme de deux contours d'IP. Cela démontre que les IP peuvent être emboîtés de manière *réursive*. En 1996, Selkirk propose le même type de violation, selon laquelle un pw peut dominer un autre pw (dans certains cas de clitiques affixaux). Plus intéressant, elle propose que le principe d'exhaustivité puisse être violé, puisqu'elle propose que les clitiques libres (de type Déterminants) sont rattachés non pas au niveau immédiatement supérieur du pw mais au niveau de la *Phonological Phrase PhP* dominant le pw (Figure 2).

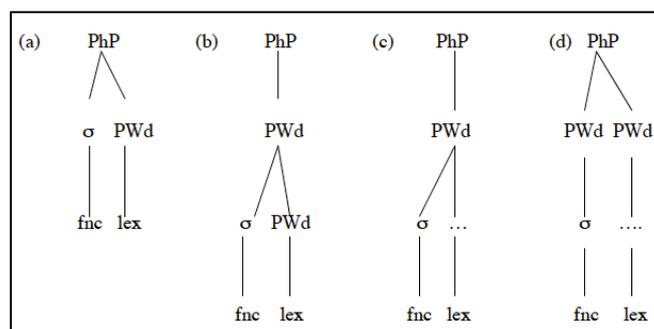


Figure 2 – Les quatre possibilités de structures prosodiques des clitiques faibles non finaux de syntagmes (clitiques inaccentués : *stressless syllables* ; Selkirk, 1996 : 19)

Nous reviendrons sur cette proposition pour ce qui concerne le français (voir § 16, p. 140). Ce dernier point soulève donc la question de la relation entre structure prosodique et structure syntaxique. Une critique de la représentation métrique émane des chercheurs

⁸ L'analyse du discours permet d'envisager un niveau encore supérieur au Syntagme Intonatif, le *paraton*, à l'interface entre prosodie et discours (Brown & Yule, 1986).

travaillant sur la parole spontanée. Puisqu'elle se fonde sur une représentation morphosyntaxique des relations de proéminences, elle s'interdit la représentation des structures prosodiques dépassant le cadre de la phrase telle qu'elle est définie en syntaxe (Couper-Kuhlen, 1993). De même, comment peut-elle rendre compte de l'organisation prosodique du message dépendant du choix du locuteur, lié à des considérations pragmatiques ? (Bolinger, 1972 ; Rossi, 1987). C'est l'interaction avec des considérations sémantiques et pragmatiques qui questionne le lien prosodie/syntaxe :

« L'intonation n'a pas de fonction syntaxique essentielle ; elle joue un rôle démarcatif qui délimite en tout état de cause des constituants bien formés ; la fonction de hiérarchisation qu'elle assume par ailleurs ressortit au niveau de l'énonciation. Ainsi, la prosodie a pour rôle de réorganiser l'énoncé, malgré la rigidité du cadre syntaxique, pour les besoins du message dans l'acte de communication. Mais tout n'est pas possible. La syntaxe impose des contraintes. C'est l'interaction de ces deux facteurs, énonciatif et syntaxique, qui est sous-jacente dans les règles proposées. » (Rossi, 1987 : 41).

On peut dépasser cette limite si le cadre métrique est appliqué non pas au *texte*, mais au *message*, supposant alors le passage par une phase de délimitation perceptive des proéminences et des groupements (Di Cristo, 2011 ; Astésano, 2001).

3.3. Sur la notion d'« Accent »

Il n'est pas du ressort de ce travail de discuter des nombreuses définitions, complexes, de l'accent. Nous renvoyons le lecteur aux travaux de Ladd (2008), Beckman (1986) et Shattuck-Huffnagel & Turk (1996) pour une revue. Mais s'entendre sur l'acception d'*Accent* permettra de mieux comprendre la position défendue dans ce travail.

Les termes 'proéminence' et 'accent' font référence soit à une entité abstraite relevant de la phonologie (proéminence dans l'organisation métrique, *i.e.* l'accent comme catégorie phonologique, tel que l'accent lexical), soit au domaine phonétique concret. Dans ce cas, la proéminence est équivalente à une saillance perceptive : une syllabe ou un mot proéminents ressortent par rapport au contexte. Au niveau phonétique, l'accent en français sera généralement qualifié d'accent de hauteur, d'accent de durée, etc. (Di Cristo, 2004). Il est donc nécessaire de toujours préciser le niveau d'analyse auquel on se situe. La difficulté terminologique est moindre en anglais qui distingue '*stress*' (niveau abstrait) de '*accent*' (niveau concret).

Pour l'anglais, '*stress*' désignera ainsi l'accent lexical (sous-jacent, propriété lexicale abstraite des syllabes individuelles), et '*pitch accent*'⁹, sa réalisation concrète (Bolinger, 1985) puisque les variations de *f0* sont fondamentales dans le marquage de l'accent en anglais. Cependant, Ladd (2008 ; 2017, reprenant les propositions de Beckman, 1986) insiste sur le fait que les manifestations phonétiques de cette proéminence phonologique peuvent être variées. D'une part, le '*pitch accent*' (ou 'accent de hauteur' en français) n'est pas forcément le seul moyen de révéler, en surface, la proéminence métrique (*stress*) : on peut rencontrer des syllabes métriquement fortes marquées par un renforcement articulatoire seul en anglais par exemple ; d'autre part, une variation mélodique (*f0*) permettant la perception d'une proéminence n'est pas toujours associée avec une syllabe lexicalement accentuée : Beckman fait alors référence à *non stress accent*, i.e. des syllabes proéminentes en surface non reliées à la structure métrique sous-jacente.

C'est à ce type de nuances complexes (quelque peu 'tord-neurones') que Di Cristo fait référence lorsqu'il dit :

« *Sans la connaissance préalable de ces subtiles distinctions terminologiques, il peut s'avérer difficile de comprendre le titre d'un ouvrage de Beckman (1986) : 'Stress and non-stress accent', ou la signification d'une expression comme : 'syllables are deaccented but not destressed' » (Di Cristo, 2004 : 87)¹⁰.*

S'il n'est sans doute pas aisé pour un prosodiste anglophone (même averti) de saisir toutes ces subtilités, il l'est encore moins pour un francophone pour lequel cette distinction n'existe pas dans son lexique. Par sécurité, on parlera tantôt d'« accent effectif » vs d'« accent sous-jacent ». Je pourrai également utiliser le terme 'accent métrique' ou le terme anglais '*stress*' pour faire référence aux proéminences métriques (accent/stress). J'emploierai également le terme de 'proéminences' pour faire référence à la qualité générique de saillance (syllabes saillantes dans leur contexte), quel que soit le niveau d'analyse (acoustique, métrique ou perceptif ; Shattuck-Hufnagel & Turk, 1996).

⁹ On remarquera une certaine contradiction de cette terminologie utilisée également dans les descriptions AM pour le français, puisque l'accent en français est généralement défini par ses caractéristiques de durée (voir Post, 2000, pour une discussion).

¹⁰ Je rajouterai une anecdote qui a suscité une certaine fébrilité chez mon doctorant et moi-même pendant quelques minutes (voire heures) de 'brain melting' suite à la lecture d'un commentaire (fort pertinent au demeurant !!) d'un 'reviewer' anonyme : « *Are the accents always aligned with stressed syllables ? It would be interesting to compare the degree of prominence perceived of a given accent when aligned with stressed versus unstressed syllables. A similar comparison would be interesting for the unaccented syllables : are unaccented stressed syllables perceived as more prominent than the unstressed ones ?* ». Je reviendrai au cours du document sur les difficultés de transmettre ce que j'essaie de démontrer, pour nous francophones n'ayant pas forcément baigné dans la culture anglophone AM, et pour lesquels l'accent n'est pas essentiellement défini sur la base des variations locales de *f0*.

3.4. Les principes de l'approche Métrique Autosegmentale (AM)

Je renvoie d'emblée le lecteur à l'ouvrage de Ladd (1996-2008) qui livre la description la plus complète de l'AM sur les plans historique, théorique et méthodologique. Le titre même de l'ouvrage (*Intonational Phonology*) met l'accent sur la prééminence de l'intonation comme fondement de la description phonologique des langues. Dans cette approche, intonation et accentuation sont intimement liées, les 'segments tonals' permettant de définir les variations tonales locales qui relèvent de l'accentuation et les variations tonales plus larges qui relèvent du marquage des frontières par l'intonation. L'AM est une adaptation des principes de phonologie autosegmentale érigés par Goldsmith (1976) à la description prosodique de l'anglais (Pierrehumbert, 1980). L'intérêt de cette approche est qu'elle utilise 2 segments tonals simples L et H (pour ton 'bas' L et ton 'haut' H) pour caractériser les contours tonals, et qu'elle associe ces éléments phonologiques discrets avec les paramètres acoustiques continus. La grammaire tonale d'une langue va donc se caractériser par les règles de combinaison dans le temps de ces 2 segments tonals. Précisons qu'à la différence des approches en contours intonatifs comme Delattre (1966a), Rossi (1985a, 1999) ou Martin (1981) pour le français, AM pose la structure métrique comme base d'analyse des événements intonatifs et non pas la structure syntaxique. Bien sûr, la structure morphosyntaxique est un préalable à l'analyse de la structure métrique, mais la prise en compte de niveaux de constituance non isomorphiques à la structure morphosyntaxique (comme le pied) ou la violation du principe de non récursivité évoqués plus haut ancrent définitivement AM dans une approche 'basée prosodie'¹¹ avec une certaine indépendance par rapport à la syntaxe.

Je reprends ici les 4 principes fondamentaux de AM tels que présentés dans Ladd (2008), et repris dans Arvaniti (2017) :

- 1 - La structure tonale (de surface) est séquentielle : les événements locaux sont associés avec certains éléments dans la chaîne segmentale. Ces éléments tonals sont discrets et le contour mélodique est interpolé entre les éléments tonals discrets. Les tons sont des autosegments contrastifs qui s'associent avec des positions dans la structure métrique : les '*pitch accents*' (marqués par le diacritique * : L* ou H*) sont associés aux syllabes proéminentes, *i.e.* aux têtes des constituants ; les '*edge tones*' sont les tons de frontières (marqués par le diacritique % : L% ou H%) associés aux bords des mouvements intonatifs aux frontières majeures. Ladd définit le *pitch accent* ainsi (ma traduction) : Le *pitch accent* est défini comme une caractéristique locale d'un contour intonatif (généralement mais pas

¹¹ Terminologie empruntée de l'anglais '*prosody based*' ou '*prosody first*'.

obligatoirement un changement de hauteur et impliquant souvent un maximum ou un minimum) qui signale que la syllabe avec laquelle il est associé est proéminente dans la phrase. Il a donc une double fonction : signaler la proéminence et participer à la construction du contour.

- 2 - Il faut distinguer ‘*pitch accent*’ et ‘*stress*’ : les *pitch accents* servent d’indices perceptuels concrets à l’accent métrique (*stress*)¹². La proéminence perçue des syllabes accentuées est affaire de poids métrique et/ou de dynamique d’accent, qui peut être distinct du *pitch accent*.

- 3 - L’analyse des *pitch accents* se fait en termes de tons de niveau : 2 tons primitifs H et L.

- 4 - Les sources de variations tonales locales (accentuation) se combinent pour révéler les tendances tonales globales (intonation): les tons sont utilisés pour rendre compte de certains événements globaux (par exemple la baisse graduelle de la dynamique tonale dans la phrase). Ils peuvent ainsi refléter l’opération de changements localisés mais itératifs dans la dimension d’amplitude de f_0 (*Scaling*).

Les tons phonologiques sont donc réalisés en tant que cibles tonales définies selon deux dimensions : leur amplitude contextuelle (*Scaling*) et leur alignement (*Alignment*), *i.e.* leur position temporelle par rapport aux repères segmentaux (Arvaniti, 2017). Ce dernier point (l’alignement) est important dans la discussion sur le statut phonologique de IA en français.

L’intérêt de l’AM réside donc dans la compositionnalité des tons, selon laquelle les tons permettent de relier les niveaux lexical et post-lexical de l’intonation : en d’autres termes, le sens émerge de leur combinaison.

A la différence de certaines approches (le modèle de Mertens par exemple, ou les approches holistiques de l’intonation), tous les segments ne sont pas spécifiés prosodiquement et le contour est dérivé par interpolation entre les points spécifiés. L’« encombrement » tonal (*tonal crowding*) est un cas inverse de *sur*-spécification tonale, qui implique parfois un ajustement segmental (allongement de la voyelle) pour pouvoir supporter l’ensemble des tons (Arvaniti, 2017). Ces cas de ‘*tonal crowding*’ seront évoqués plus loin pour démêler l’implication de IA et FA dans la démarcation des constituants en français.

Signalons ici que l’approche AM est un outil très puissant de transcription de la prosodie et de construction des grammaires tonales d’une langue. Combinée avec un ancrage métrique fort (je reviendrai précisément sur ce point en § 11.3, p.84), elle permet de manipuler

¹² Ils sont donc à l’interface entre perception et acoustique.

indépendamment les phénomènes de frontière et de *pitch accent*. Cet aspect méthodologico-conceptuel est particulièrement crucial pour le français, langue dans laquelle les phénomènes accentuels sont quelque peu fusionnés avec la démarcation intonative des frontières.

Je ne vais pas ici présenter un panorama des différentes descriptions et modèles prosodiques pour le français (voir pour cela, Lacheret-Dujour & Beaugendre, 1999). Je vais en revanche poser les questions du statut de l'accentuation en français et de son domaine d'actualisation. Ces questions sont centrales d'un point de vue fondamental, évidemment, mais aussi eu égard aux conséquences que peut avoir un certain malaise définitoire sur les applications de notre recherche, plus particulièrement sur :

- les mécanismes d'apprentissage d'une L2 et ceux à l'œuvre dans les pathologies du langage,
- notre capacité à repérer automatiquement les unités prosodiques pertinentes dans les gros corpus à notre disposition (traitement automatique des langues naturelles),
- le rôle de la prosodie (et de l'accentuation notamment) dans la segmentation de la parole et l'accès au sens. Il me semble en effet que quelques clarifications sont nécessaires pour les recherches et modèles en psycholinguistiques.

Globalement, la fameuse conception du '*français, langue sans accent*' (Rossi, 1980) a été selon moi surinterprétée¹³, ce qui a entraîné une surinterprétation en chaîne concernant l'absence de représentation cognitive et de matérialité supposées de l'accent (*stress*) en français.

Les propositions émanant des différents modèles de la prosodie du français viendront donc illustrer mon propos, et me conduire à la présentation du cœur de mon objet de recherche et de mes propositions (voir *Fugue*).

Avant cela, il me faut introduire rapidement ce qui fait du français une langue si particulière dans le paysage prosodique.

4. PARTICULARITES ACCENTUELLES DU FRANÇAIS

La particularité du français réside essentiellement dans le caractère non distinctif de son accent lexical. On accorde certes au français la présence d'un accent final (FA) sur les mots isolés mais aucune preuve n'a été apportée sur le fait qu'il soit une *propriété* du mot lexical, comme dans la plupart des autres langues (Welby, 2003). Par ailleurs (ou plutôt par voie de conséquence), son domaine d'application ne serait pas le mot mais le groupe de mots (syntagme mineur, syntagme accentuel, mot prosodique, mot phonologique ou encore *ap*, selon les modèles). Autrement dit, un syntagme comportant deux mots lexicaux, un adjectif

¹³ Mario Rossi m'avait d'ailleurs alertée sur ma propre surinterprétation de son article lors de ma soutenance de thèse.

antéposé et un nom par exemple, verrait seulement le deuxième mot lexical marqué par FA (ex : ‘les meilleurs aMIS’, où seule la dernière syllabe est métriquement forte, FA sur ‘meilleurs’ ayant disparu dans le groupe). La désaccentuation de la syllabe finale du mot interne au groupe accentuel ou groupe intonatif est donc un corollaire de la ‘*langue sans accent*’. Comble de malchance, sa position en finale de (groupe de) mots n’aide pas à en saisir clairement les propriétés accentuelles puisque ses caractéristiques acoustico-phonétiques fusionnent avec celles du marquage de la frontière de groupe (allongement pré-frontière et mouvement de f_0 de frontière). Rossi (1980) parle de synchrétisme entre accentuation et intonation, et Fónagy (1980 : 32) souligne « *l’ambiguïté inhérente à l’accent phrastique du français moderne* » imputable notamment au ‘chiasme acoustique’ aux frontières de constituants (où l’effort articulatoire de FA se ‘dissout’ dans la démarcation de la frontière intonative). Fónagy fait également à cette occasion référence à la propension en français moderne de marquer les débuts de constituants par un accent initial (IA), contribuant à la dispersion de l’effort articulatoire de l’accentuation en français et renforçant la notion de ‘chiasme’. L’absence apparente de fondement phonétique (sur un plan perceptif et acoustique) de l’accentuation finale conduit même certains auteurs à qualifier l’interprétation ‘accentuée’ de la syllabe finale « *d’imagination phonologique* » (Vihanta, 1993).

Je vais maintenant revenir tour à tour sur ces aspects plus en détails.

5. CONTEXTE THEORIQUE

Il est plus aisé de comprendre les propositions sur le statut phonologique de l’accentuation du français si l’on prend en considération les deux approches théoriques historiques de la prosodie du français.

5.1. Approches intonosyntaxiques vs métriques et AM du français

L’absence de fonction distinctive lexicale en français et le synchrétisme de FA avec les frontières intonatives expliquent sans doute que la tradition prosodique en France se soit davantage centrée sur les liens que *l’intonation* entretenait avec les autres niveaux linguistiques (syntaxe, sémantique et pragmatique, dont les travaux de Delattre, 1966a, Rossi, 1981 et suivantes, et Martin, 1981 et suivantes sont les plus représentatifs)¹⁴. Les approches intonosyntaxiques ne font référence aux prééminences accentuelles que de manière anecdotique, pour la plupart (voir néanmoins Rossi, j’y reviendrai amplement).

¹⁴ Notons cependant que l’étude de l’interface de la prosodie avec les autres niveaux linguistiques peut néanmoins tout à fait s’opérer dans une approche métrique et AM (voir les travaux de Delais-Roussarie, 2000 et 2016 notamment).

Les fameuses ‘*dix intonations de base*’ de Delattre sont la première tentative pour le français de mise en relation de l’intonation et des fonctions linguistiques : les contours intonatifs de base sont à l’interface avec le sens (fonctions pragmatiques et formes illocutoires) et la syntaxe (contours intonatifs hiérarchiques de continuation mineure (Cm) et Majeure (CM) notamment, reflétant la profondeur de la structure syntaxique). Ces contours sont définis sur la base de leur direction (montants, descendants, montants/descendants, parenthétiques plats) et de leur amplitude sur une échelle de 1 (niveau le plus bas du registre de *f0*) à 4 (plus haut niveau du registre).

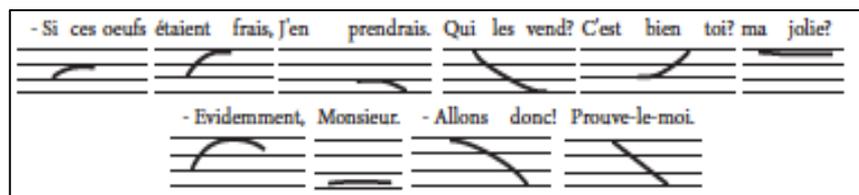


Figure 3 – Les dix intonations de base de Delattre (1966a). Les contours intonatifs reflètent les fonctions structurales et pragmatiques de l’intonation.

Dans la lignée de ces travaux, Martin (1981 et suivantes) propose des règles intonosyntaxiques sur des critères de substance (longueur des voyelles et amplitude des contours). On retiendra les deux règles intonosyntaxiques suivantes, qui nous permettront d’illustrer nos résultats sur FA (Figure 4): a) la sélection à droite : le contour se construit à partir du contour terminal (C0); b) l’inversion de pente : pour que la relation syntaxique entre deux groupes accentuels ou groupes intonatifs soit instanciée, il faut que leurs contours intonatifs soient de pentes opposées.

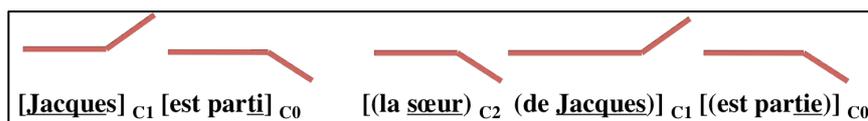


Figure 4 – Règles de sélection à droite et d’inversion de pentes intonatives indiquant la relation de dépendance entre deux groupes accentuels entre les contours C2 et C1 (Inspiré de Martin, 1981)

Cette règle d’inversion de pentes avait déjà été mentionnée par Delattre (1966a) au niveau des groupes accentuels ou Cm. Nous reviendrons sur la nécessité de définir précisément et expérimentalement les règles d’occurrences de ces contours inversés, car cela a une incidence non négligeable sur le statut phonologique de FA.

Rossi (1981 et suivantes) rajoute aux contours de Delattre de nouveaux intonèmes. Plus intéressant, il ancre les contours terminaux (porteurs de sens) sur l’accent lexical sous-jacent (FA) qui détermine la place des *morphèmes* intonatifs. Cet ancrage des contours sur l’accent primaire distingue l’approche de Rossi de celles de Delattre et Martin, car il reconnaît une représentation métrique abstraite de FA. L’avancée majeure de l’approche de Rossi concerne

la prise en compte de la structure informationnelle, ce qu'il appellera plus tard « l'intonation pragmatique » (Rossi, 2011 ; Figure 5).

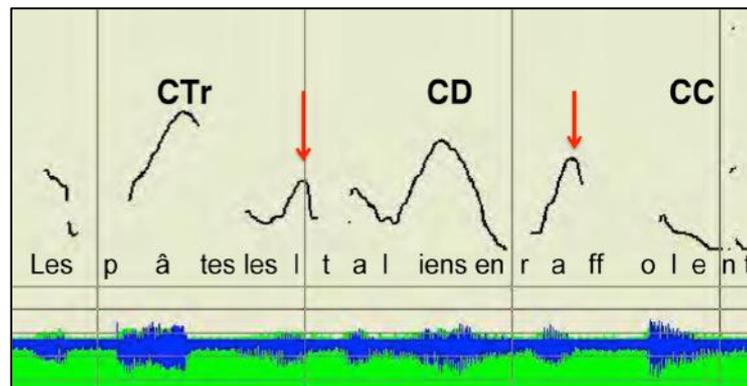


Figure 5 - Énoncé « Les pâtes, les Italiens en raffolent » réalisé avec le topique ‘Les pâtes’ accompagné de l’intonème contextuel référentiel CTr de focalisation. CD = continuatif dominant de sujet ; CC = intonème de clôture, conclusif, d’assertion. Les flèches rouges ont été rajoutées par moi-même et renvoient à ce que Rossi qualifie dans ce contexte d’Accent Focal. (Adapté de Rossi, 2011 : 119) :

Notons que l’accent initial (IA) est également pris en compte dans la description de Rossi, alors qu’il est (quasi) inexistant dans les approches de Delattre et Martin. J’ai rajouté les flèches rouges sur la figure ci-dessus, indiquant IA que Rossi qualifie d’accent focal. Elles participeront à illustrer mon propos sur la confusion entre deux types d’IA dans la littérature (voir § 5.5, p. 50).

L’autre originalité de Rossi consiste dans la distance qu’il prend avec la structure syntaxique, au fur et à mesure de l’évolution de son modèle vers la prise en compte de la pragmatique (voir sa citation de 1987 dans le § 3.2, p. 25).

A l’inverse des approches intonosyntaxiques, les approches métriques et AM de la prosodie du français sont le plus souvent indépendantes de la syntaxe (Di Cristo, 2000 et suivantes ; Delais-Roussarie, à partir de 2000 ; Jun & Fougeron, 2000 & 2002 ; Welby, 2003). Seul le modèle de Post (2000) adopte une vision *syntax first*. Dans les modèles *prosody first*, les constituants prosodiques sont définis uniquement sur la base de critères prosodiques (relatifs à l’organisation métrique, tonale et temporelle des énoncés) : les prééminences métriques correspondent aux têtes des constituants prosodiques et en sont donc définitoires. En aucun cas ne doit-on concevoir que la structure syntaxique sous-jacente va contraindre à elle seule la hiérarchie prosodique. C’est une manière d’éviter une certaine circularité dans l’interprétation des événements de surface. Comme le fait remarquer Di Cristo (2004), le piège est de ne pas créer des règles prosodiques pour correspondre à la syntaxe et inversement. La prosodie est donc considérée comme une structure totalement autonome de la syntaxe, même si on ne peut nier une forte contrainte de la structure syntaxique (Selkirk, 2011).

M'inscrivant dans ces approches descriptives, je vais maintenant présenter les niveaux de constituance proposés pour le français dans ces différents modèles métriques et AM. Toutes les autres considérations définitoires des éléments prosodiques du français se feront dans cette visée, mais des références seront faites aux autres approches, notamment aux travaux de Rossi qui sont au cœur de la problématique.

5.2. Les niveaux de constituance en français et leurs liens avec IA et FA

Les niveaux de constituance sont discutés, testés et éprouvés dans chaque langue. Le nombre de niveaux nécessaires à la description d'une langue dépendrait de la langue (Beckman, 1986). On considère de fait, en phonologie prosodique, un lien à établir entre forme et substance, entre unité grammaticale (telle que le mot) et unité prosodique (telle que la proéminence ou les tons de frontières), conditionnées l'une par l'autre.

Une tentative de représentation de ces niveaux hiérarchiques, et de leur variabilité en fonction des auteurs et des langues, est proposée par Di Cristo (2011 : 76). Il propose essentiellement une comparaison entre les modèles construits sur l'anglais (qui domine largement les descriptions prosodiques) et le français (Figure 6), et distingue les modèles *syntax first* (basés sur la syntaxe) et les modèles *prosody first*, marqués d'une étoile.

Il existe une multitude de terminologies pour décrire les niveaux de constituance prosodique, notamment en français, en fonction de l'approche : métrique, AM ou intonosyntaxique. Pour le français, deux niveaux sont largement reconnus : le niveau de *ap* (*accentual phrase*, terminologie AM que j'utiliserai essentiellement ; encore appelé groupe accentuel, syntagme accentuel, syntagme ou continuation mineure ...) et le niveau de IP (ou *Intonational Phrase* pour AM, encore appelé syntagme intonatif, groupe intonatif, continuation majeure ou unité intonative). Comme décrit plus haut (§ 3.2, p. 25), il existe pour l'anglais un niveau intermédiaire entre *ap* et IP : le syntagme intermédiaire (*ip* pour *intermediate phrase* correspondant peu ou prou au syntagme majeur des modèles de Selkirk ou à la continuation majeure). Pour le français, ce niveau intermédiaire à l'interface avec la syntaxe est encore (largement¹⁵) débattu car les modèles initiaux de AM (Jun & Fougeron, 2000 ; 2002) n'avaient pas trouvé expérimentalement la substance définitoire de ce niveau pour le français (essentiellement, *ip* serait le domaine de la règle de *downstep*¹⁶). La seule

¹⁵ Je me permets une petite réflexion : en réalité, les recherches sur la prosodie du français se font de plus en plus rares, et on est bien en peine de trouver des références de recherches expérimentales actuelles. Ce n'est plus le foisonnement de recherches des années 80 à 2000...

¹⁶ Le *downstep* correspond à un phénomène phonologique selon lequel les tons H contenus dans l'unité prosodique s'abaissent au fur et à mesure dans l'unité jusqu'au dernier ton H% associé à la frontière prosodique. Il est noté par le diacritique ! devant le ton H (!H). Cette forme phonologique donne donc une cohérence tonale à l'unité prosodique, dont on dit qu'elle est située à l'interface avec la structure syntaxique.

trace de ce niveau *ip* pour le français – défini sur la base de variations de durées essentiellement et dépendant de la taille des constituants – revient à Michéas & D’Imperio (2010a ; 2012). D’autres propositions rattachent plutôt ce domaine *ip* à des structures syntaxiques spécifiques de l’oral, du type détachement à gauche (Di Cristo & Hirst, 1996 : appelé ‘Segment d’unité intonative’ ; Jun & Fougeron, 2000 ; Delais *et al*, 2015).

Si *ip* est le domaine des règles de *downstep*, *ap* serait le domaine des règles phonologiques (assignation des accents, règles de liaisons, désaccentuation et règles rythmiques d’évitement du clash accentuel). IP est le domaine du contour intonatif et peut être composé d’une ou plusieurs *ap*.

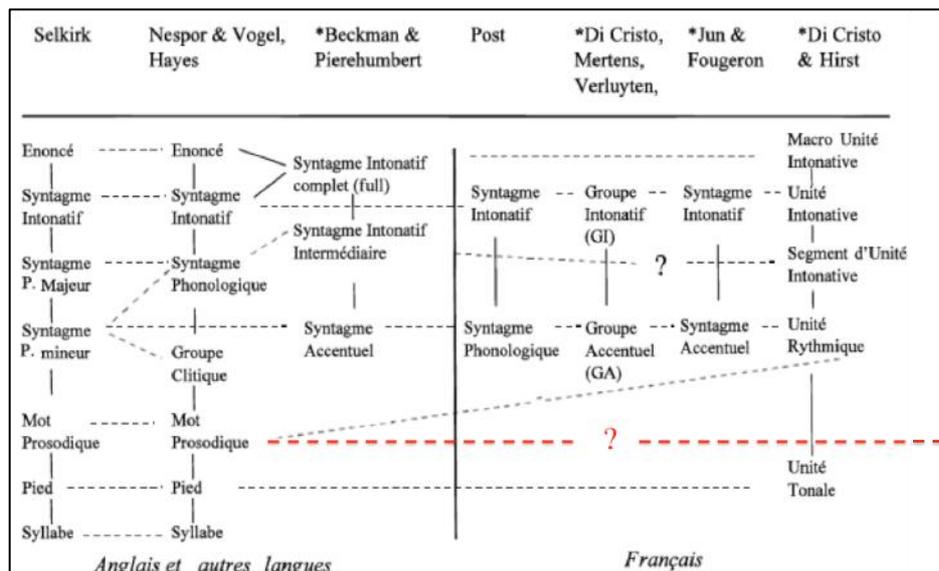


Figure 6 – Terminologie utilisée dans les différentes approches de la hiérarchie prosodique. L’astérisque devant les noms d’auteurs signale les approches basées uniquement sur des critères prosodiques (tiré de Di Cristo, 2011 : 76). Remarquons deux choses : Di Cristo note par un point d’interrogation (?) l’absence du syntagme intermédiaire dans les modèles AM du français (Post, 2000 ; Jun & Fougeron, 2000) et le fait correspondre au niveau du ‘segment d’unité intonative’ dans le modèle de Di Cristo & Hirst (1993 ; 1996). J’ai pris la liberté de rajouter un point d’interrogation en rouge (?) au niveau du Mot Prosodique (ou Phonologique) pour montrer l’absence de ce niveau interprétatif en français (sauf à équivaloir avec l’Unité Rythmique de Di Cristo & Hirst sans lui être totalement isomorphe). Le questionnement autour de l’existence de cette unité constitue le cœur de ma recherche.

La question maintenant est la suivante : quelles relations unissent ces différents niveaux de constituance pour le français avec les prééminences qu’on lui reconnaît, IA et FA ? Passons en revue ce lien, définitoire des niveaux de constituance, avant de se pencher plus précisément sur les conséquences de ces liens sur le statut de IA et FA. Je commencerai par leurs liens dans la phonologie AM avant de mentionner le modèle de Di Cristo & Hirst sur lequel mes premiers travaux portaient. Mon approche se situe en effet maintenant à la croisée des deux types d’approches.

Je ne m’attarderai précisément que sur le niveau de *ap* qui est le niveau de base dans la phonologie AM où se jouent toutes les règles phonologiques, notamment dans l’assignation

des proéminences. Le niveau de IP contient les projections accentuelles des *ap* qui le composent, mais il convient d’y ajouter le ton de frontière H% délimitant ce niveau d’actualisation des contours intonatifs (voir cependant § 5.5, p. 50, pour une discussion sur le maintien de H* à ce niveau de constituance). Il en va de même pour le niveau intermédiaire *ip* si tant est qu’on le juge pertinent dans la phonologie du français. Je renvoie donc le lecteur à Michelas & D’Imperio (2012) sur leur proposition de caractérisation tonale de ce niveau en H*H- (le diacritique – indiquant la frontière intonative intermédiaire).

Le niveau de *ap* est délimité par FA obligatoirement (accent primaire), et par IA de manière optionnelle (accent secondaire) dont l’occurrence dépend essentiellement de la taille du constituant. Dans la phonologie AM pour le français (Jun & Fougeron, 2000 ; 2002), FA est caractérisé par les tons LH*, l’étoile indiquant qu’il s’agit d’un *pitch accent* aligné sur une syllabe métriquement forte. IA est caractérisé par les tons LHi, ‘i’ indiquant que le ton H se situe à l’initiale du constituant. Nous reviendrons sur le statut alloué à IA dans cette vision ci-dessous mais notons que Post (2000) caractérise les deux accents IA et FA de métriquement forts (H* ; § 5.5, p. 50). Les réalisations tonales de *ap* sont récapitulés dans Welby (2006), qui modifie légèrement l’annotation de IA et FA, en ne se concentrant que sur leur position dans le constituant *ap* (IA y est qualifié de ‘*early rise*’ et FA de ‘*late rise*’). Ainsi, IA est caractérisé par L₁H₁ et FA par L₂H₂, ne préjugant pas du caractère métrique ou hiérarchique de ces proéminences. Welby propose plusieurs variantes de réalisations de *ap* en fonction des contraintes prosodiques qui s’y appliquent : L₁H₁-L₂H₂ (Figure 7) peut être réalisé L₁H₁-H₂ ou L₁H₁-H₂. En effet, certains tons ne vont pas être réalisés (*undershooting*) en fonction de la durée et de la taille du constituant. Notons que chez Welby, le ton L₁ de IA est toujours présent et aligné avec le début du mot lexical : il manifeste la présence sous-jacente de IA même quand une cible haute H₁ n’a pas pu être réalisée.

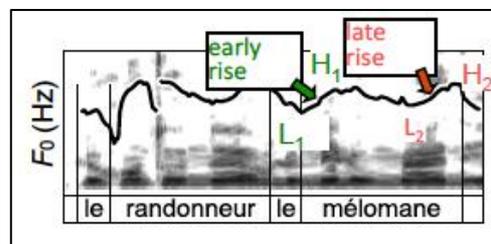


Figure 7 – Réalisation tonale canonique de l’*ap* sur le syntagme ‘*le mélomane*’ en L₁H₁-L₂H₂ (adapté de Welby et al, 2016).

Di Cristo & Hirst (1993 ; 1996) proposent des niveaux similaires à *ap* (Unité Rythmique, pas totalement isomorphe à *ap* selon Jun & Fougeron, 2000), *ip* (Segment d’Unité

Intonative¹⁷) et IP (Unité Intonative). Ils rajoutent cependant un niveau encore supérieur rendant compte d'une profondeur de frontière prosodique plus importante et d'une possible récursivité de IP (Macro Unité Intonative ; voir Ladd & Campbell, 1991, et Ladd, 2008 pour une proposition similaire pour l'anglais). Je commenterai ici plus longuement les niveaux inférieurs de leur modèle : l'Unité Tonale et l'Unité Rythmique.

Le niveau de l'Unité Tonale a été introduit plus tôt (Hirst & Di Cristo, 1984) avant d'être intégrée au modèle de 1993. Au départ, il s'agissait de proposer une unité intermédiaire entre la syllabe et le groupe intonatif, du niveau du *pied* métrique. Comme pour l'anglais, l'unité tonale est l'unité de réalisation intonative de base, mais le principe de dominance métrique à droite pour le français suppose un gabarit tonal inverse à celui de l'anglais, /BH/ (où B signifie ton bas et H, ton haut), alors que l'Unité Tonale en anglais possède un gabarit tonal /HB/. Dans ce modèle, l'Unité Tonale est définie essentiellement sur une base lexicale et délimite une unité marquée indifféremment par un accent initial ou un accent final. Elle caractérise sur le plan phonétique la substance d'une unité supposant une réalisation purement tonale de l'unité accentuable (en l'occurrence la syllabe). Elle constitue enfin une unité de traitement où se réalise la synchronisation des segments tonals et des segments phonématiques (Di Cristo & Hirst, 1993).

L'Unité Rythmique est délimitée par un accent primaire FA, et est immédiatement supérieure à l'Unité Tonale. Les Unités Rythmiques et Tonales sont congruentes lorsqu'elles sont situées à la frontière de FA. En revanche, si l'énoncé contient un IA, l'Unité Rythmique contient alors deux Unités Tonales (Figure 8).

.	+	(04)			
(.	.	+	.	+	.	.	.	+	(03)			
(.	.	+	.	+	(.	+	(.	+	(02)	
(+	+	+	(+	+	(+	+	(+	+	(01)
Ils ont dé gusté un bon vin blanc													

Figure 8 – Représentation en grille métrique parenthésée des niveaux de constituance dans les modèles de Di Cristo & Hirst (1993) pour l'énoncé 'ils ont dégusté un bon vin blanc'. Le niveau de la syllabe n'est représenté qu'au niveau supérieur dans cette représentation parenthésée (niveau 01); le niveau 01 correspond à l'Unité Tonale marquée par IA ou FA, le niveau 02 à l'Unité Rythmique marquée par FA, le niveau 03 est celui de l'Unité Intonative et le niveau 04 celui de l'énoncé (Adapté de Di Cristo & Hirst, 1997 : 91).

L'intégration de ces deux unités dans le modèle de Di Cristo & Hirst suscite plusieurs réactions et commentaires. Le premier point sur lequel j'aimerais insister est l'ancrage métrique puissant introduit par le niveau de l'Unité Tonale puisqu'il permet la prise en compte du *pied* métrique, à la base de toute description prosodique. Ce niveau du pied

¹⁷ Rappelons que pour ces auteurs, le segment d'UI ou *ip* ne caractérise que certaines formes syntaxiques de l'oral telles que le détachement à gauche du syntagme nominal.

métrique est d'ailleurs étrangement absent des modèles d'approche AM, ce qui constitue selon moi un premier flou définitoire, voire un premier acte de 'marginalisation' prosodique du français. L'absence du niveau de la syllabe est également notable, et s'explique sur la base des phénomènes de liaisons qui viendraient violer l'hypothèse de Hiérarchie Stricte de Selkirk (1980). Ceci aussi constitue selon moi une limite, voire une véritable contradiction dans tous les modèles prosodiques du français¹⁸. D'ailleurs, et j'y reviendrai dans le § 11.2, p. 79, les caractéristiques de durées infrasyllabiques des accents en début de mot lexical (IA) obtenus dans Astésano (2001), où IA est marqué par un allongement significatif de l'attaque, pourrait être une indication d'une démarcation syllabique de début de mot lexical plus nette qu'on ne peut l'imaginer, malgré les phénomènes de sandhi¹⁹.

L'ancrage métrique fort que représente l'introduction de l'UT est selon moi précieux. Ce n'est pas l'avis de Jun & Fougeron (2000) qui rejettent ce niveau de constituance. Jun & Fougeron (2000) soulève la contradiction contenue dans l'approche de Di Cristo & Hirst qui n'inclut pas le niveau de la syllabe puisqu'il viole le principe de Hiérarchie Stricte, mais qui introduit le niveau du pied à travers l'Unité Tonale qui vient, lui, violer ce même principe puisque les unités tonales ne respectent pas les frontières de mots. Plus important, Jun & Fougeron soulèvent un point crucial : deux UT marquées par IA ou FA n'ont ni le même comportement tonal, ni la même durée sous-jacente. L'UT marquée par FA donnera forcément lieu à la sensation plus forte d'une frontière puisqu'elle sera marquée également par la durée au niveau supérieur de constituance UR. Voilà donc les arguments de l'approche AM pour considérer *ap* comme le niveau le plus bas de constituance en français.

Je rejoins sur certains points les remarques de Jun & Fougeron sur l'UT. J'avais moi-même remis en cause l'UT sur la base de mes résultats de configurations tonales (Astésano, 2001). Avec une méthodologie différente, je montrais également que les pentes de *f0* et l'alignement des pics de IA et FA par rapport au contenu segmental sont tout à fait différents, suggérant plutôt une cohésion tonale interne au groupement (plutôt le Mot Prosodique pour moi, et l'*ap* pour Jun & Fougeron). Je reviendrai sur ce point dans le chapitre *Fugue*, mais il faut retenir d'ores et déjà que ma proposition rejoint celle de Jun & Fougeron (2000) sur le marquage *cohésif* par les configurations tonales (et durationnelles, pour moi) de l'unité grammaticale. La différence entre leur proposition et la mienne concerne le domaine d'application de cette

¹⁸ N'oublions pas en effet qu'on parle souvent du français comme ayant une rythmicité syllabique (réurrence régulière des syllabes) et que c'est l'un des arguments retenus par Cutler pour défendre l'idée selon laquelle l'accès au sens ne se ferait pas sur des bases métriques en français (voir § 8.1, p. 57).

¹⁹ Nos résultats sur la liaison (certes très préliminaires) présentés en § 22, p. 163, semblent confirmer cette tendance à la délimitation du mot lexical : l'attaque des syllabes initiales de mots lexicaux est plus longue que l'attaque de la consonne de liaison et les configurations tonales distinguent également la consonne de liaison de la consonne initiale de syllabe.

cohésion prosodique : la force centrifuge des proéminences s'exerce pour moi au niveau du mot prosodique *pw*, inférieur à *ap*. De fait, mon approche garde, à l'instar du modèle de Di Cristo et Hirst, la proposition de *pied métrique* en français, mais pas celle d'un marquage 'tête à droite' pur induit par le niveau de l'Unité Tonale. De plus, l'UR de Di Cristo & Hirst n'est pas selon moi équivalente au mot prosodique *pw* comme suggéré dans la Figure 6 puisqu'une UR peut contenir deux voire trois mots lexicaux selon ces auteurs, comme dans les exemples qu'ils proposent p. 93 (ex : *le GEnéral de l'ARmée de l'AIR*, comportant 3 UT et une seule UR).

Dans mon acception, *pw* est défini dans les termes de Nespor & Vogel (1986) et Selkirk (1996) et correspond au mot lexical, le clitique libre étant rattaché au niveau immédiatement supérieur de *ap* ou *pw* récursif (voir Figure 2). Ce dernier point pourrait correspondre par ailleurs à la proposition de Welby (2003 ; 2006) selon laquelle, si L_1 est doublement rattaché au mot lexical *et* à la frontière gauche de *ap*, il se 'répandrait' ('*tonal spreading*') à la frontière gauche de *ap*. Ainsi, le '*tonal spreading*' de L_1 sur le déterminant précédant le mot lexical pourrait être le signe d'une association de ce dernier au niveau supérieur de *ap*.

5.3. Fonctions de l'accentuation en français

Le point de départ de ma réflexion se trouve dans l'ambiguïté inhérente à la définition de l'accent en français. La 'faiblesse' terminologique de notre langue, qui ne permet pas clairement de distinguer sur le plan lexical les notions de *stress* (accent métrique) et *accent* (accent effectif) sans devoir constamment le qualifier, est sans doute à l'origine de ce flou conceptuel, qui engendre toute une série de conséquences notamment concernant le rôle de l'accentuation en français dans le traitement de la parole. Les conséquences du flou terminologique concernent aussi la définition du *domaine* de l'accentuation, intimement liée aux fonctions que l'on reconnaît à l'accentuation en français.

Avant de procéder à l'examen plus précis du statut de l'accentuation en français, il nous faut donc faire un détour par les *fonctions* que l'accentuation exerce dans la prosodie des langues. Or, les théories prosodiques sont largement dominées par les recherches sur les langues, comme l'anglais ou le hollandais²⁰, dont l'organisation prosodique est incontestablement basée sur l'accentuation lexicale, *i.e.* le marquage de l'unité morphologique par une syllabe métriquement forte ('*stressed*') à valeur *distinctive*. Au delà des incidences sur les descriptions phonologiques des langues du monde, cette prédominance

²⁰ Les années 60 à 90 ont été marquées par le développement de théories prosodiques portées par les écoles américaine (D. Bolinger), anglaise (M. Halliday) et hollandaise (t'Hart) qui ont un impact considérable sur les approches actuelles comme la Métrique Autosegmentale (Ladd, 2008).

descriptive a des conséquences en cascade sur l'ensemble de notre vision des processus linguistiques. Puisque la prosodie organise la matière sonore, tous les niveaux de l'organisation langagière sont impactés par la phonologie prosodique d'une langue donnée : lexicale, morphosyntaxe, sémantique, pragmatique et structure informationnelle. La recherche de processus communs aux langues, d'universaux linguistiques, pousse naturellement les scientifiques à trouver les règles sous-jacentes participant à l'encodage et au décodage du langage. Les fonctions culminatives et distinctives que l'on reconnaît principalement à l'accentuation lexicale façonnent par conséquent nos conceptions des principes sous-tendant les processus de compréhension de la parole. Parmi les langues occidentales, les mieux décrites, le français fait, à plus d'un titre, figure de 'langue exotique' puisque son accentuation n'est pas lexicalement distinctive, se porte au niveau post-lexical et est donc dépourvue, dans cette vision, de fonction culminative.

Qu'entend-on par là ? L'accentuation permet de marquer une syllabe et la rendre saillante dans son contexte (*fonction contrastive*). C'est une fonction qui relève essentiellement du rythme langagier. Dès lors qu'on lie cette fonction contrastive à une unité grammaticale (généralement le mot lexical), on parle de *fonction culminative* selon laquelle l'unité grammaticale va porter un seul accent par mot, permettant l'extraction sémantique dans la langue parlée. Il faut considérer cette fonction comme préliminaire ou présupposant la *fonction distinctive*, finalement assez marginale dans les langues à accent libre comme l'anglais, et qui permet la distinction sur le plan métrique de quelques paires minimales de mots (ex : *reCORD*, verbe signifiant 'enregistrer' ; *REcord*, nom signifiant 'le fait ou l'état d'être enregistré', l'enregistrement). Or, dans les faits, il semblerait que la *fonction distinctive* prenne le pas sur la *fonction culminative*, en phonologie mais surtout dans les modèles d'accès au sens. De fait, les langues ne présentant pas de fonction distinctive, comme le français, n'auraient pas de fonction culminative, pourtant cruciale dès lors qu'il s'agit d'envisager, notamment, le rôle de la prosodie dans la segmentation du flot de parole et l'accès au sens.

L'accentuation en français (langue à accent fixe comme le polonais, le hongrois, le tchèque ou le finnois par exemple) n'aurait donc qu'une fonction *démarcative* d'unités, permettant la signalisation des frontières de ces unités. Finalement, ce qui distingue ces langues est qu'on ne leur reconnaît pas le caractère '*stress*' au niveau du lexique. Il s'agit d'un accent intonatif ou plutôt d'une variation de hauteur de 'bord' ('*edge pitch*'), dédié seulement à signaler la frontière (Ladd, 2017, pour le français au moins). Cette vision rappelle les propositions de Vaissière (1990) et Beckman (1992) selon lesquelles le français est une *langue de frontières*,

où les indices phonétiques fonctionnent seulement comme démarcateurs d'unités sans participer au caractère accentuel (*stressed*) du mot. Remarquons que cela n'empêche pas d'envisager le rôle de l'accentuation dans la *segmentation* du flot de parole continue. Les travaux de Welby (2007, notamment) en sont l'exemple puisqu'elle montre que IA (ou 'early rise' d'un point de vue strictement phonétique) participe à la délimitation des unités lexicales dans le cadre de *ap*.

5.4. Quel domaine pour rendre compte de l'accentuation en français ?

L'accent en français est donc avant tout démarcatif d'unités. La question se pose alors de quelle(s) unité(s) il s'agit, et à quel niveau d'interprétation on se situe. Tout un tas de raisons phonologiques, phonétiques, morphophonologiques et morphosyntaxiques sont évoquées pour rendre compte du statut post-lexical de l'accentuation en français. Parmi elles, les phénomènes de Sandhi tels que la liaison et l'enchaînement constituent notamment un argument de poids contre la notion même de mot lexical en français (voir Durand & Echeynne, 2014 pour une discussion). Si l'on rajoute à cela le caractère non distinctif et non culminatif de l'accentuation décrit ci-dessus, on aboutit à une représentation lexicale 'amoindrie' favorisant cette interprétation post-lexicale de la langue.

Au niveau sous-jacent, pourtant, Rossi (1980) est le premier à reconnaître une propriété accentuelle au niveau du morphème et du lexème, *contra* les visions traditionnelles du français (Grammont, 1963 ; Garde, 1968, notamment). Mais comme « l'accentème » n'est pas générateur, en surface, d'un accent de mot, il perd cette propriété accentuelle au profit du seul « intonème » qui marque des unités supérieures au mot (*ap* ou syntagme accentuel, IP ou syntagme intonatif).

Mais qu'est-ce qu'un accent de mot ? Est-ce forcément ce qui transparaît acoustiquement en surface ? Si l'on considère, comme Rossi, des propriétés lexicales à l'accent sous-jacent, comment évacuer sa réalité métrique (sans parler pour l'instant de réalité 'cognitive', j'y viendrai dans le chapitre *Fugue*) ?

Certains indices des liens qui unissent accentuation et lexique se projettent pourtant en surface. Ce lien transparaît via l'instanciation de IA sur les mots plurimorphémiques longs (tels que *anticonstitutionnellement*). IA marquerait le début des morphèmes internes de mots (Pasdeloup, 1990 ; Delais, 1995 ; et dans une moindre mesure, Jun & Fougeron, 2002). Si tel est effectivement le cas, et que IA est sensible aux règles de constructions morpho-lexicales, ce serait bien un indice que cet accent entretient un lien avec les unités grammaticales. Autrement dit, si le trait 'accent/*stress*' se définit par sa relation avec une unité grammaticale,

alors cette instanciation de IA sur les morphèmes internes au mot serait une indication de son statut métrique.

Par ailleurs, Jankowski (1996) apporte un éclairage quantitatif du lien mot lexical / accentuation en français : 80% des mots lexicaux seraient en effet marqués par FA dans l'étude de Fant *et al* (1991) sur des corpus lus. On ne peut certes pas parler d'une oxytonie régulière (20% des mots lexicaux resteraient non marqués par FA) mais on peut raisonnablement postuler l'existence d'un accent sous-jacent pour tout mot lexical. La fréquente correspondance qu'il y a entre accent sous-jacent et mot lexical permet de penser qu'avant le syntagme lui-même, c'est le mot lexical, hautement porteur de signification, qui bénéficie des fonctions culminative et démarcative (Jankowski, 1996). Il convient de rajouter à ce constat la tendance forte, dans différents styles de parole et particulièrement en parole spontanée, de marquer le début du Mot Phonologique (ou Mot Prosodique, *pw*) par IA (Astésano, 2001). La probabilité pour un mot lexical d'être marqué en surface par IA ou FA augmente donc considérablement. Il reste cependant encore à quantifier précisément cette proportion de marquage lexical en surface, mais cela implique que l'on adopte une approche théorique et méthodologique adaptée à l'investigation de ce domaine d'analyse.

Enfin, tout est affaire de présupposés théoriques dans la mise en évidence des unités prosodiques et de leur domaine d'actualisation. Je prendrai quelques exemples piochés dans la littérature sur le français pour exemplifier mon propos. Ces exemples émanent d'approches théoriques souvent très éloignées, dont le dénominateur commun est de ne pas prendre en considération les faits accentuels en deçà du niveau de *ap* (ou groupe intonatif, selon les approches).

Un premier exemple concerne la non prise en compte très courante de IA comme marqueur potentiel de structure (voir 12.2, p. 88 pour une proposition). La figure ci-dessous illustre ce point, à l'initiale du mot 'librairie' que j'ai moi-même entouré en rouge.

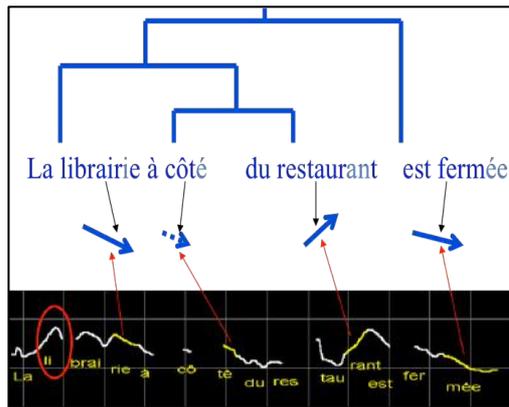


Figure 9 – Illustration de l'inversion de pentes pour instancier la hiérarchie intonosyntaxique de la phrase '*la librairie à côté du restaurant est fermée*' (tiré d'une conférence de Ph. Martin, d'après Martin, 2008). J'ai entouré en rouge l'instanciation de IA marqueur du début de l'énoncé, non prise en compte dans le modèle de Martin. Il est possible aussi que la première syllabe de 'côté' soit marquée par un IA.

La figure ci-dessous, issue de Jun & Fougeron (1995 et 2002) illustre plusieurs cas de réalisations accentuelles que je souhaiterais commenter.

Intéressons-nous tout d'abord aux instances de FA non prises en compte, selon moi, dans ces transcriptions : dans les deux figures, la courbe tonale que j'ai entourée en rouge est selon moi la réalisation de FA dans sa version soit !H, soit L* (voir § 12.3, p. 92 et § 13.2, p. 100 pour une discussion), subissant en effet la réorganisation tonale due à l'antéposition de l'adjectif (ne s'agit-il pas d'ailleurs d'un cas d'inversion de pentes comme décrit par Delattre et Martin ?). On ne peut pas parler de désaccentuation de FA ici selon moi car il y a des traces manifestes d'allongement final et/ou de contour tonal marqueurs de frontière (de mot ou d'*ap* ; voir encore § 13.2, p. 100 pour une discussion).

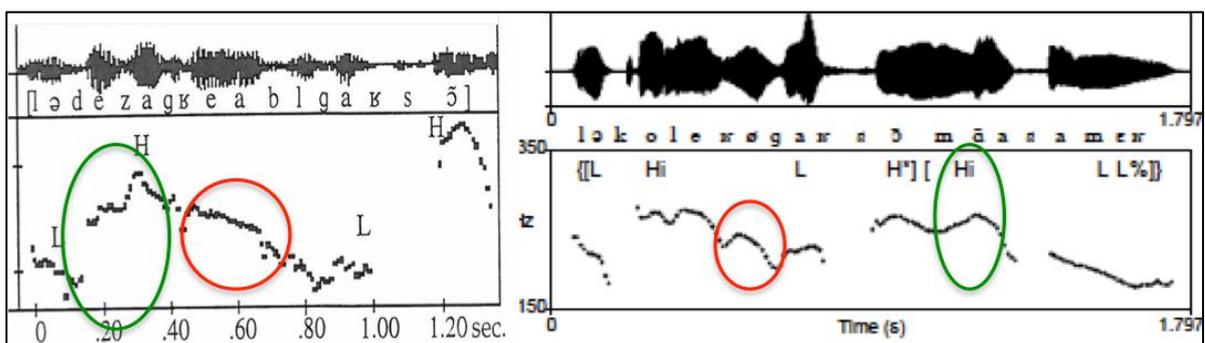


Figure 10 – Transcription tonale utilisant le formalisme AM de la phrase '*le désagréable garçon*' (à gauche) et de la phrase '*le coléreux garçon ment à sa mère*' (à droite) (d'après Jun & Fougeron, 1995 ; 2002). J'ai entouré en rouge les instances de FA selon moi non prises en compte. En vert, les occurrences de IA sur lesquelles j'apporte des commentaires dans le texte.

Viennent ensuite les transcriptions de IA (le premier LH dans la figure de gauche, Hi sans ton L dans la figure de droite). A gauche, le ton H est placé manifestement sur le point le plus haut de la courbe tonale, mais il pourrait s'agir d'un phénomène local de micro-prosodie (décrochage de f_0 dû à la fricative voisée). Ce qui me semble plus intéressant est le *saut* de f_0

entre le ton L sur le déterminant et la première syllabe du mot lexical ‘*désagréable*. Je reviendrai notamment sur cet aspect d’alignement de IA dans les propositions AM actuelles en § 13.3, p. 103. A droite, IA est placé en initiale absolue du niveau de *ap*, *i.e.* sur le mot (verbe) monosyllabique ‘*ment*’. Cela peut sembler plus incongru encore puisqu’on ne peut pas distinguer un IA d’un FA sur un mot monosyllabique, et que le contour tonal et l’allongement syllabique sur le verbe ‘*ment*’ seraient plutôt indicatifs d’un FA de niveau *ap* marqué H*.

En réalité, il s’agit tout simplement de l’impression²¹ ‘top-down’ d’un modèle sur la réalité phonologique ; autrement dit, les choix de transcription naissent d’un choix *a priori* du domaine d’application de l’accentuation que l’on choisit. Cela vaut évidemment pour tous les choix de transcription, y compris les miens, dont on aura compris qu’ils cherchent à mettre en évidence une réalité accentuelle au niveau du mot lexical ! Donc sans préjuger de la validité ou non de tel ou tel choix, je souhaitais simplement faire remarquer l’incidence d’un choix *x* ou *y* sur « l’oubli » de certains phénomènes phonologiques potentiellement pertinents dans la parole.

Je terminerai par deux autres exemples : l’un, issu des travaux de Michelas & D’Imperio (2010) visant à tester le rôle de l’accentuation dans la segmentation de la parole en français ; l’autre issu d’un exemple que j’ai construit pour (tenter d’) illustrer le phénomène de désaccentuation sur les mots internes aux *ap*.

