



Caractérisation phonétique et phonologique du syntagme intermédiaire en français : de la production à la perception.

Amandine Michelas

► To cite this version:

Amandine Michelas. Caractérisation phonétique et phonologique du syntagme intermédiaire en français : de la production à la perception.. Linguistique. Université de Provence - Aix-Marseille I, 2011. Français. NNT : . tel-00764624

HAL Id: tel-00764624

<https://theses.hal.science/tel-00764624>

Submitted on 13 Dec 2012

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

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

DOCTORAT AIX-MARSEILLE UNIVERSITÉ
Délivré par l'UNIVERSITÉ DE PROVENCE

N° attribué à la bibliothèque

Caractérisation phonétique et phonologique du syntagme intermédiaire en français : de la production à la perception

THÈSE

présentée pour l'obtention du grade de
Docteur d'Aix-Marseille Université

Champ disciplinaire : Sciences du Langage
Spécialité : Phonétique et Phonologie

par

Amandine MICHELAS

Composition du jury

Président : Albert DI CRISTO, Professeur Emérite, Département de Linguistique et Phonétique Générales, Laboratoire Parole et Langage, UMR 6057 CNRS – Université de Provence

Rapporteurs : Annie RIALLAND, Directeur de Recherche, Laboratoire de Phonétique et Phonologie, UMR 7018 CNRS – Université Sorbonne-Nouvelle, Paris 3

Caroline L. SMITH, Professeur, Département de Linguistique – Université du Nouveau Mexique, Etats-Unis

Examineur : Ulrich Hans FRAUENFELDER, Professeur, Département de Psychologie – Université de Genève, Suisse

Directeur : Mariapaola D'IMPERIO, Professeur, Département de Linguistique et Phonétique Générales, Laboratoire Parole et Langage, UMR 6057 CNRS – Université de Provence

à Flo,

REMERCIEMENTS

Me voici arrivée en fin de thèse : à la croisée des chemins entre la fin d'une aventure et le début d'une nouvelle. Je voudrais profiter de cet instant pour remercier tous ceux qui ont contribué de près ou de loin à ce travail.

En premier lieu je tiens à remercier chaleureusement Caroline L. Smith et Annie Rialland pour avoir accepté d'être mes rapporteurs. Je remercie également Albert Di Cristo et Ulrich Hans Frauenfelder pour avoir accepté de faire partie de mon jury de thèse. C'est un grand honneur pour moi que vous ayez accepté de juger mon travail. J'espère que vous apprécierez cette thèse autant que j'ai pris plaisir à l'élaborer.

Il y a une personne que je tiens tout particulièrement à remercier : Mariapaola D'Imperio. Je ne sais trop comment lui dire en quelques lignes combien je lui suis reconnaissante de tout ce qu'elle m'a apporté. Elle a su tout d'abord me couvrir telle une «mamma napolitaine» afin que je sois suffisamment solide pour exposer ma recherche et mes idées, puis m'a encouragé à m'affirmer en tant que chercheur. Elle a été mon guide scientifique pendant ces trois années. Je la remercie du fond du cœur pour toutes les longues discussions que nous avons pu avoir et qui m'ont aidée à avancer. J'ai adoré grandir scientifiquement sous son aile et j'espère qu'elle sera fière du parcours que je m'apprête aujourd'hui à accomplir.

J'ai réalisé ma thèse au sein du Laboratoire Parole et Langage, laboratoire qui accorde une grande importance à ses doctorants et au sein duquel j'ai pris tant de plaisir à réaliser cette thèse. En premier lieu, je tiens à remercier Philippe Blache pour avoir cru en mon projet et m'avoir aidée à obtenir mon allocation de recherche. Je tiens également à remercier la direction du laboratoire : Noël Nguyen et Christine Meunier pour m'avoir toujours soutenue tant humainement que financièrement. Au cours de ces trois années j'ai eu la chance de pouvoir participer à de nombreux congrès qui m'ont permis de rencontrer de nombreux chercheurs qui m'ont aidée à avancer. Je remercie grandement le LPL pour m'avoir donné cette chance. Je remercie également tout particulièrement Noël Nguyen pour m'avoir donné l'opportunité de travailler sur le projet ANR SPIM avant même d'avoir achevé ce travail de thèse.

Travailler au sein du LPL m'a permis de rencontrer beaucoup de gens que je ne pourrai malheureusement pas tous citer ici. Je tiens cependant à adresser tout particulièrement mes remerciements à quelques personnes du LPL. En premier lieu Robert Espesser pour avoir pris autant de temps pour répondre à mes questions. Je le remercie de m'avoir attirée vers les modèles mixtes qui demeureront un grand mystère pour moi s'il n'avait pas été là. Je remercie également Pauline Welby pour tous ses encouragements, le partage de son expérience et tous ses précieux conseils. Je tiens également à remercier Sophie Dufour qui a toujours été présente pour répondre à la moindre de mes questions ainsi que Cristel Portes et Roxane Bertrand pour leur soutien qui m'a fait me sentir au sein d'une véritable famille « Prosodie » à Aix. Enfin je ne

peux oublier le personnel technique du laboratoire qui m'a aidée à me servir de la plateforme technique sans laquelle les expérimentations présentées dans cette thèse auraient difficilement été réalisables. Même s'il n'appartient plus tout à fait au LPL, je tiens également à remercier James German, pour ses relectures, ses conseils et nos discussions. Je suis très heureuse de la collaboration que nous avons entamée tous les deux et suis sûre qu'elle donnera des résultats prometteurs.

Au LPL, de nombreux doctorants m'ont apporté leur soutien. Je les remercie tous de tout mon cœur, que ce soit les anciens, les actuels ou les derniers arrivés. Plus que des collègues j'ai gagné parmi eux de véritables amis et je les remercie pour l'immense joie et les nombreux fous rires qu'ils m'ont apportés. Je remercie également Francesco Cangemi et Catarina Petrone pour avoir été mon frère et ma sœur au sein de la *Mariapaola's team*. J'ai adoré nos discussions et nos « combats » sur la phonologie de l'intonation qui m'ont fait réfléchir bien plus que vous ne l'imaginez tous les deux.

Au cours de ma thèse j'ai eu l'occasion de séjourner pendant plus de quatre mois au sein du Département de Psychologie Expérimentale de Bristol. Je remercie Sven Mattys pour m'avoir si chaleureusement ouvert les portes de son équipe « Langage » et m'avoir appris que ce n'est pas parce qu'on perçoit des indices qu'on les utilise dans le but de traiter le langage. Je remercie également Laurence White pour son rire si communicatif, sa franchise, sa bonne humeur et le fait de me rappeler sans cesse que, bien que l'anglais et le français soit deux langues aux structures prosodiques bien différentes, elles se rapprochent également en bien des points. Ce séjour a été pour moi une étape importante de ma thèse et je remercie Sven Mattys, Olesya Rauch et Laurence White pour l'avoir rendu si agréable.

Ces années de thèse m'ont permis de rencontrer de nombreux chercheurs et je les remercie tous du temps qu'ils ont pris pour m'écouter et me conseiller. Je voudrais en particulier remercier Marina Vigário, Sónia Frota, Pilar Prieto et Ingo Feldhausen. Je remercie également Elizabeth Selkirk et Hubert Truckenbrodt rencontrés lors de la conférence *Interfaces Discours et Prosodie* en 2010 à Paris puis lors de leur passage à Aix-en-Provence, grâce à qui est née une des expérimentations présentées dans cette thèse. Je remercie également Anne Christophe pour ses conseils et le partage de ses stimuli.

Sur un plan privé je tiens à remercier mes amis et ma famille qui ont su me soutenir et m'entourer de tout leur amour. Un grand merci à mes parents dans les yeux desquels je peux lire une grande fierté, à ma petite sœur qui est pour moi un bel exemple de volonté et de persévérance. Merci également à mes trois amis Christine, Anne-Sophie et Marie qui ont su me pardonner mes absences et sans qui la vie ne serait pas aussi belle. Enfin (et surtout) je remercie du fond du cœur Florian sans qui rien de tout ça n'aurait été pareil. Tu as su partager tous les moments forts de cette thèse, les bons comme les mauvais. Tu as su m'entourer de tant d'amour de passion et de tendresse que j'ai pu, grâce à ton aide, surmonter tous les moments difficiles. Je suis très fière et immensément heureuse de partager ta vie et c'est donc tout naturellement que je te dédie cette thèse.

TABLE DES MATIERES

REMERCIEMENTS	i
TABLE DES MATIERES.....	iii
TABLE DES ILLUSTRATIONS.....	vi
INDEX DES TABLEAUX	ix
INDICES DE LECTURE	1
CHAPITRE 1 INTRODUCTION	2
1. PROBLÉMATIQUE	2
2. LA STRUCTURE PROSODIQUE.....	7
2.1. La hiérarchie des domaines prosodiques.....	7
2.2. Les relations entre structure prosodique et structure syntaxique.....	9
2.3. La définition et le nombre de constituants prosodiques.....	10
3. PROPRIÉTÉS PHONÉTIQUES ET PHONOLOGIQUES DES FRONTIÈRES PROSODIQUES	12
3.1. Propriétés acoustiques	12
3.2. Propriétés articulatoires	15
4. LA STRUCTURE PROSODIQUE DANS LE MODÈLE AUTO-SEGMENTAL MÉTRIQUE DE L'INTONATION.....	16
4.1. Le modèle auto-segmental métrique de l'intonation.....	16
4.2. Les structures prosodiques proposées pour le français dans le cadre du modèle auto-segmental métrique de l'intonation	19
5. OBJECTIFS DE LA THÈSE	30
CHAPITRE 2 INFLUENCE DE LA SYNTAXE ET DE LA VITESSE D'ÉLOCUTION SUR LA RÉALISATION DES FRONTIÈRES DE SYNTAGME ACCENTUEL EN FRANÇAIS (Étude 1)	32
1. PROBLÉMATIQUE	32
2. MÉTHODE	39
2.1. Matériel utilisé.....	39
2.2. Participants et procédure expérimentale	41
2.3. Recueil et analyse des données	42
2.4. Analyses statistiques	45
3. RÉSULTATS	46
3.1. Manipulation de la vitesse d'élocution.....	46
3.2. Phénomènes de restructuration prosodique observés à vitesse d'élocution rapide	48
3.3. Distribution des patrons mélodiques	49
3.4. Résultats relatifs à la durée.....	58
3.5. Résultats relatifs à la fréquence fondamentale (fo).....	61
4. DISCUSSION	63
4.1. Influence de la vitesse d'élocution sur la réalisation des frontières d'AP..	65
4.2. Facteurs influençant la réalisation du patron mélodique de l'AP.....	69
4.3. Influence de la syntaxe sur la réalisation des frontières d'AP.....	72

CHAPITRE 3 EFFETS D'ABAISSEMENT, VARIATIONS DE REGISTRE ET DÉFINITION DU SYNTAGME INTERMÉDIAIRE EN FRANÇAIS (Étude 2)	74
1. PROBLÉMATIQUE	74
2. HYPOTHÈSES.....	84
3. MÉTHODE	87
3.1. Matériel linguistique.....	87
3.2. Participants et procédure expérimentale	90
3.3. Recueil et analyse des données	90
4. RÉSULTATS	93
4.1. Manipulation de la vitesse d'élocution.....	93
4.2. Phénomènes de restructuration prosodique.....	95
4.3. Résultats relatifs à la durée.....	99
4.4. Résultats relatifs à la fo	102
5. DISCUSSION	112
5.1. Résultats principaux.....	112
5.2. Preuves de l'existence du syntagme intermédiaire.....	113
CHAPITRE 4 PERCEPTION DE LA FRONTIÈRE DE SYNTAGME ACCENTUEL ET ACCÈS AU LEXIQUE (Étude 3)	121
1. PROBLÉMATIQUE	121
2. MÉTHODE	133
2.1. Matériel expérimental	133
2.2. Participants et Procédure	136
3. RÉSULTATS	136
3.1. Analyse acoustique des stimuli	136
3.2. Analyse des temps de réaction	141
4. DISCUSSION	143
CHAPITRE 5 PERCEPTION DE LA FRONTIÈRE DROITE DE SYNTAGME INTERMÉDIAIRE ET TRAITEMENT DE LA STRUCTURE SYNTAXIQUE (Études 4 et 5)	148
1. PROBLÉMATIQUE	148
2. ÉTUDE 4: À LA DÉCOUVERTE DE L'INFLUENCE DE LA FRONTIÈRE DROITE DU SYNTAGME INTERMÉDIAIRE DANS LA CONSTRUCTION DE LA STRUCTURE SYNTAXIQUE.....	160
2.1. Méthode	162
2.2. Résultats	165
2.3. Discussion.....	172
3. ÉTUDE 5 : QUEL EST L'INDICE ACOUSTIQUE LE PLUS SAILLANT DANS LA PERCEPTION DES FRONTIÈRES DE SYNTAGME INTERMÉDIAIRE EN FRANÇAIS ?	175
3.1. Méthode	178
3.2. Résultats	181
3.3. Discussion.....	185
4. DISCUSSION GÉNÉRALE.....	186
4.1. Résultats généraux.....	186
4.2. Modèles de l'intonation proposés pour le français.....	187
4.3. Modèles de traitement syntaxique	188
CONCLUSIONS GÉNÉRALES	190
1. RAPPEL DE LA PROBLÉMATIQUE INITIALE	190
2. RÉSULTATS MAJEURS	191

2.1. Étude 1 : Influence de la syntaxe et de la vitesse d'élocution sur la structuration prosodique.....	191
2.2. Étude 2 : Caractéristiques phonétiques et phonologiques du syntagme intermédiaire en français.....	192
2.3. Étude 3 : Influence de la frontière droite de syntagme accentuel sur l'accès au lexique.....	193
2.4. Étude 4 : Perception de la frontière droite de syntagme intermédiaire et traitement syntaxique de la parole.....	194
2.5. Étude 5 : Contribution séparée des indices mélodiques et de durée dans la construction de la syntaxe.....	195
3. DIRECTIONS POUR NOS RECHERCHES FUTURES	196
3.1. Parole de laboratoire et parole spontanée.....	196
3.2. Frontière de syntagme intermédiaire et structure informationnelle.....	196
3.3. Corrélats articulatoires du syntagme intermédiaire	198
3.4. Application aux technologies de la parole.....	198
3.5. Sensibilité des jeunes enfants aux frontières prosodiques.....	199
RÉFÉRENCES	202
ANNEXES	213
1. ANNEXES A : Corpus utilisé pour l'étude 1 (Chapitre 2)	213
1.1. Stimuli.....	213
1.2. Consignes.....	215
2. ANNEXES B : Corpus utilisé pour l'étude 2 (Chapitre 3)	216
2.1. Stimuli.....	216
2.2. Consignes.....	216
3. ANNEXE C : Expérience pilote 1, Frontière d'ip et frontière d'IP en français	217
3.1. Méthodologie.....	218
3.2. Résultats	220
3.3. Conclusions.....	221
4. ANNEXE D : Expérience pilote 2, Le syntagme intermédiaire en position Objet à l'intérieur des énoncés.	223
4.1. Méthodologie	224
4.2. Résultats	225
4.3. Conclusions.....	226
5. ANNEXES E : Corpus utilisé pour l'étude 3 (Chapitre 4)	228
5.1. Stimuli.....	228
5.2. Consignes.....	229
6. ANNEXES F : Corpus utilisé pour les études 4 et 5 (Chapitre 5)	230
6.1. Stimuli.....	230
6.2. Consignes.....	231
RÉSUMÉ	233

TABLE DES ILLUSTRATIONS

Figure 1. Illustration de la séquence sonore /zãlev/ à vitesse d'élocution normale.....	4
Figure 2. Illustration de la séquence sonore /zãlev/ à vitesse d'élocution rapide.....	5
Figure 3. Schématisation du modèle proposé par Truckenbrodt pour l'allemand.	15
Figure 4. Grammaire de l'intonation de l'anglais selon la théorie auto-segmentale métrique de l'intonation.	18
Figure 5. Illustration du syntagme nominal « les chatons sages ».	20
Figure 6. Illustration du segment d'unité intonative selon Di Cristo & Hirst (1996).....	23
Figure 7. Spécifications tonales du syntagme intonatif selon Post (2000, p°155).	24
Figure 8. Hiérarchie prosodique postulée pour le français tirée de Jun & Fougeron (2000, 2002).....	26
Figure 9. Les six réalisations de surface du syntagme accentuel tiré de Jun & Fougeron (2002).	26
Figure 10. Illustration du syntagme intermédiaire selon Jun & Fougeron (2000).	28
Figure 11. Degrés d'allongement pré-frontière tirés de Jun & Fougeron (2000, p°221).....	34
Figure 12. Étude 1 : Hypothèse nulle.....	37
Figure 13. Étude 1 : Hypothèse d'un degré d'allongement pré-frontière additionnel.	38
Figure 14. Étude 1 : Exemple de phrases du corpus.	40
Figure 15. Étude 1 : Exemple de patrons mélodiques observés.	44
Figure 16. Étude 1 : Exemple de modifications de la structure prosodique à vitesse d'élocution rapide.	49
Figure 17. Étude 1 : Montée initiale et montée finale de fo.....	52
Figure 18. Étude 1 : Exemple d'un patron mélodique de type /LH*/.	54
Figure 19. Étude 1 : Probabilités d'apparition du patron mélodique de type /LH*/ estimées par le modèle MLM pour les APs situés en position non finale d'IP.	56
Figure 20. Étude 1 : Probabilités d'apparition du patron mélodique de type /LH*/ estimées par le modèle MLM pour les APs situés en position finale d'IP.....	56
Figure 21. Étude 1 : Valeurs de durées estimées par le modèle LMM.....	59
Figure 22. Étude 1 : Valeurs des pics de fo estimées par le modèle LMM.....	62
Figure 23. Étude 1 : frontière d'AP marquée uniquement par les indices de durée.	68
Figure 24. Étude 1 : réalisation de la montée initiale de fo.....	71
Figure 25. Phénomène d'abaissement des accents mélodiques successifs en anglais tiré de Beckman & Pierrehumbert (1986).....	76
Figure 26. Schématisation du downstep et de l'abaissement final.....	76
Figure 27. Schématisation de la déclinaison.....	77
Figure 28. Exemple de blocage du downstep.	78
Figure 29. Schématisation de la transition entre un accent mélodique et un accent de syntagme L-.....	79

Figure 30. Schématisation de la structure grammaticale des phrases utilisées par Ladd (1988).	80
Figure 31. Schématisation de l'abaissement entre les constituants prosodiques observés par Ladd (1988).	80
Figure 32. Phénomène de réinitialisation partielle du registre observé en hollandais.	81
Figure 33. Schématisation de l'abaissement des accents mélodiques et de l'abaissement entre les constituants prosodiques.	82
Figure 34. Schématisation du modèle proposé par Truckenbrodt (2002) pour l'allemand.	83
Figure 35. Illustration du phénomène de final upstep et de réinitialisation partielle du registre.	83
Figure 36. Étude 2 : Schématisation de la structure prosodique attendue.	87
Figure 37. Étude 2 : Structures prosodiques obtenues.	89
Figure 38. Étude 2 : Exemple de l'annotation réalisée.	92
Figure 39. Étude 2 : Vitesse d'élocution pour les énoncés.	94
Figure 40. Étude 2 : Vitesse d'élocution pour les APs cibles.	94
Figure 41. Étude 2 : Structure prosodique obtenue dans les différentes conditions.	97
Figure 42. Étude 2 : Illustration du phénomène d'effacement de frontière d'AP sur « sauna » à vitesse d'élocution rapide.	98
Figure 43. Étude 2 : Illustration d'une frontière d'AP marquée uniquement par des indices de durée.	99
Figure 44. Étude 2 : Durée de la voyelle en position finale du deuxième syntagme accentuel (V2).	100
Figure 45. Étude 2 : Durée de la syllabe en position finale du deuxième syntagme accentuel (S2).	102
Figure 46. Étude 2 : Hauteur de la voyelle en position finale du deuxième syntagme accentuel (H2).	103
Figure 47. Étude 2 : Hauteur de la voyelle en position finale du deuxième syntagme accentuel (H2) par rapport à H1.	104
Figure 48. Étude 2 : Illustration du premier pic mélodique situé après la frontière d'ip.	106
Figure 49. Étude 2 : Schématisation du phénomène de réinitialisation partielle du registre.	107
Figure 50. Étude 2 : Rapport de Hpb par H-.	108
Figure 51. Étude 2 : Rapports de H- et de Hpb par H1 estimés par le modèle LMM.	109
Figure 52. Étude 2 : Rapports de H2 et de Hpb par H1.	110
Figure 53. Étude 2 : Rapports de H2 et de Hpb par H1 estimés par le modèle LMM.	111
Figure 54. Étude 2 : Exemple de frontière d'ip.	115
Figure 55. Frontières d'ip et frontières d'IP.	116
Figure 56. Exemple d'ip en position Objet dans la phrase.	118
Figure 57. Schématisation de la réinitialisation partielle du registre sur le premier accent mélodique situé après la frontière d'ip.	118
Figure 58. Illustration d'un cas d'ambiguïté lexicale à cheval sur une frontière de syntagme intonatif.	125
Figure 59. Exemples de deux découpages prosodiques possibles pour le même syntagme nominal « le chat grincheux ».	129
Figure 60. Étude 3 : Exemples d'ambiguïtés lexicales temporaires pour les trois conditions prosodiques.	132
Figure 61. Étude 3 : Durées moyennes des voyelles finales des mots cibles.	138

Figure 62. Étude 3 : Hauteur de la voyelle finale des mots cibles (en Hz) dans les différentes conditions prosodiques.	139
Figure 63. Étude 3 : Exemple d'énoncé présentant une montée initiale de fo (LHi) sur l'AP suivant l'AP cible.	141
Figure 64. Étude 3 : Temps de réaction des participants.	142
Figure 65. Étude 3 : Temps de réaction estimés par le modèle LMM.	143
Figure 66. Deux réalisations possibles de la même phrase induisant deux structures prosodiques différentes. La phrase est tirée du corpus de Millotte (2008).	155
Figure 67. Étude 4 : exemples d'une séquence dont la voyelle finale est associée soit à une frontière d'AP, soit à une frontière d'ip.	159
Figure 68. Étude 4 : Annotations réalisées.	166
Figure 69. Étude 4 : Durées des voyelles cibles V1 et V2 (en ms) dans les deux conditions prosodiques AP et ip.	167
Figure 70. Étude 4 : Hauteur moyenne des voyelles cibles H1 et H2 (en Hz) en fonction de la condition prosodique (AP et ip).	168
Figure 71. Étude 4 : Rapports de H2 par H1 en fonction de la condition prosodique (AP et ip).	169
Figure 72. Étude 4 : Pourcentage moyen de réponses de type syntagme prépositionnel (SP) ou syntagme verbal (SV).	171
Figure 73. Étude 4 : Probabilités du type de réponses (réponse SP ou réponse SV) estimées par le modèle MLM.	172
Figure 74. Étude 5 : Exemple de stimuli pour les 6 conditions prosodiques (AP-mel, ip-mel, AP-dur, ip-dur, AP et ip).	178
Figure 75. Étude 5 : Pourcentage moyen de réponses syntagme prépositionnel (SP) données par les participants en fonction du type de stimuli entendu (AP-mel, ip-mel, AP-dur, ip-dur, AP et ip).	182
Figure 76. Étude 5 : Probabilités de réponses SP données par les participants estimées par le modèle MLM.	183
Figure 77. Étude 5 : Temps de réaction des participants (ms) pour chaque type de stimuli (AP-mel, ip-mel, AP-dur, ip-dur, AP et ip) en fonction du type de réponse donné par les participants (Réponse SP / Réponse SV).	184
Figure 78. ip et structure informationnelle.	197
Figure 79. Pilote 1 : Exemple d'annotations réalisées.	219
Figure 80. Pilote 1 : Durée des voyelles cibles exprimées en millisecondes.	220
Figure 81. Pilote 1 : Rapports des valeurs de fo des voyelles cibles par les valeurs du premier pic mélodique de l'énoncé (H1).	221
Figure 82. Pilote 2 : Exemples de phrases du corpus.	222
Figure 83. Pilote 2 : Durées des voyelles cibles exprimées en millisecondes.	225
Figure 84. Pilote 2 : Hauteur de la voyelle en position finale du deuxième syntagme accentuel (H2) exprimée par rapport à la hauteur du premier pic de fo du SN objet (H1).	226
Figure 85. Pilote 2 : Exemples de phrases du corpus.	227

INDEX DES TABLEAUX

Tableau 1. Étude 1 : Exemple de phrases utilisées dans le corpus.....	39
Tableau 2. Étude 1 : Caractéristiques de la vitesse d'élocution.....	47
Tableau 3. Étude 1 : Pourcentage de réalisation des différents patrons mélodiques pour les conditions AP-interne, AP-finale et AP/XP.....	50
Tableau 4. Étude 1 : Pourcentage de réalisation des différents patrons mélodiques pour la condition IP-finale.	50
Tableau 5. Étude 1 : Durées des voyelles (ms) estimées par le modèle pour les quatre types de frontières prosodiques (AP-interne, AP-finale, AP/XP, IP- finale).	60
Tableau 6. Étude 1 : Durées des syllabes (ms) estimées par le modèle pour les quatre types de frontières prosodiques (AP-interne, AP-finale, AP/XP, IP- finale).	61
Tableau 7. Étude 1 : Valeurs de fo (Hz) estimées par le modèle pour les quatre types de frontières prosodiques (AP-interne, AP-finale, AP/XP, IP-finale).	63
Tableau 8. Étude 2 : Phrases utilisées dans le corpus.....	88
Tableau 9. Étude 3 : Exemple d'énoncés utilisés dans le corpus pour les différentes conditions.	131
Tableau 10. Étude 3 : Moyenne des durées des voyelles cibles (en ms).	139
Tableau 11. Étude 3 : Moyenne des valeurs de fo associées aux voyelles cibles (en Hz).	140
Tableau 12. Étude 4 : Moyenne des durées (ms) des voyelles cibles (V1 et V2) en fonction de la condition prosodique (AP vs. ip).	168
Tableau 13. Étude 4 : Valeurs moyennes de fo (en Hz) pour H1 et H2 et valeurs moyennes du rapport de H2 par H1 en fonction de la condition prosodique (AP vs. ip).	170
Tableau 14. Étude 5 : Six types de stimuli de resynthèse.....	176

INDICES DE LECTURE

Les abréviations suivantes seront utilisées de façon récurrente dans ce manuscrit afin de rendre la lecture plus claire et plus synthétique.

AP	Syntagme accentuel (<i>Accentual Phrase</i>)
ip	Syntagme intermédiaire (<i>Intermediate Phrase</i>)
IP	Syntagme intonatif (<i>Intonation Phrase</i>)
AP/XP	Syntagme accentuel associé à une rupture syntaxique majeure
MP	Mot prosodique
PP	Syntagme phonologique (<i>Phonological Phrase</i>)
SN	Syntagme nominal
SV	Syntagme verbal
SP	Syntagme prépositionnel

CHAPITRE 1

INTRODUCTION

1. PROBLÉMATIQUE

Lorsqu'une personne parle, son discours est segmenté en différentes unités que l'on appelle traditionnellement *constituants prosodiques* ou *domaines prosodiques*. Ces unités sont hiérarchisées et participent à l'organisation d'une structure phonologique abstraite appelée *structure prosodique*. Dans le signal de parole (représentation de surface de la langue), les frontières de constituants prosodiques sont marquées par différents indices phonétiques concrets tels que la réalisation de pauses silencieuses, de contours mélodiques spécifiques, l'allongement de certaines syllabes ou le renforcement de certains corrélats articulatoires. Les constituants prosodiques jouent un rôle important dans la production, la perception et le traitement de la parole. Du point de vue de la production, ces unités prosodiques sont des domaines au sein desquels différents phénomènes segmentaux ou prosodiques vont avoir lieu tels que des phénomènes de sandhi, c'est-à-dire l'altération phonologique des initiales ou finales de mots, selon s'ils sont morphosyntaxiquement autonomes ou en contact avec d'autres mots. Par exemple, le syntagme phonologique (*Phonological Phrase*, PP) a souvent été considéré comme le domaine de la réalisation obligatoire de la liaison en français ou du redoublement phono-syntaxique en italien (Selkirk, 1978; Nespor & Vogel, 1986). Dans les deux exemples suivants, respectivement empruntés au français et à l'italien, les phénomènes de sandhi ont lieu dans la première séquence mais pas dans la deuxième en raison du découpage en syntagmes phonologiques qui est différent dans les deux séquences.

1a. Cette famille a [trois beaux enfants]**PP**.

→ Liaison : apparition de la consonne de liaison /z/

1b. Jean a [des livres]**PP** [asssez nouveaux]**PP**.

→ Pas de liaison : pas d'apparition de la consonne /z/

2a. [Perché Carlo]**PP** non é venuto ? « Pourquoi Carlo ne vient pas ? »

→ Doublage syntaxique : allongement de la consonne /k:/

2b. [Che c'é un perché] **PP** [Carlo lo sa.]**PP** « Carlo sait qu'il y a une raison »

→ Pas de doublage syntaxique : la consonne /k/ n'est pas allongée

Du point de vue de la perception et du traitement de la parole, il a été montré que les frontières prosodiques participent à la compréhension des énoncés puisqu'elles concourent d'une part au découpage des énoncés en mots (voir par exemple pour le français : Christophe, Peperkamp, Pallier, Block & Melher, 2004 ; Spinelli, Welby & Schaegis, 2007 ; Spinelli, Nicolas, Meunier & Welby, 2010) et d'autre part à la construction de syntagmes syntaxiques au delà de ces mots (voir par exemple pour le français : Millote, René, Wales & Christophe, 2008 ; Millotte, Wales & Christophe, 2007).

Prenons l'exemple ci-dessous:

3a. A woman without her man is nothing.

[A woman without her man] [is nothing]

« Une femme sans son homme n'est rien. »

3b. A women : without her man is nothing.

[A woman] [without her] [man is nothing]

« Une femme : sans elle, l'homme n'est rien. »

Cette phrase peut être prononcée de deux manières différentes en fonction du découpage en constituants prosodiques qui est opéré par le locuteur. Ce découpage est visible à l'écrit grâce à la ponctuation qui est à la communication écrite, une partie infime de ce que la prosodie est à la communication orale. En fonction du découpage prosodique adopté, l'énoncé n'a pas le même sens. Je laisse ici au lecteur le soin de choisir l'interprétation qui lui convient le mieux.

Prenons un autre exemple. Nous avons demandé à une locutrice de prononcer la séquence sonore suivante : /ʒɑ̃lev/. Portons maintenant attention aux figures 1a et 1b dans lesquelles nous pouvons observer l'onde sonore, le spectrogramme et la fréquence fondamentale associée à deux versions de cette séquence prononcées par le même locuteur avec un débit de parole normal. Dans la figure de droite, nous pouvons remarquer la présence d'une frontière prosodique après la voyelle /ɑ̃/ de « Jean », repérable grâce à différents indices acoustiques telles que la présence d'une pause silencieuse, la réalisation d'un contour mélodique spécifique ou l'allongement de la durée de la syllabe /ʒɑ̃/. Notons que le marquage des frontières prosodiques est

donc multiparamétrique puisqu'il est signalé par différents indices. Au contraire, dans la deuxième séquence nous n'observons pas de frontière prosodique après la voyelle /ǎ/ de « J'enlève ». Dans cet exemple, les frontières prosodiques nous renseignent sur le découpage en mots : « Jean lève » dans la séquence 1a issue de la phrase « Jean lève son verre » et « J'enlève » dans la séquence 1b issue de la phrase « J'enlève son verre ».

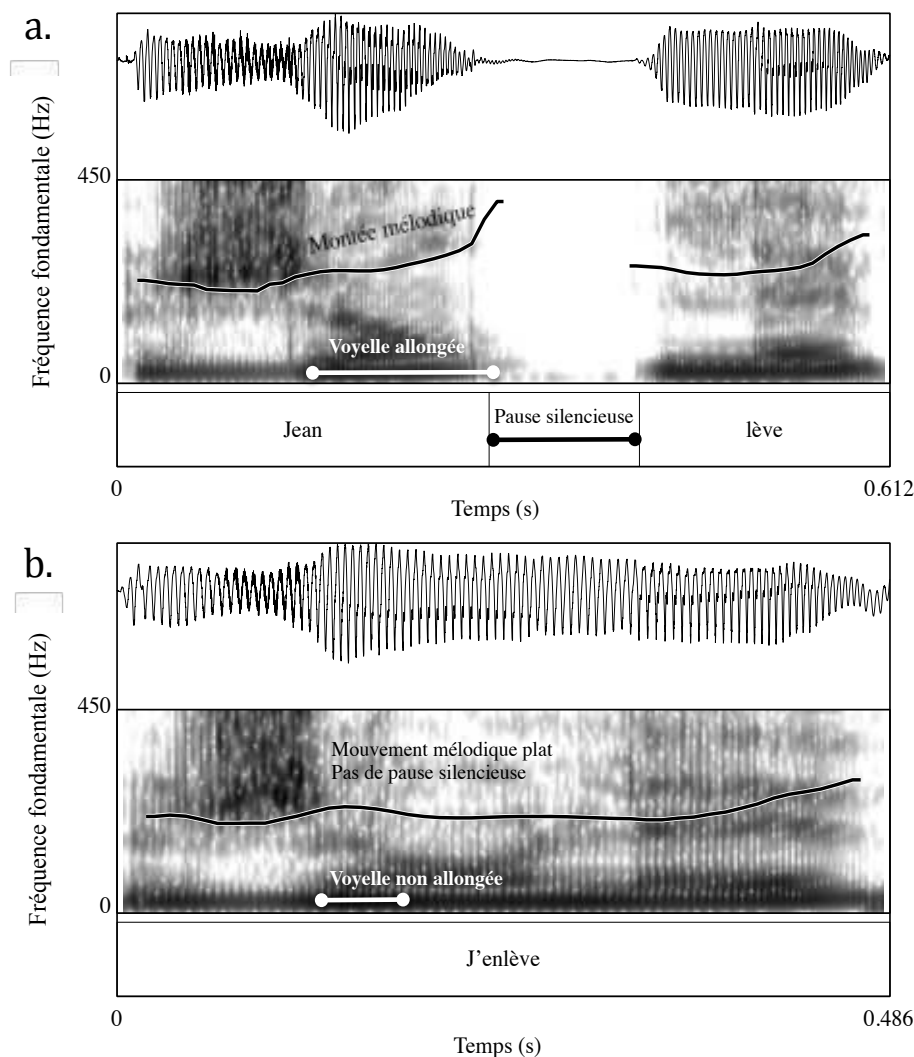


Figure 1. Illustration de la séquence sonore /zǎlev/ à vitesse d'élocution normale. Séquence lue deux fois par la même locutrice à vitesse d'élocution normale. La séquence a est extraite de la phrase « Jean lève son verre » et la séquence b est extraite de la séquence « J'enlève son verre ».

Nous avons demandé à notre locutrice de reproduire les deux séquences présentées dans la 1 avec un débit de parole plus rapide. Les séquences obtenues sont représentées dans la figure 2.

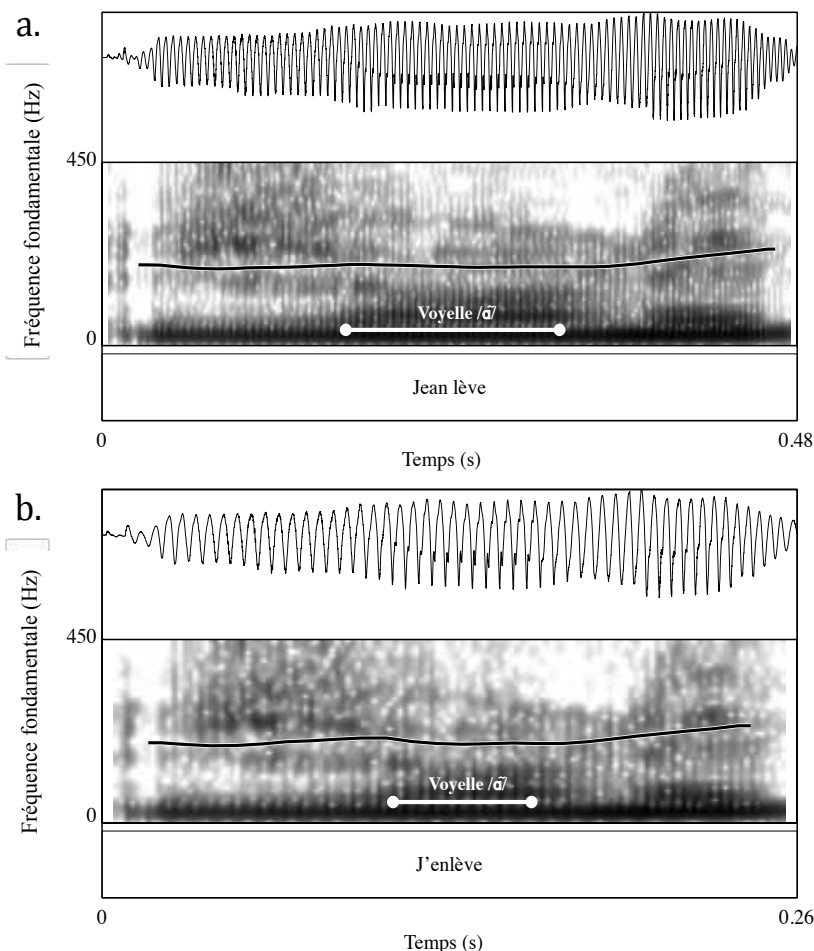


Figure 2. Illustration de la séquence sonore /ʒɑ̃lev/ à vitesse d'élocution rapide. Onde sonore, spectrogramme et courbe de fréquence fondamentale associés à la séquence sonore /ʒɑ̃lev/ extraite des phrases « Jean lève son verre » (a) « J'enlève son verre » (b) lues par la même locutrice à vitesse d'élocution rapide.

À vitesse rapide, il est beaucoup plus difficile de différencier les deux séquences. Du point de vue de la structure tonale, c'est-à-dire au regard des mouvements mélodiques se manifestant dans la parole par des variations de la fréquence fondamentale, les deux séquences sont identiques. Cependant, nous pouvons remarquer que la durée de la voyelle /ɑ̃/ est beaucoup plus importante dans la séquence 2a que dans la séquence 2b. Cet exemple illustre parfaitement une situation fréquemment rencontrée dans la parole : lorsque les différents indices marquant les frontières de constituants prosodiques ne s'accordent pas dans le marquage des frontières. Selon le type d'indice (indice mélodique ou indice de durée) qui est privilégié dans telle ou telle théorie, on va considérer qu'une frontière prosodique permettant de faire la différence entre les deux séquences est présente ou non. De plus, aucune étude menée jusqu'à ce jour en français ne nous permet de savoir si l'allongement de la voyelle sera suffisant pour que des adultes perçoivent une frontière, c'est-à-dire une rupture à l'écoute de cette séquence.

Depuis de nombreuses années les linguistes ont cherché à définir les constituants prosodiques ainsi qu'à prédire le lieu d'apparition de leurs frontières. La première idée qui a été proposée pour répondre à cette question est l'idée que les unités prosodiques correspondent aux unités syntaxiques (Cooper & Paccia-Cooper, 1980 ; Lehiste 1973). Il a été ensuite admis que les structures prosodiques et syntaxiques ne sont pas isomorphes. En effet, les travaux menés dans le cadre de la Phonologie Prosodique (Selkirk, 1978, 1984, 1986 ; Nespor & Vogel, 1986) ont permis de mettre en évidence l'existence d'une structuration phonologique de la prosodie relativement indépendante de la syntaxe bien que largement contrainte par celle-ci. Les constituants prosodiques sont des domaines au sein desquels se réalisent des phénomènes phonologiques particuliers tels que des phénomènes de Sandhi comme la liaison en français ou le redoublement phono-syntaxique en italien. Dans cette approche, malgré le fait qu'un niveau de structuration prosodique autonome soit reconnu, les constituants prosodiques sont majoritairement définis grâce à des algorithmes syntaxiques. Or nous savons aujourd'hui que des facteurs extra-syntaxiques tels que la longueur des constituants (Pascual, 1992 ; Monin & Grosjean, 1993 ; Jun, 1993, 2003 ; Watson & Gibson, 2004), le poids des constituants, c'est à dire un découpage en constituants qui ont à peu près la même taille (Selkirk, 2000, Ghini, 1993 ; Sandalo & Truckenbrodt, 2002) ou la vitesse de parole (Barthkova, 1991 ; Keller & Zeller 1995, Fougeron & Jun, 1998) influent sur le découpage prosodique. Des contraintes de nature sémantique (Breen, Gibson & Watson, 2010) ou relatives à la structure informationnelle (Féry, 2001 ; Selkirk, 2000 ; German & D'Imperio, 2010) contraignent également de façon importante le découpage en constituants prosodiques. De plus il a été montré que les prédictions de la phonologie prosodique ne sont pas toujours actualisées dans le signal de parole produit par le locuteur (Michelas & D'Imperio, 2010).

Parallèlement à la théorie de la Phonologie Prosodique, un deuxième courant de recherche s'est attaché à définir les unités prosodiques à partir de leurs manifestations physiques dans le signal de parole en faisant référence à l'organisation métrique, tonale ou temporelle des énoncés (pour l'anglais : Pierrehumbert, 1980 ; Beckman & Pierrehumbert, 1986 ; pour le français : Di Cristo, 1993 ; Jun & Fougeron, 1995, 2000, 2002). Cette approche a permis une caractérisation plus fine des indices acoustiques marquant les frontières de constituants prosodiques. En fonction des différentes approches, la priorité est accordée à un type d'indice particulier. Par exemple, en français, l'approche de Jun & Fougeron accorde la primauté aux indices de nature tonale (Jun & Fougeron, 2000), tandis que l'organisation métrique est privilégiée chez Delais-Roussarie (Delais-Roussarie, 1995). Cependant, comme nous l'avons précédemment montré, il arrive que les différents indices (en particulier les indices mélodiques et de durée) ne s'accordent pas dans le marquage des frontières. Ce problème nous amène tout naturellement à nous interroger sur la pertinence des différents indices phonétiques dans la perception des frontières prosodiques. La plupart des études menées jusqu'à ce jour sur la perception et l'utilisation des frontières prosodiques dans le

traitement de la parole en français se basent sur une définition syntaxique des constituants et ne permettent pas d'appréhender de façon satisfaisante la contribution de chaque type d'indice acoustique dans le traitement de la parole.

En prenant en compte des contraintes issues de différents niveaux de la grammaire et en admettant un marquage pluriel des frontières prosodiques, cette thèse a pour but d'aider à la compréhension des relations existant entre la structure prosodique abstraite et sa réalisation phonétique concrète dans la parole.

Nous commencerons par dresser une revue détaillée des problématiques qui se posent autour de la question de la structure prosodique. Nous définirons plus précisément ce qu'est la structure prosodique, développerons le cadre de travail que nous avons utilisé dans cette thèse (la théorie auto-segmentale métrique de l'intonation) et présenterons les trois principaux modèles proposés dans ce cadre pour rendre compte de la structure prosodique du français.

A l'issue de ce premier chapitre qui nous sert à introduire tous les concepts nécessaires à la compréhension de ce travail, notre thèse s'articulera en deux parties essentielles. Dans les chapitres 2 et 3 nous développerons plusieurs expérimentations menées en production de la parole qui nous permettront de caractériser les propriétés phonétiques et phonologiques des constituants prosodiques situés entre le mot et le syntagme intonatif. Les études admettent traditionnellement deux niveaux de structure prosodique pour le français : le syntagme accentuel et le syntagme intonatif. Au cours de ces trois chapitres, nous apporterons des preuves de l'existence d'un troisième niveau de structuration que l'on appelle syntagme intermédiaire.

Dans une deuxième partie de la thèse constituée des chapitres 4 et 5, nous nous attacherons à tester l'influence du syntagme accentuel et du syntagme intermédiaire que nous définirons à partir de leurs propriétés phonétiques et phonologiques. Ces deux constituants seront-ils validés d'un point de vue perceptif et cognitif? Une réponse positive donnerait évidemment plus de crédit à notre hypothèse de départ qui est que trois constituants prosodiques (syntagme accentuel, syntagme intermédiaire et syntagme intonatif) sont nécessaires pour rendre compte de la structure prosodique du français.

2. LA STRUCTURE PROSODIQUE

2.1. La hiérarchie des domaines prosodiques

Il est généralement admis que la structure prosodique est hiérarchiquement organisée (Liberman, 1975 ; Selkirk, 1978 ; Pierrehumbert, 1980 ; Nepor & Vogel, 1986). Les énoncés sont découpés en constituants qui sont hiérarchiquement dominés par des unités prosodiques de niveau

supérieur dans la hiérarchie prosodique. Cette hiérarchie de domaines prosodiques se base sur un étagement strict et non récursif des constituants suivant une règle appelée l'Hypothèse de l'Etagement Strict, *Strict Layer Hypothesis* (Selkirk, 1986 ; Nespor & Vogel, 1984). Selon cette règle, une unité non-terminale de la structure hiérarchique doit, d'une part, être formée d'une ou de plusieurs unités de la catégorie immédiatement inférieure et d'autre part, être immédiatement contenue dans l'unité qui la domine directement. Cette règle implique donc que les constituants d'un même niveau prosodique ne peuvent pas être emboîtés. Les systèmes d'annotations de l'intonation des langues se basent majoritairement sur ce principe de hiérarchie stricte et non récursive. Par exemple, le système anglais *Tones and Break Indices* ; Beckman & Ayers Elam, 1997) admet trois niveaux de structure prosodique (le mot, le syntagme intermédiaire et le syntagme intonatif) bien que seuls deux d'entre eux soient tonalement marqués. Dans ce système, un énoncé est composé d'au moins un syntagme intonatif qui est à son tour composé au moins d'un syntagme intermédiaire. De ce fait, un constituant d'un niveau donné ne peut être précédé que d'un constituant de même niveau prosodique.

Même si la plupart des approches s'accordent sur une vision hiérarchique de la structure prosodique, cette conception est sujette à controverse. Certaines études ont en effet suggéré que le degré d'allongement pré-frontière pourrait être proportionnel à son niveau de structure dans la hiérarchie grammaticale. Des preuves pour au moins quatre niveaux de structure ont été démontrés pour l'anglais (Wightman, Shattuck-Hufnagel, Ostendorf & Price, 1992 ; Ladd & Campbell, 1991). En effet, la question de la récursivité des domaines prosodiques, c'est-à-dire le fait qu'un constituant prosodique d'un certain niveau dans la hiérarchie prosodique pourrait être englobé dans un autre constituant de même niveau dans cette hiérarchie (Ladd, 1996), reste en suspens. L'approche proposée par Ladd (1996) conteste l'Hypothèse de l'Etagement Strict (*Strict Layer Hypothesis*) proposée par Selkirk (1986) en faveur d'une structure où la récursivité des niveaux prosodiques est possible. Ainsi par exemple un syntagme intonatif pourrait contenir d'autres syntagmes intonatifs (voir pour l'anglais : Ladd, 1986, 1988, 1996 ; Ladd & Campbell 1991 ; pour le japonais : Ito & Mester, 1992 ; pour le français : Di Cristo & Hirst, 1996). Dans cette thèse, les corpus utilisés comportent des phrases qui ne soulèvent pas le problème de la récursivité. Nous considérerons donc que les constituants prosodiques examinés suivent la *Strict Layer Hypothesis* sans pour autant ignorer les problèmes que pose l'adoption de cette règle.

Si la plupart des études menées sur le sujet s'accordent sur une vision hiérarchique des niveaux, les différents modèles qui envisagent la structure prosodique s'accordent également sur le fait que la structuration prosodique de la parole se fait selon une hiérarchie de constituants mais aussi selon une hiérarchie de proéminences. Les travaux issus de la Théorie Métrique (Liberman & Prince, 1977 ; Hayes, 1989, *inter alia*) proposent un nombre fini de catégories de proéminences qui s'organisent en quatre niveaux : la syllabe inaccentuée (*unstressed syllable*), la syllabe à accent lexical (*stressed syllable*),

la syllabe à accent mélodique (*pitched accented syllable*) et la syllabe à accent nucléaire (*nuclear pitched accented syllable*). Nous renvoyons le lecteur à Beckman & Edwards (1990) pour plus de détails sur le sujet. La hiérarchie des proéminences est reliée à la hiérarchie des constituants. La distinction entre proéminence et frontière n'est pas toujours aisée pour une langue comme le français. En effet, contrairement à l'anglais, l'accent est implémenté au niveau du syntagme et non pas au niveau du mot. Un accent final (également dit *primaire*) est donc toujours associé à une frontière de constituant prosodique. On parle du *caractère démarcatif* de l'accent final (ou primaire) qui sert à marquer une frontière de constituant. La syllabe porteuse de l'accent final en français peut être comparée à l'accent mélodique de l'anglais (*pitch accent*) malgré le fait qu'elle conserve un statut différent dans le sens où elle n'est pas uniquement un indice de proéminence mais également un indice de frontière.

2.2. Les relations entre structure prosodique et structure syntaxique

Nous savons aujourd'hui que les structures syntaxiques et prosodiques ne sont pas isomorphes (voir Shattuck-Hufnagel & Turk, 1996, pour une revue de la littérature sur ce point). Cependant malgré cette non-isomorphie, ces deux structures ne sont pas non plus complètement indépendantes l'une de l'autre. Parmi les éléments prouvant l'influence de la syntaxe sur le découpage prosodique, il apparaît que les frontières entre des constituants syntaxiques situés relativement haut dans la hiérarchie syntaxique, coïncident avec des frontières prosodiques majeures (Gee & Grosjean, 1993 ; Monin & Grosjean 1993 ; Keller, Zellner, Werner, & Blanchoud, 1993). Il a également été montré que des structures syntaxiques spécifiques telles que des parenthétiques, des clauses non restrictives, des vocatives ou des disloquées montrent une très forte corrélation avec une frontière prosodique majeure (pour l'anglais : Selkirk, 1995 ; pour le français : Delais-Roussarie, 2005 ; Delais-Roussarie & Post, 2008). Enfin des études menées sur différentes langues ont montré que certains découpages prosodiques sont impossibles en raison de fortes contraintes syntaxiques (D'Imperio, Elorideta, Frota, Prieto & Vigário, 2005 ; Frota, D'Imperio, Elordieta, Prieto & Vigário, 2007).

De plus, il a été montré que l'interface syntaxe/prosodie est influencée par des contraintes de nature prosodique telle que la taille des constituants prosodiques (pour le français : Padeloup 1992, Monin & Grosjean 1993, Delais-Roussarie, 1995 ; pour le coréen : Jun, 1993), favorisant une distribution où les unités prosodiques ont à peu près la même taille (pour l'anglais : Selkirk, 2000 ; pour l'italien : Ghini, 1993 ; pour le portugais de Brésil : Sandalo & Truckenbrodt, 2002). La vitesse de discours semble également peser sur l'interface syntaxe/prosodie dans le découpage en unités prosodiques (pour le français : Barthkova, 1991 ; Keller & Zellner, 1995 ; Fougeron & Jun, 1998). L'influence de contraintes de nature autre que syntaxique ou prosodique telles que des contraintes sémantiques ou relatives à l'organisation de la structure informationnelle ou du discours, est également

de plus en plus pris en compte (Féry, 2000 ; Selkirk, 2000 ; German & D'Imperio, 2010 ; Breen et al., 2010).

Si le fait qu'il n'existe pas d'isomorphie entre structure prosodique et structure syntaxique est aujourd'hui largement admis, la modalité de l'interface entre les deux composantes de la grammaire est encore sujette à controverse. Dans la théorie de la phonologie prosodique proposée par Selkirk (1986) et Nespor & Vogel (1984), les liens entre la structure syntaxique d'une part et la réalisation phonétique concrète de la structure prosodique dans la parole d'autre part, ne sont pas directs. Ces relations sont en effet modulées par une structure phonologique abstraite et autonome. Cette conception indirecte de l'interface entre réalisation phonétique concrète de la structure prosodique d'une part et structure syntaxique d'autre part est largement répandue dans la littérature mais a cependant été remise en question par plusieurs travaux.

Certaines études conçoivent les liens entre structure syntaxique et découpage prosodique sans postuler l'existence d'une structure prosodique hiérarchisée aussi complexe. Des précédents travaux ont en effet montré qu'un haut niveau d'attachement syntaxique ainsi que des frontières de constituants syntaxiques majeures coïncident avec des indices de rupture prosodique majeures (Cooper & Paccia-Cooper, 1980 ; Gee & Grosjean, 1983 ; Monnin & Grosjean, 1993 ; Keller, Zellner, Werner & Blanchoud, 1993). Les résultats de ces études admettent une interface directe entre la structure syntaxique d'une part et la réalisation d'indices prosodiques dans la réalisation de surface d'autre part sans médiation par une structure prosodique autonome. Une question importante reste donc en suspens : les indices observés à la frontière de constituants prosodiques sont-ils reliés à une hiérarchie des domaines prosodiques dépendante de la langue étudiée ou est-ce qu'il sont associés à une hiérarchie de constituants morphosyntaxiques (telles que la phrase, le syntagme et le mot) et à leur niveau d'emboîtement à l'intérieur de cette hiérarchie morphosyntaxique. Il est en effet légitime de se demander si les constituants de la hiérarchie prosodique postulés dans la théorie de la phonologie prosodique constituent de réelles primitives phonologiques ou s'ils n'existent que dans la réalisation de surface de la parole.

2.3. La définition et le nombre de constituants prosodiques

Une autre question importante et non résolue dans la littérature concerne le nombre et la définition des unités prosodiques. Selon les différentes approches proposées, la définition des unités prosodiques peut être soit basée sur des contraintes syntaxiques soit sur des propriétés prosodiques (voir Jun 1993, 1998, pour une comparaison de ces deux approches). L'approche dite « syntaxique » propose de définir les constituants prosodiques en se référant à des algorithmes syntaxiques. Les informations syntaxiques utilisées sont de plusieurs ordres en fonction des théories. Selkirk (1986) utilise par exemple la notion de « projection syntaxique maximale » de

la syntaxe X-barre. Dans ce cas, les frontières de constituants prosodiques coïncident avec des frontières de constituants syntaxiques. Nespor & Vogel (1986) utilisent quant à elles les relations syntaxiques « têtes compléments ». Elles proposent l'idée que la formation des groupes phonologiques s'opère en regroupant une tête lexicale (nom, verbe, adjectif) avec les mots qui lui sont rattachés sur sa gauche (pour les langues récursives droite) jusqu'à rencontrer une frontière de constituant syntaxique, donc jusqu'à rencontrer un autre tête lexicale. Une approche de même type est adoptée par Hayes (1989). L'approche dite « prosodique » s'efforce quand à elle de caractériser les constituants prosodiques en référence aux structures tonale, métrique ou temporelle des énoncés (voir par exemple Beckman & Pierrehumbert, 1986 ; Pierrehumbert, 1980 ; Di Cristo, 1993 ; Jun & Fougeron, 1995, 2000, 2002). En fonction des différentes approches, la primauté est accordée à l'une de ces structures dans la définition des unités prosodiques. Par exemple, dans le modèle proposé par Jun & Fougeron, la définition des constituants prosodiques est basée sur des critères mélodiques même si l'organisation temporelle participe à l'élaboration de la structure. Dans cette thèse nous adoptons une approche dite « prosodique » en ce sens que nous cherchons à caractériser les constituants prosodiques en fonction des indices phonétiques et phonologiques (d'ordre tonal et métrique) observés dans la réalisation de surface de la langue.

Nous savons aujourd'hui que les prédictions de la Phonologie prosodique ne sont pas toujours réalisées dans le signal de parole (Michelas & D'Imperio, 2010). Cette constatation nous a conduit à adopter une définition des constituants prosodiques basée sur les propriétés phonétiques et phonologiques des constituants prosodiques et non sur des algorithmes morpho-syntaxiques. Malgré le fait que nous défendons ce type d'approche, nous ne sommes pas sans savoir que cette définition est parfois problématique dans la mesure où les trois paramètres prosodiques (que sont la fréquence fondamentale, la durée et l'intensité) varient beaucoup d'un locuteur à l'autre ou d'un énoncé à l'autre et ce dans un même contexte prosodique. Cette grande variabilité rend parfois compliqué la caractérisation systématique des indices prosodiques. Afin de remédier en partie à cette difficulté, nous utiliserons dans nos expérimentations une parole très contrôlée nous permettant de caractériser à partir d'un grand nombre de données les propriétés phonétiques des constituants prosodiques. Nous serons également très vigilante et prendrons en compte un maximum des contraintes de différentes natures (telle que la structure syntaxique, le débit de parole, la longueur et le poids de constituants, etc.) dans nos protocoles expérimentaux afin de parvenir à une meilleure caractérisation des constituants prosodiques du français. Nous chercherons également à déterminer dans quelles mesures certaines de ces contraintes influent sur le découpage prosodique ce qui nous permettra une meilleure appréhension des relations entre structure prosodique abstraite et réalisation concrète de cette structure dans la parole.

S'il n'existe pas de consensus sur la définition des constituants prosodiques, les différentes théories sur le sujet ne s'accordent pas non plus

sur le nombre de constituants prosodiques existant. Malgré un désaccord notable sur ce point, deux niveaux de structure au dessus du mot sont généralement admis pour le français : un constituant prosodique de haut niveau dans la hiérarchie prosodique (que nous appellerons syntagme intonatif) et un constituant de niveau inférieur correspondant au domaine de réalisation de l'accent primaire (que nous appellerons syntagme accentuel). Le syntagme intonatif et le syntagme accentuel correspondent à la continuation mineure et à la continuation majeure décrites par Delattre (1966) ou au groupe accentuel et groupe intonatif de Mertens (1993). De nombreux autres termes utilisés pour désigner ces deux constituants ont également été proposés dans la littérature. Nous renvoyons le lecteur à Di Cristo (2010) pour une revue sur le sujet.

3. PROPRIÉTÉS PHONÉTIQUES ET PHONOLOGIQUES DES FRONTIÈRES PROSODIQUES

Comme nous l'avons évoqué précédemment, le marquage des frontières prosodiques est multiparamétrique, c'est à dire qu'en général, le marquage d'une catégorie de frontière ne correspond pas à un seul indice acoustique mais à un groupe d'indices principaux. L'importance et la combinaison de ces indices varie en fonction de la langue.

Plusieurs types d'indices marquant les frontières de constituants prosodiques ont été mis en évidence dans la littérature. Parmi ces indices nous retrouvons : des indices relatifs à la durée des segments (indices temporels), des indices relatifs à la structuration mélodique ou la réalisation de pauses silencieuses ainsi que des modifications affectant la manière dont les segments sont prononcés (variations spectrales et non phonémiques). Nous allons maintenant aborder plus en détail chacun de ces types d'indices.

3.1. Propriétés acoustiques

3.1.1. Propriétés temporelles

Un nombre important d'études a investi les propriétés temporelles des segments situés autour des frontières prosodiques dans les différentes langues du monde. Des études acoustiques ont montré que les syllabes situées à la fin des constituants prosodiques deviennent plus longues et de nombreux auteurs font référence à cet effet en tant que phénomène *d'allongement final* ou *d'allongement pré-frontière* (pour l'anglais : Gaitenby, 1965 ; Oller, 1973 ; Klatt, 1965 ; Cooper & Paccia-Cooper, 1980 ; Beckman & Edwards, 1990 ; Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992 ; Ferreira, 1993 ; Turk & White, 1999 ; Turk & Shattuck-Hufnagel, 2007 ; *inter alia*). Ce phénomène d'allongement pré-frontière a également été mis en évidence en français à de nombreuses reprises (Delattre, 1966 ; Padeloup, 1990 ; Fletcher, 1991 ; Jun &

Fougeron, 2000 ; Di Cristo, 2000 ; Post, 2000 ; Mertens, 2001, Astésano, 2001 ; *inter alia*). Bien que certaines études menées sur l'anglais ont mis en évidence la présence d'un allongement syllabique qui intervient non pas sur la syllabe pré-frontière mais un peu plus tôt dans le mot situé avant une frontière prosodique (Price, Ostendorf, Shattuck-Hufnagel & Fong, 1991 ; Wightman et al., 1992 ; Turk & Shattuck-Hufnagel, 1997), cet effet n'a, à n'autre connaissance, jamais été observé en français.

Un effet d'allongement de la syllabe a également été observé à l'initial de constituant prosodique chez certains auteurs (pour l'anglais : Pierrehumbert & Talkin, 1991 ; Shattuck Hufnagel & Turk, 1998 ; pour le français : Tabain, 2003) alors que d'autres auteurs n'observent aucun effet post-frontière sur la durée de la syllabe (pour l'anglais : Wightman et al., 1992).

L'existence d'une forte corrélation entre le degré d'allongement pré-frontière et la perception des ruptures prosodiques a été démontrée dans la littérature (Wightman et al., 1992 ; Breen, Gibson & Watson, 2010). Il apparaît également que le degré d'allongement pré-frontière permet de différencier plusieurs niveaux de frontières, avec des frontières plus fortes associées à un degrés d'allongement plus important (Wightman et al., 1992). L'allongement pré-final associé à la durée des pauses semblerait donc être employé pour marquer la structuration hiérarchique dans la parole puisque le degré d'allongement apparaît être proportionnel au niveau du constituant à l'intérieur de la hiérarchie des domaines prosodiques.

3.1.2. Propriétés tonales

Parmi les indices marquant les frontières de constituant prosodique, nous pouvons également compter la réalisation d'un contour mélodique spécifique ou des phénomènes de réinitialisation du registre avant et après une frontière prosodique.

Dans la théorie autosegmentale métrique de l'intonation (Goldsmith, 1976 ; Bruce, 1977 ; Pierrehumbert 1980, Ladd, 1996, *inter alia*) qui considèrent que le contour mélodique peut être décomposé en une série d'évènements intonatifs locaux, les tons périphériques (*edge tones*) sont associés aux constituants prosodiques. Ainsi l'accent de syntagme (*phrase accent*) est aligné avec la fin du syntagme intermédiaire tandis que les tons de frontière (*boundary tones*) sont alignés avec le début et/ou la fin des syntagmes intonatifs (Pierrehumbert, 1980 ; voir section 4.1 qui suit). Dans cette théorie, proposée à l'origine pour l'anglais américain, l'accent mélodique, qui a pour fonction de rendre proéminente la syllabe qui lui est associée n'est à priori rattaché à aucune frontière de constituant¹. En français, la syllabe

¹ Nous ne parlons pas ici de l'association secondaire des tons de frontières. Nous renvoyons le lecteur à Beckman & Pierrehumbert (1988) et Grice, Ladd & Arvaniti, (2000) pour plus de détails sur le sujet.

proéminente est toujours la dernière syllabe du groupe accentuel (Dell, 1984). Le français est en effet une langue à accent fixe contrairement à l'anglais où la syllabe accentuée n'est pas toujours la même (langue à accent libre). La syllabe porteuse de l'accent final en français est souvent comparée à l'accent mélodique (*pitch accent*) de l'anglais. Cependant elle conserve un statut différent dans le sens où elle n'est pas uniquement un indice de proéminence mais également un indice de frontière. En raison du caractère démarcatif de l'accent final, les indices marquant l'accentuation en français deviennent de fait des indices marqueur de la frontière de syntagme accentuel. Parmi les marqueurs de l'accentuation, plusieurs études menées en français ont mis en évidence le fait que la syllabe porteuse de l'accent finale est plus longue, associée à un mouvement montant de *f₀* et à une augmentation de l'intensité lorsqu'elle est située en position non finale dans l'énoncé (Pasdeloup, 1990; Fletcher, 1991 ; Jun & Fougeron, 2002). Si tous les auteurs qui se sont attachés à décrire le phénomène d'accentuation en français s'accordent sur les deux premiers marqueurs, tous n'observent pas une augmentation de l'intensité associée à la syllabe accentuée (Wenk & Wioland, 1991).

Des phénomènes de réinitialisation du registre observés avant et après une frontière prosodique ont également été mis en évidence en tant que marqueur des frontières prosodiques. Ladd (1996) propose de modéliser les variations du registre tonal en distinguant deux dimensions de variation indépendantes : (i) le niveau du registre correspondant à la hauteur globale à laquelle une séquence est énoncée (*level*) et (ii) l'étendue du registre correspondant à la plage fréquentielle utilisée dans une séquence (*span*).

Des phénomènes de remise à niveau du registre observés avant et après une frontière prosodique ont également été mis en évidence en tant que marqueur des frontières prosodiques. Nous savons que les accents mélodiques sont abaissés les uns par rapport aux autres ce qui est souvent désigné en tant que phénomène d'abaissement ou de *downstep*. En 1986, Pierrehumbert & Beckman ont mis en évidence le fait que le registre était réinitialisé à la frontière du syntagme intermédiaire en anglais. En 1988, Ladd a proposé que ce phénomène d'abaissement n'affecte pas uniquement les accents mélodiques mais soit englobé dans un phénomène plus large qui affecte les constituants prosodiques. Selon Ladd (1988), les constituants prosodiques sont abaissés les uns par rapport aux autres ce qui donne lieu à des phénomènes de réinitialisation partielle du registre après une frontière prosodique. En 2002, Truckenbrodt a utilisé le modèle initialement proposé par Ladd (1988) puis développé par van der Berg (1992) pour décrire le système intonatif de l'allemand. Les accents mélodiques sont abaissés au sein du syntagme intonatif en allemand et un phénomène de réinitialisation du registre qu'il appelle *upstep* a lieu sur le pic nucléaire et/ou sur le ton de frontière H% du syntagme intonatif. Les syntagmes intonatifs qui ne sont pas en position initiale sont abaissés les uns par rapport aux autres. Cet abaissement plus global donne lieu à une réinitialisation partielle du registre après la frontière de syntagme intonatif en allemand. Le modèle proposé par

Truckenbrot (2002) pour rendre compte du système intonatif de l'allemand est schématisé dans la figure 3.

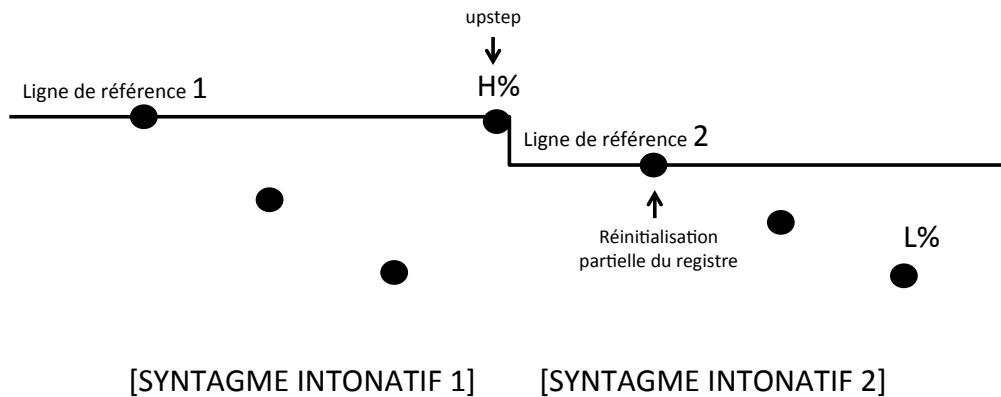


Figure 3. Schématisation du modèle proposé par Truckenbrot pour l'allemand.
Tiré de Truckenbrot (2002), adapté de Ladd (1988) et de van der Bergh, Gussenhoven & Rietveld (1992).

Nous aborderons plus en détails ces effets d'abaissement et de réinitialisation du registre au cours du chapitre 3 de cette thèse.

3.2. Propriétés articulatoires

Certains auteurs ont étudié les corrélats articulatoires du découpage prosodique. Ces études menées sur différentes langues ont montré que certains corrélats articulatoires tels que le contact entre la langue et le palais (pour le français : Fougeron & Keating, 1997 ; Fougeron, 2001 ; pour le coréen : Cho & Keating, 2001 ; Jun, 1993 ; pour le taïwanais : Hsu & Jun, 1997) ou l'articulation glottale (pour l'anglais : Pierrehumbert & Talkin, 1992 ; pour le coréen : Jun, 1993 ; Cho & Keating, 2001) sont renforcés à la frontière de constituants prosodique plutôt qu'à l'intérieur de ceux-ci. Ce renforcement des gestes articulatoires a été mis en évidence aussi bien pour la frontière droite, c'est-à-dire en final de constituant (voir par exemple Edwards & Beckman, 1998 ; Edwards, 1991 ; Tabain, 2003) que pour la frontière gauche c'est-à-dire à l'initial du constituant (voir par exemple Byrd & Saltzman, 1998 ; Byrd, 2000 ; Tabain, 2003).

Il a également été montré que les caractéristiques articulatoires observées sont à mettre en relation avec le niveau de structuration dans la hiérarchie prosodique. En d'autres termes, ces modifications articulatoire s'amplifient de façon graduelle avec le niveau du constituant dans la hiérarchie (Byrd & Saltzman, 1998 ; Byrd, 2000 ; Cho, 2005 ; Tabain, 2003 ; Tabain & Perrier, 2005, Cho & Keating 2001 ; Fougeron, 2001 ; Cho, 2005 ; Keating et al., 2004).

4. LA STRUCTURE PROSODIQUE DANS LE MODÈLE AUTO-SEGMENTAL MÉTRIQUE DE L'INTONATION

4.1. Le modèle auto-segmental métrique de l'intonation

Dans cette thèse nous adoptons le modèle auto-segmental métrique de l'intonation (Pierrehumbert, 1980 ; Beckman & Pierrehumbert 1986 ; Ladd 1996, *inter alia*). Un des présupposés de cette théorie est que le contour intonatif peut être décomposé en une série d'évènements intonatifs locaux, les tons statiques bas (L) et hauts (H) qui sont reliés par interpolation linéaire dans le processus d'implémentation phonétique. Cette conception s'oppose à l'approche holistique (Delattre, 1966 ; Rossi, 1999 ; Liberman, 1975, *inter alia*) selon laquelle le contour intonatif ne peut pas être décomposé en éléments plus petits. Selon cette théorie, le contour intonatif a donc dans son ensemble, une fonction morphémique puisque chaque contour intonatif est associé à une fonction pragmatique. Il peut par exemple être associé à la modalité interrogative (Delattre, 1966) en français ou au contour « surprise/redondance » (Liberman & Sag, 1974) en anglais. Notons que l'idée de la compositionnalité du contour, c'est-à-dire le fait de le représenter comme une séquence d'unités plus petites, n'est pas propre à la théorie auto-segmentale métrique mais se retrouve dans d'autres approches également dites « compositionnelles ». Parmi ces approches, certaines proposent d'analyser les unités qui composent le contour, non pas en terme de tons statiques (comme le fait la théorie auto-segmentale métrique), mais en terme de mouvements de *f₀*. Ce type de description est typique des approches compositionnelles dites « configurationnelles » telles que celle l'école britannique (Halliday, 1967; Crystal, 1969; Cruttenden, 1986) ou de la théorie IPO (t'hart & Cohen, 1990). Contrairement à ces dernières approches, la théorie auto-segmentale métrique néglige les trajectoires et considère uniquement les points initiaux et finaux du contour comme pertinents. Les trajectoires allant d'un ton à un autre sont considérées comme de simples transitions sans aucune importance linguistique.

L'analyse en terme de points du contour pertinents pour l'analyse n'est pas une nouveauté introduite par la théorie auto-segmentale métrique. Cette démarche avait en effet déjà été adoptée par la tradition structuraliste américaine (voir par exemple Pike, 1945) ou dans le cadre de la phonologie autosegmentale (Liberman, 1975 ; Goldsmith, 1976 ; Bruce 1977). Pike (1945) distingue quatre tons (extra-haut, haut, haut moyen et bas) dans sa description de l'anglais américain. Dans cette approche, les tons n'ont pas une fonction linguistique directe mais c'est leur combinaison qui engendre la création de morphèmes intonatifs. Dans l'approche de Bruce (1977) et de Pierrehumbert (1980) le nombre de tons jugés comme pertinent est réduit. Cette question du nombre de tons utilisés dans l'analyse est une question importante. En effet, la partition du contour en un nombre élevé de niveaux est délicate, car la valeur mélodique de chaque niveau dépend de l'étendue

tonale du locuteur. Nous renvoyons le lecteur à Bolinger (1951) pour une discussion sur ce point. Dans la théorie auto-segmentale métrique, qui admet deux niveaux de partition du contour, les différences de hauteur (L et H) sont considérées en terme relatif et les variations de leur valeur mélodique dans l'étendue tonale sont le résultat de variations paralinguistiques.

Comme nous l'avons évoqué précédemment, dans cette approche, les tons qui composent le contour mélodique se combinent en accent mélodiques (*pitch accents*) et tons périphériques (*edge tones*). Les accents mélodiques sont associés aux syllabes tandis que les tons périphériques sont associés au début ou à la fin des constituants prosodiques. Ainsi l'accent de syntagme (*phrase accent*) est associé à la fin du syntagme intermédiaire tandis que les tons de frontières (*boundary tones*) sont associés avec le début ou la fin des syntagmes intonatifs (Pierrehumbert, 1980). L'accent mélodique a pour fonction de rendre proéminente la syllabe qui lui est associée et peut être composé par un seul ton (accent monotonal L* ou H*) ou par deux tons (accent bitonal ; par exemple L*+H pour les accent montants et H*+L pour les accents descendants en anglais). Le diacritique * est utilisé pour indiquer l'association avec une syllabe métriquement forte. En cas d'accent bitonal, la liaison entre les deux tons est représentée par le signe +. Les tons périphériques sont au contraire obligatoirement monotonaux et leur statut est signalé par le diacritique - pour les accents de syntagmes (L- ou H-) ou le diacritique % pour les tons de frontière (L% ou H%). Les tons de frontière marquent le début et la fin des syntagmes intonatifs (*Intonational Phrase*, IP). Dans Pierrehumbert (1980), l'accent de syntagme est défini comme l'accent qui suit immédiatement l'accent nucléaire et qui contrôle le registre jusqu'à la fin de ce constituant. Cette définition a été corrigée par Beckman & Pierrehumbert (1986) qui admettent que l'accent de syntagme marque la frontière d'un constituant hiérarchiquement inférieur à l'IP : le syntagme intermédiaire (*Intermediate Phrase*, ip) qui est le domaine du *downstep* c'est à dire des l'abaissement des tons hauts les uns par rapport aux autres. Notons que, bien que l'accent de syntagme soit attesté dans différentes langues (pour l'anglais américain Beckman & Pierrehumbert, 1986 ; pour l'italien : D'Imperio, 2000 ; Grice, D'Imperio, Savino & Avesani, 2005 ; pour l'arabe du Caire : Helmuth, 2007), son statut est assez controversé (pour le hollandais : Gussenhoven, 1988 ; pour l'allemand : Féry, 1993).

Pour Pierrehumbert (1980), les accents mélodiques et les tons périphériques sont générés par une grammaire à états finis qui est une des trois composantes principale de l'intonation, les deux autres étant la représentation métrique du texte et les règles d'association entre contour intonatif et texte (figure 4).

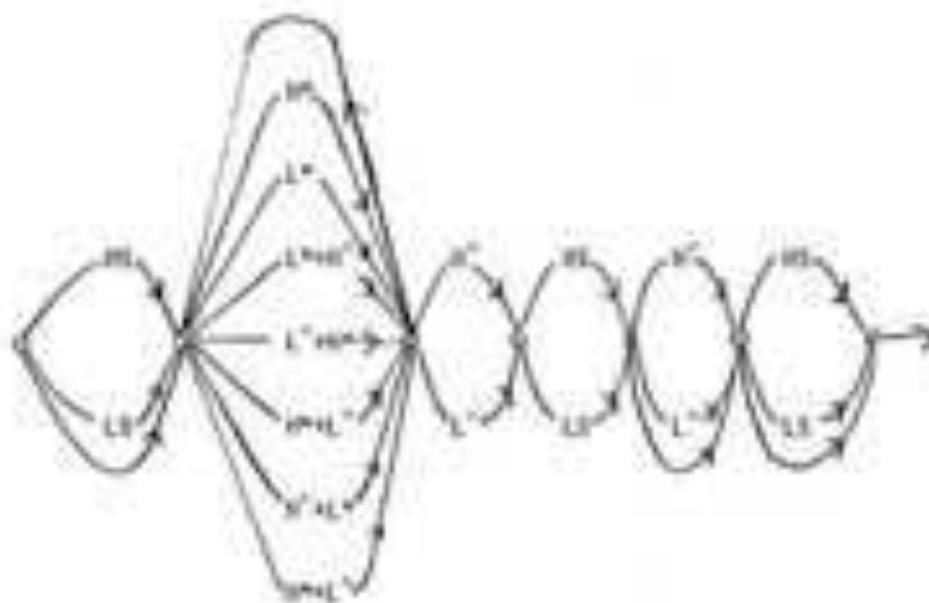


Figure 4. Grammaire de l'intonation de l'anglais selon la théorie auto-segmentale métrique de l'intonation.

Tirée de Pierrehumbert (1980), p.29. Le diacritique % désigne les tons de frontière, le diacritique * désigne les accents mélodiques, le diacritique – désigne les accents de syntagme. Les lignes noires représentent les transitions d'un état à l'autre et les flèches la direction des transitions. La présence de la flèche vers la gauche indique que l'état peut être répété plusieurs fois.

La grammaire de Pierrehumbert est dite « à états finis » car le répertoire de catégories tonales est fini, discret et spécifique à la langue étudiée. De manière plus précise, sept accents mélodiques, deux accents de phrases et deux tons de frontière ont été mis en évidence pour l'anglais américain. Cette grammaire est également dite « non-déterministe » car toutes les combinaisons entre les états du système sont possibles. En effet, tous les types d'accent mélodique peuvent être combinés avec tous les types d'accent de syntagme, tous les types d'accent de syntagme pouvant à leur tour être combinés avec tous les types de ton de frontière. La libre combinaison des tons est liée à l'idée de compositionnalité du sens intonatif qui est à la base de plusieurs théories de l'intonation (Gussenhoven, 1984 ; Pierrehumbert & Hirschberg, 1990 ; Steedman, 2003 ; *inter alia*).

Dans le cadre de la théorie auto-segmentale métrique, les constituants prosodiques sont structurés de façon hiérarchique. Deux niveaux de structure prosodique sont reconnus dans la description de l'anglais américain de Pierrehumbert : le syntagme intonatif (*Intonational Phrase*, IP) et le syntagme intermédiaire (*Intermediate Phrase*, ip). Le syntagme intonatif domine le syntagme intermédiaire selon le principe de la *Strict Layer Hypothesis* que nous avons précédemment évoqué dans ce chapitre (voir section 2.1). Selon les langues, un troisième niveau de structure, de rang inférieur au syntagme intermédiaire, appelé syntagme accentuel (*Accentual Phrase*, AP), a été mis en évidence (pour le français : Jun & Fougeron 1995, 2000, 2002 ; pour le

coréen : Jun, 1993 ; pour le japonais : Beckman & Pierrehumbert, 1986 ; pour l'arabe du Caire : Helmuth, 2004, 2007). Par exemple en arabe du Caire (Helmuth 2007), l'AP est signalé au niveau mélodique par l'abaissement des valeurs de *f₀* à la fin du constituant (*final lowering*) ainsi que la remise à niveau du registre au début du syntagme accentuel suivant (*pitch reset*)

4.2. Les structures prosodiques proposées pour le français dans le cadre du modèle auto-segmental métrique de l'intonation

Après avoir évoqué les concepts fondamentaux de la théorie auto-segmentale métrique de l'intonation qui seront utiles à la compréhension de cette thèse, nous allons désormais aborder les différentes hiérarchies prosodiques postulées pour le français dans le cadre de cette théorie. Cependant, nous souhaitons au préalable, rappeler quelques généralités sur le système intonatif du français qui seront nécessaires à la compréhension de ce qui suit.

Tous les modèles de l'intonation du français s'accordent sur l'existence d'un accent final (également dit « primaire ») sur la dernière syllabe pleine (qui ne contient pas de schwa) d'un constituant prosodique de rang supérieur au mot et inférieur au syntagme intonatif. Cet accent est marqué par un allongement de la syllabe (Pasdeloup, 1990 ; Fletcher, 1991 ; Jun & Fougeron, 2000 ; *inter alia*) ainsi qu'une montée de fréquence fondamentale si le syntagme n'est pas en position finale de l'énoncé. L'accent primaire est associé à la frontière droite d'un constituant prosodique appelé différemment selon les modèles (syntagme accentuel dans le modèle proposé par Jun & Fougeron, 1995, 2000, 2002, ou syntagme phonologique dans le modèle proposé par Post, 2000). En plus de cette montée finale de *f₀*, une montée initiale de *f₀* apparaissant au début du syntagme a été mise en évidence. Contrairement à la montée finale, la montée initiale de *f₀* n'est pas associée à un allongement de la syllabe (Pasdeloup, 1990 ; Mertens, Goldman, Wehril & Gaudinat, 2001 ; Welby, 2003 ; Astésano, 2001 ; *inter alia*). Les facteurs qui influencent l'apparition de la montée initiale ne sont pas très clairs. Il a été proposé qu'un grand nombre de syllabe et qu'un ralentissement de la vitesse d'élocution favorise son apparition (Welby, 2003, 2006 ; Astésano, 2001 ; Jun & Fougeron, 2000, *inter alia*). Un exemple du syntagme nominal « les chatons sages » extrait de la phrase « les chatons sages buvaient leur lait » réalisé soit uniquement avec une montée finale (figure 5a), soit incluant une montée initiale additionnelle (figure 5b) est illustré dans la figure suivante.

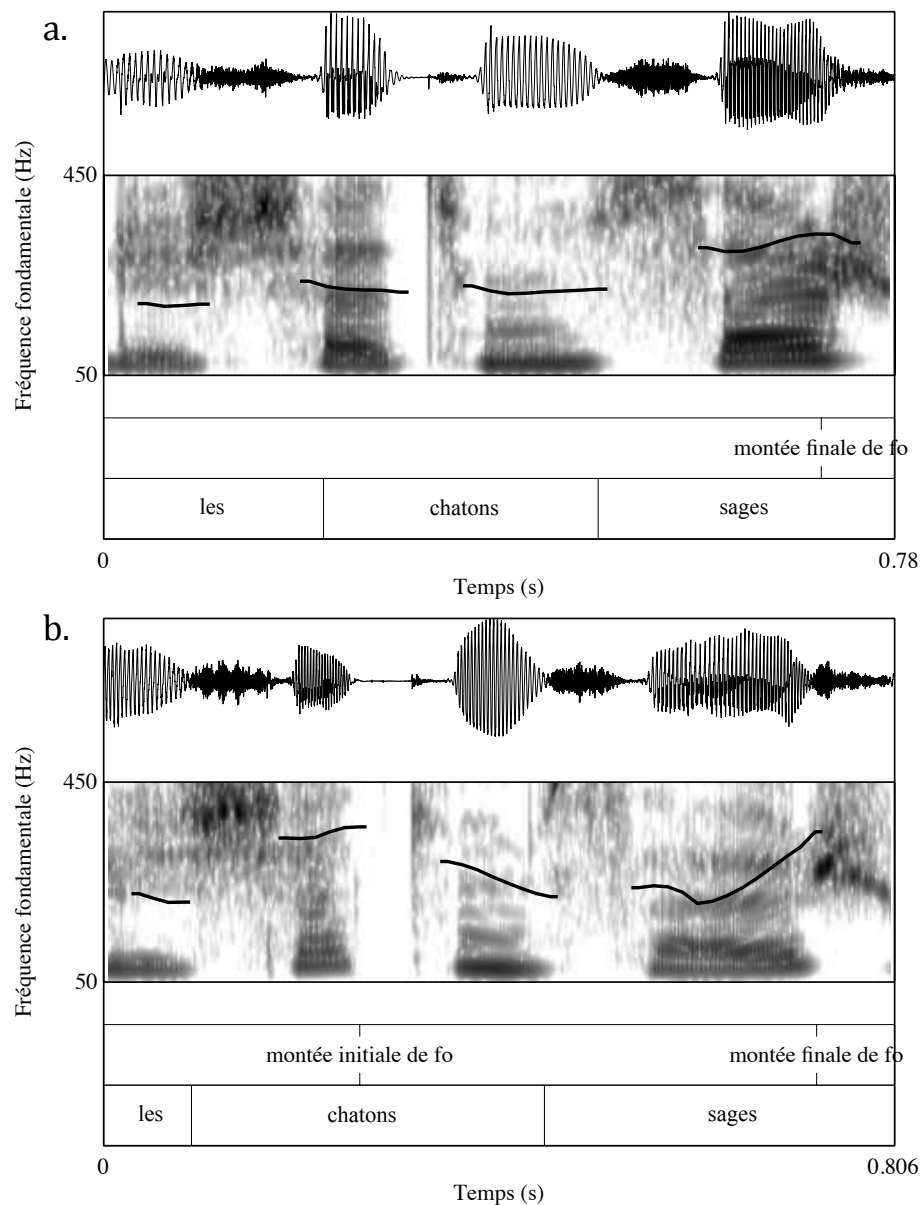


Figure 5. Illustration du syntagme nominal « les chatons sages ». Syntagme réalisé soit avec une montée finale (5a), soit avec une montée initiale optionnelle (5b).

Ces quelques précisions étant données nous allons maintenant décrire plusieurs modèles proposés dans la littérature afin de rendre compte de la structure prosodique du français. Comme nous l'avons précédemment évoqué, les approches qui se sont intéressées à la structure prosodique du français peuvent globalement être scindées en deux groupes en fonction des critères utilisés pour les définitions des constituants :

- les approches qui définissent les constituants prosodiques selon des critères morphosyntaxiques et phonologiques

- (Selkirk 1978, 1984, 1986 ; Nespor & Vogel, 1986; Hayes, 1989 ; Post, 2000),
- celles qui définissent les constituants prosodiques selon de critères purement prosodiques (Beckman & Pierrehumbert, 1988 ; Hirst & Di Cristo, 1993 ; Jun & Fougeron, 1995, 2000, 2002).

Bien que nous reconnaissons que les différents niveaux de la Grammaire (dont la syntaxe) contraignent de manière importante le découpage en constituants prosodiques, nous adopterons dans cette thèse une définition des constituants prosodiques basée sur leurs propriétés phonétiques et phonologiques observables à partir de la représentation de surface du signal de parole.

L'objectif de cette section n'est pas de présenter toutes les hiérarchies prosodiques présentées dans la littérature mais de se focaliser sur celles qui ont été proposées dans le cadre de la théorie auto-segmentale métrique, cadre théorique adopté dans cette thèse.

4.2.1. Le modèle de Hirst & Di Cristo

La première tentative de description de la structure prosodique du français dans le cadre de la théorie auto-segmentale métrique de l'intonation a été réalisée par Hirst & Di Cristo (Hirst & Di Cristo, 1984, 1996, 1998; Di Cristo & Hirst 1993, 1996 ; Di Cristo, 1998, 1999a, 1999b, 2000, 2010). Cette approche postule l'existence de trois niveaux de structure prosodique: l'unité tonale, l'unité rythmique (également appelée syntagme prosodique dans les développements les plus récents du modèle) et l'unité intonative. L'unité tonale est définie comme une unité minimale de synchronisation des segments tonals avec les séquences de phonèmes constitutifs de l'énoncé. Elle est délimitée par un ton L à sa frontière gauche et un ton H à sa frontière droite. Celle-ci est donc dotée du patron mélodique /LH/. Notons que chaque mouvement de *fo* a son propre domaine prosodique qui ne correspond pas nécessairement avec une frontière de mot. Dans ce modèle, la tête métrique de l'unité tonale peut correspondre aussi bien à l'accent initial qu'à l'accent final.

L'unité prosodique supérieure à l'unité tonale dans la hiérarchie prosodique est l'unité rythmique. Notons que cette unité n'est pas présente dans les premières versions du modèle (Hirst & Di Cristo, 1984) mais a été introduite un peu plus tardivement à l'occasion d'une révision du modèle (Di Crist & Hirst, 1993). Cette unité a également été appelée syntagme prosodique dans les développements plus récents du modèle (Di Cristo, 1998, 1999a, 1999b, 2010). L'unité rythmique a été postulée dans le but de rendre compte du fait que l'accent final (également dit *primaire*) mais pas l'accent initial (également dit *secondaire*) soit accompagné d'une marque d'allongement. L'unité rythmique n'affecte pas la structure tonale. Elle est présente uniquement pour différencier les syllabes finales d'unité tonale porteuses

d'une marque d'allongement des syllabes finales d'unité tonale qui n'en portent pas.

