



**HAL**  
open science

# Quels liens entre accentuation et niveaux de constitution en français? : une analyse perceptive et acoustique

Laury Garnier

## ► To cite this version:

Laury Garnier. Quels liens entre accentuation et niveaux de constitution en français? : une analyse perceptive et acoustique. Linguistique. Université Toulouse le Mirail - Toulouse II, 2018. Français. NNT : 2018TOU20010 . tel-02130477

**HAL Id: tel-02130477**

**<https://theses.hal.science/tel-02130477>**

Submitted on 15 May 2019

**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.



# THÈSE

En vue de l'obtention du

## DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Université Toulouse - Jean Jaurès

---

**Présentée et soutenue par :**

**Laury GARNIER**

**le** mercredi 21 février 2018

**Titre :**

Quels liens entre accentuation et niveaux de constituance en français?  
Une analyse perceptive et acoustique

---

**École doctorale et discipline ou spécialité :**

ED CLESCO : Sciences du langage

**Unité de recherche :**

Octogone-Lordat EA 4156 et CLLE-ERSS UMR 5263

**Directeur/trice(s) de Thèse :**

Corine Astésano, MCF HDR, Université Toulouse - Jean Jaurès

Anne Dagnac, MCF, Université Toulouse - Jean Jaurès

**Jury :**

Élisabeth Delais-Roussarie, DR, Université de Nantes, France (Rapporteure)

Anne-Catherine Simon, PR, Université Catholique de Louvain, Belgique (Rapporteure)

Daniel Hirst, DR émérite, Université Aix-Marseille, France (Examinateur)

Corine Astésano, MCF HDR, Université Toulouse - Jean Jaurès, France (Co-directrice)

Anne Dagnac, MCF, Université Toulouse - Jean Jaurès, France (Co-directrice)



***à Vincent,***



# Remerciements

---

Me voici à la fin de mon aventure, entre épuisement et stress mais aussi excitation et bonheur. Et comme toute grande étape dans la vie, il est important de remercier toutes les personnes qui m'ont accompagnée sur ce parcours.

Mes premiers remerciements sont adressés à Elisabeth Delais-Roussarie et Anne-Catherine Simon pour avoir accepté d'être rapporteuses de ce travail de thèse, ainsi qu'à Daniel Hirst qui a bien voulu participer à mon jury de thèse. C'est un très grand honneur pour moi que vous ayez accepté de juger mon travail.

Je tiens également à remercier mes directrices de thèse, Corine Astésano et Anne Dagnac. Je te remercie Anne d'avoir accepté cette co-direction et d'être entrée dans ce « monde » de la prosodie avec beaucoup de plaisirs. J'ai aussi une pensée toute particulière pour vous, Corine. Je vous remercie chaleureusement pour tout ce que vous m'avez apporté. Merci de m'avoir fait confiance, de m'avoir appris à surmonter mes limites, de m'avoir écoutée dans les moments plus difficiles, mais aussi de m'avoir fait grandir scientifiquement. Je suis très reconnaissante de toutes ces années partagées avec vous.

J'ai réalisé cette thèse au sein des laboratoires Octogone-Lordat et CLLE-ERSS, et je souhaiterais remercier tous les membres de ces deux laboratoires. Je les remercie de leur accueil, ainsi que de leur soutien financier. Merci également de m'avoir soutenue humainement : Barbara Köpke, Mélanie Jucla, Christiane Soum, Encarnacion Arroyo, Charlotte Alazard, Hélène Giraudo, Olga Theophanous, Vanda Marijanovic, Fabio Montermini, Thomas Verjans, et toutes les autres personnes que j'ai pu côtoyer dans les couloirs du laboratoire. Un grand merci aussi à Evelyne !

Il y a une personne que je tiens également à remercier tout particulièrement : Lorraine Baqué. Je suis tellement reconnaissante de toute l'aide précieuse que tu m'as apportée. Ce travail ne serait pas ce qu'il est sans TOI. Merci pour ta générosité et ta bienveillance constante. J'espère de tout cœur que nous continuerons à travailler ensemble.

Un grand merci à Rafèu Sichel-Bazin. Ce fut un plaisir de partager ces moments avec toi !

Un immense merci à Sebastien Déjean et à Saïd Jmel pour leur aide, leurs conseils et le temps qu'ils m'ont consacré dans le cadre de mon travail statistique.

Une pensée aussi pour Roxane Bertrand. Merci pour ta gentillesse et tes bons conseils.

Je voudrais également remercier toutes les personnes qui ont accepté de participer à mes expériences, sans qui le travail expérimental ne serait possible.

Durant mes années de doctorat, j'ai aussi eu la chance de rencontrer nombreux doctorants et jeunes docteurs, qui pour certains sont devenus de vrais amis. Merci à Francesca Cortelazzo, avec qui j'ai commencé cette aventure et avec qui je la termine. Tous ces

moments passés avec toi ont été très importants pour moi et j'espère qu'ils seront encore nombreux ! Merci également à Clara Solier, avec qui j'ai partagé plus qu'un environnement de travail. Je te remercie pour ta bienveillance et pour ton soutien quotidien. Merci à Noémie te Rietmolen et Olivier Nocaudie, mes acolytes... Je vous remercie d'avoir toujours été là pour moi. Nous avons déjà partagé de tellement bons souvenirs ! Je souhaiterais également remercier Lucille Soulier pour sa bonne humeur et sa force tranquille, si précieuses. Je pense aussi à Luce Lefeuvre et Karla Orihuela. Tous tes petits messages d'encouragement, Luce, tout au long de ma rédaction m'ont beaucoup aidée. Ta gentillesse, ta bonne humeur et ton soutien sans faille, Karla, ont été très importants pour moi. Et puis, merci à tous les autres : Kleopatra Mytara, Lyanne Ahumada Ebratt, Rogelio Méndez, Ajhar Ahmad, Aurélie Pistono, Enhao Leger-Zheng, Maxime Warnier, Lama Allan, Cécile Viollain, Marine Lasserre, Emilie Merdy. J'en oublie... Mais merci à tous !

Enfin je tiens à remercier ma famille et mes amis, qui ont su me donner beaucoup d'amour malgré mes absences durant ces derniers mois. Je dois dire que j'ai beaucoup de chance d'être aussi bien entourée. Je tiens tout d'abord à remercier mes parents, qui m'ont toujours fait confiance et qui m'ont toujours soutenue dans mes choix. J'ai tellement appris de vous : la détermination, l'optimisme, l'humilité, mais également le goût de la vie et de tout ce qu'elle peut nous offrir. Merci aussi à mes deux frères (j'espère vous rendre fière), mes petits neveux et nièces d'amour, ma tatel (un vrai modèle pour moi), mes grands-mères... et puis tout le reste de ma famille « formidable ». Je remercie également mes beaux-parents qui ont toujours su m'épauler, je sais que je peux aussi compter sur vous. Un immense merci à mes amis, Fanny, Romain, Justine, Guillaume, Marion et Pierre, qui ont contribué à rendre ma vie plus belle. J'ai bien sûr une pensée particulière pour toi, Marion. Tu étais là dès le début, puis tu as emprunté d'autres chemins, mais tu as continué à me soutenir coûte que coûte. Un énorme merci ! Tu as été très importante pour moi pendant toutes ces années.

Enfin, mes derniers remerciements te sont destinés Vincent, évidemment...Je n'ai pas de mot pour te dire combien je suis reconnaissante. Tu as partagé tous les moments importants de cette thèse, les bons mais aussi les mauvais, et tu as toujours trouvé les mots pour m'aider à surmonter les passages les plus difficiles. Merci pour tout. Tu as été d'un soutien incroyable et j'ai beaucoup de chance de t'avoir dans ma vie. C'est donc tout naturellement que je te dédie cette thèse.

# Table des Matières

---

---

Remerciements .....	3
Table des Matières .....	5
Table des Figures .....	9
Table des Tableaux .....	13
Table des Annexes .....	15
Liste des abréviations .....	17
<b>INTRODUCTION</b> .....	<b>19</b>

## Partie I - Cadre Théorique et Méthodologie

### **CHAPITRE 1 – La prosodie comme « supra-système » ..... 27**

<b>1. LES TROIS ORDRES STRUCTURELS DU PHRASE PROSODIQUE</b> .....	<b>29</b>
1.1 Ordre de structuration métrique .....	30
1.1.1 Définition de l'accent .....	31
1.1.2 Fonctions et paramètres de l'accent .....	32
1.2 Ordre de structuration temporelle .....	34
1.2.1 Définition du rythme .....	35
1.2.2 Fonctions et paramètres du rythme .....	36
1.3 Ordre de structuration tonale .....	39
1.3.1 Définition de l'intonation .....	39
1.3.2 Fonctions et paramètres de l'intonation .....	40
<b>SYNTHESE DU CHAPITRE 1</b> .....	<b>43</b>

### **CHAPITRE 2 – La question de l'identité des constituants prosodiques..... 45**

<b>1. MODELES THEORIQUES DE LA HIERARCHIE DES CONSTITUANTS PROSODIQUES</b> .....	<b>47</b>
1.1 Approche morphosyntaxique .....	47
1.1.1 La Théorie Prosodique .....	48
1.1.1.1 The Strict Layer Hypothesis (SLH) .....	50
1.1.1.2 Révision de la SLH dans le cadre de la Théorie de l'Optimalité .....	51
1.1.1.3 Construction du Syntagme Phonologique à l'interface prosodie-syntaxe ....	54
1.1.1.4 Le cas particulier du Mot Prosodique/Phonologique .....	56
1.2 Approche plus « purement » prosodique .....	59
1.2.1 La Théorie Métrique .....	60
1.2.1.1 L'arbre métrique .....	62
1.2.1.2 La grille métrique simple et parenthésée .....	63
1.2.2 La Théorie Métrique et Autosegmentale .....	64

1.2.2.1	Le choix d'une analyse à deux niveaux tonals.....	66
1.2.2.2	Le modèle standard.....	67
1.2.2.2.1	La règle de « downstep ».....	70
1.2.2.2.2	La règle d'« upstep » .....	72
1.2.2.2.3	La règle de « tonal spreading » .....	73
1.2.2.3	Révision du modèle standard.....	75
1.2.2.4	Le système de notation ToBI.....	77
<b>2.</b>	<b>REALITE PSYCHO-COGNITIVE DE LA HIERARCHIE DES CONSTITUANTS PROSODIQUES .....</b>	<b>79</b>
2.1	Travaux en production sur les indices de groupements.....	79
2.1.1	L'analyse des structures de performances .....	79
2.1.2	L'analyse en « chunks ».....	81
2.2	Travaux en perception sur les indices de frontières et de proéminences.....	83
	<b>SYNTHESE DU CHAPITRE 2 .....</b>	<b>89</b>
	 <b>CHAPITRE 3 – Les particularités du phrasé prosodique en français .....</b>	 <b>93</b>
<b>1.</b>	<b>LE FRANÇAIS EST-IL UNE LANGUE SANS ACCENT ? .....</b>	<b>95</b>
1.1	Accent non lexicalement distinctif.....	95
1.1.1	Accent final primaire.....	95
1.1.2	Accent initial secondaire.....	97
1.2	Synchrétisme entre accentuation et intonation .....	99
<b>2.</b>	<b>LES MODELES DU PHRASE PROSODIQUE EN FRANÇAIS .....</b>	<b>101</b>
2.1	L'approche Intonosyntaxique/Phonosyntaxique .....	102
2.1.1	Delattre .....	103
2.1.2	Di Cristo .....	104
2.1.3	Rossi.....	105
2.1.4	Martin.....	107
2.2	L'approche Métrique Autosegmentale .....	110
2.2.1	Hirst et Di Cristo .....	111
2.2.2	Mertens .....	116
2.2.3	Jun & Fougeron.....	121
2.2.4	Post.....	126
2.3	L'approche métrique et fonctionnaliste .....	131
2.3.1	Delais-Roussarie.....	132
2.3.2	Di Cristo .....	136
	<b>SYNTHESE DU CHAPITRE 3 .....</b>	<b>143</b>
	 <b>CHAPITRE 4 – Les aspects méthodologiques de la thèse.....</b>	 <b>147</b>
<b>1.</b>	<b>OBJECTIFS DE LA THESE .....</b>	<b>149</b>
<b>2.</b>	<b>SELECTION DU MATERIEL LINGUISTIQUE .....</b>	<b>154</b>
2.1	Le corpus de parole contrôlée : Corpus d'Edimbourg.....	154
2.2	Le corpus de parole semi-spontanée : Corpus de la Map Task Aix .....	156
2.3	Le corpus de parole spontanée : Corpus du CID.....	160

2.4	Préparation du corpus de parole sélectionné .....	162
<b>3.</b>	<b>TEST DE PERCEPTION DE LA PAROLE.....</b>	<b>164</b>
3.1	Participants .....	164
3.2	Tâches de perception.....	164
3.3	Préparation de la base de données.....	166
<b>4.</b>	<b>HYPOTHESES ET PREDICTIONS .....</b>	<b>169</b>
4.1	Hypothèse 1 : Une granularité plus fine des niveaux de constituency .....	170
4.2	Hypothèse 2 : Pas d'affaiblissement perceptif de l'AF en frontière majeure .....	174
4.3	Hypothèse 3 : Un rôle plus structurel que purement rythmique de l'Al.....	176
<b>5.</b>	<b>DEMARCHE STATISTIQUE .....</b>	<b>178</b>
5.1	Test d'indépendance du Khi deux .....	178
5.2	Modèle de régression ordinale à effet mixte .....	179

## Partie II - Analyses Expérimentales

<b>CHAPITRE 5 – Résultats et Discussion .....</b>	<b>187</b>	
<b>1. PERCEPTION DE LA FORCE DES FRONTIERES .....</b>	<b>189</b>	
1.1	Lien de dépendance entre le groupement et la condition syntaxique .....	189
1.2	Statistiques descriptives des principales variables .....	191
1.2.1	Distribution des scores de frontières par groupement.....	192
1.2.2	Distribution des scores de frontières par taille du constituant .....	197
1.3	Choix des sites d'analyse.....	199
1.4	Résultats du modèle de régression ordinale à effet mixte .....	200
1.4.1	Force de la frontière perçue après N1 .....	200
1.4.2	Force de la frontière perçue après N2 .....	202
1.4.3	Force relative entre les deux frontières perçues (N2 ADJ – N1 N2).....	208
1.5	Discussion.....	214
<b>2. PERCEPTION DE LA FORCE DES PROEMINENCES .....</b>	<b>220</b>	
2.1	Statistiques descriptives des principales variables .....	221
2.1.1	Distribution des scores de proéminences par groupement .....	222
2.1.2	Distribution des scores de proéminences par taille du constituant.....	227
2.2	Choix des sites d'analyse.....	233
2.3	Résultats du modèle de régression ordinale à effet mixte .....	234
2.3.1	Force de la proéminence perçue sur la syllabe finale de N1.....	235
2.3.2	Force de la proéminence perçue sur la syllabe initiale de ADJ.....	238
2.3.3	Force relative perçue entre la proéminence finale sur N1 et les autres proéminences initiales ou finales adjacentes.....	243
2.3.3.1	Proéminence finale sur N1 – Proéminence initiale sur N1 .....	244
2.3.3.2	Proéminence initiale sur N2 – Proéminence finale sur N1 .....	246
2.3.3.3	Proéminence finale sur N2 – Proéminence finale sur N1 .....	248

2.3.4	Force relative perçue entre la proéminence initiale sur ADJ et les autres proéminences initiales ou finales adjacentes.....	252
2.3.4.1	Proéminence finale sur ADJ – Proéminence initiale sur ADJ.....	253
2.3.4.2	Proéminence initiale sur ADJ – Proéminence finale sur N2.....	257
2.3.4.3	Proéminence initiale sur ADJ – Proéminence initiale sur N2.....	262
2.4	Discussion.....	266
<b>3.</b>	<b>CORRELATS ACOUSTIQUES DES FRONTIERES ET DES PROEMINENCES PERÇUES.....</b>	<b>274</b>
3.1	Démarche statistique.....	275
3.2	Statistiques descriptives des principales variables acoustiques.....	276
3.3	Résultats du modèle de régression ordinale à effet mixte.....	286
3.3.1	Corrélat acoustiques de la force des frontières.....	287
3.3.1.1	Choix des sites d'analyse.....	287
3.3.1.2	Force de la frontière perçue après N1.....	289
3.3.1.3	Force de la frontière perçue après N2.....	291
3.3.1.4	Discussion.....	293
3.3.2	Corrélat acoustiques de la force des proéminences.....	298
3.3.2.1	Sites d'analyse.....	298
3.3.2.2	Force de la proéminence perçue sur la syllabe finale de N1.....	302
3.3.2.3	Force de la proéminence perçue sur la syllabe initiale de ADJ.....	307
3.3.2.4	Force de la proéminence perçue sur les syllabes initiales des mots lexicaux.....	312
3.3.2.5	Discussion.....	312
	<b>CONCLUSION ET PERSPECTIVES.....</b>	<b>315</b>
	<b>BIBLIOGRAPHIE.....</b>	<b>325</b>
	<b>ANNEXES.....</b>	<b>341</b>

## Table des Figures

---

---

<b>Figure 1</b> : Règle de l'étagement strict (SLH).....	50
<b>Figure 2</b> : Phénomène de récursivité prosodique .....	52
<b>Figure 3</b> : Arbre métrique associé au mot "Alabama" .....	63
<b>Figure 4</b> : Grille métrique simple en A) et parenthésée en B) associée au mot "Alabama" .....	64
<b>Figure 5</b> : Exemple d'un stepping pattern réalisé sur l'énoncé « blueberries, bayberries, raspberries, mulberries and brambleberries ».....	66
<b>Figure 6</b> : Grammaire à état fini du modèle de Pierrehumbert (1980, p. 29).....	68
<b>Figure 7</b> : Représentation des courbes mélodiques associées à l'énoncé « only a millionaire » prononcé avec trois différents accents mélodiques .....	69
<b>Figure 8</b> : Phénomène de « downstep » (noté « d ») réalisé sur l'énoncé « there are many intermediate levels » .....	71
<b>Figure 9</b> : Phénomène d'« upstep » (noté « u ») réalisé sur l'énoncé « Can I go now ? » .....	72
<b>Figure 10</b> : Schéma des phénomènes d'« upstep » et de « partial reset » .....	73
<b>Figure 11</b> : Exemples des séquences tonales donnant lieu à la réalisation de la règle de « tonal spreading » (noté « s »).....	74
<b>Figure 12</b> : Grammaire à état fini du modèle révisé par Beckman & Pierrehumbert (1986) .....	75
<b>Figure 13</b> : Exemple d'une transcription avec le système ToBI de l'énoncé « Amelia knew him » .....	77
<b>Figure 14</b> : Exemple d'application de la règle portant sur le rattachement du Syntagme Phonologique, dont la tête est un verbe, à l'unité prosodique comportant le moins de nœuds branchants dans l'arbre.....	80
<b>Figure 15</b> : Récapitulatif des différentes hiérarchies en constituants prosodiques issues des principaux courants théoriques : approche morphosyntaxique et approche plus purement prosodique .....	90
<b>Figure 16</b> : Phénomènes d'arcs accentuels réalisés sur « amplitudes » et « divers formants » dans l'énoncé « les amplitudes des divers formants ».....	98
<b>Figure 17</b> : Dialogue illustrant les 10 intonations de base du français présentées par Delattre (1966b, p. 6).....	103
<b>Figure 18</b> : Exemple d'application de la règle d'inversion de pente dans l'énoncé « les hippopotames de Nabuchodonosor ont été regroupés en associations » .....	108
<b>Figure 19</b> : Exemple de l'application des contraintes d'eurythmie créant des cas de non-congruence entre la structure prosodique et syntaxique .....	109
<b>Figure 20</b> : Segmentations possibles de l'énoncé « la secrétaire du président a bien décoré la salle » selon le modèle de Hirst & Di Cristo (1983).....	112
<b>Figure 21</b> : Représentation arborescente et en grille métrique de l'énoncé « mon fils et son voisin se sont disputés » selon le modèle de Di Cristo & Hirst (1993, p. 6).....	113
<b>Figure 22</b> : Exemple d'une transcription sous INTSINT de l'énoncé « Oui mais la jeune chanteuse elle a de la voix et elle tient la scène » .....	115
<b>Figure 23</b> : Les différentes configurations tonales possibles associées à l'énoncé « elle est sympathique », proposées dans le modèle de Mertens (2008, p. 96).....	116
<b>Figure 24</b> : Exemple de réalisation d'un Accent Initial sur des mots clitiques .....	117
<b>Figure 25</b> : Segmentation de l'énoncé « de la voir ainsi m'attriste beaucoup », selon le modèle de Mertens (1993b, p. 3) .....	118
<b>Figure 26</b> : Cas de désaccentuation réalisée sur la syllabe finale du mot « hiver » dans l'énoncé « je crains l'hiver froid à Moscou » .....	119
<b>Figure 27</b> : Exemple d'un regroupement de 4 Groupes Intonatifs en 2 Paquets Intonatifs : [(l'institut) (de phonétique)] [(de l'université) (de Bruxelles)] .....	119
<b>Figure 28</b> : Sortie du Prosogramme associé à l'énoncé « durant toutes ces années donc de vous pencher sur le sort des femmes » .....	120
<b>Figure 29</b> : Structure prosodique hiérarchique proposée dans le modèle de Jun & Fougeron (2002, p. 152).....	121

<b>Figure 30</b> : Schéma de différents gabarits tonals possibles associés au Syntagme Accentuel, dans le cadre de modèle de Jun & Fougeron (2000, p. 216).....	123
<b>Figure 31</b> : Schéma des phénomènes d'abaissement et de réinitialisation partielle entre deux ip successifs .....	124
<b>Figure 32</b> : Graphique représentant les pourcentages associés aux différents patterns tonals mis en évidence dans l'étude en production de Welby (2003, p. 73).....	125
<b>Figure 33</b> : Grammaire tonale proposée dans le cadre du modèle de Post (2000, p. 154).....	127
<b>Figure 34</b> : Configurations tonales possibles des « boundary tones » en frontière de Syntagme Intonatif .....	128
<b>Figure 35</b> : Effet de la présence du ton « L-insertion » sur une séquence « downstepped » ....	130
<b>Figure 36</b> : Grille métrique de l'énoncé « les cours reprendront le huit septembre » réalisée après l'application des contraintes rythmiques, selon le modèle de Delais-Roussarie (1999, p. 27).....	134
<b>Figure 37</b> : Grille métrique de l'énoncé « les cours reprendront le huit septembre » réalisée après l'application des contraintes linguistiques, selon le modèle de Delais-Roussarie (1999, p. 26).....	135
<b>Figure 38</b> : Schéma illustrant le Principe de bipolarisation accentuelle en structure sous-jacente et la projection des accents en structure de surface.....	137
<b>Figure 39</b> : Grille métrique associée à l'énoncé « Au frère de François, je lui ai offert un livre de peinture » .....	138
<b>Figure 40</b> : Exemple de la variabilité de l'ancrage de l'Accent Initial non-emphatique .....	139
<b>Figure 41</b> : Cas de désaccentuation partielle des Accents Finaux positionnés sur les syllabes internes à l'arc accentuel réalisée dans l'énoncé « le président de la république »	140
<b>Figure 42</b> : Récapitulatif des différentes hiérarchies en constituants prosodiques issues des différents modèles du phrasé prosodique en français : approche Intonosyntaxique/Phonosyntaxique, approche Métrique Autosegmentale et approche Métrique et Fonctionnaliste .....	146
<b>Figure 43</b> : Exemple des repères typographiques associés aux structures syntaxiquement ambiguës issues du corpus d'Edimbourg (Astésano et al., 2007).....	157
<b>Figure 44</b> : Tableau récapitulatif de la répartition des 4 structures syntaxiquement ambiguës issues du corpus d'Edimbourg (Astésano et al., 2007) à travers les 4 dialogues (illustrés par les 4 couleurs différentes).....	158
<b>Figure 45</b> : Liste des 16 phrases retenues dans le test de perception qui ont chacune été produites dans les deux conditions syntaxiques par 4 locuteurs différents (au total : 128 phrases).....	163
<b>Figure 46</b> : Format de présentation des 3 tâches de perception appliquées à la séquence D1+N1+cc+D2+N2+ADJ .....	165
<b>Figure 47</b> : Distribution des groupements (G1, G2, G3 et G4) par condition syntaxique (Condition 1 à gauche et Condition 2 à droite) .....	190
<b>Figure 48</b> : Distribution des scores de frontières (de 0 à 3) après chaque mot du syntagme par catégorie de groupement : en G11 dans la figure (A) et en G31 dans la figure (B)	193
<b>Figure 49</b> : Distribution des scores de frontières (0 à 3) après chaque mot du syntagme par catégorie de groupement : en G22 dans la figure (A), en G32 dans la figure (B) et en G42 dans la figure (C) .....	194
<b>Figure 50</b> : Distribution des scores de frontières après N2 en G22 par le nombre de syllabes concaténé du syntagme D1+N1+CC+D2+N2+ADJ .....	198
<b>Figure 51</b> : Illustration de l'effet du groupement sur la force de la frontière perçue après N1 ....	201
<b>Figure 52</b> : Illustration de l'effet d'interaction entre le groupement et le nombre de syllabes de N sur la force de la frontière perçue après N2. ....	205
<b>Figure 53</b> : Illustration de l'effet d'interaction entre le groupement et le nombre de syllabes de N sur la force relative entre les deux frontières perçues (frontière après N2 – frontière après N1) .....	211
<b>Figure 54</b> : Distribution des scores de proéminences sur la syllabe finale de N1 par catégorie de groupement (G11, G31, G22, G32 et G42).....	223
<b>Figure 55</b> : Distribution des scores de proéminences sur la syllabe initiale de ADJ par catégorie de groupement (G11, G31, G22, G32 et G42).....	225
<b>Figure 56</b> : Distribution des scores de proéminences sur la dernière syllabe de N1 en G11 par le nombre de syllabes de N.....	229

<b>Figure 57</b> : Distribution des scores de proéminences sur la première syllabe de ADJ par le nombre de syllabes total du syntagme D+N1+CC+D+N2+ADJ en G22 .....	231
<b>Figure 58</b> : Distribution des scores de proéminences sur la première syllabe de ADJ par le nombre de syllabes total du syntagme D+N1+CC+D+N2+ADJ en G32 .....	232
<b>Figure 59</b> : Illustration de l'effet du groupement sur la force de la proéminence perçue sur la syllabe finale de N1 .....	236
<b>Figure 60</b> : Illustration de l'effet du nombre de syllabes de N sur la force de la proéminence perçue sur la syllabe finale de N1 .....	237
<b>Figure 61</b> : Illustration de l'effet du groupement sur la force de la proéminence perçue sur la première syllabe de ADJ.....	239
<b>Figure 62</b> : Illustration de l'effet d'interaction entre le nombre de syllabes de N et ADJ sur la force de la proéminence perçue sur la première syllabe de ADJ .....	242
<b>Figure 63</b> : Illustration de l'effet du groupement sur la force relative entre les deux proéminences perçues sur N1 (dernière syllabe de N1 – première syllabe de N1) .....	245
<b>Figure 64</b> : Illustration de l'effet du groupement sur la force relative entre les deux proéminences perçues sur N1 et N2 (première syllabe de N2 - dernière syllabe de N1) .....	247
<b>Figure 65</b> : Illustration de l'effet du groupement sur la force relative entre les deux proéminences perçues sur N1 et N2 (dernière syllabe de N2 – dernière syllabe de N1) .....	249
<b>Figure 66</b> : Illustration de l'effet du nombre de syllabes de N sur la force relative entre les deux proéminences perçues sur N1 et N2 (dernière syllabe de N2 – dernière syllabe de N1) .....	251
<b>Figure 67</b> : Illustration de l'effet du groupement sur la force relative entre les deux proéminences perçues sur ADJ (dernière syllabe de ADJ – première syllabe de ADJ) .....	254
<b>Figure 68</b> : Illustration de l'effet d'interaction entre le nombre de syllabes de N et ADJ sur la force relative entre les deux proéminences perçues sur ADJ (dernière syllabe de ADJ – première syllabe de ADJ) .....	257
<b>Figure 69</b> : Illustration de l'effet du groupement sur la force relative entre les deux proéminences perçues sur N2 et ADJ (première syllabe de ADJ – dernière syllabe de N2).....	259
<b>Figure 70</b> : Illustration de l'effet d'interaction entre le nombre de syllabes de N et de ADJ sur la force relative entre les deux proéminences perçues sur N2 et ADJ (première syllabe de ADJ – dernière syllabe de N2).....	261
<b>Figure 71</b> : Illustration de l'effet du groupement sur la force relative entre les deux proéminences perçues sur N2 et ADJ (première syllabe de ADJ – première syllabe de N2).....	263
<b>Figure 72</b> : Illustration de l'effet d'interaction entre le nombre de syllabes de N et de ADJ sur la force relative entre les deux proéminences perçues sur N2 et ADJ (première syllabe de ADJ – première syllabe de N2).....	265
<b>Figure 73</b> : Récapitulatif des résultats de l'effet du groupement sur les scores de proéminences perçues à l'initiale et à la finale des mots lexicaux (N1, N2 et ADJ).....	269
<b>Figure 74</b> : Moyennes de f0 (Hz) et de durée (ms) de chaque syllabe des syntagmes réalisés en Condition 1, toutes tailles du constituant confondues .....	278
<b>Figure 75</b> : f0 médiane (Hz) et durée (ms) de chaque syllabe des syntagmes réalisés en Condition 1 : « les lumières et les balises vermeilles » dans la Figure (A) et « les bagatelles et les balivernes saugrenues » dans la Figure (B) .....	279
<b>Figure 76</b> : Moyennes de f0 (Hz) et de durée (ms) de chaque syllabe des syntagmes réalisés en Condition 2, toutes tailles du constituant confondues .....	281
<b>Figure 77</b> : f0 médiane (Hz) et durée (ms) de chaque syllabe des syntagmes réalisés en Condition 2 : « les lumières et les balises verticales » dans la Figure (A) et « les bagatelles et les balivernes saugrenues » dans la Figure (B) .....	282
<b>Figure 78</b> : Moyennes de f0 (Hz) et de durée (ms) de chaque syllabe des syntagmes réalisés en Condition 1 dans la figure du haut, et en Condition 2 dans la figure du bas, toutes tailles du constituant confondues .....	284
<b>Figure 79</b> : Moyennes de f0 (Hz) associées à AI et AF sur chaque mot lexical (N1, N2 et ADJ) en Condition 1 et 2.....	285
<b>Figure 80</b> : Moyennes de f0 (Hz) et de durée (ms) sur chaque syllabe des syntagmes « les N1 et les N2 » réalisés en G11 et G31 (Condition 1), toutes tailles du constituant confondues .....	289

<b>Figure 81</b> : Moyennes de f0 (Hz) et de durée (ms) sur chaque syllabe des syntagmes « les N1 et les N2 » réalisés en G32, G42 et G22 (Condition 2), toutes tailles du constituant confondues .....	290
<b>Figure 82</b> : Moyennes de f0 (Hz) et de durée (ms) sur chaque syllabe des syntagmes « et les N2 ADJ » réalisés en G11 et G31 (Condition 1), toutes tailles du constituant confondues .....	292
<b>Figure 83</b> : Moyennes de f0 (Hz) et de durée (ms) sur chaque syllabe des syntagmes « et les N2 ADJ » réalisés en G32, G42 et G22 (Condition 2), toutes tailles du constituant confondues .....	293
<b>Figure 84</b> : Moyennes de f0 (Hz) et de durée (ms) sur chaque syllabe des syntagmes « les N1 et les N2 ADJ » réalisés en G32 (Condition 2), toutes tailles du constituant confondues .....	296
<b>Figure 85</b> : Moyennes de f0 (Hz) et de durée (ms) sur chaque syllabe des syntagmes « les N1 et les N2 ADJ » réalisés en G22 (Condition 2), toutes tailles du constituant confondues .....	296
<b>Figure 86</b> : Moyennes de f0 (Hz) et de durée (ms) sur chaque syllabe des syntagmes « les N1 et les N2 ADJ » réalisés en G42 (Condition 2), toutes tailles du constituant confondues .....	297
<b>Figure 87</b> : Moyennes de f0 (Hz) et de durée (ms) associées à la syllabe finale de N1 et initiale de N2 en Condition 1 .....	304
<b>Figure 88</b> : Moyennes de f0 (Hz) associées à la syllabe finale de N1 et initiale de N2 en Condition 2 .....	305
<b>Figure 89</b> : Moyennes de f0 (Hz) et de durée (ms) associées à la syllabe finale de N1 et initiale de N2 en Condition 1 et en Condition 2 .....	306
<b>Figure 90</b> : Moyennes de f0 (Hz) et de durée (ms) associées aux syllabes finales de N2 et ADJ en Condition 1 .....	309
<b>Figure 91</b> : Moyennes de f0 (Hz) et de durée (ms) associées à la syllabe finale de N2, et aux syllabes initiale et finale de ADJ en Condition 2 .....	311

## Table des Tableaux

---

---

<b>Tableau 1</b> : Exemple d'un résultat factice illustrant la distribution de probabilités de chaque valeur de la variable ordinale (le score de frontière : 0, 1, 2 et 3) dans chaque niveau de la variable indépendante (groupement : G1, G2, G3 et G4) .....	181
<b>Tableau 2</b> : Exemple d'un résultat factice illustrant les moyennes des différents niveaux de la variable ordinale (score de frontière : 0, 1, 2, 3) interprétées comme des valeurs numériques ordonnées de 1 à 4 dans la colonne « mean.class » et de 0 à 3 dans la colonne « mean.class (-1) », et calculées à partir des résultats de la distribution de probabilités estimée pour chaque groupement (cf. Tableau N°1) .....	182
<b>Tableau 3</b> : Résultats des comparaisons post hoc de l'effet du groupement sur la force de la frontière perçue après N1 .....	201
<b>Tableau 4</b> : Résultats des comparaisons post hoc de l'effet d'interaction entre le groupement et le nombre de syllabes de N sur la force de la frontière perçue après N2 .....	204
<b>Tableau 5</b> : Résultats des comparaisons post hoc de l'effet d'interaction entre le groupement et le nombre de syllabes de N sur la force relative entre les deux frontières perçues (frontière après N2 – frontière après N1) .....	210
<b>Tableau 6</b> : Récapitulatif des résultats de l'effet du groupement sur la force des frontières perçues après N1, après N2 et sur la force relative entre les deux frontières perçues (frontière après N2 – frontière après N1) .....	215
<b>Tableau 7</b> : Résultats des comparaisons post hoc de l'effet du groupement sur la force de la prééminence perçue sur la syllabe finale de N1 .....	235
<b>Tableau 8</b> : Résultats des comparaisons post hoc de l'effet du nombre de syllabes de N sur la force de la prééminence perçue sur la syllabe finale de N1 .....	237
<b>Tableau 9</b> : Résultats des comparaisons post hoc de l'effet du groupement sur la force de la prééminence perçue sur la première syllabe de ADJ .....	239
<b>Tableau 10</b> : Résultats des comparaisons post hoc de l'effet d'interaction entre le nombre de syllabes de N et de ADJ sur la force de la prééminence perçue sur la syllabe initiale de ADJ .....	241
<b>Tableau 11</b> : Résultats des comparaisons post hoc de l'effet du groupement sur la force relative entre les deux prééminences perçues sur N1 (dernière syllabe de N1 – première syllabe de N1) .....	245
<b>Tableau 12</b> : Résultats des comparaisons post hoc de l'effet du groupement sur la force relative entre les deux prééminences perçues sur N1 et N2 (première syllabe de N2 – dernière syllabe de N1) .....	246
<b>Tableau 13</b> : Résultats des comparaisons post hoc de l'effet du groupement sur la force relative entre les deux prééminences perçues sur N1 et N2 (dernière syllabe de N2 – dernière syllabe de N1) .....	249
<b>Tableau 14</b> : Résultats des comparaisons post hoc de l'effet du nombre de syllabes de N sur la force relative entre les deux prééminences perçues sur N1 et N2 (dernière syllabe de N2 – dernière syllabe de N1) .....	250
<b>Tableau 15</b> : Résultats des comparaisons post hoc de l'effet du groupement sur la force relative entre les deux prééminences perçues sur ADJ (dernière syllabe de ADJ – première syllabe de ADJ) .....	254
<b>Tableau 16</b> : Résultats des comparaisons post hoc de l'effet d'interaction entre le nombre de syllabes de N et de ADJ sur la force relative entre les deux prééminences perçues sur ADJ (dernière syllabe de ADJ – première syllabe de ADJ) .....	256
<b>Tableau 17</b> : Résultats des comparaisons post hoc de l'effet du groupement sur la force relative entre les deux prééminences perçues sur N2 et ADJ (première syllabe de ADJ – dernière syllabe de N2) .....	258

<b>Tableau 18</b> : Résultats des comparaisons post hoc de l'effet d'interaction entre le nombre de syllabes de N et de ADJ sur la force relative entre les deux proéminences perçues sur N2 et ADJ (première syllabe de ADJ – dernière syllabe de N2) ...	260
<b>Tableau 19</b> : Résultats des comparaisons post hoc de l'effet du groupement sur la force relative entre les deux proéminences perçues sur N2 et ADJ (première syllabe de ADJ – première syllabe de N2).....	262
<b>Tableau 20</b> : Résultats des comparaisons post hoc de l'effet d'interaction entre le nombre de syllabes de N et de ADJ sur la force relative entre les deux proéminences perçues sur N2 et ADJ (première syllabe de ADJ – première syllabe de N2) ..	264
<b>Tableau 21</b> : Mise en relation des valeurs « mean.class » des résultats perceptifs avec les rapports de durée des analyses acoustiques .....	294

# Table des Annexes

---

---

## **ANNEXE 1 :**

CORPUS DE LA MAP TASK AIX (BARD ET AL., 2013)..... 343

## **ANNEXE 2 :**

CORPUS DU CID (BERTRAND ET AL., 2008)..... 347

## **ANNEXE 3 :**

CORPUS D'EDIMBOURG (ASTESANO ET AL., 2007)..... 349

## **ANNEXE 4 :**

DOCUMENTS REMPLIS PAR LES PARTICIPANTS DE L'EXPERIMENTATION..... 351

## **ANNEXE 5:**

CONSIGNES DU TEST DE PERCEPTION ..... 357

## **ANNEXE 6 :**

RESULTATS DU LIEN DE DEPENDANCE ENTRE LE GROUPEMENT ET LA CONDITION SYNTAXIQUE. .... 361

## **ANNEXE 7 :**

RESULTATS DE LA DISTRIBUTION DES SCORES DE FRONTIERES PAR GROUPEMENT. .... 363

## **ANNEXE 8 :**

RESULTATS DE LA DISTRIBUTION DES SCORES DE FRONTIERES PAR TAILLE DU CONSTITUANT..... 369

## **ANNEXE 9 :**

RESULTATS DE L'EFFET DU GROUPEMENT ET DE LA TAILLE DU CONSTITUANT SUR LA FORCE DES FRONTIERES. .... 373

## **ANNEXE 10 :**

RESULTATS DE LA DISTRIBUTION DES SCORES DE PROEMINENCES PAR GROUPEMENT. .... 379

## **ANNEXE 11 :**

RESULTATS DE LA DISTRIBUTION DES SCORES DE PROEMINENCES PAR TAILLE DU CONSTITUANT..... 381

**ANNEXE 12 :**

RESULTATS DE LA COMPARAISON DES DEGRES DE PROEMINENCE PERÇUS SUR LES SYLLABES INITIALES ET FINALES DE N1, N2 ET ADJ..... 387

**ANNEXE 13 :**

RESULTATS DE LA COMPARAISON DES DEGRES DE PROEMINENCE PERÇUS SUR LA PREMIERE, DEUXIEME ET TROISIEME SYLLABE DE N1, N2 ET ADJ..... 389

**ANNEXE 14 :**

RESULTATS DE LA COMPARAISON DES DEGRES DE PROEMINENCE PERÇUS SUR LA DEUXIEME, TROISIEME ET DERNIERE SYLLABE DE N1, N2 ET ADJ..... 391

**ANNEXE 15 :**

RESULTATS DE LA COMPARAISON DES DEGRES DE PROEMINENCE PERÇUS SUR LES SYLLABES INTERMEDIAIRES DE N1, N2 ET ADJ, ET SUR LES SYLLABES DES MOTS GRAMMATICaux. .... 393

**ANNEXE 16 :**

RESULTATS DE L'EFFET DU GROUPEMENT ET DE LA TAILLE DU CONSTITUANT SUR LA FORCE DES PROEMINENCES..... 395

**ANNEXE 17 :**

MOYENNES DE  $f_0$  ET DE DUREE SUR LES TOUTES LES SYLLABES DES SYNTAGMES REALISES EN CONDITION 1 ET 2, TOUS LOCUTEURS CONFONDUS..... 407

**ANNEXE 18 :**

RESULTATS DE L'EFFET DES VARIATIONS DE DUREE ET DE  $f_0$  SUR LA FORCE DES FRONTIERES ET DES PROEMINENCES..... 417

**ANNEXE 19 :**

ARTICLE PUBLIE DANS LES ACTES DE LA CONFERENCE INTERNATIONALE *SPEECH PROSODY 2016*.... 421

**ANNEXE 20 :**

ARTICLE PUBLIE DANS LES ACTES DE LA CONFERENCE INTERNATIONALE *JEP-TALN-RECITAL 2016*.....427

## Liste des abréviations

---

<b>AF</b>	Accent Final	
<b>AI</b>	Accent Initial	
<b>AP</b>	Syntagme Accentuel ( <i>Accentual Phrase</i> )	
<b>cm</b>	Continuation Mineure	] Dans la terminologie de Delattre (1966b)
<b>CM</b>	Continuation Majeure	
<b>ct</b>	Continuation Mineure	] Dans la terminologie de Rossi (1981)
<b>CT</b>	Continuation Majeure	
<b>cc</b>	Conclusif Mineur	
<b>CC</b>	Conclusif Majeur	
<b>f<sub>0</sub></b>	Fréquence Fondamentale	
<b>GA</b>	Groupe Accentuel	
<b>GI</b>	Groupe Intonatif	
<b>INTSINT</b>	International Transcription System for Intonation	
<b>ip</b>	Syntagme Intermediare ( <i>Intermediate Phrase</i> )	
<b>IP</b>	Syntagme Intonatif ( <i>Intonation ou Intonational Phrase</i> )	
<b>MOMEL</b>	MOdelling MELody	
<b>N1</b>	Premier Nom	
<b>N2</b>	Deuxième Nom	
<b>ADJ</b>	Adjectif	
<b>Perceval</b>	PERCEPTION EVALuation Auditive & Visuelle	
<b>PhP</b>	Syntagme Phonologique ( <i>Phonological Phrase</i> )	
<b>PW</b>	Mot Prosodique ( <i>Prosodic Word</i> )	
<b>SLH</b>	Règle d'étagement strict ( <i>Strict Layer Hypothesis</i> )	
<b>ToBI</b>	Tone and Break Indices	
<b>UT</b>	Unité Tonale	
<b>UI</b>	Unité Intonative	
<b>UR</b>	Unité Rythmique	



## Introduction

---

La prosodie joue un rôle central dans les langues du monde, en transmettant des informations sur les attitudes, les émotions et l'identification des locuteurs, mais également en participant aux processus d'encodage et de décodage du discours. Dans le cadre de notre travail, nous choisissons de nous intéresser à cette fonction structurale, et plus précisément à l'organisation du phrasé prosodique, c'est-à-dire la manière dont les informations prosodiques sont utilisées pour segmenter et hiérarchiser le flux continu de parole. Nous choisissons par ailleurs de concentrer notre recherche sur le français, proportionnellement moins étudié sur la scène internationale que les autres langues.

Parfois qualifié de langue « exotique », le français serait la seule langue parmi toutes les autres langues romanes et germaniques sans accentuation lexicale. En effet, le français est traditionnellement considéré comme « *une langue sans accent* » (Rossi, 1980) ou « *une langue de frontière* » (Vaissière, 1991), où l'Accent Final (noté « AF ») principal est cooccurrent à la frontière prosodique majeure. Dans ce contexte, l'accent en français ne serait pas lexical mais plutôt post-lexical, en marquant le niveau du groupe de mots plutôt que le niveau du mot lui-même. Un autre type d'accent est également décrit en français. Il s'agit de l'Accent Initial (noté « AI »), dit secondaire et optionnel, qui aurait une fonction emphatique, dépendant de critères sémantico-pragmatiques, ou une fonction plus rythmique en apparaissant sur les longs constituants dépourvus d'AF. Cet AI n'interviendrait pas en revanche dans la structuration et dans la hiérarchisation du discours. Cette vision traditionnelle du système accentuel en français, que nous pouvons toutefois retrouver dans certains des modèles du phrasé prosodique plus actuels, participe donc au « flou théorique » autour du lien entre accentuation et constituance prosodique dans la langue.

Ces notions de constituance, de structuration ou de hiérarchie font par ailleurs intervenir la notion d'organisation syntaxique, et par conséquent d'interface prosodie-syntaxe. Nombreux travaux se sont intéressés au lien à établir entre les deux niveaux de structuration. Toutefois, au-delà de la question de correspondance entre la structure prosodique et syntaxique, qui peut parfois s'avérer circulaire, il apparaît difficile, dans toutes les langues en général, de trouver un consensus sur la manière dont la structure est segmentée en groupes, ou en constituants prosodiques. En effet, les auteurs ne s'accordent pas quant au choix et au nombre de niveaux de constituance reflétant la hiérarchie prosodique, et ce notamment parce qu'ils n'utilisent pas les mêmes règles de formation de

ces groupes prosodiques. Nous verrons par ailleurs que les auteurs ne s'entendent pas non plus sur la terminologie donnée aux différentes unités phonologiques.

L'objectif de notre travail est donc d'explorer l'organisation du phrasé prosodique en français, en questionnant notamment les niveaux de constituance nécessaires à l'élaboration du phrasé prosodique dans la langue. En effet, seuls deux niveaux de constituance sont actuellement admis en français : le Syntagme Intonatif (IP) et le Syntagme Accentuel (AP). Des débats sont également en cours quant à l'existence d'un Syntagme Intermédiaire (ip), situé, comme son nom l'indique, entre l'AP et l'IP. En revanche, étant donné la vision post-lexicale de l'accentuation en français, aucun des modèles du phrasé prosodique actuel ne semble intégrer le niveau du mot lexical dans la hiérarchie des constituants prosodiques. De la même manière, directement en lien avec la notion de domaine accentuel, nous tenons à rendre compte du rôle phonologique de l'accent dans le marquage de la constituance prosodique. Tout d'abord, nous remettons en question le phénomène de syncrétisme entre l'accentuation et l'intonation. En effet, dans la vision traditionnelle du système accentuel en français, mais également dans des modèles plus récents, AF s'effacerait perceptivement en frontières prosodiques majeures, et ne serait donc pas indépendant de la frontière intonative. Par ailleurs, nous nous interrogeons quant au rôle purement rythmique, secondaire et optionnel de l'AI. Nous rappelons effectivement qu'AI n'est pas envisagé dans un rôle structurel mais uniquement dans un rôle emphatique ou rythmique.

Notre thèse se compose plus précisément de deux grandes parties. La première partie de la thèse permet de définir les contours de notre cadre théorique et de mettre en place notre procédure expérimentale, tandis que la seconde partie est exclusivement consacrée à la présentation et à la discussion de l'ensemble de nos résultats.

## **Partie I – Cadre Théorique et Méthodologie**

### **Premier Chapitre : La prosodie comme « supra-système »**

Nous introduisons dans le premier chapitre les trois ordres de structuration qui composent le « *supra-système* » (Di Cristo, 2000, p. 5) qu'est la prosodie, à savoir l'organisation métrique (accentuation), temporelle (rythme) et tonale (intonation). Ces trois niveaux structurels participent conjointement à l'élaboration du phrasé prosodique. Il apparaît donc important, selon nous, de commencer notre travail de thèse par définir ces composantes phonologiques, en précisant leurs fonctions et leurs rôles dans la segmentation du flux de parole, tout en rendant en compte également de l'étroite relation qui

existe entre les trois paramètres. Par ailleurs, nous profitons de ce premier chapitre pour définir certaines notions ou certaines dichotomies fondamentales que nous retrouverons tout au long du manuscrit.

### Deuxième Chapitre : La question de l'identité des constituants prosodiques

Le deuxième chapitre de notre thèse est consacré à la question de l'identité des constituants prosodiques qui fait l'objet de discussions, liées notamment à des divergences concernant la prise en compte des différentes composantes phonologiques de la prosodie. Plus précisément, nous proposons de présenter les trois courants phonologiques qui ont contribué à la modélisation de la structure prosodique et à l'identification d'une hiérarchie des constituants : **la Théorie Prosodique** (Hayes, 1989 ; Nespor & Vogel, 1986 ; Selkirk, 1980 ; 1984 ; 1986), **la Théorie Métrique** (Halle & Vergnaud, 1987 ; Hayes, 1995 ; Liberman, 1975 ; Liberman & Prince, 1977) et **la Théorie Métrique Autosegmentale** (Beckman & Pierrehumbert, 1986 ; Pierrehumbert, 1980). Bien que ces trois courants théoriques aient été pensés initialement pour la description de l'anglais, certains des principes et des représentations prosodiques ont vocation cependant à avoir une portée universelle. En revanche, nous verrons que ces différentes théories se distinguent les unes des autres concernant notamment les règles de formation des niveaux de constituance, mais également concernant le nombre de niveaux de constituance intégrés dans la hiérarchie prosodique. En comparant par ailleurs ces différentes hiérarchies, nous verrons qu'il est difficile de mettre en correspondance certains niveaux de constituance dans les différentes approches, et que cette difficulté est renforcée lorsque les auteurs utilisent des terminologies différentes pour se référer au même niveau de constituance, ou au contraire, lorsqu'ils utilisent une même terminologie pour se référer à différents niveaux de constituance. La présentation de ces courants théoriques nous permet alors de rendre compte des différentes « affiliations » théoriques possibles, que nous retrouverons par la suite dans les modèles prosodiques appliqués plus spécifiquement au français.

Une seconde partie dans ce chapitre est quant à elle consacrée à la présentation d'une série d'études expérimentales qui ont mis en lien les prédictions théoriques précédentes avec le traitement réel du langage : c'est ce que nous avons intitulé « la réalité psycho-cognitive de la hiérarchie des constituants prosodiques ». Plusieurs études en production ou en perception sont exposées dans le cadre de cette partie, dont les différents résultats et protocoles expérimentaux sont commentés et mis en relation. La pertinence de ces différentes études sera par ailleurs discutée lors de la mise en place de notre propre procédure expérimentale.

### Troisième Chapitre : Les particularités du phrasé prosodique en français

Le troisième et dernier chapitre théorique est cette fois-ci consacré aux modèles du phrasé prosodique en français. Toutefois, avant de proposer une présentation chronologique des différentes approches, nous commençons par rappeler la vision traditionnelle du système accentuel dans la langue, en décrivant plus précisément les deux types d'accents, ainsi que le phénomène de syncrétisme entre l'accentuation et l'intonation, à l'origine du postulat de langue sans accent. Sur la base de cette remise en contexte, nous présentons par la suite les différents modèles du phrasé prosodique en français, en proposant une classification selon la manière dont les modèles sont « affiliés » par rapport aux trois courants théoriques, présentés dans le cadre du second chapitre. Nous commençons tout d'abord par **les modèles Intonosyntaxiques ou Phonosyntaxiques** (Delattre, 1966 ; Di Cristo, 1978 ; Martin, 1987 ; Rossi, 1985) qui se situent davantage dans l'approche morphosyntaxique de la Théorie Prosodique. Nous proposons par la suite des modèles dits plus « purement » prosodiques avec **l'approche Métrique Autosegmentale** (Hirst & Di Cristo, 1996 ; Jun & Fougeron, 2000 ; Mertens, 1997 ; Post, 2000) et **Métrique et Fonctionnaliste** (Delais-Roussarie, 1999 ; Di Cristo, 2000). Pour chacun de ces modèles, nous nous intéressons à la manière dont sont gérées, au vu de la vision traditionnelle du français, les composantes phonologiques de la prosodie (accentuation, rythme et intonation), les contraintes issues des autres domaines linguistiques (notamment le domaine syntaxique), mais également le nombre et le type de niveaux constituance.

Au fil de la présentation de ces différents modèles, nous verrons l'intérêt d'ancrer notre travail de recherche à la croisée de deux approches théoriques, à savoir l'approche Métrique Autosegmentale (notamment, Hirst & Di Cristo, 1996 ; Post, 2000) et l'approche Métrique et Fonctionnaliste (Delais-Roussarie, 1999 ; Di Cristo, 2000). En effet, nous considérons qu'il est nécessaire, pour rendre compte du lien entre l'accentuation et la constituance prosodique, de recourir à des modèles prosodiques qui ont un ancrage métrique fort. Nous adoptons également une méthodologie inspirée de ces deux théories, en proposant notamment une démarche acoustico-perceptive.

### Quatrième Chapitre : Les aspects méthodologiques de la thèse

Les aspects méthodologiques de la thèse sont présentés dans le quatrième chapitre : questions de recherche, matériel linguistique, procédure expérimentale, préparation de la base de données, hypothèses et analyse statistique. Nous commençons par définir les objectifs de notre recherche, au regard non seulement de notre ancrage théorique mais également au regard des précédents travaux, sur lesquels nous nous fondons pour la mise

en place de notre protocole expérimental et de nos questions de recherche. En effet, notre travail se situe dans la continuité même de précédentes études en production (Astésano et *al.*, 2007) et en perception (Astésano et *al.*, 2012 ; 2013), qui tendent à démontrer la pertinence du niveau du mot lexical en français à travers notamment les règles de réalisations des proéminences initiales.

Dans le cadre de notre travail, nous choisissons d'éprouver les résultats de ces précédents travaux. Plus précisément, nous proposons une étude perceptive, via un corpus de parole contrôlée manipulant des structures syntaxiquement ambiguës, où 80 participants ont effectué 3 tâches de perception : proéminence, frontière et groupement. Les événements perçus ont ensuite été mis en relation avec leurs réalités physiques ou acoustiques. Notre étude se distingue des précédents travaux en perception, concernant notamment le recours à une analyse syntagmatique des événements prosodiques. En effet, la plupart des travaux en perception testent les proéminences et les frontières perçues dans un énoncé sur le plan paradigmatique seulement, en s'intéressant à une catégorie d'événements prosodiques spécifique (frontière majeure, accent nucléaire, etc.). Toutefois, ces études ne semblent pas mettre en relation ces différents événements avec les autres événements prosodiques adjacents. Dans le cadre de notre travail, nous privilégions au contraire une étude testant la perception des frontières et des proéminences à la fois sur le plan paradigmatique et syntagmatique, via notamment notre tâche de groupement. Plus précisément, nous proposons de mettre en relation, au sein d'un même groupement ainsi qu'à travers les différents groupements, les proéminences avec les frontières, mais également les différents niveaux de frontière ou de proéminence entre eux. L'objectif de cette analyse est effectivement de mieux comprendre comment ces différents événements prosodiques s'organisent dans la construction même du phrasé prosodique d'un énoncé.

## **Partie II – Analyses Expérimentales**

### **Cinquième Chapitre : Résultats et Discussion**

Les résultats de l'ensemble de ces analyses sont ensuite présentés dans le cadre de ce dernier chapitre. Nous commentons tout d'abord les résultats de perception, pour ensuite décrire les analyses des corrélats acoustiques associés aux événements perçus. Etant donné l'ampleur des analyses effectuées dans le cadre de notre travail, nous proposons après chaque analyse d'un événement prosodique spécifique une discussion permettant de synthétiser les résultats obtenus. Au fil de la présentation des résultats, les discussions suivantes seront incrémentées par les discussions précédentes, pour ainsi faciliter

## Introduction

---

l'interprétation globale de l'ensemble des résultats. Nous verrons que nos résultats tendent à affiner les résultats des précédents travaux en production (Astésano et *al.*, 2007) et en perception (Astésano et *al.*, 2012 ; 2013), tout en apportant également de nouvelles perspectives sur la vision traditionnelle des phénomènes accentuels en français, et sur l'organisation de la hiérarchie des constituants prosodiques dans la langue.

Enfin, nous reviendrons en conclusion sur nos objectifs de recherche, et nous résumerons les résultats majeurs de notre travail qui ouvrent la voie à de nouvelles perspectives.

**- PARTIE I -**  
**Cadre Théorique et Méthodologie**



# CHAPITRE 1 – La prosodie comme « supra-système »

---

<b>1. LES TROIS ORDRES STRUCTURELS DU PHRASE PROSODIQUE .....</b>	<b>29</b>
1.1 Ordre de structuration métrique.....	30
1.1.1 Définition de l'accent .....	31
1.1.2 Fonctions et paramètres de l'accent.....	32
1.2 Ordre de structuration temporelle.....	34
1.2.1 Définition du rythme .....	35
1.2.2 Fonctions et paramètres du rythme.....	36
1.3 Ordre de structuration tonale.....	39
1.3.1 Définition de l'intonation .....	39
1.3.2 Fonctions et paramètres de l'intonation.....	40
 <b>SYNTHESE DU CHAPITRE 1 .....</b>	 <b>43</b>



### 1. LES TROIS ORDRES STRUCTURELS DU PHRASE PROSODIQUE

Quand nous produisons un flux de parole, notre discours est segmenté en différentes unités, appelées « *constituants prosodiques* » ou « *domaines prosodiques* », qui sont hiérarchisées et qui participent à l'élaboration du phrasé prosodique. Ces groupements jouent alors un rôle important dans le traitement de la parole, en facilitant notamment la compréhension des énoncés : le découpage prosodique influe sur le sens du message que ce soit sur le plan de l'encodage ou du décodage de la parole.

Certains auteurs tiennent par ailleurs à distinguer les notions de constituance et de domaine prosodique (Delais-Roussarie & Fougeron, 2004 ; Di Cristo, 2011), dans la mesure où elles ne se définissent pas, selon eux, sur le même plan. Effectivement, alors que le constituant prosodique renvoie à un phénomène de groupement qui s'intègre dans une hiérarchie, et qui est délimité par une frontière gauche et droite, le domaine prosodique, en revanche, concerne un « *empan* » plus flexible, sur lequel sont réalisées des règles ou des opérations spécifiques. Ainsi, les constituants prosodiques sont définis comme des domaines où opèrent des règles phonologiques (Di Cristo, 2016) : un Groupe Intonatif (ou Syntagme Intonatif, Unité Intonative selon les approches), par exemple, est le domaine de la réalisation du contour intonatif, et de la même manière, le Groupe Accentuel (ou Syntagme Accentuel, Syntagme Phonologique selon les approches) est le domaine de la réalisation de l'accent. Toutefois, il n'est pas exclu qu'un même domaine prosodique puisse aussi s'appliquer à différents types de constituants, ou au contraire, qu'un même type de constituant puisse concerner à la fois différents domaines prosodiques (Jun, 1993 cité par Delais-Roussarie & Fougeron, 2004)<sup>1</sup>.

Alors que nous verrons dans les prochains chapitres que l'identité et la taille de ces unités phonologiques font l'objet de discussion (*cf.* Chapitre 2), et plus spécifiquement en français (*cf.* Chapitre 3), nous nous intéressons ici aux différents indices qui contribuent au traitement du flux de parole ou au « *phrasage* » prosodique. En effet, la prosodie peut être considérée comme un « *supra-système* » (Di Cristo, 2000, p. 5), composé de trois niveaux structurels concourant à l'élaboration même du phrasé prosodique d'une langue : l'organisation métrique qui est le domaine de l'accent, l'organisation temporelle consacrée au rythme, et l'organisation tonale dont l'intonation est l'élément principal. Ces trois ordres structurels sont réalisés sur le plan phonétique, par trois paramètres physiques : variations

---

<sup>1</sup> Delais- Roussarie & Fougeron (2004) montrent par exemple que le domaine post-focal peut s'appliquer aussi bien au Groupe Accentuel qu'au Groupe Intonatif. Dans le même sens, Jun (1993), cité par Delais-Roussarie & Fougeron (2004), met en évidence que le Groupe Accentuel en coréen est le domaine d'application de plusieurs règles segmentales (lénition des occlusives voisées, réduction vocalique, etc.).

de la fréquence fondamentale (notée «  $f_0$  »), de l'intensité et de la durée. Nous commencerons donc par présenter l'organisation métrique, sur laquelle se fonde ensuite l'organisation temporelle, qui représente enfin le point d'ancrage de l'organisation tonale.

### 1.1 Ordre de structuration métrique

Avant d'aller plus loin dans la présentation de cette première organisation, il est nécessaire de distinguer le niveau du mètre et le niveau du rythme. En effet, alors que la structure métrique renvoie à une organisation abstraite, relativement contrainte et figée, la structure rythmique, résultant de l'organisation temporelle (cf. 1.2 du Chapitre 1), quant à elle, consiste à actualiser cette organisation abstraite lors de la production concrète de l'énoncé (Couper-Kuhlen, 1993 ; Di Cristo, 1999a ; 2003).

*« Nous proposons de différencier ces concepts sur la base d'une opposition entre « représentations sous-jacentes » et « représentations superficielles ». Selon ce point de vue, les éléments constitutifs du mètre sont représentables au niveau sous-jacent ou profond par les gabarits métriques qui définissent le système prosodique noyau de la langue. Les structures rythmiques de surface, qui sont dérivables des gabarits métriques, donnent lieu à des motifs diversifiés. » (Di Cristo, 1999a, p. 3)*

Il semblerait que cette distinction entre le niveau métrique et rythmique rappelle également une autre distinction, à savoir entre le niveau phonologique et phonétique. Effectivement, alors que le niveau phonologique correspond aux phénomènes abstraits, le niveau phonétique, en revanche, concerne les phénomènes plus concrets, impliquant les paramètres acoustiques tels que la durée, l'intensité ou la  $f_0$  (notamment Cutler & Ladd, 1983). Toutefois, selon Di Cristo (1999b), cette dichotomie « *phonologie/phonétique* » ne propose pas un cadre théorique suffisant pour rendre compte de la structuration prosodique. L'auteur envisage effectivement une approche plus complète, composée de trois à quatre niveaux de représentation prosodique (cf. 2.2.1 du Chapitre 3 : « *INTSINT* » Hirst & Di Cristo, 1998) : le niveau profond de la représentation phonologique (ou sous-jacent), consacré à l'élaboration des gabarits métriques abstraits, le niveau superficiel de la représentation phonologique (ou de surface), où sont actualisés les motifs rythmiques, le niveau phonétique qui correspond à la ligne « *mélodico-rythmique* », où les éléments prosodiques, définis dans les niveaux plus abstraits, sont actualisés concrètement par les

paramètres acoustiques, relevant enfin du niveau physique. Dans ce cadre, l'organisation métrique correspond au niveau le plus abstrait, dont l'entité primitive est représentée par un accent ou une proéminence. Cette unité phonologique se manifeste sur le plan métrique par un battement fort en comparaison aux syllabes inaccentuées, interprétées quant à elle comme un battement faible.

### 1.1.1 Définition de l'accent

Pour mieux comprendre la notion d'« *accent* », il est tout d'abord intéressant de distinguer les unités « *accentogènes* » qui sont les unités susceptibles de recevoir un accent, et les unités « *non-accentogènes* » correspondant aux unités qui, en général, ne peuvent pas recevoir d'accent. Alors que les mots lexicaux tendent à être tous « *accentogènes* », les mots grammaticaux, en revanche, sont davantage considérés comme « *non-accentogènes* ». Une autre dichotomie consiste également à distinguer « *l'unité accentuable* » qui concerne l'élément supportant l'accent, à savoir notamment la syllabe (Garde, 1968), et « *l'unité accentuelle* » qui renvoie au domaine où l'accent exerce ses fonctions, à savoir le mot, le groupe de mots ou encore l'énoncé. L'accent peut alors être « *lexical* » lorsqu'il marque le niveau du mot, ou « *post-lexical* » lorsqu'au contraire, il s'applique à des domaines dépassant le cadre du mot. Plus précisément, l'accent consiste à détacher un élément de son environnement, pour ainsi le rendre plus saillant que les autres éléments.

« *J'entends par accent l'entité linguistique (prosodique) ayant pour fonction principale la mise en relief d'une syllabe et dont la substance consiste dans un plus grand effort expiratoire et articulatoire. Cet effort se reflète sur le plan acoustique dans des modifications caractéristiques de la courbe de fréquence fondamentale, de la courbe de pression sonore et du spectre de fréquences et dans l'allongement de la syllabe accentuée.* » (Fónagy, 1980, p. 125)

Ce phénomène de mise en relief rappelle par ailleurs les principes de la Gestalt (ou « Théorie de la forme »), principalement appliqués à la perception des sons et de la musique, où il est considéré que nous percevons des « *figures* » ou des « *formes* » qui se dégagent d'un « *fond* ». Ce principe appliqué à la parole tend à considérer les syllabes inaccentuées comme un « *fond* », sur lequel se détachent les syllabes accentuées constituant alors la « *figure* » ou « *la forme* » (Astésano, 2001 ; Couper-Kuhlen, 1993 ;

Pasdeloup, 1990a ; 1990b ; Pasdeloup et *al.*, 2006). Pour finir sur ce mode dichotomique de la définition de l'accent, nous pouvons également distinguer les termes « *stress* » et « *accent* ». Alors que cette dichotomie est peu utilisée en français, qui se limite au terme « *accent* », dans d'autres langues telles que l'anglais en revanche, ces deux termes renvoient à deux unités distinctes, issues de deux niveaux d'analyse différents, dont les définitions varient d'un auteur à l'autre. Selon Bolinger (1989), le terme « *stress* » correspond à l'accent lexical potentiel défini en structure sous-jacente, ou dans la représentation phonologique, et le terme « *accent* » fait référence à la manifestation physique du « *stress* » dans la représentation phonétique. Plus précisément, l'auteur parle de « *pitch accent* » étant donné que selon lui, la réalisation physique d'un accent est essentiellement due aux variations mélodiques (*f0*). A l'inverse, la dichotomie « *stress/accent* » prend une tout autre tournure chez Beckman (1986), pour qui le terme « *accent* » renvoie cette fois-ci à la caractéristique phonologique abstraite, spécifiée au niveau du mot (« *lexical accent* » dans la terminologie de Ladd, 1996 / 2008), et le terme « *stress* » correspond à la description phonétique d'une réalisation concrète possible de l'accent. En effet, l'auteur distingue le « *stress accent* », dont la réalisation phonétique est marquée par un renforcement de l'intensité et de la durée (associé également à une variation de *f0*), et le « *non-stress accent* » qui est marqué uniquement par un mouvement mélodique ou de *f0* (similaire au « *pitch accent* »). Pour le français, Garde (1968) propose de conserver le terme « *accent* », en établissant néanmoins une distinction entre « *accent sous-jacent* » en référence à la structure profonde ou métrique, et « *accent effectif* » en référence à la structure de surface ou rythmique (dans la lignée des propositions de Bolinger, 1989). A l'inverse des autres dichotomies, Garde propose d'associer un type d'accent à un niveau d'analyse spécifique (niveau sous-jacent vs. niveau de surface), sans préjuger de la réalisation physique ou acoustique de l'accent.

### 1.1.2 Fonctions et paramètres de l'accent

Concernant plus précisément le rôle de l'accent, différentes fonctions peuvent être définies selon le type de langue où est exercée l'accentuation. Nous pouvons en effet distinguer « *les langues à accent libre* », où la position de l'accent varie d'un lexème à un autre, et « *les langues à accent fixe* », où la position de l'accent est constante et fixe. Tout d'abord, l'accent est par définition culminatif, dans la mesure où il permet de faire ressortir un élément par rapport aux autres, et plus précisément de répertorier les mots lexicaux dans un énoncé. Effectivement, cette fonction culminative sous-entend que chaque mot lexical contient au moins une syllabe accentuée. L'accent peut également avoir une fonction

distinctive, lorsque la position de l'accent a une incidence directe sur le sens du mot, et donc lorsque l'accent fait partie de la structure interne du mot (sa position figure notamment dans l'entrée du dictionnaire). Enfin, l'accent peut être démarcatif, lorsqu'il sert à délimiter le début et la fin des mots ou groupes de mots dans la chaîne sonore. Alors que les deux premières fonctions concernent davantage les langues à accent libre, la troisième fonction, en revanche, est réservée aux langues à accent fixe.

Par ailleurs, en dehors de ces trois fonctions de base, la réalisation d'un accent peut également être influencée par des contraintes de nature syntaxique, sémantique, pragmatique ou rythmique. En effet, un accent peut être réalisé pour faciliter une alternance régulière entre les syllabes accentuées et inaccentuées (accent rythmique), ou encore pour signaler qu'un élément (mot ou groupe de mots) est plus important qu'un autre (accent emphatique). La syntaxe joue également un rôle important dans la réalisation de l'accent, puisque l'accent peut être réalisé pour délimiter la fin d'un constituant syntaxique majeur : il s'agit dans ce cas-là d'un « *accent nucléaire* ». Toutefois, certains auteurs réfutent l'empreinte purement syntaxique de l'accent nucléaire, et notamment Bolinger (1972), qui considère que la position de cet accent est davantage influencée par le choix du locuteur, cherchant avant tout à optimiser sa communication en facilitant l'accès à l'information, au risque de ne pas toujours correspondre à la structure syntaxique prédite.

Sur le plan phonétique, il est admis que l'accent se manifeste par plusieurs paramètres physiques, pouvant inclure une montée de la courbe mélodique (de la  $f_0$ ), un renforcement de l'intensité et une augmentation de la durée. Cependant, la réalisation de ces différents paramètres va dépendre non seulement de la langue envisagée, mais aussi du type d'accent produit. La variation de hauteur mélodique (liée à la courbe de  $f_0$ ) tout d'abord, est généralement considérée comme l'indice prosodique le plus pertinent, quelle que soit la langue (cf. « *pitch accent* » Bolinger, 1989). Le paramètre d'intensité est également un indice prégnant, dont l'importance diffère selon les langues. En effet, alors qu'en anglais, l'intensité joue un rôle déterminant dans la réalisation accentuelle (cf. « *stress accent* » Beckman, 1986), en français, en revanche, l'intensité ne permet pas de distinguer clairement une syllabe accentuée d'une syllabe inaccentuée (Delattre, 1966a). Cependant, l'intensité peut jouer un rôle important en français, notamment lors de la réalisation d'accents emphatiques. D'un point de vue perceptif, la différence entre l'accent emphatique et non-emphatique intervient principalement au niveau de l'intensité produite sous l'accent emphatique. Enfin, concernant la durée, la perception d'un allongement est particulièrement importante dans la caractérisation de l'accentuation en français, et plus précisément pour l'accent final. Cependant, étant donné qu'en français, l'accent final est congruent aux frontières prosodiques (Rossi, 1980 : cf. 1.2 du Chapitre 3), et que l'allongement devant une frontière

prosodique est un phénomène universel, la pertinence de ce paramètre peut alors être remise en cause. Il est d'ailleurs intéressant de remarquer que les théories prosodiques majoritaires pour l'analyse du phrasé en français délaissent le paramètre de durée au profit des variations mélodiques (cf. 2.2 du Chapitre 3). Toutefois, certains travaux tendent à montrer l'intérêt de la prise en compte de la durée comme caractérisation de l'accent en français (Delais-Roussarie, 1995 ; 2000 ; Delattre, 1966b ; Di Cristo, 1999a ; 1999b ; 2004 ; Post, 2000 ; Rossi, 1981 ; 1985). Astésano (2001) montre également que la durée infra-syllabique est un indice pertinent en français permettant de distinguer les différents types d'accents dans la langue.

Enfin, en lien avec la réalisation des paramètres physiques, il est important de signaler ici l'existence de « *phénomènes microprosodiques* » (Di Cristo, 1978). En effet, les variations de  $f_0$ , mais aussi d'intensité et de durée, peuvent être soumises à des perturbations microprosodiques liées directement à la nature du matériel segmental. Plus précisément, Di Cristo (1978) rend compte de deux types de variations microprosodiques : intrinsèques et co-intrinsèques. Les phénomènes intrinsèques, tout d'abord, concernent l'influence d'un type de segment sur la configuration des paramètres physiques : une voyelle nasale, par exemple, a une durée plus longue qu'une voyelle orale, ou encore une consonne occlusive voisée présente un abaissement important de la courbe de  $f_0$  par rapport aux autres types de consonnes. Quant aux phénomènes co-intrinsèques, il s'agit cette fois-ci de l'influence de la coarticulation de deux segments sur la configuration des paramètres physiques : lorsqu'une voyelle est précédée d'une consonne sourde, par exemple, la valeur de  $f_0$  est plus élevée que lorsque cette voyelle est précédée d'une consonne sonore. Ce phénomène de microprosodie est particulièrement important dans le cadre des annotations tonales, où des erreurs d'interprétation peuvent être effectuées. Dans ce cadre, Di Cristo & Hirst (1986) proposent de séparer la composante microprosodique (ou segmentale) et macroprosodique (ou suprasegmentale) de la courbe mélodique, dans le but de pouvoir ensuite modéliser la courbe de  $f_0$  en effaçant les variations microprosodiques, permettant ainsi d'obtenir une courbe lisse et continue, plus propice, selon les auteurs, à l'analyse des variations prosodiques (cf. 2.2.1 du Chapitre 3 « *MOMEL* » Hirst & Espesser, 1993).

### 1.2 Ordre de structuration temporelle

Sur la base de cette structure métrique, l'organisation temporelle peut alors s'opérer. En effet, comme nous l'avons déjà mentionné, le mètre permet de signaler les temps forts en structure sous-jacente, tandis que le rythme organise les prééminences en structure de

surface. Selon Di Cristo (1999b, p. 6), le rythme « est fondé sur l'organisation temporelle des syllabes métriquement fortes et des syllabes métriquement faibles ». Dans ce contexte, les syllabes accentuées, résultant de l'organisation métrique, représentent le point d'ancrage de la structure rythmique.

### 1.2.1 Définition du rythme

Au sens large du terme, le rythme est avant tout biologique, du fait qu'il est inhérent à l'activité humaine : rythme des pulsations cardiaques, rythme respiratoire, rythme du sommeil, etc. La production de la parole est alors dans ce cas soumise au rythme biologique, notamment à travers l'activité respiratoire qui contraint la production des sons. Le rythme est également une capacité du système cognitif qui consiste à structurer les informations produites et perçues (notamment Fraise, 1956 ; 1967 ; 1974). En effet, Fraise (1956 ; 1967) parle de « *rhythmisation subjective* » lorsque dans un flux de parole, nous entendons une succession de groupes de sons à intervalle de temps régulier. Le rythme est donc lié à la notion de périodicité qui est elle-même en lien avec la notion de structure, dans la mesure où c'est la récurrence temporelle des groupements (périodicité) qui va permettre l'identification de la structure (Fraise, 1956). La modification de la durée d'un intervalle peut dans ce contexte affecter l'ensemble de la structure prosodique.

Cette notion d'intervalle régulier, ou de récurrence temporelle, a d'ailleurs été utilisée pour classer les différentes langues, selon la manière dont s'opère la régularité rythmique. Nous parlons d'isochronie, c'est-à-dire une alternance régulière de phénomènes identiques. En effet, Pike (1945), dont les propos ont été repris plus tard par Abercrombie (1967), propose de distinguer les langues dites « à *rythmicité accentuelle* » (ou « *stress-timed* ») et les langues dites « à *rythmicité syllabique* » (ou « *syllable-timed* »). Alors que dans la première catégorie de langues, l'isochronie est régie par l'alternance régulière des accents (notamment en anglais), dans la seconde catégorie, en revanche, c'est la régularité de la longueur des syllabes consécutives qui conduit à l'isochronie (notamment en français). Toutefois, un certain nombre d'auteurs estiment qu'il est difficile de maintenir la distinction entre les deux types d'isochronies, et montrent notamment que l'intervalle inter-accentuel en anglais n'est pas régulier puisqu'il peut varier en fonction du nombre de syllabes du constituant (entre autres, Fant et *al.*, 1991 ; Roach, 1982), et que les syllabes en français n'ont pas toujours la même durée (Roach, 1982 ; Wenk & Wioland, 1982). Dauer (1983), enfin, considère que l'isochronie ne doit pas être réservée à des langues spécifiques, puisque selon elle, il s'agit d'un phénomène universel. Cependant, l'auteur estime qu'il est possible de percevoir des différences de rythme entre les langues, qui sont liées à des

propriétés phonologiques et phonétiques, telles que la structure de la syllabe qui est plus complexe en anglais qu'en français, ou encore le phénomène de réduction vocalique qui est absent du système phonologique du français et présent dans celui de l'anglais ; d'où l'impression d'une durée relativement régulière des syllabes en français. De la même manière, l'isochronie accentuelle de l'anglais serait directement liée à la réalisation phonétique de l'accent dans la langue, qui apparaît plus prégnante perceptivement et acoustiquement qu'en français.