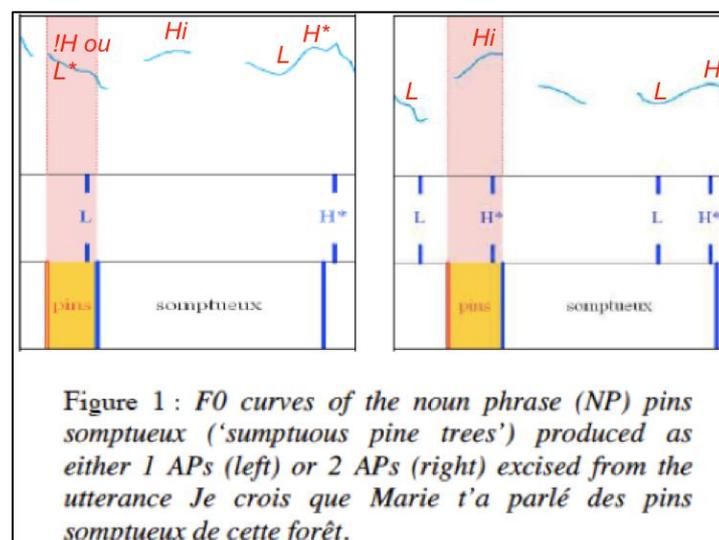


Figure 11 – Exemple de transcription AM de deux réalisations tonales d’un même syntagme ‘*pins somptueux*’ (Adapté de de Michelas & D’Imperio (2010). La transcription en rouge au dessus de la courbe tonale est ma propre proposition, commentée dans le texte ci-dessous.

Je renvoie le lecteur sur la proposition très intéressante de Michelas & D’Imperio (2010) tentant de montrer comment les indices prosodiques seuls (indépendamment de la syntaxe)

²¹ Au sens d’imprimer’.

permettent de segmenter la parole (ici : désambiguïsation entre '*pins somptueux*' et '*pinsons*'). A gauche de la figure, le syntagme '*pins somptueux*' serait réalisé en une seule *ap* marquée à gauche par L (le ton L de Hi, habilement noté L₁ chez Welby), alors qu'il serait réalisé en 2 *ap* dans la figure de droite, avec FA (LH*) sur '*pins*'. Ma proposition donne lieu à une interprétation tonale (et donc sémantique) tout à fait opposée. Selon ma propre analyse de ces deux syntagmes (donnée en rouge au dessus de la courbe), le premier syntagme (à gauche) pourrait être transcrit en 2 *ap*, mais marquée par !H ou L* sur '*pins*' et Hi + H* sur '*somptueux*'. Notons que Hi (IA) a été complètement ignoré dans la transcription des auteures alors qu'il instancie clairement une frontière entre le nom et l'adjectif (voir de nouveau § 12.2, p. 88). Le deuxième syntagme (à droite) pourrait être transcrit par LHi sur '*pins*'; dans ce dernier cas, on aurait en réalité la réalisation d'un arc accentuel reliant les deux mots lexicaux '*pins somptueux*'. Cette dernière proposition est motivée de deux manières : a) du point de vue de la substance, la pente de *f0* et l'alignement du pic sur '*pins*' correspondent davantage à la réalisation d'un IA ('*early rise*'); b) si on se situe effectivement dans la vision selon laquelle c'est le niveau de *ap* qui détermine l'occurrence de IA et FA, comme proposé par Jun & Fougeron et illustré dans la Figure 10, alors le niveau *lexical* ne vient pas contraindre l'annotation. Il y a donc dans ce dernier cas une contradiction par rapport au cadre théorique dont se réclament les auteures, motivée sans doute par les contraintes imposées par le cadre expérimental. Si mon analyse est bonne, on aurait ici en réalité un problème d'annotation des primitives tonales résultant en une inversion des conditions expérimentales à la base du test (dans cet exemple en tous cas ; je ne préjuge pas de ce problème sur l'ensemble des items).

Pour finir, je présente ci-dessous ma tentative d'illustration du phénomène de désaccentuation sur le mot lexical interne au groupe accentuel, corollaire de l'absence de l'*accent/stress* au niveau du mot lexical.

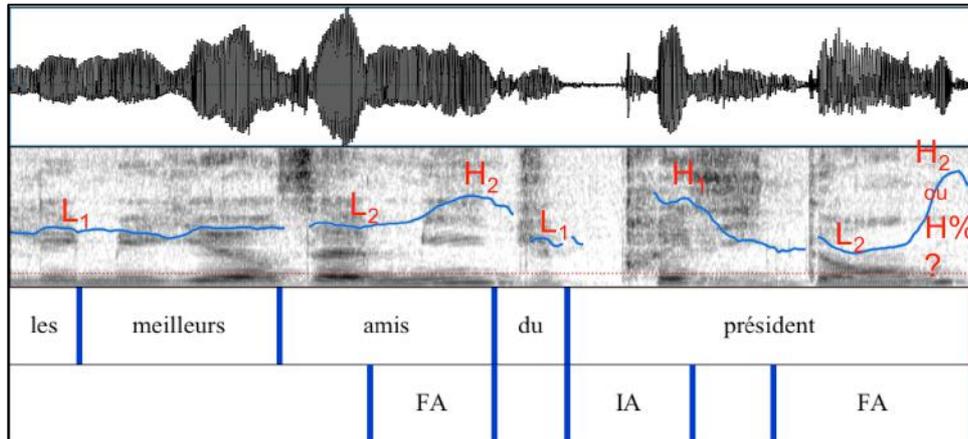


Figure 12 – Illustration du phénomène de désaccentuation sur le mot ‘meilleurs’ dans le syntagme ‘les meilleurs amis’. La transcription tonale AM est proposée en rouge. Notez que se posera la question de la ‘survivance’ de $L_2 H_2$ au niveau de constituance (IP) ; c’est pourquoi je propose l’alternative $H\%$ reflétant la frontière intonative seule, sans FA (exemple sonore n°1).

La figure ci-dessus illustre un phénomène extrêmement connu en français, justifiant de son caractère post-lexical : la désaccentuation d’un mot interne à un groupe de mots, ainsi que la liaison entre l’adjectif et le nom. On trouve beaucoup d’exemples dans la littérature prosodique utilisant la même construction Adjectif + Nom. Une autre construction également très utilisée est la construction inverse en Nom + Adjectif, donnant lieu le plus souvent à la formation d’un arc accentuel (comme dans la Figure 11, à droite, ma transcription de ‘pins somptueux’). La désaccentuation se manifeste par l’absence notable de modulation de f_0 sur ‘meilleurs’, ce qui justifie dans les approches intonatives (qu’elles soient intonosyntaxiques ou AM) que l’on considère ‘meilleurs’ comme dénué de tout accent. Ceci m’amène à deux remarques : 1) d’une part, on trouve une trace notable d’allongement sur la dernière syllabe de ‘meilleurs’ qui pourrait en elle-même justifier que l’on considère l’éventualité d’un marquage métrique de ce mot (voir § 9.2, p. 74) ; 2) D’autre part, je défie quiconque d’arriver à faire une désaccentuation naturelle sur la première *ap* de cette phrase (voir exemple sonore 1). Pour les besoins de l’illustration, j’ai dû me faire violence pour ne pas réaliser une toute autre configuration accentuelle sur l’*ap* ‘meilleurs amis’, que ce soit $L_1 H_1 - L_2 H_2$ (‘arc accentuel’), $L_1 H_{1p} - L_2 H_2$ ou encore $L_1 H_{1p} - !H/L^*$. Etant donné que la plupart des propositions phonologiques se font sur la base de phrases construites par l’expérimentateur, on peut raisonnablement se demander si les exemples de français que l’on ‘voit’ sur les représentations spectrographiques telles que celles-ci (à la base des modèles prosodiques et la plupart du temps construites par l’expérimentateur lui-même), ne sont pas pour une bonne part aussi improbables à prononcer...

En tout état de cause, ces quelques exemples étaient destinés à soulever un point : si tant est que le lecteur est d’accord avec mes propositions d’annotations, ne retrouve-t-on pas

systematiquement un marquage accentuel au niveau du mot lexical en surface, que ce soit par IA seul, FA seul, ou les deux ?

Par voie de conséquence, ne doit-on pas prendre en considération le niveau du mot lexical en français ? Encore faut-il s'entendre sur le statut de IA et de FA dans la phonologie prosodique : sont-ce seulement des manifestations tonales de surface, l'un fusionnant avec le contour intonatif (FA) et l'autre ne reflétant qu'un 'ictus mélodique' non motivé phonologiquement (IA)?

5.5. En finir avec la notion de français, langue sans accent ? Implications pour les statuts de IA et FA

Toute l'ambiguïté du français réside dans l'appellation même « d'accent ». On va parler pour FA d'accent interne, d'accent tonique, d'accent final...; on va parler d'accent externe pour l'accent initial emphatique, ou tout simplement d'accent initial ou d'accent emphatique. Or, parler « d'accent » renvoie à la propriété qu'ont certaines unités grammaticales de l'ordre du mot d'être associées à une marque prosodique. Pour Rossi (1980), l'accent final (FA) n'est pas générateur d'accentème en surface (*i.e.* de marque prosodique accentuelle en surface) mais générateur d'intonème : par conséquent, il perd sa matérialité accentuelle. Certes, on peut considérer dans cette vision que l'ancrage métrique sur la syllabe finale demeure au niveau abstrait, mais le syncrétisme entre accentuation et intonation ferait disparaître ce poids métrique en surface, dès le niveau du mot phonologique. Dans les approches intonosyntaxiques traditionnelles du français, parler d'accent final relève donc de l'oxymore.

Cette conception de FA a été partiellement reprise dans l'approche AM. Pour Jun & Fougeron (2000 ; 2002), FA est certes un *pitch accent* au niveau de *ap* (l'association entre la syllabe finale métriquement forte avec une proéminence tonale est présente) mais il perdrait son statut de *pitch accent* (H*) au profit du ton de frontière (H%) au niveau hiérarchique supérieur du Syntagme Intonatif (H% ; voir Figure 12). Aucune proposition concrète n'est faite pour l'instant sur le niveau intermédiaire de *ip* concernant le statut de FA, mais notons que la plupart des recherches expérimentales dans l'approche AM sur le français se concentrent sur le niveau de constituance *ap*, puisqu'il est central dans la caractérisation prosodique du français.

A la différence de Rossi, qui considère que le mot phonologique ou prosodique est coextensif du syntagme intonatif, donc caractérisé par l'absence de FA, l'approche AM

permet de projeter en surface une marque prosodique de FA à un niveau assez bas de la constituance, assez proche du mot lexical²².

Venons-en maintenant à l'accent initial (IA). Il n'est clairement pas pris en considération par Rossi (1980 et suivantes) comme relevant d'une association avec une unité grammaticale. Dans la lignée des qualificatifs qu'il a reçus depuis la nuit des temps (accent *rhétorique*, *expressif*, *émotionnel*, *intellectuel*²³ ...), ce type d'accent relève davantage du code para- ou extralinguistique dans la littérature. Rossi parle d'accent *externe*, ce qui l'écarte de tout rapport avec le mot, ou encore d'accent *énonciatif*. Là encore, on peut relever l'oxymore qui tend à associer la terminologie *accent* à un événement qui ne relève pas principalement du code linguistique. Il s'agit ici de la difficulté terminologique propre à notre langue évoquée en § 3.3, p. 27. Il faudrait peut-être parler ici de '*non stress accent*'. En 1985a, Rossi rajoute un autre type d'accent initial à son modèle : l'*ictus mélodique*. Il a essentiellement une fonction rythmique et a pour rôle d'organiser et d'équilibrer rythmiquement les portions d'énoncés de plus de 5 syllabes dénuées d'accent interne ou d'intonème. Cet ictus mélodique à fonction rythmique permet, pour la première fois, de mettre en évidence IA dans sa fonction non emphatique²⁴. Rossi précise qu'il marquerait les débuts de mots, mais il ne l'associe pas directement à cette unité grammaticale, préférant garder la notion d'accent *interne* pour FA à un niveau sous-jacent. L'introduction de l'ictus mélodique dans le modèle de Rossi s'inspire en quelque sorte de la proposition de Fónagy (1980) sur la bipolarisation accentuelle en français. Ce dernier remarque en effet que la stricte oxytonie, par laquelle l'accent marquerait régulièrement la dernière syllabe des groupes syntaxiques ou rythmiques, est plutôt rare ; les unités ont plutôt tendance à être délimitées à droite et à gauche par l'accentuation, donnant lieu assez souvent à la formation d'*arcs accentuels*. Cette tendance relève cependant d'une évolution diachronique de la langue française et serait plus fréquente dans la parole télévisuelle et radiophonique (voir de Mareüil *et al*, 2012 pour une description sur gros corpus radiophoniques). Fónagy ne distingue pas explicitement les deux types d'accents initiaux proposés par Rossi, à caractère emphatique ou rythmique. Cette distinction est d'ailleurs également absente dans l'approche AM : Hi ou '*early rise*' peut indifféremment signaler un accent initial emphatique ou rythmique. On peut attribuer ce choix méthodologique à une

²² *ap* est en effet juste au dessus de *pw* et souvent co-occurent à *pw*.

²³ Voir Séguinot (1976) puis Astésano (2001) pour une présentation détaillée de cet accent.

²⁴ Voir néanmoins l'interprétation de Rossi (2011) des deux accents initiaux sur la Figure 5 en accents focaux, alors que la dynamique tonale de ces accents montre clairement que le premier (sur '*Italiens*') est un IA marquant le début de la structure (Astésano *et al*, 2007) après le détachement à gauche du syntagme nominal.

certaine réalité perceptive, selon laquelle la perception de l'accentuation emphatique serait graduelle ou continue (Ladd & Morton, 1997).

L'approche AM pour le français considère le '*early rise*' comme un marqueur 'flou' de frontière dont le pic (Hi) peut être aligné sur la première, deuxième voire troisième syllabe de *ap* (Jun & Fougeron, 2000 ; Welby, 2003 et suivantes). Chez Welby (2006), '*early rise*' est cependant ancré au mot lexical par le ton L₁ mais le fait que son pic H₁ ne soit pas lui-même ancré sur le noyau vocalique de la syllabe initiale en ferait davantage un marqueur (flou : *loose*) de niveau *ap* que de mot lexical. De fait, son ancrage métrique n'est pas reconnu : il n'est donc pas un *pitch accent* au même titre que FA dans cette vision. Selon moi, et j'y reviendrai amplement, l'ancrage de IA par le ton L₁ est cependant le signe de cet ancrage métrique au niveau du mot lexical. Certes, ce n'est pas la définition canonique d'un *pitch accent*, mais on peut concevoir que les grammaires tonales des langues varient. On verra également que l'alignement du pic Hi sur le maximum absolu du début de *ap* n'est pas forcément le choix méthodologique le plus justifié pour rendre compte à la fois des réalités physique et perceptive de cet accent (§ 12.3, p. 92 et 13.3, p. 103).

Puisqu'il est unanimement considéré comme secondaire, IA est décrit dans l'approche AM comme un '*Leading tone*', *i.e.* une série de tons LHiL conduisant à l'ancrage métrique du H* de FA. Ce point de vue est remis en cause par Post (2000) dans ce même cadre AM, sur la base de toute une série d'arguments phonologiques, parmi lesquels sa possible 'starrification'²⁵ dans son expression emphatique (voir Post, 2000, pour une discussion). Post propose donc que IA et FA, malgré leur caractère secondaire et primaire différents, soient tous les deux considérés comme des *pitch accents*. Notons au passage qu'elle propose également, *contra* Jun & Fougeron, que FA garde sa caractéristique de *pitch accent* au niveau supérieur de IP (noté chez elle H*-H%).

Cette position rejoint celles de Di Cristo (2000 et suivantes) et Delais-Roussarie (2000) pour lesquels IA et FA sont tous deux métriquement ancrés au niveau du mot lexical. Dans ces approches aussi, IA et FA se projettent en surface à partir du syntagme accentuel, mais ils 'survivent' métriquement aux niveaux hiérarchiques supérieurs. Nous reviendrons beaucoup plus en détails sur ces approches métriques dans le chapitre *Fugue* car mon approche en est fortement inspirée.

Quelle que soit l'approche théorique, IA et FA sont décrits dans leurs relations comme participant de l'arc accentuel, forme prosodique courante en français. La démarcation des

²⁵ Il deviendrait H* chez Jun & Fougeron (2000) lorsqu'il a une fonction d'emphase, car aligné avec la voyelle pleine de la première syllabe, puisque les syllabes initiales contenant un schwa sont alors prononcées avec la voyelle pleine /oe/.

unités de type groupe accentuel ou *ap* par l'arc accentuel se combine donc avec un autre phénomène courant en français, coextensif à cette tendance rythmique : le phénomène de *désaccentuation*. La désaccentuation est une stratégie rythmique (supposément corollaire, selon Rossi, de la 'langue sans accent') permettant d'éviter les clashes accentuels, *i.e.* l'occurrence contiguë de deux syllabes accentuées. Selon cette stratégie rythmique, IA viendrait donc se placer sur la syllabe initiale du premier mot et FA sur la syllabe finale du second mot de l'unité, comme dans le syntagme 'MAjeure parTIE'. Fónagy attribue cette forme accentuelle à la tendance centrifuge qui pousse

« l'accent du premier membre d'un syntagme vers la première syllabe et celui du deuxième membre vers la dernière syllabe d'un syntagme », afin de « faire mieux ressortir son unité syntactique et sémantique » (Fónagy, 1980 : 143).

Fónagy fait ici référence à un cas de déplacement d'accent de la position finale à la position initiale. Ce faisant, sa définition de l'arc accentuel recouvre les deux terminologies de IA proposées par Garde (1968) : l'accent secondaire et l'écho d'accent²⁶. En réalité, on peut distinguer deux types d'arcs accentuels (Di Cristo, 1999 ; 2011) :

- l'arc accentuel purement *rythmique* destiné à éviter un clash accentuel, comme dans 'le CAfé NOIR' réalisé en une seule *ap* plutôt que dans 'le caFE au LAIT' réalisé en 2 *ap*.
- l'arc accentuel destiné à renforcer la cohésion *sémantique* entre des unités lexicales dans des séquences formulaïques du type 'les CONventions soCIALES' ou 'MAjeure parTIE'.

Le corollaire à cette tendance centrifuge de l'accentuation serait la désaccentuation de la syllabe finale sur le premier mot (Figure 13).

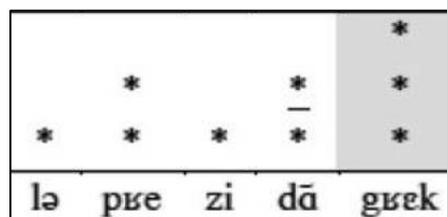


Figure 13 – Grille métrique du syntagme 'le président grec', où la syllabe initiale de 'président' est accentuée (IA) et la syllabe finale de 'président' est désaccentuée (désaccentuation illustrée par le soulignement de l'étoile au niveau du pied métrique). (Adapté de Delais-Roussarie & Di Cristo, à paraître : 12).

On parle généralement de *désaccentuation* et non pas d'*absence d'accentuation* car l'accent final survit en partie dans le groupe (Delattre, 1966b), notamment par la persistance de l'allongement syllabique à défaut de réalisation tonale. La trace métrique sur /dã/ de

²⁶ Pour Garde, l'écho d'accent est un cas particulier d'accentuation, non directement liée à l'accent initial secondaire. Il s'agit plutôt d'un *déplacement* de FA vers le début de l'unité, pour en faciliter la réalisation en surface.

'*président*' est donc toujours présente, à l'instar de ce qui a pu être proposé pour l'anglais par Selkirk (1980) et Nespor & Vogel (1986) où l'on observe la rémanence des traits métriques (durée) sur les syllabes désaccentuées dans des contextes de réorganisation rythmique. La désaccentuation en français ne serait donc pas le reflet de l'absence du trait métrique 'accent' au niveau du mot lexical mais plutôt celui d'une réorganisation accentuelle et d'un choix du locuteur d'alterner le marquage droit ou gauche des constituants. Notons au passage que cette désaccentuation peut intervenir sous le coup d'un principe centrifuge de démarcation des syntagmes, selon lequel les locuteurs auraient naturellement tendance à positionner aussi tôt et aussi tard que possible les proéminences dans le syntagme (Fraisie, 1967 et suivantes ; Hyman, 1975), principe sans doute lié à la construction de la pensée (Bolinger, 1981 ; 1985). Ce principe a également été décrit comme un phénomène de '*early accenting*' touchant le niveau du syntagme intonatif en anglais (Shattuck-Hufnagel *et al.*, 1994). Nos propres résultats sur le français montrent cette même tendance pour le français (Astésano *et al.*, 2007 ; voir § 12.2, p. 88) mais nous montrons néanmoins que IA marque préférentiellement un niveau plus bas de constituance : *ap* ou potentiellement *pw*.

Enfin, la désaccentuation peut se retrouver en position post-focale ou post-nucléaire, *i.e.* dans les énoncés où l'accent nucléaire (accent d'énoncé, le plus fort dans la hiérarchie des proéminences) est déplacé en position interne à l'énoncé, sur un mot contrasté sur le plan paradigmatique par exemple ou encore dans les cas de détachement à droite du syntagme nominal (voir Figure 18, p. 74). Ce déplacement en 'focus étroit' de l'accent nucléaire induit la réalisation en paranthèse basse du syntagme détaché post-focal, *i.e.* où le contour tonal se retrouve complètement aplati. Persistent cependant des traces métriques essentiellement portées par la durée sur les syllabes finales des mots lexicaux (Di Cristo & Jankowski, 1999). On peut se demander néanmoins, à l'instar de Ladd (2008 : 52) si notre capacité à retrouver les frontières de mots en position désaccentuée post-focale, sans aucune indication tonale donc, ne serait pas également liée à notre connaissance du contexte de communication et de la langue (processus top-down) plutôt qu'à des indices phonétiques (processus bottom-up). Que ce soit pour des raisons des persistances des marques de durées au niveau du mot lexical ou pour des raisons de plus haut niveau, l'auditeur conserve néanmoins une perception du mot lexical (voir § 10, p. 76).

J'aimerais terminer ce tour d'horizon sur le statut de l'accentuation en français par quelques considérations sur la fonction principalement rythmique que l'on reconnaît unanimement à IA. Son caractère secondaire, n'intervenant que pour marquer de longues séquences de parole non marquées par FA, affaiblit considérablement son rôle dans la

phonologie prosodique, nous l'avons vu jusque là. On dit également qu'il interviendrait dans les cas d'accentuation contiguë de FA, dans un rôle lié à sa fonction rythmique puisqu'il permettrait d'éviter les *clashes* accentuels. Il s'effacerait cependant au profit de FA dans les constituants courts. Mais les contraintes rythmiques ne sont sans doute pas la seule explication de la réorganisation entre IA et FA, et la prédominance de FA sur IA n'est pas si évidente que cela (voir Astésano *et al*, 2007, pour une discussion). Mertens (1992) a en effet montré (*contra* Dell, 1984) que, si on évite généralement l'occurrence de deux FA non séparés par une frontière majeure (IP ; Post, 2000), on trouve très souvent en parole spontanée l'occurrence contiguë de IA et FA. Il explique cette tendance par le fait que IA et FA sont différents par nature, complémentaires dans leurs traits prosodiques. On peut donc retrouver IA+FA ou FA+IA (frontières *ap*) sur des mots lexicaux ou phonologiques bisyllabiques (Figure 14).

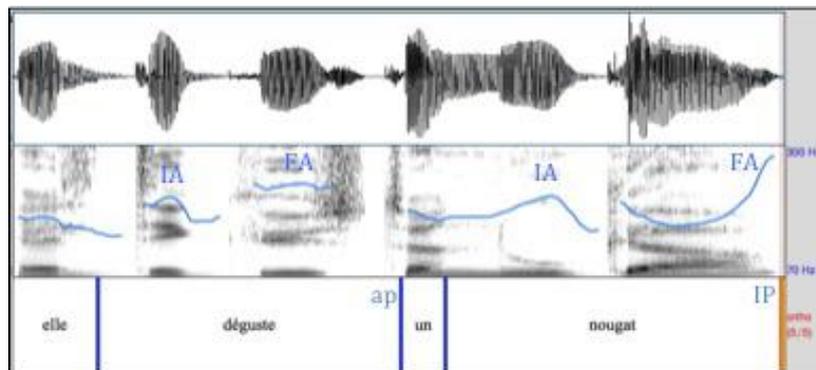


Figure 14 – Exemples de réalisation contiguë de IA et FA à deux niveaux de constituance (*ap* et *IP*). (Adapté d’Astésano, 2016).

Contrairement à ce qui a pu être dit (Portes *et al*, 2012), IA peut donc se retrouver en position contiguë avec FA non seulement en position pré-nucléaire au niveau de *ap*, mais également en position nucléaire dans les constituants *IP*. Cette proposition suggérait en effet un effacement de IA au profit de FA ($H^*-H\%$) sur des principes de hiérarchie des proéminences. En réalité, ce ne sont pas les niveaux de constituance qui présideraient à la démarcation conjointe des constituants par IA et FA. Non seulement cela, mais IA pourrait dans bien des cas et quels que soient les niveaux de constituance, marquer préférentiellement les débuts de constituants par rapport à FA. Nous y reviendrons abondamment dans la section § 12, p. 87 mais soulignons que ce cas avait été repéré par Jun & Fougeron (2002 ; Figure 15).

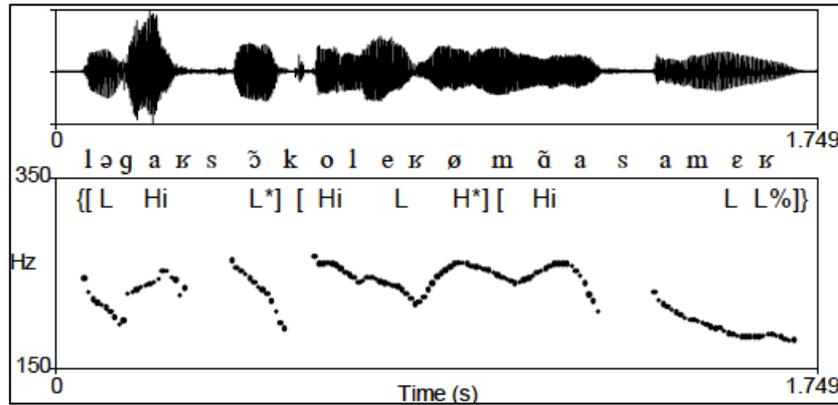


Figure 15 – Réalisation tonale en 3 *ap* de la phrase ‘le garçon coléreux ment à sa mère’, où l’introduction d’un ton L* sur ‘garÇON’ et d’un ton Hi sur ‘COLéreux’ instancient une frontière d’*ap* non prévue par la structure syntaxique entre le nom et l’adjectif (Jun & Fougeron, 2002 : 10).

Jun & Fougeron interprète cette réalisation tonale inattendue et, selon elles très rare (où le nom et l’adjectif sont séparés par une frontière *ap* par un ton L* et Hi) par l’impossibilité de produire trois tons hauts H à la suite sur ‘gar-’, ‘-çon’ et ‘co-’ (cas de *tonal crowding*, que l’on peut traduire par ‘surpopulation accentuelle’ ou ‘encombrement accentuel’). Mais elles ne proposent pas d’interprétation de ce choix tonal, favorisant IA sur FA, à la lumière du caractère secondaire et primaire de ces accents. C’est sur ces considérations que nous reviendrons dans l’interprétation de nos résultats en § 12, p. 87 et suivantes.

6. CONCLUSION

Ce chapitre était destiné à tenter d’éclaircir le statut de l’accentuation en français, à travers les différentes approches et modèles. Ce n’est pas chose aisée tant les particularités prosodiques de cette langue sont nombreuses. Le syncrétisme de l’accentuation et de l’intonation, dans un contexte théorique où la description tonale des faits prosodiques domine, est une première gageure méthodologique et théorique. La fonction non distinctive de l’accentuation en français en est une autre, tant cette fonction prend le pas sur les fonctions démarcatives et culminatives de l’accentuation. Ces deux particularités du français ont une incidence sur le domaine d’application de l’accentuation, qui est, dans toutes les approches, le niveau du groupe de mots, qu’il soit défini sur des bases accentuelles (*ap*) ou intonatives. Puisqu’en phonologie prosodique le niveau du mot lexical est coextensif de la propriété métrique de l’accent, il en résulte pour le français un certain malaise définitoire, engendrant selon moi toute une série de contradictions théoriques (la terminologie même « d’accent » pose problème, ce qui explique le titre provocateur de Rossi, 1980 : *le français, langue sans accent* ?). J’ai donc proposé de nous interroger sur les questions suivantes : les propriétés morphologiques de FA (et IA) et le lien qu’elles impliquent entre accentuation et unité

grammaticale ne permettent-elles pas de considérer le mot lexical comme domaine de l'accentuation ? Le lien entre les paramètres accentuels et le niveau du mot doit-il simplement être pris comme la conséquence d'une seule démarcation de surface (allongement final coextensif de la frontière intonative pour FA ; matérialité tonale ('ictus') seulement contrainte par la réorganisation accentuelle tonale en surface pour IA) ? N'est-ce pas antinomique de considérer le caractère primaire et secondaire de FA et IA quand on ne leur reconnaît pas de caractère métrique clair ? IA est-il vraiment secondaire, à fonction essentiellement rythmique ? J'ai proposé d'illustrer ces questionnements par quelques exemples tirés de la littérature et je pense avoir commencé à pointer que les tendances révélées sont le plus souvent davantage l'indication d'une posture théorique que d'une réalité définitoire de l'accentuation.

Intégrer le niveau du mot lexical à part entière (*pw* dans une approche métrique) dans la hiérarchie prosodique du français (non seulement comme domaine abstrait mais également comme domaine des règles accentuelles en surface) permet selon moi de lever une grande part de l'ambiguïté définitoire de la prosodie du français. Les approches métriques de Di Cristo (2000) et AM de Post (2000) et Delais-Roussarie (2000) (§ 11, p. 77) sont les seules portes d'entrée à cette proposition quelque peu osée que je me permets de soumettre dans le présent document. Et enfin, si je m'intéresse de prime abord à définir le domaine de IA et son rôle dans la prosodie du français, c'est qu'il me permet *in fine*, dans sa relation avec FA, d'éclaircir également le statut de ce dernier.

Avant de présenter mon approche et les résultats qui m'ont amenée à mes propositions théoriques, je présente rapidement les modèles de traitement du langage sur les versants neuro- et psycholinguistiques. Interroger la manière dont on traite le langage aide en effet à comprendre comment les indices prosodiques du français sont représentés. On remarquera, pour les modèles neurolinguistiques, la nécessité d'intégrer les faits prosodiques dans leurs rôles à l'interface des autres niveaux de représentation linguistique, et non pas seulement dans leur rôle paralinguistique (traitement des émotions) ou dans le traitement séparé de leurs fonctions linguistiques. On évoquera enfin les modèles pré-lexicaux de segmentation et d'accès au sens dont les principes se fondent sur une représentation métrique de l'accentuation, et la nécessité impérieuse qui en découle, d'ancrer la description prosodique du français dans une représentation de ces faits métriques.

Modèles neuro- et psycholinguistiques du traitement du langage

7. BASES NEURALES DU LANGAGE

Les neurosciences cognitives tentent de rendre compte des processus de traitement du langage par la mise en relation des architectures anatomique et fonctionnelle du cerveau.

Dans un premier temps, les modèles posaient que la fonction linguistique est *modulaire* et *indépendante* des autres fonctions cognitives, et que les différents niveaux de traitement du langage seraient organisés de manière *hiérarchique* et *sérielle* (pour une revue, voir Frazier, 1987). Dans ces modèles, la syntaxe serait traitée en premier comme structure intermédiaire vide de sens, le traitement de la sémantique et de la pragmatique n'intervenant que dans un deuxième temps. Alternativement, des modèles connexionnistes ou interactifs ont été proposés qui, tout en conservant la notion de 'modules' et de niveaux de représentation, rompent avec l'idée d'autonomie et de primauté de la syntaxe, ainsi qu'avec l'idée d'organisation hiérarchique et sérielle des niveaux de représentation. Ils proposent un traitement *interactif* et *parallèle* des niveaux linguistiques (Altmann & Steedman, 1988 ; McClelland et al, 1989).

A l'intersection de ces deux types de modèles, Friederici (2002) propose un modèle de traitement des phrases auditives selon lequel les différents niveaux (acoustico-phonologique, lexical, sémantique et syntaxique) seraient traités de manière sérielle ; seule la dernière phase d'intégration (syntaxique) serait interactive. Friederici fonde l'essentiel de ses propositions sur la méthode d'Electro-Encéphalo-Graphie (EEG) qui permet l'analyse fine du déroulement temporel des processus linguistiques. Elle est plus utilisée pour observer l'interaction possible des différents niveaux linguistiques car moins invasive et moins lourde que la méthode d'Imagerie par Résonance Magnétique fonctionnelle (IRMf), plus précise sur la localisation des activités cérébrales. La précision temporelle qu'offre l'EEG permet une observation plus fine des processus linguistiques, et de l'interaction potentielle des différents niveaux linguistiques. Avant d'aborder le modèle de Friederici, une rapide présentation de la méthode EEG est nécessaire.

L'EEG permet d'enregistrer, à l'aide d'électrodes posées sur le scalp, l'activité électrique neuronale dans le temps. On mesure précisément la réponse neuronale *évoquée* par la présentation d'un stimulus. L'activité évoquée par le stimulus est appelée Potentiel Évoqué (PE) et donne lieu à différentes composantes PE reflétant les différents types de processus (voir Figure 17). Une composante PE est définie par sa latence, *i.e.* le temps qu'elle va mettre pour se développer (en millisecondes, ms), par sa polarité (la décharge électrique des

populations de neurones est Positive ou Négative), par son amplitude (mesurée en micro-volts, μV) et enfin par sa distribution sur le scalp selon plusieurs zones. Précisons que, même si la méthode EEG ne permet pas la localisation spatiale précise des processus cognitifs, la distribution d'une composante PE est définitoire : deux composantes similaires sur le plan de leur latence, amplitude et polarité, mais qui se développent sur des régions différentes du cerveau, reflètent sans doute des processus cognitifs différents.

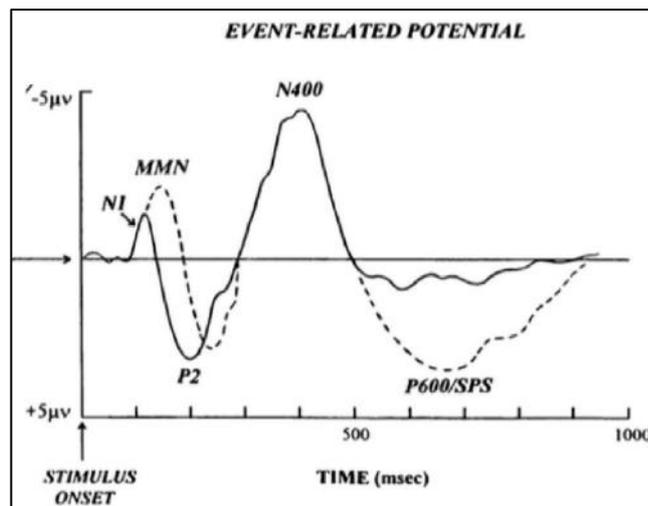


Figure 16 – Illustration de quelques unes des composantes PE les plus connues de la méthode EEG (source : *Cognitive Neuroscience of Language Lab*, http://faculty.washington.edu/losterho/erp_tutorial.htm). Par convention, la négativité est représentée dans la partie haute de la représentation axiale.

La latence à laquelle se développe la composante est proportionnelle à la précocité de son traitement : les composantes les plus précoces (N100, P200) reflètent des processus de bas niveau²⁷, les plus tardives reflétant les processus de plus haut niveau. Ainsi, les composantes les plus connues et les mieux décrites sont les composantes tardives N400, reflétant l'intégration sémantique, et P600 reflétant l'intégration syntaxique. D'autres composantes comme ELAN (Early Left Anterior Negativity) et LAN (Left Anterior Negativity) ont été décrites, qui reflèteraient des aspects morphosyntaxiques spécifiques (voir Friederici, 2011, pour une revue de littérature sur les composantes EEG).

Le modèle initial de Friederici (2002) ne permet de rendre compte que partiellement (et tardivement) des processus dynamiques, non linéaires du traitement du langage car il prend le niveau syntaxique comme socle structurel. Les aspects prosodiques ne sont pas pris en compte. Dans sa version révisée (Friederici, 2011), la composante prosodique apparaît de manière parallèle aux autres niveaux linguistiques, et seulement dans un décours temporel tardif

²⁷ La composante MMN ('*MisMatch Negativity*'), bien que précoce, reflète des processus plus complexes faisant appel également à la mémoire à long terme. Elle sera présentée plus en détails dans le § 14.5, p. 118.

d'intégration et d'interprétation du message, comme on peut le voir par le jeu de couleurs de la Figure 17.

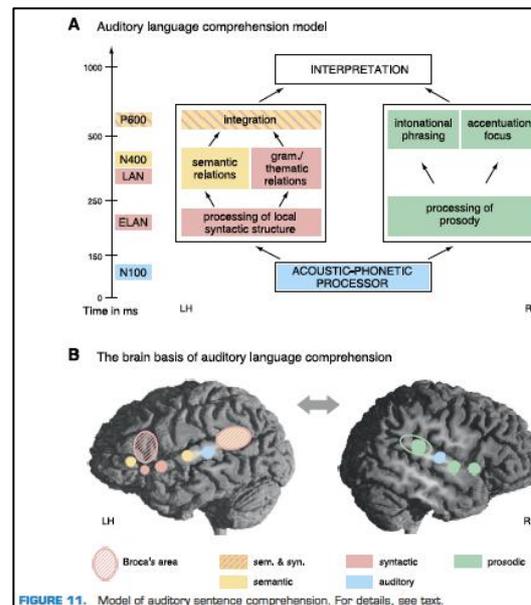


Figure 17 – A. Le modèle de compréhension du langage parlé et son déroulement temporel. B. Neuroanatomie de la compréhension du langage parlé. (Adapté de Friederici, 2011).

L'absence de composantes spécifiques dédiées au traitement de la prosodie s'explique de deux manières : d'une part, très peu d'études se sont pour l'instant penchées sur les aspects prosodiques du langage oral ; d'autre part, les dimensions linguistiques sont souvent confondues, et les composantes prosodiques sont généralement corollaires aux aspects linguistiques étudiés sans que soient précisément définies les dimensions et les relations qui les unissent. En réalité, les aspects syntaxiques sont le plus souvent étudiés, confondus avec leurs corrélats prosodiques de contours intonatifs ou de pauses, par exemple. On voit d'ailleurs dans le modèle de Friederici que l'intonation, l'accentuation et le focus sont portés au même niveau tardif d'interprétation. Le caractère lexical de l'accentuation par exemple n'est pas pris en compte. Nous reviendrons, dans la section § 14, p. 114, sur les recherches récentes concernant les bases neurales de la prosodie, illustrées également par nos propres résultats dans ce domaine.

8. LES MODELES PRE-LEXICAUX DE RECONNAISSANCE DES MOTS ET D'ACCES AU SENS

Toute la difficulté des modèles réside dans l'absence de marque explicite de frontières de mots dans la langue orale. Deux types de modèles de segmentation et d'accès au sens ont été proposés pour tenter d'expliquer les stratégies des auditeurs dans le décodage de la langue orale : les modèles linéaires post-lexicaux et les modèles pré-lexicaux (pour une revue prenant

en compte les questionnements liés au français, voir Wauquier-Gravelines, 1999, et Bagou & Frauenfelder, 2002).

Les modèles linéaires post-lexicaux (comme TRACE ou Cohorte) sont fondés sur la séquentialité (temporelle) de la parole et prennent en compte le contexte phrastique dans les processus de reconnaissance des mots. Ainsi, la vitesse et la précision avec lesquelles les auditeurs identifient un mot vont dépendre de facteurs tels que les propriétés structurales des mots (longueurs, phonologie, morphologie, règles phonotactiques ...), leurs propriétés distributionnelles (fréquence des mots et de leurs unités constitutives : phonèmes, syllabes), et le contexte phrastique (informations syntaxiques, sémantiques et pragmatiques) qui vont faciliter la prédictibilité des mots. Ces modèles présentent le désavantage d'être trop fortement influencés par la notion de 'mot' du langage écrit (particulièrement délicate en français, avec les phénomènes de liaison notamment), mais surtout, de ne pas prendre en compte l'information prosodique qui est pourtant partie intégrante de la forme sonore.

Les modèles pré-lexicaux (ou par 'points de repères' ou modèles par 'information saillante') sont fondés sur la critique des modèles linéaires basés sur l'écrit. Ils proposent de prendre en compte l'information phonologique, à la base de la langue parlée, ainsi que du contexte de communication. Ils n'excluent pas l'information séquentielle et temporelle de la parole, mais autorisent la prise en compte d'informations rétroactives. Ainsi,

« l'accès au lexique est réalisé par activation parallèle de l'information segmentale et suprasegmentale et par convergence de cette double information avec l'information lexicale descendante multilinéaire phonologiquement représentée » (Wauquier-Gravelines, 1999:149).

Ces modèles sont contraints par le système phonologique de la langue, et à ce titre, sont moins transposables d'une langue à l'autre que les modèles post-lexicaux. L'écueil de ces modèles réside dans le fait qu'il faut une connaissance très fine de la phonologie prosodique de chaque langue. Ils nécessitent un ancrage fort des propositions psycholinguistiques en linguistique théorique, notamment en phonologie prosodique. Mais, à l'instar des modèles de phonologie prosodique, les modèles pré-lexicaux métriques sont fondés essentiellement sur les présupposés phonologiques des langues à accentuation lexicale indiscutable comme l'anglais, ce qui pose quelques problèmes de conceptualisation des processus d'accès au sens pour le français.

Les modèles pré-lexicaux partent du principe que *la structure de performance* prédit mieux la perception de la parole que la structure syntaxique (Grosjean & Gee, 1987). De fait, l'organisation perceptive initiale des phrases parlées se passerait en réalité en termes de

constituants prosodiques, chaque langue utilisant ses propres règles (Mehler *et al*, 1981 ; Cutler *et al*, 1986).

A. Cutler a largement contribué à ce débat (Cutler *et al*, 1986 ; Cutler & Norris, 1988 ; Cutler & Butterfield, 1992 ; Cutler *et al*, 1997 ; Cutler, 1990 ; Cutler, 1999) et ses différentes propositions ont influencé les (trop rares) études sur le traitement de la parole sur la base des indices prosodiques. Le principe que défend Cutler est le suivant : les patrons accentuels (*stress*) seraient la porte d'entrée à la segmentation de la parole continue et l'accès au sens. De fait, les syllabes accentuées (*stressed*) seraient un point d'accès au lexique mental privilégié. Dans les langues comme l'anglais, cette stratégie serait particulièrement efficace puisque 90% des mots lexicaux sont marqués à l'initiale par un accent primaire ou secondaire. Il s'agit de l'hypothèse de *Stratégie de Segmentation Métrique* (MSS : *Metrical Segmentation Strategy* ; Cutler & Norris, 1988). C'est une élaboration de la proposition de Grosjean & Gee (1987) selon laquelle, lors du traitement de la parole, la marque '*stressed*' ou '*unstressed*' est inscrite dans la description des segments phonétiques au niveau de la représentation pré-lexicale ; l'accès au lexique serait donc initié dès que le système de traitement de la parole détecte une syllabe accentuée (*stressed*). L'interprétation de l'input de parole en représentations accentuelles / métriques abstraites pallierait le problème de la grande variabilité des indices acoustiques de l'accent.

D'autres indices sont évidemment utilisés pour segmenter le flot de parole continu, comme les probabilités phonotactiques, la connaissance du lexique, le contexte phrastique. En fonction des circonstances de la communication, ces indices peuvent être plus ou moins utilisés ; en contexte bruité par exemple, les indices prosodiques sont particulièrement utiles (Cutler & Butterfield, 1992). Une critique de la MSS est formulée par Mattys & Samuel (2000) en ce qu'elle serait inopérante pour les mots non accentués à l'initiale et source d'erreurs pour les mots polysyllabiques comportant des accents secondaires internes. Leur '*Stress-Based Strategy*' propose alors que les auditeurs seraient sensibles aux différences accentuelles subtiles sur le plan acoustico-phonologique, et ne segmenteraient que sur la base des accents primaires. Une conséquence de cette théorie concerne la charge mnésique qui serait plus lourde pour les mots ne débutant pas par une syllabe accentuée primaire.

8.1. Limites des modèles psycholinguistiques pour le français

Cutler déplore que très peu d'études aient été menées sur les stratégies de segmentation et d'accès au sens en français, comparativement aux autres langues à accentuation lexicale comme l'anglais (Cutler *et al*, 1997). Mais elle dit également que la stratégie de segmentation

métrique ne pourrait pas s'appliquer en français, car chaque syllabe serait métriquement forte (Cutler, 1990 ; 1999). C'est sans doute une conséquence de la vision traditionnelle de la prosodie du français, qui pose en effet cette langue comme une langue à rythmicité syllabique, dans laquelle l'accent final, congruent aux frontières prosodiques, n'a pas de réel statut métrique ('langue de frontière': Vaissière, 1992; Beckman, 1992 ; cf § 4, p. 33 et suivantes). La particularité de la prosodie du français dans cette conception courante est donc sans doute à l'origine de la rareté des études sur cette langue en psycholinguistique.

Faisant écho à la proposition de Cutler de segmentation syllabique en français, d'autres recherches en psycholinguistique postulent que les auditeurs francophones utiliseraient également une Heuristique (universelle) de Segmentation à l'Initiale des Syllabes (*Syllable-Onset Segmentation Heuristic* : SOSH; Content *et al*, 2001). Par voie de conséquence, les débuts de syllabes constitueraient un point d'alignement potentiel avec un début de mot lexical, l'accès au lexique pouvant se produire à l'initiale de chaque syllabe (Dumay *et al*, 2002 : expériences de 'word spotting'). Cette heuristique serait efficace puisqu'il y a une grande proportion de mots bisyllabiques en français ; les débuts de syllabes coïncideraient donc souvent avec des débuts de mots lexicaux.

L'hypothèse syllabique en français comme variante à l'hypothèse métrique en anglais ne serait cependant pas une stratégie viable en ce sens qu'elle impliquerait une stratégie strictement linéaire de la segmentation de la parole en français.

« Si les francophones développaient une stratégie de segmentation alignée sur les frontières syllabiques et considéreraient que toutes les frontières de syllabes peuvent constituer un début de mot, ils segmenteraient systématiquement de manière erronée les suites contenant une liaison ou un enchaînement » (Wauquier-Gravelines, 1999 : 145).

Il faudrait en réalité envisager un niveau intermédiaire sur lequel se calculerait la syllabation, impliquant des procédures rétroactives. Comme proposé par Grosjean & Gee (1987), ce niveau relèverait de la structure de performance et serait marqué par les informations prosodiques : cette unité de traitement serait le *Syntagme Phonologique* de Selkirk (1980), quelle que soit la langue, tenant compte de la tendance à la construction de groupes accentuels préférablement de 3 ou 4 syllabes n'excédant pas 6 ou 7 syllabes (Pascdeloup, 1990 ; Delais, 1995). Cette taille d'unité est imposée par la capacité d'encodage des informations (Levelt, 1989).

Sur cette base, on aurait pour l'anglais une combinaison des stratégies *métrique* et *prosodique*, alors que le français ne disposerait que d'une stratégie de segmentation

prosodique permettant de rendre compte des règles de Sandhi propre au français comme la liaison et la resyllabation (Wauquier-Gravelines, 1999). Cette segmentation prosodique aurait lieu au niveau du groupe intonatif, dont le contour définitoire est assez flou puisqu'il est à la fois défini par la présence d'un FA (dont on n'est pas très sûr de la réalité métrique) et des syllabes inaccentuées le précédant, et sur la présence d'un contour montant-descendant davantage définitoire du niveau de IP que du Syntagme Phonologique.

La critique que l'on peut faire à ce type de considérations est le mélange des niveaux d'analyse prosodiques : l'unité de traitement est à la fois définie par les contours intonatifs, des allongements de durée, la présence d'accents (*stress*)... Autrement dit, on ne reconnaît pas ouvertement la réalité métrique de l'accent final, seulement sa participation au traitement de l'oral à un niveau très élevé de la constituance prosodique. Cette 'hésitation' définitoire se retrouve en permanence dans toutes les approches psycholinguistiques traitant du français. Bien que sans doute partielle, cette vision selon laquelle l'allongement final de frontière intonative pourrait être utilisé en français permet cependant d'envisager une adaptation de la MSS pour le français (Kolinsky, 1998 ; Wauquier-Gravelines, 1999).

Outre une première étude montrant l'utilisation des mots fonctions (informations phonotactiques) comme marqueur gauche du début des mots lexicaux (Christophe *et al*, 1997), une première tentative d'application de la MSS au français avait été faite, par le biais d'une tâche de désambiguïsation entre 1 ou 2 mots ('bord#dur' vs 'bordure') sur la base de l'allongement final (patron fort/fort vs faible/fort ; Banel & Bacri, 1994). L'utilisation de compétiteurs lexicaux indiquait dans une autre étude (Christophe *et al*, 2002) que les auditeurs se fient aux frontières prosodiques pour segmenter ('le chat grincheux' vs 'le chat # grimpaît', où 'chagrin' est considéré comme le compétiteur lexical du premier en l'absence de frontière prosodique au dessus du mot). Considérant que l'analyse de Christophe et collègues manipulait en réalité les frontières *syntaxiques*, Michelas & D'Imperio (2010) ont adapté le protocole de Christophe et montré que les auditeurs sont capables de récupérer l'information lexicales en utilisant l'information de frontières prosodiques de niveau des *ap*, marquées par un accent final H*, indépendamment des frontières *syntaxiques*²⁸.

Un autre paradigme expérimental a donné lieu à une série de publications dans les années 1990 et 2000, où l'on montrait que les auditeurs utilisent les différents indices acoustico-phonétiques de l'accentuation dans la reconnaissance des mots d'une langue artificielle. Ces expériences concernaient essentiellement l'allongement final (Banel *et al*, 1998), la durée et

²⁸ Voir néanmoins § 5.4, p.40, notre critique de l'annotation AM proposée par Michelas & D'Imperio (2010) de ces deux phrases, qui révèlent selon nous une toute autre organisation métrique que celle annotée ici par les auteurs.

les variations de f0 sur FA (Bagou *et al.*, 2002), et enfin la présence/absence de IA et FA sur des pseudo-mots isolés (Bagou & Frauenfelder, 2006). Cette dernière étude montrait la préférence pour les patrons métriques FA seuls, avant le patron IA-FA et le patron IA seul. Le fait que le patron IA-FA ne ressorte pas préférentiellement est interprété par le fait que les pseudo-mots sont isolés et non pas contextualisés et précédés par un mot outil (déterminant). En effet, deux études avaient auparavant montré que la montée initiale de F0 (Hi) est utilisée par les auditeurs pour segmenter les mots (Welby, 2003 ; 2007), et pour délimiter les unités de sens (Rolland & Loevenbruck, 2002), laissant augurer d'une implication bien plus importante de IA dans la segmentation et l'accès au sens.

8.2. Notre proposition pour les modèles de traitement de la parole

Nous reviendrons sur ces quelques études prenant en compte IA dans les stratégies de segmentation de la parole dans le § 16, p. 140 après la présentation de nos résultats de production, de perception et de traitement de IA. Notons d'ores et déjà deux paradoxes à ce dernier type de recherches testant IA et FA dans la segmentation et l'accès au lexique, qui expliquerait pourquoi elles n'ont pas été poursuivies ni, pour la plupart, publiées en revue internationale : 1) elles testent l'utilisation de IA et FA sur des mots isolés alors même qu'elles s'inscrivent dans les théories prosodiques défendant le statut post-lexical de l'accent en français. Cette contradiction est gênante dans ce type de recherches car elle empêche de conclure aussi aisément sur le statut de l'accentuation dans les processus de traitement de la parole; 2) IA est considéré comme un phénomène *rythmique* et non métrique, et, bien que participant à la définition du niveau prosodique minimal en français (*ap*) dans l'approche AM de Jun & Fougeron, il n'est pas caractérisé comme '*pitch accent*'²⁹, *i.e.* relevant de l'identité métrique de la langue. La dissymétrie fonctionnelle entre les deux accents nuit également à la portée scientifique de ces études.

La citation ci-dessous résume selon nous la vision (encore actuelle) du rôle de la prosodie dans les modèles d'accès au sens, que ce soit du côté des linguistes ou celui des psycholinguistes.

²⁹ Seuls Post (2000) et Di Cristo (2000) considèrent l'accent initial comme métrique, mais subordonné à l'accent final. Il en va de même pour Delais (2000), qui comme Di Cristo, considère que IA marque le niveau du mot lexical en tant que démarcateur gauche (dans le cadre d'une adaptation de la théorie de l'optimalité au français). Dans ces trois propositions néanmoins, le caractère secondaire de IA amoindrit son rôle aux niveaux de constituance supérieurs.

« L'accent en français n'a pas de valeur lexicale distinctive [...] mais sa position est fixe, donc prévisible³⁰. Il est placé sur la syllabe finale des mots polysyllabiques et se marque par un allongement perceptivement saillant de cette syllabe par rapport aux précédentes. Malgré la tendance à accentuer la syllabe initiale dans certains styles de parole (ex: le style radiophonique), la structure bref-long apparaît comme la structure métrique de base (Vaissière, 1992) » (Banel et al, 1998 : 30).

En résumé, deux stratégies de segmentation différentes, sans être mutuellement exclusives, sont proposées pour le français. L'une (l'adaptation de la *MSS* au français) postule que l'accès au sens se fait sur la base de l'accent final primaire, l'auditeur attendant la fin du mot (traitement droite-gauche) pour accéder au sens. La deuxième (*SOSH*, plus proche des propositions de Cutler sur le français) postule que chaque début de syllabe est un début de mot potentiel, et en ce sens, respecte la nature séquentielle du déroulement temporel du signal. Il apparaît cependant que, dans les deux cas, ces stratégies impliquent un *surcoût* important des *ressources attentionnelles et mnésiques* lors du traitement de la parole, lié au traitement retardé du signal, avec d'un côté, stockage temporaire du signal jusqu'à l'accent final primaire, et de l'autre, cohorte de compétiteurs lexicaux, notamment pour les mots polysyllabiques. Aucune des deux stratégies n'est donc tout à fait satisfaisante, et les règles phonologiques prosodiques du français, telle qu'elles sont représentées dans ces modèles de segmentation, n'aide pas à réduire les probabilités d'alignement au lexique.

Le flou entourant la définition des caractéristiques prosodiques du français (l'accent survit-il à l'intonation ? quel niveau d'analyse pour les règles accentuelles ?) a un impact indéniable sur les modèles de segmentation et d'accès au sens. C'est dans une perspective de réconciliation des visions prosodiques et psycholinguistiques que j'ai entrepris, après ma thèse, de m'emparer modestement du sujet. Pour opérer ce rapprochement, il me semblait devoir m'écarter des approches 'intonatives basées syntaxe' de description du français et d'épouser l'approche métrique défendue par A. Di Cristo et, plus tard, l'approche AM.

L'approche métrique permet de rendre compte du caractère probabilitaire de l'accentuation en français alternant le marquage des mots et groupes de mots entre IA et FA (Fónagy, 1980 ; Di Cristo, 1999). Considérer le marquage initial et final au niveau du mot lexical (même si dans le modèle de Di Cristo, on reste au niveau abstrait de représentation sous-jacente du mot

³⁰ Notons ici deux poncifs récurrents concernant l'accentuation en français : 1) la fonction distinctive est considérée comme primordiale (voir §5.3, p. 36 pour une discussion) ; 2) l'accentuation serait fixe et prévisible, ce qui ne prend pas en compte le caractère 'probabilitaire' intéressant proposé par Fónagy (1980) introduisant le rôle de IA dans la phonologie prosodique du français.

lexical) permet en outre d'envisager une description des stratégies de segmentation en français sous un nouveau jour. Tout d'abord, il est plus aisé sur le plan théorique de tester expérimentalement la MSS en français puisque IA et FA sont clairement considérés dans leurs liens avec l'unité grammaticale 'mot lexical' comme dans les autres langues à accentuation lexicale; ensuite, considérer l'accentuation initiale nous permet par exemple de proposer que toutes les syllabes ne seraient pas des débuts de mots potentiels (Dumay *et al*, 2002; Christophe *et al*, 1997) : seules les syllabes initiales de mots lexicaux portant IA seraient des débuts potentiels de mots. Par là même, on baisse d'un cran le niveau explicatif des stratégies d'accès au sens (passage du syntagme intonatif au mot prosodique défini par deux accents) et l'accent initial nous permet de ne pas prendre toutes les syllabes comme des débuts potentiels de mots. Le coût cognitif d'un francophone lorsqu'il segmente la parole pour accéder au sens serait ainsi comparable à celui des autres locuteurs des langues du monde ! Et le français serait finalement beaucoup moins exotique qu'on ne le pense ...