L'unité intonative est une unité de rang supérieur à l'unité tonale et à l'unité rythmique. Cette unité est équivalente au syntagme intonatif que nous avons déjà évoqué précédemment. Dans la représentation phonologique de l'intonation, le patron mélodique /LH/ est attribué à l'unité intonative non terminale et /LL/ à l'unité intonative terminale. Dans les deux cas, le premier ton est associé à la frontière gauche de l'unité intonative et le second à sa frontière droite. Les tons L et H initiaux et finaux de l'unité intonative sont comparables aux tons de frontières décrits par les autres modèles (c.f. modèles de Post et de Jun & Fougeron décrits dans les deux sections à venir).

Une quatrième unité est également identifiée dans certaines versions du modèle (Di Cristo & Hirst, 1996 ; Di Cristo, 2010). Cette unité, appelée segment d'unité intonative (Segment d'UI) est située entre l'unité rythmique (ou syntagme prosodique) et l'unité intonative dans la hiérarchie prosodique. Cette unité a été postulée dans le but d'expliquer la présence de plateaux (régions plates de *fo*) de niveau haut ou bas qui ont été observés sur des structures syntaxiques particulières : des énoncés interrompus, des questions tags (*tag questions*) ou des questions avec morphème interrogatif initial (*wh-questions*), des structures disloquées, extraposées ou clivées. Plutôt que de former une unité intonative avec ses propres spécifications tonales, le segment d'unité intonative est dépourvu de ton de frontière et se résout à une mélodie uniformément plate. Dans la figure 6 suivante, la structure extraposée droite constitue un segment d'unité intonative. La frontière droite du segment d'unité intonative est dépourvue de ton spécifique. Les auteurs précisent qu'en cas de segment d'unité intonative en position finale, la réalisation du segment d'unité intonative est opérée par un phénomène de copie de ton (*tone copy*) du ton de frontière bornant la frontière droite de l'unité intonative auquel le segment d'unité intonative est associé. Pour plus d'information sur ce point, nous renvoyons le lecteur à Gussenhoven (1988) qui définit le phénomène de *tone copy* en tant que copie du ton de frontière de la spécification précédente. Le contour plat de *fo* n'est pas explicité de manière claire pour les segments d'unité intonative en position non finale d'unité intonative si ce n'est que la frontière droite du segment est dépourvue de ton spécifique.

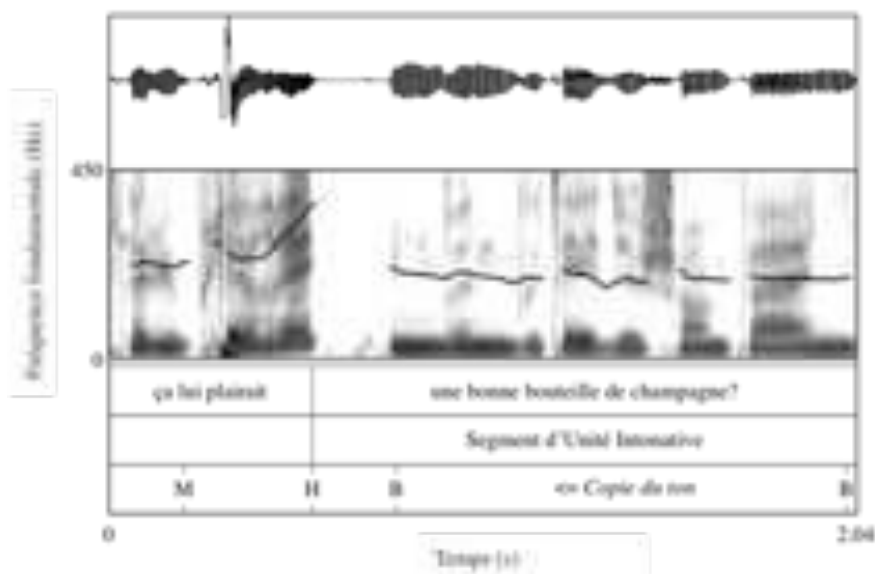


Figure 6. Illustration du segment d'unité intonative selon Di Cristo & Hirst (1996). Unité intonative « Ca lui plaisait, une bonne bouteille de champagne. » dans laquelle l'élément extraposé à droite constitue un segment d'UI. La représentation de surface est annotée [MHBB]. La réalisation du segment d'unité intonative est opérée par un phénomène de copie de ton du ton de frontière bornant la frontière droite de l'unité intonative auquel le segment d'UI est associé.

Di Cristo et collègues reconnaissent quatre niveaux de représentations prosodiques qui sont, du plus concret au plus abstrait, le niveau physique, le niveau phonétique, le niveau phonologique de surface et le niveau phonologique profond (cf. Di Cristo, 1998 pour plus d'information sur ce point). Ces auteurs admettent donc l'existence de deux niveaux phonologiques distincts. Les tons des unités tonales et des unités intonatives sont définis au niveau phonologique le plus profond (niveau de représentation phonologique le plus abstrait) puis transformé dans une représentation phonologique de surface (niveau le moins abstrait de représentation phonologique) avant d'être implémenté phonétiquement. Il a été récemment argumenté que ce niveau intermédiaire de représentation phonologique est difficilement postulable (Pierrehumbert, 2000 ; D'Imperio et al., 2005)

4.2.2. Le modèle de Post

Dans son travail de thèse, Post (2000) propose un modèle de structuration mélodique dans lequel la distribution des accents mélodiques répond à un ensemble de contraintes proposées dans le cadre de la Théorie de l'Optimalité (Prince & Smolensky, 1993). La forme de sortie est celle qui satisfait au mieux la hiérarchie des contraintes admises dans la langue. Les différences de distribution des accents observées entre les langues sont expliquées par une hiérarchisation différente des contraintes. Les contraintes prises en compte dans le modèle de Post sont de nature morphosyntaxique et

rythmique. Deux unités prosodiques sont reconnues par ce modèle : le syntagme intonatif (*Intonation Phrase*) et le syntagme phonologique (*Phonological Phrase*).

Les primitives tonales sont spécifiées au niveau phonologique de la représentation. Ces primitives sont constituées des accents mélodiques (H^* et $H+H^*$) et des tons de frontières (pour la frontière droite d'IP : $L\%$ et $H\%$; pour la frontière gauche d'IP : $L\%$ $H\%$ et 0%). Le symbole 0% ne représente aucune catégorie tonale mais est utilisé en tant que marqueur de frontière (voir Grabe, 1988 qui a proposé la spécification de la frontière 0 pour l'anglais britannique et l'allemand). La figure 7 illustre la grammaire proposée par Post pour la spécification tonale de l'IP.

$$\left\{ \begin{array}{c} \%L \\ \%H \end{array} \right\} (H * (L))_0 \left\{ \begin{array}{c} H^* \\ H + H^* \end{array} \right\} \left\{ \begin{array}{c} L\% \\ H\% \\ 0\% \end{array} \right\}$$

Figure 7. Spécifications tonales du syntagme intonatif selon Post (2000, p°155).
Les accolades indiquent les différents tons pouvant être sélectionnés et les parenthèses indiquent les tons optionnels. Le diacritique * est utilisé pour désigner les accents mélodiques tandis que le diacritique % indique les tons de frontière.

Dans le modèle proposé par Post, le syntagme intonatif commence et se termine par un ton de frontière ($L\%$ ou $H\%$). Le ton de frontière initial est suivi par un accent mélodique initial optionnel (H^*) qui est obligatoirement suivi d'un accent mélodique final (H^* ou H^*+H^*). Dans le cas où deux tons H se suivent, un ton L peut être inséré entre les deux tons hauts en fonction du choix opéré par le locuteur (*L-insertion*). Dans ce modèle la montée initiale (associée à l'accent secondaire) et la montée finale (associée à l'accent primaire) sont toutes deux traitées comme des accents mélodiques. La première cible basse du syntagme intonatif est traitée en tant que ton de frontière initial ($L\%$) alors que la seconde cible basse de l'IP est simplement décrite en terme d'un ton L inséré entre deux tons H sans que les propriétés phonologiques et d'alignement de cette cible ne soient clairement décrites.

Le modèle de Post postule également un second niveau de structuration inférieur à l'intonation Phrase dans la hiérarchie prosodique : le syntagme phonologique. La règle de formation du syntagme phonologique fait appel à la théorie de la syntaxe X-barre (Jackendoff, 1977). Cette règle est inspirée de la règle de formation des syntagmes phonologiques décrite par la théorie de la phonologie prosodique (Nespor & Vogel, 1984). Elle postule que le syntagme phonologique regroupe une tête lexicale et tous les items situés du côté non récursif de celle-ci (c'est-à-dire le côté gauche en français) à l'intérieur d'une projection syntaxique maximale. Les adjectifs pré-nominaux ne peuvent pas servir de tête lexicale pour la formation du syntagme phonologique. Un certain nombre de contraintes d'alignement de taille ou de

symétrie que nous ne détaillerons pas ici sont également proposées dans le modèle.

4.2.3. Le modèle de Jun & Fougeron

Jun & Fougeron ont à leur tour élaboré un modèle phonologique de la structure prosodique du français se référant au cadre auto-segmental métrique de l'intonation. Conformément à ce cadre, les énoncés sont constitués de séquences de tons phonologiques hauts (*Hight*, H) et bas (*Low*, L) qui sont associés aux syllabes métriquement fortes et aux frontières des unités prosodiques. Jun & Fougeron proposent deux catégories d'unités prosodiques pour le français : le syntagme intonatif (*Intonation Phrase*, IP) et le syntagme accentuel (*Accentual Phrase*, AP ; Jun & Fougeron 1995, 2000, 2002). Le syntagme accentuel est admis comme l'unité de base du système intonatif du français. Un AP contient au moins un mot de contenu qui est optionnellement précédé par un mot de contenu additionnel et/ou des mots outils. Le syntagme accentuel contient en moyenne 2.3-2.6 mots (1,2 mot de contenu) et 3.5 à 3.9 syllabes (d'après les corpus utilisés par Jun & Fougeron, 2000). Ces observations confirment ce qui avait été proposé précédemment par d'autres auteurs (Fonagy, 1979 ; Pasdeloup, 1990). D'après l'analyse de Jun & Fougeron, l'AP se caractérise par défaut par le patron mélodique /LHiLH*/. La montée initiale correspondant à l'accent initial (Hi) est décrite en tant que ton périphérique (*edge tone*) tandis que la montée finale correspondante à l'accent finale (H*) est décrite en tant qu'accent mélodique (*pitch accent*). Cette différence de traitement est justifiée par la différence de comportement des deux tons H. Les auteurs ont en effet observé que le H associé à la montée initiale n'est pas toujours réalisé sur une syllabe pleine (qui ne contient pas de schwa) et n'est pas associée à un allongement pré-final. Le H associée à la montée finale est en revanche systématiquement réalisée sur la dernière syllabe pleine de l'AP et est systématiquement associé à un allongement de la syllabe. La hiérarchie prosodique proposée dans le cadre du modèle de Jun & Fougeron est représentée dans la figure 8.

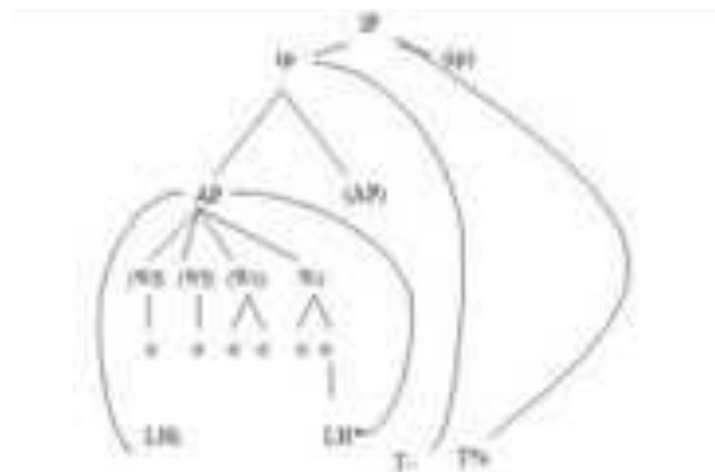


Figure 8. Hiérarchie prosodique postulée pour le français tirée de Jun & Fougeron (2000, 2002).

Les appellations « IP » sont employées pour désigner le syntagme intonatif (Intonational Phrase), « ip » pour syntagme intermédiaire (Intermediate Phrase) « AP » pour syntagme accentuel (Accentual Phrase), « Wf » pour mot outil (function word), « Wc » pour mot à contenu (content word), « σ » pour syllabe, « LHi » pour montée initiale de fo, « LH* » pour montée finale de fo, « T- » pour accent de syntagme, T% pour ton de frontière.

Le modèle de Jun & Fougeron prédit cinq variantes de surface formées par l'absence d'un ou de plusieurs tons du patron mélodique par défaut /LHiLH*/. Les facteurs influençant les variantes de surface ne sont pas explicitement évoqués. Il semble cependant que le patron mélodique par défaut ne soit pas accompli quand il n'y a pas assez de matériel segmental permettant la réalisation de tous les tons (Jun & Fougeron, 2002 : Welby, 2006) ou lorsque le locuteur n'a pas le temps, en raison d'une vitesse d'élocution rapide par exemple, de le réaliser (Fougeron & Jun, 1998). La figure 9 illustre les six réalisations de surface proposées par le modèle.

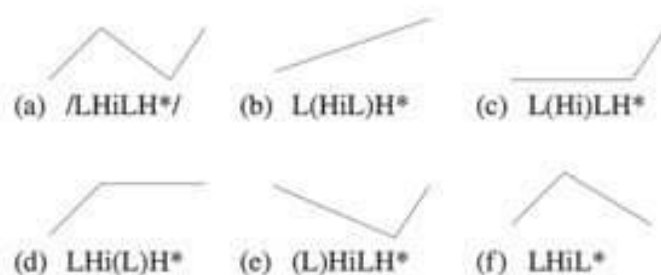


Figure 9. Les six réalisations de surface du syntagme accentuel tirée de Jun & Fougeron (2002).

Les parenthèses indiquent les tons du patron intonatif par défaut qui ne sont pas réalisés dans la réalisation de surface. Le diacritique * est utilisé pour désigner l'accent mélodique.

La version la plus récente du modèle proposée en 2002, s'attache au réexamen des règles d'associations tonales avec les syllabes au sein de l'AP (Jun & Fougeron, 2002). Les auteurs abandonnent l'idée de l'association d'un

ton avec une seule syllabe au profit de celle qui admet que deux tons puissent être associés à une même syllabe. Dans cette optique, Jun & Fougeron considèrent que la séquence bitonale /LHi/ n'est pas associée à une syllabe particulière mais avec la frontière gauche de l'AP. En revanche le patron mélodique bitonal /LH*/ est associé à une syllabe déterminée qui porte l'accent final correspondant à la dernière syllabe pleine de l'AP. Cette association lui confère, par référence au model de Pierrehumbert (1980), le statut particulier d'accent mélodique (*pitch accent*) et l'attribution de l'astérisque (*).

Le modèle inclus également une unité de rang supérieur à l'AP dans la hiérarchie prosodique : le syntagme intonatif (*Intonation Phrase*, IP). Un IP contient un AP ou plus. La frontière droite d'IP est marquée par un ton de frontière H% ou L% qui est accompagné par un allongement de la syllabe significativement plus important que celui associé à la frontière droite d'AP. La frontière d'IP est optionnellement suivie par une pause silencieuse. Les relations entre les tons de frontières et les autres tons ne sont pas explicitement évoquées par les auteurs. Elles précisent toutefois que pour les déclaratives (et quelques autres constructions telles que les questions avec morphème interrogatif initial, *wh-questions*), le ton H* en final d'AP est neutralisé par le ton de frontière. Le modèle de Jun & Fougeron est illustré dans la figure 8.

Dans une étude parue en 2000, Jun & Fougeron propose également un niveau additionnel de structuration prosodique situé entre l'AP et l'IP et appelé syntagme intermédiaire (*Intermediate Phrase*, ip). Le syntagme intermédiaire est postulé dans le but de rendre compte de contours intonatif spécifiques : un plateau de *fo* qui peut être haut, bas ou médian, observé en relation avec certaines structures syntaxiques telles que des questions tags, dislocations, questions à morphème interrogatif initial (*wh-questions*) et clivées. Le syntagme intermédiaire postulé dans ce modèle est à rapprocher du segment d'unité intonative observé par Di Cristo & Hirst (1996). Ces deux constituants (segment d'unité intonative et syntagme intermédiaire) sont en effet observés à partir du même type de structures syntaxiques et pour expliquer le même type de phénomène intonatif : un mouvement mélodique plat. Dans le cadre du modèle de Jun & Fougeron, ces plateaux sont analysés comme des séquences L-L% (dans le cas d'un plateau bas) ou H-H% (dans le cas d'un plateau haut) ou H-L% (dans le cas d'un plateau médian). Dans le cas d'un plateau médian, de manière comparable à ce qui a été observé en anglais (Ladd, 1996 ; Pierrehumbert, 1980) le H- est abaissé par rapport au dernier accent mélodique et le L% est réhaussé, ce qui donne lieu à un plateau médian dans la réalisation de surface (voir Jun & Fougeron 2000 pour plus de détails sur ce point). Dans les trois cas (L-L%, H-H% et H-L%) l'accent de syntagme (L- ou H-) s'étend du dernier accent mélodique jusqu'à la dernière syllabe de l'IP. La figure 10 reprend le syntagme intonatif « *Ca lui plairait, une bonne bouteille de champagne.* » représenté et analysé selon le modèle de Di Cristo & Hirst en figure 6. Le syntagme intonatif est cette fois-ci analysé selon le modèle proposé par Jun & Fougeron. Selon les auteurs, il est composé de deux

syntagmes intermédiaires. L'accent de syntagme L- associé à la frontière d'ip engendre un plateau bas qui s'étend du dernier accent mélodique jusqu'à la frontière d'IP.

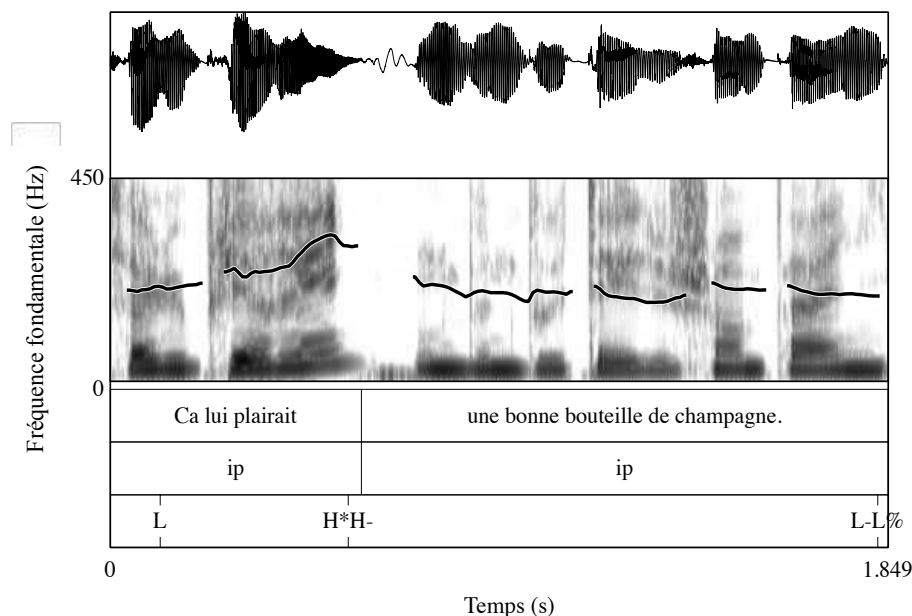


Figure 10. Illustration du syntagme intermédiaire selon Jun & Fougeron (2000). Onde sonore, spectrogramme et courbe de fréquence fondamentale associés à l'unité intonative « Ca lui plairait, une bonne bouteille de champagne. » dans laquelle à la clause principale constitue le premier syntagme intermédiaire (ip) et l'élément extraposé à droite constitue le deuxième syntagme intermédiaire de l'unité intonative. Le plateau est analysé en tant qu'accent de syntagme L- qui s'étend du dernier accent mélodique jusqu'à la frontière droite de l'unité intonative.

Dans cette thèse nous allons chercher à mieux comprendre les relations entre la structure prosodique abstraite et sa représentation concrète dans la parole. Pour cela nous allons nous attacher à fournir une meilleure caractérisation des propriétés phonétiques et phonologiques du syntagme intermédiaire en français. Notre analyse sera menée dans le cadre de la théorie auto-segmentale métrique de l'intonation. Etant donné les objectifs de notre étude, plusieurs points de divergence entre les trois modèles que nous venons de présenter, nous semblent importants à souligner.

Tout d'abord, à la différence de l'approche proposée par Jun & Fougeron, les modèles postulés par Hirst & Di Cristo et Post ne proposent pas d'établir une distinction structurelle entre montée initiale et montée finale de *fo*. Dans le modèle de Post, les deux montées mélodiques sont considérées comme des accents mélodiques et indiquent une frontière de syntagme phonologique tandis qu'elles sont traitées comme des marqueurs de l'unité tonale dans le modèle proposé par Hirst & Di Cristo. Malgré le fait que Di Cristo et ses collègues aient postulé une unité rythmique (ou syntagme prosodique) de rang supérieur à l'unité tonale pour rendre compte du fait que la montée initiale ne soit associée à aucune marque d'allongement alors que la montée finale l'est, l'unité rythmique n'est pas marquée dans la structure

intonative selon leur modèle. Récemment, les études menées par Welby (2003, 2006) ont permis de mettre en évidence le fait que les montées initiales et finales sont structurées différemment, tant du point de vue de la durée que du point de vue de leur alignement tonal. L'auteur met en évidence le fait que la montée initiale et la montée finale ont seulement en commun de pouvoir être décrites comme une séquence LH. Le ton L de la montée initiale est souvent aligné avec la frontière entre un mot outil et un mot à contenu et peut être déplacé vers la gauche de manière à marquer le début de l'AP tandis que l'alignement du ton L final semble beaucoup plus variable selon les locuteurs et même chez un même locuteur. Le ton H de la montée finale peut être réalisé sur la première ou la deuxième syllabe du premier ou second mot à contenu de l'AP (voire plus tard). En revanche, le ton H de la montée finale est systématiquement réalisé sur la dernière syllabe de l'AP. De plus, le H final est accompagné par un allongement de la syllabe alors que le H initial ne l'est pas. Enfin, du point de vue articulatoire, il apparaît que le H initial soit réalisé avec une amplitude de mouvement plus faible que le H final (Rolland & Loevenbruck, 2002). Ces différences vont l'encontre du modèle proposé par Hirst and Di Cristo qui affirme que les deux montées sont tonalement structurées de façon identique puisqu'elles forment toutes deux des unités tonales. Les résultats de Welby vont également à l'encontre du modèle proposé par Post qui postule que les deux H (initiaux et finaux) sont des accents mélodiques. L'auteur n'observe pas d'association systématique de la montée initiale avec une syllabe particulière. Or, l'association d'un ton avec une syllabe métriquement forte est une des caractéristiques principales de l'accent mélodique. Les résultats obtenus par Welby nous poussent donc à considérer la montée initiale comme un ton périphérique bitonal tandis que la montée finale est analysée en tant qu'accent mélodique bitonal, conformément à ce qui avait été postulé par Jun & Fougeron (2000).

Si les trois modèles divergent sur la question d'une structuration différente entre montée initiale et montée finale, ils ne s'accordent pas non plus sur le nombre de constituants à observer pour rendre compte de la hiérarchie prosodique du français. Le modèle de Post postule l'existence de deux constituants (le syntagme phonologique et le syntagme intonatif) sans envisager l'existence d'un constituant prosodique intermédiaire entre ces deux niveaux. Au contraire, les hiérarchies postulées par Hirst & Di Cristo et Jun & Fougeron envisagent un constituant prosodique de rang intermédiaire appelé *segment d'unité intonative* chez Di Cristo et ses collègues et *syntagme intermédiaire* chez Jun & Fougeron. À la lumière de récents résultats obtenus pour le français (Michelas & D'imperio, 2010b) qui montrent que le degré d'allongement en position finale d'AP est plus important lorsque la frontière prosodique est alignée avec une frontière de constituant syntaxique majeur (comme la frontière entre un syntagme nominal et un syntagme verbal), il nous semble intéressant d'envisager l'existence d'un niveau de structure additionnel. Cependant, les niveaux intermédiaires postulés par chacun des deux modèles ne nous semblent pas satisfaisants car restreints à des structures syntaxiques très spécifiques. De plus les propriétés phonétiques et

phonologiques de ces constituants ne sont clairement explicitées dans aucun des deux modèles. Nous nous attacherons donc dans cette thèse à voir si les propriétés phonétiques que nous observons dans la représentation de surface de la parole nous permettent de postuler l'existence d'un constituant prosodique de niveau intermédiaire situé entre l'AP et l'IP dans la structure prosodique.

Enfin, il nous paraît important de souligner que même si le modèle de Di Cristo et ses collaborateurs est le seul à postuler un constituant prosodique inférieur à l'AP, nous ne traiterons pas cette question dans la thèse². Dans cette thèse nous nous intéressons aux constituants prosodiques supérieurs au mot lexical. Nous ne nous focaliserons donc pas sur la problématique de l'unité tonale bien que nous reconnaissons que cette question pose des problèmes et mérite d'être approfondie.

5. OBJECTIFS DE LA THÈSE

Après avoir mieux défini ce qu'est la structure prosodique, décrit le cadre théorique et développé les modèles proposés dans ce cadre, nous allons proposer cinq études qui nous permettront de mieux comprendre l'articulation entre la structure prosodique abstraite d'une part et sa réalisation concrète d'autre part.

Dans une première étude, nous nous intéresserons à la question de l'interface entre syntaxe et prosodie. Si de nombreuses études s'accordent sur l'existence d'un module abstrait à l'interface entre structure syntaxique et représentation phonétique concrète de la prosodie, d'autres mettent en évidence une relation directe entre le niveau d'un constituant syntaxique dans la hiérarchie grammaticale et les indices prosodiques observés dans le signal de parole. Pour cela nous nous concentrerons sur le niveau de structure du syntagme accentuel. Nous allons observer si les indices tonals et de durée présents à la frontière du syntagme accentuel sont différents lorsque cette frontière est associée avec une frontière de constituant syntaxique majeure (comme une rupture entre un syntagme nominal et un syntagme verbal) plutôt que lorsque cette frontière est contenue à l'intérieur d'un syntagme nominal. Nous utiliserons pour cela un corpus de 1920 phrases lues par six locuteurs et manipuleront également la vitesse d'élocution des locuteurs dans le but de voir quels ont été les influences de la syntaxe d'une part et de la vitesse d'élocution d'autre part, sur la réalisation des indices mélodiques et de durée observés à la frontière du syntagme accentuel.

À la lumière des résultats obtenus dans cette première étude et grâce à la mise en place d'un nouveau corpus de parole lue, nous montrerons qu'un niveau additionnel de structure, le syntagme intermédiaire, est nécessaire pour rendre compte de la structure prosodique du français (Étude 2). Nous

² Notons que la validité de l'unité tonale constituant de la structure prosodique du français a été contestée (1995) sur la base de la violation de la *Strict Layer*.

proposerons une définition explicite des propriétés phonétiques et phonologiques de ce constituant.

Puis, au moyen de trois expérimentations menées en perception de la parole, nous proposerons d'étudier l'influence des deux niveaux de structure prosodique (le syntagme accentuel et le syntagme intermédiaire) sur le traitement perceptif de la parole.

Dans une première expérimentation (Étude 3) nous utiliserons une tâche de détection de mot afin de tester l'influence de deux types de frontières : (i) la frontière droite de syntagme accentuel associée à une rupture syntaxique majeure et (ii) la frontière droite de syntagme accentuel contenue au sein d'un constituant syntaxique majeure sur l'accès au lexique en français. Les études menées jusqu'à présent sur le français se basent sur une définition syntaxique des constituants prosodiques qui n'a pas permis de mettre en évidence le type d'indices acoustiques (indices marquant une frontière d'AP ou indices marquant une frontière d'AP associée à une rupture syntaxique majeure) influençant le découpage en mots. Les résultats obtenus nous permettront de montrer que le syntagme accentuel est le domaine de l'accès au lexique en français.

Une deuxième expérimentation de perception sera ensuite décrite afin d'étudier le rôle de la frontière droite de syntagme intermédiaire sur le traitement syntaxique des énoncés (Étude 4). Nous utiliserons une tâche de complétion de phrases dans le but de voir si la frontière de syntagme intermédiaire contraint, de manière différente de la frontière de syntagme accentuel, le traitement syntaxique de la parole. Finalement grâce à l'utilisation de stimuli de resynthèse (Étude 5), nous montrerons quelle est la contribution de chaque indice (*f₀* et durée) dans le traitement syntaxique des énoncés parlés.

Les résultats obtenus dans cette thèse nous permettront de discuter à la fois les modèles de l'intonation proposés pour le français et les modèles de perception de la parole.

CHAPITRE 2

INFLUENCE DE LA SYNTAXE ET DE LA VITESSE D'ÉLOCUTION SUR LA RÉALISATION DES FRONTIÈRES DE SYNTAGME ACCENTUEL EN FRANÇAIS (Étude 1)

1. PROBLÉMATIQUE

Une des principales questions non résolues dans la littérature est celle des relations existantes entre structure syntaxique d'une part et structure prosodique d'autre part. Nous savons que ces deux structures ne sont pas isomorphes (cf. Shattuck-Hufnagel & Turk pour un résumé). Dans le cadre de la Phonologie Prosodique, l'existence d'une structure phonologique abstraite et autonome à l'interface entre structure syntaxique et structure prosodique est postulée. Les indices marquant les frontières de constituants prosodiques observés dans le signal de parole, sont reliés à la structure prosodique et non pas directement à la structure syntaxique. Cependant les structures syntaxiques et prosodiques ne sont pas pour autant totalement indépendantes l'une de l'autre. Paccia & Cooper (1980) ont proposé un algorithme rendant compte de la distribution de l'allongement final et des pauses en anglais. Leur algorithme est basé sur la force des frontières syntaxiques (c'est-à-dire leur niveau d'emboîtement dans la hiérarchie syntaxique). La durée des voyelles et des pauses est prédite en fonction de la force de la frontière syntaxique (*boundary strength index*). Par la suite, Gee & Grosjean (1983) ont confirmé qu'en anglais, un degré majeur d'allongement final pré-frontière correspond à une rupture syntaxique majeure dans la hiérarchie des constituants morphosyntaxiques tandis qu'un allongement pré-frontière mineur est associé à une rupture syntaxique de plus bas niveau dans la hiérarchie morphosyntaxique. Ces résultats ont par ailleurs été confirmés pour le français par les travaux de Monnin & Grosjean (1993) et de Keller, Zellner, Werner & Blanchoud (1993).

Dans ces études, les auteurs remettent en question le fait que les indices observés à la frontière de constituants prosodiques sont reliés à une hiérarchie des domaines prosodiques. Selon eux, les indices marquant la frontière de constituants prosodiques pourraient être non pas rattachés à une structure prosodique abstraite et autonome mais directement reliés à

une hiérarchie de constituants morphosyntaxiques (telles que la phrase, le syntagme et le mot) et à leur niveau d'emboîtement à l'intérieur de cette hiérarchie morphosyntaxique. Différentes contraintes relatives à la taille des constituants et à leur symétrie (voir les contraintes de symétrie et de longueur égale des unités postulées pour l'anglais par Gee & Grosjean, 1983, ou pour le français par Monin & Grosjean, 1993) viennent s'ajouter à la contrainte syntaxique dans le but de générer ce que les auteurs appellent des *structures de performance* et qui correspondent selon eux à la structure hiérarchique des phrases.

Cependant, depuis les travaux menés dans le cadre de la Phonologie Prosodique, l'existence d'une structure phonologique abstraite assurant l'interface entre structure syntaxique et structure prosodique est traditionnellement admise. Cette vision n'implique pas pour autant que les deux structures, bien qu'autonomes, soient complètement indépendantes l'une de l'autre. Un argument avancé dans le cadre de cette théorie montrant les relations étroites existant entre prosodie et syntaxe est le fait que certaines structures syntaxiques spécifiques telles que des parenthétiques ou des disloquées (voir par exemple Selkirk, 1995 ; Delais-Roussarie & Post, 2008) soient fortement corrélées avec la présence d'une frontière prosodique majeure et optionnellement d'une pause. Dans la théorie prosodique, le syntagme intonatif (IP) se définit comme un domaine prosodique dont l'extension correspond à la phrase racine (Selkirk, 1978), c'est-à-dire à la séquence groupe nominal et groupe verbal. Les syntagmes extraits ou incidents, lorsqu'ils apparaissent, projettent des syntagmes intonatifs autonomes. Ainsi, dans les phrases suivantes tirées de Delais-Roussarie & Post (2008), la parenthétique « à ce qu'on m'a dit » engendre l'apparition de trois IPs autonomes (1b).

1a. [Ces jeunes enfants apprennent à parler le français]IP

1b. [Ces jeunes enfants apprennent]IP [à ce qu'on m'a dit]IP [à parler le français]IP

Cependant, dans le cadre de cette théorie, le syntagme intonatif est défini en fonction de critères morphosyntaxiques indépendamment de ce qui est réellement produit par les locuteurs. Afin de savoir à quel type de frontière prosodique nous avons réellement affaire dans ce type de constructions syntaxiques, des études expérimentales menées sur les indices acoustiques produits à ces frontières seraient donc nécessaires en français.

Nous avons précédemment évoqué le fait que deux niveaux de structure prosodique sont postulés pour le français (correspondant au syntagme intonatif et au syntagme accentuel dans le modèle proposé par Jun & Fougeron). Ces deux niveaux de structuration prosodique sont associés à des marques tonales différentes ainsi que des degrés d'allongement pré-frontières significativement différents. Ainsi les syllabes situées en position finale d'AP sont significativement plus longues que les syllabes inaccentuées

et les syllabes situées en finale d'IP sont significativement plus longues que celles situées en finale d'AP. La figure 11, tirée de Jun & Fougeron 2000, représente la durée des syllabes (exprimée en millisecondes) se trouvant dans quatre positions possibles: en position finale de syntagme intonatif (IP-final), en position finale de syntagme accentuel (AP-final), lorsque la syllabe est associée à une montée initiale de *fo* (autrement dit à un accent secondaire ; initial H) et lorsque la syllabe est inaccentuée (unaccented).

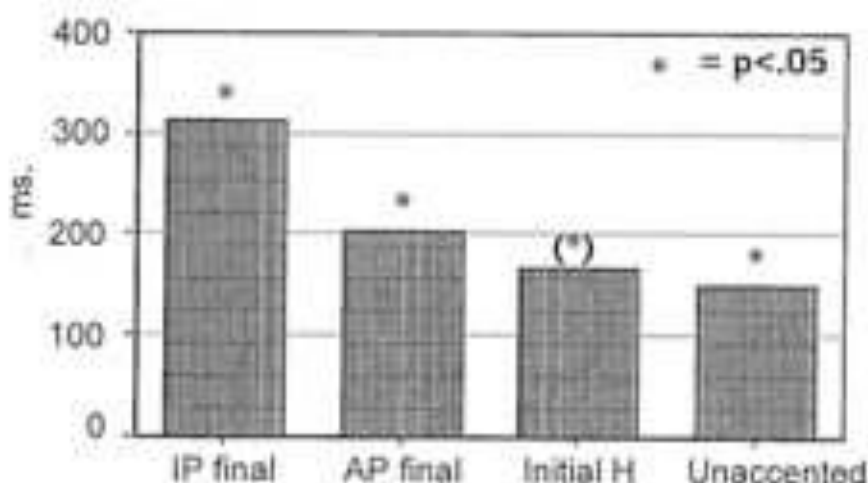


Figure 11. Degrés d'allongement pré-frontière tirés de Jun & Fougeron (2000, p°221). La durée des syllabes (en millisecondes) en position finale de syntagme intonatif (IP-final), en position finale de syntagme accentuel (AP-final), associées à une montée initiale de *fo* (Initial H) et inaccentuées (Unaccented) est représentée dans cette figure.

Bien qu'un niveau de structuration prosodique ait été proposé afin de rendre compte de contours intonatifs spécifiques (l'apparition de plateaux qui s'étendent du dernier accent mélodique jusqu'à la fin du syntagme) aucune étude menée jusqu'à ce jour n'a permis de mettre en évidence l'existence de trois niveaux de structuration prosodique sur la base de l'étude des durée segmentales pour le français. En effet, nous avons précédemment évoqué le fait que Jun & Fougeron aient postulé l'existence d'un niveau additionnel de structure située entre l'AP et l'IP. L'existence de ce syntagme intermédiaire (ip) a été postulée dans le but d'expliquer des régions plates de *fo* qui apparaissent en relation avec des structures syntaxiques spécifiques telles que des disloquées ou des vocatifs. Dans le cadre du modèle de Jun & Fougeron, et conformément à ce qui a été observé pour l'anglais américain (Beckman & Pierrehumbert, 1986), la présence de ces plateaux hauts, bas ou médians est expliquée par la réalisation d'un accent de syntagme (L- ou H-) qui est alignée avec la frontière droite du syntagme intermédiaire et dont une des propriétés est de pouvoir s'étendre du dernier accent mélodique jusqu'à la fin de l'ip. Nous renvoyons le lecteur à la figure 10 de cette thèse pour une illustration de ce phénomène. Cependant aucune étude menée jusqu'à ce jour pour le français ne s'est

intéressée aux indices de durée présents sur la syllabe finale de ce constituant intermédiaire. Si trois niveaux d'allongement pré-frontière étaient observés pour le français, cela pourrait confirmer l'existence d'un niveau additionnel de structuration prosodique.

Dans cette étude, nous nous sommes intéressées à la relation existante entre la durée de la voyelle et le niveau de frontière prosodique qui lui est associé. Au moyen d'un corpus de 1920 phrases recueillies auprès de six locuteurs, nous allons tester si l'interface syntaxe/prosodie affecte les marques de durées observées à la frontière des constituants prosodiques

Pour cela, nous sommes focalisée sur le constituant prosodique de base traditionnellement admis dans la littérature : le syntagme accentuel. Comme nous l'avons précédemment évoqué, la frontière du syntagme accentuel est marquée par un allongement de la dernière syllabe pleine du constituant (c'est-à-dire la syllabe qui ne contient pas de schwa). Les syllabes en finale d'AP sont donc significativement plus longues que les syllabes non accentuées et significativement moins longues que celles situées en finale d'IP (Pasdeloup, 1990 ; Jun & Fougeron, 2000 ; *inter alia*). La frontière d'AP est également signalée par un contour mélodique mineure (une montée de la fréquence fondamentale si l'AP n'est pas en position finale au sein de l'IP) tandis que la frontière d'IP est marquée par un contour mélodique majeur (Delattre, 1966; Vaissière, 1983; Post, 2000; Jun & Fougeron, 2000). Dans l'approche proposée par Jun & Fougeron, la frontière d'AP est associée à un accent mélodique (H*) tandis que la frontière d'IP est associée à un ton de frontière (L% ou H%). Une tendance à placer des frontières prosodiques majeures à la frontière de constituants syntaxiques majeures (comme par exemple une rupture entre un syntagme nominal, SN, et un syntagme verbal, SV) a été démontrée pour le français bien que cette tendance soit influencée par des contraintes eurythmiques (Delattre 1966 ; Rossi, 1997; Vaissière & Michaud, 2006). Cette tendance a également été observée par Delais-Roussarie (1995). Dans la grammaire prosodique que celle-ci a développée dans le cadre de la Théorie de l'Optimalité (Prince & Smolensky, 1993), les structures prosodiques sont analysées comme le résultat de l'interaction entre des contraintes sémantiques, syntaxiques et rythmiques. Dans le cadre de cette grammaire, la correspondance entre rupture SN/SV et frontière de constituant prosodique majeure est exprimée par une contrainte d'alignement appelée « rupture » (cf. Delais-Roussarie, 1995, p° 208-209).

Les effets de la vitesse d'élocution sur la réalisation des frontières prosodiques en français ont également été rapportés par plusieurs études (Bartkhova, 1991; Keller & Zellner, 1993; Fougeron & Jun, 1998). Fougeron & Jun (1998) ont montré que les locuteurs français ont recours à un certain nombre de stratégies dans le but de parvenir à un augmenter leur vitesse d'élocution. Ces stratégies peuvent être soit de nature catégorielle, soit de nature graduelle. Par exemple, plusieurs stratégies dites *catégorielles* impliquent l'effacement ou le remplacement d'une frontière d'un niveau

donné dans la hiérarchie prosodique par une frontière de plus bas niveau. En effet, les auteurs montrent qu'à vitesse d'élocution rapide, certaines frontières d'IP sont remplacées par des frontières d'AP. Ces résultats confirment ce qui avait été mis en évidence par différentes études menées sur d'autres langues que le français. Jun (1993) a montré qu'en Coréen, le nombre de syntagmes accentuels diminue à vitesse d'élocution rapide. Caspers & Van Heuven (1991, 1993, 1995 ; Caspers 1994) ont également montré que le nombre de frontières prosodiques diminue à vitesse d'élocution rapide en hollandais alors que le nombre d'accents mélodiques semble rester stable. Parmi les stratégies dites *graduelles*, Fougeron & Jun (1998) observent une réduction du nombre et de la durée des pauses silencieuses observées après une frontière d'IP et une compression globale du registre.

De plus, Fougeron & Jun (1998) observent une simplification du contour tonal qui est opéré par une non réalisation des quatre tons qui composent le patron mélodique par défaut de l'AP à vitesse rapide³. Les auteurs qualifient ces phénomènes en tant que phénomènes de *restructuration prosodique*.

Fougeron & Jun (1998) montrent également que l'organisation prosodique des unités marquées par l'allongement final pré-frontière est préservée à vitesse d'élocution rapide. Les syllabes situées en finale d'IP sont plus longues que celles situées en finale d'AP, qui sont à leur tour plus longues que les syllabes non accentuées (celles qui ne sont pas situées en finale de constituant prosodique) à vitesse d'élocution normale comme à vitesse d'élocution rapide. Cependant Fougeron & Jun ont observé que la réduction des syllabes à vitesse d'élocution rapide n'affecte pas de manière identique tous les niveaux de structuration prosodique. Cette influence varie en fonction du locuteur. Pour deux de leurs trois locuteurs, la durée des syllabes en position finale d'AP est d'avantage réduite que la durée des syllabes en finale d'IP. L'inverse se produit pour leur troisième locuteur.

L'influence de la vitesse d'élocution sur les indices phonétiques observés à la frontière des constituants prosodiques a donc largement été démontrée. La question de l'influence de l'interface syntaxe/prosodie sur ces mêmes indices reste au contraire une question non résolue. Nous allons nous attacher à combler ce vide théorique dans l'étude qui va suivre.

Pour répondre à cette question nous nous sommes basée sur le niveau de structure de l'AP et nous avons étudié l'interaction entre les indices phonétiques marquant ce niveau de frontière et la force de la frontière syntaxique (c'est à dire son niveau d'emboîtement dans la hiérarchie morphosyntaxique). Plus précisément nous avons comparé la réalisation de la frontière droite d'AP associée à une frontière de constituant

³ Rappelons que le patron mélodique par défaut est le patron mélodique /LHiLH*/ dans le modèle de Jun & Fougeron. Cinq variantes de surface formées par l'absence d'un ou de plusieurs tons du patron mélodique de base sont également prédites par le modèle (cf. figure 9)

syntagme majeure (frontière entre un syntagme nominal et un syntagme verbal ; phrase 2a) avec la réalisation d'une frontière droite d'AP qui est associée à une frontière de constituant syntaxique mineure (frontière interne au syntagme nominal ; phrase 2b).

	2a. Tes a MIS demandaient la présidente.
Structure syntaxique :	]SN
Structure prosodique :	]AP/XP
	2b. Les a MIS de Pauline adoraient ce restaurant.
Structure syntaxique :	]SN
Structure prosodique :	]AP

L'hypothèse nulle était que l'alignement entre une rupture syntaxique majeure et la frontière d'AP (condition AP/XP-finale) n'affecterait pas la durée de la voyelle comme cela est illustré dans la figure 12.

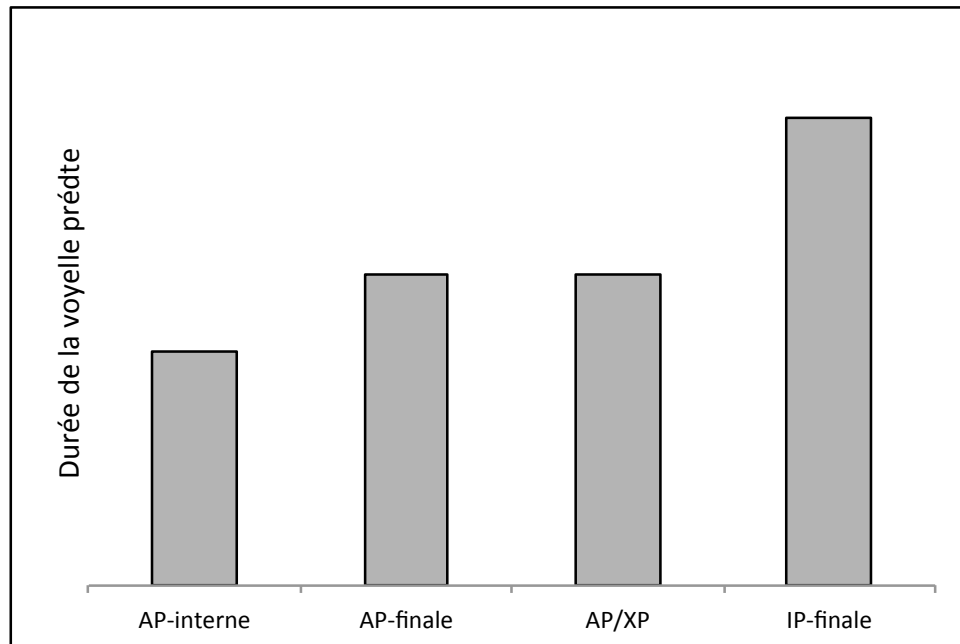


Figure 12. Étude 1 : Hypothèse nulle.

Schématisation de la durée de la voyelle prédite selon l'hypothèse nulle. Les voyelles pouvaient apparaître dans quatre positions différentes : voyelles inaccentuées (AP-interne), les voyelles situées en position finale de syntagme accentuel contenues à l'intérieur du syntagme nominal (AP-finale), les voyelles situées en position finale du syntagme accentuel associé à une rupture syntaxique majeure (AP/XP) et les voyelles situées en position finale de syntagme intonatif (IP-finale).

Contrairement au postulat nul, notre hypothèse était que l'alignement entre la frontière droite de syntagme accentuel et la rupture syntaxique majeure allait donner lieu à un degré d'allongement plus important que celui observé en position finale d'AP contenue à l'intérieur du syntagme nominal. Notre hypothèse est illustrée dans la figure 13.

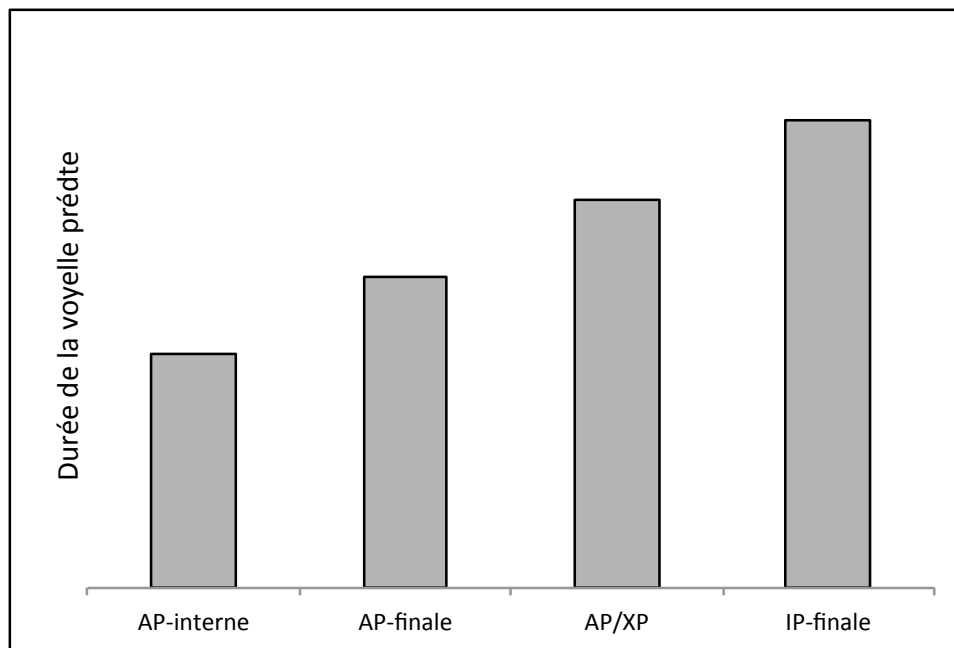


Figure 13. Étude 1 : Hypothèse d'un degré d'allongement pré-frontière additionnel. Schématisation de la durée de la voyelle prédite selon l'hypothèse d'un degré supplémentaire d'allongement pré-frontière en français. Les voyelles pouvait apparaître dans quatre position différentes : voyelles inaccentuées (AP-interne), les voyelles situées en position finale de syntagme accentuel contenues à l'intérieur du syntagme nominal (AP-finale), les voyelles situées en position finale du syntagme accentuel associé à une rupture syntaxique majeure (AP/XP) et les voyelles situées en position finale de syntagme intonatif (IP-finale).

En d'autres termes nous nous sommes attachée à tester si, un seul niveau d'allongement associé à l'AP doit être observé pour le français ou si plus d'un degré d'allongement serait nécessaire pour rendre compte de l'allongement observé en finale d'AP associé à aucune rupture syntaxique majeure et celui observée en finale d'AP associé à une rupture syntaxique mineure. Deux autres niveaux de structuration prosodique ont également été inclus dans notre étude : les syllabes non accentuées (à l'intérieur de l'AP) et les syllabes en finales d'IP. Puisque nous savons que la vitesse d'élocution affecte la réalisation des indices phonétiques marquant les frontières de constituants, nous avons manipulé la vitesse d'élocution dans notre étude. Nous nous sommes intéressée à l'influence que va avoir l'interface syntaxe/prosodie sur les indices phonétiques observés dans le signal de parole à la frontière des constituants prosodiques (et plus particulièrement à la frontière d'AP) tout en prenant en compte l'effet de la vitesse d'élocution puisque son impact sur de tels indices a largement été

démonstré. Les résultats obtenus nous ont permis de discuter les différents modèles postulés jusqu'à présent dans le but de rendre compte de la structure prosodique du français.

2. MÉTHODE

2.1. Matériel utilisé

Une série de 40 phrases a été lue quatre fois par six locuteurs de langue maternelle française (trois hommes et trois femmes) à deux vitesses d'élocution pour un total de 1920 phrases (40 phrases * 4 répétitions * 2 vitesse d'élocution * 6 locuteurs). Nous avons analysé les propriétés acoustiques (durées segmentales et valeurs de *f*₀) des voyelles et syllabes cibles associées à quatre types de frontières prosodiques pour les deux vitesses d'élocution. Deux voyelles cibles ont été choisies pour l'étude /i/ et /a/. Ces deux voyelles pouvaient apparaître dans quatre contextes différents en fonction du type de frontière prosodique associé à la syllabe: en frontière droite de syntagme intonatif (IP-final), en frontière droite de syntagme accentuel associé avec une rupture entre un syntagme nominal et un syntagme verbal (AP/XP), en frontière droite de syntagme accentuel contenu dans un syntagme nominal complexe (AP-final) et en frontière droite de syllabe contenue l'intérieur à d'un syntagme accentuel (AP-interne). Le tableau 1 fait apparaître des exemples de phrases pour les voyelles /i/ et /a/ utilisées dans notre étude. Les ondes sonores, courbes de fréquence fondamentale et spectrogrammes sont représentés dans la figure 14.

Frontière prosodique	Voyelle cible /i/
Interne à l'AP	L'après-midi de ce deux mars devenait interminable.
Finale d'AP	Les amis de Pauline adoraient ce restaurant.
Finale d'AP/XP	Tes amis demandaient la présidente.
Finale d'IP	Tes amis, d'après ce qu'on m'a dit, t'adorent vraiment.
	Voyelle cible /a/
Interne à l'AP	Le calvados de Normandie restera un alcool fort.
Finale d'AP	La diva de la pop restera Withney Houston.
Finale d'AP/XP	Cette diva deviendra incontournable.
Finale d'IP	Cette diva, d'après ce qu'on m'a dit, est la meilleure.

Tableau 1. Étude 1 : Exemple de phrases utilisées dans le corpus.

La voyelle cible (soulignée) pouvait être soit /i/ soit /a/ et pouvait être placée dans quatre conditions prosodiques différentes: à l'intérieur d'un AP (syllabe interne à l'AP) en position finale d'AP dont la frontière droite n'est associée à aucune frontière syntaxique majeure (syllabe en finale d'AP), en position finale d'AP dont la frontière droite est associée avec une rupture syntaxique majeure (syllabe en finale d'AP/XP) en position finale d'IP.

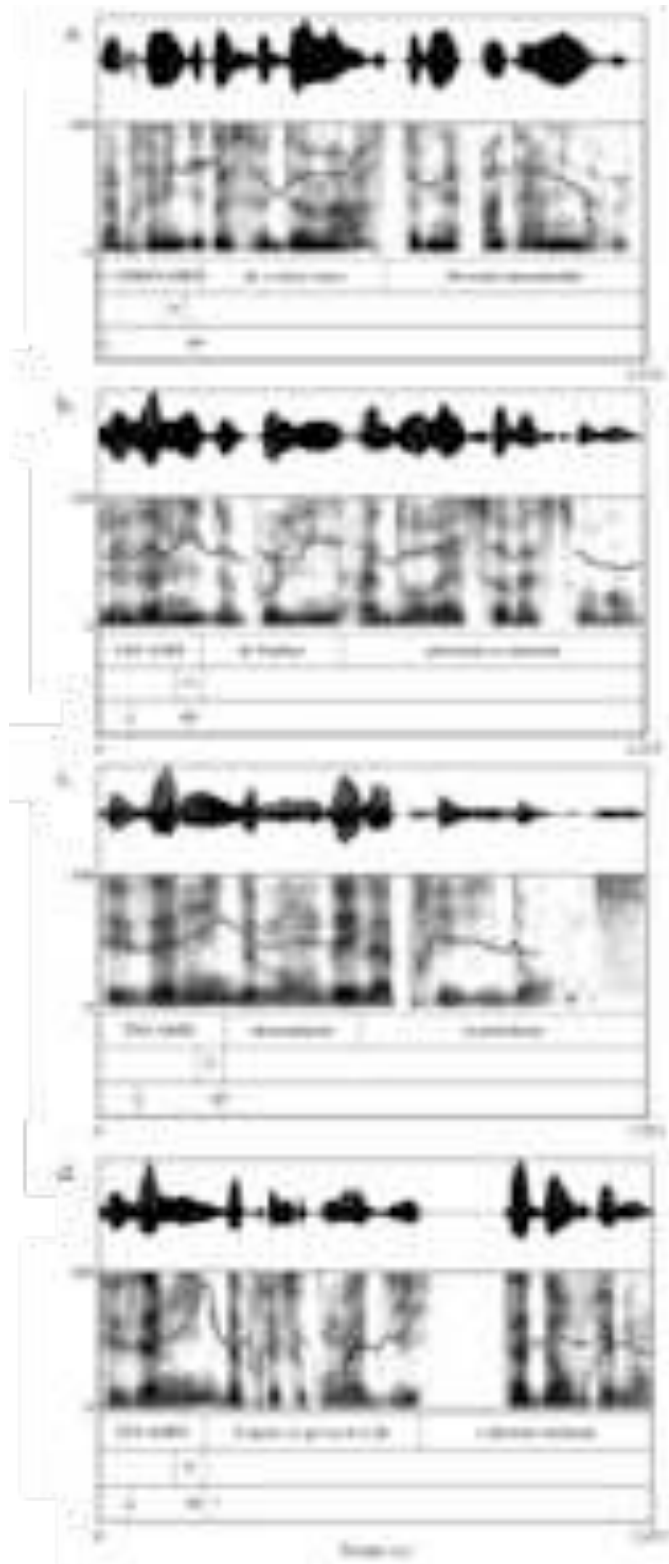


Figure 14. Étude 1 : Exemple de phrases du corpus.
Ondes sonores, spectrogrammes et courbes de fo associés à quatre différents énoncés construits pour les besoins du corpus et illustrant les quatre types de frontières prosodiques: position interne à l'AP (9a), position en finale d'AP (9b), position en finale d'AP dont la frontière droite est associée à une rupture syntaxique majeure (AP/XP, 9c) et position finale d'IP (9d).

Les consonnes qui précèdent et qui suivent les voyelles cibles étaient toujours des consonnes voisées dans le but de faciliter l'analyse de la *f*o. Dans la condition *syllabe interne à l'AP*, les voyelles cibles étaient toujours contenues dans l'avant dernière syllabe ouverte d'un mot trisyllabique ou quadrisyllabique (en fonction du déterminant utilisé : *l'* versus *le* ou *la*) du premier AP de l'énoncé. Dans la condition *syllabe finale d'AP*, les syllabes cibles étaient contenues dans la dernière syllabe ouverte d'APs composés de trois syllabes dont les frontières droites n'étaient pas associées avec des frontières de constituants syntaxiques majeurs (frontière entre un syntagme nominal, SN et un syntagme verbal, SV). Dans la condition *syllabe finale d'AP/XP*, les syllabes cibles étaient contenues dans la dernière syllabe ouverte d'APs composés de trois syllabes dont les frontières droites étaient associées avec des frontières de constituants syntaxiques majeurs. Dans la condition *syllabe finale d'IP*, les syllabes cibles étaient associées à une rupture entre une clause principale et une parenthétique. Ce type de rupture est fortement corrélé avec une frontière prosodique majeure (Selkirk, 1995 ; Delais-Roussarie & Post 2008 ; *inter alia*).

Quarante distracteurs ont été ajoutés aux quarante phrases expérimentales. Les phrases « distractrices » ont été choisies au hasard et ne présentaient aucune structure prosodique ou syntaxique particulière. Quatre listes expérimentales ont été construites de manière à ce que les participants ne puissent pas voir apparaître deux fois de suite le même mot cible. Chaque liste contenait toutes les conditions prosodiques pour un total de 40 phrases expérimentales et quarante phrases distractrices. Les phrases distractrices étaient les mêmes pour chaque liste expérimentale. Tous les participants ont été soumis à toutes les listes dans un ordre randomisé pour chaque sujet. À l'intérieur de chaque liste les énoncés étaient également randomisés.

2.2. Participants et procédure expérimentale

Six participants (trois hommes et trois femmes) ont lu les quatre listes de phrases quatre fois à vitesse d'élocution normale et quatre fois à vitesse d'élocution rapide.

Les participants étaient âgés de 17 à 29 ans. Ils étaient nés dans diverses régions de France et étaient tous titulaires du Baccalauréat. Nous les avons installés dans la chambre sourde du Laboratoire Parole et Langage, confortablement assis face à un écran d'ordinateur. Ils étaient équipés d'un micro casque et d'écouteurs. Nous leur avons demandé de lire à voix haute des phrases, d'abord avec une vitesse d'élocution normale, puis avec une vitesse d'élocution rapide. Les consignes exactes présentées aux participants sur l'écran de l'ordinateur sont données en annexe. Les phrases étaient présentées une par une sur l'écran. Elles étaient orthographiées de manière classique en incluant des signes de ponctuation et des majuscules. Une fois la phrase lue, les participants devaient cliquer sur un bouton du clavier d'ordinateur afin de faire apparaître la phrase suivante à l'écran. La

procédure expérimentale a été contrôlée grâce au logiciel DMDX (Forster & Forster, 2000). Cinq phrases « test » ont été présentées au sujet au début de l'expérience afin de procéder aux divers réglages et de vérifier que la consigne était bien comprise. Ces phrases tests ont également été présentées aux participants au début de la phase de lecture rapide. L'expérience durait environ quarante cinq minutes et se déroulait en deux phases (une phase de lecture normale et une phase de lecture rapide) avec 5 à 10 minutes de pause entre les deux phases.

2.3. Recueil et analyse des données

Les participants ont été enregistrés dans la chambre sourde du Laboratoire Parole et Langage grâce à une station d'enregistrement ZOOM HD16 en format wav 16 bit/44,1 kHz. Les fichiers sons ont été segmentés manuellement et chaque énoncé a été sauvé séparément. Chaque fichier wav a été segmenté en phonèmes à l'aide de l'aligneur Easy align (Goldman, 2007). La segmentation des syllabes et des voyelles cibles a été vérifiée manuellement grâce à l'inspection des ondes sonores et des spectrogrammes pour tous les fichiers. Les courbes de f_0 et les spectrogrammes ont été générés en utilisant PRAAT (Boersma & Weenink, 2009). Les voyelles et les syllabes cibles ont été annotées automatiquement grâce à des scripts PRAAT puis vérifiées manuellement une par une. La durée segmentale des voyelles et des syllabes cibles ainsi que la durée des énoncés ont été extraites grâce à l'utilisation de script PRAAT qui sont consultables en ligne à l'adresse suivante: <http://amandine.michelas.pagesperso-orange.fr/Praat.htm>. Les valeurs maximales de f_0 associées aux voyelles cibles ont également été automatiquement annotées puis extraites grâce à l'utilisation de scripts. L'intonation de chaque énoncé a été annotée par moi-même et vérifiée par un deuxième annotateur spécialisé en prosodie. En cas de désaccord, les deux transcripateurs ont réexaminés les fichiers et sont parvenus à un accord. L'annotation de l'intonation a été réalisée en suivant le modèle de l'intonation du français proposé par Jun & Fougeron (1995, 2000, 2002). Les cibles tonales présentes sur les syntagmes cibles ont été définies en suivant la terminologie utilisée par Welby (Welby, 2006 ; elle même basée sur le modèle proposé par Jun & Fougeron) comme suit:

H final: Un pic de f_0 associé à la dernière syllabe de l'AP (H*).

H initial: Un pic de f_0 associé à une des deux premières syllabes de l'AP (à l'exclusion de la deuxième syllabe d'un mot à contenu composé de deux syllabes, Hi).

L final: Un coude de f_0 , ou en l'absence d'un coude, un minimum de f_0 associé à une des deux dernières syllabes de l'AP (à l'exclusion de la première syllabe d'un mot à contenu de deux syllabes).

L initial: Un coude de *f*₀ ou en l'absence d'un coude, un minimum de *f*₀, situé à proximité du début de l'AP ou du premier mot à contenu de l'AP.

Cette caractérisation tonale nous a permis de classifier les syntagmes accentuels cibles en six types de patrons mélodiques selon les réalisations de surface de l'AP proposées par Jun & Fougeron (voir figure 9).

Dans certains cas, il a été difficile de différencier les patrons /LHiLH*/ des patrons /LHiH*/ en raison d'un coude plus ou moins important entre les deux tons hauts. La figure 15 illustre cette difficulté que nous avons rencontrée. Trois versions de la séquence *Le Calvados* extraite de l'énoncé *Le Calvados de Normandie restera un alcool fort*, prononcée par les trois locuteurs masculins de notre corpus, sont illustrées dans cette figure. Dans la figure 15a nous pouvons clairement identifier le patron /LHiLH*/ avec un net coude de *f*₀ entre les deux cibles H. Dans la version 15b, le patron mélodique /LHiH*/ avec une absence de ton L entre les H initial et final est clairement reconnaissable. Dans la figure 15c le minimum de *f*₀ entre les deux cibles H est beaucoup moins clair que dans la figure 15a. Dans ce type de cas, nous assumons le fait que nous sommes face au patron mélodique /LHiH*/ (comme dans la figure 15b).

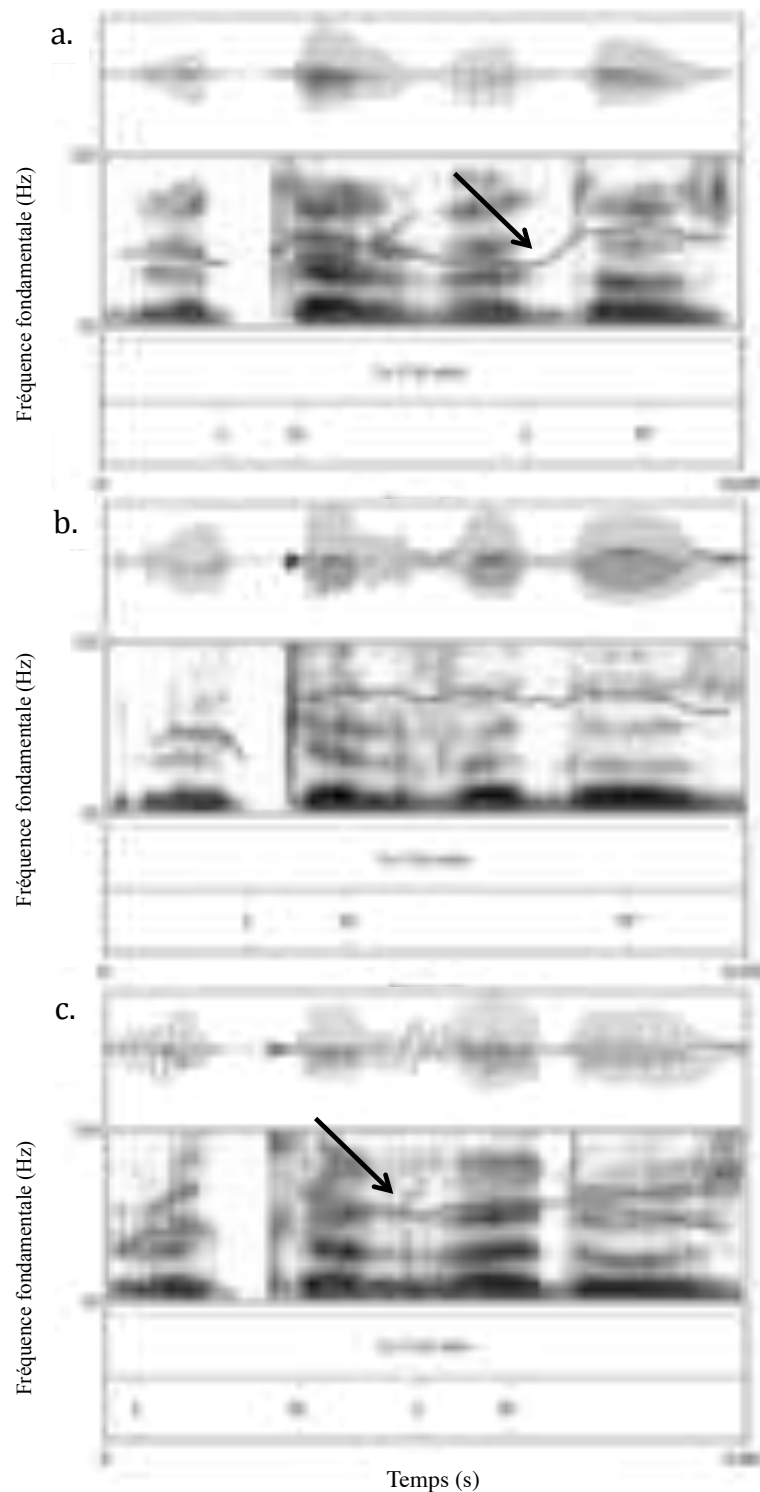


Figure 15. Étude 1 : Exemple de patrons mélodiques observés. Ondes sonores, spectrogrammes et courbes de fréquence fondamentale associés à la séquence "Le calvados" extraite de la phrase "Le Calvados de Normandie restera un alcool fort" prononcée par trois de nos locuteurs masculins. La séquence est associée au patron mélodique /LHiLH*/ dans la figure 15a avec une cible L finale clairement identifiable, au patron /LHiH*/ sans cible L finale dans la figure 15b et au patron /LHiLH*/ ou /LHiH*/ avec une cible L finale plus difficilement identifiable dans la figure 15c.

2.4. Analyses statistiques

Les analyses statistiques ont été effectuées grâce à l'environnement de travail R (R Development Core Team, 2008). Ces analyses consistent en une série de modèles de régressions linéaires multiples à effets mixtes (*Linear Mixed effects Models, LMM*) lorsque la variable étudiée était continue (par exemple la durée de la voyelle) et de modèles de régressions non linéaires (ou logit) à effets mixtes (*Mixed Logit Model, MLM*) lorsque la variable était binomiale (par exemple la présence ou l'absence du patron mélodique /LH*/). Pour cela nous avons employé le package *lme4* (Bates & Maechler, 2011).

Les modèles à effets mixtes ou plus simplement modèles mixtes incorporent des facteurs à effets fixes et des facteurs à effets aléatoires. Les effets fixes sont des paramètres associés à une population entière ou avec certains niveaux de facteurs expérimentaux reproductibles tandis que les effets aléatoires sont des paramètres considérés comme échantillonnés au hasard dans une population plus large que les données elles-mêmes. Toutefois la distinction entre effet fixe et effet aléatoire est en fait moins stricte car elle dépend aussi du type de traitement que veut effectuer l'expérimentateur sur l'effet considéré. Un facteur est dit «fixe» si l'objectif de l'expérimentateur est de comparer les niveaux de l'effet sur la variable dépendante et il est dit à effet «aléatoire» si les niveaux de l'effet ne sont pas sous le contrôle de l'expérimentateur (Pinheiro & Bates 2000, Quené & van den Berg, 2004).

Les analyses de régressions multiples à effets mixtes offrent plus d'avantage que les ANOVAs. Rappelons que l'ANOVA (*Analysis of Variance*) est un cas de régression particulier qui considère que la somme totale des carrés des écarts à la moyenne de l'ANOVA correspond à celle de la régression avec les facteurs catégoriels. Deux de ces avantages nous semblent particulièrement intéressants à souligner au regard du type de données que nous avons à traiter.

Premièrement, l'utilisation d'ANOVA sur des plans d'expérience déséquilibrés est déconseillée car cela peut affecter l'estimation des paramètres. Les régressions multiples à effets mixtes présentent l'avantage d'être plus robuste à ce type de plan expérimental. Certains des plans expérimentaux qui seront présentés dans cette thèse sont déséquilibrés. Le déséquilibre vient du fait que nous n'avons pas le même nombre de mesures par modalité. Par exemple, plusieurs de nos expériences rapportent des phrases lues. Certaines fois, nous nous sommes trouvée dans l'obligation de supprimer certaines réalisations de phrases en raison de divers phénomènes tels qu'un éternuement lors de la réalisation de la phrase, un raclement de gorge ou autre.

De plus, dans cette thèse nous aurons à évaluer les effets de facteurs sur une variable dépendante binomiale (c'est-à-dire une variable qui n'est pas graduelle comme le sont par exemple des temps de réaction mais de type

binaire comme un choix entre deux réponses possibles A ou B). Dans ce cas précis, l'utilisation d'ANOVA est déconseillée. En effet, dans le cas d'une variable dépendante binomiale où la variance dépend de la moyenne, le test statistique ANOVA n'est pas approprié. Ce type de variable relève d'un modèle de régression logit (*Mixed Logit Model*, voir par exemple Pinheiro & Bates, 2000).

Pour ces deux raisons, l'utilisation de modèles à régressions multiples à effets mixtes nous semblent donc mieux correspondre à nos données. Cependant, si ce type de tests statistiques présente de nombreux avantages, la question du calcul des degrés de liberté pour ces modèles est encore sujette à controverse dans la littérature. Pour pouvoir fournir une valeur de p fiable, nous avons eu recours à une alternative largement utilisée. Cette alternative est celle d'utiliser les p -values calculées par simulation numérique selon la procédure *Monte Carlo Markov Chain* (pMCMC). Nous tenons également à préciser que nos modèles statistiques sont construits incluent un grand nombre de données (autour de mille mesures pour chaque modèle). Ce problème d'estimation des degrés de liberté est donc moins crucial relativement à la fiabilité des p -values. Les pMCMC obtenues sont d'ailleurs très proches des p -values obtenues classiquement. Conformément ce qui est préféré dans la littérature, nous avons choisi un $\alpha < 0.05$.

3. RÉSULTATS

3.1. Manipulation de la vitesse d'élocution

Dans le but de vérifier si tous les participants ont significativement augmenté leur vitesse d'élocution à vitesse d'élocution rapide par rapport à la vitesse d'élocution normale, nous avons mesuré la vitesse d'élocution pour chaque énoncé en nombre de syllabes par secondes (syl/sec). Nous avons automatiquement extrait la durée de chaque énoncé exprimée en secondes. Nous avons ensuite compté le nombre de syllabes contenues dans chaque énoncé en fonction de la manière dont l'énoncé a été produit par le locuteur. En effet, les énoncés pouvaient contenir des schwas finaux pouvant être élidés ce qui a donné lieu à des variations du nombre de syllabes contenues dans les énoncés en fonction des locuteurs. En fonction de ce nombre ajusté de syllabes par énoncé et par locuteur nous avons calculé la vitesse d'élocution pour chaque énoncé en nombre de syllabes par seconde. Dans le but de voir si le deuxième syntagme accentuel de l'énoncé (APs cible) a été affecté de la même manière que les énoncés entiers par la manipulation de la vitesse d'élocution, nous avons également calculé la vitesse d'élocution en syllabe par seconde pour les APs cibles. Les résultats reportés dans le tableau 2 montrent que les six locuteurs ont augmenté leur vitesse d'élocution à débit rapide aussi bien pour les énoncés en entier que pour les

APs cibles. Pour les énoncés, le débit de parole à vitesse normale est compris entre 5,4 syl/sec et 6,9 syl/sec selon les locuteurs. Ces valeurs sont comparables à celles observées dans de précédentes études menées à la fois sur de la parole lue (Fougeron & Jun, 1998; Welby & Loevenbruck, 2006) et de la parole spontanée (Malécot, Johnson & Kizziar, 1972; Grosjean & Deschamp, 1975) pour le français. L'augmentation de la vitesse d'élocution du débit rapide par rapport au débit normal est donnée en pourcentage. Ce pourcentage d'augmentation va de 30,6% pour le locuteur MR à 47,6% pour le locuteur MS.

	Hommes					
	FP		VA		XL	
	Énoncés	APs cibles	Énoncés	APs cibles	Énoncés	APs cibles
Vitesse d'élocution normale	6,56	7,2	6,15	7,48	6,91	7,42
Vitesse d'élocution rapide	8,6	9,5	8,68	10,41	9,09	10,01
Pourcentage d'augmentation	30,6	31,9	41,1	39,1	31,5	36,2
	Femmes					
	MR		MS		MT	
	Énoncés	APs cibles	Énoncés	APs cibles	Énoncés	APs cibles
Vitesse d'élocution normale	5,46	5,4	5,44	6,07	5,556	6,04
Vitesse d'élocution rapide	7,8	7,5	8,03	7,7	7,74	8,45
Pourcentage d'augmentation	42,85	38,8	26,8	47,6	26,8	39,2

Tableau 2. Étude 1 : Caractéristiques de la vitesse d'élocution.

La vitesse d'élocution est exprimée en nombre de syllabes par seconde pour les énoncés et les syntagmes accentuels cibles pour les six locuteurs du corpus (FP, MR, MS, MT, VA et XL). L'augmentation du débit par rapport à la vitesse normale est donné en pourcentage.

Deux modèles linéaires à effets mixtes avec la vitesse d'élocution en syl/s comme variable dépendante ont été conduits à partir de nos données, un pour les énoncés et un pour les APs cibles. Un total de 1920 mesures a été incluse pour chaque modèle. Pour les deux modèles, nous avons incluse la manipulation de vitesse (vitesse normale vs. vitesse rapide) comme effet fixe. Nous avons observé un effet significatif du genre des locuteurs (homme vs. femme) sur la vitesse d'élocution. Nous avons donc inséré le genre en tant qu'effet fixe mais sans considérer l'interaction entre cet effet et les autres effets fixes du modèle car cela ne permettait pas une amélioration de celui-ci (Bates, 2000). Les locuteurs ont été inclus en tant qu'effet aléatoire. Le modèle obtenu fait apparaître les mêmes effets pour les énoncés que pour les APs cibles. Dans les deux cas, la manipulation de vitesse était significative (pour les énoncés: $\beta = -2.143$, $se = 0.22$, $t = -9.7$, $pMCMC < 0.001$; pour les APs cibles : $\beta = -2.211$, $se = 0.066$, $t = -33.09$, $pMCMC < 0.001$). Tous les locuteurs ont significativement augmenté leur vitesse d'élocution à débit rapide. L'effet du genre du locuteur était également significatif faisant apparaître le fait que les locutrices (MR, MS, MT) avaient une vitesse d'élocution beaucoup plus lente que les locuteurs (FP, VA and XL) (pour les énoncés: $\beta = -1.16$, $se = 0.252$, $t =$

-4.49; $pMCMC < 0.01$; pour les APs cibles: $\beta = -1.8$, $se = 0.25$, $t = -7.12$, $pMCMC < 0.01$).

3.2. Phénomènes de restructuration prosodique observés à vitesse d'élocution rapide

Un petit nombre de modification de la structure prosodique à vitesse d'élocution rapide a été observé. En effet, pour seulement 2,3 % des énoncés produits (45 énoncés) nous avons observé des restructurations prosodiques, c'est-à-dire des phénomènes d'effacement de certaines frontières prosodiques ou de remplacement d'une frontière par une frontière de niveau inférieur dans la hiérarchie prosodique (cf. Fougeron & Jun, 1998). Dans 42 énoncés le nombre d'AP a été réduit à vitesse d'élocution rapide de manière à ce qu'un plus grand nombre de syllabes soit groupé au sein du même syntagme accentuel. De plus dans trois énoncés nous avons remarqué que la frontière d'IP a été remplacée par une frontière d'AP c'est-à-dire par une frontière de niveau inférieur dans la hiérarchie prosodique. Un exemple d'effacement de frontière d'AP et un exemple de remplacement d'une frontière d'IP par une frontière d'AP sont montrés dans les figures 16a et 16b. Ces 45 énoncés ont été exclus des analyses de durée et de *fo* qui suivent (cf. sections 3.4 et 3.5).

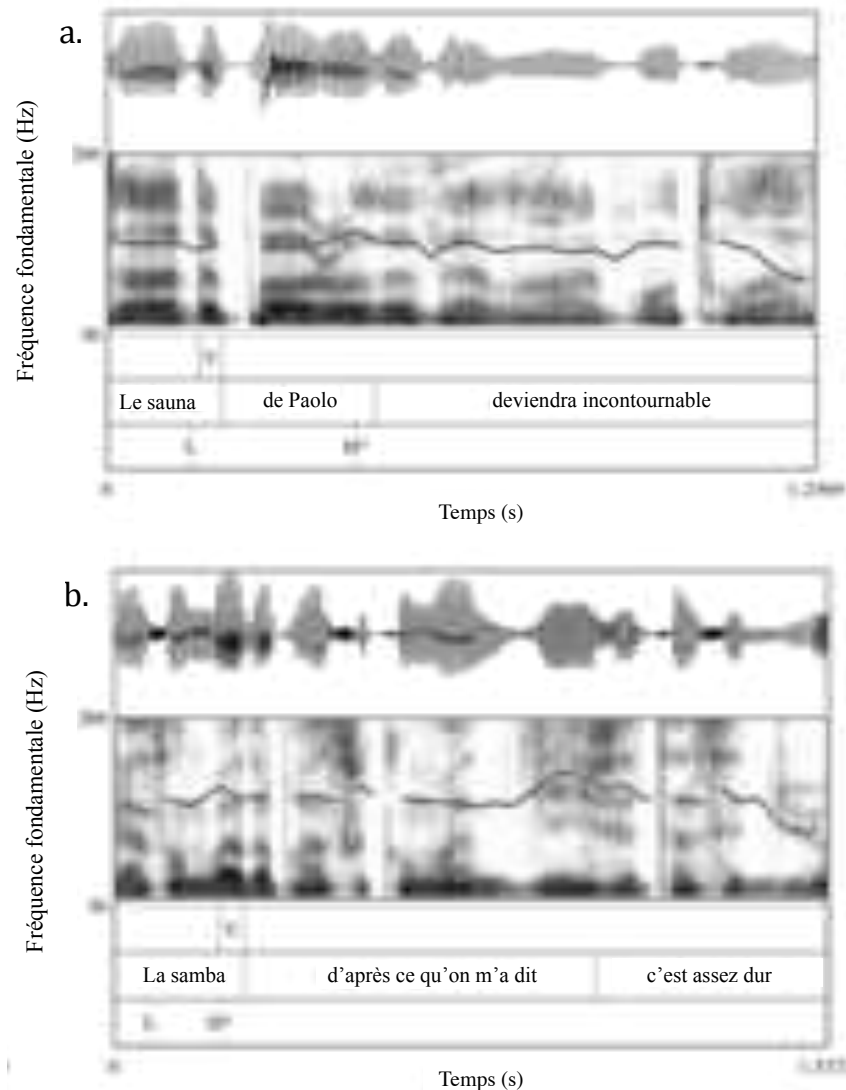


Figure 16. Étude 1 : Exemple de modifications de la structure prosodique à vitesse d'élocution rapide.

Illustration d'un effacement de frontière d'AP sur la syllabe /na/ dans la phrase "Le sauna de Paolo deviendra incontournable" (16a) et d'un remplacement de frontière d'IP par une frontière de niveau inférieur dans la hiérarchie prosodique (IP⇒AP) sur la syllabe /ba/ dans la phrase "La samba d'après ce qu'on m'a dit c'est assez dur" (16b).

3.3. Distribution des patrons mélodiques

Le pourcentage des différents patrons mélodiques observés pour les conditions *interne à l'AP*, *finale d'AP*, *finale d'AP/XP* en fonction du débit de parole (normale vs. rapide), du nombre de syllabe à l'intérieur de l'AP (3 syllabes vs. 4 syllabes) et du genre des locuteurs (homme vs. femme) est représenté dans le tableau 3.

APs en position non-finale à l'intérieur de l'IP

Vitesse d'élocution normale													
	LHiLH*		LH*		LLH*		LHiH*		HiLH*		LHiL*		
	M	F	M	F	M	F	M	F	M	F	M	F	
3 syllabes	16	46	78	47	5	3	2	2	0	0	0	0	
4 syllabes	29	60	38	21	2	2	18	6	5	8	8	4	

Vitesse d'élocution rapide													
	LHiLH*		LH*		LLH*		LHiH*		HiLH*		LHiL*		
	M	F	M	F	M	F	M	F	M	F	M	F	
3 syllabes AP	2	28	93	67	2	0	2	2	2	0	2	0	
4 syllabes AP	4	30	82	55	0	4	8	6	3	5	2	0	

Tableau 3. Étude 1 : Pourcentage de réalisation des différents patrons mélodiques pour les conditions AP-interne, AP-finale et AP/XP.

Les pourcentages sont exprimés en fonction du nombre de syllabes à l'intérieur de l'AP (3 syllabes/4 syllabes), du genre (homme/femme) et de la vitesse d'élocution (normale/rapide).

Notons que les patrons mélodiques observés pour la condition *finale d'IP* ont été analysés à part pour deux raisons essentielles: (i) nous ne disposons que d'APs cibles composés de 3 syllabes contrairement aux APs en position non finale d'IP pour lesquels nous disposons de deux niveaux (APs de 3 syllabes et APs de 4 syllabes) et (ii) car la présence de la frontière d'IP peut potentiellement influencer le patron mélodique de l'AP. Les patrons observés pour les APs en final d'IP sont représentés dans le tableau 4.

APs en position finale à l'intérieur de l'IP

Vitesse d'élocution normale												
												
	M	F	M	F	M	F	M	F	M	F	M	F
3 syllabes	31	58	47	23	14	14	1	0	8	6	0	0

Vitesse d'élocution rapide												
												
	M	F	M	F	M	F	M	F	M	F	M	F
3 syllabes	9	33	85	55	4	8	1	2	0	2	0	0

Tableau 4. Étude 1 : Pourcentage de réalisation des différents patrons mélodiques pour la condition IP-finale.

Les pourcentages sont exprimés en fonction du genre (homme/femme) et de la vitesse d'élocution (normale/rapide).

A vitesse d'élocution normale, le patron mélodique /LH*/ était le patron mélodique le plus communément observé pour les APs en position non-finale d'IP. Chez les locuteurs masculins, 78% des patrons mélodiques

observés pour les APs contenant 3 syllabes étaient de type /LH*/ et 38% des patrons mélodiques observés étaient de même type pour les APs composés de 4 syllabes. Chez les femmes, le patron mélodique de type /LH*/ est resté le patron mélodique le plus communément observé pour les APs contenant 3 syllabes bien que le pourcentage observé soit presque équivalent au patron de type /LHiLH*/ (47% de patron de type /LH*/ contre 46% de patron de type /LHiLH*). Pour les APs contenant 4 syllabes, le patron mélodique le plus observé chez les femmes n'était plus le patron mélodique /LH*/ mais le patron /LHiLH*/ (60% de patron mélodique de type /LHiLH*/ contre 21% de patron de type /LH*/). Le pourcentage de patrons mélodiques de type /LHiLH*/ a augmenté avec le nombre de syllabes à l'intérieur de l'AP. En effet, nous avons observé, une augmentation du pourcentage de patrons mélodiques de type /LHiLH*/ pour les APs contenant 4 syllabes et cette augmentation est valable aussi bien pour les hommes que pour les femmes. Ceci va dans le sens d'études précédentes qui ont montré que le nombre de syllabe à l'intérieur de l'AP était un excellent prédicteur de la réalisation ou non d'une montée initiale (LHi). Par exemple Welby (2006) a observé un nombre plus élevé de patrons mélodiques de type /LHiLH*/ pour des APs contenant 4 syllabes plutôt que pour des APs contenant un plus petit nombre de syllabes. La figure 17 montre la réalisation de l'AP *Les dromadaires* prononcé à vitesse d'élocution normale avec un patron mélodique de type /LH*/ (17a) et avec un patron mélodique de type /LHiLH*/ (17b). Les deux réalisations ont été effectuées par la même locutrice.

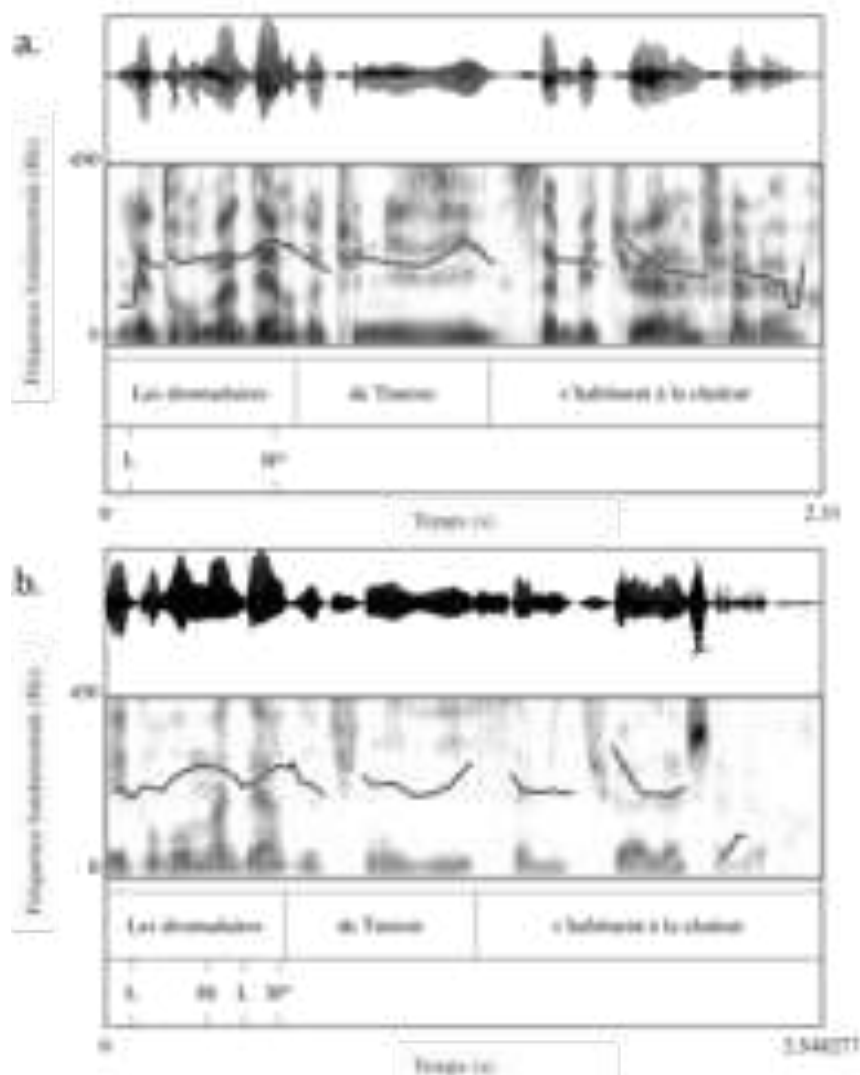


Figure 17. Étude 1 : Montée initiale et montée finale de *fo*.
Le premier AP « les dromadaires » extrait de la phrase « Les dromadaires de Tunisie s'habituent à la chaleur » est réalisé avec une montée finale /LH*/ (17a) ou avec une montée initiale additionnelle /LHi/ (17b).