« *En français, les valeurs rythmiques se sont à tel point volatilisées, et l'accent de mot est si faible, que seul le compte des syllabes s'impose impérieusement à l'oreille (...).* » (Bally, 1926, p. 262 ; cité par Astésano, 2001)

Dans ce contexte, Dauer (1983) propose de reformuler cette distinction binaire sous la forme d'un continuum, permettant de classer les langues selon leur tendance à être plus ou moins iso-accentuelles et plus ou moins iso-syllabiques.

### 1.2.2 Fonctions et paramètres du rythme

En ce qui concerne les fonctions du rythme, comme nous l'avons vu précédemment, le rythme intervient pour organiser la structure prosodique d'un énoncé, en gérant la planification des prééminences et des frontières de groupements.

« *La construction de ces motifs (en référence aux motifs rythmiques) repose sur l'organisation temporelle des entités discrètes que sont les syllabes accentuées et inaccentuées (indépendamment de la substance phonique responsable de la manifestation de l'accent et de sa quantification), ainsi que sur l'affectation des frontières et des pauses qui constituent également des entités discrètes de la structure prosodique.* »  
(Di Cristo, 1999a, p. 4)

Plus précisément, l'alternance des syllabes accentuées et inaccentuées est régie selon des principes d'eurythmie, cherchant notamment à éviter des phénomènes de « *clashes accentuels* » lorsque plusieurs syllabes consécutives sont accentuées, ou au contraire de « *vides accentuels* » lorsque plusieurs syllabes consécutives sont inaccentuées. Sur la base

de cette alternance régulière, les syllabes accentuées se regroupent avec les syllabes inaccentuées pour créer des groupements. La relation entre ces groupements est par ailleurs renforcée par différents processus rythmiques, tels que la durée syllabique, les pauses ou encore le débit de parole.

L'allongement de la durée de la syllabe, tout d'abord, participe à la délimitation des différents groupements prosodiques, et plus spécifiquement à la hiérarchie entre ces groupements. Plusieurs travaux montrent effectivement que le degré d'allongement varie en fonction du niveau de constituance (entre autres, Astésano, 1998 ; 2001 ; Delais-Roussarie & Feldhausen, 2014 ; Delais-Roussarie et *al.*, 2016 ; Rossi, 1985 ; Fletcher, 1991 ; Jankowski, 1997 ; Michélas & D'Imperio, 2010 ; Padeloup, 1991 pour le français ; Beckman & Edwards, 1990 ; Ladd & Campbell, 1991 ; Turk & Shattuck-Hufnagel, 2007 ; Wightman et *al.*, 1992 pour l'anglais). Pour le français plus spécifiquement, Fletcher (1991) montre qu'une syllabe accentuée délimitant une frontière prosodique majeure est associée à un degré d'allongement plus important qu'une syllabe accentuée en frontière prosodique mineure. Pour considérer la différence de durée comme pertinente, certains auteurs considèrent par ailleurs que seule une variation d'au moins 20% pourrait permettre une distinction catégorielle (notamment Rossi, 1972). Dans ce même sens, Astésano (1998 ; 2001) met en évidence des patrons de durée spécifiques par catégorie accentuelle dans différents styles de parole (interview, lecture, journal). L'auteur constate en effet que la durée de la syllabe finale varie en fonction du type de frontière suivant la syllabe accentuée : la frontière d'un syntagme majeur est marquée par un allongement de la durée syllabique d'environ 110%, tandis que la frontière d'un constituant mineur est marquée par un allongement de la durée syllabique d'environ 60%. Il est intéressant de préciser que les rapports de durée ont systématiquement été calculés ici entre les syllabes accentuées et les syllabes inaccentuées. Nous constatons donc, là encore, que les travaux sur le français remettent largement en question la caractéristique purement iso-syllabique de la langue, étant donné la pertinence des contrastes de durée des syllabes dans la construction du phrasé prosodique.

Directement en lien avec l'effet d'allongement de la durée syllabique, les pauses participent également à la planification du discours, en signalant notamment le degré de cohésion entre les groupements. Deux types de pauses peuvent être définis : les pauses remplies ou sonores, marquées par des « euh » ou des « hum » d'hésitation ou par un allongement de la syllabe (à distinguer de l'allongement en fin de constituant), et les pauses vides ou sourdes marquées par un silence. Concernant les fonctions des pauses, un certain nombre d'auteurs mettent en évidence le rôle majeur de la pause dans la planification ou dans la programmation du message (Duez, 1987 ; Goldman-Eisler, 1968 ; Grosjean & Deschamps, 1972). En effet, Goldman-Eisler (1968) montre que les pauses d'hésitation, qui

peuvent être marquées par une pause remplie (« euh », « hum » ou allongement de la syllabe) ou vide (silence), ont un rôle important à jouer dans l'encodage de la parole, dans la mesure où elles permettent de laisser le temps à la planification ou à la programmation du message. Par ailleurs, les pauses interviennent également dans la structuration de l'énoncé. Des études montrent par exemple que les pauses de prise de souffle (pauses vides), essentielles à la production du discours, coïncident principalement avec des frontières syntaxiques majeures, permettant ainsi de conserver la cohérence de l'énoncé (Goldman-Eisler, 1968 ; Grosjean & Deschamps, 1975 ; Léon, 1992). De la même manière, la durée et la fréquence des pauses sont étroitement liées à l'organisation de la structure en constituants (Gee & Grosjean, 1983 ; Grosjean et *al.*, 1979). Effectivement, selon le degré de cohésion souhaité entre deux groupements, la longueur et la récurrence des pauses sont modulées : pauses courtes et moins nombreuses en frontière mineure et pauses longues et plus nombreuses en frontière majeure (*cf.* 2.1.1 du Chapitre 2 « *Structures de performance* »). Par ailleurs, les pauses permettent également de maintenir l'impression d'isochronie, notamment lorsque la pause constitue un « *battement silencieux* » servant à préserver une certaine alternance régulière (entre autres, Couper-Kuhlen, 1993 ; Guaitella, 1986).

Concernant enfin le débit de parole, ce paramètre peut également avoir une incidence sur l'organisation des proéminences et des groupements. Effectivement, le nombre de groupements dans un même énoncé peut varier selon le débit de parole du locuteur (entre autres, Fougeron & Jun, 1998 ; Vaissière, 1991 pour le français ; Jun, 1993 pour le coréen). Fougeron & Jun (1998) montrent par exemple que le nombre d'unités prosodiques tend à réduire fortement à débit rapide, notamment parce qu'un plus grand nombre de mots sont regroupés dans une même unité. Plus précisément, les auteurs constatent que les frontières qui disparaissent correspondent davantage aux unités mineures qu'aux unités majeures, ce qui amène à considérer les unités mineures comme moins importantes pour l'élaboration de la structure. Par ailleurs, certaines études montrent que le débit de parole influence également la production des pulsations accentuelles et des pauses (Padeloup, 2004 ; Padeloup et *al.*, 2006). En effet, les pauses démarquant les unités majeures s'effacent lorsque le locuteur augmente son débit de parole. De la même manière, les contrastes de durée entre les syllabes accentuées et inaccentuées sont renforcés en débit lent, tandis qu'ils tendent à diminuer en débit rapide. Plus précisément, Padeloup (2004) constate que les syllabes accentuées sont plus sensibles au débit de parole que les syllabes inaccentuées (dans la lignée des résultats de Duez, 1987). L'auteur considère effectivement que les syllabes accentuées ont plus « *d'élasticité* » que les syllabes inaccentuées, dans la mesure

où la variation de durée est plus forte sur les syllabes accentuées que sur les syllabes inaccentuées.

### 1.3 Ordre de structuration tonale

L'organisation tonale, enfin, est fortement dépendante des organisations métrique et temporelle. En effet, l'intonation, qui est l'élément majeur de l'organisation tonale, s'ancre sur une structure rythmique qui est elle-même, comme nous l'avons déjà mentionné, déterminée par l'accentuation. Plus précisément, les séquences de tons s'alignent sur les syllabes métriquement fortes : l'accent, qui présuppose le rythme, est donc également prééminent à l'intonation.

#### 1.3.1 Définition de l'intonation

Avant d'aller plus loin dans la définition de l'intonation, il est important de distinguer les termes « intonation » et « ton », bien qu'ils évoquent tous les deux la notion de variation mélodique ou de variation de hauteur. Alors que l'intonation s'établit au niveau global (au niveau du groupe de mots ou de l'énoncé), le ton quant à lui, relève exclusivement du niveau local (au niveau du mot). En effet, le ton est une caractéristique propre aux langues tonales (chinois, vietnamien, thaï, etc.), qui est inhérente à la structure interne du mot, dans la mesure où le type de ton influe sur la signification du mot. Le ton peut rejoindre en ce sens la notion d'accentuation lexicale distinctive. Quant à l'intonation, décrite souvent comme la « *mélodie de la parole* » ou « *la musique de la parole* », il s'agit plutôt d'un contour intonatif porteur d'une signification qui peut être liée à la syntaxe, à la pragmatique, au discours ou encore aux émotions.

« *Intonation conveys meanings that apply to phrases or utterances as a whole, such as sentence type or speech act, or focus and information structure. By this definition, intonation excludes features of stress, accent, and tone that are determined in the lexicon, which serve to distinguish one word from another. [...] Phonetically, of course, lexical features of stress, accent, and tone interact with intonational features in many ways.*»

(Ladd, 1996 / 2008, p. 6)

L'analyse de l'intonation peut par ailleurs s'effectuer sous deux approches différentes : une approche holistique et une approche compositionnelle. La première approche consiste à traiter l'intonation comme un contour global qui exprime une signification particulière (Delattre, 1966b ; Di Cristo, 1978 ; Martin, 1987 ; Rossi, 1981), tandis que la seconde approche décompose le contour intonatif en une série d'événements tonals, liés au marquage de l'accent et des frontières, qui ont chacun une caractéristique fonctionnelle propre et qui contribuent ensemble à la construction du contour global (Beckman & Pierrehumbert, 1986 ; Hirst & Di Cristo, 1984 ; Jun & Fougeron, 1995 ; Mertens, 1987 ; Pierrehumbert, 1980 ; Post, 2000). Dans l'approche compositionnelle, l'unité primitive est le ton, appelé également « *autosegment* » ou « *segment tonal* », dont le formalisme est similaire à celui utilisé dans la description des langues tonales. Toutefois, il est important de distinguer les tons utilisés dans les langues tonales des tons utilisés dans les approches compositionnelles, dans la mesure où les autosegments n'ont pas ici de fonction lexicale distinctive. Les tons dans l'approche compositionnelle permettent effectivement de décrire la composition du contour intonatif global. Il est également intéressant de préciser que les modèles intonatifs sont encore aujourd'hui majoritaires pour l'analyse du phrasé prosodique des langues, et ne font que très rarement référence à la métrique et au rythme (*cf.* Chapitre 3 pour les modèles intonatifs appliqués au français).

### 1.3.2 Fonctions et paramètres de l'intonation

L'intonation est donc principalement marquée par des variations mélodiques (de  $f_0$ ), qui, à l'inverse de l'accentuation, relèvent plutôt du niveau global. En effet, ce sont des mouvements tonals locaux, tels que la montée mélodique associée à un accent, qui contribuent à la composition du pattern intonatif global. Quant aux fonctions de l'intonation, comme nous l'avons déjà évoqué, il est établi que ce paramètre prosodique remplit un certain nombre de fonctions, que ce soit au niveau linguistique (structuration de l'énoncé, hiérarchisation de l'information, etc.) qu'au niveau paralinguistique ou extralinguistique (émotions ou attitudes du locuteur, identification du locuteur, etc.).

[L'intonation] « désigne à la fois un système linguistique discret de structuration de l'énoncé (fonction démarcative, déterminée à la fois par la syntaxe et la pragmatique), et un système d'expressions de nuances de sens, d'attitudes et d'émotions (fonctions comportementales et émotives). » (Vaissière, 2006, p. 102)

Concernant plus précisément les fonctions linguistiques de l'intonation, les mouvements mélodiques permettent non seulement de hiérarchiser l'information (structure thème-rhème), de réguler les échanges de parole, de déterminer les modalités du discours (assertion, question, ordre, etc.), mais également de structurer l'énoncé. En effet, l'intonation contribue à linéariser la hiérarchie des constituants syntaxiques, en démarquant notamment les groupes syntaxiques majeurs ; d'où l'utilisation des termes « *Groupe Intonatif* », « *Syntagme Intonatif* » ou « *Unité Intonative* » pour caractériser l'unité prosodique majeure. L'intonation permet également de signaler les relations de dépendance entre les différents constituants (cf. 2.1 du Chapitre 3 « contour mineur » vs. « contour majeur » chez Delattre, 1966b et Rossi, 1981; « règle d'inversion de pente » chez Martin, 1987). Toutefois, certains auteurs tiennent à rendre compte de l'autonomie de l'intonation vis-à-vis de la syntaxe, en montrant que l'organisation des morphèmes intonatifs peut être contrainte par des principes d'ordre rythmique ou métrique (entre autres, Di Cristo, 1978 ; Martin, 1987 ; Post, 2000). De la même manière, bien que certains auteurs admettent le rôle démarcatif de l'intonation, ils privilégient toutefois sa fonction sémantique, et notamment son rôle de hiérarchisation thématique (entre autres, Bolinger, 1972 ; Rossi, 1985 ; Selkirk, 1984 ; Steedman, 2008). Dans ce contexte, l'intonation sert principalement à hiérarchiser le poids informatif des éléments de l'énoncé, au risque de ne pas toujours correspondre à la hiérarchie syntaxique.

« *Les théories syntaxiques de l'intonation considèrent que l'intonation est congruente à la syntaxe. Pourtant c'est dans la mesure où elle ne l'est pas qu'elle a la faculté de réorganiser l'énoncé et d'être informative.* »

(Rossi, 1985, p. 151)

Comme nous le constaterons tout au long des prochains chapitres, bien qu'il semble aujourd'hui reconnu que la structure prosodique est relativement indépendante de la structure syntaxique, nous verrons pourtant que la question du rôle de la syntaxe dans la structuration prosodique reste au cœur des débats, aussi bien dans les modèles traditionnels que dans les modèles plus actuels.



## SYNTHESE DU CHAPITRE 1

---

L'objectif de ce premier chapitre est de définir les paramètres intervenant directement dans l'élaboration de la structure prosodique. Les notions utilisées dans le cadre de cette partie sont par ailleurs fondamentales et préalables à la présentation, dans les deux prochains chapitres, des différents modèles du phrasé prosodique. En effet, nous verrons que l'identité et la taille des constituants font l'objet de discussions (*cf.* Chapitre 2), et plus spécifiquement en français (*cf.* Chapitre 3), qui sont liées à des divergences concernant la prise en compte des différentes composantes prosodiques. Nous proposons donc de reprendre la terminologie de Di Cristo (2000) qui caractérise la prosodie comme un « *supra-système* », composé de trois ordres de structuration qui participent conjointement à la construction du phrasé prosodique. Chacun des ordres structurels que nous avons présentés se compose d'une unité primitive, constituant une composante phonologique de la prosodie : l'accentuation, le rythme et l'intonation.

Plus précisément, concernant l'accentuation, nous considérons que les proéminences potentielles ou métriques se définissent au niveau sous-jacent, tandis que les proéminences effectives s'actualisent en surface sous l'impulsion de l'organisation temporelle ou rythmique (notamment, Couper-Kuhlen, 1993 ; Di Cristo, 1999a ; 2003). Les accents réalisés en structure de surface peuvent alors être soumis à un certain nombre de contraintes, issues de différents niveaux linguistiques (syntaxe, sémantique, pragmatique, discours, etc.). Des désaccords sont d'ailleurs observés entre les auteurs quant à la primauté d'un niveau linguistique sur l'autre (notamment, Bolinger, 1972). Cette définition de l'accent sur deux plans (niveau sous-jacent et niveau de surface) a également amené certains auteurs à distinguer deux types d'accent : « *stress* » et « *accent* » (Beckman, 1986 ; Bolinger, 1989). Toutefois, l'utilisation de ces deux termes varie d'un auteur à l'autre, ce qui laisse place à une large ambiguïté quant à la nature et à la fonction de ces deux types d'accents. Dans ce contexte, nous privilégions une autre terminologie, moins ambivalente, distinguant « *l'accent sous-jacent* » réalisé métriquement et « *l'accent effectif* » réalisé concrètement (Garde, 1968). Quant à l'organisation rythmique, en parallèle à la planification des proéminences, le rythme permet également de générer une structure en groupements (Di Cristo, 1999a), renforcée au moyen de différents processus temporels, tels que l'allongement de la durée syllabique, les pauses et le débit de parole. En effet, le rythme est lié à la notion de périodicité, qui permet l'identification d'une organisation structurée (notion de « *rythmisation subjective* » Fraisse, 1956 ; 1967). Cette régularité rythmique a par ailleurs été utilisée pour

classer les langues du monde, selon la manière dont elles établissent leur isochronie (alternance régulière de phénomènes) : syllabique ou accentuelle (Abercrombie, 1967 ; Pike, 1945). Cependant, cette distinction a peu résisté à l'épreuve des faits, et ce notamment en français, où la pertinence des contrastes de durée dans la hiérarchisation des constituants prosodiques n'est plus à prouver (entre autres, Astésano, 1998 ; 2001 ; Delais-Roussarie & Feldhausen, 2014 ; Fletcher, 1991). Enfin, l'intonation est également en lien avec l'accentuation et le rythme, dans la mesure où les segments tonals s'ancrent sur une structure rythmique, qui est elle-même donnée par l'accentuation. Nous remarquons que la notion d'intonation entraîne moins d'ambiguïté et de discussion que les notions liées à l'accentuation et au rythme. L'intonation est d'ailleurs le paramètre prosodique le plus privilégié, aujourd'hui encore, dans les modèles du phrasé prosodique. Toutefois, comme pour les fonctions de l'accentuation, des désaccords peuvent être observés entre les auteurs concernant l'influence des différents niveaux linguistiques sur l'intonation, et notamment l'influence de la structure syntaxique. La question du rôle de la syntaxe dans la structuration prosodique sera d'ailleurs soulevée dans les deux prochains chapitres.

## CHAPITRE 2 – La question de l'identité des constituants prosodiques

---

<b>1. MODELES THEORIQUES DE LA HIERARCHIE DES CONSTITUANTS PROSODIQUES .....</b>	<b>47</b>
1.1 Approche morphosyntaxique.....	47
1.1.1 La Théorie Prosodique .....	48
1.1.1.1 The Strict Layer Hypothesis (SLH).....	50
1.1.1.2 Révision de la SLH dans le cadre de la Théorie de l'Optimalité.....	51
1.1.1.3 Construction du Syntagme Phonologique à l'interface prosodie-syntaxe ....	54
1.1.1.4 Le cas particulier du Mot Prosodique/Phonologique .....	56
1.2 Approche plus « purement » prosodique.....	59
1.2.1 La Théorie Métrique .....	60
1.2.1.1 L'arbre métrique .....	62
1.2.1.2 La grille métrique simple et parenthésée .....	63
1.2.2 La Théorie Métrique et Autosegmentale.....	64
1.2.2.1 Le choix d'une analyse à deux niveaux tonals.....	66
1.2.2.2 Le modèle standard.....	67
1.2.2.2.1 La règle de « downstep ».....	70
1.2.2.2.2 La règle d'« upstep » .....	72
1.2.2.2.3 La règle de « tonal spreading » .....	73
1.2.2.3 Révision du modèle standard.....	75
1.2.2.4 Le système de notation ToBI.....	77
<b>2. REALITE PSYCHO-COGNITIVE DE LA HIERARCHIE DES CONSTITUANTS PROSODIQUES .....</b>	<b>79</b>
2.1 Travaux en production sur les indices de groupements.....	79
2.1.1 L'analyse des structures de performances .....	79
2.1.2 L'analyse en « chunks ».....	81
2.2 Travaux en perception sur les indices de frontières et de prééminences.....	83
<b>SYNTHESE DU CHAPITRE 2 .....</b>	<b>89</b>



### 1. MODELES THEORIQUES DE LA HIERARCHIE DES CONSTITUANTS PROSODIQUES

Trois courants phonologiques contribuant à la modélisation de la structure prosodique et à l'identification d'une hiérarchie des constituants peuvent être définis : la Théorie Prosodique (1986 Hayes, 1989 ; Nespor & Vogel, 1986 ; Selkirk, 1980 ; 1984), la Théorie Métrique (Halle & Vergnaud, 1987 ; Hayes, 1995 ; Liberman, 1975 ; Liberman & Prince 1977) et la Théorie Métrique Autosegmentale (Beckman & Pierrehumbert, 1986 ; Pierrehumbert, 1980). Ces approches théoriques proposent différentes hiérarchies se distinguant les unes des autres non seulement selon la manière dont les unités sont formées et nommées, mais également selon le nombre de niveaux de constituance qui est accepté. Par ailleurs, bien que les trois approches présentées ici soient toutes issues de la phonologie métrique, dont l'idée centrale est l'organisation hiérarchique des proéminences (cf. 1.2.1 du Chapitre 2), des différences majeures s'observent cependant entre les trois approches concernant les critères établis pour définir la structure des constituants prosodiques. Nous verrons en effet dans un premier temps que la Théorie Prosodique privilégie les critères morphosyntaxiques, tandis que la Théorie Métrique et la Théorie Métrique Autosegmentale, présentées dans un deuxième temps, se fondent sur des critères plus « purement » prosodiques (l'accentuation et l'intonation).

#### 1.1 Approche morphosyntaxique

En considérant que la notion de constituance prosodique est fortement liée à la notion même d'organisation syntaxique, certains auteurs choisissent de privilégier des critères morphosyntaxiques dans la définition des unités de la hiérarchie prosodique. En effet, dans les travaux en phonologie générative linéaire, également appelé « *modèle SPE* » (Chomsky, 1965 ; Chomsky & Halle, 1968), la composante phonologique est essentiellement dépendante de la composante syntaxique : la phonologie est conçue, dans ce modèle, comme un ensemble de règles de réécriture conduisant à l'expression sonore d'une suite générée par la syntaxe. Chomsky et Halle (1968) montrent par exemple que l'accent principal est systématiquement assigné au constituant situé le plus à droite dans la phrase (« *Nuclear Stress Rule* »). La présence ou l'absence de phénomènes phonologiques dépend donc directement, dans ce cas-là, de l'organisation syntaxique. Plus tard, la représentation phonologique des énoncés sur un seul plan (linéaire) est abandonnée au profit d'une représentation plurilinéaire, où les phénomènes phonologiques sont alors envisagés sur le plan suprasegmental, c'est-à-dire indépendamment du plan segmental. Cela permet

notamment de mettre directement en parallèle un phénomène phonologique avec un niveau morphosyntaxique spécifique.

### 1.1.1 La Théorie Prosodique

Dans ce cadre, la Phonologie Prosodique, ou la Théorie Prosodique (Hayes, 1989 ; Nespor & Vogel, 1986 ; Selkirk, 1980 ; 1984 ; 1986), propose une représentation phonologique sous la forme d'une structure arborescente hiérarchisée, composée de différents niveaux de constituance ou de domaines prosodiques. Il s'agit plus précisément d'une théorie où la structure prosodique est considérée comme une organisation abstraite des unités, propre à chaque langue, soumise à un certain nombre de contraintes métriques, tonales et syntaxiques. En effet, dans cette approche, la constituance syntaxique occupe une place centrale dans la construction des domaines phonologiques, tout en restant cependant à deux niveaux d'analyse indépendants.

« A representation (of phonological domains) that is independent of syntactic constituency but related to it by a module of syntactic-prosodic constituency correspondence constraints. » (Selkirk, 2011, p. 436)

Dans sa version standard, la théorie de la constituance prosodique propose une représentation arborescente exhaustive de la hiérarchie prosodique, composée, selon les auteurs, de 7 à 8 niveaux de constituance :

<b>Utt</b>	Enoncé (« <i>Phonological Utterance</i> »)
<b>IP ou ι</b>	Syntagme Intonatif (« <i>Intonational Phrase</i> »)
<b>PP ou φ</b>	Syntagme Phonologique (« <i>Phonological Phrase</i> »)
<b>C</b>	Groupe Clitique (« <i>Clitic Group</i> »)
<b>MP ou ω</b>	Mot Prosodique/Phonologique (« <i>Prosodic/Phonological Word</i> »)
<b>P ou Σ</b>	Pied (« <i>Foot</i> »)
<b>S ou σ</b>	Syllabe (« <i>Syllable</i> »)
<b>μ</b>	More (« <i>Mora</i> »)

Les constituants situés les plus hauts dans la hiérarchie prosodiques (Enoncé et Syntagme Intonatif) sont définis selon des critères sémantico-pragmatiques. Plus précisément, l'Enoncé

est considéré comme l'unité la plus grande dans la hiérarchie prosodique, correspondant à une ou plusieurs phrases syntaxiques, dont la frontière est marquée par une pause finale (Hayes, 1989). Quant au Syntagme Intonatif, ce constituant correspond à un contour intonatif cohérent, marqué notamment par un allongement de la durée de la syllabe finale (Selkirk, 1984). Le découpage de l'Énoncé en Syntagmes Intonatifs dépend par ailleurs de la longueur de l'Énoncé mais aussi du débit de parole.

Concernant, cette fois-ci, les constituants situés les plus bas dans la hiérarchie prosodique, la More, la Syllabe et le Pied sont définis selon des critères métriques. Plus précisément, la More est l'unité la plus petite de la hiérarchie prosodique, particulièrement utilisée dans la description de la langue japonaise, permettant d'indiquer le poids d'une Syllabe : une syllabe lourde comporte plusieurs mores. Dans le cas de certaines langues, comme le latin ou même l'anglais, les syllabes lourdes attirent l'accent (entre autres, Halle & Vergnaud, 1987 ; Prince, 1990). La Syllabe est donc une unité plus grande que la More et se compose d'un noyau vocalique qui peut être précédé d'une attaque et suivi par une coda. L'ensemble formé par le noyau vocalique et la coda est par ailleurs appelé « rime ». La structure infrasyllabique est pertinente dans la caractérisation du type d'accent. En effet, Astésano (2001) montre que les accents initiaux sont caractérisés par un allongement sur l'attaque, tandis que les accents finaux sont signalés par une rime plus allongée. Quant au Pied, il s'agit là de l'unité constitutive de la structure métrique des langues, qui est formé par une syllabe accentuée, précédée ou suivie d'une ou de plusieurs syllabes inaccentuées.

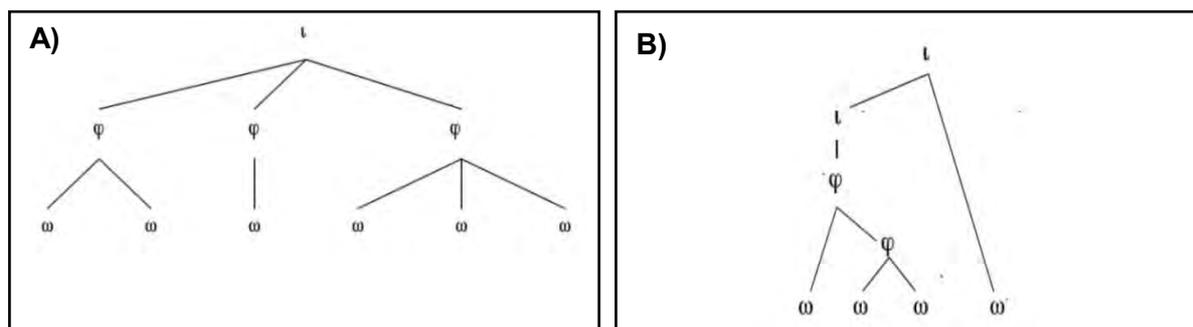
A l'inverse, les constituants situés à l'intermédiaire (Mot Prosodique/Phonologique, Groupe Clitique, Syntagme Phonologique) dépendent davantage de contraintes syntaxiques. Toutefois, seul le Syntagme Phonologique est particulièrement concerné par les contraintes liées à l'interface prosodie-syntaxe. Nous insisterons davantage sur la notion de Mot Prosodique/Phonologique et de Syntagme Phonologique dans les parties suivantes, étant donné leurs statuts particuliers (cf. 1.1.1.3 et 1.1.1.4 du Chapitre 2). En effet, nous verrons plus en détails les règles de construction de ces deux niveaux de constituance, qui dans certains cas se révèlent être peu consensuelles. Dans le cadre de notre travail, nous nous intéresserons plus spécifiquement au niveau du Mot Prosodique, dont la pertinence est aujourd'hui encore discutée dans les modèles du phrasé prosodique en français (cf. Chapitre 3). Concernant le niveau du Groupe Clitique, il est intéressant de préciser que ce niveau de constituance n'est pas pris en compte par tous les auteurs : seuls Nespor & Vogel (1986) et Hayes (1989) intègrent ce constituant entre le Mot Prosodique et le Syntagme Phonologique, dans le but de rendre compte du regroupement du mot lexical avec le ou les mots grammaticaux attachés (pronoms, déterminants, prépositions, etc.).

### 1.1.1.1 The Strict Layer Hypothesis (SLH)

Comme nous l'avons déjà mentionné, ces différents constituants prosodiques s'inscrivent dans une structure hiérarchique dont l'organisation est assurée dans la Théorie Prosodique par la règle de l'étagement strict (« *Strict Layer Hypothesis* » ou *SLH*). Alors que nous verrons par la suite les révisions majeures apportées à la *SLH*, dans sa version initiale, cependant, la règle de l'étagement strict est définie sous la forme de deux règles précises.

« a) A given non terminal unit of the prosodic hierarchy,  $X^p$ , is composed of one or more units of the immediately lower category,  $X^{p-1}$ . b) A unit of a given level of the hierarchy is exhaustively contained in the superordinate unit of which it is a part. » (Nespor & Vogel, 1986, p. 7)

Autrement dit, la structure prosodique est considérée ici comme une structure non réursive, où les constituants prosodiques sont nécessairement composés d'un ou de plusieurs constituants d'un niveau immédiatement inférieur dans la hiérarchie (un Syntagme Intonatif est composé d'un ou plusieurs Syntagmes Phonologiques), et par conséquent où chaque constituant de niveau donné est entièrement contenu dans le constituant de niveau supérieur dont il fait partie (un Syntagme Intonatif ne peut pas être contenu dans un autre Syntagme Intonatif). Afin d'illustrer cette règle de hiérarchie stricte, nous proposons la Figure N°1 issue de Selkirk (2011, pp. 437-438).



**Figure 1 :** Règle de l'étagement strict (*SLH*)

Structure prosodique respectant la règle de l'étagement strict en A) et structure prosodique ne respectant pas la règle de l'étagement strict en B) (issu de Selkirk, 2011, pp. 437-438).

Dans la Figure (A) N°1, la structure prosodique respecte la règle de l'étagement strict, où un constituant de niveau donné domine immédiatement un autre constituant de niveau

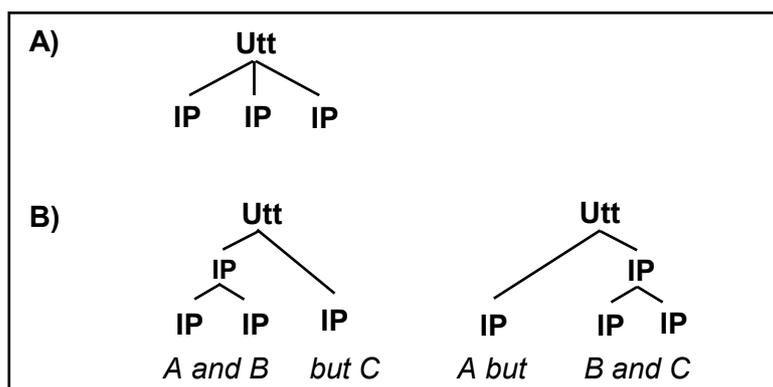
systématiquement inférieur ( $\nu\varphi$  et  $\varphi/\omega$ ), tandis que dans la Figure (B) N°1, la structure prosodique ne respecte pas la règle de l'étagement strict puisqu'un constituant d'un niveau donné domine immédiatement un constituant de même niveau dans la hiérarchie ( $\nu/\iota$  ou  $\varphi/\varphi$ ). La règle de la *SLH* suppose donc qu'à chaque niveau d'enchâssement de la structure prosodique, correspond un domaine phonologique spécifique, et de ce fait, le degré d'enchâssement dans l'arbre reste fixe.

### 1.1.1.2 Révision de la SLH dans le cadre de la Théorie de l'Optimalité

La notion de profondeur fixe ou déterminée dans l'arbre prosodique a d'ailleurs été discutée et réinterprétée, plus tard, par certains auteurs (Ladd, 1986 ; 1996 / 2008 ; Ladd & Cambell, 1991). En effet, Ladd montre que deux syntagmes de même niveau, et plus particulièrement deux Syntagmes Intonatifs, peuvent être marqués par différentes propriétés acoustiques révélant alors des variations dans la force des frontières. Plus précisément, l'auteur a effectué une étude en utilisant des phrases du type « *A and B but C* » et « *A but B and C* », où A, B et C sont des Syntagmes Intonatifs avec une structure syntaxique et accentuelle similaire (Ladd, 1986 puis repris en 1996 / 2008, p. 295) :

- a) « *Warren is a stronger campaigner, and Ryan has more popular policies, but Allen has a lot more money.* »
- b) « *Warren is a stronger campaigner, but Ryan has more popular policies, and Allen has a lot more money.* »

Selon Ladd (1986), la frontière après « *but* » serait plus forte que celle réalisée après « *and* ». En faisant lire ces phrases à 4 locuteurs, l'auteur constate effectivement des propriétés acoustiques allant dans ce sens : le pic de *f0* réalisé à l'initiale des Syntagmes Intonatifs « B » et « C » est plus important à la frontière de « *but* » qu'à la frontière de « *and* ». La pause précédant « *but* » est également nettement plus longue que celle précédant « *and* ». Cela tend à admettre effectivement que la frontière après « *but* » est plus forte que celle réalisée après « *and* » : l'auteur met en avant un phénomène de « *resetting* » en fin de Syntagme Intonatif, c'est-à-dire une remise à niveau du registre, ou une réinitialisation, qui est nécessaire avant de continuer à produire le discours. Dans ce contexte, l'auteur considère que la structure prosodique est récursive et ne respecte donc pas la règle de la *SLH*, puisqu'un Syntagme Intonatif est contenu dans un autre Syntagme Intonatif (cf. Figure N°2 (A) pour une structure non récursive et (B) pour une structure récursive).



**Figure 2 :** *Phénomène de récursivité prosodique*  
 Cas de structure prosodique non récursive respectant la SLH en A) et cas de structures récursives impliquant des « compound prosodic domains » ne respectant pas la SLH en B) (issu de Ladd, 1996 / 2008, p. 296).

L'auteur parle de « *compound prosodic domains* » (Ladd, 1996 / 2008) ou de « *superdomains* » (Ladd, 1986), c'est-à-dire une unité prosodique d'un niveau donné regroupant un constituant du même niveau : un Syntagme Intonatif contenant un autre Syntagme Intonatif. Par conséquent, Ladd propose non pas d'abandonner la SLH, mais plutôt de la réinterpréter afin qu'elle puisse s'accommoder à ce type de réalisation prosodique. Il s'agit donc de considérer le phénomène de « *compounding* » comme une sorte de récursivité limitée qui ne bloque pas pour autant les principes empiriques de la SLH.

« *If prosodic domains can be compound, then the depth of prosodic trees cannot be determinate, and this is clearly incompatible with the SLH. However, compounding is a limited kind of recursiveness. So long as the indeterminate depth of prosodic trees arise only through compounding and in no other way, then prosodic structure is still crucially different from syntactic structure, because the devices that create indeterminate depth in syntactic trees are richer and more powerful than compounding alone.* » (Ladd, 1996 / 2008, p. 298)

Dans ce même sens, Selkirk & Tateishi (1988), à travers une étude sur le japonais, proposent de diviser le niveau du Syntagme Phonologique en deux niveaux distincts : le Syntagme Phonologique Mineur et le Syntagme Phonologique Majeur. Plus récemment encore, d'autres auteurs ont également rendu compte de la récursivité de la structure prosodique dans différentes langues : la récursivité du Syntagme Intonatif et du Syntagme

Phonologique (entre autres, Gussenhoven (2005) pour l'anglais, et Féry & Schubö (2010) ; Féry & Truckenbrodt (2005) pour l'allemand), ainsi que du Mot Prosodique (entre autres, Zec (2005) pour le Serbe ; Ito et Mester (2007 ; 2009) pour le japonais). Certains auteurs constatent aussi qu'un constituant peut dominer un constituant situé à deux, voire même à trois niveaux inférieurs dans la hiérarchie prosodique : une Syllabe peut être immédiatement dominée par un Mot Prosodique (entre autres, Hayes, 1995 ; Inkelas, 1989 ; Kager, 1993). Dans ce contexte, la règle de l'étagement strict est alors considérée plus comme une tendance forte que comme une règle systématique (Shattuck-Hufnagel & Turk, 1996).

Afin de rendre compte, justement, de ces « exceptions à la règle », certains auteurs proposent de réinterpréter la règle de *SLH* dans le cadre de la Théorie de l'Optimalité (Ito & Mester, 1992 ; Kager, 1999 ; Prince & Smolensky, 1993 ; Selkirk, 1995 ; 1996). Il s'agit plus précisément de formuler une série de contraintes, conçues comme universelles, classées de la moins importante à la plus importante. A partir de là, selon le principe de la Théorie de l'Optimalité, les contraintes peuvent être transgressées sous la pression des autres contraintes classées plus haut dans la hiérarchie, dans le but d'obtenir non pas la structure parfaite respectant toutes les contraintes, mais plutôt la structure la plus optimale possible. Concernant la règle de l'étagement strict, quatre contraintes ont été formulées (adapté de Selkirk, 1996) :

- 1) Etagement: un constituant d'un niveau donné ne domine pas un constituant de niveau supérieur dans la hiérarchie (une Syllabe ne domine pas un Pied)
- 2) Dominance: un constituant d'un niveau donné domine un constituant de niveau inférieur dans la hiérarchie (un Mot Prosodique domine un Pied)
- 3) Exhaustivité: un constituant d'un niveau donné ne peut pas dominer un constituant situé à plus d'un niveau inférieur dans la hiérarchie (un Mot Prosodique ne peut pas immédiatement dominer une Syllabe).
- 4) Non-récurtivité: un constituant d'un niveau donné ne peut pas dominer un constituant de même niveau (un Pied ne peut pas dominer un Pied)

Alors que les deux premières contraintes, Etagement et Dominance, sont universellement respectées dans toutes les représentations phonologiques, les deux dernières contraintes, Exhaustivité et Non-récurtivité, peuvent être transgressées.

### 1.1.1.3 Construction du Syntagme Phonologique à l'interface prosodie-syntaxe

En dehors de ces règles gérant la hiérarchie des constituants prosodiques, des contraintes d'alignement entre les structures prosodiques et syntaxiques sont également envisagées dans la Théorie Prosodique. Comme nous l'avons déjà mentionné, bien que la structure prosodique soit vue comme autonome et indépendante de la structure syntaxique, une mise en correspondance, ou « *mapping* », entre la prosodie et la syntaxe est tout de même prise en compte, notamment lors de la construction du Syntagme Phonologique. En effet, deux types de contraintes de mise en correspondance ont été définies, dans un premier temps, en référence à la théorie X-barre (Chomsky, 1970 ; Jackendoff, 1977) qui conçoit la structure d'un syntagme selon trois types de projections syntaxiques : projection maximale (X'' : par exemple le Syntagme Nominal), projection intermédiaire (X' : par exemple le Groupe Nominal) et projection de la tête (X : par exemple le Nom).

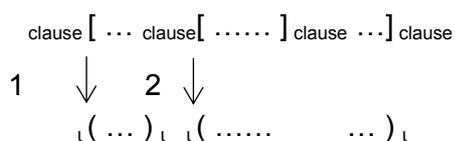
- 1) La contrainte d'alignement, Align-R (XP,  $\varphi$ ) et Align-L (XP,  $\varphi$ ), issue de Selkirk (1986 ; 1996) et dans l'esprit de la Théorie de l'Alignement (McCarthy & Prince, 1993), stipule que le bord droit (Align-R) ou le bord gauche (Align-L) de toute projection syntaxique maximale (X'' ou XP) coïncide avec le bord droit ou gauche d'un Syntagme Phonologique ( $\varphi$ ), ou Syntagme Phonologique Majeur (MaP). Cette contrainte d'alignement est également effective pour le niveau du Mot Prosodique : alignement entre la projection syntaxique intermédiaire (X') et le Mot Prosodique ( $\omega$ ).
- 2) La contrainte de cohésion, Wrap-XP, issue de Truckenbrodt (1999), vise à éviter que les syntagmes syntaxiques se séparent en plusieurs Syntagmes Phonologiques, et stipule donc que chaque projection syntaxique maximale (X'' ou XP) est contenue dans un Syntagme Phonologique ( $\varphi$ ), ou Syntagme Phonologique Majeur (MaP).

Plus tard, Selkirk (2009) propose une nouvelle contrainte qui reprend en partie les deux contraintes précédentes (Align-XP et Wrap-XP) : il s'agit de la contrainte d'adéquation ou « *Match-Theory* ». A l'inverse de la contrainte d'alignement (Align-XP), cette nouvelle contrainte prédit un alignement total entre le syntagme syntaxique et prosodique, plutôt qu'un alignement partiel du bord gauche ou droit du syntagme (Align-R ou Align-L). Par ailleurs, cette contrainte ne s'applique pas seulement au Syntagme Phonologique mais également à d'autres niveaux de la structure prosodique, à savoir au niveau du Syntagme Intonatif et du Mot Prosodique. L'auteur propose donc trois types de contraintes (adapté de Selkirk, 2009 ; 2011) :

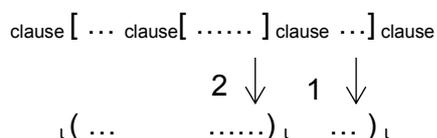
- a. Match Clause : une clause dans une structure en constituant syntaxique doit être alignée au constituant prosodique correspondant, appelé ι (Syntagme Intonatif), dans la représentation phonologique.
- b. Match Phrase : un syntagme dans une structure en constituant syntaxique doit être aligné au constituant prosodique correspondant, appelé φ (Syntagme Phonologique), dans la représentation phonologique.
- c. Match Word : un mot dans une structure en constituant syntaxique doit être aligné au constituant prosodique correspondant, appelé ω (Mot Prosodique), dans la représentation phonologique.

Cette contrainte d'adéquation promet notamment une meilleure prédiction des phénomènes de récursivité. Effectivement, l'auteur montre que seule la contrainte « *Match Theory* » rend compte de la récursivité du Syntagme Intonatif (ι), tandis que les seules contraintes d'alignement ne permettent pas d'indiquer l'enchâssement d'un Syntagme Intonatif (noté 1) dans un autre Syntagme Intonatif (noté 2). En effet, comme nous pouvons le voir dans l'exemple ci-dessous (issu de Selkirk, 2009, p. 31), les contraintes d'alignement (a) et (b) sont partielles, dans la mesure où elles prédisent soit un alignement du bord gauche (a), soit un alignement du bord droit (b), tandis que la règle « *Match Theory* » (c) permet de prédire un alignement total entre les syntagmes, à la fois des bords gauche et droit, et par conséquent de mieux rendre compte de l'enchâssement des constituants.

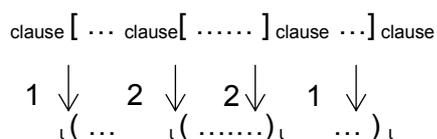
- a. Left-edge alignment of clause and intonational phrase



- b. Right-edge alignment of clause and intonational phrase



- c. Match of clause and intonational phrase



Nous comprenons donc à travers l'ensemble de ces contraintes d'alignement prosodie-syntaxe le peu d'autonomie réservée à la prosodie, et nous pouvons également nous questionner quant à la pertinence d'une telle approche dans l'étude de la parole spontanée. Certaines études sur l'analyse du discours (entre autres, Degand & Simon, 2009 ; Simon 2004), notamment, s'interrogent sur la manière dont la prosodie participe à l'interprétation du discours, mais également au lien de congruence entre la prosodie et la syntaxe dans différents styles de parole. La plupart des travaux sur l'oral spontané proposent plus précisément une analyse macro-syntaxique des énoncés (pour une revue, voir notamment Avanzi, 2007), qui serait plus adaptée à la description de configurations syntaxiques difficiles, typiques de la langue parlée. Dans le cadre de notre travail, nous privilégions également une approche laissant plus d'autonomie à la prosodie sans toutefois écarter les contraintes issues d'autres domaines linguistiques (syntaxe, sémantique, pragmatique, discours, etc.).

### 1.1.1.4 Le cas particulier du Mot Prosodique/Phonologique

Pour terminer la présentation de la Théorie Prosodique, nous choisissons de nous intéresser plus précisément au niveau du Mot Prosodique/Phonologique, central dans notre travail, et qui fera l'objet de discussion tout au long des prochains chapitres. Dans le cadre de la Théorie Prosodique, ce niveau de constituance apparaît également comme le domaine phonologique le moins consensuel de la hiérarchie. Plus précisément, le Mot Prosodique et le Mot Phonologique renvoient à la même entité, c'est-à-dire le domaine de l'accent de mot principal : Selkirk (1986) emploie le terme de Mot Prosodique, tandis que Nespor & Vogel (1986) et Hayes (1989) préfèrent utiliser le terme de Mot Phonologique. En revanche, la définition du domaine diffère d'un modèle à un autre, et notamment selon la manière dont sont envisagés les mots grammaticaux et/ou les mots lexicaux. Effectivement, selon les approches, le Mot Prosodique ou Phonologique correspond à un domaine plus ou moins restreint que le mot lexical : plus grand, plus petit ou équivalent au mot lexical.

Pour Nespor & Vogel (1986) et Hayes (1989), les mots grammaticaux ou les mots lexicaux peuvent chacun former un Mot Prosodique. Le regroupement des deux types de mots est ensuite envisagé à un niveau supérieur dans la hiérarchie prosodique, à savoir au niveau du Groupe Clitique (exemple issu de Hayes, 1989, p. 207) :

- a. « *sing it* » > ((sing)<sub>PW</sub> (it)<sub>PW</sub>)<sub>C</sub>

A l'inverse, Inkelas & Zec (1993) et Selkirk (1996) tendent davantage à considérer que seuls les mots de contenu peuvent former un Mot Prosodique. Plus précisément, dans le cadre de la Théorie de l'Optimalité, Selkirk (1996) montre que les contraintes d'Exhaustivité et de Nonrécursivité peuvent être transgressées, et donc entraîner différentes manières de regrouper les mots lexicaux et les mots grammaticaux. En effet, selon l'auteur, les mots grammaticaux dans leur forme proclitique, du type « *a message* » ou « *to come* », sont considérés comme des « *free clitics* » et n'ont pas besoin d'avoir de statut de Mot Prosodique dans la représentation phonologique (Selkirk, 1996, p. 43). Dans ce cas-là, l'auteur considère que le mot grammatical forme une Syllabe qui est directement dominée par un Syntagme Phonologique Mineur, contenant par ailleurs le Mot Prosodique constitué du mot lexical seulement (exemple adapté de Selkirk, 1996) :

- a. « *a message* » > (a (message)<sub>PW</sub>)<sub>MIP</sub>

Ce rattachement du mot grammatical implique alors la transgression de la contrainte d'Exhaustivité, puisqu'un constituant d'un niveau donné, ici le Syntagme Phonologique Mineur, domine un constituant situé à deux, voire trois niveaux inférieurs dans la hiérarchie, ici la Syllabe formée par le mot grammatical. En revanche, les formes enclitiques des mots grammaticaux, du type « *need him* », tendent à être regroupées dans une structure récursive, où le Mot Prosodique, correspondant au mot lexical, est contenu dans un autre Mot Prosodique plus large comprenant également la forme enclitique (exemple (a) issu de Selkirk, 1996, p. 35). Selkirk parle ici d'« *affixal clitic* » qui transgresse la contrainte de Nonrécursivité, où un constituant d'un niveau donné domine un constituant de même niveau : ici, un Mot Prosodique qui domine un autre Mot Prosodique. Ainsi, l'auteur semble privilégier le recours à la récursivité plutôt qu'à l'accumulation de niveaux de constituance supplémentaires, similaire au Groupe Clitique (Nespor & Vogel, 1986). Toutefois, ce phénomène de récursivité ne vaut que dans le cas de formes enclitiques « réduites » : quand « *need him* » est réalisé « *need 'm* ». Effectivement, lorsque la forme enclitique n'est pas réduite, Selkirk (1996) considère que le clitique « *him* » peut être accentué, et qu'il forme donc un Mot Prosodique à part entière, regroupé avec le mot lexical dans un Syntagme Phonologique Majeur (exemple (b) issu de Selkirk, 1996, p. 34).

- a. « *need him* » forme enclitique réduite > ((need)<sub>PW</sub>him)<sub>PW</sub>  
 b. « *need him* » forme enclitique non-réduite > (((need)<sub>PW</sub>)<sub>MIP</sub> ((him)<sub>PW</sub>)<sub>MIP</sub>)<sub>MaP</sub>

Enfin, concernant plus précisément le statut des affixes à l'intérieur d'un mot lexical, alors que le cas des suffixes est relativement clair parmi les auteurs, celui des préfixes, en revanche, est davantage discuté. En effet, pour la plupart des auteurs, le radical et le suffixe forment un seul et même Mot Prosodique. A l'inverse, pour le préfixe, Nespor & Vogel (1986) considèrent qu'il peut constituer un Mot Prosodique à part entière, mais dans le cas uniquement où le préfixe se termine par une voyelle. Lorsque le préfixe se termine par une consonne, en revanche, les auteurs estiment alors que le préfixe et le mot lexical forment ensemble un seul et même Mot Prosodique (exemples (a) et (b) de constructions lexicales italiennes issus de Nespor & Vogel, 1986, p. 129).

- a. « *asocial* » > (a)<sub>PW</sub> (sociale)<sub>PW</sub>
- b. « *disonesto* » > (disonesto)<sub>PW</sub>

De la même manière, selon Hannahs (1995), un préfixe peut correspondre à un Mot Prosodique mais uniquement dans le cas de préfixes dits « productifs ». Effectivement, les préfixes dits « non productifs », où est observé un phénomène d'assimilation phonétique de la consonne finale du préfixe devant la consonne suivante, à l'initiale du radical, forment avec le radical un seul et même Mot Prosodique (exemples (a), (b) et (c) de constructions lexicales françaises issus de Hannahs, 1995, p. 1134) :

- a. « *imbattable* » > ((im)<sub>PW</sub> (battable)<sub>PW</sub>)
- b. « *illégal* » > (illégal)<sub>PW</sub>
- c. « *immoral* » > (immoral)<sub>PW</sub>

A l'inverse, Peperkamp (1999) considère que le préfixe ne constitue pas un Mot Prosodique à lui-seul, puisqu'il ne peut pas porter l'accent de mot principal. Dans ce cas-là, l'auteur propose que le préfixe attaché au mot lexical ne s'incorpore pas, mais plutôt s'adjoigne à un Mot Prosodique plus large contenant un autre Mot Prosodique, correspondant au mot lexical, ainsi que l'élément préfixe. Cette proposition est par ailleurs dans la lignée des propositions faites par Selkirk (1996) avec les enclitiques (exemple (a) d'une construction lexicale espagnole issu de Peperkamp, 1999, p. 3) :

- a. « *inestabilidad* » > (in (estabilidad)<sub>PW</sub>)<sub>PW</sub>

Le statut particulier du Mot Prosodique semble donc dépendre davantage de questionnements morphologiques, voire phonologiques, que purement syntaxiques. Nous constatons également que le critère prosodique n'intervient pas dans la construction des Mots Prosodiques. En effet, l'argument de Hannahs (1995), notamment, sur le rattachement des préfixes « productifs » ou « non productifs » est purement phonologique, alors que les deux types de préfixes peuvent tout à fait porter un accent d'un point de vue prosodique. D'autres auteurs privilégient d'ailleurs une analyse des clitiques et des affixes, plus directement liée au poids métrique de ces éléments. Delais-Roussarie (1999), qui étudie plus spécifiquement le français, considère en effet que les proclitiques ont un comportement différent des enclitiques, étant donné la différence de poids métrique associé aux syllabes. L'auteur tend à montrer plus précisément que les enclitiques ont un poids métrique supérieur aux proclitiques, ce qui signifie qu'un mot grammatical dans sa forme enclitique peut davantage porter un accent qu'un mot grammatical dans sa forme proclitique. Delais-Roussarie (1999, pp. 21-22) montre effectivement qu'un clitique tel que « *le* » serait accentué dans sa forme enclitique comme dans « *regarde-le* », et non accentué dans sa forme proclitique comme dans « *Paul le reverra demain* », où l'accent tombe plutôt sur le préfixe « *re-verra* ». L'auteur estime d'ailleurs que les affixes sont analysés comme internes au Mot Prosodique, tandis que les clitiques ne peuvent pas l'être. Cela suppose alors que les mots grammaticaux soient considérés par l'auteur comme des éléments libres, voire comme des Mots Prosodiques pour les formes enclitiques, alors que les affixes sont directement inclus avec le mot lexical dans un seul et même Mot Prosodique.

Selon la démarche adoptée, nous constatons donc différentes propositions quant à la manière dont le Mot Prosodique est constitué et rattaché dans la hiérarchie des constituants. Cette discussion sur la structure du Mot Prosodique, mais également sur sa pertinence en tant que niveau de constituance sera par ailleurs reprise dans le cadre de l'étude du phrasé prosodique en français.

### 1.2 Approche plus « purement » prosodique

En parallèle à la Théorie Prosodique, où comme nous l'avons vu, les préoccupations morphologiques et syntaxiques dominent, d'autres approches privilégient cette fois-ci des critères essentiellement prosodiques (intonation, métrique et rythme) dans l'élaboration de la hiérarchie des constituants. En effet, bien que la Théorie Prosodique se fonde également sur une organisation métrique abstraite pour définir certains des constituants prosodiques (More, Syllabe, Pied), nous voyons également que d'autres constituants sont essentiellement

construits à partir d'informations syntaxiques (Mot Prosodique et Syntagme Phonologique), au détriment parfois des contraintes prosodiques. Ainsi, toujours dans la lignée de la représentation phonologique plurilinéaire, où les phénomènes prosodiques sont envisagés comme des représentations autonomes, certains auteurs s'inscrivent dans un paradigme métrique ou tonal, envisageant l'organisation de la structure en constituants prosodiques comme uniquement conditionnée par le phénomène accentuel pour la première approche, et par le phénomène intonatif pour la seconde approche. Par ailleurs, alors que la Théorie Prosodique envisage une structure hiérarchique exhaustive composée de 7 à 8 niveaux de constituance, dans les approches plus « purement » prosodiques, en revanche, les structures prosodiques ne dépassent pas 3 niveaux de constituance.

### 1.2.1 La Théorie Métrique

Nous commençons tout d'abord par la Théorie Métrique (Halle & Vergnaud, 1987 ; Hayes, 1995 ; Liberman, 1975 ; Liberman & Prince 1977), dont l'objectif est de définir la structure des constituants prosodiques essentiellement sur la base de l'organisation hiérarchique des proéminences. En effet, selon cette approche les proéminences contribuent à la construction de la structure prosodique des énoncés, et donc à l'élaboration des niveaux de constituance : chaque catégorie de groupement est marquée par une syllabe désignée comme la plus proéminente, constituant sa tête, qui est précédée et/ou suivi de syllabes désignées comme métriquement faibles ou non proéminentes.

La notion de tête métrique renvoie par ailleurs à la notion de Pied qui constitue l'élément fondamental de la structure métrique des langues. Effectivement, nous rappelons que le Pied est l'unité prosodique formée par une syllabe accentuée et par un nombre  $n$  de syllabes inaccentuées. Dans le cadre de la Phonologie Métrique, le Pied est donc l'unité de base signalant la tête métrique à l'origine de la formation des constituants prosodiques. Selon la langue, la position de la tête dans le Pied est différente (« *foot dominance* » : Kager, 1995 ; Liberman, 1975) : un pied trochaïque (tête à gauche) lorsque la syllabe initiale du pied est métriquement plus forte que les autres syllabes, et un pied iambique (tête à droite) lorsque c'est la syllabe finale du pied qui est métriquement plus forte que les autres syllabes. Par ailleurs, définir les syllabes métriquement fortes et métriquement faibles implique une mise en relation des différentes syllabes sur le plan syntagmatique. Autrement dit, dans la Phonologie Métrique, c'est la notion de force ou de poids métrique relatif qui est importante : une syllabe est considérée comme métriquement forte uniquement dans sa relation avec une syllabe métriquement faible (Liberman, 1975).

« The notions « strong » and « weak » which serve as node labels in metrical patterns are essentially relational. A strong position is strong only by virtue of being associated with a corresponding weak position, not because of any inherent property of strongness; a weak position can be considered weak only because, in some metrical pattern, it plays the role of weak in relation to some strong ». (Lieberman, 1975, p. 51)

Il ne s'agit donc pas d'interpréter le degré de proéminence sur chaque syllabe de manière isolée, mais plutôt de mettre en relation les syllabes pour faire ressortir celles qui sont métriquement plus fortes que les autres (en référence à une opération de comptage des syllabes et du calcul de la position de l'accent : Di Cristo, 2004). En revanche, il existe également des syllabes dites « *extramétriques* » (Hayes, 1982 ; Liberman & Prince, 1977) qui ne se soumettent pas aux règles métriques de la langue et qui ne sont pas impliquées dans la construction de la structure métrique. Autrement dit, ces syllabes extramétriques ne peuvent pas être accentuées. Ce sont en général les éléments périphériques qui sont considérés comme extramétriques, à savoir certaines syllabes initiales, notamment les syllabes initiales sans attaque (Plénat, 1994), et certaines syllabes finales contenant un « *schwa* ».

Sur la base de ces fondements théoriques, Beckman & Edwards (1990) ont proposé une approche, nommée « *Prominence-based Theory of Prosodic Structure* », où la structure prosodique est composée de trois niveaux de constituance : Pied, Syntagme Accentuel et Syntagme Intonatif. Selon les auteurs, chacun de ces niveaux de constituance est marqué par un niveau d'accentuation spécifique : la tête métrique au niveau du Pied devient un accent mélodique (ou « *pitch accent* ») au niveau du Syntagme Accentuel, qui est projeté à son tour sous la forme d'un accent nucléaire au niveau du Syntagme Intonatif. Alors que le « *pitch accent* », dont l'appellation est issue des travaux de Bolinger (1951 ; 1958 ; 1972), renvoie à la syllabe portant l'accent lexical ou l'accent de mot, l'accent nucléaire, en revanche, est l'accent pourvu du plus fort degré de proéminence positionné en fin de syntagme ou en fin de constituant prosodique majeur. Dans ce contexte, chaque niveau de constituance est formé à partir d'un type d'accent spécifique.

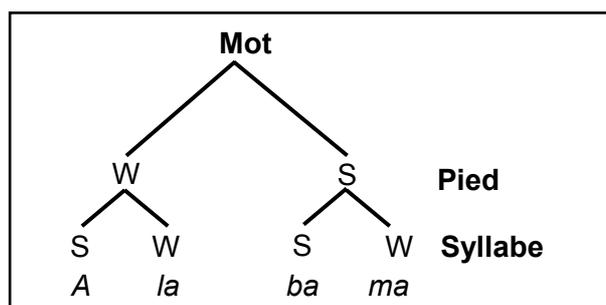
Enfin, il est important de préciser que l'accentuation est conçue, dans la Phonologie Métrique, comme étant la manifestation linguistique de la structure rythmique (Hayes, 1995). Effectivement, le rythme permet d'organiser hiérarchiquement les proéminences : il s'agit du principe d'eurythmie. Plus précisément, ce principe vise à favoriser une alternance régulière de syllabes accentuées et inaccentuées (aussi appelé « *isochronie accentuelle* »), et ainsi éviter des phénomènes de « *clash accentuel* », lorsque plusieurs syllabes accentuées sont

contiguës, et de « *vide accentuel* », lorsqu'à l'inverse, plusieurs syllabes inaccentuées sont contiguës. Il est également intéressant de préciser que le principe d'eurythmie est directement lié aux niveaux de constituance, puisqu'il ne s'applique pas sur des syllabes accentuées appartenant à deux niveaux de constituance différents.

Pour modéliser l'alternance rythmique et la relation hiérarchique entre les proéminences, deux modes de représentation sont par ailleurs conçus dans la Théorie Métrique. Alors que le premier type de représentation consiste à représenter la structure hiérarchique des niveaux de proéminences sous une forme arborescente, comparable aux arbres utilisés en syntaxe, le deuxième type de représentation, en revanche, conçoit l'organisation hiérarchique des proéminences sous la forme d'une grille composée de colonnes d'astérisques. Alors que l'arbre métrique permettrait une meilleure représentation de la relation hiérarchique entre les proéminences, la grille métrique, en revanche, faciliterait la représentation de l'organisation rythmique de l'énoncé, en rendant compte notamment de la force des proéminences.

### 1.2.1.1 L'arbre métrique

Commençons tout d'abord par la représentation en arbre métrique (Lieberman, 1975 ; Lieberman & Prince, 1977), composée de branchements binaires labélisés « *s* » pour « *strong* » (syllabe métriquement forte) et « *w* » pour « *weak* » (syllabe métriquement faible), dont l'objectif est de rendre compte de la relation entre les proéminences associées à chaque syllabe d'un énoncé. Comme nous pouvons le voir dans la Figure N°3, le niveau d'enchâssement le plus profond dans l'arbre métrique indique les syllabes métriquement plus fortes que les autres : les syllabes labélisées « *s* » sont considérées comme métriquement plus fortes que les syllabes labélisées « *w* ». Dans le cas de notre exemple, le mot « *Alabama* » est composé de 2 syllabes métriquement fortes (« *A* » et « *ba* ») et de deux syllabes métriquement faibles (« *la* » et « *ma* »). Parmi les deux syllabes considérées comme métriquement fortes, seule une syllabe sera ensuite projetée au niveau d'enchâssement supérieur pour former ainsi la tête métrique du Pied. La syllabe « *ba* » porte donc ici l'accent principal ou l'accent primaire, puisqu'elle est métriquement plus forte que la syllabe « *A* » qui porte alors l'accent secondaire.



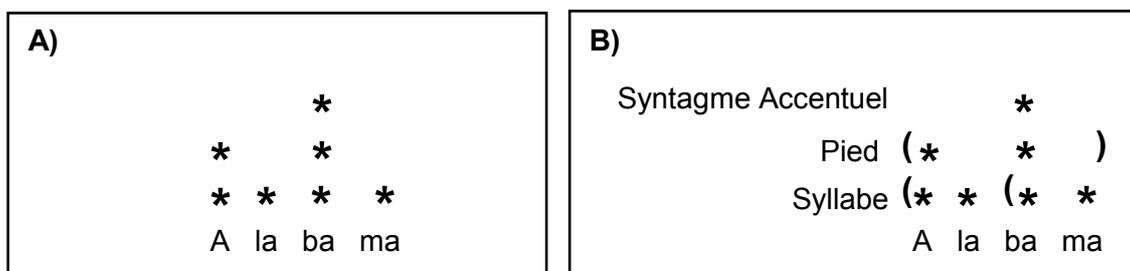
**Figure 3** : Arbres métriques associés au mot "Alabama" (issu et adapté de Kager, 1995, p. 368)

Également, comme nous l'avons mentionné plus haut, l'organisation des proéminences répond à certaines contraintes rythmiques (règles de « *clash accentuel* » et de « *vide accentuel* »), permettant de favoriser une alternance régulière de syllabes accentuées et non accentuées. Ainsi, pour éviter un cas de « *clash accentuel* », il convient d'appliquer ce qu'on appelle « *la règle des thirteen men* » visant à réorganiser la structure métrique, en inversant par exemple le pattern accentuel. Il s'agit plus précisément d'une « *stress-shift rule* » qui consiste à déplacer l'accent sur un nœud « sœur » dans l'arbre métrique (Al-Mozainy et al., 1985) : la séquence [w s s] sur « *thirteen men* » devient [s w s] en déplaçant l'accent de la syllabe finale « *teen* » sur la syllabe initiale « *thir* », donnant ainsi « *thirteen men* ».

### 1.2.1.2 La grille métrique simple et parenthésée

Suite à la représentation métrique arborescente, principalement basée sur les réseaux de relations entre les proéminences, un autre type de représentation a été proposé : il s'agit de la grille métrique simple (Lieberman & Prince, 1977) ou parenthésée (Halle & Vergnaud, 1987). En effet, ce nouveau mode de représentation permet de prendre en considération le caractère hiérarchique des proéminences, tout en rendant compte également du décours temporel et donc de la structure rythmique. Plus précisément, dans sa version simple, la grille métrique se compose de deux axes : un axe horizontal représentant l'organisation rythmique ou l'alternance des syllabes accentuées et inaccentuées, ainsi qu'un axe vertical, où la force des proéminences est indiquée à travers la superposition des marques étoilées (la rangée composée du plus grand nombre d'étoiles correspond à la syllabe la plus forte). Comme nous pouvons le voir dans la Figure (A) N° 4, la grille métrique du mot « *Alabama* » indique sur l'axe horizontal une alternance régulière de syllabes accentuées (« *A* » et « *ba* ») et inaccentuées (« *la* » et « *ma* »), et sur l'axe vertical, une prédominance de l'accent réalisé sur « *ba* » par rapport à celui réalisé sur « *A* »

(superposition d'étoiles plus importante sur « *ba* » que sur « *A* »). Une autre version de la grille métrique simple est aussi envisagée : il s'agit de la grille métrique parenthésée (Halle & Vergnaud, 1987). Dans le cadre de cette représentation métrique, l'organisation des proéminences est directement mise en correspondance avec les niveaux de constituance (cf. Figure (B) N°4). Plus précisément, à chaque strate de la grille correspond un niveau de constituance spécifique : la première ligne est associée au niveau de la syllabe, où des parenthèses indiquent les syllabes faisant partie du même Pied métrique (« *A – la* » pour le premier Pied et « *ba – ma* » pour le second Pied), puis chaque Pied projette sa tête métrique sur la seconde ligne (« *A* » et « *ba* ») où les parenthèses indiquent le début et la fin du Syntagme Accentuel (pour reprendre la terminologie de Beckman & Edwards, 1990), dont l'accent primaire, correspondant à l'accent mélodique (« *ba* »), est lui-même projeté au niveau supérieur.



**Figure 4 :** Grille métrique simple en A) et parenthésée en B) associée au mot "Alabama" (issu et adapté de Kager, 1995, p. 369)

Ces deux modes de représentation ont donc largement été utilisés dans toutes les approches présentées dans le cadre de ce chapitre. Nous constatons toutefois que la Théorie Prosodique privilégie la représentation en arbre métrique, tandis que les approches plus « purement » prosodiques semblent majoritairement adopter la grille métrique, et ce notamment dans la Théorie Métrique Autosegmentale.

### 1.2.2 La Théorie Métrique et Autosegmentale

Alors que la Théorie Métrique se fonde essentiellement sur l'organisation accentuelle, la Théorie Métrique Autosegmentale (Beckman & Pierrehumbert, 1986 ; Pierrehumbert, 1980), quant à elle, se base également sur l'organisation tonale des énoncés. Cette approche théorique, dont l'appellation a été proposée par Ladd (1996 / 2008), est issue d'une double conception de la phonologie prosodique : une approche métrique (Halle &

Vergnaud, 1987 ; Hayes, 1995 ; Liberman, 1975 ; Liberman & Prince 1977) centrée sur l'organisation rythmique des groupements et des proéminences, et une approche autosegmentale (Goldsmith, 1976) davantage orientée vers une théorie des représentations phonétiques et phonologiques des formes sonores, que ce soit sur le plan segmental ou suprasegmental. En effet, l'approche proposée par Goldsmith (1976) permet de gérer les relations entre la réalisation physique de l'intonation et la formalisation de son interprétation. Plus précisément, la Théorie Autosegmentale conçoit plusieurs « tiers » ou « lignes » associées à différents types d'autosegments qui sont alignés avec la « tier » segmentale : « *association lines* » (Goldsmith, 1976). Ces autosegments renvoient aux différents traits phonologiques, ou « *features* », que comprend un événement articulatoire, tels que les tons, les accents, les segments vocaliques ou consonantiques, la nasalité, etc.

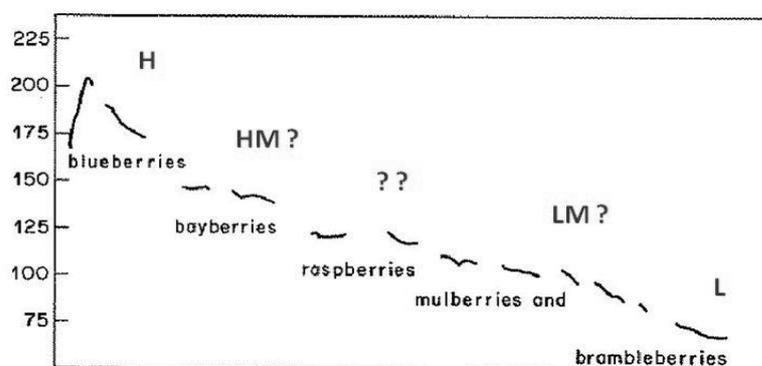
« *Just why are there autosegmental tiers? [...] The act of speech is an orchestral production, a harmonized score of several independent musical "voices". These "voices" of the analogy correspond to the distinctive features of phonology. At the most superficial level, the speech signal is broken down into a large number of independent linear parts – autosegmental tiers- with at least as many of these tiers as there are independent articulators. Thus there will minimally be such a tier for the velum, for the laryngeal gesture corresponding to pitch, and so forth. »*  
(Goldsmith, 1976, p. 158)

Ce concept d'alignement de traits phonologiques a donc été repris dans la Théorie Métrique Autosegmentale (Pierrehumbert, 1980), où il s'agit plus précisément d'aligner les tons aux syllabes métriquement fortes d'un énoncé (accents et frontières). Selon cette approche, une représentation phonologique de l'intonation satisfaisante est une représentation métrique du texte, avec des tons alignés en accord avec certaines règles. L'intonation est donc conçue ici dans une dimension compositionnelle (à distinguer de « l'approche holistique » cf. 1.3.1 du Chapitre 1), où tous les éléments constitutifs du contour intonatif, tels que les indices de proéminence et de frontière, ont une action autonome : accents et frontières constituent des morphèmes à part entière s'organisant dans des unités plus larges. La Phonologie Métrique Autosegmentale s'est nettement imposée dans le paysage prosodique actuel et incarne, encore aujourd'hui, la référence majeure pour l'analyse du phrasé prosodique.

### 1.2.2.1 Le choix d'une analyse à deux niveaux tonals

L'un des premiers points importants soulevés par Pierrehumbert (1980), dans son modèle, a été le choix concernant l'inventaire tonal privilégié pour la description de l'intonation de l'anglais. En effet, les recherches sur l'intonation ont fait naître, chez un certain nombre d'auteurs, une controverse quant à la manière dont le contour intonatif pouvait être modélisé. Alors que certains (notamment Liberman, 1975 ; Trager & Smith, 1951) ont envisagé de faire une description de l'intonation basée sur des cibles tonales, positionnées à quatre niveaux distincts (Low, MidLow, MidHigh, High), d'autres (notamment Bolinger, 1958), en revanche, ont préféré concevoir l'intonation comme des trajectoires mélodiques (montant, descendant, etc.).

En effet, Bolinger (1958) a été le premier à remettre en doute l'analyse en niveaux tonals. Selon lui, les quatre niveaux ne représentent pas des niveaux de hauteur absolus, mais plutôt des niveaux relatifs variant en fonction de l'étendue tonale disponible du locuteur (c'est ce qu'on appelle le « *pitch range* »). L'auteur considère donc qu'il y a une ambiguïté, dans la mesure où les quatre niveaux confondent réalisation tonale et sélection dans le « *pitch range* ». En effet, selon le registre tonal du locuteur, une même configuration mélodique peut être plus ou moins augmentée et plus ou moins diminuée. Pierrehumbert (1980 ; 2000) soulève également un autre inconvénient avec l'analyse à quatre niveaux, à travers l'exemple du « *stepping pattern* », que l'on pourrait traduire par « *pattern en escalier* » (cf. Figure N°5).



**Figure 5 :** Exemple d'un *stepping pattern* réalisé sur l'énoncé « *blueberries, bayberries, raspberries, mulberries and brambleberries* »

Pour chaque « *step* » un niveau tonal spécifique est associé : les « ?? » indiquent les « *steps* » pour lesquels il n'existe pas de formalisme tonal approprié (issu de Pierrehumbert, 2000, p. 16)

D'après l'auteur, la théorie à quatre niveaux n'est pas en mesure de décrire ce type de pattern. Effectivement, dès lors que l'on dépasse quatre « *steps* » dans un « *stepping pattern* », la théorie à quatre niveaux n'a plus le matériel nécessaire pour proposer la description du/des « *steps* » restants. Pour ce type de réalisation prosodique, Pierrehumbert (1980) recourt à la notion de registre tonal, et parle d'un phénomène de « *downstep* » qui consiste en la réalisation de « *steps* » successifs plus bas que le ton précédent. Selon l'auteur, une même configuration tonale est « *downstepped* » quand il y a un abaissement progressif du registre mélodique (ou « *pitch range* ») (cf. 1.2.2.2.1 du Chapitre 2 pour plus de précisions sur la notion de « *downstep* »). L'analyse en termes de configurations présente également quelques inconvénients, notamment concernant la question de l'alignement temporel avec la « tier » segmentale. Effectivement, une analyse en trajectoires de *f0* n'offre pas assez de précision pour pouvoir associer une cible tonale spécifique avec le contenu segmental. Pour cela, il est nécessaire de décomposer le mouvement mélodique en une série de repères tonals.

Dans ce contexte, Pierrehumbert (1980) choisit d'effectuer une description en niveaux tonals, avec un inventaire réduit à deux niveaux seulement (Low et High), tout en insistant également sur la notion de « *pitch range* » pour pouvoir ainsi d'une part, réduire l'ambiguïté laissée par les quatre niveaux (Bolinger, 1958), et d'autre part, expliquer les variabilités mélodiques des cibles tonales (cf. « *downstep* »).

### 1.2.2.2 Le modèle standard

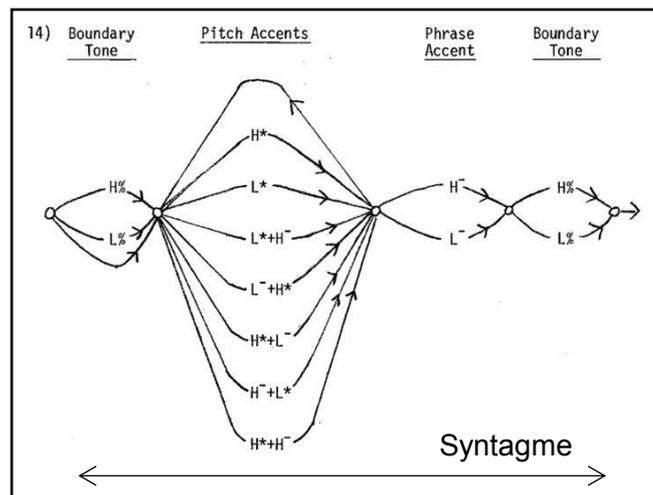
A partir de là, Pierrehumbert (1980), en se basant sur les travaux de Bruce (1977), envisage un modèle d'analyse à trois composants : une grammaire de contours intonatifs qui génère des séquences de tons hauts et bas ; une représentation métrique du texte réalisée à partir d'une grille métrique ; et des règles gérant l'alignement des tons sur le texte (« *text-to-tune association* » Liberman, 1975). Selon Ladd (1996 / 2008), combiner la phonologie métrique à la notion de « *tune-text association* » permet de distinguer les trois caractéristiques souvent confondues de l'accent<sup>2</sup> : la force métrique, la réalisation acoustique concrète et la localisation de l'événement sur le plan tonal.