Cette proposition émane de nos recherches sur différents styles de parole, incluant la parole spontanée, qui nous permet de répondre à la critique tout à fait justifiée de certains psycholinguistes, selon laquelle les modèles sont peu adaptés car fondés sur l'analyse de parole de laboratoire :

“The differences that can be observed in both the input characteristics and the listeners' processing mechanisms for continuous spontaneous speech and read speech (isolated and continuous) also raise some serious questions about the validity of the current models of spoken word recognition. Indeed, these models have been elaborated primarily to account for human recognition performance with laboratory stimuli only. In trying to adapt these models to the processing of continuous spontaneous speech, we must ask whether their structure has to change fundamentally, or whether additional modules are required?”
(Frauenfelder & Floccia, 1999 : 27).

Je propose que la segmentation et l'accès au sens se ferait sur la base du patron métrique bipolaire au niveau du *pw*, accordant un poids métrique aux deux accents IA et FA, plutôt qu'au niveau du syntagme accentuel, voire intonatif, marqué par le seul « accent » final.

Ma recherche s'est donc appliquée à démontrer la validité de cette proposition, et a été menée de front entre expériences de production et de perception sur de gros corpus de parole contrôlée dont les résultats étaient comparés à ceux issus de corpus de parole spontanée, ainsi qu'au travers d'expériences de neuroimagerie (EEG).

FUGUE

Pour un ancrage métrique des descriptions prosodiques du français

Pour rendre compte au mieux de l'accentuation du français et de son rôle dans le traitement du langage, il me semble fondamental d'adopter une approche résolument ancrée dans le *mètre*. Pour cela, il faut d'une certaine manière reconnaître le niveau du mot lexical comme domaine de l'accentuation du français, que ce soit au niveau abstrait (Di Cristo, 2000 ; Delais-Roussarie, 2000 ; Post, 2000) ou au niveau de réalisation en surface (Astésano & Bertrand, 2016).

Je commencerai par présenter ce que l'on entend par *mètre* dans sa relation au *rythme*, et la nécessité de prendre davantage en compte les paramètres de durée dans la définition métrique du français notamment. Cela me conduira ensuite à prendre en considération la notion de *proéminence cognitive* (Levelt, 1989 ; Di Cristo, 2011) et à défendre l'idée de passer par la perception comme fenêtre d'analyse de l'accentuation du français. Je terminerai par la présentation des modèles métriques qui ont inspirés ma recherche, et permis de valider expérimentalement la réalité métrique de IA et FA au niveau du Mot Prosodique Complexe (Astésano, 2001), tout en arguant de la nécessité d'adopter une approche 'mixte' entre AM et ces modèles métriques du français.

9. MÈTRE, RYTHME ET PARAMÈTRE DE DURÉE

Mètre et Rythme sont intimement liés (voir Astésano, 2001 pour une discussion). Le mètre relève définitivement du niveau abstrait des relations de proéminences entre elles. Le rythme est davantage à l'interface entre le mètre et les contraintes de surface, qu'elles soient prosodiques (taille des constituants, règles rythmiques) ou structurelles (structures syntaxique, sémantique, informationnelle). Les principes sous-tendant la constitution d'une grille métrique sont des principes rythmiques. Dans les approches métriques sur le français, on reconnaît le principe métrique de culmination (Di Cristo, 2000) ou culminativité (Delais-Roussarie, 2000) selon lequel, parmi tous les temps d'un groupement, un seul va être proéminent au niveau du mot, du syntagme accentuel ou au delà. On reconnaît également un principe d'alternance rythmique, selon lequel les syllabes accentuées vont ressortir (*forme*) par rapport aux syllabes inaccentuées (*fond*) de manière régulière. La récurrence est en effet un principe organisationnel fondateur du rythme, qu'il s'agisse du rythme linguistique, du rythme musical ou relevant de tout autre matériau sonore (Fraise, 1974). Pour Fraise, le rythme est l'élément fondateur, organisateur des proéminences. En ce sens, ce ne sont pas les

proéminences, les accents, qui constituent le rythme, mais c'est plutôt le rythme qui permettrait de révéler les proéminences.

9.1. Notion de 'battement' métrique ('beat')

En réalité, notre capacité à percevoir des éléments proéminents est corrélée avec la notion de 'battement' ('beat' en anglais). Le battement sous-jacent nous permettrait de percevoir les éléments comme appartenant à un groupement. Ce battement serait lié au tempo 'naturel' et 'spontané' de l'être humain qui se situe aux alentours de 600 ms (1,7 battements par seconde). Au delà de ce tempo, le rythme n'est plus perçu car la succession régulière des événements (à la base du rythme) n'existe plus : la cohésion est détruite. Pour être perçu, une proéminence accentuelle devra donc idéalement coïncider avec ce battement, ce rythme naturel (Handel, 1989). Il se trouve que l'analyse de plusieurs langues différentes (dont le français et l'anglais) révèle une tendance commune à placer les proéminences accentuelles en moyenne toutes les 550 ms (Fant *et al*, 1991). Les locuteurs lecteurs tendent en outre à planifier les pauses et l'allongement prépausal associé de sorte qu'ils coïncident avec un multiple entier de la valeur moyenne des pieds accentuels (550 ms). Le rythme accentuel est ainsi mis en valeur par les pauses. L'analyse de 3 styles de parole m'avait permis de confirmer ces résultats d'équilibrage interaccentuel au niveau du Mot Prosodique marqué par FA, pour les styles spontanés tout au moins (en moyenne 600 ms et 3.7 syllabes ; Astésano, 2001). Ces résultats confirmaient également les travaux antérieurs de Padeloup (1990) et Delais (1995) qui montraient cette moyenne des groupes accentuels sur des phrases isolées ou lues.

A ce stade, il convient de faire deux remarques : d'une part, la rythmicité accentuelle est bien présente en français, et le battement ne situe pas essentiellement au niveau de la syllabe dans notre langue (le '*semi beat*' syllabique de Selkirk, 1984b) mais bien au niveau du groupe accentuel ou Mot Phonologique, comme en anglais (niveau du '*beat*' de Selkirk ; voir Figure 18). De fait, les contraintes d'équilibrage rythmique se combinent avec des contraintes biologiques universelles indépendantes des contraintes rythmiques liées au code linguistique. D'autre part, dans la parole comme en musique, la proéminence acoustique et les événements (mélodiques) proéminents vont de pair avec le temps fort sous-jacent, et les mouvements de hauteur les plus forts coïncident généralement avec d'autres indices de proéminences et le mètre sous-jacent. Mais rien n'empêche une dissociation entre battements rythmiques sous-jacents et la forme acoustique de surface, où le marquage par des traits accentuels relevant de l'intonation, de la durée, ou de la force articulatoire peut ne pas forcément coïncider avec le mètre sous-jacent. L'importance de distinguer la force structurelle des événements

mélodiques et de la proéminence acoustique s'éclaire aussi par l'analogie avec la musique, comme le souligne Ladd (2008).

« By and large, syllables that are prominent in the abstract metrical structure are also phonetically stressed, but the theory does not preclude the occurrence of systematic dissociations. A syllable can be metrically strong or prominent without necessarily being stressed. Such a dissociation is comparable to those found in music: for instance, dissociations between rhythmical prominence (defined by the underlying beat) and melodic or dynamic prominence (defined by the harmonic changes, note durations, added loudness, etc.). » (Ladd, 2008 : 61-62)

Ce dernier point doit être gardé en mémoire car nous verrons que la difficulté de rendre compte de l'accentuation du français, notamment dans l'approche AM, relève aussi beaucoup de la non prise en compte des paramètres de durée, de rythme ou tout simplement de mouvements mélodiques non exclusivement représentés par des tons H.

La figure ci-dessous illustre le phénomène de battement ('beat') au niveau du groupe accentuel que l'on peut trouver y compris en parole spontanée en français. Cet exemple est issu d'un corpus d'interview collecté dans Astésano (2001)³¹. Notons que l'énoncé trouve son point d'équilibre rythmique sur la syllabe finale focalisée du mot 'abstrait'.

³¹ J'utilise très souvent cet exemple pour faire comprendre aux étudiants ce qu'est la proéminence métrique ou cognitive, et la tendance à la récurrence régulière des syllabes proéminentes en français.

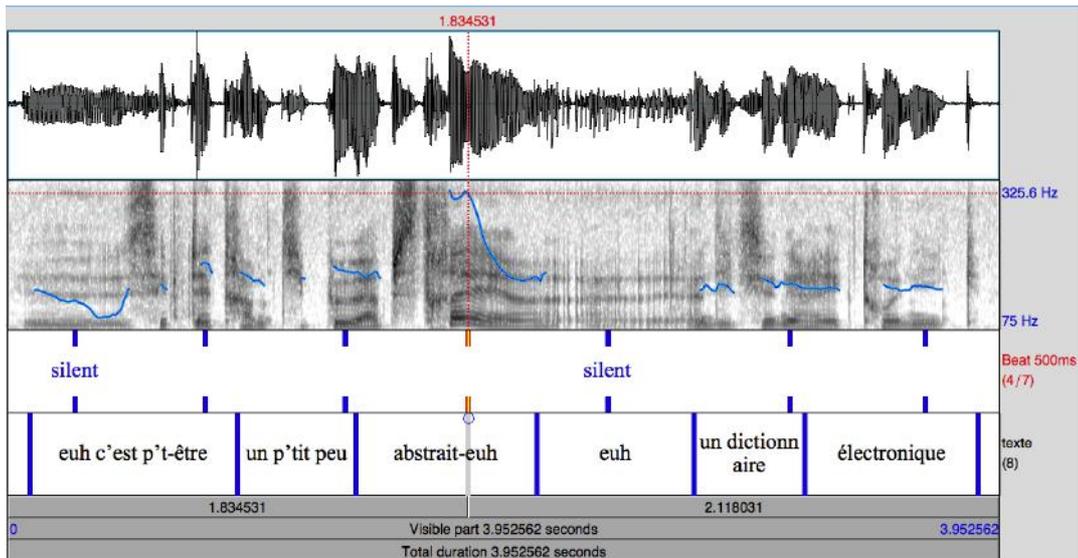


Figure 18 – Récurrence régulière des proéminences accentuelles en parole spontanée (incluant les durées des événements rythmiques tels que pauses, hésitations vocales et allongements finaux) correspondant aux battements métriques sous-jacents (extrait d’une interview radiophonique issue du corpus de Astésano, 2001 ; *exemple sonore n°2*). Notons la construction syntaxique typique en français d’extraction droite du syntagme nominal placé en position post-focale. Malgré l’aplatissement tonal typique des constructions post-focales demeure l’allongement final de mots sur ‘dictionnaire’ et ‘électronique’. Le battement (‘beat’) est placé sur les voyelles des mots prosodiques et sur la partie médiane des hésitations vocaliques ‘euh’ (notés ‘silent beat’). Ces battements interviennent entre 500 et 600 ms, ce qui correspond à l’intervalle inter-accentuel moyen dans de nombreuses langues (Fant *et al*, 1991), *i.e.* toutes les 3,7 syllabes (Padeloup, 1990 ; Delais, 1995 ; Astésano, 2001), correspondant au rythme biologique chez l’humain (Handel, 1989).

La perception des syllabes métriquement fortes en position post-focale peut s’interpréter à la fois comme l’impression (au sens d’‘imprimer’) d’une récurrence imposée par le mètre sous-jacent (battement) ou par l’impression ‘top-down’ de nos connaissances lexicales (Ladd, 2008 : voir § 5.5, p. 50). Les deux interprétations ne sont pas mutuellement exclusives.

9.2. Perception du mètre et du rythme – prise en compte nécessaire de la durée en français

La sensation de rythme et de mètre est naturellement, mais pas exclusivement, impactée par les caractéristiques physiologiques et acoustiques de l’accentuation. Sur le plan perceptif, la *position* perçue de l’élément accentué varie en fonction de la nature physique de l’accent : si l’accent se réalise par une augmentation de l’intensité et de la hauteur, il sera perçu comme débutant le groupement ; s’il est marqué par une augmentation de durée, il sera perçu comme apparaissant à la fin du groupement (Fraisse, 1974). L’anglais et le français sont ainsi généralement décrites comme étant deux langues prototypiques de langues capochrones (ou *leader-timed*) et codachrones (*trailer-timed*), correspondant sur le plan métrique à des gabarits dits ‘têtes à gauche’ et ‘tête à droite’ respectivement. Sur le plan physiologique, cette différence entre l’anglais et le français correspond aux cycles de tension et relaxation musculaire caractéristiques du rythme parolier. Pour les langues ‘*trailer-timed*’ comme le

français, les syllabes inaccentuées sont situées au début de la tension musculaire croissante jusqu'à l'accent final, impliquant une très faible centralisation des voyelles en syllabes inaccentuées. A l'inverse, les langues 'leader-timed' comme l'anglais sont caractérisées par un pic de tension musculaire sous l'accent, tandis que les syllabes inaccentuées sont situées dans la partie de relâchement articuloire et sont donc centralisées (Wenk, 1983). Si l'on se place du point de vue de la théorie de la forme, l'effet perceptif mettra davantage en exergue les syllabes accentuées de l'anglais sur fond de syllabes inaccentuées neutralisées articuloirement, alors que la régularité syllabique ressortira davantage en français, atténuant l'effet perceptif des syllabes métriquement fortes. Mais l'impression auditive d'une certaine 'régularité' syllabique n'empêche pas, nous l'avons vu, de considérer que le français est aussi marqué par la récurrence régulière des syllabes accentuées.

Le mètre iambique ('tête à droite' ou 'trailer-timed') en français induit donc naturellement un rapport particulier et étroit avec le paramètre de durée, *i.e.* d'allongement particulier des syllabes proéminentes. Contrairement à ce qui en est dit par Fry (1958) ou Allen (1975), qui placent le paramètre de f_0 comme le paramètre le plus pertinent sur le plan de la perception des proéminences, Delattre (1966b) puis Rossi (1972) ont montré que la durée joue un rôle non négligeable dans la perception de la proéminence en français.

L'importance du paramètre de durée pour le français ressort d'une analyse comparative de 4 langues (anglais, espagnol, allemand, français) de l'allongement syllabique métriquement motivé (Delattre, 1962 ; 1966b). Delattre montre que les syllabes finales d'énoncés montrent le rapport de différence de durée entre syllabes inaccentuées et accentuées le plus important (de 1 à 2). L'allongement plus important des syllabes accentuées en français est expliqué d'une part par son poids métrique, et d'autre part par sa position finale, favorisant naturellement l'allongement (pré-frontière). Puisque la proéminence accentuelle en français ne repose pas sur un surcroît d'intensité (-10% d'intensité sur les syllabes accentuées), comme c'est le cas pour les autres langues, la durée doit seule se charger de la proéminence accentuelle.

« Le rôle de la durée est très positif. La durée est le seul des trois éléments acoustiques qui soit toujours, par sa proéminence, un facteur de l'accent. C'est le seul qui puisse y varier indépendamment des deux autres. Et c'est, dans le sens positif, l'élément le plus étroitement uni à l'accent. » (Delattre, 1966b : 68).

Il faut donc noter que la *position* finale de l'accentuation dite primaire en français induit une analyse des paramètres de durée qui doit rendre compte à la fois de l'allongement final pré-frontière communs à toutes les langues (Fletcher, 1991) mais aussi d'un allongement

propre à la proéminence métrique finale (FA). Une des difficultés de description de l'accentuation du français réside donc non seulement dans le syncrétisme de l'accentuation et de l'intonation sur le plan tonal mais aussi sur le plan de la durée des faits prosodiques. Dans le cas d'un pied iambique (faible-fort), en effet, on peut considérer que les syllabes sont *métriquement* mais aussi *rythmiquement* allongées, de par leur position dans la structure métrique.

Dans une étude sur la perception de la durée sur plusieurs langues, Rossi (1972) montre que le seuil de perception de la durée est influencé par l'ordre d'apparition des stimuli. Si le stimulus étalon précède la variable, le seuil de perception de durée est plus fin. Le pied iambique du français (alternance faible/fort) favoriserait donc la perception de la proéminence marquée par la durée (typiquement, FA) puisque le poids métrique de la syllabe finale serait renforcée par la syllabe inaccentuée étalon précédente. Dans cette perspective, on peut de fait considérer que les proéminences finales en français jouissent d'une représentation cognitive chez le locuteur natif *indépendante* de la notion de frontière (Astésano 2001 ; Astésano *et al*, 2012, 2013 ; Astésano & Bertrand, 2016). Les auditeurs sont en effet capables de percevoir des différences de variations de durée syllabique des différents FA. Plus particulièrement, ils sont capables de percevoir de manière catégorielle les syllabes inaccentuées vs FA de mot (+ 50%) vs FA de syntagme intonatif majeur (+ 100% par rapport à la durée FA de mot), alors que les variations de f_0 permettent d'observer le plus souvent des jugements perceptifs graduels. Ces tendances sont issues d'une étude de perception catégorielle sur la base d'une incrémentation par pas de la durée syllabique et des valeurs de f_0 (Jankowski, 1997).

10. LA PERCEPTION COMME FENETRE SUR LA PROEMINENCE COGNITIVE

Ces considérations sur le mètre et le rythme nous invitent à adopter la conception de *proéminence cognitive*, selon laquelle les auditeurs percevraient les éléments prosodiques non seulement au travers de leurs caractéristiques acoustiques mais également en fonction de leur réalité métrique (Levelt, 1989 ; Di Cristo, 2011). Qu'entend-on par *réalité métrique* ? Outre les contraintes rythmiques que l'on évoquait ci-dessous, il s'agit aussi du lien qui unit le niveau de proéminence abstrait avec les unités grammaticales auquel il se rattache : le mot lexical, la plupart du temps, mais aussi les autres niveaux linguistiques relevant de la structuration syntaxique, sémantique et informationnelle. Autrement dit, lorsqu'un auditeur traite le langage, il traite conjointement tous ses niveaux : les informations 'bottom-up' de traitement acoustique des faits prosodiques se font en même temps que les informations 'top-down' liés à leur connaissance de la langue (Ladd, 2008 ; Cole *et al*, 2010 ; Cole *et al*, 2014).

Comme je le mentionnais dans la partie *Prélude*, il me semble important de dépasser le niveau de l'analyse des faits prosodiques par la transcription (généralement experte) des faits phonologiques, puisqu'il induit au moins deux biais principaux: 1) la tendance du transcrip-teur à transcrire les événements qu'il/elle s'attend à trouver en fonction de son prisme théorique ; 2) la lourdeur de la transcription-expert qui limite le plus souvent l'analyse à des phrases simples, construites par l'expérimentateur. Une solution est de passer par le prisme de la perception d'informants naïfs, qui ne connaissent rien à nos préoccupations de prosodistes et vont justement nous informer sur ce qui est utilisé dans le traitement de la parole continue (Cole *et al*, 2010). Les études psycho-acoustiques démontrent que les représentations mentales des événements linguistiques interviennent pour une bonne part en amont des processus perceptifs. Etudier la perception des phénomènes prosodiques est donc une fenêtre sur la représentation mentale de la prosodie d'une langue, indépendamment mais aussi en lien avec les caractéristiques acoustico-phonétiques des catégories prosodiques.

11. QUELLE APPROCHE POUR RENDRE COMPTE DE LA PROSODIE DU FRANÇAIS ?

Mon approche se situe clairement dans la lignée des travaux de Di Cristo & Hirst (1993), Di Cristo (2000) et à certains égards ceux de Delais-Roussarie (2000)³² et Post (2000). Ce qui unit ces approches est la prise en compte du niveau du mot lexical dans la représentation abstraite, marqué par les proéminences initiales et finales du français. Il s'agit donc d'approches fortement ancrées dans la métrique, à l'inverse des approches intonosyntaxiques évoquées en § 5.1, p. 34, pour lesquelles le syncrétisme de l'accentuation et de l'intonation ne permet pas d'envisager un poids métrique de l'accentuation du français. C'est sans doute pourquoi les approches intonosyntaxiques s'intéressent davantage aux fonctions de l'intonation à l'interface avec les niveaux sémantiques et pragmatiques, bien que rien n'empêche d'avoir un fort ancrage métrique pour s'intéresser à cet interfaçage (cf. Delais-Roussarie, 2016).

Qu'est-ce qui peut justifier qu'on envisage un ancrage métrique des proéminences en français ? Un argument en faveur de cet ancrage métrique est donné par Dell (1984). Dell montre que l'association du contour tonal se fait avec la dernière syllabe pleine d'un mot phonologique, mais jamais avec une syllabe contenant un schwa (voir néanmoins Sichel-Bazin, 2015 pour une discussion sur le français du midi). De fait, l'accent (FA) n'est pas

³² Mon effort de recherche est centré sur la mise en évidence du système accentuel IA-FA à un niveau de constituance très bas (le mot lexical), dans un approche expérimentale interdisciplinaire très lourde. L'interfaçage de la prosodie avec les autres niveaux linguistiques n'intervient donc dans ma recherche qu'à un niveau d'interprétation des phénomènes accentuels, et non l'inverse.

systématiquement associé avec le contour intonatif de la frontière de constituant. Le rattachement des syllabes finales contenant un schwa à la syllabe contenant une voyelle pleine précédente se fait donc dans le cadre du constituant de niveau supérieur, la Macro-Syllabe (Selkirk, 1984b). Puisque la tête du pied métrique en français ne peut être qu'une syllabe pleine ou une macro-syllabe, cela confère un poids métrique indéniable à FA (Ladd, 2008). L'argument en faveur d'un ancrage métrique de IA est moins bien défini, mais relève par exemple de son rôle dans le marquage des débuts de morphèmes dans les mots plurimorphémiques (voir discussion dans le § 5.5, p. 50).

11.1. L'approche métrique de Di Cristo (2000)

Une originalité du modèle de Di Cristo est d'appliquer le *principe de bipolarité accentuelle* (formulé empiriquement par Fónagy, 1980) à la représentation métrique du mot lexical. Ainsi, le mot lexical est marqué dans sa représentation sous-jacente par IA et FA. Le *principe de dominance à droite* suppose que la syllabe finale constitue la tête métrique du constituant : FA est donc primaire (tête) et IA secondaire (Figure 19).

			X
X			X
X	X	X	X
Fé	li	ci	té

Figure 19 – Représentation métrique du mot 'félicité' où IA et FA marquent respectivement la première et la dernière syllabe. Le poids métrique, indiqué par le nombre de croix, est plus fort pour FA (accent primaire) que pour IA (accent secondaire).

Notons que seuls les mots de contenu (noms, verbes, adverbess, adjectifs ...) peuvent recevoir ce patron accentuel : les mots grammaticaux ne sont donc pas accentuables (voir Di Cristo, 1999 ; 2000 ; Delais-Roussarie & Di Cristo, à paraître, pour une présentation détaillée).

A la différence de Rossi (1980), l'accent final sous-jacent FA est ici *générateur d'accent nucléaire* en surface. Autrement dit, en fonction des contraintes prosodiques et discursives, FA sera promu Accent Nucléaire (NA). Selon des contraintes similaires, IA pourra être le locus de l'Accent Emphatique (EMP). En surface, on retrouvera donc deux types d'accents IA et FA : les accents de niveau du mot prosodique ou *Unité Rythmique (syntagme accentuel ou ap)*³³ et les accents de niveau de l'énoncé ou *unité intonative (IP)*. Rappelons ici que le modèle de Di Cristo admet des niveaux intermédiaires modulant potentiellement la force métrique de IA et FA (voir § 5.2, p. 37).

³³ Comme dit à la fin du § 5.2 sur les niveaux de constituances en français, je ne pense pas que l'UR de Di Cristo soit comparable ni au Mot Prosodique, ni au syntagme accentuel ou *ap*, d'où l'intérêt selon moi d'introduire le niveau du *pw* en français sur la Figure 6, p. 32. J'utiliserai le terme de Mot Prosodique néanmoins pour la présentation des résultats validant le modèle de Di Cristo, pour référer à UR ou *ap*.

Aux niveaux inférieurs de constituance (l'Unité Rythmique et l'Unité Tonale), et bien que j'aie des réserves sur ces deux niveaux, l'avantage du modèle de Di Cristo est qu'il permet de prendre en compte, outre les aspects tonals de IA et FA, les aspects métriques relevant du paramètre de durée pour FA au niveau de l'Unité Rythmique. Cet aspect est selon moi central à une description complète du français, et est considéré dans une bien moindre mesure dans les approches AM du français (Jun & Fougeron, 2000, notamment).

Mon travail de thèse a consisté à proposer une validation de ce modèle pour le français, par l'analyse quantitative des différents types de proéminences, à travers les indices de durées syllabiques et infra-syllabiques, et les contours tonals.

11.2. Caractéristiques acoustiques de l'accentuation (Astésano, 2001)³⁴

Une originalité de ce travail était d'appliquer un même cadre d'analyse à différents styles de parole, allant de la lecture à la parole spontanée non contrôlée. L'analyse de la durée segmentale constituait un pôle central de ma recherche. Au-delà de l'observation des phénomènes d'allongement observés aux frontières de constituants, je proposais l'observation des phénomènes de durées segmentales au niveau infrasyllabique. Les constituants syllabiques (Attaque, Noyau et Coda) sont en effet des unités susceptibles de porter des informations prosodiques. L'allongement différentiel des éléments infrasyllabiques distingue les proéminences accentuelles initiales des proéminences finales (Campbell, 1992). Si l'on obtenait de tels résultats en français, cela constituerait un premier indice de la matérialité métrique de IA et FA.

Le repérage des proéminences s'est effectué sur la base d'un 'Gold' entre 3 experts. A l'époque déjà, nous considérons la perception comme fenêtre d'interprétation des proéminences (Di Cristo, 2000 ; 2011).

L'analyse des durées des proéminences accentuelles (Astésano, 1998 : Hirst *et al*, 1998) a révélé deux tendances: d'une part, la durée syllabique globale permet de distinguer significativement entre les accents au niveau du *mot prosodique* (IA et FA) et les accents au niveau de IP (EMP et NA), ces derniers étant plus allongés; d'autre part, l'analyse des durées infrasyllabiques permet de distinguer globalement entre les proéminences accentuelles initiales (IA, EMP), et les proéminences finales (FA, NA), les premières étant marquées par un allongement significatif de l'Attaque syllabique par rapport à la Rime (Noyau + Coda) et les deuxièmes par un allongement significatif de la Rime par rapport à l'Attaque.

³⁴ Voir Annexes : documents scientifiques 'Prosodie 1' et 'Prosodie 2'.

Globalement, les proéminences du niveau du mot prosodique ou groupe accentuel (IA et FA) ont des durées syllabiques plus courtes et des excursions de f_0 moins amples que les proéminences de niveau sémantico-pragmatique (EMP et FA nucléaire NA). IA et EMP ont tous deux des Attaques syllabiques significativement plus longues que leurs Rimes, mais la Rime de EMP se caractérise par un allongement de la Rime 80% plus important que les syllabes inaccentuées, alors que la Rime des IA de niveau du mot prosodique est plus longue que celle des syllabes inaccentuées de 25%, soit juste au dessus du seuil de discrimination des durées (*Just Noticeable Difference*) établi par Rossi (1972) et Klatt (1976). Seul l'allongement de l'Attaque distingue donc IA des syllabes inaccentuées, validant ainsi la caractérisation essentiellement tonale de IA dans les descriptions antérieures (voir notamment « l'ictus mélodique » de Rossi, 1985a). FA et FA Nucléaire NA présentent tous les deux la tendance inverse d'allongement différentiel, *i.e.* un allongement significativement plus important de la Rime par rapport à l'Attaque. La figure Figure 20 ci-dessous présente ces tendances d'allongement différentiel.

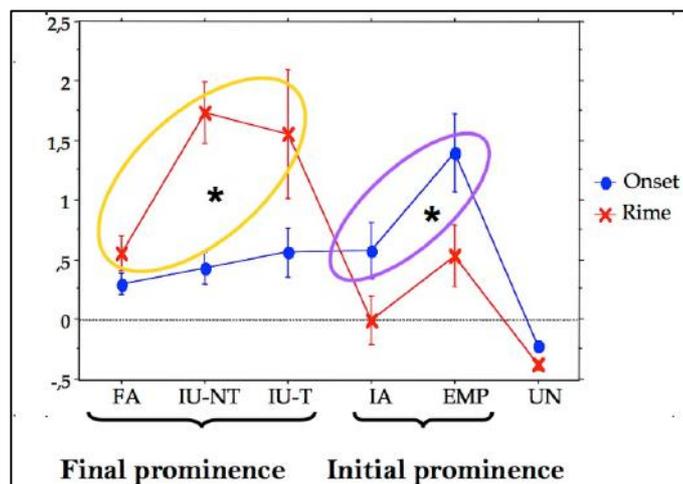


Figure 20 – Durées infra-syllabiques (Onset pour Attaque syllabique, et Rime) des proéminences initiales (IA dit 'rythmique' et IA Emphatique EMP) et des proéminences finales (FA d'ap et FA d'IP non terminale IU-NT et terminale IU-T). IU correspond à 'Unité Intonative', ou Syntagme Intonatif (majeur) dans notre terminologie actuelle. UN sont les syllabes inaccentuées. Les durées sont normalisées en *z-scores* sur l'axe des y (pour plus de détails sur la normalisation, voir Astésano, 2001).

La Figure 21 présente la manière dont les configurations tonales ont été mesurées sur ce corpus (pour plus de détails, voir Astésano *et al.*, 1997 ; 1998). Le modèle statistique montre un certain recouvrement entre les proéminences de niveau de mot prosodique (IA et FA) d'une part, et entre IA de mot et IA EMP d'autre part (voir Astésano, 2001 pour plus de détails).

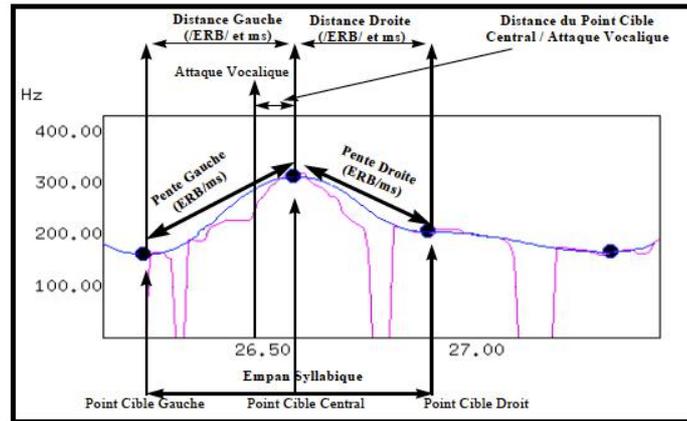


Figure 21 – Modélisation de la f_0 avec l'algorithme Momel (Hirst & Espesser, 1993) d'une proéminence perçue. 3 points cibles définissent la configuration tonale dont sont extraits les 8 paramètres (tiré de Astésano, 2001 : 241).

Il est cependant possible de conclure que, de manière générale, les proéminences initiales ont des formes de configurations tonales asymétriques (pente gauche plus abrupte que pente droite), alors que les proéminences finales sont marquées par des configurations tonales symétriques. IA et EMP ont des configurations similaires présentant une pente gauche abrupte et une pente droite s'étendant sur plusieurs syllabes. Cependant, la configuration tonale de EMP est deux fois plus ample que celle de IA de mot prosodique. Pour ce dernier également, une preuve potentielle de son caractère 'rythmique' (tel que nous le défendions à l'époque avec Di Cristo, 1999) est que sa pente droite s'étend sur un nombre de syllabes plus important que celle de EMP, interpolant ainsi le contour avec le FA suivant. Pour les accents finaux, FA et FA Nucléaire NA présentent tous les deux des configurations de pentes symétriques, celle de NA étant bien plus grande que celle de FA de mot prosodique et ne s'appliquant que sur la seule syllabe finale.

Il est intéressant de remarquer que la distinction entre IA et FA du niveau du mot prosodique ne se fait que sur la base de la pente droite, deux fois moins raide que celle de FA. La pente gauche de ces deux accents de niveau du mot est identique, comme ce que Jun & Fougeron rapporte en 1995, *i.e.* selon elles avec un mouvement ascendant entre 200 et 300 ms quel que soit le nombre de syllabes du mot. La pente droite ou mouvement descendant dépend du nombre de syllabes du mot pour IA mais est constant pour FA (entre 100 et 200 ms pour Jun & Fougeron).

Cet argument de configuration de pentes droite et gauche pour les accents IA et FA de mot constitue selon moi (dès 2001) la preuve selon laquelle IA et FA sont substantiellement et formellement marqueurs de cohésion de l'unité grammaticale qu'ils délimitent, que ce soit le mot prosodique (*pw* dans ma conception actuelle) ou le syntagme accentuel (*ap*), *i.e.* marqueurs gauche et droit de l'unité. L'unité marquée par les proéminences initiales et finales

ne doit donc pas être ‘scindée’ au niveau du pied métrique. Nos résultats (et d’autres, dont ceux de Jun & Fougeron) semblent donc ne pas aller dans le sens de cette unité tonale comme définitoire du pied métrique en français. Une autre alternative doit émerger, prenant en compte ce caractère cohésif des proéminences initiales et finales (comme décrit originellement par Jun & Fougeron, 1995) mais revisitant le statut métrique de ces deux proéminences, comme initié par les travaux de Di Cristo dans les années 90 et déjà dès le milieu des années 80 dans sa première collaboration avec D. Hirst (Hirst & Di Cristo 1984).

Le paramètre de distance du point-cible par rapport au noyau vocalique n’avait pas pu être retenu Astésano (2001) car il présentait une variabilité inter-individuelle, mais on pouvait quand même dire que les proéminences perçues les plus fortes (EMP et FA Nucléaire NA) ont des pics tardifs dans la syllabe accentuée/stressed, confirmant ainsi des résultats antérieurs sur d’autres langues. Il aurait été intéressant d’adopter ponctuellement une approche moins automatique, plus descriptive sur ce corpus, utilisant la méthodologie issue de l’AM car ce paramètre de distance ou d’alignement tonal est particulièrement pertinent dans la définition des types de proéminences.

Globalement, l’étude de plusieurs styles de parole a montré que le patron métrique bipolaire /IA-FA/, favorisant la cohésion rythmique par le marquage du début et de la fin des mots ou groupes de mot, forme une unité prosodique particulièrement pertinente en parole spontanée (Astésano, 2001). Nos résultats acoustiques pointent par ailleurs une différence structurelle entre IA et FA aussi bien sur le plan tonal que sur le plan de la durée. L’allongement significatif de l’Attaque pour IA et de la Rime pour FA, ainsi que les pentes de F0 associées à ces deux accents, sont indicatifs substantiellement de la délimitation gauche et droite des mots ou groupes de mots. Ces tendances acoustico-phonétiques de IA et FA, marqueurs gauche et droit d’unités, peuvent être résumées par la figure ci-dessous.

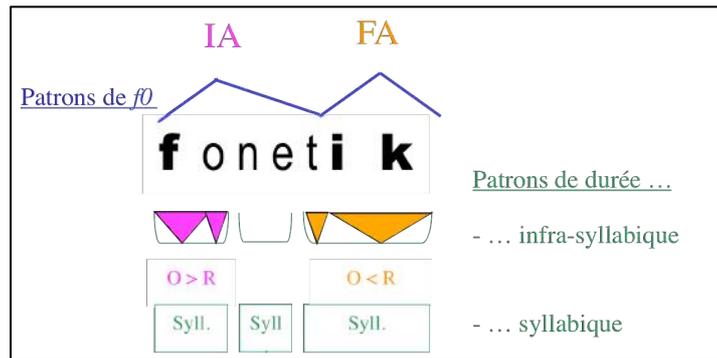


Figure 22 – Caractéristiques phonétiques schématisées de IA et FA sur le mot ‘Phonétique’, en termes de configurations tonales (en bleu) et de patrons de durée syllabique et infra-syllabique (en vert). O est l’Attaque et R la Rime. IA est caractérisé par une configuration tonale asymétrique, la pente droite interpolant le contour jusqu’au FA dont la configuration tonale est symétrique. Notons que les pentes gauches de IA et FA sont similaires. La durée syllabique de IA est 25% supérieure à celle d’une syllabe inaccentuée bien que cela ne soit imputable qu’à l’Attaque, significativement plus longue que la Rime. La durée syllabique de FA est significativement plus longue que celle de IA, l’essentiel de l’allongement étant porté par la Rime. Ensemble, ces caractéristiques phonétiques correspondent à une démarcation gauche et droite du *pw*.

Ces résultats m’avaient, à l’époque, invitée à postuler un niveau de constituance proche de *ap*, le Mot Prosodique Complexe, défini au niveau inférieur par deux pieds métriques dont le premier était marqué par IA et le deuxième par FA (Figure 23). Cette unité était la seule unité stable rythmiquement (en termes de durée et taille) que l’on retrouvait à travers les différents styles de parole : elle ressortait en effet comme une ‘figure’ métrique et rythmique commune au système quel que soit le style (Astésano, 1999).

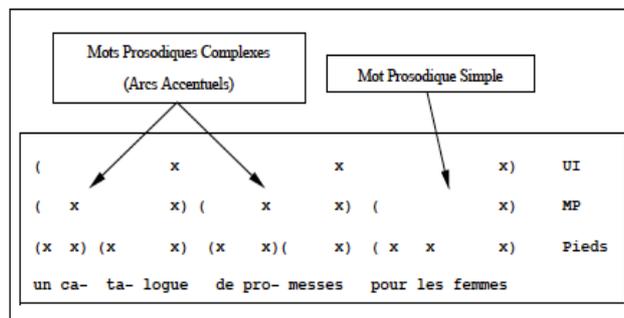


Figure 23 – Représentation des niveaux de constituance d’un extrait d’énoncé de parole spontanée (style Interview) selon la représentation en grille parenthésée de Halle & Vergnault (1987). Les Mots Prosodiques Complexes représentent le marquage accentuel bipolaire IA-FA au niveau du mot lexical, et IA est représenté comme métriquement dominé par FA. (Adapté de Astésano, 2001).

A la différence de Jun & Fougeron mais comme Di Cristo, IA y est représenté métriquement au niveau sous-jacent du mot lexical. A la différence de Di Cristo mais comme Jun & Fougeron, ce Mot Prosodique Complexe est une unité cohésive marquée à gauche et à droite, non scindable en unités de rang inférieur comme l’Unité Tonale. Je ne me prononçais pas à l’époque sur la nature du pied au rang inférieur de constituance.

Ce constituant préfigurait la proposition que je fais aujourd’hui du *pw* comme niveau d’actualisation de l’accentuation en français (§ 16, p. 140).

11.3. Une approche métrique mixte intégrant les apports de AM

Mon approche se situe résolument dans un ancrage métrique fort. On l'a vu, l'intérêt de considérer le mètre et le rythme comme centraux dans la description des faits prosodiques est que cela permet : 1) de proposer un lien avec une unité grammaticale telle que le mot lexical ou le groupe de mots, et donc de clarifier le statut de l'accentuation en français à des fins notamment d'implémentation dans les modèles de traitement du langage ; 2) de prendre en compte la notion de 'proéminence cognitive' dans ce qu'elle apporte conceptuellement aux liens unissant prosodie et autres niveaux linguistiques (connaissances lexicales et sémantiques, structuration syntaxique) ; 3) de prendre en compte le paramètre de durée dans son lien justement avec la proéminence métrique et cognitive, paramètre crucial pour le français mais trop souvent négligé au profit des paramètres tonals.

Pour cette raison, je m'inscris davantage dans les approches de Di Cristo, Delais-Roussarie et Post que dans celle de Jun & Fougeron. De l'approche de Di Cristo, je retiens aussi la quête d'invariance des paramètres accentuels à travers l'analyse systématique de différents styles de parole. Au contraire de Fónagy par exemple, ce n'est pas tant la variabilité des styles de parole qu'il m'importe de mettre à jour, mais justement ce qui, dans le système, se retrouve de manière constante. De l'approche de Di Cristo, je retiens enfin la nécessité de passer par le prisme de la perception pour rendre compte de ce système. Dans mon approche cependant, cette médiation ne passe pas seulement par l'annotation experte des faits prosodiques (trop sujette à la vision théorique du transcripteur) mais aussi et surtout par la représentation que peuvent avoir des informants naïfs de notre système.

J'ai ensuite choisi d'intégrer dans mon approche les préceptes de l'AM, pour trois raisons principales : 1) AM est l'approche dominante de description de la phonologie intonative des langues du monde : une comparaison du système du français ne peut donc que s'enrichir de la possibilité de comparaison avec les autres langues. Il faut adopter le même vocabulaire et des outils descriptifs comparables pour faire évoluer nos connaissances sur le français. 2) Les outils de transcription de l'intonation proposés pour le français (INTSINT ; Hirst *et al*, 2000) n'ont pas été assez développés pour permettre une description systématique de la phonologie intonative du français, ni pour en permettre la comparaison avec les autres langues. 3) Les outils descriptifs de AM sont très performants pour distinguer ce qui relève de la proéminence métrique de ce qui relève des phénomènes de frontière. La distinction opérationnelle que cette approche offre en termes de relation entre proéminences et frontières est au cœur de la problématique pour le français où ces deux phénomènes coexistent intimement. Cette

approche propose en outre l'utilisation très économique de 2 tons L et H dont la compositionnalité permet de relier les niveaux lexical et post-lexical de l'intonation et de faire émerger la signification par leur combinaison.

Par contre, on a pu faire la critique à l'approche AM de négliger le paramètre de durée et donc *in fine* une certaine relation au mètre. Ce paramètre de durée permet pourtant de rendre compte de certains phénomènes non explicables par le tonal seul, y compris en anglais (voir les propositions de Dilley *et al*, 2006 en ce sens). Ancrer la description des langues davantage dans le *rythme* et le mètre, comme le défendent ces auteurs pour l'anglais, justifie mon approche mixte dite 'métrique et AM' pour le français.

Comme je l'ai expliqué dans le *Prélude* (§ 4 et 5), plusieurs phénomènes en français sont problématiques. Si l'on parvient à les décrire et les expliquer, il me semble qu'on pourra considérablement avancer sur le statut de l'accentuation en français et sur son domaine d'actualisation.

Il s'agit d'une part des cas de FA non marqués par un ton H, aux niveaux de *ap*, *ip* ou même IP ; ces cas que l'on dit très rares (Jun & Fougeron, 2002) se retrouvent finalement assez souvent, on le verra, pour instancier la hiérarchie syntaxique sous-jacente (mais aussi sans doute pour éviter l'encombrement tonal). La prise en compte de ce phénomène de manière plus systématique (que l'on pourra qualifier de L* ou !H) permet pourtant de mieux circonscrire l'organisation des patrons accentuels et le jeu subtil entre IA et FA dans le marquage des constituants (§ 12.3, p. 92). Il permet en outre d'envisager la relation de syncrétisme entre accentuation et intonation de manière plus riche (§ 13.2, p. 100).

Un autre phénomène intéressant concerne les réalisations 'désaccentuées' de FA à l'intérieur des groupes de mot ou en situation post-focale. La prise en compte du paramètre de durée permet dans ces cas là de ne pas négliger certains phénomènes accentuels qui ne relèvent pas de la composante tonale seule, et d'éclairer le domaine d'actualisation en surface de l'accentuation.

Enfin, l'analyse systématique de l'alignement des tons L₁ et H₁ de IA doit être entreprise pour nous renseigner sur le statut de IA : est-il vraiment un 'marqueur flou' de frontière, à cause de l'alignement supposé 'lâche' de son ton H dans le mot ? Selon moi, Welby (2003 ; 2006) avait sans doute déjà apporté une preuve de son statut métrique sur la seule base de l'ancrage fort de son ton L au niveau du mot lexical.

Ces trois questions descriptives occupent l'essentiel de ma recherche. Selon moi, prendre en compte systématiquement IA dans l'investigation de l'accentuation en français permet

d'éclairer également le statut de FA, dans les relations qu'ils entretiennent aux différents niveaux de constituance.

Pour avancer dans cette description, j'envisage la constituance prosodique à 3 niveaux (*ap*, *ip* et *IP*) et n'exclus pas la possibilité de récursivité des constituants prosodiques pour rendre compte de la profondeur de la structure prosodique. Cependant, à la différence des approches métriques et AM dont je m'inspire (Di Cristo, Delais-Roussarie, Jun & Fougeron et Post), j'envisage le niveau inférieur à *ap* comme le plus petit domaine des règles accentuelles en français : le *pw*, que je prends ici au sens de Selkirk (1996), Nespors & Vogel (1986) et Hayes (1989) *i.e.* le Mot Phonologique. Je pose ce niveau comme potentiellement définitoire de l'accentuation en français sur la base du présupposé selon lequel la fonction culminative de l'accentuation est plus importante que la fonction distinctive dans le lien qui unit la prééminence métrique à l'unité grammaticale du mot lexical. De fait, la fonction démarcative de l'accentuation du français peut émerger en surface dès le niveau du mot phonologique.

Mon approche est enfin fondée sur l'investigation multi-dimensionnelle et interdisciplinaire des faits prosodiques. Multi-dimensionnelle car je propose de mener de front la transcription experte, l'analyse acoustique et l'analyse de la perception d'informants naïfs pour rendre compte des phénomènes prosodiques. Généralement, les descriptions prosodiques sont davantage fondées sur la transcription de phrases *ad hoc*, passant par la seule évaluation perceptive du transcripneur. Le biais expérimental potentiel est donc double : le transcripneur risque d'opérer un lien trop direct avec l'acoustique et d'imprimer son prisme théorique notamment eu égard au *domaine* d'actualisation de l'accentuation (voir les exemples donnés en § 5.4, p. 44). Mon approche est aussi multi-corpus car mes travaux se fondent sur de gros corpus de parole contrôlée, semi-spontanée et spontanée (Corpus D'Edinburgh, MapTask-Aix et CID ; voir § 12, p. 87 et § 13, p. 95). Interdisciplinaire enfin car j'emprunte également aux paradigmes expérimentaux de la psychologie expérimentale (mesures comportementales de la perception) et des neurosciences cognitives (imagerie cérébrale) pour éclairer les résultats de production.

Les résultats de mes recherches sont présentés dans la section suivante, et sont suivis par une conclusion (*Strette*) où j'expose mes propositions phonologiques et les implications qu'elles ont sur les modèles psycholinguistiques de segmentation de la parole et d'accès au sens.

12. LE ROLE DE IA DANS LA STRUCTURATION PROSODIQUE DU FRANÇAIS – ETUDES DE PRODUCTION ET DE PERCEPTION³⁵

Les résultats d'Astésano (2001) suggéraient un rôle plus important de IA dans la structuration prosodique en français que ce que les descriptions habituelles de cette langue laissent supposer. Ils permettaient également de clairement distinguer entre les fonctions structurelle (et rythmique) de IA et sémantico-pragmatique (focus) auxquelles la plupart des approches sur le français renvoient indifféremment (voir § 5.5, p. 50). Cela m'avait amenée à proposer le Mot Prosodique Complexe comme unité de base dans la métrique du français, faisant écho à l'unité prosodiquement cohésive de l'*ap* définie par IA et FA dans les premiers travaux sur le français dans le cadre de l'AM (Jun & Fougeron, 1995). A la différence de Jun & Fougeron, cependant, j'accordais un rôle métrique à IA qu'elles n'envisageaient pas dans leur modèle. Cette conception de Mot Prosodique Complexe (voir § 11.3, p. 84) constituait les prémisses d'un rôle bien plus important de IA dans la structuration prosodique, et de mes propositions métriques actuelles.

12.1. Représentation phonologique de IA : une étude exploratoire de perception (Jankowski *et al*, 1999)

Une étude contemporaine de la fin de la thèse avait été entreprise pour investiguer pour la première fois la réalité *perceptive* de IA comme marqueur de constituant (Jankowski *et al*, 1999). Sur la base des résultats de durées différentielles de IA (voir § 11.2, p. 79), nous avons créé des stimuli lexicaux variant en nombre de syllabes et en types de contexte rythmique, et manipulé alternativement la durée de l'attaque consonantique de chaque début de syllabe des mots lexicaux selon une incrémentation croissante de durée de l'attaque. Les participants devaient dire si la syllabe-cible présentée à l'écran était proéminente ou non (Paradigme Oui/Non). Les résultats indiquaient que les manipulations de durée de l'attaque des syllabes *initiales* de mot lexical induisent la perception d'une proéminence environ 23% plus souvent que dans les autres positions (médiane et finale), pour lesquelles les durées d'attaque devaient de surcroît être plus fortes pour impliquer la détection d'une proéminence. En réalité, même une incrémentation minimale de durée de l'attaque de la syllabe initiale induisait la perception d'une proéminence, nous encourageant à déduire que les auditeurs francophones ont une

³⁵ Voir Annexes : documents scientifiques 'Prosodie 3'.

attente phonologique forte de IA en début de mot lexical. Cet effet était légèrement sensible au contexte rythmique, *i.e.* à la taille du constituant (voir Figure 24).

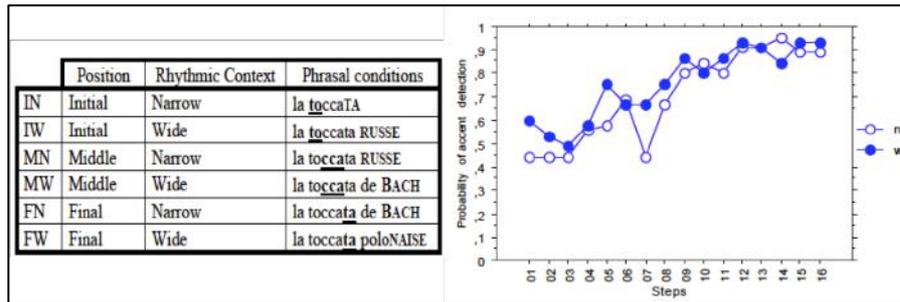


Figure 24 – A gauche, les stimuli pour lesquels la syllabe manipulée par l’allongement de l’attaque est soit en position initiale, médiane ou finale du mot-cible ‘Toccata’, en contexte rythmique étroit (*Narrow*) ou large (*Wide*). Les syllabes finales déterminent l’éloignement de la syllabe manipulée avec l’accent final, en majuscule. A droite, les résultats de détection d’une syllabe proéminente *initiale* uniquement. Le contexte rythmique a une légère incidence sur le taux de détection (5% plus élevé en moyenne en contexte rythmique large). Cependant, les proéminences initiales sont détectées autour de 50% quand bien même aucune incrémentation d’allongement de l’attaque n’est opérée, indiquant une attente phonologique de IA. (Adapté de Jankowski *et al*, 1999).

Il nous paraissait donc intéressant de poursuivre cette étude, en nous penchant de plus près sur les modèles psycholinguistiques de segmentation et d’accès au sens en français, notamment concernant les propositions de Content *et al*, 2001 et Dumay *et al*, 2002 (voir 8.1, p. 63 pour une discussion).

C’est sur cette base que nous avons monté, Ellen Gurman Bard de l’Université d’Edinburgh et moi-même, un projet de post-doctorat financé par une bourse européenne Marie Curie de 2000 à 2002, et co-encadré par Alice Turk. Nous avons rapidement estimé qu’il fallait poser une série de questions préliminaires avant de se pencher sur le rôle de IA dans l’accès au sens.

12.2. IA marqueur de structure (Astésano *et al*, 2007 ; 2003 ; 2002)

En effet, trop peu d’études existaient jusqu’alors sur IA et aucune ne se fondait sur des corpus conséquents, constitués pour tester systématiquement cet aspect structurel. Elles étaient essentiellement descriptives et empiriques (par exemple, les travaux de Fónagy) ou n’étaient pas spécifiquement dédiées à tester en particulier le rôle de IA dans la structuration prosodique, mais le plus souvent dans sa fonction expressive (accent emphatique). Les questions qu’on était donc en droit de se poser étaient les suivantes :

- Quelles règles structurelles président à l’occurrence de IA ? L’occurrence de IA est-elle davantage imposée par la structure syntaxique ou par des contraintes prosodiques (taille des constituants) ?

- Quel niveau de constituance est principalement marqué par IA ? Est-ce effectivement *ap* ou bien également les niveaux supérieurs comme le syntagme phonologique intermédiaire (*ip*) et le syntagme intonatif (*IP*) ?

De là est née la formidable aventure de la construction du corpus dit « d'Edinburgh³⁶ » qui a depuis été à la base d'un très grand nombre d'autres études de production et de perception³⁷.

Le corpus est constitué de phrases syntaxiquement ambiguës que les indices prosodiques tels que les frontières et les proéminences IA et FA doivent permettre de désambigüiser. L'ambigüité syntaxique est créée par la manipulation de la portée de l'adjectif, *i.e.* son attachement haut ou bas dans la structure syntaxique. Dans la phrase *les gants et les bas lisses* par exemple, l'adjectif A qualifie soit

- 1) le deuxième nom *bas* (N2) seul (attachement syntaxique bas) donnant lieu à la structuration suivante : [*les gants*][*et les bas lisses*];
- 2) ou les deux noms *gants* et *bas* (N1 et N2) (attachement syntaxique haut) donnant lieu à la structuration suivante : [*les gants et les bas*][*lisses*].

La structure prosodique est également manipulée en termes de taille des constituants, où N1, N2 et A varient de 1 à 4 syllabes, dans toutes les combinaisons possibles (ex : *les gants et les bas lisses* vs. *les bonimenteurs et les baratineurs fabulateurs*). N1 et N2 covariant en taille cependant. Le corpus est contrôlé phonétiquement pour éviter la variabilité segmentale lorsque la taille des constituants est incrémentée entre N2 et A (ex : *bas lisses* → *balises vertes* → *balivernes sottes...*) et ce, de sorte à ce que les variations micro-prosodiques ne viennent pas perturber la réalisation des FA et IA à cette frontière cruciale dans nos tests. C'est en effet là que nous manipulons les principales forces de frontières, dont nous supposons qu'elles donneront lieu à la production de FA et IA (voir Figure 25, illustrant également le corpus dans les expériences de perception suivantes).

³⁶ Ce corpus est mis à la disposition de la communauté scientifique sur la plateforme 'Ortholang' ; Astésano, C.; Bard, E.; Turk, A. (2011). *Corpus Edinburgh - Initial Accent in French*. oai:sldr.org:sldr000779. <http://sldr.org/sldr000779/fr>.

³⁷ La constitution de cet énorme corpus a occupé une bonne partie de mon temps de post-doctorat, mais je ne serai jamais assez reconnaissante envers Ellen et Alice de m'avoir largement encouragée dans ce processus fastidieux. Ce corpus est une vraie mine d'or pour étudier les règles d'accentuation et de phrasé en français !

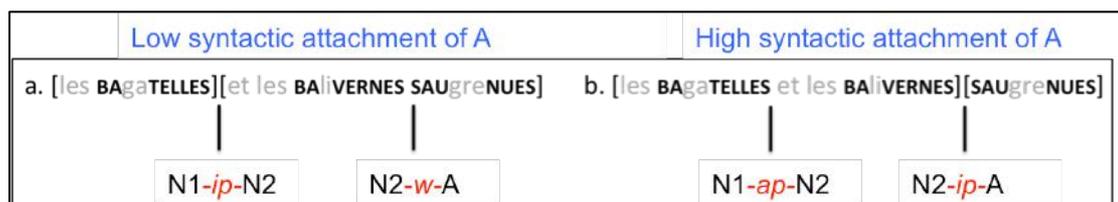


Figure 25 – Les sites prosodiques d'intérêt manipulant la force des frontières en fonction de la structure syntaxique (a. attachement syntaxique bas de l'adjectif : *exemple sonore n°3*; b. attachement syntaxique haut de l'adjectif : *exemple sonore n°4*). Les syllabes en majuscule et gras sont celles sur lesquelles IA et FA peuvent potentiellement se réaliser comme marqueurs de structure. Les structures syntaxiques a et b présentent les frontières suivantes : a) syntagme intonatif entre N1&N2 (N1-*ip*-N2) et mot (prosodique) entre N2&A (N2-*w*-A) ; b) syntagme accentuel entre N1&N2 (N1-*ap*-N2) et syntagme intonatif entre N2&A (N2-*ip*-1). Dans la première étude de production (Astésano *et al.*, 2007), seules les frontières entre N2 et A sont analysées car elles sont phonétiquement contrôlées, et la taille des constituants N2 et A varient de 1 à 4 syllabes dans toutes les combinaisons.

Le corpus est constitué de 4 ensembles ou 'sets' de phrases différents. Chaque syntagme cible N1+N2+A est extrait d'une phrase porteuse dont la structure syntaxique force la réalisation d'une frontière prosodique majeure après l'adjectif A, réalisé comme un ton de frontière non terminale d'IP (H%). 8 locuteurs ont prononcé chacun des 64 scripts originels, répétés 3 fois dans chaque condition syntaxique ($n = 2$), donnant lieu à la production de 3072 phrases. 2 locuteurs ont été écartés des analyses après un jugement sémantique de l'ensemble du corpus sur 10 auditeurs francophones, jugeant de l'adéquation du sens par rapport à la structure syntaxique attendue. Les analyses portent donc sur 2304 phrases (voir Astésano *et al.*, 2007, pour plus de détails sur la constitution et le contrôle du corpus).

Les résultats de cette première étude, centrée sur la production, indiquent que les locuteurs utilisent IA comme marqueur de structure, et ce de manière plus systématique que FA puisque les patrons accentuels IA seuls sont plus fréquents comme marqueur de frontière que les patrons FA+IA³⁸. Une plus grande proportion de IA est trouvée aux frontières syntaxiques qu'à l'intérieur des constituants (Figure 26).

³⁸ Notons cependant que ce résultat émerge d'une analyse tonale uniquement, où FA était recherché automatiquement sur le corpus par la présence d'un ton H. Nous reviendrons sur cet aspect en § 13.2, p. 93.