Le nombre de patrons mélodiques de type /LHiLH*/ varie également en fonction du sexe du locuteur. Nous observons en effet plus de patrons de type /LHiLH*/ chez les femmes que chez les hommes (à vitesse d'élocution normale: 46% des patrons mélodiques observés sont de type /LHiLH*/ chez les femmes contre 16% chez les hommes pour les APs contenant 3 syllabes ; 60% des patrons mélodiques observés sont de type /LHiLH*/ chez les femmes contre 29% chez les hommes pour les APs contenant 4 syllabes).

Concernant les APs en position non finale d'IP pour lesquels nous ne disposons que d'APs composés de 3 syllabes, le patron mélodique /LH*/ est le plus fréquemment observé chez les hommes (47% des patrons mélodiques observés) alors que le patron mélodique /LHiLH*/ est le patron

le plus communément réalisé par les femmes (58% des patrons mélodiques observés).

De manière peu surprenante, à vitesse d'élocution rapide, le patron mélodique /LH*/ devient le patron mélodique le plus fréquent, à la fois pour les hommes et pour les femmes et ce indépendamment du nombre de syllabes à l'intérieur de l'AP ainsi que de la position de l'AP l'intérieur de l'IP (position finale/position non finale). La vitesse d'élocution influe également sur la réalisation des patrons mélodiques avec une forte augmentation du patron mélodique de type /LH*/ à vitesse d'élocution rapide à la fois chez les hommes et chez les femmes (le pourcentage de patron mélodique de type /LH*/ passe de 47 à 67 % pour les APs contenant 3APs et de 21% à 55% pour les APs contenant 4 syllabes chez les femmes; le pourcentage de patron mélodique de type /LH*/ passe de 78% à 93 % pour les APs content 3 syllabes et de 38% à 82% pour les APs contenant 4 syllabes chez les hommes). Un exemple du patron mélodique /LH*/ prononcé par un locuteur masculin est montré à vitesse normale (18a) et à vitesse rapide (18c). Les figures 18b et 18d illustrent le même patron mélodique prononcé par une femme respectivement à vitesse d'élocution normale et à vitesse d'élocution rapide.

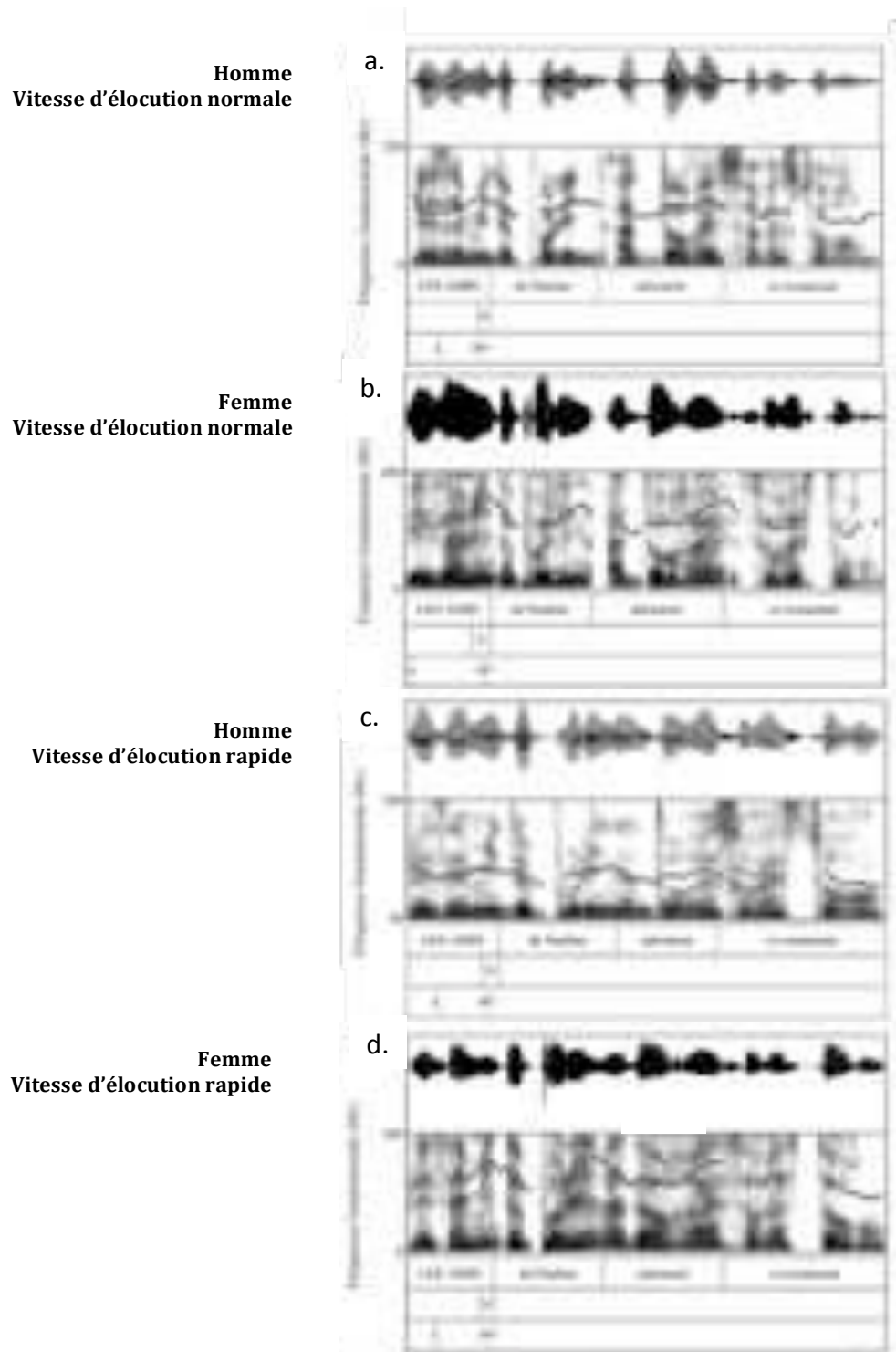


Figure 18. Etude 1 : Exemple d'un patron mélodique de type /LH*/.
 Ondes sonores, courbes de fréquence fondamentale et spectrogrammes obtenus à partir de l'énoncé « Les amis de Pauline adoraient ce restaurant » prononcé par un locuteur à vitesse normale (18a) à vitesse rapide (18c) et locutrice à vitesse normale (18b) et à vitesse rapide (18d). La premier AP « Les amis » est réalisé avec un patron mélodique de type /LH*/ dans les quatre occurrences.

Nous avons également observé des patrons mélodiques de type /LLH*/, /LHiH*/, /HiH*/ et /HiL*/. Chacun de ces patrons mélodiques ont été produits un petit nombre de fois comme en témoignent les pourcentages d'apparition de ces patrons allant de 0 à 5% seulement.

Deux modèles mixtes de régressions logit (*Mixed Logit Models, MLM*) ont été utilisés dans le but d'examiner les facteurs qui influencent l'apparition du patron mélodique /LH*/. Ces deux modèles ont été générés séparément pour les APs en position non-finale au sein de l'IP (condition interne à l'AP, finale d'AP et finale d'AP/XP) et en position finale à l'intérieur de l'IP (condition finale d'IP). Ce type de modèle est un modèle statistique qui permet d'examiner des choix discrets. En effet, la variable dépendante est une variable binaire (réalisation du patron mélodique /LH*/ ou réalisation d'un autre patron mélodique). Nous renvoyons le lecteur à la section 2.4 de ce chapitre pour plus de détails sur l'avantage de ce type de modèle. Pour les deux modèles la variable dépendante était la réalisation du patron mélodique /LH*/ qui était le patron mélodique le plus communément observé dans nos données. Les données ont été codées avec la valeur 1 attribuée lorsque le patron mélodique réalisé était de type /LH*/ et 0 lorsque le patron mélodique réalisé était de tout autre type. Dans le premier modèle MLM, examinant les facteurs influençant la réalisation du patron mélodique /LH*/ pour les APs en position non finale d'IP, les effets fixes étaient la vitesse d'élocution (vitesse normale/vitesse rapide), le genre des locuteurs (homme/femme) et le nombre de syllabes contenues dans l'AP (3 syllabes/4 syllabes). Pour le deuxième modèle MLM, examinant l'apparition du patron mélodique /LH*/ à partir des APs en position finale d'IP, les facteurs fixes étaient la vitesse d'élocution (vitesse normale/vitesse rapide) et le genre du locuteur (homme/femme). Les deux modèles incluaient les locuteurs et les énoncés en tant qu'effets aléatoires. Les probabilités d'apparition du patron mélodique /LH*/ prédites par le modèle sont échelonnées sur une échelle allant de 0 à 1 et sont illustrées dans la figure 19 pour les APs en position non finale d'IP et dans la figure 20 pour les APs en position finale d'IP. Le premier modèle incluait 1440 mesures alors que le second modèle incluait 440 mesures.

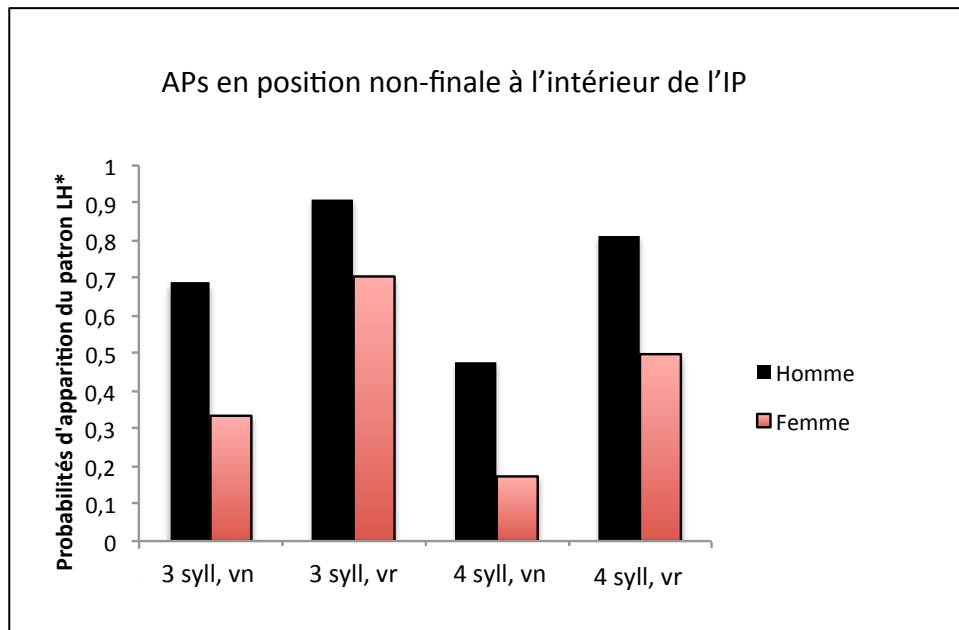


Figure 19. Étude 1 : Probabilités d'apparition du patron mélodique de type /LH*/ estimées par le modèle MLM pour les APs situés en position non finale d'IP. Les probabilités sont échelonnées sur une échelle allant de 0 à 1 en fonction du nombre de syllabes à l'intérieur de l'AP (3 syllabes/4 syllabes), de la vitesse d'élocution (vn/vr) et du genre du locuteur (homme/femme).

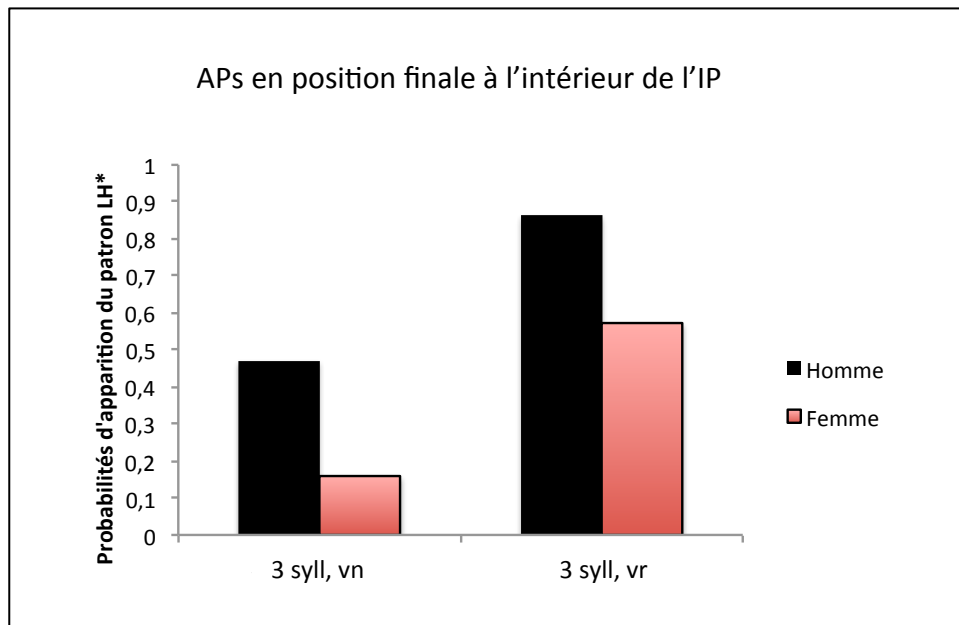


Figure 20. Étude 1 : Probabilités d'apparition du patron mélodique de type /LH*/ estimées par le modèle MLM pour les APs situés en position finale d'IP. Les probabilités sont échelonnées sur une échelle allant de 0 à 1 en fonction du nombre de syllabes à l'intérieur de l'AP (3 syllabes/4 syllabes), de la vitesse d'élocution (vn/vr) et du genre du locuteur (homme/femme).

Les résultats des deux modèles ont montré que tous les facteurs fixes avaient un effet significatif sur la variable dépendante. Pour les APs en position non finale d'IP, le modèle MLM a prédit un plus grand nombre de patrons mélodiques de type /LH*/ à vitesse d'élocution rapide plutôt qu'à vitesse d'élocution normale aussi bien pour les locuteurs masculins ($\beta=1.44$, $se=0.133$, $z=10.87$, $p<0.0001$) que pour les locuteurs féminins ($\beta=1.47$, $se=0.38$, $z=3.876$, $p<0.0001$) et ce indépendamment du nombre de syllabes à l'intérieur de l'AP. Le même effet de la vitesse d'élocution a été observé pour les APs en position finale d'IP (pour les hommes : $\beta=1.956$, $se=0.235$, $z=8.329$, $p<0.0001$; pour les femmes : $\beta=1.956$, $se=0.235$, $z=8.329$, $p<0.0001$). De plus, pour les APs en position non finale d'IP, un plus grand nombre de patrons mélodiques de type /LH*/ a été prédit par le modèle pour les APs contenant 4 syllabes que pour les APs contenant 3 APs indépendamment du sexe du locuteur et de la vitesse d'élocution (pour les femmes à vitesse normale : $\beta=-1.223$, $se=0.29$, $z=-4.23$, $p<0.0001$; pour les femmes à vitesse rapide : $\beta=-1.228$, $se=0.228$, $z=-4.142$, $p<0.0001$; pour les hommes à vitesse normale : $\beta=-1.229$, $se=0.2952$, $z=-4.143$, $p<0.0001$; pour les hommes à vitesse rapide : $\beta=-1.228$, $se=0.2952$, $z=-4.12$, $p<0.0001$). Enfin, l'influence du genre du locuteur sur le nombre de patrons mélodiques de type /LH*/ était significatif avec un plus grand nombre de patrons de type /LH*/ prédits pour les hommes que pour les femmes indépendamment de la position de l'AP à l'intérieur de l'IP et de la vitesse d'élocution (pour les APs en position non finale d'IP : $\beta=1.44752$, $se=0.4331$, $z=3.571$, $p<0.01$; pour les APs en position finale d'IP : $\beta=1.5546$, $se=0.6646$, $z=2.339$, $p<0.01$). Les probabilités estimées par le modèle sont représentées sur une échelle allant de 0 à 1. Pour les APs en position non finale d'IP, le modèle incluait 1920 mesures alors que le modèle généré pour les APs en position finale d'IP incluait 440 mesures. Pour les deux modèles, aucune des interactions entre les effets fixes n'était significative (les p values étaient toutes supérieures à 0.05).

Jusqu'ici, les analyses que nous avons menées nous ont permis de mettre en évidence le fait que la vitesse d'élocution rapide n'a pas induit de restructuration significative du découpage en constituants prosodiques. En effet, seuls 2,3% des énoncés produits par les locuteurs contenaient des phénomènes d'effacement de frontière ou de remplacement d'une frontière d'IP par une frontière d'AP.

De plus, l'analyse du type de patron mélodique produit par les locuteurs, a montré que la réalisation du patron mélodique de type /LH*/ est conditionnée par la vitesse d'élocution (avec plus de patrons mélodiques de type /LH*/ observés pour les hommes que pour les femmes), le genre des locuteurs (avec plus de patrons mélodiques de ce type observés pour les hommes que pour les femmes) ainsi que la longueur du syntagme (avec plus de patrons de ce type observés pour les APs contenant 3 syllabes que pour les APs contenant 4 syllabes).

3.4. Résultats relatifs à la durée

Dans cette section nous allons étudier la relation existante entre la durée de la voyelle et le niveau de frontière prosodique qui lui est associé. Rappelons que nos voyelles cibles pouvaient être à l'intérieur d'un AP (condition AP-interne), en position finale d'AP (condition AP-finale), en position finale d'AP associé à une rupture entre syntagme nominal et syntagme verbal (condition AP/XP-finale) ou en position finale d'IP (condition IP-finale). L'hypothèse nulle était que l'alignement entre une rupture syntaxique majeure et la frontière d'AP (condition AP/XP-finale) n'affecterait pas la durée de la voyelle. En d'autres termes nous nous sommes attachée à tester si, un seul niveau d'allongement associé à l'AP aller être observé pour le français ou si plus d'un niveau d'allongement serait nécessaire pour rendre compte de l'allongement observé en finale d'AP associé à une rupture syntaxique mineure et celui observé en finale d'AP associé à une rupture syntaxique majeure. Nous avons également testé l'influence de la vitesse d'élocution sur la force de la frontière prosodique. Un modèle de régressions linéaires multiples à effets mixtes (*Linear Mixed effects Models, LMM*) a été employé sur les valeurs logarithmiques des durées avec la type de frontière prosodique (AP-interne, AP-finale, AP/XP, IP-finale) et la vitesse d'élocution (normale/rapide) en tant qu'effet fixe. Etant donné que le genre des locuteurs avait un effet significatif sur la durée des voyelles, nous avons également inclus le genre des locuteurs en tant qu'effet fixe. Les locuteurs et les syllabes ont également été inclus en tant qu'effets aléatoires.

Les résultats montrent que la durée de la voyelle augmentait significativement en fonction de la force de la frontière prosodique quel que soit le genre du locuteur (homme/femme) ou la vitesse d'élocution (normale/rapide) bien que ces effets d'allongement aient été plus marqués à vitesse d'élocution normale plutôt qu'à vitesse d'élocution rapide. Les valeurs de durées des voyelles estimées par le modèle sont illustrées dans la figure 21.

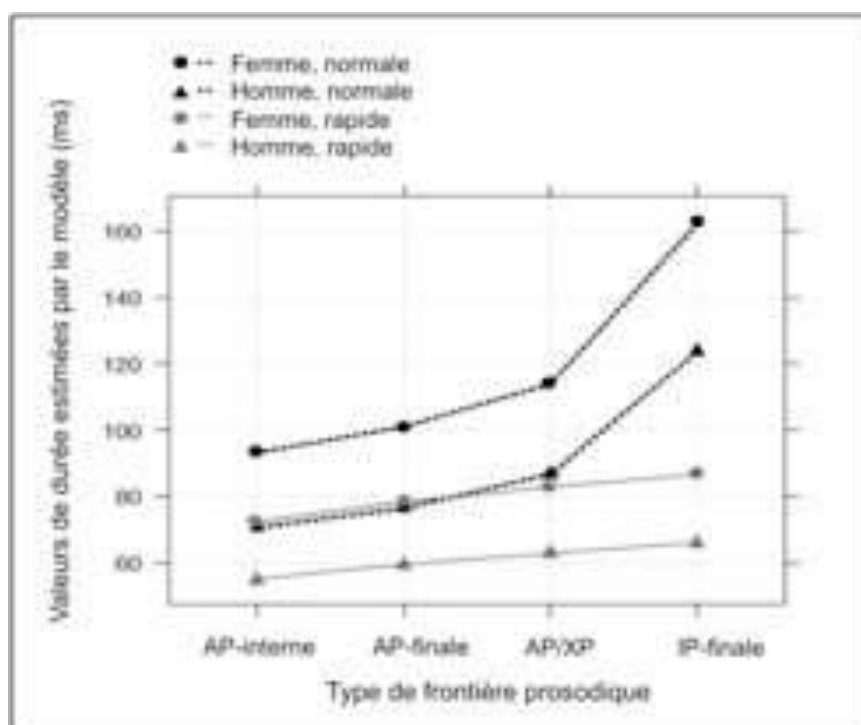


Figure 21. Étude 1 : Valeurs de durées estimées par le modèle LMM.

Les valeurs sont fonction du type de frontière prosodique (AP-interne, AP-finale, AP/XP et IP-finale), de la vitesse d'élocution (normale/rapide) et du genre du locuteur (homme/femme).

A vitesse d'élocution normale et indépendamment du genre des locuteurs (homme/femme) les voyelles situées en finale d'AP étaient significativement plus longues que les voyelles inaccentuées internes à l'AP ($\beta = -0.08$, $se = 0.022$, $t = -3.67$, $pMCMC < 0.0001$). Il est également intéressant de noter que les voyelles en position finale d'AP associées à une rupture syntaxique majeure (condition AP/XP) étaient significativement plus longues que les voyelles en finale d'AP située à l'intérieur du syntagme nominal ($\beta = -0.122$, $se = 0.228$, $t = -5.49$, $pMCMC < 0.0001$). Conformément à ce qui était attendu, les voyelles en finale d'IP étaient significativement plus longues que les voyelles dans les autres conditions (condition AP-interne $\beta = -0.5608$, $se = 0.228$, $t = -25.17$, $pMCMC < 0.0001$; condition AP-finale: $\beta = -0.47913$, $se = 0.228$, $t = -21.50$, $pMCMC < 0.0001$; condition AP/XP: $\beta = -0.3567$, $se = 0.228$, $t = -16.01$, $pMCMC < 0.0001$). Le tableau 5 montre les durées estimées par le modèle pour les différentes conditions en fonction du type de frontière prosodique, de la vitesse d'élocution et du genre des locuteurs. Les paires significativement différentes les unes des autres sont indiquées par un astérisque.

Type de frontière prosodique	Vitesse d'élocution normale		Vitesse d'élocution rapide	
	H	F	H	F
AP-interne	71	93	55	72
AP-finale	77	101	59	78
AP/XP	87	114	63	83
IP-finale	124	163	65	87

Tableau 5. Étude 1 : Durées des voyelles (ms) estimées par le modèle pour les quatre types de frontières prosodiques (AP-interne, AP-finale, AP/XP, IP-finale).

La durée des voyelles est exprimée en fonction de la vitesse d'élocution (normale/rapide) et du genre des locuteurs (homme/femme). Les paires significativement différentes les unes de autres sont indiquées par un astérisque.

Par exemple, pour les locuteurs masculins, les voyelles associées aux frontières d'APs (70.66 ms) étaient significativement plus longues que les voyelles inaccentuées (internes à l'AP) (76.63 ms). De manière similaire, pour les locuteurs féminins, les voyelles en position finale d'AP étaient plus longues (101 ms) que les voyelles inaccentuées (93 ms). Conformément à ce que nous avons observé lors des analyses sur la manipulation de la vitesse d'élocution (voir section 3.1 de ce chapitre) qui montrait une vitesse d'élocution moins rapide chez les femmes que chez les hommes, les résultats montrent que la durée des voyelles était plus longue chez les femmes que chez les hommes ($\beta = -0.2743$, $se = 0.029$, $t = -9.15$, $pMCMC < 0.001$) bien que le même nombre de degrés d'allongements significativement différents ait été observé pour les femmes et pour les hommes et ce quel que soit la vitesse d'élocution. Conformément à ce qui était attendu, l'effet de la vitesse d'élocution était significatif ($\beta = -0.2522$, $se = 0.0222$, $t = -11.32$, $pMCMC < 0.0001$).

A vitesse d'élocution rapide, la durée des voyelles étaient plus similaire de sorte qu'une différence significative de durée ait été observée uniquement pour le plus bas niveau (AP-interne) qui était significativement différent de la condition AP-finale ($\beta = 0.079$, $se = 0.0228$, $t = -3.58$, $pMCMC < 0.001$) et pour le plus haut niveau (IP-finale) qui était significativement différent de tous les autres niveaux excepté le niveau AP/XP (voyelles internes à l'AP: $\beta = -0.181$, $se = 0.228$, $t = -8.16$, $pMCMC < 0.0001$; voyelles en position finale d'AP: $\beta = -0.1022$, $se = 0.228$, $t = -4.59$, $pMCMC < 0.0001$; voyelles en position finale d'AP/XP: $\beta = -0.046$, $se = 0.228$, $t = -2$, $pMCMC < 0.05$). Notons qu'à la différence de la vitesse d'élocution normale, nous n'avons pas observé de différence significative entre le niveau AP-finale et le niveau AP/XP ($\beta = -0.044$, $t = 2$, $pMCMC = 0.05$), aussi bien qu'entre le niveau IP-finale et AP/XP ($\beta = -0.046$, $se = 0.228$, $t = -2$, $pMCMC < 0.05$) et ce quel que soit le genre des locuteurs. Le modèle incluait 1875 mesures. L'interaction entre la vitesse d'élocution et le type de frontière prosodique était significative avec un plus haut degré de réduction de la voyelle observé pour les niveaux prosodiques élevés (pour les voyelles en finale d'IP: $\beta = 0.378$, $se = 0.031$, $t = 12.02$, $p < 0.0001$; for AP/XP : $\beta = 0.06$,

se=0.031, $t=2.18$, $pMCMC<0.001$) plutôt que pour le plus bas niveau ($\beta=0.001$, se=0.031, $t=2$, $p<0.05$, $pMCMC=0.05$). Les interactions entre les autres facteurs fixes n'étaient pas significatives. En effet, les valeurs de $pMCMC$ étaient toutes supérieures à 0.05.

Les mêmes effets ont été observés pour la durée des syllabes. Les médianes des durées des syllabes estimées par le modèle sont représentées dans le tableau 6. Ces résultats ne sont pas surprenants car nous savons que l'allongement pré-frontière influence préférentiellement la rime de la syllabe plutôt que l'attaque (pour l'anglais : White 2002, Turk & Shattuck-Hufnagel, 2007 ; pour le français : Astézano, 2001).

Type de frontière prosodique	Vitesse d'élocution normale		Vitesse d'élocution rapide	
	H	F	H	F
AP-interne	113	144	90	114
AP-finale	120	153	96	122
AP/XP	134	170	100	127
IP-finale	177	225	104	132

Tableau 6. Étude 1 : Durées des syllabes (ms) estimées par le modèle pour les quatre types de frontières prosodiques (AP-interne, AP-finale, AP/XP, IP-finale). La durée des syllabes est exprimée en fonction de la vitesse d'élocution (normale/rapide) et du genre des locuteurs (homme/femme). Les paires significativement différentes les unes des autres sont indiquées par un astérisque.

3.5. Résultats relatifs à la fréquence fondamentale (f_0)

Nous avons également examiné les effets du type de frontière prosodique sur les valeurs de f_0 associées aux syllabes de nos quatre conditions prosodiques (AP-interne, AP-finale, AP/XP; IP-finale). Un modèle mixte (LMM) a été réalisé sur les logarithmes des plus hautes valeurs de f_0 associées aux syllabes cibles. De manière similaire à ce que nous avons fait pour les valeurs de durées, nous avons retenu, le type de frontière prosodique, la vitesse d'élocution et le genre en tant que facteurs fixes. Le locuteur et la syllabe ont été inclus en tant que facteurs aléatoires.

Les valeurs de f_0 estimées par le modèle sont représentées dans la figure 22.

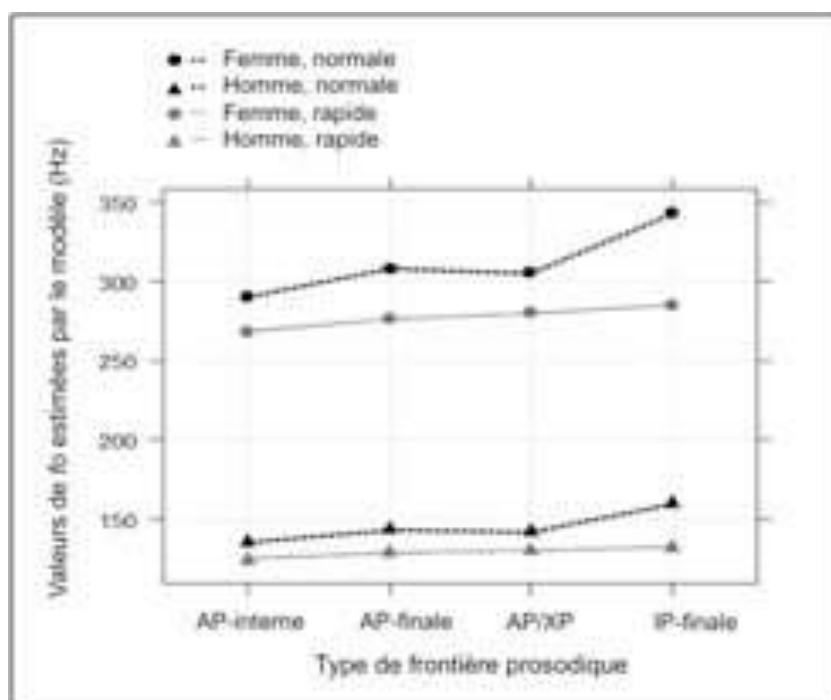


Figure 22. Étude 1 : Valeurs des pics de f_0 estimées par le modèle LMM.

Les valeurs sont fonction de la condition prosodique (AP-interne, AP-finale, AP/XP et IP-finale) de la vitesse d'élocution (normale/rapide) et du genre des locuteurs (homme/femme).

À vitesse d'élocution normale, les valeurs de f_0 associées aux syllabes pré-frontière étaient significativement plus hautes pour les syllabes en position finale d'IP que pour les syllabes associées à tous les autres niveaux de frontière (condition AP-interne : $\beta=0.165$, $se=0.007$, $t=21.9$, $pMCMC<0.001$; condition AP-finale : $\beta=0.107$, $se=0.007$, $t=14.17$, $pMCMC<0.0001$; condition AP/XP : $\beta=0.116$, $se=0.007$, $t=15.37$, $pMCMC<0.001$). Contrairement à ce qui était observé pour la durée des segments, aucune différence significative n'est observée entre la condition AP-finale et la condition AP/XP indépendamment du genre des locuteurs et de la vitesse d'élocution (à vitesse d'élocution normale : $\beta=-0.009$, $se=0.007$, $t=-1.2$, $pMCMC=0.232$; à vitesse d'élocution rapide : $\beta=0.011$, $se=0.007$, $t=-1.52$, $pMCMC=0.136$). Les valeurs de f_0 estimées par le modèle LMM pour chaque condition prosodique en fonction de la vitesse d'élocution et du genre des locuteurs sont reportées dans le tableau 7.

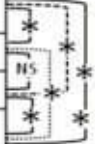
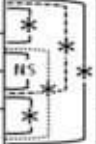
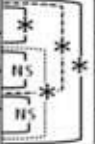
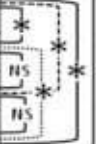
Type de frontière prosodique	Vitesse d'élocution normale		Vitesse d'élocution rapide	
	H	F	H	F
AP-interne	135 	291 	125 	268 
AP-finale	143 	308 	128 	276 
AP/XP	142 	305 	130 	279 
IP-finale	159 	343 	132 	285 

Tableau 7. Étude 1 : Valeurs de f_0 (Hz) estimées par le modèle pour les quatre types de frontières prosodiques (AP-interne, AP-finale, AP/XP, IP-finale). Les valeurs de f_0 sont exprimées en fonction de la vitesse d'élocution (normale/rapide) et du genre des locuteurs (homme/femme).

Nous pouvons également remarquer qu'à vitesse d'élocution rapide, les valeurs de f_0 des syllabes en position finale d'IP (condition IP-finale) n'étaient pas significativement plus hautes que les syllabes en position finale d'AP associé à une rupture syntaxique majeure (condition AP/XP) ($\beta=0.019$, $se=0.007$, $t=1.9$, $pMCMC=0.05$). En revanche les syllabes en position finale d'IP étaient significativement plus hautes que les syllabes en position finale d'AP ($\beta=0.030$, $se=0.007$, $t=-4.03$, $pMCMC<0.001$). Conformément à ce qui était attendu, les valeurs de f_0 étaient significativement plus hautes pour les femmes que pour les hommes ($\beta=0.765$, $se=0.0095$, $t=8.02$, $p<0.001$) bien que le même nombre de niveau prosodiques observés sur la base des valeurs de f_0 ait été obtenu pour les hommes et pour les femmes (trois niveaux distincts à vitesse d'élocution normale et quatre niveaux distincts à vitesse d'élocution rapide). L'effet de la vitesse d'élocution était également significatif ($\beta=-0.081$, $se=0.007$, $t=-10.74$, $pMCMC<0.001$). Le modèle incluait 1875 mesures. L'interaction entre la vitesse d'élocution et le type de frontière prosodique était significative avec un degré plus important d'abaissement des valeurs de f_0 associé pour les syllabes situées en position finale d'IP ($\beta=-0.789$, $se=0.031515$, $t=-12.02$, $pMCMC<0.001$) que pour les valeurs de f_0 associées aux syllabes en position finale des autres conditions prosodiques (AP/XP, AP-final, AP-interne). Les interactions entre les autres facteurs fixes n'étaient pas significatives. En effet, les $pMCMC$ étaient toutes supérieures à 0.05.

4. DISCUSSION

Dans cette expérience nous avons testé l'influence de la syntaxe et de la vitesse d'élocution sur la réalisation des frontières de syntagme accentuel en français. Pour cela nous avons mesuré les indices de durée et les indices mélodiques présents sur des syllabes en position finale d'AP qui n'étaient alignées avec aucune rupture syntaxique majeure (condition AP-finale) et des syllabes en position finale d'AP qui étaient alignées avec une rupture

entre une syntagme nominal et un syntagme verbal. L'hypothèse nulle était que, conformément aux deux niveaux de structuration prosodique traditionnellement admis pour le français, les indices de durée et de *fo* ne différaient pas dans les deux conditions AP-finale et AP/XP. Le plan expérimental incluait également deux conditions supplémentaires : la condition où les syllabes étaient en position finale d'IP (condition IP-finale) et la condition où les syllabes étaient inaccentuées (en position interne à l'AP, condition AP-interne). En plus des mesures de durées et de *fo* nous sommes également intéressée à la réalisation du patron mélodique de l'AP et aux différents facteurs influant sur la réalisation de celui-ci. De plus, étant donnée qu'il a été montré que la vitesse d'élocution affecte la structuration prosodique, nous avons également manipulée celle-ci dans notre expérience.

Les résultats obtenus nous amènent à tirer un certain nombre de conclusions sur l'influence de la structure syntaxique et de la vitesse d'élocution sur le découpage prosodique en français. Nos résultats permettent donc une meilleure appréhension des relations existantes entre la structure prosodique abstraite d'une part et sa réalisation concrète dans la parole d'autre part.

Tout d'abord, les résultats obtenus montrent que la vitesse d'élocution rapide n'a pas induit pas de restructuration significative du découpage en constituants prosodiques dans nos données contrairement à ce qui a été précédemment montré par Jun & Fougeron (1998). En effet, seulement 2,3 % des énoncés de notre corpus (c'est-à-dire 45 énoncés) ont montré soit un effacement de frontière prosodique (soit d'AP, soit d'IP) ou le remplacement d'une frontière prosodique par une frontière de niveau inférieur dans la hiérarchie prosodique (remplacement d'une frontière d'IP par une frontière d'AP).

Nos résultats montrent également que la réalisation du patron mélodique de type /LH*/ , par opposition aux autres types de patrons mélodiques (/LHiLH*/ , /LLH*/ , LHIH*/ , /HiLH*/ , /LHiL*/), est conditionnée par la vitesse d'élocution, le nombre de syllabes contenues dans l'AP ainsi que le genre du locuteur. En effet, nous avons observé plus de patrons mélodiques de type /LH*/ à vitesse d'élocution rapide qu'à vitesse d'élocution normale, pour les APs contenant 4 syllabes que pour les APs contenant 3 syllabes et chez les hommes plutôt que chez les femmes. Notons qu'un lien avait déjà été établi dans la littérature entre présence d'une montée initiale et longueur de l'AP (Welby, 2006; German & D'Imperio, 2010). Jun & Fougeron (1998) avait également mis en évidence un lien entre occurrence du patron mélodique de type /LH*/ et vitesse d'élocution. Cependant, aucune étude menée jusqu'à ce jour sur le français n'avait mis en évidence l'influence du genre des locuteurs sur la réalisation du patron mélodique de l'AP. Le fait que nous ayons observé un nombre significativement plus important de patrons mélodiques de type /LH*/ chez les hommes que chez les femmes nous a amené à nous questionner sur une

possible influence du genre des locuteurs sur la réalisation du patron mélodique de l'AP en français. Une influence de ce type a récemment été mise en évidence en anglais par Clopper & Smjlianic (à paraître). Les auteurs ont en effet observé que dans certaines variétés d'anglais américain, certains types d'accents mélodiques tels que L*+H sont plus employés par les femmes que par les hommes. Les résultats que nous avons obtenus nous ont amené à nous demander si un patron mélodique plus complexe (notamment un patron mélodique incluant une montée mélodique initiale de type /LHiLH*/ ou /LHiH*/) pourrait éventuellement être conditionnée par des facteurs indexicaux tel que le genre du locuteur. Cependant l'effet du genre que nous avons observé pourrait être une conséquence de la vitesse d'élocution et non un effet indépendant. De manière plus spécifique, étant donné que nous avons observé une vitesse d'élocution plus élevée chez les hommes que chez les femmes et que nous savons que la vitesse d'élocution conditionne la réalisation du patron mélodique de l'AP, l'effet observé pourrait être lié à la différence de vitesse d'élocution.

Enfin, les résultats obtenus montrent qu'à vitesse d'élocution normale, les syllabes en position finale d'AP alignées avec une rupture entre un syntagme nominal et un syntagme verbal étaient significativement plus longues que les syllabes situées en position finale d'AP contenues à l'intérieur d'un syntagme nominal. Notons que si des différences significatives de durée ont été observées, les valeurs de *f₀* mesurées dans les deux conditions ne montraient, quant à elles, pas de différence significative. Ces résultats posent la question du nombre de constituants prosodiques pertinents pour le français. Rappelons que deux constituants prosodiques sont traditionnellement admis pour le français: le syntagme accentuel et le syntagme intonatif. Étant donné que nous avons observé un degré d'allongement significativement différent en frontière d'AP alignée avec une rupture syntaxique majeure mais pas en frontière d'AP contenue à l'intérieur d'un constituant syntaxique majeur, la question de l'existence d'un niveau de structuration prosodique intermédiaire dont l'émergence serait contrainte par des algorithmes syntaxiques mérite d'être discutée.

Dans la section qui suit et à la vue des résultats obtenus dans cette première étude, nous allons tour à tour discuter (i) l'influence de la vitesse d'élocution sur la réalisation des frontières d'AP puis (ii) les facteurs qui influent sur la réalisation du patron mélodique de l'AP et enfin (iii) l'influence de la syntaxe sur la réalisation des frontières d'AP en français.

4.1. Influence de la vitesse d'élocution sur la réalisation des frontières d'AP

Tout d'abord, nos résultats permettent une meilleure analyse de l'influence de la vitesse d'élocution sur la structure prosodique. L'annotation des frontières prosodiques observées dans nos données nous a permis de montrer qu'une augmentation de la vitesse d'élocution ne semble pas

induire de restructuration significative des découpages prosodiques obtenus à vitesse normale contrairement à ce qui a été observé par Jun & Fougeron. En effet, seulement 2,3% des énoncés de notre corpus (c'est-à-dire 45 énoncés) présentaient des phénomènes d'effacement de frontière (d'AP ou d'IP) ou de remplacement de frontières par une frontière de niveau inférieur dans la hiérarchie prosodique (c'est-à-dire une frontière d'IP remplacée par une frontière d'AP).

Les analyses de durées ont montré que la relation hiérarchique entre unités prosodiques et allongements pré-finaux a été préservée à vitesse d'élocution rapide. En effet, nous avons observé différents degrés d'allongement entre les syllabes inaccentuées (AP-interne) et les syllabes en position finale d'AP et entre les syllabes en position d'AP et les syllabes en position d'IP conformément à ce qui avait été observé dans de précédentes études (Hirst & Di Cristo, 1984; Pasdeloup 1990; Vaissière, 1983). Notons que, contrairement à ce qui a été observé pour la vitesse d'élocution normale, aucune différence significative entre la durée des syllabes en position finale d'AP contenues à l'intérieur de syntagmes nominaux et la durée des syllabes situées à la frontière entre un syntagme nominal et un syntagme verbal a été observée. Ce résultat n'est pas surprenant étant donné qu'une réduction globale de la vitesse d'élocution se traduit par de plus petits allongements pré-finaux de sorte qu'il est plus difficile de parvenir à un seuil qui soit statistiquement significatif. En effet, même si la différence n'est pas significative, nous observons une tendance pour les syllabes de la condition AP/XP à être plus longues que les syllabes de la condition AP-finale. Cette différence était de l'ordre d'environ 5 ms.

Parmi les stratégies employées par les locuteurs du français dans le but de parvenir à augmenter leur vitesse d'élocution, Jun & Fougeron (1998) ont observé une simplification des contours mélodiques, que les auteurs ont interprété en tant que phénomènes de "restructuration prosodique" (*prosodic restructuring*). Par ces termes les auteurs ont désigné des phénomènes d'effacement de certaines frontières d'AP ou d'IP (entraînant un regroupement d'un plus grand nombre de syllabes à l'intérieur du même constituant prosodique et une réorganisation de l'association entre ton et syllabe) aussi bien que des phénomènes de remplacement d'une frontière prosodique par une frontière de niveau inférieur dans la hiérarchie prosodique (c'est-à-dire le remplacement d'une frontière d'IP par une frontière d'AP). Nos données ont très peu été affectées par ce type de phénomène qui a été observé seulement dans 2,3% des énoncés produits, c'est-à-dire, de manière plus précise, dans 45 phrases. De façon détaillée, dans 41 de ces 45 phrases nous avons observé des phénomènes de réduction de frontière (c'est-à-dire de remplacement de la frontière d'IP par une frontière d'AP). Pour 4 phrases seulement nous avons observé des phénomènes d'effacements de frontières. Deux facteurs peuvent, selon nous, permettre d'expliquer cette différence de résultat entre notre étude et l'étude de Fougeron & Jun (1998) où 19 à 29 % des frontières d'AP étaient

supprimées et 20 à 60 % des frontières d'IP étaient remplacées par des frontières d'APs. Premièrement, le matériel linguistique employé présente des différences qui nous semblent importantes. Dans notre étude, nous avons utilisé un ensemble de phrases isolées et très contrôlées dans lesquelles les syllabes cibles présentaient toujours la même position dans la phrase et dans lesquelles les structures syntaxiques et sémantiques étaient toujours les mêmes. En revanche, le matériel utilisé par Fougeron & Jun variait beaucoup plus du point de vue de plusieurs facteurs linguistiques et notamment du point de vue de la structuration du discours. En effet, Fougeron & Jun ont comparé des frontières prosodiques pouvant apparaître dans différentes parties du monologue "La Bise et le Soleil". Le monologue utilisé était divisé en deux parties, la première incluait deux phrases d'introduction et la deuxième partie consistait en une série de quatre phrases constituant le développement et la conclusion du texte. Nous savons qu'il existe une relation étroite entre discours et structure prosodique. Par exemple, des énoncés situés en début de paragraphe sont habituellement caractérisés par un registre plus haut et plus étendu que des énoncés situés en fin de paragraphe. Nous renvoyons le lecteur aux études de Hirshberg & Pierrehumbert (1986), Swerts, (1997) ou den Ouden, Noordman & Terken, (2009) pour plus d'informations sur ce point. Il a également été montré que les énoncés en début de paragraphe sont produits avec une intensité plus importante (Brown & Yule, 1983) ainsi que des pauses plus longues (Swerts, 1997; Fon, 2002) que les énoncés situés en fin de paragraphe. Nous nous demandons dans quelle mesure, la structure du texte, c'est-à-dire le fait que les auteurs comparent des phrases pouvant apparaître soit au début soit à la fin d'une section, a pu influencer leurs résultats. Par exemple, à vitesse d'élocution rapide, les auteurs observent que dans un énoncé composé de deux APs ("qui arrivaient" et "le premier"), ces deux APs avaient tendance à se combiner en un seul AP. L'énoncé était situé à la fin de leur première section qui est une région que l'on peut supposer être très affectée par la réduction des pauses et potentiellement par des effets de réduction de frontière prosodique étant donné qu'il s'agit de la fin (et non du début) de l'introduction du texte. Nous pourrions spéculer que leur locuteur aurait tendance à produire plus de phénomènes de restructuration prosodique à la fin d'un paragraphe plutôt qu'au début. Ceci pourrait être lié au fait que la force des indices prosodiques soit affectée par la position de la frontière dans le discours avec des indices plus marqués au début d'une partie du discours plutôt qu'à la fin, avec par exemple des allongements pré-finaux plus importants, des pauses plus longues et des pics de *f₀* plus hauts.

De plus, nous savons que dans certains cas, les indices de durée ne s'accordent pas avec les indices mélodiques dans le marquage des frontières prosodiques. Dans le système d'annotation de l'intonation ToBI qui a été proposé à l'origine pour l'anglais américain mais qui est aujourd'hui utilisé dans de nombreuses langues (*Tone and Break Indices* ; Beckman & Elam, 1997), les indices de rupture correspondent aux trois catégories distinguées

pour l'anglais : le mot (break indice 1), le syntagme intermédiaire (break indice 3) et le syntagme intonatif (break indice 4). En plus de ces trois indices, l'indice 2 est employé pour signaler les cas où seuls les indices de durée marque la frontière d'ip.

Dans nos données, à vitesse d'élocution rapide, nous avons rencontré un certain nombre de cas où la perception d'une frontière prosodique, pourtant très clairement marquée à l'écoute, n'était pas liée à des indices mélodiques saillants. Dans la figure 23, malgré l'absence d'indices mélodiques, en raison d'un registre très compressé dû à la vitesse d'élocution rapide, les deux annotateurs ont clairement perçu la présence d'une frontière d'AP certainement en raison de la préservation des indices de durée.

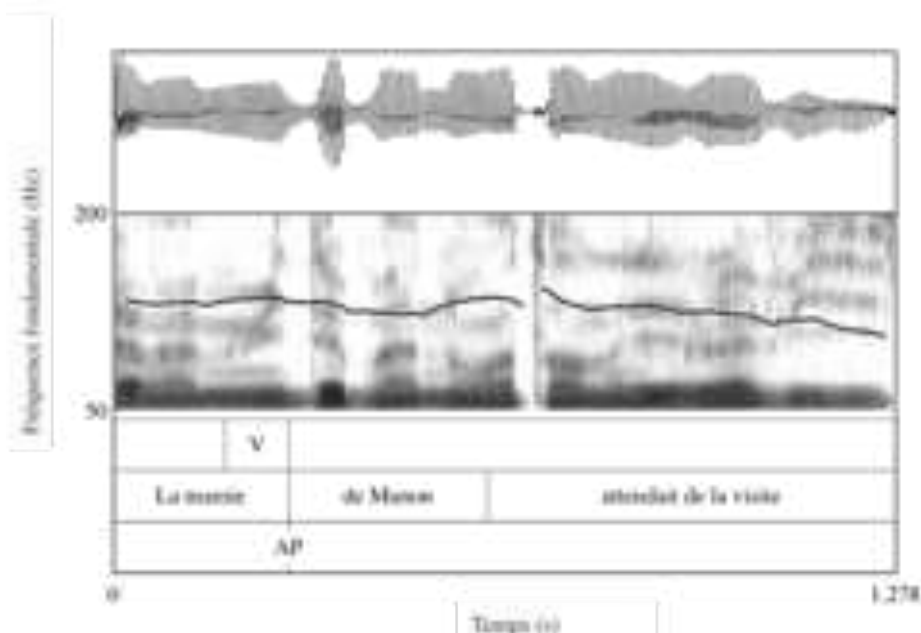


Figure 23. Étude 1 : frontière d'AP marquée uniquement par les indices de durée. La frontière d'AP associée à la syllabe "mi" du mot "mamie" à l'intérieur de l'énoncé "La mamie de Manon attendait de la visite" était perçue par les deux annotateurs malgré l'absence d'indices mélodiques.

Notons que le repérage de constituants prosodiques grâce à l'allongement pré-frontière lorsque l'étendue du registre est très compressé avait également été mis en pour le français par Di Cristo et Jankowski (1999). Les indices de durée semblent jouer un rôle important dans la perception des frontières prosodiques en anglais (Breen, Gibson & Watson, 2010). Un certain nombre d'études semblent également souligner l'importance des indices de durée dans la perception des frontières prosodiques en français (Vaissière, 1983 ; Vaissière & Michaud, 2006). Nous assumons le fait que lorsque le registre est compressé, comme c'est le cas à vitesse d'élocution rapide, les auditeurs basent leur identification des frontières prosodiques essentiellement sur des indices de durée plutôt que sur des indices mélodiques. Nos résultats suggèrent la pertinence d'un

marquage pluriel des frontières prosodiques selon lequel, la pertinence perceptive d'indices acoustiques spécifiques serait modulé par un ensemble complexe de contraintes parmi lesquelles la vitesse d'élocution ou d'autres facteurs linguistiques (tels que des facteurs sémantiques ou relatifs à la structuration du discours) ainsi qu'éventuellement des facteurs indexicaux (tel que le genre des locuteurs) auraient leur place.

4.2. Facteurs influençant la réalisation du patron mélodique de l'AP

Nous avons également analysé les patrons intonatifs réalisés dans le but de déterminer l'effet de la vitesse d'élocution, du nombre de syllabes à l'intérieur de l'AP et du genre des locuteurs sur la réalisation du patron mélodique de l'AP. Un modèle mixte de régressions logit a en effet été employé dans le but d'examiner les facteurs qui influencent l'apparition du patron mélodique de type /LH*/ par opposition à tous les autres types de patrons mélodiques qui sont des réalisations plus complexes du patron mélodique par défaut de l'AP. Rappelons que dans le modèle proposé par Jun & Fougeron, le patron mélodique par défaut de l'AP est le patron /LHiLH*/. Le modèle prédit également cinq réalisations de surface résultant de l'absence d'un ou de plusieurs tons du patron mélodique par défaut dans la réalisation de surface. Lorsque la cible initiale haute Hi et la cible suivante basse L ne sont pas réalisés, il en résulte la forme de réalisation minimale postulée pour l'AP: le patron mélodique /LH*/. Les facteurs influençant la réalisation des variantes de surface de l'AP ne sont pas explicitement décrits dans le modèle de Jun & Fougeron puisqu'il est seulement précisé que les variantes simplifiées du patron mélodique par défaut sont achevées grâce à la non réalisation d'un ou de plusieurs tons de la forme par défaut (Jun & Fougeron 1995, 2000, 2002). Le modèle statistique conduit sur nos données prédit un plus grand nombre de patron mélodique de type /LH*/ à vitesse d'élocution rapide par rapport à la vitesse d'élocution normale ce qui va dans le sens d'une réduction articulatoire globale. Plusieurs études ont en effet montré une simplification des mouvements articulatoires lorsque la vitesse d'élocution augmente. Parmi les stratégies employées par les locuteurs dans le but de réduire le temps d'articulation, les sujets peuvent réduire la magnitude spatiale de leurs mouvements articulatoires (Lindblom, 1963 ; 1964; Kent & Moll, 1972; Gay, 1981) ou peuvent simplement supprimer une partie des gestes (Munhall & Löfqvist 1992). De manière similaire à ce qui se passe pour la simplification des mouvements articulatoires, nos résultats peuvent laisser penser qu'il existe une simplification des contours mélodiques lorsque la vitesse d'élocution augmente. Fougeron & Jun (1998) ont également observé ce type de simplification du patron mélodique par défaut à vitesse d'élocution rapide. Les auteurs interprètent cette simplification en tant qu'inaptitude des locuteurs à réaliser les quatre tons du patron mélodique de base sous l'effet de contraintes temporelles.

Fougeron & Jun (1998) observent que le ton H initial de l'AP n'est jamais réalisé à vitesse d'élocution rapide pour deux de leurs trois locuteurs lors de la lecture du texte « La bise et le soleil ». Les auteurs expliquent ces résultats par un chevauchement entre le geste Haut et le geste Bas suivant, lors de la réalisation du contour mélodique. Dans nos données, le patron mélodique /LH*/ était le patron le plus communément observé pour les APs contenant 3 et 4 syllabes à vitesse d'élocution rapide (93% des patrons mélodiques réalisés pour les APs contenant 3 syllabes et 83% des patrons mélodiques réalisés pour les APs contenant 4 syllabes chez les hommes) avec des pourcentages de réalisation toujours plus élevés chez les hommes que chez les femmes. Le pourcentage de réalisation du patron mélodique /LH*/ diminuait avec l'augmentation du nombre de syllabes et ce quel que soit la vitesse d'élocution alors que le patron mélodique /LHiLH*/ augmentait avec l'augmentation du nombre de syllabes. Cela va dans le sens d'études précédentes qui prédisent un nombre plus important de patrons mélodiques de type /LHiLH*/ ou /LHiH*/ pour les APs contenant 4 syllabes et plus par rapport aux APs de moins de 4 syllabes. Conformément à ce qui a été observé par Jun & Fougeron (1995), dans les APs contenant 3 syllabes, nous avons parfois observé une trace du ton initial H. Cette petite montée initiale a été plus fréquemment observée à vitesse d'élocution normale qu'à vitesse d'élocution rapide. La figure 24 fait apparaître deux réalisations du premier AP « Le mari » montrant un patron mélodique réalisé avec seulement une montée finale (/LH*/, 24a) ou avec une petite montée initiale additionnelle (24b). Les deux réalisations ont été prononcées par une locutrice à vitesse d'élocution normale.

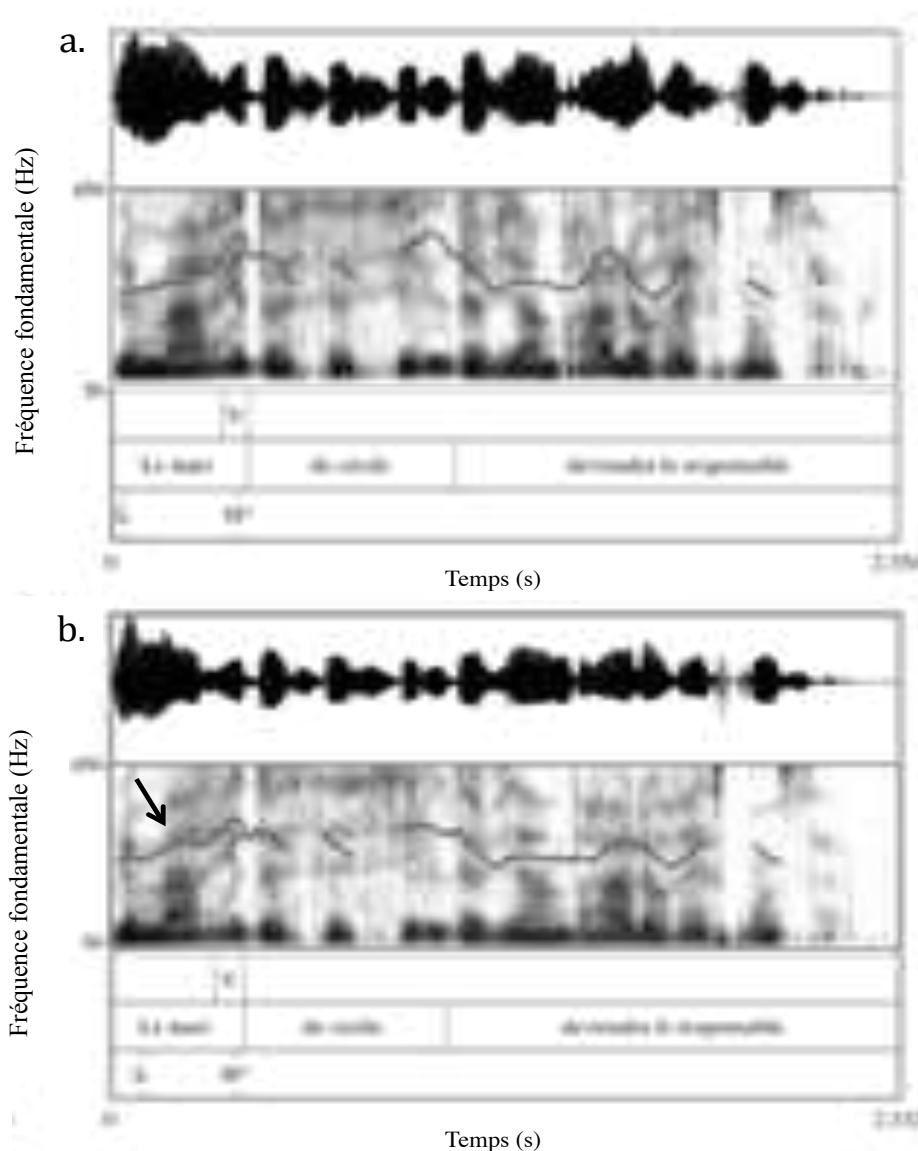


Figure 24. Étude 1 : réalisation de la montée initiale de *fo*.
Deux réalisations du premier AP "Le mari" faisant toutes deux apparaître un patron mélodique de type /LH*/ contenant uniquement une montée finale (a) une petite montée initiale additionnelle et optionnelle (b) prononcées par une de nos locutrices à vitesse d'élocution normale.

Enfin, nos résultats nous ont amené à nous questionner sur une éventuelle influence du genre des locuteurs sur la réalisation du patron mélodique /LH*/ de l'AP. En effet, notre modèle statistique a prédit un nombre plus important de patron mélodique de type /LH*/ chez les femmes que chez les hommes indépendamment de la vitesse d'élocution. Cependant l'effet « genre » obtenu pourrait simplement être une conséquence de la vitesse d'élocution qui était plus élevée chez les hommes que chez les femmes. Il a déjà été montré que le genre affecte la structure tonale en anglais (Clooper & Smijlanic, à paraître) alors qu'aucun effet de ce type n'a été mis

en évidence jusqu'ici mis en évidence pour le français. Nous pensons que la simplification du patron mélodique par défaut de l'AP pourrait peut être dépendre du genre du locuteur en français. Cependant nos résultats ne nous permettent pas de mettre en évidence ce type d'effet et d'autres expérimentations seront nécessaires dans le but de montrer ou non l'influence du genre des locuteurs sur la réalisation de l'AP en français.

4.3. Influence de la syntaxe sur la réalisation des frontières d'AP

Les résultats de deux modèles mixtes de régressions multiples conduits sur nos données ont permis de mettre en évidence le fait qu'à vitesse d'élocution normale, les syllabes en position finale d'AP alignées avec une rupture entre un syntagme nominale et un syntagme verbal, sont significativement plus longues que les syllabes situées en position finale d'AP contenu à l'intérieur d'un syntagme nominal. Par contre les valeurs de *f₀* n'étaient pas significativement différentes dans les deux conditions. Ce résultat pose la question du nombre et de la définition des constituants prosodiques en français. En effet, sur la base des indices de durée, il apparaît qu'un niveau de structuration intermédiaire situé entre le syntagme accentuel et le syntagme intonatif soit nécessaire pour rendre compte des phénomènes d'allongement observés à la frontière des constituants prosodiques en français. Sur cette base, nous pourrions émettre l'hypothèse que l'alignement entre la frontière droite d'un syntagme accentuel et la frontière droite d'un constituant syntaxique majeur conduise au placement d'une frontière droite de syntagme intermédiaire en français. Nous aborderons plus en détails ce type de contrainte d'alignement dans le chapitre suivant. Retenons pour l'instant que ce syntagme intermédiaire ne serait pas restreint à des structures spécifiques tel que cela avait précédemment été prédit par Jun & Fougeron (2000). Rappelons que dans le cadre du modèle proposé par Jun & Fougeron, l'existence d'un constituant prosodique intermédiaire situé entre l'AP et l'IP avait déjà été postulée dans le but de rendre compte de certains phénomènes intonatifs spécifiques (des phénomènes de plateaux) observés sur des structures syntaxiques particulières telles que des structures clivées, extraposées ou autres (cf. figure 10). Dans nos données les différences d'allongement pré-frontière observés entre les syllabes finales d'AP contenue à l'intérieur d'un syntagme nominal et les syllabes finales d'AP associées à une rupture entre un syntagme nominal et un syntagme verbal pourraient être expliquées par des différences de force de frontière prosodique (frontière de syntagme accentuel vs. frontière de syntagme intermédiaire). Cependant aucun indice mélodique venant s'ajouter à la différence d'allongement pré-frontière n'a été observé sur les syllabes en finale d'APs associés à une rupture syntaxique majeure. Cette absence de correspondance entre indices de durée et indices

mélodiques conduisent à deux interprétations possibles. Première interprétation possible: les indices de durées (malgré l'absence d'indices mélodiques) sont suffisants pour marquer la frontière d'un constituant intermédiaire en français. Des expériences menées en perception de la parole sont cependant nécessaires pour confirmer le fait que des auditeurs sont capables de différencier le degré d'allongement de l'AP et le degré d'allongement observé lorsque l'AP est aligné avec une rupture syntaxique majeure. Nous reviendrons sur cette question dans le chapitre 5 de cette thèse où, au moyen d'une tâche de complétion de phrases, nous fournirons des éléments de réponses à cette question. Deuxième interprétation possible : les phénomènes d'allongement observés à la frontière d'AP alignée avec une rupture syntaxique majeur ne sont pas associés avec un niveau additionnel de structuration prosodique (le syntagme intermédiaire) mais résulteraient d'une interface directe entre la structure syntaxique et le indices acoustiques observés dans le signal de parole. En effet, nous avons précédemment évoqué le fait que des travaux antérieurs sur l'interface syntaxe/prosodie en français ont postulé l'existence d'une interface directe entre les degrés d'enchâssement de la structure syntaxique et les degrés d'allongement observés (Gee & Grosjean, 1993; Monin & Grosjean, 1993; Keller, Zellner, Werner & Blanchoud, 1993; Rossi, 1999).

A la seule vue de ces résultats nous ne sommes pas encore en mesure de trancher entre les deux interprétations possibles. Cependant de nombreuses études dans la littérature montrent que la longueur des constituants contraint de manière importante le placement des frontières prosodiques. Nous renvoyons par exemple le lecteur à D'Imperio et al. (2005) pour plus de renseignements sur ce point. Nous pensons que la manipulation de la taille du constituant pourrait potentiellement engendrer l'apparition d'indices mélodiques à la frontière d'AP associée à une rupture syntaxique majeure. L'hypothèse pourrait par exemple être que l'abaissement des accents mélodiques successifs soit bloqué à la frontière d'AP associé à une rupture syntaxique majeure de manière similaire à ce qu'il se passe pour l'anglais où l'abaissement des accents mélodiques est stoppé à la frontière de syntagme intermédiaire (Beckman & Pierrehumbert, 1986 ; cf. chapitre 3 pour plus de détails sur ce point).

Nous proposons donc, en plus de manipuler la structure syntaxique, de manier la taille du constituant dans le but de voir si cette manipulation engendre ou non l'apparition d'indices mélodiques marquant la frontière d'AP associée à une rupture syntaxique majeure. La présence de tels indices mélodiques, ajoutés aux indices de durée déjà observés, nous fournirait la preuve de l'existence d'un niveau de structuration prosodique supplémentaire pour le français : le syntagme intermédiaire. C'est ce que nous allons tester dans le chapitre suivant.

CHAPITRE 3

EFFETS D'ABAISSEMENT, VARIATIONS DE REGISTRE ET DÉFINITION DU SYNTAGME INTERMÉDIAIRE EN FRANÇAIS (Étude 2)

1. PROBLÉMATIQUE

Comme nous l'avons précédemment évoqué, nous savons aujourd'hui que la structure prosodique est composée de différents niveaux de structure qui sont hiérarchisés. Cependant, même pour une langue aussi bien étudiée que l'anglais, le nombre de constituants existant au sein de cette hiérarchie prosodique est toujours sujet à controverse. Malgré le fait que deux niveaux de structure ait été proposés (le syntagme intonatif ou *intonational phrase* et le syntagme intermédiaire ou *intermediate phrase*, Beckman & Pierrehumbert, 1986), un certain nombre d'études ont suggéré que le degré d'allongement pré-frontière pourrait être proportionnel au niveau de structuration du constituant dans la hiérarchie grammaticale et l'existence d'au moins quatre niveaux définis sur la base de l'allongement pré-frontière ont été observés (Wightman, Shattuck-Hufnagel, Ostendorf & Price, 1992). Dans le chapitre précédent nous avons obtenu des preuves de l'existence d'au moins trois niveaux de structure définis sur la base de l'allongement pré-frontière pour le français. En effet, outre le degré d'allongement pré-frontière associé à la frontière du syntagme accentuel et celui associé à la frontière du syntagme intonatif déjà mis en évidence dans la littérature, nous avons observé le fait que les syllabes en position finale d'AP associées à une rupture syntaxique majeure (comme dans la phrase « Ré**MY**]SN demandait l'institutrice ») étaient significativement plus longues que les syllabes en position finale de syntagme accentuel qui étaient contenues à l'intérieur d'un syntagme nominal complexe (par exemple « Les a**MIS** de Rémy]SN »). Cela va dans le sens l'existence d'au moins trois niveaux de structuration prosodique définis sur la base de l'allongement pré-frontière en français. Cependant, ce niveau de structuration intermédiaire observé entre l'AP et l'IP lorsque la syllabe était associée à une rupture syntaxique majeure ne semblait pas marqué par la présence d'indices mélodiques particuliers. En effet, nous n'avons pas observé de contour mélodique particulier associé à la frontière d'AP/XP. De plus, les mesures de

de *fo* que nous avons réalisés ont montré que les syllabes en position finale d'AP/XP n'étaient pas plus hautes que les syllabes en finale d'AP. Nous nous posons donc la question suivante : le niveau d'allongement pré-frontière que nous avons observé est-il la marque d'un niveau de structure prosodique additionnel en français ?

Dans ce chapitre nous allons répondre à cette question en fournissant des preuves de l'existence d'un syntagme intermédiaire situé entre l'AP et l'IP pour le français. Au moyen d'un corpus de 384 phrases lues par 6 locuteurs de langue maternelle française, nous définirons les propriétés phonétiques et phonologiques de ce constituant.

Beckman & Pierrehumbert (1986) ont les premiers mis en évidence l'existence d'un constituant prosodique appelé syntagme intermédiaire pour l'anglais (*intermediate phrase*, ip). Sur la base d'une comparaison entre la structure intonative du japonais et de l'anglais, les auteurs ont défini le syntagme intermédiaire en tant que domaine de l'abaissement des accents mélodiques successifs ou *downstep* (Pierrehumbert, 1980). Ce phénomène a également été appelé *catathesis* (Pierrehumbert & Beckman 1986). La règle de *downstep* proposée pour l'anglais par Pierrehumbert (1980) puis par Liberman & Pierrehumbert (1984) est responsable de l'abaissement et de la compression du registre (plus précisément de la plage tonale utilisée par les locuteurs d'une langue ou *pitch range*)⁴ et permet de rendre compte de l'apparition de tons médians (*mid tone*) dans les contours mélodiques. Rappelons que dans le modèle de l'intonation de l'anglais américain proposé par Pierrehumbert (1980), seuls deux tons (*Low* et *High*) sont admis. La règle de *downstep* ou d'abaissement spécifie que tous les accents bitonaux sont abaissés par rapport à l'accent bitonal qui précède. Lorsque le phénomène d'abaissement apparaît de façon répétitive, cela donne lieu à une configuration tonale descendante comme celle illustrée dans la figure 25. Dans cette figure, nous pouvons observer un abaissement des accents mélodiques successifs sur chacun des mots « blubberies », « bayberries », « raspberries », « mulberries » et « brambleberries ».

⁴ Ladd (1996) distingue deux dimensions au sein du registre: (i) le niveau du registre (*level*) qui correspond à la hauteur globale à laquelle une séquence est énoncée : selon Ladd, la meilleure mesure pour cette dimension correspond à la moyenne des valeurs des cibles basses qui sont les plus stables et (ii) l'étendue du registre (*span*) qui correspond à la plage fréquentielle utilisée dans une séquence : on la mesure en soustrayant la moyenne des valeurs des cibles basses à celle des cibles hautes.

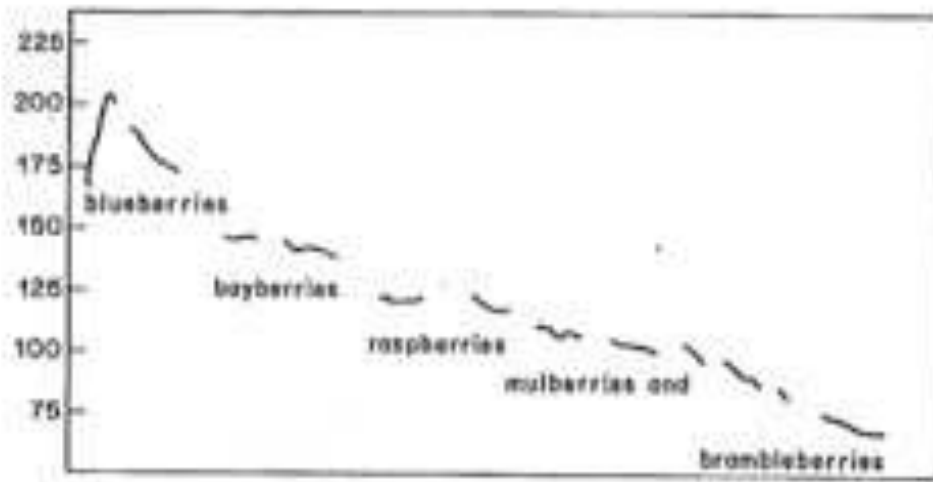


Figure 25. Phénomène d'abaissement des accents mélodiques successifs en anglais tiré de Beckman & Pierrehumbert (1986).

Selon Liberman & Pierrehumbert (1984), le degré d'abaissement diminue à chaque étape de manière à être modélisé par une décroissance exponentielle des valeurs de f_0 associées à chaque accent mélodique en position non-finale par rapport à une ligne de référence abstraite du registre. En outre, l'accent mélodique final est considéré comme plus bas que ce qui serait attendu. Ce phénomène d'abaissement final est appelé *final lowering* et est illustré dans la figure 26 empruntée à Truckenbrodt (2004). Sur cette figure nous pouvons voir que chaque accent mélodique (H) est abaissé par rapport au premier accent mélodique de l'énoncé qui marque la ligne de registre abstraite. Nous pouvons également remarquer que le degré d'abaissement par rapport à cette ligne de registre diminue à chaque accent mélodique avec moins d'abaissement pour les derniers accents mélodiques que pour les premiers. L'accent mélodique final est plus bas que ce qui est attendu illustrant le phénomène d'abaissement final. Le fait qu'un accent mélodique soit abaissé par rapport au précédent est noté grâce au signe de ponctuation « ! ».

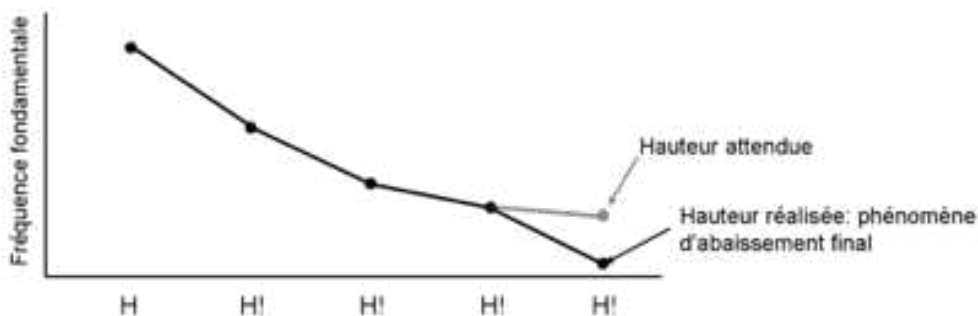


Figure 26. Schématisation du downstep et de l'abaissement final.

Figure adaptée de Truckenbrodt 2004 elle-même inspirée de Liberman and Pierrehumbert (1984).

Les phénomènes d'abaissement et d'abaissement final sont souvent discutés en relation avec le phénomène de déclinaison, phénomène également responsable de l'abaissement de la *fo*. La déclinaison est un phénomène global et dépendant du temps qui revoie à l'abaissement graduel de l'espace tonal dans lequel s'échelonne les cibles tonales au cours de l'énoncé (voir également Cohen & t'Hart, 1967 ; Vaissière, 1983 ; Ladd, 1992). Parce qu'elle affecte la fois les sommets et les creux de la *fo*, la déclinaison est schématisée à l'aide de lignes déclinantes telles qu'elles sont représentées dans la figure 27.

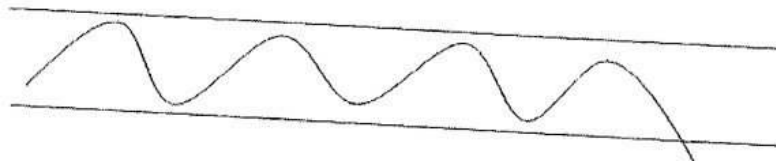


Figure 27. Schématisation de la déclinaison.

Contrairement à l'effet d'abaissement (*downstep*), l'abaissement des valeurs de *fo* entre deux pics dû à la déclinaison est égal en tout point du temps. La déclinaison est attribuée à des contraintes physiologiques telles que la diminution graduelle de la pression sous-glottique et est dépendante du temps linéaire. Dans le modèle de Pierrehumbert & Liberman et contrairement à l'effet de déclinaison, l'effet d'abaissement lié au *downstep* est indépendant de la distance entre deux tons H : la distance temporelle entre deux tons H n'a pas d'influence sur le degré d'abaissement. Dans ce modèle, la déclinaison est donc absorbée par le phénomène de *downstep*.

Bekman & Pierrehumbert (1986) ont postulé pour la première fois l'existence d'un constituant prosodique appelé syntagme intermédiaire pour le japonais et pour l'anglais. Le syntagme intermédiaire est de niveau inférieur au syntagme intonatif dans la hiérarchie prosodique. Dans ces deux langues la frontière du syntagme intermédiaire est responsable du blocage de l'abaissement successif des accents mélodiques (*downstep*). La frontière droite du syntagme intermédiaire est marquée par un accent de syntagme (L- ou H-) qui est responsable du blocage du *dowstep* après sa frontière droite. Ainsi, dans la figure 28 comportant deux syntagmes intermédiaires, les accents mélodiques sur « eleven », « and nine » « and one » sont abaissés les uns par rapport aux autres en raison du phénomène d'abaissement successif des accents mélodiques au sein de l'ip. En revanche, ce phénomène d'abaissement est bloqué en raison de la présence d'un accent de syntagme (L-) avant l'élément focalisé (« and EIGHTY »).

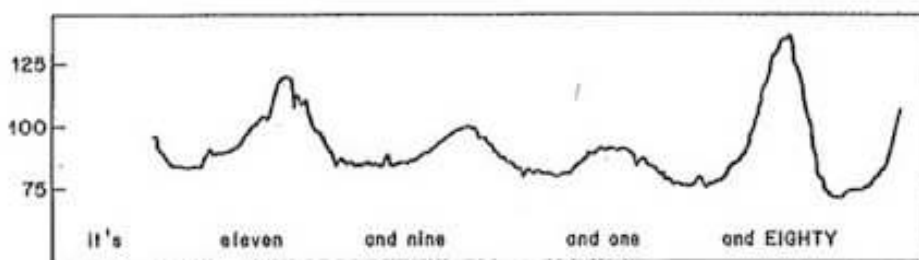


Figure 28. Exemple de blocage du downstep.

Contour de fo de la phrase « *It's eleven and nine and one and [eighty]F* » qui signifie en français « C'est onze puis neuf puis quatre vingt » où l'élément « *eighty* » est focalisé. Cet énoncé est composé de deux syntagmes intermédiaires [*It's eleven and nine and one*] et [*and eighty*].

Dans la figure 28, le premier syntagme intermédiaire « *it's eleven and nine and one* » est séparé du deuxième syntagme intermédiaire « *and eighty* » qui est un élément focalisé par un accent de syntagme L- marquant la frontière de syntagme intermédiaire. Cet accent de syntagme signale la présence d'une frontière prosodique qui serait responsable du blocage du *downstep* à la frontière de l'ip et expliquerait le fait qu'une frontière prosodique soit perçue avant l'élément focalisé (« *and EIGHTY* »). Il semble que cette frontière prosodique soit ressentie par des auditeurs de l'anglais comme une rupture moins marquée que celle perçue lors de la réalisation d'une frontière de syntagme intonatif telle que l'on pourrait l'observer dans la phrase 1 (Beckman & Pierrehumbert, 1986).

1. *But it doesn't make any sense ? Why did she do it ?*

- *Well I think she did it % for JOHN'S sake.*

Mais ça n'a pas de sens? Pourquoi a-t-elle fait ça?

- Eh bien, je crois qu'elle l'a fait % pour l'amour de JOHN.

Dans la phrase 1, l'élément focalisé « JOHN » est précédé par une rupture de syntagme intonatif dans le but de rendre cet élément plus saillant.

Une des particularités de l'accent de syntagme en anglais (à la différence du japonais) est qu'il a la possibilité de s'étendre à sa gauche jusqu'à la dernière syllabe accentuée ce qui est référencé dans la littérature sous le nom de *spreading*. Notons que cette capacité que peut avoir l'accent de syntagme à s'étaler est surtout vraie pour l'accent de syntagme L-. La transition entre un accent mélodique H* et un accent de syntagme L- est schématisée dans la figure 29.

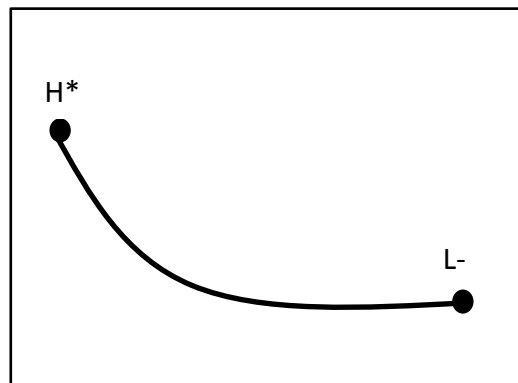


Figure 29. Schématisation de la transition entre un accent mélodique et un accent de syntagme L-

Tiré de Pierrehumbert & Beckman (1986).

Dans cette figure nous pouvons voir que l'accent de syntagme L- se déploie jusqu'à l'accent mélodique L- ce qui donne lieu à un phénomène de plateau bas.

Des phénomènes de réinitialisation du registre après une frontière sont donc utilisés dans les langues afin de marquer les constituants prosodiques. Dans le modèle de l'intonation de l'anglais américain proposé par Pierrehumbert (1980) et Beckman & Pierrehumbert (1986), des phénomènes d'abaissement des accents mélodiques et de réinitialisation du registre sont observées relations avec les frontières prosodiques. En 1988, Ladd met en évidence que certaines structures syntaxiques particulières vont donner lieu à des phénomènes de variation de registre non pas à l'intérieur des constituants prosodiques mais entre ces constituants. Pour cela, il compare des séquences présentant deux types de structure syntaxique où A, B et C sont des clauses similaires du point de vue de leur structuration rythmique et syntaxique. Les deux types de structure utilisés sont illustrés des phrases 2a et 2b dont la structure syntaxique correspond respectivement aux figures 30a et 30b.

2a. Allen is a strong campaigner, and Ryan has more popular policies, but Warren has a lot of money.

Allen est un solide compétiteur lors des campagnes, et Ryan a une politique plus populaire, mais Warren a beaucoup d'argent.

2b. Ryan has a lot more money, but Warren is a stronger campaigner, and Allen has more popular policies.

Ryan a beaucoup plus d'argent, mais Warren est un solide compétiteur lors des campagnes, et Allen a une politique plus populaire.

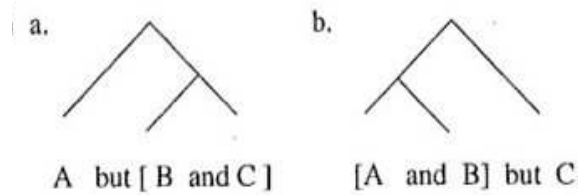


Figure 30. Schématisation de la structure grammaticale des phrases utilisées par Ladd (1988).

A, B et C sont des clauses similaires du point de vue de leur structuration rythmique et syntaxique

Ladd observe que les valeurs de *fo* observées sur le mot correspondant à la clause B sont plus élevées lorsque ce mot est suivi par la conjonction « but » que par la conjonction « and ». Celui-ci en déduit que les structures induites par les deux types de conjonction se reflètent dans l'organisation mélodique des énoncés. Ladd propose de modéliser ces résultats induits par les conjonctions « but » et « and » en postulant que les constituants A, B et C vont être organisés mélodiquement les uns par rapport aux autres avec des phénomènes d'abaissement et de réinitialisation similaires à ce qui a été observé par Beckman & Pierrehumbert (1986) à l'intérieur des constituants prosodiques. Dans le modèle proposé par Ladd (1988), l'abaissement entre les constituants est modélisé par des objets de registres abstraits hauts (h) et bas (l) tel que cela est illustré dans la figure 31.

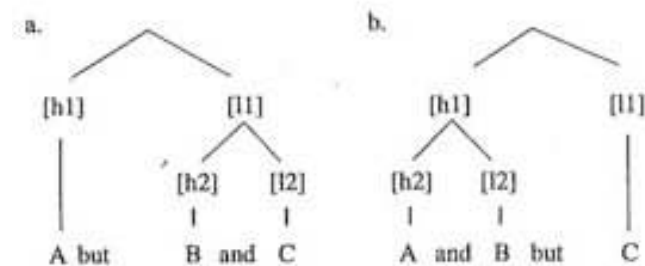


Figure 31. Schématisation de l'abaissement entre les constituants prosodiques observés par Ladd (1988).

Figure tirée de Truckenbrodt et Féry (2005, p°224).

Dans cette figure, nous pouvons observer que l'organisation globale du registre de chaque constituant dépend des autres constituants prosodiques de l'énoncé. Les phénomènes d'abaissement et de réinitialisation observés entre les constituants prosodiques sont liés à des phénomènes linguistiques et non à des facteurs paralinguistiques tel que cela avait été postulé par Beckman et Pierrehumbert (1986). Le modèle proposé par Ladd permet au phénomène d'abaissement observé à l'intérieur des constituants prosodiques d'être emboîté au sein d'un phénomène d'abaissement plus large qui a lieu non plus entre les accents mélodiques mais entre les constituants prosodiques. Cette modélisation permet à l'auteur de mettre en évidence des phénomènes de

réinitialisation partielle du registre après une frontière prosodique (*partial reset*) c'est-à-dire le retour à des valeurs de *fo* qui ne sont pas aussi hautes que celles attendues. Ce phénomène de réinitialisation du registre observé pour l'anglais par Ladd a été confirmé en hollandais par une étude de van der Berg, Gussenhoven & Rietveld. (1992). Le phénomène de réinitialisation partielle du registre est illustré dans la figure 32.

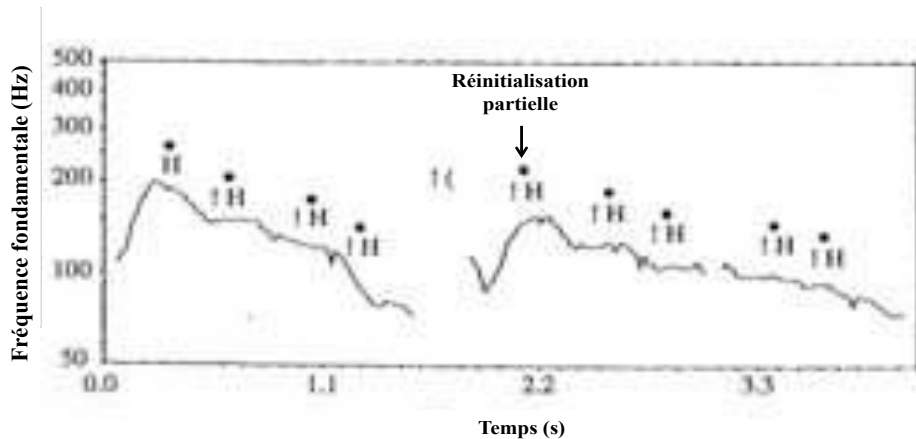


Figure 32. Phénomène de réinitialisation partielle du registre observé en hollandais. Illustration de la phrase « Merel, Nora, Leo, Remy en Nelie, Mary, Meendert, Mona en Lorna ». Figure tirée de van den Berg, Gussenhoven & Rietveld (1992), p° 334.

Dans cette figure, nous pouvons observer des phénomènes d'abaissement des accents mélodiques situés à l'intérieur des deux constituants prosodiques. Par exemple l'accent mélodique sur le mot « Nora » est abaissé relativement à l'accent mélodique associé au mot « Merel ». Le registre est réinitialisé sur le mot « Nelie », mais on peut observer que les valeurs de *fo* associées à cet accent mélodique ne sont pas aussi hautes que celles observées sur le premier accent mélodique de l'énoncé. Dans cette étude, van der Berg et al. (1992) testent empiriquement le modèle proposé par Ladd. En plus des propositions faites par celui-ci, cette étude menée sur le hollandais permet de mettre en évidence deux faits nouveaux importants. Tout d'abord, le *downstep* à l'intérieur des constituants prosodiques est plus large que le *downstep* observé entre les constituants prosodiques. Deuxièmement, le phénomène d'abaissement est réalisé au moyen d'une ligne abstraite de référence du registre qui serait marquée par le premier pic de *fo* du domaine prosodique initial. Le *downstep* entre les constituants prosodiques mis en évidence par Ladd est implémenté dans le modèle proposé par van der Berg et al. (1992) par l'abaissement de la ligne de registre de chaque constituant prosodique. Lorsqu'un constituant prosodique est abaissé par rapport au constituant qui précède, c'est en fait la ligne de référence du registre, marquée par le premier pic mélodique de ce constituant, qui est abaissée. Lorsque l'abaissement entre les constituants prosodiques est rompu, c'est la ligne de référence du registre

du dernier constituant qui est réinitialisée. Le phénomène d'abaissement entre les constituants est illustré dans la figure 33.

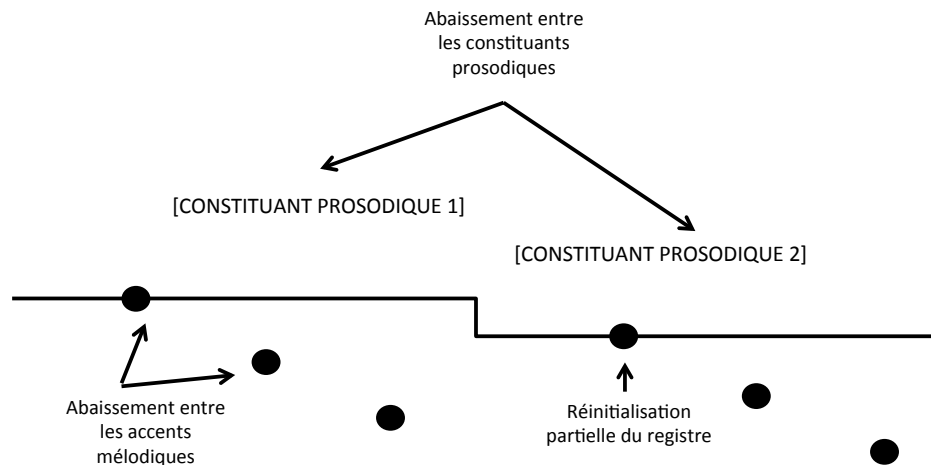


Figure 33. Schématisation de l'abaissement des accents mélodiques et de l'abaissement entre les constituants prosodiques.
D'après le model proposé par Ladd (1988) et complété par van der Berg et al. (1992).