Dans ce cadre, Pierrehumbert (1980) propose une grammaire dite à état fini, c'est-à-dire une grammaire composée de catégories tonales dont l'inventaire est « fini », et propre à chaque langue (cf. Figure N°6). Cette grammaire se compose de « *edge tones* » (tons

---

<sup>2</sup> Voir 1.1.1 du Chapitre 1 dichotomies « *stress/accent* » chez Bolinger, 1989 ; Beckman, 1986, et « *accent sous-jacent/effectif* » chez Garde, 1968.

consacrés aux bords/frontières du syntagme) et de « *pitch accents* » (accents mélodiques). Plus précisément, les « *edge tones* » comprennent deux types d'éléments tonals : le « *boundary tone* » (un ton de frontière) qui marque le début puis la fin du Syntagme Intonatif, et le « *phrase accent* » qui n'est pas associé à une syllabe particulière, et qui apparaît entre le dernier « *pitch accent* » et le « *boundary tone* ». Selon l'auteur, tout Syntagme Intonatif doit donc contenir un « *boundary tone* » en début de syntagme, un ou plusieurs « *pitch accents* » suivis potentiellement d'un « *phrase accent* », et enfin un deuxième « *boundary tone* » pour clôturer le syntagme. Concernant le formalisme tonal, Pierrehumbert (1980) choisit de suivre le système d'annotation proposé par Goldsmith (1976) : « % » pour signaler le « *boundary tone* », « \* » pour signaler le « *pitch accent* », et « - » pour signaler le ton qui suit ou précède le ton étoilé dans le cas d'un « *pitch accent* » bitonal, ou pour signaler le « *phrase accent* ».

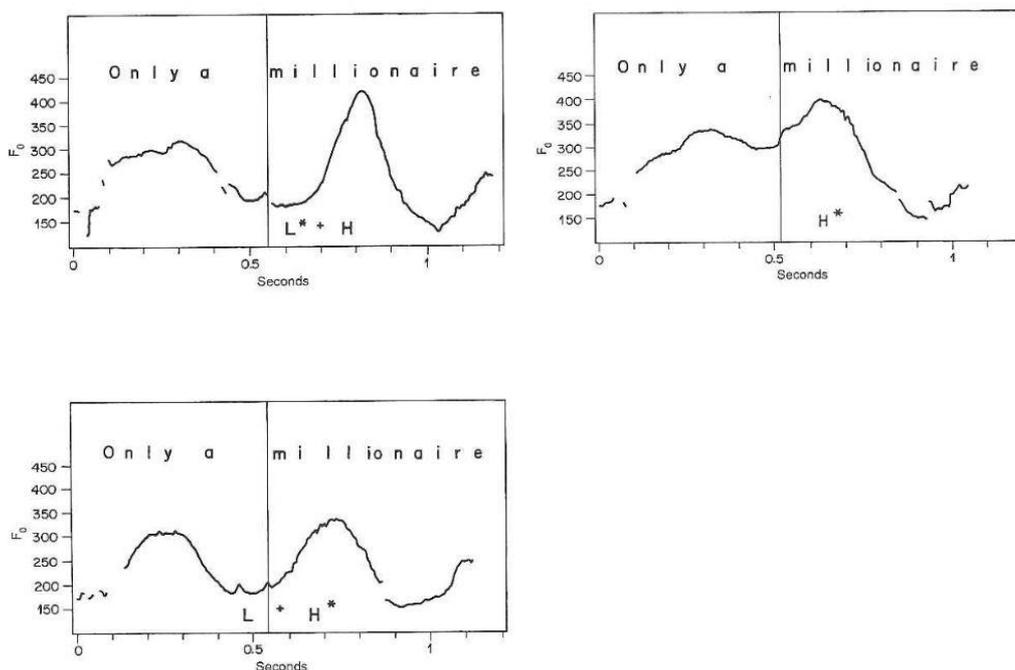


**Figure 6 :** Grammaire à état fini du modèle de Pierrehumbert (1980, p. 29)

Il faut savoir que la notion de « *phrase accent* » a initialement été proposée par Bruce (1977) pour interpréter l'accent en suédois. En effet, l'auteur montre que l'accent en suédois qui est caractérisé par un pic de  $f_0$ , correspondant au « *pitch accent* », est systématiquement suivi par un second pic de  $f_0$ , correspondant au « *phrase accent* », qui est lui-même suivi par une chute ou une montée de  $f_0$ , correspondant enfin au « *boundary tone* ». Dans le modèle de Pierrehumbert, le « *phrase accent* » sert à caractériser, comme en suédois, l'événement tonal réalisé entre le dernier « *pitch accent* » et le « *boundary tone* ». Nous verrons que le statut du « *phrase accent* » sera plus tard réformé, pour constituer le ton de frontière du Syntagme

Intermédiaire (cf. 1.2.2.3 du Chapitre 2), et prendra également une tout autre fonction dans les modèles du phrasé prosodique en français (cf. 2.2.3 du Chapitre 3 dans le modèle de Jun & Fougeron, 2000 ; 2002).

Quant à la notion de « *pitch accent* », nous rappelons que cette terminologie est issue des travaux de Bolinger (1958). En effet, selon l'auteur cet accent mélodique consiste, comme son nom l'indique, en une proéminence accentuelle qui participe aussi à la construction de la configuration tonale, d'où le terme de « *pitch* ». Nous rappelons également que l'auteur distingue « *stress* » et « *accent* » en considérant que le premier terme renvoie à la propriété lexicale abstraite de l'accent potentiel, alors que le second terme correspond, cette fois-ci, à sa réalisation concrète dans l'énoncé (Bolinger, 1989). Le « *pitch accent* » est en ce sens une composante de base du contour intonatif et un signal indispensable à la perception d'une proéminence. Dans le cadre du modèle de Pierrehumbert (1980), il existe sept types de « *pitch accents* » (cf. Figure N°6), réalisés sur les syllabes métriquement fortes, qui correspondent soit à des pics mélodiques (H\*) soit à des creux (L\*), et qui peuvent se composer d'un seul ton (H\* ou L\*) ou de deux tons (par exemple : L\*+H-) (cf. Figure N°7).



**Figure 7 :** Représentation des courbes mélodiques associées à l'énoncé « *only a millionaire* » prononcé avec trois différents accents mélodiques  
 Réalisation de l'accent mélodique bitonal L\*+H- dans la figure en haut à gauche, réalisation de l'accent mélodique bitonal L+H\* dans la figure en bas à gauche, et réalisation de l'accent mélodique simple H\* dans la figure en haut à droite (issu de Pierrehumbert & Steel, 1989, p. 182)

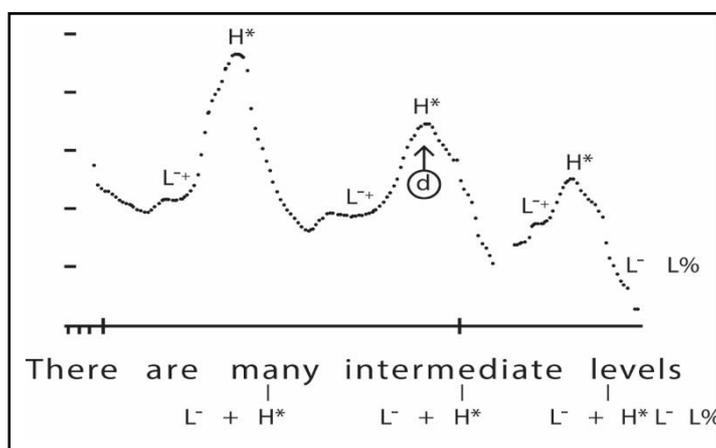
Il s'agit d'accents mélodiques bitonals, lorsque le marquage de la proéminence est associée à deux événements tonals (montée : L\*+H- ou chute : H\*+L-). Dans ce cas-là, selon l'auteur, seul le ton « le plus fort » (Pierrehumbert, 1980, pp. 76-77) porte le diacritique « \* », et l'autre élément tonal, suivant ou précédant le ton étoilé, est marqué par le diacritique « - ». Il est intéressant de préciser que la question de l'identification du ton étoilé dans le cas d'accents bitonals a été soulevée, encore récemment, par certains auteurs (entre autres Arvaniti et *al.*, 2000 ; Grice, 1995 ; Prieto et *al.*, 2005). En effet, la plupart considère la définition de l'accent bitonal proposée par Pierrehumbert (1980) comme étant ambiguë, dans la mesure où il est difficile de dire si le diacritique « \* » indique un alignement phonétique précis entre le trait phonologique et la syllabe accentuée, ou plutôt une association floue et mal établie (Prieto et *al.*, 2005). Dans ce contexte, Arvaniti et *al.* (2000) stipulent que l'utilisation du diacritique « \* » doit être exclusivement consacrée à l'identification du ton aligné à la syllabe accentuée : dans le cas de L+H\*, c'est le H qui est aligné avec la syllabe accentuée, et dans le cas de L\*+H-, c'est le L qui est aligné avec la syllabe accentuée. Par ailleurs, le ton précédant le ton étoilé est un « *leading tone* » (L+H\*), et celui suivant le ton étoilé est un « *trailing tone* » (L\*+H-). Comme nous pouvons le voir dans la Figure N°7, l'énoncé « *only a millionaire* » est produit avec trois types de « *pitch accents* », dont deux bitonals et un simple. Plus précisément, dans le pattern L+H\*, une montée de la courbe de *f0* est réalisée dès le début du mot « *millionaire* », tandis que dans le pattern L\*+H-, le début du mot est marqué par une courbe de *f0* qui forme d'abord un creux pour ensuite remonter, et enfin, dans le pattern H\*, une montée de la courbe de *f0* est également réalisée dès le début du mot « *millionaire* », qui, à l'inverse du pattern L+H\*, démarre à partir d'un registre tonal initialement élevé. Dans ce dernier cas, aucune chute de *f0* ne précède le ton H\*.

Enfin, sur la base de cette grammaire tonale, Pierrehumbert (1980), met en place un certain nombre de règles phonologiques fournissant une valeur phonétique aux tons, et façonnant ainsi le contour mélodique entre ces tons. Les règles phonologiques étant propres à chaque langue, l'auteur en propose quelques-unes pour l'anglais. Nous retrouverons dans la partie consacrée aux modèles du phrasé prosodique en français certaines des règles présentées ici, et notamment la règle de « *downstep* ».

### 1.2.2.2.1 La règle de « *downstep* »

La règle de « *downstep* », tout d'abord, est selon l'auteur un élément fondamental en anglais. D'une manière générale, le phénomène de « *downstep* » peut être présenté comme étant l'abaissement par palier d'un ton « H » ou « L ». Pierrehumbert (1980) considère par

ailleurs que le « *downstep* » est un phénomène local, s'expliquant par la diminution progressive du volume d'air au cours de la production d'un énoncé. Comme nous pouvons le voir dans la Figure N°8, l'énoncé « *There are many intermediate levels* » est prononcé avec une série de trois accents mélodiques bitonaux, dont les deux derniers sont « *downstepped* » : le dernier accent est abaissé par rapport au deuxième accent, qui est lui-même abaissé par rapport au premier accent. Dans la grammaire de Pierrehumbert (1980), les éléments tonaux susceptibles de provoquer le « *downstep* » sont les quatre accents mélodiques bitonaux suivants : «  $H^-+L^*$  » / «  $H^*+L^-$  » / «  $L^*+H^-$  » / «  $L^-+H^*$  ».



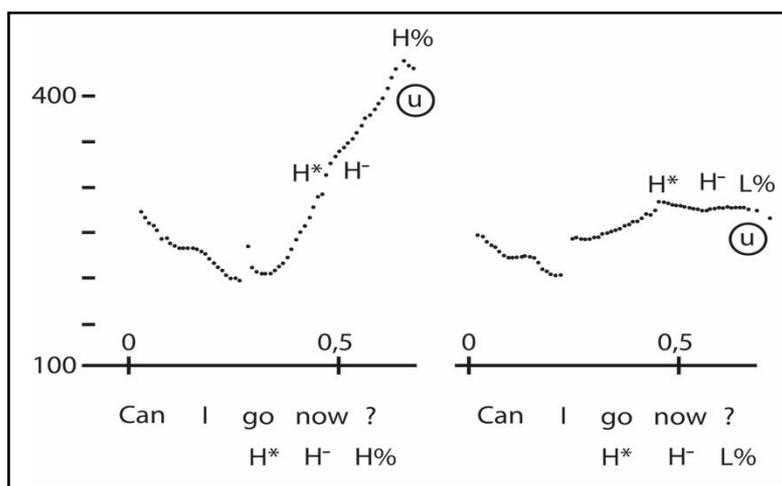
**Figure 8 :** Phénomène de « *downstep* » (noté « *d* ») réalisé sur l'énoncé « *there are many intermediate levels* » (issu de Pierrehumbert, 1980, p. 330)

Ce phénomène, bien que parfaitement intégré dans le modèle standard, a toutefois été remis en question par certains auteurs. En effet, selon Ladd (1983), le « *downstep* » ne peut pas être considéré comme étant simplement une conséquence d'un phénomène phonétique provoqué par certaines séquences de tons, mais plutôt comme un phénomène intonatif indépendant, qui peut apparaître avec tous types de tons. Plus précisément, l'auteur estime que le phénomène de « *downstep* » a une signification particulière, et traduit un sentiment de finalité ou d'achèvement (Ladd, 1996 / 2008). Dans ce contexte, Ladd (1983) suggère d'utiliser le diacritique « ! », déjà employé dans la description des langues tonales africaines, pour marquer les tons « *downstepped* », et ainsi pouvoir considérer le phénomène comme un fait phonologique autonome, qui peut apparaître ou non, et ce indépendamment du type d'accent. Par ailleurs, l'auteur propose la notion de « *downstep emboîté* » (« *nested downstep* » Ladd, 1988 ; 1996 / 2008), en considérant qu'un « *downstep* » peut apparaître aussi bien entre deux tons d'un même syntagme qu'entre deux tons de deux syntagmes

consécutifs. En analysant les phrases du type « *A and B but C* » ou « *A but B and C* », où A, B et C forment trois Syntagmes Intonatifs différents, Ladd (1988) constate effectivement un abaissement de *f0* à l'intérieur de chaque Syntagme Intonatif (A, B et C), mais également entre chaque Syntagme Intonatif : c'est-à-dire un abaissement progressif de la *f0* sur B par rapport à A, et sur C par rapport à B. Ainsi, alors que Pierrehumbert (1980) envisage l'effet de « *downstep* » uniquement à l'intérieur du Syntagme Intonatif, Ladd considère que cet effet peut s'appliquer aux tons appartenant au même domaine phonologique, mais également aux domaines phonologiques où sont inclus ces tons.

### 1.2.2.2.2 La règle d'« *upstep* »

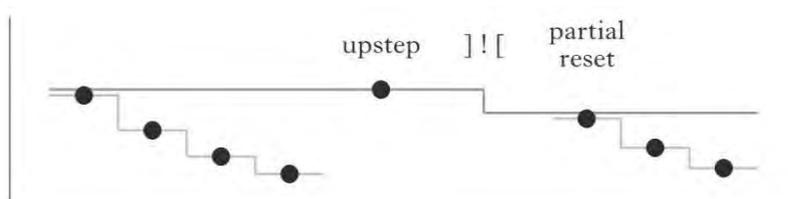
A l'opposé du phénomène de « *downstep* », il existe également, dans la grammaire de Pierrehumbert (1980), une règle d'« *upstep* » qui s'applique cette fois-ci uniquement aux tons de frontière (« *boundary tone* ») réalisés après un « *phrase accent* » haut (« *H-* »). Comme pour la règle précédente, l'« *upstep* » sert à réajuster la valeur phonétique d'un ton dans un contexte particulier. Plus précisément, cette règle consiste à maintenir le même niveau tonal après la réalisation d'un ton « *H-* » (« *phrase accent* »), voire même à augmenter le niveau tonal jusqu'à la fin du Syntagme Intonatif, notamment dans le cas des patterns associés aux « *Yes/No questions* » (cf. Figure N°9).



**Figure 9 :** Phénomène d'« *upstep* » (noté « *u* ») réalisé sur l'énoncé « *Can I go now ?* » (issu de Pierrehumbert, 1980, p. 333)

Clements (1979) estime que l'« *upstep* » peut également renvoyer à un phénomène de réinitialisation du registre tonal, c'est-à-dire un retour à la hauteur initiale du registre

mélodique, notamment après la réalisation d'une série de « *downstep* ». Huang (1985) remarque d'ailleurs que le phénomène d'« *upstep* » apparaît fréquemment en anticipation du « *downstep* ». Plus récemment, Truckenbrodt (2002) considère que ce phénomène de réinitialisation intervient dans la délimitation du Syntagme Intonatif : l'« *upstep* » permet de revenir au registre tonal du début du Syntagme Intonatif, c'est-à-dire avant que le registre ne soit « *downstepped* ». Cependant, selon l'auteur, ce phénomène d'« *upstep* » ne doit pas être confondu avec le phénomène de « *partial reset* » (Ladd, 1986 ; van den Berg et al., 1992) qui apparaît non pas à la fin du Syntagme Intonatif mais plutôt au début d'un nouveau domaine phonologique (cf. Figure N°10).



**Figure 10** : Schéma des phénomènes d'« *upstep* » et de « *partial reset* »  
(issu de van den Berg et al., 1992, puis repris par Truckenbrodt, 2002. p 87)

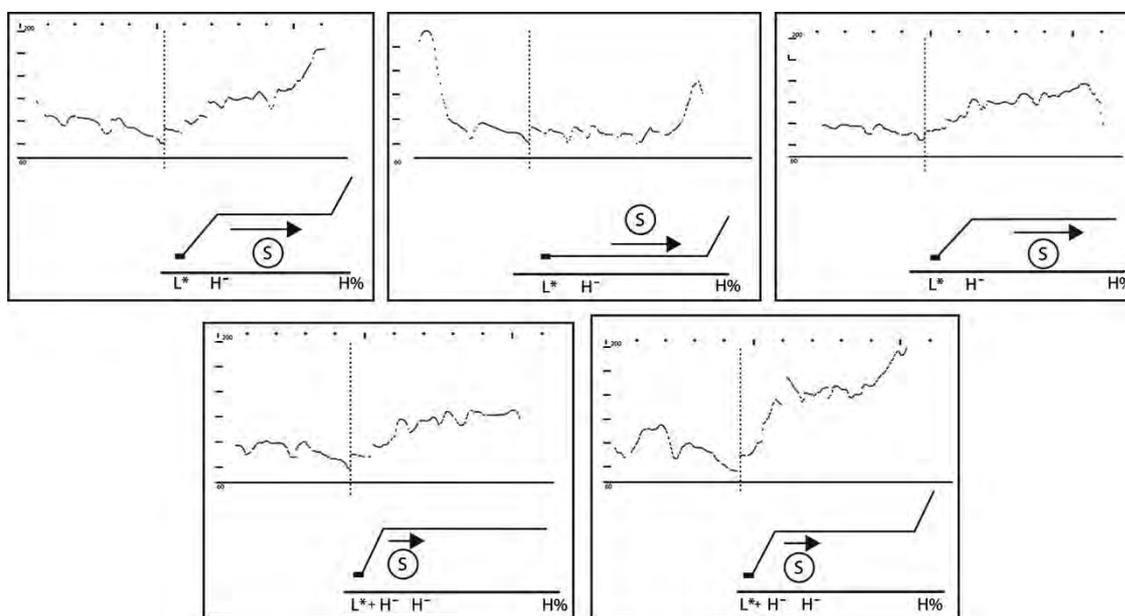
L'« *upstep* » est par ailleurs systématiquement plus haut que le « *partial reset* » suivant, ce qui pourrait correspondre selon Truckenbrodt (2002) à la réalisation d'un autre « *downstep* » : le ton « H » associé au « *partial reset* » est abaissé par rapport au ton « H » associé à l'« *upstep* ». Ce phénomène renvoie également à la notion de « *downstep emboité* » (ou « *nested downstep* » Ladd, 1988 ; 1996 / 2008), où un « *downstep* » local peut être intégré dans un « *downstep* » plus large.

### 1.2.2.2.3 La règle de « *tonal spreading* »

Enfin, directement en lien avec les deux règles précédentes, Pierrehumbert (1980) met en avant un dernier phénomène, nommé « *tonal spreading* », que l'on pourrait traduire par « *propagation tonale* ». Ce phénomène est exclusivement réservé au ton « T- », que ce soit dans le cas du « *pitch accent* » ou du « *phrase accent* ». La règle de « *spreading* » permet au contour mélodique après le « T- » de ne pas présenter une chute ou une montée graduelle, mais plutôt un maintien du niveau tonal (contour en plateau). Plus précisément, le « *tonal spreading* » concerne les tons qui ne sont pas associés à une seule syllabe spécifique mais plutôt à toutes les syllabes se trouvant avant la prochaine cible tonale : les

syllabes partageant le même ton sont donc toutes au même niveau. Ce phénomène est à distinguer des tons étoilés pour lesquels l'alignement est spécifique à une seule syllabe. Selon Pierrehumbert (1980), les patterns donnant lieu à un « *spreading tone* » sont : « H- H% » / « L- H% » / « H- L% » / « L\*+H- H- ».

Comme nous pouvons le voir dans la Figure N°11, pour qu'il y ait une propagation du « T- », il faut que le ton suivant ait une valeur phonétiquement égale ou supérieure. Pour les « T- » en contexte de « *phrase accents* », « H- » se propage jusqu'à « H% » qui est à un niveau plus haut, ou jusqu'à « L% » qui est à un niveau égal après réalisation de la règle d'« *upstep* ». De la même manière, « L- » peut se propager jusqu'à « H% » qui a une valeur phonétique supérieure. Par ailleurs, dans le cas des « T- » en contexte de « *pitch accents* », « H- » se propage dès lors qu'il est suivi par un autre ton haut.



**Figure 11** : Exemples des séquences tonales donnant lieu à la réalisation de la règle de « *tonal spreading* » (noté « s »)

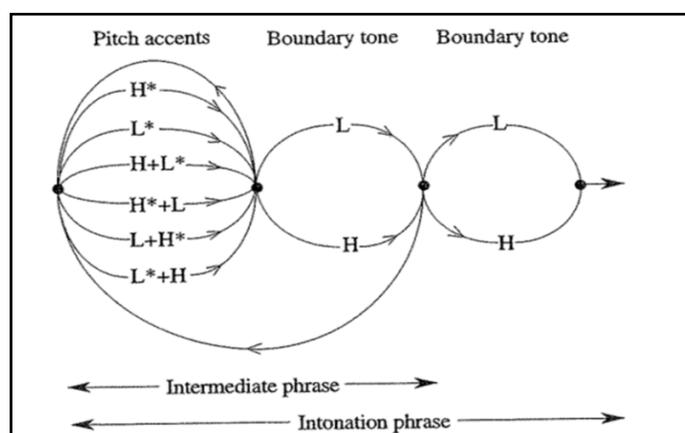
Le même syntagme « *The Uruguayan bulldozer driver's union* » (angl.) a été utilisé pour générer les contours mélodiques : le « *pitch accent* » est à chaque fois réalisé sur « *bulldozer* ». Les trois figures du haut montrent le « *tonal spreading* » pour les « T- » en contexte de « *phrase accent* » et les deux figures du bas montrent le « *tonal spreading* » pour les « T- » en contexte de « *pitch accent* » (issu de Pierrehumbert, 1980, p. 393)

Dans chacun des cas, nous n'observons pas de montées graduelles, mais plutôt des contours en plateaux qui maintiennent la valeur du ton « T- ». Enfin, dans les cas où le ton de frontière a une valeur phonétiquement supérieure (« H- H% »), nous observons, combinée au plateau, une soudaine montée rapide. En revanche, dans le cas des patterns

« L- L% », aucun « *tonal spreading* » n'est possible étant donné que la valeur phonétique de « L% » est systématiquement inférieure à celle de « L- » : il s'agit dans ce cas-là d'une interpolation entre les deux tons, marquée par une chute graduelle de la courbe de *f0*.

### 1.2.2.3 Révision du modèle standard

Par la suite, le modèle de Pierrehumbert (1980) a fait l'objet de deux révisions majeures (Beckman & Pierrehumbert, 1986 ; Liberman & Pierrehumbert, 1984). Dans le cadre de cette partie, nous choisissons de présenter uniquement les dernières révisions, proposées par Beckman & Pierrehumbert (1986), qui ont donné lieu à l'élaboration du système de notation ToBI (cf. 1.2.2.4 du Chapitre 2). Un des points centraux de cette dernière révision du modèle est la réforme du statut du « *phrase accent* » et du « *boundary tone* ». En effet, Beckman & Pierrehumbert (1986) ont choisi d'associer des niveaux de phrasé différents à ces deux éléments tonals, en ajoutant un niveau de constituance supplémentaire plus bas dans la hiérarchie que le Syntagme Intonatif : le niveau du Syntagme Intermédiaire ou « *intermediate phrase* » (noté « *ip* ») (cf. Figure N°12). Ce niveau de constituance correspond au niveau du Syntagme Phonologique dans sa version simple (Selkirk, 1986), ou au niveau du Syntagme Phonologique Majeur dans sa version récursive de la structure prosodique (Selkirk & Tateishi, 1988). Le « *phrase accent* » se voit alors attribuer une valeur démarcative et devient un véritable ton de frontière en marquant le niveau du Syntagme Intermédiaire, et le « *boundary tone* » garde quant à lui son rôle initial en marquant la frontière du Syntagme Intonatif. Cette révision implique également que le « *phrase accent* » n'est pas dans l'obligation d'être suivi par un « *boundary tone* », ce qui dans le modèle initial était nécessaire.



**Figure 12 :** Grammaire à état fini du modèle révisé par Beckman & Pierrehumbert (1986)  
(issu de Pierrehumbert, 2000, p. 22)

Toutefois, l'existence du Syntagme Intermédiaire a été discutée et remise en question par certains auteurs, et notamment par Ladd (1986 ; 1996 / 2008), qui estime qu'il est nécessaire de réduire le nombre de domaines prosodiques. Effectivement, selon l'auteur, en prenant en compte les « *compound prosodic domains* », il est possible d'éliminer la distinction entre le Syntagme Intermédiaire et le Syntagme Intonatif, et ainsi considérer le Syntagme Intermédiaire comme une unité du même type que le Syntagme Intonatif, avec cependant un degré de frontière plus faible (cf. 1.1.1.2 du Chapitre 2 « *révision de la SLH* »). De la même manière, nous verrons que ce niveau de constituance fait encore aujourd'hui l'objet de débat dans les modèles du phrasé prosodique en français (cf. Chapitre 3).

En dehors du Syntagme Intermédiaire, Beckman & Pierrehumbert (1986) ont également réfléchi à une unité prosodique de niveau encore inférieur à celui du Syntagme Intermédiaire, appelée Syntagme Accentuel ou « *Accentual Phrase* » (noté *AP*). Ce dernier serait lié à la réalisation d'un mot lexical et d'un ou de plusieurs mots grammaticaux, dont la frontière serait marquée par un « *pitch accent* ». Ce niveau de constituance peut être relié, selon les modèles, au Syntagme Phonologique Mineur (Selkirk & Tateishi, 1988) ou au Groupe Clitique (Hayes, 1989 ; Nespor & Vogel, 1989). Toutefois, Beckman & Pierrehumbert (1986) ont rendu compte de ce niveau de constituance en japonais, et ne semblent pas tout à fait en accord avec l'idée d'une telle unité en anglais, qui n'apparaît pas nécessaire dans la hiérarchie prosodique de la langue. En parallèle à la question des domaines, s'ajoute la modification de l'inventaire des « *pitch accents* », initialement prévu par Pierrehumbert (1980), et notamment la suppression d'un type d'accent mélodique bitonal : «  $H^*+H$  ». En effet, les auteurs ont rendu compte de la non-distinctivité de cet accent bitonal, et ont préféré considérer cet élément comme étant un accent mélodique simple : «  $H^*$  ».

Enfin, Beckman & Pierrehumbert (1986) ont également choisi de réviser le statut du « *downstep* » en ne le considérant plus comme un fait purement local. Effectivement, les auteurs ont montré que le « *downstep* », désormais réservé au niveau du Syntagme Intermédiaire, est un phénomène affectant l'ensemble du syntagme en contrôlant le « *pitch range* », et pas simplement un phénomène qui touche les tons de manière individuelle. Ainsi, lorsque la règle de « *downstep* » est appliquée, tous les tons suivants, jusqu'à la fin du syntagme, sont affectés. Seule une frontière forte interrompt alors l'effet du « *downstep* ». Par ailleurs, les auteurs ont choisi d'utiliser le terme « *catathesis* » pour ce phénomène, et ainsi de réserver le terme de « *downstep* » au phénomène observé dans les langues tonales africaines.

### 1.2.2.4 Le système de notation ToBI

Sur la base de la Théorie Métrique Autosegmentale (Beckman & Pierrehumbert, 1986 ; Pierrehumbert, 1980), un système de notation, nommé ToBI pour « *Tone and Break Indices* » (Beckman & Hirschberg, 1994 ; Silverman et al., 1992), a été développé. Ce système fournit quatre niveaux parallèles de transcription, sur quatre « tiers » différentes (cf. Figure N°13) : les transcriptions tonale et orthographique sur les deux premières « tiers », l'indication de la force des frontières perçues entre chaque mot (via une échelle d'évaluation allant de 0 à 4) sur la troisième « tier », et une dernière « tier » consacrée aux commentaires divers. Pour le niveau de transcription tonal, les auteurs ont choisi de reprendre le formalisme établi dans le modèle de Pierrehumbert (1980), puis révisé par Beckman & Pierrehumbert (1986). Quant à la « tier » réservée à l'indication de la force des frontières perçues entre chaque mot, le système prévoit 4 degrés de frontière permettant de rendre compte de la structure en constituants prosodiques :

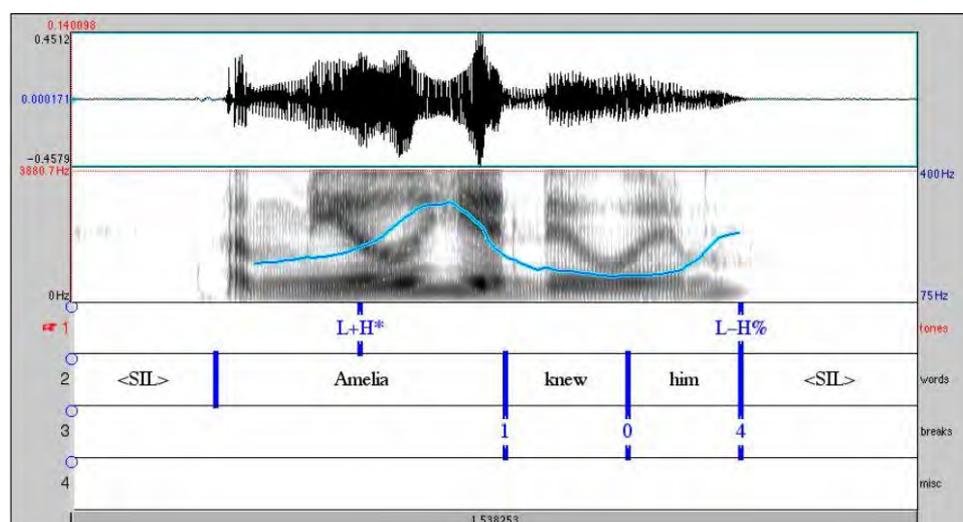
0 = Dans le cas d'une frontière entre un mot lexical et un mot clitique marquée par une trace phonétique particulière qui tend à effacer ou diminuer la force de la frontière.

1 = Frontière de mot lexical

2 = Frontière, sans marque tonale, qui n'est pas associée à un niveau de constituance spécifique, dont la force est moins importante que dans le cas d'une frontière de Syntagme Intermédiaire.

3 = Frontière de Syntagme Intermédiaire

4 = Frontière de Syntagme Intonatif



**Figure 13** : Exemple d'une transcription avec le système ToBI de l'énoncé « Amelia knew him »

(d'après MIT Open CourseWare (mit.edu), 2.5 Pitch Accents with more than one tone : Bi-tonal accents L+H\* and L\*+H, 9 août 2006)

L'objectif de ToBI est donc de fournir une transcription standardisée pouvant s'adapter non seulement à différents styles de parole, mais également à différentes variétés de langues. En effet, bien que ToBI ait été développé initialement pour la description de l'anglais, ce système d'annotation a été plus largement appliqué à d'autres langues : entre autres, K-ToBI pour le Coréen (Beckman & Jun, 1996), ToDI pour le hollandais (Gussenhoven et al., 2003), F-ToBI pour le français (Delais-Roussarie et al., 2015), G-ToBI pour l'allemand (Grice & Baumann, 2002). A chaque langue ou variété de langues, est donc associé un système ToBI particulier, puisque selon la langue décrite, le système de notation peut s'adapter, notamment par rapport à l'inventaire des catégories tonales (« *pitch accent* » et « *boundary tone* »), mais aussi par rapport à la hiérarchie des constituants prosodiques considérés : dans le cadre du hollandais, par exemple, le niveau du Syntagme Intermédiaire n'est pas pris en compte dans le système de notation ToDI (Gussenhoven et al., 2003).

Nous comprenons donc à travers la présentation du système de notation ToBI, mais aussi plus largement à travers la présentation des approches plus « purement » prosodiques, que les contraintes issues d'autres niveaux linguistiques, tels que la syntaxe ou la sémantique, sont moins centrales que les contraintes liées à l'organisation mélodiques et rythmiques. Bien que certaines approches du système ToBI prennent en compte les caractéristiques syntactico-sémantiques des énoncés (notamment Delais-Roussarie et al., 2015 pour F-ToBI), la description de la structure prosodique reste essentiellement basée sur l'observation et l'annotation du signal physique, et plus précisément sur l'annotation du mouvement tonal. Le paramètre de durée, participant également à la construction des constituants, n'est en revanche nullement considéré. Or, nous verrons dans le cadre de l'analyse du français qu'il est essentiel de prendre en compte à la fois les paramètres tonals et de durée, pour expliquer des phénomènes ne pouvant être interprétés sur la base des événements tonals seulement. Par ailleurs, il est intéressant de signaler que les deux approches plus « purement » prosodiques, présentées dans le cadre du chapitre, se distinguent l'une de l'autre concernant les unités phonologiques prises en compte dans leurs modèles prosodiques. Alors que dans certains modèles issus de la Théorie Métrique le niveau de constituance minimum est le Pied métrique, dans la Théorie Métrique Autosegmentale, en revanche, le niveau de constituance minimum reste systématiquement le niveau du Syntagme Accentuel. Ce dernier point sera également soulevé dans les prochains chapitres, à travers notamment une discussion sur la pertinence du niveau du Mot Prosodique dans la langue.

### 2. REALITE PSYCHO-COGNITIVE DE LA HIERARCHIE DES CONSTITUANTS PROSODIQUES

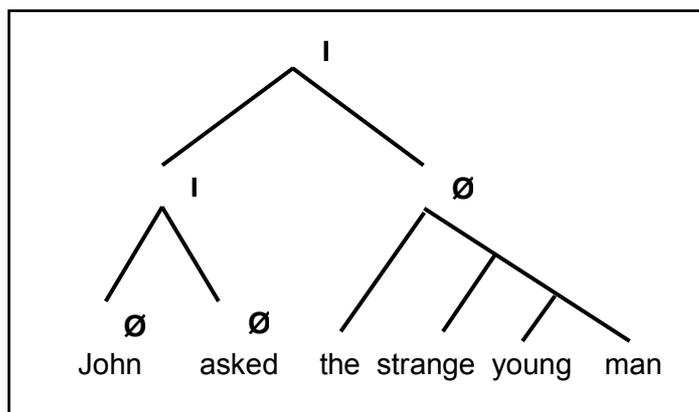
En parallèle aux modèles théoriques, dont le but est de décrire la structure prosodique d'une langue en indiquant les différents niveaux de constituance formés, selon les modèles, à partir de règles morphosyntaxiques, métriques ou tonales, d'autres travaux se sont davantage portés sur la relation entre ces théories linguistiques et le traitement du langage. Autrement dit, relier ce que produit le locuteur ou ce que perçoit l'auditeur aux prédictions théoriques. Il s'agit dans ce cas-là de la réalité psycho-cognitive de la hiérarchie des constituants prosodiques, puisqu'il s'agit de tester directement les indices prosodiques pertinents qui contribuent à l'élaboration de la structure hiérarchique. Un certain nombre d'auteurs se sont donc intéressés à cette question, avec cependant des angles d'attaque très différents selon les travaux. Alors que certains s'intéressent davantage à la manière dont les locuteurs segmentent le flux de parole en groupements, d'autres, en revanche, se positionnent à un niveau de granularité plus fin, en s'intéressant plutôt à la manière dont les auditeurs perçoivent les indices prosodiques, tels que les frontières et les proéminences, à l'intérieur et entre ces mêmes groupements. Dans la première approche nous nous situons davantage sur le plan de la production, tandis que dans la seconde approche, nous nous inscrivons plutôt sur le versant perceptif.

#### 2.1 Travaux en production sur les indices de groupements

##### 2.1.1 L'analyse des structures de performances

Cette question autour de la réalité psycho-cognitive des niveaux de constituance prosodique a principalement été abordée en psycholinguistique, et plus précisément dans les travaux effectués sur les « *structures de performances* » (Gee & Grosjean, 1983 ; Grosjean & Deschamps, 1975 ; Grosjean & Dommergues, 1983 ; Grosjean, Grosjean & Lane, 1979). En effet, grâce à différentes tâches expérimentales en production (lecture lente de texte, segmentation, jugements de liaisons, etc.), il est possible d'obtenir la structure hiérarchique d'une phrase, en testant notamment le découpage des énoncés en groupements. Grosjean & Deschamps (1975) ont montré par exemple que le nombre et la longueur des pauses sont des indices pertinents dans la construction de la hiérarchie d'une phrase. La fréquence et la durée des pauses varient, effectivement, en fonction de la relation de dépendance entre les mots ou groupes de mots : les pauses sont moins fréquentes et plus courtes entre les mots ou groupes de mots qui entretiennent une forte dépendance.

Par ailleurs, Gee & Grosjean (1983) ont mis en évidence que la position et la force des pauses, identifiées dans le discours, ne sont pas systématiquement en parfaite adéquation avec les frontières prédites par la syntaxe classique. En effet, les auteurs ont élaboré un algorithme (« *l'algorithme PHI* » pour « *Phonological Index* »), permettant de prédire les variables expliquant les « *structures de performances* », fondé sur une théorie de la structure prosodique. Plus précisément, les auteurs se sont basés sur la « Théorie Prosodique » (Hayes, 1989 ; Nespor & Vogel, 1986 ; Selkirk, 1980 ; 1984 ; 1986), en prenant en compte le niveau du Syntagme Phonologique qu'ils appellent « *∅-phrase* », dont la structure, marquée par une prééminence finale forte, se compose d'une unité tête lexicale entourée de mots grammaticaux. Gee & Grosjean (1983), puis Grosjean & Dommergues (1983) ont choisi cette unité étant donné sa position intermédiaire entre le mot lexical et la proposition. A partir de là, les auteurs ont élaboré un algorithme composé de Syntagmes Phonologiques pouvant se regrouper eux-mêmes, selon certaines règles, en Syntagmes Intonationnels (noté « I ») : une première règle consiste à regrouper dans un même Syntagme Intonational tous les Syntagmes Phonologiques faisant partie du même constituant syntaxique, et une seconde règle, portant uniquement sur le verbe, indique que le Syntagme Phonologique, dont la tête est le verbe, se rattache à l'unité prosodique comportant le moins de nœuds dans l'arbre. Cette dernière règle est illustrée dans la Figure N°14, où le verbe « *asked* » est rattaché à « *John* » plutôt qu'à « *the strange young man* », alors que la syntaxe prédit un rattachement du verbe au complément situé à droite. Cette règle tend toutefois à rendre compte des résultats obtenus dans les tests expérimentaux, où les auteurs ont observé davantage de pauses au sein d'un syntagme verbal qu'au sein d'un syntagme nominal.



**Figure 14 :** Exemple d'application de la règle portant sur le rattachement du Syntagme Phonologique, dont la tête est un verbe, à l'unité prosodique comportant le moins de nœuds branchants dans l'arbre

Le verbe « *asked* » est ici rattaché à « *John* » plutôt qu'à « *the strange young man* » (issu de Grosjean & Dommergues, 1983, p. 531)

Grosjean & Dommergues (1983) constatent effectivement que le découpage en groupements respecte un certain équilibre syllabique entre les différents groupements. Selon les auteurs, les Syntagmes Intonatifs illustrent donc mieux les productions authentiques de l'oral que les syntagmes syntaxiques classiques.

« Un algorithme fondé sur la structure prosodique des phrases est bien meilleur prédicteur des données expérimentales. Ce qui a jusqu'à présent été baptisé « structures de performance » ne serait donc en fait que le reflet des « structures prosodiques » de la langue. » (Grosjean & Dommergues, 1983, p. 533)

Ainsi, en intégrant à la fois des informations phonologiques, rythmiques et syntaxiques, « l'algorithme PHI » permet de rendre compte de 92% de la variance des pauses, contre 56% à 69% dans les autres algorithmes davantage basés sur des informations purement syntaxiques (Algorithme « GGL » Grosjean, Grosjean & Lane, 1979 et Algorithme « CPC » Cooper & Paccia-Cooper, 1980). Par ailleurs, Grosjean & Dommergues (1983) suggèrent que le Syntagme Phonologique serait l'unité de base du traitement du langage : l'auditeur isole d'abord le Syntagme Phonologique pour ensuite décomposer le syntagme en ses éléments lexicaux.

### 2.1.2 L'analyse en « chunks »

Sur la base de ces travaux en psycholinguistique, d'autres auteurs ont mis en place une technique de segmentation de la parole, qui a principalement été mise à profit dans les travaux de traitement automatique des langues : il s'agit de l'analyse en « *chunks* ». En effet, le chunk est l'unité d'analyse privilégiée et centrale de cette méthode, dont la notion est issue des travaux d'Abney (1991 ; 1992).

« I begin with an intuition: when I read a sentence, I read it a chunk at a time. » (Abney, 1991, p. 1)

Selon l'auteur, nous segmentons le discours en une série de tronçons ne correspondant pas systématiquement aux structures initialement prédites par la syntaxe classique. Dans la perspective d'Abney (1991 ; 1992), le chunk est un constituant minimal, composé d'une tête lexicale précédée ou suivie de mots grammaticaux. Nous retrouvons ici une unité similaire

au Syntagme Phonologique (Selkirk, 1986) et au « *ø-phrase* » (Gee & Grosjean, 1983 ; Grosjean & Dommergues, 1983). Le chunk est également considéré comme une unité non récursive : les chunks ne peuvent pas se chevaucher ou s'imbriquer entre eux. Il s'agit donc de constituants, dont la structure interne est relativement figée, correspondant plus ou moins aux patrons prosodiques. En effet, chaque chunk renvoie à un groupe accentuel potentiel, marqué par une proéminence forte qui peut être suivie ou non d'une pause. Le chunk correspond aussi aux unités syntaxiques : un chunk nominal, verbal ou prépositionnel correspond à un syntagme nominal, verbal ou prépositionnel. Selon Abney (1991), la correspondance entre les unités prosodiques et les unités syntaxiques n'est pas directe mais plutôt arbitrée par les chunks. Toutefois, l'approche de l'auteur n'envisage pas seulement l'unité chunk et propose trois niveaux de constituance hiérarchisés : un chunk qui peut être contenu dans une clause, qui peut elle-même être contenue dans un énoncé. Plus précisément, l'auteur conçoit deux étapes de segmentation : l'étape « *chunker* » qui sert à diviser le flux de parole en plusieurs chunks, et l'étape « *attacher* » qui sert à rassembler ces différents chunks dans des clauses. Le principal objectif de cette dernière étape est de gérer les ambiguïtés dans le rattachement d'un chunk à un autre, et donc d'établir les degrés de frontière entre les chunks. Lacheret-Dujour et Morel (2002) constatent en effet que les frontières prosodiques sont définies en fonction des relations de dépendance syntaxique entre les chunks. Les auteurs proposent d'illustrer ce phénomène par deux énoncés (issus de Lacheret-Dujour & Morel, 2002, p. 4) :

- a. (Martin)a (ne viendra pas)b
- b. (Martin)a (comme tu le sais sans doute)b (ne viendra pas)c

Dans le premier énoncé, les deux chunks (a) et (b) entretiennent une relation de dépendance linéaire contigüe, marquée prosodiquement par une proéminence finale, qui est réalisée notamment par un allongement de la durée syllabique. Dans le second énoncé, en revanche, les deux chunks (a) et (c) entrent dans une relation de contigüité non linéaire. Dans ce cas, pour comprendre la dépendance réactionnelle entre (a) et (b), une proéminence est réalisée sur la dernière syllabes de (a) avec un allongement plus prolongé, ainsi qu'une pause de longueur proportionnelle au nombre de syllabes du chunk (c).

D'autres auteurs parlent également de « *super-chunk* » (Blanc et al., 2010) qui, selon eux, est une forme améliorée du chunk pouvant contenir des unités lexicales plus complexes. Les auteurs prennent l'exemple du constituant « *marge d'exploitation* » : un chunker simple segmenterait le constituant en deux chunks, nominal (« *marge* ») et

prépositionnel (« d'exploitation), tandis qu'avec le « *super-chunk* » les deux syntagmes sont intégrés dans un seul et même constituant (« *un super-chunk* » nominal). Selon les auteurs, la prise en compte d'une telle unité facilite l'identification des unités sémantiques. Cette notion de « *super-chunk* » rappelle par ailleurs la notion de « *superdomain* » (Ladd, 1986), et permet également, à l'inverse de ce que propose Abney (1991), d'intégrer la notion de hiérarchie à travers le caractère récursif du constituant prosodique.

### **2.2 Travaux en perception sur les indices de frontières et de prééminences**

Sur le plan perceptif, maintenant, certains auteurs se sont intéressés à la manière dont les auditeurs natifs perçoivent les prééminences et les frontières à l'intérieur et entre les groupements, en mettant par ailleurs en lien la perception des auditeurs avec la réalisation physique concrète de ces événements prosodiques. Alors que dans certains travaux la perception d'auditeurs experts ou entraînés est privilégiée (entre autres, Avanzi et al., 2010 ; Buhmann et al., 2002 ; Lacheret-Dujour et al., 2014), dans d'autres, au contraire, c'est la perception d'auditeurs naïfs qui est favorisée (entre autres, Streefkerk 1997 pour le hollandais ; Roux et al., 2016 ; Smith, 2009 pour le français ; Cole et al., 2010a ; 2010b ; 2014 ; Mo et al., 2008 pour l'anglais), afin d'être plus proche des conditions dans lesquelles les auditeurs ordinaires perçoivent et interprètent la prosodie (Cole et al., 2010b).

Dans la plupart des études, il s'agit d'annoter un enregistrement de parole plus ou moins spontanée, à partir d'une transcription orthographique, en signalant la présence de mots prééminents et de frontières entre les groupes de mots. Sur la base de ces annotations, l'objectif est ensuite d'évaluer à quel point les auditeurs sont en accord quant à leur perception de la structure prosodique. Cole et al. (2010a), par exemple, ont procédé à une étude sur 97 auditeurs américains anglophones non entraînés, à qui il était demandé de surligner les mots considérés comme prééminents et d'indiquer les frontières entre les « *chunks* » à l'aide d'une barre verticale. Sur la base des résultats, les auteurs ont effectué des tests d'accords inter-annotateurs qui leur ont révélé un taux d'accord plus élevé dans le marquage des frontières prosodiques majeures que dans le marquage des prééminences. Ce résultat a également été confirmé dans d'autres travaux portant sur le français (Roux et al., 2016 ; Smith, 2009). Plus précisément, Smith (2009) montre que les francophones ont tendance à marquer les frontières prosodiques majeures, telles que les frontières de Syntagme Intonatif, plutôt que les frontières de niveau inférieur. Dans leur étude, Cole et al. (2010a) estiment que les auditeurs annotent moins fréquemment les frontières de petits niveaux, du type Syntagme Intermédiaire, du fait que les locuteurs, produisant l'énoncé, sont

moins constants dans l'implémentation phonétique lorsqu'ils encodent les petits niveaux de frontière. Par ailleurs, Smith (2009) observe que certaines proéminences sont mieux perçues que d'autres, et notamment les proéminences sur les syllabes finales précédant une frontière de groupement. Portes (2002) constatait déjà ce phénomène, puisque les résultats de son étude montrent qu'une syllabe précédant une frontière est marquée comme proéminente dans 84% des cas. De la même manière pour l'anglais, Mo et *al.* (2008) montrent un jugement plus homogène pour les accents nucléaires, en fin d'unité prosodique majeure, que pour les accents pré-nucléaires. Dans ce même sens, à travers une étude en imitation de la parole, Cole & Shattuck-Hufnagel (2011) constatent que les auditeurs perçoivent généralement la structure prosodique « basique » d'un énoncé, et la reproduisent de manière très précise, avec cependant une imitation bien plus fidèle des frontières de Syntagmes Intonatifs par rapport aux autres niveaux de frontière, et des accents nucléaires par rapport aux accents prénucléaires.

Certains auteurs s'intéressent également à l'influence que peuvent avoir les informations syntaxiques sur la perception des indices de frontières et de proéminences, mais aussi à la manière dont cette influence syntaxique est arbitrée par le signal acoustique. Cole et *al.* (2010a) ont montré dans leur étude que les auditeurs semblent davantage marquer les frontières d'unités syntaxiques majeures, du type « clause », réalisées par une durée vocalique plus importante. Wightman et *al.* (1992) observaient déjà une augmentation de la durée de la syllabe finale d'un mot en frontière d'unité prosodique majeure. Par ailleurs, les résultats de Cole et *al.* (2010a, 2010b) semblent indiquer que la contribution de la syntaxe est plus importante que celle de la durée : la perception des frontières est donc dans ce cas-là fortement corrélée à la catégorie syntaxique, qui est elle-même corrélée aux indices acoustiques. Cependant, dans une étude ultérieure, Cole et *al.* (2014) ont directement comparé l'impact des informations syntactico-sémantiques et acoustiques sur la perception des frontières et des proéminences, à travers un protocole expérimental spécifique, où les auditeurs annotent une transcription orthographique (« *Rapid Prosody Transcription* ») dans 3 contextes différents : un groupe d'auditeurs doit se baser seulement sur des critères syntactico-sémantiques, un autre groupe doit se servir uniquement de critères acoustiques, et enfin un dernier groupe ne reçoit aucune instruction spécifique. Les résultats montrent que la performance des auditeurs varie en fonction du critère d'annotation, et notamment pour le marquage des proéminences. Par ailleurs, lorsqu'aucune consigne particulière n'est donnée aux participants, les auteurs constatent que les annotateurs semblent se focaliser davantage sur les indices acoustiques. D'autres travaux ont également porté sur l'influence des indices acoustiques dans la perception des frontières et des proéminences, en utilisant cette fois-ci un paradigme expérimental évitant la

visualisation du script orthographique par l'annotateur. de Pijper & Sanderman (1994), par exemple, ont choisi d'utiliser une méthode de délexicalisation du matériel linguistique, permettant de conserver le signal prosodique tout en rendant la phrase inintelligible. Les auteurs constatent également une corrélation entre la perception des frontières prosodiques majeures et l'allongement de la durée de la syllabe finale. Plus récemment encore, Simon & Christodoulides (2016a ; 2016b) proposent de tester la perception des frontières en temps réel, également sans influence de la représentation écrite, où les auditeurs, durant l'écoute des extraits, pressent un bouton dès lors qu'ils perçoivent une frontière entre deux groupes de mots. Plus précisément, les auteurs ont sélectionné différents extraits de parole issus du corpus multigenre LOCAS-F (Martin et al., 2014), dont la moitié a été manipulée et resynthétisée afin de ne conserver que le contenu segmental. Au total, 133 participants ont chacun annoté 12 échantillons de parole naturelle et 12 échantillons de parole manipulée. A l'instar de Cole et al. (2014), les résultats de l'étude montrent que les indices acoustiques priment sur les indices syntactico-sémantiques, puisque la perception des frontières se trouve être plus consensuelle en parole manipulée (sans contenu segmental) qu'en parole naturelle. Plus précisément, les auteurs constatent que l'indice acoustique le plus pertinent dans la perception des frontières est la pause silencieuse, suivie également de l'allongement vocalique, qui reste toutefois un indice corrélé à tout type de frontière. De la même manière, Smith (2009) montre que la réalisation d'une pause, supérieure à 150ms, est corrélée à la réalisation d'une frontière d'unité prosodique majeure. Ou encore, Swerts (1998) indique que les frontières fortes sont plus fréquemment associées aux pauses pleines (« *filled pauses* ») que les frontières faibles. Concernant les corrélats associés aux mouvements mélodiques, Hermès & Rump (1994) montrent qu'un contour tonal descendant, plutôt que montant ou montant-descendant, renforcerait la perception d'une proéminence sur la syllabe. Dans ce même sens, Portes (2002) constate que les auditeurs sont plus consensuels dans leur perception des frontières conclusives, réalisées avec un contour descendant, que dans leur perception des frontières continuatives, réalisées avec un contour montant. D'autres auteurs remarquent aussi un renforcement de la voyelle, et notamment une hyper-articulation de la voyelle, corrélée à la perception de constituants prosodiques majeurs (Gendrot et al., 2016). En revanche, Petrone et al. (2017) n'observent aucune différence phonétique entre les frontières de Mots Prosodiques et les frontières de Syntagmes Phonologiques.

Enfin, certains auteurs se sont également questionnés quant à la capacité de percevoir non pas la présence ou l'absence de frontières, mais bien plus des degrés de frontière différents (Carlson et al., 2005 ; de Pijper & Sanderman, 1994 ; Krivokapic & Byrd, 2012 ; Swerts, 1997). En effet, certaines études montrent que les locuteurs sont capables de produire, et les auditeurs capables de percevoir, une granularité importante des niveaux de

frontière. de Pijper & Sanderman (1994), par exemple, ont mis en place une étude perceptive, où des auditeurs hollandais natifs doivent juger du degré de frontière sur une échelle de 1 à 10. Les résultats montrent que les auditeurs perçoivent différentes forces de frontière qui sont potentiellement associées à une seule et même catégorie phonologique. Swerts (1997), également, constate que les auditeurs perçoivent jusqu'à six forces de frontière différentes variant en fonction de la taille de la pause. De la même manière, Krivokapic & Byrd (2012) montrent, non seulement, que les locuteurs sont capables de graduer les frontières de Syntagmes Intonatifs, mais également, que les auditeurs parviennent à détecter ces différentes forces de frontière. Là encore, les auteurs considèrent que plusieurs degrés de frontière sont associés à une même catégorie de constituant prosodique. Ces observations semblent alors renvoyer aux propositions de Ladd (1986 ; 1996 / 2008) concernant la prise en compte du principe de récursivité (cf. 1.1.1.2 du Chapitre 2). Effectivement, nous rappelons que l'auteur considère qu'un constituant d'un niveau donné peut être marqué par différents indices acoustiques entraînant alors la perception de différentes forces de frontière : dans le cas de « *A and B but C* » la frontière après le Syntagme Intonatif « *B* » est plus forte que celle après le Syntagme Intonatif « *A* » (Ladd, 1986 ; 1996 / 2008). L'auteur parle alors de « *compound prosodic domains* » : une unité prosodique d'un niveau donné regroupant un constituant du même niveau. D'après Ladd, il y aurait donc deux solutions pour interpréter ces différentes forces de frontière : soit considérer qu'à chaque degré de frontière est associé un niveau de constituance spécifique, soit prendre en considération le caractère récursif de la structure prosodique. Comme nous l'avons déjà mentionné, l'auteur privilégie la seconde solution et préfère réduire l'inventaire des niveaux de constituance plutôt que d'incrémenter systématiquement en nouvelles unités la hiérarchie prosodique.

Différents protocoles expérimentaux peuvent donc être utilisés pour rendre compte de la perception des frontières et des proéminences. Des différences sont notamment observées sur le choix d'un protocole « *off-line* » ou « *on-line* », sur la visualisation ou non du script orthographique, sur l'intelligibilité ou non du contenu segmental, mais également sur le type de consignes données aux auditeurs. En revanche, l'objectif de toutes les études présentées ici est d'évaluer à quel point les auditeurs sont en accord quant à leur perception de la structure prosodique, en recourant à un accord inter-juge. De la même manière, il semblerait que dans la plupart des études, seule la perception sur le plan paradigmatique soit envisagée. Nous constatons en effet que les études testent la perception d'une catégorie d'événements prosodiques spécifique (frontière majeure, accent nucléaire, etc.), sans réellement mettre en relation ces différents événements avec les autres événements adjacents, intervenant dans la construction de la structure prosodique. Nous reviendrons

dans les prochains chapitres sur la pertinence de ces différentes études perceptives, à travers notamment la mise en place de notre propre protocole expérimental.



## SYNTHESE DU CHAPITRE 2

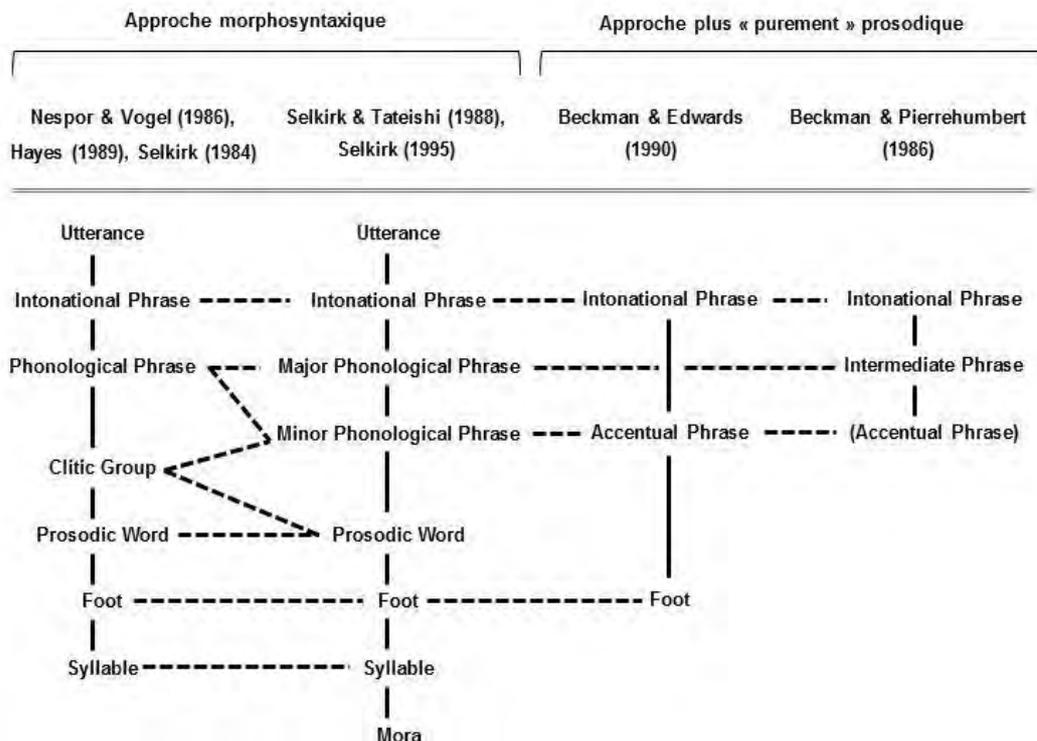
---

L'objectif de ce chapitre est de présenter dans un premier temps les grands courants théoriques à l'origine de la description du phrasé prosodique d'une langue, pour ensuite mettre en relation ces modèles théoriques avec le traitement même du langage. Cette présentation à la fois des prédictions théoriques et de leur réalité psycho-cognitive, permet ainsi de poser les fondements des deux prochains chapitres consacrés à la description du phrasé prosodique en français (*cf.* Chapitre 3), puis à la mise en place de notre protocole expérimental (*cf.* Chapitre 4).

Les trois courants théoriques que nous avons présentés se distinguent les uns des autres notamment concernant le nombre d'unités prosodiques pris en compte, mais également selon la manière dont les niveaux de constituance sont modélisés. En effet, alors que la Théorie Prosodique (Hayes, 1989 ; Nespor & Vogel, 1986 ; Selkirk, 1980 ; 1984 ; 1986) se fonde principalement sur des critères morphosyntaxiques, en établissant une hiérarchie des constituants exhaustive, la Théorie Métrique (Halle & Vergnaud, 1987 ; Hayes, 1995 ; Liberman, 1975 ; Liberman & Prince, 1977) et la Théorie Métrique Autosegmentale (Beckman & Pierrehumbert, 1986 ; Pierrehumbert, 1980), en revanche, proposent une hiérarchie moins complète, en définissant leurs unités prosodiques selon des principes plus « purement » prosodiques. En effet, dans la première théorie, ce sont les phénomènes accentuels qui permettent de générer la structure en constituants prosodiques, tandis que dans la seconde théorie, ce sont plutôt les phénomènes intonatifs, au détriment, par ailleurs, d'autres paramètres acoustiques également pertinents dans la construction de la structure prosodique (notamment le paramètre de durée). En comparant les différentes hiérarchies prosodiques proposées dans ces trois courants théoriques, nous remarquons qu'il est difficile pour certains niveaux de constituance de mettre en correspondance les différentes approches (*cf.* Figure N°15). Alors que les constituants situés aux deux extrémités de la hiérarchie semblent relativement consensuels entre les différentes théories (au niveau du Syntagme Intonatif pour le haut de la hiérarchie, et au-dessous du niveau du Mot Prosodique pour le bas de la hiérarchie), les constituants en position intermédiaire, en revanche, ne montrent pas de correspondance claire (entre le niveau du Pied et du Syntagme Intonatif). Cette confusion est d'autant plus renforcée lorsque les auteurs utilisent le même terme pour se référer à des unités prosodiques différentes (*cf.* les différentes définitions du Mot Prosodique), ou à l'inverse, lorsque les auteurs utilisent des termes

## Synthèse du Chapitre 2

différents pour se référer au même niveau de constituance (cf. Syntagme Accentuel vs. Syntagme Phonologique).



**Figure 15** : Récapitulatif des différentes hiérarchies en constituants prosodiques issues des principaux courants théoriques : approche morphosyntaxique et approche plus purement prosodique (issu et adapté de Shattuck-Hufnagel, 1996, p. 206 puis de Di Cristo, 2011, p. 76)

Le Mot Prosodique, dont la pertinence en français sera discutée dans les prochains chapitres, ne montre effectivement aucun équivalent dans les approches plus « purement » prosodiques. La Théorie Métrique ou la Théorie Métrique Autosegmentale privilégient davantage les niveaux directement supérieurs au mot. Par ailleurs, bien que le Mot Prosodique soit envisagé dans les approches morphosyntaxiques, nous avons vu que les règles de construction de ce niveau de constituance varient d'un modèle à un autre, et notamment concernant le regroupement des mots grammaticaux avec le mot lexical. Quand certains estiment que le Mot Prosodique se compose uniquement du mot lexical (entre autres, Selkirk, 1996), d'autres considèrent que les mots grammaticaux et lexicaux peuvent chacun constituer un Mot Prosodique, et que le regroupement des deux types de mots s'effectuent au niveau supérieur, à savoir au niveau du Groupe Clitique (notamment Nespor

& Vogel, 1986). Concernant le Syntagme Phonologique, décomposé en un Syntagme Phonologique Majeur et en un Syntagme Phonologique Mineur suite à la révision du principe de récursivité (notamment Ladd, 1986), nous constatons que ce niveau correspond soit au Syntagme Intermédiaire soit au Syntagme Accentuel dans les approches plus « purement » prosodiques. Toutefois, alors que le niveau du Syntagme Accentuel est pris en compte dans tous les modèles, le niveau du Syntagme Intermédiaire, en revanche, n'apparaît pas systématiquement. La pertinence de ce dernier niveau de constituance est d'ailleurs remise en question par certains auteurs (Ladd, 1986 ; 1996 / 2008).

Sur la base de ces approches théoriques, nous choisissons de terminer le chapitre en présentant les travaux qui se sont intéressés à mettre en lien les prédictions théoriques avec le traitement réel du langage : c'est-à-dire ce que produit le locuteur ou ce que perçoit l'auditeur. Que ce soit sur le plan de la production (travaux sur « *les structures de performances* » : Gee & Grosjean, 1983 ; Grosjean & Deschamps, 1975 ; Grosjean et al., 1979) ou sur le plan de la perception (entre autres, Cole et al., 2010a ; 2010b ; 2014 ; Simon & Christodoulides, 2016a ; 2016b ; Smith, 2009 ; Swerts, 1998), les auteurs constatent que la pause est l'indice acoustique le plus pertinent dans le découpage des énoncés en « *chunks* », et donc dans la construction de la structure prosodique. En effet, la plupart des auteurs s'accordent à dire que l'influence des indices acoustiques sur la perception des événements prosodiques est plus forte que celle des indices syntactico-sémantiques (Abney, 1991, 1992 ; Cole et al., 2014 ; Gee & Grosjean, 1983 ; Simon & Christodoulides, 2016a ; 2016b). En comparant également la perception des frontières et des proéminences au sein et entre les groupements, les auteurs constatent une meilleure évaluation perceptive des frontières par rapport aux proéminences (entre autres, Cole et al., 2010a ; Cole & Shattuck-Hufnagel, 2011 ; Smith, 2009), et ce principalement pour les frontières d'unités prosodiques majeures. D'autres indices acoustiques se sont également révélés intéressants dans les études en perception, notamment liés à l'allongement vocalique (de Pijper & Sanderman, 1994 ; Wightman et al., 1992) et au mouvement tonal (Hermès & Rump, 1994 ; Portes, 2002). Par ailleurs, en dépassant le cadre binaire de l'absence ou présence de tel ou tel événement prosodique, certains auteurs s'intéressent davantage à la perception graduelle des événements prosodiques. Des études montrent en effet que les auditeurs sont capables de produire et de percevoir une granularité importante des niveaux de frontière, et plus précisément, là encore, en frontière d'unités prosodiques majeures (de Pijper & Sanderman, 1994 ; Krivokapic & Byrd, 2012 ; Swerts, 1997). Nous discuterons dans les prochains chapitres de la pertinence de ces études en perception, notamment lors de l'élaboration de notre propre protocole expérimental.



## CHAPITRE 3 – Les particularités du phrasé prosodique en français

---

<b>1. LE FRANÇAIS EST-IL UNE LANGUE SANS ACCENT ?</b> .....	<b>95</b>
1.1 Accent non lexicalement distinctif.....	95
1.1.1 Accent final primaire.....	95
1.1.2 Accent initial secondaire.....	97
1.2 Syncrétisme entre accentuation et intonation.....	99
<b>2. LES MODELES DU PHRASE PROSODIQUE EN FRANÇAIS</b> .....	<b>101</b>
2.1 L'approche Intonosyntaxique/Phonosyntaxique.....	102
2.1.1 Delattre.....	103
2.1.2 Di Cristo.....	104
2.1.3 Rossi.....	105
2.1.4 Martin.....	107
2.2 L'approche Métrique Autosegmentale.....	110
2.2.1 Hirst et Di Cristo.....	111
2.2.2 Mertens.....	116
2.2.3 Jun & Fougeron.....	121
2.2.4 Post.....	126
2.3 L'approche Métrique et Fonctionnaliste.....	131
2.3.1 Delais-Roussarie.....	132
2.3.2 Di Cristo.....	136
<b>SYNTHESE DU CHAPITRE 3</b> .....	<b>143</b>



### 1. LE FRANÇAIS EST-IL UNE LANGUE SANS ACCENT ?

Une des particularités les plus marquantes dans le phrasé prosodique en français concerne l'accentuation qui fait l'objet encore aujourd'hui de controverses, concernant notamment la question de l'existence même ou non d'un accent dans la langue. Plusieurs arguments contribuent à cette vision traditionnelle de l'accentuation liée à la nature, au domaine et à la fonction de l'accent en français, mais aussi au lien entre l'accentuation et l'intonation. Bien que certains auteurs tendent à réhabiliter le rôle phonologique de l'accent en français, la vision traditionnelle reste néanmoins majoritaire, y compris dans certains des modèles du phrasé prosodique actuels.

#### 1.1 Accent non lexicalement distinctif

##### 1.1.1 Accent final primaire

A l'inverse des autres langues romanes et des langues germaniques, le français n'est pas caractérisé par un accent lexical distinctif, contribuant à l'identification du mot, mais plutôt par un accent post-lexical dont la fonction serait purement démarcative. Plus précisément, le français est considéré comme une langue oxytonique à accent fixe, où l'accent est systématiquement réalisé sur la dernière syllabe pleine (hors schwa) d'un groupe de mots (entre autres, Delattre, 1966b ; Dell, 1984 ; Garde, 1968). Dans ce contexte, il apparaît alors inadapté de considérer le mot en français comme une unité phonologique pertinente. Garde (1968) estime d'ailleurs que le français est caractérisé par une certaine « *élasticité de l'unité accentuelle* », dans la mesure où il est difficile de définir précisément le type d'unité accentogène, puisque l'accent final (noté « AF ») affecte des limites qui varient d'un énoncé à un autre. L'auteur parle également « *d'unité accentuelle virtuelle* » car l'accent sur cette unité peut être réalisé, ou non, et va jusqu'à questionner l'utilisation même du terme « accent » pour le français.

« Cette particularité [en se référant au phénomène d'élasticité de l'unité accentuelle, ndr] de notre langue crée l'impression, assez répandue chez les francophones, que le français n'a pas d' « accent tonique ». Et de fait, on peut se demander si une mise en relief qui se fait dans le cadre d'une unité qui n'est pas grammaticalement définissable mérite encore le nom

|| d'accent. Nous pensons que oui, puisqu'il subsiste la possibilité de définir grammaticalement l'unité accentuelle virtuelle. » (Garde, 1968, p. 96)

Toutefois, nous rappelons que Garde (1968) propose la dichotomie « *accent sous-jacent* » vs. « *accent effectif* », sous-entendant alors l'existence d'un accent abstrait en structure métrique. Cette reconnaissance métrique de l'accent en français semble dans ce contexte contradictoire avec la notion « *d'élasticité de l'unité accentuelle* » proposée par l'auteur. En effet, définir l'unité accentogène en français ne devrait pas être problématique, dans la mesure où nous considérons que tous les mots lexicaux dans la langue portent un accent au niveau sous-jacent. Cette notion d'accent métrique est par ailleurs centrale dans notre travail. Nous tenons en effet à rendre compte du statut phonologique de l'accent en français, en position finale et initiale.

Beckman (1986) considère par ailleurs que le français n'a pas d'accent au sens « *stress* »<sup>3</sup>, mais seulement des indices démarcatifs de syntagmes ou de mots phonologiques. Dans ce même sens, Vaissière (1991) distingue les « *langues à frontières* », où l'accent a un rôle démarcatif comme en français, des « *langues à accent* », où la fonction culminative de l'accent domine comme en anglais. Il serait intéressant de mettre en lien cette distinction avec celle établie entre les « *langues à rythmicité syllabique* » et les « *langues à rythmicité accentuelle* » (Abercrombie, 1967 ; Pike, 1945). En effet, Pike estime que les langues iso-accentuelles permettent de mieux mettre en valeur les syllabes accentuées que les langues iso-syllabiques comme le français : nous rappelons que les syllabes accentuées et non accentuées tendent à avoir une longueur comparable dans les langues à rythmicité syllabique, tandis que la variation de durée syllabique est nettement plus importante dans les langues à rythmicité accentuelle. De la même manière, Fant & Kruckenberg (1995) montrent que le contraste entre les syllabes accentuées et inaccentuées est moins important en français qu'en suédois. Cependant, comme nous l'avons déjà mentionné (cf. 1.2.1 du Chapitre 1), la distinction entre les langues à rythmicité syllabique et accentuelle a largement été discutée et remise en question (entre autres, Dauer ; 1983 ; Duez, 1989 ; Fant et al., 1991 ; Roach, 1982 ; Wenk & Wioland, 1982). En français, plus spécifiquement, des études ont montré la pertinence du contraste de durée syllabique dans le marquage des accents (entre autres, Astésano, 1998 ; 2001 ; Delattre, 1966b ; Fletcher, 1991 ; Rossi, 1981 ; 1985), mais également du mouvement mélodique, notamment lorsque l'accent démarque une unité prosodique qui n'est pas en fin d'énoncé (Post, 2000 ; Welby, 2003).

---

<sup>3</sup> Voir 1.1.1 du Chapitre 1 pour la distinction faite par l'auteur entre « *stress accent* » et « *non stress accent* ».

Enfin, des études en psycholinguistique montrent que les auditeurs francophones ont du mal à percevoir les contrastes accentuels (Dupoux et *al.*, 1997 ; 2008 ; 2010 ; Peperkamp & Dupoux, 2002). En effet, en comparant la perception des auditeurs francophones et hispanophones, c'est-à-dire une langue à accent fixe et non distinctif vs. une langue à accent libre et distinctif, les auteurs constatent que les francophones sont insensibles aux variations de l'accent lexical réalisé dans la langue étrangère (ici l'espagnol) et parlent donc de « *surdité accentuelle* » (Dupoux et *al.*, 2008 ; 2010). Selon eux, l'accent lexical en français n'est pas un trait pertinent utilisé pour distinguer les mots et n'est donc pas codé mentalement chez les locuteurs francophones, à l'inverse des locuteurs de langues à accentuation lexicale distinctive, telles que l'espagnol ou l'anglais. Toutefois, l'hypothèse selon laquelle les francophones seraient sourds à l'accent a été remise en question, notamment par Tremblay (2009) qui montre une évaluation perceptive de l'accent similaire chez les francophones et les anglophones. De la même manière, Michelas et *al.* (2016) suggèrent que les francophones ne sont pas totalement sourds à l'accent, puisqu'ils sont sensibles aux contrastes accentuels, quand ils doivent distinguer deux mots monosyllabiques avec et sans accent primaire.

### 1.1.2 Accent initial secondaire

En dehors de l'accent final, le français est également caractérisé par un accent initial (noté « AI ») décrit comme optionnel, dont la fonction est encore aujourd'hui discutée. En effet, certains auteurs se sont opposés au caractère strictement oxytonique de l'accentuation en français, et notamment Fónagy (1980) qui parle d'une « *accentuation probabilitaire* », où la position de l'accent est plus fluctuante que fixe. L'auteur considère effectivement que la réalisation de l'accent dépend de différents facteurs tels que le genre de parole, la taille et le type de mot, la structure syntaxique, les états émotifs du locuteur ou encore la structure phonétique de la syllabe. Fónagy rend également compte d'une tendance de plus en plus importante en français à accentuer la syllabe initiale des mots et des groupes de mots. Cette accentuation initiale est cependant davantage envisagée dans un rôle discursif, pragmatique et sémantique que purement structurel. En effet, selon l'auteur, AI est utilisé dans une réalisation accentuelle plus complexe, appelée « *arc accentuel* », qui consiste à accentuer le début et la fin d'une unité dans le but de renforcer la cohésion sémantique à l'intérieur de cette même unité, et ainsi favoriser une démarcation nette du groupe de sens.