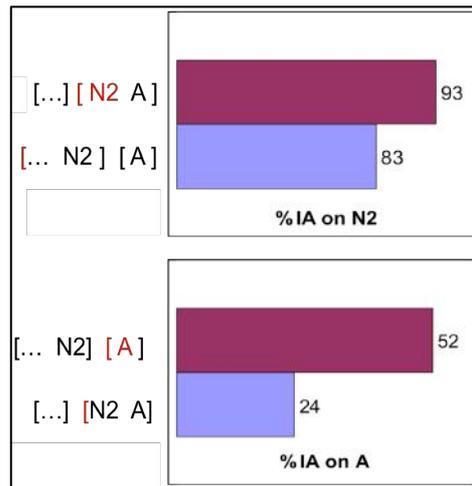


Figure 26 – Proportion d’IA (en pourcentage) aux frontières syntaxiques (indiquées par la police rouge à gauche de la figure). Le pourcentage d’IA est significativement supérieur sur N2 (graphe du haut) lorsqu’il est situé à la frontière syntaxique (frontière d’*ip*) que lorsqu’il est interne au constituant (frontière d’*ap*). Le pourcentage d’IA est significativement supérieur sur A (graphe du bas) lorsqu’il est situé à la frontière syntaxique (ici, frontière d’*ip*) que lorsqu’il est interne au constituant (frontière de niveau du mot *w*). Notons que la proportion d’IA est plus élevée sur N2 que sur A, à frontière syntaxique égale (ici, frontière d’*ip* dans les deux cas).

L’effet de la structure syntaxique sur la proportion d’IA est également modulé par l’effet de taille des constituants prosodiques. En effet, la probabilité d’occurrence d’IA augmente avec la taille des constituants prosodiques, les constituants les plus longs ayant une plus grande proportion d’IA. Cependant, l’influence de la structure prime sur la taille des constituants : IA est donc davantage lié à la démarcation de structure qu’à un effet purement rythmique. Ainsi, lorsqu’on trouve IA sur A à l’intérieur du constituant syntaxique (cas b de rattachement syntaxique bas de A), ce n’est pas tant un effet rythmique dû aux variations de la taille des constituants, mais plutôt un effet de structuration interne puisqu’il marque également dans une proportion non négligeable des A courts. Grâce à l’analyse indépendante des effets d’allongement sur le N2 précédent, nous avons pu déduire que, dans ces cas là, IA sur A était corollaire de l’introduction d’une frontière d’*ap* entre N2 et A, non prédite par la syntaxe. Enfin, nos résultats nous permettent de conclure que, dans ce dessin expérimental précis, IA est plus souvent marqueur du niveau d’*ap* que du niveau du mot phonologique (lexical), même si IA est présent de manière non négligeable à ce niveau le plus bas de constituance.

Cette étude était réalisée sur la base de la détection des pics de *f0* trouvés à l’initiale des mots lexicaux. Sur 2304 phrases (* 3 mots lexicaux par phrase), il s’agissait de procéder au repérage automatique d’IA : nous considérons la présence d’IA dès lors qu’était repéré un changement hauteur de *f0* de plus de 10% sur le début du mot lexical. Il s’agit du seuil minimal de variation de hauteur à partir duquel on peut considérer que se trouve une proéminence de type ‘*pitch accent*’ (Ladd, 2008). En aucun cas n’avons nous procédé à

l'époque à l'annotation manuelle des patrons tonals des phrases cibles, bien trop nombreuses. Nous avons entrepris cela dans une étude exploratoire ultérieure sur une toute petite partie du corpus.

12.3. Réalisations d'IA en plateau (Welby *et al*, 2016)

Il était intéressant de se poser la question de la *variabilité* de production d'IA. Fónagy parle d'une tendance se développant dans des styles de parole particuliers, comme le style journalistique ou télévisuel, qui se répandrait en français contemporain dans la parole de tous les jours. Certes, nous avons remarqué une tendance plus importante au style journalistique à accentuer à l'initiale (Astésano, 2001), mais nous avons montré que, dans sa version non emphatique, IA se retrouve dans tous les styles de parole. L'étude décrite dans le paragraphe précédent (Astésano *et al*, 2007) a ensuite constitué une grande avancée puisque nous avons montré de manière systématique que IA non emphatique n'est pas juste un événement mélodique rythmique, mais bien un phénomène motivé par la structure. Petit à petit, la motivation phonologique d'IA se profilait. Pour autant, seule une investigation de la systématisme de son utilisation à travers les styles de parole et les locuteurs garantirait sa validité phonologique.

Le corpus d'Edinburgh a donc été utilisé dans une étude exploratoire pour tester cette variabilité (Welby *et al*, 2016). Une annotation tonale d'une petite partie du corpus (32 phrases) a été réalisée par 4 annotatrices sur 6 locuteurs du corpus. Il s'agissait déjà de s'entendre sur les primitives utilisées, dépendant en grande partie de l'entente sur un niveau d'analyse (*ap* ou mot lexical ? Voir § 5.4, p. 44 pour une discussion).

L'étude s'est concentrée sur la première *ap* des syntagmes d'Edinburgh, *i.e.* sur N1 dans le cas d'attachements bas de A et sur N1+N2 dans les cas d'attachements hauts de A. Contrairement à ce que décrit la littérature AM sur la réalisation tonale de l'*ap* ($L_1H_1-L_2H_2$ pour les *ap* plus longues; L_1H_2 ou $L_1H_1-H_2$ pour les *ap* plus courtes ; voir Welby, 2006 pour une présentation détaillée), nos résultats d'annotation montrent que les formes tonales de IA et FA (également appelés dans AM '*early rise*' et '*late rise*') sont systématiquement différentes. La variabilité de FA est quasi nulle à travers les locuteurs, qui le réalisent à 93% dans sa forme convexe terminée par un seul pic. En revanche, les formes tonales de IA varient davantage : de forme tonale plutôt concave, son pic était à 49% des cas réalisé en 'plateau' ou 'bosse' plutôt qu'en simple pic. Cette étude exploratoire montre par exemple qu'aucun des 6 locuteurs ne montre une tendance à réaliser plus de configurations à 2 pics à mesure que la taille de *ap* augmente. Nous nous attendions aussi à ce que la forme 'plateau' de IA (L_1H_{1p} -

L₂H₂) soit particulièrement présente dans les longs constituants mais cela ne se vérifie que dans 50% des cas. Si deux locuteurs montrent cette tendance, deux autres n'utilisent en revanche pas la forme de plateau initial quelle que soit la taille de l'*ap*, et les deux derniers locuteurs utilisent alternativement le 'plateau' ou la forme plus canonique en pic de IA (Figure 27).

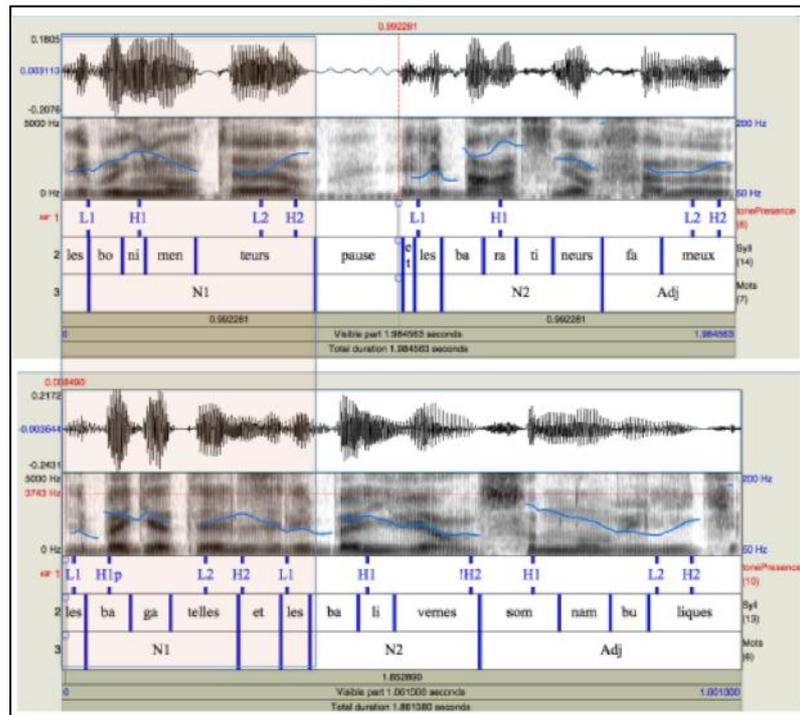


Figure 27 – En haut, réalisation canonique d'une *ap* longue (sur N1, surligné), où IA se réalise avec un pic (L₁H₁-L₂H₂). En bas, réalisation en plateau de IA sur N1 (L₁H₁p-L₂H₂). La figure du haut représente le cas d'attachement syntaxique bas de A à N2 ; la figure du bas représente le cas d'attachement syntaxique haut de A à N1+N2. (Adapté de Welby *et al*, 2016).

La différence de forme a déjà été montrée dans la littérature (Welby, 2003) et il est admis que ces variations reflètent un même événement. Cependant, les facteurs conditionnant la distribution en simple pics ou en plateaux hauts ne sont pas encore connus. Le corpus d'Edinburgh se prête donc à une investigation systématique de ces facteurs, qui, de par la nature contrôlée du corpus, ne concernent que les aspects structurels et prosodiques.

Ces résultats nous invitent à poursuivre l'investigation des facteurs responsables de ces variations de formes, notamment sur la parole spontanée. A titre d'exemple, la Figure 28 montre la réalisation de plateaux tonals en interaction conversationnelle (exemples issus du Corpus CID, voir sa description §13.7, p. 112), encore différents de ceux observés dans l'étude de Welby *et al* (2016), et dans un contexte « d'arc accentuel » entre un nom et un adjectif.

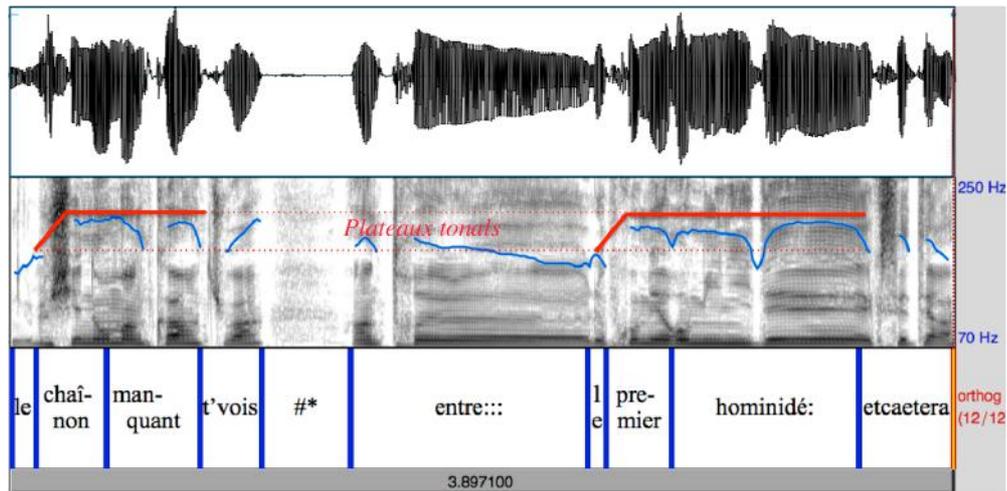


Figure 28– Réalisation de plateaux tonals (symbolisés par les droites rouges) au niveau des *ip* ‘chaînon manquant’ et ‘premier hominidé’ consistant en une forme particulière du marquage bipolaire IA-FA sur chaque mot lexical (*pw*), et extrait de la phrase « *ouais ils recherchaient des ... le chaînon manquant tu vois entre le premier hominidé etcætera* » (*exemple sonore n°5*). Ces plateaux sont une alternative à la réorganisation sur le plan tonal en arcs accentuels avec désaccentuation des FA internes de groupes. Chaque plateau est délimité par un saut tonal entre le déterminant et la première syllabe du groupe de mots s’étendant *recto-tono* jusqu’à la dernière syllabe du groupe de mots (*exemple sonore n°6*). Des traces d’allongement infra-syllabique sont observés sur les mots lexicaux internes de groupes (Adapté de Astésano & Bertrand, 2016)

Notons qu’il s’agit d’une réalisation peu canonique de l’arc accentuel, non concave, et où subsistent des traces métriques des FA et IA internes à l’arc accentuel (Figure 29). IA sur la première syllabe /*ʃɛ*/ du nom ‘*chaînon*’ présente un allongement typique de l’Attaque, alors que FA sur la dernière syllabe /*kã*/ de l’Adjectif ‘*manquant*’ présente aussi un allongement typique de sa Rime /*ã*/. Cependant, on note des traces d’allongement de la Rime de la deuxième syllabe /*nõ*/ du nom ‘*chaînon*’. Le plateau tonal, qui soutient un ton H s’étalant sur l’ensemble de l’*ap*, permet en fait de maintenir tonalement la présence de IA au début de l’adjectif ‘*manquant*’. De fait, les indices métriques de durée infrasyllabique et de tons permettent de manière complémentaire de garder traces de la démarcation du niveau du *mot lexical*, à l’intérieur de l’arc accentuel. Nous reviendrons sur cet aspect dans la section § 16, p. 140 concernant notre proposition du pied spondéique au niveau du *pw*.

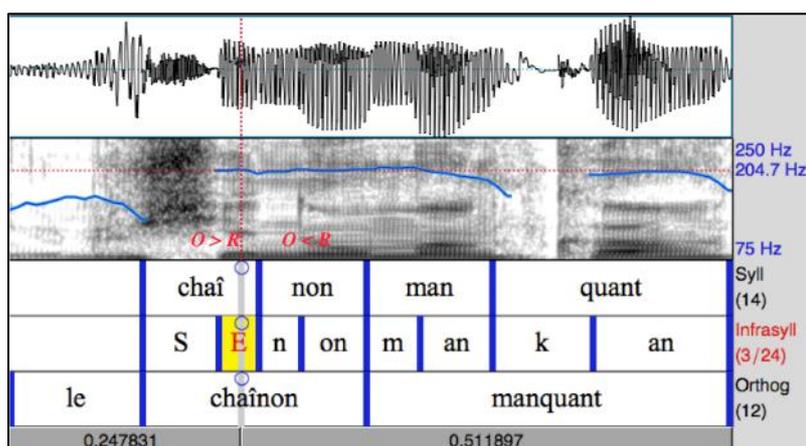


Figure 29 – Exemple d’allongement infra-syllabique sur le premier mot du groupe de mots ‘*chaînon manquant*’ où l’attaque syllabique (O) /f/ est plus longue que la rime (R) /ɛ/ sur la première syllabe ‘chaî-’, et où la rime (R) /õ/ est plus longue que l’attaque (O) /n/ sur la deuxième syllabe du mot ‘-non’. L’information de durée segmentale indique donc les limites du mot lexical à l’intérieur du groupe de mots marqué par le plateau tonal.

Des études ultérieures sur des corpus de parole spontanée nous permettront de repérer des constructions syntaxiques similaires à celle du corpus d’Edinburgh pour appréhender les facteurs responsables d’une variabilité des réalisations tonales de IA sur *ap* en parole spontanée similaires à celles observées dans un corpus contrôlé (voir sections § 13.6, p. 110 et § 13.7, p. 112).

13. PERCEPTION DES FRONTIÈRES, DES PROEMINENCES ET DU PHRASE PROSODIQUE³⁹

Suite à ces travaux, nous voulions en apprendre plus sur l’interaction entre IA et FA dans la démarcation des constituants en français, et leurs liens avec les frontières prosodiques. Nous pensons avoir montré avec les études décrites ci-dessus que IA doit avoir un statut plus important dans les modèles pour le français, ancré dans la métrique par sa fonction de structuration au delà de sa fonction d’équilibrage purement rythmique qu’on lui concède trop souvent dans les approches classiques. Concernant FA, rappelons que les modèles sur le français, qu’ils soient basés ‘syntaxe’ (modèles intonosyntaxiques) ou issus de la mouvance Métrique-Autosegmentale (AM), postulent que FA disparaît au profit des frontières intonatives au delà du groupe accentuel ou *ap* (voir § 5.5, p. 50). Les questions que l’on s’est posées dans le cadre de l’ANR *PhonlaCog* avec Roxane Bertrand, et qui ont donné lieu également à des travaux de thèses et de post-doctorats, étaient les suivantes :

- FA survit-il métriquement au delà de *ap* ?
- Quels sont les liens entre IA et FA dans le marquage des constituants ? Autrement dit, IA (secondaire) cède-t-il sa place à FA en fonction des contraintes prosodiques ?

³⁹ Voir Annexes : documents scientifiques ‘Prosodie 3’.

L'étude d'Astésano *et al* (2007) centrée essentiellement sur IA ne permettait en effet de répondre à cette question que partiellement.

- Peut-on trouver des niveaux de prédilection du marquage par IA et/ou FA ? Autrement dit, quels liens entretiennent-ils avec le phrasé prosodique ?
- Existe-il une variabilité de réalisations tonales de FA, comme celle que l'on peut trouver pour IA ?
- Les patrons accentuels décrits sur de la parole de laboratoire (ici, le Corpus d'Edinburgh) se retrouvent-ils en parole spontanée ?
- Quid de leurs caractéristiques acoustiques ? Retrouve-t-on dans différents types de gros corpus systématiquement les mêmes tendances que celles trouvées dans Astésano (2001) sur l'allongement infra-syllabique notamment ?

Pour répondre à la plupart de ces questions, il nous semblait pertinent de passer par le prisme de la perception des paramètres prosodiques (voir § 10, p. 76). Cela nous permet au passage de tester la supposée 'surdité' des francophones *à leur propre système phonologique*. La perception constitue une interface entre représentation phonologique et réalisation acoustique (Di Cristo, 2004 ; Cole *et al*, 2010). Elle permet de surcroît de contourner le problème de la variabilité ou '*flexibilité*' des indices acoustiques instanciant les constituants de même niveau dans la hiérarchie prosodique, et très souvent observée en français (Rossi, 1997, 1999 ; Delais-Roussarie & Felhausen, 2014).

13.1. Perception de IA et FA et leurs liens avec les frontières prosodiques (Astésano *et al*, 2012 ; 2013 ; en cours)

La première étude exploratoire a été menée sur 32 phrases du Corpus d'Edinburgh lues par la locutrice CP, qui avait obtenu les meilleurs scores de reconnaissance du sens des phrases (portée syntaxique de A) lors du pré-test sur 20 auditeurs mené préalablement à l'étude d'Astésano *et al* (2007).

Il s'agissait d'une tâche de perception 'offline' où 18 sujets devaient effectuer 2 tâches différentes (l'ordre de présentation des tâches était contrôlé à travers les sujets).

- *Tâche de jugement de la force de proéminence :*

Les sujets devaient juger de la proéminence relative de chaque syllabe sur une échelle de 0 à 3 ('0' étant la ligne de base 'non proéminent' et 3 la syllabe la plus proéminente).

- *Tâche de jugement de la force des frontières :*

Les sujets devaient juger de la force relative de frontière entre chaque mot, sur la même échelle de 0 à 3, '0' représentant aucune frontière et '3' la frontière la plus forte.

Ils pouvaient écouter chaque phrase jusqu'à 5 fois et chaque tâche était précédée d'une phase d'entraînement. La Figure 30 illustre la feuille de réponse pour chaque tâche.

Figure 30 – A gauche, la tâche de jugement de la force de frontières entre chaque mot. A droite, la tâche de jugement de prééminence relative sous chaque syllabe. Les jugements se font sur une échelle de 0 (pas de prééminence ou pas de frontière) à 3 (syllabe la plus prééminente ou frontière la plus forte).

Les sites d'intérêts pour nos analyses statistiques sont représentés dans la Figure 25 p. 90.

Les analyses sur chaque site et par type de tâche ont été réalisées avec un modèle logit mixte ordinal⁴⁰.

Avant de voir le lien entre les frontières et les prééminences, il fallait déjà voir dans quelle mesure les auditeurs sont capables de percevoir les prééminences et dans quelle mesure les forces de frontière perçues reflètent les frontières prédites par la manipulation syntaxique.

Afin de renforcer la puissance statistique, les résultats sont donnés pour une taille d'adjectif fusionnée (de 1 à 4 syllabes). Nous avons testé la différence entre les scores successifs de chaque syllabe. Seules les syllabes initiales et finales de mots ou de groupes ont été perçues comme significativement plus prééminentes ; les syllabes internes de mots et les mots outils 'et' et 'les' n'ont pas été perçues comme prééminentes. La Figure 31 présente la perception des syllabes pour les phrases de 3 et 4 syllabes uniquement, les plus représentatives des résultats.

⁴⁰ Plus précisément appelé : 'modèle de régression logistique polytomique ordinaire à effets mixtes'. Notez que les résultats donnés pour cette première étude sont issus d'Astésano *et al* en cours de réécriture, à soumettre dès que possible à *Language and Speech*. Ils utilisent des modèles statistiques (et des étapes de constructions de cette démonstration complexe) différents de ceux présentés dans Astésano *et al* (2012 ; 2013). Les résultats sont donc légèrement différents, mais l'interprétation de dissociation entre FA et les frontières restent identiques, ainsi que ceux concernant la perception plus forte de IA par rapport à FA. Nous remercions ici Alice Turk, Ellen Gurman Bard et Lorraine Baqué pour leurs retours très fructueux sur une première version de cet article à soumettre.

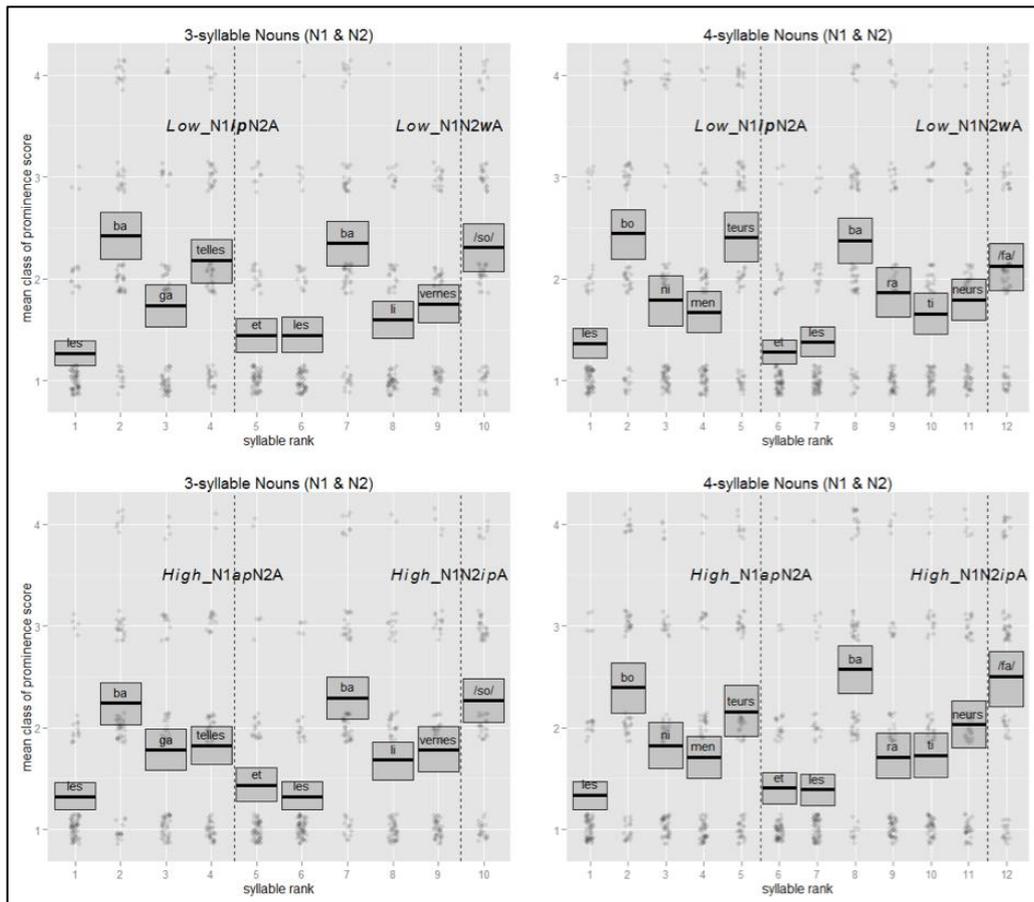


Figure 31 – Scores de prééminences moyen sur chaque syllabe en fonction du rang d’occurrence de la syllabe dans la phrase, dans la condition d’attachement bas de l’adjectif (en haut) et la condition d’attachement haut de l’adjectif (en bas) pour les noms de 3 syllabes (à gauche) et de 4 syllabes (à droite). Notez que la taille de l’adjectif est fusionnée pour augmenter la puissance statistique : la première syllabe (phonétiquement contrôlée) est donnée en phonétique. La hauteur des boîtes indique l’intervalle de confiance du score moyen. Les points indiquent les réponses individuelles avec jitter. Des lignes verticales en pointillé indiquent les sites de frontières. Notons cependant que l’échelle perceptive de 0 à 3 est ici (et dans la figure suivante) rapportée à une échelle de 1 à 4. (Adapté de Astésano *et al*, en cours).

Notons toutefois que les syllabes initiales IA sont systématiquement perçues comme prééminentes dans leur contexte alors qu’il n’en est pas de même pour les syllabes finales (FA) ; certaines sont en effet perçues comme moins prééminentes que les syllabes précédentes (par exemple FA sur N1 et N2 de trois syllabes, ou encore FA sur N2 de trois et quatre syllabes). Cela est sans doute imputable à la réalisation acoustique de FA qui varie parfois (réalisation tonale L ou !H) sur laquelle nous reviendrons ci-dessous. Ce comportement particulier de FA (variabilité de réalisation acoustique tantôt H tantôt L ou !H), à l’inverse du comportement très régulier de IA, est en partie ce qui questionne le lien entre prééminences et frontières en français, ainsi que le syncrétisme accentuation/intonation.

Pour rendre compte du lien entre prééminences et frontières, nous devons procéder à une démonstration en 3 étapes : 1) les auditeurs sont-ils capables de percevoir les différents types de frontières ? C’est l’étape de ‘Séparabilité’ entre *w*, *ap* et *ip* telles que prédites par la manipulation syntaxique du corpus ; 2) Les auditeurs sont-ils capables de distinguer entre les

phénomènes de frontières et les phénomènes de proéminences ? C'est l'étape de 'Dissociation' ; 3) Alors seulement peut-on commenter le rôle respectif de IA et FA dans la démarcation des frontières, en mesurant les 'Pentes' *i.e.* la modulation des scores de frontières par les scores des proéminences IA et FA.

1). Séparabilité des plans : nos résultats indiquent que les frontières les plus fortes (N1-*ip*-N2 en condition de 'Low attachment' syntaxique de A ; N2-*ip*-A en condition de 'High attachment' syntaxique de A) sont en effet perçues comme plus fortes que les frontières prédites plus faibles (frontières de *w* et d'*ap*). Cependant, les frontières *ip* en condition de 'low attachment' entre N1 et N2 sont perçues plus fortes que les mêmes frontières prédites d'*ip* en condition 'high attachment' entre N2 et A. Par ailleurs, les frontières faibles ne sont pas perçues dans la hiérarchie escomptée puisque *w* est perçue plus forte que *ap*, même si cette tendance ne permet pas de les séparer clairement. Les frontières sont donc séparées en 3 plans : N1-*ip*-N2 > N2-*ip*-A > N2-*w*-A ≥ N1-*ap*-N2.

2). Dissociation proéminences/frontières : Globalement, il n'y a pas de correspondance entre la force perçue de IA et la force perçue des différents niveaux de frontières. De fait, on peut en déduire que IA est dissocié des frontières qu'il marque. On voit d'ailleurs qu'il est toujours perçu avec une force comparable dans la Figure 31. La tendance est la même pour FA où la corrélation entre la force de perception de la proéminence et la force de frontière perçue n'est pas démontrée, sauf aux frontières N1-*ip*-N2 dans la condition 'Low attachment'. Dans ce cas là, la force perçue de FA est la plus haute et corrélée avec la plus forte perception de frontière. Notons qu'une analyse auditive et visuelle suite à l'annotation tonale dans Welby *et al*, 2016 sur le même corpus, indique que cette frontière est réalisée par cette locutrice CP avec une longue pause entre N1 & N2, renforçant sans doute la perception d'une plus forte frontière et permettant d'autre part une réalisation tonale 'canonique' de FA avec un ton H. La dissociation entre frontières et proéminences est donc observée, mais partiellement puisque non avérée dans ce dernier cas.

3) Modulation des forces de frontières par les forces de proéminences perçues : les 4 plans de frontières présentent la même inclinaison de modulation par IA et FA. Les différences de forces de perception de IA et FA n'ont donc pas d'impact sur le changement catégoriel (séparabilité) de forces de frontières.

Ces résultats nous amène globalement à déduire tout d'abord que les auditeurs francophones ne sont pas 'sourds' aux proéminences en français, qu'elles soient initiales (IA) ou finales (FA). IA est perçu systématiquement plus fort que FA, et ce à tous les niveaux de

frontière, y compris le niveau *w*. Il est toujours perçu aussi fort, même à ce niveau minime de la constituance prosodique, indiquant donc que le plus petit niveau de frontière qu'il marque est le *mot lexical* (Figure 32). Cela est en contradiction avec les descriptions habituelles de cet accent, qui le qualifient de secondaire par rapport à FA, et n'intervenant concrètement qu'au niveau de *ap*.

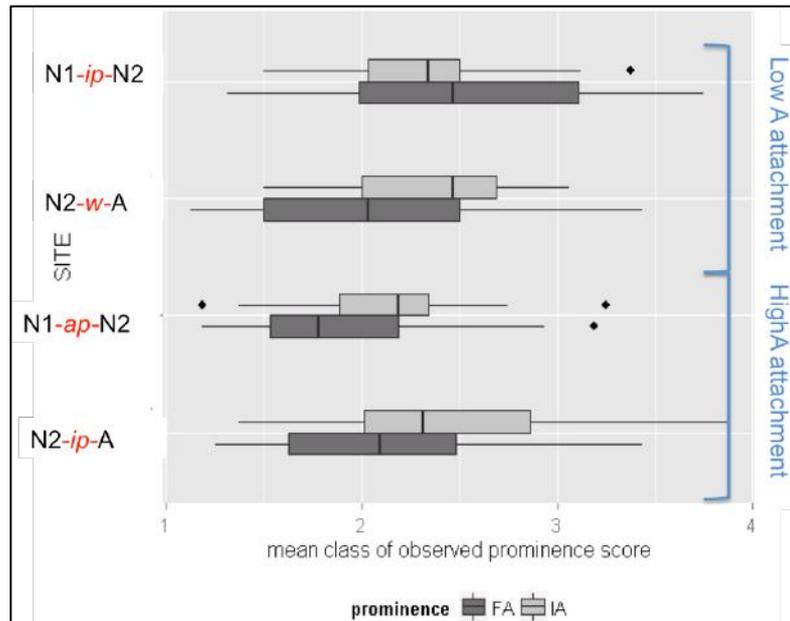


Figure 32 – Scores de force de perception de IA et FA aux 4 sites de frontières prédits. L'intervalle interquartile montre la dispersion ; la barre verticale dans les boîtes indique la moyenne des scores.

Nous pensons également avoir démontré que FA est perceptible comme métriquement proéminent quel que soit le type de frontière auquel il est associé, notamment aux niveaux de frontières supérieurs *ip* où la littérature dit qu'il disparaît au profit de la frontière intonative. Nous pensons au contraire avoir montré qu'il 'survit' à ce niveau de la constituance, et que son poids métrique renforce même la perception de *ip*. Notons enfin que, comme IA, FA est perçu avec une (faible) proéminence au niveau du mot *w*.

D'autres résultats viennent renforcer ces premières propositions sur le rôle de IA et FA.

13.2. Perception de FA produit comme L ou !H à l'intérieur d'un constituant (Mendez & Astésano, 2017 ; Mendez, en cours).

La littérature mentionne sporadiquement une inversion de pente (Delattre, 1966a ; Martin, 1981) à l'intérieur de l'énoncé (généralement au niveau du contour mineur, *i.e.* l'*ap*), indiquant certains types de relations syntaxiques entre les constituants. Ces descriptions intonosyntaxiques ne rendent pas compte cependant du lien entre FA et le contour intonatif, et ne mentionnent pas d'autres niveaux de la hiérarchie prosodique où ces inversions de pentes peuvent être rencontrées.

Le corpus d'Edinburgh offre une variabilité de réalisation des formes tonales de IA et FA (Welby *et al*, 2016). La manipulation de la structure syntaxique, où le rattachement de A se fait avec N2 seul ou N1&N2, induit naturellement des relations syntaxiques variables entre les composants lexicaux à divers niveaux de frontières. Ce corpus se prête donc parfaitement à l'analyse systématique de la variation tonale de FA.

Jusque là, une réalisation 'downstep' !H de FA n'a été proposée que dans certains cas de contours nucléaires à valeur épistémique (Delais-Roussarie *et al*, 2015) et nous ne sommes pas convaincus que sa réalisation dans le corpus contrôlé d'Edinburgh relève d'un phénomène de 'Downstep' (Welby *et al*, 2016). Il pourrait s'agir d'un ton L permettant d'accentuer le rattachement tonal au constituant supérieur IP (Figure 33).

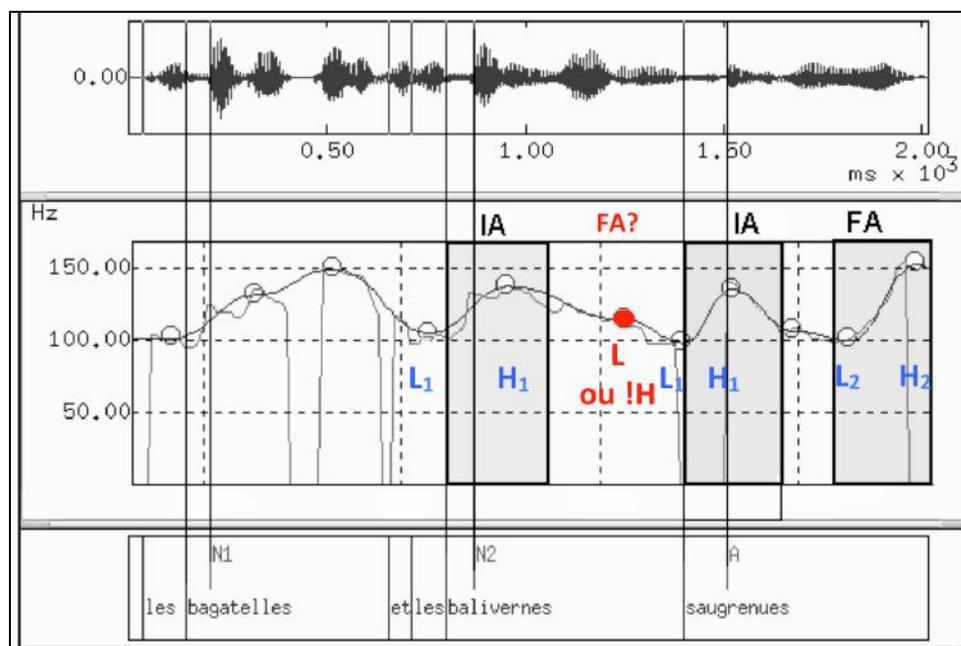


Figure 33 – Réalisation d'un FA en ton L ou !H en frontière d'*ip* entre N2-*ip*-A dans le cas de rattachement syntaxique haut (High attachment) de A. FA n'est pas perçu comme proéminent par rapport à la syllabe précédente /li/, mais significativement plus proéminent que les mots grammaticaux précédents 'et' 'les' (voir Figure 31). La modélisation de la courbe de *f0* en points-cibles est réalisée avec l'algorithme MoMEL (Hirst & Espesser, 1993). (Adapté d'Astésano *et al*, 2007).

Nous avons donc sélectionné 32 phrases (mots tri- et quadrisyllabiques) du Corpus d'Edinburgh prononcées par 4 locuteurs différents et présentant pour moitié des réalisations L ou !H de FA, pour moitié des réalisations plus canoniques de FA en H*. Ces réalisations étaient pour moitié situées au niveau de frontière *ap* et pour moitié au niveau *ip*, afin de pouvoir évaluer la modulation de la force de frontière par les deux réalisations de FA. 20 auditeurs ont effectué deux tâches : une tâche de jugement de frontière et une tâche de jugement de proéminence. La procédure expérimentale est la même que dans l'étude précédente (§ 13.1, Figure 30). Nos analyses statistiques pour la tâche de jugement des

proéminences montrent la hiérarchie de force de proéminences suivante : IA = FA > DFA > UN (où DFA est le Downstep FA et UN les syllabes inaccentuées médianes de mots et mots grammaticaux). DFA est certes perçue moins proéminente que IA et FA ‘canonique’, mais tout de même significativement plus proéminente que les syllabes inaccentuées. Il y a donc bien une force métrique perçue sur un ton L ou !H, et la représentation phonologique de FA demeure malgré cette réalisation tonale descendante. Si l’on croise maintenant la perception de FA et DFA avec la perception de la force des frontières, nos résultats indiquent que les deux types d’accent final sont perçus moins forts aux frontières *ap* qu’aux frontières *ip*. Cependant, si les deux types d’accent final sont perçus quasiment avec la même force en frontière d’*ap* ($p=.0057$), DFA est perçue plus forte que FA avec ton H en frontière d’*ip* ($p<.001$). De fait, la variation phonétique (tonale) module le poids métrique de l’accent final et la relation entre poids métrique et force de frontière : le poids métrique de DFA est renforcé en frontière *ip*.

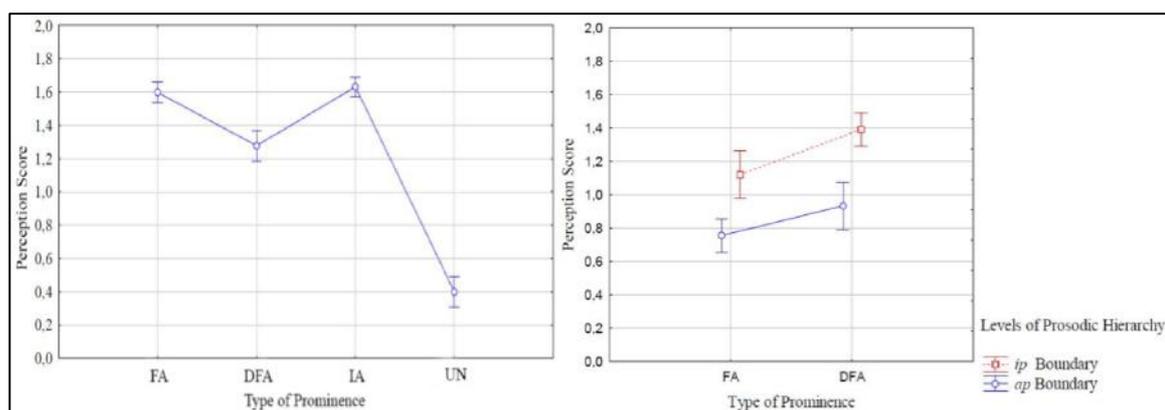


Figure 34 – A gauche, les scores de perception des proéminences FA (ton H), DFA (ton L ou !H), IA (H) et syllabes inaccentuées (UN) sur une échelle de 0 à 3. A droite, scores de perception de FA et DFA en fonction des types de frontières *ap* et *ip*. FA et DFA sont perçus moins forts aux frontières *ap* que *ip*. Le niveau supérieur de frontière *ip* renforce le poids métrique de DFA.

Ces résultats, encore préliminaires et en cours de développement (Mendez, en cours), confirment nos précédents résultats : les francophones ne sont décidément pas sourds aux proéminences, et la perception des proéminences finales est partiellement dissociée des frontières prosodiques puisqu’on observe une modulation distinguant les deux types d’accent final par la frontière, seulement pour les frontières de niveau supérieur, *ip*. Ils pointent en revanche deux nouvelles tendances très importantes pour la suite de notre démonstration phonologique :

- 1) IA et FA (marqué par un ton H) sont perçus *avec la même force métrique*. Nos résultats antérieurs (Astésano *et al*, 2012, 2013, et en cours) pointaient une force

métrique plus importate de IA car nous ne distinguions pas entre les deux réalisations tonales de FA. Or, DFA est perçu globalement moins fort que FA et IA ;

- 2) Non seulement ces résultats montrent que FA ‘survit’ phonologiquement aux frontières de niveaux supérieurs (ici *ip*) mais en plus, il est perçu métriquement plus fort à ce niveau de frontière lorsqu’il est réalisé avec un patron tonal descendant (L ou !H). Nous observons de fait un comportement radicalement inverse de celui proposé dans la littérature (Rossi, 1980 ; Jun & Fougeron, 2000 ; 2002) : plutôt que de disparaître au profit du contour intonatif de frontière, le contour tonal descendant de FA s’additionne avec le contour intonatif d’*ip* pour renforcer la sensation de force de frontière.

Le syncrétisme entre FA et frontière intonative complexifie considérablement l’analyse tonale de ces deux phénomènes : seules des expériences de perception permettent de mettre au jour la réalité métrique de FA en dépit de son assimilation tonale au contour de frontière. A défaut, pour l’instant, d’un formalisme d’annotation AM consensuel permettant cette distinction entre mètre et intonation, passer par la perception semble la meilleure voie pour décrire la phonologie du français.

Notons que nos expériences perceptives vont dans le sens de la proposition d’annotation tonale de Post (2000), où FA conserve sa force métrique aux niveaux supérieurs de constituance (H*-H%).

13.3. Perception de IA et alignement tonal de H₁ ; Perception de DFA en lien avec ses corrélats acoustiques (Astésano & Bertrand, 2016 ; Garnier, en cours ; Mendez, en cours)

Nous rapportons maintenant les tendances préliminaires (en cours de traitement systématique dans Garnier, en cours) sur la mise en relation des proéminences perçues dans le Corpus D’Edinburgh avec leurs corrélats acoustiques de durée et de *f0*. La Figure 35 illustre la démonstration phonologique que nous cherchons à établir statistiquement sur 80 participants auditeurs et 4 locuteurs du corpus. Elle permet de comparer terme à terme, sur un exemple de phrase comportant des N1&N2 de trois syllabes (*‘les bagatelles et les balivernes’*), la force de proéminence perçue de chaque syllabe et leurs corrélats acoustiques.

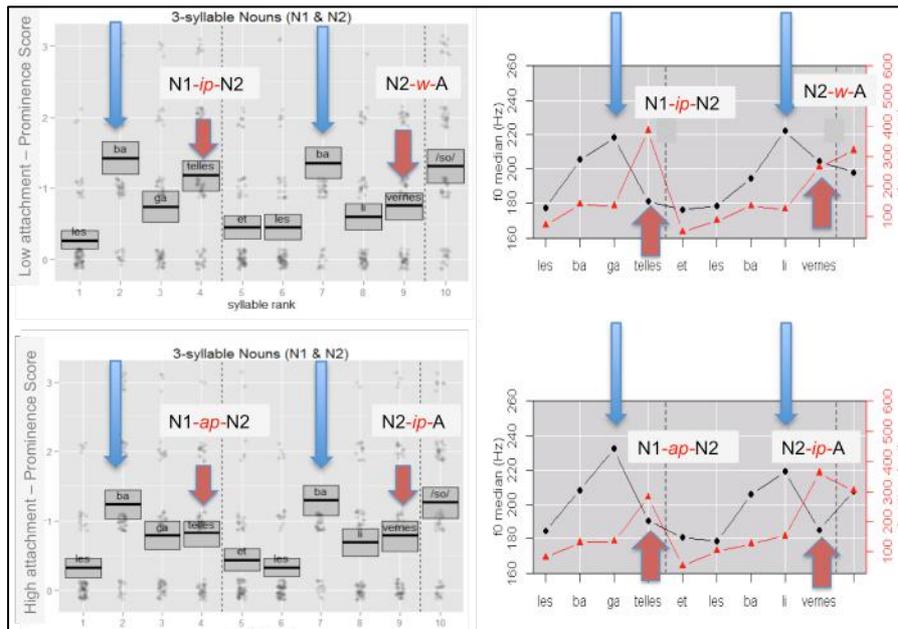


Figure 35 – A gauche, les scores de proéminences perçues de la séquence ‘les baguettes et les balivernes’ dans les conditions de ‘low attachment’ de A (en haut) et de ‘High attachment’ de A (en bas). A droite, les mesures acoustiques brutes de ces mêmes phrases en durées syllabiques (en rouge) et en médianes de f_0 (en noir). 4 sites de frontières sont représentés en fonction de la manipulation syntaxique de rattachement haut ou bas de A (*ip* entre N1 & N2 dans les figures du haut ; *ip* entre N2 et A en bas ; *ap* entre N1 & N2 en bas ; *w* entre N2 et A en haut). Les flèches rouges à droite indiquent les corrélats acoustiques de FA, à comparer avec la perception de ces mêmes FA représentée par les flèches rouges à gauche ; les flèches bleues à droite représentent le pic de f_0 de IA sur la deuxième syllabe, à comparer avec les flèches bleues à gauche indiquant la perception de la proéminence IA sur la première syllabe (voir exemples sonores associés à la Figure 25).

Dans cette configuration prosodique, FA est réalisé sur chaque mot dans sa version L ou !H précédemment décrite, quelle que soit la structure syntaxique (attachement syntaxique bas de A en haut, attachement syntaxique haut de A en bas) ou le niveau de constituance (*w*, *ap*, *ip*). Sur N1, ce qui distingue acoustiquement les deux FA dans les deux configurations syntaxiques est la durée syllabique, plus longue en frontière *ip* que *ap* ; la hauteur tonale est proche de celle des syllabes inaccentuées (L) dans les deux conditions syntaxiques. Sur N2, la durée de FA est plus longue en frontière d’*ip* que de *w*, mais il faut cependant noter une durée non négligeable de FA en frontière de mot, largement plus importante que celle des syllabes inaccentuées. La hauteur tonale est également comparable à celle des syllabes inaccentuées (L) sauf dans le cas de la frontière *w* puisqu’elle se trouve dans le cadre de « l’arc accentuel » produit entre IA sur N2 et FA sur A : la forme ‘canonique’ de l’arc accentuel est celle d’un arc de cercle concave reliant les deux extrémités gauche et droite de deux mots. La perception significativement plus proéminente de FA sur N1 en contexte d’attachement bas de A est donc dû au renforcement métrique de cet accent en frontière d’*ip*. Notons que dans ce cas, la frontière d’*ip* est le plus souvent instanciée également par la présence d’une longue pause, rappelant la conception selon laquelle allongement final et phénomènes de frontières marqués par des pauses relèveraient d’un même phénomène sous-jacent (Duez, 1987).

13.4. Conclusion intermédiaire sur le statut phonologique de IA

La comparaison de la perception de IA et de sa réalisation acoustique est particulièrement intéressante au regard du défi que cela soulève par rapport à la littérature. Rappelons-nous que IA est considéré comme un événement prosodique secondaire, dont le statut métrique est loin d'être reconnu dans la littérature. Notamment, les descriptions AM du français (Jun & Fougeron, 2000 ; 2002 ; Welby, 2003 et suivantes) décrivent un alignement 'lâche' du pic de IA intervenant sur la deuxième ou parfois troisième syllabe d'un mot long de contenu qu'il délimite ('loose boundary marker'⁴¹). Cette irrégularité de l'ancrage du pic sur la première syllabe de IA serait une indication de sa nature rythmique, non métrique. Il est donc tout à fait remarquable de noter que les auditeurs perçoivent systématiquement la première syllabe des mots comme étant, et de loin, la plus proéminente, indépendamment de la réalisation plus tardive du pic de f_0 (systématiquement sur la deuxième syllabe de chaque mot trisyllabique dans la Figure 35).

Une fois encore, la perception vient à l'aide de l'interprétation phonologique des phénomènes accentuels complexes en français. Pour ce qui est particulièrement de IA, deux interprétations sont possibles (et non mutuellement exclusives) : 1) les auditeurs perçoivent systématiquement IA sur la première syllabe, indépendamment de l'alignement du pic, parcequ'ils se fondent sur le 'saut' de f_0 entre le ton bas du déterminant précédent et la première valeur de f_0 sur la première syllabe (largement au dessus des 10% perceptibles pour la détection d'un accent ; Ladd, 2008). En cela, nos résultats font écho aux propositions de Fry (1958) ou 't Hart *et al* (1990) sur la pertinence de prendre en compte la *dynamique* des changements tonals dans l'interprétation phonologique de la mélodie des langues, et non pas seulement les pics ou les creux les plus importants d'une courbe mélodique. C'est en ce sens aussi que nous émettons des réserves sur le parti-pris de AM de se fonder sur une analyse essentiellement visuelle des variations de la courbe mélodique en termes de tons discrets, pour déterminer les règles phonologiques d'une langue. 2) Le deuxième point vaut pour IA aussi bien que pour FA : les auditeurs perçoivent les syllabes initiales et finales du mot *w* comme proéminentes en dehors de toute autre syllabe et malgré les corrélats acoustiques contradictoires ; il est donc possible qu'une reconnaissance 'top-down' des mots lexicaux soit appliquée à défaut de marques acoustiques claires. Dans une interprétation comme dans l'autre, les auditeurs ont donc bien une représentation phonologique du mot lexical en

⁴¹ Welby (2006) parle cependant d'un alignement gauche de IA à la frontière du *mot lexical* par le ton L, beaucoup plus stable.

français... Et pourtant, ce niveau n'est jamais pris en compte, dans aucun modèle, comme domaine de l'accentuation et comme unité de performance pertinent.

13.5. Perception du phrasé prosodique (Garnier *et al.*, 2016 a&b ; Garnier, en cours)

Jusque-là, nos investigations tentaient de démêler les relations entre proéminences et frontières, la perception étant une fenêtre d'observation nous permettant de tester nos hypothèses sur l'indépendance relative de FA par rapport aux frontières prosodiques, et sur la pertinence phonologique de IA et FA dans le système français. Mais les tâches de jugement des proéminences et des frontières ne nous donnent qu'une vision partielle du *phrasé* prosodique, le phrasé étant la manière dont la prosodie aide au morcelage en groupements ('chunks') du flot de parole. Certes, proéminences et frontières ont un lien avec le phrasé, mais l'information de groupement est complémentaire car elle suppose un lien plus étroit avec l'information syntaxique et syntagmatique. D'autre part, percevoir une frontière, une proéminence ou un groupement ne relève pas du même processus cognitif sous-jacent. Les deux premières tâches sont davantage 'musicales', faisant appel à une analyse peu naturelle du signal perçu. La perception du groupement est finalement la résultante des deux premières, et se rapproche davantage de ce que fait un auditeur lorsqu'il décode la parole : peu importe ce qui permet la construction d'un groupement, le principal est de diviser le flot de parole continue en 'chunks'.

Une troisième tâche de perception a donc été imaginée, proposant aux auditeurs plusieurs choix de groupements pour ces mêmes phrases issues du Corpus d'Edinburgh (Figure 36).

Le croisement des informations de perception des frontières d'un côté, des proéminences de l'autre, avec celle du groupement devraient nous éclairer plus précisément sur les niveaux de constituance prosodique nécessaires pour rendre compte du phrasé prosodique en français.

Cependant, une telle mise en relation est extrêmement complexe à modéliser statistiquement. A l'époque, R. Espesser, co-auteur de nos premiers travaux et spécialiste (également !) de statistiques, s'était refusé à l'envisager car aucun modèle n'existait à sa connaissance pour rendre compte des interactions à trois dimensions. De plus, notre corpus de l'époque était trop réduit (18 auditeurs et 32 phrases) pour nous permettre d'avoir assez de puissance statistique. Je salue ici L. Garnier qui a embrassé cette tâche colossale dans le cadre de sa thèse. Ce n'est qu'au prix de longues recherches et de longs efforts qu'elle a pu trouver la solution à cette exploration, à notre connaissance jamais entreprise jusqu'ici⁴².

⁴² J'en profite pour remercier Prof. Lorraine Baqué de l'UAB qui s'est emparée de ce problème avec Laury et qui y a passé également de très nombreuses heures de recherche, à démêler les modèles statistiques pouvant rendre compte de telles interactions.

Nous présentons ici les premiers résultats de ce lourd travail en cours d'achèvement, sans rentrer dans les détails notamment statistiques. Pour cela, nous renvoyons à la thèse de L. Garnier qui sera soutenue début 2018 et qui devrait donner lieu à des articles de revue passionnants.

80 auditeurs francophones ont passé les 3 tâches de perception : proéminences, frontières et groupements. 32 phrases ont été utilisées (16 par condition syntaxique), lus par 4 locuteurs, donnant lieu à un corpus de 128 phrases. Les auditeurs ont été séparés en 4 groupes et devaient juger 32 phrases, répétées 3 fois dans un ordre aléatoire, dans les 3 tâches. Les phrases étaient réparties dans les 4 groupes de sorte à ce que ce ne soit pas toujours les mêmes phrases jugées par le même groupe de sujets. Mais chaque sujet écoutait le même nombre de phrase de chaque condition syntaxique.



Figure 36 – Les 4 groupes proposés aux auditeurs pour révéler le phrasé prosodique des phrases du corpus d'Edinburgh. Les groupements G1 et G2 correspondent au phrasé prédit par les structures syntaxiques manipulées (G1 = 'Low attachment' de A ; G2 = 'High attachment' de A). G3 représente une alternative de phrasé où les syntagmes nominaux et adjectivaux constituent chacun un groupe. Pour G4 à l'inverse, l'ensemble de la séquence est perçu comme un seul groupe. (Garnier *et al*, 2016b))

Nous décrivons ici les hypothèses liées à cette nouvelle tâche : G1 et G2 correspondent respectivement à la première condition syntaxique ('Low attachment' de A, où la syntaxe prédit une frontière d'*ip* entre N1 & N2 et de *w* entre N2 et A) et à la seconde condition syntaxique ('High attachment' de A, où la syntaxe prédit une frontière d'*ap* entre N1 & N2 et d'*ip* entre N2 & A). G3 et G4 ont été choisis pour proposer une alternative à ces deux groupements, en permettant de faire ressortir des degrés de granularité de niveaux de frontières plus ou moins fins. G3 satisfait à la fois les deux prédictions syntaxiques. On attend cependant différentes forces relatives de frontières en condition de 'low attachment' (N1 || N2 | A) et en condition de 'high attachment' (N1 | N2 || A). G4 peut également refléter les deux conditions syntaxiques, mais ne permet pas de mettre en évidence des frontières internes graduées à l'intérieur de la séquence entière (frontière IP de fin de séquence après A).

Comme dans l'étude précédente, la taille de A est fusionnée pour augmenter la puissance statistique et interprétative des résultats. Nous procédons d'abord à l'analyse du jugement des groupements avant de faire le lien entre ces résultats et les jugements de proéminences et de

frontières, séparément. Les sites d'analyses sont N1| N2 (FA sur N1, IA sur N2, autour de la frontière) et N2 | A (FA sur N2 et IA sur A, autour de la frontière).

Les résultats pour les jugements des phrases en groupements sont les suivants : les phrases 'Low attachment' de A sont massivement perçues comme relevant de G1 (93,5%), le reste relevant de G3, alors que les phrases 'High attachment' de A se distribuent de manière plus égales sur les groupes (G4 (42.9%), G2 (32.1%) et G3 (21.0%), marginal en G1). Les cas marginaux de perception de 'low attachment' hors G1, et de 'High attachment' en G1, sont donc exclus des analyses. Si G1 et G2 correspondent aux prédictions syntaxiques, les phrases 'High attachment' sont en revanche aussi distribuées sur G4 et G3 de manière non négligeable, indiquant que différents degrés de granularités de phrasé ont donc été perçus dans cette condition syntaxique. Notons enfin que la condition 'low attachment' de A est très souvent réalisée avec une pause après N1, induisant sans doute un fort biais vers ce choix de G1. Afin de rendre compte au mieux de cette granularité dans le phrasé, nous avons recodé les groupements perçus en fonction des conditions syntaxiques attendues ('low attachment' = 1 ; 'High attachment' = 2) donnant ainsi : G11, G31, G22, G32 et G42 (où le dernier chiffre représente la condition syntaxique attendue).

Nous ne présentons ici que les résultats concernant la granularité des niveaux de constituance (Figure 37), révélée par l'analyse des forces de frontières perçues par type de groupement et l'analyse de la *force relative des frontières* sur le plan syntagmatique (*i.e.* en soustrayant la force de frontière entre N2 | A et N1| N2).

A)			
Groupement	Frontière perçue après N1	Frontière perçue après N2	Force relative entre les frontières perçues (N2 ADJ - N1 N2)
G11	2,4	0,9	-1,5
G31	2,5	1,6	-1,0
G32	1,1	1,5	0,3
G42	0,6	0,9	0,4
G22	0,5	1,5	0,7
B)			
Groupement	Frontière perçue après N1	Frontière perçue après N2	Force relative entre les frontières perçues (N2 ADJ - N1 N2)
G11	[les N1] 	[et les N2]  ADJ]	-1,5 N1 N2 > N2 ADJ
G31	[les N1] 	[et les N2]  [ADJ]	-1,0 N1 N2 > N2 ADJ
G32	[les N1] 	[et les N2]  [ADJ]	0,3 N1 N2 < N2 ADJ
G42	[les N1] 	et les N2  ADJ]	0,4 N1 N2 < N2 ADJ
G22	[les N1] 	et les N2  [ADJ]	0,7 N1 N2 < N2 ADJ

Figure 37 – En haut, les forces de frontières perçues en fonction des groupements par ordre croissant, sur une échelle de 0 à 3 (rappel : les 4 groupes ont été séparés en fonction de la condition syntaxique que leur choix reflétait en majorité) ; en bas, représentation imagée des forces relatives de prééminences (la couleur et la largeur des bandes verticales reflètent les résultats statistiques). (Adapté de Garnier, en cours).