Dans cette figure, nous pouvons voir un abaissement des accents mélodiques au sein des constituants prosodiques et un phénomène d'abaissement plus global au niveau des constituants prosodiques. Conformément aux prédictions de ven der Berg et al. 1992, la ligne de référence du registre du deuxième constituant prosodique de l'énoncé est abaissée relativement au premier. Cette ligne de référence du registre est marquée par le premier pic mélodique du constituant prosodique. Nous pouvons également remarqué un phénomène de réinitialisation partielle du registre sur l'accent mélodique initial du deuxième constituant prosodique. Ce phénomène de réinitialisation partielle est induit par la présence d'une frontière prosodique qui bloque l'abaissement entre les accents mélodiques. Cependant, cette réinitialisation du registre n'est pas totale mais partielle : les valeurs de *f₀* observées sont plus basses que les valeurs de *f₀* observées sur le premier accent mélodique du premier constituant prosodique de l'énoncé. Dans la figure 33, nous pouvons également observer que le *downstep* entre les accents mélodiques est plus important que le *downstep* entre les deux constituants prosodiques.

En 2002, Truckenbrodt a utilisé le modèle initialement proposé par Ladd (1988) puis développé par van der Berg (1992) pour décrire un phénomène intonatif en allemand. L'auteur met en évidence des phénomènes de réinitialisation du registre qu'il appelle *final upstep* observé dans certaines variétés d'allemand du sud. Les accents mélodiques sont abaissés au sein du syntagme intonatif en allemand mais un phénomène de réinitialisation du

registre a lieu sur le pic nucléaire et/ou sur le ton de frontière H% du syntagme intonatif et donc juste avant la frontière et non pas après. Le modèle proposé par Truckenbrodt (2002) pour rendre compte du système intonatif de l'allemand est schématisé dans la figure 34.

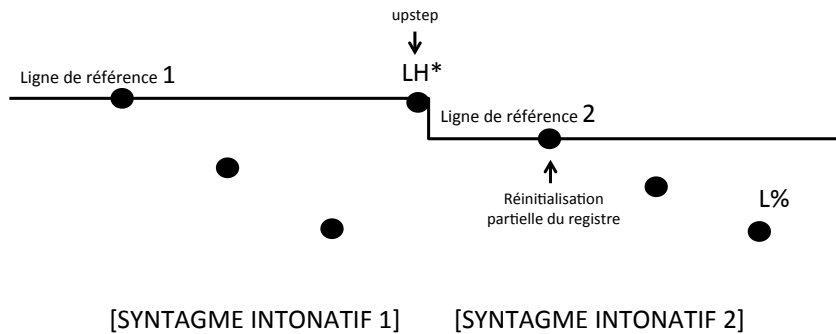


Figure 34. Schématisation du modèle proposé par Truckenbrodt (2002) pour l'allemand. Adapté de Ladd (1988) et van der Berg et al. (1992).

La courbe de *fo* de la phrase « [Der Maurer un sein Lehrling wollen der Werner in Kamerun en Lama malen] [und der Maler will im Janner in Murnau wohnen] » est représentée dans la figure 35. Dans cette phrase nous pouvons observer les phénomènes de réinitialisation totale du registre (*final upstep*) observée avant la frontière de syntagme intonatif et de réinitialisation partielle du registre (*partial reset*) décrits par Truckenbrodt.

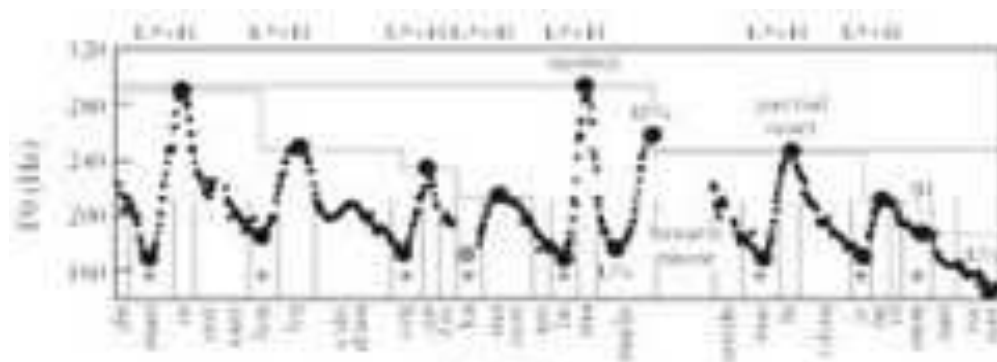


Figure 35. Illustration du phénomène de *final upstep* et de réinitialisation partielle du registre.

Courbe de *fo* de la phrase [Der Maurer un sein Lehrling wollen der Werner in Kamerun en Lama malen] [und der Maler will im Janner in Murnau wohnen]. Tiré de Truckenbrodt (2002).

Enfin en 2005, Truckenbrodt et Féry ont confirmé le fait que l'abaissement entre les accents mélodiques est plus important que l'abaissement entre les constituants prosodiques pour l'allemand.

A la lumière des modèles proposés dans le but de modéliser les effets d'abaissement entre accents mélodiques et entre constituants prosodiques pour l'anglais, le hollandais et l'allemand, nous allons maintenant étudier la structure prosodique du français. Dans l'étude qui va suivre, nous cherchons à déterminer dans quelle mesure les phénomènes d'abaissement entre accents mélodiques et entre constituants prosodiques peuvent nous aider à comprendre le marquage des frontières prosodiques en français.

2. HYPOTHÈSES

Au cours du chapitre 2, nous avons vu que l'alignement entre une frontière droite de syntagme accentuel et une rupture syntaxique majeure donne lieu à un degré d'allongement de la syllabe pré-frontière plus important que celui associé aux syllabes en position finale d'AP contenue au sein d'un constituant syntaxique majeur. Par exemple dans des phrases telles que celles présentées en 3a et 3b, la durée de la syllabe cible (en lettres capitales) était plus importante dans la phrase 3b que dans la phrase 3a en raison de l'association entre frontière droite d'AP et frontière droite de constituant syntaxique majeur.

3a. Les a MIS de Pauline adoraient ce restaurant.		
Structure syntaxique :		]SN
Structure prosodique :	]AP	]AP
3b. Tes a MIS demandaient la présidente.		
Structure syntaxique :		]SN
Structure prosodique :	]AP	

Dans l'étude que nous présentons maintenant, nous avons manipulé la taille du syntagme nominal dans le but d'obtenir des phrases dont le syntagme nominal contenait soit 2 APs (4a), soit 3 APs (4b).

4a. La mamie de Ré <u>MY</u> demandait l'institutrice.		
Structure syntaxique :		]SN
Structure prosodique :	]AP	]AP
4b. La mamie des a <u>MIS</u> de Rémy demandait l'institutrice.		
Structure syntaxique :		]SN
Structure prosodique :	]AP	]AP

En raison de l'alignement entre la frontière droite du syntagme accentuel et une rupture syntaxique majeure et grâce à des syntagmes nominaux suffisamment longs (composés soit de 2 APs soit de 3APs), nous nous attendions à voir émerger, dans ces deux phrases, une frontière de syntagme intermédiaire alignée avec la rupture entre le syntagme nominal et le syntagme verbal. En effet, une tendance à placer des frontières prosodiques majeures à la frontière de constituants syntaxiques majeures (comme par exemple une rupture entre un syntagme nominal, SN, et un syntagme verbal, SV) a été démontrée pour le français bien que cette tendance soit influencée par des contraintes eurythmiques (Delattre 1966 ; Delais-Roussarie, 1995 ; Rossi, 1997 ; Vaissière & Michaud, 2006). Nous pouvons émettre l'hypothèse que la frontière prosodique présente à la rupture entre le syntagme nominal et le syntagme verbal soit souvent associée à une frontière de niveau intermédiaire dans la hiérarchie prosodique (situé entre l'AP et l'IP).

Rappelons que Jun & Fougeron (2000) avait déjà proposé pour le français, un niveau additionnel de structuration prosodique situé entre l'AP et l'IP et appelé syntagme intermédiaire (*intermediate phrase*, ip). Cependant, dans ce modèle, le syntagme intermédiaire est postulé dans le but de rendre compte de contours intonatifs spécifiques : l'existence de régions plates de *fo* qui peuvent être hautes, basses ou médianes et qui sont en étroite relation avec des structures syntaxiques spécifiques telles que des questions tags, dislocations, questions à morphème interrogatif initial (*wh-questions*) et clivées. Dans le cadre du modèle de Jun & Fougeron, ces plateaux sont analysés comme des séquences L-L% (dans le cas d'un plateau bas) ou H-H% (dans le cas d'un plateau haut) ou H-L% (dans le cas d'un plateau médian) où L- et H- sont des accents de syntagme marquant la frontière du syntagme intermédiaire. Une des propriétés de l'accent de syntagme associé à la frontière d'ip est de s'étendre sur la gauche ce qui donne lieu à des phénomènes de plateau. Dans l'hypothèse que nous faisons, et contrairement à ce qui a été proposé par Jun & Fougeron, le syntagme intermédiaire ne serait pas restreint à des structures syntaxiques spécifiques mais pourrait apparaître sur tout énoncé dès lors que les contraintes syntaxiques et prosodiques le permettent.

Notre hypothèse était que, conformément à ce qui était observé en anglais, le syntagme intermédiaire pourrait être le domaine de l'abaissement des accents mélodiques en français. Par contre et de façon similaire à l'allemand, nous nous attendions donc à ce que le registre soit réinitialisé sur le dernier accent mélodique de l'ip. Au regard des résultats obtenus lors de l'étude 1 menée précédemment, nous nous attendions également à ce que la syllabe située en position finale d'ip soit plus longue que la syllabe en position finale d'AP. Enfin, sur le modèle de ce qui a été observé pour le hollandais et pour l'allemand, nous nous attendions à ce qu'un phénomène d'abaissement moins important que celui qui intervient entre les accents mélodiques au sein de l'ip affecte la ligne de référence du registre des syntagmes intermédiaires.

Dans des phrases telles que celles illustrées en 4a et 4b où les syllabes cibles étaient toujours les syllabes finales du deuxième syntagme accentuel et pouvaient être associées soit à une frontière d'ip (SN composés de 2 APs), soit à une frontière d'AP, nos prédictions étaient les suivantes (voir également figure 36 suivante):

- En raison de l'association de la syllabe avec la frontière d'ip, les syllabes cibles seraient plus longues dans la condition 2 APs que dans la condition 3 APs.
- En raison du blocage du *downstep* des accents mélodiques que nous nous attendons à observer à la frontière droite d'ip, le registre serait réinitialisé sur la syllabe associée à la frontière d'ip. De ce fait, les valeurs de *f₀* seraient plus hautes dans la condition 2 APs que dans la condition 3 APs.
- Enfin, en raison d'un phénomène d'abaissement entre les lignes de référence du registre des ips, nous nous attendons à ce que le premier accent mélodique qui suit la frontière d'ip ne soit pas aussi haut que l'accent mélodique associé à la frontière d'ip. De ce fait nous nous attendons à observer des valeurs de *f₀* de la syllabe cible plus basse sur l'accent mélodique qui suit la frontière d'ip (sur « demandait » dans la figure 36) que sur l'accent mélodique associé à la frontière d'ip (sur « my » dans la figure 36).

Nos prédictions sont schématisées dans la figure suivante pour les phrases dont le SN sujet était composé de 2 APs (36a) et pour les phrases dont le SN sujet était composé de 3 APs (36b).

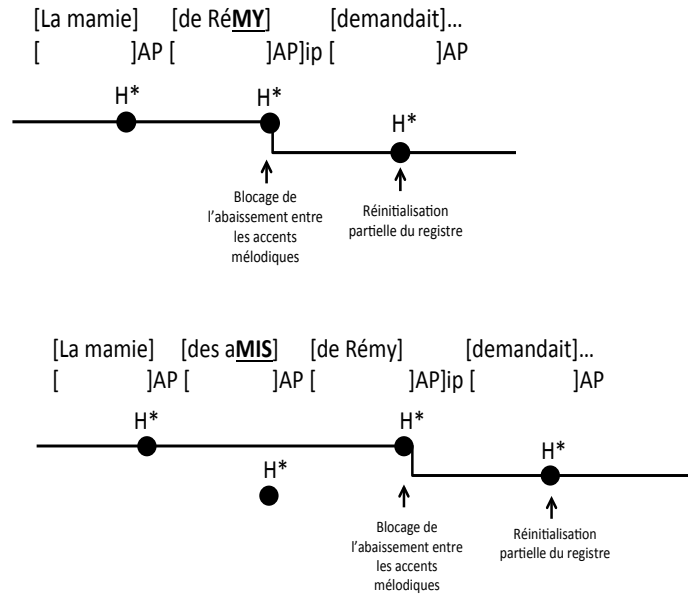


Figure 36. Étude 2 : Schématisation de la structure prosodique attendue.
Structure prosodique attendue pour les phrases dont le SN sujet était composé soit de 2 APs (en haut), soit de 3 APs (en bas).

3. MÉTHODE

3.1. Matériel linguistique

Un corpus de huit phrases lues à quatre reprises, à deux vitesses d'élocution, par quatre locuteurs de langue maternelle française, a été construit pour un total de 256 énoncés (8 phrases x 4 répétitions x 2 vitesses d'élocution x 4 locuteurs). Dans ce corpus nous avons analysé les propriétés acoustiques (durée segmentale et valeurs de *fo*) des voyelles et syllabes cibles associées à la frontière du deuxième syntagme accentuel de chaque énoncé (syllabes en lettres capitales dans la figure 36). Nous avons fait varier la longueur du syntagme nominal en terme de nombre de mots dans le but d'obtenir des syntagmes nominaux Sujets composés soit de deux syntagmes accentuels (phrase 5a), soit de trois syntagmes accentuels (phrase 5b).

5a. [La mamie]AP [de RéMY]AP demandait l'institutrice.

5b. [La mamie]AP [des aMIS]AP [de Rémy]AP demandait l'institutrice.

Dans la phrase 5a, la syllabe cible est alignée avec une rupture syntaxique majeure entre un syntagme nominal et un syntagme verbal. De ce fait nous supposons qu'elle soit associée à une frontière de syntagme intermédiaire (ip). Au contraire dans la phrase 5b, la syllabe cible est contenue à l'intérieur du syntagme nominal et est donc associée avec une frontière de syntagme

accentuel à l'intérieur d'un syntagme intermédiaire. Les syllabes cibles pouvaient donc être associées à deux types de frontière prosodique : soit à une frontière de syntagme intermédiaire, soit à une frontière de syntagme accentuel. Les syllabes cibles étaient toutes des syllabes ouvertes (de type CV). Elles étaient composées d'une voyelle antérieure non arrondie haute /i/ pour la moitié des phrases et d'une voyelle basse /a/ pour l'autre moitié. La consonne précédant la voyelle était toujours voisée mais son mode d'articulation pouvait varier (consonne liquide /l/ ou /r/ ou nasale /m/ ou /n/). Dans le but de faciliter l'analyse de la fréquence fondamentale (*f₀*), la consonne qui suivait la syllabe cible était toujours une consonne voisée /d/. L'ensemble des phrases utilisées dans le corpus est illustré dans le tableau 8.

	Voyelle	SN sujet contenant 2APs	SN sujet contenant 3APs
1	haute	La mamie de Ré <u>MY</u> demandait l'institutrice.	La mamie des amis de Ré <u>MY</u> demandait l'institutrice.
2	haute	Le mari de Car <u>RRY</u> deviendra un grand docteur.	Les carries du mari de Ca <u>RRY</u> demandaient des soins urgents.
3	basse	Le sauna d'Hél <u>éNA</u> deviendra la plus connu.	La nana du sauna d'Hél <u>éNA</u> devenait vraiment méchante.
4	basse	Le trois-mâts de Thom <u>MAS</u> devance tous les autres bateaux.	Le schéma du trois-mâts de Tho <u>MAS</u> devenait vraiment brouillon.

Tableau 8. Étude 2 : Phrases utilisées dans le corpus.
Les noyaux vocaliques des syllabes cibles (en lettres capitales) pouvaient être constitués d'une voyelle haute ou basse. Les syllabes cibles pouvaient être associées soit à une frontière de syntagme intermédiaire (SN sujet contenant 2APs) soit à une frontière de syntagme accentuel (SN sujet contenant 3 APs).

Un exemple de la structure prosodique des phrases obtenues pour les deux types de SN (SN Sujet contenant 2 APs et SN Sujet contenant 3 APs) est illustré dans la figure 37.

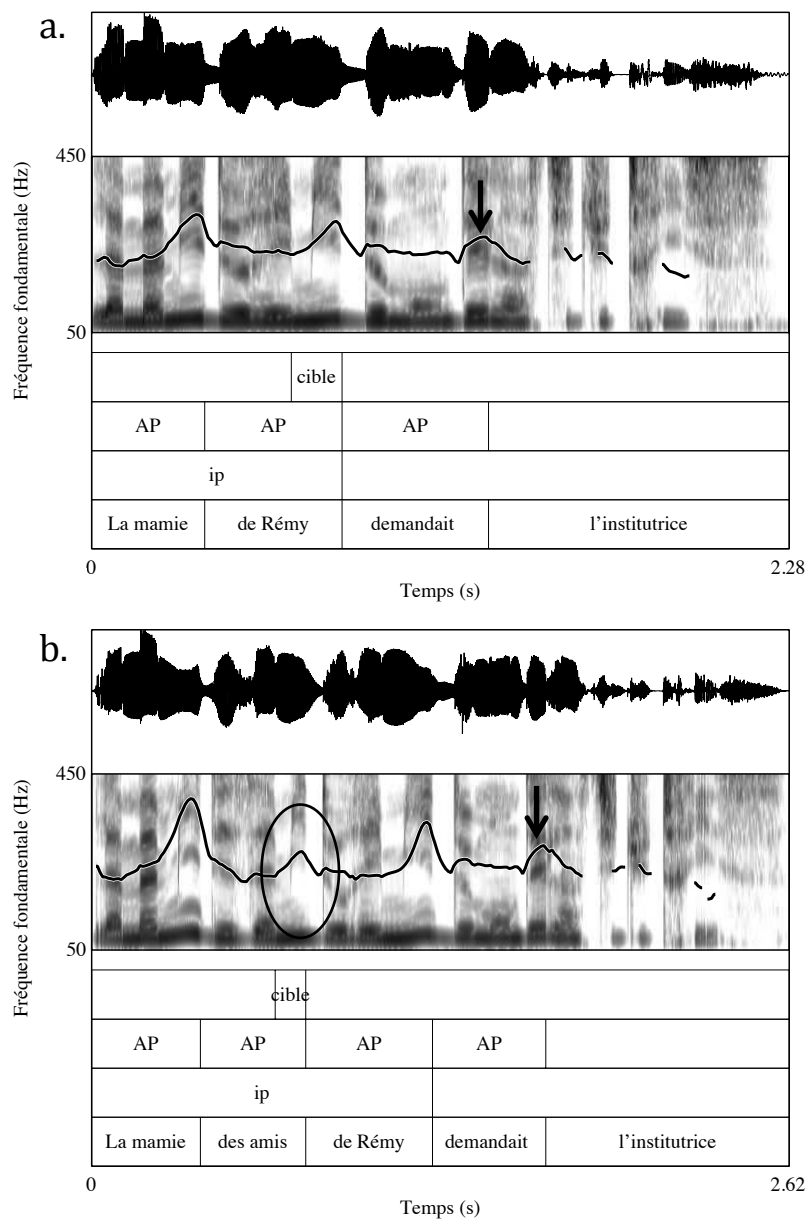


Figure 37. Étude 2 : Structures prosodiques obtenues.

Phrases « La mamie de Rémy demandait l'institutrice » où le SN Sujet est composé de deux APs (37a) et « La mamie des amis de Rémy demandait l'institutrice » où le SN sujet est composé de 3 APs (37b).

16 phrases distractrices ont été ajoutées aux 8 phrases expérimentales. Ces phrases ont été choisies au hasard parmi les phrases distractrices utilisées lors de l'étude 1. Elles ne présentaient aucune structure syntaxique ou prosodique particulière. Deux listes expérimentales ont été construites de manière à ce que les participants ne puissent pas voir apparaître deux fois de suite le même mot cible. Chaque liste contenait les deux types de syntagmes nominaux (composés de 2 APs et composés de 3

APs) pour un total de quatre phrases expérimentales et 8 phrases distractrices. Les phrases distractrices étaient différentes dans les deux listes expérimentales. Tous les participants ont été soumis aux deux listes expérimentales à 8 reprises dans un ordre randomisé pour chaque participant (quatre fois avec pour la vitesse d'élocution normale et quatre fois pour la vitesse d'élocution rapide). A l'intérieur de chaque liste, les énoncés étaient également randomisés.

3.2. Participants et procédure expérimentale

4 locuteurs de langue maternelle française (deux hommes et deux femmes) ont lu les phrases randomisées à 4 reprises en vitesse d'élocution normale et à quatre reprises à vitesse d'élocution rapide. Nous savons que la vitesse d'élocution affecte de manière significative le découpage en constituants prosodiques, c'est pourquoi nous avons manipulé celle-ci (cf. chapitre 2). Les participants étaient âgés de 22 à 29 ans avec une moyenne d'âge de 25.5. Ils étaient originaires de différentes régions de France et étaient tous titulaires du baccalauréat. Nous les avons confortablement installés dans la chambre sourde du Laboratoire Parole et Langage. Ils étaient assis face à un écran d'ordinateur, équipés d'un micro casque. Nous leur avons demandé de lire à voix haute les phrases, d'abord avec une vitesse d'élocution normale, puis avec une vitesse d'élocution rapide. Les consignes exactes présentées aux participants sur l'écran d'ordinateur sont données en annexe. Les phrases étaient présentées une par une sur l'écran. Elles étaient orthographiées de manière classique en incluant les lettres majuscules et la ponctuation. Une fois la phrase lue, les participants devaient cliquer sur un bouton pour faire apparaître la phrase suivante. La procédure expérimentale a été contrôlée grâce au logiciel DMDX (Forster & Forster, 2000). Quatre phrases d'entraînement similaires aux phrases expérimentales et distractrices ont été présentées au sujet avant que l'expérimentation ne commence. Ces phrases tests ont également été présentées aux participants au début de la phase de lecture rapide. L'expérience durait environ quinze minutes et se déroulait en deux phases (une phase de lecture normale et une phase de lecture rapide) avec quelques minutes de pause entre les deux phases.

3.3. Recueil et analyse des données

Les enregistrements ont été réalisés dans la chambre sourde du laboratoire Parole et Langage grâce à une station d'enregistrement ZOOM HD16 en format wav 16bit/44.1 Hz. Les fichiers sons ont été segmentés manuellement et chaque fichier son a été sauvé séparément. Tous les fichiers ont été segmentés en phonèmes à l'aide de l'aligneur Easy Align (Goldman, 2007). La segmentation des syllabes et voyelles cibles a été vérifiée manuellement grâce à l'inspection des ondes sonores et des spectrogrammes.

Les courbes de *f₀* ont été générées grâce au logiciel PRAAT (Boersma & Weenik, 2009). Les voyelles et les syllabes cibles ont été annotées automatiquement grâce à des scripts Praat puis vérifiées manuellement une par une. Nous avons respectivement codé :

- V1 et S1 : durée de la voyelle et de la syllabe en position finale du premier syntagme accentuel de l'énoncé
- V2 et S2 : durée de la voyelle et de la syllabe en position finale du deuxième syntagme accentuel de l'énoncé qui pouvait être associé soit à une frontière d'AP (phrases dont le SN Sujet contenaient 3 APs) soit à une frontière d'ip (phrases dont le SN Sujet contenant 2 APs)
- V3 et S3 : durée de la voyelle et de la syllabe du troisième syntagme accentuel pour les phrases dont le SN sujet était composé de 3 APs.

Nous avons étiqueté automatiquement les valeurs maximales de *f₀* associées à chaque syllabe annotée (respectivement codées H1, H2 et H3). Nous avons également codé Hpb (*post boundary H*) le maximum de *f₀* situé après la frontière d'ip. La figure 38 illustre l'étiquetage réalisé pour les deux types de syntagme nominaux (SN composés de 2 APs et SN composés de trois APs).

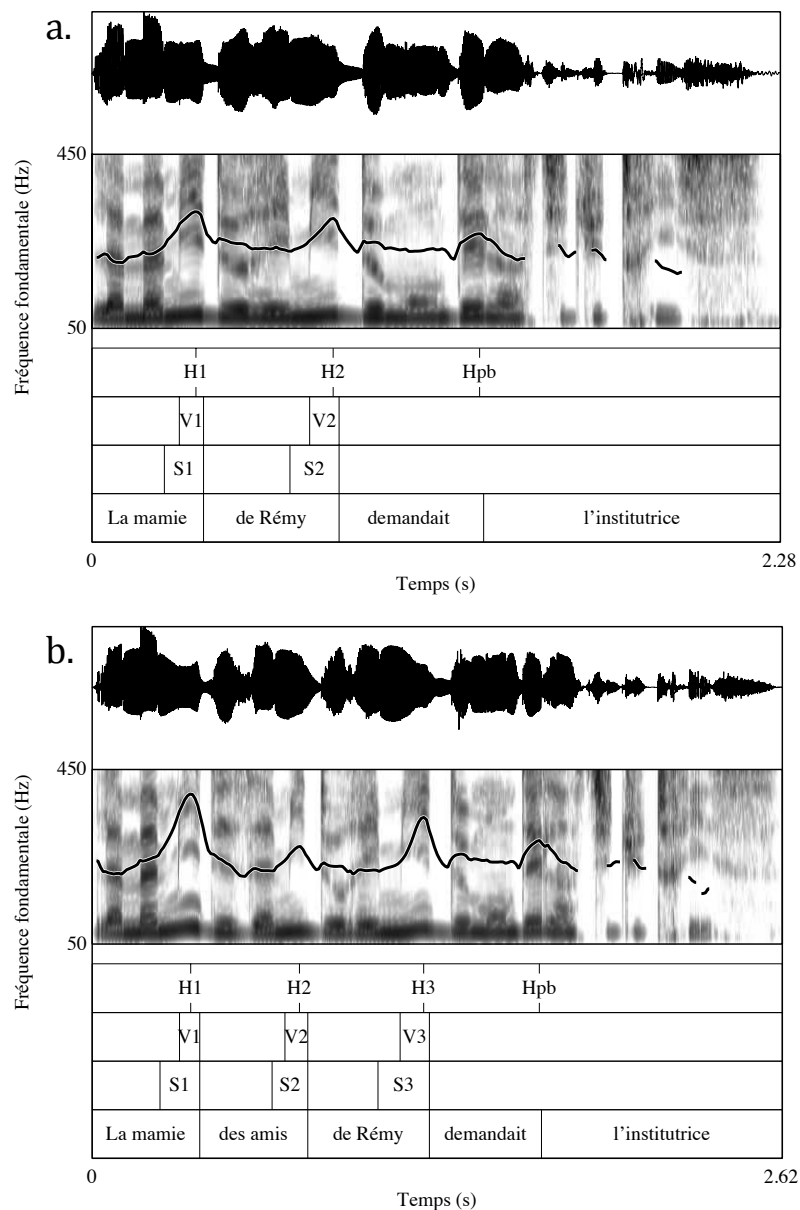


Figure 38. Étude 2 : Exemple de l'annotation réalisée.
Phrase dont le SN Sujet contenait 2 APs (38a) et phrase dont le SN Sujet contenait 3 APs (38b).

La durée segmentale des voyelles et des syllabes étiquetés, ainsi que la durée des énoncés, ont été extraites grâce à l'utilisation de scripts PRAAT qui sont consultables en ligne à l'adresse suivante: <http://amandine.michelas.pagesperso-orange.fr/Praat.htm>. Les valeurs maximales de f_0 associées à V1, V2 et Vpb ont également été annotées automatiquement puis extraites grâce à l'utilisation de scripts.

L'intonation de chaque énoncé a été annotée par nous-même et un deuxième annotateur. Cette annotation a été réalisée dans le but de voir si la structure prosodique attendue (SN composés de 3APs et SN composés de 3

APs) a effectivement été produite par les locuteurs et ce pour les deux vitesses d'élocution. Le premier annotateur a annoté tous les fichiers qui ont été ensuite vérifiés par le second annotateur. En cas de désaccord, les deux transcripteurs ont réexaminés les fichiers et sont parvenus à un accord. L'annotation de l'intonation a été réalisée en suivant le modèle de l'intonation du français proposé par Jun & Fougeron (1995, 2000, 2002).

4. RÉSULTATS

4.1. Manipulation de la vitesse d'élocution

Dans le but de vérifier si tous les participants ont significativement augmenté leur vitesse d'élocution à débit rapide par rapport au débit normal, nous avons mesuré la vitesse d'élocution pour chaque énoncé en nombre de syllabes par seconde (syl/sec). Nous avons automatiquement extrait la durée de chaque énoncé exprimée en secondes. Nous avons ensuite compté le nombre de syllabes contenues dans chaque énoncé en fonction de la manière dont l'énoncé a été produit par le locuteur. En effet, les énoncés pouvaient contenir des schwas finaux pouvant être élidés ce qui a donné lieu à des variations du nombre de syllabes contenu dans les énoncés en fonction des locuteurs. En fonction de ce nombre ajusté de syllabes par énoncé et par locuteur nous avons calculé la vitesse d'élocution pour chaque énoncé en nombre de syllabes par seconde. Dans le but de voir si le deuxième syntagme accentuel de l'énoncé, que nous appelons *AP cible*, a été affecté de la même manière que les énoncés entiers par la manipulation de la vitesse d'élocution, nous avons calculé la vitesse d'élocution en syllabe par seconde pour les APs cibles. La vitesse d'élocution ainsi obtenue est représentée dans la figure 39 pour les énoncés et 40 pour les APs cibles.

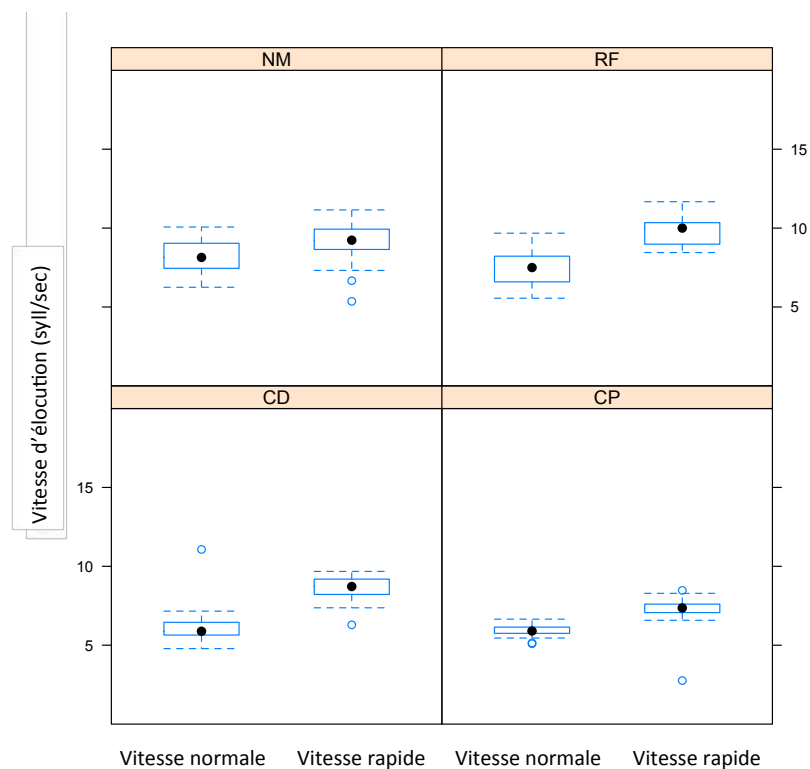


Figure 39. Étude 2 : Vitesse d'élocution pour les énoncés.
La vitesse d'élocution est donnée en nombre de syllabes par sec (Syll/sec) pour les hommes (RF et NM) et pour les femmes (CD et CP).

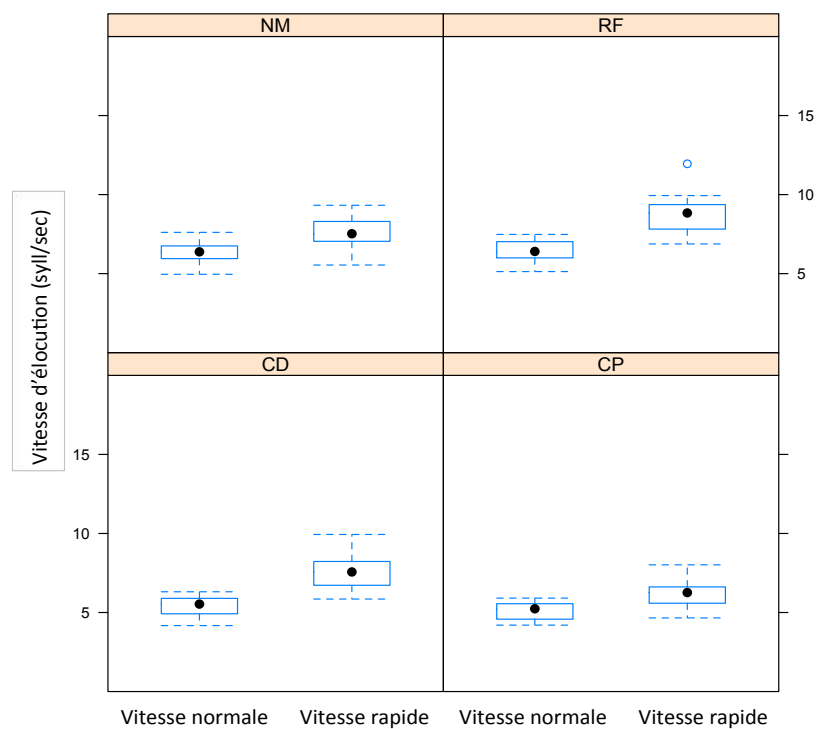
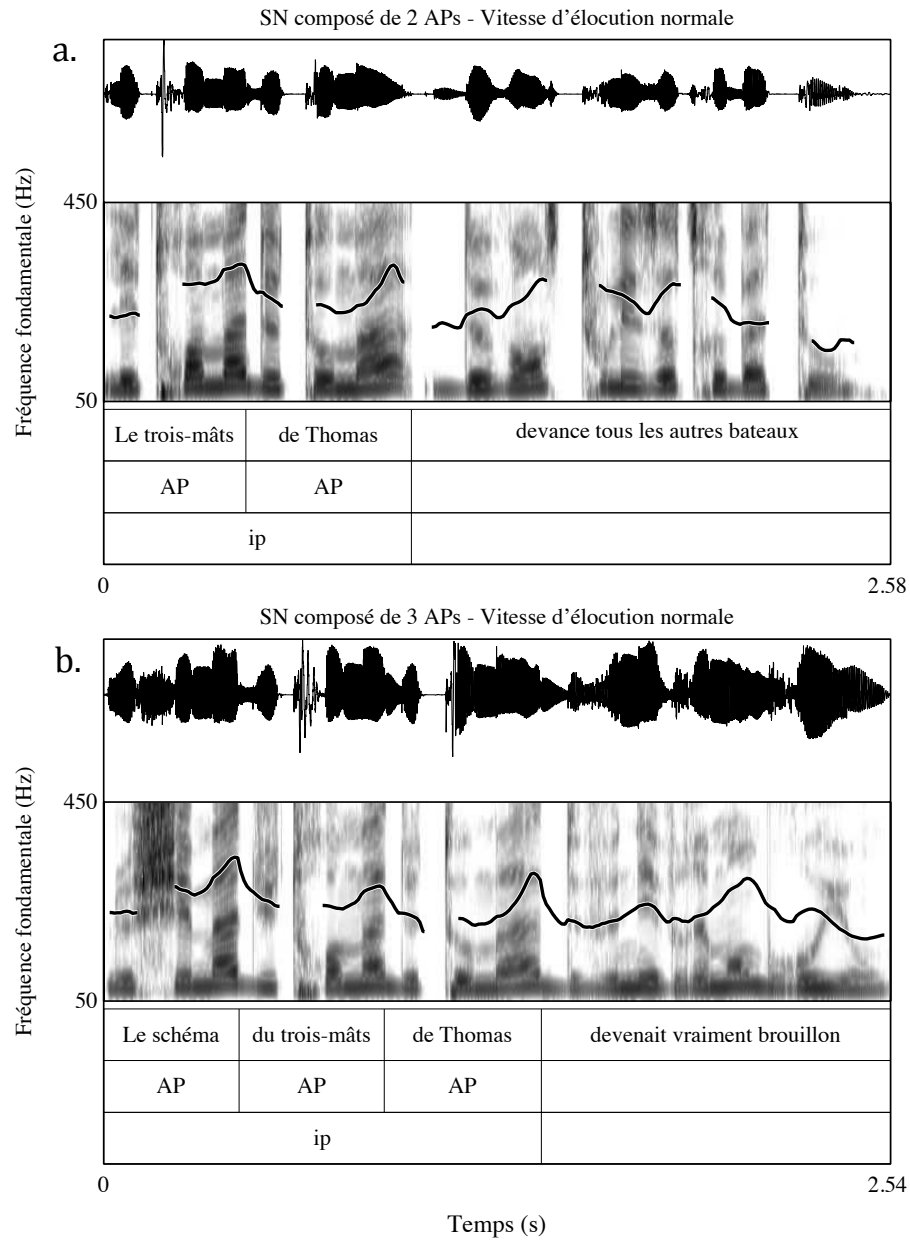


Figure 40. Étude 2 : Vitesse d'élocution pour les APs cibles.
La vitesse d'élocution est donnée en nombre de syllabes par sec (Syll/sec) pour les hommes (RF et NM) et pour les femmes (CD et CP).

Deux modèles linéaires à effets mixtes avec la vitesse d'élocution en nombre de syllabes par seconde comme variable dépendante ont été conduits sur nos données, un pour les énoncés et un pour les APs cible. Un total de 256 mesures a été inclus pour chaque modèle. Ces deux modèles incluait la manipulation de la vitesse d'élocution ainsi que le genre des locuteurs en tant qu'effets fixes. Les locuteurs et les énoncés étaient inclus en tant qu'effets aléatoires. Les deux modèles montrent que la manipulation de la vitesse d'élocution avait un effet significatif sur le nombre de syllabe/sec produites par les locuteurs que ce soit pour les énoncés en entier ou pour les APs cibles (énoncés : $\beta=1.6878$, $se=0.1031$, $t=16.365$, $pMCMC<0.001$; APs cibles : $\beta=1.8958$, $se=0.1737$, $t=10.917$, $pMCMC<0.001$). Les locuteurs ont donc significativement augmenté leur vitesse d'élocution à vitesse rapide par rapport à la vitesse normale. Nous avons également observé un effet significatif du genre des locuteurs avec une vitesse d'élocution plus élevée chez les hommes que chez les femmes (pour les énoncés : $\beta=1.2190$, $se=0.1737$, $t=10.917$, $pMCMC<0.001$; pour les APs cibles : $\beta=1.7723$, $se=0.4114$, $t=4.308$, $pMCMC<0.001$). L'interaction entre les deux effets fixes n'était pas significative (pour les énoncés : $\beta=0.1507$, $se=0.2065$, $t=0.730$, $pMCMC=0.4662$; pour les APs cibles : $\beta=-0.2169$, $se=0.2456$, $t=-0.883$, $pMCMC=0.3883$). La vitesse moyenne d'élocution observée pour le débit normal était comprise entre 5.2 et 7 syll/sec selon les locuteurs. Elle est comparable à ce que nous avons observé lors de l'étude 1 présentée dans cette thèse et également comparable à ce qui a été observé dans d'autres études menées à la fois sur de la parole lue (Fougeron & Jun, 1998 ; Welby & Loevenbruck, 2006) et de la parole spontanée (Malécot, Johnson & Kizziar, 1972 ; Grosjean & Deschamps, 1975) pour le français.

4.2. Phénomènes de restructuration prosodique

L'annotation de l'intonation effectuée sur les 256 énoncés montre qu'à vitesse d'élocution normale, la totalité des énoncés a été produit avec la structure prosodique attendue, c'est-à-dire avec des syntagmes nominaux composés de 2APs pour la première condition et composés de 3APs pour la troisième condition. À vitesse d'élocution rapide, la majorité des énoncés ont été produits avec la même structure prosodique que celle observée à vitesse d'élocution normale. Un exemple typique de la structure prosodique obtenue pour les deux conditions (syntagmes nominaux composés de 2 APs et syntagme nominal composé de 3 APs) et pour les deux vitesses d'élocution est représenté dans la figure 41.



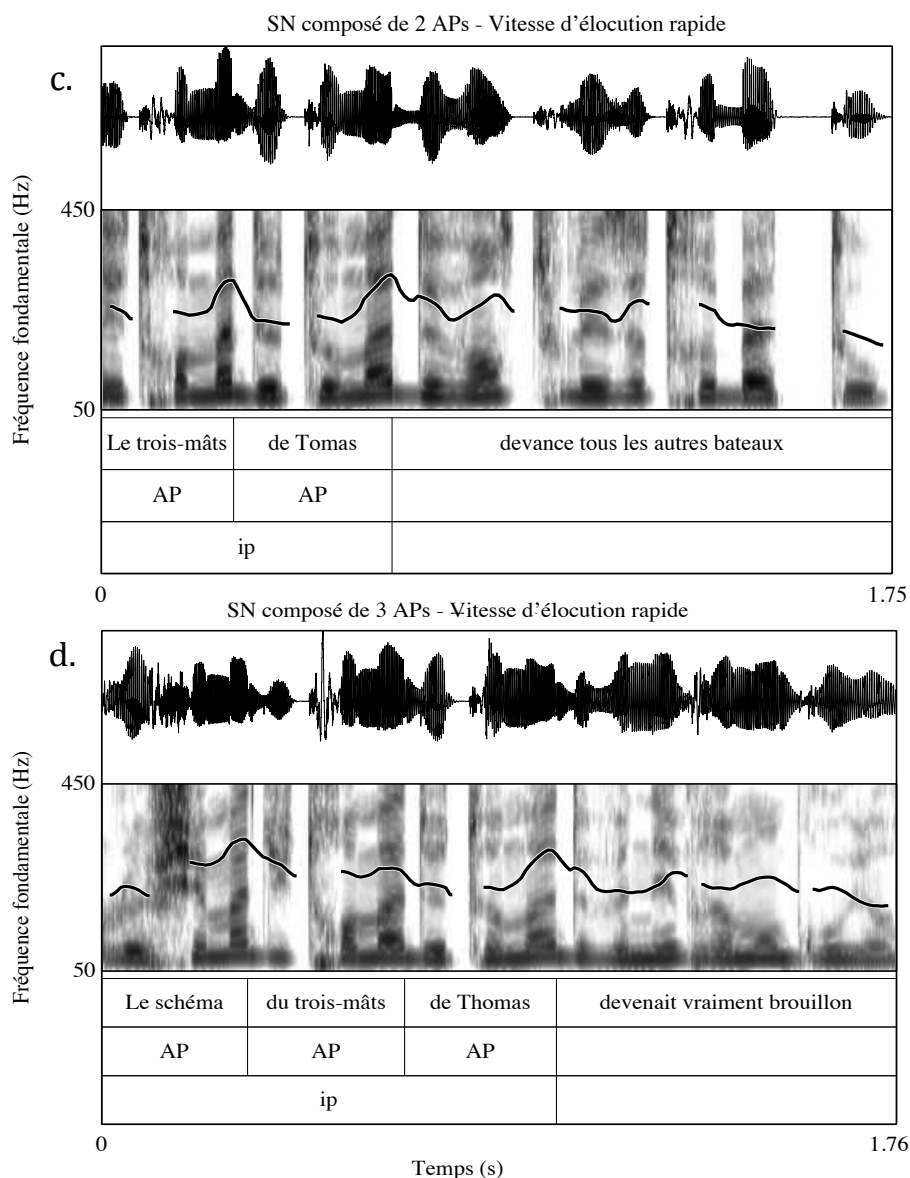


Figure 41. Étude 2 : Structure prosodique obtenue dans les différentes conditions. Structure prosodique pour la phrase « Le trois-mâts de Thomas devance tous les autres bateaux » prononcée à vitesse d'élocution normale (41a) et à vitesse d'élocution rapide (41c) et pour la phrase « Le schéma du trois-mâts de Thomas devenait vraiment brouillon » prononcée à vitesse d'élocution normale (41b) et à vitesse d'élocution rapide (41d).

À vitesse d'élocution rapide, nous avons observé un petit nombre de phénomènes de restructuration prosodique. En effet, dans 12 énoncés c'est à dire pour 4,6% des phrases, nous avons observé des phénomènes d'effacement de frontière d'AP de sorte qu'au lieu d'obtenir 2 APs contenant chacun 3 syllabes (exemple : [la mamie]AP [de Rémy]AP), un seul AP contenant un plus grand nombre de syllabes ait été produit (exemple : [la mamie de Rémy]AP). Un phénomène d'effacement de frontière d'AP est illustré dans la figure 42.

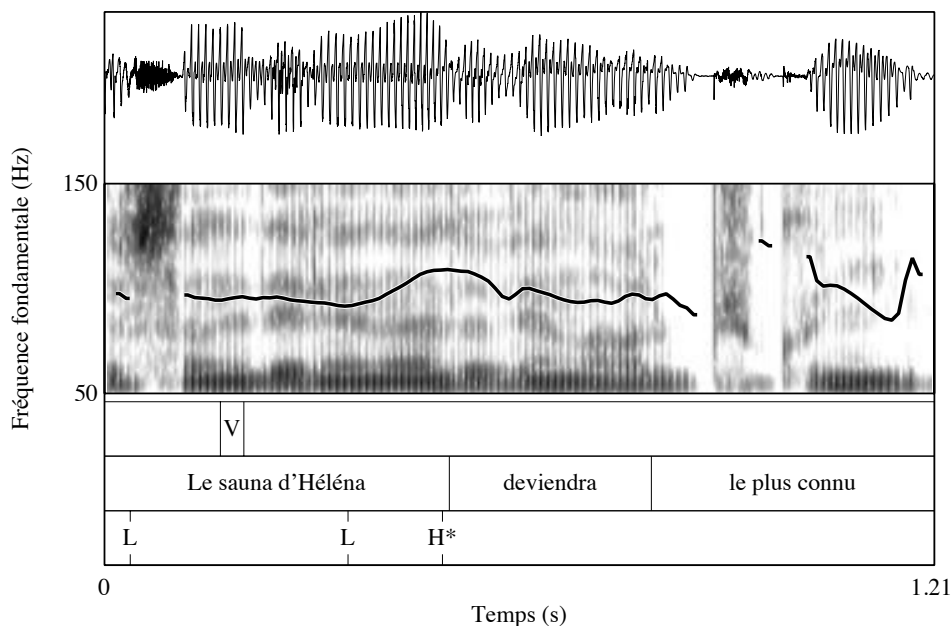


Figure 42. Étude 2 : Illustration du phénomène d'effacement de frontière d'AP sur « sauna » à vitesse d'élocution rapide.

À vitesse d'élocution rapide, nous avons également rencontré un petit nombre de cas (5 énoncés) où, malgré l'absence d'indices tonals marquant la frontière d'AP, l'allongement pré-frontière associé à la syllabe était suffisant pour que les deux annotateurs perçoivent une frontière d'AP. Dans ces cas là et conformément à la position que nous avons adopté dans l'étude précédente nous n'avons pas considéré ces cas comme des effacements de frontière étant donné que la frontière était clairement perçue par les deux annotateurs. Un exemple de ce type de cas est illustré dans la figure 43.

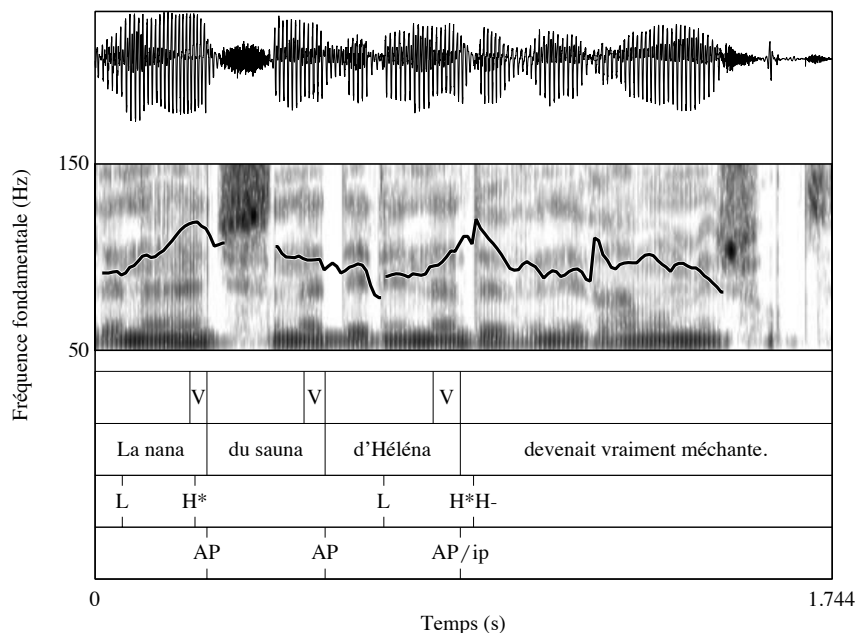


Figure 43. Étude 2 : Illustration d'une frontière d'AP marquée uniquement par des indices de durée.

Malgré l'absence d'indices mélodiques sur « sauna » cette frontière a été clairement perçue par les deux annotateurs.

Etant donné que nos énoncés étaient tous prononcés de façon isolée (ils apparaissaient les uns après les autres sur l'écran de l'ordinateur dès que le participant appuyait sur la barre espace), qu'ils étaient tous constitués d'une clause principale et qu'ils ont tous été produits comme un seul IP, nous n'avons pas observé de phénomène de réduction de frontière c'est à dire de remplacement d'une frontière d'IP par une frontière de plus bas niveau (cf. Fougeron & Jun, 1998).

4.3. Résultats relatifs à la durée

Dans cette section nous allons nous intéresser à la durée de la syllabe en position finale du deuxième syntagme accentuel des énoncés. Rappelons que cette syllabe pouvait être associée soit à une frontière de syntagme intermédiaire (SN composés des 2 APs) soit à une frontière de syntagme accentuel à l'intérieur d'un syntagme intermédiaire (SN composés de 3 APs). Notre hypothèse était que la durée de la voyelle et de la syllabe seraient supérieures dans la condition SN composé de 2 APs que dans la condition SN composé de 3 APs en raison de l'association de la syllabe avec la frontière d'ip dans la première condition. Si de tels résultats étaient obtenus cela nous indiquerait que les indices de durée associés à la frontière d'ip sont significativement plus importants que ceux associés à la frontière d'AP. Cela

renforcerait l'hypothèse de l'existence d'un syntagme intermédiaire situé entre l'AP et l'IP en français.

La durée des voyelles en position finale du deuxième syntagme accentuel (V2) exprimée en millisecondes en fonction du type de syntagme nominal (SN composé de 2APs, SN composé de 3APs) de la vitesse d'élocution (vitesse normale, vn et vitesse rapide, vr) et du genre des locuteurs (hommes et femmes) est représentée dans la figure 44.

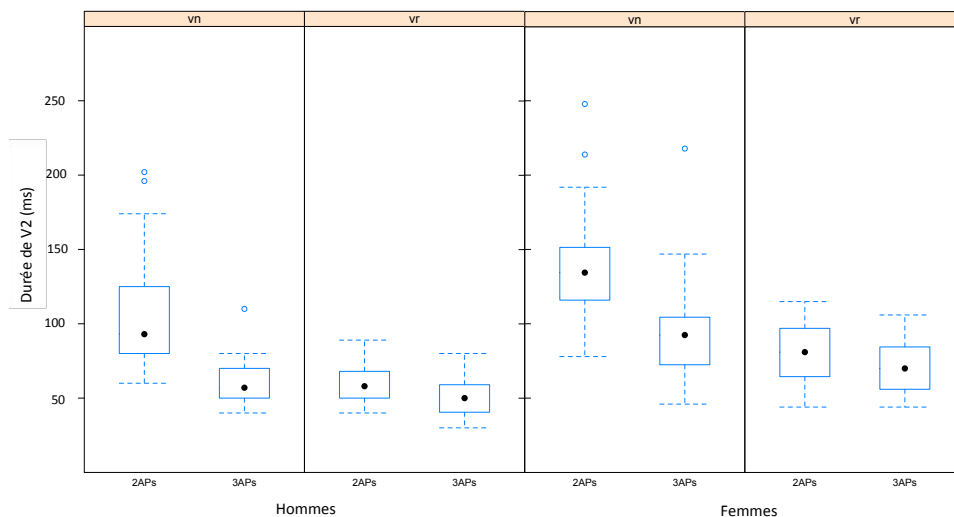


Figure 44. Étude 2 : Durée de la voyelle en position finale du deuxième syntagme accentuel (V2).

La durée est exprimée en millisecondes (ms) en fonction de la vitesse d'élocution (vitesse normale, vn et vitesse rapide, vr) du genre des locuteurs (hommes et femmes) et du type de syntagme nominal (SN composé de 2APs et SN composé de 3APs).

Nous avons conduit un modèle de régressions linéaires multiples à effets mixtes (*Linear Mixed effects Models, LMM*) sur les valeurs logarithmiques des durées des voyelles en position finale du deuxième syntagme accentuel avec le type syntagme nominal (SN composé de 2 APs et SN composé de 3 APs), la vitesse d'élocution (vitesse normale et vitesse rapide) et le genre des locuteurs (homme et femme) en tant qu'effets fixe. Rappelons que le genre des locuteurs influait significativement sur la vitesse d'élocution. Nous avons donc inclus le genre des locuteurs (homme et femme) en tant qu'effet fixe. Les locuteurs (CD, CP, RF, NM) et le type de syllabe (mi, ri, na, ma) ont été inclus en tant qu'effets aléatoires. Le modèle incluait 244 mesures.

Les résultats montrent que quel que soit la vitesse d'élocution, la voyelle est significativement plus longue dans la condition SN composé de 2APs que dans la condition SN composé de 3APs avec un effet plus fort à vitesse d'élocution normale qu'à vitesse d'élocution rapide (vitesse d'élocution normale : $\beta = -0.39304$, $se = 0.05578$, $t = -7.046$, $pMCMC < 0.00001$; à vitesse d'élocution rapide : $\beta = -0.14182$, $se = 0.05586$, $t = -2.539$, $pMCMC = 0.0117$). Le

genre du locuteur n'était pas significatif bien que la pMCMC value était très proche du seuil de significativité confirmant le fait que la durée des voyelles est plus courte chez les hommes que chez les femmes ($\beta=-0.27599$, $se=0.114423$, $t=-1.914$, $pMCMC=0.0568$). L'interaction entre le type de SN et la vitesse d'élocution est significative montrant un plus petit degré de réduction de la voyelle à vitesse rapide qu'à vitesse normale ($\beta=0.1507$, $se=0.2065$, $t=3.182$, $pMCMC=0.4662$). L'interaction entre le genre des locuteurs et le type de SN était également faiblement significative ($\beta=-0.16674$, $se=0.07888$, $t=-2.114$, $pMCMC=0.0355$) avec un plus haut degré de réduction des voyelles pour les SN composés de 3APs chez les hommes plutôt que chez les femmes ($\beta=0.25122$, $se=0.07894$, $t=3.182$, $pMCMC=0.0016$). L'interaction entre la vitesse d'élocution et le genre des locuteurs n'était pas significative indiquant que l'augmentation de la vitesse d'élocution affectait de façon identique la durée des voyelles chez les hommes que chez les femmes ($\beta=-0.02583$, $se=0.07888$, $t=-0.327$, $pMCMC=0.436$). Aucune autre interaction n'était significative (les pMCMC values étaient toutes supérieures à 0.05).

Des résultats similaires ont été observés pour la durée des syllabes. Ces résultats ne sont pas surprenants dans la mesure où nous savons que l'allongement pré-frontière affecte préférentiellement la rime de la syllabe plutôt que l'attaque (en anglais : White 2002, Turck & Shattuck-Hufnagel, 2001 ; en français : Astézano, 2001). La figure 45 montre la durée de la syllabe en position finale du deuxième syntagme accentuel (S2) exprimée en millisecondes en fonction du type de syntagme nominal (SN composé de 2APs, SN composé de 3APs) de la vitesse d'élocution (vitesse normale, vn et vitesse rapide, vr) et du genre des locuteurs (hommes et femmes).

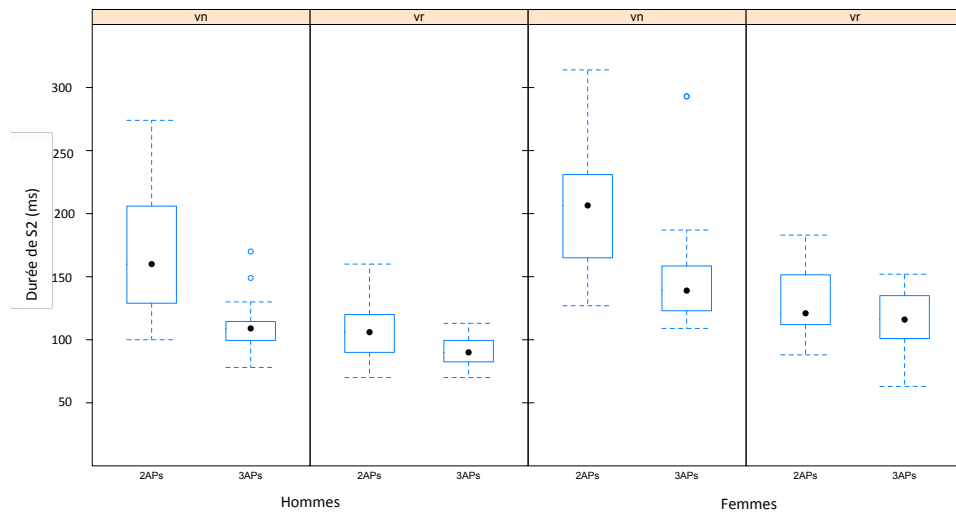


Figure 45. Étude 2 : Durée de la syllabe en position finale du deuxième syntagme accentuel (S2).

La durée est exprimée en millisecondes (ms) en fonction de la vitesse d'élocution (vitesse normale, vn et vitesse rapide, vr) du genre des locuteurs (hommes et femmes) et du type de syntagme nominal (SN composé de 2APs et SN composé de 3 APs).

4.4. Résultats relatifs à la *fo*

Nous avons également examiné les effets du type de syntagme nominal et de la vitesse d'élocution sur la hauteur de la voyelle située en position finale du deuxième syntagme accentuel des énoncés (H2). Les valeurs de *fo* exprimées en Hz en fonction du type de syntagme nominal (SN composé de 2APs ou SN composé de 3APs) de la vitesse d'élocution (vitesse normale, vn ou vitesse rapide, vr) et du genre des locuteurs (hommes ou femmes) est représentée dans la figure 46.

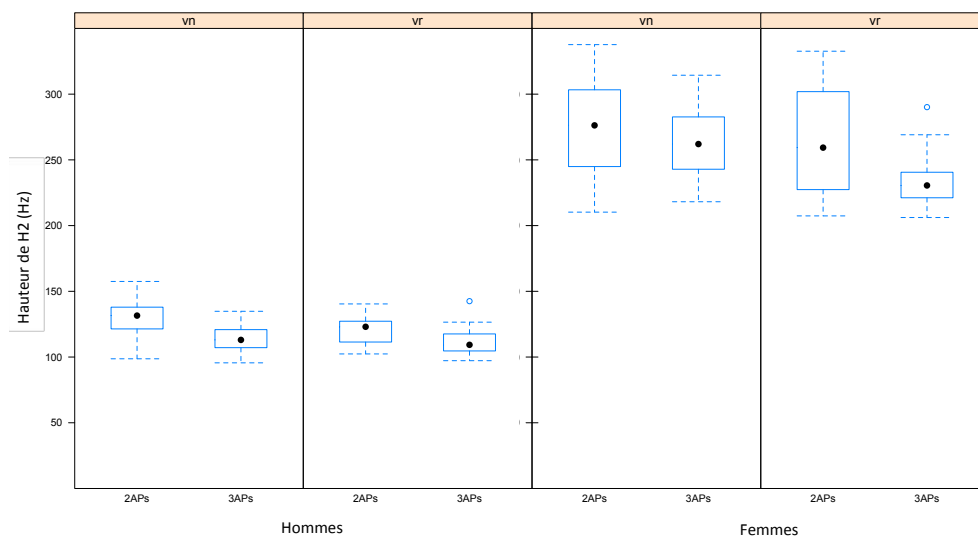


Figure 46. Étude 2 : Hauteur de la voyelle en position finale du deuxième syntagme accentuel (H2).

La hauteur est exprimée en millisecondes (ms) en fonction de la vitesse d'élocution (vitesse normale, vn et vitesse rapide, vr) du genre des locuteurs (hommes et femmes) et du type de syntagme nominal (SN composé de 2 APs et SN composé de 3 APs).

Nous avons conduit un modèle de régression linéaires multiples à effets mixtes (*Linear Mixed effects Models, LMM*) sur les valeurs logarithmiques de *fo* associées à la voyelle finale du second syntagme accentuel des énoncés avec le type syntagme nominal (SN composés de 2 APs et SN composés de 3 APs), la vitesse d'élocution (vitesse normale et vitesse rapide) et le genre des locuteurs (homme et femme) en tant qu'effets fixes. Les locuteurs (CD, CP, RF, NM) et le type de syllabe (mi, ri, na, ma) ont été inclus en tant qu'effets aléatoires. Le modèle incluait 244 mesures.

Les résultats montrent que le pic mélodique en position finale du deuxième syntagme accentuel des énoncés (H2) était significativement plus bas dans la condition 3 APs que 2APs à vitesse d'élocution rapide ($\beta = -0.12069$, $se = 0.02106$, $t = -5.73$, $pMCMC < 0.001$) mais pas à vitesse d'élocution normale ($\beta = -0.03962$, $se = 0.02105$, $t = -1.88$, $pMCMC = 0.061$). Malgré cette absence de significativité nous pouvons observer que H2 tend à être abaissé dans la condition 3APs par rapport à la condition 2APs. L'effet de la vitesse d'élocution n'était pas significatif ($\beta = -0.03665$, $se = 0.02105$, $t = -1.74$, $pMCMC = 0.08$) tandis que l'effet du genre des locuteurs l'était ($\beta = -0.74272$, $se = 0.09432$, $t = -8.22$, $pMCMC < 0.0001$). L'interaction entre le genre des locuteurs et le type de SN était significatif montrant un plus grand abaissement de la hauteur de la voyelle chez les hommes que chez les femmes au regard des deux vitesses d'élocution ($\beta = -0.08107$, $se = 0.02977$, $t = -3.28$, $pMCMC < 0.01$). La triple interaction type de SN, la vitesse d'élocution et le genre des locuteurs qui était significative indique un abaissement de H2 moins important à vitesse d'élocution rapide qu'à vitesse d'élocution normale

chez les hommes que chez les femmes ($\beta=0.13165$, $se=0.04211$, $t=3.13$, $pMCMC<0.01$). Aucune autre interaction n'était significative (les $pMCMC$ value étaient supérieures à 0.05).

Dans le but de prendre en compte les variations de registre observés entre les locuteurs mais aussi (et surtout) chez un même locuteur, nous avons mesuré le hauteur de H2 par rapport à la hauteur du premier pic de *fo* de l'énoncé (H1). Ainsi nous avons obtenus de rapports entre H2 et H1. Cette méthodologie est en lien avec et l'idée de van der Berg et al. (1992) qui a montré que le phénomène d'abaissement entre les accents mélodiques et entre constituants prosodiques est réalisé au moyen d'une ligne abstraite de référence du registre marquée par le premier pic de *fo* du domaine prosodique initial. Lorsque le rapport était égal à 1, cela signifiait donc que H2 avait une valeur équivalente à celle de H1. Au contraire, lorsque le rapport était inférieur à 1 cela signifiait que la hauteur de H2 était plus basse que celle de H1. Les rapports entre H2 et H1 sont illustrés dans la figure 47 en fonction du type de syntagme nominal (SN composé de 2APs ou SN composé de 3APs) de la vitesse d'élocution (vitesse normale, vn ou vitesse rapide, vr) et du genre des locuteurs (hommes ou femmes).

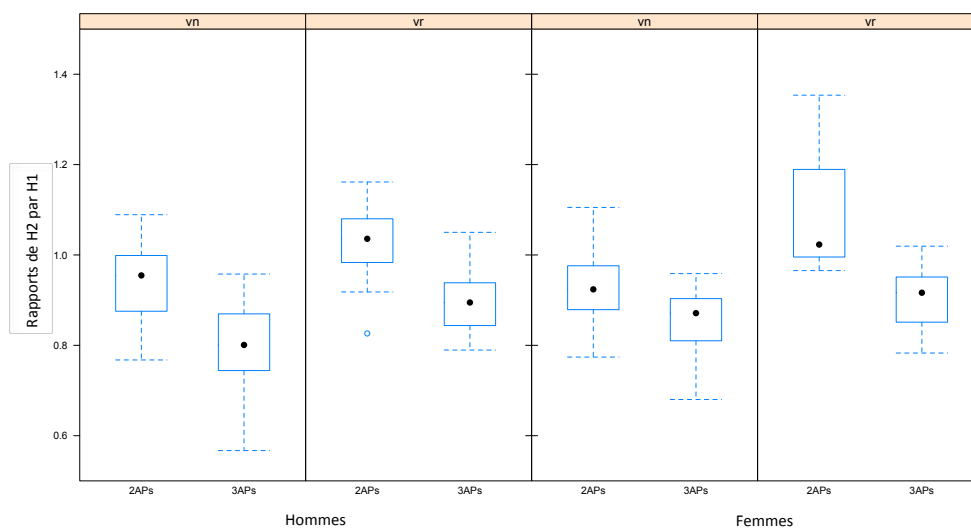


Figure 47. Étude 2 : Hauteur de la voyelle en position finale du deuxième syntagme accentuel (H2) par rapport à H1.

La hauteur de la voyelle est exprimée par rapport à la hauteur du premier pic de *fo* de l'énoncé (H1) en fonction de la vitesse d'élocution (vitesse normale, vn ou vitesse rapide, vr) du genre des locuteurs (hommes et femmes) et du type de syntagme nominal (SN composé de 2APs ou SN composé de 3APs). Le rapport 1 indique que H1 et H2 sont à une hauteur équivalente tandis qu'un rapport inférieur à 1 indique H2 est abaissé par rapport à H1.

Nous avons conduit un modèle de régressions linéaires multiples à effets mixtes (*Linear Mixed effects Models, LMM*) sur les rapports de H2 par H1 avec le type syntagme nominal (SN composés de 2 APS ou SN composés de 3

APs), la vitesse d'élocution (vitesse normale ou vitesse rapide) et le genre des locuteurs (homme et femme) en tant qu'effets fixes. Les locuteurs (CD, CP, RF, NM) et le type de syllabe (mi, ri, na, ma) ont été inclus en tant qu'effets aléatoires. Le modèle incluait 244 mesures.

Les résultats montrent que les rapports de H2 par H1 étaient significativement plus bas dans la condition 2APs que dans la condition 3APs quel que soit la vitesse d'élocution avec un effet plus marqué à vitesse d'élocution rapide qu'à vitesse d'élocution normale (à vitesse d'élocution normale : $\beta = -0.65543$, $se = 0.09432$, $t = -6.949$, $pMCMC < 0.01$; à vitesse d'élocution rapide : $\beta = -0.194413$, $se = 0.01979$, $t = -9.82$, $pMCMC < 0.0001$). L'effet de la vitesse d'élocution était significatif (vitesse d'élocution : $\beta = -0.3466$, $se = 0.09432$, $t = -63.675$, $pMCMC < 0.0001$) tandis que l'effet du genre du locuteur ne l'était pas ($\beta = 0.17867$, $se = 0.135$, $t = 1.323$, $pMCMC = 0.8655$). L'interaction entre la vitesse d'élocution et le type syntagme nominal était significative montrant que les rapports entre H1 et H2 étaient significativement plus élevés à vitesse d'élocution rapide qu'à vitesse d'élocution normale dans la condition 2APs que dans la condition 3APs ($\beta = 0.39643$, $se = 0.13344$, $t = 2.971$, $pMCMC < 0.001$). Aucune autre interaction n'était significative (les pMCMC-values étaient toutes supérieures à 0.05).

Nous avons également mesuré la hauteur du premier accent mélodique situé après la frontière d'ip. Ce pic mélodique (noté Hpb) était toujours situé sur la dernière syllabe du verbe de la phrase comme cela est illustré dans la figure 48.

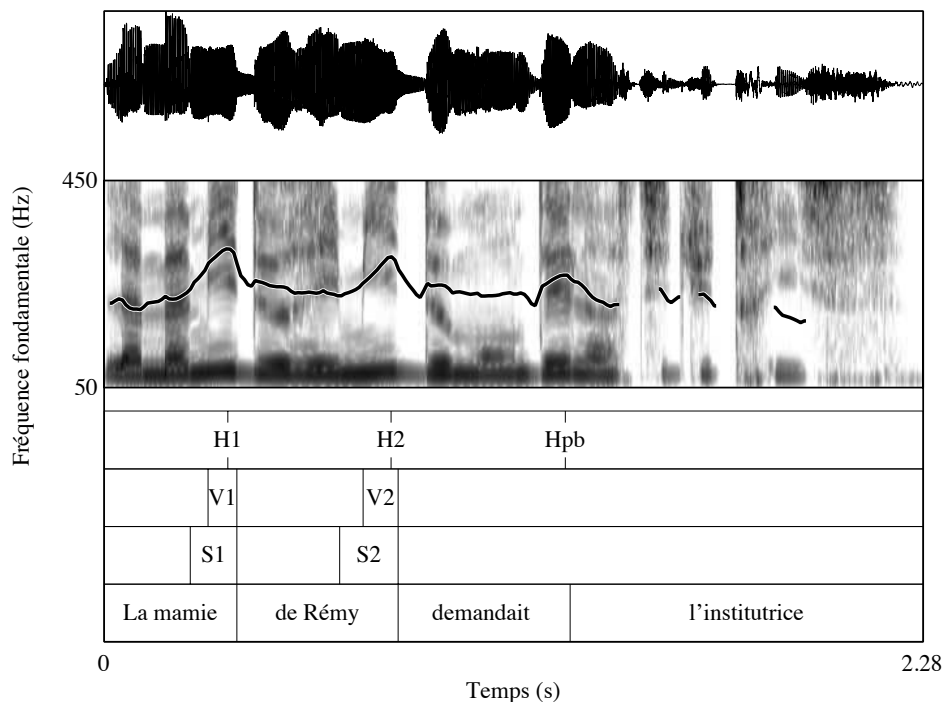


Figure 48. Étude 2 : Illustration du premier pic mélodique situé après la frontière d'ip. Onde sonore, spectrogramme et courbe de fréquence fondamentale associés à la phrase « La mamie de Rémy demandait l'institutrice » par une de nos locutrices à vitesse d'élocution normale. Le verbe « demandait » constitue un syntagme accentuel autonome de type LH*. Hpb correspond à la valeur de fo associée au premier accent mélodique situé après la rupture syntaxique entre le syntagme nominal et le syntagme verbal.

Dans cette figure, nous pouvons voir que le verbe « demandait » extrait de la phrase « la mamie de Rémy demandait la musicienne » est réalisé comme un syntagme accentuel de type LH*. L'accent mélodique sur le verbe (Hpb) correspond donc au premier accent mélodique observé après la frontière entre le syntagme nominal et le syntagme verbal, c'est à dire après la frontière d'ip. Nous n'avons pas observé de cas où le premier accent mélodique suivant la rupture syntaxique entre le syntagme nominal et le syntagme verbal n'était pas placé sur le verbe, y compris à vitesse d'élocution rapide. Notre hypothèse était qu'en raison d'un phénomène de *downstep* entre les deux ips de la phrase, la hauteur de Hpb serait plus basse que la hauteur d'un dernier accent mélodique de l'ip. En effet, nous avons émis l'hypothèse qu'un phénomène de reset partiel allait intervenir sur Hpb car ce pic mélodique serait associé à une nouvelle ligne de registre pour le deuxième ip et que cette ligne de registre serait abaissée par rapport à la ligne du registre du premier ip. Autrement dit, de manière similaire à ce qui a été observé pour le hollandais (van der Berg et al., 1992) et pour l'allemand (Truckenbrodt, 2002 ; Truckenbrodt & Féry, 2005) nous nous attendons à ce que qu'un phénomène de *downstep* intervienne entre les syntagmes intermédiaires en français. Le phénomène de réinitialisation partielle du registre que nous nous attendons donc à observer sur Hpb est schématisé dans la figure 49a pour les phrases dont les SN Sujet

étaient composés de deux APs et dans la figure 49b pour les phrases dont les SN Sujet étaient composés de trois APs.

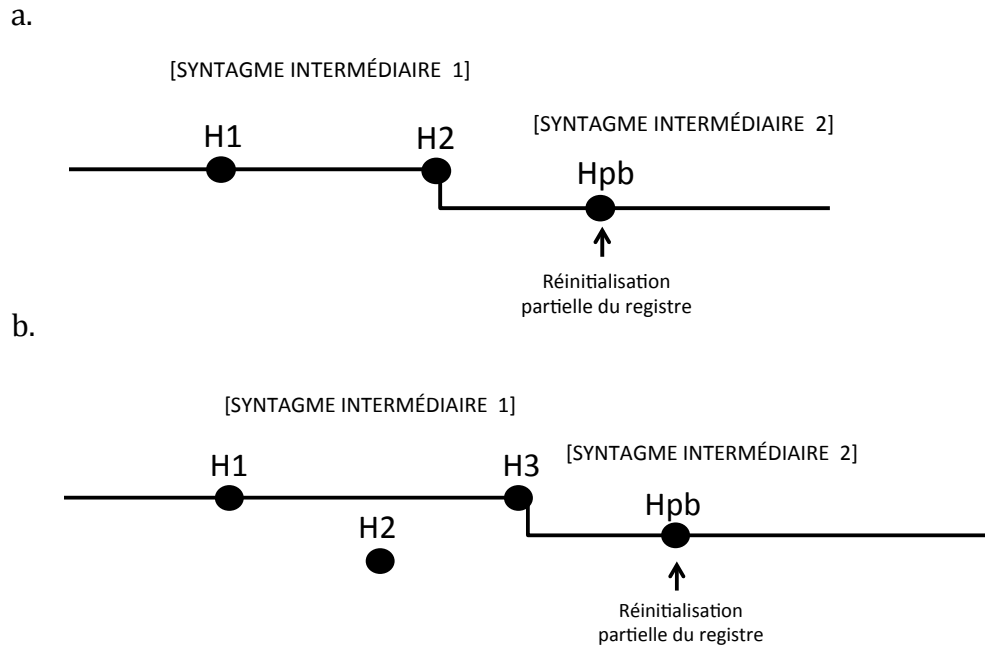


Figure 49. Étude 2 : Schématisation du phénomène de réinitialisation partielle du registre.

Nous nous attendons à observer un phénomène de réinitialisation partielle du registre sur Hpb en raison d'un phénomène d'abaissement entre les syntagmes intermédiaires en français pour les phrases dont les SN Sujet étaient composés de deux syntagmes accentuels (49a) ou de trois syntagmes accentuels (49b). Les points noirs symbolisent les accents mélodiques associés aux frontières de syntagme accentuels en français.

La valeur de fo de Hpb exprimée par rapport au dernier mouvement mélodique du premier syntagme intermédiaire que nous appelons H- est représentée dans la figure 50. Notons que H- était toujours associée à une frontière d'ip. Ce pic est le deuxième de l'énoncé (H2) dans la condition SN composés de 2 APs et le troisième de l'énoncé dans la condition SN composés de 3 APs. Rappelons que le rapport 1 indique que la hauteur de Hpb est équivalente à la hauteur de H-, tandis qu'un rapport plus petit que 1 indique que la hauteur de Hpb est inférieure à celle de H-.

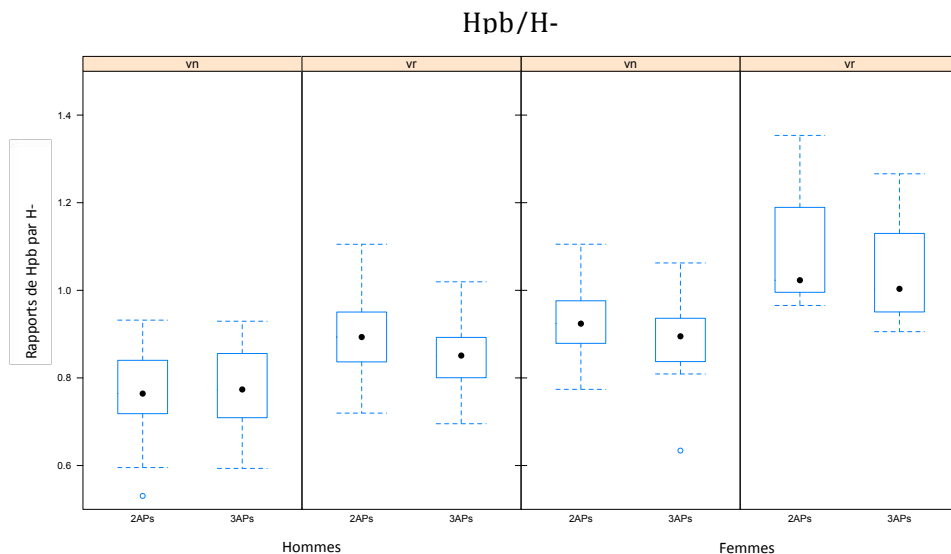


Figure 50. Étude 2 : Rapport de Hpb par H-. Valeurs du pic de fo pour de la voyelle associée au premier accent mélodique observé après la rupture entre le syntagme nominal et le syntagme verbal (Hpb) exprimée par rapport à la hauteur du pic mélodique précédent (H-) en fonction de la vitesse d'élocution (vitesse normale, vn et vitesse rapide, vr) du genre des locuteurs (hommes et femmes) et du type de syntagme nominal (SN composé de 2 APs et SN composé de 3 APs).

Dans cette figure nous pouvons voir que les rapports de H- par Hpb étaient toujours inférieurs à 1 ce qui signifie que Hpb était toujours plus bas que H-. Cependant, afin de vérifier que cet abaissement de Hpb était significatif, nous avons conduit un modèle de régression linéaires multiples à effets mixtes (*Linear Mixed effects Models, LMM*) sur les rapports de Hpb par H1 et de H- par H1- avec le type de pic (Hpb ou H-), le type syntagme nominal (SN composé de 2 APS et SN composé de 3 APs), la vitesse d'élocution (vitesse normale et vitesse rapide) et le genre des locuteurs (homme et femme) en tant qu'effets fixes. Les locuteurs (CD, CP, RF, NM) et le type de syllabe (mi, ri, na, ma) ont été inclus en tant qu'effets aléatoires. Le modèle incluait 244 mesures. Les rapports moyens de H- par H1 et Hpb par H1 estimés par le modèle sont représentés dans la figure suivante en fonction du type de pic (H- ou Hpb), de la vitesse d'élocution (normale ou rapide) du genre des locuteurs (homme ou femme) ainsi que du type de SN (composés de 2 ou 3 APs).

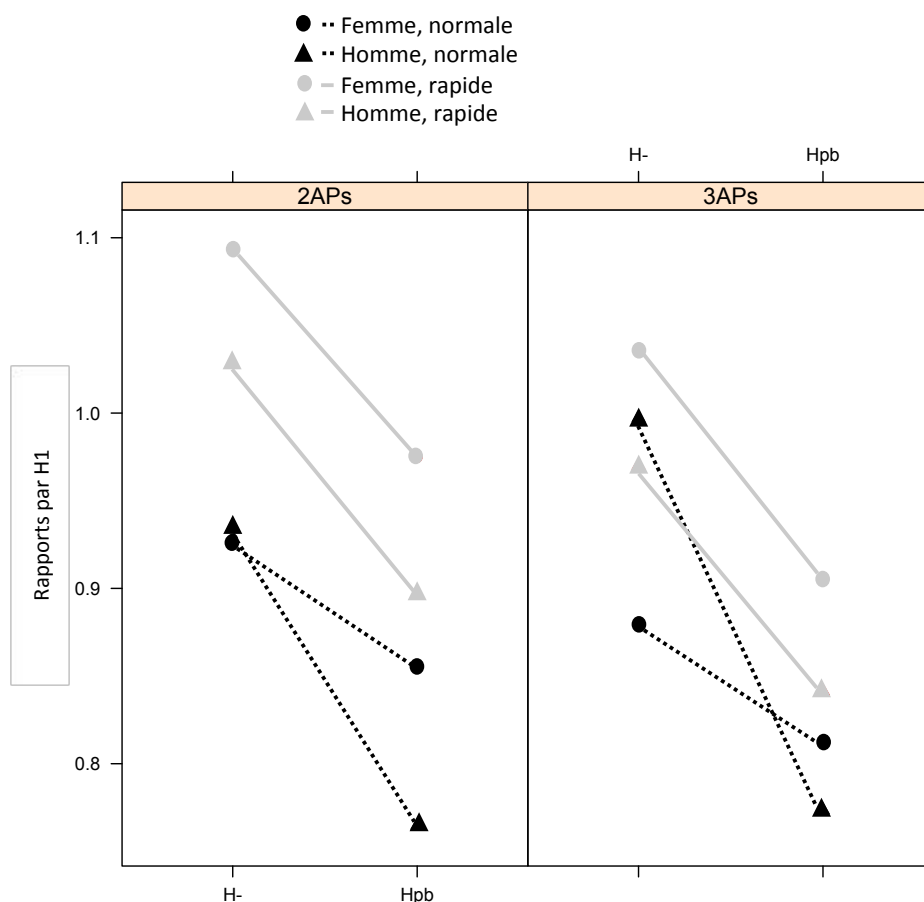


Figure 51. Étude 2 : Rapports de H- et de Hpb par H1 estimés par le modèle LMM. Les rapports sont fonction du genre du locuteur (homme ou femme) de la vitesse d'élocution (normale ou rapide) ainsi que du type de SN (SN composés de 2 APs ou SN composés de 3 APs).

Les résultats du modèle LMM montrent que les rapports de Hpb par H1 étaient significativement plus bas que les rapports de H- par H1 quel que soit la vitesse d'élocution, quel que soit le nombre d'APs et quel que soit le genre des locuteurs (à vitesse d'élocution normale pour les SN composés de 2 APs : $\beta = -0.073042$, $se = 0.0293$, $t = -3.49$, $pMCMC < 0.001$; à vitesse d'élocution normale pour les SN composés de 3 APs : $\beta = -0.067613$, $se = 0.02093$, $t = -3.23$, $pMCMC < 0.001$; à vitesse d'élocution rapide pour les SN composés de 2 APs : $\beta = -0.118249$, $se = 0.02093$, $t = -5.65$, $pMCMC < 0.001$; à vitesse d'élocution rapide pour les SN composés de 3 APs : $\beta = -0.130346$, $se = 0.02093$, $t = -6.25$, $pMCMC < 0.001$). Le type de SN et la vitesse d'élocution avait un effet significatif sur la hauteur des ratios tandis que le genre des locuteurs n'en avait pas ($\beta = 0.005724$, $se = 0.037176$, $t = 0.15$, $pMCMC = 0.877$). Les rapports de H- et de Hpb par H1 était plus élevés à vitesse d'élocution rapide qu'à vitesse d'élocution normale ($\beta = 0.164661$, $se = 0.02093$, $t = 7.87$, $pMCMC < 0.001$) et étaient moins élevés dans la condition 3 APs que dans la condition 2 APs ($\beta = -0.048383$, $se = 0.0293$, $t = -0.15$, $pMCMC < 0.001$). La double interaction entre le

type de pic (Hpb et H-) et le genre des locuteurs était significative montrant que Hpb était significativement plus bas chez les hommes que chez les femmes ($\beta=-0.096295$, $se=0.02960$, $t=-3.25$, $pMCMC<0.001$). La double interaction entre le type de SN (composés de 2 APs ou de 3 APs) et le genre des locuteurs était également significative avec des rapports de H- par H1 et de Hpb par H1 plus élevés pour la condition 3 APs chez les hommes que chez les femmes ($\beta=0.111577$, $se=0.0296$, $t=3.77$, $pMCMC=0.0001$). Enfin, aucune autre interaction entre les facteurs fixes n'était significative (les $pMCMC$ values étaient toutes supérieures à 0.005).

Afin de voir si le phénomène d'abaissement entre les accents mélodiques au sein de l'ip était plus fort que le phénomène d'abaissement entre les ips conformément à ce qui a été observé en hollandais (van der Berg, 1992) et en allemand (Truckenbrodt & Féry, 2005), nous avons également comparé la hauteur de Hpb et de H2 dans les phrases dont le SN Sujet contenait 3 APs. Notre hypothèse était que dans ces phrases, H2 serait plus bas que Hpb et donc que les rapports entre H2 et H1 seraient plus bas que les rapports entre Hpb et H1. Les rapports de H2 par H1 et de Hpb par H1 sont représentés dans la figure 52 en fonction de la vitesse d'élocution (normale ou rapide) et du genre des locuteurs.

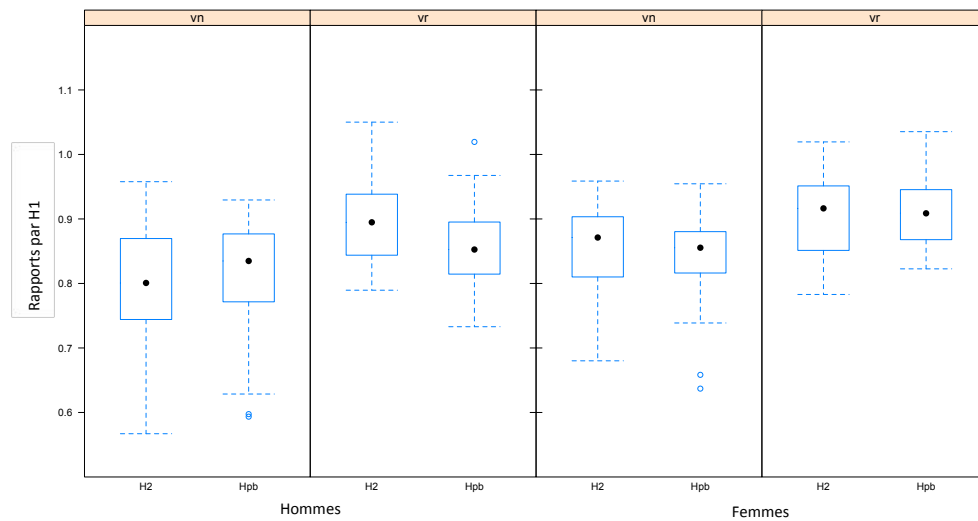


Figure 52. Étude 2 : Rapports de H2 et de Hpb par H1.
Les rapports sont donnés pour les phrases dont le SN contenait 3 APs en fonction de la vitesse d'élocution (vitesse normale, vn, ou vitesse rapide, vr) et le genre des locuteurs (hommes ou femmes).

Nous avons conduit un modèle de régressions linéaires multiples à effets mixtes (*Linear Mixed effects Models, LMM*) sur les rapports de H2 par H1 et de Hpb par H1- avec le type de pic (H2 ou Hpb) et la vitesse d'élocution (vitesse normale et vitesse rapide) et le genre des locuteurs (homme et femme) en tant qu'effets fixes. Les locuteurs (CD, CP, RF, NM) et le type de

syllabe (mi, ri, na, ma) ont été inclus en tant qu'effets aléatoires. Le modèle incluait 244 mesures.

Les résultats montrent que contrairement à ce qui était attendu les rapports H1 ne sont pas significativement différents pour H2 que pour Hpb et ce quel que soit la vitesse d'élocution et le genre des locuteurs (à vitesse d'élocution normale : $\beta = -0.01416$, $se = 0.01416$, $t = -0.888$, $pMCMC = 0.375$; à vitesse d'élocution rapide : $\beta = 0.01188$, $se = 0.015958$, $t = 0.713$, $pMCMC = 0.4763$). L'effet du genre des locuteurs n'était pas non plus significatif ($\beta = -0.05867$, $se = 0.04972$, $t = -1.180$, $pMCMC = 0.2391$). L'effet de la vitesse d'élocution était par contre significatif montrant des rapports par H1 plus élevé à vitesse rapide qu'à vitesse normale ($\beta = 0.0458$, $se = 0.01595$, $t = 2.871$, $pMCMC = 0.01$). Les interactions entre les facteurs fixes n'étaient pas significatives (les pMCMC values étaient toutes supérieures à 0.05). Enfin, les rapports de H2 par H1 et de Hpb par H1- estimés par le modèle en fonction de la vitesse d'élocution (vitesse normale et vitesse rapide) et le genre des locuteurs (homme et femme) sont représentés dans la figure 53.