« Une tendance centrifuge poussant l'accent du premier membre d'un syntagme vers la première syllabe et celui du deuxième membre vers la

|| dernière syllabe d'un syntagme, qui pourrait faire mieux ressortir son unité syntactique et sémantique. » (Fónagy, 1980, p.143)

Cet arc accentuel peut être réalisé sur un seul mot (composé au moins de trois syllabes) ou sur un groupe de mots. Comme nous pouvons le voir dans la Figure N°16, l'énoncé, produit en parole spontanée, est réalisé par un premier arc accentuel sur le mot « *amplitudes* » et un par second arc accentuel sur le groupe de mots « *divers formants* ». Il est également intéressant de remarquer que les syllabes internes aux arcs accentuels sont marquées par un poids accentuel moyen de 0, excepté la syllabe « *vers* » dans l'arc accentuel « *divers formants* », associée à un poids accentuel moyen de 0,25. Cela pourrait alors correspondre à une trace métrique de l'AF sur le mot lexical « *divers* », qui selon Di Cristo & Jankowski (1999) est en général essentiellement portée par le paramètre de durée.

les	<b>am-</b>	pli-	<b>tudes</b>	des	<b>di-</b>	vers	for-	<b>mants</b>
0	<b>0,85</b>	0	<b>0,95</b>	0	<b>0,85</b>	0,25	0	<b>0,95</b>

**Figure 16 :** Phénomènes d'arcs accentuels réalisés sur « *amplitudes* » et « *divers formants* » dans l'énoncé « *les amplitudes des divers formants* »

Les valeurs associées à chaque syllabe correspondent au poids accentuel moyen déterminé à partir des valeurs : 0= absence d'accent, 1= accent principal, 0,5= accent secondaire (issu de Fonagy, 1980, p. 155)

D'autres auteurs décrivent également AI comme un accent emphatique ou d'insistance. Lucci (1983), notamment, parle d'un « *accent didactique* » comme le produit d'une couche socio-culturelle spécifique, qui est utilisé dans le cadre de conférences, de discours, d'interviews ou encore de lectures de texte. Pour certains auteurs, en revanche, AI aurait un rôle plus rythmique que sémantique. En effet, Rossi (1980 ; 1985 ; 1999) utilise le terme « *d'ictus mélodique* » et Padeloup (1990a) « *d'accent secondaire* » pour parler d'un accent rythmique, dont le but est d'éviter une suite trop longue de syllabes inaccentuées. Cette fonction rythmique renvoie au phénomène de « *stress-shift rule* » (ou « *règle des thirteen men* » cf. 1.2.1.1 du Chapitre 2) qui vise à favoriser une alternance régulière de syllabes accentuées et inaccentuées (principe d'eurythmie), en évitant ainsi des collisions d'accents ou des vides accentuels. Dans ce contexte, AI est alors envisagé comme un déplacement de

l'accent. Garde (1968) suggérait déjà l'existence d'un « *écho d'accent* », à savoir un accent supplémentaire réalisé sur la seconde ou troisième syllabe avant AF. Padeloup (1990b, p. 513) considère plus précisément qu'AI joue un rôle de régulateur dans la structuration accentuelle, afin que cette organisation corresponde « *aux normes biologiques de production du rythme* » : des contraintes dues au système audio-perceptif évitant que deux accents consécutifs soient trop rapprochés, au risque de les percevoir comme la réalisation d'un seul et même accent ; et des contraintes dues à la taille de la mémoire immédiate imposant alors une taille limite aux groupes accentuels, qui doivent être composés au maximum de 7 à 8 syllabes. Enfin, plus récemment, certains auteurs attribuent une fonction plus structurelle à AI et montrent une volonté de l'intégrer dans les modèles phonologiques du français (Di Cristo 1999a ; 1999b ; Jun & Fougeron, 2000 ; 2002 ; Post, 2000). En effet, Astésano et al. (2007) rendent compte qu'AI tend à marquer de manière plus systématique qu'AF les niveaux de constituance inférieurs dans la hiérarchie. De la même manière, des études en neuro-imagerie (Aguilera et al., 2014 ; te Rietmolen et al., 2016 ; 2017), et plus spécifiquement en utilisant la méthode d'Electro-Encéphalo-Graphie (EEG), mettent en évidence un véritable statut phonologique de l'AI en français. Plus précisément, les résultats d'une première étude montrent que les auditeurs préfèrent les mots marqués par un pattern accentuel comprenant un AI (Aguilera et al., 2014 ; te Rietmolen et al., 2016), et les résultats d'une seconde étude révèlent qu'AI est encodé au niveau lexical et facilite le traitement sémantique (te Rietmolen et al., 2017). Ces derniers travaux remettent par ailleurs en doute la question de « *surdité accentuelle* » des francophones (entre autres, Dupoux et al., 1997 ; 2008 ; 2010).

### 1.2 Synchrétisme entre accentuation et intonation

Au-delà des questionnements concernant les différents types d'accents ou leur fonction en français, d'autres problèmes se posent quant au rapport entretenu entre la syllabe accentuée et le contour intonatif. En effet, étant donné que les syllabes accentuées servent de point d'ancrage des mouvements tonals, un lien étroit est donc établi entre ces deux paramètres prosodiques. Rossi (1980), notamment, met en évidence un phénomène de synchrétisme entre l'accentuation et l'intonation. Selon l'auteur, tous les lexèmes peuvent potentiellement porter un accent (il parle « *d'accentème* »), mais étant donné le domaine post-lexical de l'accent en français, l'accent sur un mot isolé tend à être désaccentué lorsque le mot est inclus dans un groupe de mots. Un cas de congruence entre la syllabe accentuée et la frontière de groupe de mots est dans ce cas-là observé.

« *L'opposition des accentèmes ne joue aucun rôle dans les lexèmes en français ; mais les lexèmes possèdent tout de même une propriété accentuelle : ils sont tous générateurs d'accent ; ils sont accentogènes. Etant donné que par définition, le domaine de réalisation de l'accent est le groupe intonatif où accent et intonème sont syncrétisés et que l'intonème domine l'accent, les lexèmes sont en réalité intonogènes.* » (Rossi, 1980, p.21)

Selon Garde (1968) le trait « accent » serait donc redondant. Là encore, nous constatons l'ambiguïté laissée par l'auteur quant au statut métrique de l'accent en français. En effet, l'auteur considère-t-il que le trait « accent » renvoie à la réalisation de surface de l'accent ou bien à sa réalisation métrique ? Par ailleurs, Delattre (1966b) et Rossi (1980) ne considèrent pas que le mot perd totalement sa propriété accentuelle lorsqu'il est compris dans un groupe de mots plus large, et préfèrent utiliser le terme de « *désaccentuation incomplète* » ou « *d'accentuation partielle* », pour rendre compte du fait que « *l'accent du mot peut en partie survivre dans le groupe* » (Delattre, 1966b, p.72). Cette proposition semble également rappeler les exemples précédents d'arcs accentuels (Figure N°16), où nous observions une trace métrique de l'accent sur la syllabe finale du mot lexical situé à l'intérieur de l'arc accentuel. De la même manière, Ladd (1996 / 2008) contre-argumente cet état syncrétique de l'accent et de l'intonation, en insistant sur l'existence phonologique de l'accent en français, et donc en se positionnant dans une approche métrique de l'accentuation. Selon l'auteur, il est important de distinguer les faits phonologiques et phonétiques, à savoir le poids métrique d'une syllabe et sa réalisation perceptive et acoustique : une syllabe peut être métriquement forte sans pour autant être saillante sur le plan phonético-perceptif. Ainsi, même si l'accent final est congruent à la frontière du groupe de mots, et donc moins saillant perceptivement et acoustiquement, il n'y a aucune raison de penser que cette syllabe n'est pas pour autant métriquement forte. Ladd (1996 / 2008) se base également sur les propositions de Dell (1984), qui montre que la proéminence métrique est pertinente dans l'assignation du mouvement tonal. En effet, Dell constate que le contour tonal est systématiquement associé avec la dernière syllabe pleine d'un mot, et de ce fait, lorsque la dernière syllabe contient un « schwa », le contour tonal est alors associé à la syllabe précédente qui ne correspond pas à la frontière du constituant. Dans ce cas-là, l'accent n'est

pas congruent à la frontière intonative. Ladd suggère alors la possibilité d'une dissociation entre le niveau métrique et le niveau phonétique<sup>4</sup>.

« *By and large, syllables that are prominent in the abstract metrical structure are also phonetically stressed, but the theory does not preclude the occurrence of systematic dissociations. A syllable can be metrically strong or prominent without necessarily being stressed.* » (Ladd, 1996 / 2008, p.61)

Une étude en perception sur le français (Astésano et al., 2012 ; 2013 ; Astésano & Bertrand, 2016) montre également que les auditeurs francophones sont capables de percevoir les proéminences finales indépendamment des frontières prosodiques, que ce soit en frontière majeure ou en frontière de mot à l'intérieur même des groupes prosodiques. Selon les auteurs, l'Accent Final serait alors perçu métriquement fort quel que soit le type de frontière auquel il est associé.

Il apparaît donc difficile de trouver un consensus parmi les auteurs sur le sujet de l'accentuation en français, dont le rôle phonologique est encore aujourd'hui mal établi. En effet, des ambiguïtés semblent encore persister concernant le statut métrique de l'accent en français. Cela implique par ailleurs des questionnements dans les théories prosodiques actuelles quant au lien entre l'accent et le phrasé, et plus précisément sur la participation de l'accent dans le marquage de la constituance prosodique de la langue.

## **2. LES MODELES DU PHRASE PROSODIQUE EN FRANÇAIS**

Plusieurs modèles du phrasé prosodique ont été proposés pour le français au cours de ces 50 dernières années. Chaque modèle a ses propres hypothèses sur la manière de construire la structure prosodique de la langue et sur les unités phonologiques qui la composent. L'intérêt de la présentation de ces différents modèles est d'avoir une vue globale des méthodes de représentation et d'étiquetage du continuum de parole, tout en s'intéressant à la manière dont les différents modèles, au vu de la vision traditionnelle du français, gèrent les composantes phonologiques de la prosodie (accentuation, rythme et

---

<sup>4</sup> A mettre en lien avec les dichotomies « *accent sous-jacent/effectif* » de Di Cristo (1999a ; 2003) ; Garde (1968) et « *stress/accent* » de Bolinger (1989) cf. 1.1.1 du Chapitre 1.

intonation), les contraintes issues des autres domaines linguistiques (notamment le domaine syntaxique), mais également, le nombre et le type de niveaux de constituance. Nous rappelons que nous nous intéressons plus spécifiquement aux niveaux de constituance situés entre le niveau du Mot Prosodique et le niveau du Syntagme Intonatif. En effet, alors que nous savons que les niveaux du Syntagme Intonatif et du Syntagme Accentuel sont relativement consensuels, le niveau du Mot Prosodique, en revanche, apparaît plus ambigu. Nous verrons notamment que les modèles du phrasé prosodique en français ne semblent pas envisager le mot lexical comme une unité de planification, ou de réalisation des règles accentuelles, de surface.

Nous proposons dans ce cadre une présentation chronologique des modèles du français, en commençant tout d'abord par l'approche Intonosyntaxique/Phonosyntaxique traditionnelle. Nous présenterons ensuite les modèles du français inscrits dans l'approche Métrique Autosegmentale, qui se trouvent être majoritaires actuellement, et nous terminerons enfin avec l'approche Métrique et Fonctionnaliste, moins majoritaire mais particulièrement pertinente pour l'étude du français.

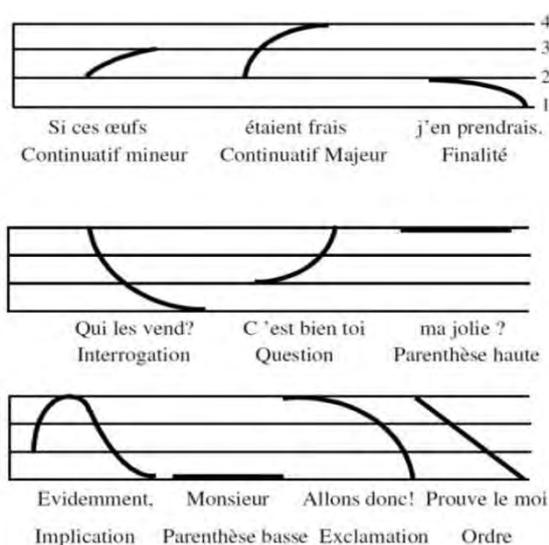
### **2.1 L'approche Intonosyntaxique/Phonosyntaxique**

A l'image de la Théorie Prosodique (Hayes, 1989 ; Nespors & Vogel, 1986 ; Selkirk, 1980 ; 1984 ; 1986), les modèles inscrits dans une approche intonosyntaxique, ou phonosyntaxique, se fondent principalement sur la structure syntaxique pour segmenter le continuum de parole et définir les différents constituants prosodiques. Par ailleurs, alors que certains modèles proposent exclusivement une modélisation de l'intonation, d'autres prennent davantage en compte l'accentuation, en s'inscrivant toujours, toutefois, dans une vision traditionnelle de l'accent en français. Il s'agit plus précisément d'une approche holistique de l'intonation, en opposition à une approche compositionnelle (cf. 1.3.1 du Chapitre 1). La structure intonative est considérée ici comme étant composée d'une suite de contours intonatifs, dont la forme globale a une signification précise. Il n'y a donc pas de prédiction sur l'alignement entre le ton et le texte, ni d'ancrage précis des contours intonatifs sur les syllabes, puisqu'ici le contour intonatif est envisagé comme un mouvement global, réalisé sur un empan plus large que la syllabe. Dans ce contexte, l'accentuation est interdépendante du contour intonatif. Enfin, en ce qui concerne les niveaux de constituance, à l'inverse des approches morphosyntaxiques traditionnelles proposant une hiérarchie exhaustive composée de 6 à 7 niveaux de constituance, tous les modèles présentés dans cette première partie envisagent un système composé uniquement de deux unités

phonologiques, associées pour l'une au marquage accentuel, et pour l'autre au marquage du contour intonatif.

### 2.1.1 Delattre

Le modèle de Delattre (1966b), tout d'abord, considéré comme le modèle précurseur des travaux sur l'intonation du français, décrit les 10 contours intonatifs de base de la langue. Plus précisément, l'auteur se place dans un contexte d'apprentissage du français, et cherche donc à formaliser clairement les différents contours intonatifs utilisés dans la langue pour en faciliter ainsi l'apprentissage. Pour cela, Delattre propose une représentation de l'intonation sous forme de contours globaux qui ont chacun une direction (contour montant, descendant ou plat), un niveau de hauteur variant entre les niveaux 1 et 4, et une forme concave ou convexe (cf. Figure N°17).



**Figure 17 :** Dialogue illustrant les 10 intonations de base du français présentées par Delattre (1966b, p. 6)

Ces 10 contours intonatifs ont été répertoriés suite à une analyse spectrographique des variations de  $f_0$ , d'intensité et de durée, suivie de tests auditifs, basés sur la seule intonation, servant à discriminer tel ou tel contour. Selon l'auteur, chaque intonème correspond à un groupe de sens, ou à un groupe rythmique, auquel est associé un contour intonatif spécifique correspondant à une fonction pragmatique particulière. Cependant, seuls les deux intonèmes « *Continuation Mineure* » (noté « cm ») et « *Continuation Majeure* » (noté « CM ») permettent de rendre compte de la structuration syntaxique, et notamment d'actualiser les

relations de dépendance ou d'enchâssement entre les groupes de sens. En effet, la Continuation Majeure marque un niveau de frontière plus fort que la Continuation Mineure. L'auteur exemplifie l'utilisation de ces deux contours continuatifs avec des énoncés syntaxiquement ambigus (exemple a. issu de Delattre, 1966b, p. 9) ou des énoncés de type « *série en échelon* » (exemple b. issu de Delattre, 1966b, p. 10). Ces deux exemples permettent de montrer, selon l'auteur, le rôle essentiel en français du niveau 4 (CM) par opposition au niveau 3 (cm) : le rôle de la Continuation Majeure étant de réunir de plus petites unités de sens, marquées par une Continuation Mineure, dans une unité plus large.

- a. Il a demandé /<sub>cm</sub> qui parlait //<sub>CM</sub> à Marie (= Il a demandé à Marie qui est-ce qui parle)  
Il a demandé //<sub>CM</sub> qui parlait /<sub>cm</sub> à Marie (= Il a demandé qui est-ce qui parle à Marie)
- b. Si Jean-Marie /<sub>cm</sub> nous invitait /<sub>cm</sub> à déjeuner /<sub>cm</sub> à Montparnasse /<sub>cm</sub> avec son frère  
//<sub>CM</sub> on s'y rendrait

Egalement, alors que la Continuation Majeure est systématiquement marquée par une configuration intonative montante, la Continuation Mineure, en revanche, n'a pas de direction fixe et peut être configurée en contour montant ou descendant. Delattre (1966b, p. 10) considère que cette variation de direction permet de « *briser la monotonie, pour donner de la variété à la ligne mélodique* », sans pour autant avoir une signification particulière.

#### 2.1.2 Di Cristo

A l'inverse du modèle proposé par Delattre, fondé exclusivement sur la modélisation de l'intonation, Di Cristo (1976 ; 1978) envisage un modèle mixte alliant à la fois l'accentuation et l'intonation. Il est toutefois intéressant de préciser que ce modèle est, pour ainsi dire, présenté ici dans sa version initiale. Nous verrons en effet au fur et à mesure de ce chapitre l'évolution des positions théoriques de l'auteur. Ainsi, Di Cristo propose tout d'abord une étude permettant de tester le rapport entre le système prosodique et syntaxique, notamment pour voir si la segmentation prosodique est conditionnée ou non par la syntaxe. Sur la base d'une analyse acoustique et perceptuelle, l'auteur établit une typologie de frontières intonatives permettant de signaler la nature des relations entre les unités syntaxiques :

- Frontière terminale / . /
- Frontière non terminale majeure / ,, /

- Frontière non terminale mineure / , /

Ces trois types de frontières peuvent être associés respectivement aux contours de Finalité, de Continuation Majeure et de Continuation Mineure issus de la typologie de Delattre (1966b). Le choix de ces trois types de frontières est par ailleurs effectué selon une série de règles phonosyntaxiques. Ainsi, en s'inscrivant dans le cadre syntaxique du modèle Générativiste et Transformationnel (Dubois, 1969), l'auteur émet des règles d'alignement entre la prosodie et la syntaxe. Alors que la première règle stipule que les constituants syntaxiques dominés par « P » sont marqués par une frontière prosodique terminale, correspondant à une frontière de Groupe Intonatif (noté « GI »), la seconde règle stipule quant à elle que les constituants qui sont dominés par les symboles catégoriels non terminaux, du type « SN » ou « SV », sont bornés par une frontière prosodique non terminale majeure, correspondant à la frontière d'un Groupe Accentuel (noté « GA »). Par ailleurs, cette dernière règle se trouve amendée lorsqu'un syntagme verbal (noté « SV ») est constitué d'un verbe (hors auxiliaire) suivi d'un syntagme nominal (noté « SN »), auquel cas le SV est alors démarqué par une frontière prosodique non terminale mineure (*cf.* Exemples issus de Di Cristo, 1976, p. 31) :

- a. Les enfants / , , / sont contents / . /
- b. Marie / , , / a redécoré / , / la maison / . /

Di Cristo considère néanmoins que cette troisième règle est principalement conditionnée par des facteurs phonétiques (nombre de syllabes, débit de parole, etc.). En effet, l'auteur rend compte des aspects rythmiques du groupement prosodique, en évitant notamment de créer des groupes prosodiques monosyllabiques. Des analyses ultérieures (Di Cristo, 1978) montrent que les frontières de GI sont marquées par un contour intonatif montant ou descendant, et par un allongement de la syllabe finale suivi éventuellement d'une pause, tandis que les frontières de GA sont associées à un pic ou un creux mélodique sans allongement de la durée de la syllabe finale.

### 2.1.3 Rossi

Directement en lien avec le modèle de Delattre, Rossi (1981 ; 1985 ; 1999) propose d'apporter quelques modifications à la typologie intonative (Delattre, 1966b), en ajoutant notamment de nouveaux intonèmes. Par ailleurs, à l'instar du modèle de Di Cristo, l'auteur

envisage une segmentation de l'énoncé en une succession de Groupes Intonatifs (noté « GI »), auxquels sont associés des contours terminaux spécifiques. Selon Rossi, c'est la position de l'accent lexical qui détermine la place des morphèmes intonatifs et qui délimite le niveau du Groupe Accentuel (noté « GA »). Nous rappelons en effet que l'auteur considère que tous les mots lexicaux en français sont accentués. Autrement dit, Rossi reconnaît ici le statut métrique de l'accent final en français. Concernant l'accent initial, en revanche, l'auteur considère que la prééminence initiale n'a pas de valeur de morphème et ne peut donc pas contribuer à la construction d'intonèmes. Ainsi, sur la base d'une étude pluriparamétrique de la prosodie ( $f_0$ , intensité et durée) puis perceptive, l'auteur établit une typologie de 6 niveaux intonatifs (notamment, Rossi & Chafcouloff, 1972), ainsi qu'une série de différents intonèmes comprenant essentiellement des contours de continuation et de finalité (Rossi, 1981 ; 1985) :

- Suraigu
- Aigu
- Infra-aigu
- Medium
- Grave
- Infra-grave

**CT** = Continuation Majeure marquée par un contour montant dans le niveau infra-aigu, accompagné d'une augmentation de 50% de la durée de la voyelle par rapport à la moyenne des voyelles inaccentuées.

**ct** = Continuation Mineure marquée par une rupture tonale ou un glissando non perceptible dans le niveau médium, associée à un allongement de 50% de la durée vocalique.

**CC** = Conclusif Majeur marqué par chute mélodique dans les niveaux grave ou infra-grave, accompagnée d'un glissando d'intensité de -10 dB et d'un allongement vocalique de 100% par rapport aux syllabes inaccentuées.

**cc** = Conclusif Mineur marqué par une rupture tonale vers le bas, d'un ton en moyenne, utilisé pour marquer la parenthèse.

Rossi propose donc de reformuler la dichotomie « CM » et « cm » (Delattre, 1966b) en « CT » et « ct », respectivement « *Continuation Majeure* » et « *Continuation Mineure* », et de décomposer le contour Conclusif en deux types de contours intonatifs : « *Conclusif Mineur* » (noté « cc ») et « *Conclusif Majeur* » (noté « CC »). Alors que les deux premiers contours

permettent d'unir un énoncé à un autre, les deux derniers contours, en revanche, rompent le lien entre deux unités. Il s'agit donc de rendre compte de la présence ou l'absence d'une relation de dépendance syntaxique entre deux ou plusieurs unités (*cf.* Exemples issus de Rossi, 1985, p.143).

- a. [Il est arrivé]<sub>CC</sub> [mon oncle]<sub>CC</sub> [d'Amérique]<sub>CC</sub>
- b. [Il est arrivé]<sub>CC</sub> [mon oncle]<sub>CT</sub> [d'Amérique]<sub>CC</sub>

Comme nous pouvons le voir dans l'exemple (a), le contour Conclusif Mineur après « *mon oncle* » permet de comprendre que « *d'Amérique* » est le complément du verbe « *il est arrivé* » et non du nom « *mon oncle* ». En revanche, un contour Continuatif Mineur après « *mon oncle* » dans l'exemple (b) permet de comprendre au contraire que « *d'Amérique* » est le complément du nom « *mon oncle* ». Enfin, Rossi (1981) suggère également la prise en compte de nouvelles unités intonatives, telles que le « *Continuatif Appellatif* » qui est réalisé par un glissando dans le haut niveau infra-aigu, associé à un allongement de la durée vocalique similaire au CT, ou la « *Parenthèse* » qui est marquée par une intonation plate dans le niveau infra-grave. Ces deux derniers contours concernent cependant davantage l'acte énonciatif.

### 2.1.4 Martin

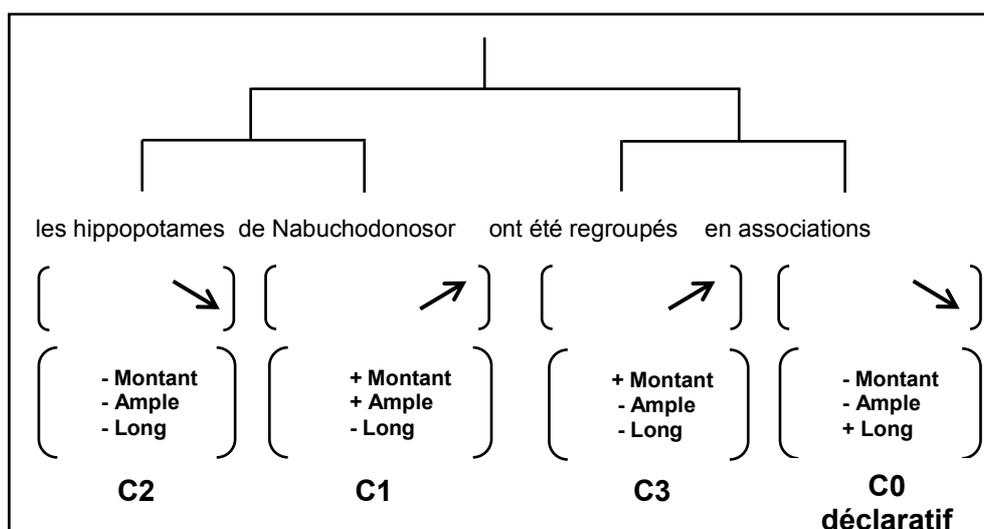
A l'instar du modèle précédent, le modèle de Martin (1975 ; 1981 ; 1987) se trouve également dans la continuité du modèle de Delattre. Selon l'auteur, il s'agit en effet de constituer une séquence de contours intonatifs en se basant non seulement sur les relations de dépendance syntaxique et sur les différentes modalités du discours, mais également sur des principes d'eurythmie. Plus précisément, sur la base de critères formels et structurels seulement, l'auteur envisage un ensemble de traits binaires permettant d'indiquer la direction, l'amplitude et la forme d'un contour. Selon lui, c'est la configuration phonétique de la voyelle accentuée finale d'une unité, appelée « *Mot Prosodique* » (similaire au Groupe Accentuel chez Di Cristo, 1976 et Rossi 1981), qui détermine le type de contour intonatif. Là encore, seul l'accent final permet de générer un intonème, tandis que l'accent initial n'intervient pas dans la construction des morphèmes intonatifs.

- +/- Long : durée de la voyelle accentuée finale et de la consonne voisée éventuelle qui la suit.

- +/- Montant : contours intonatifs montants ou descendants.
- +/- Ample : amplitude relative de la variation tonale (la variation de  $f_0$ ).

On retrouve donc ici certaines des caractéristiques tonales proposées par Delattre (1966b), et plus précisément concernant la direction et le niveau de hauteur du contour intonatif. Toutefois, à l'inverse de Delattre ou même de Di Cristo (1976) et de Rossi (1981), seuls les corrélats phonétiques et acoustiques sont pris en compte, tandis que le phénomène perceptuel n'est pas envisagé. Ainsi, sur la base de l'ensemble des traits binaires, Martin (1981 ; 1987) parvient à faire émerger 9 types de contours intonatifs, dont 3 contours terminaux (noté « C0 ») se distinguant les uns des autres en fonction de la modalité du discours (déclarative, impérative et interrogative), et 6 contours non terminaux (noté « C1 », « C2 », « C3 », « C4 », « C5 » et « C6 ») s'opposant en fonction des traits acoustiques qui les composent. Par ailleurs, l'auteur propose deux règles conditionnant le regroupement des Mots Prosodiques :

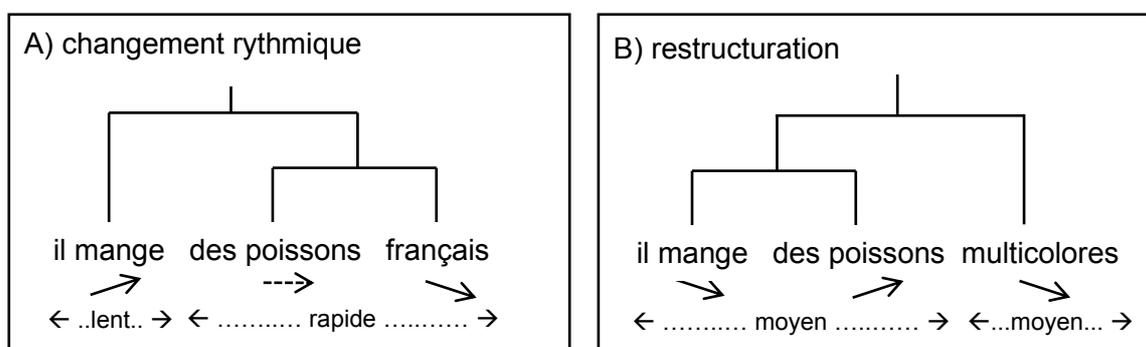
- **Règle de sélection à droite** : l'interprétation de la structure intonative est conditionnée par le contour terminal, ce qui signifie que les 6 contours non terminaux dépendent directement du type de C0.
- **Règle d'inversion de pente** : tout Mot Prosodique qui se joint à un Mot Prosodique à sa droite pour former un groupe plus grand doit être marqué par un contour intonatif de sens opposé à celui du Mot Prosodique auquel il se joint.



**Figure 18** : Exemple d'application de la règle d'inversion de pente dans l'énoncé « les hippopotames de Nabuchodonosor ont été regroupés en associations » (issu et adapté de Martin, 2006, p.152)

Ainsi, comme nous pouvons le voir dans la Figure N°18, le contour terminal est associé à une modalité déclarative et donc à un contour descendant. De là, dépendent alors les autres contours non terminaux (« Règle de sélection à droite »), dont le regroupement est assuré par un système de contraste de pente : la pente montante sur « *ont été regroupés* » permet le regroupement avec « *en associations* » marqué par une pente descendante, puis la pente descendante sur « *les hippopotames* » permet le regroupement avec « *de Nabuchodonosor* » marqué par une pente montante. Ce phénomène d'inversion de pente semble rappeler les propositions de Delattre (1966b) concernant la possibilité d'avoir un contour de Continuation Mineure descendant plutôt que montant, impliquant de ce fait un contraste de pente avec le contour de Continuation Majeure montant. Toutefois, Delattre n'associe pas directement ce phénomène avec une contrainte de dépendance syntaxique, puisque selon lui, il ne s'agit que d'une variante possible destinée à « *briser la monotonie* » (Delattre, 1966b, p. 10).

Enfin, Martin (1987) conçoit également des contraintes d'eurythmie pouvant créer des cas de non-congruence entre la structure prosodique et syntaxique. En effet, le principe d'eurythmie, consistant à équilibrer le plus possible le nombre de syllabes dans chaque groupe prosodique, peut engendrer un découpage ne respectant pas forcément celui de la syntaxe. Martin (2006, p. 149) propose l'exemple de deux énoncés : « *il mange des poissons français* » et « *il mange des poissons multicolores* » (cf. Figure N°19).



**Figure 19 :** Exemple de l'application des contraintes d'eurythmie créant des cas de non-congruence entre la structure prosodique et syntaxique  
 Cas d'une structure prosodique congruente à la structure syntaxique en A) et non congruente à la structure syntaxique suite à l'application de la contrainte d'eurythmie en B) (issu et adapté de Martin, 200, p. 149)

L'exemple (A) montre une structure prosodique respectant le découpage syntaxique, où le locuteur procède à un changement rythmique afin de compenser le déséquilibre en nombre de syllabes entre les deux groupes : une diction plus lente du premier groupe « *il mange* » composé de moins de syllabes, et une diction plus rapide du second groupe « *des poissons*

*français* » composé d'un plus grand nombre de syllabes. Dans l'exemple (B), en revanche, le locuteur privilégie une structure eurhythmique au détriment de la congruence avec la structure syntaxique, et procède donc à une restructuration totale de l'énoncé afin d'équilibrer le nombre de syllabes dans les deux groupes : le premier Mot Prosodique « *il mange* » est cette fois-ci regroupé avec « *des poissons* », tandis que le Mot Prosodique complément « *multicolores* » est isolé, puisqu'il est composé à lui-seul d'un nombre de syllabes quasiment similaire au premier groupe.

Ce dernier point est particulièrement intéressant puisqu'il permet de rendre compte que même si la syntaxe est centrale dans ces approches pour construire les constituants prosodiques, certains auteurs envisagent toutefois la possibilité de structures prosodiques non congruentes aux structures syntaxiques, et ce notamment pour respecter les contraintes liées à l'organisation rythmique. Par ailleurs, nous comprenons à travers la présentation de ces différents modèles que l'intonation reste le paramètre le plus pertinent dans l'élaboration des constituants prosodiques. Bien que l'accentuation soit envisagée dans tous les modèles, le rôle de l'accent reste cependant non déterminant. Il est toutefois important de rappeler que le modèle de Rossi (1980) semble être ici le seul à accorder une propriété métrique à l'accent final. Concernant d'ailleurs le domaine accentuel en français, nous constatons que tous les modèles présentés dans le cadre de cette partie envisagent le groupe de mots comme le domaine de réalisation des règles accentuelles. Martin (1987) parle du niveau du Mot Prosodique, en faisant cependant référence au niveau du groupe de mots. Dans le cadre de notre travail, nous choisissons de nous inscrire dans une approche différente, où l'accent est conçu comme plus indépendant de l'intonation. Nous tenons par ailleurs à ne pas considérer la structure syntaxique comme prédominante dans la détermination de la hiérarchie prosodique, sans pour autant nier une forte dépendance entre les deux structures.

### **2.2 L'approche Métrique Autosegmentale**

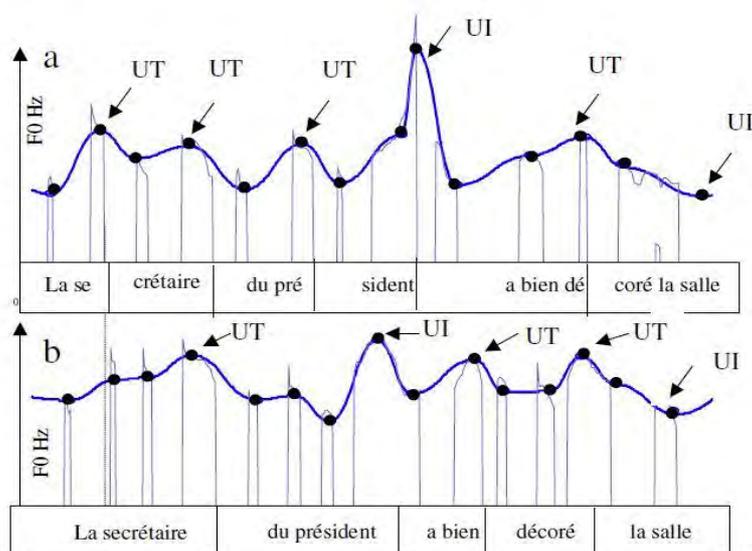
A l'inverse des modèles précédents, où l'intonation est envisagée dans une approche holistique, et où la syntaxe est utilisée pour définir les unités de la structure prosodique, les modèles présentés dans cette seconde partie accordent une place plus importante aux aspects prosodiques, et proposent une analyse compositionnelle du contour intonatif. En effet, directement en lien avec les approches plus « purement » prosodiques (cf. 1.2 du Chapitre 2), les modèles inscrits dans l'approche Métrique Autosegmentale représentent les travaux les plus utilisés actuellement pour l'analyse du français. Il s'agit donc ici de définir les constituants prosodiques sur la base des événements tonals constitutifs du contour intonatif.

Alors que dans l'approche Intonosyntaxique, l'intonation est conçue comme une séquence de morphèmes intonatifs ou comme un contour global, dans l'approche présentée ici, en revanche, l'intonation est élaborée à partir d'une séquence linéaire de tons Hauts ou Bas, ancrés sur les syllabes métriquement fortes (accents et frontières) de l'énoncé. Les modèles tendent alors à considérer l'accentuation comme un phénomène plus indépendant de l'intonation, en se distanciant plus ou moins, selon les approches, de la vision traditionnelle de l'accentuation en français. De la même manière, certains auteurs ne s'inscrivent pas pleinement dans la tradition du modèle standard (Beckman & Pierrehumbert, 1986 ; Pierrehumbert, 1980), et proposent des modèles à la croisée d'autres approches plus morphosyntaxiques ou plus métriques, tout en conservant néanmoins la méthodologie propre à l'approche Métrique Autosegmentale. Concernant enfin les niveaux de constituance, à l'instar du modèle standard, la plupart des auteurs envisagent un système prosodique composé de 2 à 3 niveaux de constituance.

### 2.2.1 Hirst et Di Cristo

Bien que nous ayons choisi d'intégrer le modèle de Hirst & Di Cristo (1984 ; 1996 ; 1998) ou de Di Cristo & Hirst (1993 ; 1996) dans l'approche Métrique Auto-segmentale, il est important de préciser que les auteurs se distinguent de la tradition de Pierrehumbert (1980) et Beckman & Pierrehumbert (1986), en ne proposant pas un modèle purement tonal, puisqu'ils privilégient également les aspects rythmiques de la langue. Dans leur modèle initial (Hirst & Di Cristo, 1984), les deux auteurs admettent l'existence de deux unités phonologiques : l'Unité Tonale (noté « UT ») et l'Unité Intonative (noté « UI »). L'UT est considérée comme l'unité minimale qui est principalement contrainte par la structure métrique de la langue, et délimitée par un ton bas (« L ») à sa gauche et par un ton haut (« H ») à sa droite, produisant alors le gabarit tonal « LH ». Une originalité de cette unité est qu'elle ne présente pas d'équivalent dans les autres modèles du phrasé prosodique en français, puisqu'il s'agit là d'une unité permettant d'intégrer pour la première fois l'accent initial (noté « AI »). En effet, la tête métrique de l'UT peut être marquée aussi bien par un AI (UT initiale) que par un AF (UT finale). Plus précisément, les auteurs considèrent qu'AF marque obligatoirement une frontière d'UT, tandis qu'AI est plus optionnel et peut générer une frontière d'UT dès lors que cela ne crée pas de phénomène de « *clash accentuel* ». Ces UT sont ensuite regroupées dans une unité plus large, à savoir l'UI, qui peut être marquée par deux types de patterns tonals : « LL » lorsqu'il s'agit d'une UI terminale, et par le pattern « LH » lorsque l'unité est cette fois-ci non terminale ou dans une modalité

interrogative. Cette unité majeure correspond par ailleurs au Groupe Intonatif proposé dans les modèles précédents (Di Cristo, 1976 ; Martin, 1981 ; Rossi, 1981).



**Figure 20 :** Segmentations possibles de l'énoncé « la secrétaire du président a bien décoré la salle » selon le modèle de Hirst & Di Cristo (1983)  
(issu de Di Cristo, 2016, p. 115)

La Figure N°20 montre un exemple de deux courbes de  $f_0$  modélisées, associées au même énoncé « la secrétaire du président a bien décoré la salle ». Sur chaque courbe, sont indiqués les moments où la  $f_0$  change de direction (signalés par les cercles noirs), ainsi que la délimitation des UT et des UI (signalée par les flèches). Nous remarquons ici qu'un même énoncé peut être « phrasé » de différentes manières, selon si nous prononçons ou non toutes les UT initiales.

Plus tard, Di Cristo & Hirst (1993) ajoutent une troisième unité à leur modèle et proposent l'Unité Rythmique (noté « UR »), pour pouvoir ainsi rendre compte, en plus du critère tonal, d'un critère relatif à l'organisation temporelle et rythmique des énoncés. Dans ce contexte, les UT ne sont pas directement comprises dans une UI plus large, mais d'abord dans une UR qui est elle-même ensuite contenue dans une UI. Plus précisément, comme nous l'avons mentionné plus haut, les auteurs considèrent que « la réalisation de l'accent final présente un caractère plus systématique que celui de l'accent initial » (Di Cristo & Hirst, 1993, p. 5), étant donné qu'AI est davantage soumis à des contraintes phonotactiques ou rythmiques. Dans ce contexte, les deux auteurs estiment que la prééminence initiale présuppose la réalisation d'une prééminence finale : AI est alors un accent secondaire et AF

un accent primaire. Sur le plan phonétique, les deux accents n'ont pas la même réalisation, puisqu'AF est réalisé à la fois par un mouvement tonal et par allongement de la durée de la syllabe, tandis qu'AI est marqué seulement par un mouvement tonal.

Par conséquent, les deux auteurs proposent un traitement différent des groupes délimités par un AI et par un AF, en considérant que seules les UT marquées par AF peuvent générer une frontière d'UR. Enfin, à l'inverse des autres unités, l'UR (similaire au Groupe Accentuel chez Di Cristo, 1976 ; Rossi, 1981 ou au Mot Prosodique chez Martin, 1975) n'est pas associée à un pattern tonal particulier, puisque cette unité est seulement métriquement motivée, et ne présente pas de correspondance physique spécifique sur le plan tonal (similaire aux propositions de dissociation entre le niveau métrique et phonétique : Ladd, 1996 / 2008) (cf. Figure N°21).

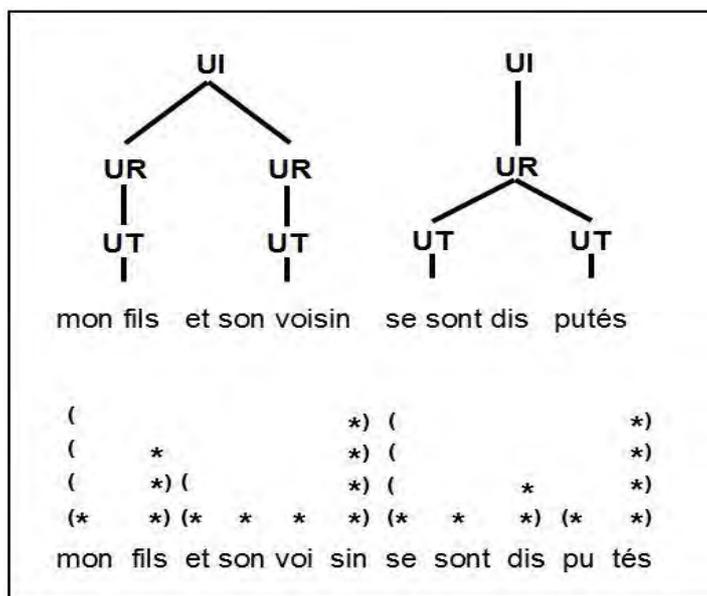


Figure 21 : Représentation arborescente et en grille métrique de l'énoncé « mon fils et son voisin se sont disputés » selon le modèle de Di Cristo & Hirst (1993, p. 6)

Comme nous pouvons le voir dans la Figure N°21, l'énoncé « mon fils et son voisin se sont disputés » est décomposé en 4 UT, dont 3 UT finales (« mon fils », « et son voisin », « -putés ») et 1 UT initiale (« se sont dis- »). Seul le groupe « se sont disputés » se prononce avec deux UT (se sont dis)<sub>UT</sub> putés)<sub>UT</sub>), tandis que le groupe « et son voisin » n'est prononcé qu'avec une seule UT finale (et son voisin)<sub>UT</sub>). En effet, une UT initiale dans ce dernier groupe provoquerait un cas de « *clash accentuel* » entre les deux syllabes accentuées : et son voi)<sub>UT</sub> sin)<sub>UT</sub>. Au niveau supérieur de la hiérarchie, nous constatons ensuite un

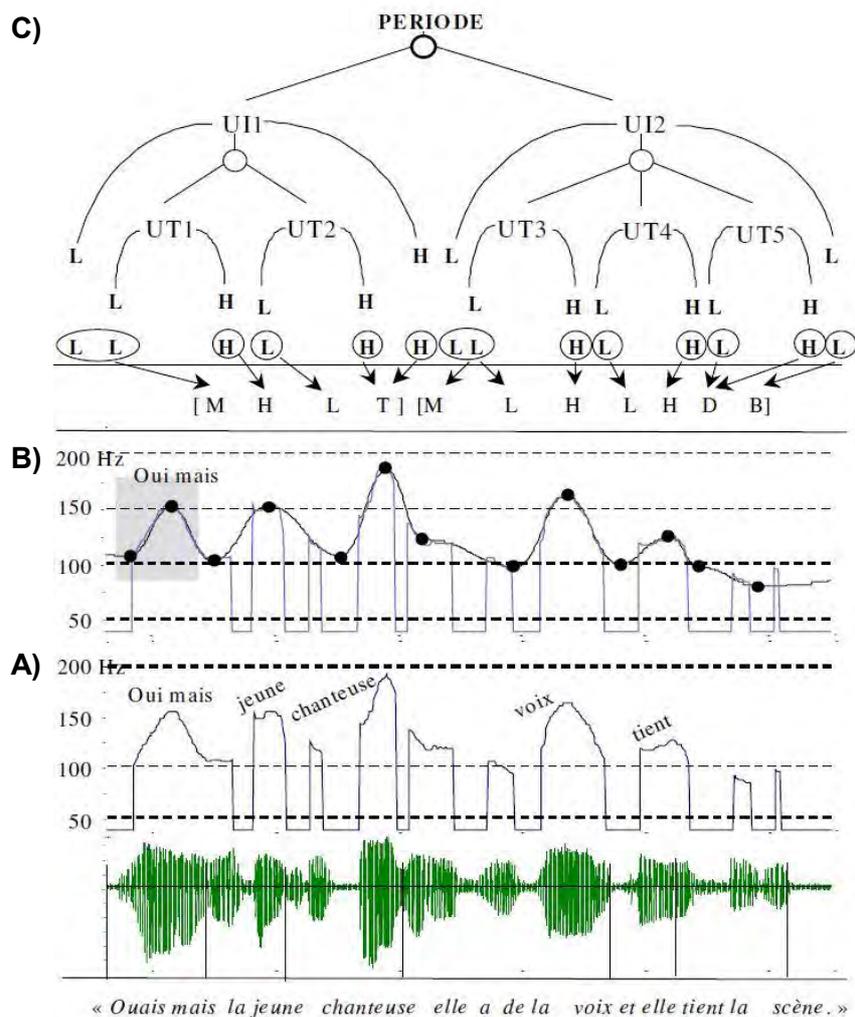
regroupement des différentes UT en 3 UR, dont la dernière qui contient 2 UT étant donné le marquage initial du participe passé « disputés ». Ces 3 UR sont ensuite elles-mêmes comprises dans 2 UI différentes.

Dans une version encore plus récente du modèle (Di Cristo & Hirst, 1996), les auteurs proposent également un autre niveau de constituance supplémentaire, appelé « *Segment d'Unité Intonative* » (noté « SUI »), qui est réservé, cette fois-ci, à des structures syntaxiques spécifiques telles que les structures disloquées ou les focus, dont la réalisation tonale est caractérisée par un plateau haut ou bas :

- a. Ça lui plairait<sub>UI</sub>, une bonne bouteille de champagne<sub>SUI</sub>.
- b. Ça lui plairait<sub>UI</sub>, une bonne bouteille de champagne<sub>SUI</sub> ?

En effet, comme nous pouvons le voir dans les exemples précédents (issus de Di Cristo, 2016, p. 119), la partie soulignée correspond à l'élément en position focale qui est délimité par une frontière d'UI conclusive en (a) et interrogative en (b). Cette partie est suivie d'une séquence post-focale pouvant être identifiée comme un Segment d'Unité Intonative.

Enfin, à la suite de toutes leurs propositions théoriques, Hirst & Di Cristo (1998) développent une méthode de transcription automatique de l'intonation, nommée « *INTSINT* » pour « *International Transcription System for Intonation* ». Ce système se compose de 3 niveaux de représentation : physique, phonétique et phonologique. Le niveau physique (cf. Figure (A) N°22) est le niveau de représentation le plus concret correspondant au tracé de la courbe de *f0* brute. De là, les auteurs proposent une modélisation de cette courbe, générée par l'algorithme MOMEL, pour « *MOdelling MELody* » (Hirst & Espesser, 1993), permettant d'éliminer les perturbations microprosodiques et de proposer une courbe de *f0* lisse et continue, qualifiée de « *représentation phonétique* » (cf. Figure (B) N°22). Sur cette courbe, des points cibles indiquent les changements de direction de la courbe pertinents, pouvant faire l'objet d'une interprétation phonologique (cf. Figure (C) N°22). Plus précisément, l'annotation au niveau phonologique s'opère via l'alphabet INTSINT composé de 8 symboles tonals : 3 tons de référence indiquant l'étendue tonale du locuteur (« T » pour le maximum haut, « B » pour le maximum bas et « M » pour le niveau médian) et 5 tons signalant les variations tonales par rapport aux trois tons de référence (« H » pour Haut, « L » pour Bas, « S » pour « *Same* » ou équivalent, « U » pour « *Upstep* » et « D » pour « *Downstep* »).



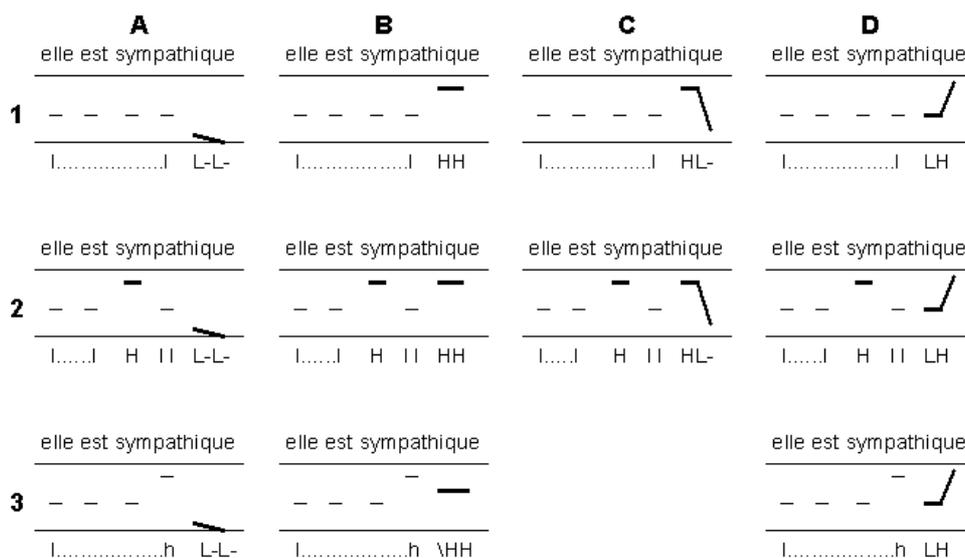
**Figure 22 :** Exemple d'une transcription sous INTSINT de l'énoncé « Oui mais la jeune chanteuse elle a de la voix et elle tient la scène » La partie grisée sur la figure renvoie au signal associé à la formule d'introduction « oui mais » qui n'est pas pris en compte dans la description de la figure par l'auteur (issu de Di Cristo, 2016, p. 162)

Enfin, selon Di Cristo (2013 ; 2016), le système d'annotation INTSINT se différencie du système ToBI, directement issu de l'approche Métrique Autosegmentale (Beckman & Hirschberg, 1994 ; Silverman et al., 1992), en proposant notamment un niveau de représentation intermédiaire, appelé « *représentation phonologique de surface* ». En effet, le système ToBI propose un codage phonologique, basé sur l'observation et l'annotation du signal physique, où chaque ton remplit une fonction phonologique spécifique. Ce mode de transcription nécessite un opérateur humain, dont les connaissances/compétences phonologiques influencent directement la qualité de la transcription. L'approche INTSINT, en

revanche, s'affranchit de l'intervention humaine et privilégie davantage une description purement formelle de l'intonation sans préjuger des aspects fonctionnels.

### 2.2.2 Mertens

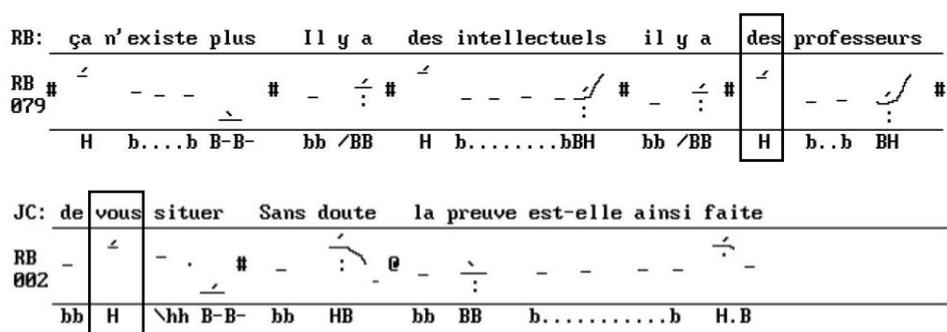
A l'instar du modèle de Hirst & Di Cristo, Mertens (1987 ; 1992 ; 1993a ; 1993b 1997 ; 2004) propose une modélisation de l'intonation qui est aussi relativement différente de la tradition de Pierrehumbert (1980) et Beckman & Pierrehumbert (1986), avec notamment un attachement particulier à l'influence de la structure syntaxique. En effet, ce modèle semble être à la croisée de l'approche Intonosyntaxique et de l'approche Métrique Autosegmentale, en proposant une description compositionnelle de l'intonation, tout en rendant également compte du lien entre la structure prosodique et la structure syntaxique. Concernant tout d'abord le formalisme tonal, Mertens (1987) propose une typologie de tons composée de 15 symboles (*cf.* Figure N°23).



**Figure 23 :** Les différentes configurations tonales possibles associées à l'énoncé « elle est sympathique », proposées dans le modèle de Mertens (2008, p. 96)

Les deux tons Hauts ou Bas peuvent être notés en minuscule pour signaler une syllabe inaccentuée (« h » et « l »), ou en majuscule pour signaler une syllabe accentuée (« H » et « L »). Des diacritiques associés à ces 4 types de tons (« H », « h », « L », « l ») permettent également de spécifier les variations de hauteur : « L- » ou « l- » pour infra-bas et « H+ » ou « h+ » pour suraigu. Enfin, les tons hauts ou bas peuvent aussi être rehaussés (noté /) ou

abaissés (noté \) : « /H » ou « /h » pour haut rehaussé, « \H » ou « \h » pour haut abaissé, « /L » ou « /l » pour bas rehaussé et « \L » ou « \l » bas abaissé. Selon Mertens (1992), l'organisation tonale est directement dérivée de la structure accentuelle, qui elle-même dépend de contraintes morphosyntaxiques. Dans ce contexte, l'auteur considère que les relations hiérarchiques ou de dépendances instanciées par l'intonation doivent refléter la structuration syntaxique. Plus précisément, Mertens (1993a ; 1997) envisage une structure prosodique hiérarchique composée de trois niveaux de constituance : le Groupe Accentuel (noté « GA »), le Groupe Intonatif (noté « GI ») et le Paquet Intonatif. Les niveaux du GA et du GI correspondent à des niveaux de constituance déjà observés dans les modèles précédents (Di Cristo, 1976 ; Rossi, 1981 ; Martin, 1981). Mertens (1993a ; 1997) définit le GA comme une unité composée d'un élément tonique (accentuable), à savoir le mot lexical, et d'un ou plusieurs mots clitiques (unité similaire également à l'Unité Rythmique chez Di Cristo & Hirst, 1993). Selon l'auteur, la syntaxe intervient dès le niveau du GA, dans la mesure où l'organisation des mots clitiques autour des mots non-clitiques s'opère sur la base des relations syntaxiques : seuls les mots clitiques régis par le mot non-clitique forment ensemble un Groupe Accentuel. Selon Mertens (1992), seul AF permet de délimiter un GA, tandis que la réalisation d'un AI est facultative, puisque liée notamment aux aspects sémantico-pragmatiques. Mertens (1993a) montre d'ailleurs qu'AI est indifférent à la nature clitique ou non clitique d'un mot, puisque l'auteur observe la réalisation de proéminences initiales sur des déterminants, des pronoms personnels ou encore des adjectifs possessifs. Ainsi, une syllabe non accentuée sur le plan métrique peut malgré tout être accentuée en structure de surface (cf. Figure N°24).



**Figure 24** : Exemple de réalisation d'un Accent Initial sur des mots clitiques  
 Réalisation d'un Accent Initial sur un déterminant dans la figure du haut ("des professeurs") et sur un pronom personnel relatif dans la figure du bas ("de vous situer") (issu de Mertens, 1993a, p. 17)

Un ou plusieurs GA sont ensuite regroupés dans un GI plus large, dont l'organisation peut varier d'un locuteur à un autre pour un même énoncé. Toutefois, là encore, le regroupement des GA en GI dépend largement de facteurs syntaxiques : les GA réunis dans un même GI doivent être régis par un nœud syntaxique commun, c'est-à-dire faire partie d'une même unité syntaxique (Mertens, 1993b). Prenons l'exemple de l'énoncé « *de la voir ainsi m'attriste beaucoup* » (cf. Figure N°25) qui montre différentes segmentations en GI possibles (a, b et c). Sur la première ligne, désignée « WS » pour « *Word Stress* » ou « *Mot Lexical* », les mots clitiques sont signalés par « . » et les mots non clitiques par un « o ». Au niveau supérieur, indiqué « SG » pour « *Stress Group* » ou « *Groupe Accentuel* », les parenthèses signalent le début et la fin d'un GA délimité par une syllabe accentuée « - ». Ces GA sont ensuite regroupés en GI, au niveau encore supérieur, signalé « IG » pour « *Intonation Groupe* » ou « *Groupe Intonatif* », où les parenthèses indiquent le début et la fin du groupe, et les symboles « + » signalent la réalisation d'un AF. Ainsi, le locuteur peut choisir d'associer un GI à chaque GA comme dans la réalisation (a), ou de regrouper plusieurs GA en un seul GI comme dans la réalisation (b), uniquement dans le cas où ce regroupement ne transgresse pas les règles de dépendance syntaxique comme dans la réalisation (c) : « *ainsi* » ne fait pas partie de la même unité syntaxique que « *m'attriste* ». Par ailleurs, l'auteur parle de la notion « *d'accentuation virtuelle* », c'est-à-dire les AF (« - »), indiqués sur la ligne « SG », qui peuvent ou non être réalisés au niveau du GI : dans la réalisation (b), les AF sur « *voir* » et « *m'attriste* » ne sont pas réalisés au niveau du GI. Cette notion rappelle celle « *d'unité accentuelle virtuelle* » proposée par Garde (1968), selon laquelle l'accent sur un mot peut ou non être réalisé, mais également les propositions faites par Ladd (1996 / 2008) concernant la prise en compte de la réalité métrique d'une proéminence potentielle.

(3)		de	la	voir	ainsi	m'attriste	beaucoup
	WS	.	.	o	.	o	.
	SG	( . . - )	( . - )	( . - )	( . - )	( . - )	( . - )
(a)	IG	( . . + )	( . + )	( . + )	( . + )	( . + )	( . + )
(b)	IG	( . . . + )	( . + )	( . - )	( . - )	( . + )	( . + )
(c) *	IG	( . . + )	( . . - + )	( . - )	( . - )	( . + )	( . + )

Figure 25 : Segmentation de l'énoncé « *de la voir ainsi m'attriste beaucoup* », selon le modèle de Mertens (1993b, p. 3)

Mertens (1992) discute également le phénomène de désaccentuation lorsque deux syllabes finales contiguës sont accentuées. Dans ce cas, la syllabe désaccentuée est celle correspondant à la frontière syntaxique inférieure. En effet, comme nous pouvons le voir dans la Figure N°26, AF sur « *hiver* » est potentiellement accentué au niveau du GA « - »,

mais se trouve être désaccentué au niveau supérieur du GI « . », étant donné le risque de « *clash accentuel* » entre les deux syllabes accentuées « *hiver froid* ». Dans ce cas, c'est AF sur « *hiver* » qui est désaccentué, puisque la frontière syntaxique est ici à un niveau d'enchâssement inférieur à celle réalisée après « *froid* » : (*l'hiver (froid)*).

	( je crains ) ( l'hiver ( froid ) ) ( à Moscou )							
GA	.	-	.	-	-	.	-	avant désaccentuation
GI	.	-	.	-	.	.	-	après désaccentuation

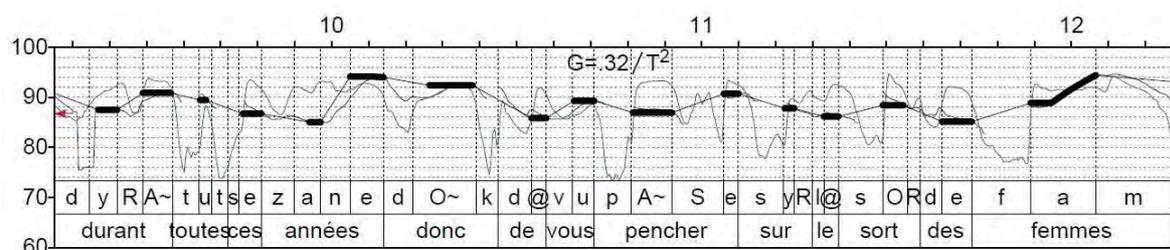
**Figure 26 :** Cas de désaccentuation réalisée sur la syllabe finale du mot « *hiver* » dans l'énoncé « *je crains l'hiver froid à Moscou* » (issu de Mertens, 1992, p. 163)

Par ailleurs, Mertens (1997) prend également en compte le caractère récursif de la structure prosodique en proposant la possibilité de regrouper un GI dans un autre GI. Plus précisément, l'auteur parle de « Paquet Intonatif » consistant à regrouper deux GI successifs lorsque la valeur de la frontière associée au ton AF du premier GI est plus faible que celle du deuxième GI. Comme nous pouvons le voir dans la Figure N°27, lorsqu'une frontière de GI faible est suivie d'une frontière de GI forte, les deux groupes peuvent ensemble former un Paquet Intonatif : le GI « *l'institut* » délimité par un ton bas rehaussé (« /B ») est rattaché au GI « *de phonétique* » délimité par un ton haut (« H ») pour former une unité plus large. En revanche, lorsque deux GI successifs sont marqués par un même degré de frontière, ou lorsque le premier GI est délimité par une frontière plus forte que le GI suivant, aucun regroupement n'est alors possible : le GI « *de phonétique* » est marqué par un ton haut (« H ») qui est plus fort que le ton bas (« B ») délimitant le second GI « *de l'université* », ce qui implique une simple juxtaposition des deux GI plutôt qu'un regroupement. Il est intéressant de mettre en lien ce processus d'empaquetage avec le phénomène « *d'inversion de pente* » proposé par Martin (1981), qui consiste à réunir deux groupes prosodiques dès lors que le premier groupe est marqué par un contour de pente inverse à celui réalisé sur le second groupe (cf. 2.1.4 du Chapitre 3). Nous retrouvons ici en effet le même phénomène puisque qu'un groupe délimité par un ton bas, donc descendant, peut être inclus dans un groupe marqué par un ton haut, donc montant.

[l'institut]	[de phonétique]	[de l'université]	[de Bruxelles]
( /BB	HH	) ( /BB	B-B- )

**Figure 27 :** Exemple d'un regroupement de 4 Groupes Intonatifs en 2 Paquets Intonatifs : [(l'institut) (de phonétique)] [(de l'université) (de Bruxelles)] (issu de Mertens, 2008, p. 100)

Enfin, sur la base de ce système de notation, Mertens (2004) propose une méthode de transcription de l'intonation semi-automatique, appelée « *prosogramme* ». Il s'agit plus précisément d'une représentation graphique simulant la perception de hauteur mélodique. En effet, en se basant notamment sur les travaux de Rossi (1999) concernant la perception des glissandos de  $f_0$ , Mertens (2004) propose un algorithme donnant une estimation de la hauteur perçue par un auditeur moyen. La sortie du prosogramme (cf. Figure N°28) se compose de transcriptions orthographique et phonétique, au-dessus desquelles des lignes parallèles en pointillés, distantes l'une de l'autre de 2 demi-tons, permettent d'indiquer non seulement la différence de hauteur entre deux voyelles signalées par un trait épais sur la courbe de  $f_0$ , mais aussi la variation de hauteur réalisée à l'intérieur même d'une voyelle : il y a une différence de hauteur entre les deux voyelles du mot « *durant* » sans variation interne à chacune des voyelles, signalées alors par un trait plat, alors que pour la voyelle du mot « *femme* », l'inclinaison du trait signale une variation de hauteur interne à la voyelle.



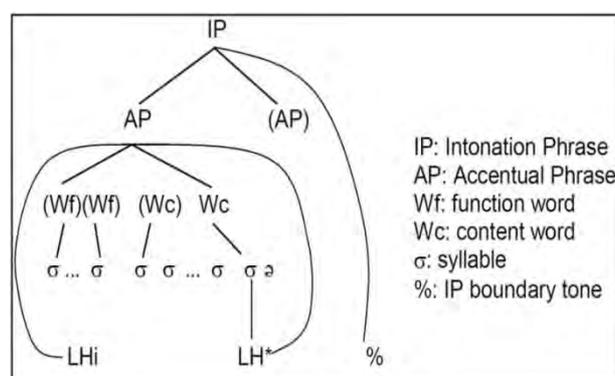
**Figure 28:** Sortie du Prosogramme associé à l'énoncé « *durant toutes ces années donc de vous pencher sur le sort des femmes* » (issu de Mertens, 2004, p. 10)

Comme pour le système INTSINT (Hirst & Di Cristo, 1998), le prosogramme se distingue du système ToBI (Beckman & Hirschberg, 1994 ; Silverman et al., 1992), notamment sur le mode de représentation essentiellement acoustique dans ToBI et qui tend à être plus perceptif dans le prosogramme. En effet, selon Mertens (2004), à l'inverse de la représentation acoustique, la représentation de l'intonation telle qu'elle est perçue correspond mieux aux attentes du linguiste, dans la mesure où elle donne un meilleur aperçu de l'image auditive que l'auditeur reçoit. Concernant plus précisément l'accentuation, la perception permet également de mieux révéler le caractère proéminent ou non d'une syllabe, y compris lorsque l'accent est réalisé au même endroit que la frontière prosodique (en référence aux propositions de Ladd (1996 / 2008) concernant la dissociation entre le niveau métrique et le niveau phonétique cf. 1.2 du Chapitre 3). Quant à la comparaison entre INTSINT et le prosogramme, nous constatons également une différence dans le mode de représentation, puisque INTSINT, à l'instar de ToBI, se base sur une réalisation phonétique.

En revanche, alors que ToBI nécessite des connaissances phonologiques à associer avec les réalisations acoustiques, INTSINT et le prosogramme consistent principalement en une description formelle, relativement indépendante de toute théorie.

### 2.2.3 Jun & Fougeron

En comparaison aux deux précédents modèles, celui de Jun & Fougeron (1995 ; 2000 ; 2002) est un modèle perpétuant davantage la tradition de Pierrehumbert (1980) et Beckman & Pierrehumbert (1986). En effet, comme dans le modèle standard, les deux auteurs considèrent que chaque ton composant le contour mélodique est lié de manière spécifique à la chaîne segmentale : un ton associé à une syllabe métriquement forte constitue un accent mélodique (« *pitch accent* »), et un ton associé à une frontière droite ou gauche d'une unité prosodique constitue un ton de frontière (« *boundary tone* »). Nous retrouvons par ailleurs la même terminologie des niveaux de constituance et le même formalisme tonal. Plus précisément, dans leur modèle initial, Jun & Fougeron (1995) envisagent une structure prosodique composée de deux niveaux de constituance (cf. Figure N°29) : le Syntagme Accentuel (noté « AP ») et le Syntagme Intonatif (noté « IP »).



**Figure 29 :** Structure prosodique hiérarchique proposée dans le modèle de Jun & Fougeron (2002, p. 152)

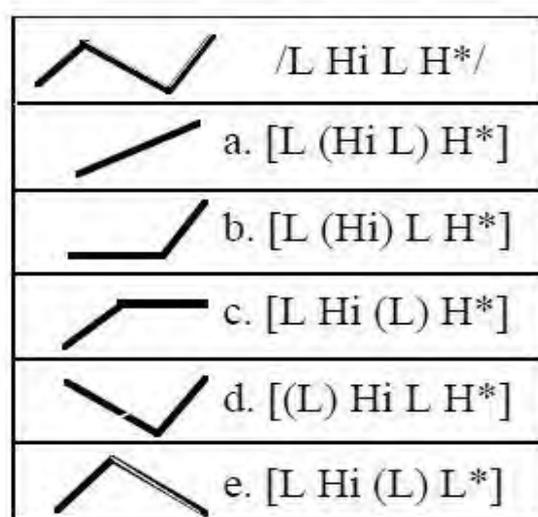
Ces deux unités phonologiques correspondent respectivement au Groupe Accentuel et au Groupe Intonatif proposés dans les modèles précédents (Di Cristo, 1976 ; Mertens, 1993a ; Rossi, 1981). L'AP peut également être assimilé au Mot Prosodique (Martin, 1975) ainsi qu'à l'Unité Rythmique (Di Cristo & Hirst, 1993). En revanche, nous remarquons qu'à l'inverse du modèle de Hirst & Di Cristo, Jun & Fougeron (1995 ; 2000) n'envisagent pas de prendre en compte une unité similaire à l'Unité Tonale (noté « UT »). En effet, les deux auteurs ont

largement contesté la validité de cette unité, en considérant notamment que la segmentation en UT ne respecte pas la règle de l'étagement strict (cf. 1.1.1.1 du Chapitre 2), dans la mesure où les frontières d'UT ne tiennent pas compte des limites de mots. Hirst & Di Cristo (1996) considèrent néanmoins que l'UT peut tout à fait être assimilée à une unité du type Pied Métrique, dont la construction n'est pas remise en question, alors que le Pied ne respecte pas non plus les frontières de mots. Astésano (2001) désapprouve également l'existence d'une UT en français, et notamment le fait qu'une même unité phonologique puisse être construite à partir de deux types d'accents différents, aussi bien sur le plan formel que fonctionnel : l'UT initiale est marquée par une montée mélodique tandis que l'UT finale associe une montée mélodique et un allongement de la durée syllabique.

Sur le plan tonal, l'AP de Jun & Fougeron (1995) est délimité par un « *pitch accent* » bitonal sur la syllabe finale, noté « LH\* » (pour AF), et par un « *phrase accent* » bitonal facultatif sur la syllabe initiale, noté « LHi » (pour AI), donnant le pattern tonal par défaut /LHiLH\*/. L'emploi de « *phrase accent* » est ici différent du modèle de Beckman & Pierrehumbert (1986), où le terme renvoie plutôt à un ton de frontière délimitant le niveau du Syntagme Intermédiaire (cf. 1.2.2.3 du Chapitre 2). Jun & Fougeron (2000) considèrent par ailleurs AI comme un « *phrase accent* » et non comme un « *pitch accent* », étant donné que la réalisation de l'AI n'est pas systématique, et que sa position dans le mot peut varier d'un syntagme à un autre. En effet, selon les auteurs, la position de l'accent secondaire dépend principalement de la présence du mot clitique : lorsque l'AP commence par plusieurs clitiqes monosyllabiques, AI tend à être placé après tous les mots clitiqes, tandis que lorsqu'un ou plusieurs de ces clitiqes sont bisyllabiques, AI est alors réalisé sur l'un d'entre eux. Dans ce contexte, les deux auteurs estiment que seul AF, en ayant une position fixe et constante (dernière syllabe pleine de l'AP), est un marqueur de frontière à part entière.

Par ailleurs, Jun & Fougeron (2000) montrent que le pattern tonal /LHiLH\*/ peut se décliner en cinq variantes potentielles en structure de surface. Plus précisément, cette variation tonale dépend de contraintes temporelles (nombre de syllabes, débit de parole, etc.) qui peuvent parfois entraîner la suppression de certains des 4 tons par défaut : il s'agit dans ce cas-là d'un « *undershoot* ». En effet, lorsque l'AP est constitué, par exemple, d'une ou de deux syllabes, seul AF est réalisé. Dans ce cas-là, le ton « Hi » ainsi que le ton « L » précédant initialement « H\* » sont supprimés (cf. Figure (A) N°30). De la même manière, dans le pattern /LHiH\*/ (cf. Figure (C) N°30), le ton bas associé à la montée finale n'est pas réalisé, ce qui génère alors une configuration en plateau entre AI et AF. Ce type de configuration rappelle par ailleurs la règle de « *tonal spreading* », proposée par Pierrehumbert (1980), qui consiste à maintenir le niveau tonal entre deux tons. Nous constatons également ce phénomène dans le pattern /LLH\*/ (cf. (B) Figure N°30), où cette

fois-ci le maintien du niveau tonal est réalisé entre les deux tons bas. Ces AP peuvent ensuite être contenus dans un IP plus large. Sur le plan tonal, l'IP est marqué par un « *boundary tone* » haut (« H% ») ou bas (« L% ») sur la dernière syllabe. Par ailleurs, lorsque la syllabe finale de l'AP coïncide avec la syllabe finale de l'IP, Jun & Fougeron (2002) considèrent que le ton associé à AF « H\* » s'efface au profit du seul ton de frontière « H% », donnant alors le pattern tonal « LHiLH% » ou « LHiLL% ». Ce phénomène d'effacement de la prééminence finale rappelle la notion de syncrétisme entre l'accentuation et l'intonation (Rossi, 1980), et semble alors contribuer à l'idée selon laquelle le français serait « *une langue sans accent* ».



**Figure 30 :** Schéma de différents gabarits tonals possibles associés au Syntagme Accentuel, dans le cadre de modèle de Jun & Fougeron (2000, p. 216)

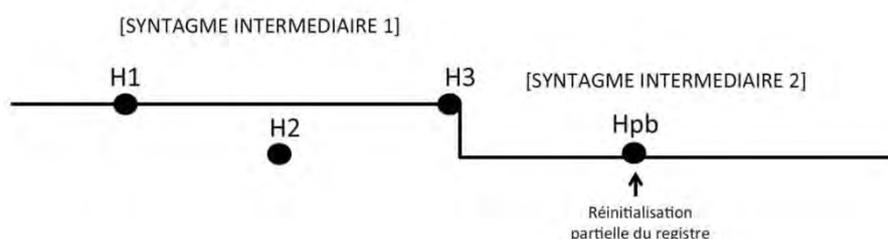
Plus tard, Jun & Fougeron (2000 ; 2002) proposent d'ajouter une troisième unité à leur hiérarchie prosodique, à savoir le Syntagme Intermédiaire (noté « ip »), située entre l'AP et l'IP. D'après les auteurs, ce constituant est réalisé uniquement dans les cas de structures syntaxiques spécifiques, telles que les « *tag-questions* » ou les structures « *thème-rhème disloquée* » (cf. Exemples issus de Jun & Fougeron, 2000, p. 232) :

- a. [(une bonne bouteille de champagne,)ip (ça lui plairait ?)ip ]IP
- b. [(Mon voisin)ip (il est toujours malade)ip]IP

Ce syntagme intermédiaire est délimité par un « *phrasal tone* » (noté « H- » ou « L- ») qui n'est associé à aucune syllabe particulière : le « *phrasal tone* » s'étend depuis le mot « *focus* » en forme de plateau. Nous remarquons que cette unité correspond en tout point au Segment d'Unité Intonative proposé par Di Cristo & Hirst (1996). En revanche, l'ip de Jun & Fougeron ne semble pas conforme à l'ip de Beckman & Pierrehumbert (1986), pour qui l'unité n'est pas réservée aux structures syntaxiques atypiques. De la même manière, Michelas & D'Imperio (2010) considèrent que l'ip en français peut apparaître sur tout type d'énoncé, dès lors que les structures syntaxiques et prosodiques le permettent. Plus précisément, sur la base d'une étude en production, les deux auteurs montrent que l'ip est associé à un niveau de frontière marqué par un allongement de la durée de la voyelle finale, qui est à la fois significativement plus important que celui observé en frontière d'AP, mais également moins important que celui observé en frontière d'IP (cf. Exemple issu de Michelas & D'Imperio, 2010, p. 1) :

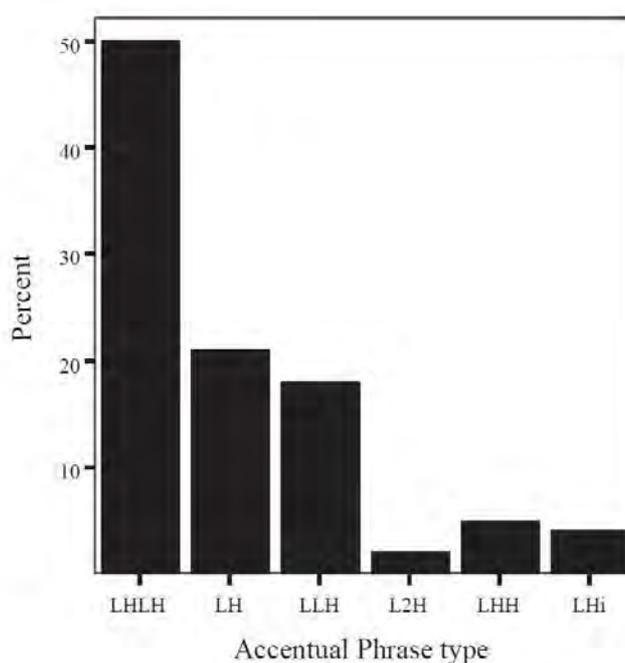
- a. Les grenadiers)<sub>ap</sub> de marrakesh)<sub>ap</sub>]ip ne poussent pas bien vers chez nous}]<sub>IP</sub>

Comme nous pouvons le voir dans l'exemple précédent, l'ip est donc associé à une frontière syntaxique majeure lorsque le syntagme est suffisamment long, c'est-à-dire lorsque le syntagme est composé d'au moins deux AP. Par ailleurs, Michelas (2011) montre également l'existence d'un phénomène d'abaissement (ou « *downstep* ») entre les ip : les ip sont abaissés les uns par rapport aux autres au sein de l'IP. Ce phénomène d'abaissement implique alors une réinitialisation partielle du registre tonal sur le premier ton « H » situé après la frontière droite d'ip (similaire aux propositions de Ladd, 1986 et de van der Berg et al., 1992) (cf. Figure N°31).



**Figure 31 :** Schéma des phénomènes d'abaissement et de réinitialisation partielle entre deux ip successifs (issu de Michelas, 2011, p. 118)

Enfin, sur la base du modèle de Jun & Fougeron, Welby (2003) propose de tester expérimentalement les montées mélodiques associées à l'AP, en s'intéressant notamment à la question de l'alignement entre le mouvement mélodique et la syllabe. Plus précisément, sur la base d'une étude en production (tâche de lecture avec deux débits de parole différents), l'auteur met en évidence 6 types de réalisations tonales possibles pour le niveau de l'AP (cf. Figure N°32), dont la configuration /L<sub>2</sub>H/ qui n'a pas été recensée précédemment dans le modèle de Jun & Fougeron (2000 ; 2002).