Il faut retenir de ces résultats que les auditeurs ajustent sur le plan syntagmatique la force relative des frontières pour conserver le sens de la structure syntaxique sous-jacente : la condition syntaxique 1 est toujours perçue avec une suite de frontière Forte + faible, l'inverse étant vrai pour la condition syntaxique 2. L'analyse par site nous indique que les auditeurs perçoivent jusqu'à 4 niveaux de frontières en dessous de l'IP de la séquence entière, ce qui nous amène à poser que la plus forte frontière perçue en G11 et G31 entre N1 et N2 (N1|N2 > N2|A) est sans doute également une IP (voir la notion de récursivité des frontières majeures proposée par Ladd, 2008 et la notion de 'macro-unité intonative' de Di Cristo, 2011). En descendant dans la hiérarchie, *ip* est perçue après N2 en G31, G32 et G22, puis *ap* après N2 en G11 et G42, et après N1 en G32, et enfin *w* après N1 en G42 et G22. Aucun de ces résultats ne correspond exactement aux prédictions de la structure syntaxique. Nous n'attendions pas forcément la perception de niveau du mot *w* entre N1 et N2 dans la condition syntaxique 2 mais davantage dans la condition syntaxique 1 après N2, perçue pourtant plus forte (cela rappelle cependant nos résultats en § 13.1, p. 96).

La mise en relation avec la taille des constituants doit encore être démêlée pour avoir une vision plus claire de ces résultats. Cela pourrait expliquer notamment pourquoi la frontière prédite *w* entre N2 & A dans la condition syntaxique 'low attachment' est perçue systématiquement plus forte qu'entre N1 & N2 en condition syntaxique 'high attachment' (prédite *ap*). Selon toute vraisemblance, la construction morphosyntaxique (conjonction 'et' entre N1 et N2) induit la perception d'un lien syntaxique fort entre N1 et N2, réduisant l'impact des indices prosodiques de frontières de mots et de constituants entre les deux noms.

Cela montre sans doute aussi que le marquage par IA et FA sur les mots est attendu quel que soit le niveau de constituance (rappelant nos résultats en § 13.1, p. 96) et ne module pas la perception de frontières de constituants.

Pour ce qui concerne les proéminences justement, les résultats issus de cette étude montrent que FA est modulé partiellement par le type de groupement et le niveau de constituance, notamment sur N1 en G11 et G31. En revanche, IA est perçu avec la même force quel que soit le niveau de constituance et le type de groupement ; autrement dit, ils confirment nos résultats précédents sur le caractère ‘immuable’ de cet accent, essentiellement marqueur de structure et ce, dès le plus bas niveau de constituance. Aucun effet de la taille des constituants ne ressortent pour moduler ces résultats, confirmant encore davantage sa fonction essentiellement structurelle, au moins dans ce type de corpus contrôlé.

13.6. Analyse des syntagmes dits « d’Edinburgh » dans une tâche de MapTask (Bard et al, 2013 ; Gorish et al, 2014)⁴³

Avant de passer à la présentation des premiers résultats d’une étude similaire menée sur de la parole spontanée interactionnelle, il faut mentionner que la présente étude (Garnier, en cours) est en cours d’adaptation sur le Corpus MapTask-Aix⁴⁴ (Gurman Bard et al, 2013 ; Gorish et al, 2014). Ce corpus a été constitué par E. Bard lors de son séjour sabbatique au LPL à Aix en Provence, et financé par la bourse Marie Curie obtenue pour mon séjour de recherche à Edinburgh. Ce corpus a été constitué pour présenter, entre autres mots-cibles, certains syntagmes-cibles issus du Corpus d’Edinburgh (ex : ‘*Les lumières et les balises vertigineuses*’ : voir exemples sonores n°3 et 4) prononcés par des sujets en situation de parole spontanée semi-guidée. La MapTask a été élaborée dans les années 90 par des chercheurs de l’Université d’Edinburgh (Anderson et al, 1991). L’idée est de mettre deux locuteurs face à face (en visuel ou non) possédant chacun une version légèrement différente d’une carte sur laquelle des points de repères (‘*landmarks*’) sont dessinés. Un locuteur endosse le rôle de ‘guide’ et l’autre doit, sur la base de ces instructions, retrouver le tracé d’un chemin le menant d’un point A à un point B sur la carte. Les deux locuteurs engagent donc une conversation spontanée pour effectuer la tâche, mais ils doivent prononcer certains mots-cibles dessinés sur la carte : la conversation est donc considérée comme ‘semi-spontanée’. L’intérêt de cette tâche est qu’elle permet d’élucider des négociations, des questions, des

⁴³ Voir Annexes : documents scientifiques ‘Ressources linguistiques et corpus’.

⁴⁴ Ce corpus est (partiellement) mis à disposition de la communauté scientifique sur la plateforme ‘Ortholang’ ; Bard, E.; Astésano, C.; Frenck-Mestre, Ch.; D’Imperio, M-P.; Turk, A.; Nguyen, N. (2011). *MapTask-Aix*. oai:sldr.org:sldr000732. <http://sldr.org/sldr000732/fr>.

accords ou désaccords etc... *i.e.* toutes sortes de situations linguistiques interactionnelles difficiles à contrôler dans une situation totalement spontanée (voir pour comparaison la description du Corpus CID, ci-dessous). Les MapTasks sur d'autres langues que l'anglais sont rares, et celle-ci constitue la seule version d'une MapTask en français (voir Bard *et al.*, 2013 pour une description détaillée du corpus ; voir Gorish *et al.*, 2014, pour une adaptation en 'face à face' filmé de la version originale de la MapTask française).

16 paires de locuteurs ont produit 8 tâches de cartes dans l'un ou l'autre des rôles (guide ou suiveur). Les cartes ont été dessinées par une artiste⁴⁵ qui avait la lourde tâche d'illustrer certains syntagmes-cibles du Corpus d'Edinburgh (Figure 38).

Si le corpus entier a permis d'éliciter plus de 5000 occurrences de noms et/ou adjectifs dans différentes configurations syntaxiques, seules 79 correspondent aux structures 'Dét + N1 + Conjonction + Dét + N2 + A' du Corpus d'Edinburgh. Une analyse auditive de ces occurrences révèle que la plupart a été prononcé avec des groupements reflétant la structure syntaxique 1 de 'High attachment', sans doute la lecture par défaut de ce type de structure syntaxique. Des analyses perceptives sont en cours pour déterminer dans quelle mesure les locuteurs ont effectivement tenté d'effectuer la tâche de désambiguïsation syntaxique dans la MapTask sur les syntagmes-cibles. Cela nous permettra de comparer 'terme à terme' le marquage prosodique par IA et FA de la structure sous-jacente dans les corpus de parole contrôlée (Edinburgh) et semi-spontanée (MapTask-Aix).

Notons cependant que ce corpus est extrêmement riche et que s'il ne nous permet pas *in fine* de procéder à cette comparaison terme à terme, il comporte tout de même assez d'occurrences autres de structures Nom+Adjectif permettant d'envisager des poursuites d'analyses prosodiques.

⁴⁵ Je salue ici ma très chère amie Luce Monier ('Luce of Escalette') qui a mis son immense talent au service de notre corpus improbable !

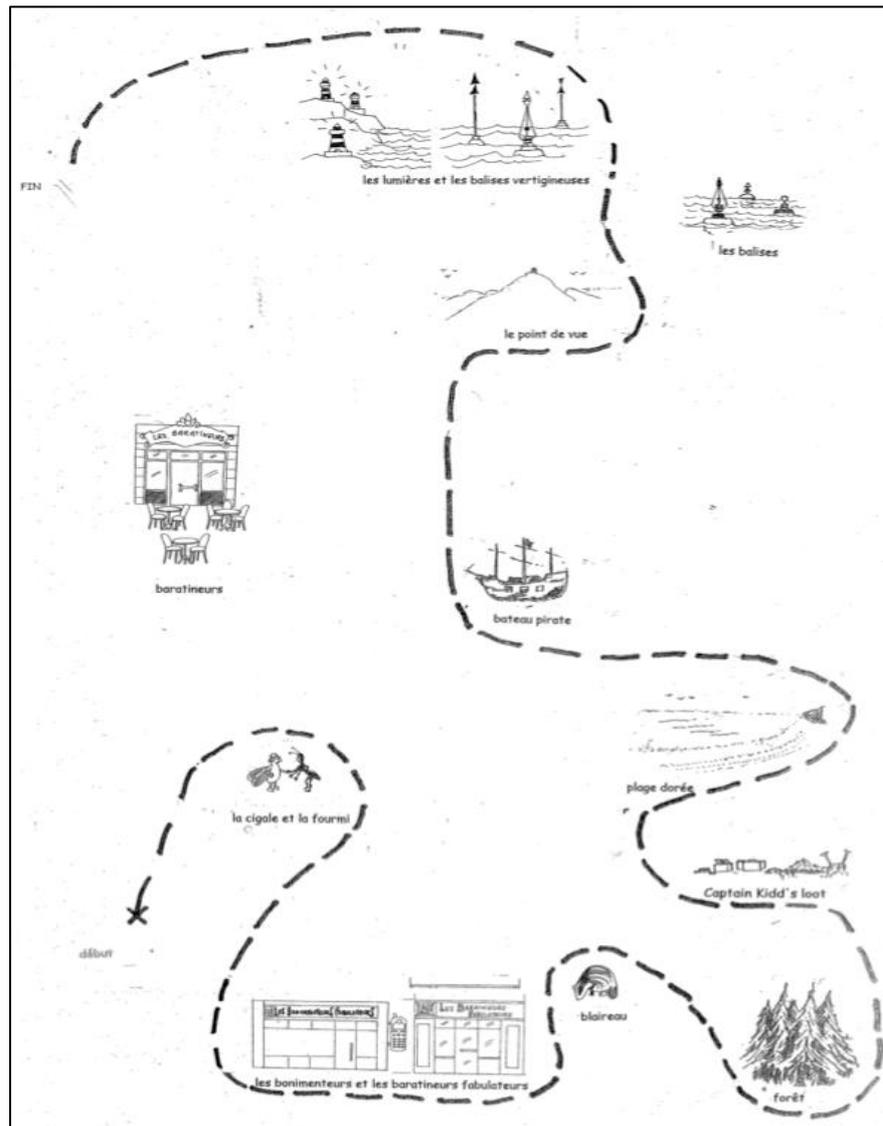


Figure 38 – Exemple de carte de la MapTask-Aix, où apparaissent à divers endroits de la route à suivre les illustrations issues de notre corpus d'Edinburgh (en haut : 'les lumières et les balises vertigineuses' ; en bas : 'les bonimenteurs et les baratineurs et les baratineurs fabulateurs'). Certaines parties de syntagmes cibles sont dessinées sans contexte de manipulation syntaxique, comme distracteurs (ex : 'bonimenteurs' ou 'les balises'). *Exemple sonore n°7*.

13.7. Perception de IA et FA, et liens avec les frontières en parole spontanée (Roux et al., 2016 ; en cours)

Dans le même ordre d'idées, nous avons entrepris de rechercher des structures N+N+A dans le corpus de parole spontanée interactionnelle 'CID' (Bertrand et al., 2008). Ce corpus est constitué de plus de 8 heures d'enregistrements de parole interactionnelle, avec 8 paires de locuteurs parlant d'événements marquants et/ou drôles de leur vie.

Le corpus est entièrement aligné et annoté phonétiquement (algorithme *SPPAS* ; Bigi, 2012) et morphosyntaxiquement (*MarsaTag*, Rauzy et al., 2014). G. Roux, alors post-doctorant sur l'ANR *PhonIACog*, a recherché grâce au tagger morphosyntaxique les 3 patrons

morphosyntaxiques les plus fréquents présents dans le CID : N de N ($n = 430$) ; A + N ($n = 462$) et N + A ($n = 403$). Après correction du tagger, il a procédé à l'enrichissement morphosyntaxique d'une partie du corpus en proposant d'annoter les fonctions syntaxiques de ces patrons, leur position dans l'énoncé, la fonction sujet ou prédicat du syntagme nominal SN, et la complexité du SN. La fonction syntaxique la plus fréquente de ces patrons est Complément d'Objet Direct ($n = 514$), plus fréquente en position finale ($n = 149$) et intermédiaire ($n = 138$) d'énoncés. C'est donc cette dernière fonction syntaxique que nous avons choisie pour les tests de perception comparables à ceux élaborés sur le Corpus d'Edinburgh. Les énoncés les plus audibles et les plus fluents ont été sélectionnés ($n = 133$), puis annotés prosodiquement par G. Roux et R. Bertrand. 73 auditeurs naïfs ont passé les deux tâches de perception : perception des proéminences et des frontières, dont l'ordre de présentation était réparti équitablement entre les sujets. A la différence du Corpus d'Edinburgh, une échelle à 3 niveaux uniquement a été proposée aux auditeurs, allant de 'rien' (pas de proéminence ou de frontière) à 'fort' (proéminence ou frontière la plus forte). La présentation du protocole de test a été adaptée aux contraintes de la parole spontanée dans le logiciel PERCEVAL (Ghio *et al*, 2003 ; André *et al*, 2003).

Figure 39 – A gauche, un exemple de la feuille réponse pour la tâche de jugement de la force des proéminences ; à droite, un exemple de la tâche de jugement de la force des frontières. Les deux tâches sont similaires à celles effectuées sur le Corpus d'Edinburgh. (Adapté de Roux *et al*, 2016).

Au total, les analyses portent sur 3395 syllabes (tâche de proéminences) et 2617 intervalles (tâche de frontières). Globalement, les scores de proéminences et de frontières révèlent un bon taux d'accord entre les 73 auditeurs, et les proéminences sont jugées globalement plus fortes que les frontières mais le taux d'accord est moins fort pour les proéminences. L'annotation experte sur les forces de frontières révèle un très fort taux d'accord (en moyenne .91 sur les 3 forces de frontières) permettant de la comparer avec la perception des auditeurs. Alors que l'annotation experte révèle une bonne distinction entre 4 forces de frontières (IP, *ip*, *ap*, rien), et un meilleur accord au niveau de *ip*, les auditeurs naïfs ne perçoivent une différence de force de frontières qu'entre IP et *ip/ap*, ces deux dernières n'étant pas distinguées. Ce résultat rappelle la proposition de Delais-Roussarie *et al* (2015) selon laquelle la frontière *ip* peut être déduite de la construction morphosyntaxique et ne nécessite pas le traitement des indices prosodiques (voir également Cole *et al*, 2010 et Cole *et*

al, 2014 sur l'influence des informations syntaxiques et sémantiques dans le traitement online de la parole spontanée). Pour ce qui concerne la perception des proéminences, les résultats indiquent que les auditeurs perçoivent de manière similaire IA et FA, *i.e.* qu'ils les jugent forts ou faibles indépendamment de leur position dans le mot ou le syntagme. En d'autres termes, FA n'est pas perceptuellement ou métriquement plus fort que IA. Plus intéressant, les jugements de force métrique sont plus variables pour FA que pour IA, indiquant que IA est plus 'stable' dans sa représentation phonologique que FA. Cela fait écho aux résultats que nous avons déjà obtenus sur le Corpus d'Edinburgh. L'accord entre les auditeurs naïfs et l'annotation experte de IA, FA et inaccentuée est bon, indiquant que les naïfs ne perçoivent pas de proéminences sur les syllabes inaccentuées.

Cette première investigation de la perception des proéminences et des frontières en parole spontanée va encore dans le sens d'une représentation phonologique de IA et FA chez les francophones, et confirment même sur de la parole spontanée 'décousue' que les auditeurs francophones ne sont pas sourds au mètre (*stress*).

Je présente maintenant les résultats que nous avons obtenus sur le traitement de IA par les méthodes d'imagerie cérébrale (EEG). Pour cela, je reviens sur les bases neurales des aspects *prosodiques* du langage, et je présente d'abord les 3 études effectuées avec Mireille Besson sur d'autres aspects prosodiques que le traitement du mètre. La dernière m'a encouragée à envisager l'analyse actuelle des phénomènes métriques fins (traitement de IA) dont les résultats sont présentés dans la foulée.

14. BASES NEURALES DE LA PROSODIE⁴⁶

C'est assez récemment que quelques études de neurosciences ont pris en compte les aspects prosodiques dans le traitement du langage. Elles concernent par exemple le traitement des ruptures temporelles (N2P3 : Besson *et al*, 1997, pour le français), le traitement des frontières prosodiques (Closure Positive Shift ou CPS : Steinhauer *et al*, 1999, pour l'allemand), ou encore le traitement des phénomènes de focalisation (Négativité Frontale et CPS : Hruska *et al*, 2001, pour l'allemand; P300 et N400 : Magne *et al*, 2005a, pour le français, voir § 14.3, p. 119). La grande majorité des études s'intéressent donc aux aspects prosodiques globaux (pauses, contours intonatifs aux frontières prosodiques) ou relevant de l'interprétation sémantico-pragmatique du langage, et relativement peu aux aspects plus précoces en lien avec la représentation lexicale. A part l'étude de Böcker *et al* (1999),

⁴⁶ Voir Annexes : documents scientifiques 'Bases neurales 1, 2, 3 et 4'.

précurseur dans le domaine, cet aspect crucial de la fonction prosodique faisant le lien avec les propositions des modèles psycholinguistiques n'est étudiée que depuis très peu de temps. Nous reviendrons sur ce point particulier en § 14.4, p. 121.

14.1. Spécialisation hémisphérique du traitement de la prosodie

Traditionnellement, on considère que le traitement de la prosodie est localisé dans l'hémisphère droit (voir d'ailleurs le modèle de Friederici, 2011, qui reprend ce point de vue). Il a cependant été proposé que la latéralisation du traitement prosodique s'opérerait sur une base fonctionnelle: la prosodie linguistique (intonation, accentuation, rythme) serait traitée par l'hémisphère gauche (HG), et la prosodie non linguistique (émotions, attitudes) par l'hémisphère droit (HD). Ces deux propositions ont été remises en cause récemment (Kotz *et al*, 2003), et d'autres hypothèses ont été avancées. La première pose que la latéralisation du traitement de la prosodie serait fonction des indices acoustiques (traitement des informations de durée dans HG et des informations de F0 dans HD; Zatorre & Belin, 2001). La deuxième repose sur l'idée que cette latéralisation dépendrait du *domaine* d'application (traitement prosodique au niveau du mot dans HG et au niveau de la phrase dans HD ; Meyer, *et al*, 2002 ; cf. Baum & Pell, 1999, pour une revue).

Présente à tous les niveaux d'intégration linguistique et paralinguistique, il nous semble plus pertinent d'envisager le traitement de la prosodie à travers ses différentes fonctions (fonction de structuration morpho-syntaxique, fonction lexicale, fonction illocutoire et de modalité, fonction pragmatique et discursive, fonction affective et émotionnelle...).

On peut alors émettre l'hypothèse que HG pourrait être impliqué dans la structuration et la hiérarchisation prosodique (marquage lexical, organisation syntaxique, structuration sémantico-pragmatique...), alors que l'appréhension de la pragmatique dans son lien avec le contexte discursif (fonction phatique notamment) et des émotions relèveraient de HD.

Nous présentons maintenant les 3 études que nous avons menées sur différents aspects et fonctions de la prosodie, également documentées dans le chapitre récapitulatif de Besson *et al* (2009). La troisième de ces études constitue la pierre angulaire de ma recherche, puisqu'elle m'a encouragée à tester la réalité métrique de IA à travers les 3 dernières études que nous présentons. Globalement, ces investigations neurolinguistiques viennent confirmer nos propositions sur le statut métrique de IA (et FA) au niveau du *pw* issus de nos expériences en production et en perception.

14.2. Traitement des modalités (Astésano et al, 2004a)

Cette étude constituait notre toute première tentative de décrire le traitement d'une des fonctions les plus aisément appréhendables et consensuelles de la prosodie : la fonction modale de la prosodie permettant aux locuteurs de traduire leur intention de communication. Cette fonction est typiquement décrite par Delattre (1966a) et permet de distinguer souvent sur la base de la prosodie seule, des intentions différentes comme l'assertion (« *Tu vas à Paris.* »), l'interrogation (« *Tu vas à Paris ?* ») ou encore l'ordre, l'injonction (« *Tu vas à Paris !* », sous-entendu, tu ne discutes pas cette décision). Le contour intonatif seul, dans ces cas-là, produit ces changements de sens sur la base d'une même construction syntaxique.

Dans cette première expérience (Astésano et al, 2001a&b ; Alter et al, 2001 ; Astésano et al, 2002 ; 2004a), nous utilisons les modalités d'assertion et d'interrogation totale. Le contour intonatif de l'assertion en français est produit typiquement par un contour descendant jusqu'à la dernière syllabe nucléaire de la phrase. Un contour pré-nucléaire montant est souvent réalisé en fonction de la structure syntaxique sous-jacente (ici sur le nom du groupe nominal, puisque les phrases sont toutes construites sur la forme syntaxique SN+SV).

Le contour intonatif de l'interrogation totale est construit avec un échelonnement descendant ('*downstep*' de f_0) jusqu'à la dernière syllabe nucléaire de la phrase, qui porte un contour intonatif montant. Nos phrases étaient construites sur la même base, *i.e.* des phrases de 6 syllabes (2 groupes accentuels de forme syntaxique Syntagme Nominal SN + Syntagme Verbal SV) assez canoniques en français, du type « *le cycliste pédalait. / ?* ». Un exemple des contours intonatifs utilisés dans cette expérience est donné dans la Figure 40.

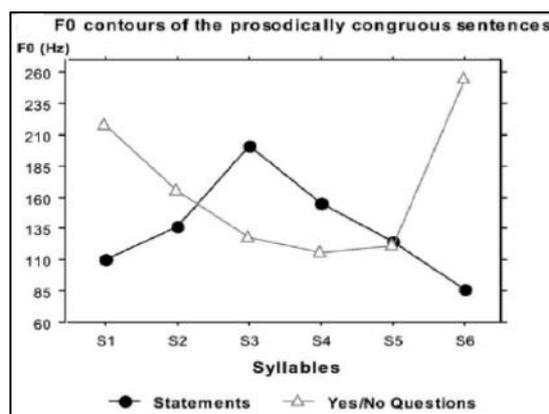


Figure 40 – Contours intonatifs des phrases déclaratives ('Statements') et interrogatives ('Yes/No Questions') du corpus. La f_0 (Hz) est sur l'axe des ordonnées, les valeurs de f_0 sont données pour chaque syllabe de la phrase sur l'axe des abscisses. (Adapté d'Astésano et al, 2004).

L'incongruité prosodique a été créée par la manipulation des contours intonatifs. Chaque phrase ayant été produite en modalité assertive et interrogative, nous avons 'collé' (méthode

dite du ‘cross-splicing’ acoustique) le début d’une modalité assertive avec la fin d’une modalité interrogative et vice-versa. Ce ‘cross-splicing’ acoustique se faisait après la 3^{ème} syllabe de chaque phrase, à la frontière entre le SN et le SV (chaque verbe commençait par une occlusive sourde afin de faciliter le cross-splicing et pour qu’il ne soit pas perceptible sur le plan segmental).

Les phrases étaient également construites de sortes à présenter pour moitié une incongruité sémantique (par exemple : « *le bilan traversait* »). Le dessin expérimental (carré latin) présentait donc 4 conditions expérimentales : 1) Phrases prosodiquement et sémantiquement congruentes ; 2) Phrases prosodiquement et sémantiquement incongrues ; 3) Phrases prosodiquement congruentes et sémantiquement incongrues ; 4) Phrases prosodiquement incongrues et sémantiquement congruentes.

L’intérêt de manipuler orthogonalement la dimension prosodique (modalité) et la dimension sémantique est triple: d’une part, le traitement sémantique est très bien documenté dans la littérature des Potentiels Evoqués, et la composante N400 que nous devrions observer sur les phrases incongrues sémantiquement nous sert de base de comparaison avec le traitement de la modalité prosodique, jusqu’alors jamais documenté dans la littérature : ce traitement intervient-il avec la même latence ou plus tôt, et avec la même polarité ou non ? D’autre part, sachant que la prosodie est intimement liée avec les autres niveaux de traitement linguistique, il était intéressant de voir dans quelle mesure l’information de modalité est indépendante de l’information sémantique, ou si, au contraire, elle interagit avec elle. Enfin, nous voulions savoir si cette information prosodique est tout simplement traitée en temps réel, au même titre que les autres informations linguistiques (sémantique, puis syntaxe étant les niveaux linguistiques les plus documentés dans la littérature des Potentiels Evoqués). Comme cela est souvent le cas dans ce type d’études, l’attention des participants était également manipulée (attention au sens de la phrase d’un côté, à la mélodie de la phrase de l’autre) afin de déterminer le caractère automatique ou non du traitement des informations prosodiques.

Les résultats montrent une plus grande facilité de traitement des aspects sémantiques que des aspects prosodiques (temps de réaction et pourcentages d’erreur). Au delà de la composante N400 attendue pour le traitement de la sémantique, cette étude nous a permis de mettre en évidence une composante EEG liée au traitement de la modalité prosodique : la P800, un traitement plus tardif par conséquent de cet aspect prosodique, reflétant sans doute une difficulté supérieure de la tâche prosodique. Il faut noter en effet qu’il est difficile de créer une incongruité prosodique nette, car les contours intonatifs admettent une certaine variabilité individuelle alors que le sens d’un mot non congruent dans son contexte est plus

facilement repérable (et manipulable expérimentalement). Il est tout à fait possible que le ‘cross-splicing’ mélodique n’ait pas permis de créer des ‘monstres prosodiques’ parfaits, et que certains cas aient pu être interprétés comme des variations possibles d’une question ou d’une assertion, voire qu’ils aient donné lieu à l’interprétation d’un troisième type de contour modal. Le deuxième résultat concerne le caractère non automatique apparent du traitement de la modalité, puisque la composante P800 était absente lorsque les sujets focalisaient leur attention sur la sémantique ; la N400 était en revanche présente même lorsque les sujets se focalisaient sur le contour mélodique, preuve que la sémantique est traitée de manière automatique, comme rapporté multiples fois dans la littérature. Enfin, l’interaction entre sémantique et modalité prosodique est asymétrique en ce sens que seule la sémantique module la composante P800 (voir Figure 41) alors que l’incongruence prosodique n’influence pas l’amplitude de la N400.

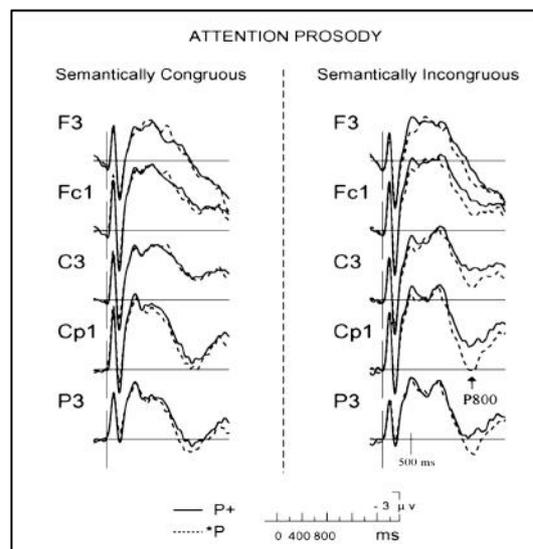


Figure 41 – Composante P800 reflétant le traitement de la modalité lorsque les participants focalisent leur attention sur la prosodie (contour mélodique de la phrase). La P800 est plus ample lorsque la sémantique est également incongrue, reflétant une interaction (asymétrique) entre traitement sémantique et prosodique, la sémantique étant prédominante. Notez le traitement centro-pariétal gauche de la modalité prosodique. (Adapté d’Astésano *et al.*, 2004).

Ces résultats montrent une prédominance du traitement sémantique sur le traitement prosodique mais des études supplémentaires contrôlant d’avantage la manipulation prosodique ou présentant un autre type de manipulation devraient être entreprises pour confirmer ou moduler ces résultats, qui doivent être pris avec précaution au regard des résultats que nous avons obtenus par la suite en neuroimagerie sur d’autres dimensions du traitement prosodique.

14.3. Traitement du focus (Magne *et al*, 2004 ; 2005a)

Lors de mon deuxième post-doctorat dans l'équipe de Mireille Besson, cette dernière a eu la gentillesse de m'associer à une recherche en cours qu'elle avait entreprise avec Anne Lacheret-Dujour, sur le traitement de la fonction prosodique de focalisation. Le corpus avait déjà été construit par les linguistes de Caen et j'ai aidé Cyrille Magne et Mireille Besson dans la passation et l'interprétation des résultats EEG, ainsi qu'à l'analyse acoustique de l'accent focal (Astésano *et al*, 2004b).

Dans l'interaction langagière, la prosodie aide à la structuration du message en mettant en exergue certains mots du message (figures) par rapport aux autres (fond), de sorte à faire ressortir l'information nouvelle dans le contexte interactionnel, à laquelle l'interlocuteur doit prêter attention. Dans ce processus de focalisation d'un mot ou d'une partie d'énoncé, les paramètres acoustiques vont se modifier : augmentation de la durée et de l'intensité, concomitante à un mouvement tonal la plupart du temps ascendant/descendant sur le mot focalisé. En phonologie prosodique, on parle de déplacement de l'accent nucléaire sur le début du mot focalisé ou le mot entier, donnant lieu, lorsque celui-ci est placé en milieu d'énoncé, à une parenthésisation de la suite de l'énoncé (contour plat de *f0* post-focal, mais où la structuration prosodique perdure à travers l'allongement de durée sur les syllabes métriquement fortes ; Di Cristo & Jankowski, 1999).

Comme pour l'étude précédente, il s'agissait de voir si cette caractéristique prosodique est traitée en temps réel par les auditeurs ; contrairement à l'étude précédente cependant, la sémantique n'était pas manipulée orthogonalement. Seules la congruité prosodique et la *position* du focus prosodique, en position médiane ou finale d'énoncé, étaient manipulées. La congruité prosodique consistait à manipuler la portée de l'accent focal sur l'un ou l'autre des mots de l'énoncé, en réponse à une question demandant un choix paradigmatique. Par exemple, à la question :

« *Qu'a bu Pierre ce matin : du café ou du chocolat?* », la réponse attendue doit proposer un choix paradigmatique entre '*café*' et '*chocolat*' en focalisant l'un ou l'autre ; par exemple, « *Pierre a bu du café ce matin* », où le mot souligné figure l'accent focal ('*ce matin*' étant ici post-focalisé). Si à cette même question, la réponse est « *Pierre a bu du café ce matin* », la portée focale est incongrue dans le contexte discursif donné par la question. Cette réponse serait congruente à la question suivante : « *Quand Pierre a-t-il bu du café : ce matin ou ce midi ?* ».

Les phrases ont été construites de sorte à obtenir des paires Questions/Réponses congruentes ou incongrues sur le plan pragmatique. Chaque mot ('café' et 'matin') peut donc se retrouver soit congruent soit incongru en fonction du contexte, et constitue de ce fait son propre contrôle. La contre-partie était que le traitement du focus était analysée en fonction de sa position dans l'énoncé : en position médiane ou finale, sans que l'on ait d'hypothèse *a priori* sur l'influence de ce facteur sur le traitement du focus.

Les résultats sont intéressants à deux titres : d'une part, les accents focaux incongrus en milieu ou fin d'énoncé ne donnent pas lieu à la même composante EEG. L'incongruité prosodique intervenant au milieu de la phrase élicite une composante positive se développant entre 300 et 800 ms, de la famille de la P300 que nous interprétons comme l'expression de la surprise de l'auditeur qui ne s'attend pas à trouver ce mot-là focalisé. L'incongruité prosodique intervenant à la fin de l'énoncé élicite en revanche une négativité se développant autour de 400 ms, que nous interprétons comme une N400 révélant une difficulté d'intégration par rapport au contexte précédant l'accent focal. La Figure 42 illustre ces principales tendances.

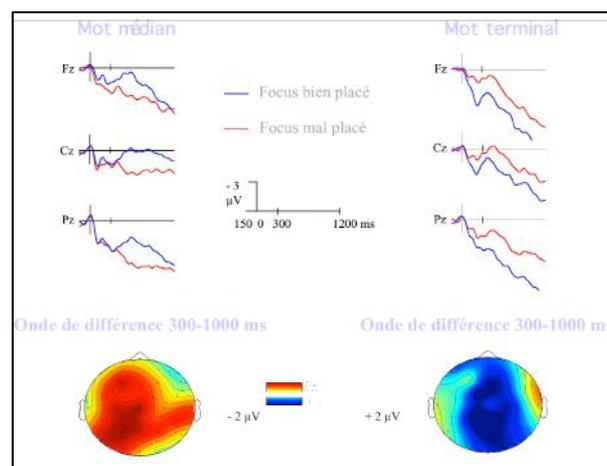


Figure 42 – Composante P300, à gauche, associée au traitement de l'accent focal sur le mot en position médiane dans l'énoncé ; Composante N400, à droite, associée au traitement de l'accent focal sur le mot en position finale d'énoncé. En rouge, les composantes plus amples associées à l'incongruité de l'accent focal. Les représentations sur les têtes illustrent la positivité (couleur rouge) et la négativité (couleur bleue). Notons que la négativité a la même distribution centro-pariétale que la N400 'classique' trouvée dans les cas d'incongruité sémantique. (Adapté de Magne *et al*, 2005a, et 2004).

Plus intéressant encore, nos résultats indiquent que l'absence d'accent focal sur le mot attendu dans le contexte discursif élicite quand même une composante positive en milieu d'énoncé et négative en fin d'énoncé. Certes, la *latence* de cette composante élicitée en l'absence d'accent focal est moins rapide que lorsque l'accent focal incongru est réalisé, mais la réponse cérébrale est tout de même similaire. En d'autres termes, que l'incongruité pragmatique soit instanciée par la *présence d'un accent focal inattendu* ou par l'absence d'un

accent focal attendu, la composante PE associée est similaire. Cela signifie que la réponse cérébrale n'est pas *phonétiquement* motivée mais *pragmatiquement* motivée.

Ces résultats apportent un éclairage enrichissant dans le domaine de l'interprétation des données linguistiques par la neuroimagerie : nous avons, par cette étude, mis au jour le traitement de la *représentation cognitive* du focus pragmatique, et non pas son traitement purement *acoustique*. Plutôt, ce dessin expérimental en carré latin nous a permis de montrer que *l'absence d'indices acoustiques* est également (et évidemment) traité, nous permettant de sortir de l'interprétation de surface qu'ont parfois trop souvent ce type d'études.

Il faut garder en tête ce dernier point pour l'interprétation des prochains résultats, notamment pour ce qui concerne le traitement de l'accentuation initiale.

14.4. Traitement du mètre (Astésano et al, 2004c ; Magne et al, 2007)

Cette étude constitue en fait la pierre angulaire de ma recherche. C'est la première fois en effet que l'on montrait le traitement en temps réel du *mètre* en français⁴⁷, m'invitant à développer mes investigations en ce sens dans toutes les études suivantes en perception de l'accentuation du français (versants comportementaux et de neuroimagerie) : elle préfigurait l'obsession qu'allait devenir pour moi le fait de démontrer que l'accentuation en français a bel et bien une réalité phonologique (donc métrique).

Nous avons vu plus haut que les patrons métriques sont utilisés pour segmenter le flot de parole continue et accéder au sens (voir § 8, p. 61). Les quelques recherches sur le français montrent essentiellement l'utilisation de l'allongement final dans ces stratégies de segmentation, mais au niveau du groupe accentuel et indépendamment de l'utilisation de cet indice métrique dans l'accès au sens (segmentation).

Notre étude EEG proposait de créer une incongruité métrique sur des mots trisyllabiques, en utilisant un algorithme de 'time-stretching' conservant les propriétés tonales de chaque syllabe, développé par notre collègue G. Pallone (voir Ystad et al, 2003 pour une présentation détaillée de l'algorithme). L'incongruité métrique ne visait pas la dernière syllabe du mot trisyllabique mais la *syllabe médiane*. Il s'agissait de rapporter l'allongement de la syllabe médiane à l'allongement 'naturel' moyen des syllabes accentuées en français, d'un rapport de 1.7 (Astésano, 2001)⁴⁸. L'allongement final naturel de fin d'énoncé était en revanche conservé. La moitié des phrases comportait cette incongruité métrique. Parallèlement, nous avons

⁴⁷ Voir néanmoins les résultats comportementaux obtenus auparavant dans Jankowski et al (1999) sur le traitement de l'accent initial, malheureusement restés confidentiels.

⁴⁸ Notons qu'une partie de ce corpus, impliquant une manipulation 'écologique' et extrêmement subtile des indices prosodiques sur le plan perceptif, a également servi à l'évaluation et la validation d'un système audio performant utilisable dans le contexte très bruyé de l'IRMf (Ghio et al, 2004).

manipulé l'incongruité sémantique orthogonalement à l'incongruité métrique, obtenant ainsi 4 conditions expérimentales : 1) Phrases métriquement et sémantiquement congruentes ; 2) Phrases métriquement et sémantiquement incongrues ; 3) Phrases métriquement incongrues et sémantiquement congruentes ; 4) Phrases métriquement congruentes et sémantiquement incongrues (Figure 43).

	Semantically congruous (S+)	Semantically incongruous (S-)
Metricaly congruous (M+)	<i>Le concours a regroupé mille <u>can</u>didats</i> "The competition hosted a thousand <u>can</u> didates"	<i>Le concours a regroupé mille <u>bigoudis</u></i> "The competition hosted a thousand <u>bigoudis</u> "
Metricaly incongruous (M-)	<i>Le concours a regroupé mille <u>can</u>didats</i> "The competition hosted a thousand <u>can</u> didates"	<i>Le concours a regroupé mille <u>bigoudis</u></i> "The competition hosted a thousand <u>cur</u> liers"

Figure 43 – Exemples de phrases manipulant orthogonalement l'incongruité métrique (M+ et M- ; en M-, l'allongement est d'un rapport de 1.7 sur la syllabe médiane ; l'allongement est naturel de la syllabe finale) et l'incongruité sémantique (S+ et S-). L'allongement syllabique (naturel ou manipulé) est indiqué par le soulignement. (Adapté de Magne *et al*, 2007).

Les participants devaient focaliser leur attention tantôt sur les aspects métriques (irrégularités rythmiques) tantôt sur les aspects sémantiques, de sorte à pouvoir observer le traitement automatique ou non du mètre et d'évaluer les effets d'interaction avec le traitement sémantique.

Les résultats concernant la tâche sémantique révèlent des tendances habituellement observées dans ce type d'études : l'incongruité sémantique élicite une composante N400 'classique', et ce même lorsque les participants portent leur attention sur la composante métrique seule, confirmant encore le caractère automatique du traitement sémantique. L'incongruité métrique élicite quant à elle une composante positive autour de 700 ou 800 ms en centro-pariétal, rappelant ainsi le traitement des effets de modalité observés dans notre étude antérieure (Astésano *et al*, 2004) : les aspects métriques sont donc bien traités en temps réel. Comme auparavant, cette positivité n'est observée que lorsque les participants prêtent attention aux aspects métriques, mais jamais lorsqu'ils prêtent attention à la sémantique. Dans ce dernier cas cependant, il est très intéressant de noter que, contrairement à l'étude précédente, l'effet d'incongruité métrique induit une modulation de la N400 alors même que les participants se concentrent sur la sémantique (Figure 44). De fait, l'incongruité métrique semble nuire à l'accès au sens, ce qui se retrouve également dans les résultats comportementaux puisque les phrases métriquement incongrues ralentissent la détection de l'effet de congruité sémantique.

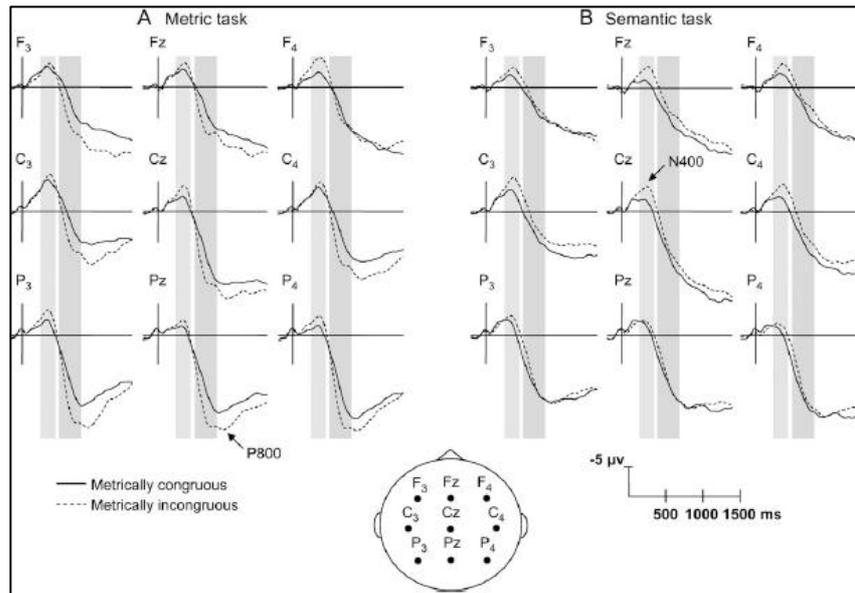


Figure 44 – Effet de l'incongruité métrique dans les deux conditions d'attention : les participants portent leur attention sur la métrique (A) ou sur la sémantique (B). On observe une composante positive P800 en Pz dans la tâche d'attention à la métrique, alors qu'une composante N400 se développe en Cz dans la tâche d'attention à la sémantique. (Adapté de Magne *et al*, 2007).

L'article de Magne *et al* (2007) ne s'est pas attardé sur les effets de double interaction entre sémantique et métrique, *i.e.* dans quelle mesure l'effet d'incongruité métrique module l'effet d'incongruité sémantique. Nous avons particulièrement insisté sur ces aspects dans Astésano *et al* (2004c) car ils nous semblaient particulièrement fondamentaux pour démontrer l'importance de la représentation métrique *au niveau lexical*. En effet, nous avons montré que, dans la tâche d'attention portée à la sémantique, l'incongruité métrique avait une incidence y compris sur les mots *sémantiquement congruents* (ex : « *Le concours a regroupé mille candidats* », où '*candidats*' est métriquement incongru mais sémantiquement congruent). Une composante N400 était élicitée quand bien même le mot était cohérent dans son contexte (Figure 45), ce qui, à notre connaissance, n'avait jamais été montré.

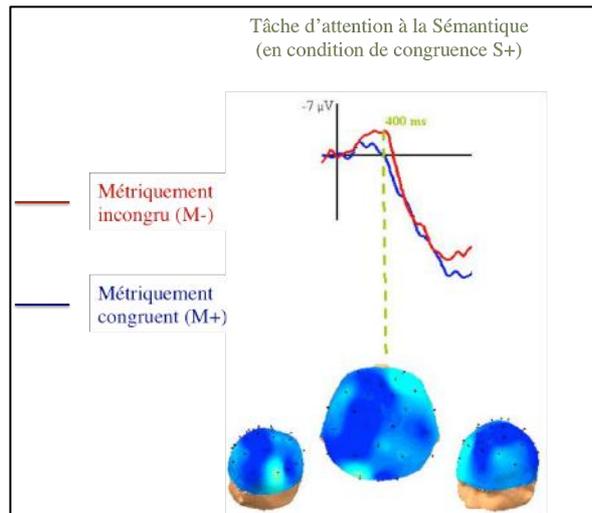


Figure 45- Effet de l'incongruité métrique sur les mots sémantiquement *congruents* (S+) dans la tâche où les participants font attention à la Sémantique. La composante N400 pour les mots métriquement incongrus (en rouge) révèle que, même si le mot ne pose pas de problème d'interprétation dans le contexte de la phrase, le patron métrique illégal entrave et ralentit son traitement. (Adapté de Astésano *et al*, 2004c ; *exemples sonores n°8 M+ et n° 9 M-*).

Cela nous avait amenés à conclure que *la composante métrique en français joue un rôle crucial dans l'accès au sens*. En d'autres termes, si un mot congruent dans son contexte sémantique est reconnu moins facilement quand son patron métrique est inadéquat (ou illégal dans le système phonologique), cela signifie que la représentation du patron métrique en français est intimement liée à la représentation du mot lexical.

14.5. Traitement de l'accent Initial IA (Astésano *et al*, 2013; Aguilera *et al*, 2014 ; te Rietmolen *et al*, 2016 a, b, c)⁴⁹

Quelques années se sont écoulées depuis ce résultat saisissant, augurant le meilleur pour ma quête du 'graal' métrique. L'installation dans mon poste à Toulouse et les diverses collaborations qui en ont découlé ont agréablement retardé la poursuite de ce travail.

Ma délégation CNRS entre 2011 et 2013 m'a heureusement permis de pouvoir me recentrer un temps sur ma recherche. C'est à cette occasion que j'ai pu affiner mon projet, et le faire financer par l'ANR (de janvier 2013 à décembre 2016). C'est également là que j'ai pu nouer toutes les collaborations nécessaires à l'avancement de ce projet avec le LPL qui m'accueillait à Aix-en-Provence, et reprendre ma collaboration avec Ellen Gurman Bard sur l'Accent Initial (voir § 13.6, p. 110 notamment). C'est enfin grâce à ce temps-recherche que j'ai pu développer une de mes collaborations locales à Toulouse les plus fructueuses⁵⁰, avec Radouane el Yagoubi à CLLE-LTC.

⁴⁹ Voir Annexes : documents scientifiques 'Bases neurales 4'.

⁵⁰ Mais également les plus frustrantes si l'on considère que nos emplois du temps respectifs nous empêchent trop souvent de pouvoir concrétiser, par des publications en revues internationales, nos résultats pourtant très élégants sur le traitement de l'accent initial.

Il nous semblait important de se poser la question centrale du *traitement de l'accentuation en français par des francophones*, afin de pouvoir faire évoluer le débat selon lequel les français sont tout simplement 'sourds' (phonologiquement) à l'accentuation (voir § 17, p. 142 pour une discussion). Nous n'avons pas choisi la voie de la facilité en décidant de nous pencher d'emblée sur le traitement de IA, type d'accent dont le statut est encore moins consensuel en français (voir § 5.5, p. 50). Mais pour les raisons évoquées notamment au § 11.3, p. 84, selon lesquelles IA permet d'éclairer le statut de l'accentuation en français, nous avons pris ce risque. Certes, IA est marqué par des indices prosodiques plus subtils que FA, et sans doute moins facilement perceptible, mais d'une part, nous avons déjà quelques indices qu'il était marqueur systématique de la structure prosodique (Astésano *et al*, 2007) et d'autre part, nous avons montré qu'il pouvait être perçu par les auditeurs aussi bien que FA (Astésano *et al*, 2012). Il fallait donc persévérer en ce sens, en utilisant une méthode encore plus sensible au traitement de cet accent : l'EEG.

Nous avons pris la décision de commencer l'investigation du traitement de IA en partant des processus les plus précoces (processus de détection de la saillance accentuelle et d'attente contextuelle, révélant le traitement d'un événement par la composante MMN), et de 'remonter' petit à petit vers les traitements plus tardifs, en relation avec le traitement lexical puis sémantique.

Traitement à court et long termes de IA dans un paradigme dit de 'oddball' (Astésano et al, 2013 ; Aguilera et al, 2014).

Tout a donc commencé avec notre co-encadrement d'un Master 1&2 de Sciences du Langage.

Il s'agissait de construire des stimuli permettant de tester le traitement de IA dans la mémoire à court terme (traitement phonétique) et à long terme (traitement phonologique). Le paradigme de Oddball se prête à ce type d'investigation puisqu'il consiste à présenter aux auditeurs des stimuli 'déviants' dans un contexte de stimuli 'standards', plus nombreux. Il permet donc de mesurer la trace mnésique que la répétition des stimuli standards laisse en mémoire à court terme. Les stimuli déviants vont éliciter une négativité précoce révélant leur traitement. Cette composante négative se développe autour de 200 ms et révèle un défaut d'appariement dans la perception du stimulus déviant par rapport au stimulus standard. La composante de MisMatch Negativity (MMN) est donc une composante PE reconstruite que l'on obtient en soustrayant les ondes cérébrales élicitées par les stimuli déviants aux ondes cérébrales élicitées par les stimuli standards. La composante MMN est d'autant plus ample

que le stimulus-étalon (standard) laisse une trace mnésique importante. En d'autres termes, si le stimulus étalon appartient à l'expérience linguistique des auditeurs, ici à la représentation phonologique qu'il a de son système, la MMN sera d'autant plus ample puisqu'elle reflètera le traitement des représentations en mémoire à long terme. Le paradigme Oddball permet donc à la fois de mesurer la discrimination auditive d'un stimulus, mais aussi de voir si cette discrimination dépasse le traitement de bas niveau (Näätänen *et al*, 2007). Si les francophones sont sourds à l'accentuation, que ce soit d'un point de vue phonétique ou phonologique, alors on ne devrait pas obtenir de MMN montrant une distinction de traitement entre les stimuli avec et sans IA.

Ce protocole expérimental ne demande pas la construction d'un grand nombre de stimuli. Au contraire, c'est la répétition d'un même stimulus qui va engendrer l'effet attendu. La sélection des stimuli a été délicate, ainsi que leur manipulation. Nous avons choisi de partir sur la base de stimuli trisyllabiques portant naturellement un IA (voir la description du corpus ayant servi de base à l'extraction des stimuli dans le paragraphe suivant décrivant l'étude de décision lexicale). Nous avons sélectionné certains stimuli qui nous paraissaient les plus naturels puis nous leur avons appliqué un algorithme de resynthèse permettant de supprimer la variation tonale portée par IA (puisque l'essentiel de l'information prosodique de IA est portée par le contour de f_0). Les stimuli ont ensuite tous été resynthétisés pour ne pas créer de biais potentiel entre les stimuli avec et sans IA. Une vingtaine de participants naïfs ont ensuite écouté ces stimuli et donné un score évaluant leur aspect 'naturel'. Aussi bien les stimuli non modifiés avec IA et ceux resynthétisés sans IA ont ainsi été jugés. Nous n'avons retenu que 2 des meilleurs stimuli ('*candidat*' et '*diffusion*'), nécessaires à notre étude. La Figure 46 illustre la manipulation acoustique d'un de ces stimuli.

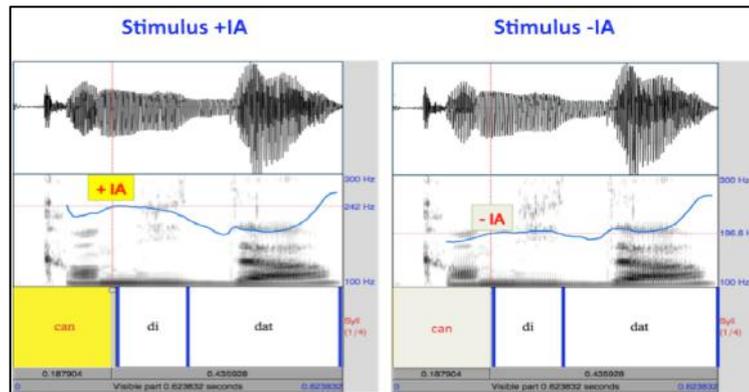


Figure 46 – Exemple de resynthèse de la f_0 (à droite) utilisant un algorithme d’interpolation quadratique entre la valeur de f_0 du déterminant précédent le mot ‘candidat’ et la dernière syllabe portant FA. Les variations micro-prosodiques sont ainsi préservées pour plus de naturel. A gauche, le stimulus naturellement prononcé avec IA, mais resynthétisé pour que les conditions expérimentales soient comparables (Adapté de Aguilera *et al*, 2014 ; *exemples sonores n°10 +IA et n°11 -IA*).

Afin de véritablement tester l’hypothèse selon laquelle IA appartient au système phonologique, nous avons créé deux groupes de participants de sorte à ne pas induire le résultat que nous espérions trouver : un groupe ($n = 14$) écoutait la série de stimuli pour lesquels +IA était le stimulus étalon et -IA le déviant, alors que l’autre groupe ($n = 16$) écoutait la série de stimuli où -IA était le stimulus étalon et +IA le déviant. Nous ne présumons pas, ainsi, de l’existence ou non de IA dans la représentation de la mémoire à long terme. Les deux groupes ont passé chacun deux types de tâche : une tâche passive dans laquelle ils écoutaient simplement le flot continu de stimuli et une tâche active dans laquelle ils devaient appuyer sur un bouton dès qu’ils entendaient un stimulus déviant dans le contexte.

Les résultats indiquent que les auditeurs perçoivent bien la différence entre les mots trisyllabiques portant +IA et ceux qui n’en portent pas. En effet, s’ils avaient été ‘sourds’ à l’accentuation, les deux groupes de participants auraient présenté les mêmes tendances, *i.e.* pas de différence entre les stimuli déviants +IA et -IA. Au contraire, nos résultats indiquent que non seulement ils traitent l’information de saillance acoustique mais qu’ils semblent préférer un patron accentuel avec +IA. Les résultats pour chaque groupe de participants (*‘within-participants’*) montrent en effet que les participants ayant écouté les stimuli où -IA est le déviant présentent une MMN beaucoup plus ample que les participants ayant écouté les stimuli où +IA est le déviant. Si le traitement n’était qu’un traitement de bas niveau (acoustique), l’onde de différence (*‘difference wave’*) de la MMN aurait été similaire dans les deux cas. De plus, étant donné que la MMN mesure la rareté d’un stimulus dans un contexte, le fait que les stimuli +IA déviants aient élicité une MMN moins ample signifie que les auditeurs n’ont pas pu faire appel à leur mémoire à long terme, car les standards -IA ne laissent pas de trace suffisamment forte dans la mémoire à court terme.

Afin de s'assurer de ce qui relève de la rareté du stimulus indépendamment des phénomènes acoustiques, nous avons procédé à une comparaison inter-participants ('*between participants*'). Nous avons comparé la MMN résultant de la soustraction des stimuli -IA entre eux (déviant et standard) et +IA entre eux. Les résultats montrent clairement que les stimuli -IA élicitent une MMN bien plus ample que les stimuli +IA, nous permettant d'écartier l'interprétation purement acoustique de ces résultats. Globalement, nos résultats montrent non seulement que les stimuli avec ou sans IA sont traités en mémoire à court terme (*l'absence de saillance prosodique est surprenante*) mais aussi que les stimuli +IA sont encodés en mémoire long terme puisque les mots portant des patrons métriques -IA ne correspondent pas à la représentation à long terme des patrons métriques. La Figure 47 illustre ces résultats.

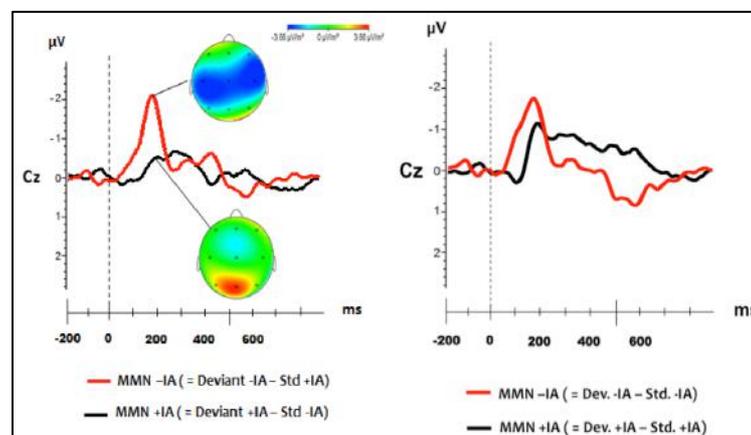


Figure 47 – A gauche, MMN pour les stimuli déviants -IA (en rouge) et déviants +IA (en noir) intra-participants ('within participants') indiquant que les stimuli -IA sont plus inattendus que les stimuli +IA. A droite, recalcul de la MMN pour les stimuli déviants -IA comparés aux même stimuli -IA standards (en rouge) et déviants +IA comparés aux même stimuli +IA (en noir). Ce résultat, inter-participants ('between participants'), montre qu'au delà des caractéristiques purement acoustiques, les stimuli -IA restent des patrons métriques plus inattendus que les stimuli +IA. (Adaptés d'Aguilera *et al.*, 2014).

Les résultats comportementaux issus de la tâche active sont également fort intéressants et viennent compléter dans le même sens les résultats EEG. Ils montrent déjà que les stimuli déviants -IA sont moins rapidement détectés que les stimuli déviants +IA, et qu'ils génèrent davantage d'erreurs de détection. Plus intéressant encore, l'analyse ultérieure des résultats comportementaux entre les 2 blocs de passation⁵¹ montre que les temps de réaction se raccourcissent et les taux d'erreurs baissent entre les deux blocs uniquement pour les stimuli +IA (voir Figure 48), ce qui montre un effet d'apprentissage généralement observé seulement pour les items représentés en mémoire à long terme. De fait, les résultats comportementaux assoient encore davantage notre interprétation selon laquelle le patron métrique attendu en français comporte IA.

⁵¹ Les participants avaient 5 minutes de pause entre 2 blocs de présentation de la tâche, plus fatigante que la tâche passive.