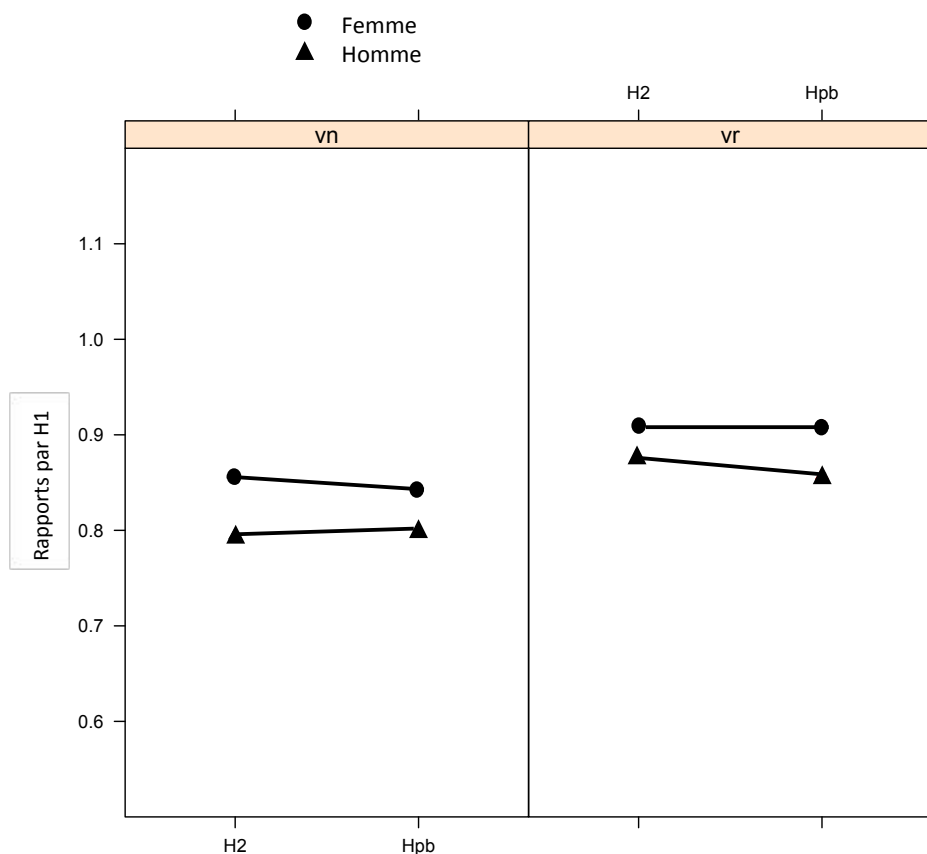


Figure 53. Étude 2 : Rapports de H2 et de Hpb par H1 estimés par le modèle LMM. Les rapports sont donnés pour les phrases dont le SN était composé de 3 APs en fonction du genre du locuteur (homme ou femme) de la vitesse d'élocution (vitesse d'élocution normale, vn, ou vitesse d'élocution rapide, vr).

5. DISCUSSION

5.1. Résultats principaux

L'étude que nous avons menée nous a permise de montrer que les indices acoustiques (durée et hauteur de la voyelle) associés à la voyelle finale du deuxième AP dans des phrases dont le syntagme nominal pouvait contenir soit 2 APs soit trois APs sont significativement différents. En effet, nous avons observé que la durée de la voyelle était significativement plus longue dans les phrases contenant 2 APs que dans les phrases contenant 3 APs. De plus la hauteur de la voyelle était significativement plus haute dans les phrases contenant 2 APs que dans les phrases contenant 3 APs. Ces résultats confirment notre hypothèse qui était que, dans la condition 2 APs, la voyelle finale était associée à une frontière de syntagme intermédiaire tandis qu'elle était associée à une frontière de syntagme accentuel dans la condition 3 APs telle que cela est illustré dans les phrases suivantes :

6a. [[La mamie]AP [**de RéMY**]AP]ip demandait l'institutrice.

6b. [[La mamie]AP[**des aMIS**]AP[de Rémy]AP]ip demandait l'institutrice.

Nous avons mené deux types d'analyse afin d'étudier la hauteur des voyelles cibles : une analyse sur les valeurs brutes de *f₀* et une analyse sur les rapports des valeurs de *f₀* associées à la syllabe cible (codées H2) par les valeurs de *f₀* du premier pic mélodique de l'énoncé (codées H1). A vitesse d'élocution normale et au regard des valeurs brutes de *f₀*, la différence de hauteur de la voyelle cible APs n'était pas significative entre la condition 2 APs et la condition 3 APs. Par contre, l'analyse menée sur les rapports de H2 par H1 montre que la voyelle cible était significativement abaissée dans la condition 2 APs relativement à la condition 3 APs. Cette analyse en terme de rapports nous a permise de normaliser les variations de registres observées non seulement entre les locuteurs mais surtout au sein d'un même locuteur et d'un même énoncé. Ces résultats montrent l'importance d'opérer des analyses normalisées dans le but d'étudier les variations de registre. L'idée d'une ligne de référence de registre marquée par le pic initial de *f₀* du constituant (cf. van den Berg et al., 1992), nous a paru particulièrement pertinente dans le but d'étudier la contribution des variations de registre dans le marquage des frontières prosodiques.

Les résultats que nous avons obtenus au cours de cette étude sont les premiers à montrer l'existence d'un syntagme intermédiaire pour le français qui ne serait pas restreint à des structures intonatives spécifiques tel que cela avait été précédemment postulé. À la lumière de ces résultats, il apparaît que le syntagme intermédiaire peut être réalisé dans tout énoncé dès lors que les structures syntaxiques et prosodiques le permettent.

Les résultats que nous avons obtenus montrent également que le premier accent mélodique suivant la frontière d'ip n'est jamais aussi haut que le dernier accent mélodique de l'ip. Ces résultats vont dans le sens d'un phénomène de réinitialisation partielle du registre apparaissant sur l'accent mélodique suivant la frontière d'ip pour les ips en position non-initiale dans l'IP. En effet, conformément à ce qui a été observé pour le hollandais ou l'allemand, ces résultats mettent en évidence un phénomène d'abaissement entre les syntagmes intermédiaires en français.

Enfin, nos résultats laissent apparaître une compensation très intéressante des indices de durée et des indices mélodiques dans le marquage des frontières de syntagme intermédiaire à vitesse d'élocution rapide. Nous avons en effet observé qu'à vitesse d'élocution rapide, tout comme à vitesse d'élocution normale, la frontière de syntagme intermédiaire était marquée à la fois par des indices mélodiques et de durée. Cependant, nous avons observé que si le degré d'allongement de la voyelle entre le niveau d'AP et d'ip diminue à vitesse d'élocution rapide, les indices mélodiques observés à la frontière d'ip semblent, eux, renforcés.

Ces résultats ont plusieurs implications théoriques importantes dans la mesure où (i) ils confirment l'existence d'un syntagme intermédiaire pour le français dont nous sommes désormais capable de mieux définir les propriétés phonétiques et phonologiques, (ii) ils montrent l'existence d'une structuration de l'IP en français avec l'existence d'un phénomène d'abaissement entre le ips et (iii) ils laissent entrevoir une compensation très intéressante de différents types d'indices acoustiques à vitesse d'élocution rapide dans le but de marquer les frontières d'ip. Nous allons maintenant développer plus en détails chacun de ces trois points.

5.2. Preuves de l'existence du syntagme intermédiaire

Dans le chapitre 2 de cette thèse, nous avons montré que la voyelle en position finale du deuxième syntagme accentuel était plus longue et plus haute si elle était associée à une rupture syntaxique majeure (telle qu'une rupture entre un syntagme nominal et un syntagme verbal) que si elle était contenue à l'intérieur d'un constituant syntaxique majeure. Au regard de ces résultats, nous postulons l'existence d'un niveau de structure prosodique additionnel pour le français qui serait situé entre l'AP et l'IP. À la différence de ce qui a été postulé par Jun & Fougeron (2000), nous considérons que l'existence du syntagme intermédiaire peut être attestée sur tout type d'énoncé dès lors que les structures syntaxiques et prosodiques le permettent. Rappelons que dans

le modèle de Jun & Fougeron (2000), un syntagme intermédiaire a été postulé dans le but de rendre compte de phénomènes de plateaux analysés comme des séquence L-L% (dans le cas d'un plateau bas) ou H-H% (dans le cas d'un plateau haut) ou H-L% (dans le cas d'un plateau médian). Selon cette proposition, l'accent de syntagme L- ou H- qui a comme propriété de pouvoir s'étendre de la frontière du syntagme intermédiaire jusqu'à l'accent mélodique qui le précède marquerait la frontière du syntagme intermédiaire.

Dans l'analyse que nous proposons, une frontière droite de syntagme intermédiaire peut apparaître sur tout énoncé. La frontière droite d'ip serait alignée avec la frontière droite d'une projection syntaxique majeure dès lors que le syntagme intermédiaire contient au minimum deux APs (voir également la contrainte de longueur MAX BIN postulé pour le Catalan par Prieto, à paraître, et par Feldhausen, 2010). L'émergence d'une frontière d'ip alignée avec la rupture syntaxique entre un syntagme nominal et un syntagme verbal est représentée dans la phrase 7. Notons que dans cette phrase le syntagme nominal sujet contient 2 APs ce qui est suffisamment long pour permettre à l'accent de syntagme marquant la frontière d'ip d'être réalisé.

7. La mamie	de Rémy	demandait l'institutrice.
Structure syntaxique :	[	]SN
Structure prosodique :	[[[	]AP[]AP]ip

Nous proposons donc que la frontière droite du syntagme intermédiaire soit marquée par un allongement pré-frontière significativement plus important que celui associé à la syllabe en position finale de l'AP et par accent de syntagme (L- ou H-) responsable du blocage de l'abaissement des accents mélodiques successifs à l'intérieur de l'ip. En raison de cet accent de syntagme, le registre est réinitialisé sur le dernier accent mélodique de l'ip. Dans la figure 54 montrant la courbe de *f₀* de la phrase « la mamie de Rémy demandait l'institutrice », nous pouvons remarquer une frontière de syntagme accentuel alignée avec la rupture entre le syntagme nominal et le syntagme verbal. Cette frontière est marquée à la fois par un allongement significatif de la syllabe et par un accent de syntagme H-responsable de la réinitialisation du registre sur l'accent mélodique situé juste avant la frontière du syntagme intermédiaire. De ce fait les valeurs de *f₀* observées sur la syllabe /mi/ du mot « Rémy » sont aussi hautes que les valeurs de *f₀* observées sur la syllabe /mi/ du mot « mamie ».

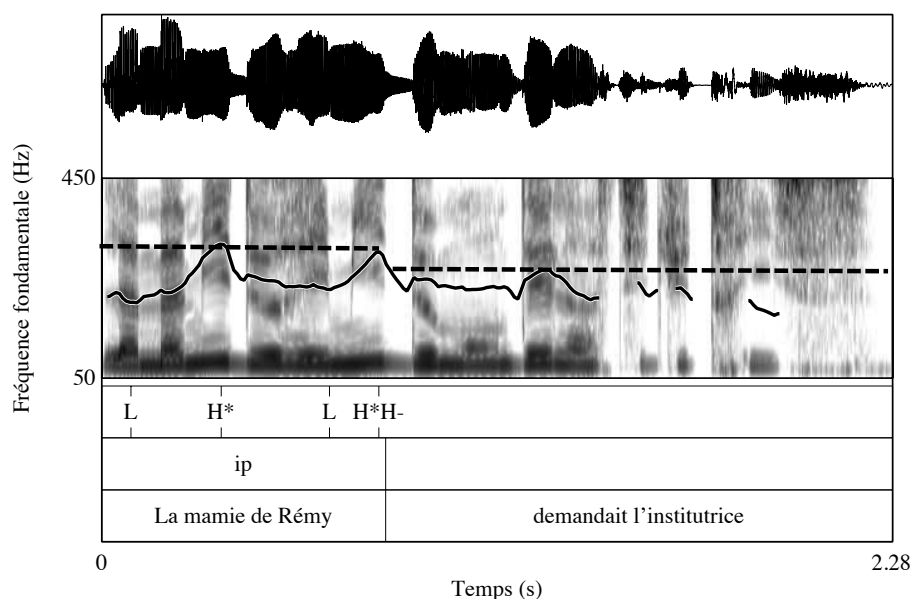


Figure 54. Étude 2 : Exemple de frontière d'ip.
Onde sonore, spectrogramme et courbe de fréquence fondamentale associée à la phrase « La mamie de Rémy demandait l'institutrice » prononcée par une de nos locutrices à vitesse d'élocution normale. Le syntagme nominal Sujet de la phrase est réalisé comme un syntagme intermédiaire. La frontière droite du syntagme intermédiaire est marquée par un accent de syntagme H- responsable de la réinitialisation du registre.

Si les données obtenues dans ce chapitre nous ont permis de mettre en évidence un degré d'allongement significativement plus important pour la frontière d'AP que pour la frontière d'ip, nous ne disposons pas du matériel nécessaire permettant de comparer le degré d'allongement associé aux frontières d'ip et d'IP. Étant donné que nous avons postulé l'existence d'un niveau additionnel de structure nous nous attendions en effet à ce que le degré d'allongement pré-frontière observé en position finale d'IP soit significativement plus important que celui associé à la frontière d'ip. Dans une étude pilote que nous avons menée par ailleurs et qui est présentée en annexe (Annexe C), nous avons mis en évidence que les indices acoustiques présents à la frontière d'IP sont significativement différents que ceux observés à la frontière d'ip. En effet, cette étude a montré (i) que le degré d'allongement observé à la frontière d'IP est significativement différent de celui associé à la frontière d'ip et que (ii) le ton de frontière marquant la frontière d'IP est responsable de l'élévation des valeurs de f_0 qui dépassent la ligne de référence du registre. Deux phrases dont la syllabe finale du troisième AP est associée soit à une frontière d'AP (55a) soit à une frontière d'IP (55b) sont illustrées dans la figure 55. Ces phrases sont extraites du corpus utilisé pour l'étude pilote présentée en Annexe C.

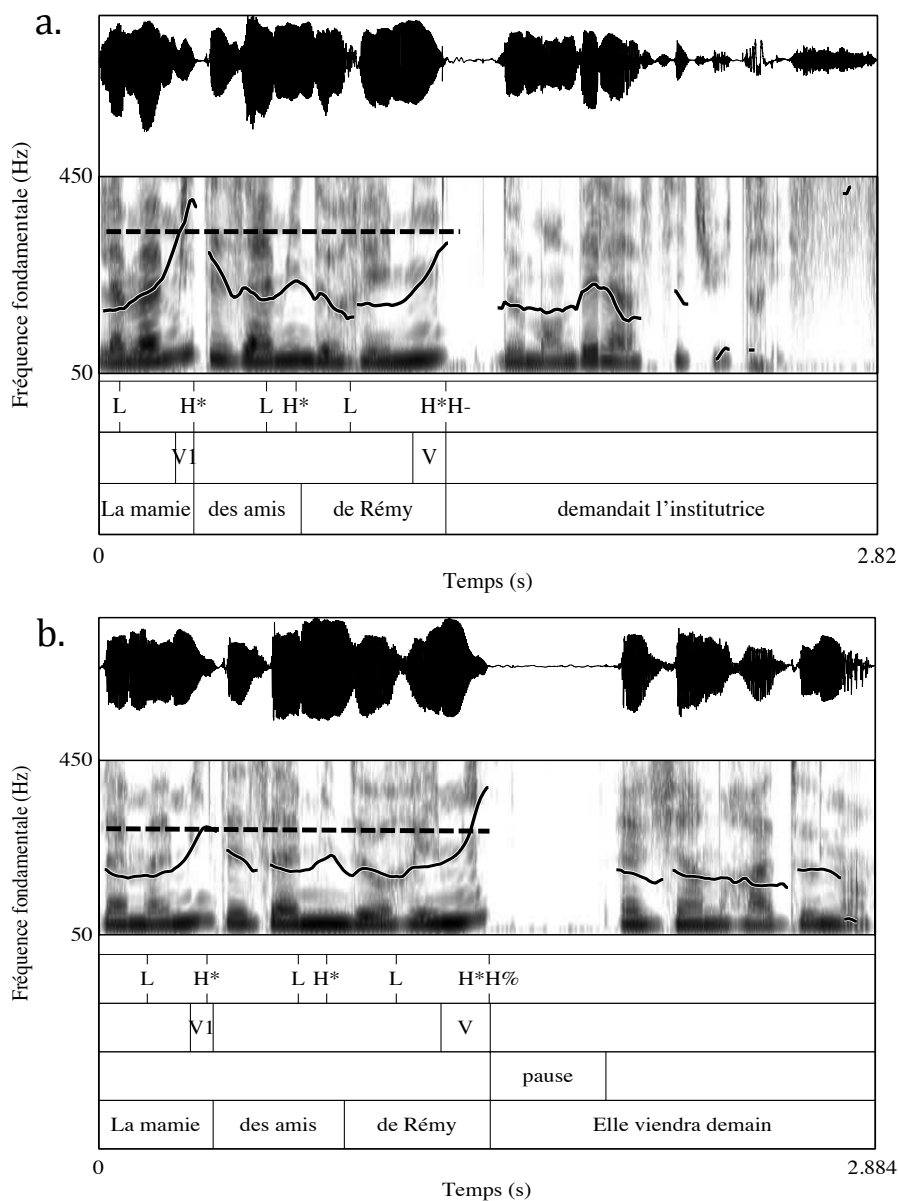


Figure 55. Frontières d'ip et frontières d'IP.

Contours mélodiques obtenus pour deux phrases dont la voyelle cible (V) pouvait être associée soit à une frontière d'ip (55a) marquée par un accent de syntagme H-responsable de la réinitialisation du registre sur le dernier accent mélodique de l'ip soit à par une frontière de syntagme intonatif (55b) marquée par un ton de frontière H% responsable de valeurs de *fo* qui dépassent la ligne de référence du registre marquée par le premier pic mélodique de l'énoncé.

Dans la figure 55a, nous pouvons observer que les valeurs de *fo* associées à la frontière d'ip (H1) sont similaires à celles observées sur le premier pic mélodique de l'énoncé en raison de la présence d'un accent de syntagme (H-) avant la frontière. Par contre dans la figure 55b, les valeurs de *fo* associées à la syllabe en position finale de l'IP dépassent celles observées

sur le premier pic mélodique de l'énoncé en raison de la présence d'un ton de frontière H% aligné avec la frontière droite d'IP

Dans nos données l'accent de syntagme associé à la frontière d'ip était toujours de type H- (jamais de type L-) et était toujours responsable de la réinitialisation du registre. Cependant notre analyse n'exclue pas que dans certains cas particuliers cet accent de syntagme puisse s'étendre tel que cela a été proposé par Jun & Fougeron (2000). En effet, nous pouvons envisager une analyse similaire à celle faite par Beckman & Pierrehumbert pour l'anglais (1986) où le syntagme intermédiaire est le domaine du *downstep* en anglais et où une des particularités de ce syntagme est qu'il peut s'étendre de la frontière du syntagme jusqu'à l'accent mélodique qui le précède. Des analyses supplémentaires seront nécessaires dans le but de déterminer les contraintes de différentes natures qui peuvent engendrer le déploiement (*spreading*) de l'accent de syntagme associé à la frontière d'ip en français.

Si nos données mettent en évidence l'existence d'un niveau additionnel de structuration pour le français, elles permettent également de montrer que l'émergence du syntagme intermédiaire est conditionnée par une contrainte de longueur en français. En effet, au cours du chapitre précédent, nous avons observé que la voyelle en position finale d'un AP, alignée avec une rupture syntaxique majeure (frontière AP/XP) était plus longue que la voyelle en position finale d'un AP contenu à l'intérieur d'un syntagme nominal sujet (frontière d'AP). Cependant les indices mélodiques associés au deux types de voyelles (en position finale d'AP ou en position finale d'AP/XP) n'étaient pas différents. Une des interprétations possibles de ce résultat était que l'allongement observé à la frontière AP/XP était à associer à un niveau de structure prosodique additionnel dans la hiérarchie prosodique : le syntagme intermédiaire. Cependant, en raison du fait que le syntagme nominal Sujet était trop court (composé uniquement d'un AP), l'accent de syntagme associé à la frontière d'ip, qui est responsable de l'abaissement des accents mélodiques successifs à l'intérieur de l'ip, n'était pas visible dans le signal de parole. Des expériences en perception de la parole sont en fait nécessaires dans le but de déterminer si la frontière d'AP/XP que nous avons observée dans l'étude 1 (cf. Chapitre 2) est perçue comme une frontière d'ip ou si la seule présence des indices de durée n'est pas suffisant pour que les auditeurs du français perçoivent la frontière d'AP/XP en tant que frontière d'ip. Nous apporterons des éléments de réponse un peu plus loin dans cette thèse (cf. Chapitre 5).

Enfin, dans les analyses que nous avons menées jusqu'ici le syntagme intermédiaire était toujours observé en position Sujet dans la phrase. Cependant, des résultats pilotes obtenus à partir d'un corpus de 64 phrases lues que nous avons étudié par ailleurs (Annexe D), montrent que le syntagme intermédiaire peut également apparaître en position Objet dans les phrases en français. Un exemple d'ip en position Objet dans la phrase « Marie a rencontré la mamie des amis de Rémy dans la journée » est illustré dans la figure 56.

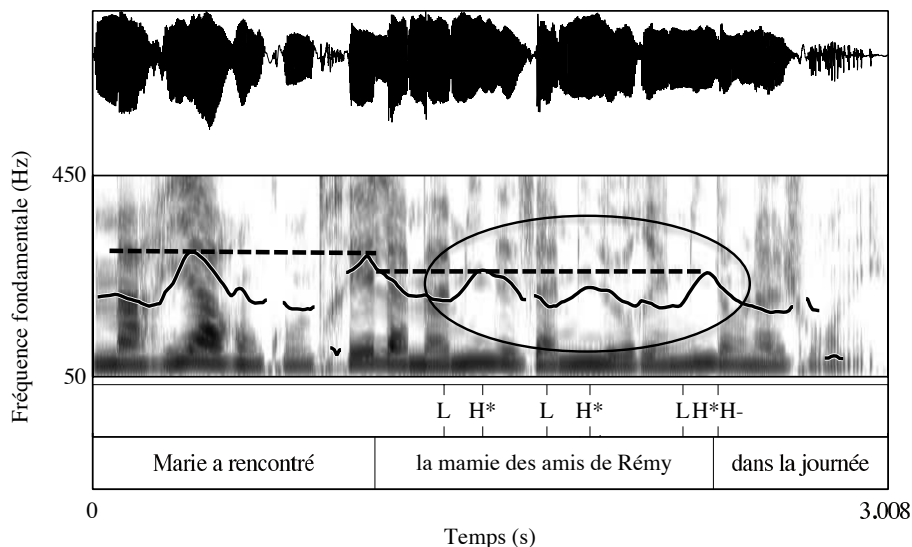


Figure 56. Exemple d'ip en position Objet dans la phrase

5.3. Preuves de l'existence d'un phénomène d'abaissement entre les ips

Dans l'étude présentée au cours de ce chapitre, nous avons également montré que le premier accent mélodique suivant la frontière d'ip n'est jamais aussi haut que le dernier accent mélodique de l'ip. Notons que Rialland & Aborobongi (2010) ont également suggéré un abaissement du registre après une frontière prosodique en français. Ces résultats confirment l'existence d'un phénomène de réinitialisation partielle du registre. Ce type de phénomène a déjà été montré pour l'anglais (Ladd, 1988), le hollandais (van der Berg et al. 1992) et l'allemand (Truckenbrodt, 2002 ; Truckenbrodt & Féry 2005). En français, ce phénomène de réinitialisation partielle du registre intervient après la frontière d'ip comme cela est schématisé dans la figure 57.

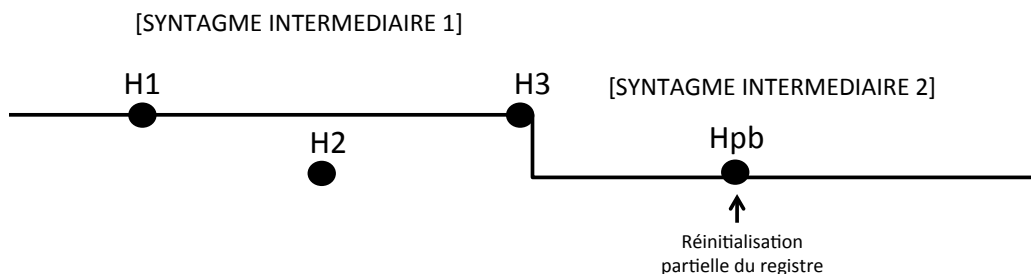


Figure 57. Schématisation de la réinitialisation partielle du registre sur le premier accent mélodique situé après la frontière d'ip.

Ce phénomène de réinitialisation du registre fait apparaître un phénomène d'abaissement qui intervient entre deux syntagmes intermédiaires successifs en français. En effet, comme nous pouvons le voir dans la figure 57, le deuxième syntagme intermédiaire est abaissé relativement au premier syntagme intermédiaire. Ce phénomène d'abaissement se manifeste par un abaissement de la ligne de registre du syntagme intermédiaire marqué par le premier accent mélodique du syntagme. Ces résultats confirment ce qui avait été observé pour l'anglais, le hollandais ou l'allemand et vont à l'encontre de ce qui a été proposé par Beckman & Pierrehumbert (1986) pour qui les phénomènes d'abaissement et de réinitialisation observés entre les constituants seraient purement liés à des phénomènes paralinguistiques ou d'ordre discursif.

Par contre, nos résultats ont également permis de mettre en évidence que les valeurs de *fo* de Hpb n'étaient pas significativement différentes des valeurs de *fo* de H2 dans les phrases dont le SN sujet contenait 3 APs. Ces résultats semblent indiquer que le phénomène d'abaissement des accents mélodiques à l'intérieur de l'ip n'est pas plus fort que le phénomène d'abaissement des ips entre eux. Ce résultat va à l'encontre de ce qui a été observé pour l'allemand ou le hollandais où le *downstep* à l'intérieur du constituant prosodique est plus fort que le *downstep* entre constituants prosodiques (van der Bergh, 1992 ; Truckenbrodt & Féry, 2005). Cependant, dans nos données, la syllabe porteuse de l'accent mélodique située après la frontière d'ip (Hpb) n'était pas toujours la même et n'avait pas toujours la même structure ce qui a pu potentiellement affecter les valeurs de *fo* que nous avons mesuré sur cette syllabe. Des analyses supplémentaires seront donc nécessaires dans le but de déterminer si, effectivement, le phénomène d'abaissement entre les accents mélodiques à l'intérieur de l'ip est de taille équivalente à celui du phénomène d'abaissement entre les ips en français.

5.4. Effets de la vitesse d'élocution sur les indices acoustiques marquant la frontière d'ip

Enfin, les résultats développés au cours de cette étude nous ont permis d'observer un phénomène de compensation très intéressant entre les différents types d'indices acoustiques marquant les frontières de syntagme intermédiaire à vitesse d'élocution rapide. Rappelons que nous admettons un marquage pluriel de la frontière d'ip en français dont les indices acoustiques principaux semblent être l'allongement de la syllabe pré-frontière et la présence d'un accent de syntagme. De manière peu surprenante, nous avons observé, que la différence du degré d'allongement associé à la frontière d'AP et à la frontière d'ip était moins importante à vitesse d'élocution rapide, qu'à vitesse d'élocution normale. En effet, en raison d'une réduction globale des syllabes, liée à l'augmentation de la vitesse d'élocution, le marquage des frontières par l'allongement devient difficile surtout pour les constituants

prosodiques situés à un haut niveau dans la hiérarchie qui réclament un degré d'allongement important. Par contre, de manière inattendue, nous avons observé, qu'à l'inverse de l'allongement pré-frontière, les indices mélodiques présents à la frontière d'ip semblent renforcés à vitesse d'élocution rapide. En effet, la différence de hauteur entre la voyelle en position finale du deuxième syntagme accentuel (H2) dans les conditions 2 APs et 3 APs était plus importante à vitesse d'élocution rapide qu'à vitesse d'élocution normale. Ces résultats montrent que, si la contribution conjointe des indices mélodiques et de durée semble nécessaire dans le marquage des frontières d'ip en français, les indices mélodiques semblent être plus importants à vitesse d'élocution rapide. Ces derniers viennent en quelque sorte combler une efficacité moins grande des indices de durée en raison d'une réduction globale des syllabes.

CHAPITRE 4

PERCEPTION DE LA FRONTIÈRE DE SYNTAGME ACCENTUEL ET ACCÈS AU LEXIQUE (Étude 3)

1. PROBLÉMATIQUE

Le processus de perception de la parole peut être envisagé comme une série d'opérations et de représentations qui permettent, à terme, d'associer une signification conceptuelle à un signal physique. Il nous semble particulièrement important de souligner deux caractéristiques fondamentales de ces procédures de reconnaissance des mots. Tout d'abord, l'extrême fluidité de ces procédures est particulièrement étonnante. En effet, quand nous entendons un mot de notre langue, nous le comprenons de façon presque immédiate alors que notre capacité à comprendre ces mots implique des procédures très complexes. Notons également que ces procédures sont hautement automatisées : lorsqu'un mot ou une phrase nous sont présentés nous l'identifions instantanément et nous ne pouvons pas ne pas le faire. L'efficacité de ces procédures de reconnaissance des mots est d'autant plus surprenante étant donné la nature continue du signal de parole. Cette caractéristique du signal acoustique implique que, contrairement au système écrit où les espaces blancs marquent les frontières de mots, aucun indice ne marque clairement et de façon systématique le début et la fin des mots à l'oral. Comment l'auditeur est-il donc capable d'analyser un signal essentiellement continu en une séquence d'unités significatives discrètes ?

La reconnaissance d'un mot peut être envisagée comme l'étape terminale du processus enclenché par la présentation du signal acoustique. Il est possible de distinguer trois étapes distinctes dans le processus de reconnaissance des mots : celle du contact initial, celle de la sélection et enfin celle de la reconnaissance à proprement dite (voir par exemple Frauenfelder & Dufour, 1991). La phase de contact permet, à l'écoute d'un mot, d'activer un certain nombre de candidats phonologiquement compatibles avec le mot entendu. La phase de sélection permet ensuite l'émergence d'un candidat unique, qui est le candidat le plus compatible avec la forme d'entrée, c'est-à-

dire le signal acoustique reçu par l'auditeur. Enfin le candidat sélectionné sera reconnu par l'auditeur lors de la phase de reconnaissance.

Depuis longtemps les modèles psycholinguistiques de traitement de la parole comme le modèle TRACE (McClelland & Elman, 1986) et Shortlist (Norris, 1994) se sont intéressés à la question du découpage en mots. Dans ces modèles, la segmentation est considérée comme le résultat d'une compétition entre différents candidats lexicaux. Par exemple le modèle TRACE (McClelland & Elman, 1986) envisage ce processus complexe comme une activation des candidats qui entrent en compétition les uns avec les autres. À n'importe quel point de la chaîne parlée, tous les mots compatibles avec l'information phonémique sont activés. Dans ce modèle, tous les candidats qui partagent un ou plusieurs phonèmes vont être activés, indépendamment du fait qu'ils partagent ou non les segments initiaux avec les mots entendus. Par exemple, à l'écoute du mot « rondelle » des mots tels que « ronde » ou « rongeur » seront potentiellement activés mais également des mots tels que « onde » ou « elle » (cf. Dufour & Frauenfelder, 2007, pour plus de détails sur ce point). Dans ce modèle, la compétition est réalisée au moyen d'un mécanisme d'inhibition latérale entre les mots. L'inhibition entre les compétiteurs permet à ceux dont le niveau d'activation est le plus élevé de prédominer et d'éliminer ceux dont l'activation est le moins élevé (Frauenfelder & Peeters, 1990 ; Norris, 1994).

Il arrive qu'une chaîne des phonèmes puisse être découpée de plusieurs façons différentes. Par exemple, l'information segmentale /ʒɑ̃levsɔ̃vʁ/ peut être compatible avec plusieurs hypothèses de découpage en mots : « Jean lève son verre » ou « J'enlève son verre ». L'auditeur d'une langue est fréquemment confronté à ce type de phénomène que l'on appelle des ambiguïtés lexicales, c'est-à-dire le fait qu'à un moment donné du signal plusieurs candidats soient compatibles avec les formes d'entrée.

Alors que certains ont mis en évidence le rôle prépondérant des régularités phonotactiques et probabilités transitionnelles dans la segmentation (Saffran, Newport & Aslin, 1996 ; McQueen, 1998), de nombreuses études accordent aujourd'hui un rôle important à l'information prosodique. Dans le modèle Shortlist, les informations relatives à l'accent lexical ont été incorporées (cf. Norris, 1994, p°229). En effet, les candidats qui commencent par une syllabe métriquement forte reçoivent un taux d'activation plus important si le signal d'entrée est lui même doté d'une syllabe métriquement forte au début du mot. Nous savons depuis longtemps que les auditeurs d'une langue sont capables d'utiliser les indices métriques dans le découpage en mots. En 1988, Cutler & Norris avait montré que les auditeurs de l'anglais ont recours à l'accent de mot en tant qu'indice de la frontière initiale des mots. D'autres études également menées sur l'anglais ont permis de confirmer l'influence de l'accent de mot sur la segmentation y compris chez les enfants. En effet, il apparaît que des bébés âgés de neuf mois, utilisent les syllabes métriquement fortes et faibles de mots bi-syllabiques en tant qu'indice de la frontière de mot (Mattys, Jusczyk, Luce & Morgan, 1999). Cependant, la stratégie de segmentation métrique ainsi mise en évidence n'est

pas applicable telle quelle à toutes les langues. Rappelons qu'en anglais, l'accent lexical est distinctif et a tendance à être placé au début des mots à contenu. Toutes les langues ne sont pas dotées d'un système prosodique où la position de l'accent lexical est distinctive. En français il n'est pas possible de trouver des paires minimales de mots qui ne diffèrent que par la position de l'accent. Les locuteurs du français ont donc recours à d'autres indices acoustiques dans le but de guider la segmentation lexicale. Nous avons vu précédemment, qu'en français, la dernière syllabe d'un syntagme accentuel est porteuse de l'accent mélodique. Plusieurs études ont mis en évidence le fait que les auditeurs du français sont capables d'utiliser ce type d'indice pour segmenter la chaîne de parole en mots. Un des moyens le plus employé pour étudier ce type de phénomène est le recours à des séquences comprenant une ambiguïté lexicale locale (c'est-à-dire temporaire). Par exemple en 1980, Rietveld a examiné les différences phonétiques existantes entre des paires telles que « le combat saccagé » et « le comte a saccagé ». L'auteur met en évidence des différences de *f₀*, d'intensité et de durée correspondant à un accent final sur la dernière syllabe du syntagme nominal. Les résultats d'un test de perception montrent que les auditeurs du français sont capables d'identifier les deux séquences sur la base de leurs différences prosodiques. Grâce à l'utilisation d'un modèle statistique à régressions multiples, l'auteur montre que la durée est le meilleur prédicteur permettant d'expliquer la variance. L'auteur interprète prudemment ce résultat en disant que la durée semble l'indice le plus pertinent dans la détection des frontières de mots en français. L'auteur propose également que le « groupe rythmique » soit le domaine de l'accès au lexique étant donné que seule la dernière syllabe du groupe est accentuée en français.

En 1991, Banel & Bacri ont montré que dans des séquences telles que « bagages » et « bas gages » des auditeurs perçoivent préférentiellement deux mots lorsque la première syllabe est allongée et un seul mot lorsque la deuxième syllabe est allongée. Plus récemment, Bagou, Fougeron & Frauenfelder (2006) ont également montré que, dans une langue artificielle, les auditeurs du français suisse utilisent la montée finale de *f₀* ainsi que l'allongement associé à la fin des mots dans le but d'inférer la présence d'une frontière lexicale. Contrairement à ce qui avait été observé par Rietveld (1980) qui considère que l'allongement est un meilleur indice que les marques mélodiques pour guider la segmentation lexicale, les auteurs ont observé que lorsqu'un seul indice était pris en compte, les indices de durée n'induisaient pas un meilleur taux d'identification des frontières de mots que les indices mélodiques. Une explication possible de ce résultat est liée au fait que la montée mélodique associée à chaque fin de mot dans l'expérience était induite par un contour de liste. Le fait que ces montées successives de *f₀* apparaissaient de manière récurrente a potentiellement pu favoriser la segmentation en mots. En effet, les participants de l'expérience ont pu deviner qu'ils étaient en train d'écouter une liste de mots plutôt que des phrases ce qui aurait pu les aider à repérer les frontières lexicales.

D'autres études ont plus spécifiquement mis en évidence la contribution des indices mélodiques dans la segmentation des mots en français. Dans sa thèse de doctorat, Welby (2003), en suivant des suggestions faites dans la littérature par Vaissière (1997) et Di Cristo (2000), a montré que les auditeurs du français utilisent la présence de la montée initiale de *fo* (LHi), ou même la présence d'un point d'inflexion dans la courbe de *fo*, comme un indice de segmentation du début des mots à contenu. Par exemple, des auditeurs du français interprétaient des séquences de logatomes telles que « mélamondine » comme un seul mot lorsque la montée initiale de *fo* était associée à la première syllabe et comme deux mots lorsque la montée initiale de *fo* était associée à la deuxième syllabe du logatome. Plus récemment et grâce à différentes études menées à la fois sur de la parole naturelle et de la parole de resynthèse, Spinelli, Welby & Shaegis (2007) et Spinelli, Nicolas, Meunier & Welby (2010) ont montré que des participants utilisaient en temps réel la montée initiale pour discriminer des séquences telles que « c'est l'affiche » et « c'est la fiche ». Cependant, dans une étude menée grâce à un paradigme d'acquisition de langue artificielle, Bagou & Frauenfelder (2002) ont observé qu'alors qu'un patron mélodique de type LHiLH* (c'est-à-dire contenant à la fois une montée initiale et une montée finale de *fo*) semblait faciliter l'identification des mots, la montée initiale seule (c'est-à-dire sans la présence d'une montée finale) ne semblait pas pouvoir être utilisée en tant qu'indice marquant le début des mots. Cependant, les unités de la langue artificielle utilisées étaient des pseudo-mots à contenu et aucun pseudo-mot grammatical n'était présent. Or Welby (2003) a montré que la montée mélodique initiale caractérise la frontière entre un mot grammatical et un mot à contenu. Etant donné l'absence mot grammatical dans le corpus utilisé par Bagou & Frauenfelder (2002), et comme le concède eux-mêmes les auteurs dans la discussion de leur article, il est possible que la montée mélodique initiale n'ait pas pu jouer son rôle d'indice de frontière entre mot grammatical et mot à contenu dans cette expérience.

Parallèlement aux études menées sur le français, des études menées sur d'autres langues semblent confirmer le rôle des indices mélodiques dans la segmentation lexicale. Kim (2004) et Kim & Cho (2009) ont par exemple montré qu'en coréen, la présence des indices mélodiques seule (en l'absence de tout indice de durée) guide les auditeurs dans le processus de segmentation lexicale. Grâce à deux tâches de localisation de mot à l'intérieur de l'AP (*word spotting*) les auteurs ont également démontré que, dans cette langue, la présence d'indices temporels n'améliorait les performances des locuteurs que dans le cas où les frontières prosodiques étaient marquées par des configurations mélodiques peu fréquentes. Ces résultats sont très intéressants et soulignent le fait que l'importance du type d'indice marquant les frontières prosodiques et la combinaison de ces indices semble varier en fonction des langues.

Si la contribution des indices prosodiques dans les processus de segmentation lexicale semble admise par tous, le domaine à l'intérieur duquel les stratégies de segmentation lexicale opèrent, n'est pas encore clairement

établi pour le français. Nous avons vu que la parole produite par un locuteur est organisée en constituants prosodiques de différents niveaux qui sont hiérarchisés. Il semble raisonnable d'assumer que, lors du processus de reconnaissance des mots, l'activation des candidats se fait à l'intérieur des frontières de syntagme. Par exemple deux syllabes situées de part et d'autre d'une frontière de syntagme intonatif pouvant notamment être marquée par la présence d'une pause silencieuse, ne pourront pas donner lieu à l'activation d'un candidat composé de deux syllabes séparées par la pause silencieuse. Prenons l'exemple 1 suivant.

1. « Jean voudrait arracher ce très vieux **pin.** **IP** **Son** père n'est pas d'accord ».
Présence d'une ambiguïté lexicale temporaire : compétiteur /pēsɔ̃/

On pourrait supposer que cette séquence soit temporairement ambiguë en raison de la présence côte à côte des syllabes « pin » et « son » qui forment le mot « pinson » en français. Cependant nous savons que dans la plupart des cas, le candidat potentiel « pinson » ne sera pas activé en raison de la frontière de syntagme intonatif présent entre ces deux syllabes qui serait interprétée comme une frontière de mots. La frontière de syntagme intonatif qui sépare l'énoncé « Jean voudrait arracher ce très vieux pin » de l'énoncé « Son père n'est pas d'accord » est représentée dans la figure 58.

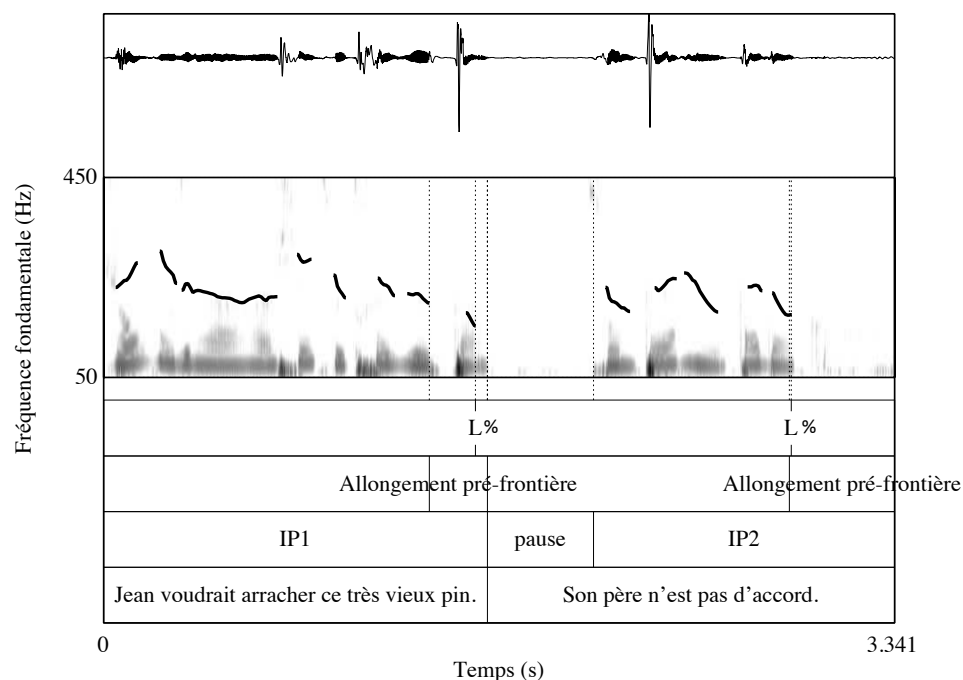


Figure 58. Illustration d'un cas d'ambiguïté lexicale à cheval sur une frontière de syntagme intonatif.

Onde sonore, spectrogramme et courbe de fréquence fondamentale associés aux deux unités intonatives «[Jean voudrait arracher ce très vieux pin]IP. [Son père n'est pas d'accord.]IP» prononcées par une locutrice de langue maternelle française à vitesse d'élocution normale.

Comme nous pouvons le voir dans cette figure, la frontière de syntagme intonatif séparant les deux énoncés est repérable grâce à la réalisation d'un ton de frontière présent sur la dernière syllabe (L%), de la présence d'un allongement pré-frontière ainsi que de la réalisation d'une pause silencieuse.

Rietveld (1980) avait le premier suggéré qu'un constituant prosodique de niveau inférieur au syntagme intonatif, qu'il appelle « groupe rythmique », soit le domaine d'accès au lexique en français étant donné que seule la dernière syllabe du groupe est accentuée dans cette langue. Suivant cette idée, Christophe et ses collaborateurs (Christophe et al. 2004) ont montré que les frontières de syntagme phonologique (*Phonological Phrase*, PP) semblent contraindre l'accès au lexique en français. Dans une étude menée en 2004, les auteurs ont étudié l'effet d'ambiguïtés lexicales temporaires qui pouvaient être situées soit à l'intérieur d'un syntagme phonologique comme dans la phrase 2a soit à la frontière d'un syntagme phonologique comme c'est le cas dans la phrase 3a. Notons en effet la présence d'une ambiguïté lexicale dans les phrases 2a et 3a en raison de la compétition entre deux candidats lexicaux « chat » et « chagrin ». À l'inverse, les phrases 2b et 3b ne contiennent aucune ambiguïté lexicale car aucun mot ne commence en français par la séquence phonémique /ʃadʁ/.

2a. Le livre racontait l'histoire [d'un **chat grincheux**]PP qui avait mordu un facteur.

Présence d'une ambiguïté lexicale sur /ʃagʁɛ̃/ : compétiteur « chagrin »

2b. Le livre racontait l'histoire [d'un **chat drogué**]PP qui dormait tout le temps.

Pas d'ambiguïté lexicale sur /ʃadʁ/ : aucun mot ne commence par /ʃadʁ/ en français

3a. D'après ma sœur, [le gros **chat**]PP [**grimpait** aux arbres]PP.

Présence d'une ambiguïté lexicale sur /ʃagʁɛ̃/ : compétiteur « chagrin »

3b. D'après ma sœur [le gros **chat**]PP [**dressait** l'oreille]PP.

Pas d'ambiguïté lexicale sur /ʃadʁ/ : aucun mot ne commence par /ʃadʁ/ en français

Christophe et al. (2004) ont observé que l'accès au lexique était retardé lorsqu'une ambiguïté lexicale intervenait à l'intérieur d'un syntagme phonologique (2a) alors qu'il ne l'était pas lorsque cette même ambiguïté intervenait à la frontière entre deux syntagmes phonologiques (3a). Ces résultats semblent indiquer que le processus de segmentation lexicale interviendrait donc à l'intérieur du syntagme phonologique en français.

Remarquons que dans cette étude, les frontières prosodiques (syntagme phonologique et mot prosodique) sont définies selon les algorithmes syntaxiques proposés par la Phonologie Prosodique. Pour Nespor & Vogel (1986) les mots outils et les mots à contenus forment des mots prosodiques. Ces mots prosodiques sont regroupés au sein d'un constituant prosodique plus large : le syntagme phonologique. Nespor & Vogel (1986) suggèrent que la formation des groupes phonologiques se déroule en regroupant une tête lexicale (nom, verbe, adjectif) avec les modifieurs de celle-ci situés du côté non récursif (à gauche pour le français)⁵. Prenons comme exemple la phrase « le livre racontait l'histoire d'un chat ». La structure syntaxique de cette phrase est représentée dans la phrase 4a (les têtes lexicales sont soulignées) tandis que sa structure syntaxique est représentée en 4b :

4a. **Structure syntaxique** : S[SN[Le livre]SV[racontait SN[l'histoire SP[d'un chat grincheux]]]]

4b. **Structure prosodique** : [Le livre]PP [racontait] [l'histoire]PP [d'un chat] [grincheux]PP

La structure prosodique est obtenue de la manière suivante : le français étant une langue réursive à droite, nous partons de la dernière tête lexicale « grincheux ». Un syntagme phonologique (PP) est construit en intégrant les éléments situés à gauche de la tête lexicale (du côté non récursif), jusqu'à atteindre la tête lexicale suivante donc le mot « chat » qui constitue à son tour un PP en incluant les modifieurs situés du côté non récursif de la tête lexicale (à gauche en français) et ainsi de suite.

Nespor & Vogel (1986) ont discuté la possibilité d'une règle de restructuration des syntagmes phonologiques qui inclurait le premier modifieur non-branchant du côté récursif de la tête lexicale. La règle de restructuration des groupes phonologiques énoncée par Nespor & Vogel (1986, p°173) est détaillé ci-dessous :

« A nonbranching phonological phrase which is the first complement of X on its recursive side is joined into the phonological phrase that contains X ».

Selon Nespor & Vogel cette règle est appliquée pour l'anglais et l'italien tel que ce la est illustrée dans la phrase 5 suivante (tirée de Nespor & Vogel, 1986, p°175).

5a. **Structure syntaxique** : S[SN[Le ciTTA NORrdiche] SV[non mi piacciono]]

5b. **Structure prosodique** : [Le ciTTA NORrdiche]PP [non mi piacciono]PP

⁵ Les langues sont dites réursives à droite quand on peut ajouter des éléments à la fin (comme en français ou en anglais) et réursives à gauche dans le cas contraire (comme en japonais ou en turc).

Dans cet exemple, nous pouvons voir qu'il n'y a pas de frontière de groupe phonologique entre le nom « citta » et l'adjectif « nordiche » en raison de la règle de restructuration des groupes phonologiques énoncée précédemment.

Selon Nespor & Vogel cette règle de restructuration des groupes phonologiques n'est pas applicable en français tel que cela est illustré dans les phrases 6 et 7 (tirées de Nespor & Vogel, 1986, p°179)⁶:

6. [Les maisons]PP [italiennes]PP coûtent chères.

7. Le garçon [les aidait]PP [activement]PP.

Dans leur étude proposée en 2004, Christophe et al. considèrent que le syntagme nominal « le chat grincheux » constitue un PP et assument donc (contrairement à ce qui est prédit par Nespor & Vogel (1986) mais en accord avec la proposition de Post 2000) la règle de restructuration des syntagmes phonologiques. Cependant Christophe et ses collaborateurs assument ce découpage en PP sans avoir vérifié que c'est celui-ci a effectivement été produit par le locuteur qui a lu les phrases. Il apparaît en effet que cette séquence « le chat grincheux » puisse être prononcé avec deux découpages prosodiques différents : soit comme un seul syntagme phonologique (ou un seul syntagme accentuel selon le modèle théorique considéré) soit comme deux syntagmes phonologiques (ou deux syntagmes accentuels). En effet cette séquence peut être prononcé soit avec une seule montée finale de *f₀* associée à un allongement de la syllabe (comme c'est le cas dans la figure 59a), soit avec deux montées finales de *f₀* associées à un allongement (comme cela est illustré dans la figure 59b). Ainsi dans la figure 59a. le découpage en syntagme phonologique prédit par Christophe (notée SP dans la figure) a été effectivement réalisé par le locuteur alors que ce n'est pas le cas dans la figure 59b.

⁶ Cette question a également été traitée par Post (2000) qui avance l'hypothèse que deux syntagmes phonologiques peuvent être restructurés (de manière optionnelle) si les conditions suivantes sont respectées : (i) le second PP est le premier complément du côté récursif de la première tête lexicale du premier PP et (ii) la tête lexicale du premier PP est flexionnelle (Post, 2000, p°42 et p°56). Notons que les facteurs qui induisent la restructuration du PP ne sont pas explicités.

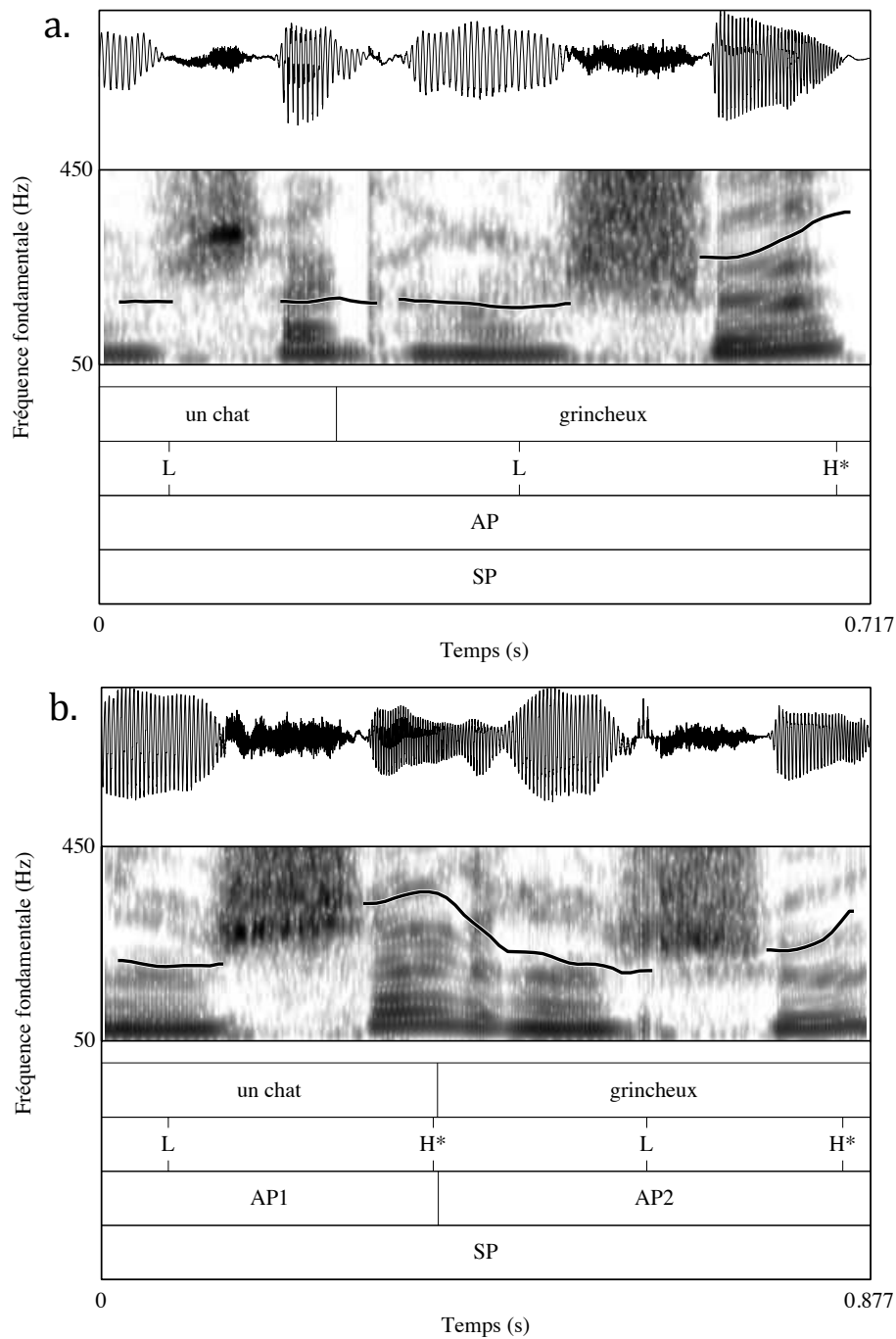


Figure 59. Exemples de deux découpages prosodiques possibles pour le même syntagme nominal « le chat grincheux ».

Dans la figure 59a, le syntagme phonologique prédit par Christophe et ses collaborateurs (noté SP dans la figure) a été réalisé comme un seul syntagme accentuel (AP) alors qu'il a été réalisé comme deux syntagmes accentuels dans la figure 59b. Les deux réalisations ont été extraites de la phrase « Le livre racontait l'histoire d'un chat grincheux qui avait mordu un facteur » lue par une locutrice de langue maternelle française à vitesse d'élocution normale.

De plus, nous avons précédemment montré que lorsque la frontière de syntagme accentuel coïncide avec une frontière de constituant syntaxique

majeur (frontière d'AP/XP⁷), le degré d'allongement observé à la frontière était plus important que celui observé à la frontière de syntagme accentuel contenu à l'intérieur d'un constituant syntaxique majeur (cf. chapitre 2). Dans l'expérimentation menée par Christophe et al. le domaine prosodique observé était un syntagme accentuel associé à une rupture syntaxique majeure. En effet dans le corpus utilisé, la frontière de syntagme phonologique était systématiquement associée à une rupture syntaxique majeure (rupture entre un syntagme nominal et un syntagme verbal). Ainsi lorsqu'une ambiguïté lexicale était contenue à l'intérieur d'un AP/XP de type [le **chat grincheux**]AP/XP, les participants mettaient plus de temps à détecter le mot « chat ». Par contre lorsque cette ambiguïté apparaissait à la frontière d'AP/XP (que Christophe et ses collègues appellent frontière de PP, ex. [ce gros **chat**]AP/XP [**grimpait**]), les temps de réaction des participants n'étaient pas retardés. Alors que Christophe et ses collaborateurs ont démontré l'influence de la frontière d'AP/XP sur les stratégies de segmentation lexicale, nous ignorons toujours si une frontière de syntagme accentuel contenue à l'intérieur d'un constituants syntaxique majeur (par exemple un syntagme nominal) est suffisante pour lever une ambiguïté lexicale en français. En d'autres termes nous demandons si c'est l'alignement entre la frontière d'AP et la rupture syntaxique majeure qui contraint l'accès au lexique en français conformément à ce qui a été observé par Christophe et ses collaborateurs. Par exemple, nous nous demandons si dans la séquence « le chat grincheux » prononcée comme deux syntagmes accentuels [[le **chat**]AP [**grincheux**]AP]XP (séquence illustrée dans la figure 59b) la frontière d'AP contenue à l'intérieure du syntagme nominal serait suffisante pour lever l'ambiguïté lexicale temporaire présente à cette frontière. Afin de mieux appréhender le rôle de la prosodie dans les stratégies de segmentation lexicale nous proposons donc à notre tour une expérience afin de déterminer le domaine à l'intérieur duquel les stratégies de segmentation lexicales opèrent. Notre but est de voir si la frontière de syntagme accentuel qui n'est pas associée à une rupture syntaxique majeure est suffisante pour lever une ambiguïté lexicale temporaire présente à cette frontière ou si seule la frontière de syntagme accentuel associée à une rupture syntaxique majeure (AP/XP) permet de lever une ambiguïté lexicale locale en français. En d'autres termes nous nous demandons si la frontière de syntagme accentuel contraint l'accès au lexique en français ou si c'est l'alignement entre la frontière d'AP avec une rupture syntaxique majeure qui joue ce rôle.

Pour cela nous avons demandé à des auditeurs du français de faire une tâche de détection de mots dans des phrases qui contenaient des ambiguïtés lexicales temporaires. Prêtons attention aux phrases 8a et 8b.

⁷ Nous appelons frontière d'AP/XP toute frontière de syntagme accentuel qui est associée à une rupture syntaxique majeure (cf. chapitre 2 de cette thèse). Notons que cette frontière ne correspond pas à un niveau additionnel de structure dans la hiérarchie des constituants prosodiques mais indique l'alignement de la frontière droite d'un syntagme accentuel avec la frontière droite d'un constituant syntaxique majeure.

8a. Le livre racontait l'histoire [d'un **chat grincheux**]AP qui avait mordu un facteur.

Présence d'une ambiguïté lexicale sur /ʃagʁɛ̃/ : compétiteur chagrin

8b. Le livre racontait l'histoire [d'un **chat drogué**]AP qui avait mordu un facteur.

Pas d'ambiguïté lexicale sur /ʃadʁ/ : aucun mot ne commence par la séquence /ʃadʁ/ en français

La phrase 8a est temporairement ambiguë en raison de la présence des syllabes « cha » et « grin » côte à côte qui forment le mot « chagrin » en français. De ce fait au point d'ambiguïté les deux candidats « chat » et « chagrin » sont potentiellement activés. En revanche la phrase 8b ne contient aucune ambiguïté lexicale car aucun mot commençant par la séquence phonémique /ʃadʁ/ n'est attesté en français. Dans cette étude, nous nous sommes focalisée sur des ambiguïtés lexicales temporaires pouvant apparaître (i) à la frontière d'un mot prosodique (MP) contenu à l'intérieur d'un syntagme accentuel, (ii) à la frontière d'un syntagme accentuel contenu au sein d'un constituant syntaxique majeur ou (iii) à la frontière d'un syntagme accentuel associée à une rupture syntaxique majeure (AP/XP). Un exemple de phrases utilisées dans le corpus en fonction de l'ambiguïté (phrases ambiguës/phrases non ambiguës) et de la position de l'ambiguïté (ambiguïté pouvant apparaître à la frontière d'un mot prosodique ou à la frontière d'un AP/XP) sont illustrées dans le tableau 9 et la figure 60.

	Phrases ambiguës	Phrases non ambiguës
Frontière de MP	...[un CHAT grincheux]AP... L H*	...[un CHAT dro*gué]AP... L H*
Frontière d'AP	...[un CHAT]AP [grincheux]AP... L H* L H*	...[un CHAT] [dro*gué]AP... L H* L H*
Frontière d'AP/XP	...[son CHAT]AP/XP[grimpait]AP... L H* L H*	...[son CHAT]AP/XP[dré*ssait]AP... L H* L H*

Tableau 9. Étude 3 : Exemple d'énoncés utilisés dans le corpus pour les différentes conditions.

L'ambiguïté lexicale (phrases ambiguës vs. phrases non ambiguës) est croisée avec le type de frontière prosodique où elle pouvait apparaître (frontière de MP, frontière d'AP ou frontière d'AP/XP).

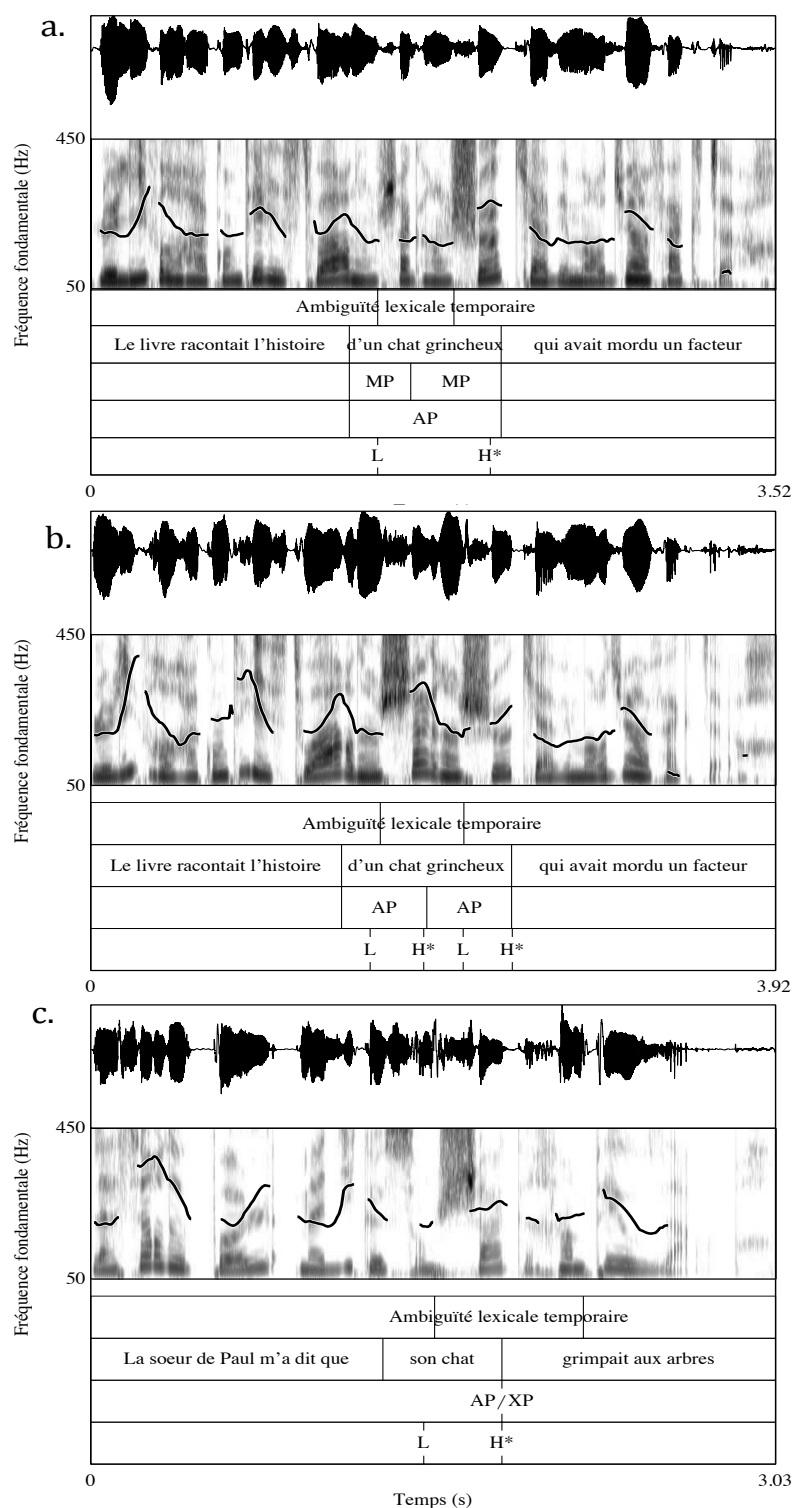


Figure 60. Étude 3: Exemples d'ambiguïtés lexicales temporaires pour les trois conditions prosodiques.

Les ambiguïtés lexicales sont liées à la présence des syllabes « cha » et « grin » côte à côte qui forment le mot « chagrin » en français et pouvaient apparaître soit à une frontière de mot prosodique (MP) au sein d'un syntagme accentuel (60a) soit à une frontière de syntagme accentuel contenue à l'intérieur d'un constituant syntaxique majeur (AP, 60b) soit à une frontière de syntagme accentuel associée à une rupture syntaxique majeure (AP/XP ; 60c).

Conformément à ce qui avait été observé par Christophe qui avait montré que la frontière de syntagme phonologique permettait de lever une ambiguïté lexicale temporaire en français, nous nous attendions à ce que la frontière d'AP/XP permette de lever le même type d'ambiguïté. Notre prédiction était que, lors d'une tâche de reconnaissance de mots, l'accès au lexique serait retardé lorsque l'ambiguïté était contenue à l'intérieur d'un syntagme phonologique indépendamment de la présence d'une frontière droite d'AP mais pas lorsque cette même ambiguïté apparaissait à la frontière d'un syntagme accentuel associé à une rupture syntaxique majeure (AP/XP).

De plus nous avons testé si la frontière d'AP qui était contenue au sein d'un constituant syntaxique majeur allait induire le même effet que la frontière d'AP/XP. Autrement dit nous avons testé si seule la frontière de syntagme accentuel était suffisante pour lever une ambiguïté syntaxique temporaire ou si c'était l'alignement entre la frontière syntaxique et la frontière prosodique qui contraignait l'accès au lexique en français. Tandis que l'étude menée par Christophe et al. (2004) a permis de montrer l'influence sur l'accès au lexique d'un constituant qui a été prédit sur la base de contraintes morphosyntaxiques, nos données ont permis d'étudier la contribution des frontières prosodiques définies à partir des indices phonétiques et phonologiques réellement produits par les locuteurs.

2. MÉTHODE

2.1. Matériel expérimental

12 paires de phrases expérimentales ont été construites. Pour chaque paire de phrases, la première contenait une ambiguïté lexicale temporaire (de type « chat grincheux » où la séquence phonémique /ʃagʁɛ̃/ constitue un mot en français) alors que la seconde n'en contenait pas (séquence de type « chat drogué » qui est une séquence non-ambiguë étant donné qu'aucun mot ne commence par les phonèmes /ʃadʁ/ en français). L'ambiguïté lexicale pouvait apparaître soit à la frontière entre un adjectif et un nom (9) soit à la frontière entre un syntagme nominal et un syntagme verbal (10).

9. Le livre racontait l'histoire d'un **chat grincheux** qui avait mordu un facteur.

a. []AP

b. []AP[]AP

10. La sœur de Paul m'a dit que son **chat** **grimpait** aux arbres.

[]AP/XP[]AP

En fonction de la manière dont les phrases du type 9 ont été prononcées, l'ambiguïté lexicale pouvait apparaître soit à une frontière de mot prosodique au sein d'un syntagme accentuel (9a) soit à une frontière de syntagme accentuel contenue au sein d'un syntagme nominal (9b). En effet, le

syntagme nominal cible pouvait être prononcé soit comme un seul syntagme accentuel soit comme deux syntagmes accentuels (cf. figure 60a et 60b). Dans les phrases du type de la phrase 10, l'ambiguïté lexicale apparaissait toujours à une frontière de syntagme accentuel associée à la rupture entre le syntagme nominal entre le syntagme verbal (frontière AP/XP, voir également la figure 60c). Ainsi les ambiguïtés lexicales temporaires pouvait apparaître dans trois positions différentes : soit à la frontière d'un mot prosodique au sein d'un syntagme accentuel (frontière de MP), soit à la frontière d'un syntagme accentuel contenu au sein d'un constituant syntaxique majeur (frontière d'AP) soit à la frontière d'un syntagme accentuel aligné avec une rupture syntaxique majeure (frontière d'AP/XP).

Le mot cible était toujours un mot monosyllabique et apparaissait toujours en milieu de phrase. Le début des phrases a été construit de manière à ce que les informations sémantiques ne favorisent pas l'activation du mot cible plutôt que l'activation du mot compétiteur. Pour cela nous avons utilisé des séquences sémantiquement neutres du type « On m'a raconté que... ». Les paires de phrases étaient complètement identiques jusqu'à l'apparition du mot cible et contenaient le même nombre de syllabes. Les mots suivant le mot cible étaient différents selon si les phrases étaient ambiguës ou non mais étaient appariés en nombre de syllabes et en fréquence dans les différentes conditions prosodiques (moyenne de fréquence d'apparition des mots suivant les mots cibles pour la condition MP et la condition AP: 5.9 dans la condition ambiguë contre 6.6 dans la condition non ambiguë, $t(31) < 1$; moyenne des fréquences d'apparition des mots cibles pour la condition AP/XP : 6.7 pour la condition ambiguë contre 7.4 dans la condition non ambiguë, $t(31) < 1$). Les fréquences lexicales ont été obtenues grâce à la base de données Lexique 2 (New, Pallier, & Ferrand, <http://www.lexique.org>).

Nous avons également réalisé des statistiques sur les diphtongues présents aux frontières prosodiques observées (par exemple le diphtongue /ag/ dans « chat grincheux »). Ces statistiques ont été réalisées dans le but de vérifier si les diphtongues présents à la frontière des constituants prosodiques étaient susceptibles d'apparaître plus fréquemment à la frontière d'un mot plutôt qu'à l'intérieur d'un mot ce qui aurait pu influencer sur le temps de réaction des sujets. Pour cela nous avons utilisé les fréquences de diphtongue calculées par Espesser et Rauzy (2008, non publié) obtenues à partir d'une base de données de 10 millions de mots et extrait du journal *Le Monde* publié entre 1990 et 1996. Les statistiques effectuées montrent que les diphtongues obtenus apparaissent plus fréquemment en frontière de mot plutôt qu'à l'intérieur des mots et ce, quelque soit la condition prosodique (la valeur $1.00E+07$ est la valeur numérique donnée pour l'infini, moyenne des fréquences d'apparition des diphtongues pour les conditions frontière de MP ou frontière d'AP: 0.002 à l'intérieur d'un mot contre 0.0037 en frontière de mot, $t(31) > 1$; moyenne d'apparition des diphtongues pour la condition frontière d'AP/XP : 0.0018 contre 0.0045, $t(31) > 1$). Ceci a potentiellement pu amoindrir l'effet d'ambiguïté que nous avons pu observer. Nous reviendrons sur ce point dans la discussion.

Afin de s'assurer de la plausibilité des phrases construites ainsi que du fait que les compétiteurs pouvaient être activés dans les phrases contenant des ambiguïtés, nous avons, pour chaque paire de phrases, construit une troisième phrase contenant le compétiteur (par exemple le compétiteur « chagrin » pour les phrases contenant la séquence « chat grincheux »). Ces phrases étaient strictement identiques aux deux autres membres du trio de phrases jusqu'à l'apparition du mot cible et contenaient le même nombre de syllabes.

Un groupe de 10 participants a jugé de la plausibilité des phrases sur une échelle de 0 à 10 (0 indiquant une phrase totalement invraisemblable en français et 10 indiquant une phrase très plausible). Les phrases contenant les compétiteurs ont été jugées aussi plausibles (moyenne de 6.2 pour les compétiteurs en frontière de MP ou d'AP et 6.1 pour les compétiteurs en frontière d'AP/XP) que celles qui contenaient les mots cibles (moyenne de 6.4 pour les mots cibles en frontière de MP et 6.0 pour les mots cibles en frontière d'AP/XP). Les phrases ambiguës (moyenne de 6.5 pour les mots cibles en frontière de MP ou d'AP et de 6.1 pour les mots cibles présentés en frontière d'AP/XP) ont également été jugées aussi plausibles que les phrases non-ambiguës (moyenne de 6.3 pour les mots cibles en frontière de MP et d'AP et de 6.0 pour les mots cibles en frontière d'AP/XP).

Nous avons demandé à une locutrice de langue maternelle française, de produire ces phrases de manière la plus naturelle possible avec une vitesse d'élocution normale. Pour chaque énoncé nous avons observé en direct le découpage prosodique produit par la locutrice à l'aide du logiciel PRAAT. Grâce à l'observation des spectrogrammes et de la courbe de *f₀*, nous avons sélectionné les énoncés correspondant au mieux au découpage prosodique attendu dans les différentes conditions prosodiques. De cette manière nous avons obtenu 24 phrases expérimentales où le mot cible était associé à une frontière de mot prosodique, 24 phrases où le mot cible était associé à une frontière d'AP et 24 phrases où le mot cible était associé à une frontière d'AP/XP pour un total de 72 phrases expérimentales. Une analyse acoustique des stimuli est présentée dans la section 3.1 suivante.

24 phrases distractrices ont été additionnées aux 72 phrases expérimentales ainsi obtenues. Parmi ces 24 distracteurs, 12 contenaient le mot cible que les participants avaient à identifier (6 d'entre eux contenaient le mot cible au début et les 6 autres contenaient le mot cible à la fin de la phrase). Ces phrases ne présentaient aucune ambiguïté. Les 12 autres phrases distractrices ont été conçues comme des pièges utilisés afin de contraindre les sujets à répondre au mot et non à la syllabe. Les distracteurs ne contenaient pas les mots cibles que les sujets avaient à identifier mais d'autres mots contenant des syllabes homophones au mot cible. Par exemple les participants devaient identifier le mot chat dans la phrase « Il était installé comme un **pacha** dans son nouvel appartement ». Pour la moitié de ces phrases, les syllabes homophones étaient situées en position initiale du mot (comme par exemple dans la séquence « un uniforme **cha**marré »), pour l'autre moitié les syllabes homophones étaient situées en position finale de mot (comme par

exemple dans la séquence « comme un pacha »). 6 listes de phrases ont été construites de manière à ce que chaque mot cible apparaisse dans une seule liste expérimentale. Chaque liste expérimentale contenait des phrases ambiguës et des phrases non ambiguës issues des trois conditions prosodiques (frontière de MP, frontière d'AP et frontière d'AP/XP). Chaque liste était composée de 12 phrases expérimentales et des 24 distracteurs. Les distracteurs étaient les mêmes pour toutes les listes de phrases. À l'intérieur de chaque liste les phrases apparaissaient dans un ordre aléatoire différent pour chaque sujet.

En plus des phrases expérimentales et distractrices, 8 phrases ont été créées et utilisées au début de l'expérience comme entraînement. Ces phrases étaient similaires aux phrases expérimentales et aux distracteurs

2.2. Participants et Procédure

60 participants de langue maternelle français ont pris part à l'expérience (10 pour chaque liste de phrases). Ces participants étaient âgés de 17 à 29 ans avec une moyenne de 24.2 ans. Ils étaient tous étudiants de l'Université de Provence et étaient originaires de différentes régions de France. Chaque participant a été testé individuellement dans la chambre sourde du Laboratoire Parole et Langage. Les participants étaient confortablement installés face à un écran d'ordinateur avec un casque sur les oreilles. La tâche était une tâche de reconnaissance de mot. Avant de commencer chaque participant remplissait un formulaire l'assignant à un groupe qui déterminait la liste de stimuli auquel il allait être soumis. Des instructions écrites sur l'écran informaient les participants qu'ils allaient voir apparaître un mot sur l'écran puis entendre une phrase et que leur tâche était d'identifier le mot dans la phrase le plus rapidement possible. Les instructions exactes données à chaque participant sont présentées en annexe.

Une cible visuelle était présentée à l'écran pendant 1.5s (par exemple le mot « chat »). L'écran était laissé blanc pendant une seconde puis un énoncé était présenté auditivement. La cible visuelle suivante était automatiquement présentée 2.5 secondes après la réponse du participant. La procédure expérimentale était contrôlée par le logiciel E-Prime. Les temps de réaction des sujets ont été mesurés à partir du début et de la fin du mot cible. Les participants devaient traiter les 36 stimuli (12 phrases expérimentales et 24 phrases distractrices) pour que l'expérience se termine. L'expérience durait environ 5 minutes.

3. RÉSULTATS

3.1. Analyse acoustique des stimuli

Avant de tester tous les participants nous avons réalisé une série d'analyses sur les stimuli afin de nous assurer que les ambiguïtés lexicales

temporaires apparaissaient soit (i) à la frontière d'un mot prosodique contenu au sein d'un syntagme accentuel (frontière de MP), soit en frontière de syntagme accentuel contenue à l'intérieur d'un syntagme nominal (frontière d'AP), soit (iii) à la frontière d'un syntagme accentuel associée à une rupture majeure (frontière d'AP/XP). La structure prosodique de chaque énoncé a été annotée par des annotateurs (moi-même et un deuxième annotateur spécialisé en prosodie). Cette annotation a été réalisée selon le modèle de l'intonation du français proposé par Jun & Fougeron (1995, 2000, 2002). Les annotations réalisées ont montré que, dans les phrases de la condition MP (phrases ambiguës et phrases non ambiguës), le nom et l'adjectif, à la frontière desquels pouvait potentiellement être localisée l'ambiguïté, formaient un seul syntagme accentuel de type /LH*/. Dans les phrases de la condition AP (phrases ambiguës et phrases non ambiguës), le nom et l'adjectif étaient structurés en deux syntagmes accentuels. Enfin dans les phrases de la condition AP/XP (phrases ambiguës et phrases non ambiguës), le syntagme nominal et le syntagme verbal étaient séparés par une frontière de syntagme accentuel qui était alignée avec une frontière syntaxique entre un syntagme nominal et un syntagme verbal. Les structures prosodiques typiquement obtenues pour les trois conditions prosodiques sont illustrées dans la figure 60.

Nous avons également mesuré la durée des voyelles associées à la frontière prosodique où pouvaient potentiellement apparaître les ambiguïtés lexicales temporaires. Les durées moyennes des voyelles qui étaient associées aux frontières prosodiques de MP, d'AP et d'AP/XP sont représentées dans la figure 61.

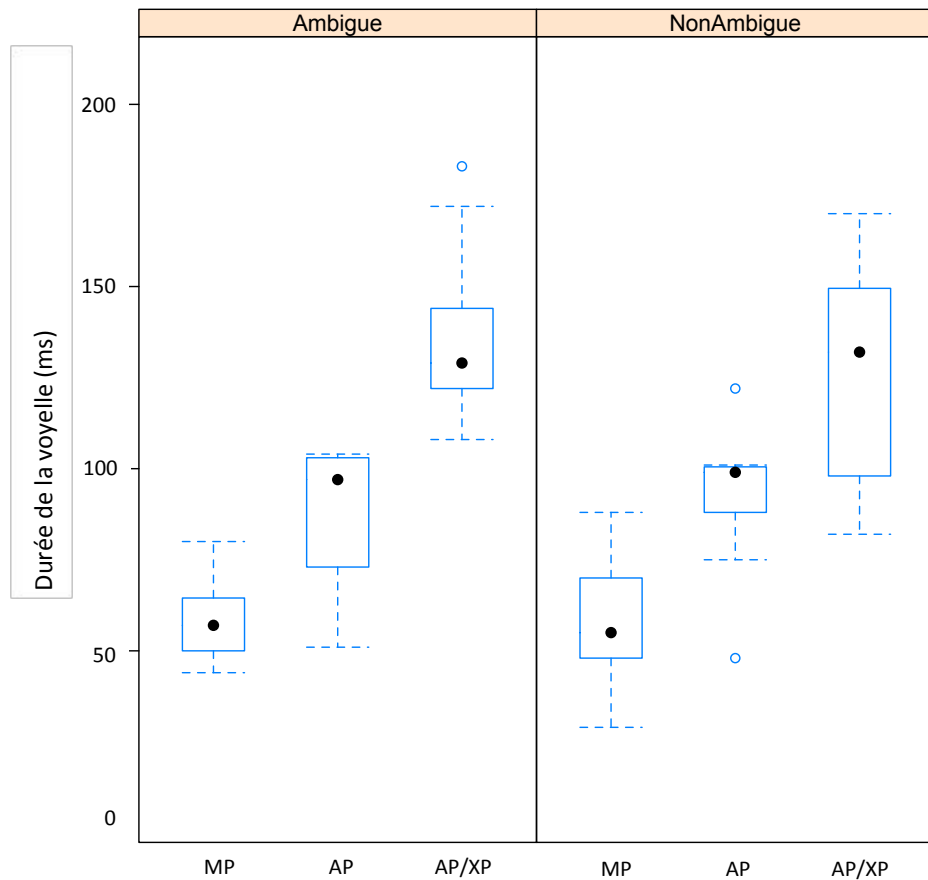


Figure 61. Étude 3 : Durées moyennes des voyelles finales des mots cibles.
 Les durées sont fonction du type de frontière prosodique où les ambiguïtés lexicales pouvaient apparaître (frontière de mot prosodique, MP, frontière de syntagme accentuel, AP, ou frontière de syntagme accentuel associée à une rupture syntaxique majeur, AP/XP).

Nous avons effectué une série de t-tests dans le but de vérifier si la durée des voyelles était différente dans les trois conditions prosodiques. Conformément à ce qui était attendu, la durée des voyelles était plus longue dans la condition AP/XP, que dans la condition AP et ce, quelque soit l'ambiguïté (pour les phrases ambiguës : $t(21.8)=3.65$, $p<0.01$, pour les phrases non-ambiguës : $t(21.908)=3.933$, $p<0.01$) et plus longue dans la condition AP que dans la condition MP (pour les phrases ambiguës : $t(21.8)=4.28$, $p<0.01$, pour les phrases non-ambiguës : $t(21.9)=3.19$, $p<0.001$). Ces résultats confirment l'annotation de la structure prosodique qui a été réalisée par les deux annotateurs. Ils sont également en lien avec les résultats de l'étude 1 présentée dans cette thèse qui montre que le degré d'allongement de la voyelle en position finale d'AP est plus important lorsque la frontière d'AP est alignée avec à une rupture syntaxique majeure plutôt que lorsqu'elle est contenue au sein d'un constituant syntaxique majeure.

Les moyennes de durée des voyelles dans les trois conditions prosodiques (frontière de MP, frontière d'AP et frontière d'AP/XP) sont représentées dans le tableau 10.

	Ambigües	Non-Ambigües
frontière de MP	62	57
frontière d'AP	102	98
frontière d'AP/XP	132	120

Tableau 10. Étude 3 : Moyenne des durées des voyelles cibles (en ms).

Les durées sont exprimées en fonction de l'ambiguïté (phrases ambiguës vs. phrases non ambiguës) et du type de frontière (frontière de MP, frontière d'AP ou frontière d'AP/XP).

Nous avons également mesuré les valeurs de f_0 associée à la syllabe où pouvaient apparaître les ambiguïtés. Les valeurs de f_0 (en Hz) sont représentées dans la figure 62 en fonction du type de frontière prosodique.

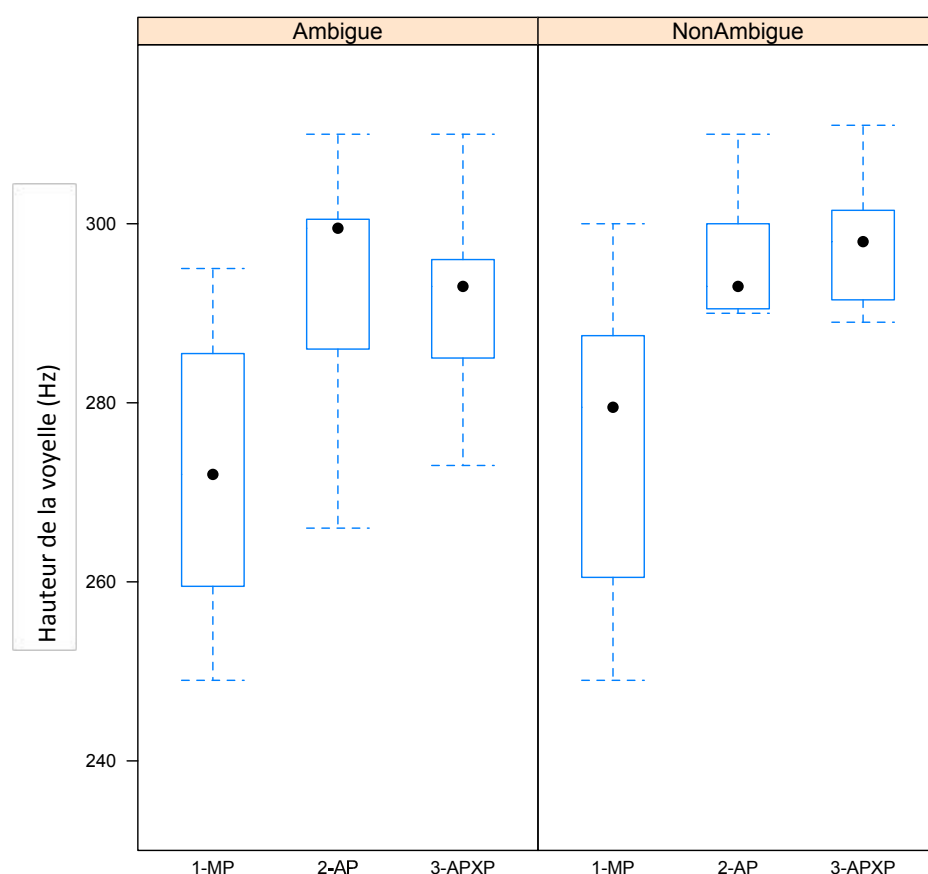


Figure 62. Étude 3 : Hauteur de la voyelle finale des mots cibles (en Hz) dans les différentes conditions prosodiques.

Conformément à ce que nous avons observé dans l'étude 1, les valeurs de f_0 présentes à la frontière d'AP et à la frontière d'AP/XP n'étaient pas

significativement différentes (pour les phrases ambiguës : $t(21.01)=0.12$, $p=0.9$; pour les phrases non ambiguës : $t(21.07)=1.4$, $p=0.16$).

Les moyennes de hauteur des voyelles (en Hz) dans les trois conditions prosodiques (frontière de MP, frontière d'AP et frontière d'AP/XP) sont représentées dans le tableau 11.

	Ambigües	Non-Ambigües
frontière de MP	273	276
frontière d'AP	292	293
frontière d'AP/XP	294	285

Tableau 11. Étude 3 : Moyenne des valeurs de f_0 associées aux voyelles cibles (en Hz). Les valeurs de f_0 sont exprimées en fonction de l'ambiguïté (phrases ambiguës vs. phrases non ambiguës) et du type de frontière (frontière de MP, frontière d'AP ou frontière d'AP/XP).

Dans le but d'étudier la contribution de la frontière droite d'AP et d'AP/XP, nous nous sommes également assurée que dans les phrases de la condition frontière d'AP et les phrases de la condition AP/XP, le patron mélodique des deux APs cibles était de type /LH*/. En effet, la présence d'un accent initial sur le deuxième AP aurait pu favoriser l'activation du mot cible plutôt que du mot compétiteur. Il a été montré que l'accent initial est exploité par les auditeurs du français comme un indice marquant la frontière gauche des mots (Welby, 2003 ; Spinelli et al. 2007 ; Spinelli et al. 2010). Nous avons donc supprimé tous les énoncés dont le syntagme accentuel suivant le syntagme accentuel cible contenait une montée initiale. Un exemple d'énoncés dont l'AP suivant l'AP cible contenait une montée initiale additionnelle de f_0 (LHi) est illustré dans la figure 63. Cet énoncé n'a pas été utilisé comme stimulus pour notre expérimentation.