**Figure 32 :** Graphique représentant les pourcentages associés aux différents patterns tonals mis en évidence dans l'étude en production de Welby (2003, p. 73)

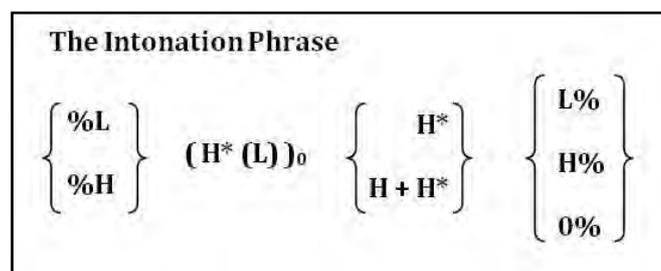
En effet, le pattern /L<sub>2</sub>H/ est observé par l'auteur dans les cas d'un AP en position médiane (un AP au milieu d'autres AP), marqué par une chute de *f0* entre le ton « H » du syntagme précédent et le ton « L *late* » du syntagme suivant. Le ton « L *late* » correspond au ton bas associé à la montée finale (AF) et doit donc être distingué du ton « L *early* » qui est associé à la montée initiale (AI), d'où le formalisme « L<sub>2</sub> » choisi par l'auteur pour signaler le deuxième ton bas. La typologie des configurations tonales mise en évidence par Welby permet également de rendre compte des différences structurelles du « *early rise* » et du « *late rise* », et notamment de la variabilité de l'ancrage du ton « H » associé à la montée initiale, et du ton « L » associé à la montée finale. Le ton « H » peut en effet se réaliser soit sur la première ou deuxième syllabe du mot lexical, et le ton « L » peut être aligné avec

n'importe quelle syllabe suivant le ton « H » initial et précédant le ton « H » final. En revanche, les tons « L » de la montée initiale et « H » de la montée finale sont associées à des entités segmentales particulières. Le ton « L » est aligné à la frontière entre le mot grammatical et le mot lexical, et le ton « H » est réalisé sur la dernière syllabe pleine de l'AP. Cette variabilité de l'ancrage tonal de « H » initial, correspondant à AI, renforce alors l'idée, déjà avancée par Jun & Fougeron (2000 ; 2002), selon laquelle AI et AF ne renvoient pas au même phénomène. Selon Welby (2003), seul AF peut être considéré comme un « *pitch accent* », tandis qu'AI est envisagé comme un « *marqueur flou de frontière* ».

### 2.2.4 Post

Enfin comme pour le modèle précédent, Post (2000) s'inscrit dans la tradition de Pierrehumbert (1980) et Beckman & Pierrehumbert (1986), avec par ailleurs un attachement particulier à la Théorie Prosodique (Hayes, 1989 ; Nespor & Vogel, 1986 ; Selkirk, 1980 ; 1984 ; 1986). En effet, l'auteur propose une analyse compositionnelle de l'intonation associée à une série de règles de bonne formation, établies dans le cadre de la Théorie de l'Optimalité (Prince & Smolensky, 1993), permettant de gérer non seulement l'alignement entre la structure morphosyntaxique et la structure prosodique, mais également l'assignation des accents mélodiques ou « *pitch accents* » (Post, 2000, p. 78). Plus précisément, le modèle se compose de deux unités hiérarchiques : le « Syntagme Phonologique » ou « *Phonological Phrase* » (noté « PhP ») et le Syntagme Intonatif ou « *Intonation Phrase* » (noté « IP »). Le PhP correspond ici au Groupe Accentuel (Di Cristo, 1976 ; Mertens, 1993a ; Rossi, 1981), au Syntagme Accentuel (Jun & Fougeron, 1995) ou encore à l'Unité Rythmique (Di Cristo & Hirst, 1993) proposés dans les modèles précédents. De la même manière, l'IP peut être associé au Groupe Intonatif (Di Cristo, 1976 ; Martin, 1975 ; Mertens, 1993a ; Rossi, 1981). En revanche, alors qu'on associait l'AP de Jun & Fougeron au Mot Prosodique de Martin (1975), étant donné la similitude dans la définition de l'unité en question, Post envisage quant à elle le niveau du « Mot Prosodique » ou « *Prosodic Word* » (noté « PW ») comme un domaine inférieur au PhP, qui correspond au mot lexical seulement. Toutefois, le PW n'est pas considéré ici comme une unité de planification, ou de réalisation des règles accentuelles, à part entière. En effet, le domaine de réalisation des règles accentuelles reste, selon Post, le niveau du PhP. Autrement dit, l'auteur n'inclut pas le PW dans sa hiérarchie des constituants prosodiques. Ce dernier point est par ailleurs observé dans la plupart des modèles du français présentés jusqu'à présent. Le mot est effectivement envisagé comme l'unité accentogène abstraite, et non comme une unité de planification de surface, servant à la structuration de l'énoncé, au même titre que le

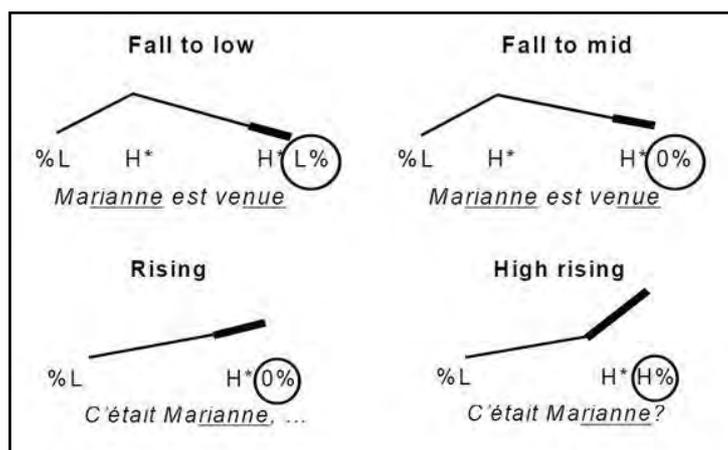
Syntagme Phonologique, ou Accentuel, et le Syntagme Intonatif. Nous reconnaissons donc dans le modèle de Post la terminologie employée dans la Théorie Prosodique concernant les niveaux de constituance, ainsi que le formalisme tonal utilisé dans le modèle standard de l'approche Métrique Autosegmentale (cf. Figure N°33).



**Figure 33 :** Grammaire tonale proposée dans le cadre du modèle de Post (2000, p. 154)

En revanche, l'auteur ne prend pas en compte une unité similaire au « Syntagme Intermédiaire » (Jun & Fougeron, 2000 ; 2002) ou au « Segment d'Unité Intonative » (Di Cristo & Hirst, 1996), puisque selon elle, ce type d'unité n'est pas phonologiquement motivée et complexifie donc inutilement l'analyse de l'intonation de la langue.

Post (2000) propose une grammaire générale résumant le système intonatif du français, où les parenthèses indiquent les éléments facultatifs, et les crochets représentent l'inventaire des éléments tonals possibles. Selon l'auteur, le constituant le plus élevé dans la hiérarchie prosodique est l'IP, marqué par deux « *boundary tones* » en frontière gauche (« %L » et « %H ») et en frontière droite (« L% », « H% », « 0% ») : une frontière tonale haute ou basse pour les deux bords du syntagme, mais aussi une frontière non spécifiée (« 0% ») uniquement pour le bord droit (cf. Figure N°33). En effet, ce ton de frontière « 0% » correspond à un mouvement intonatif montant ou descendant, qui reste cependant à un niveau médian dans le « *pitch range* » (« *middle of the range* »), c'est-à-dire ni au maximum haut de l'étendue tonale (« H% »), ni au minimum bas de l'étendue tonale (« L% ») (cf. Figure N°34).



**Figure 34 :** Configurations tonales possibles des « boundary tones » en frontière de Syntagme Intonatif  
 Comparaison des tons « L% » et « 0% » dans les deux figures du haut, et des tons « 0% » et « H% » dans les deux figures du bas (issu de Post, 2000, pp. 156-157)

Post (2000) envisage également la réalisation de « *pitch accents* » monotonal (« H\* ») ou bitonal (« H+H\* ») (cf. Figure N°33), marquant la frontière du PhP, plus bas dans la hiérarchie que l'IP. Selon l'auteur, le PhP est marqué par un allongement de la durée syllabique pertinent, mais également par un « *pitch accent* » en position finale et initiale du syntagme. Plus précisément, l'auteur propose un certain nombre de règles gérant la distribution et la localisation des « *pitch accents* », et par extension, la construction du PhP (règles similaires à celles proposées par Delais-Roussarie (1995) cf. 2.3.1 du Chapitre 3). Une série de contraintes concerne tout d'abord la réalisation des accents mélodiques finaux, et notamment l'alignement du PhP avec la structure morphosyntaxique, l'attribution du « *pitch accent* », le phénomène de « *clash accentuel* », ou encore la longueur maximale des PhP. Puis une autre série de contraintes est cette fois-ci réservée à la réalisation de l'accent mélodique initial, qui régit l'alignement du « *pitch accent* » sur le bord gauche du PhP, le phénomène de « *vide accentuel* », ainsi que la position de la proéminence initiale dans le mot. Afin de résumer l'ensemble de ces contraintes, l'auteur propose une prononciation « par défaut » du PhP (issu et traduit de Post, 2000, pp. 84-85) :

- I. Les PhP sont toujours marqués par un accent mélodique final, comme en (11).
- II. Des accents mélodiques initiaux et finaux supplémentaires peuvent également être réalisés sur les mots lexicaux tant que cela ne crée pas de phénomène de « *clash accentuel* » à l'intérieur du PhP (12a), où :
  - les accents mélodiques finaux priment sur les accents mélodiques initiaux (12b).

- les accents de syntagmes initiaux (« *phrase accent* ») priment sur les accents internes au syntagme (12c).
- III. Des accents secondaires peuvent être réalisés sur la première syllabe de mots lexicaux composés de plus de 3 syllabes, à moins que le mot commence par une voyelle, comme en (13).

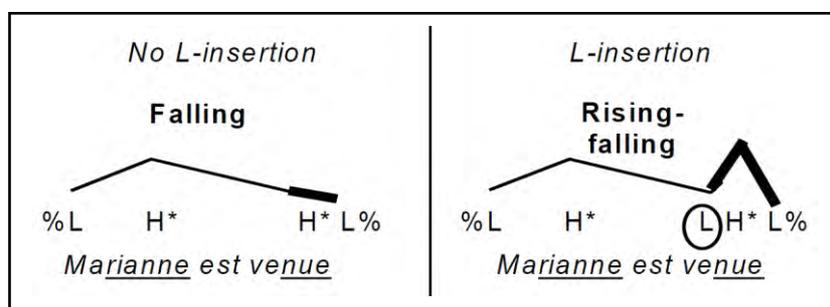
(11) a. (des enfants sages)<sub>PhP</sub> b. (Il disait :)<sub>PhP</sub> (Prends-le)<sub>PhP</sub>

(12) a. (une jolie jupe)<sub>PhP</sub> b. (de petits enfants)<sub>PhP</sub> c. (une assez jolie jupe)<sub>PhP</sub>

(13) a. (la nécessité)<sub>PhP</sub> b. (l'impossibilité)<sub>PhP</sub>

Toutefois, Post (2000) constate une certaine variabilité dans cette prononciation « par défaut ». Il arrive en effet que la hiérarchie des contraintes établie initialement par l'auteur puisse être réorganisée pour correspondre à la structure de surface. Par exemple, lorsque le « *pitch accent* » final est réalisé sur la syllabe initiale plutôt que sur la syllabe finale d'un PW, et que cela donne lieu à ce qu'on appelle un « *arc accentuel* » (Fónagy, 1980) du type « *Beaujolais Nouveau* » (issu de Post, 2000, p. 107), la contrainte d'attribution d'un accent mélodique sur le bord gauche d'un PW prime alors sur celle attribuant un accent mélodique sur le bord droit d'un PW. Ainsi, à l'inverse de Jun & Fougeron (2000 ; 2002) ou même de Welby (2003), Post considère AI comme un accent mélodique au même titre qu'AF. En effet, Jun & Fougeron (2000) estiment qu'AI (« Hi ») peut être promu au rang d'accent mélodique (« H\* ») dans le cas uniquement où AI marque un élément focus, du type « *non, j'ai pas dit déboucher j'ai dit reboucher* » (issu de Post, 2000, p. 178). Autrement dit, la différence entre « Hi » et « H\* » n'est valable que s'il est possible de distinguer phonologiquement l'accent focal et non-focal, ce que Post considère inutile étant donné que, selon elle, ces deux types d'accents présentent la même structure tonale. Plus précisément, l'auteur estime que la différence entre les deux accents, focals et non-focals, est plus graduelle que purement catégorielle. Ainsi AI et AF peuvent être plus proéminents dans les discours emphatiques (augmentation des paramètres accentuels), sans pour autant renvoyer à deux types de tons différents. Autrement dit, l'auteur choisit de ne pas distinguer les fonctions d'ordre pragmatique et structurel des accents en français, et conçoit plutôt un continuum de l'accentuation. Par ailleurs, il est intéressant de remarquer qu'à l'inverse, là encore, de Jun & Fougeron (2000 ; 2002), Post n'envisage pas d'effacement du « *pitch accent* » final lorsque la syllabe finale du PhP coïncide avec la syllabe finale de l'IP, mais plutôt une combinaison de tons associant celui de l'accent mélodique avec celui du ton de frontière, du type « H\*-H% ». Cette conception de l'accentuation tend alors à conforter l'idée selon laquelle l'accent final, bien qu'il soit congruent à la frontière intonative et donc potentiellement moins saillant

perceptivement et acoustiquement, reste malgré tout métriquement fort (Astésano et *al.*, 2012 ; 2013 ; Astésano & Bertrand, 2016 ; Ladd, 1996 / 2008). Le modèle de Post semble donc privilégier une approche avec une visée plus métrique que le modèle de Jun & Fougeron (2000 ; 2002), en s’émancipant davantage de la description traditionnelle de l’accentuation en français. Enfin, un dernier élément tonal est proposé par l’auteur : il s’agit du « *L-insertion* ». Ce ton renvoie à une cible tonale basse réalisée entre deux accents mélodiques, dont la fonction est de bloquer l’abaissement graduel entre les deux accents, dû à un effet de « *downstep* », et ainsi créer un contour montant-descendant où le deuxième « *H\** » présente une valeur égale voire supérieure au premier « *pitch accent* » (cf. Figure N°35).



**Figure 35 :** Effet de la présence du ton « *L-insertion* » sur une séquence « *downstepped* » (issu de Post, 2000, p. 155)

Selon Post (2000), la présence d’un ton « *L-insertion* » est morphologiquement déterminée, dans la mesure où le locuteur, qui choisit d’insérer un ton bas entre deux cibles hautes, tend à réaliser une séparation plus explicite entre les deux items marqués par l’accent mélodique. Nous pouvons toutefois nous questionner quant à l’intérêt structurel de cette réalisation tonale, qui semble par ailleurs essentiellement pragmatique.

Au vu de l’ensemble des modèles présentés dans le cadre de ce chapitre, il est alors aisé de voir comment les modèles inscrits dans la Théorie Métrique Autosegmentale se sont imposés dans le paysage prosodique actuel. En effet, ces modèles proposent une méthodologie très opérationnelle, basée sur l’annotation de la matière sonore, qui permet, à l’inverse des approches Intonosyntaxiques, de distinguer les phénomènes liés à l’accentuation et aux frontières prosodiques. Comme nous l’avons vu précédemment, cette distinction est particulièrement pertinente pour l’étude du phrasé prosodique en français. Nous choisissons donc dans le cadre de notre travail de nous inscrire dans cette approche Métrique Autosegmentale. Plus précisément, nous privilégions un modèle proposant un

système d'annotation à deux tons (Hirst & Di Cristo, 1998 ; Jun & Fougeron, 2000 ; Post, 2000), au lieu d'un modèle plus exhaustif, comme chez Mertens (1997) qui semble tendre vers une trop grande surinterprétation. Nous nous inscrivons par ailleurs davantage dans les modèles de Hirst & Di Cristo et de Post que dans le modèle de Jun & Fougeron, étant donné la visée un peu plus métrique de ces premiers modèles. Nous constatons effectivement une conscience plus importante de la réalité métrique des proéminences potentielles, à travers notamment une meilleure dissociation des phénomènes accentuels et de frontières, et avec également une prise en compte de la durée comme paramètre pertinent dans la construction des constituants prosodiques (*cf.* Unité Rythmique chez Di Cristo & Hirst, 1996). Le modèle de Post est également intéressant du point de vue des autres domaines linguistiques. L'auteur prend notamment en compte les contraintes syntaxiques, sans pour autant considérer que la structure syntaxique contraigne à elle seule la hiérarchie prosodique. En ce qui concerne les niveaux de constituance, à l'instar des modèles Intonosyntaxiques, la plupart des modèles inscrits dans l'approche Métrique Autosegmentale considèrent le niveau du groupe de mots comme l'unité de planification minimale. Seul le modèle de Hirst & Di Cristo propose une unité inférieure au groupe de mots, similaire au Pied tonal (*cf.* Unité Tonale). En revanche, aucun des modèles ne propose d'intégrer le niveau du Mot Prosodique, en tant que mot lexical, dans leur hiérarchie prosodique.

Enfin, bien que nous comprenions l'intérêt de nous situer dans le cadre de la Théorie Métrique Autosegmentale, nous tenons toutefois à ancrer notre travail dans une démarche encore plus métrique, où l'accentuation est considérée comme plus indépendante de l'intonation. Nous choisissons dans ce contexte de nous intéresser également à l'approche Métrique et Fonctionnaliste.

### **2.3 L'approche Métrique et Fonctionnaliste**

Bien que l'approche Métrique et Fonctionnaliste fasse également partie des approches plus « purement » prosodiques, nous constatons en effet que ces modèles accordent une place plus importante à l'organisation métrique qu'à l'organisation tonale de la langue (similaire à la Théorie Métrique : Halle & Vergnaud, 1987 ; Hayes, 1995 ; Liberman, 1975 ; Liberman & Prince 1977). L'accentuation a ici un rôle privilégié et représente même l'élément essentiel et préalable, à partir duquel sont ensuite définis les constituants prosodiques. Il s'agit donc d'une approche purement métrique visant à rendre compte de l'organisation rythmique des proéminences accentuelles, en privilégiant, outre le paramètre tonal, les phénomènes liés à la durée. Nous rappelons en effet que la durée est également

un paramètre pertinent dans la construction des constituants prosodiques, ce que les modèles issus de la Théorie Métrique Autosegmentale semblent, pour certains, négliger. Par ailleurs, à l'instar des approches morphosyntaxiques (Hayes, 1989 ; Nespor & Vogel, 1986 ; Selkirk, 1980 ; 1984 ; 1986) et Intonosyntaxiques (notamment Delattre, 1966b ; Di Cristo, 1976 ; Martin, 1987 ; Rossi, 1981), les modèles inscrits dans une approche Métrique et Fonctionnaliste accordent également une place importante aux informations syntaxiques. Des contraintes d'alignement prosodie-syntaxe, similaires aux propositions faites dans la Théorie Prosodique, sont d'ailleurs envisagées dans un des modèles présentés ici (Delais-Roussarie, 1995 ; 1999 ; 2000). D'autres contraintes liées, cette fois-ci, à la sémantique, au discours ou à la pragmatique sont également prises en compte. En effet, l'intention de communication du locuteur est ici envisagée comme un phénomène pouvant également réorganiser la structure prosodique de surface. Quant aux niveaux de constituance, les auteurs inscrits dans la présente approche envisagent une structure composée entre 2 et 3 niveaux de constituance prosodique.

### 2.3.1 Delais-Roussarie

Dans l'approche de Delais-Roussarie (1995 ; 1999 ; 2000), tout d'abord, la structure des constituants prosodiques est directement déduite de l'organisation accentuelle. En effet, l'auteur considère que l'accentuation joue un double rôle, dans la mesure où l'accent permet non seulement de faire ressortir la structure morphosyntaxique de surface, en démarquant les frontières des constituants linguistiques, mais également d'aider à la construction du rythme en évitant notamment les phénomènes de « *clash accentuel* » ou de « *vide accentuel* » (en référence à la notion de « structure eurythmique »). Dans le cadre de son modèle, Delais-Roussarie considère par ailleurs que des règles ou contraintes de bonne formation sont nécessaires afin de gérer la distribution accentuelle, et par conséquent la construction des constituants. Pour cela, l'auteur propose une approche, dite « *grille métrique seule* », dont les principes relevant de la Théorie de l'Optimalité (Prince & Smolensky, 1993) permettent d'encoder les contraintes prosodiques à la fois de nature morphosyntaxique (ou linguistique) et rythmique, qui interviennent en parallèle dans la construction des constituants prosodiques. L'auteur conçoit en effet une structure prosodique constituée de deux niveaux de constituance : le Syntagme Phonologique, noté « PhP » (similaire au Groupe Accentuel chez Di Cristo, 1976 ; Mertens, 1993a ; Rossi, 1981, à l'Unité Rythmique chez Di Cristo & Hirst, 1993, au Syntagme Accentuel chez Jun & Fougeron, 1995, et au Syntagme Phonologique chez Post, 2000) et le Groupe Intonatif, noté « GI » (similaire au Groupe Intonatif chez Di Cristo, 1976 ; Martin, 1975 ; Mertens, 1993a ; Rossi,

1981, et au Syntagme Intonatif chez Post, 2000). Nous constatons donc ici une similitude entre le modèle de Post (2000), présenté précédemment, et celui de Delais-Roussarie (1995), du fait notamment que Post se soit fondée sur les mêmes règles de bonne formation que Delais-Roussarie. De la même manière, nous reconnaissons ici certains principes de la Théorie Prosodique (*cf.* 1.1.1 du Chapitre 2), qui se distingue toutefois du modèle de Delais-Roussarie, concernant le fait que les constituants prosodiques ne soient pas ici préalables à l'analyse des phénomènes accentuels. En effet, nous rappelons que Delais-Roussarie considère l'accentuation comme primaire.

En ce qui concerne les contraintes de bonne formation, Delais-Roussarie (1995 ; 1999 ; 2000) propose, comme nous l'avons déjà mentionné, deux séries de contraintes. La première concerne l'organisation rythmique de l'énoncé et se fonde sur le principe d'alternance rythmique garantie par le « *principe du contour obligatoire* » (PCO). Dans ce contexte, les syllabes accentuées, qui servent de point d'ancrage pour la construction des groupements, doivent apparaître à intervalles réguliers.

- a. **PCO<sub>Rythme Inf</sub>** : aux niveaux inférieurs de la grille métrique, le rythme doit être binaire ou ternaire afin d'éviter une succession de deux syllabes accentuées et de trois syllabes inaccentuées. Il semblerait que l'énoncé est à ce stade décomposé en mots, qui ne correspondent pas toutefois aux mots lexicaux seuls. Delais-Roussarie et *al.* (2016) parlent dans ce cas-là de Mot Prosodique, comme le regroupement du mot lexical, délimité par une syllabe proéminente, et des mots grammaticaux attachés. Le niveau du Mot Prosodique est donc ici équivalent au Syntagme Accentuel (notamment Jun & Fougeron, 2000 ; 2002), et il ne semble pas, à l'instar du modèle de Post (2000), intégré à la hiérarchie des constituants prosodiques.
- b. **PCO<sub>Rythme Struct</sub>** : aux niveaux supérieurs de la grille métrique, où se construisent les constituants (PhP et GI), le rythme doit être binaire pour ainsi éviter que les constituants émergeant au niveau supérieur contiennent plus de deux constituants générés au niveau inférieur.

A partir de ces deux contraintes, plusieurs grilles métriques différentes peuvent donc être proposées selon le choix de l'alternance rythmique. Comme nous pouvons le voir dans la Figure N°36, l'énoncé « *les cours reprendront le huit septembre* » peut être décomposé en quatre Mots Prosodiques, marqués chacun par une proéminence au niveau inférieur de la

grille métrique<sup>5</sup> (les cours)<sub>Binaire</sub> (reprendront)<sub>Ternaire</sub> (le huit)<sub>Binaire</sub> (septembre)<sub>Binaire</sub>. Parmi ces quatre proéminences disponibles, seules deux sont plus fortement accentuées et projetées au niveau supérieur de la grille métrique (les cours reprendront)<sub>Binaire</sub> (le huit septembre)<sub>Binaire</sub>.

				*				*	
	*			*		*		*	
*	*	*	*	*	*	*	*	*	<*>
le	kUR	RƏ	prã	drõ	lə	ɥi	sep	tã	<brƏ>

**Figure 36 :** Grille métrique de l'énoncé « les cours reprendront le huit septembre » réalisée après l'application des contraintes rythmiques, selon le modèle de Delais-Roussarie (1999, p. 27)

Une deuxième série de contraintes concerne cette fois-ci les contraintes linguistiques gérant l'alignement entre les constituants morphosyntaxiques et prosodiques, mais également le rapport de dominance entre toutes les proéminences, ou encore la force relative entre les différents accents.

- Démarcation<sub>LexD</sub> ou Aligner (Lex, D, σ)** : toute frontière droite d'un mot lexical doit être délimitée par une syllabe accentuée.
- Démarcation<sub>LexG</sub> ou Aligner (Lex, G, σ)** : dans le cas uniquement de certains styles de parole (journalistique, politique, etc.), la frontière gauche d'un mot lexical peut être délimitée par une syllabe accentuée.
- Démarcation<sub>X<sub>Tête</sub></sub> ou Aligner (X<sub>tête</sub>, D, σ)** : la frontière droite des têtes lexicales de projections syntaxiques intermédiaires (X'), correspondant au PhP, doit être délimitée par une syllabe accentuée.
- Démarcation<sub>X<sub>Max</sub></sub> ou Aligner (X<sub>Max</sub>, D, σ)** : la frontière droite des projections syntaxiques maximales (X''), correspondant au GI, doit être délimitée par une syllabe accentuée.
- Dominance<sub>Lex</sub>** : la syllabe finale d'un mot lexical doit être marquée par la proéminence la plus forte de tout le lexème.

<sup>5</sup> Nous remarquons l'extramétricalité du schwa en position finale de « septembre » dans la figure N°36 (voir 1.2.1 du Chapitre 2).

- f. **Dominance<sub>XTête</sub>** : la syllabe finale d'un domaine délimité par une tête lexicale de projection syntaxique intermédiaire (X'), à savoir le PhP, doit être marquée par la prééminence la plus forte.
- g. **Dominance<sub>XMax</sub>** : la syllabe finale d'une projection syntaxique maximale (X''), à savoir le GI, doit être marquée par la prééminence la plus forte.

Ainsi l'énoncé « *les cours reprendront le huit septembre* » (cf. Figure N°37) doit être composé au niveau inférieur de la grille métrique (niveau 2) de quatre accents lexicaux démarquant chacun un Mot Prosodique, en référence aux contraintes Démarcation<sub>LexD</sub> et Dominance<sub>Lex</sub> : (les cours) (reprendront) (le huit) (septembre). Au niveau directement supérieur, seules trois prééminences, sur les quatre disponibles, se projettent pour démarquer les têtes lexicales des PhP, en référence aux contraintes Démarcation<sub>XTête</sub> et Dominance<sub>CTête</sub> : (Les cours) (reprendront) (le huit septembre). Enfin, au dernier niveau de la grille métrique (niveau 1), deux prééminences, sur les trois disponibles, se projettent pour démarquer les frontières droites des GI (en référence aux contraintes Démarcation<sub>XMax</sub> et Dominance<sub>XMax</sub> : (Les cours) (reprendront le huit septembre).

	*							*	
	*			*				*	
	*			*		*		*	
*	*	*	*	*	*	*	*	*	<*>
<i>le</i>	<i>kUR</i>	<i>RƏ</i>	<i>prã</i>	<i>drɔ̃</i>	<i>lə</i>	<i>ʒi</i>	<i>sep</i>	<i>tã</i>	<brƏ>

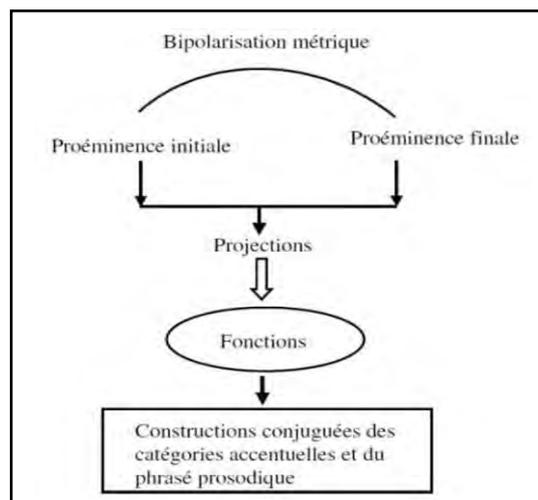
**Figure 37** : Grille métrique de l'énoncé « *les cours reprendront le huit septembre* » réalisée après l'application des contraintes linguistiques, selon le modèle de Delais-Roussarie (1999, p. 26)

Les deux grilles métriques générées à partir des deux séries de contraintes proposent alors différentes organisations accentuelles abstraites qui doivent être appariées, en mettant en interaction les deux hiérarchies de contraintes, afin de faire ressortir une structure de surface respectant au maximum les différentes contraintes imposées (le candidat le plus optimal). Sur le plan phonétique, d'ailleurs, l'auteur montre que les différents niveaux de constituance sont systématiquement marqués par un allongement de la durée de la syllabe accentuée, qui varie proportionnellement en fonction de la force de la frontière (Delais-Roussarie & Feldhausen, 2014 ; Delais-Roussarie et al., 2016). Plus précisément, Delais-Roussarie et al. (2016) mettent en évidence les taux d'allongement qui correspondent à la différence de durée entre les syllabes inaccentuées et accentuées, associés à tous les niveaux de la

hiérarchie prosodique. Les résultats indiquent une augmentation moyenne de la durée de 20% en frontière de Mot Prosodique, de 40% en frontière de PhP et de 98% en frontière de GI. Ces résultats tendent par ailleurs à confirmer ceux obtenus dans des études antérieures (notamment Astésano, 1998 ; 2001).

### 2.3.2 Di Cristo

Dans la continuité de précédents travaux (Di Cristo & Hirst, 1993 ; 1996 ; Hirst & Di Cristo, 1984 ; 1996 ; 1998), Di Cristo (1999a ; 1999b ; 2000 ; 2004 ; 2011) propose un modèle plus spécifiquement centré, cette fois-ci, sur l'organisation métrique et rythmique de la langue. En effet, à l'inverse des modèles inscrits dans une approche Métrique Autosegmentale (notamment Jun & Fougeron, 2000 ; 2002), l'auteur privilégie les composantes prosodiques liées aux phénomènes de durée. Par ailleurs, Di Cristo se distingue du modèle de Delais-Roussarie (1995 ; 1999 ; 2000) et plus largement des approches morphosyntaxiques (Hayes, 1989 ; Nespor & Vogel, 1986 ; Selkirk, 1980 ; 1984 ; 1986), notamment concernant la prise en compte des composantes syntaxiques qui est ici envisagée, mais de manière beaucoup plus modérée. En revanche, à l'instar du modèle de Delais-Roussarie, l'accent est considéré comme l'élément primaire à partir duquel s'élaborent les constituants prosodiques. Selon l'auteur, la projection d'un accent engendre effectivement l'émergence d'un domaine prosodique, dont l'accent projeté constitue la tête. Dans ce contexte, à un constituant prosodique particulier est associé un type d'accent spécifique (similaire au modèle de Beckman & Edwards, 1990). Plus précisément, Di Cristo (1999a ; 1999b ; 2000) conçoit un système composé de trois unités prosodiques de base : l'Unité Tonale (noté « UT »), le Syntagme Prosodique (noté « SP »), qui est le nouveau nom donné à l'Unité Rythmique, et l'Unité Intonative (noté « UI »). Di Cristo propose donc une analyse du lien entre l'accentuation et le phrasé prosodique, en envisageant une modélisation de l'accentuation fondée sur deux niveaux : un niveau sous-jacent (abstrait), où sont déterminées les syllabes potentiellement accentuables, et un niveau de surface (concret), où seules les syllabes effectivement accentuées, en accord avec des règles eurhythmiques ou fonctionnelles, sont projetées (*cf.* Figure N°38).



**Figure 38 :** Schéma illustrant le Principe de bipolarisation accentuelle en structure sous-jacente et la projection des accents en structure de surface  
(issu de Di Cristo, 2016, p. 125)

Cette vision de l'accent rappelle non seulement la distinction proposée par Garde (1968) entre « *accent sous-jacent* » et « *accent effectif* » (en lien avec celle proposée par Bolinger, 1958 entre « *stress* » et « *accent* »), et tend également à rendre compte du caractère « *probabilitaire* » de l'accentuation en français (Fónagy, 1980). Plus précisément, au niveau sous-jacent, correspondant à la représentation phonologique de l'énoncé, deux principes majeurs sont établis par l'auteur : le « *Principe de Bipolarisation* » accentuelle et le « *Principe de Dominance Droite* ». Alors que le premier principe consiste à attribuer une double accentuation potentielle (AI + AF) aux mots lexicaux, le second principe, quant à lui, attribue un statut primaire à AF et secondaire à AI. Le mot est donc considéré comme l'unité accentuelle minimale, ce qui par ailleurs fait écho à la notion « *d'accentème* » proposée par Rossi (1980).

Ces proéminences potentielles peuvent donc ensuite être projetées en structure de surface sous forme de différents types d'accents. L'auteur tient à prendre en compte effectivement la variabilité de la parole, notamment concernant la projection des proéminences qui peut varier d'un locuteur à un autre selon son intention de communication, ou selon la manière dont il prévoit de planifier son discours (similaire à Bolinger, 1972). La proéminence finale, tout d'abord, peut se projeter en surface sous la forme d'un accent nucléaire ou non-nucléaire. L'accent nucléaire sert à segmenter les énoncés en constituants syntaxiques de haut rang, ou en groupes de sens majeurs. Ces accents sont par ailleurs souvent observés en frontière droite de focus ou de structures disloquées (cf. Figure N°39).

Cet accent démarque alors le niveau de constituance prosodique le plus élevé, à savoir l’UI, et représente le point d’ancrage du contour intonatif. Sur le plan métrique, l’accent nucléaire est doté du degré de proéminence le plus élevé au sein de l’UI, et sur le plan phonétique, cet accent est marqué par un allongement important de la syllabe finale qui peut également être suivi d’une pause. Nous constatons alors ici le choix du maintien de l’accent final plutôt que de son effacement en frontière prosodique majeure (en référence à la notion de « *langue sans accent* » Rossi, 1980, soutenue par Jun & Fougeron, 2000 ; 2002). En revanche, nous aurions pu attendre dans l’exemple ci-dessous un niveau d’accent plus élevé sur la syllabe du mot lexical « *livre* ». Il semble effectivement difficile de prononcer cet énoncé sans rehausser le niveau accentuel de cette syllabe. Nous pouvons par ailleurs nous questionner quant à la possibilité d’observer ici la réalisation d’un arc accentuel (cf. Fónagy, 1980) sur le groupe de mots « *livre de peinture* ».

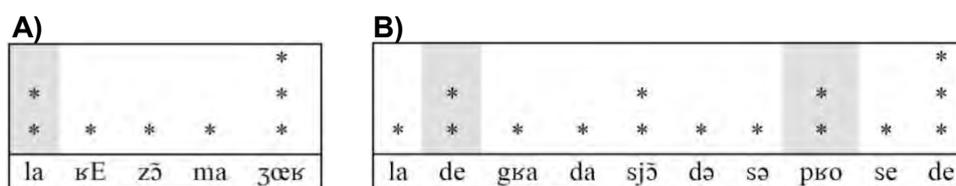
				*											*
	*			*					*						*
	*		*	*				*	*					*	*
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
o	fʁɛʁ	də	fʁɑ̃	swa	ʒə	lʁi	E	O	fɛʁ	ē	liv(ɛ)	də	pē	tyʁ	

**Figure 39 :** Grille métrique associée à l’énoncé « *Au frère de François, je lui ai offert un livre de peinture* »  
 Les frontières d’Unités Intonatives marquées par un accent nucléaire sont indiquées dans les cadres gris (issu de Delais-Roussarie & Di Cristo, à paraître, p. 9)

Quant à l’accent non-nucléaire, il s’agit de l’accent primaire qui marque la frontière droite des UT et des SP. Sur le plan phonétique, l’accent non-nucléaire en frontière d’UT se distingue de celui en frontière de SP, notamment par rapport à l’allongement de la durée syllabique : la syllabe finale en frontière d’UT a une durée dite « normale » ou non marquée (en référence aux syllabes inaccentuées), tandis que la syllabe finale en frontière de SP a une durée plus allongée qui reste cependant inférieure à celle en frontière d’UI. A l’inverse de l’UI, également, l’émergence des frontières de SP a un caractère davantage probabilitaire du fait qu’elle ne soit pas uniquement due à la structure syntaxique ou informationnelle de l’énoncé, mais bien plus à l’intention de communication du locuteur. Ainsi il est possible d’observer, pour une même structure syntaxique, différents types de découpages en SP (cf. Exemples issus de Di Cristo, 2016, p. 132).

- a. J’ai appris<sub>SP</sub> que la politique<sub>UT</sub> était l’art<sub>UT</sub> du compromis<sub>UI</sub>
- b. J’ai appris<sub>UT</sub> que la politique<sub>SP</sub> était l’art<sub>UT</sub> du compromis<sub>UI</sub>

Concernant maintenant la projection en surface des proéminences initiales, il peut s'agir d'un accent emphatique ou non-emphatique. Le premier type d'accent a une fonction de mise en relief sémantique et sert à démarquer le niveau du SP. Sur le plan phonétique, cet accent est marqué par un renforcement de l'intensité, par un allongement de la consonne initiale, mais également par la présence d'une pause précédant la syllabe accentuée. Astésano (2001) montre également que la montée mélodique sur AI emphatique est deux fois plus ample que sur AI non-emphatique. Concernant d'ailleurs la projection de l'AI sous la forme d'un accent non-emphatique, il s'agit d'un accent dont le rôle, principalement rythmique, est de favoriser une structure eurythmique (éviter les phénomènes de « *clash accentuel* » ou « *vide accentuel* »). Sur le plan du phrasé prosodique, AI non-emphatique délimite la frontière gauche de l'UT. Par ailleurs, à l'inverse de l'AF qui est toujours positionné sur la dernière syllabe pleine du mot lexical, ou même de l'AI emphatique qui est systématiquement réalisé sur la syllabe initiale du mot lexical, AI non-emphatique peut s'ancrer sur l'une des syllabes initiales de l'unité : position initiale absolue du syntagme (cf. Figure (A) N°40) ou syllabe initiale du mot accentuable (cf. Figure (B) N°40). Cette variation de l'ancrage de l'AI rappelle notamment les propositions de Jun & Fougeron (2000 ; 2002) et de Welby (2003) concernant le statut de « *phrase accent* » ou de « *marqueur flou de frontière* » de la proéminence initiale en français. Il est également intéressant de signaler dans l'exemple (A) de la Figure N°40 que la proéminence finale sur le mot lexical « *raison* » nécessite un rehaussement du niveau accentuel. En prononçant nous-même le syntagme, nous ancrons en effet naturellement une proéminence à cet endroit, afin d'améliorer l'équilibre rythmique.



**Figure 40 :** Exemple de la variabilité de l'ancrage de l'Accent Initial non-emphatique  
 Réalisation de l'Accent Initial non-emphatique sur le déterminant « *la* » dans l'énoncé « *la raison majeure* » en A), et sur la syllabe initiale des mots lexicaux « *dégradation* » et « *procédé* » dans l'énoncé « *la dégradation de ce procédé* » en B)(issu de Delais-Roussarie & Di Cristo, à paraître, p. 14)

Par ailleurs, ces deux types d'AI, emphatique et non-emphatique (respectivement exemples (a) et (b) issus de Di Cristo, 2011, p. 84), peuvent également être utilisés dans la construction d'arcs accentuels (Fónagy, 1980) qui, nous le rappelons, servent à renforcer la

cohésion sémantique au sein d'un mot ou d'un groupe de mots : Di Cristo (1999a, p. 44) parle de « *principe de bipolarisation à des fins pragmatiques* ». Ce phénomène s'accompagne alors également d'une désaccentuation des syllabes intermédiaires à l'intérieur de l'arc (les syllabes désaccentuées dans l'exemple suivant sont soulignées).

- a. « pour # {**TRANS**porter de la marchan**DISE**}<sub>SP</sub> »
- b. « c'était une # {calamité}<sub>SP</sub> ! »

Delais-Roussarie & Di Cristo (à paraître) considèrent néanmoins que les AF sur les syllabes intermédiaires, internes à l'arc accentuel, ne peuvent qu'être partiellement désaccentuées lorsque la séquence de syllabes est trop importante entre les AI et AF des deux extrémités (cf. Figure N°41). Cette notion renvoie par ailleurs aux phénomènes de « *désaccentuation incomplète* » (Delattre, 1966b) ou « *d'accentuation partielle* » (Rossi, 1980) visant à rendre compte de la « survie » d'un accent lexical sur le plan métrique. Selon Di Cristo & Jankowski (1999), cette trace métrique de l'AF serait essentiellement portée par le paramètre de durée.

			↓			↓			*
	*		*			*			*
*	*	*	*	*	*	*	*	*	*
lə	pʁe	zi	dã	də	la	ʁe	py	blik	

**Figure 41 :** Cas de désaccentuation partielle des Accents Finaux positionnés sur les syllabes internes à l'arc accentuel réalisée dans l'énoncé « le président de la république »  
Les Accents Finaux désaccentués sont signalés par une flèche (issu de Delais-Roussarie & Di Cristo, à paraître, p. 13)

Plus récemment encore, Di Cristo (2016) semble remettre en question son modèle, en discutant notamment la possibilité d'intégrer un niveau de constituance de rang intermédiaire entre l'UT et le SP, qui coïnciderait avec la réalisation d'un accent lexical final. Il pourrait s'agir ici d'un Mot Prosodique ou d'un Syntagme Prosodique Mineur (en référence au Syntagme Phonologique Mineur de Selkirk & Tateishi, 1988). Par ailleurs, l'auteur questionne également l'existence du Syntagme Intermédiaire (Jun & Fougeron, 2000 ; 2002), entre les niveaux de l'AP et de l'IP (respectivement SP et UI dans le présent modèle), en avançant l'idée, selon laquelle il pourrait s'agir d'un constituant similaire à l'IP, ou plus précisément d'une variante de l'IP, réalisé sous une forme récursive (un IP qui contient un

autre IP). En effet, l'auteur met en évidence le fait que les corrélats phonétiques décrits pour l'ip ne diffèrent pas réellement de ceux proposés pour l'IP. Le phénomène « *d'upset* » et de « *partial reset* » associés à l'ip, par exemple, ont été initialement proposés pour le niveau de l'IP (Ladd, 1986 ; Truckenbrodt, 2002 ; van den Berg et al., 1992). Cette hypothèse de récursivité rappelle donc la proposition de Ladd (1986 ; 1996 / 2008) concernant la réduction du nombre de domaines prosodiques, et donc la prise en compte d'unités récursives ou de « *compound prosodic domains* ». Selon Ladd, l'ip peut en effet être considéré comme une unité du même type que l'IP, avec cependant un degré de frontière plus faible. Un certain nombre de travaux en perception ont aussi montré la possibilité de percevoir une granularité importante des niveaux de frontière (notamment de Pijper & Sanderman, 1994 ; Krivokapic & Byrd, 2012 ; Swerts, 1997), potentiellement associés à une seule et même catégorie de constituant prosodique. Le prisme de la perception est aussi un élément fondamental dans le modèle de Di Cristo, qui selon lui, permet « *d'extraire du signal physique les informations les plus pertinentes* » (Di Cristo, 2004, p. 83) servant à l'analyse du phrasé prosodique.

Sur la base de cette approche, Astésano & Bertrand (2016) mettent également en évidence l'importance d'adopter un angle métrique dans l'analyse du phrasé prosodique en français, en réhabilitant notamment le rôle de la durée, en complément de l'analyse tonale de l'accentuation. En effet, Astésano et al. (2012 ; 2013) montrent par exemple qu'un AF peut être marqué seulement par un allongement de la durée de la syllabe qui n'est pas forcément associé à une montée mélodique. De la même manière, les auteurs rendent compte de la perception d'une certaine dissociation des frontières et des proéminences, à travers le fait, notamment, qu'AF reste métriquement fort y compris en frontière prosodique majeure. Cela tend alors à contrecarrer la notion de « *langue sans accent* » (Rossi, 1980) ou « *langue de frontière* » (Vaissière, 1991) selon laquelle l'accent final du groupe de mots en français s'effacerait au profit du seul ton de frontière (Jun & Fougeron, 2000 ; 2002). Leurs résultats indiquent par ailleurs que les syllabes marquées par un AF sont perçues comme métriquement plus fortes que les syllabes inaccentuées à l'intérieur de l'ap (donc en frontière de Mot Lexical), ce qui semble confirmer également le phénomène de « *survie* » de l'accent lexical (cf. notion de « *désaccentuation incomplète* » ou « *d'accentuation partielle* » Delattre, 1966b ; Rossi, 1980). Enfin, concernant AI, Astésano et al. (2012 ; 2013) affirment qu'AI est perçu presque toujours plus proéminent qu'AF, et plus intéressant encore, qu'AI est toujours perçu sur la première syllabe des mots lexicaux, y compris lorsque le pic de la montée mélodique est réalisé sur la seconde ou la troisième syllabe du mot (à l'inverse des propositions de Jun & Fougeron, 2000 ; 2002 et Welby, 2003 sur la variabilité de la position de l'AI). Il s'agit donc ici d'un véritable ancrage métrique de la proéminence initiale au niveau du mot lexical.

Nous comprenons donc, à travers la présentation de ces derniers modèles, l'intérêt pour nous d'adopter une approche à la croisée de la Théorie Métrique Autosegmentale et de la Théorie Métrique et Fonctionnaliste. Comme nous l'avons déjà mentionné, il est particulièrement pertinent dans l'étude du français de s'inscrire dans une approche avant tout métrique. Nous avons en effet besoin de rendre compte du statut phonologique de l'accent dans la langue, et de l'organisation rythmique des prééminences accentuelles en structure de surface, pour ainsi faciliter la dissociation entre les phénomènes de prééminences et de frontières, et par conséquent de mieux rendre compte du lien entre l'accentuation et la constituance prosodique. Nous rappelons par ailleurs que nous tenons à envisager à la fois les paramètres de durée et de  $f_0$  dans l'élaboration des constituants. D'un point de vue méthodologique, également, il est intéressant de pouvoir associer les différentes méthodes utilisées dans les deux types d'approches : recourir à la perception des événements prosodiques, tout en procédant à l'observation et à l'annotation du signal physique, pour ainsi mettre en lien les événements perçus avec leur réalité acoustique. Un événement perçu comme saillant, peut effectivement ne pas être corrélé à des variations acoustiques « typiques » (l'exemple d'un accent final qui n'est pas réalisé par une montée de  $f_0$  et qui peut être perçu comme fort). Enfin, la prise en compte des contraintes sémantico-pragmatiques, syntaxiques ou encore discursives dans les modèles issus de l'approche Métrique et Fonctionnaliste nous semble particulièrement intéressante, et notamment pour l'étude de la parole spontanée. Concernant par ailleurs les niveaux de constituance, les modèles inscrits dans l'approche Métrique et Fonctionnaliste, comme la plupart des modèles présentés dans le cadre de ce chapitre, envisagent le Mot Prosodique comme une unité relevant du niveau sous-jacent, et non pas comme une unité de planification en structure de surface, où s'actualisent les règles accentuelles. Nous verrons dans le cadre de notre étude que nous nous intéresserons particulièrement à cette question de domaine accentuel minimal, et notamment à la pertinence du niveau du Mot Prosodique dans la hiérarchie des constituants prosodiques du français.

## SYNTHESE DU CHAPITRE 3

---

Après avoir présenté dans le Chapitre 2 les grands courants théoriques à l'origine de la description du phrasé prosodique d'une langue, mais également les études destinées à rendre compte de la réalité psycho-cognitive des niveaux de constituance, l'objectif de ce présent chapitre est alors de s'intéresser, plus précisément, aux travaux consacrés à la description du phrasé prosodique en français. Nous commençons par décrire, dans une première partie, les particularités liées à l'accentuation de la langue qui font encore aujourd'hui l'objet de débats parmi les auteurs, et qui contribuent, comme nous pouvons le voir dans la seconde partie du chapitre, à un « flou théorique » concernant le lien entre accentuation et phrasé prosodique.

En effet, le rôle phonologique de l'accent dans la langue est encore aujourd'hui mal établi, que ce soit concernant la nature, le domaine ou la fonction de l'accent, mais également concernant le lien entre l'accentuation et l'intonation. Le français est traditionnellement décrit comme « *une langue sans accent* » (Rossi, 1980) ou « *une langue de frontières* » (Vaissière, 1991) du fait que l'accent en français, à l'inverse des autres langues romanes, n'est pas caractérisé par un accent lexical, mais plutôt par un accent post lexical démarcatif du niveau du groupe de mots. L'accent lexical final est dans ce cas-là désaccentué au profit d'une accentuation de groupe (ou post-lexical), cooccurrence à la frontière prosodique majeure. Certains auteurs, tels que Ladd (1996 / 2008), réfutent cependant le caractère syncrétique de l'accentuation et de l'intonation, en insistant notamment sur le poids métrique des proéminences. Un autre type d'accent est également décrit en français : il s'agit de l'accent initial. Cet accent, décrit comme facultatif, est conçu dans un rôle principalement sémantico-pragmatique (entre autres, Fónagy, 1980 ; Lucci, 1983) mais également rythmique (Garde, 1968 ; Padeloup, 1990a ; Rossi, 1980 ; 1999). D'autres travaux récents en production (Astésano et al., 2007), en perception (Astésano et al., 2012 ; 2013) mais aussi en neuro-imagerie (Aguilera et al., 2014 ; te Rietmolen et al., 2016 ; 2017), tendent en revanche à attribuer un rôle plus structurel à l'accent initial en lui accordant également un véritable statut phonologique, au même titre que l'accent final (Astésano, 2017 ; Astésano & Bertrand, 2016).

Sur la base de cet aperçu des particularités accentuelles de la langue, nous proposons, dans la deuxième partie de ce chapitre, une présentation chronologique des différents modèles du phrasé prosodique en français. L'objectif de cette partie est de recenser les différentes méthodes de représentation et d'étiquetage du continuum de parole,

utilisées par les différents auteurs, mais également de comparer la manière dont les différents modèles envisagent la participation de l'accent dans le marquage de la structure prosodique, gèrent les contraintes issues des autres domaines linguistiques (syntaxe, sémantique, discours, etc.), et construisent leur hiérarchie en constituants prosodiques (nombre et types de niveaux de constituance).

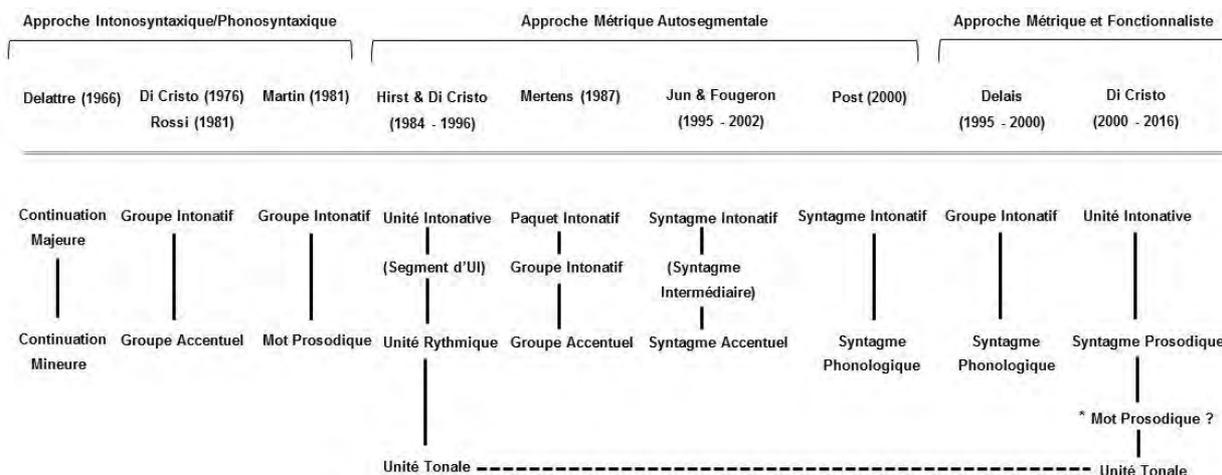
Concernant tout d'abord la prise en compte des autres domaines linguistiques, nous constatons que sur les trois approches présentées dans le cadre de cette partie, deux approches envisagent concrètement la prise en compte de la structure syntaxique, mais également des aspects plus sémantico-pragmatiques. En effet, dans les approches Intonosyntaxiques ou Phonosyntaxiques, la syntaxe occupe une place importante puisqu'elle contribue à définir les différents constituants prosodiques. Toutefois, certains auteurs inscrits dans cette même approche envisagent également la prise en compte d'autres types de contraintes, notamment liées à l'organisation rythmique (Di Cristo, 1976 ; Martin, 1987). De la même manière, l'approche Métrique et Fonctionnaliste, bien que principalement fondée sur la structure métrique et rythmique de la parole, rend tout de même compte de l'influence des informations syntaxiques, mais aussi sémantiques, pragmatiques et discursives, sur la réorganisation prosodique en structure de surface. Enfin, les modèles issus de l'approche Métrique Autosegmentale n'adhèrent pas tous pleinement aux traditions du modèle standard (Beckman & Pierrehumbert, 1986 ; Pierrehumbert, 1980), et proposent pour certains des démarches à la croisée d'autres approches plus morphosyntaxiques et métriques (Mertens, 1987 ; Post, 2000).

Quant au rôle des composantes phonologiques de la prosodie, et plus précisément de l'accentuation, seuls les modèles inscrits dans une approche Métrique et Fonctionnaliste rendent réellement compte de l'organisation accentuelle de la parole, à travers notamment la prise en compte du rôle phonologique de la durée dans la construction des constituants prosodiques. En effet, selon cette approche, les différents constituants sont directement déduits de l'accentuation : un type d'accent spécifique est associé à un niveau de constituance particulier (Di Cristo, 2000). Plus intéressant encore, étant donné les perspectives métriques de cette approche, la force métrique de l'accent final est ici privilégiée au profit de sa réalisation acoustique, facilitant ainsi une description dissociée de l'accentuation et de l'intonation (notamment Astésano et *al.*, 2012 ; 2013). L'accentuation est par ailleurs considérée aussi bien dans sa réalisation initiale que finale (*cf.* « *Principe de Bipolarisation* » : Di Cristo, 2000). L'approche Métrique Autosegmentale tient également compte du phénomène accentuel, avec toutefois une interprétation de l'accent plus tonale que métrique. En effet, les modèles inscrits dans cette approche proposent une analyse du phrasé prosodique principalement basée sur la modélisation de l'intonation, dans une

perspective compositionnelle, en privilégiant le mouvement mélodique au profit d'autres aspects, tels que la durée, pourtant pertinents dans le marquage accentuel du français. Bien que certains auteurs rendent davantage compte du rôle autonome et indépendant de l'accent en français (notamment Hirst & Di Cristo, 1984 et Post, 2000), nous retrouvons en revanche dans certains modèles prosodiques actuels une conception de l'accent encore très proche des descriptions traditionnelles (Jun & Fougeron, 1995 ; 2000 ; 2002). Pour ces derniers, l'accent final tend en effet à s'effacer au profit du ton de frontière, et l'accent initial n'a pas de statut phonologique clair, dans la mesure où il est considéré comme « *un marqueur flou de frontière* » (Welby, 2003). Quant à l'approche Intonosyntaxique ou Phonosyntaxique, étant donné leur approche holistique ou globalisante de la modélisation de l'intonation, l'accentuation est alors totalement soumise et interdépendante au contour intonatif. Certains auteurs envisagent néanmoins de prendre en compte le phénomène accentuel, notamment à travers l'application de contraintes rythmiques (Di Cristo, 1976 ; Rossi, 1985 ; 1999 ; Martin, 1987). Toutefois, seul l'accent final participe dans ce cas-là à la construction des morphèmes intonatifs, tandis que l'accent initial ne peut pas générer d'intonèmes.

Enfin, concernant les niveaux de constituance, nous remarquons que les modèles proposent majoritairement un système prosodique composé de 2 niveaux de constituance, dont les terminologies diffèrent d'un modèle à un autre (*cf.* Figure N°42) : l'un associé à la réalisation de l'accent final (Groupe Accentuel, Syntagme Accentuel, Syntagme Phonologique) et l'autre associé à la réalisation du contour intonatif (Groupe Intonatif, Syntagme Intonatif, Unité Intonative). D'autres auteurs proposent également une unité intermédiaire entre ces deux niveaux de constituance, destinée à rendre compte de structures syntaxiques spécifiques : Segment d'Unité Intonative (Di Cristo & Hirst, 1996) et Syntagme Intermédiaire (Jun & Fougeron, 2000). Toutefois, il reste aujourd'hui difficile de trouver un consensus quant à l'existence de cette unité intermédiaire dans la structure prosodique de la langue (Di Cristo, 2016 ; Ladd, 1996 / 2008 ; Michelas & D'Imperio, 2010). Par ailleurs, certains modèles choisissent d'intégrer un niveau de constituance inférieur au Groupe Accentuel, correspondant au Pied Métrique (pour l'UT de Hirst & Di Cristo, 1984), qui permet de rendre compte de l'accent initial. L'existence d'un niveau entre le Pied et le Syntagme Accentuel n'est en revanche nullement considérée. Bien que la plupart des auteurs conçoivent le mot lexical en français comme l'unité accentogène abstraite, aucun des modèles traditionnels ou plus actuels n'envisage le mot comme une unité de planification de surface à part entière. Par conséquent, le mot lexical n'est jamais intégré dans la hiérarchie des constituants prosodiques. Seul Di Cristo (2016) semble cependant questionner la pertinence de ce niveau. Pour tous les auteurs, le domaine de réalisation des

règles accentuelles reste le groupe de mots (Syntagme Accentuel, Groupe Accentuel, Syntagme Phonologique, Mot Prosodique). Enfin, des propositions ont également été faites sur des unités récursives, et plus précisément par Mertens (1997) qui propose le niveau du « Paquet Intonatif », correspondant à un IP qui contient un autre IP. Di Cristo (2016) envisage également cette récursivité en remettant notamment en doute la validité du Syntagme Intermédiaire.



**Figure 42 :** Récapitulatif des différentes hiérarchies en constituants prosodiques issues des différents modèles du phrasé prosodique en français : approche Intonosyntaxique/Phonosyntaxique, approche Métrique Autosegmentale et approche Métrique et Fonctionnaliste

Le niveau du Mot Prosodique chez Di Cristo (2000 - 2016) est signalé avec une étoile et un point d'interrogation pour indiquer que l'auteur n'intègre pas « officiellement » ce niveau de constituance dans sa hiérarchie prosodique, mais qu'il en questionne toutefois la pertinence (issu et adapté de Di Cristo, 2011, p. 76)

Dans le cadre de notre travail, nous choisissons de nous inscrire dans une approche plus « purement » prosodique, à la croisée de la Théorie Métrique Autosegmentale et de la Théorie Métrique et Fonctionnaliste. Nous tenons en effet à adopter une approche métrique, favorisant une analyse dissociée des phénomènes accentuels et intonatifs. De plus, en clarifiant le statut phonologique de l'accent, cet ancrage métrique nous permet de mieux rendre compte du rôle du mot lexical dans la construction du phrasé prosodique, et de questionner par conséquent son utilité dans la hiérarchie des constituants. Nous choisissons par ailleurs de privilégier, outre le paramètre tonal, les phénomènes liés à la durée qui se révèlent pertinents dans la construction des constituants prosodiques du français. Enfin, nous adoptons également une méthodologie inspirée de ces deux théories, en proposant notamment une démarche acoustico-perceptive.

## CHAPITRE 4 – Les aspects méthodologiques de la thèse

---

<b>1. OBJECTIFS DE LA THESE .....</b>	<b>149</b>
<b>2. SELECTION DU MATERIEL LINGUISTIQUE .....</b>	<b>154</b>
2.1 Le corpus de parole contrôlée : Corpus d'Edimbourg .....	154
2.2 Le corpus de parole semi-spontanée : Corpus de la Map Task Aix .....	156
2.3 Le corpus de parole spontanée : Corpus du CID.....	160
2.4 Préparation du corpus de parole sélectionné .....	162
<b>3. TEST DE PERCEPTION DE LA PAROLE.....</b>	<b>164</b>
3.1 Participants .....	164
3.2 Tâches de perception.....	164
3.3 Préparation de la base de données.....	166
<b>4. HYPOTHESES ET PREDICTIONS .....</b>	<b>169</b>
4.1 Hypothèse 1 : Une granularité plus fine des niveaux de constituance .....	170
4.2 Hypothèse 2 : Pas d'affaiblissement perceptif de l'AF en frontière majeure .....	174
4.3 Hypothèse 3 : Un rôle plus structurel que purement rythmique de l'AI.....	176
<b>5. DEMARCHE STATISTIQUE .....</b>	<b>178</b>
5.1 Test d'indépendance du Khi deux .....	178
5.2 Modèle de régression ordinale à effet mixte .....	179



### 1. OBJECTIFS DE LA THESE

L'objectif de notre travail est d'explorer l'organisation du phrasé prosodique en français, à travers notamment le lien entre l'accentuation et la constituance prosodique. Nous avons vu en effet qu'il est difficile de trouver un consensus, dans toutes les langues en général, non seulement sur le choix et le nombre des niveaux de constituance reflétant la hiérarchie prosodique, sur les règles de formation de ces groupes prosodiques, mais également sur la manière dont les unités sont nommées.

En français, plus spécifiquement, seuls deux niveaux de constituance entre le niveau de la Syllabe et l'Énoncé sont communément admis : le Syntagme Intonatif (IP) contenant un ou plusieurs Syntagmes Accentuels (AP). L'existence d'un Syntagme Intermédiaire (ip) a également été débattue, mais là encore sans consensus. Certains auteurs vont même jusqu'à remettre en doute la validité de cette unité intermédiaire dans la langue, et parlent plutôt de récursivité de la structure prosodique, en considérant que l'ip n'est rien de plus qu'un IP contenu dans un autre IP (Di Cristo, 2016 ; Ladd, 1996 / 2008). Enfin, aucune des hiérarchies proposées pour le français n'envisage de prendre réellement en compte le Mot Lexical comme une unité de planification, ou de réalisation des règles accentuelles, de surface. Dans la plupart des modèles, le Mot Prosodique, constitué soit du mot lexical seulement, soit du mot lexical et des mots grammaticaux attachés, n'est jamais intégré dans les hiérarchies prosodiques. Cette question autour de la validité du niveau du mot dans la planification de la structure prosodique est par ailleurs étroitement liée aux particularités accentuelles du français.

Nous rappelons en effet que l'accent final (AF) principal en français est considéré comme cooccurrent à la frontière prosodique majeure, ce qui a d'ailleurs entraîné certains auteurs à questionner le rôle phonologique de l'AF, et à caractériser le français comme une « *langue sans accent* » (Rossi, 1980) ou une « *langue de frontière* » (Vaissière, 1991). Ce phénomène de syncrétisme entre accentuation et intonation s'observe également dans les théories prosodiques plus récentes, notamment dans le cadre de l'approche Métrique Autosegmentale, où AF est décrit comme un accent mélodique, marqueur de la plus petite unité de la hiérarchie prosodique, s'effaçant au profit du ton de frontière lorsqu'il se trouve en frontière d'unité prosodique majeure (Jun & Fougeron, 2002). En dehors de la prééminence finale, un accent initial (AI), supposé secondaire et optionnel, a également été décrit pour marquer le niveau du groupe de mots en français. Là encore, les auteurs ne s'accordent pas quant au rôle phonologique de cet accent. Quand certains considèrent AI comme un accent emphatique ou énonciatif, dépendant de critères sémantico-pragmatiques (Fónagy, 1980 ; Lucci, 1983), d'autres auteurs estiment qu'il s'agit d'un accent rythmique apparaissant sur les

longs constituants (cf. « *écho d'accent* » de Garde, 1968 ; « *ictus mélodique* » de Rossi, 1980 ou « *accent secondaire* » de Pasdeloup, 1990a). Aujourd'hui encore, AI est considéré comme un marqueur accentuel optionnel du niveau de l'AP, influencé par différents facteurs rythmiques, discursifs ou pragmatiques. Cet accent optionnel n'est par ailleurs pas considéré comme un accent mélodique, au même titre qu'AF, car il n'est pas lié à une syllabe spécifique (cf. notion de « *marqueur flou de frontière* » Welby, 2003).

Toutefois, certains auteurs s'accordent à dire que les mots lexicaux en français sont tous générateurs d'accents, et possèdent donc tous une propriété accentuelle (notion « *d'accentème* » Rossi, 1980). Autrement dit, ces auteurs reconnaissent la propriété métrique de l'accent en français (notamment l'accent final ici), en envisageant qu'une syllabe peut être métriquement forte sans être forcément accentuée (Ladd, 1996 / 2008). Delattre (1966b, p. 72) considère d'ailleurs que « *l'accent du mot peut en partie survivre dans le groupe* », et parle même de « *désaccentuation incomplète* » (ou « *d'accentuation partielle* » Rossi, 1980). Plus récemment encore, dans les modèles issus de l'approche Métrique Autosegmentale majoritaire, Post (2000) montre qu'AF conserve son statut d'accent mélodique même lorsqu'il est cooccurrent à une frontière prosodique majeure. Di Cristo (1999a ; 1999b ; 2000), quant à lui, estime que l'accent initial et l'accent final coexistent selon « *le Principe de Bipolarisation* » accentuelle dans la représentation de la structure sous-jacente de tous les lexèmes du français. Concernant plus précisément la prééminence initiale, une étude en production sur de la parole contrôlée (Astésano et al., 2007) met en évidence un rôle plus structurel que purement rythmique de l'AI. Selon les auteurs, la prééminence initiale marque la structure prosodique de manière plus privilégiée que la prééminence finale, spécifiquement à des niveaux inférieurs de la hiérarchie prosodique, proches du mot lexical. Nous comprenons donc que le mot lexical est envisagé par ces auteurs comme l'unité accentogène abstraite. Cependant, malgré la reconnaissance du statut métrique de l'accent, aucun des modèles du phrasé prosodique en français ne conçoit le mot comme une unité de planification de surface, servant à la structuration de l'énoncé, au même titre que le Syntagme Accentuel et Intonatif.

Dans ce contexte, nous choisissons d'adopter dans notre travail une approche à la croisée de la Théorie Métrique Autosegmentale et de la Théorie Métrique et Fonctionnaliste. Il s'agit plus précisément d'une démarche avant tout métrique, en complément d'une méthode centrée davantage sur la réalisation physique ou acoustique des indices prosodiques. Comme nous l'avons vu, seule une analyse métrique de l'accentuation permet de rendre compte du rôle phonologique de l'accent en français, et par conséquent du lien entre accentuation et constituance. D'un point de vue méthodologique, l'approche Métrique Autosegmentale propose une méthode opérationnelle basée sur l'observation et l'annotation

de la matière sonore, tandis qu'une approche plus purement Métrique, ne se mesurant pas physiquement, nécessite une méthodologie basée sur une démarche plus abstraite de la relation entre l'accentuation et le phrasé prosodique. Nous rappelons par ailleurs que nous choisissons de prendre en compte le rôle du paramètre de durée dans la construction des constituants prosodiques, ce que les modèles issus de la Théorie Métrique Autosegmentale semblent pour la plupart négliger. Ainsi, nous envisageons de nous fonder sur la réalité perceptive des événements prosodiques, en dirigeant l'attention des auditeurs non plus sur la réalité phonétique de surface, mais davantage sur le poids métrique de ces événements. Autrement dit, « *extraire du signal physique les informations les plus pertinentes* » (Di Cristo, 2004, p. 83) intervenant dans l'élaboration du phrasé prosodique. Nous proposons par la suite de mettre en relation les événements prosodiques perçus avec leur réalité physique ou acoustique. La perception est donc utilisée ici comme une interface entre le niveau acoustique concret et le niveau phonologique abstrait. A l'instar de Mertens (2004), nous considérons effectivement que la perception donne un meilleur aperçu de l'image auditive que l'auditeur reçoit, et donc une meilleure idée de ce que l'auditeur utilise pour segmenter et hiérarchiser le flux de parole.

Notre travail est également dans la continuité de précédentes études en perception sur de la parole contrôlée (Astésano et *al.*, 2012 ; 2013) et spontanée (Roux et *al.*, 2016), où les auteurs se sont intéressés à la manière dont les auditeurs francophones perçoivent les proéminences et les frontières à chaque niveau de la hiérarchie prosodique. Les résultats de l'étude sur la parole contrôlée (Astésano et *al.*, 2012 ; 2013) montrent tout d'abord que les francophones ne sont pas « *sourds* » aux accents, puisque les auditeurs ont systématiquement perçu comme proéminentes les syllabes à l'initiale et à la finale des mots. Les auteurs rendent également compte d'une « dissociation partielle » entre la perception des frontières et des proéminences, dans la mesure où un même score de proéminence peut correspondre à différents niveaux de frontière, à l'exception de la frontière d'ip où un renforcement de l'accent final est observé. Plus précisément, l'analyse du poids relatif des deux types d'accents montrent qu'AI est perçu plus fort qu'AF lorsque le lexème se trouve en frontière d'AP, et plus faible qu'AF lorsque le lexème se trouve en frontière d'IP. En revanche, pour les frontières d'ip ou inférieures à l'AP, proches du Mot Lexical, la participation de l'AI et de l'AF semble être équivalente. Ces résultats vont donc à l'encontre de l'hypothèse du syncrétisme entre l'accentuation et l'intonation en frontière prosodique majeure (Rossi, 1980), et valident, en revanche, l'hypothèse de « *désaccentuation incomplète* » (Delattre, 1966b ; Di Cristo, 1999a ; 1999b ; 2011 ; Rossi, 1980), où le lexème garde une trace de l'existence d'un accent de mot. Enfin, concernant AI, les résultats de l'étude confirment l'ancrage métrique de la proéminence initiale en français, ce qui a par

ailleurs été conforté via des études en neuro-imagerie (Aguilera et *al.*, 2014 ; te Rietmolen et *al.*, 2016 ; 2017). Sur la base de l'ensemble de ces travaux, Astésano (2017) propose qu'Al et AF soient les marqueurs d'une unité prosodique cohésive, qui à l'inverse des propositions actuelles, ne se situerait pas au niveau du groupe de mots (Syntagme Accentuel ou Syntagme Phonologique) mais au niveau du mot lui-même. L'auteur envisage dans ce contexte la possibilité d'intégrer le mot lexical à la hiérarchie des constituants prosodiques. Enfin, l'étude en perception sur la parole spontanée (Roux et *al.*, 2016) montre que les auditeurs francophones perçoivent les frontières et les proéminences, avec cependant, un jugement plus consensuel pour la perception des frontières. Ce résultat confirme également ceux obtenus dans de précédentes études (Cole et *al.*, 2014 ; Cole et *al.*, 2010a ; Cole & Shattuck-Hufnagel, 2011 pour l'anglais ; Smith, 2009 pour le français).

Dans le cadre de notre étude, nous nous distinguons des précédents travaux en perception quant à la manière dont nous envisageons l'analyse des frontières et des proéminences. La plupart des études actuelles testent en effet la perception des proéminences et des frontières dans un énoncé sur le plan paradigmatique seulement, en s'intéressant à une catégorie d'événements prosodiques spécifique (frontière majeure, accent nucléaire, etc.), sans réellement mettre en relation ces différents événements avec les autres événements prosodiques adjacents. Nous privilégions donc dans notre travail une étude testant la perception des frontières et des proéminences à la fois sur le plan paradigmatique et syntagmatique. Plus précisément, nous proposons de mettre en relation, au sein d'un même énoncé ainsi qu'à travers les différents énoncés, les proéminences avec les frontières, mais également les différents niveaux de frontière ou de proéminence entre eux. L'objectif de cette analyse est effectivement de mieux comprendre comment ces différents événements prosodiques s'organisent dans la construction même du phrasé prosodique d'un énoncé. A l'instar également d'autres études en perception (Carlson et *al.*, 2005 ; de Pijper & Sanderma, 1994 ; Krivokapic & Byrd, 2012 ; Swerts, 1997), nous nous intéressons à la capacité des auditeurs à percevoir non seulement la présence et l'absence des frontières et des proéminences, mais surtout leur capacité à percevoir une granularité des degrés de proéminence et de frontière.

En ce qui concerne nos questions de recherche, nous nous interrogeons, comme nous l'avons vu, sur les niveaux de constitution perçus en français, et notamment sur le niveau du Mot Prosodique, en tant que mot lexical, jusque-là écarté des analyses phonético-perceptives :

- **Combien de niveaux de constitution sont perçus ?**

- **Est-ce que chaque niveau de frontière est associé à un niveau de constituance spécifique ? Ou alors, est-ce qu'un même niveau de constituance est marqué par différents degrés de frontière ?**
- **Enfin, est-ce que le Mot Prosodique a une utilité particulière dans l'élaboration du phrasé prosodique ?**

Nous nous interrogeons également sur le rôle phonologique de l'accent dans la langue. Concernant tout d'abord AF, nous questionnons en effet sa contribution au marquage de la structure indépendamment de la frontière prosodique :

- **AF est-il vraiment « écrasé » perceptivement par la frontière prosodique majeure ?**
- **La force de l'AF varie-t-elle en fonction du niveau de constituance ?**

Quant à l'AI, nous nous intéressons à sa fonction structurale, en remettant en question son rôle purement rythmique mais aussi son statut d'accent secondaire :

- **AI est-il réalisé et perçu uniquement sur les longs constituants ?**
- **A quel niveau de constituance AI contribue-t-il spécifiquement à la segmentation du flux de parole ?**

Nous tenons par ailleurs à tester la position de l'AI en français, au regard de son statut de « *marqueur flou de frontière* » (Jun & Fougeron, 2000 ; 2002 ; Welby, 2003) :

- **La position de l'AI varie-t-elle en fonction de la taille du constituant ?**

Enfin, sur la base des résultats perceptifs, nous proposons de mettre en correspondance les événements perçus avec leur réalisation acoustique, pour ainsi éclairer certains de nos résultats perceptifs au regard de l'acoustique. Les travaux sur la réalité psycho-cognitive de la constituance prosodique montrent effectivement que l'allongement vocalique et la présence de pauses sont souvent corrélés à la perception d'une frontière prosodique majeure (Cole et al., 2014 ; Delais-Roussarie & Feldhausen, 2014 ; Simon & Christodoulides, 2016a ; 2016b ; Smith, 2009 ; Swerts, 1998). De la même manière, concernant les contours intonatifs, certains auteurs constatent que les contours tonals

descendants renforcent la perception d'une proéminence, et par conséquent la perception d'une frontière forte, tandis que les frontières plus faibles sont associées aux contours tonaux montants (Hermès & Rump, 1994 ; Portes, 2002). Dans ce cadre, nous proposons de tester l'influence des indices acoustiques, tels que la durée de la syllabe finale précédant les frontières et la valeur de  $f_0$ , associées aux syllabes initiales et finales autour des frontières, sur la force des frontières et des proéminences perçues. Nous rappelons en effet que nous tenons à rendre compte du statut métrique des proéminences, et de leur rôle structural en tant que marqueurs gauches et droites des constituants prosodiques.

## **2. SELECTION DU MATERIEL LINGUISTIQUE**

Afin de répondre à ces questions de recherche, nous avons procédé à un travail de sélection du matériel linguistique. Nous présentons en effet dans cette partie les trois corpus de parole différents, auxquels nous nous sommes intéressée pour mettre en place notre protocole expérimental : le corpus d'Edimbourg (Astésano et *al.*, 2007) en parole contrôlée, le corpus de la Map Task Aix (Bard et *al.*, 2013) en parole semi-spontanée, et le corpus du CID (Bertrand et *al.*, 2008) en parole spontanée. Dans le cadre de notre étude du phrasé prosodique en français, nous choisissons plus précisément de manipuler des structures syntaxiquement ambiguës. La levée des ambiguïtés syntaxiques a fait l'objet d'un certain nombre de travaux, dont l'objectif était de mieux comprendre la relation de congruence ou de non-congruence entre la structure syntaxique et la structure prosodique (entre autres Fodor, 1998 ; Lehiste, 1973 ; Martin, 2009 ; Millotte et *al.*, 2007 ; Price et *al.*, 1991 ; Schafer et *al.*, 1996). Lors de notre étude, cependant, l'utilisation des ambiguïtés syntaxiques ne vise pas à rendre compte d'une certaine congruence entre les deux structures, mais plutôt à faire ressortir différents phrasés prosodiques possibles pour une même structure syntaxique, et par conséquent une granularité des niveaux de constituance. Nous verrons à travers la présentation de ce matériel linguistique que nous pourrons effectuer notre étude seulement sur le corpus d'Edimbourg, étant donné le manque de structures syntaxiques comparables à travers les trois corpus.