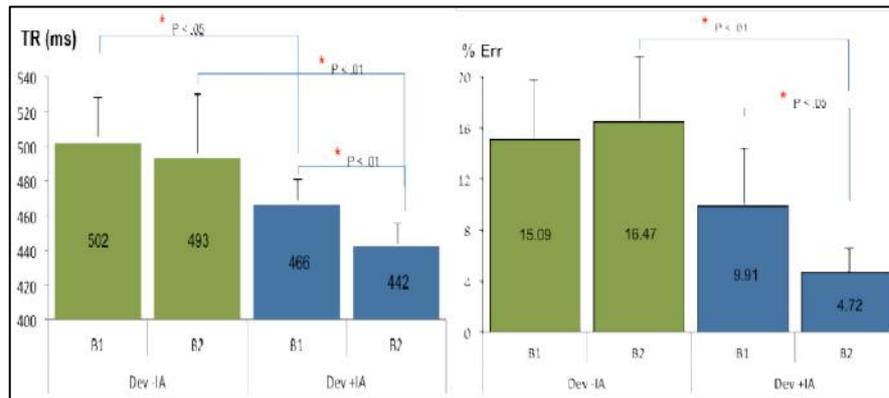


Figure 48 – Temps de réaction (à gauche) et taux d’erreurs (à droite) entre les deux blocs (bloc 1 : B1 ; bloc 2 : B2) de passation de la tâche active de détection des stimuli déviants pour les déviants –IA (histogrammes verts à gauche) et les déviants +IA (histogrammes bleus à droite). Les déviants +IA sont significativement reconnus plus rapidement que les stimuli –IA dans chaque bloc et entre les blocs. Les taux d’erreurs montrent un effet significatif des blocs pour les déviants +IA (effet d’apprentissage) mais également moins d’erreurs dans le bloc2 par rapport aux déviants –IA, confortant encore cet effet d’apprentissage. (Astésano & el Yagoubi, en cours).

Ces résultats doivent être approfondis et sont en cours de rédaction. Tout d’abord, nous allons appliquer une méthode statistique plus conservatrice (tests non paramétriques) afin de tester plus avant la localisation peu canonique de cette MMN, la MMN étant généralement plus frontale que centrale dans la littérature. Ensuite, nous allons mettre en parallèle les résultats comportementaux (effets d’apprentissage avec les stimuli +IA) avec les résultats EEG comparant ce même effet d’apprentissage à travers les différents blocs de passation.

Traitement pré-lexical de IA dans une tâche de décision lexicale (te Rietmolen, en cours ; te Rietmolen et al, 2016a&b).

On sait maintenant que IA semble être le patron accentuel attendu en français. Il nous fallait donc tester plus avant son lien avec la représentation lexicale. Encore une fois, très peu d’études utilisant les potentiels évoqués PE concernent ce lien entre accentuation et lexique. A notre connaissance, seule l’étude de Böcker *et al* (1999) sur le néerlandais a permis de mettre à jour une nouvelle composante EEG reflétant la difficulté dans l’extraction des patrons accentuels : la composante N325. Cette composante négative, se développant à partir de 325 ms (donc avant la composante reflétant le traitement du sens, la N400) en frontal (Fz), a été révélée dans une tâche de discrimination de patrons accentuels sur des mots bisyllabiques présentant soit le patron métrique courant en néerlandais (F-f) soit le patron métrique moins courant (f-F). Les résultats indiquaient que la composante N325 était plus ample pour les patrons accentuels moins fréquents (f-F). La N325 est donc interprétée comme révélant le traitement du mètre.

Nous avons donc testé la préférence du patron métrique +IA en français sur des mots lexicaux et des pseudo-mots trisyllabiques, manipulant la présence et l'absence de IA à l'instar de ce que nous avons fait sur l'étude MMN précédente.

Le corpus est constitué de 120 mots et 120 pseudo-mots prononcés par une locutrice naïve dans des phrases porteuses complexes forçant l'introduction de pauses avant et après l'incise afin de s'assurer de la réalisation claire et naturelle d'un IA et d'un FA (marqueurs de syntagme intonatif) sur le mot cible. Un exemple de phrases porteuses est donné ici :

Mot cible 'candidats' : *L'ensemble des présidents du jury, dit-il, et les candidats, pourraient ne pas être reçus dans la même salle.*

Pseudo-mot cible 'piralos' : *Les jeux de hasard et d'argent, dit-il, et les piralos, peuvent être une addiction très dangereuse pour lui.*

Le corpus a été contrôlé phonétiquement afin de ne comporter que des mots équilibrés sur le plan de la complexité syllabique et sur le plan segmental (proportion équivalente de mots et pseudo-mots dont la syllabe initiale débute par une plosive sourde, une liquide etc..., comportant des voyelles nasales ou hautes etc...)

La resynthèse permettant d'enlever les IA sur la moitié des stimuli utilise la méthode d'interpolation quadratique décrite dans l'étude de MMN précédente. Les 240 stimuli étaient donc répartis dans un carré latin entre mots/pseudo-mots et stimuli +IA et -IA. Les 26 auditeurs avaient pour tâche de dire si le stimulus entendu était un mot ou un pseudo-mot. De fait, contrairement à l'étude de Böcker *et al* (1999), nos participants n'avaient pas conscience de la manipulation métrique.

Les résultats comportementaux ne révèlent qu'un effet (attendu) de la lexicalité mais pas d'effet du patron accentuel. Autrement dit, la présence ou l'absence de IA n'intervient pas dans la rapidité et la précision avec laquelle ils vont donner leur réponse.

Les résultats EEG en revanche sont plus intéressants (Figure 49).

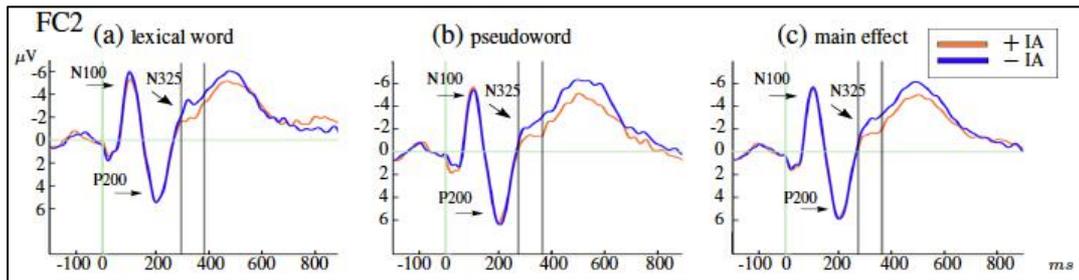


Figure 49 – La composante N325 révélant la difficulté d’extraction des patrons métriques est plus ample pour les stimuli –IA que +IA en (c) pour l’effet principal, et en (a) pour les mots lexicaux. Cet effet n’atteint cependant pas le niveau de significativité pour la condition ‘pseudo-mots’ (b). Notons au passage que les composantes N1-P2 ne diffèrent pas selon le patron + ou –IA, indiquant un processus descendant de traitement du mètre. (Adapté de te Rietmolen *et al*, 2016).

Les stimuli –IA élicitent une N325 plus ample que les stimuli +IA, révélant une plus grande difficulté de traitement des stimuli sans IA. Même si l’amplitude de la composante est faible, elle est tout à fait comparable à celle trouvée chez Böcker *et al* (1999), alors même que dans notre étude, l’attention des participants n’était pas attirée sur le patron métrique. Ceci montre donc que IA est traité à un niveau pré-lexical, de manière automatique. Un autre résultat intéressant concerne la composante P200, connue pour refléter les processus de bas niveau d’extraction des paramètres physiques et acoustiques. Notre manipulation de la présence de IA n’a pas modulé la P200, indiquant que le patron accentuel relève davantage d’un processus descendant (top-down ; pour plus de détails sur l’interprétation de ce résultat, voir te Rietmolen *et al*, 2016). Globalement, ces résultats sont l’indication d’une représentation phonologique *métrique* de IA, au même titre que dans les langues à accentuation lexicale reconnues comme le néerlandais.

Il faut évidemment aller plus loin dans l’investigation des corrélats neuronaux de IA en particulier, et de l’accentuation bipolaire IA-FA, pour obtenir une cartographie plus détaillée nous permettant de saisir avec plus de nuances la différence de *niveau* de traitement phonologique de l’accentuation (non lexicalement distinctive) en français.

Le rôle de IA dans le traitement de la sémantique (te Rietmolen, en cours ; te Rietmolen *et al*, 2017).

En analysant le traitement de IA au niveau de l’énoncé et non plus seulement au niveau du mot isolé, cette dernière étude nous permet d’investiguer le lien entre mètre et représentation sémantique, à une étape plus tardive du traitement de la parole. Cette étude a utilisé une partie du corpus d’Astésano *et al* (2004c) et Magne *et al* (2007) décrite plus haut qui montrait un traitement retardé des mots portant un patron métrique illégal (allongement de la syllabe médiane). La présente étude vise à compléter ces études précédentes, en testant plus

précisément la réalité métrique de IA. Nous postulons que les mots ne portant pas +IA éliciteront également une N400 plus ample. Autrement dit, nous testons le fait que les patrons métriques –IA représentent une déviance à la représentation métrique des francophones. Le niveau cognitif testé est donc plus subtil que dans Magne *et al* (2007), passant de patrons totalement illégaux dans les études précédentes, à des patrons simplement moins attendus.

Nous avons sélectionné 80 phrases du corpus originel où les mots-cibles (mots trisyllabiques de fin d'énoncé ; voir § 14.4, p. 121) portaient naturellement +IA dans leur contexte. Le même algorithme d'interpolation quadratique a été utilisé pour créer les stimuli –IA, mais a dû être considérablement adapté aux contraintes de la parole continue (te Rietmolen, en cours). Les mots étaient pour moitié sémantiquement congruents. La manipulation métrique et sémantique était orthogonale (x4), donnant lieu à 320 stimuli. 16 sujets ont écouté chacun 80 phrases uniques.

Les résultats comportementaux ne révèlent pas d'effet de la manipulation métrique, seulement celui de la manipulation sémantique. Les résultats EEG donnent des résultats intéressants, mais moins nets que dans l'expérience principale sur des mots isolés. L'effet principal de la sémantique est évidemment significatif, les phrases sémantiquement incongrues présentant une N400 plus ample. Il est intéressant également de noter que ce n'est pas tant l'amplitude mais la latence de la N400 qui ressort statistiquement de l'analyse des stimuli –IA (voir Figure 50 : (b) à gauche), montrant une latence de N400 plus tardive pour les mots sémantiquement incongrus –IA. L'amplitude de la N400 seule explique (marginale) l'effet de la sémantique pour les mots +IA (voir Figure 50 : (c) à gauche). On peut donc dire sur la base de ces résultats testant l'effet sémantique que les mots –IA retardent la résolution du conflit sémantique, alors que les mots +IA sont modulés par l'incongruité sémantique 'typique', ce qui tendrait à indiquer que les patrons +IA sont plus naturels et attendus phonologiquement.

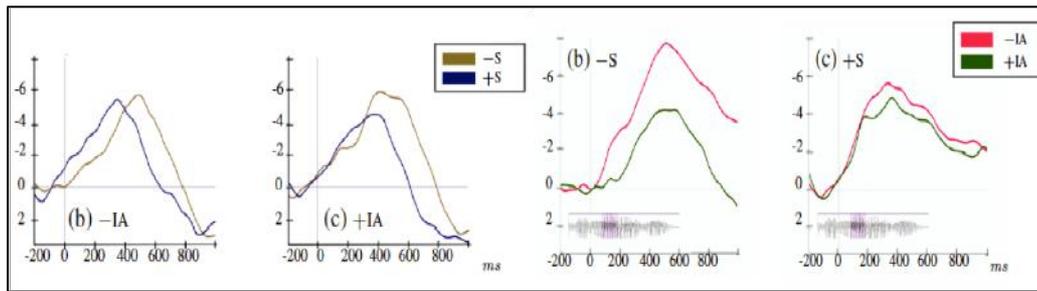


Figure 50 – A gauche, l’effet de congruïté sémantique sur les stimuli –IA et + IA ; à droite, la modulation de l’effet sémantique par la présence/absence de IA (Adapté de te Rietmolen *et al.*, 2017).

L’effet principal de la métrique n’atteint pas le seuil de significativité. En revanche, les patrons –IA génèrent une N400 plus ample sur les mots sémantiquement incongrus que les patrons +IA (Figure 50 : (b) à droite). Le traitement des mots sémantiquement congruents n’est en revanche pas modulé par \pm IA (Figure 50 : (c) à droite ; l’effet n’atteint pas le seuil de significativité). Il se pourrait que nos analyses statistiques (et notre puissance statistique) ne parviennent pas à révéler cette modulation par le patron supposé attendu +IA, ou que seul le double coût de traitement (–IA –S) soit assez fort pour ressortir dans ces analyses.

Ces résultats vont tout de même globalement dans le même sens que les résultats obtenus sur des mots isolés, à des niveaux plus bas de traitement linguistique. Ils sont explorés plus avant dans la thèse en cours de N. te Rietmolen.

Ces études sur l’accent initial sont maintenant étendues à FA, toujours dans le cadre de cette même thèse, et devrait nous permettre de vérifier si FA et IA relèvent bien d’une même représentation métrique (voir notre proposition, § 16, p. 140).

Les résultats de ces deux dernières études (décision lexicale et traitement de la sémantique) me font m’interroger sur le *niveau* d’interprétation de l’accentuation en français. Le fait que les résultats comportementaux (impliquant un traitement émergent à la conscience de l’auditeur dans la plupart de nos études) ne permettent pas de révéler les processus sous-jacents au traitement du mètre est en soi, sans doute, un indice de la subtilité d’un tel processus en français, un indice d’un traitement moins lié précisément à l’*identité* du mot lexical. Il est possible que la différence de résultats que nous obtenons entre les résultats comportementaux (non significatifs) et EEG (montrant, à l’inverse, les mêmes composantes associées au traitement des patrons accentuels que dans les langues à accentuation lexicale) soit l’indice même d’une différence de *niveau* de représentation de l’accentuation en français, pressentie par Di Cristo (2000). Ainsi, si on admet que la représentation de l’accentuation en français se situe à *un certain niveau du mot lexical* (Di Cristo), voire quelle est plus

profondément ancrée dans la représentation du mot lexical encore (notre proposition), il n'en demeure pas moins que l'accentuation ne définit pas le caractère lexical du mot. En tant que marqueur 'pur' des limites du mot, le niveau de représentation métrique est donc moins profondément imbriqué avec le niveau de représentation lexicale qu'il ne l'est pour les langues à accentuation lexicale comme l'anglais ou le néerlandais. Selon notre proposition métrique, l'alternance possible entre IA et FA pour marquer les différents niveaux de constituance en surface (alternance nécessaire pour éviter la 'surpopulation' tonale dans certains cas) est l'indication même de cette moindre imbrication avec le lexique.

Nos résultats EEG sont donc extrêmement encourageants et nous avons hâte de voir ce qui pourra ressortir de la comparaison entre le traitement de IA et de FA pour avancer encore dans notre réflexion.

STRETTE

Quelles implications pour les modèles phonologiques et psycholinguistiques ?

Résumons rapidement les principales tendances révélées par nos résultats de production, de perception et d'imagerie cérébrale. Elles concernent IA, au centre de ces études, FA dans ses différentes réalisations, et enfin le niveau de constituance auxquels ils semblent s'appliquer.

15. PRINCIPALES TENDANCES

Dans la plupart de nos études de perception, IA est perçu plus fort que FA, ou en tous cas, à la même force que FA H* et plus fort que FA L* ou !H, y compris en parole spontanée où les auditeurs appliquent des choix similaires de forces de prééminences à FA et IA. En outre, IA est toujours perçu avec la même force quel que soit le niveau de constituance ou sa position dans la hiérarchie (en constituant pré-nucléaire ou nucléaire) et ce, dès le niveau du *pw*, renforçant le point de vue selon lequel il relève de ce niveau grammatical. D'autres auteurs l'avaient noté bien avant nous (Pasdeloup, 1990 ; Delais, 1995 ; Vaissière, 1997 ; Welby, 2003 et suivantes), mais n'avaient jamais jugé cet argument suffisant pour justifier de la prise en compte systématique de IA dans les descriptions prosodiques (voir cependant Sichel-Bazin, 2016, pour une proposition d'accent lexical initial en français méridional). Cet 'oubli' est sans doute lié à deux phénomènes : d'une part, le mot lexical n'est pas considéré comme niveau d'analyse en français : le rôle lexical de IA ne rentre donc pas dans les modèles pré-établis ; d'autre part, IA est indifféremment considéré dans sa fonction 'rythmique' et 'emphatique' dans la plupart des approches. Ces deux fonctions sont mêlées et ne sont pas considérées comme pertinentes dans la distinction fonctionnelle de IA. Son implication dans la structuration du français, indépendamment de sa fonction emphatique, n'a été expérimentalement testée que dans Astésano *et al* (2007), où nous trouvons que IA était davantage structurel et secondairement rythmique.

Une rapide investigation de l'alignement du pic de IA (H_i ou H₁) montre en effet que le pic est situé au centre du mot lexical long lorsqu'on prend en considération la hauteur maximale de H. Les données acoustiques (tonales et de durée) le confirment. Pour autant, il est *systématiquement* perçu sur la première syllabe du mot lexical, soit que les auditeurs projettent de manière 'top-down' la représentation lexicale et adaptent la réalisation acoustique 'floue' à cette dernière ; soit qu'ils se basent sur le 'saut' de *f0* entre le déterminant précédent et la première syllabe, généralement bien au dessus des 10% nécessaires à la

perception d'une proéminence tonale. Ce dernier point constituerait une preuve que les *changements* de f_0 sont plus importants dans la perception d'une proéminence tonale que leur *amplitude* (Fry, 1958 ; Ladd, 2008). Cette question demande à être testée expérimentalement. Notons en outre que H_1 est parfois réalisé comme H_{1p} , *i.e.* en un plateau tonal interpolant avec H^* . Cette tendance a été observée sur un corpus de parole contrôlée et en parole spontanée mais doit encore être systématiquement éprouvée.

Enfin, nos résultats de neuroimagerie indiquent que IA semble être le patron attendu par les auditeurs : il est représenté en mémoire à court terme (les auditeurs le perçoivent), en mémoire à long terme (il est préféré à la version des mots sans IA), et participe à la représentation du mot lexical. Par ailleurs, son absence ralentirait l'accès au sens. La faiblesse de nos résultats comportementaux, qui peinent à atteindre la puissance statistique nécessaire pour confirmer les résultats EEG, pourrait révéler que l'accentuation n'effleure pas à la conscience des participants francophones, comme ce serait le cas pour l'accent lexical en anglais par exemple. Cela pourrait donc être interprété comme l'indication que IA, au moins, est représenté à un autre niveau du mot lexical que l'accent lexical dans les langues à accentuation distinctive, comme le propose par ailleurs Di Cristo (2000).

Pour ce qui concerne FA, nos expériences de perception montrent également qu'il est perçu. En réalité, il est perçu plus saillant que les syllabes inaccentuées environnantes même en situation de désaccentuation interne d'*ap*. Reste à déterminer, comme pour IA, si cette perception est due à la connaissance du lexique des auditeurs (processus 'top-down') ou si les traces de durée observées sur FA dans cette position suffisent à la perception de son poids métrique (ou encore s'il est perçu en raison du contexte tonal environnant). Aux frontières de constituants majeurs, FA est également perçu comme métriquement fort, contredisant les visions de Rossi (1980) ou Jun & Fougeron (2000) de sa disparition au profit du seul contour intonatif. Par ailleurs, nous avons montré que le type de configuration tonale sur FA (H^* , L^* ou $!H$) module sa perception au niveau de l'IP uniquement. Autrement dit, au niveau *ap*, FA H^* ou L^* sont perçus avec le même poids métrique, alors que FA est perçu plus fort métriquement en frontière d'IP lorsqu'il est réalisé en L^* ou $!H$. La dissociation que l'on observe entre proéminence finale et frontière, permettant d'arguer du statut métrique indépendant de FA par rapport aux frontières intonatives, est donc à moduler aux niveaux les plus hauts de constituance. Cette dissociation partielle est intéressante à plus d'un titre : 1) FA survit aux frontières d'IP ; 2) à l'inverse même des propositions antérieures à la base du questionnement sur le statut de l'accentuation en français, nos résultats montrent au contraire un renforcement métrique aux niveaux de constituance supérieurs et une interaction entre

frontière et proéminence à ce niveau. L'inversion de pente décrite par Delattre (1966a) et Martin (1981) pourrait certes relever d'une contrainte intono-syntaxique mais elle pourrait aussi prendre appui sur une règle de *réorganisation métrique et tonale* de l'accent final en fonction du contexte tonal (présence de IA et « encombrement tonal »). Il s'agirait donc d'une réinterprétation métrique de cette proposition intonative initiale. Une étude sur l'alignement de ce ton L* ou !H avec la syllabe accentuée doit être entreprise (Mendez, en cours) pour déterminer le type d'accent auquel on a affaire, et sur le lien qui l'unit aux contours intonatifs.

Enfin, on observe incontestablement une relation entre proéminences et unités grammaticales en français. Les questions qu'on est en droit de se poser, surtout dans le contexte théorique entourant les descriptions du français, sont les suivantes : quelle est la nature du lien qui unit les proéminences à l'unité grammaticale ? Est-ce que la proéminence est représentée au niveau de l'unité grammaticale, et à quel niveau ? Doit-elle être représentée ? Si on considère que les locuteurs analyseraient le niveau grammatical préalablement aux proéminences, alors aucune spécialisation des proéminences n'est nécessaire. La proéminence serait redondante par rapport à l'unité grammaticale. Or, il y a non seulement des phénomènes phonétiques donnant une substance à la forme phonologique qui permettent d'envisager les proéminences en tant qu'unités, mais il existe aussi des contraintes syntagmatiques et hiérarchiques qui pèsent sur les proéminences et les constituants qui les régissent. Ces phénomènes structuraux apparaissent comme une confirmation de la représentation de la proéminence au niveau de l'unité grammaticale. Des études antérieures ont montré en outre que les indices acoustico-phonétiques sont utilisés pour distinguer la hiérarchie des proéminences même en contexte désémantisé (Jankowski, 1997). Nos résultats sur la perception du groupement mettent en évidence les contraintes syntagmatiques influant sur l'occurrence des proéminences. Ils indiquent que ces proéminences interviennent dès le niveau du mot (Garnier, en cours). Ces tendances sont confirmées par la capacité qu'ont les auditeurs à percevoir IA et FA comme métriquement plus forts que les syllabes inaccentuées alentours dès le niveau du mot, que ce soit en parole contrôlée ou spontanée. Enfin, on a pu montrer par la méthode EEG que IA au moins est représenté au niveau du mot et utilisé par les auditeurs pour accéder au sens. Des études sont en cours sur FA utilisant la même méthode.

Tous ces résultats nous amènent à considérer le niveau du *pw*, tel que défini par Selkirk (1996) et Nespors & Vogel (1986), comme domaine d'actualisation de l'accentuation en français.

16. PROPOSITION DE PIED SPONDEÏQUE MARQUEUR DU *PW*⁵²

Je soumets donc cette proposition avec une certaine fébrilité⁵³ : le niveau du *pw* serait marqué en français par IA et FA de même force métrique. Cette proposition peut sembler étrange à plusieurs égards. D'une part, je propose que l'accentuation soit considérée comme lexicale en français (évidemment pas dans une fonction distinctive). D'autre part, le pied spondeïque (F-F) est considéré comme rare et irrégulier en poétique. Sur le plan de la phonologie prosodique, il viole le principe de dominance selon lequel le pied est marqué par une seule tête (à droite en français ou à gauche en anglais). Ce faisant, la fonction de culminativité de l'accentuation est quelque peu mise à mal à ce niveau abstrait, puisque l'unité grammaticale ne porte pas une marque de proéminence mais deux. Ce pied est donc hybride puisqu'il est défini par deux accents primaires⁵⁴ et qu'il ne se situe pas au niveau habituel d'un pied (la syllabe) mais au niveau du mot lexical lui-même. Je définis donc ce pied comme un 'pied interne au *pw*', tel que rapidement évoqué par Shattuck-Huffnagel & Turk (1996 : 218), *i.e.* un pied qui se définirait par rapport à l'unité grammaticale qui le contient (notion de 'Within-Prosodic Word Foot'). Je rappelle en effet qu'à l'instar de Jun & Fougeron, je ne suis pas d'accord avec le niveau inférieur d'Unité Tonale de Di Cristo & Hirst car elle est définie par deux accents de substance différente (tonale essentiellement pour IA ; tonale et durationnelle pour FA). Je propose en revanche que IA et FA instancient une unité prosodique cohésive (tendance centrifuge de l'accentuation) que je ne situe pas au niveau de *ap* comme Jun & Fougeron ou du Mot Prosodique Complexe comme je le proposais en 2001, mais au niveau du mot.

Finalement, cette proposition étrange de pied spondeïque au niveau du mot renvoie, sous un autre angle, à la notion de 'langue exotique' que je dénonçais au début de ce travail...

Me serais-je également laissée piéger par la difficulté de description de cette langue ?

Quoiqu'il en soit, voici ce que donnerait une telle proposition, avec toutes les améliorations qu'on doit sans doute pouvoir lui apporter (Figure 51 ; exemple issu de Delais-Roussarie & Di Cristo, à paraître : « *le président grec* »).

⁵² Voir Annexes : documents scientifiques 'Prosodie 1'.

⁵³ J'espère que cette proposition suscitera beaucoup de débats me permettant de l'affiner ou de la réviser. Elle en est pour l'instant à un stade prototypique.

⁵⁴ Voir Ladd (2008) pour une critique de la notion d'accent primaire et secondaire, et de hiérarchie des proéminences.

contenue dans une *ap* incluant le proclitique. C'est ce que j'appelle (peut-être un peu abusivement) les cas d'enclitisation prosodique de l'adjectif au niveau du syntagme, à défaut d'une autre terminologie.

Une toute dernière remarque sur cette proposition. Je disais ci-dessus qu'il reste encore à tester expérimentalement le niveau de projection de ces deux accents 'primaires' et les règles d'effacement de l'un ou de l'autre pour éviter l'encombrement tonal. Cela m'évoque la notion d'*accent probabilitaire* de Fónagy (1980) : peut-être avait-il à l'esprit ce marquage alternatif par IA et/ou FA des différents niveaux hiérarchiques. Cette proposition permet peut-être d'apporter un éclairage métrique à la notion de Fónagy.

Tournons-nous maintenant vers les implications que de telles propositions ont sur notre conception de la représentation de l'accentuation en français et sur les modèles psycholinguistiques.

17. LA SURDITE ACCENTUELLE EN QUESTION

Comme énoncé dans l'introduction, la confusion terminologique qui règne autour de la phonologie prosodique du français, et les mésententes sur le niveau de représentation auquel se porte l'accentuation, ont des effets en cascade sur la conception des psycholinguistes désireux de comparer les stratégies de traitement de la prosodie dans plusieurs langues. C'est le cas de Dupoux *et al* (1997, et suivantes) dont il me semble que les conclusions sur la supposée surdité accentuelle des francophones ont été surinterprétées.

Généralement, ce type d'études est essentiellement centré sur la conception selon laquelle le français possède un seul accent FA, fixe, prévisible et congruent aux frontières prosodiques. Il n'est jamais fait référence à IA, ou si c'est le cas, jamais dans sa version de représentation *métrique*. Son rôle est par conséquent peu ou pas pris en considération. Globalement, nul cas n'est fait non plus de la potentielle *probabilité* d'occurrence des accents en français, telle que proposée en son temps par Fónagy (1980). Globalement aussi, on néglige le caractère métrique de l'accentuation (finale), sur la base de la non distinctivité de ladite accentuation. Bref, les préceptes issus des langues à accentuation lexicale comme l'anglais, qui dominent les descriptions prosodiques, sont calqués sur les autres types de langue, dont cette fameuse langue exotique qu'est le français. Ce qui nous avait fait dire que :

« *l'interprétation (...) de surdité accentuelle persistante des francophones initiée par Dupoux et collègues (depuis) 1997 pourrait, de fait, être davantage liée aux prémisses théoriques et aux outils conceptuels sous-tendant ces recherches* »
(Astésano & Bertrand, 2016 : 25).

Rappelons que cette interprétation émane de tests de perception où les sujets devaient, dans différentes tâches, percevoir la place de l'accent sur des non-mots trisyllabiques prononcés par des locuteurs néerlandophones. Les caractéristiques de l'accentuation sont donc celle d'une langue étrangère. D'autre part, l'accent était porté sur la première, deuxième ou troisième syllabe des non-mots. Les résultats montraient que les francophones percevaient moins bien l'accent que les locuteurs hispanophones, même si la différence entre les deux groupes n'était pas colossale. Pourtant, les auteurs avaient conclu à la surdit  accentuelle (phonologique) des francophones, due à l'absence de r alit  m trique de l'accentuation dans leur langue. Cette surdit  phonologique  tait sans doute pr suppos e, et on a sans doute affaire à une d monstration 'à double bande', quelque peu circulaire. Finalement, plus qu'une surdit  (phonologique), ne devrait-on pas conclure que les auditeurs francophones ont en r alit  une attente phonologique diff rente de celle pr suppos e dans cette  tude ? Cette attente serait  quivalente sur les sites initiaux et finaux des mots lexicaux (mais pas m dian comme  a peut  tre le cas pour les hispanophones) : ils transf reraient donc cette strat gie m trique en langue  trang re. Michaux (2016) montre d'ailleurs que les francophones per oivent plus facilement les accents initiaux en n erlandais que les patrons m triques f-F.

De plus, ne valait-il pas mieux avant tout tester cette 'surdit  accentuelle' sur la mani re dont les francophones traitent *leur propre syst me accentuel* ? C' tait le but des recherches qui ont  t  men es dans le cadre de ce travail. Nos r sultats de perception *ET* de traitement de l'accentuation (m thode EEG) contredisent compl tement cette notion de surdit  accentuelle, et m'ont m me permis de faire cette proposition m trique incongrue que je ne pouvais jusque-là qu'intuiter. En esp rant qu'elle permettra de mener d'autres exp riences de perception et de traitement de l'accentuation sur les francophones avec des pr misses th oriques plus propices, à mon sens, à la discussion scientifique.

18. LE *PW* COMME BASE METRIQUE POUR LES MODELES PSYCHOLINGUISTIQUES

Nous renvoyons le lecteur à nos remarques sur les limites des mod les m triques propos s pour le fran ais (§ 8.1, p. 63). Rappelons simplement que la prise en compte de IA dans la phonologie m trique du fran ais permet selon moi de proposer une alternative conceptuelle à la suppos e faiblesse m trique et perceptuelle de FA en fronti res de constituants intonatifs (nos r sultats ont d'ailleurs montr  qu'elle n' tait finalement pas si effective que cela puisque FA survit perceptuellement aux fronti res intonatives). La prise en compte de IA permet  galement de diminuer la probabilit  de segmentation à l'initiale de chaque syllabe (Content *et al*, 2001) aux seules syllabes de *pw* marqu es par IA, si tant est que l'on adh re à la

conception selon laquelle la ‘métrique’ syllabique dominerait en français (Cutler *et al*, 1986, et suivantes). Je pense avoir montré avec ces résultats qu’on peut au moins prendre en compte IA et FA comme deux entrées possibles à une stratégie de segmentation métrique accentuelle en français. Reste encore à déterminer dans quelle mesure ils interviennent et à quel niveau ils s’appliquent l’un et l’autre préférentiellement en surface. On n’est peut-être pas obligé de croire au fait qu’ils soient marqueurs *lexicaux* au sens canoniquement entendu. Mais tout porte à croire qu’ils sont tout de même ancrés, spécifiés quelque part dans la représentation lexicale en français. Je ne vais pas reprendre tous nos résultats abondamment présentés dans ce document pour étayer ma proposition. Je voudrais juste signaler un dernier point : Cutler (1999) propose que, au delà de la notion de force métrique (*stress*), la notion de groupement jouerait également un rôle important dans l’accès lexical. C’est pourquoi d’ailleurs j’avais intégré cette dimension dans les études de perception que nous menons actuellement. C’est en effet une dimension différente des notions de proéminences et de frontières prosodiques, intégrant davantage la manière dont les auditeurs ‘coupent’ les ‘morceaux’ de parole (notion assez polysémique de ‘*chunks*’) pour décoder le sens. Notre proposition de marquage cohésif des constituants par IA et FA dès le niveau de *pw* pourrait aller dans le sens des propositions de Cutler sur les stratégies d’accès au sens. En tous cas, elle mérite d’être testée systématiquement sur le plan expérimental.

Voilà qui en termine avec cette *Fugue*, cœur de ma recherche. Les propositions qui émanent de ce travail seront sans doute amenées à être révisées, mais elles permettent de fixer l’état de réflexion auquel j’ai abouti tout au long de ces années.

Les *Variations* autour de ce thème sont présentées ci-après.

VARIATIONS

La prosodie a des applications évidentes pour les recherches en L2 et pathologies du langage, puisqu'à l'interface avec tous les niveaux de représentations du langage qu'elle organise. Son étude interdisciplinaire permet également d'apporter un éclairage théorique relevant de la (macro-)syntaxe ou de la phonologie (liaison par exemple). J'ai donc volontiers participé à des recherches interdisciplinaires plus éloignées de mon cœur d'expertise, en essayant néanmoins d'insuffler ma vision métrique de la prosodie du français. Par ailleurs, pour en revenir à la genèse de mon parcours, j'ai également construit une partie de ce parcours sur la comparaison des systèmes linguistiques et musicaux via la prosodie. Puisque les recherches applicatives (L2 et pathologies du langage notamment) font souvent appel au concept de 'Prosodie, Musique de la langue' dans leurs pratiques et leurs méthodes, la cohérence de ces apartés scientifiques est quelque peu respectée.

19. COGNITION COMPAREE DU LANGAGE ET DE LA MUSIQUE⁵⁶

Historiquement, mon parcours m'a conduit à m'intéresser aux relations que peuvent entretenir deux systèmes typiquement humains que sont le Langage et la Musique. C'est ainsi que j'ai entamé une longue collaboration avec Mireille Besson s'étalant sur plusieurs années et post-docs. En réalité, cette collaboration aura été l'occasion pour moi de me familiariser avec les principes de la méthode de neuroimagerie, l'Electro-EncéphaloGraphie (EEG), permettant de suivre le décours temporel de l'activité électrique du cerveau lors du traitement des informations linguistiques, et plus particulièrement prosodiques. L'investigation des bases neurales de la prosodie (voir § 14, p. 114) a en effet davantage été au cœur de mes travaux avec l'équipe de Mireille Besson. J'ai cependant participé à plusieurs études qui ont donné lieu à une dizaine de publications et communications sur les liens musique/langage, ainsi que des conférences invitées dans des manifestations en sciences cognitives et sciences du cerveau.

La prosodie est naturellement au cœur de la comparaison langage/musique ; elle est d'ailleurs souvent définie comme la 'musique de la langue' puisqu'elle partage avec la musique les paramètres organisateurs que sont le rythme et le mètre, et avec elle aussi les paramètres acoustiques qui participent de ces principes organisateurs sonores : variations de hauteur (*f*0), de durée et de force (intensité). Cette similitude naturelle entre prosodie et musique explique d'ailleurs pourquoi, historiquement, les linguistes-prosodistes ont souvent utilisé la transcription musicale pour illustrer les règles d'organisation prosodique, à des fins de recherche fondamentale ou pour des applications en L2 par exemple : l'école britannique ;

⁵⁶ Voir Annexes : documents scientifiques 'Bases neurales 5'.

Bolinger et ses phrases annotées sur plusieurs hauteurs (sans doute un casse-tête à l'époque des machines à écrire !), Fónagy pour le français, Faure & Di Cristo (1977) avec leur livre à destination des apprenants étrangers... Ladd (2008) et Patel (2010) en font rapidement mention dans leurs ouvrages respectifs.

C'est pourquoi ce thème de recherche entre en résonance avec les préoccupations des chercheurs en remédiations linguistiques (neurologues, orthophonistes), des enseignants de langues étrangères et des spécialistes du développement langagier, qui voient dans les similitudes entre ces deux systèmes un outil précieux pour renforcer l'acquisition ou le recouvrement des éléments linguistiques.

19.1. Rythme musical et linguistique

Ce premier volet de recherche concernait la comparaison entre le traitement du rythme en langue et en musique (Ystad *et al*, 2003 ; Magne *et al*, 2005b) sur la base d'une manipulation orthogonale de la congruité sémantique pour le langage, harmonique pour la musique, et la manipulation de l'allongement final dans les deux modalités linguistique et musicale. Nous avons décrit la première étude linguistique dont est issue cette comparaison avec la musique dans la section § 14.4, p. 121 (Astésano *et al*, 2004 ; Magne *et al*, 2007). Les manipulations sur le corpus musical ont été créées pour épouser les manipulations faites sur le corpus linguistique. Autrement dit, nous avons construit des phrases musicales qui présentaient des violations harmoniques ('fausses notes') sur la dernière triade de notes et/ou des violations rythmiques, en supprimant l'allongement final naturel se trouvant à la fin des phrases musicales.

L'allongement final est un phénomène musical à part entière, et se manifeste dans la performance des musiciens en tant que déviations d'un battement régulier dans la structure musicale sous-jacente (Repp, 1992). Comme pour la parole, l'allongement final a une fonction première de démarcation des frontières structurelles. Cependant, il n'intervient pas uniquement aux frontières majeures (où le compositeur a prévu un 'ritardando') mais également à la fin des sous-sections et phrases mélodiques individuelles. Le degré d'allongement tend à être proportionnel à la force structurelle de la frontière (phénomène identique en parole : Swerts & Geluykens, 1994).

Les résultats de cette étude montrent une certaine similitude dans le traitement du rythme pour le langage et la musique, puisque l'incongruité rythmique élicite dans les deux cas une composante positive autour de 600 et 700 ms. Comme pour le langage, l'incongruité harmonique dans l'expérience musicale élicite une composante négative, mais plus précoce

(N300) que l'incongruité sémantique dans le langage (N400). La principale différence réside dans le fait que ce processus est modulé par l'attention pour le langage, *i.e* que la composante positive n'est élicitée que lorsque les participants portent leur attention sur le rythme de la phrase, alors que le processus se retrouve indépendamment de l'attention dans la tâche musicale. On peut en déduire sans doute que le traitement de l'harmonie et du rythme relève d'un processus similaire ou en tous cas indissociable (la construction harmonique participe du rythme musical) et qu'ils sont traités de manière automatique, où que soit portée l'attention du participant.

Ce type de recherche doit être développé et reste bien difficile à mettre en place. De nombreux chercheurs se sont confrontés au problème épineux de la mise en correspondance entre les niveaux interprétatifs en musique (organisations harmonique, temporelle et métrique) et les niveaux interprétatifs en langage (typiquement, sémantique, syntaxe). La prosodie est visiblement transversale aux niveaux linguistiques. Mais qu'en est-il de la structure harmonique d'une pièce musicale ? Est-elle l'équivalent musical de la composante syntaxique de la langue, ou relève-t-elle de la sémantique ? Le débat reste encore très ouvert et les pistes de recherche sont encore nombreuses (voir Patel, 2010, pour une revue, et notamment la proposition de *Shared Syntactic Integration Resource Hypothesis* selon laquelle, dans une vision générative de la langue et de la musique, les deux systèmes partageraient les mêmes ressources d'intégration syntaxique).

19.2. Interactions Langage & Musique

Alors que la nature même de leur lien reste encore mystérieuse à bien des égards (Bernstein, 1976 : *the unanswered question* ; cité dans Patel 2010), il n'en demeure pas moins que Langage et Musique sont intimement liés. Ce lien verrait même ses origines dans l'évolution du langage humain : selon le paléanthropologue Jean-Jacques Hublin (2005), il semblerait que la première communication entre les hommes ait été musicale. Au fil du temps, cette 'voix chantée' se serait transformée en proto-langage chanté qui donnerait ensuite le langage. Une forme de communication intermédiaire entre le Langage moderne et la Musique pourrait donc être à l'origine des capacités communicatives de notre espèce (Darwin, 1871 ; cité par Hublin, 2005). De là à en déduire que la prosodie serait une réminiscence de ce proto-langage disparu...

Un certain nombre de recherches sont donc dédiées à tester ce lien intime entre musique et langage en neurosciences cognitives, afin d'une part, de déterminer le recouvrement des réseaux neuronaux impliqués dans les deux systèmes (architecture cérébrale, latéralisation ou

spécialisation hémisphérique des processus linguistiques et musicaux), et d'autre part de déterminer dans quelle mesure les capacités/habilités musicales et linguistiques sont transférables d'un système à l'autre. J'ai modestement pris part à certaines de ces études dans le cadre du large réseau de recherche de Mireille Besson.

Partage des réseaux neuronaux entre Langage & Musique

Un pan de ces travaux visait à étudier l'impact des dimensions phonologiques en parole chantée et l'interaction entre les niveaux phonologique et mélodique (Lidji *et al*, 2004 ; 2005). Un autre, plus développé, consistait à étudier, en utilisant la méthode d'Imagerie par Résonance Magnétique fonctionnelle (IRMf) en complément de l'EEG, les bases neurales de la voix chantée (Gordon *et al*, 2004), et les ressources cérébrales partagées par le chant, la parole et la musique (Gordon *et al*, 2010 ; Schön *et al*, 2010). Ces expériences devaient nous permettre d'une part de déterminer la nature de la relation (indépendante ou interactive) entre les dimensions musicales et linguistiques des mots chantés, et, d'autre part, de voir comment l'attention module la dynamique de cette relation. Les stimuli étaient constitués de mots trisyllabiques chantés sur une mélodie de 3 notes présentés par paire (amorces + cible). Les stimuli étaient orthogonalement manipulés selon deux axes (sémantique et mélodique) et les sujets devaient décider si la paire était constituée de deux stimuli différents ou identiques. Leur attention était attirée soit sur le traitement sémantique soit sur le traitement musical des stimuli. A la variation sémantique était associée une N400 attendue en traitement de la parole et attestée ici sur la parole chantée (voir Figure 52), et présentant la distribution centrale classique de la N400.

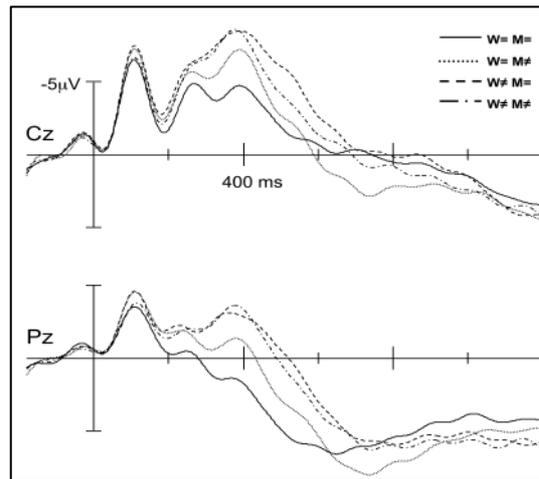


Figure 52 – Potentiels Evoqués associés aux 4 conditions expérimentales pour les électrodes Cz et Pz. Trait plein : même mot, même mélodie ; pointillés : même mot, mélodie différente ; trait discontinu: mot différent, même mélodie ; trait discontinu et pointillé: mot et mélodie différents. (Adapté de Gordon *et al*, 2010).

De plus, cette N400 demeure à la même latence et même distribution lorsque l'attention des sujets n'est pas portée sur le sens des mots mais sur la musique seule, montrant ainsi le caractère automatique du traitement du sens des mots chantés. Un résultat intéressant de ces études consiste en la présence d'une composante N400, sous les deux conditions attentionnelles, associée à la présentation d'un stimulus chanté sur une mélodie différente que l'amorce. Ce résultat pourrait être interprété comme la preuve d'un lien entre les deux systèmes. En effet, il suggère que la forme du contour tonal sur les mots chantés donné par la mélodie se rapproche d'un contour 'intonatif', de sorte que les variations des effets 'semblables à la prosodie' sur les mots chantés viendraient contrarier l'intégration et le traitement de ces mots chantés. En d'autres termes, un patron mélodique sur un mot chanté serait associé à ce mot à un niveau 'lexical' ou 'sémantique'⁵⁷. Ce résultat fait écho aux résultats que nous avons obtenus sur le traitement de l'incongruité métrique (voir § 14.4, p. 121), montrant une augmentation de l'amplitude de la N400 pour les mots (parlés) comportant un patron métrique illégal.

Il semble donc que les processus musicaux et linguistiques interagissent fortement, et ce, dès un stade précoce du traitement cognitif. La même expérience a d'ailleurs été conduite en IRMf, et montre les mêmes effets robustes d'interaction entre mots et mélodie en voix chantée (voir Figure 53).

⁵⁷ A ce sujet, qui n'a jamais remarqué que les paroles d'une chanson reviennent subitement en mémoire dès lors qu'on se souvient de la mélodie ? Comme si cette dernière était un intégrateur du sens des mots chantés.

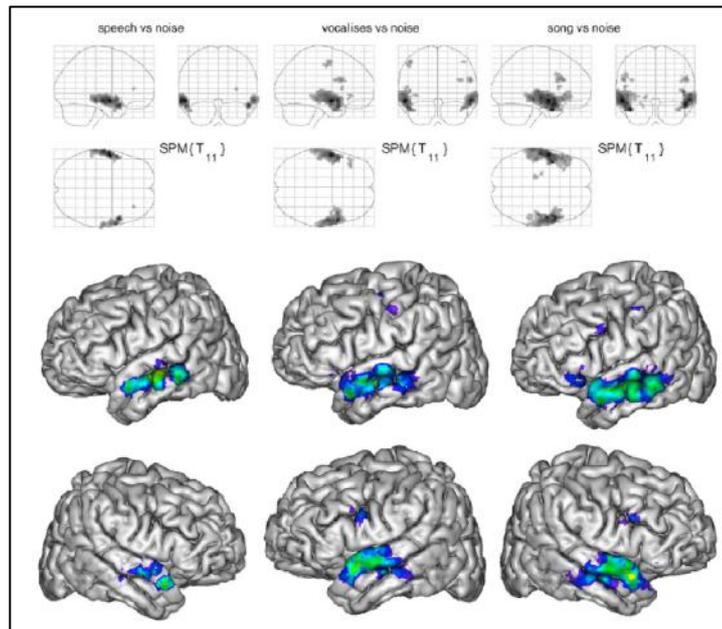


Figure 53 – Zones d’activité maximale émergeant de la soustraction entre les conditions expérimentales (*speech* = mots parlés ; *vocalises* ; *song* = mots chantés) et la condition contrôle (*noise* = bruit), indiquées sur des cerveaux transparents et des cerveaux reconstruits. (Adapté de Schön *et al.*, 2010).

L’ensemble de ces résultats tend à montrer que Langage et Musique partagent des ressources neuronales communes lors des traitements interactifs entre les dimensions phonologico-sémantique et mélodico-harmonique.

Toujours dans le même esprit, j’ai participé à une étude comportementale montrant l’impact de l’entraînement moteur à une tâche musicale sur le traitement de la parole (Cason *et al.*, 2015). Plus précisément, il s’agissait de voir dans quelle mesure une amorce métrique musicale (*beats* musicaux structurés métriquement) a une incidence sur la rapidité de traitement phonologique. Les participants étaient divisés en deux groupes : un groupe ayant bénéficié d’un entraînement audio-moteur préalable sur les amorces musicales, et un groupe n’ayant pas eu d’entraînement spécifique. Tous les participants étaient non musiciens. Les stimuli musicaux amorces présentaient 4 structures métriques (temps forts et faibles) différentes, correspondant aux 4 structures métriques des stimuli linguistiques cibles. Les amorces étaient soit métriquement congruentes soit non congruentes avec les phrases cibles suivantes (voir Figure 54). Les deux groupes de sujets devaient procéder à une tâche de détection de phonème dans la dernière syllabe du dernier mot de la phrase.

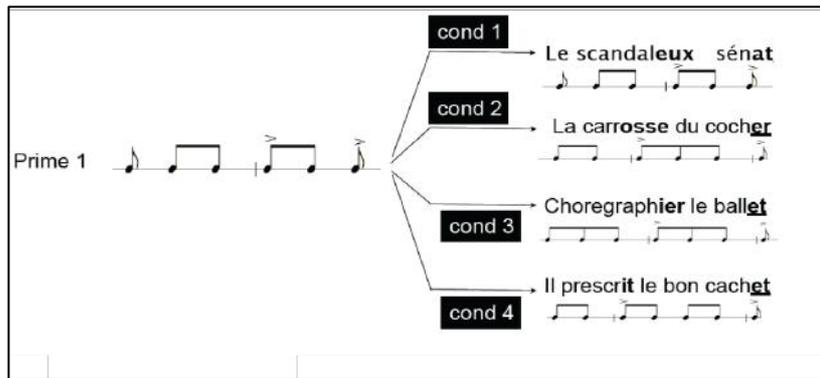


Figure 54 – Les 4 conditions expérimentales des cibles linguistiques, avec des structures métriques différentes. L’amorce 1 (Prime 1) correspond sur le plan métrique à la condition linguistique 1 mais pas aux trois autres. Le participant doit décider si le phonème qu’on lui présente visuellement entre les présentations auditives de l’amorce musicale et de la cible linguistique est bien effectivement prononcé dans la dernière syllabe (soulignée) du dernier mot. (Adaptée de Cason *et al*, 2015).

Les résultats indiquent, d’une part, qu’une amorce musicale congruente à la structure métrique de la phrase cible facilite la détection phonémique, et d’autre part, que le groupe de participants ayant bénéficié d’un entraînement audio-moteur sur les amorces musicales avaient des temps de réaction plus rapides et un meilleur taux de réponses correctes dans la tâche de détection phonémique que les participants de l’autre groupe. Le fait de pouvoir anticiper le rythme de la phrase cible (entraînement rythmique induisant une attente métrique) faciliterait donc le traitement linguistique. Notons enfin un résultat intéressant relatif au rythme linguistique : l’amorce musicale présentant la structure métrique $((xxX)(xxX)/$ induisait de meilleurs effets d’amorçage phonologique que les autres types d’amorces musicales. D’une part, il s’agit d’une structure répétitive, la répétition de structure étant favorisée dans le rythme moteur de la parole (Pasdeloup, 2005) ; d’autre part, les mots bisyllabiques (et donc les groupes accentuels déterminant + nom, trisyllabiques) sont plus fréquents en français (Wenk & Wioland, 1982 ; Delais, 1994 ; 1995).

Transferts de compétences entre Langage & Musique

Il est naturel de se demander quelles conséquences peuvent être tirées, en pratique, de ces recouvrements neuronaux entre les deux systèmes. Un tel partage de ressources neuronales doit bien avoir une incidence sur les capacités linguistiques des musiciens qui entraînent régulièrement des zones cérébrales communes par leur pratique musicale, notamment pour ce qui est de leur capacité à discriminer des sons (perception acoustique) et peut-être à les catégoriser plus rapidement. On sait qu’il existe des différences morphométriques entre les cerveaux des non musiciens et des musiciens, ces derniers présentant une plus grande densité de matière grise dans l’aire de Broca. De plus, la densité de matière grise décroît beaucoup

moins rapidement chez les musiciens avec l'âge (Sluming *et al*, 2002). La pratique régulière de la musique entraînerait donc une plasticité (le traitement des mots étant plus bilatéral chez les musiciens) et un renforcement cérébral dont on peut naturellement entrevoir l'intérêt pour les pathologies langagières notamment : a) dans le cas de la dyslexie, pour améliorer le déficit de catégorisation des sons ; b) pour les aphasies non fluentes où la dysarthrie ne permet pas d'articuler des phrases liées, il est possible que l'entraînement à chanter ce que l'on veut dire pourrait améliorer grandement leur communication.

La thématique du transfert de compétences entre les deux systèmes est au cœur des travaux de Mireille Besson (voir Besson *et al*, 2011 pour une revue). Les populations étudiées sont des adultes, des enfants en âge scolaire (apprentissage de la lecture), sains ou présentant des troubles du développement (dyslexie notamment). La comparaison va s'opérer d'une part entre participants qui ont une formation musicale solide et ceux qui ne sont pas musiciens, mais aussi sur des groupes d'adultes ou d'enfants non musiciens soumis à un entraînement musical spécifique le temps de l'expérience (longitudinale, par conséquent ; Moreno *et al*, 2004). Les résultats indiquent une nette amélioration des capacités linguistiques pour les sujets musiciens ou ayant suivi un entraînement musical qui laissent penser qu'au delà des transferts de bas niveau d'un domaine sur l'autre, on a bien affaire à des transferts de plus haut niveau permettant la construction de percepts abstraits.

Il est intéressant de noter que ces transferts peuvent s'effectuer dans l'autre sens, du Langage vers la Musique. Il est dit que les musiciens parlant une langue à tons (comme le chinois par exemple) auraient une meilleure oreille musicienne (voire qu'il y aurait une plus grande proportion de musiciens à l'oreille absolue) que les musiciens natifs d'une langue à accentuation. Les recherches utilisant l'EEG que l'on a pu entreprendre sur des non musiciens thaïlandais indiquaient en outre des composantes PE plus amples et des temps de réactions plus courts pour la détection de déviations infimes de hauteurs, de durée et d'intensité (sons purs), par rapport aux non musiciens français et finnois (Besson *et al*, 2015), laissant penser que la pratique d'une langue à tons renforce davantage les capacités de discriminations auditives des sons non linguistiques.

Depuis ces travaux effectués au contact de Mireille Besson, je prends d'ailleurs toujours soin de bien sélectionner les participants à mes expériences sur le traitement de la prosodie, qu'elles soient comportementales ou par imagerie (EEG), afin d'écartier autant que possible les musiciens. Il m'est en effet souvent arrivé de constater que certains participants avaient des temps de réponse plus rapides et des taux d'erreurs moins importants dans des tâches de perception des caractéristiques prosodiques : 9 fois sur 10, il s'agissait de musiciens !

J'ai dû m'écarter de ce thème de recherche (cognition comparée du Langage et de la Musique) depuis plusieurs années dans mon parcours, alors que c'est paradoxalement ce qui m'a conduit à la recherche ! J'étais trop occupée à développer mon pan de recherche interdisciplinaire sur les caractéristiques prosodiques du français. Mais l'avenir me réserve peut-être de pouvoir m'y replonger, avec grand bonheur ; un Master de Sciences du Langage va d'ailleurs débiter sur ce thème en septembre 2017.

20. PROSODIE ET PATHOLOGIES DE LA PAROLE⁵⁸

En intégrant l'URI Octogone-Lordat, j'ai été naturellement sollicitée pour contribuer aux recherches historiques sur les pathologies du langage, qui sont le cœur de recherche de son créateur, Jean-Luc Nespoulous. Etrangement, ce ne sont pas les recherches sur l'aphasie qui ont occupé ces collaborations, Jean-Luc nous quittant quelques années après mon arrivée pour se retirer (vraiment ??) dans sa grange. La seule collaboration effective que j'ai eue avec Jean-Luc concernait une pathologie assez similaire à l'aphasie puisque relevant d'une agnosie auditive entraînant des symptômes proches de ceux de l'aphasie. Les autres collaborations concernaient davantage la neurochirurgie et la phoniatrie. Une multitude d'étudiants en Master se sont tout de même penchés sur des questions de pathologie, que j'ai eu plaisir à encadrer sur des aspects relevant souvent de la di(y)sfluence en parole (*cf. Curriculum Vitae*). C'est donc tout naturellement que j'ai eu le plaisir de co-éditer un ouvrage en l'honneur du parcours hors norme de Jean-Luc Nespoulous, ce qui m'a permis de me rendre compte de l'étendue de son travail et de son impact sur la communauté internationale, mais aussi dans le paysage de la linguistique en France (Astésano & Jucla, 2015).

Avant de présenter rapidement ces différents travaux, qui consistaient essentiellement à construire des batteries de tests permettant de tester les différentes fonctions de la prosodie dans la communication orale, il convient de rappeler quelle est la place de ce domaine de la linguistique dans les recherches en pathologies de la parole.

Nous avons mentionné (voir § 1, p.19) la place marginale qu'a longtemps occupée l'étude de la prosodie en linguistique, dominée par la syntaxe, la sémantique et la langue écrite (la « *Cendrillon de la linguistique* » de Bolinger (1986) ; Fónagy (1989) et Di Cristo (2012). Comme le souligne Di Cristo (2012), ce sont sans doute les sciences cognitives qui ont permis à la prosodie de 'prendre sa revanche' et d'occuper de plus en plus de terrain dans les

⁵⁸ Voir Annexes : documents scientifiques 'Prosodie et Pathologies Langagières'.

descriptions linguistiques. Certains psycholinguistes reconnaissent que la prosodie intervient dans la représentation même que l'on peut avoir de l'écrit (« *prosodie silencieuse* », Fodor, 2002a) et qu'elle influence le découpage syntaxique et la représentation même que l'on peut avoir des structures syntaxiques (« *Psycholinguistics cannot escape prosody* », Fodor, 2002b). J'aurais envie de rajouter, comme un clin d'œil qu'il ne faut pas prendre pour une provocation, « *linguistics cannot escape prosody* ».