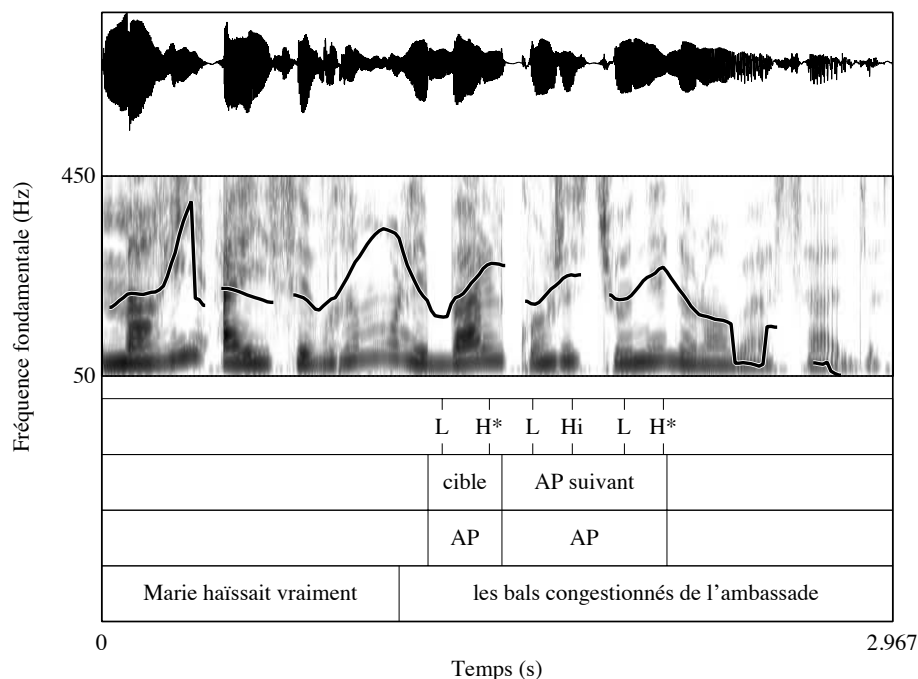


Figure 63. Étude 3 : Exemple d'énoncé présentant une montée initiale de fo (LHi) sur l'AP suivant l'AP cible.

3.2. Analyse des temps de réaction

Nous avons mesuré les temps de réaction à partir du début du mot cible. Pour chaque participant, les temps de réaction supérieurs et inférieurs à 2,5 écart-types de la moyenne ont été exclus des analyses. Adoptant ce critère seulement 2% des données ont été rejetées. Les temps de réaction en fonction de l'ambiguïté (phrases ambiguës et phrases non ambiguës) et du type de frontière prosodique (frontière de mot prosodique, frontière de syntagme accentuel et frontière de syntagme accentuel associée à une rupture syntaxique majeure) sont présentés dans la figure 64.

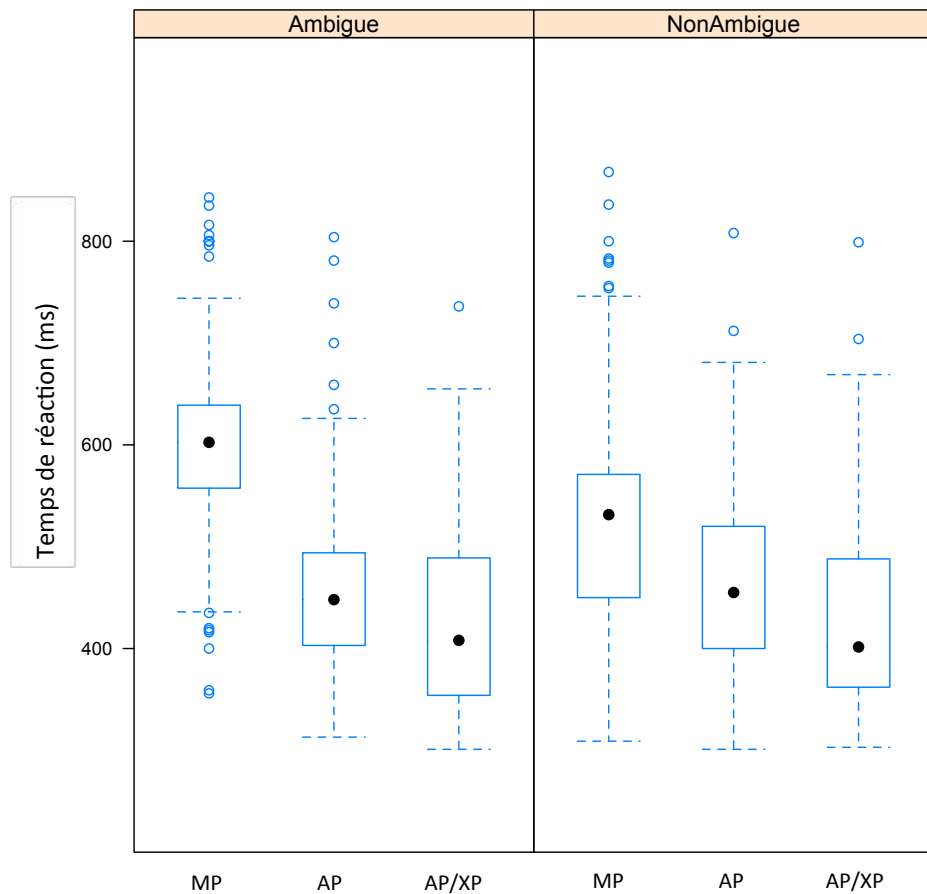


Figure 64. Étude 3 : Temps de réaction des participants.
 Les temps de réaction (en millisecondes) sont exprimés en fonction de l'ambiguïté (phrase ambiguës ou phrases non ambiguës) et du type de frontière prosodique où apparaissaient l'ambiguïté (frontière de syntagme accentuel, AP, ou frontière de syntagme accentuel alignée avec une rupture syntaxique majeure, AP/XP)

Un modèle de régressions linéaires multiples à effets mixtes (*Linear Mixed effects Models, LMM*) a été conduit sur les temps de réaction. Les effets fixes étaient l'ambiguïté (phrases ambiguës ou phrases non ambiguës) et le type de frontière prosodique (frontière de MP, frontière d'AP ou frontière d'AP/XP). Les participants et les mots cibles ont été inclus en tant qu'effets aléatoires. Le modèle incluait 706 mesures.

Les résultats montrent que les temps de réaction étaient significativement plus courts dans la condition frontière d'AP que dans la condition frontière de MP ($\beta = -175.83$, $se = 27.09$, $t = -6.683$, $pMCMC < 0.0001$) et dans la condition frontière d'AP/XP que dans la condition AP ($\beta = -108.60$, $se = 27.86$, $t = -3.911$, $pMCMC < 0.0001$) indépendamment de l'ambiguïté. L'effet d'ambiguïté n'était significatif que pour les frontières de MP avec des temps de réaction significativement plus courts pour les phrases non ambiguës que pour les phrases ambiguës ($\beta = -101.60$, $se = 21.84$, $t = -4.651$, $pMCMC < 0.0001$). L'effet d'ambiguïté n'était pas significatif dans les deux autres conditions prosodiques MP (pour la condition frontière d'AP : $\beta = -16.04$, $se = 22.02$, $t = -$

0.725, $pMCMC=0.4685$; pour la condition frontière d'AP/XP ($\beta=22.29$, $se=21.48$, $t=1.038$, $pMCMC=0.3072$). L'interaction entre l'ambiguïté et le type de frontière prosodique était significatif en raison de l'effet d'ambiguïté pour la condition frontière de MP ($\beta=85.57$, $se=31.08$, $t=-2.753$, $pMCMC<0.05$).

Les temps de réaction moyens estimés par le modèle en fonction de l'ambiguïté et du type de frontière prosodique sont représentés dans la figure 65.

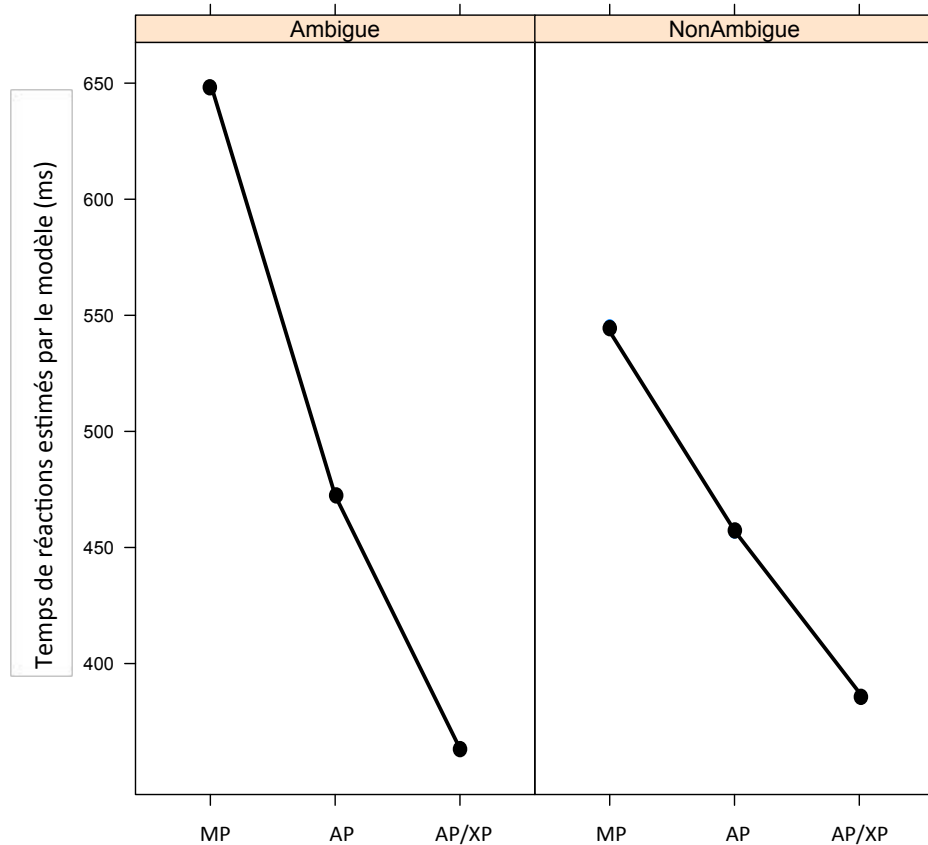


Figure 65. Étude 3 : Temps de réaction estimés par le modèle LMM.
Les temps de réaction sont fonction de l'ambiguïté (phrases ambiguës ou phrases non-ambiguës) et du type de frontière prosodique (frontière de MP, frontière d'AP et frontière d'AP/XP).

4. DISCUSSION

L'expérimentation menée au cours de ce chapitre nous a permis de mettre en évidence plusieurs résultats importants.

Tout d'abord, nous avons observé un effet significatif de l'ambiguïté lorsque celle-ci apparaissait à la frontière d'un mot prosodique. En effet, les temps de réaction des participants étaient significativement plus longs lorsqu'une ambiguïté lexicale temporaire apparaissait à la frontière d'un mot

prosodique plutôt que lorsqu'aucune ambiguïté n'était présente. Notons que nous avons observé un effet d'ambiguïté malgré que des effets de fréquence des diphtonges auraient pu, selon nos statistiques, diminuer cet effet. En effet, les diphtonges présents à la frontière des mots prosodiques (ex. /ag / dans « chat]MP grincheux » ; voir section 2.1 de ce chapitre) étaient plus fréquents en frontière de mot qu'à l'intérieur des mots. Ceci aurait pu potentiellement favoriser l'activation du mot cible plutôt que du mot compétiteur. Malgré cela, nous avons observé un effet d'ambiguïté lorsque celle-ci apparaissait à la frontière d'un mot prosodique.

Notre hypothèse de départ était que la frontière de syntagme accentuel induirait le même effet que la frontière d'AP/XP. Conformément à ce postulat, les temps de réaction des participants n'étaient pas significativement plus longs dans la condition AP pour les phrases ambiguës que pour les phrases non ambiguës. Ceci montre que la frontière d'AP contenue au sein d'un constituant syntaxique majeure a le même effet que la frontière d'AP/XP. Pour la première fois nos résultats permettent donc de montrer que la montée finale de la fréquence fondamentale ainsi que l'allongement pré-final associé à la frontière d'AP permettent de lever une ambiguïté lexicale temporaire en français et ce indépendamment de l'alignement de cette frontière avec une rupture syntaxique majeure. Contrairement à Christophe et ses collaborateurs qui ont mis en évidence l'influence de la frontière d'AP/XP définie à partir de prédictions morphosyntaxiques, notre expérimentation atteste donc de l'influence de la frontière prosodique du syntagme accentuel (définie à partir des indices acoustiques observés dans le signal de parole). Le syntagme accentuel est donc le domaine prosodique au sein duquel l'accès au lexique se réalise en français.

Nous avons également remarqué que l'ambiguïté n'avait aucune influence sur le temps de réaction des participants lorsque celle-ci apparaissait soit à la frontière d'un syntagme accentuel contenue à l'intérieur d'un syntagme nominal, soit à la frontière d'un syntagme accentuel associé à une rupture syntaxique majeure. Ces résultats vont dans le sens de ceux obtenus par Christophe et al. qui ont montré que la frontière de PP permettait de lever des ambiguïtés lexicales temporaires en français.

Enfin, indépendamment de l'effet d'ambiguïté, nous avons remarqué que les temps de réaction des participants étaient significativement plus courts dans la condition frontière d'AP/XP que dans la condition frontière d'AP. Ces résultats suggèrent que l'effet de la frontière de syntagme accentuel était renforcé lorsque cette frontière était alignée avec une rupture syntaxique majeure. Ces résultats sont en lien avec ceux obtenus dans l'étude 1 où nous avons montré que les indices de durée présents à la frontière du syntagme accentuel étaient significativement plus importants lorsque cette frontière était associée à une rupture syntaxique majeure. Il se peut en effet qu'un degré d'allongement plus important soit associé à un niveau de structure supérieur dans la hiérarchie prosodique (et donc un indice de rupture plus important) qui permettrait de lever plus efficacement l'ambiguïté lexicale temporaire.

Les résultats que nous avons obtenus au cours de cette étude ont une implication à la fois pour les études sur la structure prosodique du français mais aussi pour les modèles d'accès au lexique.

En premier lieu, seule une différence d'allongement pré-frontière permet de distinguer la frontière d'AP de la frontière d'PA/XP. Le fait que nous ayons observé des temps de réaction significativement plus courts dans la condition AP/XP que dans la condition AP montre que cette différence de degré d'allongement observée dans les conditions frontière d'AP et frontière d'AP/XP a permis aux auditeurs de segmenter plus rapidement la parole en mots. Ces résultats soulignent l'importance des indices de durée dans le marquage des constituants prosodiques en français. Cependant dans cette expérimentation la frontière d'AP était marquée à la fois par des indices de durée et des indices mélodiques. Une nouvelle série d'expérimentations où les indices mélodiques et de durée seraient manipulés de façon indépendante (par de la parole de resynthèse), permettrait de déterminer quels types d'indices acoustiques (indices mélodiques ou indices de durée) sont les plus importants dans la segmentation lexicale en français. Si nos résultats semblent aller dans le sens de Rietveld (1980) qui considérait la durée comme l'indice le plus saillant dans le but de segmenter en mots en français, notre étude ne nous permet pas de mettre en évidence un rôle plus important des indices de durée par rapport aux indices mélodiques. Une des perspectives que nous envisageons est donc de reproduire l'expérimentation que nous venons de présenter mais en manipulant de manière indépendante les indices mélodiques et les indices de durée présents à la frontière du syntagme accentuel.

Les résultats obtenus représentent également un apport significatif pour les modèles d'accès au lexique. Le fait que nous ayons obtenu des temps de réaction plus courts dans la condition AP/XP que dans la condition AP et dans la condition AP que dans la condition MP, nous permet d'affirmer que les indices acoustiques présents aux frontières d'AP et d'AP/XP ont été activés en temps réel dans le but de segmenter la parole en mots. Ces résultats nous renseignent sur le moment où les informations prosodiques sont utilisées dans le traitement lexical. Deux propositions ont été faites dans la littérature à ce sujet. Il a tout d'abord été envisagé que la prosodie soit exploitée comme une stratégie « de dernier recours » utilisée parmi un ensemble d'autres stratégies de différentes natures lorsque le processus de segmentation lexicale échoue. Selon cette proposition, les informations prosodiques n'interviendraient pas lors de l'activation des candidats lexicaux mais seraient utilisées uniquement après ce processus lorsque la suite des informations lexicales disponibles est incompatible avec le candidat sélectionné. Dans ce cas, l'information prosodique ne serait impliquée dans le processus de segmentation uniquement lorsque l'information lexicale est insuffisante pour parvenir au découpage en mots. La deuxième solution envisagée est que la prosodie soit utilisée lors du processus d'activation lexicale et qu'elle puisse influencer celui-ci. Selon cette deuxième proposition, l'information prosodique contribuerait à l'activation d'un candidat lexical plutôt que d'un autre. Les

résultats que nous avons obtenus suggèrent que l'influence des compétiteurs dépendait de s'ils étaient situés à cheval sur une frontière de MP ou d'AP. Ces résultats suggèrent donc que la prosodie peut influencer l'activation des candidats lors du processus de segmentation. Ils sont en lien avec ceux obtenus par Christophe et al. (2004) qui ont montré que les indices prosodiques présents à la frontière d'un AP/XP participent à l'activation des candidats lexicaux. Cependant, notre étude ne permet pas de déterminer quel type de rôle joue la frontière d'AP dans l'activation lexicale. En effet, deux types de rôles peuvent être envisagés. La première possibilité est que les informations prosodiques contribuent à renforcer l'activation d'un candidat plutôt que d'un autre. Dans ce cas, le mot cible (ex. « chat ») et le mot compétiteur (ex. « chagrin ») sont activés et les indices acoustiques liés à la présence de la frontière d'AP viennent, aux côtés de nombreuses autres sources d'information, renforcer l'activation de la cible plutôt que celle du compétiteur. Dans la deuxième solution envisageable, les informations prosodiques sont directement intégrées aux représentations mentales des mots. Selon cette proposition, seul le mot cible (et non le compétiteur) serait activé car les indices acoustiques présents sur la syllabe /ja/ rendent impossible l'activation du mot « chagrin » (dont la première syllabe est plus courte et n'est pas associée à une montée finale de *fo*). Ce deuxième raisonnement implique que le mot « chat » sans frontière d'AP et le mot chat associé à une frontière d'AP activeraient deux représentations mentales différentes. Les données issues de cette expérience ne nous permettent pas de pencher en faveur de l'un ou de l'autre type de relations possibles entre information lexicale et information prosodique lors du processus d'activation des candidats lexicaux. Une des solutions que nous envisageons afin d'apporter des éléments de réponse à cette question est la réalisation d'une expérience qui fait appel au paradigme d'amorçage de répétition. Cette technique consiste à présenter un premier bloc de mots à des participants sur lequel ils doivent réaliser une tâche (par exemple une tâche de répétition ou une tâche de décision lexicale). Ensuite, un second bloc de mots est présenté. Ce bloc contient à la fois des mots ayant déjà été présenté dans le premier bloc et d'autres pas. Typiquement les mots répétés sont reconnus plus rapidement que les mots non répétés en raison de l'activation de la même représentation en mémoire (cf. McLennan & Luce, 2005 ; Dufour, Nguyen & Frauenfelder, 2007). L'atténuation de cet effet lors de la modification d'une dimension particulière (ex. la présence ou non d'une frontière d'AP sur la dernière syllabe du mot) indiquerait que le même mot active deux représentations lexicales différentes. Par exemple, le mot « chat » sans frontière d'AP et le mot « chat » associé à une frontière d'AP activeraient deux représentations mentales différentes.

Enfin, les résultats que nous avons obtenus au cours de ce chapitre montrent qu'il est nécessaire de prendre en compte les informations prosodiques (à la fois les informations temporelles et mélodiques) dans la définition des constituants prosodiques. En effet, on ne peut se contenter de prédire l'apparition des frontières prosodiques en fonction d'algorithmes

morphosyntaxiques sans étudier les indices acoustiques réellement présents dans le signal de parole. Cette prise en compte devient d'autant plus nécessaire lorsqu'il s'agit d'étudier la contribution des indices prosodiques dans le traitement de la parole. Dans son étude menée en 2004, Christophe et ses collaborateurs avaient proposé que le syntagme phonologique contraigne l'accès au lexique en français. Cependant dans cette étude, le syntagme phonologique était défini en fonction de prédictions morphosyntaxiques indépendamment de ce qui a été réellement produits par les locuteurs. En observant les indices mélodiques et temporels réalisés à l'intérieur des frontières de syntagme phonologique prédites par Christophe et al., nous nous sommes aperçue que la même séquence de phonème (qui constituait un seul et même PP selon les auteurs) pouvait en fait être produite avec deux découpages en syntagmes accentuels différents selon la manière dont cette séquence était prononcée. De plus, une étude approfondie des relations entre syntaxe et prosodie nous a permis de montrer que la frontière d'AP alignée avec une rupture syntaxique majeure est marquée par un degré d'allongement plus important que la frontière d'AP contenue à l'intérieur d'un constituant syntaxique majeure. Ce constat nous a poussé à reconsidérer la proposition faite par Christophe et ses collaborateurs. Grâce à l'étude présentée dans ce chapitre, nous sommes désormais en mesure d'affirmer que c'est le syntagme accentuel (et non le syntagme accentuel associé à une rupture syntaxique majeur) qui est le domaine de l'accès au lexique en français. Nos résultats mettent donc en lumière l'importance de prendre en compte (i) la structure tonale et métrique des énoncés et (ii) l'interface syntaxe/prosodie dans la définition des constituants prosodiques et l'évaluation de leur rôle dans le traitement perceptif de la parole.

CHAPITRE 5

PERCEPTION DE LA FRONTIÈRE DROITE DE SYNTAGME INTERMÉDIAIRE ET TRAITEMENT DE LA STRUCTURE SYNTAXIQUE (Études 4 et 5)

1. PROBLÉMATIQUE

Au cours du chapitre précédent, nous avons insisté sur le fait que les locuteurs d'une langue doivent découper les phrases qu'ils entendent en mots afin de leur assigner un sens. Nous avons montré que la frontière de syntagme accentuel (et donc la structure prosodique des énoncés) contribue de manière importante à la segmentation lexicale. Cependant, au delà du découpage en mots, les auditeurs d'une langue doivent également associer ces mots les uns aux autres au sein d'une structure syntaxique. Une des questions fondamentales qui se pose aujourd'hui en psycholinguistique, concerne la manière dont cette structure syntaxique est construite. De manière plus précise, de nombreuses recherches s'intéressent aux facteurs qui vont influencer sur la construction de celle-ci et du moment où ces facteurs interviennent au cours du processus de construction. Dans ce chapitre, nous allons traiter de l'influence des frontières prosodiques sur le traitement syntaxique. Nous proposerons deux études qui montrent que certains indices acoustiques présents aux frontières de syntagme accentuel d'une part et de syntagme intermédiaire d'autre part contraignent de manière très importante la construction de la structure syntaxique des énoncés et ce, dès les toutes premières étapes de l'analyse syntaxique.

L'analyse syntaxique des phrases est généralement envisagée comme un processus composé de trois étapes : (1) la génération d'une structure ou de plusieurs structures syntaxiques possibles (selon les théories la génération d'une seule ou de plusieurs structures initiales sont admises), (2) la sélection d'une structure syntaxique et (3) la ré-analyse de la phrase si la structure

sélectionnée s'avère être incorrecte (voir par exemple Boland & Blodgett, 2001 pour plus de détails sur le sujet).

Les premiers modèles proposés dans le but de rendre compte de la construction de la structure syntaxique admettent que seules des informations de nature syntaxique sont utilisées afin de construire la structure syntaxique d'une phrase. Dans ces modèles dits modèles sériels ou modulaires (voir par exemple Frazier, 1978 ; Frazier & Fodor 1978), la construction initiale de la représentation syntaxique est élaborée au fur et à mesure que les mots de la phrase ont reconnus et repose sur les catégories grammaticales des mots qui sont traités. Une seule structure syntaxique est construite à partir des catégories grammaticales des mots, y compris si la phrase présente une ambiguïté syntaxique. Il n'y a donc pas de différence entre l'étape de génération et l'étape de sélection de la structure dans la mesure où une seule structure est générée. Si la structure syntaxique sélectionnée ne permet pas l'intégration d'un mot qui suit, la structure initialement sélectionnée est alors ré-analysée. En cas d'ambiguïté rencontrée à un moment donné dans la phrase, des principes syntaxiques généraux permettent d'opter pour la structure syntaxique la plus économique (Frazier, Clifton & Randall, 1983 ; Frazier & Rayner, 1987). Les stratégies syntaxiques générales proposées sont des principes tels que le principe d'attachement minimal ou de clôture tardive. La stratégie d'attachement minimal consiste à construire la structure syntaxique la plus simple (c'est-à-dire celle qui contient le moins de nœuds syntaxiques ; Britt, Perfetti, Garrod & Rayner, 1992 ; Feirreira & Clifton, 1986 ; Rayner, Carlson & Frazier, 1983) tandis que le principe de clôture tardive (*late closure*) conduit à choisir l'attachement le plus bas (c'est-à-dire à attacher les items lexicaux au dernier constituant ; Ferreira & Henderson, 1991 ; Frazier, 1978 ; Frazier & Rayner, 1982 ; Mitchell & Holmes, 1985). Afin d'illustrer ces principes prenons par exemple la phrase suivante :

1. *The horse raced past the barn fell.*

Le cheval qui avait été monté devant l'écurie est tombé.

Cette phrase empruntée à Bever (1970) présente une ambiguïté syntaxique temporaire liée au fait qu'au moment où le verbe « *raced* » est traité, il peut être interprété soit comme un verbe au passé (phrase active « le cheval court ») soit comme un participe passé (phrase passive « le cheval est monté »). Notons que l'ambiguïté syntaxique est dite temporaire puisqu'une seule interprétation est possible lorsque tous les mots de la phrase sont traités. Selon le principe d'attachement minimal, la structure syntaxique qui sera initialement générée est celle *du cheval qui court* (Interprétation 1) puisque cette interprétation donne la structure syntaxique la plus simple. Hors, lorsque l'auditeur rencontrera le mot « *fell* » la phrase devra être ré-analysée car seule la structure passive est compatible avec l'intégration de ce nouveau mot dans la structure syntaxique.

D'autres types de modèles postulés afin de rendre compte de la construction syntaxique des énoncés ont été proposés. Ces modèles présentent des différences majeures avec les modèles sériels précédemment évoqués (voir par exemple Jurafsky, 1996 ; MacDonald, Pearlmutter & Seidenberg, 1994 ; Tabor, Juliano & Tanenhaus, 1997). Selon ces modèles, différentes structures sont générées lorsque l'analyseur syntaxique rencontre un point d'ambiguïté dans une phrase. Les différentes structures générées reçoivent des activations différentes et la structure sélectionnée est celle ayant reçu le plus haut taux d'activation. Les facteurs entrant en jeu dans l'activation des structures syntaxiques sont des informations pouvant être de différentes natures. En effet, selon ces modèles, les informations lexicales (Ford, Bresnan & Trueswell 1989), l'attachement des verbes (Shapiro, Nagel & Levine, 1993 ; Boland, 1997), les informations relatives au contexte d'apparition de la phrase (Crain & Steedman, 1985) ou au contexte visuel (Tanenhaus, Spilvey-Knowlton, Ebhard & Sedivy, 1995), les informations relatives à la fréquence d'apparition des structures dans la langue (Jurafsky, 1996 ; MacWhinney, 1987 ; Mitchell & Cuetos, 1991) ou également les informations prosodiques seraient utilisées par les auditeurs dans le but de favoriser l'activation de l'une ou l'autre des structures prosodiques générées.

Quelque soit le modèle théorique adopté, la méthode la plus largement utilisée dans le but d'étudier la contribution des indices prosodiques dans la construction de la structure syntaxique des énoncés consiste à recourir à des phrases syntaxiquement ambiguës. Ces ambiguïtés syntaxiques peuvent être globales ou temporaires en fonction du fait que l'ambiguïté intervient tout au long de la phrase ou à un moment donné dans la phrase (appelé point d'ambiguïté). En effet, une ambiguïté syntaxique est considérée comme globale si plusieurs interprétations syntaxiques sont possibles y compris après avoir perçu l'énoncé dans sa totalité: les seules informations lexicales ne permettront pas de choisir parmi les différentes interprétations possibles. La phrase 2 est un exemple d'ambiguïté dite « globale ».

2. Jean expédie un vase de Chine

Cette phrase est syntaxiquement ambiguë car on ne sait pas si le groupe prépositionnel « de Chine » fait partie du groupe nominal « un vase » ou qualifie le verbe « expédie ». Dans le premier cas *Jean envoie un vase que vient de Chine alors que* dans la deuxième interprétation *Jean envoie un vase de provenance inconnu à partir de la Chine* (où il est supposé se situer en ce moment). Même arrivé en fin de phrase, aucune information syntaxique ne nous permet de choisir une interprétation plutôt que l'autre si ce n'est le principe de clôture tardive précédemment évoqué qui conduit à choisir préférentiellement la première interprétation (où le groupe prépositionnel appartient au groupe nominal Objet) plutôt que la seconde. Au contraire, dans la phrase 3, l'ambiguïté syntaxique est dite « temporaire » car en traitant uniquement la phrase jusqu'au point d'ambiguïté (c'est à dire avant que toutes

les informations ne soient disponibles) plusieurs interprétations syntaxiques sont possibles.

3. Jean expédie un vase de Chine où il passe actuellement ses vacances.

Par contre, seule la deuxième interprétation syntaxique reste possible lorsque tous les mots de la phrase ont été traités. La phrase 3 n'est donc pas complètement ambiguë mais revêt une ambiguïté temporaire ou locale à un point donné de la phrase.

De nombreuses études ont permis de montrer que la résolution d'ambiguïtés syntaxiques pouvant être globales ou temporaires est facilitée par la présence de frontières prosodiques au point d'ambiguïté (voir Cutler, Dahan & Van Donselaar, 1997 pour une revue détaillée sur le sujet). Par exemple, en 1991 Beach a utilisé des phrases présentant une ambiguïté syntaxique temporaire du type de la phrase 4. Dans cette phrase, l'ambiguïté est liée au fait que le groupe nominal « the gossip about the neighbors » peut être soit un complément du verbe « believed » (comme dans la phrase 4a), soit le groupe nominal sujet d'une deuxième proposition (comme dans la phrase 4b).

4a. *Jay believed the gossip about the neighbors right away.*

Jay a tout de suite cru les commérages à propos de voisins.

4b. *Jay believed the gossip about the neighbors wasn't true.*

Jay croyait que les commérages à propos des voisins n'étaient pas vrais.

L'auteur a montré qu'à l'écoute du début de ces deux séquences, dont les indices de durée et de *f₀* étaient manipulés par resynthèse, des participants devant identifier de quelle phrase venait l'extrait, prédisaient plus fréquemment que le début de phrase était un complément d'objet quand la prosodie était appropriée plutôt que quand elle ne l'était pas. En 2000, grâce à l'utilisation de phrases temporairement ambiguës telles que celles illustrées en 5a et 5b, Schepman & Rodway ont montré que des participants qui entendaient des séquences temporairement ambiguës tronquées juste avant l'auxiliaire désambiguïsant (« was » ou « were ») complétaient préférentiellement ces phrases par le pluriel (« were », cf. phrase 5b) plutôt que par la complétion « singulier » (« was », cf. phrase 5a) lorsqu'ils entendaient un contour prosodique correspondant à la séquence plurielle.

5a. *The lawyer greeted the powerful barrister and the wise judge who was walking to the courtroom*

Le juriste salua le puissant avocat et le sage juge qui traversait le tribunal.

5b. *The lawyer greeted the powerful barrister and the wise judge who were walking to the courtroom.*

Le juriste salua le puissant avocat et le sage juge qui traversaient le tribunal

De plus en plus d'études, qui ont également eu recours à des ambiguïtés syntaxiques, utilisent des tâches « on-line » pour tester si les indices prosodiques sont utilisés en temps réel dans le but de désambiguïser la structure syntaxique. Grâce à la technique du suivi des mouvements oculaires (*eye tracking*), méthode permettant de suivre la position et le mouvement du regard, Krajlic et Brennan (2005) ont par exemple montré que les participants d'une expérience utilisent les indices prosodiques dans le but de désambiguïser des séquences du type. En effet, le groupe nominal *in the basket* peut se rapporter au verbe (« Met le chien, dans le panier sur l'étoile ») ou modifier le nom précédent (« Met le chien dans le panier, sur l'étoile »). Différentes images apparaissaient à l'écran dont deux qui contenaient les deux interprétations syntaxiques possibles : une image faisait apparaître un chien dans un panier et une étoile seule tandis qu'une autre montrait un panier sur une étoile. Les résultats obtenus montrent que les participants regardaient préférentiellement une image ou l'autre en fonction de la structure prosodique avec laquelle le stimulus audio était prononcé.

Dans une étude menée en 2007 sur le français, Millotte, Wales & Christophe (2007) ont montré que les frontières de syntagmes phonologiques sont utilisées dans le but de désambiguïser des séquences syntaxiquement ambiguës. Nous renvoyons le lecteur à la section 1 du chapitre précédent pour une définition précise du syntagme phonologique tel qu'elle a été proposée dans le cadre de la phonologie prosodique. Notons que conformément à ce qui a été observé par Christophe et al. en 2004, les auteurs assument (contrairement à ce qui est prédit par Nepor & Vogel (1986) mais en accord avec la proposition de Post 2000) la règle de restructuration des syntagmes phonologiques. Les auteurs ont utilisé des paires de mots homophones provenant de deux catégories différentes telles que l'adjectif «dur» ou le verbe conjugué «dure» dans le but de créer des phrases contenant des ambiguïtés syntaxiques temporaires. Le type de phrase utilisé est illustré dans les exemples 6a et 6b.

6a. Les pommes durent plus longtemps que les bananes si on les conserve à l'abri de la lumière.

6b. Les pommes dures font de meilleures tartes que les golden par exemple.

Dans les phrases suivantes le mot «dur» est temporairement ambigu car il peut être soit un nom, soit un verbe. Les deux séquences sont strictement identiques jusqu'au point d'ambiguïté du point de vue segmental mais différent du point de vue de leur structure prosodique ([Les pommes]PP [durent...]PP, cf. phrase 6a, vs. [les pommes durent]PP, cf. phrase 6b). Grâce à une tâche de complétion où les participants entendaient les débuts des deux séquences tronquées juste après le mot ambigu et devaient les compléter à l'écrit, les auteurs ont montré que les participants utilisent la frontière de syntagme phonologique (PP) dans le but de désambiguïser ces séquences et d'assigner correctement au mot ambigu sa catégorie grammaticale. En 2008, Millote, René, Wales & Christophe ont manipulé la saillance des indices présents à la frontière de syntagme phonologique en demandant à un locuteur expert soit de maximiser l'information prosodique présente à la frontière de syntagme phonologique, soit de minimiser celle-ci :

« Elle essayait soit de produire des indices prosodiques forts pour désambiguïser le mieux possible des phrases expérimentales, soit de minimiser les indices prosodiques entre ces couples de phrases ambiguës. Dans tous les cas les stimuli résultant étaient possibles : aucune prosodie n'était agrammaticale » (Millote 2005, p° 48).

Dans une tâche de détection de mots, les participants entendaient les séquences syntaxiquement ambiguës, tronquées juste après le mot ambigu (séquences illustrées en 7a et 7b) et d'autres séquences non ambiguës (8a, 8b).

7a. Le petit chien mord la laisse qui le retient.

7b. Le petit chien mort sera enterré demain.

8a. Parfois on se mord la langue quand on mange trop vite.

8b. Maintenant qu'il est mort, les batailles d'héritage vont commencer.

Pour les deux types de séquence (ambiguës et non ambiguës) la tâche était de détecter des entrées lexicales abstraites telle qu'un verbe à l'infinitif pour les phrases 7a et 8a et de courtes séquences pronom-copule-adjectif telles que « il est mort » pour les phrases 7b et 8b. Conformément à l'hypothèse de départ, les temps de réaction des participants étaient plus élevés pour les phrases ambiguës que pour les phrases non ambiguës. Les indices de frontière de groupe phonologique ne permettaient donc pas de désambiguïser totalement les phrases expérimentales. Cependant, la différence d'ambiguïté était modulée par la saillance des indices prosodiques produits. Autrement dit la différence de temps de réaction entre les phrases ambiguës et non ambiguës était plus importante pour les phrases où les indices prosodiques étaient « maximisés » que pour les phrases où les indices prosodiques étaient

minimisés. L'effet d'ambiguïté a donc été réduit pour les phrases où la prosodie était « plus marquée » que pour les phrases où elle l'était moins. Les auteurs en ont conclu que même si les indices prosodiques présents à la frontière de syntagme phonologique ne sont pas totalement désambiguïsant, ils sont clairement exploités par les participants. Une analyse de temps de réponses données avant la fin du mot ambiguë montre que dans 39% des cas, les participants sont capables d'assigner une catégorie morphosyntaxique au mot ambiguë avant même d'avoir entendu la fin du mot. Les auteurs interprètent ce résultat en disant que les locuteurs sont capables de décider si le mot ambigu est un nom ou un adjectif avant même d'avoir étendu ce mot.

Si cette étude nous paraît intéressante car elle pose la question de la contribution des indices prosodiques dans la construction de la structure syntaxique en français, nous pensons que la méthodologie employée pose plusieurs problèmes. Tout d'abord, les frontières prosodiques sont définies en fonction des prédictions morphosyntaxiques indépendamment de la manière dont les énoncés sont réellement prononcés. Nous avons déjà évoqué au cours du chapitre précédent le fait qu'une frontière de syntagme phonologique telle qu'elle est définie par Christophe et al. (2004) ou Millote et al. (2005, 2007, 2008) ne coïncide pas forcément avec les indices mélodiques effectivement produits dans le signal acoustique. Plus spécifiquement un syntagme nominal en français, composé d'un nom et d'un adjectif, peut être réalisé comme un ou plusieurs syntagmes accentuels. Les frontières qui sont observées par Millote et ses collègues (frontières de syntagme phonologique) sont donc des frontières syntaxiques et les indices prosodiques observés à ces frontières sont des indices qui peuvent être rattachés à différents niveaux de structure dans la hiérarchie prosodique. Afin de mieux cerner le problème, nous avons demandé à une locutrice experte de prononcer plusieurs fois la phrase 9a. tirée du corpus utilisé par Millote et ses collègues (2005, 2007, 2008) :

9a. Ce grand écrivain célèbre la naissance de Victor Hugo avec son nouveau livre.

9b. Ce grand écrivain célèbre fera une dédicace demain dans une grande librairie.

Dans ces phrases le mot « célèbre » est syntaxiquement temporairement ambiguë car il peut être un verbe (9a) ou un nom (9b).

Les ondes sonores, spectrogrammes et courbes de fréquence fondamentale obtenues pour deux versions de la phrase 9a. sont représentés dans la figure 66.

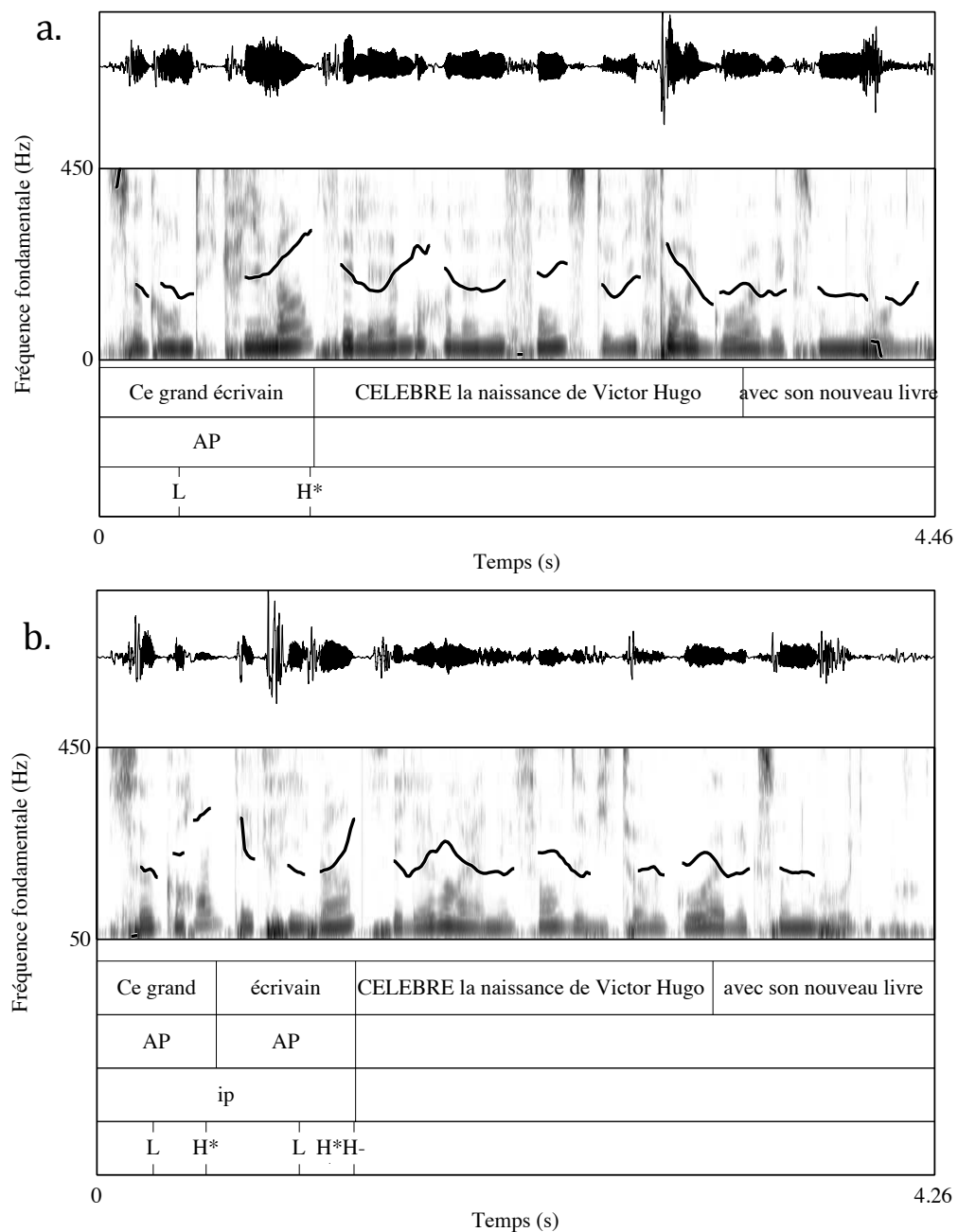


Figure 66. Deux réalisations possibles de la même phrase induisant deux structures prosodiques différentes. La phrase est tirée du corpus de Millotte (2008). Deux réalisations possibles de la phrase « Ce grand écrivain célèbre la naissance de Victor Hugo avec son nouveau livre » où le syntagme nominal « Ce grand écrivain » est réalisé soit comme un seul syntagme accentuel (66a), soit comme deux syntagmes accentuels (66b). Les deux phrases ont été lues par la même locutrice de langue maternelle française à vitesse d'élocution normale.

Dans la figure 66 nous pouvons observer que le syntagme nominal « Ce grand écrivain » peut être prononcé soit comme un seul AP (66a)⁸ soit comme deux APs (66b). La frontière prosodique située juste avant le mot syntaxiquement ambigu « célèbre » peut être soit une frontière de syntagme accentuel associée à une rupture syntaxique majeure si le syntagme phonologique est réalisé comme un seul AP (figure 66a) soit une frontière de syntagme intermédiaire (figure 66b) si le syntagme accentuel est réalisé comme deux APs. En effet dans le deuxième cas, le syntagme phonologique est suffisamment long (il contient deux APs) dans le but de faire émerger une frontière d'ip marquée par une accent de syntagme (H- dans la figure 66b) alors que ce n'est pas le cas dans la figure 66a où le syntagme phonologique est constitué d'un seul AP dont la frontière droite est associée à un accent (H* dans la figure 66a). Nous renvoyons le lecteur au chapitre 3 pour plus de détails sur la définition de la frontière d'ip. Notons que dans la figure 66b nous pouvons remarquer un retour à la ligne de registre se produisant sur le deuxième accent mélodique de l'énoncé en raison de la présence de la frontière d'ip alors que ce retour à la ligne de référence du registre n'est pas possible dans la figure 66a en raison de la présence d'un seul AP au sein du syntagme phonologique.

Dans la mesure où Millotte et ses collègues (Millotte et al., 2005, 2007, 2008) observent les indices acoustiques présents à des frontières définies sur la base de critères morphosyntaxiques et dans la mesure où ces indices acoustiques observés à ces frontières peuvent marquer des frontières prosodiques de différents niveaux dans la hiérarchie prosodique, les résultats obtenus dans ces études sont à nuancer. En effet, nous ne savons toujours pas à ce jour quel est le domaine prosodique permettant de lever une ambiguïté syntaxique en français. Serait-ce le syntagme accentuel, le syntagme intermédiaire? De plus, la manipulation de la saillance de ces indices prosodiques observés par Millotte et al. (2008) a pu entraîner des phénomènes de restructuration prosodique qui confondent également les résultats obtenus. En effet, en demandant à une locutrice de minimiser ou de maximiser les indices présents aux frontières de syntagme phonologique, cela a pu induire des phénomènes de restructuration prosodique (des phénomènes d'effacement de frontière ou de remplacement de frontière par une frontière de niveau différent dans la hiérarchie prosodique) qui ont pu avoir un impact sur les résultats. Par exemple, lors de la prononciation de la phrase illustrée dans la figure 66a, si on demande à une locutrice de minimiser les indices prosodiques, la frontière de syntagme accentuel réalisée sur la syllabe finale du mot « écrivain » pourrait être effacée ce qui fait que les deux syntagmes accentuels « ce grand écrivain » et « célèbre » pourraient être

⁸ Remarquons que dans la figure 66b nous considérons que la syllabe « grand » est associée à un accent mélodique final et non à un accent initial. En effet, la voyelle est de durée similaire à la voyelle finale du mot « écrivain » (qui est associée à un accent final) et il est clairement possible d'identifier deux montées mélodiques qui sont de forme similaire se rapprochant plus de deux montées LH* que de la séquence tonale LHiLH*. Nous renvoyons le lecteur à Welby (2006) pour plus de détails sur les différences entre montées initiales et finales en français.

réalisés comme un seul syntagme accentuel « ce grand écrivain célèbre ». De façon similaire, en demandant à une locutrice de maximiser les indices prosodiques, la frontière de syntagme intermédiaire associée à la syllabe finale du mot « écrivain » (voir figure 66b) pourrait être remplacée par une frontière prosodique de niveau supérieur dans la hiérarchie (une frontière de syntagme intermédiaire). Pour d'autres langues que le français nous savons que les frontières de syntagme intonatif influent de manière importante sur la construction syntaxique des énoncés (voir par exemple Price, Ostendorf, Stattuck-Hufnagel & Wong, 1991 ; Kjergard & Speer, 1999 ; Schepman & Rodway, 2000). Nous sommes donc en droit de nous demander quel type de frontière prosodique était réellement présent dans le corpus de Millotte et al. (frontière de syntagme accentuel, frontière de syntagme intermédiaire ou frontière de syntagme intonatif) et donc quel est le domaine prosodique permettant de lever une ambiguïté syntaxique temporaire en français.

Si nous nous ignorons encore à ce jour quel est le type de frontière prosodique (c'est à dire le niveau de hiérarchisation de frontière dans la structure prosodique) permettant de contraindre l'analyse syntaxique des phrases en français, la question du moment où les indices prosodiques interviennent dans le traitement syntaxique est également une question qui demeure non-résolue. Contrairement à ce qui avait été proposé par les premiers modèles sur la construction de la structure syntaxique, nous savons aujourd'hui que les informations syntaxiques ne sont pas les seules à guider les premiers choix des structures syntaxiques. En effet, l'utilisation d'indices prosodiques peut prédominer et annuler les préférences d'interprétation induites par les principes syntaxiques tels que les principes de clôture tardive ou d'attachement minimal. Cependant le fait que les indices prosodiques contribuent à la construction de la structure syntaxique n'implique pas nécessairement le fait qu'ils soient utilisés lors de la phase initiale de structuration syntaxique. Rappelons que dans les cas où l'analyseur syntaxique fait face à une ambiguïté, plusieurs structures syntaxiques sont générées. La structure sélectionnée sera celle ayant reçu le plus haut taux d'activation. Deux solutions sont en fait possibles : soit l'information prosodique va contribuer à renforcer l'activation d'une structure particulière et dans ce cas, le rôle de l'information prosodique serait donc, au côté de diverses sources d'informations, de renforcer le niveau d'activation d'une structure plutôt que d'une autre, soit l'information prosodique va entraîner la génération d'une seule structure prosodique possible ce qui reviendrait à admettre que la prosodie intervient dès le tout début du traitement syntaxique.

Comme nous l'avons évoqué tout au long de ce chapitre, la majorité des travaux dans le domaine ont recours à des phrases ambiguës dans le but d'analyser la contribution des indices prosodiques dans la construction de la structure syntaxique. Le problème de cette méthode est qu'il est difficile de savoir si les indices prosodiques sont utilisés dans le but de renforcer l'activation d'une structure syntaxique particulière à un point d'ambiguïté ou

si les indices prosodiques agissent très tôt et entraînent la génération d'une seule structure. Dans les deux études que nous proposons dans ce chapitre, nous abordons le problème sous un angle différent. En effet, nous avons utilisé des énoncés ne présentant aucune ambiguïté syntaxique. Notre hypothèse est que les auditeurs d'une langue, vont utiliser les indices présents à la frontière des constituants prosodiques dans le but d'inférer les catégories morphosyntaxiques des mots qui suivent et ce, avant même d'avoir entendu ces mots. De manière plus précise, nous supposons que les différents types de frontière prosodique, c'est à dire le niveau de hiérarchie du constituant auquel les frontières sont rattachées (frontière de syntagme accentuel ou frontière de syntagme intermédiaire) permettra à l'auditeur d'inférer la catégorie morphosyntaxique du constituant à venir. Prenons un exemple. Voici deux syntagmes nominaux qui sont strictement identiques du point de vue segmental mais dont la syllabe finale est associée soit à une frontière de syntagme accentuel (67a), soit à une frontière de syntagme intermédiaire (67b).

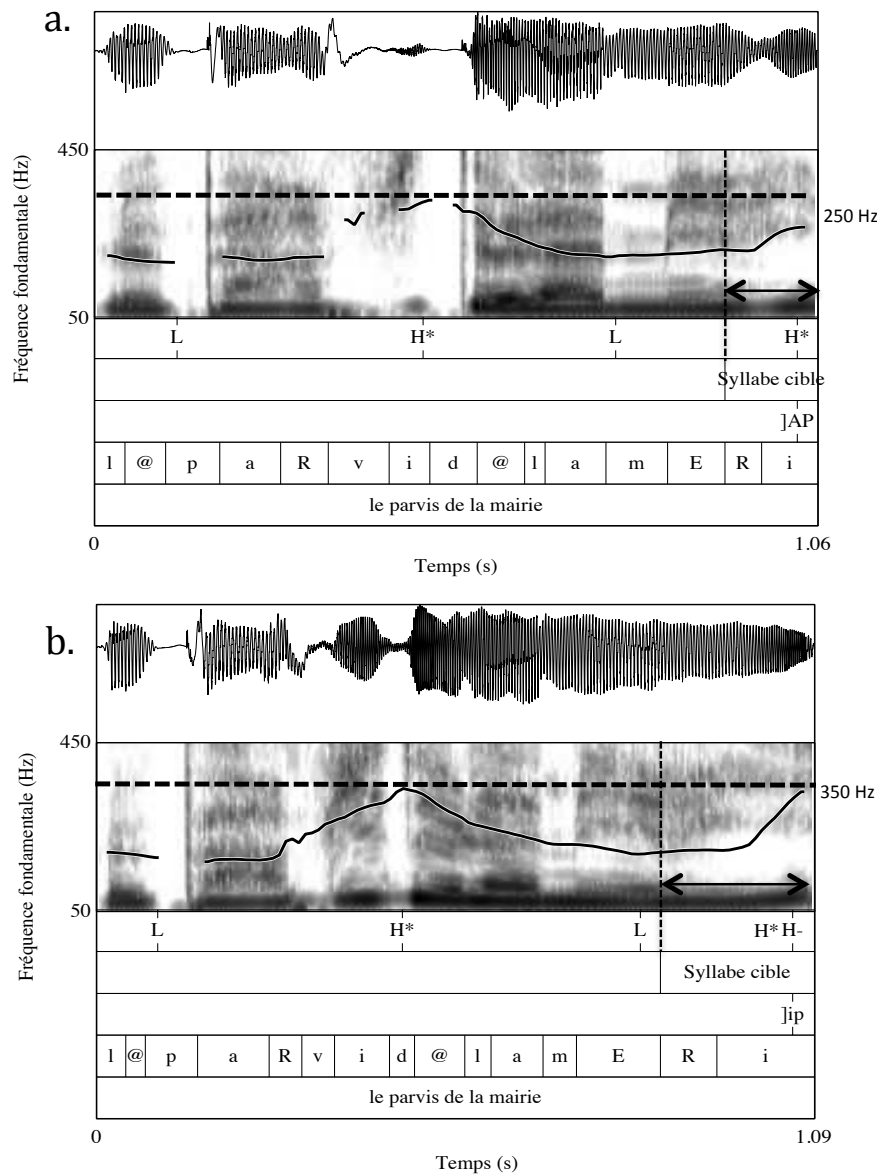


Figure 67. Étude 4 : exemples d'une séquence dont la voyelle finale est associée soit à une frontière d'AP, soit à une frontière d'ip.

Deux réalisations de la séquence « le parvis de la mairie » extraites des phrases « Le parvis de la mairie de Paris devrait être entièrement restauré l'an prochain » dans la figure 67a et « Le parvis de la mairie devrait être entièrement restauré l'an prochain » dans la figure 67b. Les deux séquences sont identiques du point de vue de leur matériel segmental mais diffèrent du point de vue de leur structure prosodique : la syllabe finale /Ri/ de la séquence est associée à une frontière de syntagme accentuel dans la figure 67a et à une frontière de syntagme intermédiaire dans la figure 67b.

Les indices acoustiques associés à la dernière syllabe des deux séquences sont différents. Dans la séquence 67a nous pouvons remarquer que la voyelle en position finale est associée à un accent mélodique (LH*) qui est abaissé relativement au premier accent mélodique de la séquence et présente

un niveau d'allongement comparable à celui observé pour les syllabes en position finale d'AP. Dans la séquence 67b, la voyelle en position finale de la séquence est associée à un accent de syntagme (H-) responsable du fait que la voyelle présente une hauteur similaire à la hauteur du premier accent mélodique de l'énoncé et présente un niveau d'allongement comparable à celui observé pour les voyelles en position finale d'ip. Notre hypothèse est que, si nous demandons à des auditeurs du français de compléter ces deux séquences, ils interpréteront la frontière de syntagme accentuel en tant que rupture syntaxique mineure et complèteront préférentiellement la séquence par un syntagme prépositionnel (comme par exemple « de Paris ») tandis qu'ils interpréteront la frontière de syntagme intermédiaire en tant que rupture syntaxique majeure et complèteront préférentiellement la séquence par un syntagme verbal (de type « devrait être rénové »). De plus, nous savons que les frontières de syntagme accentuel et de syntagme intermédiaire sont marquées par plusieurs indices tels que les indices de durée et les indices mélodiques qui ne coïncident pas toujours dans le marquage des frontières. Par exemple, au cours du chapitre 2 de cette thèse nous avons remarqué que, à vitesse d'élocution rapide, certaines frontières de syntagme accentuel n'étaient marquées que par des indices de durée et que malgré l'absence d'indices mélodiques, cela était suffisant pour que les deux annotateurs perçoivent la frontière. Nous proposons donc dans ce chapitre de voir si, lorsque les frontières de syntagme accentuel et de syntagme intermédiaire sont marquées par un seul type indice acoustique (soit le contour mélodique, soit l'allongement pré-frontière), cela suffit à guider les participants dans la construction de la structure syntaxique des phrases.

Deux études ayant pour objectifs (i) de voir si la frontière de syntagme intermédiaire contraint le traitement syntaxique des phrases et (ii) d'étudier la contribution séparée des indices mélodiques et des indices de durée dans la construction de cette structure, seront présentées. Ces deux études, qui sont les premières à adresser la question de l'influence des frontières de syntagme intermédiaire sur le traitement syntaxique des énoncés pour le français, nous permettront d'asseoir l'existence du syntagme intermédiaire tant du point de vue perceptif que du point de vue de sa validité cognitive en français.

2. ÉTUDE 4: À LA DÉCOUVERTE DE L'INFLUENCE DE LA FRONTIÈRE DROITE DU SYNTAGME INTERMÉDIAIRE DANS LA CONSTRUCTION DE LA STRUCTURE SYNTAXIQUE

Dans le but d'étudier si la frontière droite de syntagme intermédiaire est interprétée par les auditeurs français comme indiquant une rupture

syntaxique majeure (telle qu'une rupture entre un syntagme nominal et un syntagme verbal) ou une rupture syntaxique mineure (telle qu'une rupture à l'interne au syntagme nominal), nous avons utilisé des phrases dans lesquelles nous avons fait varier la taille du syntagme nominal en termes de nombre de mots. Les syntagmes nominaux étaient strictement identiques du point de vue de leur matériel segmental jusqu'à la syllabe cible, qui pouvait être associée soit à une frontière du syntagme accentuel, soit à une frontière de syntagme intermédiaire. Prêtons attention aux phrases 4a et 4b.

4a. Le parvis de la mairie]**AP** de Paris devrait être entièrement restauré l'an prochain.

4b. Le parvis de la mairie]**ip** devrait être entièrement restauré l'an prochain.

Ces deux phrases sont strictement identiques du point de vue segmental jusqu'à la syllabe cible, en gras (« ri » du mot « mairie ») qui peut être associée soit à une frontière de syntagme accentuel (comme c'est le cas dans la phrase 4a), soit à une frontière de syntagme intermédiaire (comme c'est le cas dans la phrase 4b). Les ondes sonores, spectrogrammes et courbes de fréquence fondamentale de ces deux phrases tronquées après la syllabe cible sont illustrés dans la figure suivante.

Dans la figure 67 présentée plus haut, nous pouvons remarquer que, bien que les deux énoncés soient strictement identiques du point de vue de leurs segments jusqu'à la syllabe cible, leur structure prosodique est différente. En effet, tandis que la syllabe cible « rie » du mot « mairie » est associée à une frontière de syntagme accentuel dans la figure 67a, cette même syllabe est associée à une frontière de syntagme intermédiaire dans la figure 67b. De ce fait, les propriétés acoustiques des indices observés à la frontière du second syntagme accentuel sont différentes. La durée de la syllabe cible est plus importante dans l'énoncé en 67b que dans l'énoncé en 67a. De plus, l'accent mélodique associé à la syllabe cible est abaissé par rapport au premier accent mélodique de l'énoncé en 67a alors que ce n'est pas le cas en 67b.

Les phrases du corpus ont été coupées juste après la syllabe cible (cf. annexe F pour la série complètes des phrases utilisées). Nous avons demandé à des participants qui entendaient ces séquences tronquées de les compléter au travers d'une tâche de complétion de phrase à choix forcé. Deux choix possibles étaient présentés aux participants : une complétion constituée d'un syntagme prépositionnel (tel que « de Paris » comme dans la phrase 4a) et une complétion constituée d'un syntagme verbal (tel que « devait être » comme dans la phrase 4b). Nous nous attendions à obtenir plus de réponses de type syntagme verbal (SV) lorsqu'une frontière d'ip était présente dans le stimulus présenté auditivement plutôt que lorsqu'une frontière d'AP était présente. En d'autres termes nous nous attendions à ce que les participants infèrent la catégorie syntaxique des mots situés après la voyelle cible en fonction du type

de frontière prosodique (d'AP ou d'ip) qui était présente dans le stimulus audio qui leur était présenté.

2.1. Méthode

2.1.1. Matériel linguistique

20 paires de phrases ont été construites. Chaque paire de phrases contenait un syntagme nominal dont nous avons fait varier la longueur en terme de nombre de mots afin de faire varier le nombre d'APs produits au sein de ce syntagme. Les syntagmes nominaux pouvaient contenir soit un seul syntagme prépositionnel (comme par exemple « le parvis de la mairie ») soit deux syntagmes prépositionnels (comme par exemple « le parvis de la mairie de Paris »). Les deux énoncés de chaque paire de phrases étaient strictement identiques du point de vue du segmental jusqu'à la sixième syllabe des énoncés (qui constituait la syllabe cible) et ne différaient que par leur structure prosodique : la syllabe cible était associée à une frontière de syntagme accentuel dans la première phrase de chaque paire et à une frontière de syntagme intermédiaire dans la deuxième phrase de chaque paire. La syllabe cible était identique pour chaque paire de phrases et était toujours une syllabe ouverte (de type CV). Elle était toujours suivie par un syntagme prépositionnel commençant par la préposition « de » (par exemple « de Paris » ou « de Thomas ») pour le premier membre de chaque paire de phrases (condition AP) et par un syntagme verbal commençant par le phonème /d/ (tel que « devenait » ou « demandait ») pour la deuxième phrase de chaque paire (condition ip). De ce fait quelque soit la condition prosodique, la syllabe cible était toujours suivie par le phonème /d/.

En plus de ces 40 phrases expérimentales, 40 phrases distractrices ont été construites. Ces phrases ne présentaient pas de structure prosodique ou syntaxique particulière (par exemple : le nouveau professeur arrivera demain).

Pour chaque paire de phrases, nous nous sommes assurée que les continuations situées après les séquences cibles étaient plausibles. Pour cela, nous avons demandé à un groupe de 10 participants de langue maternelle française de lire les phrases des deux conditions prosodiques (frontière d'AP et frontière d'ip) et de juger la plausibilité de ces phrases sur une échelle allant de 0 (complètement invraisemblable) à 7 (complètement plausible). Les résultats de ces questionnaires montrent que les phrases de la condition AP sont jugées aussi plausibles que les phrases de la condition ip (moyenne obtenue pour les phrases de la condition AP : 6.1 ; moyenne obtenue pour la condition ip : 6.2 ; $t(31) > 1$).

Nous avons tronqué les phrases après les syllabes cibles et avons demandé à 10 participants de langue maternelle française de compléter à l'écrit ces phrases en leur demandant de choisir parmi les deux suites créées lors de la constitution du corpus. Par exemple les participants lisaient « Le

parvis de la mairie » et devaient compléter la séquence en choisissant soit la continuation en 10a soit la continuation en 10b.

10a. devrait être restauré.

10b. de Paris devrait être restauré.

Les résultats obtenus montrent qu'à l'écrit, les participants complètent préférentiellement les séquences par la continuation de type 10a c'est-à-dire par un syntagme verbal (choisie dans 81% des cas) que par la continuation de type 10b c'est-à-dire par un syntagme prépositionnel. Ce résultat n'est pas surprenant et est en accord avec la stratégie d'attachement minimal qui consiste à construire la structure syntaxique la plus simple (c'est-à-dire celle qui contient le moins de nœuds syntaxiques ; Britt, Perfetti, Garrod & Rayner, 1992 ; Feirreira & Clifton, 1986 ; Frazier & Rayner, 1972 ; Rayner et al., 1983 ; voir section 1 de ce chapitre). La question est donc de savoir si la structure prosodique des séquences cibles va influencer le choix des participants à tel point que ceux-ci choisiront préférentiellement la continuation avec un syntagme prépositionnel à l'écoute des séquences cibles associées à une frontière de syntagme accentuel.

Nous avons demandé à une locutrice française native de produire l'ensemble des phrases du corpus de la manière la plus naturelle possible et avec une vitesse d'élocution normale. Le corpus a été enregistré dans la chambre sourde du Laboratoire Parole et Langage grâce à une station d'enregistrement ZOOM HD16 en format wav 16 bit/44,1 kHz. Les énoncés ainsi produits ont ensuite été isolés manuellement et chaque énoncé sauvé séparément. Chaque fichier wav a été segmenté en phonèmes à l'aide de l'aligneur Easy align (Goldman, 2007). La segmentation a été vérifiée manuellement grâce à l'inspection des ondes sonores et des spectrogrammes pour tous les fichiers. Les courbes de *f₀* et les spectrogrammes ont été générés en utilisant PRAAT (Boersma & Weenink, 2009). Les syllabes cibles ainsi que la syllabe associée au premier accent mélodique des énoncés ont été annotées automatiquement grâce à des scripts PRAAT puis vérifiées manuellement une par une. La durée segmentale des syllabes ainsi que la durée des énoncés ont été extraites grâce à l'utilisation de scripts PRAAT qui sont consultables en ligne à l'adresse suivante: <http://amandine.michelas.pagesperso-orange.fr/Praat.htm>. Les valeurs maximales de *f₀* associées aux syllabes cibles ainsi qu'aux syllabes associées au premier accent mélodique des énoncés ont également été automatiquement annotées puis extraites grâce à l'utilisation de scripts. La structure prosodique de chaque énoncé a été vérifiée afin de s'assurer que le locuteur ait produit les structures prosodiques attendues. Nous avons ainsi obtenu des séquences cibles qui étaient toutes composées de deux APs. Par contre, la dernière syllabe de chaque séquence cible était soit associée à une frontière d'AP, soit à une frontière d'ip. Une analyse acoustique des stimuli est présentée dans les résultats.

Tous les énoncés produits ont été coupés à la fin de la cible, à un passage à zéro de la courbe d'amplitude. Les phrases distractrices ont elle aussi été coupées à la fin d'un mot qui pouvait être placé soit au début (pour 10 d'entre elles), soit au milieu (pour 10 d'entre elles), soit à la fin (pour 20 d'entre elles).

Sur la base des 40 séquences cibles (20 pour chaque condition prosodique) et 40 phrases distractrices ainsi produites par notre locutrice, deux listes de phrases ont été construites de manière à ce que chaque membre d'une paire de phrases apparaisse dans l'une ou l'autre liste. Chaque liste contenait des phrases issues des deux conditions prosodiques. La moitié des sujets a été soumise à la liste 1 et l'autre moitié à la liste 2. Cette procédure a été adoptée afin que chaque participant ne soit soumis qu'une seule fois à la même séquence cible. Ainsi chaque liste de phrases contenait les 20 séquences cibles et les 20 phrases distractrices. Les phrases distractrices étaient les mêmes dans les deux listes de phrases. À l'intérieur des deux listes de phrases, les stimuli étaient présentés de façon aléatoire.

En plus des phrases expérimentales et distractrices, huit phrases ont été créées et utilisées au début de l'expérience comme entraînement.

2.1.2. Procédure

Chaque participant était testé individuellement dans la chambre sourde du Laboratoire Parole et Langage. Les participants étaient confortablement assis face à un écran d'ordinateur avec des écouteurs. Avant de commencer chaque participant remplissait un formulaire l'assignant à un groupe qui déterminait la liste des stimuli à laquelle il allait être soumis. Des instructions écrites sur l'écran informaient les participants qu'ils allaient entendre des débuts de phrases coupées à la fin d'un mot, et que leur tâche était de compléter ces phrases en choisissant parmi deux solutions qui leur étaient proposées à l'écran. Les instructions exactes données à chaque participant sont présentées en annexe. Les participants étaient informés que les deux séquences étaient grammaticalement correctes mais qu'ils devaient choisir celle qui leur semblait la plus appropriée en fonction de la séquence qu'ils entendaient.

Un essai commençait avec la présentation auditive d'un stimulus, puis, 2 secondes après la fin de ce stimulus, les deux complétions apparaissaient à l'écran (l'une à droite de l'écran et l'autre à gauche). Les participants devaient appuyer sur le bouton droit ou gauche d'un boîtier de réponse afin d'indiquer leur choix. La position droite ou gauche des deux stimuli à l'écran a été manipulée de manière à ce que le même type de complétion apparaisse tantôt du côté droit, tantôt du côté gauche de l'écran. Les participants n'entendaient qu'une seule fois chaque stimulus. Ils devaient traiter les 60 stimuli (20 phrases expérimentales et 40 phrases distractrices) pour que l'expérience se termine. L'expérience durait environ 15 minutes.

2.1.3. Participants

36 participants de langue maternelle française ont participé à l'expérience (18 par liste expérimentale). Ces participants étaient âgés de 18 à 25 ans avec une moyenne de 22.6 ans. Ils étaient tous étudiants de l'Université de Provence et étaient originaires de différentes régions de France. Aucun participant n'a rapporté de problèmes auditifs ou de troubles du langage. A la fin de l'expérience nous avons demandé aux participants s'ils avaient compris le but de l'expérimentation. Certains avaient repéré qu'en fonction de la manière dont était prononcé l'item, ils choisissaient une réponse plutôt qu'une autre mais ne semblaient pas avoir mis en place de stratégie de réponse particulière. Ils disaient souvent avoir sélectionné la première suite qui leur venait à l'esprit.

2.2. Résultats

2.2.1. Analyse acoustique des stimuli

Avant de tester les participants, nous avons réalisé des analyses acoustiques sur les stimuli afin de nous assurer qu'ils contenaient les structures prosodiques attendues. C'est ce que nous présentons maintenant.

Dans un premier temps, l'intonation de chaque énoncé a été annotée par nous-même et une autre spécialiste en prosodie. Le premier annotateur a annoté tous les fichiers qui ont été ensuite vérifiés par le second annotateur. Les deux annotateurs n'ont été en désaccord sur aucun des fichiers. L'annotation de l'intonation a été réalisée en suivant le modèle de l'intonation du français proposé par Jun & Fougeron (1995, 2000, 2002). Les 20 stimuli de la condition AP présentaient la structure intonative suivante : deux syntagmes accentuels présentant tous deux le patron mélodique /LH*/. Les 20 stimuli de la condition ip étaient composés de deux syntagmes accentuels de type /LH*/, le dernier ton H de la séquence étant superposé à un accent de syntagme (H-) marquant une frontière de syntagme intermédiaire (cf. figure 67b).

Nous avons également effectué différentes mesures de durée et de hauteur de la fréquence fondamentale. Nous avons tout d'abord mesuré la durée des voyelles situées (i) en position finale des stimuli et (ii) associées au premier accent mélodique des stimuli. Nous avons concentré nos analyses sur la durée des voyelles et non sur la durée des syllabes cibles car nous savons que l'allongement pré-frontière affecte principalement la rime de la syllabe plutôt que l'attaque (en anglais : White, 2002 ; Turk & Shattuck-Hufnagel, 2007 ; en français : Astézano, 2001). Nous avons également relevé le maximum de *f*₀ associé à chacune de ces deux voyelles pour chaque stimulus. Nous avons respectivement codé V1 et H1, la durée et la valeur maximale de *f*₀ de la voyelle associée au premier accent mélodique des stimuli et V2 et H2, la durée et la valeur maximale de *f*₀ associées à la voyelle finale des stimuli des

deux conditions prosodiques (AP et ip) tel que cela est illustré dans la figure 68.

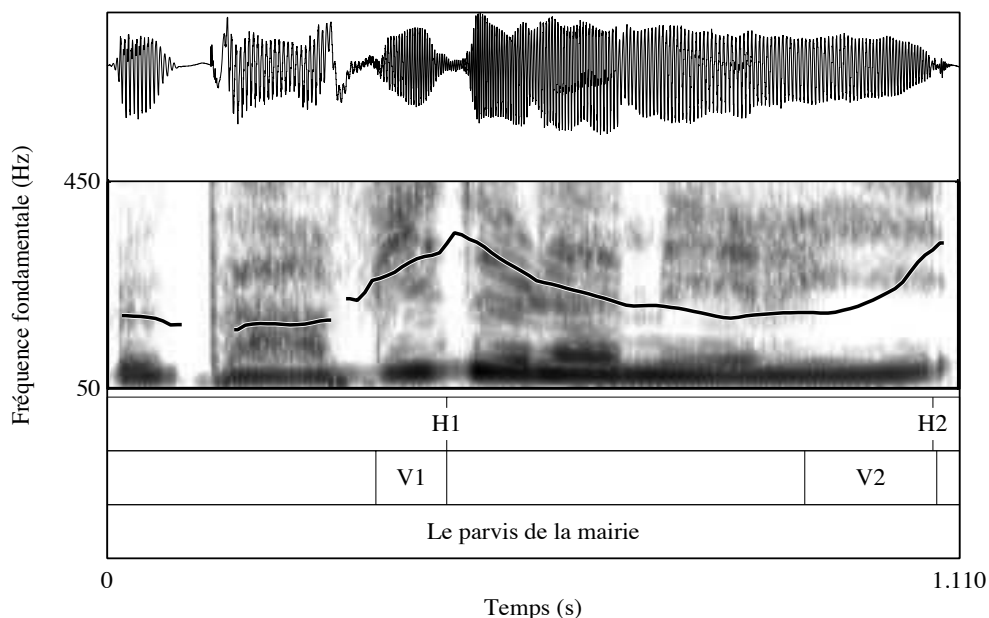


Figure 68. Étude 4 : Annotations réalisées.

Annotations des voyelles cibles (V1 et V2) et des valeurs maximales de f0 (H1 et H2) pour la séquence « le parvis de la mairie de paris » extraite de la phrase « Le parvis de la mairie devrait être entièrement restauré l'an prochain ». V1 correspond à la voyelle associée au premier accent mélodique de la séquence tandis que V2 correspond à la voyelle finale de la séquence pouvant être associée soit à une frontière de syntagme accentuel (condition AP), soit à une frontière de syntagme intermédiaire (condition ip).

Les durées moyennes de V1 et V2 en fonction de la condition prosodique (AP et ip) sont illustrées dans la figure 69.

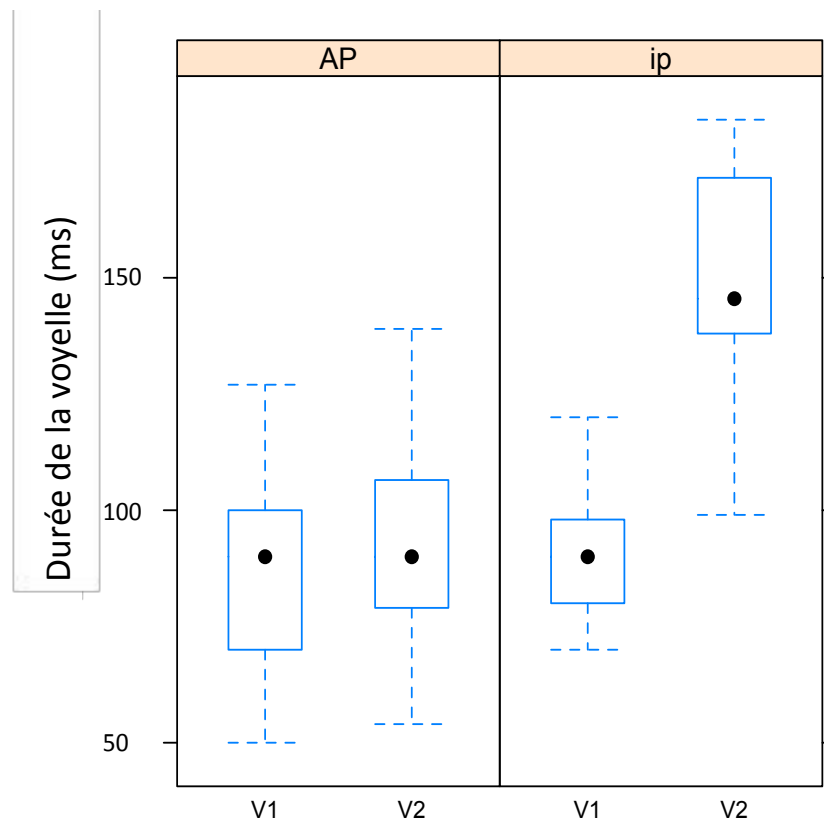


Figure 69. Étude 4 : Durées des voyelles cibles V1 et V2 (en ms) dans les deux conditions prosodiques AP et ip.

V1 correspond à la voyelle associée au premier accent mélodique de l'énoncé (exemple « i » de « vi » dans la séquence « le parvis de la mairie ») et V2 correspond à la dernière voyelle de la séquence (exemple « i » de « mairie » dans la séquence « le parvis de la mairie »). V1 est toujours associée à une frontière de syntagme accentuel quelque soit la condition prosodique (AP et ip) tandis que V2 peut être associée soit à une frontière de syntagme accentuel (condition AP) soit à une frontière de syntagme intermédiaire (condition ip).

Nous avons effectué une série de t-tests dans le but de vérifier si la durée des voyelles cibles (V1 et V2) diffère dans les deux conditions prosodiques (AP et ip). Conformément à ce qui était attendu, la durée de la voyelle en position finale des stimuli (V2) était significativement plus longue dans la condition ip que dans la condition AP ($t(37.86)=8.2258$, $p>0.001$). Ceci confirme l'annotation réalisée par les deux experts faisant apparaître le fait que V2 était associée à une frontière de syntagme intermédiaire dans la condition ip tandis qu'elle était associée à une frontière de syntagme accentuel dans la condition AP. De plus, nous pouvons noter que la durée de V1 n'était pas significativement différente dans les deux conditions ($t(34.317)=0.251$, $p=0.803$). Ceci confirme que la voyelle associée au premier accent mélodique de l'énoncé était associée à une frontière d'AP quelque soit la condition prosodique. Nous pouvons également remarquer que les durées de V1 et de V2 n'étaient pas statistiquement différentes au sein de la condition AP ($t(37.994)=0.7175$, $p=0.4775$). Ceci confirme le fait que V2 soit associée à

une frontière de syntagme accentuel seulement dans la condition AP. Les moyennes de durées obtenues pour les voyelles cibles (V1 et V2) dans les deux conditions prosodiques (AP et ip) sont représentées dans le tableau 12.

	V1	V2
AP	87.8	86.5
ip	89.4	135.2

Tableau 12. Étude 4 : Moyenne des durées (ms) des voyelles cibles (V1 et V2) en fonction de la condition prosodique (AP vs. ip).

En plus des analyses de durée, nous avons également mesuré la hauteur des voyelles cibles. Les valeurs moyennes de *fo* associées aux voyelles cibles dans les deux conditions prosodiques (AP et ip) sont illustrées dans la figure 70.

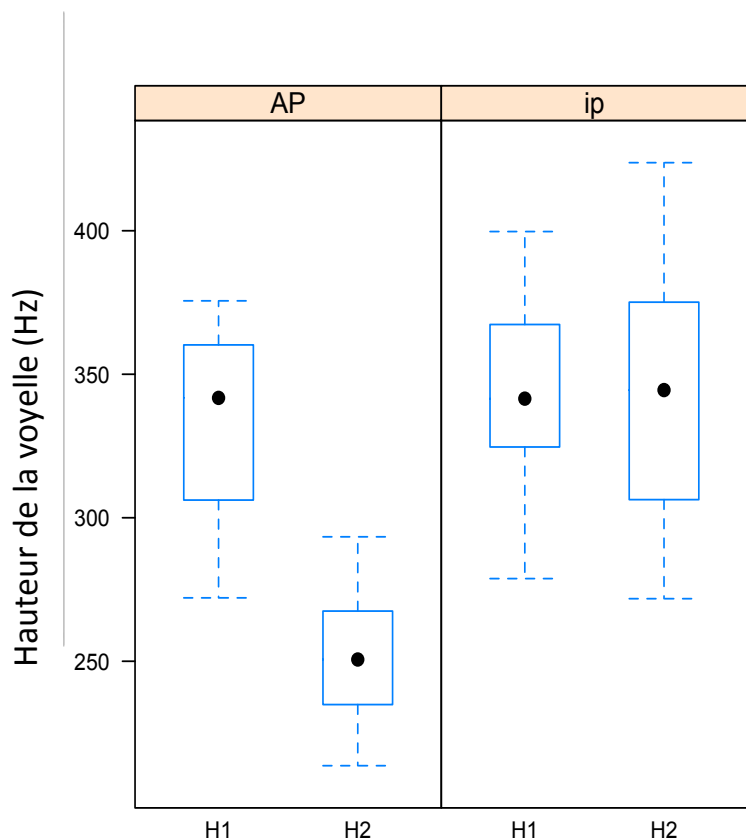


Figure 70. Étude 4 : Hauteur moyenne des voyelles cibles H1 et H2 (en Hz) en fonction de la condition prosodique (AP et ip).

H1 correspond à la voyelle associée au premier accent mélodique de l'énoncé (exemple « i » de « vi » dans la séquence « le parvis de la mairie ») et H2 correspond à la dernière voyelle de la séquence (exemple « i » de « mairie » dans la séquence « le parvis de la mairie »). V1 est toujours associée à une frontière de syntagme accentuel quelque soit la condition prosodique (AP et ip) tandis que V2 peut être associée soit à une frontière de syntagme accentuel (condition AP) soit à une frontière de syntagme intermédiaire (condition ip).

Au regard des moyennes de *f₀* et conformément à ce qui était attendu, les valeurs de *f₀* de H2 sont plus basses que les valeurs de *f₀* de H1 dans la condition AP ($t(27.886)=8.3916$, $p<0.001$). Au contraire, dans la condition ip, les valeurs de *f₀* de H1 et H2 sont similaires ($t(36.922)=1.057$, $p=2.974$). Ceci semble confirmer le fait que le deuxième accent mélodique (H2) est abaissé relativement au premier dans la condition AP mais pas dans la condition ip. Toutefois, afin de prendre en compte les variations de registre observées chez notre locutrice et de vérifier que ces variations de registre n'ont pas affecté nos mesures, nous avons également effectué des rapports entre H2 et H1 dans les deux conditions. Ces rapports par H1 nous permettent d'évaluer la hauteur de la syllabe cible en relation avec le premier pic mélodique de l'ip qui marque la ligne de référence du registre. Nous renvoyons le lecteur au chapitre 2 de cette thèse pour plus de détails sur les avantages de cette méthode. Nous nous attendons à ce que H2 soit abaissé relativement à H1 dans la condition AP tandis que nous nous attendons à ce que H2 soit aussi haut que H1 dans la conditions ip en raison de l'accent de syntagme (H-) aligné avec la frontière droite du syntagme intermédiaire.

Les rapports de H2 par H1 en fonction de la condition prosodique sont illustrés dans la figure 71. Le rapport 1 indique les valeurs de *f₀* de H1 sont similaires à celles de H2 tandis qu'une valeur inférieure à 1 indique que les valeurs de *f₀* de H2 sont inférieures à celles de H1.

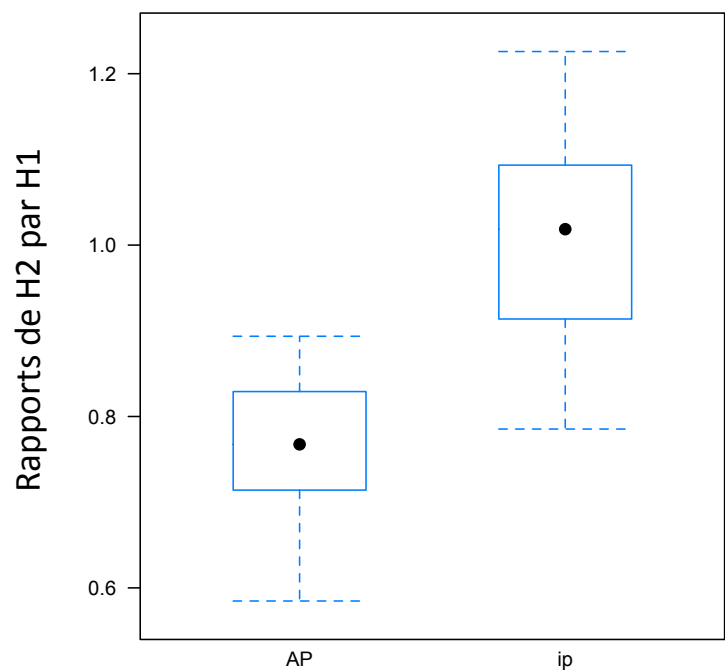


Figure 71. Étude 4 : Rapports de H2 par H1 en fonction de la condition prosodique (AP et ip).
H2 peut être associée soit à une frontière de syntagme accentuel (condition AP) soit à une frontière de syntagme intermédiaire (condition ip).

Les rapports de H2 par H1 vont également dans le sens d'un abaissement du second accent mélodique (H2) dans la condition AP mais pas dans la condition ip puisque le ratio entre H1 et H2 est inférieur à 1 pour la condition AP et à peu près équivalent à 1 pour la condition ip. Les valeurs moyennes de *fo* ainsi que les valeurs moyennes des rapports entre H2 et H1 en fonction de la condition prosodique sont illustrées dans le tableau 13.

	H1	H2	H2/H1
AP	333	251	0.75
ip	343	344	1.3

Tableau 13. Étude 4 : Valeurs moyennes de *fo* (en Hz) pour H1 et H2 et valeurs moyennes du rapport de H2 par H1 en fonction de la condition prosodique (AP vs. ip).

Conformément à ce qui avait été observé par les annotateurs, les analyses de durée et de *fo* menées sur V1 et V2 nous permettent donc d'affirmer que la dernière voyelle de chaque stimulus était associée à une frontière de syntagme accentuel dans la condition AP et à une frontière de syntagme intermédiaire dans la condition ip. Ces stimuli ont constitué la base de notre expérience de perception.

2.2.2. Résultats de la tâche de complétion de phrases

Rappelons qu'à l'écoute des séquences tronquées après la voyelle cible (V2), les participants devaient compléter la phrase en choisissant parmi deux propositions qui étaient présentées à l'écran. Pour chaque stimulus auditif, les deux choix possibles étaient constitués soit d'un syntagme prépositionnel, soit d'un syntagme verbal. Par exemple, à l'écoute du stimulus « le parvis de la mairie » les participants devaient compléter soit par « de Paris » (réponse syntagme prépositionnel, SP), soit par « devrait être restauré » (réponse syntagme verbal, SV).

La figure 72 présente le pourcentage moyen de réponses syntagme prépositionnel (SP) et de réponses syntagme verbal (SV) pour les stimuli de la condition AP et de la condition ip. Notre hypothèse était que les participants choisiraient préférentiellement (c'est-à-dire dans plus de 50 % des cas qui constitue le seuil du hasard) la réponse SV lorsqu'ils entendraient une séquence dont la dernière voyelle était associée à une frontière de syntagme intermédiaire plutôt que lorsque cette frontière était associée à une frontière de syntagme accentuel. Conformément à nos hypothèses, dans la figure 72, nous pouvons remarquer que les participants ont donné plus de réponses de type SP (67,2%) que de type SV lorsque les stimuli contenaient une frontière d'AP. Au contraire, lorsque les stimuli contenaient une frontière d'ip, ils ont donné plus de réponse de type SV (78,3%) que de type SP.

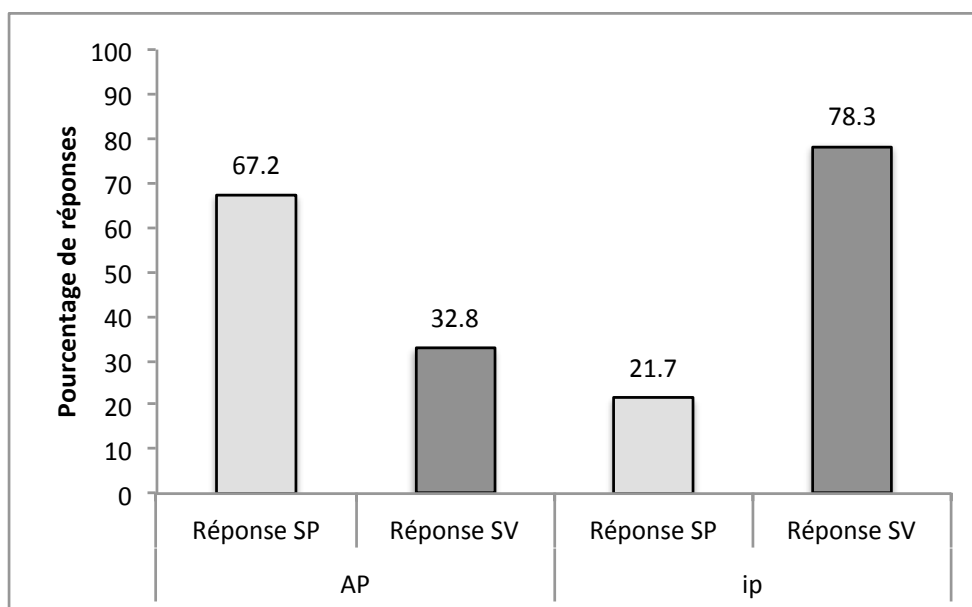


Figure 72. Étude 4 : Pourcentage moyen de réponses de type syntagme prépositionnel (SP) ou syntagme verbal (SV).