### **2.1 Le corpus de parole contrôlée : Corpus d'Edimbourg**

Le corpus d'Edimbourg (Astésano et *al.*, 2007) est un corpus de parole contrôlée composé de structures syntaxiquement ambiguës (structures syntaxiques comparables au

paradigme « *old men and women* » en anglais de Lehiste, 1973). Les deux structures syntaxiques possibles, constituées de deux noms coordonnés (N1 et N2) et d'un adjectif (ADJ), sont créées en manipulant la portée de l'adjectif :

« **Les gants et les bas lisses** » en Condition 1 où ADJ qualifie uniquement le dernier nom (seuls les bas sont lisses).

« **Les gants et les bas lisses** » en Condition 2 où ADJ qualifie l'ensemble des deux noms (les gants et les bas sont tous les deux lisses).

Sur la base de ces deux structures syntaxiques, différents phrasés prosodiques peuvent être prédits, variant en fonction des niveaux de constituance. Plus précisément, une frontière d'ip entre N1 et N2, et de PW entre N2 et ADJ sont prédites en Condition 1, tandis qu'une frontière d'AP entre N1 et N2, et d'ip entre N2 et ADJ sont attendues en Condition 2.

**Condition 1** : { [ les ( gants ) ] }<sub>ip</sub> { [ et les ( bas )<sub>pw</sub> ( lisses ) ] }

**Condition 2** : { [ les ( gants ) ]<sub>ap</sub> [ et les ( bas ) ] }<sub>ip</sub> { [ ( lisses ) ] }

Le corpus manipule également la taille des constituants en incrémentant le nombre de syllabes des N et ADJ d'une à quatre syllabes dans toutes les combinaisons possibles, excepté pour N1 et N2 dont les longueurs covarient systématiquement. Chaque longueur de N apparaît donc avec chacune des 4 longueurs possibles de ADJ. Le corpus est constitué au total de 3072 syntagmes répartis dans quatre sets de syntagmes différents : 32 séquences D1+N1+cc+D2+N2+ADJ (16 par condition) \* 4 sets \* 3 répétitions \* 8 locuteurs. Le contenu segmental de chaque syntagme a été soigneusement contrôlé pour éviter une variabilité segmentale. Au sein de chaque set de syntagmes, les N2 de différentes longueurs commencent par le même contenu segmental (bas > balises > balivernes > baratineurs). Ce principe s'applique également à l'adjectif (lisses > licites > licencieux > libérateurs). De la même manière, les syllabes supplémentaires incrémentées en N2 reprennent le même contenu segmental que celui de la syllabe initiale de la série de ADJ précédente (bas lisses, licites, licencieux, libérateurs > balises vertes, vermeilles, verticales, vertigineuses > balivernes sottes, etc.). Pour l'enregistrement du corpus, les séquences D1+N1+cc+D2+N2+ADJ ont été placées dans un contexte phrastique plus large forçant la réalisation d'une frontière prosodique forte (frontière d'*IP*) après le syntagme cible : « les

*gants et les bas lisses, en fait, sont relativement rares* ». Afin de marquer l'ambiguïté syntaxique attendue en condition 1 et 2, les 8 locuteurs s'aidaient d'une instruction écrite pour chaque syntagme facilitant ainsi la lecture du contexte syntaxique :

Pour la Condition 1 : « ***si les bas sont lisses, mais les gants ne sont pas lisses, vous dites...*** », et pour la Condition 2 : « ***si à la fois les gants et les bas sont lisses, vous dites...*** ».

En revanche, aucune des frontières syntaxiques et prosodiques n'était explicitement indiquée dans les scripts que les locuteurs devaient lire à haute voix. Un test de jugement sémantique a ensuite été réalisé sur tous les syntagmes, afin de ne garder que les locuteurs parvenant à instancier les deux conditions syntaxiques. Nous rappelons que ce corpus a été élaboré à l'origine dans le but de tester le rôle des indices prosodiques (plus spécifiquement celui de l'AI) dans la désambiguïsation de structures syntaxiques sur le plan de la production (Astésano et *al.*, 2007) et de la perception (Astésano et *al.*, 2012 ; 2013).

### **2.2 Le corpus de parole semi-spontanée : Corpus de la Map Task Aix**

A la suite du corpus d'Edimbourg (Astésano et *al.*, 2007), les auteurs ont élaboré un second corpus de parole, en reproduisant certaines des structures syntaxiquement ambiguës du précédent corpus dans un contexte de parole moins contrôlée. Plus précisément, ils ont choisi de constituer un corpus de parole semi-spontanée ou semi-guidée, basé sur le dessin expérimental des dialogues de tâches d'orientation, où les consignes données par l'expérimentateur permettent de contrôler plus ou moins l'interaction. En effet, la Map Task (Anderson et *al.*, 1991) est une tâche conversationnelle entre deux locuteurs, durant laquelle les locuteurs doivent se donner des indications à partir d'une carte pour aller d'un point à un autre de la carte. Les deux locuteurs n'ont pas exactement le même type de carte sous les yeux, selon si le locuteur donne ou reçoit les instructions, ce qui donne lieu à une certaine négociation rendant alors l'interaction riche d'un point de vue linguistique. Le corpus de la Map Task Aix (Bard et *al.*, 2013) a été enregistré par 8 locuteurs (4 hommes et 4 femmes) qui ont chacun produit 8 Map Tasks (4 Map Tasks dans le rôle de celui qui donne les instructions, 4 Map Tasks dans le rôle de celui qui reçoit les instructions : cf. Annexe N°1-A). Le contenu lexical du corpus reprend 8 structures syntaxiquement ambiguës, issues du corpus d'Edimbourg (Astésano et *al.*, 2007) :

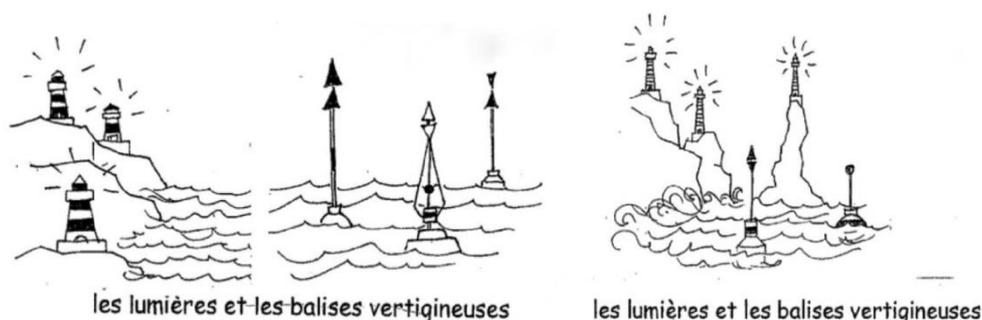
« **Les lumières et les balises vermelles** » en Condition 1 et 2

« **Les lumières et les balises vertigineuses** » en Condition 1 et 2

« **Les bonimenteurs et les baratineurs fameux** » en Condition 1 et 2

« **Les bonimenteurs et les baratineurs fabulateurs** » en Condition 1 et 2

Sur chaque carte, une de ces 8 structures syntaxiquement ambiguës a été représentée sous forme d'un repère typographique illustrant le sens et la condition syntaxique de chaque structure. Le syntagme est écrit avec une orthographe classique sous chaque repère, et selon la condition syntaxique associée, l'illustration du repère varie. Prenons l'exemple du repère typographique associé au syntagme « *les lumières et les balises vertigineuses* » (cf. Figure N°43).



**Figure 43 :** Exemple des repères typographiques associés aux structures syntaxiquement ambiguës issues du corpus d'Edimbourg (Astésano et al., 2007)  
Le repère de gauche correspond à la structure syntaxique produite en Condition 1, où seules les balises sont vertigineuses, et le repère de droite correspond à la structure syntaxique produite en Condition 2, où à la fois les lumières et les balises sont vertigineuses (issu de Bard et al., 2013, p. 16)

Pour illustrer la Condition 1 (où seules les balises sont vertigineuses), le repère est décomposé en deux parties constituées, pour la première, de phares (les lumières) dessinés sur des falaises et, un peu en marge, pour la seconde partie, des bouées (les balises) dessinées au milieu d'une mer agitée (en référence à l'adjectif « vertigineuses »). Quant à la Condition 2 (où les lumières et les balises sont toutes les deux vertigineuses), le repère montre les phares et les bouées dessinés ensemble au milieu d'une mer agitée.

Les différents repères typographiques associés aux 8 structures syntaxiquement ambiguës ont ensuite été distribués équitablement entre les cartes (cf. Figure N°44), selon si le locuteur donne (IG) ou reçoit (IF) les instructions, et si les repères associés aux deux

conditions syntaxiques sont coordonnés entre les deux cartes (IG=IF : BB =Condition 2 dans les deux cartes, NN=Condition 1 dans les deux cartes), ou alternés d'une carte à l'autre (IG≠IF : BN= Condition 2 dans une carte et Condition 1 dans l'autre carte, NB=Condition 1 dans une carte et Condition2 dans l'autre carte). Les différentes configurations sont produites par chaque locuteur dans chacun des 4 dialogues<sup>6</sup> : nous rappelons qu'un même locuteur réalise 4 Map Tasks dans le rôle de celui qui donne les instructions et 4 Map Tasks dans le rôle de celui qui reçoit les instructions.

		noun 4-syllables + adjective 2-syllables <i>les bonimenteurs et les baratineurs fameux</i>				noun 4-syllables + adjective 4-syllables <i>les bonimenteurs et les baratineurs fabulateurs</i>				
		IG=IF		IG≠IF		IG=IF		IG≠IF		
		BB	NN	BN	NB	BB	NN	BN	NB	
noun 2-syllables + adjective 2-syllables <i>les lumières et les balises vermeilles</i>	IG=IF	BB			Ai1	Aiii1			Bi6	Biii6
		NN			Aii1	Aiv1			Bii6	Biv6
	IG≠IF	BN	Av5	Avii5			Bv2	Bvii2		
		NB	Avi5	Aviii5			Bvi2	Bviii2		
noun 2-syllables + adjective 4-syllables <i>les lumières et les balises vertigineuses</i>	IG=IF	BB			Ci3	Ciii3			Di8	Diii8
		NN			Cii3	Civ3			Dii8	Div8
	IG≠IF	BN	Cv7	Cvii7			Dv4	Dvii4		
		NB	Cvi7	Cviii7			Dvi4	Dviii4		

**Figure 44 :** Tableau récapitulatif de la répartition des 4 structures syntaxiquement ambiguës issues du corpus d'Edimbourg (Astésano et al., 2007) à travers les 4 dialogues (illustrés par les 4 couleurs différentes)

IG correspond au locuteur donnant les instructions, IF correspond au locuteur recevant les instructions ; BB ou NN indiquent que les repères associés aux deux conditions syntaxiques sont coordonnés entre les deux cartes ; BN ou NB indiquent que les repères associés aux deux conditions syntaxiques sont alternés entre les deux cartes (issu de Bard et al., 2013, p. 19)

Les enregistrements ont été effectués dans une chambre sourde ou anéchoïque. Ce corpus est par ailleurs transcrit orthographiquement, facilitant ainsi le repérage des mots-cibles. Afin de vérifier l'exploitabilité de ce corpus, nous avons effectué un comptage des différentes structures syntaxiques, afin de faire ressortir notamment le nombre de syntagmes issus du corpus d'Edimbourg disponibles. Au total, nous avons relevé 5145 structures constituées de noms et/ou d'adjectifs dans différentes configurations syntaxiques (cf. Annexe N°1-B). En

<sup>6</sup> Les 4 couleurs utilisées dans la Figure N°44 renvoient aux 4 dialogues réalisés dans le corpus.

revanche, sur les 5145 structures, seules 79 correspondent aux structures D1+N1+cc+D2+N2+ADJ du corpus d'Edimbourg (Astésano et *al.*, 2007) :

« **les lumières et les balises vermeilles** » : 24 occurrences produites par les 8 locuteurs, dont 14 occurrences associées aux repères correspondant à la Condition 2 (10 produites par celui donnant les instructions et 4 produites par celui recevant les instructions), et 10 occurrences associées aux repères correspondant à la Condition 1 (7 produites par celui donnant les instructions et 3 produites par celui recevant les instructions).

« **les lumières et les balises vertigineuses** » : 41 occurrences produites par les 8 locuteurs, dont 29 occurrences associées aux repères correspondant à la Condition 2 (27 produites par celui donnant les instructions et 2 produites par celui recevant les instructions), et 12 occurrences associées aux repères correspondant à la Condition 1 (10 produites par celui donnant les instructions et 2 produites par celui recevant les instructions).

« **les bonimenteurs et les baratineurs fameux** » : 2 occurrences produites par 1 locuteur, dont 1 occurrence associée au repère correspondant à la Condition 2 (produite par celui donnant les instructions), et 1 occurrence associée au repère correspondant à la Condition 1 (produite par celui donnant les instructions).

« **les bonimenteurs et les baratineurs fabulateurs** » : 12 occurrences produites par 4 locuteurs, dont 4 occurrences associées aux repères correspondant à la Condition 2 (produites par celui donnant les instructions), et 8 occurrences associées aux repères correspondant à la Condition 1 (7 produites par celui donnant les instructions et 1 produite par celui recevant les instructions).

Nous avons relevé également 4 autres structures non issues du corpus d'Edimbourg (Astésano et *al.*, 2007) correspondant néanmoins à une structure D1+N1+cc+D2+N2+ADJ :

« **les girafes et la montagne nue** » : 1 occurrence pouvant correspondre à une Condition 1 (seule la montagne est nue) produite par celui recevant les instructions.

« **les bonimenteurs et la terre labourée** » : 1 occurrence pouvant correspondre à une Condition 1 (seule la terre est labourée) produite par celui donnant les instructions.

« *les lumières et le cadeau promis* » : 1 occurrence pouvant correspondre à une Condition 1 (seul le cadeau est promis) produite par celui recevant les instructions.

« *la ruine et les bonimenteurs baratineurs* » : 1 occurrence pouvant correspondre à une Condition 1 (seuls les bonimenteurs sont baratineurs) produite par celui donnant les instructions.

Sur la base de cette sélection, nous avons construit un tableau Excel répertoriant chacune de ces structures en y associant systématiquement des informations sur : le type de carte (couleur), le locuteur (numéro et code du locuteur mais aussi rôle du locuteur), le contexte discursif dans lequel se trouve la structure syntaxique cible, ainsi que la condition syntaxique dans laquelle a été réalisé le syntagme cible (*cf.* Annexe N°1-C). Enfin, bien qu'on puisse associer les différentes occurrences avec la condition syntaxique prédite, nous ne pouvons cependant assurer l'adéquation entre cette condition prédite et la production effective des locuteurs. Pour cela, un test de jugement sémantique est nécessaire afin de vérifier, via la perception de différents auditeurs, que les locuteurs ont produit correctement les deux conditions syntaxiques prédites.

### **2.3 Le corpus de parole spontanée : Corpus du CID**

Le corpus du CID, « *Corpus of Interactionnal Data* » (Bertrand et *al.*, 2008), est un corpus de parole spontanée, dans lequel 16 participants (10 femmes et 6 hommes) ont enregistré 8 dialogues non mixtes. Les enregistrements ont été effectués dans une chambre sourde ou anéchoïque. Les deux locuteurs étaient munis d'un micro-casque permettant d'enregistrer chaque voix sur des pistes séparées. Les participants ont également été filmés. Par ailleurs, il est intéressant de préciser que les locuteurs, étant des collègues de travail, ont une certaine familiarité favorisant ainsi des dialogues naturels et spontanés. Dans chaque dialogue, les deux locuteurs avaient pour consigne de discuter librement soit à propos de conflits professionnels, soit à propos de situations insolites dont ils avaient été témoins. Le corpus du CID se compose au total de 8h de dialogues qui ont donné lieu à une annotation multimodale, permettant ainsi de travailler sur différents niveaux linguistiques et extralinguistiques (syntaxe, prosodie, phonologie, sémantique, phénomènes mimogestuels, etc.). Dans le cadre de notre travail, nous nous sommes appuyés sur l'annotation morphosyntaxique permettant de rechercher facilement des structures syntaxiques D1+N1+cc+D2+N2+ADJ, comparables à celles observées dans le corpus d'Edimbourg (Astésano et *al.*, 2007) et de la MapTask Aix (Bard et *al.*, 2013). Aucune structure de ce type

n'a cependant été trouvée dans l'ensemble du corpus du CID. En revanche, nous avons relevé d'autres structures reprenant certains composants de D1+N1+cc+D2+N2+ADJ (cf. Annexe N°2) :

**ADJ+ADJ+N** du type « *une belle petite histoire* » : 11 occurrences

**ADJ+N** du type « *le vrai personnel* » : 495 occurrences

**ADJ+N+ADJ** du type « *un grand panneau lumineux* » : 18 occurrences

**N+cc+N** du type « *les frères et sœurs* » : 24 occurrences

**N+de+N** du type « *la farine de châtaignes* » : 217 occurrences

**N+de+N+ADJ** du type « *une tenue de moine bouddhiste* » : 15 occurrences

**N+ADJ** du type « *la situation inverse* » : 415 occurrences

Nous nous sommes ensuite plus particulièrement intéressés à la structure N+ADJ (cf. Annexe N°2-B), disposée à illustrer des cas d'arcs accentuels (Fónagy, 1980) comparables à ce qu'on pourrait observer dans le corpus d'Edimbourg (Astésano et *al.*, 2007). Pour cela, nous avons procédé à un travail de sélection et de description de cette structure, en proposant une catégorisation en 5 parties, afin de ne garder que les « bons candidats » : présence d'un déterminant précédant le N, absence de particule modale et de disfluence avant, pendant ou après la structure.

**Catégorie 1** « bons candidats » : 239 occurrences du type « *des relations conflictuelles* » ou « *le code secret* ».

**Catégorie 2** « absence du déterminant » : 34 occurrences du type « *en sens inverse* » ou « *∅ mesure disciplinaire* ».

**Catégorie 3** « précédé/suivi d'une particule » : 63 occurrences du type « *les cartes bleues effectivement ouai* » ou « *le code secret donc* ».

**Catégorie 4** « précédé/suivi d'une disfluence » : 56 occurrences du type « *des des relations conflictuelles* » ou « *les euh fonctions publiques euh* ».

**Catégorie 5** « cumul des types » : 23 occurrences « *des des expés anciennes quoi* » (disfluence + particule modale) ou « *euh ∅ restaurant antillais* » (disfluence + absence de déterminant).

Pour chaque structure dans chacune des catégories, nous avons répertorié la référence du stimulus (code du locuteur + temps en ms indiquant la position de la structure dans le dialogue) et le nombre de syllabes de chaque composant du constituant (isolé ou cumulé). Nous avons également déterminé les structures plus ou moins figées. Enfin, nous rappelons que ce corpus a été utilisé dans le cadre d'une étude en perception (Roux et *al.*, 2016), élaborée dans le but de tester, non pas l'organisation du phrasé prosodique, mais plutôt la capacité des auditeurs francophones à percevoir les frontières et les proéminences, indépendamment des autres événements prosodiques environnants. Plus précisément, les auteurs ont utilisé dans le cadre de leur étude les trois structures syntaxiques les plus fréquentes réalisées dans le corpus : N+ADJ ; ADJ+N ; N+de+N.

### 2.4 Préparation du corpus de parole sélectionné

Dans le cadre de notre travail, nous avons dû élaborer notre test de perception sur le corpus de parole contrôlée seul (Corpus d'Edimbourg : Astésano et *al.*, 2007). En effet, les structures syntaxiques qui nous intéressent (D1+N1+cc+D2+N2+ADJ) n'ont pas été observées dans le corpus de parole spontanée, et ne sont pas assez contrôlées sur le plan expérimental dans le corpus de parole semi-guidée<sup>7</sup>. Nous proposons donc de commencer notre exploration du phrasé prosodique sur un corpus *ad hoc*, afin de faire ressortir les primitives prosodiques pertinentes que nous pourrions par la suite tester sur de la parole moins contrôlée. Par ailleurs, étant donné notre travail de préparation des corpus de la Map Task Aix et du CID (sélection et transcription des structures syntaxiques plus ou moins comparables au corpus d'Edimbourg), nous pouvons d'ores et déjà envisager la mise en place de futures expérimentations.

Concernant plus précisément le corpus d'Edimbourg, nous avons sélectionné pour notre étude en perception un seul des quatre sets lus par 4 locuteurs (3 femmes / 1 homme), donnant lieu à 128 séquences D1+N1+cc+D2+N2+ADJ (*cf.* Figure N°45).

---

<sup>7</sup> Nous rappelons qu'il est nécessaire de procéder à un test de jugement sémantique sur les syntagmes du corpus de la Map Task Aix (Bard et *al.*, 2013) afin de vérifier que les locuteurs ont produit correctement les deux conditions syntaxiques prédites.

<b>les gants et les bas</b>	<b>lisses licites licencieux libérateurs</b>
<b>les lumières et les balises</b>	<b>vertes vermeilles verticales vertigineuses</b>
<b>les bagatelles et les balivernes</b>	<b>sottes saumâtres saugrenues sommambuliques</b>
<b>les bonimenteurs et les baratineurs</b>	<b>fades fameux fabuleux fabulateurs</b>

**Figure 45 :** Liste des 16 phrases retenues dans le test de perception qui ont chacune été produites dans les deux conditions syntaxiques par 4 locuteurs différents (au total : 128 phrases)

Les 4 locuteurs choisis ont été classés comme parvenant le mieux à instancier les deux conditions syntaxiques (via le test de jugement sémantique). Quant au choix des 128 syntagmes, nous avons sélectionné la meilleure version des trois répétitions de chaque séquence D1+N1+cc+D2+N2+ADJ. Parmi les critères de sélection, nous éliminons les phrases réalisées avec des disfluences ou des hésitations, avec un débit de parole trop rapide, ou encore avec une exagération des paramètres prosodiques. Nous avons ensuite extrait le syntagme cible (D1+N1+cc+D2+N2+ADJ) à l'aide du logiciel Praat (Boersma et *al.*, 2017), à partir duquel nous avons également effectué une transcription phonétique et syllabique de chaque séquence.

Enfin, en vue de notre protocole expérimental, nous avons procédé à la répartition de ces 128 syntagmes dans 4 groupes expérimentaux (*cf.* Annexe N°3), composés chacun de 32 séquences D1+N1+cc+D2+N2+ADJ différentes. Ces 32 syntagmes ont été distribués équitablement sur les 4 locuteurs, soit 8 syntagmes différents par locuteur dans un même groupe expérimental, à travers les deux conditions syntaxiques et la taille des constituants.

### **3. TEST DE PERCEPTION DE LA PAROLE**

Sur la base du corpus de parole sélectionné, nous proposons de procéder à un test de perception de la parole. Nous présentons dans le cadre de cette partie les participants de notre expérimentation, les tâches expérimentales proposées dans notre étude, mais également la préparation de notre base de données.

#### **3.1 Participants**

Les séquences D1+N1+cc+D2+N2+ADJ choisies pour l'étude ont été jugées perceptivement par 80 auditeurs francophones natifs (71 femmes, âge moyen= 23ans, plage d'âge: 18-43), sans problème d'audition et avec une vue correcte ou corrigée. Le recrutement des participants a été effectué parmi les étudiants de l'université Toulouse 2 Jean-Jaurès et via les réseaux sociaux. Avant de débiter l'expérience, tous les participants signaient un formulaire de consentement éclairé (*cf.* Annexe N°4-A) et complétaient une fiche de renseignement (*cf.* Annexe N°4-B), donnant des informations générales sur l'informant (nom, prénom, sexe, âge, région de naissance, etc.), mais aussi des indications sur son niveau en langues étrangères et sur sa pratique musicale et des jeux vidéo (*cf.* récapitulatif des renseignements obtenus en Annexe N°4-C). A l'issue de l'expérience, les participants signaient un reçu d'indemnisation (*cf.* Annexe N°4-D) faisant preuve de leur rémunération, à hauteur de 20 euros, pour leur participation à l'expérimentation (financement obtenu dans le cadre de l'ANR PhonIACog<sup>8</sup>).

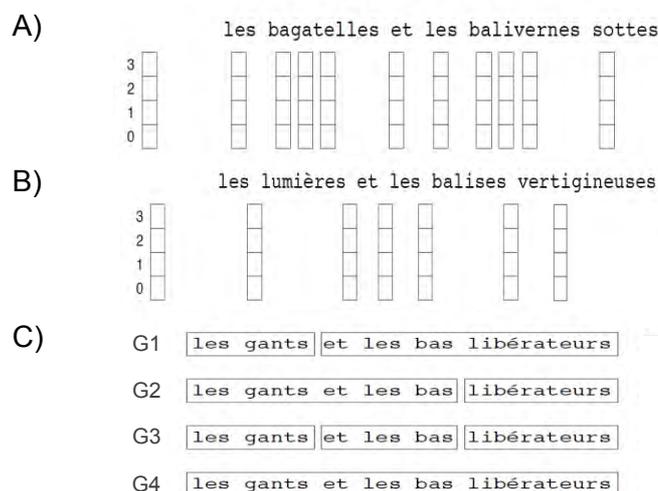
#### **3.2 Tâches de perception**

Les 80 sujets ont été répartis équitablement sur les 4 groupes expérimentaux pour que chaque séquence D1+N1+cc+D2+N2+ADJ soit jugée par 20 auditeurs différents. Chaque participant écoutait 32 syntagmes, à 3 reprises, dans 3 tâches de perception différentes (même série de 32 syntagmes par tâche de perception): une tâche de prééminence où ils devaient juger le degré de mise en relief de chaque syllabe du syntagme (variant de 6 à 15 sites potentiels selon la combinaison du nombre de syllabes des constituants), sur une échelle d'évaluation allant de 0 à 3 ; une tâche de frontière où les auditeurs devaient juger le degré de rupture entre chaque mot (6 sites potentiels) sur une

---

<sup>8</sup> Projet ANR-1-BSH2-0001-01 (IP : Corine Astésano)

échelle d'évaluation allant de 0 à 3 ; une tâche de groupement où ils devaient juger la manière dont les mots se regroupaient les uns avec les autres, parmi 4 propositions de groupements (cf. Figure N°46).



**Figure 46 :** Format de présentation des 3 tâches de perception appliquées à la séquence D1+N1+cc+D2+N2+ADJ  
 A) Tâche de proéminence, B) Tâche de frontière, C) Tâche de groupement

Il est intéressant de préciser que contrairement à la plupart des précédentes études en perception qui utilisent un paradigme binaire pour recueillir les réponses perceptives (absence vs. présence d'un événement prosodique), nous avons fait le choix ici de recueillir les scores de frontières et de proéminences sur la base d'une échelle à 4 niveaux, dans le but de ne pas seulement juger l'existence ou non d'un phénomène, mais aussi de faire ressortir le degré de perception de ces phénomènes prosodiques.

Ces tâches de perception ont été effectuées à partir du logiciel d'évaluation perceptive « *Perceval* », pour « *PERception EVALuation Auditive & Visuelle* », développé au Laboratoire Parole et Langage d'Aix-en-Provence (Ghio et al., 2003). Plus précisément, nous avons utilisé l'interface « *Lancelot* » permettant de recueillir plusieurs réponses à des stimuli auditifs par l'intermédiaire d'un formulaire HTML classique. Chaque participant était installé seul dans une salle silencieuse, muni d'un casque audio haute-fidélité<sup>9</sup> et confortablement assis devant un ordinateur. Les consignes exactes pour chacune des tâches (cf. Annexe N°5-A) étaient présentées aux participants sur l'écran de l'ordinateur. Les consignes,

<sup>9</sup> Beyer Dynamic DT 770 pro

orthographiées classiquement, incluaient des exemples illustrant précisément les différentes possibilités de réponses. Avant chaque tâche, les participants effectuaient une phase d'entraînement, comprenant 4 séquences D1+N1+cc+D2+N2+ADJ issues d'un set de syntagmes différent de celui utilisé dans cette étude. Durant cette phase d'entraînement, nous effectuions les réglages nécessaires (volume sonore, lisibilité, etc.) tout en nous assurant de la bonne compréhension des consignes. Les participants prenaient le temps qu'ils souhaitaient pour effectuer les tâches, et pouvaient écouter jusqu'à 5 fois le même syntagme en cliquant sur le bouton d'un haut-parleur visible à l'écran. De la même manière, les participants devaient cliquer sur le bouton « *suivant* » pour faire apparaître le syntagme suivant à l'écran. Enfin, l'ordre de présentation des 3 tâches de perception était contrebalancé entre tous les participants. L'expérimentation durait en moyenne 50 minutes et comprenait 3 phases expérimentales avec 5 minutes de pause entre chaque phase, correspondant aux 3 tâches de perception. Les données de notre expérimentation ont été recueillies automatiquement par l'intermédiaire du logiciel « *PERception EVALuation Auditive & Visuelle* » dans trois tableaux Excel différents, chacun étant associé à une tâche de perception spécifique.

### 3.3 Préparation de la base de données

Nous avons par la suite mis en commun les résultats issus des trois tâches de perception (proéminence, frontière et groupement) pour ne former qu'une seule et même base de données. Le tableau Excel global est composé de 2560 lignes correspondant aux réponses perceptives de chaque participant pour chaque syntagme (80 participants ont chacun testé perceptivement 32 syntagmes différents :  $80 \times 32 = 2560$ ), et de 30 colonnes, initialement, correspondant aux différentes variables de notre protocole expérimental.

Une première série de variables donnent des indications sur les participants (code du sujet et groupe expérimental) et sur le matériel linguistique (condition syntaxique, locuteur, taille du constituant, etc.). Pour les indications sur le matériel linguistique, et notamment sur la taille du constituant, il est important de préciser que nous avons choisi de distinguer le nombre de syllabes additionné et concaténé de D1+N1+cc+D2+N2+ADJ, pour ainsi rendre compte non seulement de la taille totale du constituant, mais aussi des « configurations syllabiques » possibles pour le même nombre de syllabes d'un constituant. Le terme « configuration » renvoie en effet ici à la manière dont les mots à l'intérieur du syntagme D1+N1+cc+D2+N2+ADJ se répartissent le nombre de syllabes. Le constituant composé de 11 syllabes, par exemple, peut être configuré de 2 manières différentes :

$D1+N1+cc+D2+N2+ADJ = 1+2+1+1+2+4$  ou  $1+3+1+1+3+2$ . La concaténation du nombre de syllabes nous donne dans ce cas-là une information supplémentaire sur la taille du constituant. D'autres variables sont ensuite plus spécifiquement associées aux scores de perception : groupement, frontière et proéminence. Nous rappelons que les scores de frontières ont été renseignés sur chaque mot du syntagme (mots lexicaux et grammaticaux), et que les scores de proéminences ont été recueillis sur chaque syllabe du syntagme (syllabes accentuées et inaccentuées).

Sur la base de ce tableau de données, nous avons ensuite choisi d'exclure les syntagmes composés de constituants N ou ADJ monosyllabiques ne permettant pas de distinguer la proéminence initiale et finale, et qui pourront faire l'objet d'une étude spécifique ultérieure (les **gants** et les **bas lisses** ; les lumières et les balises **vertes** ; les bagatelles et les balivernes **sottes** ; les bonimenteurs et les baratineurs **fade**s).

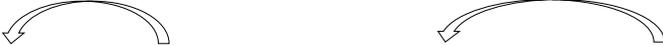
Enfin, nous avons créé de nouvelles variables, calculées à partir des scores de perception des frontières et des proéminences, indiquant la force relative entre deux frontières ou deux proéminences perçues. Cette force relative permet notamment de signaler quelle frontière ou quelle proéminence est perçue plus ou moins forte que l'autre. L'information de force relative est particulièrement centrale dans le cadre de notre travail, étant donné notre approche syntagmatique de l'analyse du phrasé prosodique. Nous avons en effet besoin de tester les forces relatives entre les frontières et les proéminences perçues au sein d'un même énoncé, pour ainsi faire ressortir les relations hiérarchiques établies entre les différentes frontières et proéminences. Pour créer ces variables, nous avons soustrait le score d'un événement prosodique au score d'un autre événement prosodique, en respectant l'ordre d'apparition de ces événements dans le syntagme. Si nous souhaitons par exemple calculer le score de force relative entre la frontière perçue après N1 et après N2, nous devons alors soustraire le score de la frontière apparaissant en première position dans le syntagme, à savoir N1, au score de la frontière apparaissant en seconde position dans le syntagme, à savoir N2 :



Les N1 | et les N2 | ADJ

A la suite de ce calcul, nous obtenons alors 7 scores de force relative possibles : -3, -2, -1, 0, 1, 2 ou 3. Un score de force relative négatif signifie que l'événement prosodique apparaissant en première position dans le syntagme est perçu comme plus fort que l'événement prosodique apparaissant en seconde position dans le syntagme. A l'inverse, un

score de force relative positif montre que l'événement prosodique apparaissant en première position dans le syntagme est perçu comme moins fort que l'événement prosodique apparaissant en seconde position dans le syntagme. Enfin un score de force relative égale à 0 indique une force équivalente entre les événements prosodiques. Pour reprendre notre exemple précédent sur la force relative entre la frontière perçue après N1 et après N2, si le score est négatif cela signifie que la frontière après N1 est perçue plus forte que celle après N2, si le score est positif cela signifie que la frontière après N1 est perçue moins forte que celle après N2, et si le score est égal à 0 cela signifie que les frontières sont perçues avec une force équivalente. Nous avons donc procédé à ce calcul pour les scores de frontières après N1 et N2, et pour les scores de proéminences perçues sur chaque syllabe de N1, N2 et ADJ. Plus précisément, nous avons calculé la force relative entre les proéminences perçues au sein d'un même constituant, mais également entre différents constituants, dans toutes les combinaisons possibles. De la même manière, nous avons calculé la force relative entre deux types de proéminences différents (AI vs. AF), ainsi qu'entre deux proéminences du même type (AI vs. AI ou AF vs. AF). Nous tenons en effet à avoir une vue générale de l'organisation hiérarchique des frontières et des proéminences, en testant la relation entre des événements prosodiques plus ou moins à proximité l'un de l'autre, dans le but de mieux comprendre l'organisation globale du phrasé prosodique.

  
Les bo ni men teurs et les ba ra ti neurs fa bu la teurs

Nous souhaitons par ailleurs contrôler au mieux dans notre étude la position de la proéminence initiale, traditionnellement décrite comme variable (Jun & Fougeron, 2000 ; 2002 ; Post, 2003). Dans ce contexte, nous choisissons de calculer la force relative entre les proéminences, en incluant également les syllabes intermédiaires. Toutefois, il est important de préciser que le score de la proéminence perçue sur la deuxième syllabe renvoie systématiquement aux N ou ADJ tri- et quadrisyllabiques, tandis que le score de la proéminence perçue sur la troisième syllabe correspond uniquement aux N ou ADJ quadrisyllabiques :

  
Les bo ni men teurs et les ba ra ti neurs fa bu la teurs

La création de l'ensemble de ces nouvelles variables ajoute alors à notre tableau de données initial 55 variables supplémentaires. Sur la base de l'ensemble des données

recueillies, nous pouvons alors émettre un certain nombre d'hypothèses et de prédictions quant à la manière dont les événements prosodiques sont perçus, et s'organisent dans l'élaboration du phrasé prosodique.

### **4. HYPOTHESES ET PREDICTIONS**

Nous nous intéressons plus précisément à la manière dont les frontières et les proéminences ont été perçues dans chaque groupement. En effet, l'objectif de notre étude est de questionner les niveaux de constituance qui sont nécessaires à l'élaboration du phrasé prosodique en français. Nous questionnons notamment l'utilité du niveau du Mot Lexical dans la hiérarchie prosodique de la langue. Directement en lien avec la notion de domaine accentuel, nous tenons également à rendre compte du rôle phonologique de l'accent en français, et plus spécifiquement dans le marquage de la constituance prosodique. Nous remettons par ailleurs en question l'affaiblissement perceptif de l'AF en frontières prosodiques majeures, ainsi que le rôle purement rythmique, secondaire et optionnel de l'AI.

Pour répondre à ces questionnements, nous proposons un test de perception composé de 3 tâches expérimentales. Dans les deux premières tâches, proéminences et frontières sont volontairement jugées séparément. Comme nous l'avons mentionné dans notre cadre théorique, le français est traditionnellement caractérisé par un syncrétisme entre la proéminence finale et la frontière prosodique. Il est donc intéressant de pouvoir tester séparément ces deux phénomènes prosodiques pour mieux rendre compte du rôle de chacun dans le marquage de la constituance. De la même manière, l'idée de juger la mise en relief de toutes les syllabes du syntagme dans la tâche de proéminence, vise à mieux distinguer les syllabes accentuées par rapport aux syllabes inaccentuées, toujours dans le but d'interpréter plus distinctement le rôle et les fonctions des proéminences initiales et finales en français, mais aussi à mieux délimiter la position de l'AI, traditionnellement considérée comme variable. Enfin, la troisième tâche de groupement sert plus spécifiquement à questionner les stratégies du phrasé prosodique : comment les auditeurs ont-ils découpé le flux de parole en groupes de mots ? Cette dernière tâche représente par ailleurs le point de départ de notre travail d'exploration. Nous supposons en effet que les auditeurs segmentent le flux de parole en groupements (en référence à l'analyse en « *chunks* » d'Abney, 1991), et que le choix de ce découpage serait conjugué aux forces des frontières et des proéminences perçues entre et au sein de ces groupements. Dans ce

contexte, nous nous intéressons à la manière dont les frontières et les proéminences ont été perçues au sein d'un énoncé, en fonction du type de groupement choisi par l'auditeur.

Par ailleurs, nous souhaitons également mettre en relation les scores de frontières et de proéminences avec la taille du constituant. Il est en effet intéressant de tester l'influence que peut avoir le nombre de syllabes du constituant sur la perception de la force des frontières et des proéminences. Nous rappelons que la taille des N et des ADJ est incrémentée d'une à quatre syllabes, dans toutes les combinaisons possibles, excepté pour N1 et N2 dont les longueurs covarient systématiquement. Chaque longueur de N apparaît donc avec chacune des 4 longueurs possibles de ADJ. Afin d'avoir une vue globale de l'influence potentielle du nombre de syllabes, nous choisissons de contrôler à la fois la taille des N et des ADJ de manière isolée, mais aussi de prendre en compte le nombre de syllabes total du syntagme (D1+N1+cc+D2+N2+ADJ) dans deux versions différentes : une première version où le nombre de syllabes est additionné et une seconde version où le nombre de syllabes est concaténé. Cette dernière version permet notamment de rendre compte de la variabilité des « configurations syllabiques » possibles pour une même taille de constituant.

### 4.1 Hypothèse 1 : Une granularité plus fine des niveaux de constituance

Concernant tout d'abord le lien entre les groupements et les scores de frontières, nous faisons l'hypothèse que le groupement est fortement corrélé à la condition syntaxique, et qu'il permet plus précisément d'affiner les prédictions initiales faites sur les niveaux de frontière ou de constituance, associés à chacune des deux conditions syntaxiques :

**Condition 1 :** { [ les ( N1 ) ] }<sub>ip</sub> { [ et les ( N2 )<sub>pw</sub> ( ADJ ) ] }

**Condition 2 :** { [ les ( N1 ) ]<sub>ap</sub> [ et les ( N2 ) ] }<sub>ip</sub> { [ ( ADJ ) ] }

Nous supposons tout d'abord que G1 et G2 correspondent respectivement à la Condition 1 et à la Condition 2. Plus précisément, nous attendons une frontière plus forte après N1 (//) qu'après N2 (/) en G1, et à l'inverse, une frontière plus forte après N2 (//) qu'après N1 (/) en G2. Selon nous, la frontière est donc perçue plus forte lorsque le mot est situé à une limite de groupement (signalé par « ] ») : après N1 en G1 et après N2 en G2.

**G1** [ les N1 ] // [ et les N2 / ADJ ]

**G2** [ les N1 / et les N2 ] // [ ADJ ]

Quant à G3 et G4, nous avons choisi de proposer ces deux groupements supplémentaires, afin de laisser aux auditeurs une alternative possible aux deux groupements principaux. Plus précisément, nous faisons l'hypothèse que G3 et G4 peuvent correspondre aussi bien à la Condition 1 qu'à la Condition 2, en admettant toutefois, que selon la condition syntaxique à laquelle sont rattachés les deux groupements, les auditeurs ne perçoivent pas la même hiérarchie entre les frontières après N1 et N2. En effet, si G3 et G4 correspondent à la Condition 1, nous attendons une frontière plus forte après N1 (//) qu'après N2 (/), et si G3 et G4 correspondent à la Condition 2, nous attendons alors une frontière moins forte après N1 (/) qu'après N2 (//).

Condition 1 :

**G3** [ les N1 ] // [ et les N2 ] / [ ADJ ]

**G4** [ les N1 // et les N2 / ADJ ]

Condition 2 :

**G3** [ les N1 ] / [ et les N2 ] // [ ADJ ]

**G4** [ les N1 / et les N2 // ADJ ]

Nous supposons en G1 et G2 que la frontière est perçue plus forte lorsque le mot est situé à une limite de groupement. Pour G3 et G4, en revanche, nous faisons l'hypothèse qu'une frontière de type « // » peut ne pas être associée à une limite de groupement (Condition 1 : après N1 en G4 ; Condition 2 : après N2 en G4), et à l'inverse, qu'une frontière de type « / » peut être associée à une limite de groupement (Condition 1 : après N2 en G3 ; Condition 2 : après N1 en G3). En effet, bien que G3 présente deux limites de groupement après N1 et N2, et que G4 ne montre au contraire aucune limite de groupement après N1 et N2, la hiérarchie entre les deux niveaux de frontière serait toutefois systématiquement maintenue. Dans ce contexte, une frontière notée « // », qui est considérée comme plus forte sur le plan syntagmatique que la frontière notée « / », peut être perçue avec une granularité différente des niveaux de frontière sur le plan paradigmatique, lorsque nous la comparons avec les autres frontières de type « // ». Nous attendons par ailleurs le même phénomène pour la frontière notée « / », qui peut également être plus ou moins forte qu'une autre frontière de type « / ». Selon nous, la différence de force entre deux frontières du même type (// ou /) est liée à la présence ou non d'une limite de groupement.

Sur la base de ces hypothèses, nous pouvons alors prédire les niveaux de constituance associés à chaque groupement. En Condition 1 tout d'abord, si nous admettons une frontière d'IP à la fin du syntagme<sup>10</sup> (ADJ)], nous pouvons alors supposer que la frontière de type « // » réalisée après N1 en G1 et G3, présentant une limite de groupement, corresponde à une frontière d'ip. Le même type de frontière perçue en G4, en revanche, serait marqué par une frontière d'AP, étant donné l'absence de limite de groupement. Quant aux frontières de type « / », nous supposons que la frontière perçue en G3, présentant une limite de groupement, renvoie à une frontière d'AP, permettant ainsi de respecter la hiérarchie sur le plan syntagmatique entre les deux frontières après N1 et N2. Enfin, les frontières du même type perçues en G1 et G4, qui ne sont pas situées à la limite d'un groupement, seraient alors associées à une frontière de PW. En mettant en comparaison la prédiction des frontières par groupement avec celle de la condition syntaxique (Condition 1), nous constatons que la frontière d'AP entre N2 et ADJ n'est en aucun cas envisagée. Cela peut éventuellement s'expliquer par la taille du constituant, impliquant alors un rehaussement du niveau de frontière. Nous pouvons en effet supposer que l'augmentation du nombre de syllabes du constituant N2, ou même N2+ADJ, renforce le degré de la frontière perçue entre N2 et ADJ. Par ailleurs Astésano *et al.* (2007) rendent également compte sur le même corpus de la possibilité d'avoir une frontière d'AP entre N2 et ADJ, à travers notamment l'analyse des occurrences de l'AI sur ADJ. Plus précisément, les auteurs constatent que la probabilité d'avoir un AI sur ADJ est plus grande en Condition 1 qu'en Condition 2, ce qui supposerait selon eux une promotion du niveau de frontière à cet endroit. L'occurrence des AI sur ADJ pourrait alors renforcer le niveau de frontière après N2, en passant du niveau du PW au niveau de l'AP. Dans leur étude en production, Astésano *et al.* (2007) n'ont toutefois pu directement tester cette hypothèse. Au-delà de l'analyse des frontières et des proéminences, la notion de groupement semble en effet nécessaire pour pouvoir rendre compte des différents phrasés prosodiques possibles pour une même structure syntaxique.

**Condition 1 :** { [ les ( N1 ) ] }<sub>ip</sub> { [ et les ( N2 ) ] }<sub>pw</sub> ( ADJ ) }

**G1** [ les N1 ]<sub>ip</sub> // [ et les N2 ]<sub>pw</sub> / ADJ ]<sub>IP</sub>

**G3** [ les N1 ]<sub>ip</sub> // [ et les N2 ]<sub>AP</sub> / [ ADJ ]<sub>IP</sub>

**G4** [ les N1 ]<sub>AP</sub> // et les N2 ]<sub>pw</sub> / ADJ ]<sub>IP</sub>

---

<sup>10</sup> Nous rappelons qu'une frontière d'IP a été forcée lors de l'enregistrement du corpus.

Cette granularité des degrés de frontière est également attendue en Condition 2. Nous faisons l'hypothèse en effet que les frontières de type « // » perçues après N2 en G2 et G3, présentant une limite de groupement, correspondent à une frontière d'ip. La frontière du même type perçue en G4, qui ne présente pas cette fois-ci de limite de groupement, pourrait correspondre à une frontière d'AP. Concernant les frontières du type « / », nous supposons que la frontière présentant une limite de groupement en G3 renvoie à une frontière d'AP, pour ainsi respecter la hiérarchie sur le plan syntagmatique entre les deux frontières après N1 et N2. Enfin, les frontières du même type perçues en G2 et G4, qui ne se situent pas à une limite de groupement, seraient marquées par une frontière de PW. En mettant en comparaison la prédiction des frontières par groupement avec la prédiction de la condition syntaxique (Condition 2), nous voyons que tous les niveaux de frontière supposés sont envisagés. Seul le niveau du PW après N2 n'est pas prédit dans les groupements. Cela pourrait s'expliquer par la hiérarchie établie entre les frontières après N1 et N2, qui est nécessaire pour correspondre à la condition syntaxique.

**Condition 2 :** { [ les ( N1 ) ]<sub>ap</sub> [ et les ( N2 ) ] ]<sub>ip</sub> { [ ( ADJ ) ] }

**G2** [ les N1 <sub>PW</sub> / et les N2 ]<sub>ip</sub> // [ ADJ ] <sub>IP</sub>

**G3** [ les N1 ]<sub>AP</sub> / [ et les N2 ]<sub>ip</sub> // [ ADJ ] <sub>IP</sub>

**G4** [ les N1 <sub>PW</sub> / et les N2 <sub>AP</sub> // ADJ ] <sub>IP</sub>

Nous pouvons également mettre en relation tous les groupements associés aux deux conditions syntaxiques à la fois. A partir de là, nous supposons que le niveau de l'AP peut être associé aussi bien à une frontière de type « // », comme après N1 en G4 associé à la Condition 1, ou de type « / », comme en G3 associé à la Condition 2, tant que la hiérarchie entre les deux frontières (après N1 et N2) est respectée. Par ailleurs, nous pouvons également faire l'hypothèse que ce niveau de l'AP peut être perçu avec un degré de frontière plus ou moins fort, selon la présence, là encore, d'une limite de groupement. Nous rappelons en effet que nous nous questionnons quant à la possibilité qu'un même niveau de constituance puisse être marqué par différents degrés de frontière. La frontière d'AP après N1 en G4 associé, à la Condition 1, pourrait dans ce cas-là être marquée par un degré de frontière un peu plus faible que la frontière d'AP après N1 en G3 associé à la Condition 2, qui présente une limite de groupement. De la même manière, la frontière d'AP après N2 en G4, associé à la Condition 2, pourrait être réalisée avec un degré de frontière un peu plus faible que la frontière d'AP après N2 en G3, associé à la Condition 1, étant donné la

présence à cet endroit d'une limite de groupement. Le groupement permettrait, selon nous, de mettre en évidence une granularité plus fine des niveaux de constituance, tout en rendant compte également de la diversité des degrés de frontière pour un même niveau de constituance. Dans le cadre d'une seconde analyse, nous mettrons également en lien ce marquage des niveaux de frontière avec l'occurrence des proéminences, pour lesquelles nous proposons également plusieurs hypothèses.

### **4.2 Hypothèse 2 : Pas d'affaiblissement perceptif de l'AF en frontière majeure**

La perception des scores de proéminences finales peut également varier en fonction du type de groupement. Selon nous, la force des proéminences finales sur les syllabes des mots lexicaux serait en effet nuancée selon si le mot est situé ou non à la limite d'un groupement (signalé « ] »). Plus précisément, nous supposons que la proéminence finale devrait être perçue sur tous les mots lexicaux (N1, N2 et ADJ), avec cependant une force plus importante lorsqu'elle est positionnée avant la limite d'un groupement, supposant par ailleurs à cet endroit une force de frontière également plus importante (*cf.* Hypothèse 1). En effet, la proéminence finale resterait, selon nous, métriquement forte y compris en frontières prosodiques majeures. En structure de surface, la proéminence finale serait également corrélée à l'augmentation de la force des frontières. Les résultats de l'étude en perception d'Astésano et *al.* (2012) montrent une corrélation entre le degré de l'AF le plus fort et le niveau de frontière le plus élevé. Une autre étude en perception (Méndez, en cours ; Méndez & Astésano, 2017) rend compte par ailleurs d'une modulation de la force des frontières d'ip par l'AF descendant ou montant. Selon les auteurs, un AF marqué par une chute de  $f_0$ , plutôt que par une montée de  $f_0$ , tendrait à renforcer le degré de la frontière d'ip auquel il est associé. Dans ce contexte, nous supposons, tout d'abord en Condition 1, un renforcement de la proéminence finale sur N1 en G1 et G3, présentant une limite de groupement, par rapport à la proéminence finale perçue en G4, qui n'est pas située cette fois-ci à la limite d'un groupement. De la même manière, nous attendons un renforcement de la proéminence finale sur N2 en G3, qui présente une limite de groupement, par rapport aux proéminences finales perçues en G1 et G4. Concernant la Condition 2, nous faisons l'hypothèse d'un renforcement de la proéminence finale sur N1 en G3, et d'un renforcement de la proéminence finale sur N2 en G2 et G3 par rapport aux autres groupements, étant donné là encore la présence d'une limite de groupement.

**Condition 1 :**

**G1** [ les AI **AF** ]<sub>ip</sub> // [ et les AI AF **PW** / AI **AF** ] IP

**G3** [ les AI **AF** ]<sub>ip</sub> // [ et les AI **AF** ]<sub>AP</sub> / [ AI **AF** ] IP

**G4** [ les AI AF **AP** // et les AI AF **PW** / AI **AF** ] IP

**Condition 2 :**

**G2** [ les AI AF **PW** / et les AI **AF** ]<sub>ip</sub> // [ AI **AF** ] IP

**G3** [ les AI **AF** ]<sub>AP</sub> / [ et les AI **AF** ]<sub>ip</sub> // [ AI **AF** ] IP

**G4** [ les AI AF **PW** / et les AI AF **AP** // AI **AF** ] IP

Par ailleurs, lorsqu'il y a un renforcement des proéminences finales à la fois sur N1 et N2 (sur le plan syntagmatique), comme en G3 dans les deux conditions, nous faisons l'hypothèse que la proéminence finale associée au niveau de frontière le plus élevé est perçue plus forte que la proéminence finale associée au niveau de frontière le moins élevé. Enfin, étant donné la présence d'une frontière d'IP à la fin du syntagme, nous attendons également la présence d'un renforcement de la proéminence finale sur ADJ.

En mettant en relation tous les groupements associés aux deux conditions syntaxiques, nous supposons qu'une frontière d'AP peut ou non être associée à un renforcement de la proéminence finale. La frontière d'AP après N1 serait effectivement marquée par un renforcement de la proéminence finale en G3, et non pas en G4. Nous pouvons mettre en lien cette hypothèse avec celle faite sur les forces des frontières, où nous attendons différents degrés de frontière pour le niveau de l'AP (*cf.* Hypothèse 1). Dans ce contexte, nous faisons l'hypothèse que la frontière d'AP marquée par un degré de frontière plus élevé (comme après N1 en G3 associé à la Condition 2) serait associée à un renforcement de la proéminence finale précédant la frontière, tandis que la frontière d'AP marquée par un degré de frontière plus faible (comme après N1 en G4 associé à la Condition 1) ne serait pas associée à un renforcement de la proéminence finale précédant la frontière. L'accentuation finale ne s'effacerait donc pas en frontières prosodiques majeures, et viendrait au contraire « supporter » le renforcement du degré de frontière.

### 4.3 Hypothèse 3 : Un rôle plus structurel que purement rythmique de l'AI

Comme pour la proéminence finale, nous supposons que les scores de proéminences initiales varient en fonction du type de groupement. Selon nous, AI serait perçu sur tous les mots lexicaux (N1, N2 et ADJ), avec cependant une force plus élevée lorsque la proéminence est positionnée après la limite d'un groupement, supposant par ailleurs à cet endroit une force de frontière également plus élevée (cf. Hypothèse 1). A l'instar de Fougeron & Keating (1997), nous supposons en effet un renforcement articulatoire en début de constituant qui s'incrémente en fonction du niveau de frontière. Il est en effet important de préciser que le renforcement du degré de la proéminence initiale par rapport au niveau de frontière est uniquement effectué, selon nous, sur la proéminence suivant la frontière : la force de l'AI sur ADJ varierait en fonction de la frontière réalisée après N2, et la force de l'AI sur N2 varierait en fonction de la frontière réalisée après N1. Ainsi, comme pour AF, AI serait métriquement fort à tous les niveaux de constituance, et sa force en structure de surface augmenterait en fonction du niveau de frontière marqué avant la syllabe accentuée. Par ailleurs, en admettant qu'AI serait plus structurel que purement rythmique, le degré des proéminences initiales ne devrait pas varier en fonction de la taille du constituant.

Dans ce contexte, nous supposons, tout d'abord en Condition 1, un renforcement de la proéminence initiale sur N2 en G1 et G3, présentant une limite de groupement, par rapport à la proéminence initiale perçue en G4, ne présentant pas cette fois-ci de limite de groupement. De la même manière, AI sur ADJ serait perçu plus fort en G3, qui est situé à la limite d'un groupement, qu'en G1 et G4. Quant à la Condition 2, nous faisons l'hypothèse d'un renforcement de la proéminence initiale sur N2 en G3 par rapport à la proéminence initiale perçue en G2 et G4, étant donné, là encore, la présence d'une limite de groupement en G3. En ce qui concerne la proéminence initiale sur ADJ, AI serait alors perçu plus fort en G2 et G3, qui présentent une limite de groupement, qu'en G4, qui n'est pas associé à une limite de groupement.

#### Condition 1 :

**G1** [ les AI AF ]<sub>ip</sub> // [ et les AI AF PW / AI AF ]<sub>IP</sub>

**G3** [ les AI AF ]<sub>ip</sub> // [ et les AI AF ]<sub>AP</sub> / [ AI AF ]<sub>IP</sub>

**G4** [ les AI AF AP // et les AI AF PW / AI AF ]<sub>IP</sub>

**Condition 2 :**

**G2** [ les AI AF PW / et les AI AF ]<sub>ip</sub> // [ AI AF ]<sub>IP</sub>

**G3** [ les AI AF ]<sub>AP</sub> / [ et les AI AF ]<sub>ip</sub> // [ AI AF ]<sub>IP</sub>

**G4** [ les AI AF PW / et les AI AF ]<sub>AP</sub> // AI AF ]<sub>IP</sub>

Par ailleurs, lorsqu'il y a un renforcement des prééminences initiales à la fois sur N2 et ADJ (sur le plan syntagmatique), comme en G3 dans les deux conditions syntaxiques, nous faisons l'hypothèse que la proéminence initiale, associée au niveau de frontière le moins élevé, est perçue plus forte que la proéminence initiale, associée au niveau de frontière le plus élevé. En référence aux résultats obtenus dans l'étude en perception d'Astésano et al. (2012 ; 2013), nous nous attendons effectivement à percevoir AI moins fort qu'AF en frontière d'IP, et plus fort qu'AF en frontière d'AP. Dans ce contexte, nous supposons alors qu'AI serait un marqueur privilégié des petits niveaux de constituance, tandis qu'AF serait un marqueur privilégié des grands niveaux de constituance. Enfin, nous faisons également l'hypothèse d'un renforcement systématique de la proéminence initiale sur N1, étant donné sa position initiale d'IP. Il se pourrait même que la première démarcation initiale soit globalement perçue plus forte que les autres proéminences initiales.

Enfin, nous rappelons que nous tenons à contrôler la position de la proéminence initiale, étant donné la vision selon laquelle AI serait un marqueur flou de frontière (Jun & Fougeron, 2000 ; 2002 ; Welby, 2003). Dans ce contexte, nous testerons la force de la proéminence initiale dans chaque groupement et dans chaque taille du constituant, en prenant en compte systématiquement ses trois positions possibles (première, deuxième et troisième syllabe). Toutefois, nous faisons l'hypothèse que la perception des proéminences sur les syllabes intermédiaires ne varie ni en fonction du type de groupement, ni en fonction du nombre de syllabes du constituant. Nous attendons en effet une force des proéminences quasi-équivalente sur les syllabes intermédiaires et sur les syllabes inaccentuées des mots grammaticaux, quel que soit le niveau de constituance ou le nombre de syllabes du constituant.

### 5. DEMARCHE STATISTIQUE

Afin de répondre à l'ensemble de ces hypothèses, nous proposons d'effectuer une analyse statistique spécifique nous permettant de tester la manière dont les proéminences et les frontières ont été perçues, en fonction du type de groupement choisi par les auditeurs, mais également en fonction du nombre de syllabes des constituants. Sur le plan statistique, nous proposons donc de mettre en relation une variable dépendante avec une ou plusieurs variables indépendantes (ou variables explicatives), dans le but de tester la manière dont la variable dépendante est influencée par la ou les variables indépendantes. Nous choisissons plus précisément d'effectuer un modèle de régression ordinale à effet mixte. Toutefois, avant de présenter plus en détails ce modèle statistique, nous proposons d'expliquer dans un premier temps les analyses préliminaires que nous avons effectuées, et qui nous ont permis notamment de sélectionner les variables pertinentes à intégrer plus tard au modèle mixte.

#### 5.1 Test d'indépendance du Khi deux

Avant chaque modèle de régression ordinale à effet mixte, nous effectuons en effet un test du Khi deux, comme une analyse préliminaire à caractère descriptif, visant à contrôler, sur le plan symétrique, l'existence ou non d'une relation de dépendance entre deux variables. Nous rappelons en effet que notre étude vise à mettre en relation les scores de frontières ou de proéminences avec les scores de groupements et la taille des constituants. Il est donc intéressant, avant de mettre en œuvre un modèle plus complexe, de s'assurer que cette relation de dépendance existe bien. Par ailleurs, nous rappelons que les scores de frontières et de proéminences ont été recueillis sur plusieurs sites au sein d'un même énoncé (après chaque mot pour les frontières, et sur toutes les syllabes pour les proéminences). Chaque site correspond donc à un score de perception, et donc à une variable spécifique. Dans ce contexte, le test du Khi deux nous permet également de sélectionner, en amont, les variables présentant ou non un lien de dépendance, et donc de retenir les variables les plus pertinentes à intégrer au modèle de régression ordinale à effet mixte. Il est par ailleurs important de préciser que le test du Khi deux permet de contrôler la relation de dépendance entre deux variables uniquement sur le plan symétrique, sans contrôler le sens de la dépendance ou le rapport de causalité entre les deux variables. C'est en recourant au modèle mixte que nous pourrions induire un sens plutôt qu'un autre.

Dans le cas de l'analyse des scores de frontières, nous utiliserons également le test du Khi deux pour estimer la relation de dépendance entre le groupement et la condition

syntactique. Nous rappelons en effet que la relation de dépendance supposée entre le groupement et les scores de frontières fait intervenir par ailleurs la condition syntaxique, que nous supposons fortement corrélée à la catégorie de groupement (*cf.* Hypothèse 1). Il est donc intéressant de vérifier, dans un tout premier temps, l'adéquation des prédictions associées aux deux conditions syntaxiques avec les résultats de perception des groupements, avant même d'examiner la relation de dépendance entre le groupement et les scores de frontières.

### **5.2 Modèle de régression ordinale à effet mixte**

Les variables présentant un lien de dépendance sont alors ensuite incluses dans un modèle plus complexe permettant, comme nous l'avons déjà signalé, de tester la manière dont la variable dépendante est influencée par la ou les variables indépendantes. Plus précisément, dans le cadre de notre étude, nous choisissons de considérer les scores de frontières ou de proéminences comme étant les variables dépendantes, et les scores de groupements et la taille du constituant comme les variables indépendantes. En dehors des facteurs fixes (groupement et taille du constituant), l'avantage d'un modèle mixte est également de pouvoir inclure dans le modèle des facteurs aléatoires, permettant de prendre en compte la variation due à certains facteurs. Par définition, les facteurs, ou variables, aléatoires sont considérés comme des paramètres dont les niveaux n'ont pas été fixés par l'expérimentateur, et qui représentent un échantillon aléatoire d'une population plus large. Dans le cadre de notre étude, nous choisissons donc de prendre en compte la variabilité due aux participants, aux locuteurs et aux types de syntagmes.

Nous aurions pu également proposer d'inverser les variables dépendantes et indépendantes de notre modèle, dans la mesure où nous pouvons supposer que les scores de frontières et de proéminences influencent la manière dont nous regroupons les mots les uns avec les autres, et non pas l'inverse. Nous choisissons cependant de conserver l'information de groupement comme variable indépendante. Selon nous, le groupement explique autant les scores de frontières et de proéminences, que les scores de frontières et de proéminences expliquent le groupement. Le fait que l'auditeur segmente un énoncé en deux groupements différents, par exemple, pourrait expliquer la perception d'une frontière forte plutôt que d'une frontière faible entre les deux groupements. De même, le fait que l'auditeur perçoive une frontière forte plutôt qu'une frontière faible entre deux constituants, pourrait expliquer la segmentation de l'énoncé en deux groupements. Le choix de garder l'information de groupement comme variable indépendante, dépend par ailleurs de la

formulation de nos questions de recherches et de nos hypothèses. Comme nous l'avons signalé au début de cette partie, nous tenons en effet à expliquer la perception de tel ou tel niveau de frontière et de proéminence par le choix du groupement, mais aussi par la taille du constituant.

Il est également intéressant de préciser que nous choisissons de procéder à un modèle de régression ordinale à effet mixte<sup>11</sup>, étant donné la nature de notre variable dépendante. En effet, nous rappelons que les scores de frontières et de proéminences, qui sont les variables dépendantes de nos modèles, ont été recueillis via une échelle d'évaluation allant de 0 à 3. Ces scores de perception correspondent donc à des données ordinales, et non pas à des données numériques continues. Pour ces raisons, nous privilégions un modèle linéaire généralisé plutôt qu'un modèle linéaire général, plus classique, qui convient uniquement aux variables dépendantes continues. Plus précisément, quand la variable dépendante est ordinale, la plupart des auteurs, tels que Liao (1994), suggèrent de recourir à un modèle de probabilité permettant, dans notre cas, d'établir la probabilité que les scores de frontières ou de proéminences prennent une valeur de 0, 1, 2 ou 3, en fonction du choix du groupement et de la taille du constituant. L'interprétation des résultats est alors ici plus difficile que dans le cas d'un modèle linéaire « traditionnel », où la relation entre la variable dépendante et indépendante est directement établie à partir de la valeur numérique de la variable indépendante.

Pour ce faire, nous effectuons nos modèles sur le logiciel R (R Core Team, 2016), en utilisant plus spécifiquement le package « ordinal » (Christensen, 2015), qui permet l'exécution des modèles de régression ordinale à effet mixte (clmm : « a cumulative link mixed models »). Pour l'analyse post hoc de nos résultats, nous avons également utilisé le package « lsmeans : Least-Squares Means » (Lenth, 2016), permettant de calculer les contrastes deux à deux des niveaux de la variable réponse Y (variable dépendante) pour chaque niveau des facteurs (variable indépendante) entrant dans le modèle considéré, et de tester ainsi la significativité des différences entre ces contrastes. Plus précisément, dans le cadre du calcul de ces contrastes, nous avons d'abord utilisé la commande « lsmeans » associée au mode « prob », permettant d'estimer la distribution de probabilités de chacune des valeurs de la variable ordinale dans chaque niveau des facteurs. Pour mieux comprendre, nous proposons l'exemple d'un résultat factice (cf. Tableau N°1), où la probabilité que la variable dépendante (scores de frontières) prenne une valeur de niveau 0, 1, 2 ou 3 est estimée pour chaque niveau de la variable indépendante (scores de

---

<sup>11</sup> Nous tenons à remercier Lorraine Baqué pour son aide dans la mise en place de ce modèle statistique. Nous remercions également Sébastien Déjean et Saïd Jmel pour leurs conseils avisés.

groupements : G1, G2, G3 et G4). Nous observons, par exemple, une probabilité de 60,3% de percevoir un score de niveau 0 en G1.

**Tableau 1** : Exemple d'un résultat factice illustrant la distribution de probabilités de chaque valeur de la variable ordinale (le score de frontière : 0, 1, 2 et 3) dans chaque niveau de la variable indépendante (groupement : G1, G2, G3 et G4)

Distribution de probabilités des scores de frontières pour chaque groupement				
score de frontière	groupement			
	G1	G2	G3	G4
<b>0</b>	60,3%	50,6%	26,1%	1,7%
<b>1</b>	30,9%	36,9%	44,6%	8,8%
<b>2</b>	7,2%	10,2%	22,9%	31,2%
<b>3</b>	1,6%	2,3%	6,4%	58,3%

Afin de garder la simplicité d'interprétation du modèle linéaire « traditionnel », nous avons ensuite utilisé la commande « lsmeans » associée cette fois-ci au mode « mean.class », indiquant les moyennes des différents niveaux de la variable ordinale, interprétées comme des valeurs numériques, ordonnées de 1 à 4, et calculées à partir des résultats de la distribution de probabilités estimée pour chaque niveau des facteurs (cf. Tableau N°2). Cette dernière commande ne nous permet pas cependant de retrouver notre échelle d'évaluation, à partir de laquelle nous avons recueilli nos scores de perception, à savoir une échelle allant de 0 à 3. De fait, nous proposons de présenter les valeurs « mean.class » en modifiant l'échelle des valeurs numériques ordonnées. Pour cela, nous avons soustrait la valeur « 1 » à toutes les valeurs « mean.class », interprétées comme des valeurs numériques ordonnées de 1 à 4, pour ainsi obtenir les valeurs « mean.class (-1) », c'est-à-dire des valeurs numériques ordonnées cette fois-ci de 0 à 3, qui correspondent donc mieux à notre échelle d'évaluation (cf. Tableau N°2).

Modification de l'échelle des valeurs numériques ordonnées :

$$4 - 1 = 3$$

$$3 - 1 = 2$$

$$2 - 1 = 1$$

$$1 - 1 = 0$$

Cependant, dans le cas des valeurs « mean.class » associées aux variables des scores de force relative des frontières ou des proéminences, les moyennes des différents niveaux de la variable ordinale sont interprétées comme des valeurs numériques ordonnées de 1 à 7. Nous rappelons en effet qu'en calculant la force relative entre des événements prosodiques, nous obtenons non plus 4 scores, mais 7 scores différents (cf. la partie consacrée à la préparation de notre base de données) : -3, -2, -1, 0, 1, 2, 3. Dans ce cas, au lieu de soustraire la valeur « 1 » à toutes les valeurs « mean.class », nous proposons de soustraire la valeur « 4 », pour ainsi obtenir les valeurs « mean.class (-4) », correspondant à notre échelle initiale.

### Modification de l'échelle des valeurs numériques ordonnées en cas de force relative :

$$7 - 4 = 3$$

$$6 - 4 = 2$$

$$5 - 4 = 1$$

$$4 - 4 = 0$$

$$3 - 4 = -1$$

$$2 - 4 = -2$$

$$1 - 4 = -3$$

Il faut savoir par ailleurs que ce sont les valeurs « mean.class » rééchelonnées (« mean.class (-1) » ou « mean.class (-4) ») qui ont été utilisées dans l'élaboration de tous nos graphiques.

**Tableau 2 :** Exemple d'un résultat factice illustrant les moyennes des différents niveaux de la variable ordinale (score de frontière : 0, 1, 2, 3) interprétées comme des valeurs numériques ordonnées de 1 à 4 dans la colonne « mean.class » et de 0 à 3 dans la colonne « mean.class (-1) », et calculées à partir des résultats de la distribution de probabilités estimée pour chaque groupement (cf. Tableau N°1)

La colonne « .group » indique les contrastes significatifs : les groupements partageant une même lettre dans la colonne ne présentent pas de différence significative.

Moyennes des scores de frontières calculées à partir des résultats de la distribution de probabilités			
groupement	mean.class	mean.class(-1)	.group
G1	1,50	0,50	a
G2	1,64	0,64	
G3	2,09	1,09	b
G4	3,39	2,39	c

Pour ensuite interpréter la significativité des différences entre chaque contraste, et donc entre chaque valeur « mean.class(-1) » ou « mean.class(-4) » associées aux différents niveaux de la variable ordinale, nous utilisons la fonction « cld » (Piepho, 2004) permettant d'associer une série de lettres à chacune des comparaisons (*cf.* colonne « .group » du Tableau N°2). Ces séries de lettres renvoient plus précisément aux *p-values* associées à chaque comparaison des valeurs « mean.class » (G1 vs G2 ; G1 vs. G3 ; G1 vs. G4 ; G2 vs. G3 ; G2 vs. G4 ; G3 vs. G4). Lorsque la *p-value* dépasse le seuil de significativité (0,05), cela signifie alors que les deux valeurs « mean.class » ont au moins une lettre en commun. Autrement dit, les groupes partageant une même lettre dans la colonne « .group » ne présentent pas de différence significative. Dans l'exemple de notre résultat factice (*cf.* Tableau N°2), la frontière est perçue significativement plus forte en G4 (mean.class : 2,39) qu'en G3 (mean.class : 1,09), et les frontières en G4 et G3 sont également perçues significativement plus fortes qu'en G1 et G2 (mean.class : 0,50 et 0,64), qui ne s'opposent pas quant à eux significativement.

Enfin, nous avons appliqué la même procédure pour chacun de nos modèles : nous démarrions avec le modèle le plus complet et nous excluons progressivement du modèle final les variables et interactions non-significatives. Dans le cadre de la présentation de nos résultats (*cf.* Chapitre 5), nous commentons uniquement les tableaux composés des valeurs « mean.class » et des indices de significativité (les lettres dans la colonne « .group »). Les tableaux de la distribution des probabilités estimées pour chacune des valeurs de la variable ordinale dans chaque niveau des facteurs, figurent quant eux dans les Annexes N°9 et N°16.



**- PARTIE II -**  
**Analyses Expérimentales**



## CHAPITRE 5 – Résultats et Discussion

---

<b>1. PERCEPTION DE LA FORCE DES FRONTIERES .....</b>	<b>189</b>
1.1 Lien de dépendance entre le groupement et la condition syntaxique .....	189
1.2 Statistiques descriptives des principales variables .....	191
1.2.1 Distribution des scores de frontières par groupement.....	192
1.2.2 Distribution des scores de frontières par taille du constituant .....	197
1.3 Choix des sites d'analyse.....	199
1.4 Résultats du modèle de régression ordinale à effet mixte .....	200
1.4.1 Force de la frontière perçue après N1 .....	200
1.4.2 Force de la frontière perçue après N2 .....	202
1.4.3 Force relative entre les deux frontières perçues (N2 ADJ – N1 N2).....	208
1.5 Discussion.....	214
<b>2. PERCEPTION DE LA FORCE DES PROEMINENCES .....</b>	<b>220</b>
2.1 Statistiques descriptives des principales variables .....	221
2.1.1 Distribution des scores de proéminences par groupement .....	222
2.1.2 Distribution des scores de proéminences par taille du constituant.....	227
2.2 Choix des sites d'analyse.....	233
2.3 Résultats du modèle de régression ordinale à effet mixte .....	234
2.3.1 Force de la proéminence perçue sur la syllabe finale de N1.....	235
2.3.2 Force de la proéminence perçue sur la syllabe initiale de ADJ.....	238
2.3.3 Force relative perçue entre la proéminence finale sur N1 et les autres proéminences initiales ou finales adjacentes.....	243
2.3.3.1 Proéminence finale sur N1 – Proéminence initiale sur N1 .....	244
2.3.3.2 Proéminence initiale sur N2 – Proéminence finale sur N1 .....	246
2.3.3.3 Proéminence finale sur N2 – Proéminence finale sur N1 .....	248
2.3.4 Force relative perçue entre la proéminence initiale sur ADJ et les autres proéminences initiales ou finales adjacentes.....	252
2.3.4.1 Proéminence finale sur ADJ – Proéminence initiale sur ADJ.....	253
2.3.4.2 Proéminence initiale sur ADJ – Proéminence finale sur N2 .....	257
2.3.4.3 Proéminence initiale sur ADJ – Proéminence initiale sur N2 .....	262
2.4 Discussion.....	266
<b>3. CORRELATS ACOUSTIQUES DES FRONTIERES ET DES PROEMINENCES PERÇUES .....</b>	<b>274</b>
3.1 Démarche statistique .....	275
3.2 Statistiques descriptives des principales variables acoustiques .....	276
3.3 Résultats du modèle de régression ordinale à effet mixte .....	286
3.3.1 Corrélat acoustiques de la force des frontières.....	287
3.3.1.1 Choix des sites d'analyse.....	287
3.3.1.2 Force de la frontière perçue après N1 .....	289
3.3.1.3 Force de la frontière perçue après N2 .....	291

3.3.1.4	Discussion .....	293
3.3.2	Corrélat acoustique de la force des prééminences .....	298
3.3.2.1	Sites d'analyse .....	298
3.3.2.2	Force de la prééminence perçue sur la syllabe finale de N1.....	302
3.3.2.3	Force de la prééminence perçue sur la syllabe initiale de ADJ.....	307
3.3.2.4	Force de la prééminence perçue sur les syllabes initiales des mots lexicaux 312	
3.3.2.5	Discussion .....	312

### 1. PERCEPTION DE LA FORCE DES FRONTIERES

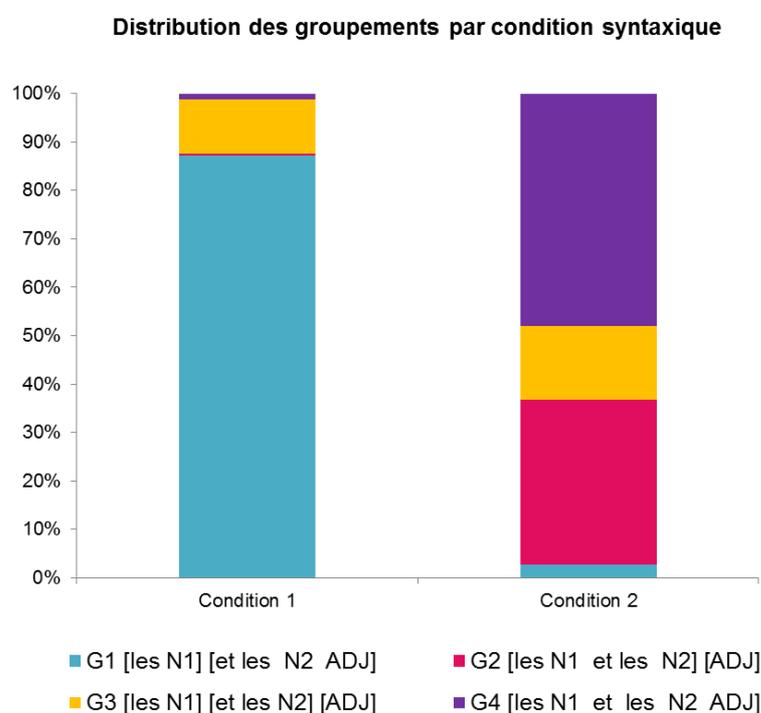
Nous proposons de commencer ce chapitre par la présentation des résultats associés à la perception de la force des frontières. Dans un deuxième temps, nous mettrons en relation ces analyses avec les résultats des scores de proéminences, pour pouvoir ainsi rendre compte de la manière dont la force des proéminences varie en fonction du niveau de frontière. Enfin, nous présenterons dans un troisième temps les corrélats acoustiques des différents événements perçus.