Cet engouement pour la prosodie en sciences cognitives est relativement récent (années 2000), et, comme décrit dans la section § 14, p. 114 (bases neurales de la prosodie), de nombreuses hypothèses sur la spécialisation hémisphérique du traitement de la prosodie ont été proposées, aucune n'étant pleinement satisfaisante. Même si c'est déjà un net progrès épistémologique (la prosodie a longtemps été considérée uniquement sous l'angle de l'expression des émotions), la prise en compte de la distinction entre prosodie linguistique et prosodie émotionnelle ne justifie pas forcément que l'on considère simplement pour la première un traitement dans l'hémisphère gauche et pour l'autre dans l'hémisphère droit. En réalité, il nous semble plus 'écologique' (plus proche de la réalité des fonctions multiples de la prosodie avec ses interfaces) de postuler que le traitement de la prosodie ne se fait pas *per se* mais en relation avec ses interfaces : prosodie lexicale, prosodie morphosyntaxique, prosodie et pragmatique (focus) ou structure informationnelle (contexte), prosodie émotionnelle... Dans mon optique, le rythme (comme 'organisation temporelle des prééminences', donc en lien avec le mètre ; Astésano, 2001) est l'interface avec tous les niveaux linguistiques. Nous avons donc pris soin, dans nos différentes collaborations, de placer l'accentuation au centre de nos protocoles d'évaluation, tout en n'omettant pas également les autres dimensions prosodiques (organisation temporelle et intonative, force articulatoire) pour rendre compte à la fois des aspects linguistiques et émotionnels que recouvre la prosodie. Il est naturel que les conceptions qu'ont les praticiens (orthophonistes, phoniatres, neurologues...) de la prosodie soient partielles⁵⁹. Ces conceptions sont d'ailleurs souvent retrouvées chez certains collègues linguistes, et c'est bien naturel aussi. En fait, la prosodie est très complexe à appréhender : les mêmes paramètres (intensité, *fo*, durée) vont intervenir à tous les niveaux de construction du matériel sonore, pour exprimer une multitude de fonctions, en interfaçage avec tous les niveaux de signification du contenu. Il est difficile, même entre prosodistes, de s'entendre sur les niveaux d'analyse auquel on doit se situer pour

⁵⁹ J'ai rencontré un jour une neurologue avec mon étudiante de Master travaillant sur le syndrome de l'accent étranger ; nous cherchions des patients. Lorsque nous lui avons présenté nos hypothèses, fondées sur une description du système prosodique du français, elle s'est exclamée : « *Mais il n'y a pas de prosodie en français !* »

faire sens de l'expression prosodique (voir § 5.4, p. 44). C'est toute la difficulté de notre travail : démêler les niveaux d'analyse (Hirst, 1983 ; Hirst *et al*, 2000). Un exemple intéressant nous vient de ce protocole incontournable pour l'évaluation des troubles du langage : le protocole MEC (Joanette *et al*, 2004). Si c'est un des premiers protocoles à prendre en compte la dimension linguistique de la prosodie (aux côtés de l'incontournable fonction émotionnelle), la prosodie dite linguistique dans ce protocole consiste en la production et répétition des 3 principales modalités (question, assertion, ordre). Rien n'est prévu concernant l'évaluation de la prosodie en lien avec la structure informationnelle, le focus, ou encore, sa fonction de structuration syntaxique.

La collaboration interdisciplinaire entre linguistes et praticiens est donc le gage d'un meilleur contrôle des protocoles d'évaluation clinique. Mais ce travail n'est pas aisé : nos contraintes, nos motivations et nos buts respectifs sont très différents. Il est déjà difficile pour un linguiste d'être limité à une étude de cas ; les protocoles que l'on a l'habitude de construire ne sont pas forcément adaptés. Il faut de fait s'adapter méthodologiquement, et trouver par exemple des sujets contrôles appariés en âges et niveaux d'étude⁶⁰. Trouver les patients relève souvent de la gageure ; limiter notre protocole pour l'insérer aux protocoles cliniques déjà lourds que doivent passer ces patients en est une autre ; l'adapter enfin aux situations d'enregistrement sur place dans les hôpitaux (blocs opératoires, salles d'hôpital non adaptées à la passation d'expériences linguistiques ...) complète le tableau.

1.1 La prosodie à l'aide de l'accès au sens dans les cas d'aphasie développementale

Le premier protocole de test de la prosodie que nous avons mis en place était destiné à étudier le rôle de la prosodie dans un cas spécifique d'aphasie développementale, le syndrome de Landau-Kleffner (SLK : Kim-Dufor *et al*, 2012 ; 2010 ; 2009a ; 2009b). Ce type de pathologie se manifeste par des troubles langagiers comparables aux troubles du spectre autistique. En réalité, il s'agit de cas très rares d'épilepsie induisant des troubles de la perception, *i.e.* une incapacité à traiter les informations phonologiques et sémantiques (surdité lexicale ou agnosie auditive). Jusque-là, les rares études sur ces cas d'aphasie développementale testaient le vocabulaire dans des tâches de mémoire phonologique à court terme. Etant donné que la prosodie est le socle de l'acquisition du langage (Konopczynski, 1986 ; de Boisson-Bardies, 1996), il nous semblait important de tester l'utilisation des indices

⁶⁰ Une contrainte que nous n'avons pas en tant que chercheurs à l'université travaillant sur le sujet tout-venant. Nos cohortes d'étudiants nous permettent de trouver notre quarantaine de participants par expérience. Cela constitue d'ailleurs un biais méthodologique puisque nos travaux ne sont basés que sur une population homogène sur le plan du niveau d'étude et socio-culturel, peu représentative de l'ensemble de la population.

prosodiques par ces enfants au développement atypique. Par ailleurs, la prosodie affective est utilisée dans le langage adressé à l'enfant et constitue de fait un élément moteur de son apprentissage en favorisant l'acquisition précoce du système intonatif de la langue maternelle (Di Cristo, 2012). Nous postulons donc que la prosodie affective serait mieux préservée que la prosodie linguistique, plus intimement liée aux aspects lexicaux du langage.

Nous avons donc créé un protocole testant d'un côté la prosodie linguistique via la manipulation du rythme, élément fondamental de la prosodie, en manipulant la durée de chaque syllabe ramenée à la durée moyenne des syllabes de la phrase (condition R-), l'autre moitié des phrases conservant les informations d'alternance des syllabes accentuées (condition R+); d'un autre côté, la prosodie affective consistait en des phrases neutres (condition N) pour moitié et des phrases dites sur le ton de la colère (condition A, pour 'anger'). Enfin, les phrases dans chacune de ces conditions étaient soit désémantisées (condition S- : transformation du signal par sinusoïde) soit elles conservaient leur sens (condition S+).

Les participants (3 cas d'enfants SLK ; 22 enfants contrôles et 12 adultes contrôles) avaient pour tâche d'associer la bonne image (choix de 4 images pour les conditions de prosodie linguistique ; 2 images pour la prosodie affective) au stimulus auditif. Nous renvoyons le lecteur à Kim-Dufor *et al* (2012) pour plus de détails sur l'expérience.

Les principaux résultats sont les suivants : les enfants à développement typique utilisent moins les informations prosodiques que les enfants SLK, et ont plus de difficulté avec les phrases désémantisées. Les SLK s'appuient donc davantage sur les informations prosodiques pour pallier leur déficit sémantique. Pour un cas de SLK sur les 3, la prosodie affective est mieux reconnue que pour les autres enfants à développement typique, nous incitant à formuler l'hypothèse que, pour ce type d'aphasie, la prosodie affective pourrait être utilisée en remédiation pour renforcer l'accès au sens. Il s'agit bien évidemment de résultats très partiels qui demandent à être confirmés par des études plus étendues. Ce type de pathologie étant tout de même très rares, il serait bon d'adapter ce protocole à d'autres cas d'aphasie et de troubles du spectre autistique.

1.2 Spécialisation hémisphérique du traitement de la prosodie chez des patients atteints de tumeurs cérébrales envahissantes⁶¹

Les travaux initiaux réalisés en aphasiologie ont révélé l'importance de l'hémisphère gauche (HG) dans le traitement de la production (aire de Broca) et de la compréhension (aire de Wernicke) du langage. Avec le développement de cette approche localisationniste et l'avènement de l'approche modulariste en psycholinguistique, le rôle prépondérant de l'HG est confirmé notamment dans le traitement des aspects phonologiques, lexico-sémantiques et syntaxiques (pour une revue de littérature, voir Aura, 2012). Or, d'autres travaux ont montré que des lésions droites pouvaient provoquer des altérations aux niveaux pragmatique, discursif et prosodique (Joanette *et al*, 1990). Cependant, même si la prosodie tend à être prise en compte dans les recherches en pathologies du langage, elle ne l'est que partiellement, *i.e.* sous l'angle de ses fonctions paralinguistiques (prosodie émotionnelle).

Certaines tumeurs cérébrales (tumeurs envahissantes, ou gliomes de bas grade) sont opérées en chirurgie éveillée lorsqu'elles sont localisées dans HG, afin que le neurochirurgien puisse contrôler en temps réel si les excrèses qu'il doit pratiquer dans le cortex pour évacuer la tumeur touchent des zones impliquant le contrôle moteur et les zones langagières du patient. Traditionnellement, ces protocoles éveillés ne sont pratiqués que sur les patients atteints de tumeurs dans HG, alors que ceux dont les tumeurs se trouvent dans HD sont opérés sous anesthésie, sans contrôle des conséquences de l'exérèse. Or, les orthophonistes qui suivent ces derniers patients après opération notent 9 fois sur 10 une dégradation des facultés langagières, notamment dans le traitement des inférences pragmatiques, et souvent une certaine dysprosodie. L'une d'elle (K. Prodhomme, ancienne doctorante de Jean-Luc Nespoulous), en accord avec le neurochirurgien V. Lubrano du CHU de Toulouse, a fait appel à moi pour mettre au point un protocole prosodique complet, permettant à la fois de tester les patients HG *et* ceux dont la tumeur se trouve dans HD. Le bénéfice était double : pour les patients HG, développer des tâches linguistiques plus fines que les traditionnelles tâches de dénomination (qui testent seulement le versant lexical et sémantique du langage) ; permettre aux patients HD de pouvoir bénéficier également de la neurochirurgie éveillée et limiter ainsi les troubles langagiers post-opératoires. Sachant que le traitement des aspects linguistiques est seulement *spécialisé* dans un hémisphère ou l'autre, et, qu'en réalité, la plupart des niveaux langagiers

⁶¹ Nous avons un temps envisagé de faire passer ce protocole aux patients atteints de tumeurs focales (gliomes de haut grade) avec lesquels les conclusions scientifiques auraient été plus évidentes à tirer pour la localisation des différentes zones cérébrales impliquées dans le traitement des différents niveaux linguistiques. L'idée a cependant vite été abandonnée pour des raisons éthiques, ce genre de tumeurs ayant un taux de morbidité extrêmement élevé. Le bénéfice-patient pèse en effet lourdement dans la balance dans les protocoles de chirurgie.

sont traités de manière *bilatérale*, ce protocole présente en outre l'avantage de tester les dommages contra-latéraux potentiels de ce type de chirurgie.

Ce travail a d'abord été entrepris dans le cadre d'un Master 1 & 2 (Jiaa, 2009-2011) et d'un mémoire d'orthophonie (Le Cam, 2010 ; Le Cam *et al.*, 2009) dédiés à la mise en place de ce protocole, et sa normalisation sur des sujets contrôles. Les passations sur des patients ($n = 2$) utilisant ce protocole ont été réalisées dans le cadre de la thèse de K. Aura (2012)⁶². Les tests ne concernaient que le versant compréhension des tâches prosodiques. Deux tâches ont été mises au point afin de pouvoir tester aussi bien les patients HG que HD, avec l'hypothèse que les caractéristiques prosodiques permettant d'actualiser ces fonctions linguistiques sont traités préférentiellement dans le même hémisphère que le-dit niveau linguistique: une tâche de désambiguïsation syntaxique (tâche dite 'Prosodie-Syntaxe') permettant de tester les patients HG ; et une tâche de compréhension de la fonction pragmatique de la prosodie (tâche dite 'Prosodie-Pragmatique') permettant de tester les patients HD (Aura *et al.*, 2011).

Tâche 'Prosodie-Syntaxe'

Nous nous sommes fondés sur les travaux d'Astésano *et al.* (2007) qui montrent que IA est plus systématiquement réalisé comme marqueur du niveau Syntagme Phonologique (*ap*) que FA, dans la production (et supposément l'interprétation) de phrases syntaxiquement ambiguës (voir § 12.2, p. 88). Il nous fallait cependant des items-cibles imageables puisque les participants allaient devoir reconnaître un sens ou l'autre sur la base d'une image de la phrase entendue. Le plus lourd travail du Master de M. Jiaa a donc consisté en la création de ce corpus, en concertation avec deux étudiantes en arts plastiques (rémunérées) pour la création des dessins.

Nous avons construit un corpus initial de 73 scripts pour chaque condition (un seul nom N et leurs homologues deux noms NN). Afin de s'adapter aux sujets pathologiques, nous avons fait passer un pré-test auprès de sujets sains de langue maternelle française, entre 25 et 63 ans, qui devaient juger de leur compréhensibilité et leur acceptabilité sémantique. A l'issue de ce pré-test, seuls 46 scripts ont été conservés. Deux locuteurs (un homme et une femme) ont enregistré ces scripts sous les deux modalités syntaxiques ; chaque script a été répété trois fois, donnant lieu à un total de 552 stimuli. Les trois chercheurs menant cette expérience (K. Aura,

⁶² Ce lourd travail de constitution de protocole n'ayant malheureusement pas été publié dans une revue internationale, nous renvoyons le lecteur à la version de la thèse de K. Aura téléchargeable sur ResearchGate.

M. Jiaa et moi-même) ont alors sélectionné les meilleures instances de ces stimuli, en prenant soin d'équilibrer le nombre de stimuli prononcés par les locuteurs de chaque sexe pour instaurer une certaine variabilité du corpus et se rapprocher ainsi de conditions plus naturelles. Les illustratrices ont alors dessiné chacun des scripts retenus mais 4 scripts trop difficiles à illustrer ont dû être éliminés. Enfin, un nouveau pré-test a été passé auprès de 25 sujets de 25 à 61 ans pour sélectionner, sur la base de leur capacité à choisir la bonne illustration en fonction du stimulus audio, les scripts les plus aisément reconnus pour cette tâche. A l'issue de ce test, 20 stimuli ont été retenus (4 pour la phase d'entraînement ; 16 pour le test en lui-même : 8 stimuli N et 8 stimuli NN).

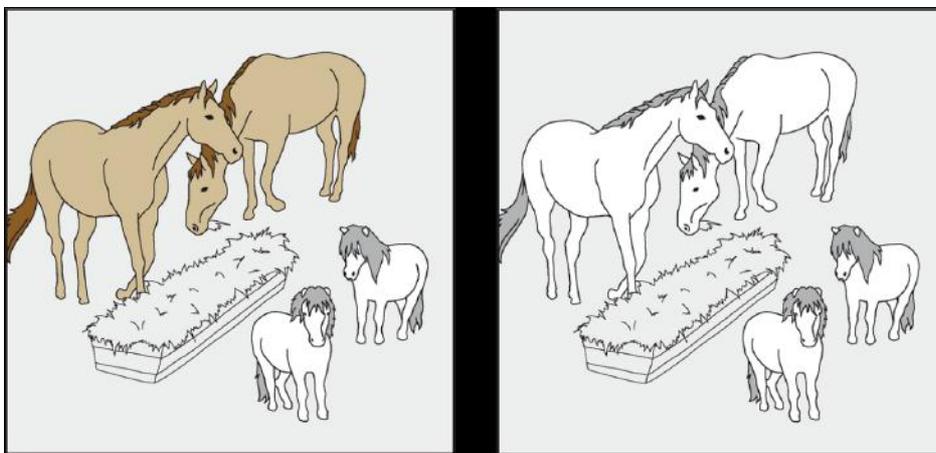


Figure 55 – Exemple de planche de dessins illustrant les deux conditions syntaxiques pour le script ‘les chevaux et les poneys blancs’ (Condition N ‘les chevaux// et les poneys blancs’ à gauche ; Condition NN ‘les chevaux et les poneys // blancs’ à droite). Les participants doivent sélectionner l’image correspondant au stimulus auditif à l’aide d’un boîtier-réponse.

Tâche ‘Prosodie-Pragmatique’

Nous adaptons ici le corpus initialement construit par Magne et *al* (2005a)⁶³ (voir § 14.3, p. 119, pour une présentation détaillée du corpus). Ce corpus permet de tester la compréhension du focus, en faisant écouter un couple de phrase Question + Réponse, induisant une attente pragmatique sur la portée du focus. Par exemple, à la question : Q ‘A qui Pierre a-t-il offert un foulard : à sa femme ou à sa grand-mère ?’, la réponse attendue est celle pour laquelle le focus se porte sur l’un des compléments d’objet indirect, comme par exemple: R ‘Pierre a offert un foulard à sa GRAND-MERE’, alors qu’une réponse où le focus porte sur le complément d’objet direct R ‘Pierre a offert un FOULARD à sa grand-mère’ serait incongrue. Cette dernière réponse serait congruente avec la question : Q ‘Qu’a offert Pierre à sa grand-mère ?’.

⁶³ Nous profitons de cette occasion pour remercier encore une fois Anne Lacheret-Dujour et Mireille Besson de nous avoir autorisées à réutiliser en partie ce corpus.

Le corpus initial (Magne *et al.*, 2005a), composé de 240 dialogues questions-réponses, avait déjà été validé par les auteurs. Nous avons cependant procédé à un tri afin de réduire le nombre d'énoncés à 80. Afin de garder la même cohérence que pour la tâche Prosodie-Syntaxe (variabilité), nous avons fait réenregistrer les questions et les réponses par un locuteur masculin, et avons associé la voix masculine avec la voix féminine du corpus originel sur les 80 couples Questions-Réponses, en prenant soin d'avoir autant des deux voix sur les Questions que sur les Réponses.

Nous avons soumis les 80 dialogues à un groupe de 48 sujets (de 21 à 60 ans, sans trouble, langue maternelle française), qui devait choisir si le couple Question/Réponse était congruent ou non, en cliquant soit sur la touche  d'un boîtier si la Réponse leur semblait congruente avec la Question, soit sur la touche  lorsqu'elle ne l'était pas. Sur la base de ce pré-test, 24 des stimuli les plus facilement compris ont été sélectionnés (4 stimuli d'entraînements et 20 stimuli tests comportant 10 couples congruents et 10 couples incongrus).

Les deux tâches ont été réalisées sous le logiciel PercEVAL (Ghio *et al.*, 2003 ; André *et al.*, 2003). Nous présentons ici seulement les résultats sur les participants-contrôles, et renvoyons le lecteur à l'interprétation des résultats pour les deux sujets pathologiques (étude de cas) à Aura (2012). Une fois encore, l'intérêt pour nous est de valider des protocoles à destination des études en pathologies langagières, de les concevoir, en espérant qu'ils serviront à l'étude de telle ou telle pathologie langagière. Il s'agit ici d'établir les seuils de référence qui seront utilisés pour interpréter les résultats des patients. Les deux tâches ne révèlent pas de difficulté particulière pour les participants-contrôles, avec près de 70% de taux d'identification correcte des stimuli audio sur la base des dessins pour la tâche Prosodie-Syntaxe, et près de 82% pour la tâche Prosodie-Pragmatique. Cependant, cette dernière apparaît beaucoup plus aisée que la tâche Prosodie-Syntaxe, ce que révèlent les temps de réaction moyens, considérablement plus élevés pour la tâche Prosodie-Syntaxe (environ 2 secondes en moyenne contre 1 seconde pour la tâche Prosodie-Pragmatique). Il convient donc de tenir compte de cette difficulté plus importante pour la tâche Prosodie-Syntaxe lors de la passation et l'évaluation des patients en situation très inconfortable de chirurgie éveillée. Un entraînement pré-opératoire est d'ailleurs prévu, et les patients sont ensuite réévalués après l'opération (Aura, 2012). Globalement, nous observons un effet de l'âge des participants dans les deux tâches sur les scores de réussite, les participants âgés de moins de 50 ans ayant de meilleurs scores que les participants plus âgés. Nous observons un effet du niveau scolaire uniquement pour la tâche Prosodie-Pragmatique,

les participants avec un niveau supérieur au bac présentant un meilleur taux de réussite. Le niveau d'étude ne transparaît pas pour la tâche Prosodie-Syntaxe car elle est sans doute plus difficile pour l'ensemble des sujets.

Tournons-nous maintenant vers le dernier type de protocole, développé dans le cadre du projet 'C2SI : *Carcinologic Speech Severity Index*' porté par V. Woisard (CHU Toulouse) et que je coordonne localement pour Octogone-Lordat. Ce protocole est une déclinaison du protocole précédent, développé pour s'adapter aux contraintes et déficits spécifiques des patients relevant de la phoniatry.

20.1. Conservation des fonctions prosodiques après traitement des cancers de la cavité buccale et du pharynx

Le projet C2SI (*Carcinologic Speech Severity Index*) a pour objectif d'obtenir une mesure de l'impact des traitements des cancers de la cavité buccale et du pharynx sur la production de la parole. Cette mesure permet d'évaluer un indice de sévérité des troubles de la production de la parole, et est envisagée à la fois par le biais de méthodes perceptives et par un traitement automatique de la parole. Plus que la compréhensibilité, c'est d'abord l'intelligibilité qu'il faut être capable de quantifier. En effet, les tests de compréhensibilité prennent en compte le contexte dans lequel les participants interagissent et également tous les moyens paralinguistiques que les patients vont utiliser pour se faire mieux comprendre (gestes et mimiques notamment). Or, ce projet s'intéresse à la quantification de la sévérité des troubles de la parole, dans le but, à terme, d'automatiser l'indexation de sévérité des pathologies de la parole. Pour éviter tout processus de restauration phonémique dû à la connaissance du lexique par le participant, une partie des tests élaborés au LPL (Aix-en-Provence) consiste en la création d'une batterie de logatomes ou pseudo-mots, prononcés par les patients, et que les auditeurs doivent décoder (processus ascendant de décodage acoustico-phonétique) avant de les transcrire (Ghio *et al*, 2016 ; 2017).

La perte d'intelligibilité est une plainte très importante chez les patients présentant un trouble de la parole car elle nuit considérablement à leur communication quotidienne et a un impact délétère sur leur qualité de vie (Woisard *et al*, 2013). L'intelligibilité peut se définir comme « *le degré de précision avec lequel un message est compris par un individu* » (Ghio *et al*, 2016 : 590).

L'altération de la production de la parole chez les patients souffrant de cancers de la cavité buccale et/ou pharyngale touche évidemment le niveau segmental, puisque le traitement

carcinologique peut mutiler ces organes, mais également les aspects prosodiques. Les niveaux segmental et prosodique interagissent en effet, le niveau segmental pouvant ‘altérer’ ou en tous cas modifier les composantes prosodiques (tonales notamment ; Welby & Niebuhr, 2016) et inversement, la prosodie a une incidence sur la réalisation des segments (débit de parole, accentuation etc...). Dans la partie du projet développée à Octogone-Lordat, il s’agit dans un premier temps de voir quel impact les différents types de lésions post-traitements carcinologiques ont sur l’intelligibilité de la prosodie (Nocaudie *et al*, 2017). Les tâches de prosodie nous permettent de tester 3 fonctions organisationnelles primordiales de la prosodie (Di Cristo, 2012): la fonction structurale, la fonction de modalité, et la fonction pragmatique.

Dans un second temps, il s’agira également de voir l’impact de l’entraînement prosodique sur l’amélioration et la récupération potentielle des aspects segmentaux altérés. Nous participons également avec le LPL à la passation de tests de perception dans une tâche dite de SVT (Sentence Verification Task ; Pisoni *et al*, 1987) permettant de manière assez similaire à la tâche de logatomes d’éviter l’impact trop important du contexte dans l’intelligibilité, puisque dans la moitié des phrases, le dernier mot est totalement inadéquat par rapport au contexte. Dans ce dernier cas, l’auditeur ne peut pas se baser sur le contexte pour décider si la portée sémantique de la phrase est vraie ou fausse. Rappelons que, dans cette tâche, l’auditeur a une tâche binaire à effectuer, *i.e.* cliquer sur le bouton ‘Vrai’ (ex : « *L’essence est un excellent carburant* ») ou ‘Faux’ (ex : « *L’essence est un excellent apéritif* »).

Le recueil de données se fait au CHU Rangueil auprès de patients carcinologiques sur la base des tests construits au LPL et à Octogone-Lordat (tâches prosodiques), puis les données sont recueillies pour être ensuite évaluées par un panel d’auditeurs. A Octogone-Lordat, nous faisons passer les 3 tâches de prosodie des patients, ainsi que la tâche de SVT (voir Figure 56), dans la plateforme de mesures comportementales liées à la perception des données langagières (plateforme CCU⁶⁴ Cloé : Responsable Hélène Giraudo).

⁶⁴ CCU : Cognition, Comportement, Usages

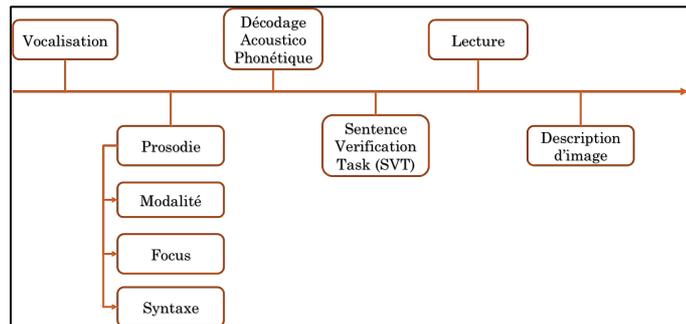


Figure 56 – Recueil de données du corpus C2SI sur les patients atteints de cancers de la cavité buccale et/ou pharyngale. La vocalisation consiste en la production d'un /a/ tenu, test très utilisé en phoniatry à l'instar de la Lecture ('la chèvre de M. Seguin') et la description d'images. Le test de Décodage Acoustico-Phonétique est construit et réalisé au LPL, alors que la SVT a été conçue au LPL mais partiellement passée à Octogone-Lordat. La passation des 3 tâches de prosodie incombe totalement à Octogone-Lordat.

Le travail de recueil des données d'intelligibilité (jugements par des auditeurs des différentes tâches enregistrées par les patients recueillis avec le logiciel PercEVAL) et l'analyse des résultats sont en cours. Les résultats (partiels) sur 30 patients évalués par 42 auditeurs naïfs ont été présentés aux Journées de Phonétique Clinique à Paris en mai 2017. Voici une rapide présentation des 3 tâches 'prosodie' développées à Octogone-Lordat : les tâches 'prosodie-syntaxe' (désambiguïstation syntaxique par les indices prosodiques) et 'prosodie-pragmatique' (traitement du focus dans les couples Question-Réponse) ont été adaptées d'Aura (2012 ; voir 1.2, p. 159). Les deux tâches comptent respectivement 26 et 20 items, lus par 30 locuteurs (25 patients, 5 témoins) soit 658 et 510 items exploitables pour le jugement d'intelligibilité. La tâche de modalité a été créée pour ce projet ; elle compte 30 items répartis sur les modalités Ordre, Question, Assertion, soit environ 764 items exploitables pour le jugement perceptif. La SVT compte 50 phrases, donc 1153 items pour la tâche de jugement. Chaque auditeur ($n = 42$; rémunérés) procède à 3 écoutes de chaque item, soit en tout 9255 jugements et temps de réponse sur les 4 tâches. Nous émettons l'hypothèse que les patients dont la source prosodique (larynx) est théoriquement non affectée, auront une bonne conservation des *fonctions* prosodiques mais l'altération du contenu segmental induira tout de même un temps de réponse plus long pour les patients lors des jugements d'intelligibilité.

Les résultats sont présentés dans la Figure 57 ci-dessous.

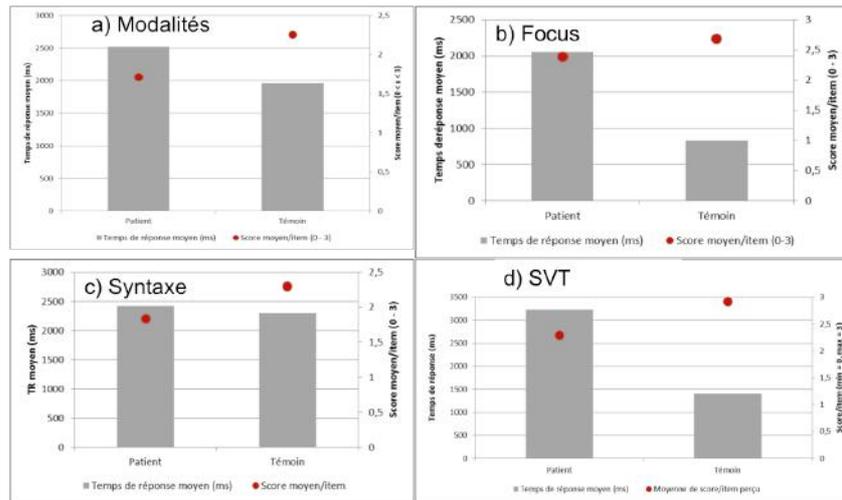


Figure 57 – Résultats des jugements d’intelligibilité (42 auditeurs naïfs ; 25 locuteurs patients et 5 témoins) dans les 4 tâches (3 tâches Prosodie : a) Modalités ; b) Focus ; c) Syntaxe ; et d) SVT). (Adapté de Nocaudie *et al*, 2017).

Nos hypothèses sont (partiellement) validées : les auditeurs jugent plus rapidement et avec de meilleurs scores les témoins que les patients. Les résultats sont cependant surprenants pour la tâche de Modalités (a), tâche la plus couramment utilisée par les praticiens hospitaliers car sensément la plus aisée ; elle montre en effet des temps de réaction très longs et un score de reconnaissance faible, y compris pour les participants Témoins, comparables à ceux de la tâche Syntaxe (c) qui est, sans surprise, plus difficile. Les tâches de Focus (b) et SVT (d) sont les plus discriminantes entre les patients et les témoins, particulièrement la SVT pour laquelle l’information segmentale est encore plus primordiale que pour le Focus.

Ces résultats ont ceci d’intéressant qu’ils nous permettront de sélectionner les tâches prosodiques à intégrer dans les protocoles finaux de mesure de l’index de sévérité carcinologique. Les passations se poursuivent actuellement pour atteindre 90 sujets pathologiques et environ 180 auditeurs-juges. Les suites de ce travail vont dans deux directions : d’une part, nous allons procéder à la mesure objective et subjective de la fluence des patients, comme indicateur sans doute plus informatif de la sévérité du trouble parolier ; d’autre part, nous allons corrélérer les présents résultats de jugements avec les données acoustiques dans les 4 tâches, notamment en ce qui concerne les contours intonatifs de la tâche Modalité dont les résultats perceptifs sont surprenants.

La soumission d’une ANR sur ce projet est en cours.

21. PROSODIE ET APPRENTISSAGE D'UNE LANGUE SECONDE⁶⁵

Un dernier volet de ces variations autour de la prosodie concerne une autre thématique qui se développe de plus en plus à Octogone-Lordat ces dernières années: l'acquisition/apprentissage d'une langue seconde (Axe 2 de l'URI).

21.1. Impact de l'entraînement prosodique et MVT sur la fluence en langue étrangère

Un des mes premiers co-encadrement avec M. Billières sur ce thème concernait la mise au point d'un protocole expérimental visant à comparer l'efficacité de la méthode verbo-tonale (pour une revue de cette méthode, voir Billières, 2005), dans une comparaison longitudinale avec un entraînement phonétique plus traditionnel, relevant de la méthode dite 'articulatoire' (Alazard, 2013). Il s'agissait de la toute première tentative d'investigation expérimentale et contrôlée testant l'efficacité de cette méthode de remédiation phonétique, jusque là cantonnée à une pratique très marginale et empirique.

Dans l'enseignement du français langue étrangère, l'apprentissage des aspects phonétiques est largement laissé de côté. On impute cette lacune dans les manuels et dans la pratique à la difficulté qu'auraient les enseignants non formés à la phonétique à s'approprier les connaissances de cette discipline... Ne mentionnons même pas la place de la prosodie dans les manuels, quasi-inexistante et (là encore) parent pauvre de la didactique des langues, alors que, nous pensons l'avoir illustré à de multiples reprises dans ce document, la prosodie est à la base même de la maîtrise langagière, depuis le développement jusqu'à la remédiation linguistique. Ceci conduit au paradoxe selon lequel, d'après le *Cadre Européen Commun de Références pour les Langues (CECRL)*, un apprenant de niveau B2 (Avancé) aura acquis «une prononciation et une intonation claires et naturelles », et ce, sans entraînement spécifique (par inoculation mystérieuse ?).

Ce travail s'inscrit donc dans une visée didactique assez confidentielle, élaborée par Guberina à partir des années 50, et défendue par quelques rares spécialistes au niveau européen (dont M. Billières, en France, est l'un des plus illustres représentants), selon laquelle l'apprentissage d'une L2 doit être centré au départ sur le rythme et l'intonation. Cet enseignement des caractéristiques prosodiques de la langue-cible sera d'autant plus bénéfique s'il est proposé dès le début de l'apprentissage. La priorité à l'enseignement de la prosodie constitue un point fort de la méthode verbo-tonale d'intégration phonétique (MVT). Dans ce courant didactique, l'écrit est théoriquement introduit après plusieurs dizaines d'heures (au

⁶⁵ Voir Annexes : documents scientifiques 'Prosodie et Apprentissage d'une L2'.

moins 6 semaines de cours) exclusivement consacrées à l'acquisition de la compétence orale⁶⁶ et à l'amélioration de la production des sonorités parolières grâce à une « rééducation de l'oreille » obtenue par les procédés de remédiation employés en MVT (Billières, 2005). Le rythme et l'intonation font l'objet d'une sensibilisation systématique (Billières, 2008). Les exercices de pratique orale prévoient en effet des répétitions de logatomes ('dadada') sensées imprégner les apprenants de l'ossature rythmique de la langue-cible. Le travail sur l'intonation⁶⁷ permet d'éclaircir ou d'assombrir le timbre des sons prononcés par l'apprenant, afin de lui faire atteindre la cible segmentale (par exemple, si un apprenant espagnol ne parvient pas à prononcer le phonème /y/ et tend vers un /u/ appartenant à son système, l'enseignant mettra le son cible /y/ sur la dernière syllabe d'une modalité ascendante afin d'éclaircir le timbre du /u/ de l'apprenant vers le /y/).

Le travail d'Alazard (2013) consistait en outre à mesurer l'impact de l'entraînement verbo-tonal sur les capacités de décodage de l'écrit des apprenants. L'hypothèse sous-tendant ce travail est celle d'un transfert prosodique entre l'oral et l'écrit en langue étrangère, sur la base de ce qui a pu être observé en langue maternelle (Fodor, 2002 a&b). Ainsi, nous postulons que la maîtrise de la prosodie est indispensable en parole spontanée comme en parole lue pour assurer l'encodage et le décodage de la langue étrangère.

Nous avons testé cette hypothèse sur un groupe linguistique homogène (anglophones), la difficulté ayant été de trouver (et garder !) des apprenants qui seraient disponibles sur une période de 8 semaines, à raison de 2 heures de cours par semaine. Seuls 8 apprenants anglophones ont finalement suivi l'ensemble de l'étude : 4 apprenants de niveau élémentaire (niveau A) et 4 apprenants de niveau indépendant (niveau B), équitablement répartis en quatre groupes. Deux groupes contrôles (niveau A et niveau B) ont suivi des cours de prononciation selon la méthode articulatoire, et deux groupes tests (niveau A et niveau B) ont suivi des séances de correction phonétique via la MVT. Les apprenants ont été testés en lecture oralisée avant, pendant et en fin de formation. Les analyses phonétiques portant sur leurs 3 lectures étaient destinées à quantifier la fluence en LE : durée totale des productions, nombre, durée et types de pauses, d'hésitations, équilibrage syllabe accentuées et inaccentuées (taille des groupes rythmiques) et densité accentuelle.

⁶⁶ Se rapprochant ainsi des étapes ontogénétiques du développement langagier où la prosodie est acquise en premier, avant la mise en place de (et comme socle à) l'acquisition du lexique et de la syntaxe.

⁶⁷ Je pense qu'il s'agit en réalité d'un travail sur l'accentuation et sur les contours tonaux locaux au niveau des mots isolés, utilisant parfois les contours globaux (contours intonatifs) dans le jeu des différentes modalités... Nous nous sommes souvent 'empoignés' en toute amitié à ce sujet avec M. Billières qui travaille sur le rythme et l'accentuation à son corps défendant.

Ces analyses quantitatives ont ensuite été mises en regard d'une évaluation auditive de la fluence des apprenants par 12 auditeurs natifs sur une échelle de magnitude de 1 (non fluent) à -7 (très fluent). Il était demandé aux auditeurs de justifier leurs choix afin de déterminer les critères les plus prégnants dans l'appréciation de la fluence en LE.

Les résultats (Alazard *et al*, 2009 ; 2010 ; 2011 ; 2012a,b,c) montrent très clairement que les apprenants ayant suivi la MVT sont plus fluents (nombre et durée des pauses moins importants, taille des groupes rythmiques plus homogènes ...), et que les auditeurs jugent leur lecture plus fluente que les apprenants ayant suivi un entraînement 'classique' de correction phonétique (voir Figure 58). Cette tendance est particulièrement prégnante pour les apprenants de niveau inférieur (niveau A), où une amélioration très nette de leur fluence a été observée avant et après entraînement, justifiant ainsi le précepte de la MVT selon lequel l'apprentissage de la prosodie doit se faire au tout début de l'apprentissage, sans passer par l'écrit dans un premier temps.

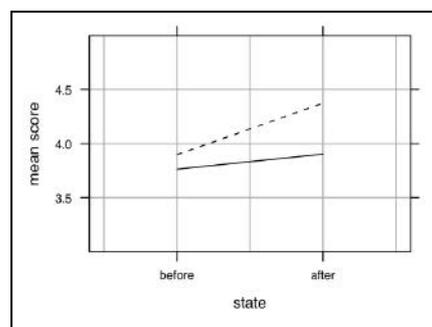


Figure 58 – Moyenne des scores de jugement de fluence des auditeurs natifs (échelle de 1 à 7), sur les lectures effectuées avant (before) entraînement et après (after) entraînement. La ligne pleine concerne les apprenants du groupe de correction phonétique 'classique', la ligne en pointillée représente les apprenants du groupe MVT. Les niveaux A et B sont mélangés dans cette figure.

Ce travail a donné lieu également à la création de deux ressources très innovantes et précieuses dans ce domaine de l'enseignement du FLE: une ressource pédagogique à l'Université Ouverte des Humanités (Billières *et al*, 2013, en ligne : <http://w3.uohprod.univ-tlse2.fr/UOH-PHONETIQUE-FLE/>), et un corpus multimodal (audiovisuel) de cours filmés de correction phonétique à la disposition de la communauté scientifique et présenté à LREC (Alazard *et al*, 2012d).

21.2. Mesure de la capacité d'imitation prosodique en langue maternelle et langue étrangère

Une deuxième collaboration autour de la prosodie et de l'apprentissage d'une langue étrangère concernait un sujet tout à fait passionnant, à la croisée des disciplines en sciences cognitives : l'utilisation d'un processus automatique et irrépressible chez l'humain

(l'imitation) à des fins de remédiation phonétique en langue étrangère (Nocaudie, 2016). L'idée de départ était de 'profiter' de cette tendance naturelle à l'imitation qu'ont les locuteurs pour adapter leur discours à leur interlocuteur (convergence phonétique) pour la mettre au service d'une méthode d'enseignement de la phonétique, en l'occurrence la MVT. Cette idée (brillante) d'O. Nocaudie consistait à tester expérimentalement ce que la MVT utilise 'naïvement' et sans contrôle dans sa pratique de classe. Il y a en effet une sorte de boucle de répétition/imitation entre les productions du professeur et l'apprenant, au niveau segmental et prosodique. Outre la voix, le corps est également très sollicité par ce mécanisme de répétition/imitation en classe de langue.

Les travaux sur la convergence phonétique sont relativement rares de manière générale, mais évidemment encore plus en ce qui concerne l'apprentissage d'une L2 (voir cependant les travaux de Lewandowski, 2012). Le premier travail d'O. Nocaudie, en M2, a été d'utiliser une petite partie du corpus d'Edinburgh (8 scripts de phrases syntaxiquement ambiguës ; voir § 12.2, p. 88) pour vérifier, dans un premier temps, que les auditeurs/imitateurs *francophones* étaient en effet capables de répliquer la structuration prosodique entendue (IA, FA, pauses), *i.e.* leur capacité à percevoir/traiter des 'détails phonétiques fins' (*Fine phonetic details*, Nguyen *et al*, 2009). A la différence des rares autres études sur le sujet (voir par exemple Michelas & Nguyen, 2011), nous proposons à un panel de 8 auditeurs francophones naïfs 3 tâches différentes mesurant différents degrés d'imitation : de la simple répétition de structures syntaxiques à l'exagération (très utilisée en MVT) en passant par l'imitation traditionnelle. Cette étude pilote (Nocaudie & Astésano, 2012) a montré que les auditeurs sont capables de reproduire des détails phonétiques fins (durées de FA et durée des pauses) dès le niveau de la tâche d'imitation ; la tâche d'exagération n'implique pas forcément une meilleure approximation de ces détails phonétiques, sauf quand les corrélats acoustiques sont contradictoires (insertion d'une pause non attendue dans le contexte syntaxique et raccourcissement de FA), les auditeurs/imitateurs reproduisant alors très précisément ces indices uniquement dans cette tâche d'exagération. Ce résultat est encourageant pour la pratique de la MVT puisque la répétition d'éléments phonétiques 'déformés' est à l'œuvre pour permettre aux apprenants d'atteindre *in fine* la cible phonémique ou prosodique exacte. C'est un premier pas pour montrer expérimentalement que l'imitation/exagération est un procédé pertinent que l'on peut entraîner chez les naïfs.

Cette lourde étude pilote n'a pas été poursuivie dans la thèse de Nocaudie (ni dans une publication de revue) car ce dernier était insatisfait de la méthode de test parcellaire (tester les composantes prosodiques séparément) qu'elle offrait. Il préférerait se concentrer sur la

recherche d'une mesure 'globale' de la prosodie, plus proche de ce que les auditeurs/imitateurs sont amenés à faire en situation réelle de communication. Il s'agissait aussi de participer au questionnement actuel concernant le type de paramètres acoustiques pertinents à corrélérer à l'évaluation perceptive de l'imitation (Pardo, 2013). Ce faisant, la recherche se dirigeait d'une part sur la recherche d'une *métrique prosodique* de l'imitation, et d'autre part, elle s'éloignait de l'apprenant pour se concentrer sur la capacité du professeur de langue à répéter à l'identique les patrons rythmiques et intonatifs qu'il tente de faire percevoir à ses apprenants. Tester la cohérence des (re)productions de l'enseignant devenait donc une étape cruciale de cette recherche, dont la visée ultime (outre la portée théorique en recherche fondamentale et appliquée) serait la création d'un outil d'évaluation du contrôle prosodique de l'enseignant de L2, utile à sa formation dans le domaine de la correction phonétique. Pour être en mesure de quantifier et modéliser les patrons rythmiques et intonatifs, il fallait trouver une mesure objective permettant de comparer les formes prosodiques. La méthode d'interpolation non linéaire *Dynamic Time Warping* (DTW ; Rillard *et al*, 2011) permet de mesurer la (dis)similarité entre les contours tonals après normalisation tonale et alignement temporel des pics et des creux des contours (voir Nocaudie, 2016 ; Nocaudie & Astésano, 2016a&b, pour plus de détails). Sur la base du même corpus que Nocaudie & Astésano (2012), une comparaison des contours d'imitation des phrases syntaxiquement ambiguës a été effectuée sur la base de cette mesure objective DTW, et mise en relation avec une tâche de jugement perceptif de type AX avec 15 auditeurs naïfs. Une très forte corrélation entre les mesures objectives de (dis)similarités prosodiques et les mesures subjectives de la tâche de perception a été observée (voir Figure 59), impliquant que l'algorithme peut être implémenté en toute confiance dans un outil d'entraînement à l'imitation prosodique.

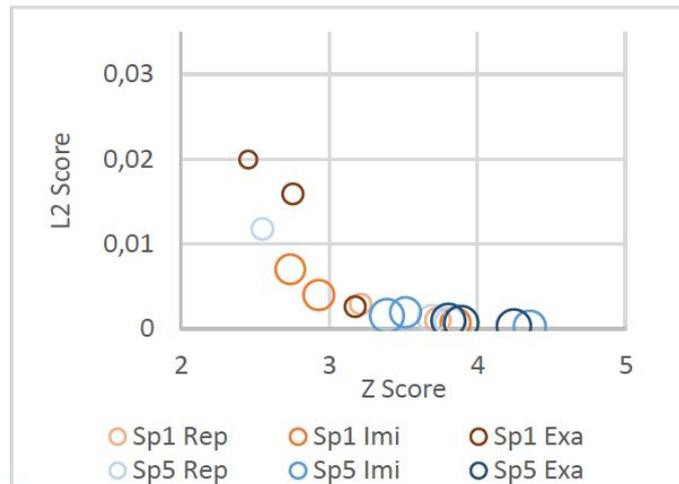


Figure 59 – Illustration, en trois dimensions, de l'évaluation de la performance de 2 locuteurs (Sp1 et Sp5) imitant la phrase « *Les bagatelles et les balivernes sottes* » au cours des trois tâches de production REP (répétition), IMI (imitation) & EXA (exagération). Les scores Zr et L2 sont respectivement sur les axes des x et des y (plus le score L2 est haut, moins la paire est similaire ; plus le score Zr est haut, plus la paire est similaire). La taille des bulles représente le score moyen au test AX. Les 3 types de mesures (objectives et subjectives) sont très corrélés. (Adapté de Nocaudie & Astésano, 2016a).

Ces deux études en prosodie et L2 m'auront permis, par ma modeste contribution à la mise en place de protocoles expérimentaux, de tester objectivement les processus sous-tendant la MVT, et ainsi, de participer à lever une (toute petite) partie du mystère relatif à l'efficacité étonnante de la méthode MVT de remédiation phonétique.

22. AUTRES RECHERCHES EN PHONÉTIQUE⁶⁸

Ces dernières années m'auront également permis de contribuer à quelques collaborations intéressantes, légèrement en marge de mon sujet de recherche, et n'ayant malheureusement pas abouti à des publications en revues, mais uniquement en actes de conférences internationales francophones.

Une collaboration inattendue avec un étudiant de M2 FLE⁶⁹ portait sur l'utilisation des indices audio-visuels dans l'identification des tons du chinois (Roulet-Guiot & Astésano, 2012). Les résultats de cette petite (mais très lourde) étude préliminaire montraient que, sans la présentation auditive des stimuli, les participants semblaient avoir deux types de stratégies de décodage visuel : une partie des participants utilisant essentiellement le cadrage visuel sur les lèvres pour décoder le ton, et l'autre partie utilisant le cadrage plus complet lèvre-cou, comme généralement trouvé dans la littérature (Burnham *et al*, 2000). Les événements de la vie ayant transporté cet étudiant étonnant sous d'autres cieux, cette recherche en est restée là.

⁶⁸ Voir Annexes : documents scientifiques 'Autres Travaux de Phonétique'.

⁶⁹ Encore un étudiant, comme O. Nocaudie, que j'aurai fièrement dévoyé de son Master Pro FLE (comme j'ai pu l'être moi-même un jour par A. Di Cristo) afin qu'il se plonge avec passion dans la recherche en phonétique!

Une deuxième collaboration s'est faite autour d'une question de macro-syntaxe, en collaboration avec N. Rossi-Gensane, ma collègue syntacticienne au département à l'époque. Nous voulions voir dans quelle mesure les indices prosodiques pouvait nous aider à trancher entre plusieurs propositions théoriques concernant les cas de détachements à gauche en français (ex : *Mon père, il vient demain*), et leur construction en une ou deux clauses (Astésano *et al.*, 2008). Les paramètres prosodiques généralement retenus dans ce type d'étude sont les pauses (présence et durée) et les contours tonals, et ne suffisent généralement pas pour aider à trancher ce débat. Il nous semblait intéressant d'ajouter à ces paramètres tonals la mesure de la durée des mots cibles autour de la frontière présumée de la clause (autrement dit, le dernier mot '*père*' du premier segment syntaxique et le premier clitique '*il*' du second segment syntaxique). Nos résultats montrent en effet l'utilisation inattendue de la durée aux frontières des segments pour instancier la relation référentielle qu'entretient le clitique avec l'élément détaché dans les 4 cas de détachement étudié. Comme dit dans la section § 2.5, p. 23, les variations de durée (et de rythme) ne sont que marginalement analysées, y compris en français, alors qu'ils font partie des paramètres importants permettant de rendre compte des relations de phrasé prosodique, y compris dans des cas ambigus de constructions macro-syntaxiques. Cette collaboration n'a malheureusement pu se poursuivre avec cette collègue qui a dû également partir sous d'autres cieux plus propices au développement de sa carrière.

Une troisième collaboration s'est faite dans le cadre du projet ANR *PhonLex* porté par Jacques Durand et collègues (*De la phonologie aux formes lexicales: liaison et cognition en français contemporain*, ANR blanc 2007), avec Annelise Coquillon alors post-doctorante dans ce projet.

La question de la resyllabation en contexte de liaison est une question fondamentale en phonologie, car elle pose la question du statut du 'mot' en français (Durand & Eychenne, 2014). Les auteurs défendent le point de vue selon lequel '*le découpage orthographique (du mot) est en grande partie motivé, et (...) que les unités de type « mot », à divers ajustements près, correspondent à des unités valides sur le plan cognitif et jouent un rôle central dans le système linguistique*' (p. 228). Ce point de vue me séduit bien évidemment puisque depuis le début de ma carrière, je teste l'hypothèse selon laquelle le mot (phonologique, prosodique et donc lexical) est une unité relevant de la hiérarchie prosodique en français. Ce 'mot prosodique' est marqué dans sa représentation sous-jacente par un IA et un FA, et sur le plan de la substance, il est instancié par une démarcation acoustique gauche et droite (voir entre autre § 11.2, p. 79). Cette étude me tenait à cœur mais est malheureusement restée à l'état de 'pilote', n'ayant donné lieu qu'à des communications internes au projet, lors des colloques et

jours d'étude PFC⁷⁰. Nous avons commencé à collecter avec Annelise Coquillon une série de données orales très prometteuses, posant certaines contraintes méthodologiques non négligeables (Astésano *et al*, 2008), sur des populations de locuteurs différentes (méridionaux : Coquillon & Astésano, 2008a&b ; ou cajuns, locuteurs mais *non lecteurs* de langue maternelle française aux Etats-Unis : Coquillon & Astésano, 2009).

Voici les principales tendances, pour le français dit 'standard'. Notons que les résultats acoustiques sont assez différents dans la littérature d'une étude à l'autre, montrant tantôt un rattachement de la consonne de liaison (CL) au mot suivant, tantôt au mot précédent (voir Durand & Lyche, 2008, pour une revue). Nous voulions mesurer la réalisation acoustique (durée, f_0 , intensité) des consonnes de liaison (CL – ex : *un avion*) et des consonnes initiales (Ci – ex : *un navet*) des mots lexicaux, et voir dans quelle mesure ces résultats pouvaient nous éclairer sur le rattachement potentiel de CL au mot précédent (clitique) ou au mot lexical suivant. Selon nous, les résultats d'allongements infrasyllabiques et de configurations tonales tels que décrits dans Astésano (2001) pouvaient nous permettre de déterminer de trancher sur ce rattachement. En effet, Astésano (2001) montrait que la pente de f_0 symétrique (avec alignement tonal précoce sur la voyelle) et l'allongement significatif de l'attaque consonantique sont caractéristiques du début d'un mot lexical. La modulation de telles configurations observées sur CL par rapport à Ci pourraient être interprétées comme le signe que CL n'est pas rattachée phonologiquement au mot suivant (*un Navion*) et que le mot lexical préserve son caractère d'unité grammaticale indépendant en contexte de liaison.

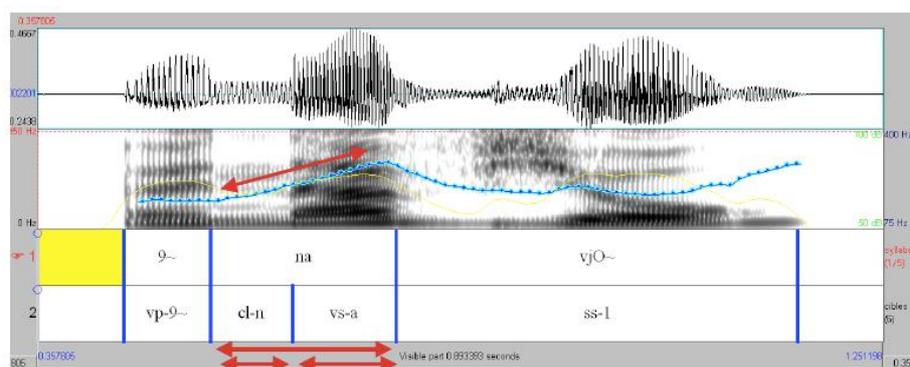


Figure 60 – Paramètres acoustiques analysés pour l'étude de la liaison sur le syntagme 'un avion'. Les mesures de durées se faisaient sur la voyelle précédant (vp) la consonne de liaison (cl) et la voyelle la suivant (vs). Les mesures de f_0 prenaient en compte la pente du contour montant à l'initiale du mot cible ('avion'), i.e. l'écart de f_0 entre cl et vs. Les résultats en contexte de liaison comme ici étaient comparés avec les cas d'attaque consonantique de début de mot (ex : 'un navet').

Nos résultats (très préliminaires, certes) étaient encourageants en ce sens qu'ils nous ont permis de distinguer CL et Ci sur le plan acoustique : Ci a une pente de f_0 plus abrupte que

⁷⁰ PFC : Phonologie du Français Contemporain. <http://www.projet-pfc.net/> (Durand *et al*, 2009)

CL (donc un alignement plus précoce dans la voyelle suivante, *vs*) et une attaque significativement plus longue que CL, confirmant la démarcation temporelle et tonale du mot lexical décrite dans Astésano (2001). Egalement, la voyelle suivant CL (*vs*) est plus longue que la voyelle après attaque consonantique de début de mot lexical (Coquillon & Astésano, 2008a&b), confirmant encore les résultats de Astésano (2001) selon lesquels l'attaque consonantique de début de mot lexical porte l'essentiel de la durée de la syllabe initiale d'un mot lexical (voir § 11.2, p. 79).

Voilà qui clôt ce chapitre sur les variations les plus importantes autour de mon thème de prédilection.

CODA

Un travail tel que celui-ci est un exercice de style à mi-parcours dans une carrière, permettant de faire un état d'avancement des travaux à un instant t . Il est donc par définition non abouti (mais aboutit-on jamais dans la recherche ?). La difficulté particulière de ce travail-ci est la tentative de rapprochement interdisciplinaire, à la croisée de la phonologie prosodique et des sciences cognitives. L'interdisciplinarité est selon moi absolument nécessaire à l'avancement des connaissances, mais elle engendre une certaine lourdeur expérimentale et conceptuelle. La visibilité du travail entrepris dans ce cadre-là s'en ressent.

Si j'éprouve une certaine frustration par rapport à la lourdeur d'une telle entreprise, je suis en revanche rassurée de voir qu'elle peut susciter un dynamisme scientifique inspirant de jeunes chercheurs. Parmi les doctorants constituant la 'fine équipe' (Olivier Nocaudie, Laury Garnier, Noémie te Rietmolen et Rogelio Mendez), l'émulation est telle qu'elle m'étonne moi-même. Je me surprends parfois à devoir tempérer leur enthousiasme à essayer de démontrer expérimentalement l'ancrage de l'accentuation au niveau du mot phonologique. Je mesure ma chance de les avoir à mes côtés dans cette aventure. Si j'ai osé faire ces propositions phonologiques pour le français, je le dois également à l'enthousiasme des post-doctorants Guillaume Roux et Rafeu Sichel-Bazin, et de mes collègues Lorraine Baqué et Roxane Bertrand notamment. Il reste cependant encore beaucoup à faire.

Mes recherches à venir seront donc essentiellement centrées sur la poursuite des investigations entreprises jusque-là. Si on part du principe que le pied est effectivement spondéique au niveau du pw , il faut maintenant montrer systématiquement quelles règles président à l'occurrence de IA seul, IA + FA, FA seul etc... Il y a visiblement des règles de réorganisation tonale qui s'opèrent pour marquer alternativement les niveaux de constituance par IA et FA. Laquelle des prééminences IA ou FA résiste le mieux au '*tonal crowding*' ? Il semblerait que FA puisse se réaliser en tons L^* ou $!H$ pour laisser la place à IA de se réaliser, mais cela demande à être systématiquement éprouvé. Si le corpus d'Edinburgh est idéal pour observer ces règles, l'investigation de la MapTask-Aix et du corpus CID demande à se développer dans les mois et années à venir. Des résultats sont en cours d'analyse sur des sous-parties de ces deux corpus, sur le versant perceptif et acoustico-phonétique (thèse de L. Garnier).

Je m'attacherai également à développer l'investigation systématique de l'allongement différentiel (Attaque et Rime) de IA et FA dès le niveau du mot, afin de consolider notre interprétation de ce niveau de constituance. L'analyse du corpus d'Edinburgh est en cours et elle devra être étendue à la MapTask et au CID. Ces investigations acoustico-phonétiques plus fines nous donneraient la possibilité d'éclairer également nos résultats très préliminaires sur

les phénomènes de liaison en français, et le rattachement de la consonne de liaison au mot précédant ou suivant (Durand & Eychenne, 2014).