Le type de réponse donné est fonction du type de stimuli entendu par les participants et dont la voyelle finale était associée soit à une frontière de syntagme accentuel (AP), soit à une frontière de syntagme intermédiaire (ip).

Un modèle de régressions logit (*Mixed Logit Model*, MLM) a été employé dans le but d'examiner l'influence du type de stimuli présenté au locuteur (AP ou ip) sur les réponses données. La variable dépendante était le type de réponse (réponse SV ou SP) donné par les participants. Nous rappelons que dans cette expérience nous avons à évaluer l'effet d'un facteur sur une variable dépendante binomiale. Dans ce cas précis, l'utilisation d'ANOVA est déconseillée. En effet, dans le cas d'une variable dépendante binomiale où la variance dépend de la moyenne, le test statistique ANOVA n'est pas approprié. Ce type de variable relève d'un modèle de régressions logit (*Mixed Logit Model*, voir par exemple Pinheiro & Bates, 2000). Nous renvoyons le lecteur à la section 2.4 du chapitre 2 pour plus de détails sur l'avantage de ce type de modèle. Les données ont été codées avec la valeur 1 attribuée lorsque la réponse donnée était de type SV et 0 lorsque qu'elle était de type SP. Le facteur fixe était le type de stimuli entendu par les participants (AP ou ip). Les participants ainsi que les items ont été inclus en tant que facteurs aléatoires. Le modèle incluait 720 mesures.

Les résultats du modèle MLM montrent que le type de stimuli présenté (AP vs. ip) a eu un effet significatif sur le type de réponse donné par les participants plus de réponses SV données lorsque le stimulus entendu était de type ip ($\beta=2.2731$, $se=0.1828$, $z=12.438$, $p<0.0001$) que lorsque le stimulus était de type AP. Aucun autre facteur n'a eu d'effet significatif sur la variable dépendante (les p-values étaient toutes supérieures à 0.05). Les probabilités de réponses de type SV et de type SP estimées par le modèle sont échelonnées sur une échelle allant de 0 à 1 et sont illustrés en figure 73. Notons

qu'indépendamment du type de stimulus entendu par les locuteurs, les participants ont toujours donné plus de réponses de type SV que de type SP. Ceci va dans le sens de ce que nous avons observé à l'écrit où les participants complétaient préférentiellement (dans 81% des cas) par une réponse de type SV. Ce résultat n'est pas surprenant et est en accord avec la stratégie d'attachement minimal qui consiste à construire la structure syntaxique la plus simple (c'est-à-dire celle qui contient le moins de nœuds syntaxiques ; Britt, Perfetti, Garrod & Rayner, 1992 ; Feirreira & Clifton, 1986 ; Frazier & Rayner, 1972 ; Rayner et al., 1983 ; voir section 1 de ce chapitre). Toutefois, le résultat crucial de cette expérience est le fait que la structure prosodique des stimuli entendus influe de manière significative sur la réponse donnée par les participants.

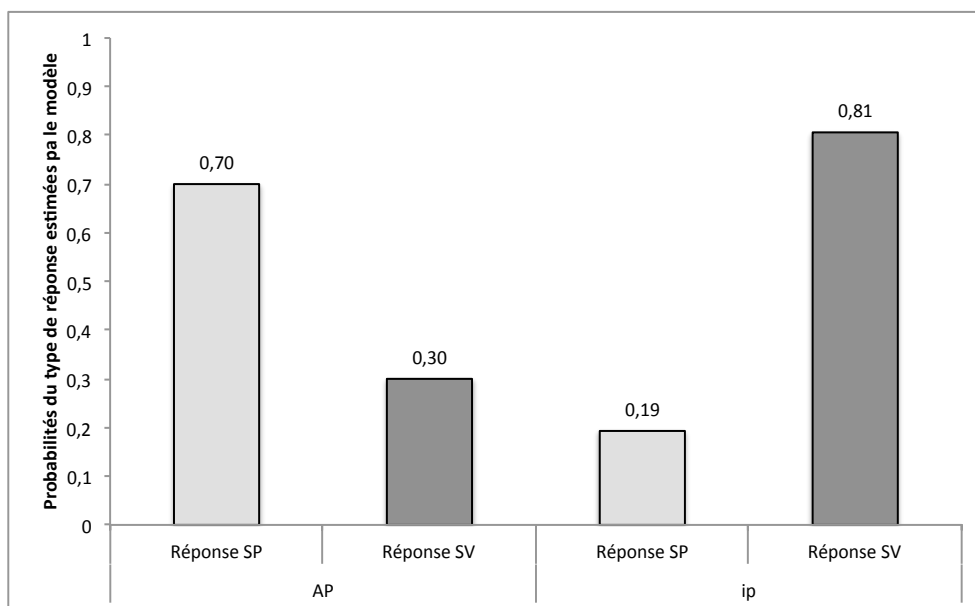


Figure 73. Étude 4 : Probabilités du type de réponses (réponse SP ou réponse SV) estimées par le modèle MLM.

Le type de réponse donné est exprimé en fonction du type de frontière prosodique associé à la dernière voyelle des stimuli entendus par les locuteurs (frontière d'AP ou frontière d'ip). Les probabilités estimées par le modèle sont échelonnées sur une échelle allant de 0 (improbable) à 1 (très probable).

2.3. Discussion

Cette expérience montre que les participants ont été capables d'utiliser le type de frontière prosodique c'est-à-dire son niveau de hiérarchisation dans la structure prosodique (frontière de syntagme accentuel vs. frontière de syntagme intonatif) dans le but d'inférer la catégorie grammaticale du constituant suivant la frontière sans même avoir entendu ce

constituant. En effet, les participants ont donné plus de réponses de type SV (78,3% des réponses) quand la frontière associée à la dernière voyelle du stimulus était une frontière de syntagme intermédiaire plutôt que lorsque cette frontière était typique d'un syntagme accentuel (32,8% des réponses). Ces résultats apportent la preuve que les participants ont interprété la frontière de syntagme intermédiaire en tant qu'indice de frontière syntaxique majeure (comme une rupture entre un syntagme nominal et un syntagme verbal) alors qu'ils ont interprété la frontière de syntagme accentuel comme un indice de frontière syntaxique mineure (telle qu'une rupture au sein du syntagme nominal). Ces résultats ont une double implication : ils nous permettent de discuter (i) les modèles proposés dans le but de rendre compte de la structure intonative du français et (ii) le rôle des frontières prosodiques dans la construction de la structure syntaxique.

Nous avons précédemment évoqué le fait que les modèles traditionnellement postulés dans le but de rendre compte de l'intonation du français admettent deux niveaux de structure prosodique : un niveau majeur que nous pouvons assimiler au syntagme intonatif et un niveau de structuration mineure que nous pouvons assimiler au syntagme accentuel (cf. Jun & Fougeron, 1995, 2000, 2002). Au cours des chapitres précédents, nous avons fourni des preuves de l'existence d'un niveau additionnel de structuration (que nous avons appelé le syntagme intermédiaire) au moyen de plusieurs expérimentations menées en production de la parole. Nous avons en effet montré que le syntagme intermédiaire peut apparaître dans tous types d'énoncés (et non uniquement pour expliquer des phénomènes de *tonal spreading* observés sur des structures syntaxiques spécifiques tel que cela avait été précédemment postulé) dès lors que les structures syntaxiques et prosodiques le permettent. L'alignement d'une frontière droite de syntagme accentuel avec une rupture syntaxique majeure entraîne l'émergence d'une frontière droite de syntagme intermédiaire si le syntagme est suffisamment long (composé au minimum de deux APs). Dans les stimuli auditifs utilisés au cours de l'étude que nous venons de présenter, la dernière voyelle pouvait être associée soit à une frontière de syntagme accentuel soit à une frontière de syntagme intermédiaire. Le fait que les sujets aient donné plus de réponses SV quand ils entendaient un stimulus dont la voyelle finale était associée à une frontière d'ip plutôt que lorsque cette voyelle était associée à une frontière d'AP, suggère que les sujets perçoivent et utilisent les indices acoustiques présents à la frontière d'AP et d'ip. Si, conformément aux modèles traditionnels de l'intonation du français, on accepte l'existence de deux niveaux de structuration prosodique, comment peut-on rendre compte de cette différence dans le type de réponse donnée par les participants ? Une des explications possible est que les participants de notre expérience ont perçu la différence entre les indices acoustiques présents à la frontière d'AP et d'ip et ont utilisé cette différence pour inférer la catégorie morphosyntaxique du mot suivant. Les résultats issus de cette expérience soutiennent donc, pour la première fois et sur la base de résultats obtenus en perception de la parole,

l'hypothèse de l'existence d'un niveau de structuration additionnel, le syntagme intermédiaire.

De plus, les résultats obtenus grâce à cette étude nous permettent d'affirmer que les indices présents à la frontière des constituants prosodiques influencent les premières étapes du traitement syntaxique. La question du moment où les auditeurs utilisent les indices prosodiques durant le traitement syntaxique des énoncés était jusqu'ici une question sujette à débat dans la littérature. Nous avons précédemment évoqué le fait que la plupart des études qui se sont intéressées à cette question ont recours à des ambiguïtés syntaxiques. Lorsque l'auditeur d'une langue fait face à une ambiguïté syntaxique qui induit la génération de plusieurs structures syntaxiques possibles, deux types de rôle des indices prosodiques peuvent être envisagés. Soit les indices prosodiques sont mobilisés dès les premières étapes du traitement, soit ils sont utilisés en dernier recours comme moyen de résoudre l'ambiguïté syntaxique lorsque celle-ci est détectée par l'analyseur. Il est en effet possible d'imaginer que, dès que l'analyseur syntaxique rencontre des indices prosodiques, ceux-ci soient directement utilisés par les auditeurs afin de renforcer l'activation de la structure syntaxique appropriée. Dans ce cas là, les informations prosodiques seraient utilisées très rapidement afin de guider les premières étapes et les premiers choix de l'analyse. Il est également possible d'envisager le fait qu'au moment où l'auditeur fait face à une ambiguïté syntaxique (et en l'absence d'informations désambiguïsantes), plusieurs structures syntaxiques soit générées et les informations disponibles (dont les informations prosodiques font parties) seraient utilisées pour permettre de faire un choix parmi les différentes structures générées. Les informations prosodiques seraient donc utilisées plus tardivement dans le but de faire un choix parmi différentes structures syntaxiques possibles. De précédentes études menées en temps réel sur le français semblent indiquer que les indices prosodiques soient impliqués dès les premières étapes syntaxiques du traitement (Millotte et al., 2007, 2008). Cependant, comme nous l'avons précédemment évoqué, ces études ne permettent pas de résoudre de façon précise cette question étant donné que le type de frontière prosodique (c'est-à-dire le niveau de structuration prosodique) influençant la construction de la structure syntaxique reste toujours inconnu. Dans notre expérience et étant donné que nous voulions étudier le moment où les indices prosodiques interviennent, nous avons utilisé des phrases ne présentant aucune ambiguïté syntaxique. Nous demandions aux participants d'inférer la catégorie grammaticale de mots situés après un début de phrase tronqué en faisant l'hypothèse que la frontière prosodique qu'ils percevaient à la fin des stimuli allait guider leur choix. Le fait que les participants aient complété majoritairement les séquences tronquées par un syntagme verbal lorsqu'ils percevaient une frontière d'ip plutôt qu'une frontière d'AP nous permet d'affirmer que le type de frontière prosodique (AP ou ip) leur permettaient d'inférer la catégorie morphosyntaxique du mot suivant sans même avoir entendu ce mot. Au vu de ces résultats nous pouvons donc conclure que les

indices prosodiques ont été mobilisés par les participants non pas dans le but de résoudre une quelconque ambiguïté (puisque aucune ambiguïté n'est présente dans les stimuli) mais pour guider les tous premiers choix de l'analyse syntaxique et ce, avant même d'avoir entendu le mot qu'ils devaient insérer dans la structure syntaxique.

Nos résultats nous renseignent également sur le domaine prosodique qui contraint l'analyse syntaxique des énoncés en français. En effet, l'étude précédente a montré que les frontières de syntagmes accentuels et que les frontières de syntagme intermédiaires influencent toutes deux la construction de la structure syntaxique : les premières sont interprétées par les auditeurs du français comme indiquant une rupture syntaxique mineure (telle qu'une rupture au sein d'un syntagme nominal sujet) tandis que les secondes sont interprétées en tant qu'indice de rupture syntaxique majeure (telle que la frontière entre un syntagme nominal et un syntagme verbal). Les différents indices acoustiques présents à la frontière d'AP et d'ip (indices de durée et indices mélodiques) sont donc utilisés par les auditeurs du français dans le but d'inférer la catégorie syntaxique des mots situés après ces frontières. Cependant, à ce stade de notre analyse, nous ne savons pas si un seul type d'indice acoustique serait suffisant afin de marquer les frontières d'AP ou d'ip ou si c'est la superposition des indices mélodiques et de durée qui permet l'utilisation des frontières prosodiques dans la construction de la syntaxe. Dans l'étude présentée précédemment, les frontières prosodiques étaient marquées à la fois par des indices mélodiques et de durée. Nous pourrions imaginer que des frontières d'AP et d'ip marquées par un seul type d'indice (par exemple la durée) pourraient être suffisantes pour influencer le traitement syntaxique des phrases. Dans ce cas là les indices de durée influeraient-ils de manière plus importante sur les réponses des participants que les indices mélodiques ? Dans le but d'étudier de façon plus précise la contribution de chaque type d'indice acoustique (indices de durée et indices mélodiques) dans la perception de la frontière d'ip et d'AP, nous proposons une deuxième expérience dans laquelle nous avons manipulé ces indices indépendamment les uns des autres.

3. ÉTUDE 5 : QUEL EST L'INDICE ACOUSTIQUE LE PLUS SAILLANT DANS LA PERCEPTION DES FRONTIÈRES DE SYNTAGME INTERMÉDIAIRE EN FRANÇAIS ?

Cette expérience a été réalisée afin de déterminer l'influence séparée des indices de durée et des indices mélodiques situés à la frontière de l'AP et de l'ip dans la construction de la structure syntaxique en français.

Nous savons qu'en anglais, les indices de durée contraignent de façon très importante la perception des frontières prosodiques (Wightman, Shattuck-Hufnagel, Ostendorf, Price, 1992; Breen, Gibson & Watson, 2010). En

français les choses sont beaucoup moins claires et bien que de nombreuses études aient souligné l'importance des indices de durée dans la perception des frontières prosodiques (voir par exemple Rossi, 1997; Vaissière & Michaud, 2006), cette contribution de la durée étudiée indépendamment de l'intonation a rarement été expérimentalement testée. Au cours du chapitre précédent, nous avons mentionné le fait que Rietveld (1980) considère que l'allongement est un meilleur indice que les indices mélodiques dans le but de guider la segmentation lexicale en français. Contrairement à ce dernier, Bagou, Fougeron & Frauenfelder (2002) ont montré que, dans une langue artificielle et lorsqu'un seul indice est pris en compte, les indices mélodiques induisent un meilleur taux d'identification des frontières de mots que les indices de durée. De plus, aucune étude testant séparément la contribution des indices de durée et des indices mélodiques dans la construction de la structure syntaxique n'a, à notre connaissance, été menée jusqu'à ce jour pour le français. C'est donc ce que nous allons présenter au cours de la section suivante.

Dans cette étude, les participants devaient compléter des séquences tronquées dont la voyelle finale pouvait être associée soit à une frontière d'AP, soit à une frontière d'ip. Comme pour l'étude 4 présentée précédemment, les participants pouvaient choisir entre une réponse de type syntagme verbal (SV) ou une réponse de type syntagme prépositionnel (SP).

Grâce à l'utilisation du logiciel de resynthèse PSOLA nous avons manipulé les indices acoustiques associés à la voyelle finale des séquences qui pouvaient être associées soit à une frontière de syntagme accentuel, soit à une frontière de syntagme intermédiaire. Nous avons donc testé la contribution séparée des indices de durée et des indices mélodique dans la construction de la structure syntaxique des phrases.

Dans cette expérience, les participants étaient confrontés à 6 types de stimuli possibles qui sont présentés dans le tableau 14:

	Indices mélodiques	Indices de durée
AP-mel	LH*	Indices de durée ambigus
ip-mel	LH*LH-	Indices de durée ambigus
AP-dur	Indices mélodiques ambigus	Allongement pré-frontière d'AP
ip-dur	Indices mélodiques ambigus	Allongement pré-frontière d'ip
AP	LH*	Allongement pré-frontière d'AP
ip	LH-	Allongement pré-frontière d'ip

Tableau 14. Étude 5 : Six types de stimuli de resynthèse.

Nous avons manipulé les indices de durée et les indices mélodiques présents sur la voyelle finale des stimuli ce qui a donné lieu à 6 types de stimuli.

Dans les conditions d'AP-mel et ip-mel, les frontières prosodiques associées à la voyelle finale des stimuli n'étaient marquées que par des indices

mélodiques, les indices de durée étant rendus ambigus. Au contraire dans les conditions AP-dur et ip-dur, les frontières prosodiques associées à la voyelle finale des stimuli n'étaient marquées que par des indices de durée, les indices mélodiques étant rendus ambigus.

Enfin dans les conditions AP et ip, les frontières prosodiques associées à la voyelle finale des stimuli étaient marquées à la fois par les indices mélodiques et de durée comme dans les stimuli naturels.

Un exemple des 6 types de stimuli obtenus à partir de la séquence « le gala de fin d'année » est illustré dans la figure 74. Cette manipulation des indices mélodiques et de durée nous a permis d'étudier la contribution séparée de chaque type d'indice acoustique sur la réponse des participants. Nous avons analysé le type de réponse donné par les participants (Réponse SP et Réponse SV) ainsi que leurs temps de réaction.

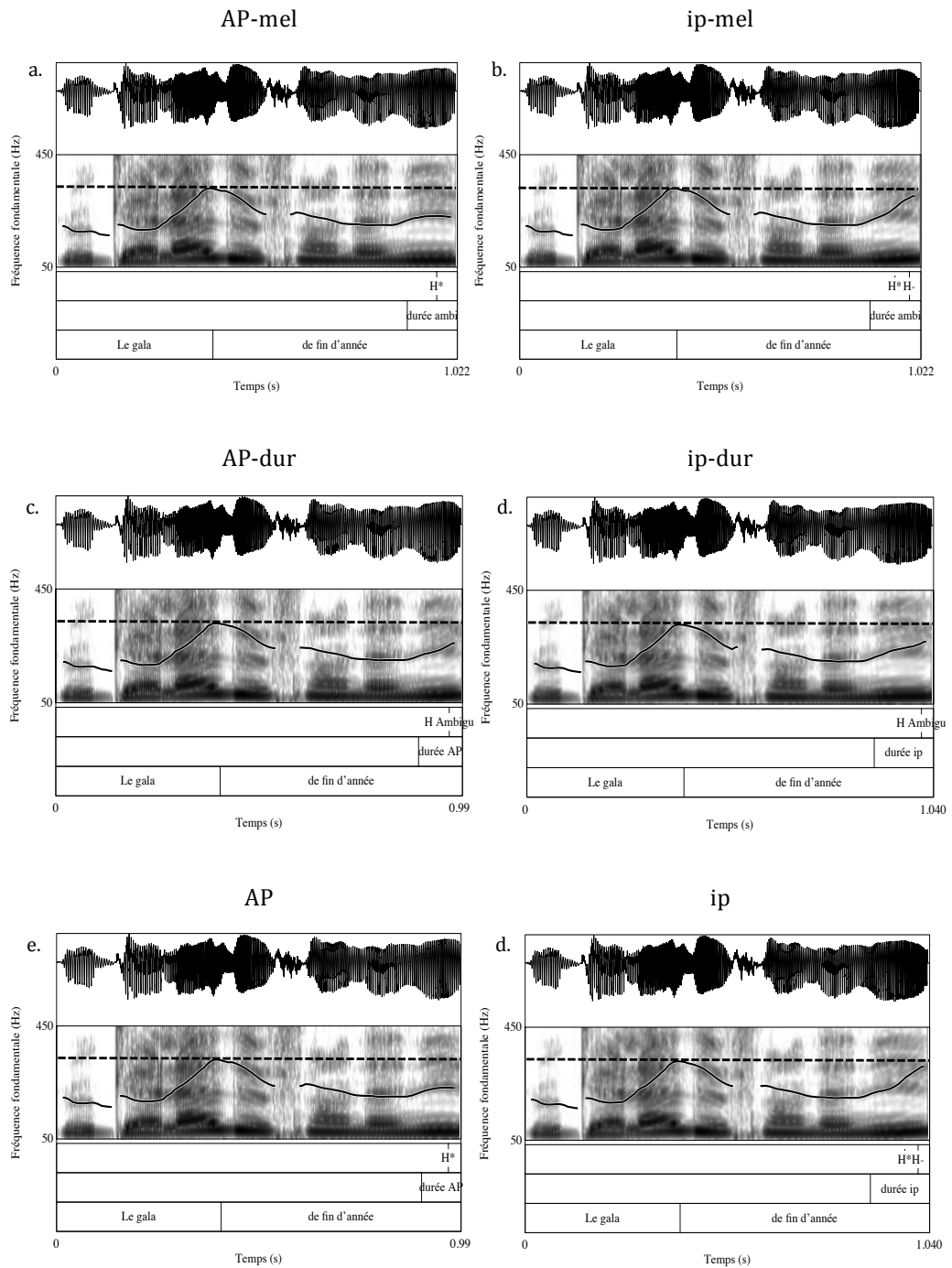


Figure 74. Étude 5 : Exemple de stimuli pour les 6 conditions prosodiques (AP-mel, ip-mel, AP-dur, ip-dur, AP et ip)

3.1. Méthode

3.1.2. Matériel

Les phrases de la condition AP enregistrées lors de la précédente expérience (étude 4) ont été utilisées comme base pour la construction de stimuli. Nous avons modifié les indices de durée et les indices mélodiques

présents sur la dernière voyelle de chacun de ces stimuli. La manipulation des indices mélodiques et de durée a été réalisée grâce à l'utilisation du plug-in PSOLA implémenté dans le logiciel PRAAT.

Pour les deux conditions « frontières mélodiques » (*AP-mel* et *ip-mel*) nous avons rendu les indices de durée ambigus afin que les frontières prosodiques d'AP et d'ip ne soient marquées que par la mélodie. Dans le but de rendre la durée ambiguë, nous avons remplacé la durée de la voyelle finale du stimulus par la moyenne des durées obtenues dans les conditions ip et AP lors de l'analyse des stimuli acoustiques réalisés lors de l'étude 6 (valeur correspondant à 110 ms). Pour la condition ip-mel nous avons également manipulé la hauteur de la voyelle en position finale du stimulus dans le but d'introduire un accent de syntagme (H-) responsable du blocage de l'abaissement des accents mélodiques successifs. Autrement dit nous avons manipulé la hauteur de la voyelle en position finale du stimulus de manière à ce qu'elle soit aussi haute que le premier accent mélodique du stimulus. Cette manipulation est illustrée dans la figure 74b. En effet, dans la figure 74a la voyelle en position finale du stimulus est abaissée relativement au premier accent mélodique (LH*) alors que ce n'est pas le cas dans la figure 74b en raison de la frontière d'ip qui se manifeste par un retour à la ligne de référence du registre marquée par le premier accent mélodique de l'énoncé.

Pour les deux conditions AP et ip, notre objectif était de rendre les indices mélodiques ambigus de manière à ce que les deux types de frontière (ip et AP) ne soient marqués que par des indices de durée. Pour cela nous avons rendu les indices mélodiques ambigus en modifiant la hauteur de la voyelle en position finale du stimulus. Afin d'obtenir une hauteur ambiguë, nous avons calculé pour chaque stimulus utilisé lors de l'étude 4, la différence de hauteur entre la voyelle finale en position finale dans la condition AP et la voyelle finale dans la condition ip. Nous avons remplacé la hauteur de la voyelle par la moitié de cette différence. Prenons un exemple : Dans le stimulus « Le gala de fin d'année », la voyelle finale « é » a été réalisée avec une hauteur de 231Hz dans la condition AP (c'est à dire lorsque le dernier accent mélodique était abaissé relativement au premier). Le premier accent mélodique a quant à lui été réalisé avec une hauteur de 328 Hz. Nous avons donc remplacé la valeur du dernier accent mélodique par une hauteur correspondant à 279 Hz ($231 + ((328 - 231) / 2)$). De cette manière nous avons obtenu des indices mélodiques ambigus puisque ni vraiment abaissés par rapport au premier accent mélodique ni vraiment réinitialisés. Les valeurs mélodiques ambiguës ainsi obtenues sont illustrées dans la figure 74c et 74d où les deux « H ambigus » ont une valeur correspondant à 279 Hz. Dans le cas de la condition AP-dur, la durée de la voyelle en position finale du stimulus a été remplacée par la moyenne des durées de la condition AP de l'étude 4 (90 ms) alors que dans la condition ip-dur, la durée de la voyelle en position finale du stimulus a été remplacée par la moyenne des durées de la voyelle finale dans la condition ip (130 ms).

Enfin pour les deux conditions AP et ip, les frontières étaient marquées à la fois par les indices mélodiques (c'est à dire l'abaissement du second accent mélodique par rapport au premier accent mélodique pour la condition AP et la même hauteur pour les deux accents mélodiques pour la condition ip) et les indices de durée (une durée correspondant à 90 ms pour la condition AP et à 130 ms pour la condition ip). Ces deux conditions sont illustrées dans les figures 74e et 74f précédemment présentées.

Afin de s'assurer du fait que les modifications conduites sur les stimuli ne leur donnaient pas un aspect trop mécanique ou artificiel, nous avons demandé à 5 auditeurs de langue maternelle française de juger du naturel de ces phrases sur une échelle de 1 à 7 (1 correspondant à une phrase totalement mécanique et 7 correspondant à une phrase totalement naturelle). Les sujets entendaient les 120 stimuli synthétisés ainsi que les 20 phrases naturelles utilisées lors de l'étude 4. Les auditeurs ont jugé les phrases naturelles de l'étude 4 aussi naturelles que nos stimuli synthétisés (moyenne pour les phrases naturelles : 5.1 ; moyenne pour les phrases synthétisées : 5.158 ; $t(290.855)=0.4311$, $p=0.6667$). Ce test nous a permis de nous assurer de la qualité de la resynthèse que nous avons effectué sur nos stimuli. De plus, les phrases synthétisées ont obtenu des scores similaires dans les six conditions (moyenne pour la condition AP-dur : 5.85 ; moyenne pour la condition AP-mel : 5.4 ; moyenne pour la condition AP : 5.15 ; moyenne pour la condition ip-dur : 4.895 ; moyenne pour la condition ip-mel : 5 ; moyenne pour la condition ip : 4.7).

120 stimuli ont ainsi été construits à partir de la manipulation de 20 séquences de base.

Les 120 stimuli ont été répartis en 6 listes de phrases. Chaque liste de stimuli comportait donc 20 stimuli expérimentaux auxquels nous avons ajouté 40 phrases distractrices correspondant aux phrases distractrices utilisées pour l'étude 4. Les phrases distractives étaient les mêmes pour chaque liste de stimuli. Chaque sujet a été soumis à une seule liste expérimentale de manière à n'entendre qu'une seule fois la même séquence. La présentation des stimuli à l'intérieur des listes était aléatoire.

En plus des phrases expérimentales et distractrices, 8 ont été créées et utilisées au début de l'expérience comme entraînement.

3.1.2. Procédure

La même procédure expérimentale que celle présentée lors de l'étude 4 a été utilisée. Les participants devaient traiter 60 stimuli (20 phrases expérimentales et 40 phrases distractrices) pour que l'expérience se termine. L'expérience durait environ 15 minutes.

3.1.3. Participants

60 participants, locuteurs natifs du français ont participé à l'expérience (10 par liste). Ces participants étaient âgés de 17 à 24 ans avec une moyenne de 22.1 ans. Ils étaient tous étudiants de l'Université de Provence et étaient originaires de différentes régions de France. Aucun participant n'a rapporté de problèmes auditifs ou de troubles du langage. Tout comme pour l'étude 4, à la fin de l'expérience nous avons demandé aux participants s'ils avaient compris le but de l'expérimentation. Certains ont repéré qu'en fonction de la manière dont était prononcé l'item ils choisissaient une réponse plutôt qu'une autre. Deux d'entre eux nous ont rapporté s'être fiés à l'intonation mais ne pouvaient pas nous préciser s'ils avaient adopté une stratégie de réponse particulière.

3.2. Résultats

3.2.1. Analyse du type de réponse

Rappelons qu'à l'écoute des stimuli, les participants devaient compléter la phrase qu'ils avaient entendue en choisissant parmi deux propositions qui étaient présentées à l'écran. Tout comme lors de l'étude 4, les participants devaient compléter les stimuli en choisissant parmi deux suites possibles constituées soit d'un syntagme prépositionnel soit d'un syntagme verbal. Par exemple, à l'écoute de la séquence « le parvis de la mairie » les participants devaient compléter soit par « de Paris » (réponse syntagme Prépositionnel) soit « devrait être restauré » (réponse syntagme verbal).

Notre hypothèse était que les sujets choisiraient préférentiellement la réponse syntagme prépositionnel (SP) lorsqu'ils entendraient une séquence de la condition AP plutôt qu'une séquence de la condition ip. À l'inverse, nous nous attendions à ce que les sujets choisissent préférentiellement la réponse SP lorsqu'ils entendraient un stimulus de la condition ip plutôt que de la condition AP. Par contre nous ne savions pas si les stimuli des conditions ip-mel et ip-dur, ainsi que les stimuli des conditions AP-mel et AP-dur suffiraient à influencer de manière significative leur type de réponse.

La figure 75 présente le pourcentage moyen de réponses syntagme prépositionnel (SP) pour les stimuli des six conditions prosodiques (AP-mel, ip-mel, AP-dur, ip-dur, AP et ip).

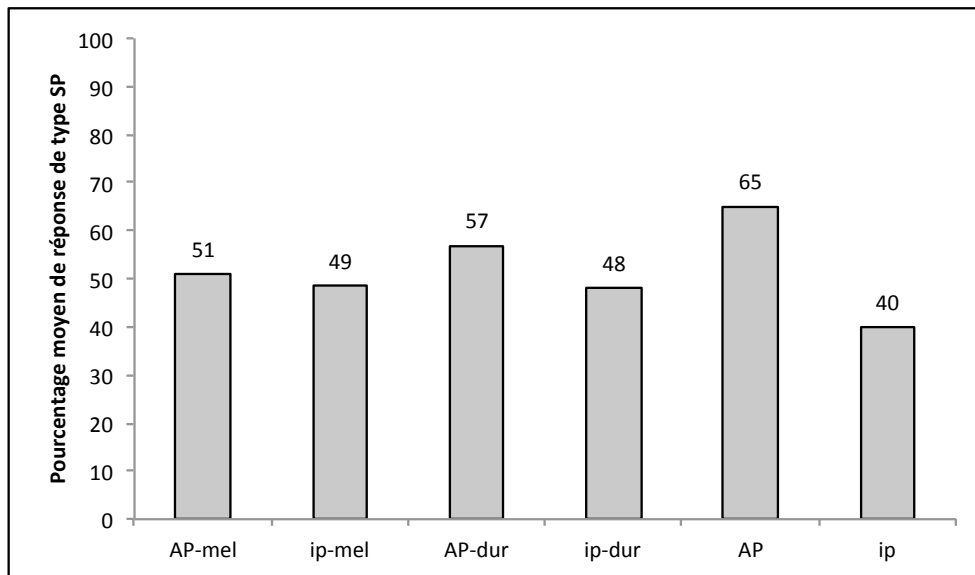


Figure 75. Étude 5 : Pourcentage moyen de réponses syntagme prépositionnel (SP) données par les participants en fonction du type de stimuli entendu (AP-mel, ip-mel, AP-dur, ip-dur, AP et ip).

Un modèle de régressions logit (*Mixed Logit Model*, MLM) a été employé dans le but d'examiner l'influence du type de stimuli présenté au locuteur sur le type de réponse qu'ils ont donné. La variable dépendante était le type de réponse donné par les participants (réponse SV ou réponse SP). Les données ont été codées avec la valeur 1 attribuée lorsque la réponse donnée était de type SV et 0 lorsque qu'elle était de type SP. Le facteur fixe était le type de stimuli entendu par les (AP-mel, ip-mel, AP-dur, ip-dur, AP et ip). Les participants ainsi que les items ont été inclus en tant que facteurs aléatoires. Le modèle incluait 1200 mesures. Les résultats du modèle MLM montrent que la condition *ip* influait significativement sur le type de réponse donné par les participants plus de réponses de type SV que de type AP ($\beta=0.5306$, $se=0.1914$, $z=2.2374$, $p<0.01$). Les conditions *ip-mel* et *ip-dur* n'ont, par contre, eu aucun effet significatif sur le type de réponse donné par les participants (conditions *ip-mel*: $\beta=0.06352$, $se=0.18724$, $z=0.339$, $p=0.73442$; condition *ip-dur*: $\beta=0.01652$, $se=0.1871$, $z=0.88$, $p=0.92965$). Les conditions AP et AP-dur ont également eu un effet significatif sur le type de réponse donné avec significativement moins de réponses SV que de réponses SP (condition AP: $\beta=-0.6598$, $se=0.1926$, $z=-2.427$, $p<0.001$; condition AP-dur: $\beta=-0.412$, $se=0.1892$, $z=-2.174$, $p<0.01$). Enfin la condition AP-mel n'a eu aucun effet sur le type de réponse donné par les participants ($\beta=-0.04935$, $se=0.18725$, $z=-0.2864$, $p=0.792135$). Aucun autre facteur n'avait d'effet significatif sur la variable dépendante (les p-values étaient toutes supérieures à 0.05). Notons qu'indépendamment du type de stimulus entendu, les participants ont toujours donné plus de réponses de type SP que de type SV

Les probabilités de réponses de type SP estimées par le modèle en fonction des 6 types de stimuli sont illustrées dans la figure 76. Ces probabilités sont échelonnées sur une échelle allant de 0 à 1.

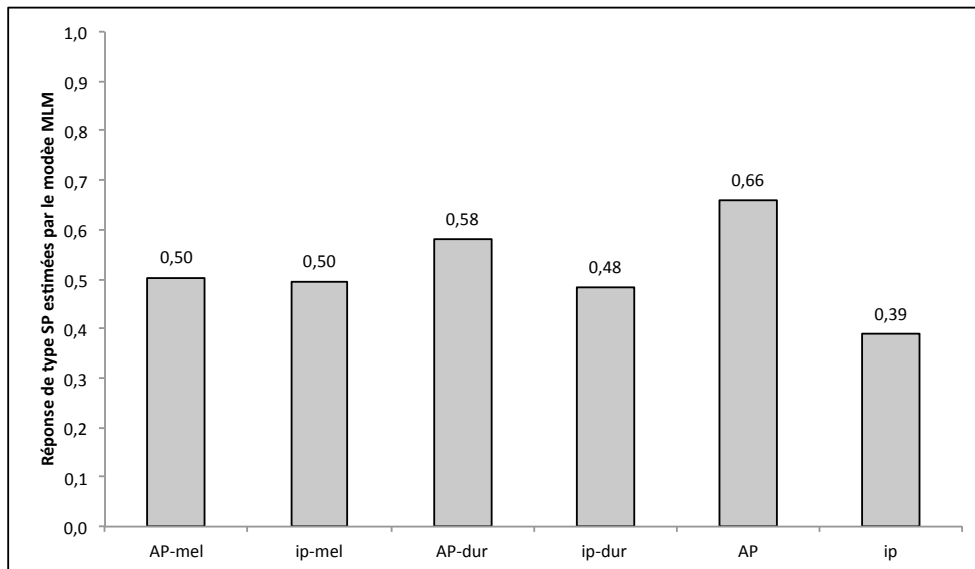


Figure 76. Étude 5 : Probabilités de réponses SP données par les participants estimées par le modèle MLM.

Les probabilités sont fonction du type de stimulus auditif présenté (AP-mel, ip-mel, AP-dur, ip-dur, AP et ip). Les probabilités estimées par le modèle sont estimées sur une échelle allant de 0 (improbable) à 1 (très probable).

3.2.2. Analyse des temps de réaction

Nous avons également mesuré les temps de réaction avec lesquels les participants ont répondu à partir de la fin des stimuli auditifs. Pour chaque participant, les temps de réaction supérieurs et inférieurs à 2,5 écart-types de la moyenne ont été exclus des analyses. Adoptant ce critère, seulement 4.08% des données ont été rejetées. Les temps de réaction pour chaque type de stimuli sont présentés dans la figure 77.

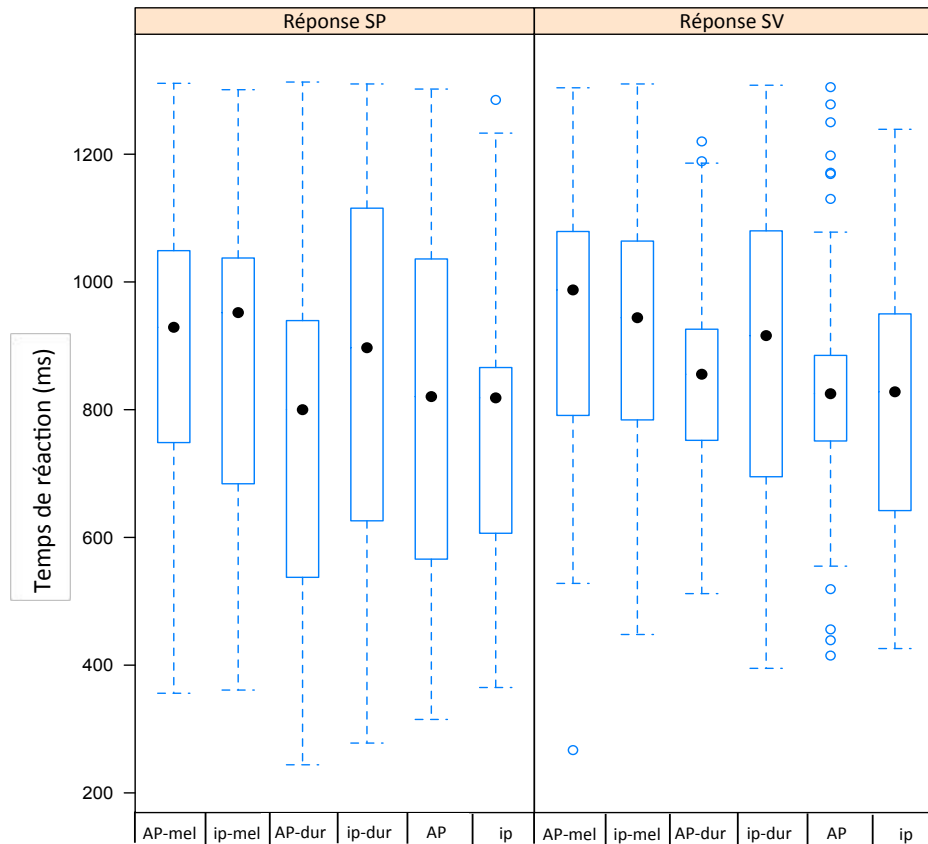


Figure 77. Étude 5 : Temps de réaction des participants (ms) pour chaque type de stimuli (AP-mel, ip-mel, AP-dur, ip-dur, AP et ip) en fonction du type de réponse donné par les participants (Réponse SP / Réponse SV).

Un modèle de régressions linéaires multiples à effets mixtes (*Linear Mixed effects Models, LMM*) a été employé sur les temps de réaction avec le type de stimuli (AP-mel, ip-mel, AP-dur, ip-dur, AP et ip) et le type de réponse donné par les participants (SP/SV) en tant qu'effets fixes. Les participants et les items ont été inclus en tant qu'effets aléatoires.

Les résultats montrent que le type de stimuli présenté a eu une influence significative sur les temps de réaction des participants. En effet, les temps de réactions étaient significativement plus longs dans les conditions AP-mel ($\beta=77.731$, $se=30.987$, $t=2.508$, $pMCMC<0.05$), ip-mel ($\beta=64.7463$, $se=29.447$, $t=2.189$, $pMCMC<0.05$), et ip-dur ($\beta=59.720$, $se=29.889$, $t=2.089$, $pMCMC<0.05$), que dans la condition AP. Par contre les temps de réaction n'étaient pas significativement plus longs dans la condition AP-dur ($\beta=-79.870$, $se=28.787$, $t=-0.343$, $pMCMC=0.738$) et ip ($\beta=-53.325$, $se=30.276$, $t=-1.675$, $pMCMC=0.07$) que dans la condition AP. Les participants mettaient donc plus de temps à répondre lorsque la frontière prosodique était indiquée uniquement par la mélodie que lorsque la frontière prosodique était marquée à la fois par les indices mélodiques et les indices de durée et ce, quelque soit le

type de frontière (frontière d'AP et frontière d'ip). Les résultats montrent également que les participants mettaient autant de temps à répondre lorsque la frontière d'AP était marquée à la fois par la mélodie et la durée que lorsque la frontière d'AP était marquée uniquement par la durée. Notons que le même résultat n'a pas été observé pour la frontière d'ip : les participants mettaient plus de temps à répondre lorsque la frontière d'ip était marquée uniquement par la durée plutôt que lorsqu'elle était marquée à la fois par les indices mélodiques et de durée.

Ni l'effet du type de réponse ($\beta=-1.893$, $se=31.147$, $t=-0.057$, $pMCMC=0.9543$) ni l'interaction entre les deux facteurs expérimentaux n'étaient pas significatifs (les $pMCMC$ values étaient toutes supérieures à 0.05). Le modèle incluait 1151 mesures.

3.3. Discussion

L'analyse du type de réponse nous indique que les participants ont donné significativement plus de réponses de type SV que de type AP lorsqu'ils entendaient un stimulus contenant une frontière d'ip marquée à la fois par des indices mélodiques et des indices de durée. Les participants ont également donné significativement plus de réponses de type SP lorsqu'ils entendaient un stimulus contenant une frontière d'AP que d'ip marquée à la fois par les indices mélodiques et les indices de durée. Ces résultats confirment les résultats que nous avons obtenus lors de l'étude 4 en montrant que les participants ont interprété la frontière d'AP comme une rupture syntaxique mineure alors qu'ils ont interprété la frontière d'ip en tant que rupture syntaxique majeure.

L'analyse du type de réponse nous permet également d'étudier la contribution des indices de durée indépendamment des indices mélodiques et inversement, dans le marquage des frontières prosodiques. Le type de réponse donné par les participants n'était pas influencé par le type de stimuli qu'ils entendaient lorsque les frontières prosodiques (que ce soit des frontières d'ip ou d'AP) n'étaient marquées que par les indices mélodiques. Par contre, les participants ont donné significativement plus de réponses de type SP que de type SV lorsqu'ils entendaient une frontière d'AP marquée uniquement par les indices de durée. Ce résultat semble indiquer que les indices de durée suffisent à marquer la frontière d'AP en français. Par contre, il semble que la présence conjointe des indices mélodiques et de durée soit nécessaire pour marquer la frontière d'ip en français puisque le type de réponse des participants n'était pas influencé par le type de stimuli qu'ils entendaient dans la condition ip-dur.

L'analyse des temps de réaction va dans le même sens que les résultats obtenus pour le type de réponse donné. En effet, cette analyse a permis d'observer que le temps de réaction des participants étaient significativement plus longs lorsqu'ils entendaient une frontière prosodique marquée uniquement par la mélodie que lorsqu'ils entendaient une frontière marquée

conjointement par la mélodie et la durée et ce quelque soit le type de frontière (frontière d'AP ou frontière d'ip). De plus nous pouvons également observer que les participants ne mettaient pas plus de temps à répondre lorsqu'ils entendaient une frontière d'AP marquée par la durée plutôt que lorsqu'ils entendaient une frontière d'AP marquée par la mélodie et la durée. À l'inverse, ils mettaient significativement plus de temps à répondre lorsqu'ils entendaient une frontière d'ip marquée uniquement par la durée.

Ces résultats nous renseignent de manière très intéressante sur le type d'indice acoustique le plus saillant dans la perception des frontières prosodiques en français. Nos données vont dans le sens d'une importance plus grande des indices de durée par rapport aux indices mélodiques dans le marquage des frontières prosodiques pour le plus bas niveau postulé dans la hiérarchie prosodique du français : l'AP. En effet, lorsque la frontière d'AP était marquée uniquement par la durée, nos participants répondaient de manière similaire et aussi rapidement que lorsque cette frontière d'AP était marquée à la fois par la mélodie et par la durée. L'allongement pré-frontière semble donc suffisant en français pour percevoir une frontière d'AP. Il semble, par contre, que les indices de durée ne soient pas suffisamment saillants pour percevoir une frontière d'ip en français. Le marquage conjoint de la frontière d'ip par les indices mélodiques et de durée semble donc nécessaire dans la perception de ce niveau de frontière.

4. DISCUSSION GÉNÉRALE

4.1. Résultats généraux

Les expériences décrites dans ce chapitre suggèrent que les auditeurs du français exploitent les indices acoustiques présents à la frontière du syntagme accentuel et du syntagme intermédiaire dans le but d'inférer les catégories grammaticales de mots qu'ils doivent insérer au sein de la structure syntaxique. Grâce à deux tâches expérimentales où nous avons demandé à des participants de compléter des phrases tronquées dont la structure segmentale était identique mais dont la structure prosodique était différente, les participants ont été capables d'inférer la catégorie morphosyntaxique des mots suivants les séquences tronquées sans même avoir entendu le début de ces mots. Afin de parvenir à cette tâche les participants ont utilisé le type de frontière prosodique associé à la syllabe finale des phrases tronquées, c'est à dire le niveau de hiérarchie de cette frontière dans la structure prosodique. Les participants ont interprété la frontière de syntagme accentuel comme une rupture syntaxique mineure alors qu'ils ont analysé la frontière de syntagme intermédiaire en tant que rupture syntaxique majeure. Dans la deuxième étude présentée dans ce chapitre, nous avons manipulé les indices acoustiques présents aux deux types de frontières (AP et ip). Les résultats obtenus montrent que dans le cas où la frontière de syntagme accentuel était

marquée uniquement par des indices de durée, cela était suffisant pour que les participants perçoivent et utilisent cette frontière dans la construction de la structure syntaxique. La présence conjointe des indices mélodiques et de durée était au contraire nécessaire pour que les participants perçoivent et utilisent la frontière de syntagme intermédiaire de façon significative.

Ces résultats novateurs ont des implications à la fois pour les modèles de l'intonation du français et pour la modélisation du traitement syntaxique des phrases parlées.

4.2. Modèles de l'intonation proposés pour le français

Les modèles traditionnellement postulés dans le but de décrire l'intonation du français admettent deux niveaux de structuration prosodique : le syntagme accentuel et le syntagme intermédiaire. Les différents résultats présentés au début de cette thèse ont permis d'apporter des preuves de l'existence, en production de la parole, d'un niveau de structuration additionnel situé entre l'AP et l'ip. Ce niveau de structure a été appelé syntagme intermédiaire (Jun & Fougeron, 2000). À la vue des résultats que nous avons obtenu dans les études 4 et 5, nous pouvons désormais affirmer que les auditeurs du français perçoivent la différence entre une frontière de syntagme accentuel et une frontière de syntagme intermédiaire et utilisent cette différence dans l'analyse de la structure syntaxique des énoncés. Ces résultats nous conduisent à admettre l'existence de trois niveaux de structure prosodique en français : le syntagme accentuel, le syntagme intermédiaire et le syntagme intonatif. Si les expériences menées permettent d'affirmer que les auditeurs perçoivent la différence entre frontière droite d'AP et frontière droite d'ip, des études supplémentaires seront cependant nécessaires dans le but de confirmer que les auditeurs du français sont capables de faire la différence entre frontière droite de syntagme intermédiaire et frontière droite de syntagme intonatif.

Les expériences menées au cours de ce chapitre nous renseignent également sur le rôle des différents types d'indices acoustiques dans le marquage des frontières prosodiques. Les recherches menées en prosodie dans les différentes langues du monde s'accordent aujourd'hui sur un marquage pluriel (ou multiparamétrique) des frontières de constituants prosodiques. Dans de nombreux cas, les différents indices marquant les frontières prosodiques s'accordent dans le marquage d'un niveau de structure prosodique. Par exemple, nous savons que la syllabe finale d'un syntagme accentuel est associée à un accent mélodique (indices mélodiques) qui est généralement accompagné par un allongement significatif de la syllabe (indices de durée). Cependant, il arrive que les différents indices marquant les frontières de constituants prosodiques ne s'accordent pas dans le marquage de ces frontières. Jusqu'à présent aucun modèle postulé pour le français ne prend en compte ce type de cas. Les expérimentations menées dans ce chapitre apportent de nouvelles informations sur la contribution de chaque

type d'indice dans le marquage des frontières. Ces résultats peuvent enrichir les modèles proposés afin de rendre compte de la structure prosodique du français. En effet, nous avons observé que les indices de durée sont suffisants dans le but de percevoir la frontière de syntagme accentuel en français. Il semble par contre que la présence conjointe des indices mélodiques et de durée soient nécessaires pour marquer la frontière de syntagme intermédiaire. De manière générale il semble donc que les indices de durée soient plus saillants en français que les indices mélodiques dans le marquage des frontières prosodiques. Nous savons que l'importance et la combinaison des indices acoustiques varient en fonction de la langue et en particulier de la façon dont chaque langue va gérer le nombre de phénomènes linguistiques à marquer et le nombre de paramètres acoustiques disponibles (Lehiste, 1967 ; Vaissière, 1989). Dans la mesure où la plupart des modèles admettent un marquage multiparamétrique des frontières prosodiques, il nous semble très important de déterminer la contribution spécifique de chaque indice dans le but de mieux comprendre comment la structure prosodique abstraite se manifeste de manière concrète dans le signal parole et ce dans les différentes langues du monde.

4.3. Modèles de traitement syntaxique

Très peu d'expériences avaient permis d'étudier jusqu'ici le rôle des constituants prosodiques dans la construction de la structure syntaxique en français. Millotte et ses collègues (Millotte et al., 2007, 2008) ont montré que les auditeurs du français utilisent la frontière de syntagme phonologique dans le but de lever des ambiguïtés syntaxiques temporaires présentes à cette frontière. Cependant nous avons précédemment évoqué le fait que les frontières observées par ces auteurs étaient définies en fonction de critères morphosyntaxiques pouvant coïncider avec différents niveaux de frontières prosodiques en fonction de la manière dont les phrases étaient prononcées par les locuteurs. Dans les deux études que nous avons proposées, les frontières prosodiques ont été définies à partir des indices acoustiques observés dans le signal de parole. Ces deux études nous ont permis d'affirmer que les indices acoustiques présents à la frontière du syntagme accentuel contraignent la construction de la structure syntaxique de manière différente des indices présents et à la frontière des syntagmes intermédiaire en français. En effet, les participants ont interprété la frontière de syntagme accentuel en tant que rupture syntaxique mineure (telle qu'une rupture au sein d'un syntagme nominal sujet) tandis que la frontière de syntagme intermédiaire a été interprétée en tant que rupture syntaxique majeure (telle que la rupture entre un syntagme nominal sujet et un syntagme verbal).

De plus, la plupart des études qui se sont intéressées au traitement syntaxique des énoncés ont recours à des ambiguïtés syntaxiques temporaires. Le problème d'une telle méthodologie et que bien souvent, elle ne permet pas (ou permet indirectement) de répondre à la question du

moment où les indices prosodiques interviennent dans le processus de construction de la structure syntaxique. En effet, si l'utilisation d'ambiguïtés syntaxiques permet de montrer l'importance des indices prosodiques dans la désambiguïsation de la structure syntaxique, il est difficile de déterminer dans ce type d'expérimentation, si les indices prosodiques sont utilisés afin de générer une seule structure syntaxique ou s'ils ont été utilisés dans le but de renforcer l'activation d'une structure syntaxique particulière. Dans les expérimentations que nous avons proposées, les corpus utilisés ne présentaient pas d'ambiguïtés syntaxiques. Nous avons demandé à des participants d'inférer la catégorie grammaticale de mots suivant des séquences auditives tronquées. Ces séquences étaient identiques du point de vue segmental mais différaient du point de vue de leur structure prosodique : la dernière syllabe de chaque séquence était associée soit à une frontière d'AP, soit à une frontière d'ip. Les résultats obtenus montrent que les participants ont utilisé le type de frontière prosodique (AP vs. ip) présent dans les stimuli auditifs pour inférer la catégorie grammaticale des mots suivants. Sur la base des résultats que nous avons obtenus, nous pouvons donc affirmer que les indices prosodiques ne sont pas seulement mobilisés dans le but d'activer une structure particulière lorsque l'analyseur syntaxique rencontre une ambiguïté. De plus, le fait que les participants aient été capables d'inférer la catégorie morphosyntaxique d'un mot avant même d'avoir entendu ce mot, prouve que les indices prosodiques sont utilisés très tôt dans la construction du traitement syntaxique. Ce résultat apporte la preuve de l'implication des indices prosodiques lors de la génération de la structure syntaxique, c'est-à-dire dès l'étape initiale du processus de traitement syntaxique. En fonction des frontières prosodiques que les auditeurs d'une langue perçoivent et de leur niveau de hiérarchie dans la structure prosodique, ceux-ci construisent des attentes sur la catégorie morphosyntaxique des mots qu'ils auront à traiter. De plus, le fait que les participants ont mis plus de temps à répondre en fonction du type d'indice acoustique qui marquait les frontières prosodiques dans notre étude confirme le fait que les indices prosodiques sont utilisés en temps réel dans la construction de la structure syntaxique.

Les résultats obtenus au cours de ce chapitre nous permettent donc d'affirmer que les frontières de syntagme accentuel et les frontières de syntagme intermédiaire sont utilisées par les auditeurs du français en temps réel dans le but de construire la structure syntaxique des phases et ce, dès les étapes précoces l'analyse syntaxique. Nos études soulignent le rôle primordial que joue la prosodie dans la construction de la structure syntaxique et insistent sur la nécessité de définir les frontières prosodiques en se basant sur des indices acoustiques présents dans le signal pour les futures recherches menées sur le traitement syntaxique de la parole.

CONCLUSIONS GÉNÉRALES

1. RAPPEL DE LA PROBLÉMATIQUE INITIALE

Malgré les nombreuses études qui ont été menées au cours de ces dernières années, la question de la définition et du nombre des constituants prosodiques reste encore sujette à controverse dans la littérature. Selon les différentes approches proposées, la définition des unités prosodiques peut être basée soit sur des contraintes morphosyntaxiques, soit sur des propriétés prosodiques et plus particulièrement des propriétés tonales (voir Jun 1993, 1998 pour une comparaison de ces deux approches).

Nous savons aujourd'hui que les prédictions de la Phonologie Prosodique (Nespor & Vogel, 1986 ; Selkirk 1984 ; *inter alia*) ne sont pas toujours réalisées dans le signal de parole. Cette constatation nous a conduit à adopter une définition des constituants prosodiques basée sur les propriétés phonétiques et phonologiques de ces constituants et non exclusivement sur des algorithmes morphosyntaxiques. Cependant cette définition peut parfois poser des problèmes dans la mesure où le marquage des constituants prosodiques est multiparamétrique. En effet, différents types d'indices, tels que des indices acoustiques (indices de durée et indices mélodiques) ou des indices articulatoires (tel que le renforcement du contact entre la langue et la palais), marquent les frontières prosodiques. Ces différents indices ne concordent pas toujours dans le marquage des frontières. Par exemple, en l'absence de marquage tonal, la présence d'un allongement pré-frontière peut parfois suffire à percevoir une frontière prosodique. Très peu d'études menées jusqu'à ce jour pour le français se sont attachées à l'analyse de la contribution séparée des indices mélodiques et de durée dans la parole. Dans ce travail de thèse, nous nous sommes focalisée sur l'étude de deux types d'indices acoustiques : les indices mélodiques et les indices de durée. L'observation fine de ces deux catégories d'indices en production de la parole, nous a permis de mieux comprendre l'influence de différents facteurs tels que la structure syntaxique ou la vitesse d'élocution sur le marquage des frontières prosodiques. L'étude de la manière dont ces indices sont perçus par les auditeurs du français nous a également permis une meilleure appréhension du marquage des frontières prosodiques. En effet, dès lors que l'on admet un marquage pluriel des frontières de constituants prosodiques, l'étude de la saillance des différents types d'indices devient très importante. La durée, en tant que corrélat de l'allongement pré-final, est-elle suffisante pour percevoir une frontière prosodique en français ? Les indices mélodiques seuls peuvent-ils suffire à des auditeurs du français pour percevoir les différents niveaux de frontière prosodique en français ?

Dans cette thèse nous avons observé ces indices dans le but de mieux comprendre l'articulation entre la structure prosodique abstraite et sa réalisation concrète. Les différentes études réalisées nous ont conduites à mieux motiver l'existence d'un niveau additionnel de structure déjà postulé pour le français : le syntagme intermédiaire (cf. Jun & Fougeron, 2000). Les résultats obtenus nous ont permis de caractériser les propriétés phonéto-phonologiques de ce constituant ainsi que d'attester de sa validité cognitive.

2. RÉSULTATS MAJEURS

2.1. Étude 1 : Influence de la syntaxe et de la vitesse d'élocution sur la structuration prosodique

Dans une première étude, nous nous sommes intéressée à la question de l'interface entre syntaxe et prosodie. Nous avons testé si les indices acoustiques présents à la frontière d'AP contenue à l'intérieur d'un constituant syntaxique majeur (comme dans le syntagme nominal « Les a**MIS**JAP de Pauline ») et les indices acoustiques présents à la frontière d'un AP associée à une rupture syntaxique majeure (telle que la rupture entre un syntagme nominal et un syntagme verbal « Tes a**MIS**JAP/XP demandaient l'institutrice ») étaient différents. Les résultats obtenus dans le chapitre 2 nous ont permis de montrer que l'alignement de la frontière droite d'AP avec la rupture entre un SN et un SV n'avait pas d'influence sur les indices mélodiques dans des phrases déclaratives simples à focalisation large. À l'inverse, nous avons observé que le degré d'allongement pré-frontière associé à la frontière de syntagme accentuel était significativement plus important si cette frontière était associée à une rupture syntaxique majeure.

Nos résultats ont également montré que la réalisation du patron mélodique du syntagme accentuel est conditionnée par la vitesse d'élocution et par le nombre de syllabes contenues dans l'AP. En effet, nous avons observé plus de patrons mélodiques de type /LH*/ par opposition aux autres types de patrons mélodiques (/LHiLH*/, /LLH*/, LHiH*/, /HiLH*/, /LHiL*/) à vitesse d'élocution rapide et pour les APs contenant 4 syllabes que pour les APs contenant 3 syllabes. Un lien avait déjà été établi dans la littérature entre vitesse d'élocution et type de patron mélodique réalisé d'une part (Fougeron & Jun, 1998) et entre longueur du syntagme et réalisation de la montée initiale (LHi) d'autre part (Welby, 2006). Nos résultats confirment une simplification du patron mélodique produit à vitesse d'élocution rapide.

Enfin, les résultats obtenus au cours de cette première étude nous ont permis de montrer qu'à vitesse d'élocution rapide la relation hiérarchique entre unités prosodiques et allongements pré-finaux est presque toujours préservée. En effet, nous avons observé différents degrés d'allongement entre les syllabes inaccentuées et les syllabes en position finale d'AP d'une part et entre les syllabes en position finale d'AP et les syllabes en position finale d'IP d'autre part, conformément à ce qui avait été observé dans d'autres études

(Hirst & Di Cristo, 1984; Padeloup 1990; Vaissière, 1983). Notons que, contrairement à ce qui a été observé pour la vitesse d'élocution normale, aucune différence significative entre la durée des syllabes en position finale d'AP et la durée des syllabes en position finale d'AP/XP n'a été observée. Nos résultats permettent également de montrer qu'une augmentation de la vitesse d'élocution ne semble pas induire de restructuration significative des découpages prosodiques. En effet, seulement 2,3% des énoncés de notre corpus présentaient des phénomènes d'effacement de frontière (d'AP ou d'IP) ou de remplacement de frontières par une frontière de niveau inférieur dans la hiérarchie prosodique (c'est-à-dire une frontière d'IP remplacée par une frontière d'AP). Notons qu'à la différence de certaines dernières où les frontières prosodiques sont définies sur la base des indices tonals (Fougeron & Jun 1998), nous avons observé plusieurs cas où la seule présence de l'allongement était suffisante pour que les deux annotateurs perçoivent la frontière d'AP.

2.2. Étude 2 : Caractéristiques phonétiques et phonologiques du syntagme intermédiaire en français

Au cours de cette deuxième expérimentation, nous avons fourni des preuves phonologiques de l'existence d'un syntagme intermédiaire en français. Contrairement à ce qui avait été proposé précédemment par Jun & Fougeron (2000) nous avons montré qu'une frontière de syntagme intermédiaire peut apparaître en français dans tout énoncé dès lors que les structures prosodiques et syntaxiques le permettent. Dans cette expérience nous avons manipulé la longueur du syntagme nominal Sujet de phrases de façon à obtenir des syntagmes composés soit de 2 APs (« la mamie]AP de Rémy]AP[ip) » soit de 3 APs (« la mamie]AP des amis]AP de Rémy]AP[ip) »). Nous avons montré que les indices mélodiques et de durée associée au deuxième syntagme accentuel étaient significativement différents dans les deux types de séquences. Cette différence est liée à l'association de la syllabe cible à une frontière d'ip dans le premier type de syntagme, frontière qui est absente dans le second type de syntagme.

Nous avons donc proposé que la frontière droite de syntagme intermédiaire soit alignée avec une rupture syntaxique majeure dès lors que le syntagme est suffisamment long (composé au minimum de deux syntagmes accentuels) en français. Cette frontière est marquée par un accent de syntagme H- responsable de la réinitialisation du registre sur le dernier accent mélodique de l'ip. Dans nos données nous n'avons pas observé de cas où l'accent de syntagme s'étend du dernier accent mélodique jusqu'à la frontière droite de l'ip donnant lieu à des plateaux de *f₀* (cf. Jun & Fougeron, 2000). Cependant, nous n'excluons pas que, tout comme en anglais (Beckman & Pierrehumbert, 1986), l'accent de syntagme puisse s'étendre sur la gauche (d'une région allant du dernier accent mélodique à la fin de l'ip) dans les déclaratives en français.

Nous avons également mis en évidence un phénomène d'abaissement plus large que celui qui affecte les accents mélodiques en français. En effet,

nous avons montré que les syntagmes intermédiaires sont abaissés les uns par rapport aux autres au sein de l'IP. Ce phénomène d'abaissement entre les ips se manifeste par l'abaissement de la ligne de référence du registre marquée par le premier pic mélodique de l'ip. De ce fait un phénomène de réinitialisation partielle du registre est observé sur le premier pic mélodique situé après la frontière droite de l'ip.

D'autre part et contrairement à ce qui a été observé en allemand (Truckenbrodt & Féry, 2005) et en hollandais (van den Berg et al., 1992), il semblerait que le phénomène d'abaissement qui affecte les accents mélodiques ne soit pas plus fort que le phénomène d'abaissement qui intervient entre les syntagmes intermédiaires en français.

Dans cette deuxième étude nous avons également manipulé la vitesse d'élocution. Cette manipulation nous a permis de montrer que les indices acoustiques marquant la frontière d'ip sont préservés à vitesse d'élocution rapide. De plus nous avons observé une compensation très intéressante des indices mélodiques et de durée à vitesse d'élocution rapide. Alors que la différence de degré d'allongement entre les frontières d'AP et à la frontière d'ip était moins importante à vitesse d'élocution rapide qu'à vitesse d'élocution normale, les indices mélodiques présents à la frontière d'ip étaient eux renforcés avec l'augmentation de la vitesse d'élocution.

2.3. Étude 3 : Influence de la frontière droite de syntagme accentuel sur l'accès au lexique

Dans cette étude nous nous sommes intéressée à l'influence de la frontière droite de syntagme accentuel sur la segmentation de la parole en mots. De précédentes études (Christophe et al., 2004) ont montré que l'accès au lexique était retardé lorsqu'une ambiguïté lexicale temporaire apparaissait au sein d'un syntagme phonologique ou PP (ex. [le **chat** **grin**cheux] où /ʃa/ et /gʁẽ/ forme le mot « chagrin » en français) mais pas lorsque cette même ambiguïté apparaissait à une frontière de syntagme phonologique (ex. [le **chat**]PP [**grim**pait]PP). Cependant, dans cette étude les frontières de syntagmes phonologiques (prédites sur la base d'algorithmes syntaxiques sans vérification du découpage prosodique réellement produit par le locuteur) étaient toujours associées à une rupture syntaxique majeure. Or nous avons démontré dans cette thèse que les indices de durée présents à la frontière d'un syntagme accentuel associée à une rupture syntaxique majeure (frontière d'AP/XP) sont plus importants que les indices présents à la frontière de syntagme accentuel contenu dans un syntagme nominal. Nous nous sommes donc interrogée sur le niveau de frontière prosodique permettant de lever une ambiguïté lexicale en français. De manière plus précise, nous nous sommes demandée si une frontière de syntagme accentuel contenue à l'intérieur d'un syntagme nominal (« le **chat**]AP **grin**cheux ») serait suffisante pour lever une ambiguïté lexicale où si c'est l'alignement entre frontière de syntagme accentuel et rupture syntaxique majeure qui permet à des auditeurs du français d'identifier les frontières de mots (« le **chat**]AP/XP **grim**pait »). Pour cela nous avons demandé à des auditeurs du français de réaliser une tâche de

détection de mot. Les phrases contenaient des ambiguïtés lexicales temporaires pouvant être situées soit (i) à l'intérieur d'AP (« le **chat grincheux** »), soit (ii) à une frontière d'AP contenue à l'intérieur d'un syntagme nominal (« le **chat**]AP **grincheux** »), soit (iii) à une frontière de syntagme accentuel alignée avec une rupture syntaxique majeure (« le **chat**]AP/XP **grimpait** »). Les résultats obtenus ont montré que l'accès au lexique était retardé lorsque l'ambiguïté lexicale apparaissait au sein du syntagme accentuel mais pas à la frontière d'AP ou d'AP/XP. De plus nous avons observé que les temps de réaction des participants étaient significativement plus courts lorsque l'ambiguïté lexicale apparaissait à la frontière d'AP/XP plutôt qu'à la frontière d'AP. Ces résultats suggèrent que le syntagme accentuel (et non pas le syntagme accentuel/phonologique associé à une rupture syntaxique majeure contra Christophe et al., 2004) est le domaine de l'accès au lexique en français. Ils suggèrent également que le degré d'allongement associé à la frontière d'AP/XP (qui est significativement plus important que celui associé à la frontière d'AP) facilite l'accès au lexique en français.

2.4. Étude 4 : Perception de la frontière droite de syntagme intermédiaire et traitement syntaxique de la parole

Dans le but d'accéder au sens, les auditeurs d'une langue doivent découper la chaîne parlée en mots. Au delà de ce découpage, ils doivent également assigner les mots au sein d'une structure syntaxique. Nous savons que les indices présents à la frontière des constituants prosodiques sont exploités par les auditeurs du français dans le but de construire la structure syntaxique. Nous avons réalisé donc une tâche de complétion de phrases, où des participants devaient compléter des phrases tronquées. Ces séquences étaient identiques sur le plan segmental mais différentes du point de vue de leurs structure prosodique : la syllabe finale pouvait être associée soit à une frontière d'AP (LE PARVIS DE LA MAIRIE]AP de Paris devrait être...) soit à une frontière d'ip (LE PARVIS DE LA MAIRIE]ip devrait être....) La tâche était à choix forcé: les participants avaient le choix entre un syntagme prépositionnel (SP) ou un syntagme verbal (SV). Les résultats obtenus montrent que les participants ont préférentiellement complété les séquences (plus que le seuil du hasard) par un syntagme verbal lorsque les stimuli contenaient une frontière d'ip. A l'inverse ils ont préférentiellement complété les séquences par un syntagme prépositionnel lorsque la syllabe finale des séquences était une frontière d'AP. Les participants ont interprété la frontière de syntagme accentuel comme une rupture syntaxique mineure tandis qu'ils ont analysé la frontière de syntagme intermédiaire en tant que rupture syntaxique majeure. Ces résultats montrent que les auditeurs du français exploitent le niveau des frontières prosodiques dans la hiérarchie prosodique (frontière d'AP vs. frontière d'ip) dans le but d'inférer les catégories grammaticales des mots qui suivent ces frontières. Les indices prosodiques seraient donc exploités très tôt dans le processus d'analyse de la structure syntaxique et permettraient aux

auditeurs du français de construire des attentes sur la structure syntaxique des phrases entendues.

2.5. Étude 5 : Contribution séparée des indices mélodiques et de durée dans la construction de la syntaxe

Dans la dernière étude présentée dans cette thèse, nous avons manipulé de façon indépendante les différents types d'indices acoustiques (indices mélodiques et de durée) présents à la frontière du syntagme accentuel et du syntagme intermédiaire. Grâce à de la parole de resynthèse, nous avons, pour chaque niveau de frontière (AP et ip), créé trois conditions prosodiques : une où les frontières prosodiques étaient marquées uniquement par des indices mélodiques, une où les frontières étaient marquées uniquement par des indices de durée et enfin, une où les frontières étaient marquées à la fois par les indices mélodiques et de durée. Nous avons utilisé le même protocole expérimental que celui présenté lors de l'étude 4 afin de tester quel type d'indice acoustique était le plus saillant dans la perception des frontières prosodiques. Dans une tâche de complétion de phrase à choix forcé (complétion SP ou complétion SV), le type de réponse donné par les participants n'était pas influencé par le niveau de frontière prosodique présent dans les stimuli lorsque les frontières étaient marquées uniquement par des indices mélodiques. Par contre, les participants ont donné significativement plus de réponses de type SP que de type SV lorsqu'ils entendaient une frontière d'AP marquée uniquement par les indices de durée. Ce résultat semble indiquer que les indices de durée suffisent à marquer la frontière d'AP en français. A l'inverse, la présence conjointe des indices mélodiques et de durée semble nécessaire dans le but de signaler la frontière d'ip en français puisque nous n'avons pas observé plus de réponse de type SV lorsque les participants entendaient une frontière d'ip marquée uniquement par la durée. L'analyse des temps de réaction des participants a confirmé ces résultats et a montré que les participants utilisaient les indices prosodiques en temps réel dans le traitement syntaxique des énoncés parlés.

3. DIRECTIONS POUR NOS RECHERCHES FUTURES

3.1. Parole de laboratoire et parole spontanée

Une limite des expérimentations que nous avons menées est que nous avons eu recours à de la parole lue. Cependant, l'utilisation de cette parole très contrôlée nous a permis de mettre en évidence les propriétés phonétiques et phonologiques du syntagme intermédiaire en français. Grâce aux différents corpus que nous avons mis en place pour les besoins spécifiques de chaque étude, nous avons pu mesurer de façon très fine des phénomènes d'allongements ou de variations de registre parfois plus difficile à mesurer en parole spontanée où les informations de nature interactionnelle et discursive entrent en jeu.

Nous savons que la parole lue et la parole spontanée diffèrent sur de nombreux points et notamment du point de vue de la structure intnative. On admet couramment dans la littérature qu'il est difficile de retrouver les unités prosodiques définies sur la base de leurs propriétés phonétiques et phonologiques dans les corpus de parole spontanée. Parce qu'il permet une meilleure compréhension de l'influence de différents facteurs tels que la structure syntaxique ou de la vitesse d'élocution et parce que nous cherchons à mieux comprendre la contribution séparée de chaque type d'indice acoustique, notre travail de thèse peut s'avérer très utile pour tous les travaux qui s'intéressent à l'identification des unités prosodiques en interaction.

A l'inverse, nous pensons également que des études menées sur de la parole spontanée permettrait une meilleure compréhension de la structure prosodique abstraite du français. La question des unités prosodiques inachevés ou même la construction conjointe d'une unité prosodique par deux inter actants, nous paraissent par exemple des phénomènes particulièrement intéressants à étudier.

3.2. Frontière de syntagme intermédiaire et structure informationnelle

Si nous pensons qu'il est non seulement possible d'identifier dans le discours spontané les unités prosodiques telles que nous les avons définies dans cette thèse (c'est-à-dire sur la base de leurs propriétés phonétiques et phonologiques) nous pensons également que ces unités jouent un rôle organisateur dans le discours. Le fait que le découpage prosodique soit influencé par la structure informationnelle des phrases a été montré pour de nombreuses langues dont le français (voir Gussenhoven 2004 pour une discussion sur le sujet). Féry (2001) a notamment proposé que les locuteurs du français utilisent le découpage prosodique dans le but de signaler la focalisation en français. Selon cette proposition, une focalisation contrastive étroite engendre l'apparition d'un ton frontière H%, plus haut que ce qu'il ne l'est habituellement. Nous pensons quant à nous, qu'une des stratégies

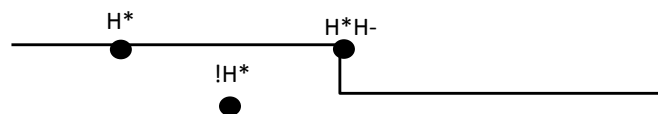
adoptée par les locuteurs français pour de marquer un élément focalisé est de remplacer un niveau de frontière donné dans la hiérarchie prosodique, par un niveau de frontière supérieur dans la hiérarchie prosodique. Par exemple une séquence de syllabes qui serait structurée comme un syntagme accentuel au sein d'un syntagme intermédiaire pourrait donner lieu à un phénomène de restructuration prosodique et être associée à une frontière de syntagme intermédiaire (donc à un niveau plus de hiérarchie prosodique plus élevé). Ainsi la structure prosodique de la phrase 1a réalisée avec une focalisation large pourrait être restructurée et donner lieu à celle illustrée en 1b en raison d'une focalisation étroite sur le second AP du syntagme nominal Sujet « de diamants ».

1a. [Le collier de diamants de mamie a été vendu la semaine dernière]F
]AP

1b. Le collier [de diamants]F de mamie a été vendu la semaine dernière]F
]ip

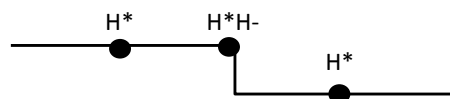
Cette restructuration entraînerait un retour à la ligne de registre sur la syllabe « mants » de diamants en 1b mais pas en 1a tel que cela est schématisé dans la figure 78.

a. Focalisation large



Le collier]AP de diamants]AP de mamie]AP]ip a été vendu la semaine dernière.

b. Focalisation étroite sur diamants



Le collier]AP de diamants]AP]ip de mamie]AP a été vendu la semaine dernière.

Figure 78. *ip et structure informationnelle.*

Schématisation de la structure intonative attendue pour les phrases « Le collier de diamants de mamie a été vendu la semaine dernière » prononcée avec une focalisation large (78a) où une focalisation étroite sur « diamants » (78b). Le registre est réinitialisé sur le troisième accent mélodique de l'énoncé dans la phrase 78a. tandis qu'il est réinitialisé sur le deuxième accent mélodique de l'énoncé dans la phrase 78b en raison d'un phénomène de restructuration prosodique liée à la focalisation étroite sur « diamants ».

Nous pensons donc que la caractérisation du syntagme intermédiaire telle qu'elle a été proposée dans cette thèse pourrait être utile dans le but d'étudier de quelle façon des contraintes relatives à la structure

informationnelle interagissent avec les structures prosodiques et syntaxiques dans le découpage en constituants prosodiques en français.

3.3. Corrélats articulatoires du syntagme intermédiaire

Dans cette thèse nous nous sommes focalisée sur l'étude des indices acoustiques marquant les frontières de constituants prosodiques et plus particulièrement sur la caractérisation de la frontière de syntagme intermédiaire. Cependant d'autres types d'indices participent au marquage des constituants prosodiques. Parmi ces indices, des corrélats articulatoires du découpage prosodique ont été mis en évidence pour différentes langues dont le français (pour le français : Fougeron & Keating, 1997 ; Fougeron, 2001 ; pour le coréen : Cho & Keating, 2001 ; Jun, 1993 *inter alia*). Ces études ont montré que certains corrélats articulatoires sont renforcés à la frontière de constituants prosodiques plutôt qu'à l'intérieur de ceux-ci. Ce renforcement des gestes articulatoires a été mis en évidence aussi bien pour la frontière droite, c'est-à-dire en position finale de constituant (voir par exemple Tabain 2003) que pour la frontière gauche c'est-à-dire en position initiale du constituant (voir par exemple Byrd & Saltzman, 1998 ; Byrd, 2000 ; Tabain, 2003).

Il a également été montré que les caractéristiques articulatoires observées sont à mettre en relation avec le niveau de structuration dans la hiérarchie prosodique. En d'autres termes, ces modifications articulatoires s'amplifient de façon graduelle avec le niveau du constituant dans la hiérarchie (Byrd & Saltzman, 1998 ; Byrd, 2000 ; Cho, 2005 ; Tabain, 2003 ; Tabain & Perrier, 2005 ; Cho & Keating, 2001 ; Fougeron, 2001 ; Cho, 2005 ; Keating et al., 2004 ; Tabain, 2003). Etant donné que nous avons mis en évidence un niveau additionnel de structuration prosodique, il serait pertinent de tester si l'existence de ce constituant se reflète dans les modifications articulatoires observées à la frontière des constituant. Notre hypothèse est que conformément à ce qui a été observé pour les indices de durée pour lesquels nous observons une augmentation graduelle du degré d'allongement associé à la frontière prosodique en fonction de son niveau dans la hiérarchie prosodique, les corrélats articulatoire tels que le contact entre la langue et le palais (cf. Fougeron & Keating, 1997) seraient significativement plus importants en position initiale de l'ip qu'en position initiale de l'AP. De manière similaire nous pouvons postuler que le contact linguo-palatal serait plus important à l'initiale de l'IP qu'à l'initiale de l'AP en français.

3.4. Application aux technologies de la parole

Nos résultats ont également des applications potentielles pour les technologies de la parole. La prosodie de manière générale et l'intonation en particulier représentent une source supplémentaire de connaissance pour le domaine de la reconnaissance de la parole. L'identification des frontières de constituants prosodiques peut notamment permettre une meilleure segmentation de la parole. Dans cette thèse nous avons montré que la montée

finale de *fo* et l'allongement pré-frontière sont des indices qui facilitent le repérage de la frontière droite des mots en français. Ce type d'indice mériterait d'être intégré à des systèmes de reconnaissance vocale.

De plus grâce à une approche multiparamétrique des frontières prosodiques et à l'étude de la contribution séparée de chaque type d'indice acoustique, nous avons montré que les indices de durée étaient suffisants pour percevoir la frontière d'AP en français. Depuis longtemps les études menées en technologie de la parole s'intéressent à la saillance des différents types d'indices acoustiques (Vaissière, 1983 ; Ostendorf & Veilleux, 1994). Notre thèse peut donc être une contribution intéressante pour tout système cherchant à repérer de façon automatique les frontières de constituant prosodique.

Enfin, la prise en compte de phénomènes d'abaissement et de variation de registre que nous avons mis en évidence dans cette thèse permettrait également d'améliorer à la fois les systèmes de reconnaissance mais aussi les systèmes de synthèse vocale dans le but d'aboutir à une parole toujours plus « naturelle ».

3.5. Sensibilité des jeunes enfants aux frontières prosodiques

Nous savons aujourd'hui que les jeunes enfants sont sensibles aux frontières prosodiques. Par exemple grâce à la technique de succion non-nutritive, Christophe, Dupoux, Bertoncini & Melher (1994) ont montré que les nouveau-nés français pouvaient discriminer deux syllabes qui ne diffèrent que par la présence d'une frontière prosodique (comme les syllabes /ma/ dans « mathé**MA**ticien » ou panora**MA** typique »). Dans son travail de thèse, Millotte (2005) a également montré que des enfants de 16 mois peuvent utiliser les frontières de syntagmes phonologiques pour segmenter la parole continue.

Dans cette thèse, nous avons étudié la contribution de chaque type d'indices acoustiques dans la perception des frontières prosodiques droites. Nous avons montré que les indices de durée étaient suffisants pour que des adultes perçoivent la frontière droite de syntagme accentuel et s'en servent dans le traitement syntaxique des phrases. Il serait intéressant d'étudier si de la même manière, les enfants perçoivent la différence entre des syllabes contenues à l'intérieur d'un syntagme accentuel et des syllabes associées à une frontière d'AP marquées uniquement par des indices de durée.

D'autres études ont également montré qu'en plus de percevoir les frontières prosodiques, les enfants ont des attentes quant à l'apparition de ces frontières dans des phrases. Par exemple Hirsh-Pasek et al. (1987) ont montré que des enfants américains âgés de 9 mois préféraient écouter des phrases où les frontières syntaxiques correspondaient à des frontières de syntagmes intonatifs qui étaient marquées à la fois par des pauses silencieuses, un allongement pré-frontière et un ton de frontière (par exemple « Cindarella lived in a big house]IP but it was dark] because she had a mean stepmother]IP ») plutôt que des phrases dans lesquelles les frontières de groupes intonatifs étaient insérés à l'intérieur de propositions et où

l'apparition des pauses n'étaient pas corrélées avec des variations de durée et d'intonation (« Cindarella lived (pause) in a big house but it was (pause) dark because she had a mean stepmother. (cf. également et Jusczyk et al. 1992, Gerken et al. 1994).

Nous pouvons également nous demander si les enfants construisent des attentes quant au niveau de frontière (AP, ip, IP) apparaissant dans les phrases en français. Dans cette thèse nous avons montré que les adultes se servent du niveau de frontière (AP/ip) afin d'inférer la catégorie morphosyntaxique des mots qu'ils doivent insérer au sein de la structure syntaxique des phrases. Plus précisément nous avons montré que les frontières de syntagme accentuel sont interprétées par des adultes comme des ruptures syntaxiques majeures tandis que les frontières de syntagmes accentuels sont interprétées en tant que rupture syntaxique mineure. Nous nous demandons si de tels indices peuvent être perçus par les enfants et utilisés dans le but de construire une représentation syntaxique des phrases. Notre idée est que tout comme les adultes, les indices de frontières prosodiques pourraient leur permettre de construire des attentes quant à la catégorie grammaticale des mots qui suivent les frontières prosodiques. Nous pourrions par exemple postuler que des enfants préféreraient entendre des phrases dans lesquelles le niveau de frontière prosodique (AP ou ip) serait en adéquation avec les contraintes syntaxiques pesant sur le découpage prosodique. Par exemple dans des phrases telles que la phrase 2, nous prédisons que des enfants préféreront entendre les phrases produites avec la structure présentée en 2a. (où une frontière d'ip est associée à la rupture entre un syntagme nominal et un syntagme verbal) qu'avec la structure présentée en 2b. (où une frontière de plus bas niveau est associée avec la rupture entre le SN et le SV).

2. La grand-mère du petit chaperon rouge était très malade.

- a. [[La grand-mère]AP [du petit chaperon rouge]AP]**ip** était très malade.
- b. [[La grand-mère]AP [du petit chaperon rouge]AP] était très malade.

Notre travail de thèse ouvre donc de nouvelles et d'intéressantes perspectives de recherche. Grâce à l'étude des relations existantes entre la structure prosodique abstraite et sa réalisation concrète, nous avons proposé l'existence d'un niveau additionnel de structuration prosodique pour le français : le syntagme intermédiaire. Nous avons défini ce syntagme à partir de ses propriétés phonétiques et phonologiques. Nous avons également comparé les propriétés de ce constituant avec les deux autres niveaux de structure attesté en français : le syntagme accentuel et le syntagme intermédiaire. Notre travail a permis de fournir une base d'analyse plus explicite des indices acoustiques marquant les frontières des constituants prosodiques en français. Cette base d'analyse pourra permettre à différents domaines de parvenir à une meilleure définition des constituants prosodiques particulièrement pour les études menées sur de la parole spontanée ou les

recherches en psycholinguistique, où il est indispensable d'adopter une définition des unités prosodiques basée sur les indices phonétiques et phonologiques réellement produits par les locuteurs.

Dans ce travail de thèse nous avons cherché à mieux comprendre les relations existantes entre la structure prosodique abstraite et sa réalisation concrète dans la parole. Nous avons développé l'hypothèse de l'existence d'un niveau de structuration prosodique additionnel situé entre les deux niveaux de structure traditionnellement postulés pour le français. Grâce à cinq expérimentations nous avons attesté l'existence de ce niveau à la fois en production et en perception de la parole.

RÉFÉRENCES

- Astésano, C. (2001). *Rythme et accentuation en français. Invariance et variabilité stylistique*. Paris : Editions L'Harmattan, Collection Langue et Parole.
- Baayen, R. H. (à paraître). *Analyzing Linguistic Data : A practical introduction to Statistics*. Cambridge : Cambridge University Press.
- Bagou, O., & Frauenfelder, U.-H. (2006). Stratégie de segmentation prosodique: rôle des proéminences initiales et finales dans l'acquisition d'une langue artificielle, *In Actes des XXVIe Journées d'études sur la Parole*, Dinard, France.
- 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 In *Proceedings of the 1st Conference on Speech prosody*, Aix-en-Provence, France, 159-162.
- Banel, M.-H., & Bacri, N. (1994). On metrical patterns and lexical parsing in French. *Speech Communication*, 15, 115-126.
- Barthkova, K. (1991). Speaking rate modelization in French: Application to speech synthesis. In *Proceedings of the XIIth International Congress of Phonetic Sciences*, 3, 482-485.
- Bates D, Maechler M, Bolker B (2011). lme4: Linear Mixed-Effects Models Using S4 Classes.R package version 0.999375-38, URL <http://CRAN.R-project.org/package=lme4>.
- Beach, C.-M. (1991). The interpretation of prosodic patterns at point of syntactic ambiguity : evidence for a cue trading relations. *Journal of Memory and Language*, 30, 664-663.
- Beckman, M.-E., & Pierrehumbert, J. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, 3, 255-309.
- Beckman, M., Edwards, J., & Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics, In *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, Docherty G., & Ladd, D.-R. (Eds.), 68-86. Cambridge : Cambridge U.P.
- Beckman, M.-E., & Edwards, J. (1990). Lengthenings, shortenings, and the nature of prosodic constituency. In Kingston, J., & Beckman, M.-E. (Eds.), *Papers in Laboratory Phonology I : Between the Grammar and the Physics of Speech*, 152-178. Cambridge : Cambridge University Press.
- Beckman, M.-E., & Ayers Elam, G.-A. (1997). Guidelines for ToBI labelling, version 3.0, unpublished ms. (available online at http://ling.ohio-state.edu/~tobi/ame_tobi/labelling_guide_v3.pdf; date of last accessed: 2011-01-26).
- Bever, T. (1970). The cognitive basis for linguistic structure. In Hayse, J.-R. (Eds.), *Cognition and Developpement of Language*. New York : Villey.
- Boersma, P. & Weenink, D. (2009). Praat: doing phonetics by computer (Version 5.1). www.praat.org.