Nous nous intéressons donc tout d'abord à la manière dont les frontières ont été perçues en fonction du type de groupement choisi par les auditeurs, mais également en fonction de la taille du constituant. Toutefois, avant d'exposer les résultats liés à la perception des scores de frontières, nous choisissons de débiter cette partie en présentant dans un premier temps les résultats du test du Khi deux, contrôlant la relation de dépendance entre le groupement et la condition syntaxique. Nous rappelons que nos hypothèses sur le lien entre le groupement et les scores de frontières supposent par ailleurs une relation de dépendance entre le groupement et la condition syntaxique (*cf.* Hypothèse 1). Après avoir commenté les résultats de ce premier test du Khi deux, nous pourrions alors présenter les statistiques descriptives, visant à contrôler cette fois-ci l'existence ou non d'une relation de dépendance entre les scores de frontières et les groupements, puis entre les scores de frontières et la taille du constituant. Comme nous l'avons déjà mentionné dans notre démarche statistique, seules les variables présentant un lien de dépendance seront ensuite incluses dans notre modèle de régression ordinaire à effet mixte, dont les résultats seront présentés par la suite. L'ensemble des résultats seront enfin discutés dans le cadre d'une dernière partie.

#### 1.1 Lien de dépendance entre le groupement et la condition syntaxique

Nous supposons en effet un lien de dépendance entre les catégories de groupements et les deux conditions syntaxiques. Selon nous, G1 et G2 correspondraient respectivement à la Condition 1 et 2, tandis que G3 et G4 pourraient s'appliquer équitablement aux deux conditions, avec cependant une différence dans la hiérarchie entre les frontières perçues après N1 et N2, selon la condition associée au groupement : si G3 et G4 correspondent à la Condition 1, la frontière après N1 devrait être perçue plus forte que la frontière après N2, et si G3 et G4 correspondent à la Condition 2, la frontière après N1 devrait être perçue moins forte que la frontière après N2.

Les résultats du test du Khi deux confirment le lien de dépendance entre la condition syntaxique et le groupement (Khi-deux= 1139, 84, ddl = 3,  $p < 0.001$ ). Nous voyons plus précisément (cf. Figure N°47 et Tableau de contingence en Annexe N°6) une répartition des groupements nettement différente entre les deux conditions syntaxiques. Nous constatons en Condition 1 une importante majorité de G1 (87,2%) ainsi qu'un certain nombre de G3 (11,4%). En revanche, presque aucun G4 (1,1%) ou G2 (0,3%) n'a été perçu en Condition 1. Concernant la Condition 2, nous observons une distribution davantage répartie entre les différents groupements. Cette condition syntaxique semble en effet avoir été massivement perçue comme un G4 (48,1%) ou un G2 (34,2%), et un peu moins comme un G3 (15,1%). Par ailleurs, presque aucun G1 (2,6%) n'a été perçu pour cette condition.



**Figure 47 :** *Distribution des groupements (G1, G2, G3 et G4) par condition syntaxique (Condition 1 à gauche et Condition 2 à droite)*

Les résultats du test confirment donc nos hypothèses de départ (cf. Hypothèse 1): G1 n'est quasiment pas associé à la Condition 2 et G2 n'est pas perçu pour la Condition 1. En revanche, alors que nous pouvions nous attendre à une répartition relativement équivalente de G3 et G4 sur les deux conditions, seul G3 a été associé, équitablement, aux Conditions 1 et 2, et G4 a été massivement perçu uniquement pour la Condition 2. Nous supposons que

la pause relativement longue réalisée en Condition 1 entre N1 et N2 pourrait expliquer le fait que G4 n'ait pas été perçu pour cette condition syntaxique.

A l'issue de ce test, nous décidons d'exclure de nos futures analyses les cas trop marginaux où la Condition 1 a été perçue comme G2 (0,3%) et G4 (1,1%), et les cas où la Condition 2 a été perçue comme G1 (2,6%). De la même manière, nous choisissons de recoder notre actuelle base de données pour ajouter l'information de la condition syntaxique à chaque catégorie de groupement, et permettre ainsi de distinguer G3 associé à la Condition 1 et G3 associé à la Condition 2. Ce codage a par ailleurs été appliqué aux autres catégories de groupements (G1, G2 et G4) afin de tout uniformiser. Notre base de données est alors désormais composée d'une nouvelle variable « *group\_cond* » constituée de 5 groupements différents :

<b>G1</b> * Condition 1	→	<b>G11</b>	[les N1] [et les N2 ADJ]
<b>G3</b> * Condition 1	→	<b>G31</b>	[les N1] [et les N2] [ADJ]
<b>G2</b> * Condition 2	→	<b>G22</b>	[les N1 et les N2] [ADJ]
<b>G3</b> * Condition 2	→	<b>G32</b>	[les N1] [et les N2] [ADJ]
<b>G4</b> * Condition 2	→	<b>G42</b>	[les N1 et les N2 ADJ]

Il semblerait que les participants de notre expérimentation soient alors capables de distinguer les deux conditions syntaxiques, en percevant notamment différents phrasés prosodiques possibles pour une même structure syntaxique. Dans le cadre de notre travail, nous nous intéressons plus spécifiquement à ces différents phrasés prosodiques, en testant l'existence ou non d'une relation de dépendance entre les scores de frontières et les groupements, illustrant les phrasés prosodiques possibles pour chaque condition syntaxique.

### **1.2 Statistiques descriptives des principales variables**

Sur la base de ces résultats, nous pouvons en effet nous intéresser plus directement à la relation entre les différents groupements, chacun associé à une condition syntaxique spécifique, et les scores de frontières perçues après chaque mot du syntagme. Nous rappelons que nous nous intéressons également au lien de dépendance entre le nombre de syllabes des constituants et les scores de frontières. Pour l'ensemble des tests du Khi deux suivants, nous choisissons de commenter les histogrammes illustrant les différentes distributions. Nous renvoyons le lecteur aux Annexes N°7 et N°8, où les tableaux de

contingence avec les effectifs et les pourcentages associés aux différents résultats sont présentés.

### 1.2.1 Distribution des scores de frontières par groupement

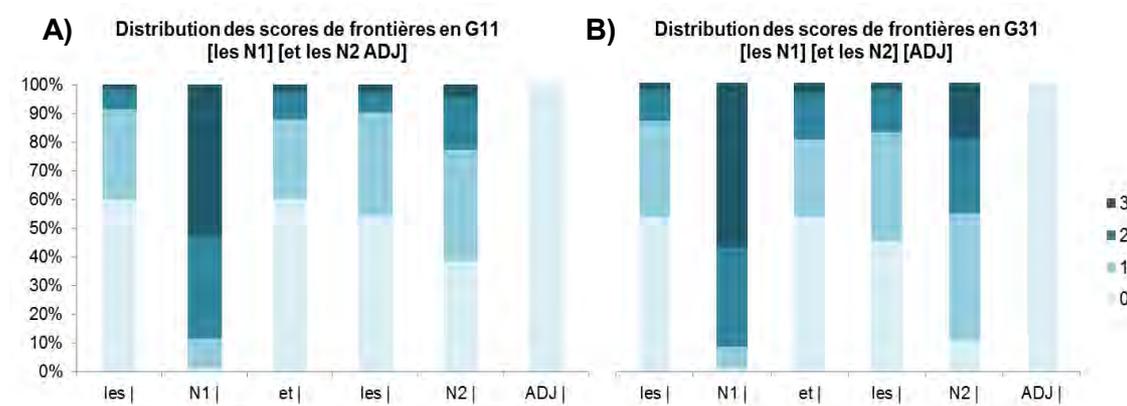
Les résultats du test du khi deux confirment en effet un lien de dépendance entre la force des frontières et le type de groupement. Nous observons plus précisément un lien de dépendance entre le groupement et la force des frontières perçues après N1 (Khi-deux= 899,917, ddl= 12,  $p < 0.001$ ) et N2 (Khi-deux= 197,499, ddl= 12,  $p < 0.001$ ), mais aussi entre le groupement et la force des frontières perçues après la conjonction de coordination, notée « cc » (Khi-deux= 45,933, ddl= 12,  $p < 0.001$ ) et le deuxième déterminant, noté « D2 » (Khi-deux= 33,199, ddl= 12,  $p = 0.001$ ) (cf. Tableaux de contingence en Annexe N°7). Aucune relation n'est en revanche observée entre la catégorie de groupement et la force des frontières perçues après le premier déterminant et après ADJ ( $p > 0,05$ ). Dans le cadre de nos hypothèses, nous n'attendions de lien de dépendance spécifique ni entre les groupements et les frontières perçues après les mots grammaticaux (D1, cc, D2), ni entre les groupements et la frontière perçue après ADJ. Nous supposons en revanche une relation de dépendance forte avec les scores de frontières perçues après N1 et N2. Les résultats de cette relation de dépendance montrent plus précisément deux profils distributionnels différents: en G11 et G31 d'un côté, et en G22, G32 et G42 de l'autre. Il est intéressant de remarquer que ces deux profils respectent par ailleurs la distinction entre les deux conditions syntaxiques de départ.

#### Frontières après N1 et N2 en Condition 1 → G11 et G31 (cf. Figure N°48):

La frontière après N1 montre une majorité de score de niveau 3 (G11 : 52,5% et G31 : 57%) mais aussi une grande proportion de score de niveau 2 (G11 : 36% et G31 : 34%). En revanche, peu de scores de niveau 0 (G11 : 1,3% et G31 : 1%) et de niveau 1 ont été perçus (G11 : 10,2% et G31 : 7%). Quant à la frontière après N2, nous observons un degré de frontière globalement moins élevé qu'après N1, qui reste toutefois plus élevé que le degré de frontière après les mots grammaticaux (cc et D2). Une augmentation globale des scores de niveau 2 et 3 est en effet observée après N2 en G11 (22,7% contre 9,7% à 12,1% après cc et D2), et en G31 (46% contre 17% à 19,6% après cc et D2) (cf. Tableaux de contingence en Annexes N°7-B et N°7-E).

Il semblerait donc que la force de frontière après N1 soit globalement plus importante que la force de frontière après N2. Par ailleurs, l'augmentation des scores de frontières

après N2 se révèle être plus importante en G31 qu'en G11, tandis que le degré de frontière après N1 apparaît relativement équivalent entre les deux groupements. Ce résultat rappelle les hypothèses que nous avons faites quant aux niveaux de constituance perçus par les auditeurs (cf. Hypothèse 1). Nous supposons en effet un niveau équivalent entre les deux groupements après N1 (frontières d'ip), et un niveau plus élevé en G31 (frontière d'AP) qu'en G11 (frontière de PW) après N2. Concernant enfin les scores de frontières après les mots grammaticaux, nous observons des scores très faibles dans les deux groupements, avec toutefois des scores un peu plus élevés en G31 qu'en G11. La répartition du score de niveau 0 est effectivement moins importante en G31 (45% à 54%) qu'en G11 (54,3% à 60,4%), tandis que la répartition du score de niveau 2 est plus importante en G31 (11% à 16%) qu'en G11 (7% à 9% en G11).



**Figure 48 :** Distribution des scores de frontières (de 0 à 3) après chaque mot du syntagme par catégorie de groupement : en G11 dans la figure (A) et en G31 dans la figure (B)

Frontières après N1 et N2 en Condition 2 → G22, G32 et G42 (cf. Figure N°49) :

La frontière après N2 semble être perçue plus forte que la frontière après N1. Une augmentation générale des scores de niveau 2 et 3 est en effet observée après N2 en G22 (56,9% contre 15,5% après N1), en G32 (48,7% contre 26,6% après N1) et en G42 (22,8% contre 8,1% après N1). Par ailleurs, bien que le niveau de frontière après N1 apparaisse moins important que le niveau de frontière après N2, la frontière après N1 semble cependant être perçue plus forte que les frontières après les mots grammaticaux (cc et D2). Les scores de niveau 2 et 3 tendent en effet à être plus nombreux après N1 qu'après les mots grammaticaux en G22 (15,5% contre 8,5% à 9,3% après cc et D2), en G32 (26,6% contre

5,5% à 6,4% après cc et D2), et en G42 (8,1% contre 5,2% à 5,8% après cc et D2) (cf. Tableaux de contingence en Annexes N°7-B et N°7-E).

La force de frontière après N2 semble alors globalement plus élevée que la force de frontière après N1. Par ailleurs, l'augmentation des scores de frontières après N2 apparaît plus importante en G22, un peu moins importante en G32, et encore moins importante en G42, tandis que le degré de frontière après N1 semble plus important en G32 qu'en G22 et G42. Ces résultats correspondraient alors à nos prédictions sur les niveaux de constituance perçus par les auditeurs (cf. Hypothèse 1). Nous supposons en effet un niveau plus élevé en G32 (frontière d'AP) qu'en G22 et G42 (frontières de PW) après N1, et un niveau plus élevé en G22 (frontière d'ip) qu'en G42 (frontière d'AP) après N2. En revanche, alors que les résultats semblent indiquer des scores de frontières plus élevés après N2 en G22 qu'en G32, nos hypothèses prédisent quant à elles un niveau de constituance similaire entre les deux groupements (frontières d'ip). Enfin, les scores de frontières après les mots grammaticaux se révèlent être très faibles dans les trois groupements, avec notamment une majorité de score de niveau 0 en G22 (61,8% à 72,8%), en G32 (58,7% à 79,8%) et en G42 (65,6% à 71,4%).

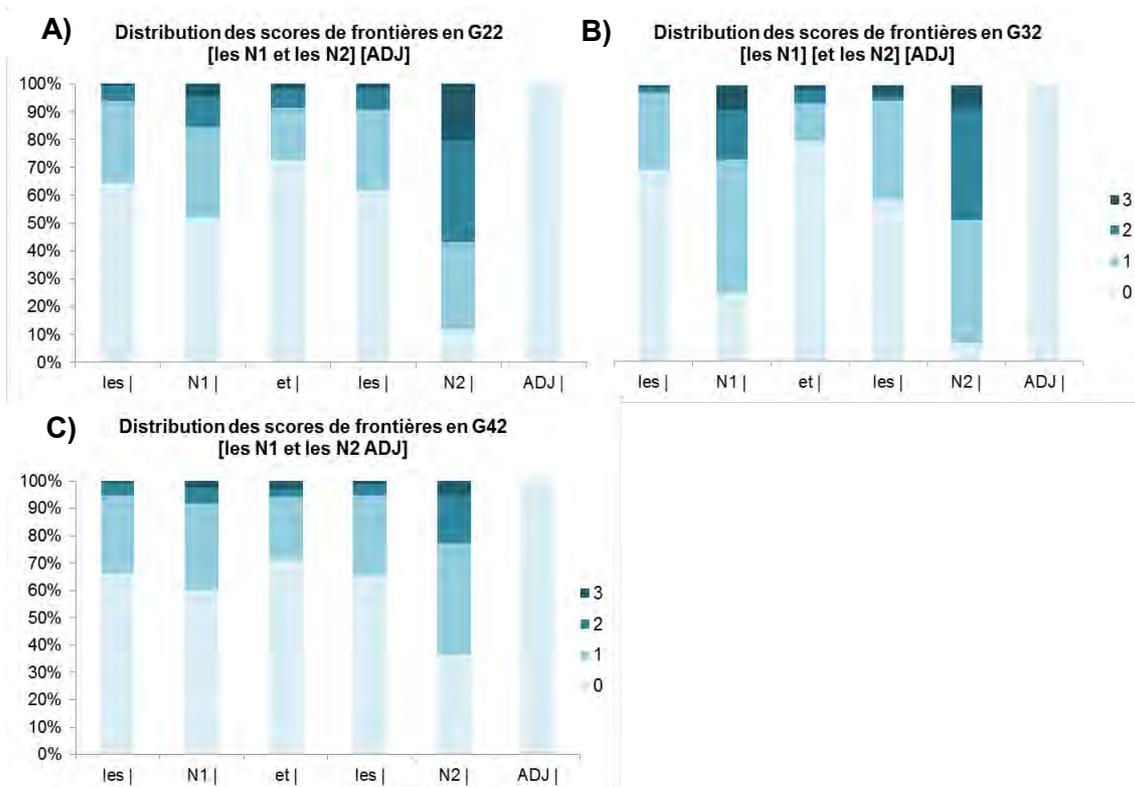


Figure 49 : Distribution des scores de frontières (0 à 3) après chaque mot du syntagme par catégorie de groupement : en G22 dans la figure (A), en G32 dans la figure (B) et en G42 dans la figure (C)

Synthèse des résultats :

Deux profils de force des frontières semblent donc avoir été perçus en G11 et G31 d'une part puis en G22, G32 et G42 d'autre part. La différence entre ces deux distributions est notamment liée aux variations des scores de frontières perçues après N1 et N2 : en G11 et G31, la frontière après N1 est perçue plus forte que la frontière après N2, tandis qu'en G22, G32 et G42, c'est la frontière après N2 qui est perçue plus forte que la frontière après N1. En mettant en relation tous les groupements, nous observons également une hiérarchie des niveaux de frontières :

$$N1 \left| \begin{array}{l} \text{en G11} \\ \text{en G31} \end{array} \right\} > N2 \left| \begin{array}{l} \text{en G22} \end{array} \right\} > N2 \left| \begin{array}{l} \text{en G31} \\ \text{en G32} \end{array} \right\} > N2 \left| \begin{array}{l} \text{en G42} \\ \text{en G11} \end{array} \right\} > N1 \left| \begin{array}{l} \text{en G32} \end{array} \right\} > N1 \left| \begin{array}{l} \text{en G22} \\ \text{en G42} \end{array} \right\}$$

Alors que nous supposons dans nos prédictions de départ un même niveau de frontière (frontière d'ip) après N1 en G11 et G31, et après N2 en G22 et G32, il semblerait qu'il y ait ici une variation plus importante des niveaux de frontière. La frontière après N1 en G11 et G31 semble en effet avoir été perçue globalement plus forte que la frontière après N2 en G22, qui elle-même apparaît plus élevée que la frontière après N2 en G32. A l'inverse, la frontière après N2 semble avoir été perçue, au vu de la distribution des scores, aussi forte en G32 qu'en G31, alors que nos hypothèses prédisent un niveau de frontière plus élevé en G32 (frontière d'ip) qu'en G31 (frontière d'AP).

Rappel de nos prédictions :

**Condition 1 :**

**G1** [ les N1 ]<sub>ip</sub> // [ et les N2 **PW** / ADJ ] **IP**

**G3** [ les N1 ]<sub>ip</sub> // [ et les N2 ]<sub>AP</sub> / [ ADJ ] **IP**

**Condition 2 :**

**G2** [ les N1 **PW** / et les N2 ]<sub>ip</sub> // [ ADJ ] **IP**

**G3** [ les N1 ]<sub>AP</sub> / [ et les N2 ]<sub>ip</sub> // [ ADJ ] **IP**

Les résultats des distributions montrent également un degré de frontière après N2 plus important en G32 et G31 qu'en G42 et G11. Ce résultat confirme nos prédictions sur

l'existence d'un niveau de frontière plus important en G32 (frontière d'ip) qu'en G42 (frontière d'AP) et G11 (frontière de PW), mais également sur l'existence d'un niveau de frontière plus important en G31 (frontière d'AP) qu'en G11 (frontière de PW). En revanche, alors que nous supposons un niveau de frontière équivalent en G31 et G42 (frontières d'AP), les résultats de la distribution des scores tendent à montrer un niveau de frontière plus élevé en G31 qu'en G42. Il est toutefois intéressant de mettre en lien ce résultat avec notre hypothèse quant à la possibilité qu'un même niveau de constituance puisse être marqué par différents degrés de frontière. Nous supposons que la frontière d'AP en G31 pourrait être perçue plus forte que la frontière d'AP en G42, étant donné la présence en G31 d'une limite de groupement (signalée par « ] »).

Rappel de nos prédictions :

**Condition 1 :**

**G1** [ les N1 ]<sub>ip</sub> // [ et les N2 <sub>PW</sub> / ADJ ]<sub>IP</sub>

**G3** [ les N1 ]<sub>ip</sub> // [ et les N2 ]<sub>AP</sub> / [ ADJ ]<sub>IP</sub>

**Condition 2 :**

**G3** [ les N1 ]<sub>AP</sub> / [ et les N2 ]<sub>ip</sub> // [ ADJ ]<sub>IP</sub>

**G4** [ les N1 <sub>PW</sub> / et les N2 <sub>AP</sub> // ADJ ]<sub>IP</sub>

Enfin, nos prédictions sur les niveaux de frontière après N1 en G32, G22 et G42 correspondraient aux résultats de la distribution des scores de frontières. La frontière après N1 en G32 (frontière d'AP prédite) apparaît effectivement plus importante que la frontière après N1 en G22 et G42 (frontières de PW prédites).

Rappel de nos prédictions :

**Condition 2 :**

**G2** [ les N1 <sub>PW</sub> / et les N2 ]<sub>ip</sub> // [ ADJ ]<sub>IP</sub>

**G3** [ les N1 ]<sub>AP</sub> / [ et les N2 ]<sub>ip</sub> // [ ADJ ]<sub>IP</sub>

**G4** [ les N1 <sub>PW</sub> / et les N2 <sub>AP</sub> // ADJ ]<sub>IP</sub>

L'ensemble de ces résultats semblent alors montrer que la distribution des scores de frontières serait encore plus nuancée en fonction du groupement, que ce que nous avons pu

prédire initialement. Le modèle de régression ordinale à effet mixte nous permettra de préciser davantage les différences de force entre les frontières, en exprimant notamment la significativité des contrastes deux à deux des différentes catégories de groupements à chaque niveau des scores de frontières. Par ailleurs, bien que les résultats montrent un lien de dépendance entre le type de groupement et les frontières perçues après cc et D2, nous observons peu de variabilité entre les différents groupements. Une très faible frontière (quasi-majorité de score de niveau 0) est en effet systématiquement perçue. Enfin, Nous comprenons l'absence de dépendance entre la catégorie de groupement et la frontière perçue après ADJ qui est systématiquement associée à un score de niveau 0, quel que soit le groupement. Ce résultat pourrait effectivement être lié à la consigne donnée dans la tâche de rupture, où il était demandé aux participants de juger du degré de frontière entre chaque mot. ADJ étant en fin d'énoncé, les participants semblent avoir considéré qu'il ne s'agissait pas d'un site de rupture potentiel.

### 1.2.2 Distribution des scores de frontières par taille du constituant

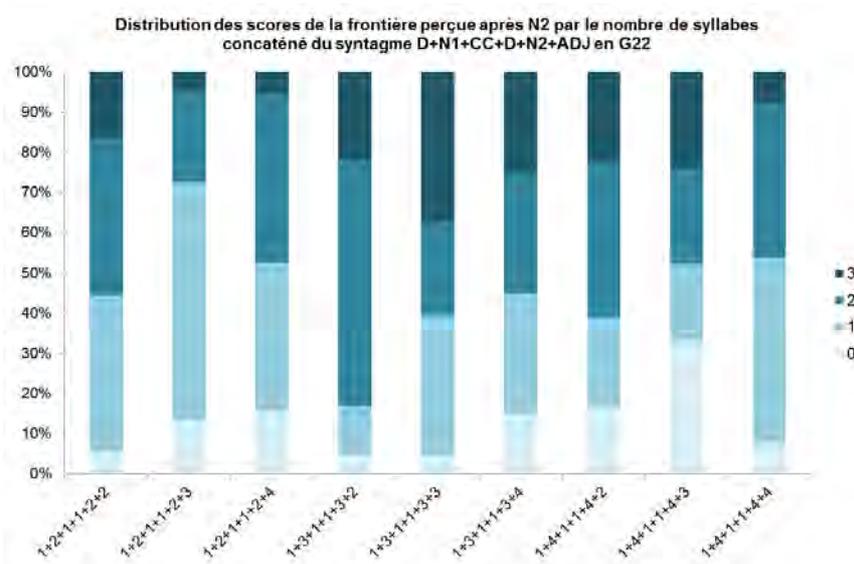
Nous supposons également un lien de dépendance entre le nombre de syllabes du constituant et les scores de frontières. Pour avoir une idée plus précise du lien de dépendance potentiel entre les deux variables, nous avons effectué plusieurs tests du Khi deux en proposant à chaque fois une variable différente liée à la taille du constituant. En effet, nous avons testé le lien de dépendance entre les scores de perception et le nombre de syllabes de N, le nombre de syllabes de ADJ, mais aussi le nombre de syllabes total du syntagme (D1+N1+cc+D2+N2+ADJ), dans une première version où le nombre de syllabes est additionné, puis dans une seconde version où le nombre de syllabes est concaténé<sup>12</sup>. Nous proposons par ailleurs de tester cette relation de dépendance entre les scores de frontières et la taille du constituant en prenant en compte chacune des catégories de groupements, pour faciliter ainsi l'interprétation des résultats au regard toujours des différents phrasés prosodiques possibles pour les deux conditions syntaxiques.

Les résultats du test du Khi deux montrent une relation de dépendance entre la frontière perçue après N2 en G22 et la taille du constituant, que ce soit le nombre de syllabes de N (Khi deux= 18,855, ddl= 6, p=0,004) et de ADJ (Khi deux= 17,066, ddl= 6, p=0,009), ou le nombre de syllabes total du syntagme dans sa version additionnée (Khi deux= 46,007, ddl= 18, p<0,001) et concaténée (Khi deux= 53,486, ddl= 24, p<0,001). Aucune autre relation de dépendance n'a par ailleurs été observée dans les autres

---

<sup>12</sup> Pour plus de détails sur la variable « taille du constituant », nous renvoyons le lecteur à la partie consacrée à la préparation de notre base de données (cf. 3.3 du Chapitre 4).

groupements et sur les autres sites de frontières ( $p > 0,05$ ) (cf. Tableaux de contingence en Annexe N°8). Plus précisément, les résultats montrent que le degré de frontière après N2 semble être plus élevé lorsque le syntagme est composé d'un N de 3 syllabes (cf. Figure N°50). La distribution du score de niveau 3 est en effet plus importante lorsque N est composé de 3 syllabes (22% à 37,2%) plutôt que de 2 ou 4 syllabes (4,5% à 23,8%), quelle que soit la taille de l'ADJ suivant. Nous constatons également une augmentation importante du score de niveau 2 lorsque N est composé de 3 syllabes, mais seulement dans le cas où N est suivi par un ADJ de 2 syllabes (61% contre 22,7% à 42,1% pour les autres « configurations syllabiques ») (cf. Tableau de contingence en Annexe N°8-D).



**Figure 50 :** Distribution des scores de frontières après N2 en G22 par le nombre de syllabes concaténé du syntagme D1+N1+CC+D2+N2+ADJ

Il est intéressant de remarquer que cette relation de dépendance entre les scores de frontières et la taille du constituant apparaît uniquement à la frontière de deux constituants, entre N2 et ADJ, qui varient systématiquement en nombre de syllabes. Nous rappelons en effet que les tailles de N1 et N2 covarient en nombre de syllabes, et ne présentent donc pas de variabilité en termes de « configurations syllabiques » pour une même taille de constituant. Par ailleurs, cette relation de dépendance entre les scores de frontières et la « configuration syllabique » de N2+ADJ est uniquement établie en G22, où le niveau de frontière après N2 est perçu le plus fort par rapport aux autres groupements sur le même site (cf. Figure N°50). La frontière après N2 semble en effet être perçue plus forte en G22 qu'en

G31 et G32, où le niveau de frontière apparaît également plus élevé qu'en G11 et G42. Cela tendrait alors à supposer que le renforcement de la frontière après N2 en G22 soit lié à la taille du constituant. Nous avons en effet supposé dans nos hypothèses un renforcement du niveau de frontière dû à l'augmentation du nombre de syllabes du constituant N2+ADJ. Bien que cette hypothèse ne concerne pas spécifiquement G22, nous pouvons tout de même nous questionner quant à l'influence de la taille du constituant sur la perception du niveau de frontière dans ce groupement.

Le modèle de régression ordinaire à effet mixte nous permettra de mieux répondre à cette question, en testant plus précisément l'effet du nombre de syllabes sur la perception des scores de frontières. Nous pourrions également vérifier si le rehaussement du niveau de frontière est dû à l'augmentation même du nombre de syllabes du constituant, ou plutôt à une taille spécifique du constituant.

### **1.3 Choix des sites d'analyse**

Nos analyses confirment donc tout d'abord l'existence d'une relation de dépendance entre le groupement et la condition syntaxique. Après avoir recodé nos données pour ainsi avoir les informations liées aux conditions syntaxiques associées à chaque catégorie de groupement, nous avons contrôlé plus directement le lien de dépendance entre les scores de frontières et le groupement d'une part, et entre les scores de frontières et le nombre de syllabes des constituants d'autre part. Alors que les scores de frontières montrent une forte relation de dépendance avec les catégories de groupements, la relation de dépendance avec la taille du constituant, en revanche, ne semble pas aussi pertinente. Un seul site de frontière (après N2), dans une catégorie de groupement spécifique (en G22) est effectivement concerné. Toutefois, comme nous l'avons déjà signalé, il est intéressant d'inclure la taille du constituant dans une analyse de régression ordinaire à effet mixte, qui permettra de mieux tester l'influence du nombre de syllabes du constituant sur la perception des frontières. Plus précisément, nous choisissons de tester l'effet de la taille de N et de ADJ, mais aussi l'effet d'interaction entre la taille de N et ADJ, pour ainsi contrôler à la fois la taille et la « configuration syllabique » du constituant. Concernant plus précisément la relation de dépendance entre le groupement et les scores de frontières, comme nous l'avons prédit dans nos hypothèses de départ, la force des frontières après N1 et N2 varie en fonction du type de groupement. La distribution des scores de frontières serait même encore plus nuancée en fonction du groupement, que ce que nous avons pu prédire initialement. Il est donc intéressant de procéder à un modèle de régression ordinaire à effet

mixte, nous permettant de mieux définir la granularité des niveaux de frontière perçus dans chaque catégorie de groupement.

Dans ce contexte, nous choisissons d'inclure dans le modèle mixte trois types de variables dépendantes associées aux scores de frontières. Nous tenons en effet à tester la manière dont les frontières après N1 et N2 ont été perçues en fonction du type de groupement choisi par l'auditeur, mais également en fonction de la taille des constituants. Par ailleurs, afin d'avoir une analyse plus syntagmatique que purement paradigmatique, nous proposons également de tester la manière dont les frontières après N1 et N2 sont perçues l'une par rapport à l'autre, en fonction toujours du type de groupement et du nombre de syllabes du constituant. Cette dernière variable dépendante correspond aux variables que nous avons créées (*cf.* Préparation de la base de données 3.3 du Chapitre 4) visant à rendre compte de la force relative entre deux événements prosodiques. Enfin, dans chacun des modèles effectués, nous conservons systématiquement les trois variables aléatoires définies au préalable, à savoir les variables liées au participant, au locuteur et au type de syntagme.

### **1.4 Résultats du modèle de régression ordinale à effet mixte**

Nous proposons de présenter tout d'abord les résultats du modèle sur chaque site de frontière isolé, pour ensuite présenter les résultats sur la force relative entre les frontières perçues. Pour chaque modèle réalisé, nous avons également analysé les composantes de la variance, qui montrent systématiquement une valeur de la variance liée au participant plus élevée que celle liée au locuteur ou au syntagme. Cela signifierait alors que les résultats obtenus dans le cadre de notre analyse varieraient davantage en fonction de l'auditeur qu'en fonction du locuteur produisant les différents syntagmes. Il semblerait en effet que tous les participants n'aient pas forcément les mêmes critères d'attribution des scores de perception, et qu'ils n'estiment pas non plus de la même manière l'échelle d'évaluation (un score de même niveau pourrait renvoyer à une réalité différente selon les participants). En revanche, il est intéressant de se dire que les résultats obtenus dans nos analyses pourraient être valables quel que soit le locuteur prononçant le syntagme.

#### **1.4.1 Force de la frontière perçue après N1**

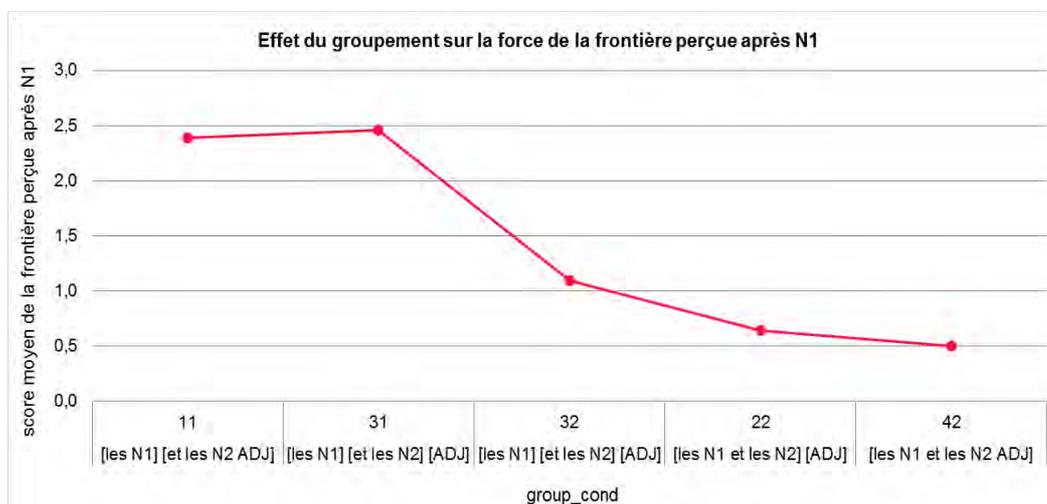
Seul le groupement montre un effet significatif sur la force de la frontière perçue après N1 (LR Chisq= 139.13, df= 4,  $p < 0,001$  *cf.* Annexe N°9-A). Aucun effet de la taille du constituant n'est par ailleurs observé. Plus précisément, les analyses post hoc indiquent une

différence significative entre tous les groupements, excepté entre G11 et G31 d'un côté, et entre G22 et G42 de l'autre (cf. Tableau N°3 et Figure N°51).

**Tableau 3 :** Résultats des comparaisons post hoc de l'effet du groupement sur la force de la frontière perçue après N1

Le tableau présente les moyennes des scores de frontières interprétées comme des valeurs numériques ordonnées de 1 à 4 pour les valeurs « mean.class » et de 0 à 3 pour les valeurs « mean.class(-1) », calculées à partir des résultats de la distribution de probabilités estimée pour chaque catégorie de groupement (cf. Annexe N°9-A). Les groupes partageant une même lettre dans la colonne ".group" ne présentent pas de différence significative.

FRONTIERE PERCUE APRES N1			
Moyennes des scores de frontières calculées à partir des résultats de la distribution de probabilités			
group_cond	mean.class	mean.class(-1)	.group
42	1,50	0,50	a
22	1,64	0,64	
32	2,09	1,09	b
11	3,39	2,39	c
31	3,46	2,46	



**Figure 51:** Illustration de l'effet du groupement sur la force de la frontière perçue après N1. Les points reliés par la ligne correspondent aux valeurs « mean.class (-1) » du tableau N°3 indiquant les moyennes des scores de frontières perçues pour chaque groupement, interprétées comme des valeurs numériques ordonnées de 0 à 3

Trois niveaux de frontière seraient alors perçus par les auditeurs : un premier niveau avec le score de frontière le plus fort en G11 et G31 (score moyen : entre 2,39 et 2,46), un second

niveau avec un score de frontière intermédiaire en G32 (score moyen : 1,09), et enfin un troisième niveau avec le score de frontière le plus faible en G22 et G42 (score moyen : entre 0,50 et 0,64). Cependant, la différence de force entre le niveau de frontière intermédiaire, en G32, et le niveau de frontière le plus faible, en G22 et G42, apparaît nettement moins importante que la différence de force observée entre le niveau de frontière le plus fort, en G11 et G31, et le niveau de frontière intermédiaire, en G32.

Ces résultats confirmeraient alors nos hypothèses de départ quant à la prédiction d'une frontière d'ip en G11 et G31, d'une frontière d'AP en G32 et d'une frontière de PW en G22 et G42.

### Rappel de nos prédictions :

**G11** [ les N1 ]<sub>ip</sub> // [ et les N2 / ADJ ]<sub>IP</sub>

**G31** [ les N1 ]<sub>ip</sub> // [ et les N2 / [ ADJ ] ]<sub>IP</sub>

**G32** [ les N1 ]<sub>AP</sub> / [ et les N2 // [ ADJ ] ]<sub>IP</sub>

**G42** [ les N1 ]<sub>PW</sub> / et les N2 // ADJ ]<sub>IP</sub>

**G22** [ les N1 ]<sub>PW</sub> / et les N2 // [ ADJ ] ]<sub>IP</sub>

En revanche, la différence de force de frontière plus importante observée entre G11, G31 et G32 pourrait supposer une différence entre les deux niveaux de constituance également plus importante que celle prédite initialement entre la frontière d'ip en G11 et G31 et la frontière d'AP en G32. Il sera intéressant de voir si nous retrouvons cet effet dans l'analyse de la force des frontières après N2, où nous avons également des frontières d'ip et d'AP prédites.

### 1.4.2 Force de la frontière perçue après N2

Les résultats montrent un effet significatif du groupement (LR Chisq= 85,196, df=4, p<0,001) et du nombre de syllabes de N (LR Chisq= 7,709, df= 2, p=0,02), ainsi qu'un effet d'interaction entre le groupement et le nombre de syllabes de N (LR Chisq=17,943, df=8, p=0,02) (cf. Annexe N°9-B). Toutefois, bien qu'un effet simple du nombre de syllabes de N soit observé, les analyses post hoc ne présentent aucun contraste significatif entre les trois tailles de N. Ce résultat pourrait être dû à l'implication du nombre de syllabes de N dans son interaction significative avec le groupement.

### Effet simple du groupement :

Les analyses post hoc de l'effet simple du groupement indiquent que la frontière après N2 est perçue plus forte en G31, G32 et G22 (score moyen : entre 1,48 et 1,60) qu'en G11 et G42 (score moyen : entre 0,89 et 0,91). Deux niveaux de frontière seraient donc perçus : un premier niveau de frontière avec le plus faible score de perception en G11 et G42, puis un second niveau de frontière avec le score de perception le plus fort en G31, G32 et G22 (cf. Annexe N°9-B).

### Effet d'interaction entre le groupement et le nombre de syllabes de N :

Le résultat de l'effet simple du groupement peut toutefois être modulé par l'effet d'interaction entre le groupement et le nombre de syllabes de N, qui permet de tester à quel point la relation entre les scores des frontières perçues après N2 et le groupement varie en fonction de la taille du constituant.

En regardant les contrastes post hoc entre les différents groupements dans chaque taille de N (cf. Tableau (B) N°4), il semblerait effectivement que les oppositions initialement prédites par l'effet simple du groupement soient quelque peu nuancées. En effet, seul G22 s'oppose systématiquement à G11 et G42, quelle que soit la taille du constituant. Plus précisément, la distinction entre G11 et G42 d'un côté (scores moyens : 0,83 à 0,94), puis G31, G32 et G22 de l'autre (scores moyens : 1,40 à 1,63), n'apparaît que dans le cas où N est composé de 2 syllabes. Lorsque N est trisyllabique, seul G22 (score moyen : 1,89) s'oppose à G11 et G42 (scores moyens : 0,89 à 0,99), puis lorsque N est quadrisyllabique, seuls G32 et G22 (scores moyens : 1,52 à 1,54) s'opposent à G11 et G42 (scores moyens : 0,85 à 0,90). Aucune différence n'est en revanche observée ni entre G42 et G11 ( $p > 0,05$ ), ni entre G31, G32 et G22 ( $p > 0,05$ ).

Ces résultats ne semblent donc pas correspondre à nos hypothèses de départ, où un même niveau de frontière est attendu en G22 et G32 (frontières d'ip), et où un niveau de frontière encore inférieur est attendu en G31 (frontière d'AP). Dans nos analyses descriptives précédentes, à l'inverse, la distribution des scores semble indiquer des scores de frontières plus élevés après N2 en G22 qu'en G31 et G32. Cela tendrait alors à supposer que la différence entre ces trois groupements soit trop ténue perceptivement pour être significative.

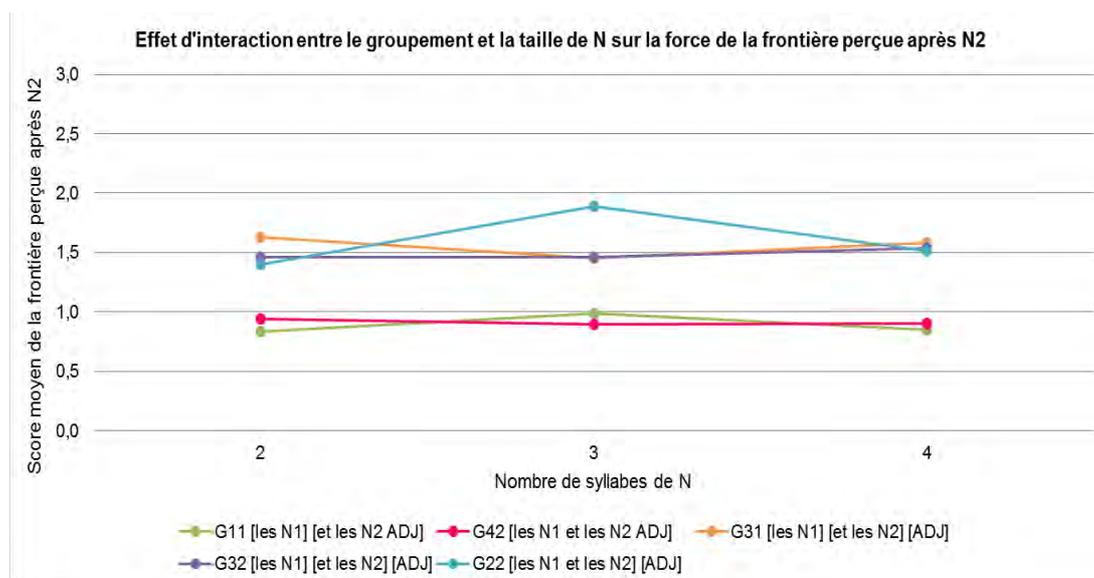
**Tableau 4 :** Résultats des comparaisons post hoc de l'effet d'interaction entre le groupement et le nombre de syllabes de N sur la force de la frontière perçue après N2

A) Lecture des contrastes en colonne pour une même catégorie de groupement et B) Lecture des contrastes en ligne pour un même nombre de syllabes de N. Dans chaque tableau, les moyennes des scores de frontières sont interprétées comme des valeurs numériques ordonnées de 1 à 4 pour les valeurs « mean.class » et de 0 à 3 pour les valeurs « mean.class(-1) », qui sont calculées à partir des résultats de la distribution de probabilités estimée pour chaque catégorie de groupement dans chaque taille de N (cf. Annexe N°9-B). Les groupes partageant une même lettre dans la colonne ".group" ne présentent pas de différence significative.

FRONTIERE PERCUE APRES N2										
Moyennes des scores de frontières calculées à partir des résultats de la distribution de probabilités										
A) Lecture des contrastes en colonne pour une même catégorie de groupement										
nombre de syllabes du N	G11		G42		G31		G32		G22	
	mean.class	.group								
	mean.class(-1)		mean.class(-1)		mean.class(-1)		mean.class(-1)		mean.class(-1)	
2	1,83	a	1,94	ab	2,63	de	2,46	cde	2,40	cd
	0,83		0,94		1,63		1,46		1,40	
3	1,99	abc	1,89	ab	2,45	bcde	2,46	bcde	2,89	e
	0,99		0,89		1,45		1,46		1,89	
4	1,85	a	1,90	ab	2,58	abcde	2,54	de	2,52	de
	0,85		0,90		1,58		1,54		1,52	
B) Lecture des contrastes en ligne pour un même nombre de syllabes de N										
nombre de syllabes du N	G11		G42		G31		G32		G22	
	mean.class	.group								
	mean.class(-1)		mean.class(-1)		mean.class(-1)		mean.class(-1)		mean.class(-1)	
2	1,83	a	1,94	ab	2,63	de	2,46	cde	2,40	cd
	0,83		0,94		1,63		1,46		1,40	
3	1,99	abc	1,89	ab	2,45	bcde	2,46	bcde	2,89	e
	0,99		0,89		1,45		1,46		1,89	
4	1,85	a	1,90	ab	2,58	abcde	2,54	de	2,52	de
	0,85		0,90		1,58		1,54		1,52	

Par ailleurs, en regardant cette fois-ci les contrastes post hoc entre les différentes tailles de N dans chaque groupement (cf. Tableau (A) N°4), il semblerait que la taille de N n'influence pas de manière significative la force de la frontière au sein d'un même groupement. La seule variation significative est observée en G22, où la frontière après N2 est perçue plus forte lorsque N est composé de 3 syllabes plutôt que de 2 syllabes (scores moyens : 1,89 contre 1,40).

Ce résultat a également été observé dans les analyses descriptives précédentes, où le degré de frontière après N2 en G22 semble augmenter lorsque N est composé de 3 syllabes. Nous avons alors supposé que le renforcement du niveau de frontière pouvait être lié à l'augmentation du nombre de syllabes du constituant. Toutefois, les résultats ne montrent pas d'effet similaire pour les N composés de 4 syllabes. Il semblerait dans ce contexte que seuls les N de 3 syllabes impactent la perception des frontières, mais là encore uniquement en G22 (cf. Figure N°52).



**Figure 52 :** Illustration de l'effet d'interaction entre le groupement et le nombre de syllabes de N sur la force de la frontière perçue après N2.

Les points reliés par des lignes correspondent aux valeurs « mean.class (-1) » du tableau N°4 indiquant les moyennes des scores de frontières perçues pour chaque taille du N, interprétées comme des valeurs numériques ordonnées de 0 à 3.

Au vu de l'ensemble de ces résultats, les scores de frontières après N2 varieraient davantage en fonction du type de groupement qu'en fonction de la taille du constituant. La granularité des niveaux de frontière après N2 apparaît cependant moins variable que celle observée après N1. Deux niveaux de frontière seraient en effet perçus : un premier niveau avec le score de frontière le plus fort en G22 et un second niveau avec le score de frontière le plus faible en G11 et G42. Le niveau de frontière perçue en G31 et G32 se situerait, en revanche, à l'intermédiaire entre les deux autres niveaux de frontière, en G22 d'un côté et en G11 et G42 de l'autre. Toutefois, alors que les frontières en G31 et G32 peuvent parfois, selon la « configuration syllabique », être perçues plus fortes que les frontières en G11 et G42, aucune différence n'est cependant observée entre G31, G32 et G22. Dans ce contexte, la frontière perçue en G31 et G32 pourrait être considérée comme une frontière de même niveau que la frontière perçue en G22.

En mettant en relation ces résultats avec nos prédictions sur les niveaux de constituance perçus dans chaque groupement, nous constatons toutefois peu de correspondance. Alors que nous supposons au départ trois différents degrés de frontière après N2, seuls deux niveaux de constituance semblent avoir été perçus par les auditeurs.

### Rappel de nos prédictions :

**G11** [ les N1 ]<sub>ip</sub> // [ et les N2 <sub>PW</sub> / ADJ ] IP

**G31** [ les N1 ]<sub>ip</sub> // [ et les N2 ]<sub>AP</sub> / [ ADJ ] IP

**G32** [ les N1 ]<sub>AP</sub> / [ et les N2 ]<sub>ip</sub> // [ ADJ ] IP

**G42** [ les N1 <sub>PW</sub> / et les N2 <sub>AP</sub> // ADJ ] IP

**G22** [ les N1 <sub>PW</sub> / et les N2 ]<sub>ip</sub> // [ ADJ ] IP

En comparant les scores moyens des frontières perçues après N1 et après N2, il semblerait que les scores associés à la frontière perçue la plus forte après N2 en G22 (score moyen : 1,40 à 1,89) soient globalement plus élevés que le score associé à la frontière d'AP après N1 en G32 (score moyen : 1,1), mais tout de même moins élevé que les scores moyens associés à la frontière d'ip après N1 en G11 et G31 (scores moyens : 2,4 à 2,5). Nous pourrions alors supposer un niveau de frontière intermédiaire après N2 en G22, qui serait à la fois supérieur à la frontière perçue après N1 en G32, et à la fois inférieur à la frontière perçue après N1 en G11 et G31. Ainsi, en admettant que la frontière après N1 en G32 relève du niveau de l'AP, la frontière après N2 en G22, qui est supposée supérieure, pourrait alors relever du niveau de l'ip. Ce niveau serait également observé après N2 en G31 et G32, qui comme nous l'avons déjà mentionné correspondrait au même niveau que G22, avec cependant des degrés de frontière qui peuvent parfois être moins forts selon les « configurations syllabiques ».

**G31** [ les N1 ]<sub>ip</sub> // [ et les N2 ]<sub>AP</sub> ou <sub>ip</sub> ? / [ ADJ ] IP

**G32** [ les N1 ]<sub>AP</sub> / [ et les N2 ]<sub>ip</sub> // [ ADJ ] IP

**G22** [ les N1 <sub>PW</sub> / et les N2 ]<sub>ip</sub> // [ ADJ ] IP

La frontière après N1 en G11 et G31, dont les scores moyens sont nettement plus élevés que ceux associés à la frontière après N2 en G22, G32 et G31, pourrait quant à elle relever d'un niveau de frontière encore supérieur à l'ip, correspondant potentiellement au niveau de l'IP. Cette supposition permettrait également de mieux interpréter le résultat obtenu dans l'analyse de la force de frontière après N1, où nous observons une différence de force entre le niveau de frontière intermédiaire, en G32, et le niveau de frontière le plus faible, en G22 et G42, nettement moins importante que la différence de force entre le niveau de frontière le plus fort, en G11 et G31, et le niveau de frontière intermédiaire, en G32. Cette différence

plus marquée pourrait en effet être liée à la présence non pas d'une frontière d'ip après N1 en G11 et G31, mais plutôt d'une frontière d'IP, qui aurait alors pour effet de créer une différence plus importante avec les frontières d'AP et de PW sur le même site de frontière (après N1) dans les trois autres groupements.

**G11** [ les N1 ]<sub>ip</sub> ou IP ? // [ et les N2 PW / ADJ ] IP

**G31** [ les N1 ]<sub>ip</sub> ou IP ? // [ et les N2 ]<sub>AP</sub> / [ ADJ ] IP

Quant aux scores moyens associés cette fois-ci à la frontière perçue la plus faible après N2 en G11 et G42 (scores moyens : 0,83 à 0,99), nous constatons un score plus élevé que celui associé à la frontière de PW après N1 en G42 et G22 (scores moyens : 0,50 à 0,64), mais qui reste cependant relativement équivalent au score moyen associé à la frontière d'AP après N1 en G32 (score moyen : 1,09). Il se pourrait dans ce contexte que la frontière perçue après N2 en G11 et G42 relève du même niveau que la frontière perçue après N1 en G32, à savoir une frontière d'AP.

**G11** [ les N1 ]<sub>ip</sub> // [ et les N2 PW ou AP ? / ADJ ] IP

**G42** [ les N1 PW / et les N2 AP // ADJ ] IP

Toutefois, le choix de ces niveaux de constituance implique également la relation hiérarchique entre les frontières sur le plan syntagmatique. Nous rappelons en effet que nous supposons pour chaque catégorie de groupement une hiérarchie spécifique entre les deux frontières, après N1 et N2, qui impacterait par ailleurs le choix des niveaux de constituance sur chaque site. Nous prédisons en effet des niveaux de constituance en se référant notamment aux scores moyens associés à chaque groupement, mais il est cependant nécessaire de vérifier que la différence entre ces scores moyens sur le plan syntagmatique est réellement significative. En G32, par exemple, nous supposons la présence d'une frontière d'AP après N1 et d'ip après N2, sur la base d'une comparaison des scores moyens associés aux deux frontières, après N1 et N2, qui semblent montrer un score plus élevé après N2 qu'après N1. Il est toutefois nécessaire de vérifier sur le plan syntagmatique si la force entre les deux frontières est significativement différente. La même problématique se retrouve également dans tous les autres groupements. Nous avons donc besoin de recourir à une analyse de la force relative des frontières, testant la relation

hiérarchique entre les frontières non plus sur le plan paradigmatique mais plutôt sur le plan syntagmatique, pour ainsi compléter notre interprétation du phrasé prosodique.

### 1.4.3 Force relative entre les deux frontières perçues (N2|ADJ – N1|N2)

Les analyses précédentes nous renseignent en effet sur la flexibilité du degré de frontière perçue au sein d'un même site, mais ne nous permettent pas de rendre compte du rapport hiérarchique entre les frontières des deux sites à la fois. Dans ce cadre, nous proposons une analyse syntagmatique de la force relative des frontières, en soustrayant le score de frontière perçue après N1 à celui perçue après N2 (N2|ADJ – N1|N2). Nous obtenons ainsi une force relative indiquant la manière dont la frontière après N1 est perçue par rapport à la frontière après N2. Une force relative négative signifie que la frontière après N1 est perçue comme plus forte que la frontière après N2, tandis qu'une force relative positive montre que la frontière après N1 est perçue comme plus faible que la frontière après ADJ<sup>13</sup>.

Les résultats montrent un effet significatif du groupement (LR Chisq= 180,942, df= 4,  $p < 0,001$ ) et du nombre de syllabes de N (LR Chisq= 9,552, df= 2,  $p = 0,008$ ), ainsi qu'un effet d'interaction entre le groupement et le nombre de syllabes de N (LR Chisq= 16,689, df= 8,  $p = 0,03$ ) (cf. Annexe N°9-C). Toutefois, bien qu'un effet simple du nombre de syllabes de N soit observé, les analyses post hoc ne présentent aucun contraste significatif entre les trois tailles de N. Ce résultat pourrait être dû à l'implication du nombre de syllabes de N dans son interaction significative avec le groupement.

#### Effet simple du groupement :

Les analyses post hoc de l'effet simple du groupement montrent que la force relative des frontières varie entre tous les groupements, excepté entre G42 et G32 qui présentent une force relative des frontières équivalente (cf. Annexe N°9-C). Plus précisément, la force relative en G11 et G31 est significativement plus faible que la force relative en G22, G32 et G42. Deux profils de force relative seraient perçus. Étant donné la valeur négative de la force relative en G11 et G31, cela suppose que la frontière après N1 soit perçue plus forte que la frontière après N2. À l'inverse, étant donné la valeur positive de la force relative en G22, G32 et G42, cela suppose que la frontière après N1 soit perçue moins forte que la frontière après N2 en G22, G32 et G42. Cependant, les comparaisons post hoc montrent également des distinctions entre les groupements présentant le même profil de force

---

<sup>13</sup> Pour plus de détails sur la construction des variables de force relative, nous renvoyons le lecteur à la partie consacrée à la préparation de notre base de données (cf. 3.3 du Chapitre 4).

relative. En effet, la force relative entre les frontières apparaît encore plus faible en G11 qu'en G31 (scores moyens respectifs : -1,52 contre -0,97), et encore plus forte en G22 qu'en G32 et G42 (scores moyens respectifs : 0,93 contre 0,39 et 0,40). La différence entre les deux frontières pourrait alors être plus ou moins marquée dans tel ou tel groupement.

### Effet d'interaction entre le groupement et le nombre de syllabes de N :

Le résultat de l'effet simple du groupement peut toutefois être modulé par l'effet d'interaction entre le groupement et le nombre de syllabes de N, qui permet de tester à quel point la relation entre le score de la force relative des frontières perçues après N1 et N2 et le groupement varie en fonction de la taille du constituant. En regardant tout d'abord les contrastes post hoc entre les différents groupements pour un même nombre de syllabes de N (cf. Tableau (B) N°5), nous constatons toujours l'opposition principale entre G11 et G31 d'un côté et G32, G42 et G22 de l'autre. Les deux profils de force relative observés dans l'effet simple du groupement semblent donc maintenus, quelle que soit la taille du constituant. En revanche, concernant les distinctions entre les groupements présentant le même profil de force relative, les oppositions initialement observées dans l'effet simple du groupement sont relativement nuancées. G11 et G31 présentent effectivement des différences significatives uniquement dans le cas du N composé de 2 syllabes, pour lequel la force relative des frontières est plus élevée en G31 (score moyen : -0,82) qu'en G11 (score moyen : -1,53). Quant à G42, G32 et G22, des différences apparaissent seulement lorsque N est composé de 3 syllabes, où G22 (score moyen : 1,33) montre une force relative plus importante qu'en G42 et G32 (score moyen : entre 0,34 et 0,43).

Ces résultats semblent donc correspondre à nos hypothèses de départ qui prédisent en G11 et G31 un niveau de frontière après N1 (frontières d'ip) supérieur au niveau de frontière après N2 (frontières d'AP et de PW), et à l'inverse en G22, G32 et G42, un niveau de frontière après N1 (frontières d'AP et de PW) inférieur au niveau de frontière après N2 (frontières d'ip et d'AP). Les différences entre les groupements présentant le même profil de force relative des frontières sont également envisagées dans nos hypothèses. En G11 et G31, tout d'abord, nous attendons une même frontière d'ip après N1 dans les deux groupements, et deux niveaux de frontière différents après N2. G11 serait marqué par une frontière de PW, tandis que G31 serait associé à une frontière d'AP, ce qui impliquerait effectivement une différence entre les deux frontières après N1 et N2 plus marquée en G11 qu'en G31. La différence entre une frontière d'ip et de PW est effectivement plus grande que la différence entre une frontière d'ip et d'AP. Nous supposons également cet effet en G22, qui est marqué par une frontière de PW après N1 et par une frontière d'ip après N2. G32 et G42, en revanche, ne sont pas réalisés avec une différence entre les deux frontières aussi

importante, d'où la possibilité de percevoir une différence moins marquée qu'en G22. Nous rappelons toutefois que cet effet ne semble pas systématique, dans la mesure où il intervient dans le cas de « configurations syllabiques » spécifiques.

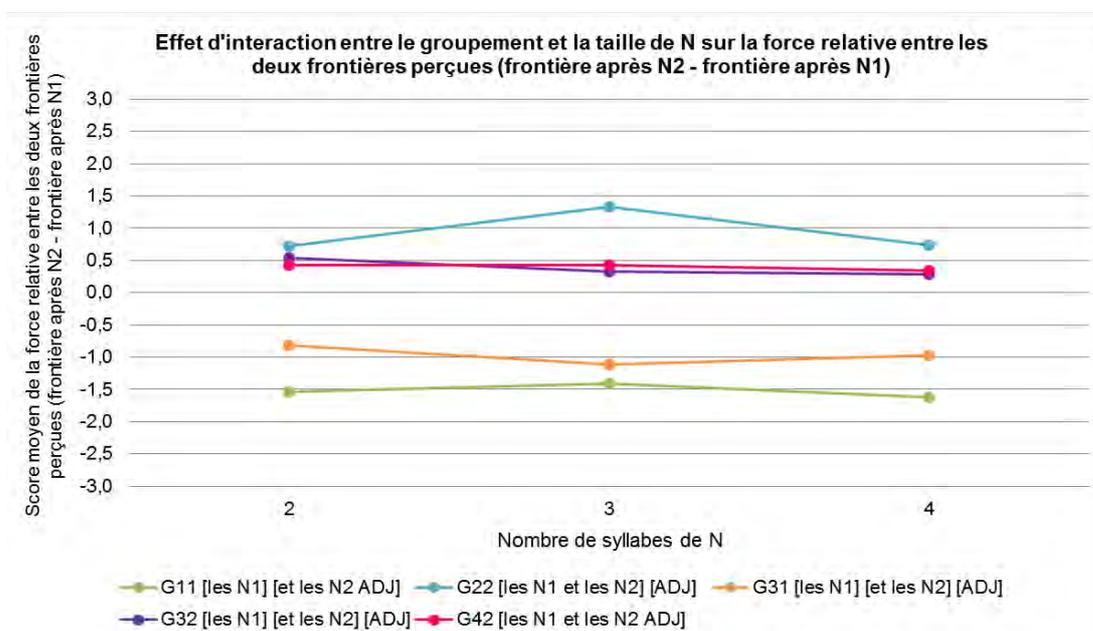
**Tableau 5 : Résultats des comparaisons post hoc de l'effet d'interaction entre le groupement et le nombre de syllabes de N sur la force relative entre les deux frontières perçues (frontière après N2 – frontière après N1)**

A) Lecture des contrastes en colonne pour une même catégorie de groupement et B) Lecture des contrastes en ligne pour un même nombre de syllabes de N. Dans chaque tableau, les moyennes des forces relatives des frontières sont interprétées comme des valeurs numériques ordonnées de 1 à 7 pour les valeurs « mean.class » et de -3 à 3 pour les valeurs « mean.class(-4) », qui sont calculées à partir des résultats de la distribution de probabilités estimée pour chaque catégorie de groupement dans chaque taille de N (cf. Annexe N°9-C). Les groupes partageant une même lettre dans la colonne ".group" ne présentent pas de différence significative.

FORCE RELATIVE ENTRE LES DEUX FRONTIÈRES PERÇUES : FRONTIÈRE APRES N2 - FRONTIÈRE APRES N1										
Moyennes des scores de force relative des frontières calculées à partir des résultats de la distribution de probabilités										
A) Lecture des contrastes en colonne pour une même catégorie de groupement										
Nombre de syllabes du N	G11		G31		G32		G42		G22	
	mean.class	.group								
	mean.class(-4)		mean.class(-4)		mean.class(-4)		mean.class(-4)		mean.class(-4)	
2	2,47	a	3,18	b	4,54	c	4,43	c	4,72	c
	-1,53		-0,82		0,54		0,43		0,72	
3	2,60	ab	2,89	ab	4,34	c	4,43	c	5,33	d
	-1,40		-1,11		0,34		0,43		1,33	
4	2,38	a	3,02	ab	4,29	c	4,35	c	4,74	c
	-1,62		-0,98		0,29		0,35		0,74	
B) Lecture des contrastes en ligne pour un même nombre de syllabes de N										
Nombre de syllabes du N	G11		G31		G32		G42		G22	
	mean.class	.group								
	mean.class(-4)		mean.class(-4)		mean.class(-4)		mean.class(-4)		mean.class(-4)	
2	2,47	a	3,18	b	4,54	c	4,43	c	4,72	c
	-1,53		-0,82		0,54		0,43		0,72	
3	2,60	ab	2,89	ab	4,34	c	4,43	c	5,33	d
	-1,40		-1,11		0,34		0,43		1,33	
4	2,38	a	3,02	ab	4,29	c	4,35	c	4,74	c
	-1,62		-0,98		0,29		0,35		0,74	

Par ailleurs, en regardant cette fois-ci les contrastes post hoc entre les différentes tailles de N et pour une même catégorie de groupement (cf. Tableau (A) N°5), il semblerait que la taille de N n'influence pas de manière significative la force relative entre les deux frontières perçues (N2|ADJ – N1|N2) au sein d'un même groupement : la seule variation significative observée est en G22 où la force relative est plus élevée lorsque N est composé de 3 syllabes (score moyen : 1,33) plutôt que de 2 et 4 syllabes (score moyen : entre 0,72 et 0,74). Cela signifierait alors qu'en G22, la frontière après N2 serait perçue d'autant plus forte que la frontière après N1 lorsque N est composé de 3 syllabes plutôt que de 2 et 4 syllabes.

Il est intéressant de remarquer que ce résultat concernant l'effet du N de 3 syllabes a déjà été observé lors des analyses sur la perception du score de frontière après N2. Comme nous l'avons déjà mentionné, nous supposons au départ un renforcement du niveau de frontière lié à l'augmentation du nombre de syllabes du constituant. Toutefois, nos résultats ne montrent pas de renforcement perceptif pour les N composés de 4 syllabes. Il semblerait donc que seuls les N de 3 syllabes impactent la perception des frontières, mais là encore uniquement en G22 (cf. Figure N°53).



**Figure 53** : Illustration de l'effet d'interaction entre le groupement et le nombre de syllabes de N sur la force relative entre les deux frontières perçues (frontière après N2 – frontière après N1)  
 Les points reliés par des lignes correspondent aux valeurs « mean.class (-4) » du tableau N°5 indiquant les moyennes de force relative des frontières perçues pour chaque taille du N, interprétées comme des valeurs numériques ordonnées de -3 à 3

Au vu de l'ensemble des résultats, il semblerait que les auditeurs aient perçu deux profils de force relative des frontières différents : un premier profil où la frontière après N1 est perçue plus forte que la frontière après N2, en G11 et G31, et un second profil où, à l'inverse, la frontière après N2 est perçue plus forte que la frontière après N1, en G42, G32 et G22. Des oppositions sont également observées au sein d'un même profil, entre G11 et G31, mais aussi entre G32, G42 et G22. Cela signifierait que dans certains cas de « configurations syllabiques », la différence entre les deux frontières serait plus ou moins marquée dans tel ou tel groupement : la différence entre les deux frontières serait plus marquée en G11 qu'en

G31, et plus marquée en G22 qu'en G32 et G42. Il est par ailleurs intéressant de remarquer que cette différence plus importante entre les deux frontières est perçue en G11 et G22, qui correspondent aux groupements « canoniques » représentant par défaut les deux conditions syntaxiques. Nous rappelons en effet que G1 (G11) n'a quasiment pas été associé à la Condition 2 et que G2 (G22) n'a pas été perçu pour la Condition 1. Il se pourrait donc que les auditeurs aient choisi massivement ces deux groupements pour représenter chacune des deux conditions syntaxiques, étant donné l'ambiguïté peut-être moins grande quant à la hiérarchie établie entre les deux frontières en G11 et G22, par rapport aux autres catégories de groupements.

En mettant en relation ces résultats avec nos hypothèses sur les niveaux de constituance prédits, mais également avec les résultats obtenus sur les deux sites de frontières isolés, nous pouvons dès lors affiner notre interprétation des phrasés prosodiques perçus. Nous rappelons que la frontière après N1, tout d'abord, est réalisée par trois niveaux de frontière différents qui pourraient correspondre aux trois niveaux de constituance prédits : une frontière d'ip en G11 et G31, une frontière d'AP en G32 et une frontière de PW en G22 et G42. La frontière après N2, en revanche, montre que seuls deux niveaux de frontière ont été perçus, tandis que trois niveaux de constituance différents étaient initialement prédits.

<b>G11</b>	[ les N1 ] <b>ip</b> ou <b>IP</b> ?	//	[ et les N2 <b>PW</b> ou <b>AP</b> ?	/	ADJ ] <b>IP</b>
<b>G31</b>	[ les N1 ] <b>ip</b> ou <b>IP</b> ?	//	[ et les N2 ] <b>AP</b> ou <b>ip</b> ?	/	[ ADJ ] <b>IP</b>
<b>G32</b>	[ les N1 ] <b>AP</b>	/	[ et les N2 ] <b>ip</b>	//	[ ADJ ] <b>IP</b>
<b>G42</b>	[ les N1 <b>PW</b>	/	et les N2 <b>AP</b>	//	ADJ ] <b>IP</b>
<b>G22</b>	[ les N1 <b>PW</b>	/	et les N2 ] <b>ip</b>	//	[ ADJ ] <b>IP</b>

En comparant les scores moyens des frontières perçues après N1 et après N2, nous avons pu émettre quelques hypothèses quant à la répartition des niveaux de constituance. Nous avons besoin toutefois des résultats de l'analyse de la force relative pour mieux interpréter ces différents niveaux. Effectivement, en sachant que la frontière après N2 est perçue plus forte que la frontière après N1 en G22, G32 et G42, cela suppose alors que le niveau de frontière après N2 en G32 et G22 relève bien du niveau de l'ip et non de l'AP, pour ainsi respecter la supériorité hiérarchique par rapport à la frontière d'AP après N1 en G32, et par rapport à la frontière de PW après N1 en G22. De la même manière, le niveau de frontière après N2 en G42 relèverait bien du niveau de l'AP et non du PW, pour ainsi respecter la hiérarchie entre les deux frontières. Par ailleurs, étant donné que la frontière après N2 en

G31 ne se distingue pas significativement de la frontière après N2 en G32 et G22, cela suppose alors que la frontière après N2 en G31 relève également du niveau de l'ip et non de l'AP comme prédit initialement. Dans le même sens, la frontière après N2 en G42 ne se différencie pas significativement de la frontière après N2 en G11, ce qui tend alors à supposer que la frontière après N2 en G11 relèverait également du niveau de l'AP et non du niveau du PW comme prédit initialement.

**G11** [ les N1 ]<sub>ip</sub> ou IP ? // [ et les N2 ]<sub>AP</sub> / ADJ ] IP

**G31** [ les N1 ]<sub>ip</sub> ou IP ? // [ et les N2 ]<sub>ip</sub> / [ ADJ ] IP

**G32** [ les N1 ]<sub>AP</sub> / [ et les N2 ]<sub>ip</sub> // [ ADJ ] IP

**G42** [ les N1 ]<sub>PW</sub> / et les N2 ]<sub>AP</sub> // ADJ ] IP

**G22** [ les N1 ]<sub>PW</sub> / et les N2 ]<sub>ip</sub> // [ ADJ ] IP

Enfin, nous avons émis une hypothèse quant à la possibilité d'avoir une frontière d'IP plutôt que d'ip après N1 en G11 et G31. L'analyse de la force relative nous permet également de trancher en faveur d'une frontière d'IP, étant donné le rapport hiérarchique entre les deux frontières sur le plan syntagmatique. Nous rappelons en effet qu'en G11 et G31, la frontière après N1 est perçue plus forte que la frontière après N2. Il serait alors nécessaire que la frontière après N1 en G31, notamment, soit associée à un niveau de frontière supérieur à l'ip, donc potentiellement équivalent à l'IP. Cette organisation des niveaux de constituance n'impacte pas par ailleurs l'effet sur les différences plus ou moins marquées entre les groupements présentant le même profil de force relative. La différence entre les deux frontières est toujours plus importante en G11 qu'en G31, et en G22 qu'en G32 et G42.

**G11** [ les N1 ]<sub>IP</sub> // [ et les N2 ]<sub>AP</sub> / ADJ ] IP

**G31** [ les N1 ]<sub>IP</sub> // [ et les N2 ]<sub>ip</sub> / [ ADJ ] IP

**G32** [ les N1 ]<sub>AP</sub> / [ et les N2 ]<sub>ip</sub> // [ ADJ ] IP

**G42** [ les N1 ]<sub>PW</sub> / et les N2 ]<sub>AP</sub> // ADJ ] IP

**G22** [ les N1 ]<sub>PW</sub> / et les N2 ]<sub>ip</sub> // [ ADJ ] IP

Nos résultats sur la perception des scores de frontières mettent donc en évidence une granularité des niveaux de frontière plus fine que ce nous prédisons initialement, incluant

non pas trois niveaux mais plutôt quatre niveaux de constituance différents (IP, ip, AP et PW).

### 1.5 Discussion

Les scores de frontières perçues après N1 et N2 sont sans nul doute nettement plus influencés par le groupement que par la taille des constituants. Alors que nous envisagions dans nos hypothèses de départ la possibilité que l'augmentation du nombre de syllabes d'un constituant renforce la perception de la frontière, nos résultats montrent que la force de la frontière augmente uniquement dans le cas d'un nombre de syllabes spécifique. Seuls les N de 3 syllabes semblent effectivement montrer un renforcement du degré de frontière. Ce résultat pourrait être lié à des contraintes d'ordre rythmique, impliquant un renforcement du degré de frontière pour signaler la fin d'un groupe prosodique. En effet, il semblerait selon Delais-Roussarie (1995) que les groupes de 3 syllabes représente la taille « canonique » des groupes prosodiques, et que les constituants composés de plus de 3 syllabes soient plutôt rares. Il apparaît toutefois que cet effet ne puisse être généralisé à tous les niveaux de frontières, puisque seule la frontière perçue après N2 en G22, qui renvoie potentiellement à une frontière d'ip, présente ce résultat. D'autres frontières du même niveau ne sont en revanche aucunement impliquées. Nous proposons donc dans le cadre de cette discussion de nous intéresser plus spécifiquement aux résultats de l'effet du groupement sur la force des frontières, que nous proposons de résumer dans un tableau récapitulatif (cf. Tableau N°6).

Pour constituer ce tableau, nous avons tout d'abord repris les scores moyens (selon les cas, les valeurs « mean.class (-1) » ou « mean.class (-4) ») associés à chaque groupement sur tous les sites d'analyse. Alors que dans l'analyse de la frontière perçue après N1, chaque groupement montre un seul score moyen, dans le cas des effets d'interaction observés pour la frontière après N2, et pour la force relative entre les deux frontières, en revanche, plusieurs scores moyens peuvent être associés à un même groupement. Dans ce contexte, nous avons choisi de retenir la médiane de l'ensemble de ces valeurs. Nous avons ensuite transformé les mesures chiffrées, associées aux frontières perçues après N1 et N2, en une barre verticale représentant la frontière, dont l'épaisseur de trait (mesuré en point ou « pt ») varie proportionnellement en fonction de la valeur chiffrée (ajout de la constante \*4 à chaque valeur : 2,4 = 9,6 pt ; 2,5 = 10 pt ; 1,1 = 4,4 etc.).

**Tableau 6 :** Récapitulatif des résultats de l'effet du groupement sur la force des frontières perçues après N1, après N2 et sur la force relative entre les deux frontières perçues (frontière après N2 – frontière après N1)

A) Valeurs « mean.class (-1) » ou « mean.class (-4) » des scores de frontières perçus sur chaque site d'analyse associées aux différents groupements. Dans le cas des effets d'interaction c'est la médiane des valeurs « mean.class (1) » ou « mean.class (4) » qui a été retenue. B) Transformation des valeurs « mean.class » en barres verticales dont l'épaisseur de trait varie en fonction de la valeur chiffrée. Les couleurs ou les nuances de couleurs permettent d'identifier les groupements ayant le même niveau de frontière.

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

Au vu du tableau récapitulatif, il semblerait alors que pour rendre compte de la granularité des niveaux de frontière perçus par nos auditeurs, quatre différentes catégories d'épaisseurs de trait soient nécessaires. Trois niveaux de frontière sont en effet perçus après N1 : la frontière la plus forte en G11 et G31 (rouge), intermédiaire en G32 (jaune) et la plus faible en G42 et G22 (bleu). Nous avons également remarqué que la différence entre G11 et G31 d'un côté et G32 de l'autre, apparaît nettement plus importante que la différence entre G42 et G22 d'un côté et G32 de l'autre. Les résultats sur la perception de la frontière après N2, en revanche, montrent que la force des frontières perçues après N2 est beaucoup moins variable entre les différents groupements : seul G22 s'oppose systématiquement à G11 et G42, tandis que G31 et G32 se différencient de G11 et G42 uniquement dans le cas de « configurations syllabiques » spécifiques. Ces résultats sont illustrés dans le tableau, où trois degrés d'épaisseur de trait s'opposent de manière significative uniquement entre G22 (vert) d'un côté et G11 et G42 (jaune pâle) de l'autre. La frontière perçue en G31 et G32, quant à elle, n'est pas illustrée comme un troisième niveau de frontière à part entière, mais

plutôt comme une nuance de force de frontière se situant entre les deux autres degrés de frontière. Toutefois, la frontière en G31 et G32 reste plus proche du niveau de frontière perçu en G22 qu'en G11 et G42, d'où la même nuance de couleur en G31, G32 (vert pâle) et G22 (vert). Nous avons tenu par ailleurs à rendre compte dans le tableau de la similarité entre le niveau de frontière perçu après N1 en G32 (jaune) et celui perçu après N2 en G11 et G42 (jaune pâle). Au vu des scores moyens associés à ces trois groupements (scores moyens : 0,9 à 1,1), nous supposons en effet qu'il s'agit ici d'un seul et même niveau de frontière. Enfin, des différences d'épaisseur entre les barres verticales sur le plan syntagmatique cette fois-ci, permettent de rendre compte des résultats de l'analyse de la force relative entre les deux frontières perçues (N2|ADJ – N1|N2). Cette dernière analyse permet effectivement d'indiquer la hiérarchie établie entre ces deux frontières en fonction du groupement. Les résultats montrent que deux profils de force relative sont perçus : un premier profil où la frontière après N1 est perçue plus forte que la frontière après N2, en G11 et G31, et un second profil où, à l'inverse, la frontière après N2 est perçue plus forte que la frontière après N1, en G22, G32 et G42. Ces deux profils sont donc illustrés dans notre tableau : en G11 et G31 la barre verticale est plus épaisse après N1 qu'après N2, alors qu'en G32, G42 et G22, la barre est plus épaisse après N2 qu'après N1. De la même manière, comme nous l'avons observé dans nos résultats, pour certains cas de « configurations syllabiques », la différence entre les deux frontières est plus ou moins marquée dans tel ou tel groupement, faisant néanmoins partie du même profil de force relative des frontières : la différence entre les deux frontières est plus marquée en G11 qu'en G31, mais également plus marquée en G22 qu'en G32 et G42. Ce constat est par ailleurs observé dans le tableau, où la différence entre les deux degrés d'épaisseur de trait est plus importante en G22 qu'en G32 et G42, et également plus importante en G11 qu'en G31.