La thèse en cours de R. Mendez devrait nous permettre d'aller beaucoup plus loin dans notre compréhension de la réorganisation tonale aux différents niveaux de constituance puisqu'elle est centrée sur la réalisation L* ou !H de FA à tous les niveaux. Nous mettrons ainsi à l'épreuve les outils puissants de AM pour le français, permettant de distinguer les phénomènes de proéminences métriques (*) et de frontières intonatives (%). C'est un des nœuds de la réponse à la question de l'indépendance de la métrique et de l'intonation en français. Delattre (1966b) et Fletcher (1991) soulignaient déjà que l'allongement final observé en frontière de constituants était à la fois imputable à l'accent et à la position pré-frontière. C'est sans doute parce que ces effets s'additionnent que l'on considère que le mètre du français est iambique. Toute la difficulté réside dans la mise en évidence expérimentale de l'implication de l'un et de l'autre. Nous l'entreprendrons sur le versant de la production (transcription prosodique) et perceptif. Les résultats seront particulièrement intéressants à mettre en regard de la conception de '*langue sans accent*'.

L'alignement des tons L₁ et H₁ de IA sera également systématiquement testé pour déterminer l'ancrage de IA au niveau du mot lexical. Ces résultats d'annotation seront mis en regard de la perception de IA.

Nous devrions également aborder (dans la thèse de Mendez ou d'autres travaux) la question cruciale de la perception de l'accentuation par des bilingues (français/espagnol notamment), qui nous donnerait un éclairage comparatif précieux sur le *niveau* de représentation de l'accentuation en français. De même, je trouverais intéressant de procéder à de telles comparaisons avec des langues comme le polonais ou le hongrois qui présentent certaines similitudes avec le système accentuel du français.

Dans les années à venir, je porterai un intérêt particulier à la parole spontanée. Pour l'instant, les résultats qui m'ont conduite à faire cette proposition phonologique sont issus essentiellement de parole contrôlée (corpus d'Edinburgh essentiellement) mais nous avons vu que la démarcation accentuelle des constructions spontanées en Nom+Adjectif (Figure 28, p.94) se caractérise (souvent ?) par un plateau tonal, que l'on retrouve d'une certaine manière en parole contrôlée (Welby *et al*, 2016). Un grand pan de ma recherche à venir, en collaboration toujours je l'espère avec Roxane Bertrand et Pauline Welby, sera d'investiguer ces formes particulières d'arcs accentuels en parole spontanée, sur toutes formes de constructions syntaxiques (N+N ; N+A ; N de N etc... ; Corpus CID et MapTask-Aix).

Les analyses de neuroimagerie (EEG) seront également poursuivies. Bien qu'encore trop peu d'études existent sur le traitement de l'accentuation (voir cependant Honbolygó et collègues sur le hongrois, justement), elles nous permettent vraiment d'affiner notre compréhension du niveau d'interprétation de l'accentuation en français, et son interaction avec les autres niveaux linguistiques (sémantique notamment). Quatre pans de recherche vont intervenir à court et moyen terme : d'une part, l'investigation de FA est en cours (thèse de N. te Rietmolen) selon les mêmes protocoles que ceux développés pour IA ; d'autre part, te Rietmolen veut terminer son investigation de IA et FA par l'investigation de la synchronisation cérébrale sur ces deux types de proéminences. Si ces deux proéminences devaient présenter effectivement un même poids dans le marquage du niveau du mot phonologique, alors l'onde cérébrale Delta (intervenant entre 400 et 600 ms ; travaux de Poeppel : Ding *et al*, 2016 par exemple) devrait se synchroniser avec l'occurrence de IA et FA.

A moyen terme, enfin, il faudrait répliquer ses tendances sur des mots bisyllabiques, puisque nous n'avons pour l'instant testé que les mots trisyllabiques offrant un marquage IA-FA (tonals) beaucoup plus fréquent. Enfin, j'aimerais investiguer l'occurrence de la composante CPS (*Closure Positive Shift*) mise en évidence par Steinhauer *et al* (1999), reflétant le traitement des frontières prosodiques, sur le corpus d'Edinburgh qui permet d'observer tous les niveaux de constituance dès le *pw*. Cette composante permettrait de tester la validité de notre proposition de spondée au niveau du mot phonologique, si elle ne se développait qu'au niveau de *pw* et non pas au niveau inférieur de l'unité tonale. Il serait en outre également intéressant de voir dans quelle mesure son amplitude est modulée par les niveaux de constituance.

Toutes ces investigations du rôle phonologique de IA et FA vont occuper encore une bonne partie de ma carrière. Mais je ne voudrais pas terminer ce document sans faire référence à la thématique qui m'a amenée, initialement, à mon parcours actuel : la comparaison entre Langage et Musique. Cela fait de trop nombreuses années que je voudrais m'y pencher sérieusement, tant ces deux systèmes sont comparables à bien des égards et peuvent nous éclairer sur la cognition humaine. J'inscris pour la première fois un Master SDL dans cette thématique avec une jeune étudiante très sérieuse et musicienne avertie. Gageons que cela sera le début d'une nouvelle aventure scientifique...

BIBLIOGRAPHIE

- Aguilera, M. ; El Yagoubi, R. ; Espesser, R. ; Astésano, C. (2014). Event Related Potential investigation of Initial Accent processing in French. *Speech Prosody 2014*, Dublin, U.K., May 20-23 2014 : 383-387.
- Alazard, C. (2013). *La Prosodie au cœur de la problématique oral/écrit – Etude de l'influence de la pratique orale sur les compétences en lecture chez des apprenants anglophones de Français Langue Etrangère*. Thèse de Doctorat, Université de Toulouse2-Jean Jaurès.
- Alazard, C.; Astésano, C.; Billières, M. (2012a). From prosodic skills to L2 reading proficiency – An experimental approach to L2 phonetics teaching methods. *Methodological Perspectives on Second Language Prosody Papers from ML2P 2012*. Maria Grazia Busà & Antonio Stella (Eds), Cleup, Padova: 55-59.
- Alazard, C.; Billières, M. ; Astésano, C. (2012b). Méthode Verbo-Tonale et passage à l'écrit : la prosodie, un recours nécessaire. In : Calvo-Medina, M.-V. ; Murillo-Puyal, J. (Eds.), *Perception phonique et Parole*. Percepción fónica, Habla y Hablar, Mons, CIPA, pp. 157-166.
- Alazard, Ch.; Astésano, C.; Billières, M. (2012c). Évaluer les transferts prosodiques de l'oral vers l'écrit en langue étrangère: Mesures quantitatives et qualitatives en production et en perception de la fluence en lecture oralisée chez des apprenants de Français Langue Étrangère. *L2 Proficiency Assessment Workshop*, Montpellier, 24-25 février 2012.
- Alazard, Ch.; Astésano, C.; Billières, M. (2012d). MULTIPHONIA: a MULTImodal database of PHONetics teaching methods in classroom InterActions. *LREC 2012*, Istanbul, 23-25 mai 2012: 2578-2583.
- Alazard, C.; Astésano, C.; Billières, M.; Espesser, R. (2011). Rôle de la prosodie dans la structuration du discours - Proposition d'une méthodologie d'enseignement de l'oral vers l'écrit en Français Langue Etrangère. In : Yoo, H-Y & Delais-Roussarie, E. (eds), *IDP 2009 Proceedings*, Paris, Sept 2009, ISSN 2114-7612: 49-61. http://makino.linguist.jussieu.fr/idp09/actes_fr.html
- Alazard, C. ; Astésano, C. ; Billières, M. (2011) : MULTIPHONIA (MULTImodal database of PHONetics teaching methods in classroom InterActions). oai:sldr.org:sldr000780. <http://sldr.org/sldr000780/fr>

- Alazard, C.; Astésano, C.; Billières, M. (2010). The importance of explicit and implicit prosody in Foreign Language Learning. *Speech Prosody 2010*, 100648:1-4, ISBN:978-0-557-51931-6, Chicago, 11-13 mai 2010.
- Alazard, C., Billières, M.; Astésano, C. (2009). Méthode Verbo-Tonale et passage à l'écrit : la prosodie, un recours nécessaire. *Colloque International de Verbo-Tonale*, Barcelone, 8-9 mai 2009.
- Allen, G.D. (1975). Speech rhythm: its relation to performance universals and articulatory timing, *Journal of Phonetics*, (3) : 75-86.
- Alter, K.; Astésano, C.; Besson, M. (2001). Influence des Modalités Intonatives sur le Traitement Sémantique: une Etude Electrophysiologique. In *Oralité et Gestualité - Actes d' Orage 2001*, 13-17 Juin, Aix-en-Provence, Editions L'Harmattan, Paris: 411-415.
- Altmann, G & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, 30 (3), 191-238.
- Anderson, A. ; Bader, M. ; Bard, E. ; Boyle, E. ; Doherty, G. ; Garrod, S. ; Isard, S. ; Kowtko, J. ; McAllister, J. ; Miller, J. ; Sotillo, C. ; Thompson, H. & R. Weinert (1991). The HCRC Map Task corpus. *Language and Speech*, 34(4) : 351–366.
- André, C. ; Ghio, A. ; Cavé, C. ; Teston, B. (2003). PERCEVAL: a Computer-Driven System for Experimentation on Auditory and Visual Perception. In Proceedings of XVth ICPHS (p. 1421-1424). Barcelone, Espagne.
- Arvaniti, A. (2017). The Autosegmental-Metrical model of intonational phonology. In: Shattuck-Hufnagel, Stefanie and Barnes, Jonathan, eds (In press). *Prosodic Theory and Practice*. MIT Press.
- Astésano, C. (2016). The Prosodic Characterization of Reference French. In : *Varieties of Spoken French: a source book*. Detey, S.; Durand, J.; Laks, B.; Lyche, Ch. (Eds.), Oxford University Press: 68-85.
- Astésano, C. & Bertrand, R. (2016). Accentuation et niveaux de constituance en français: enjeux phonologiques et psycholinguistiques. *Langue française*, (2016/3 ; n°191 : *La Prosodie du français : accentuation et phrasé*), 11-30.

- Astésano, C. & Jucla, M. (Eds.) (2015) *Neuropsycholinguistic perspectives on Language cognition – Essays in honour of Jean-Luc Nespoulous*. Hove: Psychology Press. ISBN 9780203797365.
- Astésano, C., El-Yagoubi, R., Aguilera, M., Espesser, R. (2013). Processing of the Initial Accent by French listeners: A MisMatch Negativity Investigation. *AMLAP 2013, Architectures and Mechanisms for Language Processing*, Marseille, France, Sept.2-4.
- Astésano, C. ; Boutin, B.A. ; Coquillon, A. (2008) Les problèmes méthodologiques des enquêtes de terrain avec des locuteurs francophones non lecteurs. *Journées PFC décembre 2008: Phonologie du français contemporain: variation, interfaces, cognition*. 11-13 décembre 2008, Paris.
- Astésano, C. ; Espesser, R. ; Rossi-Gensane, N. (2008). Quelques cas particuliers de détachement à gauche-ou la prosodie à l'aide de la syntaxe. *Journées d'Etude sur la Parole (JEP)* : 109-112.
- Astésano, C., Bard, E. G. ; Turk, A. (2007). Structural influences on Initial Accent placement in French. *Language and speech*, 50(3), 423-446.
- Astésano, C.; Besson, M.; Alter, K. (2004a). Brain potentials during semantic and prosodic processing in French. *Cognitive Brain Research*, 18 (2): 172-184.
- Astésano, C.; Lacheret-Dujour, A.; Magne, C.; Morel, M.; Coquillon, A.; Espesser, R.; Besson, M. (2004b). Marquage acoustique du focus contrastif non codé syntaxiquement en français. *JEP 2004*, Fès, 19-22 Avril: 41-44.
- Astésano, C.; Magne, C.; Besson, M. (2004c). Influence du rythme sur le traitement sémantique en français: approches comportementale et électrophysiologique. *JEP 2004*, Fès, 19-22 Avril: 37-40.
- Astésano, C.; Alter, K.; Besson, M. (2002). Processing of Prosodic and Semantic Information during spoken language comprehension: an ERP Investigation. *ISCA Workshop on Temporal Integration in the Perception of Speech*, Aix-en-Provence, 8-10 Avril 2002.
- Astésano, C.; Alter, K.; Besson, M. (2001a). Brain Potentials during Semantic and Prosodic Processing: an ERP investigation. *Conference on Human Brain Mapping*, Bristol, Juin 2001, in *NeuroImage*, 13(6), Part 2, S506.

- Astésano, C.; Alter, K.; Besson, M. (2001b). Brain Potentials Investigations of Semantic and Prosodic Processing during Spoken Language Comprehension. *2nd ESCOP & 18th BPS Cognitive Section Conference*, Edinburgh, 5-8 Sept: 36.
- Astésano, C.; Espesser, R.; Flachaire, E.; Nicolas, P. (1997). Tonal configurations of prominence in French. *ESCA Workshop on Intonation*, Athènes: 29-32.
- Astésano, C.; Espesser, R.; Flachaire, E.; Nicolas, P. (1998). Analyse multi-paramétrique des configurations tonales des proéminences en français. *XXII ièmes Journées d'Etude sur la Parole*, Martigny : 95-98.
- Aura, K. (2012). *Protocole d'évaluation du langage fondé sur le traitement de fonctions prosodiques: étude exploratoire de deux patients atteints de gliomes de bas grade en contexte péri-opératoire*. Université Toulouse le Mirail-Toulouse II. Consulté à l'adresse <https://hal.archives-ouvertes.fr/tel-00798667/>
- Aura, K. ; Astésano, C. ; Lubrano, V. ; Prod'homme, K. ; Köpke, B. (2011) Spécialisation hémisphérique du traitement de la prosodie chez des sujets cérébro-lésés atteints de gliomes de haut grade. *Quatrième Journée de Phonétique Clinique (JPC)*, 19-21 mai 2011, Strasbourg :
- Aura, K ; Duvignau, K ; Astésano, C., Nespoulous, J-L. (2009) Production de la prosodie linguistique : étude du continuum autistique, de l'autisme de haut-niveau au syndrome d'Asperger. *3ièmes Journées de Phonétique Clinique*, 4-5 Décembre 2009, Aix-en-Provence : 17
- Bagou, O. ; Frauenfelder, U.H. (2006). Stratégie de segmentation prosodique: rôle des proéminences initiales et finales dans l'acquisition d'une langue artificielle. *JEP*, Dinard : 571-574.
- Bagou, O.; Fougeron, C.; Frauenfelder, U. H. (2002). Contribution of Prosody to the Segmentation and Storage of "Words" in the Acquisition of a New Mini-Language. *Speech Prosody 2002*, Aix en Provence.
- Bagou, O. & Frauenfelder, U. H. (2002). Alignement lexical et segmentation de la parole. *Revue française de linguistique appliquée*, 7(1) : 67-82.
- Banel, M.H.; Frauenfelder, U.H.; Perruchet, P. (1998). Contribution des indices métriques à l'apprentissage d'un langage artificiel. *XIIèmes Journées d'Études sur la Parole*, Martigny, Suisse : 29-32.

- Banel, M.H.; Bacri, N. (1994). Reconnaissance de la parole et indices de segmentation métriques et phonotactiques. *L'Année psychologique*, 97 : 77-112.
- Baum, S.R.; Pell, M.D. (1999). The neural bases of prosody : insights from lesion studies and neuroimaging. *Aphasiology*, 13(8) : 581-608.
- Beckman, M.E. (1992). Evidence for Speech Rhythms across Languages. In *Speech Perception, Production and Linguistic Structure*, Y. Tohkura; Vatikiotis-Bateson, E.; Sagisaka, Y. (eds.), Tokyo : 457-463.
- Beckman, M. E. (1986). *Stress and non-stress accent*. Walter de Gruyter.
- Bertrand, R. ; Blache, P. ; Espesser, R. ; Ferré, G. ; Meunier, C. ; Priego-Valverde, B. ; Rauzy, S. (2008). « Le CID – Corpus of Interactional Data – Annotation et exploitation multimodale de parole conversationnelle », *Traitement Automatique des Langues*, 49 (3) :105-134.
- Besson, M. ; Chobert, J. ; François, C. ; Astésano, C. ; Marie, C. (2015). Influence of musical expertise on the perception of pitch, duration and intensity variations in speech and harmonic sounds. In Astésano, C. & Jucla, M. (eds.), *Neuropsycholinguistic Perspectives on Language Cognition – Essays in honour of Jean-Luc Nespoulous*. Hove, Psychology Press : 88-102.
- Besson, M. ; Chobert, J. ; Marie, C. (2011). Transfer of training between music and speech: common processing, attention, and memory. *Frontiers in Psychology* 2(94):94 · DOI: 10.3389/fpsyg.2011.00094 ·
- Besson, M. ; Marie, C. ; Astésano, C. ; Magne, C. (2009). Electrophysiological studies of prosody. In Alter, K. ; Horne, M. ; Lindgren, M. ; Roll, M. & von Koss Torkildsen, J. (Eds.), *Brain Talk*. Lund University, Media Tryck : 123-137.
- Besson, M.; Faïta, F.; Czernasty, C. ; Kutas, M. (1997). What's in a pause: event-related potential analysis of temporal disruptions in written and spoken sentences. *Biological Psychology*, 46 : 3-23.
- Bigi, B. (2012). SPPAS: a tool for the phonetic segmentation of speech. *Language Resource and Evaluation Conference*, Istanbul :1748–1755.
- Billières, M.; Alazard, C; Astésano, C.; Nocaudie, O. (2013). *Phonétique Corrective en Français Langue Etrangère: Méthode Verbo-Tonale*. Université Ouverte des Humanités (UOH): <http://w3.uohprod.univ-tlse2.fr/UOH-PHONETIQUE-FLE/>

- Billières, M. (2008). Le statut de l'intonation dans l'évolution de l'enseignement/apprentissage de l'oral en FLE. *Numéro spécial: quel oral enseigner, cinquante ans après le Français fondamental*, 27-37.
- Billières, M. (2005). Les pratiques du verbo-tonal. Retour aux sources. In : M. Berré (Ed.), *Linguistique de la parole et apprentissage des langues. Questions autour de la méthode verbo-tonale de P. Guberina*. Mons : 67-87.
- Böcker, K.B.E.; Bastiaansen, M.C.M.; Vroomen, J.; Brunia, C.H.M.; de Gelder, B. (1999). An ERP correlate of metrical stress in spoken word recognition. *Psychophysiology*, 36 : 720-760.
- Bolinger, D. (1986). *Intonation and its parts: Melody in spoken English*. Stanford University Press.
- Bolinger, D. (1985). Two views of accent. *Journal of linguistics*, 21(1), 79-123.
- Bolinger, D.L. (1981). *Two kinds of vowels, Two kinds of rhythm*. Indiana University Linguistics Club.
- de Boysson-Bardies, B. (1996). *Comment la parole vient aux enfants*. Odile Jacob.
- Brown, G. & Yule, G. (1983). *Discourse Analysis*. Cambridge University Press.
- Burnham, D. ; Ciocca, V. ; Lauw, C. ; Lau, S. ; Stokes, S. (2000). Perception of visual information for Cantonese tones. *8th Australian International Conference on Speech Science and Technology* : 86-91.
- Butterfield, S.; Cutler, A. (1988). Segmentation errors by human listeners: evidence for a prosodic segmentation strategy. *7th FASE Symposium*, Edinburgh, 3, 827-833.
- Campbell, W.N. (1992). Syllable-based segmental duration. In *Talking Machines: Theories, Models and Designs*. G. Bailly; C. Benoit & T.R. Sawallis (eds.), Elsevier Science Publishers, North-Holland, Amsterdam, 211-224.
- Campbell, W.N. (1995). Loudness, spectral tilt, and perceived prominence in dialogues. *ICPh'S 95*, Stockholm, Suède, 3, 676-679.
- Cason, N. ; Astésano, C. ; Schön, D. (2015). Bridging music and speech rhythm: Rhythmic priming and audio-motor training affect speech perception. *Acta psychologica*, 155 : 43-50.

- Christophe, A. (2002). Discovering words in the continuous speech stream: the role of prosody. *TIPS workshop*, Aix-en-Provence, 22.
- Christophe, A. ; Guasti, T. ; Nespors, M ; Dupoux, E. ; Van Ooyen, B. (1997). Reflections on phonological bootstrapping: Its role for lexical and syntactic acquisition. *Language and Cognitive Processes*, 12(5-6) : 585-612.
- Cole, J. ; Mahrt, T. ; Hualde, J. I. (2014). Listening for sound, listening for meaning: Task effects on prosodic transcription. *Speech Prosody*, Dublin : 858–863.
- Cole, J. ; Mo, Y. ; Hasegawa-Johnson, M. (2010). Signal-based and expectation-based factors in the perception of prosodic prominence. *Laboratory Phonology*, 1(2) : 425-452.
- Content, A.; Kearns, R.; Frauenfelder, U. H. (2001). Boundaries versus onsets in syllabic segmentation. *Journal of Memory and Language*, 45(2), 177-199.
- Coquillon, A.; Astésano, C. (2009) Etude comparative des consonnes de liaison en français cadien et de métropole. *Phonologie du Français Contemporain: 10 ans de PFC – Bilan et perspectives – 10-12 Décembre 2009*, Paris.
- Coquillon, A.; Astésano, C. (2008a). A multiparametric phonetic investigation of liaison in standard and meridional French. *Colloque PFC 2008, la Nouvelle-Orléans : Structures des français en contact*, 25-28 juin, Nouvelle-Orléans, Etats-Unis.
- Coquillon, A. & Astésano, C. (2008b) Investigation multiparamétrique de la liaison en français standard et méridional. *Journée d'études Phonlex*, 4 septembre 2008, Toulouse.
- Cutler, A. (1999). Prosodie Structure and Word Recognition. In *Language Comprehension : a biological perspective* (second edition), A.D. Friederici (Ed.), Berlin : 41-70.
- Cutler, A. (1990). Exploiting Prosodic Probabilities in Speech Segmentation. *Cognitive Models of Speech Processing. Psycholinguistic and Computational Perspectives*, Gerry T.M. Altmann (Ed.), The MIT Press : Cambridge, Mass. : 105-121.
- Cutler, A.; Dahan, D.; van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, 40 :141-201.
- Cutler, A. & Butterfield, S. (1992). Rhythmic cues to speech segmentation: evidence from juncture misperception. *Journal of Memory and Language*, 31 : 218-236.

- Cutler, A.; Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 14 (1) : 113-121.
- Cutler, A. ; Mehler, J. ; Norris, D. ; Segui, J. (1986). The Syllable's Differing Role in the segmentation of French and English. *Journal of Memory and Language*, 25 : 385-400.
- Delais-Roussarie, E. & Di Cristo, A. (à paraître). L'Accentuation. In : A. Abeillé & D. Godard (eds.), *Grande Grammaire du Français*, Paris: CNRS.
- Delais-Roussarie, E. (2016). Prosodic phonology and its interfaces. In : Fischer, S & C. Gabriel (eds.): *Grammatical Interfaces in Romance Linguistics, série Manuals of Romance Linguistics (MRL)*, De Gruyter.
- Delais-Roussarie, E. ; Post, B.; Avanzi, M.; Buthke, C.; Di Cristo, A.; Feldhausen, I.; Jun, S.-A.; Martin, P.; Meisenburg, T.; Rialland, A.; Sichel-Bazin, R. ; Yoo, H. (2015). Intonational phonology of French: Developing a ToBI system for French. In : S. Frota & P. Prieto (eds), *Intonational Variation in Romance*, Oxford: Oxford University Press : 63-100.
- Delais-Roussarie, E. & Feldhausen, I. (2014). Variation in Prosodic Boundary Strength: a study on dislocated XPs in French. *Speech Prosody 2014*, Dublin : 1052–1056.
- Delais-Roussarie, E. (2000). Vers une nouvelle approche de la structure prosodique. *Langue française*, (126) : 92-112.
- Delais, E. (1995). *Pour une approche parallèle de la structure prosodique. Étude de l'organisation prosodique et rythmique de la phrase française*. Thèse de Doctorat, Université de Toulouse-Le Mirail.
- Delais, E. (1994). Prédiction de la variabilité dans la distribution des accents et les découpages prosodiques en français. *XX ièmes Journées d'Études sur la Parole*, Trégastel : 379-384.
- Delattre, P. (1966a). Les dix intonations de base du français. *French Review*, 41 (3) : 326-339.
- Delattre, P. (1966b). A comparison of syllable length conditioning among languages. *International Review of Applied Linguistics*, 4 : 183-198.
- Delattre, P. (1962). Some Factors of Vowel Duration and their Cross-Linguistic Validity. *Journal of the Acoustic Society of America*, 34 : 1141-1143.

- Dell, F. (1984). L'accentuation dans les phrases en français. In : *Formes sonores du langage*, F. Dell; D.J. Hirst; Vergnaud, J.R. (eds.), Hermann :65-122.
- Di Cristo, A. (2012). Le pouvoir de la prosodie ou la revanche de Cendrillon. In : *L'Homme Communicant (CIPA, p. 95-114)*. Mons, Belgique: Baqué, L. & Estrada, M.
- Di Cristo, A. (2011). Une approche intégrative des relations de l'accentuation au phrasé du français. *Journal of French Language Studies*, 21(2) :73-95.
- Di Cristo, A. (2004). La prosodie au carrefour de la phonétique, de la phonologie et de l'articulation formes-fonctions. *TIPA*, 23 : 67–211.
- Di Cristo, A. (2000). Vers une modélisation de l'accentuation en français. Deuxième partie : le modèle. *Journal of French Language Studies*, 10(1) : 27-44.
- Di Cristo, A. (1999). Le cadre accentuel du français contemporain: essai de modélisation. *Langues 2* (3) : 184-205 et *Langues 2* (4) : 258-267.
- Di Cristo, A. (1998). Intonation in french. *Intonation systems: A survey of twenty languages*, D. Hirst & A. Di Cristo (Eds), Cambridge University Press :195-218.
- Di Cristo, A. (1985). *De la Microprosodie à l'Intonosyntaxe*. Publications de l'Université de Provence.
- Di Cristo, A.; Hirst, D.J. (1997). L'accentuation non emphatique en français: stratégies et paramètres. *Polyphonie pour Ivan Fónagy*, J. Perrot (ed.), Paris, l'Harmattan : 71-101.
- Di Cristo, A. et Hirst, D.J. (1996). Vers une typologie des unités intonatives du français. *JEP Avignon* : 219–222.
- Di Cristo, A. et Hirst, D.J. (1993). Rythme syllabique, rythme mélodique et représentation hiérarchique de la prosodie du français. *Travaux de l'Institut de Phonétique d'Aix*, 15: 9- 24.
- Di Cristo, A., & Jankowski, L. (1999). Prosodic organisation and phrasing after focus in French. *ICPhS*, San Francisco : 1565-1568.
- Dilley, L. ; Breen, M. ; Bolivar, M. ; Kraemer, J. & E. Gibson (2006). A comparison of inter-coder reliability for two systems of prosodic transcriptions: RaP (Rhythm and Pitch) and ToBI (Tones and Break Indices). *Interspeech 2006 –ICSLP* : 277-312.

- Ding N.; Melloni L.; Zhang H.; Tian X.; Poeppel D. (2016) Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, 19(1):158-64
- Duez, D. (1987). Hiérarchisation des paramètres acoustiques et identification des frontières. *JEP* : 221-223.
- Dumay, N.; Frauenfelder, U.; Content, A. (2002). The role of the syllable in lexical segmentation in French: Word-spotting data. *Brain and Language*, 81 : 144-161.
- Dupoux, E. ; Pallier, C. ; Sebastian, N. ; Mehler, J. (1997). A distressing “deafness” in French?. *Journal of Memory and Language*, 36(3), 406-421.
- Durand, J. & Eychenne, J. (2014). Mot et phonologie en français: de la persistance d’une illusion. In : *La Phonologie du français: normes, périphéries, modélisation*. Nanterre, Presses universitaires de Paris Ouest : 227-260.
- Durand, J. ; Laks, B. ; C. Lyche (2009). Le projet PFC: une source de données primaires structurées. In J. Durand, B. Laks et C. Lyche (eds)(2009) *Phonologie, variation et accents du français*. Paris: Hermès : 19-61.
- Durand, J. & Lyche, C. (2008). French liaison in the light of corpus data. *Journal of French Language Studies*, 18(1) : 33-66.
- Fant, G.; Kruckenberg, A.; Nord, L. (1991). Durational correlates of stress in Swedish, French and English, *Journal of Phonetics*, 19 : 351-365.
- Faure, G. & Di Cristo, A. (1977). *Le français par le dialogue*. Hachette.
- Fletcher, J. (1991). Rhythm and final lengthening in French. *Journal of Phonetics*, 19(2) : 193-212.
- Fodor, J. (2002a). Prosodic Disambiguation in Silent Reading. *NELS 32*, Amherst, MA: GLSA, University of Massachusetts.
- Fodor, J. (2002b). Psycholinguistics cannot escape prosody. *Speech Prosody*, Aix-en-Provence : 83-88.
- Fónagy, I. (1989). La Cendrillon de Dwight Bolinger. *Semiotica*, 76 (3-4) : 217-264.
- Fónagy, I. (1980). L'accent en français : accent probabilitaire. In *L'accent en français contemporain (Studia Phonetica)*, 15, I. Fónagy & P. Léon (eds.) : 123-233.

- Fraisse, P. (1974) *Psychologie du Rythme*, Presse Universitaire de France.
- Frazier, L. (1987). Sentence processing: a tutorial review. In: M. Coltheart (Ed.), *Attention and Performance*, N.J. Erlbaum, Hillsdale, 12 : 559– 586.
- Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiological reviews*, 91(4) : 1357-1392.
- Friederici, A.D. (2002). Towards a neural basis of auditory sentence processing. *Trends Cogn. Sci.*, 6 : 78-84.
- Frauenfelder, U. H., & Floccia, C. (1999). The recognition of spoken words. In : *Language comprehension: A biological perspective*. Springer Berlin Heidelberg : 1-40.
- Fry, D.B. (1958). Experiments in the perception of stress, *Language and Speech* (1) : 126-152.
- Garde, P. (1968). *L'Accent*. Paris: Presses Universitaires de France.
- Garnier, L. (en cours). *Analyse acoustico-perceptive de la constituance prosodique en français*. Thèse de Doctorat, Université de Toulouse2-Jean Jaurès.
- Garnier, L. ; Astésano, C. ; Baqué, L. ; Dagnac, A. (2016). Constituance et phrasé prosodique en français : une étude perceptive. *JEP-TALN-RECITAL 2016*, Paris : 146-154.
- Garnier, L. ; Baqué, L. ; Dagnac, A. ; Astésano, C. (2016). Perceptual investigation of prosodic phrasing in French. *Speech Prosody 2016*, Boston : 1153-1157.
- Ghio, A. ; Giusti, L. ; Blanc, E. ; Pinto, S. ; Lalain, M. ; Robert, D. ; Fredouille, C. ; Woisard, V. (2016). Quels tests d'intelligibilité pour évaluer les troubles de production de la parole ? *Journées d'Etude sur la Parole*, : 589-596.
- Ghio, A.; Teston, B.; Frenck-Mestre, C.; Astésano, C.; Schön, D.; Nazarian, B.; Roth, M.; Anton, J.L. (2004). Perception de la parole et IRM: réalisation, évaluation et validation d'un système permettant une stimulation sonore de qualité en cours de séquence IRMf. *JEP 2004*, Fès, 19-22 Avril : 241-244.
- Ghio, A. ; André, C. ; Teston, B. ; Cavé, C. (2003). PERCEVAL : une station automatisée de tests de PERception et d'EVALuation auditive et visuelle. *Travaux Interdisciplinaires du Laboratoire Parole et Langage*, 22 : 115-133.
- Goldsmith, J. A. (1976). *Autosegmental phonology*. PhD dissertation, MIT.
- Grammont, M. (1963). *Traité de phonétique*. Paris: Delagrave.

- Grosjean, F. & Gee, J. P. (1987). Prosodic structure and spoken word recognition. *Cognition*, 25(1) : 135-155.
- Gussenhoven, C. (2002). Intonation and interpretation. *Speech Prosody 2002* : 47-57.
- Halle, M. & Vergnaud, J.R. (1987). *An Essay On Stress*, MIT Press, Cambridge, Mass.
- Handel, S. (1989). *Listening. An Introduction to the Perception of Auditory Events*. MIT Press, Cambridge.
- 't Hart, J.; Collier, R.; Cohen, A. (1990) *A perceptual study of intonation : an experimental-phonetic approach to speech melody*. Cambridge : Cambridge University Press.
- Hayes, B. (1989). The prosodic hierarchy in meter. In : *Phonetics and Phonology : Rhythm and Meter*, Vol. 1, Academic Press, INC., Paul Kiparsky & Gilbert Youmans (eds.), San Diego, U.S.A., 201-260.
- Hermes, D.J. & Van Gestel, J.C. (1991). The frequency scale of speech intonation. *Journal of the Acoustic Society of America*, Vol. 90(1) : 97-102.
- Hirst, D. J.; Di Cristo, A.; Espesser, R. (2000). Levels of description and levels of representation in the analysis of intonation. In : M. Horne (ed.) *Prosody : Theory and Experiment*. Dordrecht, Pays-Bas: Kluwer. 2000 : 51-87.
- Hirst, D. & Espesser, R. (1993). Automatic Modelling Of Fundamental Frequency Using A Quadratic Spline Function. *T.I.P.A.*, 15 : 71-85.
- Hirst, D.J. & Di Cristo, A. (1984). French Intonation: a Parametric Approach. *Die Neueren Sprachen*, 83(5) : 554-569.
- Hirst, D.J. (1983). Structures and Categories in Prosodic Representations. In : *Prosody: Models and Measurements*, A. Cutler & D.R. Ladd (eds.), Heidelberg : 93-109.
- Honbolygó, F.; Csépe, V. (2012). Saliency or template? ERP evidence for long-term representation of word stress. *International Journal of Psychophysiology*, 87 : 165-172.
- Honbolygó, F.; Csépe, V.; Ragó, A. (2004). Suprasegmental speech cues are automatically processed by the human brain: a mismatch negativity study. *Neurosci. Lett.*, 363 : 84-88.
- House, D.; Hermes, D.; Beaugendre, F. (1997). Temporal-alignment categories of accent-leading rises and falls. *5th European Conference On Speech Communication and Technology*, 2, Rhodes, Greece : 879-882.

- Hruska, C., Steinhauer, K.; Alter, K. (2001). Misleading dialogues: Human's brain reaction to prosodic information. *ORAGE'2001*, Aix-en-Provence : 425-430.
- Hublin, J.-J. (2005) La langue des premiers hommes. In: J.-M. Hombert (ed.), *Aux Origines des Langues et du Langage*. Fayard : 102-117.
- Hyman, L.M. (1975) *Phonology: Theory and Analysis*, Holt, Rinehart & Winston.
- Jankowski, L. (1996). *Le marquage prosodique des mots*, Mémoire de maîtrise, Institut de Phonétique, Université de provence, (132 pp).
- Jankowski, L. (1997). *Approche linguistique, acoustique et perceptive des proéminences du français*. Mémoire de DEA, Institut de Phonétique, Université de Provence, (235 pp).
- Jankowski, L.; Astésano, C.; Di Cristo, A. (1999). The initial rhythmic accent in French: acoustic data and perceptual investigation. *ICPhS'99*, San Francisco, 1 : 257-260.
- Jiaa, M. (2011). *Elaboration d'un protocole de test en perception et production de la prosodie en contexte péri-opératoire pour les patients atteints de gliomes de bas-grades*. Mémoire de Master 2, Université Toulouse2.
- Joanette, Y.; Ska, B.; Côté, H. (2004). *Protocole Montréal d'Évaluation de la Communication*. Isbergues, France : Ortho Édition.
- Joanette, Y. ; Goulet, P. ; Hannequin, D. (1990). *Right hemisphere and verbal communication*. New York : Springer-Verlag.
- Jun, S. A. & Fougeron, C. (2002). Realizations of accentual phrase in French intonation. *Probus*, 14(1) : 147-172.
- Jun, S. A & Fougeron, C. (2000). A phonological model of French intonation. In *Intonation*, A. Botinis (ed.) : 209-242.
- Jun, S.A. & Fougeron, C. (1995). The accentual phrase and the prosodic structure of French. *ICPhS*, Stockholm, 2 : 722-725.
- Kager, R. (1994). The Metrical Theory of Word Stress. In : *The Handbook of Phonological Theory*, John Goldsmith (ed.), Blackwell, Oxford : 367-402.
- Kim-Dufor, D-H. ; Ferragne, E. ; Dufor, O. ; Astésano, C.; Nespoulous, J-L. (2012) A novel prosody assessment test: Findings in three cases of Landau–Kleffner syndrome. *Journal of Neurolinguistics*, 25 (3): 194–211.

- Kim-Dufor, D-H., Ferragne, E., Dufor, O., Astésano, C.; Nespoulous, J-L. (2009a). From the perception of emotions to their expression. A study on linguistic and affective prosody problems in French-speaking High functioning autistic children. *Current Issues on Linguistic Interfaces, 1*, The SICOLI: 502-514.
- Kim-Dufor, D-H.; Ferragne, E.; Dufor, O.; Astésano, C.; Nespoulous, J-L. (2010). Perception and comprehension of linguistic & affective prosody in children with Landau-Kleffner syndrome. *Speech Prosody 2010*, 100885:1-4, ISBN:978-0-557-51931-6, Chicago.
- Kim-Dufor, D-H.; Ferragne, E.; Dufor, O.; Astésano, C.; Nespoulous, J-L. (2009b). Batterie de tests de perception/compréhension orale: Prosodie linguistique vs prosodie émotionnelle – étude comparative entre le syndrome de Landau-Kleffner et les troubles du spectre autistique. *MultiMod 2009*, Toulouse, 9-11 Juillet 2009
- Klatt, D.H. (1976). Linguistic uses of segmental duration in English: acoustic and perceptual evidence. *Journal of the Acoustic Society of America*, 59 (5) : 1208-1221.
- Konopczynski, G. (1986). *Du prélangage au langage: acquisition de la structuration prosodique*. Thèse de Doctorat d'État, Université de Strasbourg.
- Kolinsky, R. (1998). Spoken word recognition: a stage processing approach to language differences. *European Journal of Cognitive Psychology*, 10 : 1-40.
- Kotz, S. A. ; Meyer, M. ; Alter, K. ; Besson, M. ; von Cramon, D. Y. ; Friederici, A. D. (2003). On the lateralization of emotional prosody: an event-related functional MR investigation. *Brain and language*, 86(3) : 366-376.
- Lacheret-Dujour, A. & Beaugendre, F. (1999). *La prosodie du français*. Paris, Editions du CNRS.
- Ladd, D.R. (2017). Practical introduction to stress, accent and tone. *Conférence donnée lors du 23rd ABRALIN Summer Institute, UFF Niterói, March 2017*.
- Ladd, D. R. (2008). *Intonational Phonology*. Cambridge: Cambridge University Press (Première Edition 1996).
- Ladd, D.R. & Morton, R. (1997). The perception of intonational emphasis: continuous or categorical? *Journal of Phonetics*, 25 : 313-342.
- Ladd, D.R. & Campbell, W.N. (1991). Theories of prosodic structure: evidence from syllable duration. *ICPhS*, Aix-en-Provence, 2 : 290-293.

- Le Cam, E. *Les troubles de la communication chez les patients atteints de gliomes de bas grade au niveau de l'hémisphère droit*. Mémoire d'orthophonie, Université Toulouse1.
- Le Cam, E. ; Prodhomme, K ; Jiaa, M. ; Lubrano, V. ; Astésano, C. (2009) Elaboration d'un protocole de test en perception et production de la prosodie en contexte péri-opératoire pour les patients atteints de gliomes de bas-grades. *3ièmes Journées de Phonétique Clinique*, 4-5 Décembre 2009, Aix-en-Provence : 41.
- Levelt, W.J.M. (1989). *Speaking. From intention to articulation*. MIT Press, Cambridge, Mass.
- Lewandowski, N. (2012). *Talent in non-native phonetic convergence*. PhD Thesis, Universität Stuttgart.
- Liberman, M. (1975). *The Intonational System of English*. PhD. Dissertation, Cambridge, MIT.
- Liberman, M. & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry* (2) : 249-336.
- Lindblom, B. (1990). Explaining Phonetic Variation: A Sketch of the H&H Theory. In : *Speech Production and Speech Modelling*. Kluwer Academic Publishers, W.J. Hardcastle and A. Marchal (eds.), Dordrecht : 403-439.
- Magne, C.; Astésano, C.; Aramaki, M.; Ystad, S.; Kronland-Martinet, R.; Besson, M. (2007) Influence of Syllabic Lengthening on Semantic Processing in Spoken French: Behavioral and Electrophysiological Evidence. *Cerebral Cortex* 2007, 17(11): 2659-2668. doi: 10.1093/cercor/bhl174.
- Magne, C; Astésano, C.; Lacheret-Dujour, A.; Morel, M.; Alter, K. ; Besson, M. (2005a) On-line processing of “pop-out” words in spoken french dialogues. *Journal of Cognitive Neuroscience*, 17 (5): 740-756.
- Magne, C.; Aramaki, M.; Astésano, C.; Gordon, R.L.; Ystad, S.; Farner, S.; Kronland-Martinet, R.; Besson, M. (2005b). Comparison of rhythmic processing in language and music: an interdisciplinary approach. *JMM: The Journal of Music and Meaning* (3), [<http://www.musicandmeaning.net/issues/showArticle.php?artID=3.5>], sec.5.1.
- Magne, C.; Astésano, C.; Lacheret-Dujour, A.; Morel, M.; Besson, M. (2004). Traitement de l'accent de focalisation en français : approche électrophysiologique. *JEP 2004*, Fès, 19-22 Avril: 329-332.

- de Mareüil, P. B. ; Rilliard, A. ; Allauzen, A. (2012). A diachronic study of initial stress and other prosodic features in the French news announcer style: corpus-based measurements and perceptual experiments. *Language and Speech*, 55(2) : 263-293.
- Martin Philippe (2008). L'intonation du français : le vilain petit canard parmi les langues romanes ? *Language Design, Special Issue*, 2: 1-13.
- Martin, Ph. (1981). Pour une théorie de l'intonation – L'intonation est-elle une structure congruente à la syntaxe? In : Rossi, M.; Di Cristo, A.; Hirst, D.; Martin, Ph.; Nishinuma, Y. (eds.). *L'intonation: de l'acoustique à la sémantique*. Klincksieck, Paris : 234-271.
- Mattys, S. & Samuel, A. G. (2000). Implications of stress-pattern differences in spoken-word recognition. *Journal of Memory and Language*, 42(4) : 571-596.
- McClelland, J.L. ; St John, M. ; Taraban, R. (1989). Sentence comprehension: a parallel distributed processing approach. *Language and Cognitive Processes*, 4 : 287-335.
- Méndez, R. (en cours). *Perception de la structure métrique du Français: étude comparative entre des francophones et des hispanophones*. Thèse de Doctorat, Université de Toulouse2-Jean Jaurès.
- Méndez, R. & Astésano, C. (2017). Perception of the Downstepped Final Accent in French. *Proceedings of Phonetics and Phonology in Europe (PaPE) 2017*. Cologne, Germany. 12-14 June 2017.
- Mertens, P. (1992). L'accentuation de syllabes contiguës. *ITL-Review of Applied Linguistics* (95-96) : 145-165.
- Mertens, P. (1987). *L'intonation du français. De la description linguistique à la reconnaissance automatique*. Thèse de doctorat non publiée, Université de Leuven.
- Meyer, M.; Alter, K.; Friederici, A.D. ; Lohmann, G.; von Cramon, D.Y. (2002). Functional MRI reveals brain regions mediating slow prosodic modulations in spoken sentences. *Human Brain Mapping*, 17 : 73–88.
- Michaux, M. C. (2016). La perception de l'accent lexical néerlandais par les apprenants francophones. *Langages*, (2) : 47-74.
- Michelas, A. & D'Imperio, M. (2012). When syntax meets prosody: Tonal and duration variability in French Accentual Phrases. *Journal of Phonetics*, 40(6) : 816-829.

- Michelas, A. & Nguyen, N. (2011). Uncovering the effect of imitation on tonal patterns of French Accentual Phrases. *Interspeech 2011*, Florence, Italy : 973–976.
- Michelas, A. & D'Imperio, M. (2010a). Durational cues and prosodic phrasing in French: Evidence for the intermediate phrase. In *Speech Prosody 2010*, Chicago, 100881:1-4.
- Michelas, A., & D'Imperio, M. (2010b). Accentual phrase boundaries and lexical access in French. In *Speech Prosody 2010*, Chicago, 100882:1-4.
- Moreno, S.; Magne, C.; Schön, D.; Astésano, C.; Besson, M. (2004). Influence of musical training on fundamental frequency and pitch processing: Event-Related brain Potentials studies of adults and children. *International Conference on Tone and Intonation*, Santorini, 9-11 Sept.
- Näätänen R.; Paavilainen P.; Rinne T.; Alho K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology*, 118: 2544-2590.
- Nespor, M. & Vogel, I. (1986) *Prosodic Phonology*. Dordrecht: Foris.
- Nguyen, N. ; Wauquier, S. ; Tuller, B. (2009). The dynamical approach to speech perception: from fine phonetic detail to abstract phonological categories. In: Pellegrino, F., Marsico, E., Chitoran, I., & Coupé, C. (eds), *Approaches to Phonological Complexity*. Berlin: Mouton de Gruyter : 193–217.
- Nocaudie, O. (2016). *Imitation et contrôle prosodique dans l'entraînement à la remédiation phonétique : évaluation, mesure et applications pour l'enseignant en langue étrangère*. Thèse de Doctorat, Université de Toulouse2-Jean Jaurès.
- Nocaudie, O. ; Astésano, C. ; Woisard, V. (2017). Conservation des fonctions prosodiques post traitement des cancers de la cavité buccale et du pharynx. *7èmes Journées de Phonétique Clinique*, Paris, 29-30 juin 2017 : 59-60.
- Nocaudie, O. & Astésano, C. (2016a). Quelle mesure de la similarité prosodique comme évaluation de l'imitation ? *Journées d'Etude de la Parole 2016*, Paris : 580-588.
- Nocaudie, O. & Astésano, C. (2016b). Evaluating prosodic similarity as a means towards L2 teachers prosodic control training. *Speech Prosody 2016*, Boston : 26-30.

- Nocaudie, O. & Astésano, C. (2012). Prosodic structuring imitation in L1 French, a first step towards correcting phonetic-prosodic features in L2 French. *International Symposium on Imitation and Convergence in Speech*, Aix-en-Provence, 3-5 septembre.
- Pardo, J. S. (2013). Reconciling diverse findings in studies of phonetic convergence. *Meetings on Acoustics* (19) : 060140. Acoustical Society of America.
- Pasdeloup, V. (2005). Figures et fond dans la scène prosodique: leur résistance face aux variations du débit de parole. *Symposium Interface Discours-Prosodie-IDP05* (Aix-en-Provence, France).
- Pasdeloup, V. (1990). *Modèle de règles rythmiques du français appliqué à la synthèse de la parole*. Thèse de Doctorat, Université de Provence.
- Passy, P. (1892). *Les sons du français*, 3e édition, Paris, Firmin-Didot.
- Patel, A. D. (2010). *Music, language, and the brain*. Oxford university press.
- Portes, C. ; D'Imperio, M. ; Lancia, L. (2012). Positional constraints on the initial rise in French. *Speech Prosody* : 563-566.
- Post, B. (2000). *Tonal and Phrasal Structures in French Intonation*. PhD Thesis, The Hague.
- Pierrehumbert, J. (1980). *The Phonology and Phonetics of English Intonation*. PhD Thesis, Cambridge: MIT, publiée (1988): Indiana University Linguistic Club.
- Pisoni, D. B. ; Manous, L. M. ; Dedina, M. J. (1987). Comprehension of natural and synthetic speech: Effects of predictability on the verification of sentences controlled for intelligibility. *Computer speech & language*, 2(3-4) : 303-320.
- Rauzy, S. ; Montcheuil, G. ; Blache, P. (2014). MarsaTag, a tagger for French written texts and speech transcriptions. *Proceedings of Second Asian Pacific Corpus linguistics Conference* : 220.
- Repp, B. (1992). Probing the cognitive representation of musical time: structural constraints on the perception of timing perturbations. *Haskins Laboratories Status Report on Speech Research*, 111/112 : 293-320.
- Rilliard, A. ; Allauzen , A. ; Boula de Mareüil , P. (2011). Using Dynamic Time Warping to Compute Prosodic Similarity Measures. *Interspeech*, Florence : 2021– 2024.

- Rolland, G. & Loevenbruck, H. (2002). Characteristics of the Accentual Phrase in French: an acoustic, articulatory and perceptual study. *Speech Prosody 2002*, Aix en Provence : 611-614.
- Rossi, M. (2011). L'intonation modale. *Modèles linguistiques*, 63 :117-129.
- Rossi, M. (1999). *L'intonation, le système du français : description et modélisation*. Paris, France: OPHRYS.
- Rossi, M. (1997). Is Syntactic Structure Prosodically Retrievable. *Proceedings of 5th European Conference of Speech Communication and Technology*, Greece, 1 : 1-8.
- Rossi, M. (1985a). L'intonation et l'organisation de l'énoncé. *Phonetica*, 42 : 135-153.
- Rossi, M. (1985b). De la quiddité des variables. In : *Variabilité et spécificité du locuteur*, Société Française d'Acoustique, Marseille-Luminy : 11-27.
- Rossi, M. (1981a). Intonation, énonciation, syntaxe. In : *L'intonation: de l'acoustique à la sémantique*, Rossi, M.; Di Cristo, A.; Hirst, D.; Martin, Ph.; Nishinuma, Y. (eds.), Klincksieck, Paris : 184-233.
- Rossi, M. (1980). Le français, langue sans accent ?. In *L'accent en français contemporain (Studia Phonetica)*, 15, I. Fónagy & P. Léon (eds.) : 13-51
- Rossi, M. (1977). L'intonation et la troisième articulation. *Bulletin de la Société de Linguistique de Paris*, 72(1) : 55-68.
- Rossi, M. (1972). Le seuil différentiel de durée. *Papers in Linguistics and Phonetics to the memory of Pierre Delattre*, 54 : 435.
- Roulet-Guiot, G.; Astésano, C. (2012). Lire les tons sur les lèvres : perception(s) visuelle(s) des tons lexicaux en chinois mandarin. *JEP-TALN-RECITAL 2012*, Grenoble, 4-8 juin 2012: 577-584.
- Roux, G. ; Bertrand, R. ; Ghio, A. ; Astésano, C. (2016) Naïve listeners' perception of prominence and boundary in French spontaneous speech. *Speech Prosody 2016*, Boston : 912-916.
- Séguinot, A. (1976). L'accent d'insistance en français standard. *Studia Phonetica*, 12, 1-58.
- Selkirk, E. (1996). The prosodic structure of function words. In : K. Demuth and J. Morgan (eds.): *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah, NJ: Lawrence Erlbaum : 187-213.

- Selkirk, E. (1986). On derived domains in sentence phonology. *Phonology Yearbook*, (3) : 371-405.
- Selkirk, E. (1984a). *Phonology and Syntax. The Relation between Sound and Structure*. Cambridge, MA: MIT Press.
- Selkirk, E. (1984b). On the Major Class Features and Syllable Theory. In : *Language Sound Structure*, M. Aronoff and R.T. Oehrle (eds.), MIT Press : 107-136.
- Selkirk, E. (1980). Prosodic domains in phonology: Sanskrit revisited. In M. Aronoff and M.-L. Kean (eds), *Juncture*, Anna Libri, PO Box 876, Saratoga, Calif. : 107-129.
- Selkirk, E. (1978). On Prosodic Structure and its Relation to Syntactic Structure. *Nordic Prosody*, 2 : 111-140.
- Shattuck-Hufnagel, S. ; Ostendorf, M. ; Ross, K. (1994). Stress shift and early pitch accent placement in lexical items in American English. *Journal of Phonetics*, 22 : 357 – 388.
- Shattuck-Hufnagel, S. & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of psycholinguistic research*, 25(2) : 193-247.
- Sichel-Bazin, R. (2015). *Prosodic systems in contact: Occitan and French*. PhD thesis, Universität Osnabrück / Univ. Pompeu Fabra.
- Sichel-Bazin, R. (2016). Le français du Midi: une variété de français à accent lexical. *Langue française*, (3) : 47-66.
- Sluming, V. ; Barrick, T. ; Howard, M. ; Cezayirli, E. ; Mayes, A. ; Roberts, N. (2002). Voxel-based morphometry reveals increased gray matter density in Broca's area in male symphony orchestra musicians. *Neuroimage*, 17(3) : 1613-1622.
- Steinhauer, K.; Alter, K.; Friederici, A.D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Neuroscience*, 2 (2) : 191-196.
- Swerts, M. & Geluykens, R. (1994). Prosody as a marker of information flow in spoken discourse. *Language and Speech*, 37(1) : 21-43.
- te Rietmolen, N. ; El Yagoubi, R. ; Ghio, A. ; Astésano, C. (2017). The phonological status of the French initial accent and its role in semantic processing: an Event-Related Potentials study, *Interspeech 2017*, Stockholm.
- te Rietmolen (en cours). Etudes neurophysiologiques du traitement de l'accentuation en français. Thèse de Doctorat, Université de Toulouse2-Jean Jaurès.

- te Rietmolen, N. ; El Yaboubi, R. ; Espesser, R. ; Magnen, C. ; Astésano, C. (2016a). Investigating the phonological status of the initial accent in French : an ERP study, *Speech Prosody 2016*, Boston : 1181-1185.
- te Rietmolen, N. ; El Yaboubi, R. ; Espesser, R. ; Magnen, C. ; Astésano, C. (2016b). Investigating the phonological status of the Initial Accent in French: an Event-Related Potentials study. Poster session presented at the *AMLaP 2016 conference*, Bilbao, Spain: 70.
- te Rietmolen, N. ; Astésano, C. ; El Yaboubi, R. ; Thorpe, S. ; Kopke, B. (2016c). Project NeuroPros: EEG investigation of prosodic cues processing in French. Poster session presented at the *Nijmegen Lectures 2016*, Nijmegen, the Netherlands.
- te Rietmolen, N. & Astésano, C. (2016d). EEG entrainment to the prosodic structure in spoken language. Poster session presented at the *AMLaP 2016 conference*, Bilbao, Spain: 71.
- Traverso, V. (2012). Analyses interactionnelles: repères, questions saillantes et évolution. *Langue française*, (3) : 3-17.
- Vaissière, J. (1997). Ivan Fónagy et la notation prosodique. In : *Polyphonie pour Ivan Fónagy*, Jean Perrot (ed.), L'Harmattan, Paris : 479-487.
- Vaissière, J. (1992). « Rhythm, accentuation and final lengthening in French. ». In : Sundberg, J. et al. (eds.). *Music, Language, Brain and Speech, Wenner-Gren International Symposium Series*, 59 : 108-120.
- Vaissière, J. (1991). Perceiving rhythm in French?. *ICPhS'91*, 4, Aix-en-Provence, 258-261.
- Wauquier-Gravelines, S. (1999). Segmentation lexicale de la parole continue: la linéarité en question. *Recherches linguistiques de Vincennes*, (28), 133-156.
- Welby, P. ; Bertrand, R. ; Portes, C. ; Astésano, C. (2016). Realization of the French initial accent: Stability and individual differences. *TIE Conference 2016 – 7th conference on Tone and Intonation in Europe (TIE)*. Canterbury. 1-3 September 2016.
- Welby, P. & Niebuhr, O. (2016). The influence of F0 discontinuity on intonational cues to word segmentation: A preliminary investigation. *Speech Prosody 2016*, Boston : 40-44.

- Welby, P. (2007). The role of early fundamental frequency rises and elbows in French word segmentation. *Speech Communication*, 49(1), 28-48.
- Welby, P. (2003). *The Slaying of Lady Mondegreen, being a study of French tonal association and alignment and their role in speech segmentation*. Thèse de Doctorat, Université Ohio.
- Wenk, B.J. (1983). Effets de rythme dans le français parlé. *Recherches sur le français parlé*, 5, Publications de l'Université de Provence, Aix-en-Provence : 147-162.
- Wenk, B.J. & Wioland, F. (1982). Is French really syllable-timed? *Journal of Phonetics*, 10(2), 193-216.
- Woisard, V.; Espesser R.; Ghio A.; Duez, D. (2013). De l'intelligibilité à la compréhensibilité de la parole, quelles mesures en pratique clinique ? *Revue de laryngologie, otologie, rhinologie*, 1(134) : 27-33.
- Ylinen, S.; Strelnikov, K.; Huotilainen, M.; Näätänen, R. (2009). Effects of prosodic familiarity on the automatic processing of words in the human brain. *International Journal of Psychophysiology*, 73 : 362-368.
- Ystad, S.; Magne, C.; Farner, S.; Pallone, G.; Astésano, C.; Padeloup, V.; Kronland-Martinet, R.; Besson, M. (2003). Influence of rhythmic, harmonic and semantic violations in language and music on the electrical activity of the brain. *The Stockholm Music Acoustics Conference*, Stockholm, Août 6-9, 2003.
- Zatorre, R. J. & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cereb. Cortex*, 11 : 946-953.