- Boland, J.-E., & Blodgett, A. (2001). Understanding the constraints on syntactic generation : Lexical bias and discourse congruency. *Journal of Memory and Language*, 45, 391-411.
- Bolinger, D. (1951). *Intonation : levels versus configurations*. *Word*, 7, 199-210.
- Breen, M., Gibson, E. & Watson, D. (2010). Intonational phrasing is constrained by meaning, not balance. *Language and Cognitive Processes*, First published on: 20 September 2010 (iFirst).
- Britt, M.-A., Perfetti, C.-A., Garrod, S. & Rayner, K. (1992). Parsing in discourse : context effects and their limits. *Journal of Memory and Language*, 31, 293-314.
- Brown, G., & Yule G. (1983), *Discourse Analysis*, Cambridge University Press.
- Bruce, G. 1977. *Swedish Word Accents in Sentence Perspective*. Ph.D. Dissertation, Lund University, Lund, Sweden.
- Byrd, D., & Saltzman, E. (1998). Intragestural dynamics of multiple phrasal boundaries. *Journal of Phonetics*, 26, 173-199.
- Byrd, D., & Saltzman, E. (2003). The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, 31, 149-180.
- Caspers, J. (1994) *Pitch movements under time pressure: effects of speech rate on the melodic marking of accents and boundaries in Dutch*. Ph.D. Dissertation, Holland Institute of Generative Linguistics, University of Leiden.
- Caspers, J. & van Heuven, V. J. (1991) Phonetic and linguistic aspects of pitch movements in fast speech in Dutch. In *Proceedings of the XIIIth International Congress of Phonetic Sciences*, 5, 174-177.
- Caspers, J. & van Heuven, V. J. (1993) Effects of time pressure on the phonetic realization of the Dutch accent-lending pitch rise and fall, *Phonetica*, 50, 161-171.
- Caspers, J. & van Heuven, V. J. (1995) Effects of time pressure on the choice of accent-lending and boundary-marking pitch configuration in Dutch. In *Proceedings of Eurospeech*, 2, 1001-1004.
- Cho, T., & P. Keating (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics*, 29(2), 155-190.
- Cho, T. (2005) Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /a,i/ in English. *Journal of the Acoustical Society of America*, 117, 3867-3878.
- Chomsky, N., & Halle, M. 1968. *The Sound Pattern of English*. New York : Harper and Row.
- Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (1994). Do infants perceive word boundaries ? An empirical study of the bootstrapping of lexical acquisition. *Journal of Acoustical Society of America*, 95(3), 1570-1580.
- Christophe, A., Peperkamp, S., Pallier, C., Block, E., & Mehler, J. (2004). Phonological phrase boundaries constrain lexical access : I. Adult data. *Journal of Memory and Language*, 51, 523-547.
- Clopper, C., & Smiljanic, R. (à paraître). Effects of gender and regional dialect on prosodic patterns in American English. *Journal of Phonetics*.
- Cohen, A., & t'Hart, J. (1967). On the anatomy of intonation. *Lingua*, 19, 177-192.
- Cooper, W.-E., & Paccia-Cooper, J. (1980). *Syntax and Speech*. Cambridge MA : Harvard University Press.
- Crain, S., & Steedman, M.-J. (1985). On not being led up the garden path : the use of context by the psychological parser. In Dowty, D. Karttune, L., & Zwicky, A. (Eds.), *Natural Language Parsing : Psychological, Computational, and Theritical Perspectives*, 320-368, Cambridge

- University Press.
- Cruttenden, A. (1986). *Intonation*. Cambridge: CUP.
- Crystal, D. 1969. *Prosodic Systems and Intonation in English*. Cambridge : Cambridge University Press.
- Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 113-121.
- Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the sentence comprehension of spoken sentences. Special Issue : Sentence Processing III. *Journal of Psycholinguistic Research*, 22, 109-131.
- D'Imperio, M. (2000). *The Role of Perception in Defining Tonal Targets and their Align-ment*. Ph.D. Dissertation, The Ohio State University, Ohio State, USA.
- D'Imperio, M., Elordieta, G., Frota, S., Prito, P. Vigário, M. (2005). Intonational phrasing and constituent length in Romance. In Frota, S., Vigário, M., Freitas, M.-J. *Prosodies* (Selected papers from the Phonetics and Phonology in Iberia Conference, 2003). [Phonetics and Phonology Series], 95-98. The Hague: Mouton de Gruyter.
- Dufour, S., Nguyen, N., & Frauenfelder, U. (2007). The perception of phonemic contrasts in a non-native dialect. *Journal of the Acoustical Society of America* (JASA), 121(4), EL131-EL136.
- Delais-Roussarie, E. (1995). *Pour une approche parallèle de la structure prosodique: Etude de l'organisation prosodique et rythmique de la phrase française*, Thèse de Doctorat, Université de Toulouse le Mirail.
- Delais-Roussarie, E., & Post, B. (2008). Unités prosodiques et grammaire de l'intonation : vers une nouvelle approche. In *Actes des XXVIIème Journées d'Études sur la Parole*, Avignon, France.
- Delattre, P. 1966. Les dix intonations de base du français. *The French Review*, 40(1), 1-14.
- Dell, F.; Hirst, D.; Vergnaud, J.-R. (1984). *La forme sonore du langage. Structure des représentations en phonologie*. Paris: Hermann.
- Den Ouden, H., Noordman, L., & Terken, J. (2009). Prosodic realizations of global and local structure and rhetorical relations in read aloud news reports. *Speech Communication*, 51, 116-129.
- Di Cristo, A. (1998). Intonation in French. In Hirst, D., & Di Cristo, A. (Eds.), *Intonation systems: A survey of twenty languages*, 195-218. Cambridge: Cambridge University Press.
- Di Cristo, A. (1999b). Vers une modélisation de l'accentuation du français: première partie. *French Language Studies*, 9, 143-179.
- Di Cristo, A. & Jankowski L. (1999). Prosodic organization and phrasing after focus in French. In *Proceedings of XIVth International Congress of Phonetic Science*, San Francisco, Etats-Unis.
- Di Cristo, A. (2000). Vers une modélisation de l'accentuation du français: seconde partie. *French Language Studies*, 10, 27-44.
- Di Cristo, A. (2010). Accentuation et phrasé prosodique en français. *Journal of French Language Studies*.
- Di Cristo, A. & D. Hirst (1993). Rythme syllabique; Rythme mélodique et représentation hiérarchique de la prosodie du français. *Travaux de l'Institut de Phonétique d'Aix*, 15, 9-24.
- Di Cristo, A. & Hirst, D. (1996). *Vers une typologie des unités intonatives du français*. In *Actes des XXIème Journées d'Études sur la Parole*, Avignon, France, 219-22.
- Di Cristo, Albert. (1999a). Le cadre accentuel du français contemporain: Essai de mod élisation (première partie). *Langues*, 2(3), 184-205.

- Dufour, S., Frauenfelder, U.H. (2007). L'activation et la sélection lexicales lors de la reconnaissance des mots parlés: modèles théoriques et données expérimentales. *L'Année Psychologique*, 107, 87-112.
- Edwards J., & Beckman M., (1988). Articulatory Timing and the Prosodic Interpretation of Syllable Duration, *Phonetica*, 45, 156-174.
- Feldhausen, I. (2010). *Sentential Form and Prosodic Structure of Catalan*. Amsterdam: John Benjamins.
- Feirreira, F., & Clifton, Jr.-C. (1986). The independance of syntactic processing. *Journal of Memory and Language*, 25, 348-368.
- Ferreira, F., & Henderson, J.-M. (1991). Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language*, 25, 348-368.
- Ferreira, F. (1993). Creation of prosody during sentence production. *Psychological Review*, 100(2), 233-253.
- Féry, C. (1993). *German Intonational Patterns*. Tübingen, Germany : Niemeyer.
- Féry, C. (2001). Intonation of focus in French. In Féry, C. & Sternefeld, W. (Eds.), *Audiatur Vox Sapientes: A Festschrift for Arnim von Stechow*. Akademi Verlag, 153-181.
- Fletcher, J. (1991) Rhythm and final lengthening in French, *Journal of Phonetics*, 19(2), 193-212.
- Fon, Y. J. (2002). *A cross-linguistic study on syntactic and discourse boundary cues in spontaneous speech*. PhD. Dissertation, The Ohio State University.
- Fónagy, I. (1979). Fonction prédictive de l'intonation. In Léon, P., & Rossi, M. (Eds.), *Problèmes de prosodie*, 113-122, *Studia Phontica*, 18.
- Ford, M., Bresnan, J., & Kaplan, R.-M. (1982). A competence-based theory of syntactic closure. In Bresnan, J. (Eds.), *The mental representation of Grammar relations*, 727-796, Cambridge, MA : The MIT Press.
- Forster, K., & Forster, J. C. (2000). DMDX version 2. Retrived from the internet on 12/2000 from <http://www.ii.urizona.C'l»i/-iforster/dmdx.htm>.
- Fougeron, C. (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics*, 29, 109-135.
- Fougeron, C., & Jun, S.-A. (1998) Rate effects on French intonation: prosodic organization and phonetic realization. *Journal of Phonetics*, 26, 45-69.
- Fougeron, C., & Keating, P.A. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of Acoustical Society of America*, 101(6), 3728-3740.
- Frauenfelder, U. H., & Peters, G. (1990). On lexical segmentation in TRACE : An exercise in Simulation. In Altmann, G.-T.-M. (Eds.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives*, 50-86, Cambridge, MA: MIT Press,
- Frauenfelder, U.H. (1991). *Une introduction aux modèles de reconnaissance de mots parlés*. In Kolinski, R., Morais, J., & Segui, J. (Eds), *La reconnaissance des mots: études psycholinguistiques*, Paris, Presses Universitaires, 7-36.
- Frazier ; L., & Fodor, J.-D. (1978). The sausage machine : A new two-stage parsing model. *Cognition*, 6, 291-325.
- Frazier, L. (1978). *On comprehending sentences : Syntactic parsing strategies*. Bloomington : Indiana University Linguistics Club.
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: eye movments in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14, 178-210.
- Frazier, L., Clifton, C., & Randall, J. (1983). Filling gaps : decision principles and structure in sentence comprehension. *Cognition*, 13, 187-222.

- Frota, S., D'Imperio, M., Elordieta, G., Prieto, P., & Vigário, M., (2007). The phonetics and phonology of intonational phrasing in Romance. In Prieto, P. Mascaró, J. Solé, M.-J. (Eds.) *Segmental and prosodic issues in Romance Phonology*, 131-154. Amsterdam: John Benjamins.
- Gaitenby, J. (1965). The elastic word. *Haskins Laboratories Status Report on Speech Research*, SR-2, 3.1-3.12.
- Gee, J.P., & Grosjean, F. (1983). Performances structures : a psycholinguistic and linguistic appraisal. *Cognitive Psychology*, 15, 411-458.
- German, J., & D'Imperio, M. (2010). Focus, phrase length, and the distribution of phrase-initial rises in French. In *Proceedings of the Vth Speech Conference*, Chicago, United-States.
- Ghini, M. (1993). Φ -formation in Italian: a new proposal. *Toronto Working Papers in Linguistics*, 12, 41- 79.
- Goldman, J.P. (2007), EasyAlign : a semi-automatic phonetic alignment tool under Praat. <http://latcui.unige.ch/phonetique>.
- Goldsmith, J. (1976). *Autosegmental Phonology*. Ph.D. Dissertation, Massachusetts Institut Of Technology, Cambridge, Massachussets.
- Grice, M., D'Imperio, M., Savino, M., & Avesani, C. (2005). Towards a strategy for la- belling varieties of Italian. In Jun, S.-A. (Eds), *Prosodic Typology and Transcription : A Unided Approach*, 55-83. Oxford : Oxford University Press.
- Grice, M., Ladd, D.-R., & Arvaniti, A. (2000). On the place of phrase accents in intonational phonology, *Phonology*, 17, 143-185.
- Gussenhoven, C. (1984). On the Grammar and Semantics of Sentence Accents. Foris.
- Gussenhoven, C. (1988). Autosegmental Studies on Pitch Accent. In van der Hulst, H., & Smith, N. (Eds.), *Adequacy in intonation analysis : the case of Dutch*, 95-121. Dordrecht : Foris.
- Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge: Cambridge University.
- Halliday, M. A. K. (1967). *Intonation and Grammar in British English*. The Hague : Mouton.
- Hayes, Bruce. (1989). The prosodic hierarchy in meter. In Kiparsky, P., & Youmans, G. (Eds.), *Rhythm and Meter*, 201-260. Orlando: Academic Press.
- Hellmuth, S. (2004). *Prosodic weight and phonological phrasing in Cairene Arabic*. Ph.D. Dissertation, School of Oriental & African Studies, University of London.
- Hellmuth, S. (2007). The relationship between prosodic structure and pitch accent distri-bution : Evidence from Egyptian Arabic. *The Linguistic Review*, 24, 291-316.
- Hirschberg, J., & Pierrehumbert, J. (1986) Intonational Structuring of Discourse, *Proceedings of the 24th Meeting of the Association for Computational Linguistics*, 136-144.
- Hirsh-Pasek, K., Neslson, D.-G.-K, Jusczyk, P.W, Cassidy, K.W., Druss, B., & Kennedy, L. (1987). Caluses are perceptual units for young infants. *Cognition*, 26, 269-286.
- Hirst, D., & Di Cristo, A. (1984). French intonation: A parametric approach. *Die Neueren Sprachen*, 83(5), 554-569.
- Hirst, D., & Di Cristo, A. (1996). Ya-t-il des unités tonales en français? In *Actes des XXIèmes journées d'étude sur la parole*, Avignon, France, 223-226.
- Hirst, D., & Di Cristo, A. (1998). A survey of intonation systems. In Hirst, D., & Di Cristo, A. (Eds.), *Intonation Systems: A survey of twenty lan- guages*,

- Cambridge: Cambridge University Press, 1–44.
- Ito, J., & Mester, A. (1992). Weak Layering and Word Binarity. *Linguistic Research Center, LRC-92-09*, University of California, Santa Cruz.
- Jackendoff R (1977) *X-Bar Syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press.
- Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, 20, 137-194.
- Jun, S.-A. (1993). The phonetics and phonology of Korean prosody. PhD. Dissertation, The Ohio State University.
- Jun, S.-A. (1998). The Accentual Phrase in the Korean prosodic hierarchy. *Phonology*, 15, 189-226.
- Jun, S.-A. (2003). Prosodic Phrasing and Attachment Preferences, *Journal of Psycholinguistic Research* 32(2), 219-249.
- Jun, S.A., & Fougeron, C. (1995). The accentual phrase and the prosodic structure of French. In *Proceedings of the XIIIth International Congress of Phonetic Sciences*, 2, 722-725.
- Jun, S.A., & Fougeron, C. (2000). A phonological model of French intonation. In A. Botinis (Eds.), *Intonation: Analysis, modelling and technology*. Kluwer Academic Publishers: Boston, MA, 209-242.
- Jun, S.A., & Fougeron, C. (2002). Realizations of accentual phrase in French. *Probus*, 14, 147-172.
- Jusczyk, P.-W., Kemler-Nelson, D.-G., Hirs-Pasek, K., Kennedy, L., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, 24, 252-293.
- Keating, P., Cho, T., Fougeron, C., & Hsu, C. (2004). Domain-initial articulatory strengthening in four languages. In: *Phonetic Interpretation (Papers in Laboratory Phonology VI)*. Edited by J. Local, R. Ogden & R. Temple, 143-161. Cambridge, UK: Cambridge University Press,.
- Keller, E., & Zellner, B. (1995). A statistical timing model for French. In *Proceedings of XIIIth International Congress of Phonetic Sciences*, 3, 302-305.
- Keller, E., Zellner, B., Werner, S., & Blanchoud, N. (1993): The prediction of prosodic timing: Rules for final syllable lengthening in French. In House, D., Touati, P. (Eds.), *Proceedings of the ESCA Workshop on Prosody*, Lund, Sweden, 212–215.
- Kim, S. (2004). The role of prosodic phrasing in Korean word segmentation. Ph.D. Dissertation, University of California, Los Angeles.
- Kim, S., & Cho, T. (2009). The use of phrase-level prosodic information in lexical segmentation: Evidence from word-spotting experiments in Korean. *Journal of the Acoustical Society of America*, 125(5), 3373-3386.
- Kjelgaard, M., & Speer, S.-R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language*, 40, 153–194.
- Klatt, D.H. (1975). Vowel lengthening is syntactically determined in connected discourse. *Journal of Phonetics*, 3, 129-140.
- Krajilic, T., & Brennan, S.E. (2005). Prosodic disambiguation of syntactic structure : for the speaker or for the addressee ? *Cognitive Psychology*, 50(2), 194-231.
- Ladd, D.-R. (1988). Declination, «Reset » and the hierarchical organization of utterances. *Journal of Acoustical Society of America*, 84, 530-534.
- Ladd, D.-R. (1986). Intonational phrasing: The case for recursive prosodic structure. *Phonology Yearbook* 3, 311-340.
- Ladd, D.-R. (1992). An introduction to intonational phonology. Docherty, Ladd

- (Eds.), *Papers in Laboratory Phonology. II : Gesture, segment, prosody*, 321-334.
- Ladd, D.-R. (1996). *Intonational Phonology*. Cambridge : Cambridge University Press.
- Ladd, D.-R., & Campbell, N. (1991). Theories of prosodic structure: Evidence from syllable duration. *Proceedings of the XIIth International Congress of Phonetic Sciences*, 2, 290-293.
- Lehiste, I. (1973). Phonetic disambiguation of syntactic ambiguity. *Glossa*, 7, 107-122.
- Liberman, M. 1975. *The Intonational System of English*. Ph.D. Dissertation, Massachusetts Institut Of Technology, Cambridge, Massachussets.
- Liberman, M., & Pierrehumbert, J. (1984). Intonational Invariance under Changes in Pitch Range and Length. In Aronoff, M., & Oehrl, R.-T. (Eds), *Language Sound Structure*, 157-234, Cambrige, MA : MIT Press.
- Liberman, M., & Prince, A. 1977. On stress and linguistic rhythm. *Linguistic Inquiry*, 8, 249-336.
- Liberman, M., & Sag, I. 1974. Prosodic form and discourse function. *CLS*, 10, 487-497.
- MacDonald, M.-C., Pearlmutter, N.J., & Seidenberg, M.S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101(4), 676-703.
- MacWhinney, B. (1987). The completion model. In MacWhinney, B. (Eds.), *Mechanisms of language acquisition* (249-308). Hillsdale, New Jersey : Lawrence Erlbaum Associates.
- Mattys, S.-L., Jusczyk, P.-W, Luce, P.-A, & Morgan, J.-L. (1999). Phono- tactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*, 38, 465-494.
- McClelland, J.L., & Elman, J.L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18, 1-86.
- McLennan, C.T., & Luce, P.A. (2005). Examining the time course of indexical specificity effects in spoken word recognition. *Journal of Experimental Psychology : Learning, Memory and Cognition*, 31 :306-321.
- Mertens, P. (1993). Intonational grouping, boundaries and syntactic structure in French. In House, D., & Touati, P. (Eds.), In *Proceedings of the ESCA Workshop on Prosody*, Lund, Sweden, 155-159.
- Mertens, P., Auchlin, A., Goldman, J., & Grobet, A. (2001). L'intonation du discours : une implémentation par balises; motifs et premiers résultats. *Journées Prosodie 2001*.
- Mertens, P., Goldman, J.-P, Wehril, E., & Gaudinat, A. (2001). La synthèse de l'intonation à partir de structures syntaxiques riches. *Traitement automatique des langues*, 42 (1), 145-192.
- Michelas, A. & D'Imperio, M. (2010b). Durational cues and prosodic parsing in French: evidence for the intermediate phrase. In *Proceedings of the Vth Speech Prosody Conference*, Chicago, Unites-States.
- Michelas, A., D'Imperio, M. (2010a). Accentual Phrase boundaries and lexical access in French. In *Proceedings of the Vth Speech Prosody Conference*, Chicago, Unites-States.
- Millotte, S. (2005). *Le rôle de la prosodie dans le traitement syntaxique*
- Millotte, S., Wales, R., & Christophe, A. (2007). Phrasal prosody disambiguates syntax. *Language and Cognitive Processes*, 22, 898-909.
- Millotte, S., René, A., Wales, R., & Christophe, A. (2008). Phonological phrase boundaries constrain the on-line syntactic analysis of spoken sentences. *Journal of Experimental Psychology: Learning Memory and Cognition*, 34, 874-885.

- Mitchell, D.-C., & Holmes, V.-M. (1985). The role of specific information about the verb in parsing sentences with local structural ambiguity. *Journal of Memory and Language*, 24, 542-559.
- Mitchell ; D.C., & Cuetos, F. (1991). The origins of parsing strategies. *Paper presented at the Current issues on natural language processing*, University of Texas, Austin. TX.
- Monnin, P., & Grosjean, F. (1993). Les structures de performance en français: caractérisation et prédiction. *L'Année Psychologique*, 93, 9-30.
- Munhall, K. & Löfqvist, A. (1992) Gestural aggregation in speech: Laryngeal gestures, *Journal of Phonetics*, 20, 111-126.
- Nakatani, L. H., & Schaffer, J. A. (1978). Hearing words without words: Prosodic cues for word perception. *Journal of the Acoustical Society of America*, 63(1), 234- 245.
- Nespor, M., & Vogel, I. (1986). *Prosodic Phonology*. Dordrecht: Foris.
- Norris, D. (1994). SHORTLIST: a connectionist model of continuous speech recognition. *Cognition*, 52, 189-234.
- Oller, K.D. (1973). The effect of position in utterance on speech segment duration in English. *Journal of Acoustical Society of America*, 54, 1235-1247.
- Ostendorf, M. & Veilleux, N. (1994). A hierarchical stochastic model for automatic prediction of prosodic boundary location. *Computational Linguistics*, 20, 1-27.
- Pasdeloup, V. (1990). *Modèle de règles rythmiques du français appliqué à la synthèse de la parole*, Thèse de Doctorat, Université de Provence.
- Pasdeloup, V. (1992). A prosodic model for French text-to-speech synthesis : A psycholinguistic approach. In G. Bailly, C. Benoît & T.R. Sawallis (Eds.), *Talking Machines: Theories, Models and Designs*. Elsevier Science Publisher, 335-348.
- Pierrehumbert, J. (1980). *The Phonetics and Phonology of English Intonation*. Ph. D. Dissertation, Massachusetts Institute of Technologies.
- Pierrehumbert, J. & Hirschberg J. (1990). The Meaning of Intonation in the Interpretation of Discourse. In *Intentions in Communication*, Cohen, P., Morgan, J., & Pollack, M. Cambridge, MA: MIT Press, 271-311.
- Pierrehumbert, J., & Beckman, M. (1988) *Japanese Tone Structure*. Cambridge, MA: MIT Press.
- Pierrehumbert, J., & Talkin, D. (1992). Lenition of /h/ and glottal stop. In G. Docherty & D.R. Ladd (Eds.), *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*. Cambridge University Press, 90-116.
- Pierrehumbert, J. (2000). The phonetic grounding of phonology. *Les Cahiers de l'ICP, Bulletin de la Communication Parlée*, 5, 7-23.
- Pike, K. (1945). *The Intonation of American English*. Ann Arbor : University of Michigan Press.
- Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in s and s-plus*. New York: Springer-Verlag.
- Post, B. (2000). *Tonal and phrasal structures in French intonation*. The Hague: Holland Academic Graphics.
- Prieto, P. (à paraître). The Intonational Phonology of Catalan. In Jun, S.-A. (Eds.), *Prosodic Typology*, 2, Oxford University Press: Oxford.
- Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. *Journal of Acoustical Society of America*, 90, 2956-2970.

- Prince, A., & Smolensky, P. (1993). *Optimality Theory : Constraint Interaction in Generative Grammar*. Ms. Rutgers University and Johns Hopkins University : Zanichelli.
- Quené, H., & Van den Bergh, H. (2004). On Multi-Level Modeling of data from repeated measures designs: A tutorial, *Speech Communication*, 43 (1-2), 103-121.
- R Development Core Team. 2008. R : A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantic during sentence processing : Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior*, 22, 358-374.
- Rialland, A., & Aborobongui, M.E. (2010). The intonational system of Embosi Bantu (C25), a 2 tone language without downstep. In *Proceedings of the 4th Conference on Tone and Intonation in Europe*, Stockholm, Suède.
- Rietveld, A.C.M. (1980). French Word Boundaries. *Language and Speech* 23(3), 289- 296.
- Rolland, G., & Løevenbruck, H. (2002). Characteristics of the accentual phrase in French: An acoustic, articulatory and perceptual study. In B. Bel, & I. Marlien (Eds.), *Proceedings of the 1st Speech Prosody Conference*, Aix-en-Provence, France, 611-614.
- Rossi, M. (1997). Is syntactic structure prosodically recoverable ? In *Proceedings of Workshop Intonation: Theory, Models and Applications*, Athens, Greece, 1-8.
- Rossi, M. (1999). *L'intonation, le système du français : description et modélisation*. Paris : Ophrys.
- Saffran, J.-R., Newport, E.-L., & Aslin, R.-N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 35, 606-621.
- Sandalo, F. & Truckenbrodt, H. (2002). Some notes on phonological phrasing in Brazilian Portuguese. *MIT Working Papers in Linguistics*, 42, 285-310.
- Schepman, A., & Rodway, P. (2000). Prosody and parsing in coordination structures. *The Quarterly Journal of Experimental Psychology*, 53(A), 377-396.
- Selkirk, E.-O. (1978). On prosodic structure and its relation to syntactic structure. In T. Fretheim (Eds.), *Nordic Prosody II*, TAPIR, Trondheim, 111-140.
- Selkirk, E.-O. (1984). *Phonology and Syntax : the Relation between Sound and Structure*. Cambridge Mass : MIT Press.
- Selkirk, E.-O. (1986). On derived domains in sentence phonology. *Phonology*, 3, 371-405.
- Selkirk, E.O. (1995). The prosodic structure of function words. In Beckman, J. Walsh Dickey, L., & Urbanczyk, S. (Eds.), *Papers in Optimality Theory*, 439-470. GLSA Publications.
- Selkirk, E. (2000). The interaction of Constraints on prosodic phrasing. In G. Bruce (Eds.), *Prosody: Theory and experiment. Studies*. 231- 261. Dordrecht: Kluwer.
- Shapiro, L.-P., Nagel, H.-N., & Levine, B.-E. (1993). Preferences for a verb's complements and their use in sentence processing. *Journal of Memory and Language*, 32, 96-114.

- Shattuck-Hufnagel, S., & Turk, A. (1996) A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, 25(2), 193-247.
- Spinelli, E., Welbi, P., Schaegis, A.-L. (2007). Fine-grained access to targets and competitors in phonemically identical spoken sequences: the case of French elision. *Language and Cognitive Processes*, 22, 828-859.
- Spinelli, E., Nicolas, G., Meunier, F., & Welby, P. (2010). An Intonational Cue to Word Segmentation in Phonemically Identical Sequences. *Attention, Perception and Psychophysics*, 72(3).
- Steedman, M. (2003). Information-structural semantics for English intonation. (available at <http://www.cogsci.ed.ac.uk/~steedman/>).
- Swerts, M. (1997). Prosodic features at discourse boundaries of different strength, *Journal of the Acoustical Society of America*, 101(1), 514-521.
- t' Hart, J., Collier, R., & Cohen, A. (1990). *A Perceptual Study of Intonation : An Experimental-Phonetic Approach*. Cambridge : Cambridge University Press.
- Tabain, M. (2003), Effects of prosodic boundary on /aC/ sequences: articulatory results, *Journal of the Acoustical Society of America*, 113, 2834-2849.
- Tabain, M., & P. Perrier (2005). Articulation and acoustics of /i/ in pre-boundary position in French. *Journal of Phonetics*, 33, 77-100.
- Tabor, W., Juliano, C., & Tanenhaus, M.-K. (1997). Parsing in a dynamical system : an attractor-based account of the interaction of lexical and structural constraints in sentence processing. *Journal and Cognitive Processes*, 12(2/3), 211-271.
- Tanenhaus, M.-K, Spivey-Knowlton, M.-J., Ebhard, K.M., & Sedivy, J.E. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, 1632-1634.
- Truckenbrodt, H. (2002). Embedded register levels, *Phonology* 19, pp.77-120.
- Truckenbrodt. (2004) Final lowering in non-final position. *Journal of Phonetics* 32: 313-348.
- Truckenbrodt, H. & C. Féry (2005) Sisterhood and tonal scaling. Horne M., Van Oostendorp, M. (Eds.), Boundaries in intonational phonology, *Studia Linguistica* 59.2/3, 223-243.
- Turk, A. and White, L. (1999) Structural effects on accentual lengthening in English. *Journal of Phonetics*, 27, 171-206.
- Turk, A., & Shattuck-Hufnagel, S. (2007) Phrase-final lengthening in American English. *Journal of Phonetics*, 35(4), 445-472.
- Vaissière, J. (1983), Language-Independent Prosodic Features. In Cutler A., & Ladd, D.-R (Eds.), *Prosody: Models and Measurements*. Berlin : Springer-Verlag, 53-66.
- Vaissière, J. (1997). Langues, prosodies et syntaxe. *Traitement Automatique des Langues*, 38(1), 53-82.
- Vaissière, J. (1989). *Contribution à l'analyse des phénomènes de parole continue lue*. Habilitation à diriger des recherches, Université de Strasbourg (unpublished).
- Vaissière, J., & Michaud, A. (2006). Prosodic constituents in French : A data-driven approach. In Prosody and syntax : Cross-linguistic perspectives. Amsterdam : J. Benjamins, 47-63.
- van den Berg, R., C. Gussenhoven, & T. Rietveld (1992). Downstep in Dutch: Implications for a model. Docherty, G.J., & Ladd, D.R. (Eds.), *Papers in laboratory phonology II: Gesture, segment, prosody*, 335-367. Cambridge: Cambridge University Press.

- Watson, D., & Gibson, E. (2004). The relationship between intonational phrasing and syntactic structure in language production. *Language and Cognitive Processes*, 19, 713-755.
- Welby, P. (2003). *The slaying of Lady Mondegreen, being a study of French tonal association and alignment and their role in speech segmentation*. PhD. Dissertation, The Ohio State University.
- Welby P., & Løevenbruck H. (2006). Anchored down in Anchorage: Syllable structure and segmental anchoring in French. *Italian Journal of Linguistics*, 18, 74-124.
- Welby, P. (2006). French intonational structure: Evidence from tonal alignment. *Journal of Phonetics*, 34(3): 343–371.
- Wenk, B.-J., & Wioland, F. (1982). Is French really syllable-timed?. *Journal of Phonetics* 10, 193-216.
- White, L. (2002). *English speech timing: A domain and locus approach*. PhD. Dissertation. University of Edinburgh.
- Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, F., & Price, P.-J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America*, 91, 1707-1717.

ANNEXES

1. ANNEXES A : Corpus utilisé pour l'étude 1 (Chapitre 2)

1.1. Stimuli

Les syllabes cibles, en lettres capitales, pouvaient apparaître dans quatre contextes différents en fonction du type de frontière prosodique associé à la syllabe: en frontière droite de syntagme intonatif (IP-final), en frontière droite de syntagme accentuel associé avec une rupture entre un syntagme nominal et un syntagme verbal (AP-XP), en frontière droite de syntagme accentuel contenu dans un syntagme nominal complexe (AP-final) et en frontière droite de syllabe contenue à l'intérieur d'un syntagme accentuel (AP-interne).

Phrases pour les syllabes cibles contenues à l'intérieur d'un syntagme accentuel.

<u>Item</u>	<u>Phrase</u>
-------------	---------------

- | | |
|-----|--|
| 1. | Le coRRIdor de cette maison nécessite un grand ménage. |
| 2. | L'après-MIdi de ce deux mars devenait interminable. |
| 3. | La coRRIda de Barcelone servira de coup d'essai. |
| 4. | Tout l'aMIdon de pomme-de-terre a été assimilé. |
| 5. | Les féLIIdés du Nigeria représentent un grand danger. |
| 6. | Le calVAdos de Normandie restera un alcool fort. |
| 7. | Ce trouBAdour de Vénétie restera très populaire. |
| 8. | Les greNAdiers de Marrakech ne poussent pas bien vers chez nous. |
| 9. | Les droMAdaires de Tunisie s'habituent à la chaleur. |
| 10. | Ton baLAdeur de grande qualité deviendra à la mode. |
| 11. | Le coRRIdor de cette maison nécessite un grand ménage. |
| 12. | L'après-MIdi de ce deux mars devenait interminable. |
| 13. | La coRRIda de Barcelone servira de coup d'essai. |
| 14. | Tout l'aMIdon de pomme-de-terre a été assimilé. |
| 15. | Les féLIIdés du Nigeria représentent un grand danger. |
| 16. | Le calVAdos de Normandie restera un alcool fort. |
| 17. | Ce trouBAdour de Vénétie restera très populaire. |
| 18. | Les greNAdiers de Marrakech ne poussent pas bien vers chez nous. |
| 19. | Les droMAdaires de Tunisie s'habituent à la chaleur. |
| 20. | Ton baLAdeur de grande qualité deviendra à la mode. |

Phrases pour les syllabes cibles associées à la frontière droite d'un syntagme accentuel contenu dans un syntagme nominal complexe (frontière d'AP).

Item Phrase

1. Le maRI d'Amanda réclamait sa bicyclette.
2. Les aMIS de Pauline adoraient ce restaurant.
3. Le maRI de Cécile deviendra le responsable.
4. La maMIE de Manon attendait de la visite.
5. Les coLIS de Camille n'arrivaient jamais à l'heure.
6. La diVA de la pop restera Witney Houston.
7. La samBA d'Aurélie nécessite beaucoup de place.
8. Le sauNA de Paolo deviendra incontournable.
9. Le coMA de Julien inquiétait tous les docteurs.
10. La viLLA de Simon ressemblait à un château.
11. Le maRI d'Amanda réclamait sa bicyclette.
12. Les aMIS de Pauline adoraient ce restaurant.
13. Le maRI de Cécile deviendra le responsable.
14. La maMIE de Manon attendait de la visite.
15. Les coLIS de Camille n'arrivaient jamais à l'heure.
16. La diVA de la pop restera Witney Houston.
17. La samBA d'Aurélie nécessite beaucoup de place.
18. Le sauNA de Paolo deviendra incontournable.
19. Le coMA de Julien inquiétait tous les docteurs.
20. La viLLA de Simon ressemblait à un château.

Phrases pour les syllabes cibles associées à la frontière droite d'un syntagme accentuel qui est aligné avec la rupture entre le syntagme nominal et le syntagme verbal (frontière d'AP/XP).

Item Phrase

1. GrégoRY demandait la musicienne.
2. Tes aMIS demandaient la présidente.
3. Jean-MaRIE deviendra le responsable.
4. Tante EMMY demandait le directeur.
5. AmÉLIE demandait la musicienne.
6. Cette diVA deviendra inoubliable.
7. Ta samBA deviendra inoubliable.
8. Ce sauNA deviendra incontournable.
9. Saint ThoMAS représente la clairvoyance.
10. NicoLAS devenait insupportable.
11. GrégoRY demandait la musicienne.
12. Tes aMIS demandaient la présidente.
13. Jean-MaRIE deviendra le responsable.
14. Tante EMMY demandait le directeur.
15. AmÉLIE demandait la musicienne.
16. Cette diVA deviendra inoubliable.
17. Ta samBA deviendra inoubliable.
18. Ce sauNA deviendra incontournable.
19. Saint ThoMAS représente la clairvoyance.
20. NicoLAS devenait insupportable.

Phrases pour les syllabes cibles associées à la frontière droite d'un syntagme intonatif (frontière d'IP).

Item	Phrase
1.	Ton maRI, d'après ce qu'on m'a dit, était marin.
2.	Tes aMIS, d'après ce qu'on m'a dit, t'adorent vraiment.
3.	Son maRI, d'après ce qu'on m'a dit, parle le russe.
4.	Tante EMMY, d'après ce qu'on m'a dit, dort énormément.
5.	AmÉLIE, d'après ce qu'on m'a dit, parlait roumain.
6.	Cette diVA, d'après ce qu'on m'a dit, est la meilleure.
7.	La samBA, d'après ce qu'on m'a dit, c'est assez dur.
8.	Le sauNA, d'après ce qu'on m'a dit, n'est pas très loin.
9.	Saint ThoMAS, d'après ce qu'on m'a dit, est un apôtre.
10.	Sa viLLA, d'après ce qu'on m'a dit, n'est pas très loin.
11.	Ton maRI, d'après ce qu'on m'a dit, était marin.
12.	Tes aMIS, d'après ce qu'on m'a dit, t'adorent vraiment.
13.	Son maRI, d'après ce qu'on m'a dit, parle le russe.
14.	Tante EmMY, d'après ce qu'on m'a dit, dort énormément.
15.	AmÉLIE, d'après ce qu'on m'a dit, parlait roumain.
16.	Cette diVA, d'après ce qu'on m'a dit, est la meilleure.
17.	La samBA, d'après ce qu'on m'a dit, c'est assez dur.
18.	Le sauNA, d'après ce qu'on m'a dit, n'est pas très loin.
19.	Saint ThoMAS, d'après ce qu'on m'a dit, est un apôtre.
20.	Sa viLLA, d'après ce qu'on m'a dit, n'est pas très loin.

1.2. Consignes

1.2.1. Phase de lecture à vitesse d'élocution normale

L'expérimentation va se dérouler en deux parties.

Au cours de ces deux parties vous allez devoir lire de phrases qui vont apparaître une par une sur l'écran de l'ordinateur.

Une fois que vous avez lu la phrase vous devez appuyer sur la barre espace pour faire apparaître la phrase suivante.

Si vous vous trompez ou si vous bafouillez vous devez répéter la phrase en entier avant d'appuyer sur la barre espace.

Dans la première partie de cette expérience vous allez devoir lire les phrases le plus naturellement possible comme si vous parlez avec un ami.

À la fin de la première partie un écran blanc apparaîtra afin que vous puissiez vous reposer quelques minutes et l'expérimentateur viendra vous expliquer la suite de l'expérience.

Vous allez démarrer l'expérience avec une phase d'entraînement, ceci afin de vous familiariser avec la tâche.

Êtes-vous prêt ?

1.2.2. Phase de lecture à vitesse d'élocution rapide

Vous allez maintenant démarrer la deuxième partie de l'expérience.

Tout comme lors de la première partie vous devez lire les phrases qui vont apparaître une par une à l'écran.

Cette fois-ci vous devez lire les phrases le plus rapidement possible. Parlez aussi vite que possible, comme si vous étiez extrêmement pressé.

Attention, les phrases doivent rester compréhensibles.

Veillez à bien appuyer sur la touche espace une fois que vous avez terminé de lire la phrase. Veillez également à reprendre la phrase en entier si vous bafouillez ou vous vous trompez. Vous pouvez répéter la phrase autant de fois que nécessaire.

Êtes-vous prêt ?

2. ANNEXES B : Corpus utilisé pour l'étude 2 (Chapitre 3)

2.1. Stimuli

Les syllabes cibles, en lettres capitales, sont toujours la syllabe finale du deuxième syntagme accentuel de l'énoncé et peuvent apparaître dans deux contextes différents : soit elles sont associées à une frontière de syntagme intermédiaire (phrase dont le syntagme nominal sujet contient deux syntagmes accentuels), soit elles étaient associées à une frontière de syntagme accentuel (phrases dont le syntagme nominal sujet contient trois syntagmes accentuels).

<u>Item</u>	<u>Phrase</u>
1a.	La mamie de RÉMY demandait l'institutrice.
1b.	La mamie des aMIS de Rémy demandait l'institutrice.
2a.	Le mari de CarRY deviendra un grand docteur.
2b.	Les carries du maRI de Carry demandaient des soins urgents.
3a.	Le sauna d'HÉLÉNA deviendra le plus connu.
3b.	La nana du sauNA d'Hélène devenait vraiment méchante.
4a.	Le trois-mâts de ThoMAS devance tous les autres bateaux
4b.	Le schéma du trois-MÂTs de Thomas devenait vraiment brouillon.

2.2. Consignes

2.2.1. Phase de lecture à vitesse d'élocution normale

L'expérimentation va se dérouler en deux parties.

Au cours de ces deux parties vous allez devoir lire de phrases qui vont apparaître une par une sur l'écran de l'ordinateur.

Une fois que vous avez lu la phrase vous devez appuyer sur la barre espace pour faire apparaître la phrase suivante.

Si vous vous trompez ou si vous bafouillez vous devez répéter la phrase en entier avant d'appuyer sur la barre espace.

Dans la première partie de cette expérience vous allez devoir lire les phrases le plus naturellement possible comme si vous parlez avec un ami.

À la fin de la première partie un écran blanc apparaîtra afin que vous puissiez vous reposer quelques minutes et l'expérimentateur viendra vous expliquer la suite de l'expérience.

Vous allez démarrer l'expérience avec une phase d'entraînement, ceci afin de Vous familiariser avec la tâche.

Êtes-vous prêt ?

2.2.2. Phase de lecture à vitesse d'élocution rapide

Vous allez maintenant démarrer la deuxième partie de l'expérience.

Tout comme lors de la première partie vous devez lire les phrases qui vont apparaître une par une à l'écran.

Cette fois-ci vous devez lire les phrases le plus rapidement possible. Parlez aussi vite que possible, comme si vous étiez extrêmement pressé.

Attention, les phrases doivent rester compréhensibles.

Veillez à bien appuyer sur la touche espace une fois que vous avez terminé de lire la phrase. Veillez également à reprendre la phrase en entier si vous bafouillez ou vous vous trompez. Vous pouvez répéter la phrase autant de fois que nécessaire.

Êtes-vous prêt ?

3. ANNEXE C : Expérience pilote 1, Frontière d'ip et frontière d'IP en français

Le but de cette étude est d'étudier les indices acoustiques présents à la frontière d'ip et à la frontière d'IP en français.

Nous avons demandé à deux locutrices de langue maternelle française de lire des phrases dont la syllabe cible était associée soit à une frontière d'ip, soit à une frontière d'IP. Prêtons attention aux deux phrases suivantes :

1a. La mamie des amis de RéMY attendait l'institutrice.

1b. La mamie des amis de RéMY ? Elle viendra demain.

Dans la phrase 1a nous nous attendons à ce que la syllabe cible soit associée à une frontière d'ip. En effet, la syllabe cible est alignée avec une rupture syntaxique majeure et le syntagme est suffisamment long (probablement réalisé en tant que 3 syntagmes accentuels) pour permettre l'émergence d'un accent de syntagme H- qui serait réalisé sur la syllabe cible. Dans la phrase 1b, nous nous attendons à ce que la syllabe cible soit associée à une frontière de syntagme intonatif. En effet, ce type de rupture est fortement corrélée avec une frontière prosodique majeure (Selkirk, 1995 ; Delais-Roussarie & Post 2008 ; *inter alia*). Nous avons mesuré les indices de durée et les indices mélodiques des syllabes cibles qui pouvaient être associées soit à une frontière d'ip, soit à une frontière d'IP. Notre hypothèse était que les valeurs de durée et de *fo* observées sur les voyelles cibles seraient plus importantes dans les phrases 9b (lorsque la syllabe cible était associée à une frontière d'IP) que dans les phrases 9a (lorsque la syllabe cible était associée à une frontière d'ip).

3.1. Méthodologie

Un corpus de huit phrases lues à quatre reprises, par deux locutrices de langue maternelle française a été construit pour un total de 64 énoncés (8 phrases x 4 répétitions x 2 locutrices = 64 phrases). Nous avons analysé les propriétés acoustiques (durée segmentale et valeurs de *f₀*) des voyelles cibles. La voyelle cible était toujours la voyelle en position finale du troisième syntagme accentuel de l'énoncé et pouvait être associée soit à une frontière d'ip soit à une frontière d'IP. Cette voyelle pouvait être soit /i/, soit /a/ et était toujours le noyau vocalique d'une syllabe ouverte en position finale du troisième syntagme accentuel de l'énoncé.

20 phrases ne présentant aucune structure prosodique ou syntaxique particulière ont été ajoutées aux 8 phrases expérimentales. Le corpus de 28 phrases ainsi construit été présenté 4 fois pour chaque locuteur. L'ordre des phrases était aléatoire et différent pour chaque présentation et pour chaque locuteur avec pour contrainte que les deux membres d'une paire de phrase ne pouvaient pas apparaître à la suite. Les deux locutrices étaient âgées de 22 et 26 ans et étaient originaires du Sud de la France. Les enregistrements ont eu lieu dans la chambre sourde du laboratoire Parole et Langage grâce à une station d'enregistrement ZOOM HD16 en format wav 16bit/44.1 Hz. Les phrases étaient présentées une par une sur un écran d'ordinateur. Elles étaient orthographiées de manière classique en incluant les lettres majuscules et la ponctuation. Une fois la phrase lue, les participants devaient cliquer sur un bouton pour faire apparaître la phrase suivante. Les participants ont eu pour consigne de lire les phrases le plus naturellement possible. La totalité des phrases utilisées ainsi que la consigne exacte présentée à l'écran aux deux locutrices sont présentées en annexe. La procédure expérimentale a été contrôlée grâce au logiciel DMDX (Forster & Forster, 2000).

Les fichiers sons ont été segmentés manuellement et chaque fichier son a été sauvé séparément. Tous les fichiers ont été segmentés en phonèmes à l'aide de l'aligneur Easy align (Goldman, 2007). La segmentation des voyelles cibles a été vérifiée manuellement grâce à l'inspection des ondes sonores et des spectrogrammes. Les courbes de *f₀* ont été générées grâce à l'utilisation du logiciel PRAAT (Boersma & Weenik). Nous avons automatiquement étiqueté les voyelles cibles qui étaient toujours les voyelles en position finale du troisième syntagme accentuel de l'énoncé. La durée segmentale des voyelles ainsi étiquetées, ainsi que les valeurs maximales de *f₀* qui leur étaient associées ont été extraites automatiquement grâce à l'utilisation de scripts PRAAT. Dans le but de mesurer la hauteur de la voyelle cible par rapport au premier pic mélodique de l'énoncé nous avons également automatiquement extrait les valeurs de *f₀* associées au premier pic mélodique de chaque énoncé (H1). Un exemple du codage ainsi effectué pour les phrases où la voyelle cible était associée à une frontière d'ip et pour les phrases dont la voyelle cible était associée à une frontière d'AP est illustré dans la figure 79.

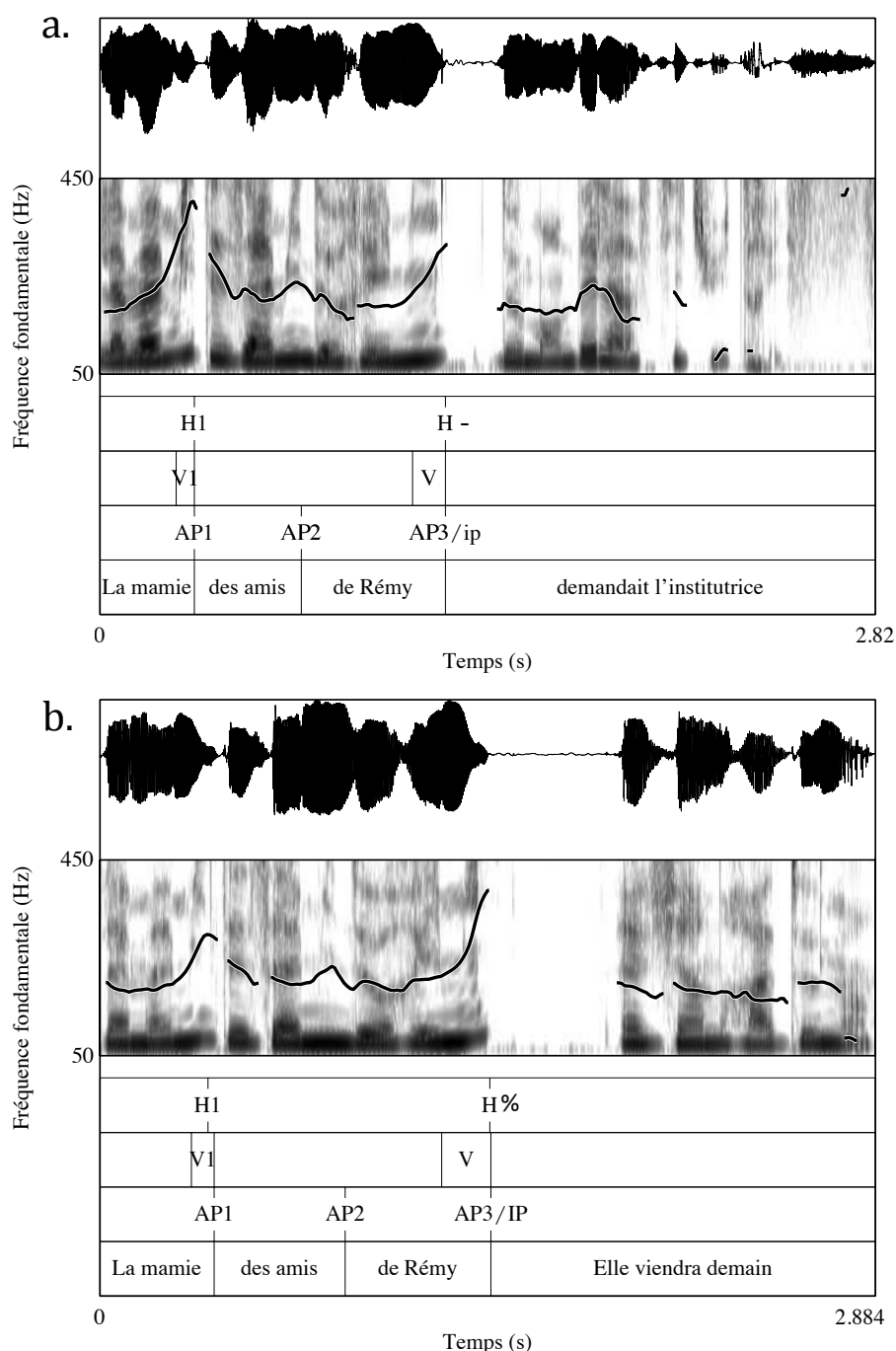


Figure 79. Pilote 1 : Exemple d'annotations réalisées.

Exemple d'annotations pour les phrases « La mamie des amis de Rémy demandait l'institutrice » où la voyelle cible (V) est associée à une frontière d'ip (79a) et de la phrase « La mamie des amis de Rémy ? Elle viendra demain » où la voyelle cible (V) est associée à frontière d'IP (79b). La voyelle V1 représente la voyelle associée au premier pic mélodique de l'énoncé. Dans les deux phrases, le syntagme nominal sujet a été réalisé comme trois syntagmes accentuels.

Pour chaque énoncé, nous avons vérifié que le syntagme nominal a été réalisé comme trois syntagmes accentuels ce qui a été le cas pour la totalité

des 64 phrases expérimentales produites par nos deux locutrices comme cela est illustré dans la figure 79.

3.2. Résultats

La durée des voyelles cibles exprimée en millisecondes en fonction du type de frontière prosodique à laquelle étaient associées les voyelles (frontière d'ip ou frontière d'IP) et du locuteur (AM ou CD) est représentée dans la figure 80.

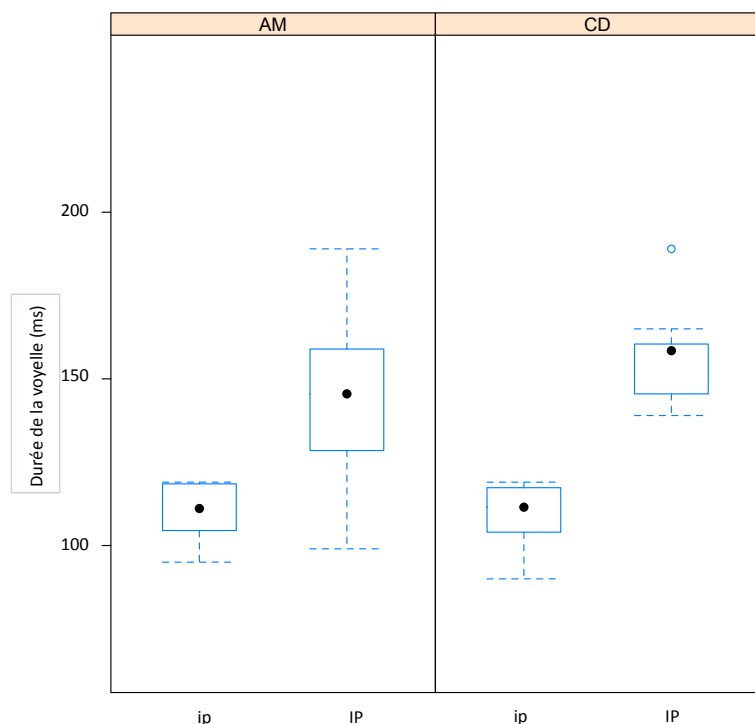


Figure 80. Pilote 1 : Durée des voyelles cibles exprimées en millisecondes. La durée des voyelles est fonction du type de frontières auxquelles elles étaient associées (frontière de syntagme intermédiaire, ip, ou frontière de syntagme intonatif, IP) et du locuteur (AM ou CD).

Les résultats d'un t-test montre que la voyelle était significativement plus longue lorsqu'elle était associée à une frontière d'IP que lorsqu'elle était associée à une frontière d'ip ($t(55.495)=6.1698$, $p < 0.0001$).

Nous avons également mesuré la hauteur en terme de rapport par le premier pic mélodique de l'énoncé (H1). Les rapports des voyelles cibles par H1 en fonction du type de frontière prosodique auxquelles étaient associées les voyelles (frontière d'ip ou frontière d'IP) et du locuteur (AM ou CD) est représenté dans la figure 81.

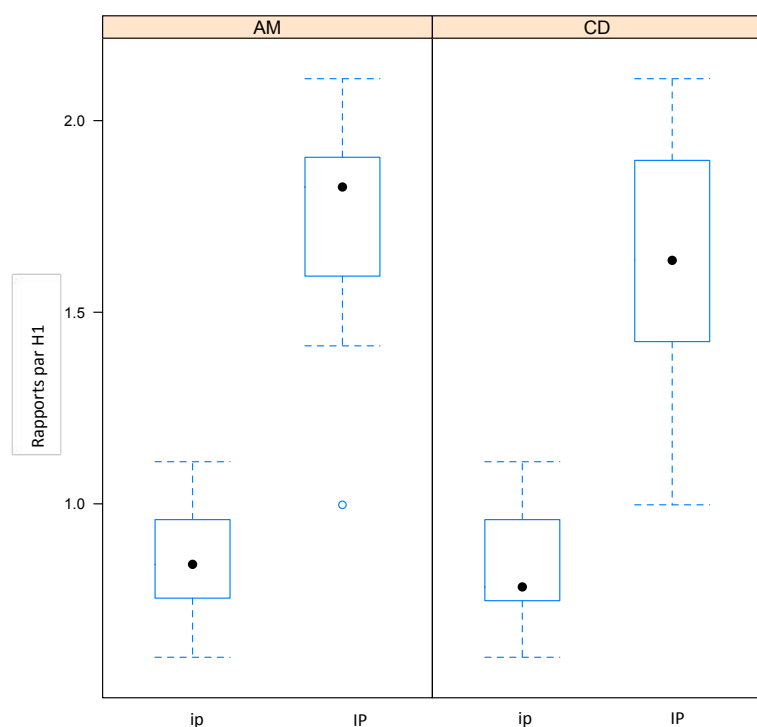


Figure 81. Pilote 1 : Rapports des valeurs de f_0 des voyelles cibles par les valeurs du premier pic mélodique de l'énoncé (H1).

Les rapports sont exprimés en fonction du type de frontière auxquelles étaient associées les voyelles cibles (frontière de syntagme intermédiaire, ip, ou frontière de syntagme intonatif, IP) et du locuteur (AM ou CD).

Un t-test réalisé sur les rapports des voyelles cibles par H1 montre que les voyelles étaient significativement plus hautes lorsque les voyelles cibles étaient associées à une frontière d'IP que lorsqu'elles étaient associées à une frontière d'AP ($t(43.352)=14.7862$, $p < 0.0001$).

3.3. Conclusions

Les analyses menées sur ce corpus de 64 énoncés nous ont permis de montrer que les indices acoustiques présents à la frontière d'ip sont significativement différents de ceux présents à la frontière d'IP. Conformément à notre hypothèse, nous avons observé que la durée de la voyelle était significativement plus importante lorsqu'elle était associée à une frontière d'IP que lorsqu'elle est associée à une frontière d'ip. De plus la hauteur de la voyelle était également plus importante lorsque cette voyelle était associée à une frontière de syntagme intonatif plutôt qu'à une frontière de syntagme intermédiaire. Notons que la hauteur des voyelles qui étaient associées à une frontière d'IP, dépassait la ligne de référence du registre marquée par le premier pic mélodique de l'ip tandis que la hauteur des voyelles qui étaient associées à une frontière de syntagme intermédiaire était similaire à celle observée sur le premier pic mélodique de l'énoncé. En effet,

nous avons obtenus des rapports par H1 supérieurs à 1 lorsque la voyelle était associée à une frontière d'IP et des rapports par H1 similaires à 1 lorsque la voyelle était associée à une frontière d'ip. Cette différence de hauteur entre les voyelles associées aux frontières d'ip et d'IP est illustrée dans la figure 82.

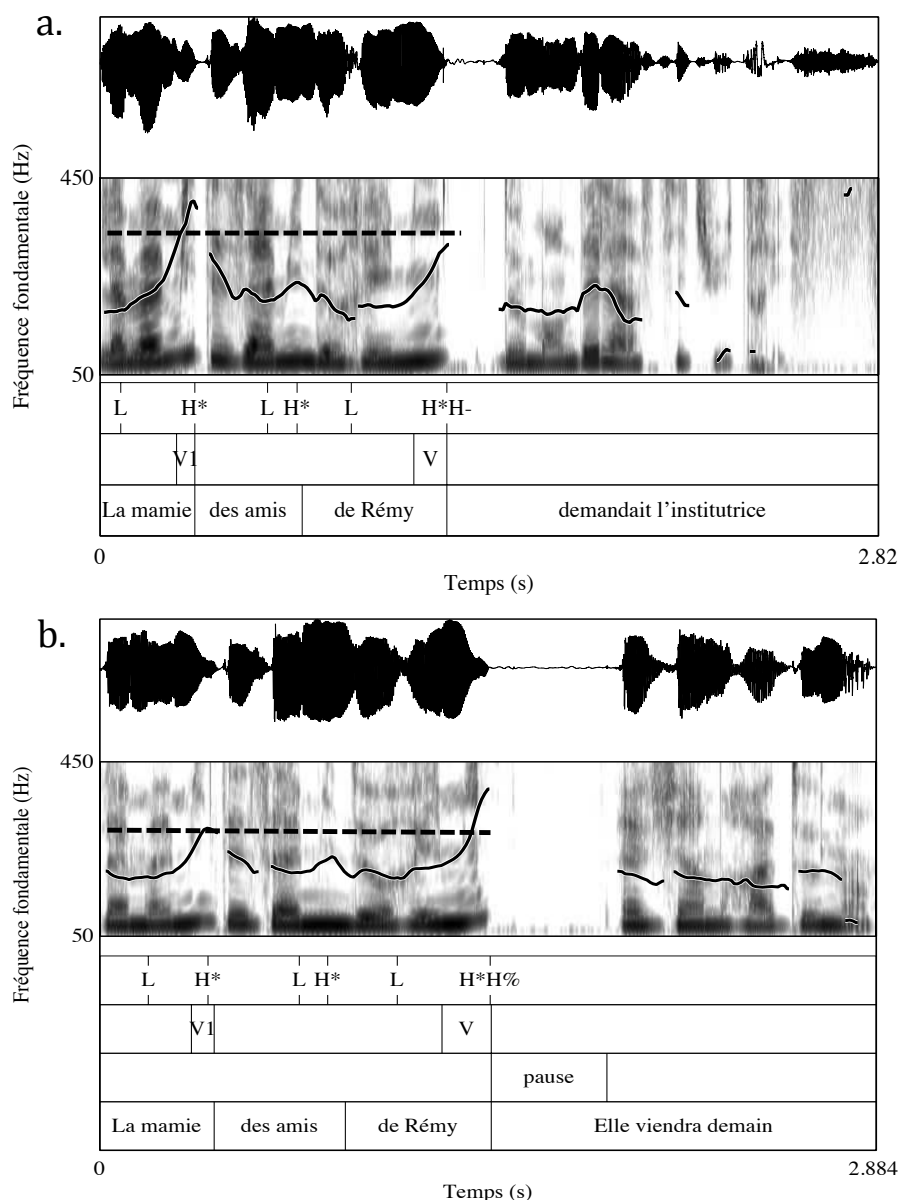


Figure 82. Pilote 2 : Exemples de phrases du corpus.

Contours mélodiques obtenus pour deux phrases dont la voyelle cible (V) pouvait être associée soit à une frontière d'ip (82a) marquée par un accent de syntagme H-, soit par une frontière d'IP (82b) marquée par un ton de frontière H%.

Dans cette figure, nous pouvons observer que les valeurs de *fo* associées à la frontière d'ip (H1) sont similaires à celles observées sur le premier pic mélodique de l'énoncé (H1, cf. 82a) tandis que les valeurs de *fo* associées à une frontière d'IP dépassent celles observées sur le premier pic mélodique de l'énoncé (cf. 82b). De plus, dans cette figure, nous pouvons

noter la présence d'une pause silencieuse après la frontière d'IP mais pas après la frontière d'ip. Cet exemple est représentatif de ce que nous avons observé dans nos données où une pause silencieuse était très souvent observée après la frontière d'IP (dans 97% des cas) alors qu'elle était moins fréquemment observée après une frontière d'ip (19% des cas). Conformément à ce que nous avons affirmé dans l'étude 2 précédemment présentée dans ce chapitre, nous proposons que l'accent de syntagme (H-) présent sur le dernier accent mélodique de l'ip soit responsable de la réinitialisation du registre à la frontière du syntagme intermédiaire. La présence d'un ton de frontière H% est quant à lui responsable de l'élévation des valeurs de *fo* qui dépassent la ligne de référence du registre marquée par le premier accent mélodique de l'énoncé. Des analyses supplémentaires seront nécessaires dans le but d'observer ce qui se passe sur le premier pic mélodique situé après la frontière d'IP. Deux solutions sont en effet envisageables : soit le registre est complètement réinitialisé après la frontière d'ip en français (ce qui serait un indice supplémentaire permettant de différencier frontière d'ip et frontière d'AP), soit tout comme après une frontière d'ip, le registre est partiellement réinitialisé après la frontière d'IP. De nouvelles expérimentations menées en production permettront d'apporter une réponse à cette question.

4. ANNEXE D : Expérience pilote 2, Le syntagme intermédiaire en position Objet à l'intérieur des énoncés.

Dans les études que nous avons menées jusqu'ici, nous avons toujours étudié la frontière de syntagme intermédiaire en position Sujet dans la phrase. Afin de confirmer le fait que le syntagme intermédiaire puisse être également observé en position Objet en français, nous avons réalisé un corpus constitué de 64 phrases. Ce corpus est construit sur le même principe que l'étude 2 présenté dans le chapitre 3 de cette thèse. Nous avons utilisé des paires de phrases dans laquelle nous avons fait varier la longueur du syntagme nominal Objet de la phrase, en terme de nombre de mots, de façon à obtenir un syntagme nominal Objet composé soit de 2 APs (phrase 2a), soit de 3 APs (phrase 2b).

2a. Marie a rencontré la mamie de RÉMY dans la journée.			
Structure syntaxique :	[		]SN
Structure prosodique :	[	]AP[	]AP/ip
2b. Marie a rencontré la mamie des amis de Rémy dans la journée.			
Structure syntaxique :	[		]SN
Structure prosodique :	[	]AP[	]AP[]AP/ip

La seule différence avec l'étude 2 présentée dans le chapitre 3 est que le syntagme nominal auquel nous nous intéressons est situé en position Objet et non en position Sujet dans la phrase. Nous supposons qu'une frontière de syntagme intermédiaire va émerger à la rupture entre le syntagme nominal Objet et le syntagme prépositionnel qui occupe la fonction de complément circonstanciel de temps dans la phrase. Nous avons mesuré les indices de durée et les valeurs de *fo* associées à la voyelle cible qui était associée soit à une frontière d'AP (2a), soit à une frontière d'ip (2b). La voyelle cible (en lettres capitales dans les phrases 2a et 2b était toujours la voyelle finale du deuxième syntagme accentuel du syntagme nominal Objet. Notre hypothèse était que, conformément à ce que nous avons observé dans l'étude 2, la voyelle cible serait plus longue dans les phrases de la condition 2 APs que dans les phrases de la condition 3 APS. De manière similaire nous nous attendions à observer des valeurs de *fo* plus hautes dans la condition 2 APs que dans la condition 3APs en raison de la présence de la frontière de syntagme accentuel dans la phrase.

4.1. Méthodologie

Un corpus de 4 paires de phrases dont la voyelle cible était associée soit à une frontière d'ip (phrases dont le SN Objet contenait 2 APs), soit à une frontière d'AP (phrases dont le SN sujet contenait 3 APs) a été construit. Nous avons demandé à deux locutrices de langue maternelle française de lire ces phrases à quatre reprises pour un total de 64 énoncés (8 phrases x 4 répétitions x 2 locutrices = 64 phrases). Nous avons analysé les propriétés acoustiques (durée segmentale et valeurs de *fo*) des voyelles cibles en position finale du deuxième syntagme accentuel du SN Objet. La voyelle cible pouvait être soit /i/, soit /a/ et était toujours le noyau vocalique d'une syllabe ouverte. Les deux locutrices étaient âgées de 22 et 26 ans et étaient originaires du Sud de la France. Les enregistrements ont eu lieu dans la chambre sourde du laboratoire Parole et Langage grâce à une station d'enregistrement ZOOM HD16 en format wav 16bit/44.1 Hz. La même procédure expérimentale que celle présentée dans l'étude 2 (cf. Chapitre 3) a été utilisée.

Nous avons annoté les voyelles cibles qui étaient toujours les voyelles en position finale du deuxième syntagme accentuel du syntagme nominal Objet. Les voyelles cibles pouvaient être associées soit à une frontière d'ip (phrases dont le SN objet contenant 2 APs), soit à une frontière d'AP (phrases dont le SN sujet contenant 3 APs). Nous avons codé V2 la durée des voyelles cibles et H2 les valeurs maximales de *fo* associées aux syllabes cibles. Dans le but de mesurer la hauteur de la voyelle cible par rapport au premier pic mélodique du SN Objet, nous avons annoté H1 les valeurs de *fo* associées au premier pic mélodique de chaque énoncé (H1).

Les valeurs de durée et de *fo* ont été automatiquement extraites grâce à l'utilisation de scripts PRAAT.

Conformément à ce que nous avons observé pour les syntagmes nominaux en position Sujet dans la phrase (cf. étude 2), nous nous attendions à observer des valeurs de durée et de *fo* supérieures dans la condition 2 APs que dans la condition 3 APs.

4.2. Résultats

La durée des voyelles cibles (pouvant être associées soit à une frontière d'ip dans les SNs Objets composés de 2 APs, soit à une frontière d'AP dans les SNs Objets composés de 3 APS) exprimée en millisecondes en fonction du type de SN (contenant 2 APs ou contenant 3 APs) et du locuteur (AL ou CP) est représentée dans la figure 83.

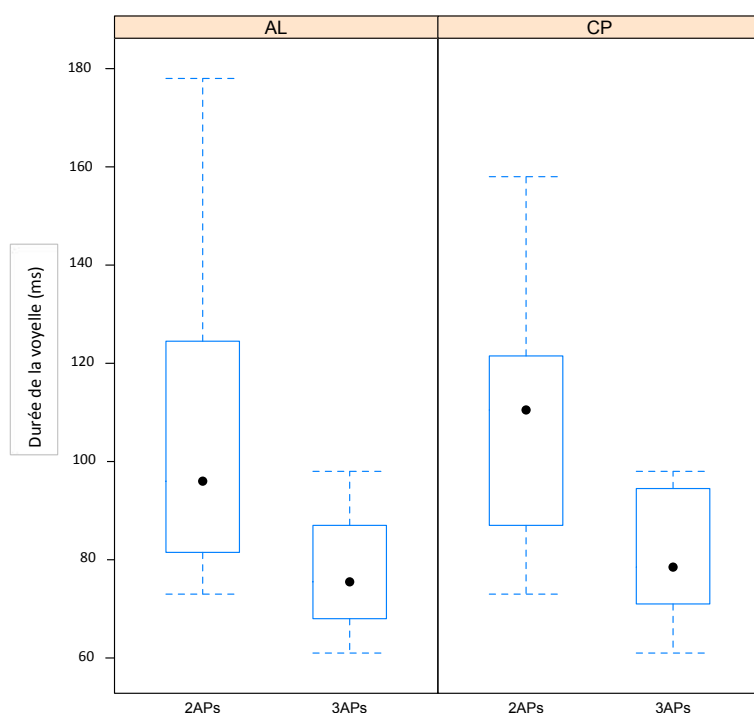


Figure 83. Pilote 2 : Durées des voyelles cibles exprimées en millisecondes. Les durées sont représentées en fonction du type de SN objet (composés de 2 syntagmes accentuels ou de 3 syntagmes accentuels) et du locuteur (AM ou CD).

Les résultats d'un t-test ont montré que la durée des voyelles cibles était significativement supérieure lorsque le syntagme nominal Objet était composé de 2 APs que lorsque le syntagme nominal Objet était composé de 3 APs ($t(51.754)=4.2.37, p<0.0001$).

Nous avons également mesuré la hauteur des voyelles cibles (H2) par rapport à la hauteur du premier pic mélodique du syntagme nominal Objet. Cette mesure en terme de rapports a été effectuée dans le but de prendre en compte les variations de registre observées chez un même locuteur. Les rapports de H2 par H1 en fonction du type de SN objet (composé de 2 APs ou

composé de 3 APs) et du locuteur (AL ou CP) sont représentés dans la figure 84.

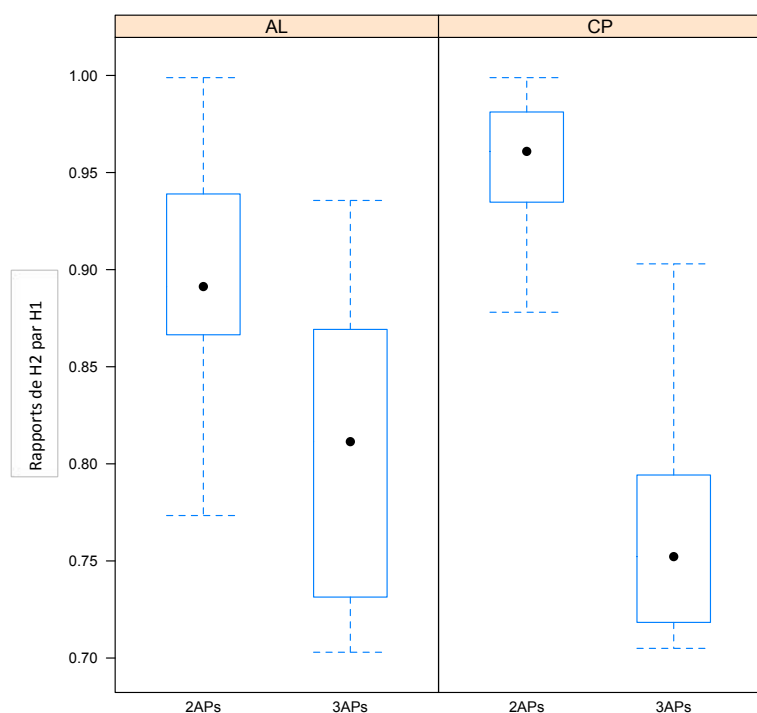


Figure 84. Pilote 2 : Hauteur de la voyelle en position finale du deuxième syntagme accentuel (H2) exprimée par rapport à la hauteur du premier pic de fo du SN objet (H1). Les rapports sont représentés en fonction du type de syntagme nominal (SN composé de 2APs et SN composé de 3 APs) et du locuteur (AL ou CP).

Les résultats d'un t-test conduit sur les rapports des voyelles cibles par H1 ont montré que la voyelle était significativement plus haute dans la condition 2APs que dans la condition 3 APs ($t(60.734)=8.5858$, $p < 0.001$).

4.3. Conclusions

L'analyse d'un corpus de 64 phrases, nous avons montré que le syntagme intermédiaire pouvait être observé aussi bien en position Objet qu'en position Sujet dans les énoncés en français. Nous avons en effet observé que la durée et la hauteur des voyelles cibles étaient supérieures lorsque les voyelles étaient associées à une frontière d'ip (dans la condition SN Objet composé de 2 APs) que lorsqu'elles étaient associées à une frontière d'AP (dans la condition SN Objet composé de 3 APs). Ces résultats confirment la présence d'un accent de syntagme (H-) aligné avec la frontière droite de l'ip et responsable d'un retour à la ligne de référence du registre avant la frontière d'ip. Un exemple de phrase extrait du corpus est illustré dans la figure 85. Dans cette figure, nous pouvons observer un phénomène de réinitialisation totale du registre avant la frontière d'ip. Notons également que la syllabe du

deuxième syntagme accentuel (syllabe cible) est abaissée dans la figure 85b mais pas dans la figure 85a.

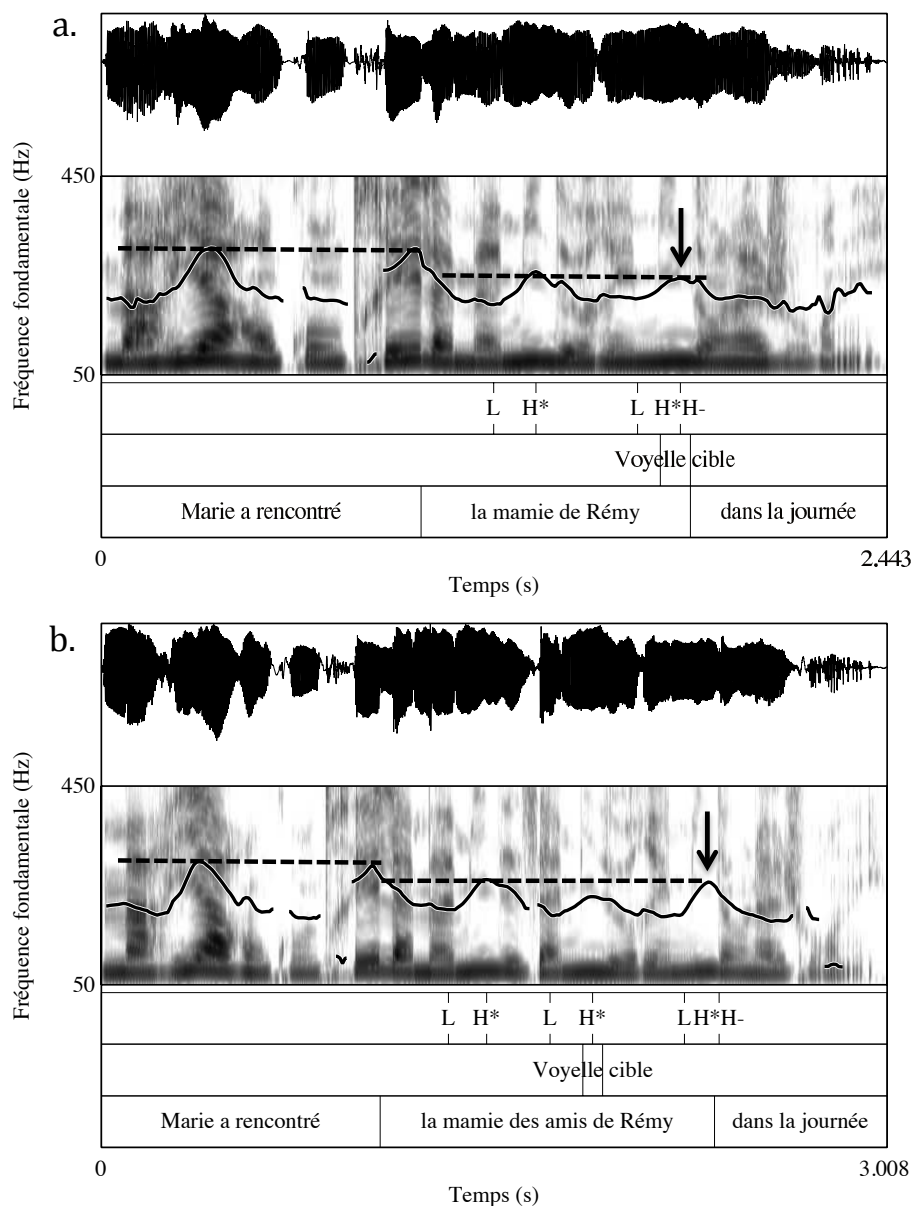


Figure 85. Pilote 2 : Exemples de phrases du corpus.

La voyelle cible pouvait être associée soit à une frontière d'AP (85a), soit à une frontière d'ip (85b). La ligne de référence du registre marquée par le premier accent mélodique de l'ip, est signalée par la ligne en pointillés.

Ces résultats confirment donc l'existence d'un syntagme intermédiaire en français dont la frontière droite serait marquée par un accent de syntagme responsable de la réinitialisation totale du registre à la frontière d'ip.

5. ANNEXES E : Corpus utilisé pour l'étude 3 (Chapitre 4)

5.1. Stimuli

Pour chaque mot cible, écrit en lettre capitales, deux phrases expérimentales ont été construites : une contenant une ambiguïté lexicale temporaire (a) et l'autre non (b).

Paires de phrases avec une frontière de mot prosodique ou de syntagme accentuel (en fonction de la manière dont la phrase est produite) au point d'ambiguïté.

<u>Item</u>	<u>Phrase</u>
1a.	Les villageois étaient très fiers des PINS somptueux du grand bosquet.
1b.	Les villageois étaient fiers des PINS luxuriants du bord de la rivière.
2a.	Pierre m'a longtemps parlé d'un BANC différent qui semblait être un peu plus à l'ombre.
2b.	Pierre m'a longtemps parlé d'un BANC majestueux qui était entouré d'une foule admiratrice.
3a.	Dans le salon il y a un BANC doré qui date du règne de Louis XIV.
3b.	Dans le salon il y a un BANC massif sur lequel les enfants aiment jouer.
4a.	Le livre racontait l'histoire d'un CHAT grincheux qui avait mordu un facteur.
4b.	Le livre racontait l'histoire d'un CHAT drogué qui dormait tout le temps.
5a.	Marie haïssait vraiment les BALS congestionnés de l'ambassade.
5b.	Marie haïssait vraiment les BALS majestueux du palais.
6a.	Dans la quiétude du bois le CERF pensif contemplait la marre.
6b.	Dans la quiétude du bois le CERF docile suivait la petite fille.
7a.	J'ai tout de suite remarqué la TOUX pitoyable de ce patient.
7b.	J'ai tout de suite remarqué la TOUX inquiétante de ce patient.
8a.	Il se souvenait bien de ce PORT trépidant où il avait passé douze années de son enfance.
8b.	Il se souvenait bien de ce PORT primitif qu'il préférait à celui qu'on venait de construire.
9a.	Il veut des renseignements sur le FORT magnifique de cette citée antique.
9b.	Il veut des renseignements sur le FORT légendaire de cette citée antique.
10a.	Mon cousin me parle sans cesse du SOL dallé qu'il doit choisir.
10b.	Mon cousin me parle sans cesse du SOL marbré qu'il doit choisir.
11a.	Nicolas m'a beaucoup parlé du PHARE doré qu'il aime observé.
11b.	Nicolas m'a beaucoup parlé du PHARE lointain qu'il aime observé.
12a.	Je crois que ma sœur t'a parlé des DENTS géniales qu'on doit me poser.
12b.	Je crois que ma sœur t'a parlé des DENTS gâtées qu'on doit m'enlever.

Paires de phrases avec une frontière de syntagme accentuel associée à une rupture syntaxique majeure au point d'ambiguïté.

<u>Item</u>	<u>Phrase</u>
1a.	J'ai remarqué que le très grand FORT magnifiait le génie architectural de cette époque.
1b.	J'ai remarqué que le très grand FORT paraissait totalement imprenable aux envahisseurs.
2a.	De manière surprenante tous les JEUX dissipaient les enfants.
2b.	De manière surprenante tous les JEUX passionnaient les élèves .
3a.	Elle racontait que le mauvais TEMPS bourdonnait à ses oreilles.
3b.	Elle racontait que le mauvais TEMPS malmenait son parapluie.
4a.	Une importante quantité de BOUE giclait au passage de la diligence.
4b.	Une importante quantité de BOUE nivelait le relief du paysage.
5a.	La sœur de Rémy m'a dit que son beau CHAT léchait tous les invités.
5b.	La sœur de Rémy m'a dit que son beau CHAT devenait très paresseux.
6a.	Il est évident que ce vieux MÂT touchera bientôt le fond.
6b.	Il est évident que ce vieux MÂT vibrera à chaque secousse.
7a.	Vincent est très heureux que son PERE milite pour les droits de la femme.
7b.	Vincent est très heureux que son PERE bataille pour gagner son procès.
8a.	La sœur de Paul m'a dit que son CHAT grimpait aux arbres.
8b.	La sœur de Paul m'a dit que son CHAT dressait l'oreille.
9a.	Tu remarqueras bien vite que ce PHARE domine toute la côte napolitaine.
9b.	Tu remarqueras bien vite que ce PHARE surplombe la côte napolitaine.
10a.	Alexandre m'a raconté que ce CHAR bombardait les villages résistants pendant la seconde guerre mondiale.
10b.	Alexandre m'a raconté que de CHAR attaquait les villages résistants lors de la seconde guerre mondiale.
11a.	Il est évident que ces somptueux BALS contribuent à l'image de marque de l'ambassadeur.
11b.	Il est évident que ces somptueux BALS accroissent l'image de marque de l'ambassadeur.
12a.	Elle affirmait que le grand BANC donnait à la pièce un caractère singulier.
12b.	Elle affirmait que le drôle de BANC penchait dangereusement sur la gauche.

5.2. Consignes

Vous allez voir apparaître un mot à l'écran puis entendre une phrase.

Votre tâche est de détecter, le plus rapidement possible, le mot que vous avez vu à l'écran dans la phrase que vous allez entendre.

Dès que vous entendez le mot, vous devez appuyer sur le bouton rouge du boîtier réponse.

Par exemple vous allez voir apparaître à l'écran le mot « PILES » puis entendre la phrase « Les piles microscopiques sont très difficiles à remplacer ».

Vous devez appuyer sur le bouton rouge si vous entendez le mot « piles » dans la phrase.

Vous devez donner votre réponse le plus rapidement possible.

Vous allez démarrer l'expérience avec une phase d'entraînement, ceci afin de vous familiariser avec la tâche.

Êtes-vous prêt ?

6. ANNEXES F : Corpus utilisé pour les études 4 et 5 (Chapitre 5)

6.1. Stimuli

Pour chaque séquence cible, en lettres capitales, deux phrases expérimentales ont été construites. La voyelle finale de chaque séquence cible est associée soit à une frontière de syntagme accentuel (a), soit à une frontière de syntagme intermédiaire (b).

Item Phrase

- 1a. LA NANA DU SAUNA d'Hélène déconseille vivement l'utilisation du savon noir.
- 1b. LA NANA DU SAUNA déconseille vivement l'utilisation du savon noir.
- 2a. LE SCHÉMA DU TROIS-MATS de Nicolas doit être terminé avant demain.
- 2b. LE SCHÉMA DU TROIS MATS doit être terminé avant demain.
- 3a. L'AGENDA DU PATRON d'Hélène demeure introuvable.
- 3b. L'AGENDA DU PATRON demeure introuvable.
- 4a. LE PARVIS DE LA MAIRIE de Paris devrait être entièrement restauré l'an prochain.
- 4b. LE PARVIS DE LA MAIRIE devrait être entièrement restauré l'an prochain.
- 7a. LA CORVÉE DE CLASSEMENT des documents durera environ deux heures.
- 7b. LA CORVÉE DE CLASSEMENT durera environ deux heures.
- 5a. LA MAISON DE LA CHIMIE de Bordeaux dispose de gros moyens.
- 5b. LA MAISON DE LA CHIMIE dispose de gros moyens.
- 6a. LES AMIS DE RÉMY de Dijon décideront des monuments qu'ils souhaitent visiter.
- 6b. LES AMIS DE RÉMY décideront des monuments qu'ils souhaitent visiter.
- 7a. LES PROJETS DU CINÉMA de Bordeaux déclenchent la polémique.
- 7b. LES PROJETS DU CINÉMA déclenchent la polémique.
- 8a. LE NIVEAU DE JAPONAIS de Marina devient meilleur chaque année.
- 8b. LE NIVEAU DE JAPONAIS devient meilleur chaque année.
- 9a. LES CADEAUX DE BIENVENUE du maire deviennent de plus en plus originaux.
- 9b. LES CADEAUX DE BIENVENUE deviennent de plus en plus originaux.
- 10a. LA VILLA DE LOLA de Marseille domine tout le paysage environnant.
- 10b. LA VILLA DE LOLA domine tout le paysage environnant.
- 11a. LE MUSÉE DE L'ARMÉE de Nancy dérange en raison de son exposition sur les nazis.
- 11b. LE MUSÉE DE L'ARMÉE dérange en raison de son exposition sur les nazis.
- 12a. LE PALAIS DES CONGRÈS du Pharo doit accueillir.
- 12b. LE PALAIS DES CONGRÈS doit accueillir 200 visiteurs demain à l'occasion des journées de la mode.

- 13a. LE GÉRANT DU RESTAURANT de Paula discute avec le président depuis plus d'une heure.
- 13b. LE GÉRANT DU RESTAURANT discute avec le président depuis plus d'une heure.
- 14a. LE DÎNER DE SARAH de ce soir débutera vers 19h.
- 14b. LE DÎNER DE SARAH débutera vers 19 heures.
- 15a. LES DÉTENUS DE LA PRISON de Versailles désirent obtenir de meilleures conditions de détention.
- 15b. LES DÉTENUS DE LA PRISON de Versailles désirent obtenir de meilleures conditions de détention.
- 16a. LE PRÉSIDENT DE L'OPÉRA de Paris décide de la programmation.
- 16b. LE PRÉSIDENT DE L'OPÉRA décide de la programmation.
- 17a. LES INTÉRÊTS DU CRÉDIT de la maison diminuent d'année en année.
- 17b. LES INTÉRÊTS DU CRÉDIT diminuent d'année en année.
- 18a. LE PRÉSIDENT DE L'ASSEMBLÉE de Tunisie démissionnera sous peu.
- 18b. LE PRÉSIDENT DE L'ASSEMBLÉE démissionnera sous peu.
- 19a. LE GALA DE FIN D'ANNÉE de Nina dure généralement deux heures.
- 19b. LE GALA DE FIN D'ANNÉE dure généralement deux heures.

6.2. Consignes

Nous avons demandé à une personne de lire des phrases.

Vous allez entendre un extrait de ces phrases puis voir apparaître à l'écran deux suites possibles pour l'extrait que vous venez d'entendre.

Pour chaque extrait, votre tâche est de deviner, le plus rapidement possible, quelle est la suite la plus probable dans les phrases telles qu'elles ont été prononcées par notre lecteur.

Si vous choisissez les mots présentés à droite de l'écran, vous devez appuyer sur la touche « droite », si vous choisissez les mots présentés à gauche de l'écran, vous devez cliquer sur la touche « gauche ».

Par exemple, vous allez entendre l'extrait : « j'ai entendu parler des enfants », puis vous verrez apparaître à l'écran deux suites possibles : suite 1 « de Josie », suite 2 « qui jouaient ». Vous devez choisir celle qui vous semble compléter au mieux la séquence que vous avez entendue.

Attention en aucun cas vous ne devez chercher à terminer les phrases mais bien à deviner quels étaient les mots qui faisaient suite aux extraits que vous entendez.

Vous devez répondre le plus vite possible.

Vous allez démarrer l'expérience avec une phase d'entraînement, ceci afin de vous familiariser avec la tâche.

Êtes-vous prêt ?

RÉSUMÉ

Le travail présenté ici est sous-tendu par deux observations majeures. Premièrement, la plupart des modèles proposés pour le français s'accordent sur l'existence de deux niveaux de structure prosodique: le syntagme accentuel et le syntagme intonatif. Deuxièmement, bien que l'existence d'un niveau additionnel de structure situé entre ces deux niveaux ait été proposé pour le français, les propriétés phonétiques et phonologiques de ce constituant n'ont pas clairement été définies. Dans cette thèse nous avons fourni des preuves de l'existence du syntagme intermédiaire (ip) à la fois en production et en perception de la parole. Grâce à cinq expérimentations menées dans le cadre de la phonologie de laboratoire, nous avons caractérisé les propriétés phonético-phonologiques de ce constituant et attesté de son rôle dans le traitement perceptif du langage. Les résultats obtenus en production montrent que l'ip est le domaine de l'abaissement des accents mélodiques successifs en français. Sa frontière droite est marquée par un allongement pré-frontière ainsi qu'un accent de syntagme responsable du retour à la ligne de référence du registre. Un retour partiel du registre a également été observé après la frontière d'ip. Par ailleurs, nous avons également testé l'influence de la frontière droite de ce constituant dans le traitement de la parole. Plus particulièrement, nos analyses en perception ont montré que les frontières droites du syntagme accentuel et du syntagme intermédiaire sont utilisées très tôt dans le processus de traitement syntaxique. Les indices phonétiques et phonologiques présents à ces frontières permettent aux auditeurs du français de construire des attentes sur la structure syntaxique des énoncés perçus. Une analyse séparée des différents types d'indices acoustiques a également montré qu'en l'absence de marquage tonal, les indices de durée suffisent à marquer la frontière de syntagme accentuel. Au contraire, le marquage conjoint de la frontière droite d'ip par les indices mélodiques et l'allongement pré-frontière, semble nécessaire pour que les auditeurs du français perçoivent et utilisent cette frontière dans le traitement du langage.

Grâce à une approche pluri-paramétrique des frontières prosodiques, cette thèse permet une meilleure appréhension des relations existantes entre la structure prosodique abstraite et sa réalisation concrète dans la parole. Elle offre également une meilleure compréhension du rôle des constituants prosodiques dans le traitement perceptif.

Mots clefs : découpage prosodique, syntagme accentuel, syntagme intermédiaire, intonation, accès au lexique, traitement syntaxique, phonologie de laboratoire, français.