En mettant en relation ces résultats avec nos prédictions sur les niveaux de constituance perçus, nous pouvons dès lors supposer la présence de quatre niveaux différents, au lieu des trois niveaux prédits initialement. Ce sont plus précisément les frontières prédites en G11 et G31 (signalés par les flèches) qui ne semblent pas tout à fait correspondre aux résultats perceptifs.

<b>G11</b>	[ les N1 ] <sub>ip</sub> ?	//	[ et les N2 ] <sub>PW</sub> ?	/	ADJ ] IP	←
<b>G31</b>	[ les N1 ] <sub>ip</sub> ?	//	[ et les N2 ] <sub>AP</sub> ?	/	[ ADJ ] IP	←
<b>G32</b>	[ les N1 ] <sub>AP</sub>	/	[ et les N2 ] <sub>ip</sub>	//	[ ADJ ] IP	
<b>G42</b>	[ les N1 ] <sub>PW</sub>	/	et les N2 ] <sub>AP</sub>	//	ADJ ] IP	
<b>G22</b>	[ les N1 ] <sub>PW</sub>	/	et les N2 ] <sub>ip</sub>	//	[ ADJ ] IP	

Les résultats montrent effectivement que la frontière après N2 en G31 ne se différencierait pas significativement de la frontière après N2 en G32 et G22. De la même manière, la frontière après N2 en G11 ne se distinguerait pas significativement de la frontière après N2 en G42. Au vu des scores moyens associés à chacun de ces groupements, tout en respectant par ailleurs la hiérarchie entre les deux frontières sur le plan syntagmatique, nous avons alors supposé que la frontière après N2 en G31 relèverait d'un niveau supérieur à l'AP, correspondant davantage au niveau de l'ip, tandis que la frontière après N2 en G11 pourrait être associée à une frontière d'AP.

**G11** [ les N1 ]<sub>ip</sub> ? // [ et les N2 <sub>AP</sub> / ADJ ] <sub>IP</sub> ←  
**G31** [ les N1 ]<sub>ip</sub> ? // [ et les N2 ]<sub>ip</sub> / [ ADJ ] <sub>IP</sub> ←  
**G32** [ les N1 ]<sub>AP</sub> / [ et les N2 ]<sub>ip</sub> // [ ADJ ] <sub>IP</sub>  
**G42** [ les N1 <sub>PW</sub> / et les N2 <sub>AP</sub> // ADJ ] <sub>IP</sub>  
**G22** [ les N1 <sub>PW</sub> / et les N2 ]<sub>ip</sub> // [ ADJ ] <sub>IP</sub>

Enfin, les frontières d'ip prédites après N1 en G11 et G31 semblent renvoyer, au vu de nos résultats, à un niveau de frontière encore supérieur. Nous rappelons en effet que les résultats sur la force des frontières après N1 montrent une différence de force entre le niveau de frontière intermédiaire, en G32, et le niveau de frontière le plus faible, en G22 et G42, nettement moins importante que la différence de force entre le niveau de frontière le plus fort, en G11 et G31, et le niveau de frontière intermédiaire, en G32. Cette différence plus marquée pourrait en effet être liée à la présence non pas d'une frontière d'ip, comme prédit initialement après N1 en G11 et G31, mais plutôt d'une frontière d'IP, qui aurait alors pour effet de créer une différence plus importante avec les frontières d'AP et de PW sur le même site de frontière (après N1) dans les trois autres groupements. Par ailleurs, les résultats de l'analyse de la force relative des frontières montre que la frontière après N1 est perçue plus forte que la frontière après N2 en G11 et G31. La frontière d'IP après N1 en G31, notamment, serait alors nécessaire pour respecter la relation hiérarchique entre les deux frontières.

**G11** [ les N1 ]<sub>IP</sub> // [ et les N2 <sub>AP</sub> / ADJ ] <sub>IP</sub> ←  
**G31** [ les N1 ]<sub>IP</sub> // [ et les N2 ]<sub>ip</sub> / [ ADJ ] <sub>IP</sub> ←  
**G32** [ les N1 ]<sub>AP</sub> / [ et les N2 ]<sub>ip</sub> // [ ADJ ] <sub>IP</sub>  
**G42** [ les N1 <sub>PW</sub> / et les N2 <sub>AP</sub> // ADJ ] <sub>IP</sub>

G22 [ les N1 PW / et les N2 ]<sub>ip</sub> // [ ADJ ] IP

Ainsi, pour mieux rendre compte des événements perçus par les auditeurs de notre expérimentation, il semblerait nécessaire d'envisager une hiérarchie des constituants prosodiques composée de 4 unités phonologiques différentes : IP, ip, AP et PW.

A travers ces résultats, nous mettons par ailleurs l'accent sur deux phénomènes majeurs, qui font encore aujourd'hui l'objet de débats dans la phonologie prosodique du français. Il s'agit plus précisément des questions autour de la pertinence du niveau du Syntagme Intermédiaire et de l'utilité du niveau du Mot Prosodique, en tant que mot lexical, dans la hiérarchie des constituants prosodiques.

En effet, le français étant traditionnellement considéré comme une langue à accentuation post-lexicale, le domaine accentuel minimal dans la langue serait par conséquent le groupe de mots et non le mot lui-même. Dans ce contexte, bien que la plupart des auteurs considèrent le mot comme une unité accentogène abstraite (notamment Delais-Roussarie, 2000 ; Post, 2000 ; Di Cristo ; 2000 ; 2004), aucun des modèles du phrasé prosodique traditionnels ou plus actuels, n'envisagent cependant de prendre en compte le mot comme une unité de planification, ou de réalisation des règles accentuelles, en structure de surface. De fait, le Mot Prosodique, en tant que mot lexical, n'est jamais inclus dans la hiérarchie des constituants prosodiques. Seules des études en production (Astésano et *al.*, 2007) et en perception (Astésano et *al.*, 2012 ; 2013) tentent de rendre compte du niveau du mot en français, à travers notamment la description de l'Accent Initial. Nous rappelons en effet que les auteurs montrent que l'AI tend à marquer de manière plus systématique qu'AF les niveaux de constituance inférieurs dans la hiérarchie (Astésano et *al.*, 2007). Leurs résultats indiquent par ailleurs que les syllabes marquées par un AF sont perçues comme métriquement plus fortes que les syllabes inaccentuées à l'intérieur de l'AP (donc en frontière de Mot Lexical), ce qui semble confirmer également le phénomène de « survie » de l'accent au niveau du mot (Astésano et *al.*, 2012 ; 2013). D'autres travaux en neuro-imagerie (Aguilera et *al.*, 2014 ; te Rietmolen et *al.*, 2016 ; 2017) semblent également mettre en évidence un patron accentuel spécifique au niveau du mot lexical, qui serait marqué par une prééminence initiale et finale. Sur la base de l'ensemble de ces résultats, Astésano (2017) propose donc de prendre en compte le niveau du mot lexical, ou Mot Prosodique, comme l'unité métrique de base en français. Toutefois, en proposant une analyse sur le plan paradigmatique, uniquement, ces travaux ne permettent pas de démontrer l'intérêt du niveau de mot dans l'organisation même du phrasé prosodique. Ainsi, en proposant une analyse statistique des relations syntagmatiques des niveaux de frontière perçus, nos résultats

permettraient de démontrer définitivement la pertinence du niveau du PW dans la hiérarchie des constituants prosodiques. Les auditeurs de notre expérimentation semblent en effet utiliser le niveau du mot pour marquer un niveau de frontière inférieur au Syntagme Accentuel, pourtant considéré dans la majorité des modèles du phrasé comme l'unité phonologique de base du français (notamment Jun & Fougeron, 2000). Le Mot Prosodique, en tant que mot lexical, aurait donc ici une utilité dans la segmentation du flux de parole et dans l'organisation même du phrasé prosodique. Comme nous l'avons déjà mentionné, G22 et G42 ont été massivement choisis par les auditeurs pour correspondre à la Condition 2. Nous remarquons par ailleurs que ces deux groupements sont tous les deux marqués par une frontière de PW après N1. Il se pourrait donc dans ce contexte que le PW permette d'accéder plus facilement à la hiérarchie syntaxique entre les deux frontières, en créant peut-être moins d'ambiguïté que les autres niveaux de frontière. En outre, le PW serait nécessaire pour rendre compte de la finesse de la granularité des niveaux de frontière perçus en français.

**G42** [ les N1 **PW** / et les N2 **AP** // ADJ ] **IP**

**G22** [ les N1 **PW** / et les N2 **]ip** // [ ADJ ] **IP**

La pertinence du PW dans la hiérarchie des constituants prosodiques apparaît donc au cœur de nos résultats, et nécessite également la prise en compte des résultats sur la perception des proéminences, pouvant nous aider à mieux comprendre le marquage lexical en structure de surface. Il est en effet intéressant de voir comment les proéminences métriques, établies en structure sous-jacente, se projettent au niveau du mot en structure de surface. En référence notamment au marquage bipolaire des mots lexicaux en français (Astésano, 2017 en structure de surface ; Astésano & Bertrand, 2016 ; Di Cristo, 1999a ; 1999b 2000 en structure sous-jacente), il est intéressant de tester la manière dont les proéminences initiales et finales s'organisent les unes par rapport aux autres dès le niveau du Mot Prosodique, mais également à tous les niveaux de constituance.

Le second résultat majeur de notre étude concerne le niveau de l'ip, dont la pertinence en français est encore aujourd'hui mal établie. Quand certains auteurs envisagent l'ip comme une unité destinée à rendre compte de structures syntaxiques spécifiques (Segment d'Unité Intonative chez Di Cristo & Hirst, 1996 et Syntagme Intermédiaire chez Jun & Fougeron, 2000), d'autres considèrent que cette unité peut apparaître sur tout type d'énoncé, dès lors que les structures syntaxiques et prosodiques le permettent (Michelas & D'Imperio, 2010). Certains auteurs vont même jusqu'à remettre en doute la validité de cette

unité intermédiaire dans la langue, et parlent plutôt de récursivité de la structure prosodique, en considérant que l'ip n'est rien de plus qu'un IP contenu dans un autre IP (Di Cristo, 2016 ; Ladd, 1996 / 2008). Dans le cadre de notre étude, les résultats semblent montrer une réelle utilité du niveau de l'ip dans l'organisation du phrasé, en marquant notamment un niveau de frontière spécifique, à l'intermédiaire du Syntagme Accentuel et du Syntagme Intonatif. Nous n'envisageons pas par ailleurs l'ip comme un IP contenu dans un autre IP, dans la mesure où nous observons déjà dans nos résultats ce phénomène de récursivité en G11 et G31, qui implique selon nous un niveau de frontière encore différent de celui de l'ip.

**G11** [ les N1 ]**IP** // [ et les N2 **AP** / ADJ ] **IP**

**G31** [ les N1 ]**IP** // [ et les N2 ]**ip** / [ ADJ ] **IP**

Il sera intéressant de mettre ce résultat en lien avec l'analyse perceptive de la force des prééminences, mais également avec les corrélats acoustiques, pouvant rendre compte potentiellement d'une réalisation accentuelle et acoustique spécifiques au Syntagme Intermédiaire. Michelas & D'Imperio (2010) montrent notamment que l'ip est associé à un niveau de frontière marqué par un allongement de la durée de la voyelle finale, qui est à la fois significativement plus important que celui observé en frontière d'AP, et moins important que celui observé en frontière d'IP. Michelas (2011) montre également l'existence d'un phénomène de « *downstep* » entre les ip réalisés au sein d'une même IP, impliquant par ailleurs une réinitialisation partielle du registre tonal sur le premier ton « H » situé après la frontière droite de l'ip abaissé.

## 2. PERCEPTION DE LA FORCE DES PROÉMINENCES

Après avoir présenté les résultats sur la perception des scores de frontières, nous proposons maintenant de commenter les résultats associés à la perception de la force des prééminences. Nous nous intéressons plus précisément à la manière dont les prééminences ont été perçues en fonction du type de groupement choisi par les auditeurs, mais également en fonction de la taille du constituant. A l'instar des résultats sur la perception des frontières, nous proposons de présenter dans un premier temps les résultats des tests du Khi deux permettant de contrôler l'existence ou non d'une relation de dépendance entre les scores de prééminences et les groupements, puis entre les scores de prééminences et la taille du

constituant. Sur la base de ces résultats, seules les variables présentant un lien de dépendance seront ensuite incluses dans notre modèle de régression ordinale à effet mixte, dont les résultats seront présentés par la suite. L'ensemble des résultats seront enfin discutés dans le cadre d'une dernière partie.

### **2.1 Statistiques descriptives des principales variables**

Nous commençons par présenter la distribution des scores de proéminences par catégorie de groupement, pour ensuite commenter les résultats de la relation de dépendance entre les scores de proéminences et la taille du constituant. Il est important de signaler que tous les scores de proéminences sur toutes les syllabes du syntagme ont été testés dans le cadre de nos analyses. Nous tenons en effet à contrôler la relation de dépendance entre toutes les proéminences potentielles et le groupement ou la taille du constituant. Pour l'ensemble des tests du Khi deux présentés dans le cadre de cette partie, nous choisissons de commenter les histogrammes illustrant les différentes distributions. Nous renvoyons le lecteur aux Annexes N°10 et N°11, où les tableaux de contingence avec les effectifs et les pourcentages associés aux différents résultats sont présentés.

Il est par ailleurs intéressant de préciser que nous avons procédé à une analyse comparative de la force des proéminences, en utilisant le test de Stuart-Maxwell (Stuart, 1955 ; Maxwell, 1970) qui permet de comparer la concordance inter-observateurs pour deux variables qualitatives appariées, présentant chacune plus de deux modalités. Dans le cadre de notre travail, nous avons comparé deux à deux les distributions des scores de proéminences recueillis via une échelle d'évaluation (de 0 à 4) composée de 4 modalités différentes : 0 qui indique l'absence de proéminence, 1 qui renvoie au degré de proéminence le plus faible, 2 qui correspond à un degré de proéminence un peu plus élevé, et enfin 3 qui renvoie au degré de proéminence le plus élevé. Les résultats de cette analyse montrent que les proéminences sur les syllabes initiales et finales de tous les mots lexicaux sont systématiquement associées à des scores plus élevés que les syllabes intermédiaires des mots lexicaux ou que les syllabes des mots grammaticaux. Nous choisissons cependant de ne pas présenter dans le cadre de cette partie les résultats de cette analyse, qui selon nous seraient redondants par rapport aux résultats du modèle de régression ordinale à effet mixte. Il est par ailleurs plus intéressant de commenter les résultats du modèle mixte qui sont plus robustes que nos résultats descriptifs. Nous renvoyons toutefois le lecteur aux Annexes N°12 à N°15, où les tableaux de contingence issus des résultats du test de Stuart-Maxwell sont présentés.

### 2.1.1 Distribution des scores de proéminences par groupement

Les résultats du test du Khi deux montrent un lien de dépendance entre le groupement et la force des proéminences, mais uniquement avec les proéminences perçues sur la dernière syllabe de N1 (Khi deux= 34,480, ddl= 12,  $p=0,001$ ) et sur la première syllabe de ADJ (Khi deux= 36,208, ddl= 12,  $p<0,001$ ). Les scores de proéminences perçues sur les autres syllabes du syntagme ne montrent en revanche aucune correspondance avec le groupement ( $p>0,05$ ).

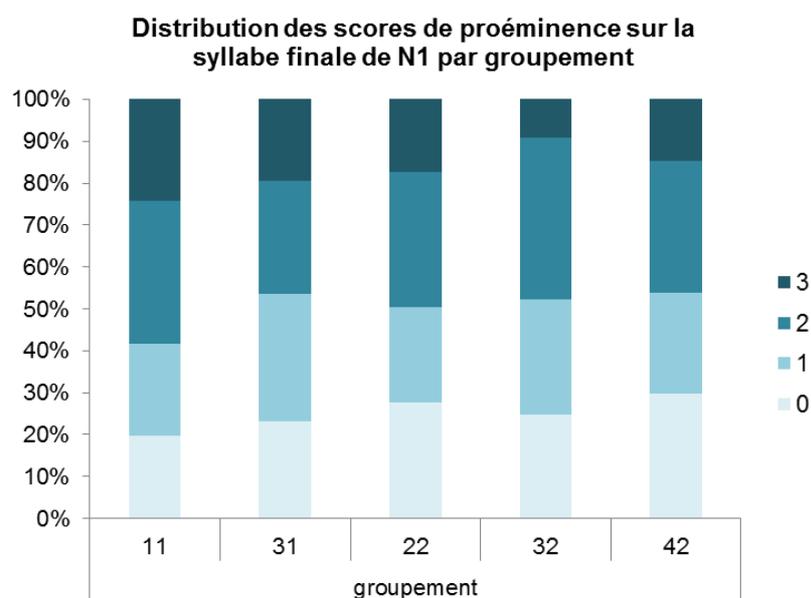
Alors que nous n'attendions pas de relation de dépendance avec les scores de proéminences sur les syllabes intermédiaires et les syllabes des mots grammaticaux, nous supposons cependant une relation de dépendance entre le groupement et les scores de proéminences sur la première et dernière syllabe de N2. Toutefois, les résultats tendent à montrer que la variation de la force des proéminences sur N2 ne dépendrait pas directement des différentes catégories de groupements. Cela pourrait alors supposer que le marquage des groupements, et par conséquent des différents niveaux de frontières, soit davantage lié à la proéminence finale sur N1 et à la proéminence initiale sur ADJ. Par ailleurs, nous pouvons supposer que les deux mots grammaticaux qui précèdent N2 impactent la relation de dépendance entre la proéminence initiale et le groupement. Selon Christophe et *al.* (1997), en effet, les auditeurs pourraient utiliser les mots grammaticaux comme indices pour savoir que la syllabe suivante serait un début potentiel de mot. Dans ce contexte, AI sur N2 ne covarierait pas avec le niveau de frontière, dans la mesure où les auditeurs s'attendent déjà à un début de mot. A l'inverse, en l'absence d'indices de ce type, comme sur ADJ, un renforcement de la proéminence initiale serait alors nécessaire pour signaler le début du mot, mais aussi pour marquer la structure prosodique.

Nous présentons donc dans un premier temps les résultats des scores de proéminences perçues sur la dernière syllabe de N1, et nous commenterons par la suite les résultats des scores de proéminences perçues sur la première syllabe de ADJ.

#### Proéminences sur la dernière syllabe de N1 (cf. Figure N°54) :

La distribution des scores sur la dernière syllabe de N1 reste relativement stable à travers les différents groupements (cf. Tableau de contingence en Annexe N°10-A). Seuls G11 et G32 semblent montrer un profil distributionnel un peu différent. En effet, G11 présente des degrés de proéminence plus élevés que dans les autres groupements, avec plus précisément une répartition du score de niveau 3 plus importante (24,2% en G11 contre 9,2% à 19,5%), et une répartition du score de niveau 0 moins importante (19,7% en G11 contre 23,2% à 29,8%) que dans les autres groupements. A l'inverse, des degrés de

proéminence un peu plus faibles sont observés en G32, et notamment une répartition du score de niveau 3 moins importante (9,2% en G32 contre 14,7% à 24,2%), et une répartition du score de niveau 2 plus importante (38,5% en G32 contre 26,8% à 34,2%) que dans les autres groupements. Peu de différence est cependant observée entre tous les groupements au niveau de la répartition des scores de niveau 0 et 1.



**Figure 54 :** *Distribution des scores de proéminences sur la syllabe finale de N1 par catégorie de groupement (G11, G31, G22, G32 et G42).*

Il semblerait donc que la proéminence sur la syllabe finale de N1 soit perçue avec un degré de proéminence un peu plus élevé en G11, un peu plus faible en G32 et avec un degré de proéminence intermédiaire en G31, G22 et G42. Ces différents degrés de proéminence doivent cependant être affinés dans le modèle de régression ordinaire à effet mixte, qui permet de mieux définir les contrastes entre les groupements à chaque niveau des scores de proéminences. Il est toutefois intéressant de mettre en lien ces résultats non seulement avec nos hypothèses de départ (*cf.* Hypothèse 2), mais également avec les résultats de l'analyse de la force des frontières. Nos hypothèses prédisent en effet un renforcement de la proéminence finale sur N1 en G11, G31 et G32, étant donné la présence d'une limite de groupement après la syllabe accentuée. G22 et G42, en revanche, seraient réalisés par un degré de proéminence sur la syllabe finale de N1 inférieur aux autres groupements, étant donné cette fois-ci l'absence de limite de groupement après la syllabe accentuée.

Rappel de nos prédictions :

### Condition 1 :

**G1** [ les AI **AF** ] // [ et les AI AF / AI AF ]

**G3** [ les AI **AF** ] // [ et les AI AF ] / [ AI AF ]

### Condition 2 :

**G2** [ les AI AF / et les AI AF ] // [ AI AF ]

**G3** [ les AI **AF** ] / [ et les AI AF ] // [ AI AF ]

**G4** [ les AI AF / et les AI AF // AI AF ]

Les résultats de la distribution des scores de proéminences montreraient toutefois un renforcement de la proéminence finale uniquement en G11. La force de la proéminence finale en G32 semblerait au contraire diminuer par rapport aux autres groupements, tandis que la proéminence finale en G31, mais aussi en G22 et G42, relèverait d'un niveau intermédiaire entre le niveau le plus fort en G11 et le niveau le plus faible en G32. Ce résultat laisserait alors penser qu'une proéminence finale est perçue sur N1 quel que soit le niveau de frontière marqué après la syllabe accentuée, et que la force de cette proéminence ne varierait pas strictement en fonction du niveau de constituance. En prenant également en compte nos résultats sur la force des frontières et sur les niveaux de constituance supposés, il semblerait qu'un même degré de proéminence puisse être perçu pour différents niveaux de constituance, mais également qu'une frontière plus faible puisse être marquée par un degré de proéminence plus élevé qu'une frontière plus forte. La frontière d'ip après N1 en G31 serait en effet marquée par le même degré de proéminence que la frontière de PW après N1 en G22 et G42. De la même manière, la frontière d'AP après N1 en G32 serait réalisée par un degré de proéminence inférieur à celui réalisé à la frontière de PW après N1 en G42 et G22.

**G11** [ les N1 **IP** ] // [ et les N2 **AP** / ADJ ] **IP**

**G31** [ les N1 **IP** ] // [ et les N2 **ip** ] / [ ADJ ] **IP**

**G32** [ les N1 **AP** ] / [ et les N2 **ip** ] // [ ADJ ] **IP**

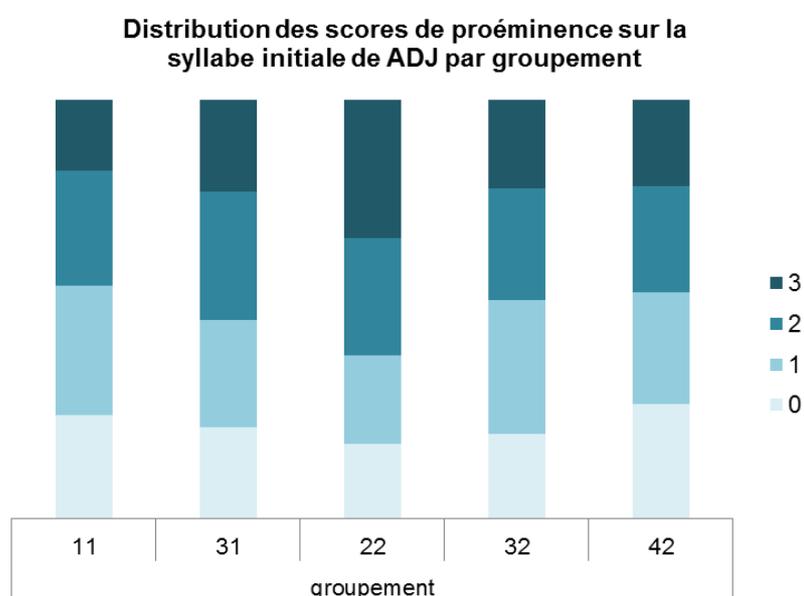
**G42** [ les N1 **PW** / et les N2 **AP** // ADJ ] **IP**

**G22** [ les N1 **PW** / et les N2 **ip** ] // [ ADJ ] **IP**

Nous savons par ailleurs que les syntagmes réalisés en Condition 1, donc en G11 et G31, sont caractérisés par une pause relativement longue après N1. Toutefois, les quatre locuteurs produisant les syntagmes ne réalisent pas forcément la même durée de pause. Il se pourrait donc que les syntagmes perçus en G11 soient associés à des pauses plus longues que les syntagmes perçus en G31, ce qui expliquerait alors la perception d'une proéminence sur la syllabe finale de N1 plus forte en G11 que dans d'autres groupements, et en revanche aussi forte en G31 que dans les autres groupements, malgré la présence d'un même niveau de frontière en G11 et G31 (frontières d'IP). Nous pourrions explorer cette hypothèse dans une future étude, en contrôlant directement les réalisations prosodiques pour chacun des locuteurs.

Proéminences sur la première syllabe de ADJ (cf. Figure N°55) :

Comme pour la proéminence finale sur N1, peu de variabilité distributionnelle est observée entre les différentes catégories de groupements (cf. Tableau de contingence en Annexe N°10-B). Seul G22 semble montrer des degrés de proéminence plus élevés, et notamment une répartition du score de niveau 3 plus importante (32,9% en G22 contre 16,9% à 22%) et une répartition du score de niveau 0 moins importante (17,9% en G22 contre 20,2% à 27,5%) que dans les autres groupements (cf. Figure N°55).



**Figure 55 :** *Distribution des scores de proéminences sur la syllabe initiale de ADJ par catégorie de groupement (G11, G31, G22, G32 et G42)*

Il semblerait également que G31 et G32 soient marqués par des degrés de proéminence moins élevés que G22, mais toutefois un peu plus élevés que G11 et G42. La différence entre les groupements s’observerait plus particulièrement sur la distribution du score de niveau 0 qui apparaît un peu moins important en G31 et G32 (20,2% à 22,0%) qu’en G11 et G42 (24,7% à 27,5%) (cf. Figure N°55).

La proéminence sur la première syllabe de ADJ serait donc perçue plus forte en G22 que dans les autres groupements. Il semblerait également que la proéminence soit perçue un peu plus forte en G31 et G32 qu’en G11 et G42. Comme pour les résultats de la proéminence finale sur N1, les différents degrés de proéminence observés ici doivent cependant être affinés dans le modèle de régression ordinaire à effet mixte, qui permet de mieux définir les contrastes entre les groupements à chaque niveau des scores de proéminences. Nous pouvons toutefois discuter ces résultats au regard de nos hypothèses de départ (cf. Hypothèse 3) et des résultats sur la force des frontières. Nos hypothèses prédisent en effet un renforcement de la proéminence initiale sur ADJ en G31, G22 et G32, étant donné la présence d’une limite de groupement avant la syllabe accentuée. Nous rappelons en effet que le renforcement du degré de la proéminence initiale par rapport au niveau de frontière est uniquement effectué, selon nous, sur la proéminence suivant la frontière. Les autres groupements, G11 et G42, ne devraient donc pas montrer de renforcement de la proéminence initiale sur ADJ, dans la mesure où aucune limite de groupement n’est réalisée avant la syllabe accentuée.

Rappel de nos prédictions :

### Condition 1 :

**G1** [ les AI AF ] // [ et les AI AF / AI AF ]

**G3** [ les AI AF ] // [ et les AI AF ] / [ AI AF ]

### Condition 2 :

**G2** [ les AI AF / et les AI AF ] // [ AI AF ]

**G3** [ les AI AF ] / [ et les AI AF ] // [ AI AF ]

**G4** [ les AI AF / et les AI AF // AI AF ]

Les résultats de la distribution des scores de proéminences semblent cependant indiquer un renforcement uniquement en G22. Les autres groupements seraient en effet marqués par un

degré de prééminence relativement équivalent, voire un peu plus fort en G31 et G32 qu'en G11 et G42. Il se pourrait dans ce contexte que la prééminence sur la première syllabe de ADJ dépende du niveau de constituance. Les résultats de l'analyse des scores de frontières montrent en effet que la frontière après N2 serait perçue la plus forte en G22, la plus faible en G11 et G42 et avec un degré intermédiaire en G31 et G32. Ces deux derniers groupements pourraient se distinguer selon la « configuration syllabique » de G11 et G42, mais ne se différencieraient jamais de G22. Nous supposons dans ce contexte que G31 et G32 relèvent du même degré de frontière que G22 (frontière d'ip), avec cependant des degrés de frontière qui peuvent être perçus moins forts dans le cas de « configurations syllabiques » spécifiques.

<b>G11</b>	[ les N1 ] <b>IP</b>	//	[ et les N2 <b>AP</b>	/	ADJ ] <b>IP</b>
<b>G31</b>	[ les N1 ] <b>IP</b>	//	[ et les N2 ] <b>ip</b>	/	[ ADJ ] <b>IP</b>
<b>G32</b>	[ les N1 ] <b>AP</b>	/	[ et les N2 ] <b>ip</b>	//	[ ADJ ] <b>IP</b>
<b>G42</b>	[ les N1 <b>PW</b>	/	et les N2 <b>AP</b>	//	ADJ ] <b>IP</b>
<b>G22</b>	[ les N1 <b>PW</b>	/	et les N2 ] <b>ip</b>	//	[ ADJ ] <b>IP</b>

Le renforcement du degré de frontière en G22 serait alors lié à la perception d'une prééminence sur la syllabe initiale de ADJ également plus forte. De la même manière, G31 et G32 pourraient être marqués par un renforcement plus ou moins important du degré de prééminence dans le cas de « configurations syllabiques » spécifiques, impliquant alors un degré de frontière également plus ou moins important. Il semble donc qu'il y ait une certaine corrélation entre l'augmentation du degré de frontière et du degré de prééminence sur la syllabe initiale de ADJ. Il est toutefois intéressant de prendre également en compte la relation de dépendance entre les scores de prééminences et la taille du constituant pouvant potentiellement venir affiner notre interprétation.

### 2.1.2 Distribution des scores de prééminences par taille du constituant

Nous supposons en effet un lien de dépendance entre le nombre de syllabes du constituant et les scores de prééminences. Comme pour l'analyse descriptive des scores de frontières, nous avons effectué plusieurs tests du Khi deux en proposant à chaque fois une variable différente liée à la taille du constituant, permettant ainsi d'avoir une idée plus précise du lien de dépendance potentiel entre les deux variables. Nous avons en effet testé le lien de

dépendance entre les scores de perception et le nombre de syllabes de N, le nombre de syllabes de ADJ, mais aussi le nombre de syllabes total du syntagme (D1+N1+cc+D2+N2+ADJ) dans une première version où le nombre de syllabes est additionné, puis dans une seconde version où le nombre de syllabes est concaténé<sup>14</sup>. Nous proposons par ailleurs de tester cette relation de dépendance entre les scores de prééminences et la taille du constituant en prenant en compte chacune des catégories de groupements, pour ainsi faciliter l'interprétation des résultats au regard toujours des différents phrasés prosodiques possibles pour les deux conditions syntaxiques.

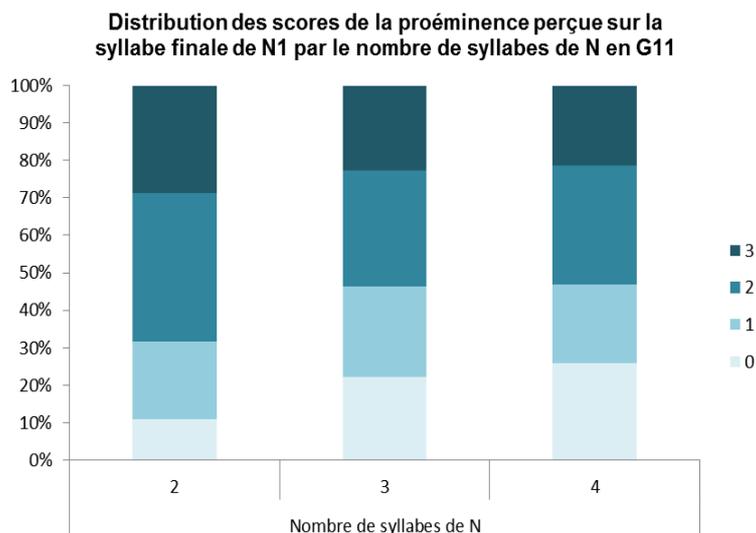
Les résultats du test du Khi deux montrent une relation de dépendance entre les scores de prééminences et la taille du constituant, mais uniquement dans le cas des prééminences sur la dernière syllabe de N1 en G11 et la première syllabe de ADJ en G22 et G32 (cf. Tableaux de contingence en Annexe N°11). Aucune correspondance n'est par ailleurs observée sur les autres sites de prééminences ( $p > 0,05$ ). Il est intéressant de remarquer que les prééminences intervenant dans la relation de dépendance avec la taille du constituant correspondent également aux prééminences intervenant dans la relation de dépendance avec le groupement. Il se pourrait qu'il y ait alors une étroite relation entre le groupement et la taille du constituant.

### Prééminences sur la syllabe finale de N1 en G11 (cf. Figure N°56) :

Les résultats montrent une relation de dépendance entre les scores de la prééminence sur la dernière syllabe de N1 et la taille du constituant, seulement en G11 (cf. Tableaux de contingence en Annexe N°11-A et N°11-B). Nous observons plus précisément un lien de dépendance entre la force de la prééminence et le nombre de syllabes de N (Khi deux= 19,090, ddl=6,  $p=0,004$ ), ainsi que le nombre de syllabes total du syntagme, dans la version où le nombre de syllabes est additionné uniquement (Khi deux= 30,552, ddl=18,  $p=0,032$ ). La dépendance serait donc due au nombre de syllabe de N quel que soit le nombre de syllabes de ADJ. La prééminence sur la syllabe finale de N1 en G11 semble effectivement perçue plus forte sur un N de 2 syllabes que sur un N de 3 ou 4 syllabes. Plus précisément, lorsque N est composé de 2 syllabes, les scores de niveau 2 et 3 tendent globalement à augmenter (68,4 contre 51,7% à 53,2% pour les autres tailles de N), tandis que le score de niveau 0 tend plutôt à diminuer (11% contre 22,2% à 25,9% pour les autres tailles de N).

---

<sup>14</sup> Pour plus de détails sur la variable « taille du constituant », nous renvoyons le lecteur à la partie consacrée à la préparation de notre base de données (cf. 3.3 du Chapitre 4).



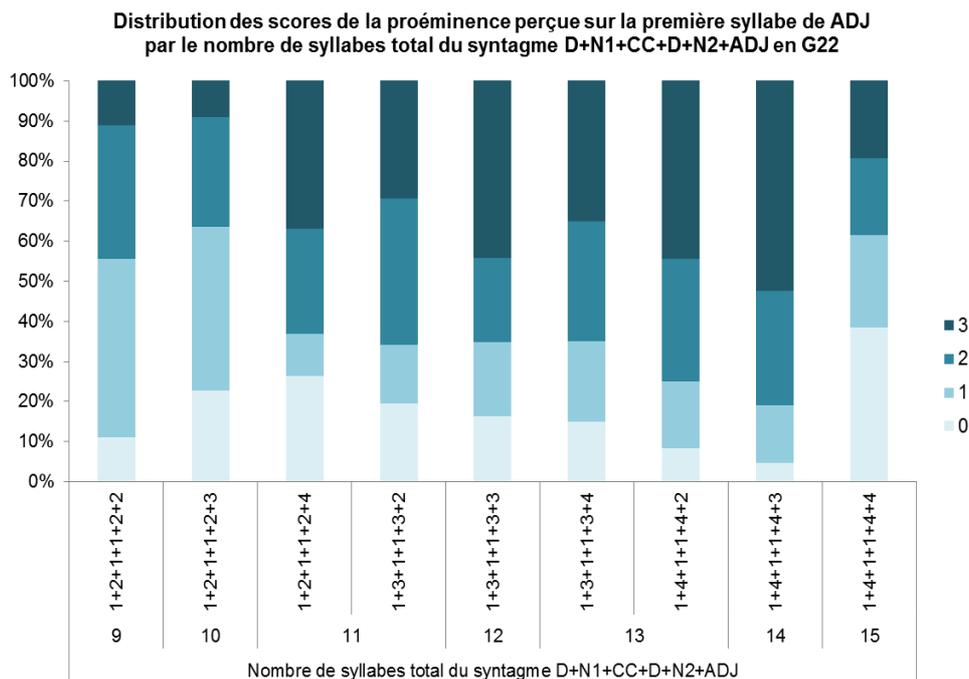
**Figure 56 :** *Distribution des scores de proéminences sur la dernière syllabe de N1 en G11 par le nombre de syllabes de N*

L'augmentation du nombre de syllabes de N pourrait alors avoir comme effet une diminution du degré de proéminence sur la syllabe finale de N1. Ce résultat serait lié par ailleurs à la réalisation potentielle d'un accent sur la syllabe initiale de N1 lorsque N est composé de plus de 2 syllabes, pouvant ainsi entraîner un abaissement du degré de la proéminence finale. Toutefois, ce résultat ne vaut que pour une seule catégorie de groupement, ce qui ne permet donc pas de généraliser le lien entre la réalisation d'une proéminence initiale et la diminution du niveau de proéminence finale à tous les autres groupements. Il est également intéressant de mettre en lien ce résultat avec celui obtenu dans l'analyse de la relation de dépendance entre les scores de proéminences et le groupement, où nous observons un renforcement du degré de proéminence sur la syllabe finale de N1 en G11 uniquement. Il se pourrait donc que le renforcement de la proéminence sur la syllabe finale de N1 observée en G11 par rapport aux autres groupements ne soit dû qu'au renforcement du degré de proéminence sur les N de 2 syllabes. Autrement dit, la proéminence sur la syllabe finale de N1 composé de 2 syllabes serait perçue plus forte en G11 que dans les autres groupements. En revanche, lorsque N est composé de 3 ou 4 syllabes, la proéminence sur la syllabe finale de N1 ne serait pas perçue plus forte en G11 que dans les autres groupements. Cela semble confirmer le fait que la force de cette proéminence ne varierait pas strictement en fonction du niveau de constituance.

### Proéminences sur la syllabe initiale de ADJ en G22 et G32 (cf. Figures N°57 et N°58) :

Les résultats indiquent une relation de dépendance, uniquement en G22 et G32, entre la proéminence perçue sur la première syllabe de ADJ et le nombre de syllabes total du syntagme, dans la version où le nombre de syllabes est additionné (G22 : Khi deux=37,922, ddl=18, p=0,004 ; G32 : Khi deux=29,450, ddl=18, p=0,043) ou concaténé (G22 : Khi deux=39,813, ddl=24, p=0,022 ; G32 : Khi deux= 37,419, ddl=24, p=0,040) (cf. Tableaux de contingence en Annexes N°11-C, N°11-D, N°11-E, N°11-F). Il est en revanche intéressant de remarquer que N et ADJ ne présentent pas de relation de dépendance avec les scores de proéminences lorsque leur taille est envisagée de manière isolée. Alors que pour la proéminence sur la syllabe finale de N1, la relation de dépendance serait principalement due au nombre de syllabes de N, quel que soit le nombre de syllabes de ADJ, la relation de dépendance entre la taille du constituant et la proéminence sur la première syllabe de ADJ, en revanche, serait due à la taille du syntagme en entier et donc à la « configuration syllabique » de N et ADJ ensemble.

Concernant tout d'abord G22, la proéminence sur la première syllabe de ADJ serait perçue plus forte lorsque le syntagme est composé d'un N de 2 syllabes suivi d'un ADJ de 4 syllabes (11 syllabes), ou d'un N de 3 syllabes suivi d'un ADJ de 2, 3 ou 4 syllabes (11, 12 et 13), ou encore lorsque le syntagme est composé d'un N de 4 syllabes suivi d'un ADJ de 2 ou 3 syllabes (14 syllabes) (cf. Figure N°57 et Tableaux de contingence en Annexe N°11-C et N°11-D). Nous observons en effet sur ces syntagmes une nette augmentation du score de niveau 3 (29,3% à 52,4 contre 9,1% à 11,1 pour les autres tailles) et une diminution du score de niveau 1 (10,5% à 20% contre 40,9% à 44,4% pour les autres tailles). Nous pourrions alors supposer que l'incrémentation en nombre de syllabes de N2+ADJ ait pour effet une augmentation de la force de la proéminence. Toutefois, lorsque N2+ADJ sont chacun composés du plus grand nombre de syllabes, c'est-à-dire lorsqu'un N de 4 syllabes est suivi d'un ADJ de 4 syllabes, la force de la proéminence sur la première syllabe de ADJ semble diminuer à nouveau. Nous observons en effet une diminution du score de niveau 3 (19,2%) et une augmentation du score de niveau 0 (38,5%).

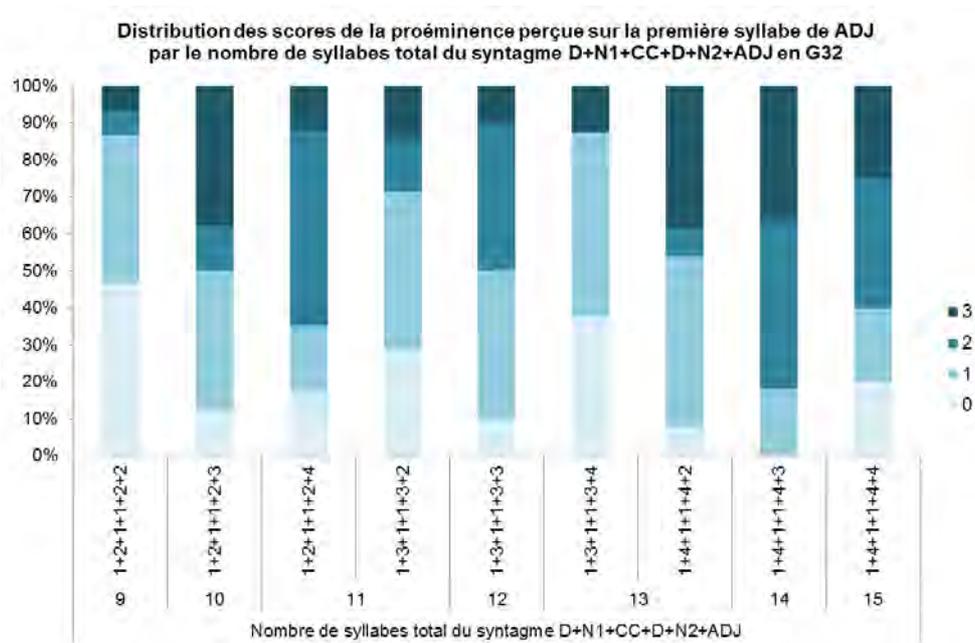


**Figure 57 :** *Distribution des scores de proéminences sur la première syllabe de ADJ par le nombre de syllabes total du syntagme D+N1+CC+D+N2+ADJ en G22*

Le renforcement de la proéminence sur la première syllabe de ADJ serait alors davantage lié à des « configurations syllabiques » de N2+ADJ spécifiques, qu'à la taille même du constituant. Il est par ailleurs intéressant de mettre en lien ce résultat avec celui obtenu dans l'analyse de la relation de dépendance entre les scores de proéminences et le groupement. Les résultats montrent en effet que la proéminence sur la première syllabe de ADJ serait perçue plus forte en G22 que dans les autres groupements. Dans ce contexte, il se pourrait que le renforcement de la proéminence sur la syllabe initiale de ADJ observée en G22 par rapport aux autres groupements ne s'observe que dans le cas de « configurations syllabiques » spécifiques. Ainsi, la corrélation que nous supposions précédemment entre l'augmentation du degré de frontière et du degré de proéminence sur la syllabe initiale de ADJ (*cf.* Hypothèse 2) ne peut être envisagée, dans la mesure où le renforcement du degré de proéminence en frontière forte ne soit valable que dans le cas de « configurations syllabiques » particulières. Comme pour la proéminence finale sur N1, la force de la proéminence sur la première syllabe de ADJ ne varierait donc pas strictement en fonction du niveau de constituance.

De la même manière, en G32, les résultats montrent que les scores de proéminence sur la première syllabe de ADJ semblent être plus élevés lorsque le constituant est composé

d'un N de 4 syllabes suivi d'un ADJ de 2, 3 ou 4 syllabes (correspondant aux constituants de 13, 14 et 15 syllabes) ou d'un N de 2 syllabes suivi d'un ADJ de 3 syllabes (correspondant au constituant de 10 syllabes) (cf. Figure N° 58 et Tableaux de contingence en Annexes N°11-E et N°11-F). Nous remarquons en effet dans ces syntagmes une distribution plus importante du score de niveau 3 (25% à 38,5%) et une distribution moins importante du score de niveau 0 (0% à 20%), par rapport aux autres tailles du constituant. A l'inverse, la prééminence sur la première syllabe de ADJ semble être perçue plus faible sur les constituants composés d'un N de 2 syllabes suivi d'un ADJ de 2 syllabes (au total 9 syllabes), ou d'un N de 3 syllabes suivi d'un ADJ de 4 syllabes (au total 13 syllabes) (cf. Figure N° 58 et tableaux de contingence en Annexes N°11-E et N°11-F). Nous observons effectivement une distribution plus importante du score de niveau 0 (37,5% à 46,7%), et une distribution moins importante des scores de niveau 2 (0% à 6,7%) et 3 (6,7% à 12,5%).



**Figure 58 :** *Distribution des scores de prééminences sur la première syllabe de ADJ par le nombre de syllabes total du syntagme D+N1+CC+D+N2+ADJ en G32*

Il semblerait qu'il y ait donc, comme en G22, une large variabilité des scores de prééminences en fonction du nombre de syllabes du syntagme, et plus spécifiquement en fonction de la manière dont N et ADJ se répartissent le nombre de syllabes à l'intérieur du syntagme. Nous observons effectivement un degré de prééminence qui peut être plus ou moins important pour une même taille de constituant. Les syntagmes composés de 13

syllabes, par exemple, montrent une distribution des scores nettement différente selon la « configuration syllabique » possible. Là encore, la manière dont N et ADJ se répartissent le nombre de syllabes du syntagme influencerait davantage la force de la proéminence que la taille seule du syntagme.

### **2.2 Choix des sites d'analyse**

Nos analyses confirment donc tout d'abord l'existence d'une relation de dépendance entre les scores de proéminences et le groupement, puis entre les scores de proéminences et la taille du constituant. Cependant, alors que nous attendions une relation de dépendance avec les scores de proéminences perçues sur les syllabes initiales et finales de tous les mots lexicaux, seules les proéminences perçues sur la syllabe finale de N1 et initiale de ADJ interviennent dans ces relations. Il est par ailleurs intéressant d'observer que ce sont toujours les mêmes sites de proéminences qui sont concernés par les deux relations de dépendance (groupement et taille du constituant). Les syllabes intermédiaires des mots lexicaux et les syllabes des mots grammaticaux ne montrent en revanche aucune relation de dépendance avec le groupement et la taille du constituant. Le marquage des groupements, et par conséquent des différents niveaux de frontière, serait dans ce contexte davantage lié à la proéminence finale sur N1 et à la proéminence initiale sur ADJ. L'analyse de la relation de dépendance entre les proéminences et la taille du constituant viennent toutefois moduler ces premiers résultats, dans la mesure où la différence de force des proéminences entre les groupements s'observerait dans le cas de « configurations syllabiques » spécifiques. La force des proéminences pourrait effectivement ne pas être strictement dépendante des niveaux de frontière, et par conséquent dépendante des niveaux de constituance marqués avant ou après la syllabe accentuée. Il est donc intéressant de procéder à un modèle de régression ordinaire à effet mixte, nous permettant d'affiner l'ensemble de ces résultats, où nous testerons directement l'effet de l'interaction entre le nombre de syllabes et le groupement sur la perception des scores de proéminences.

Dans ce contexte, nous choisissons d'inclure dans le modèle mixte 8 variables dépendantes associées aux scores de proéminences. Nous tenons en effet à tester la manière dont les proéminences sur la syllabe finale de N1 et initiale de ADJ ont été perçues en fonction du type de groupement choisi par l'auditeur, mais également en fonction de la taille des constituants. Plus précisément, nous choisissons de tester l'effet de la taille de N et de ADJ, mais aussi l'effet d'interaction entre la taille de N et ADJ, pour ainsi contrôler à la fois la taille et la « configuration syllabique » du constituant. Par ailleurs, afin d'avoir une

analyse plus syntagmatique que purement paradigmatique, nous proposons également de tester la manière dont les différentes proéminences dans le syntagme sont perçues par rapport à la proéminence sur la dernière syllabe de N1 d'une part, et par rapport à la proéminence sur la première syllabe de ADJ d'autre part. Nous nous intéressons plus précisément à la relation entre les proéminences positionnées autour des deux frontières prosodiques, après N1 et N2. Nous choisissons donc de tester la manière dont la proéminence sur la dernière syllabe de N1 est perçue par rapport aux proéminences sur la première syllabe de N1 et N2, mais également par rapport à la proéminence sur la dernière syllabe de N2. De la même manière, nous tenons à tester la manière dont la proéminence sur la première syllabe de ADJ est perçue par rapport à la proéminence sur la première syllabe de N2, mais également par rapport aux proéminences sur la dernière syllabe de N2 et ADJ. Ces dernières variables dépendantes correspondent aux variables que nous avons créées (cf. Préparation de la base de données 3.3 du Chapitre 4) visant à rendre compte de la force relative entre deux événements prosodiques. Enfin, dans chacun des modèles effectués, nous conservons systématiquement les trois variables aléatoires définies au préalable, à savoir les variables liées au participant, au locuteur et au type de syntagme.

### **2.3 Résultats du modèle de régression ordinale à effet mixte**

Nous proposons de présenter les résultats du modèle sur chaque site de proéminence isolé, pour ensuite commenter les résultats sur la force relative entre les proéminences perçues. Pour chaque modèle réalisé, nous avons par ailleurs analysé les composantes de la variance. A l'instar des analyses sur la perception des frontières, les résultats montrent systématiquement une valeur de la variance liée au participant plus élevée que celle liée au locuteur ou au syntagme. Les résultats obtenus dans le cadre de notre analyse varieraient donc davantage en fonction de l'auditeur qu'en fonction du locuteur produisant les différents syntagmes. Comme nous l'avons déjà supposé, il est en effet possible que tous les participants n'aient pas les mêmes critères d'attribution des scores de perception, et qu'ils n'estiment pas non plus de la même manière l'échelle d'évaluation (un score de même niveau pourrait renvoyer à une réalité différente selon les participants). Il est en revanche intéressant de se dire que les résultats obtenus dans nos analyses pourraient être valables quel que soit le locuteur prononçant le syntagme.

### 2.3.1 Force de la proéminence perçue sur la syllabe finale de N1

Les résultats montrent un effet significatif du groupement (LR Chisq= 28,64, df= 4,  $p < 0,001$ ) et du nombre de syllabes de N (LR Chisq= 19,82, df= 2,  $p < 0,001$ ) (cf. Annexe N°16-A).

#### Effet simple du groupement :

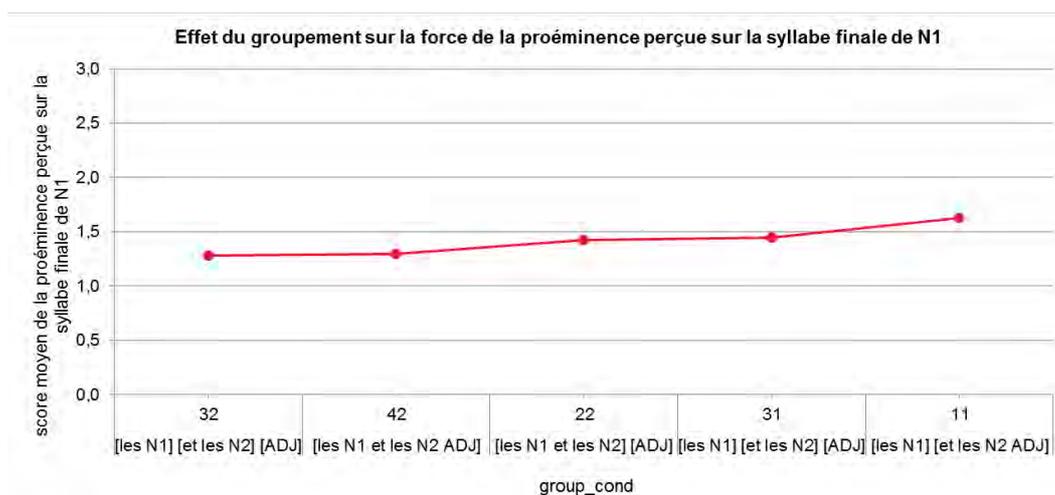
Les comparaisons post hoc de l'effet du groupement indiquent que la proéminence est perçue plus forte en G11 (score moyen : 1,63) qu'en G32 et G42 (score moyen : entre 1,28 et 1,30). En revanche, aucune différence n'est observée ni entre G31, G22, G32 et G42 ( $p < 0,05$ ), ni entre G31, G22 et G11 ( $p < 0,05$ ) (cf. Tableau N°7).

**Tableau 7 :** Résultats des comparaisons post hoc de l'effet du groupement sur la force de la proéminence perçue sur la syllabe finale de N1

Le tableau présente les moyennes des scores de proéminences interprétées comme des valeurs numériques ordonnées de 1 à 4 pour les valeurs « mean.class » et de 0 à 3 pour les valeurs « mean.class(-1) », calculées à partir des résultats de la distribution de probabilités estimée pour chaque catégorie de groupement (cf. Annexe N°16-A). Les groupes partageant une même lettre dans la colonne ".group" ne présentent pas de différence significative.

PROEMINENCE PERCUE SUR LA SYLLABE FINALE DE N1			
Moyennes des scores de proéminences calculées à partir des résultats de la distribution de probabilités			
group_cond	mean.class	mean.class(1)	.group
32	2,28	1,28	a
42	2,30	1,30	
22	2,42	1,42	ab
31	2,45	1,45	
11	2,63	1,63	b

Il semblerait donc que la proéminence sur la syllabe finale de N1 soit perçue avec un degré de proéminence un peu plus élevé en G11, un peu plus faible en G32 et G42 et avec un degré de proéminence intermédiaire en G31 et G22 (cf. Figure N°59).



**Figure 59 :** Illustration de l'effet du groupement sur la force de la proéminence perçue sur la syllabe finale de N1

Les points reliés par des lignes correspondent aux valeurs « *mean.class (-1)* » du tableau N°7 indiquant les moyennes des scores de proéminences perçues pour chaque groupement, interprétées comme des valeurs numériques ordonnées de 0 à 3

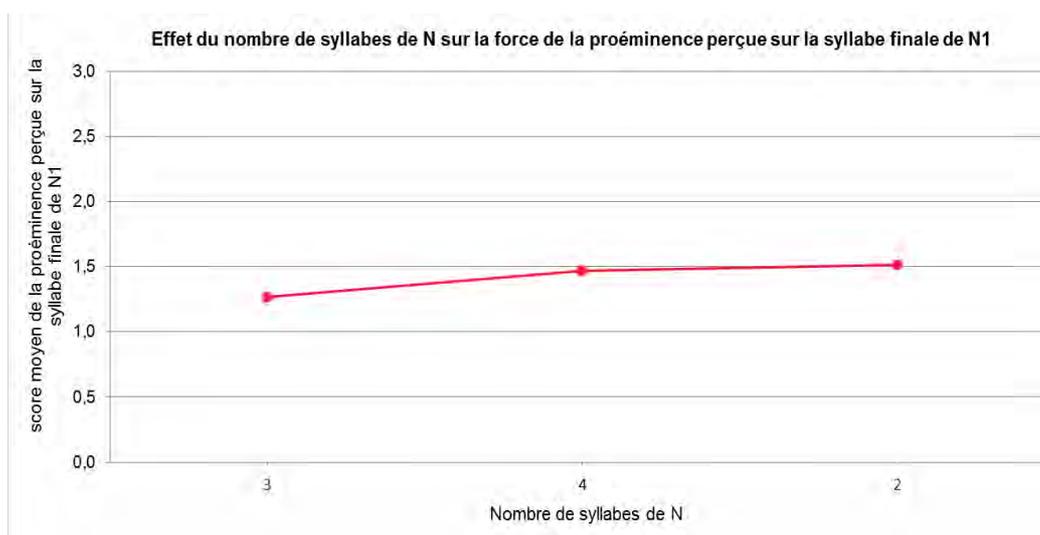
Ce résultat ne validerait donc pas nos hypothèses de départ, où une corrélation entre le renforcement de la proéminence finale et le niveau de frontière marqué après la syllabe accentuée était supposée. En effet, les résultats tendraient plutôt à admettre que la force de cette proéminence ne varierait pas strictement en fonction du niveau de constituance. Nous avons déjà supposé par ailleurs ce résultat dans l'analyse descriptive de la relation de dépendance entre les scores de proéminences et le niveau de frontière. Une même force de proéminence peut donc être perçue à des niveaux de frontière différents. La proéminence est perçue aussi forte en G31 qu'en G22, alors que les niveaux de frontière marqués dans ces deux groupements relèveraient respectivement du niveau de l'IP et du PW. La proéminence finale sur N1 serait perçue toujours avec la même force dès le niveau du PW, et n'augmenterait donc pas en fonction du niveau de frontière. Bien que ce résultat ne corresponde pas tout à fait à nos hypothèses de départ, il semblerait toutefois que le fait de percevoir une proéminence finale dès le niveau du PW, qui ne varie pas en fonction du niveau de constituance, soit en quelque sorte la preuve d'une certaine dissociation entre les scores de proéminences et de frontières. En effet, nos résultats tendent à montrer que l'accent final en français ne s'effacerait pas perceptivement en frontière prosodique majeure (*contra* Jun & Fougeron, 2000 ; 2002 ; Rossi, 1980), mais au contraire qu'il y aurait toujours une trace de l'accent final de mot perçue (Astésano, 2017 ; Astésano et al., 2012 ; 2013).

Effet simple du nombre de syllabes de N :

Concernant l'effet simple du nombre de syllabes de N, les résultats de l'analyse post hoc montrent que la proéminence sur la syllabe finale de N1 est perçue plus forte lorsque N est composé de 2 syllabes (score moyen : 1,51) plutôt que de 3 syllabes (score moyen : 1,27). Aucune différence n'est en revanche observée ni entre les N de 4 et 3 syllabes ( $p > 0,05$ ), ni entre les N de 4 et 2 syllabes ( $p > 0,05$ ) (cf. Tableau N°8 et Figure N°60).

**Tableau 8 :** Résultats des comparaisons post hoc de l'effet du nombre de syllabes de N sur la force de la proéminence perçue sur la syllabe finale de N1  
Le tableau présente les moyennes des scores de proéminences interprétées comme des valeurs numériques ordonnées de 1 à 4 pour les valeurs « mean.class » et de 0 à 3 pour les valeurs « mean.class(-1) », calculées à partir des résultats de la distribution de probabilités estimée pour chaque taille de N (cf. Annexe N°16-A). Les groupes partageant une même lettre dans la colonne ".group" ne présentent pas de différence significative.

PROEMINENCE PERÇUE SUR LA SYLLABE FINALE DE N1			
Moyennes des scores de proéminences calculées à partir des résultats de la distribution de probabilités			
Nombre de syllabes de N	mean.class	mean.class(-1)	.group
3	2,27	1,27	a
4	2,47	1,47	ab
2	2,51	1,51	b



**Figure 60 :** Illustration de l'effet du nombre de syllabes de N sur la force de la proéminence perçue sur la syllabe finale de N1  
Les points reliés par des lignes correspondent aux valeurs « mean.class (-1) » du tableau N°8 indiquant les moyennes des scores de proéminences perçues pour chaque taille de N, interprétées comme des valeurs numériques ordonnées de 0 à 3

Cela pourrait supposer que la force de la proéminence finale diminue sur les mots composés de plus de 2 syllabes, étant donné l'apparition d'un accent potentiel sur la première syllabe des mots composés de 3 et 4 syllabes. Toutefois, nos résultats n'indiquent pas de différence entre les N de 2 et 4 syllabes. Nous avons également observé cet effet dans l'analyse descriptive de la relation de dépendance entre les scores de proéminences et la taille du constituant, mais uniquement dans le cas de G11. Nous avons alors supposé à ce moment-là que le renforcement de la proéminence sur la syllabe finale de N1, observée en G11 par rapport aux autres groupements, n'était dû qu'au renforcement du degré de proéminence sur les N de 2 syllabes. Cependant, nos résultats du modèle de régression ordinaire à effet mixte ne montrent pas d'effet d'interaction entre la catégorie de groupement et le nombre de syllabes sur la perception de la proéminence. Il semblerait donc que le renforcement de la proéminence finale sur les N de 2 syllabes par rapport aux N de 3 syllabes soit davantage marqué en G11, tout en restant cependant effectif quel que soit le groupement. Il n'y aurait donc pas de corrélation entre le marquage du niveau de frontière après N1 et le nombre de syllabes des constituants.

### 2.3.2 Force de la proéminence perçue sur la syllabe initiale de ADJ

Les résultats indiquent un effet significatif du groupement (LR Chisq= 11,41, df= 4, p=0,02), du nombre de syllabes de N (LR Chisq= 7,12, df= 2, p=0,02), ainsi qu'un effet d'interaction entre le nombre de syllabes de N et ADJ (LR Chisq= 16,92, df= 4, p=0,002) (cf. Annexe N°16-B).

#### Effet simple du groupement :

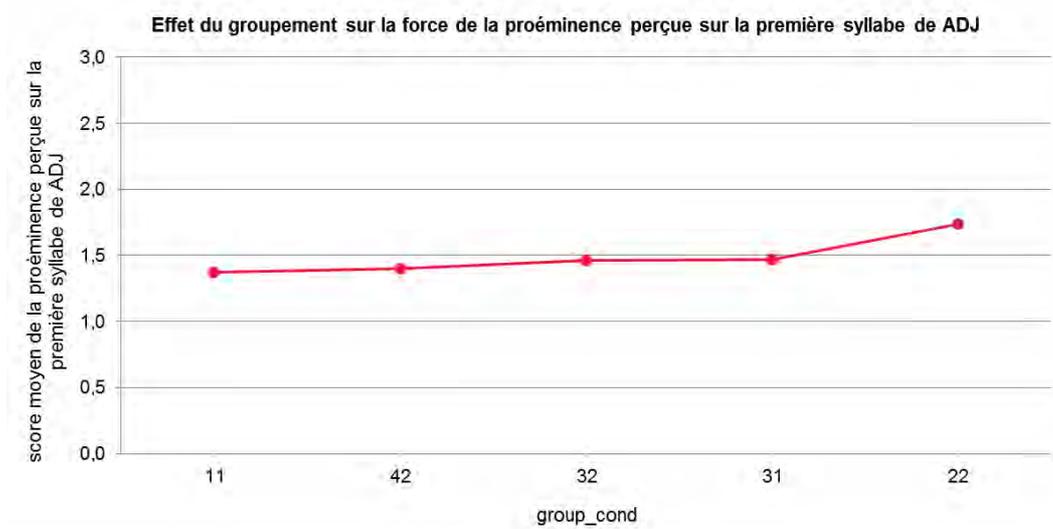
Les analyses post hoc de l'effet du groupement montrent que la proéminence sur la première syllabe de ADJ est perçue plus forte en G22 (score moyen : 1,74) qu'en G11 et G42 (score moyen : entre 1,37 et 1,40). Aucune différence entre les autres groupements n'est par ailleurs observée (p>0,05) (cf. Tableau N°9 et Figure N°61).

Il semblerait donc que la proéminence initiale sur ADJ soit perçue avec un degré de proéminence un peu plus élevé en G22, un peu plus faible en G11 et G42 et avec un degré de proéminence intermédiaire en G31 et G32.

**Tableau 9 :** Résultats des comparaisons post hoc de l'effet du groupement sur la force de la proéminence perçue sur la première syllabe de ADJ

Le tableau présente les moyennes des scores de proéminences interprétées comme des valeurs numériques ordonnées de 1 à 4 pour les valeurs « mean.class » et de 0 à 3 pour les valeurs « mean.class(-1) », calculées à partir des résultats de la distribution de probabilités estimée pour chaque groupement (cf. Annexe N°16-B). Les groupes partageant une même lettre dans la colonne ".group" ne présentent pas de différence significative.

PROEMINENCE PERCUE SUR LA SYLLABE INITIALE DE ADJ			
Moyennes des scores de proéminences calculées à partir des résultats de la distribution de probabilités			
group_cond	mean.class	mean.class(1)	.group
11	2,37	1,37	a
42	2,40	1,40	
32	2,46	1,46	ab
31	2,47	1,47	
22	2,74	1,74	b



**Figure 61 :** Illustration de l'effet du groupement sur la force de la proéminence perçue sur la première syllabe de ADJ

Les points reliés par des lignes correspondent aux valeurs « mean.class (-1) » du tableau N°9 indiquant les moyennes des scores de proéminences perçues pour chaque groupement, interprétées comme des valeurs numériques ordonnées de 0 à 3

Ce résultat confirmerait donc notre hypothèse quant à la corrélation entre le renforcement de la proéminence initiale sur ADJ et le marquage du niveau de frontière réalisée avant la

syllabe accentuée. Les résultats de l'analyse des scores de frontières montrent en effet que la frontière après N2 est perçue la plus forte en G22, la plus faible en G11 et G42 et avec un degré intermédiaire en G31 et G32. Ces deux derniers groupements peuvent en effet se distinguer de G11 et G42 selon la « configuration syllabique » du syntagme, mais ne se différencient jamais de G22. Nous supposons dans ce contexte que G31 et G32 relèvent du même niveau de frontière que G22 (frontière d'ip), avec cependant des degrés de frontière qui peuvent être perçus moins forts dans le cas de « configurations syllabiques » spécifiques. Le renforcement du degré de frontière en G22 pourrait alors être lié à la perception d'une proéminence sur la syllabe initiale de ADJ également plus forte. De la même manière, G31 et G32 pourraient être marqués par un renforcement plus ou moins important du degré de proéminence dans le cas de « configurations syllabiques » spécifiques, impliquant alors un degré de frontière également plus ou moins important. Ces résultats confirmeraient donc l'idée, selon laquelle AI joue ici un rôle pertinent dans le marquage de la structure prosodique (Astésano et al., 2007).

### Effet simple du nombre de syllabes de N :

Les résultats de l'analyse post hoc montrent que la proéminence sur la première syllabe de ADJ est perçue plus forte lorsque ADJ est précédé d'un N de 4 syllabes plutôt que d'un N de 2 syllabes (scores moyens : 1,64 contre 1,36). Aucune différence entre les autres tailles de N n'est par ailleurs observée ( $p > 0,05$  cf. Annexe N°16-B). Cela supposerait alors que la force de la proéminence initiale sur ADJ augmente dès lors que cet ADJ est précédé d'un N composé d'un grand nombre de syllabes.

### Effet d'interaction entre le nombre de syllabes de N et le nombre de syllabes de ADJ :

L'effet simple du nombre de syllabes de N peut toutefois être modulé par l'effet d'interaction entre le nombre de syllabes de N et le nombre de syllabes de ADJ, qui permet de tester à quel point la relation entre le score de la proéminence perçue sur la syllabe initiale de ADJ et le nombre de syllabes de N varie en fonction de la taille de ADJ. En regardant les contrastes post hoc (cf. Tableau (A) et (B) N°10), il apparaît que la proéminence initiale sur ADJ est perçue plus forte lorsque ADJ est précédé d'un N de 4 syllabes (score moyen : 1,82) plutôt que d'un N de 2 ou 3 syllabes (scores moyens : 1,21 à 1,37), mais uniquement dans le cas où ADJ est lui-même composé de 2 syllabes.

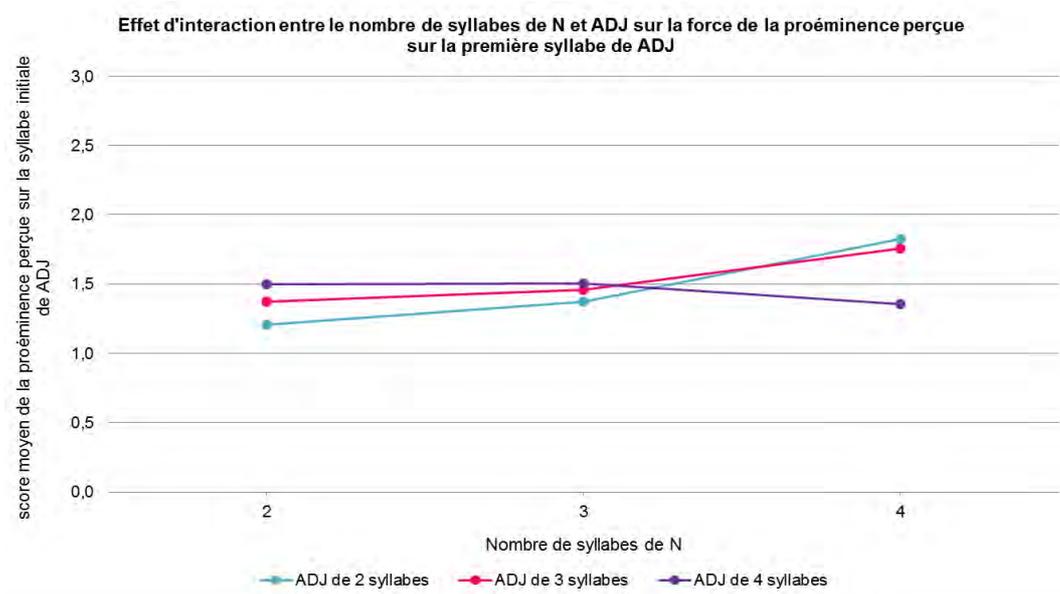
**Tableau 10 :** Résultats des comparaisons post hoc de l'effet d'interaction entre le nombre de syllabes de N et de ADJ sur la force de la proéminence perçue sur la syllabe initiale de ADJ

A) Lecture des contrastes en colonne pour un même nombre de syllabes de ADJ et B) Lecture des contrastes en ligne pour un même nombre de syllabes de N. Dans chaque tableau, les moyennes des forces de proéminences sont interprétées comme des valeurs numériques ordonnées de 1 à 4 pour les valeurs « mean.class » et de 0 à 3 pour les valeurs « mean.class(-1) », qui sont calculées à partir des résultats de la distribution de probabilités estimée pour chaque taille de N et de ADJ (cf. Annexe N°16-B). Les groupes partageant une même lettre dans la colonne ".group" ne présentent pas de différence significative.

PROÉMINENCE PERÇUE SUR LA SYLLABE INITIALE DE ADJ						
Moyennes des scores de proéminences calculées à partir des résultats de la distribution de probabilités						
A) Lecture des contrastes en colonne pour un même nombre de syllabes de ADJ						
Nombre de syllabes de N	Nombre de syllabes de ADJ					
	2		3		4	
	mean.class	.group	mean.class	.group	mean.class	.group
	mean.class(1)		mean.class(1)		mean.class(1)	
2	2,21	a	2,37	ab	2,50	abc
	1,21		1,37		1,50	
3	2,37	ab	2,46	abc	2,51	abc
	1,37		1,46		1,51	
4	2,82	c	2,76	bc	2,36	ab
	1,82		1,76		1,36	
B) Lecture des contrastes en ligne pour un même nombre de syllabes de N						
Nombre de syllabes de N	Nombre de syllabes de ADJ					
	2		3		4	
	mean.class	.group	mean.class	.group	mean.class	.group
	mean.class(1)		mean.class(1)		mean.class(1)	
2	2,21	a	2,37	ab	2,50	abc
	1,21		1,37		1,50	
3	2,37	ab	2,46	abc	2,51	abc
	1,37		1,46		1,51	
4	2,82	c	2,76	bc	2,36	ab
	1,82		1,76		1,36	

La force de la proéminence initiale sur ADJ n'augmenterait donc pas automatiquement lorsque cet ADJ est précédé d'un N composé d'un grand nombre de syllabes, dans la mesure où cet effet n'est valable que pour les ADJ composés de 2 syllabes. D'un point de vue rythmique, il semblerait que le renforcement de la proéminence initiale soit davantage attendu sur les ADJ de plus 2 syllabes, dans la mesure où le constituant N qui précède est déjà composé d'un grand nombre de syllabes. En admettant effectivement un rôle rythmique de l'AI, plus le constituant est long plus la probabilité d'occurrence de l'AI augmente. Il apparaît toutefois que l'influence du nombre de syllabes du constituant sur la perception de la proéminence initiale sur ADJ soit davantage due à la manière dont N2 et ADJ se

répartissent le nombre de syllabes, qu'à la taille seule de N2+ADJ. Nous avons également observé ce résultat dans l'analyse descriptive de la relation de dépendance entre les scores de prééminences et la taille du constituant. Toutefois, cette relation de dépendance n'était valable qu'en G22. Dans notre présente analyse, aucun effet d'interaction entre le groupement et le nombre de syllabes du constituant sur la perception de la prééminence initiale de ADJ n'est observé. Dans ce contexte, le renforcement de la prééminence initiale sur ADJ de 2 syllabes précédé d'un N de 4 syllabes pourrait donc être d'autant plus marqué en G22, mais également effectif dans tous les autres groupements (*cf.* Figure N°62).



**Figure 62 :** Illustration de l'effet d'interaction entre le nombre de syllabes de N et ADJ sur la force de la prééminence perçue sur la première syllabe de ADJ

Les points reliés par des lignes correspondent aux valeurs « mean.class (-1) » du tableau N°10 indiquant les moyennes des scores de prééminences perçues pour chaque taille de N et ADJ, interprétées comme des valeurs numériques ordonnées de 0 à 3

Comme pour AF sur N1, il semblerait donc qu'il n'y ait pas de corrélation spécifique entre le marquage du niveau de frontière après N2 et le nombre de syllabes des constituants. Ce résultat semble également confirmer qu'AI est perçu sur tous les ADJ, quel que soit le niveau de constituance, mais également quel que soit le nombre de syllabes de ADJ. Il est en effet intéressant de remarquer qu'AI est perçu dès les constituants composés de 2 syllabes. AI n'interviendrait donc pas uniquement dans le cas de constituants composés d'un grand

nombre de syllabes. Cela tendrait à consolider le postulat selon lequel AI serait plus structurel que purement rythmique (Astésano et *al.*, 2007).

Au vu de l'ensemble de ces résultats, il apparaît tout d'abord que la proéminence initiale sur ADJ serait davantage influencée par les catégories de groupements, et donc par les niveaux de frontière, que la proéminence finale sur N1. Les résultats montrent effectivement peu de variation du degré de proéminence finale en fonction des niveaux de frontière, et en revanche une variation un peu plus importante du degré de proéminence initiale. Cela tendrait alors à montrer un rôle de l'AI plus pertinent dans le marquage de la structure prosodique. Quant à l'influence de la taille du constituant, il semblerait que la manière dont les N et ADJ se répartissent le nombre de syllabes ait une influence plus importante que la taille seule du syntagme. Enfin, pour les deux types de proéminences, nous ne constatons aucune interaction entre le marquage du niveau de frontière et le nombre de syllabes des constituants. Ces résultats nécessitent toutefois de prendre en compte les autres sites de proéminences adjacents, pouvant nous aider à affiner davantage notre interprétation.

### **2.3.3 Force relative perçue entre la proéminence finale sur N1 et les autres proéminences initiales ou finales adjacentes**

Nous proposons en effet une analyse syntagmatique de la force relative des proéminences, nous permettant de rendre compte du rapport hiérarchique entre les différentes proéminences dans le syntagme. Nous nous intéressons tout d'abord à la manière dont la proéminence finale sur N1 est perçue par rapport à la proéminence initiale de N1, mais aussi par rapport à la proéminence initiale et finale de N2. Pour cela, plusieurs calculs ont été effectués.

Dans le cas de la force relative entre la proéminence finale et initiale de N1, nous avons soustrait le score de la proéminence perçue sur la syllabe finale de N1 à celui de la proéminence perçue sur la syllabe initiale de N1. Nous obtenons ainsi une force relative indiquant la manière dont la proéminence sur la syllabe finale est perçue par rapport à la proéminence sur la syllabe initiale. Une valeur de force relative négative signifierait que la proéminence sur la syllabe initiale de N1 serait perçue comme plus forte que la proéminence sur la syllabe finale de N1, tandis qu'une valeur de force relative positive montrerait que la proéminence sur la syllabe initiale de N1 serait perçue comme plus faible que la proéminence sur la syllabe finale de N1. De la même manière, pour calculer la force relative entre la proéminence finale de N1 et la proéminence initiale ou finale de N2, nous avons

soustrait le score de la proéminence perçue sur la syllabe initiale ou finale de N2 à celui de la proéminence perçue sur la syllabe finale de N1. Nous obtenons ainsi une force relative indiquant la manière dont la proéminence sur la syllabe finale de N1 est perçue par rapport à la proéminence sur la syllabe initiale ou finale de N2. Plus précisément, une valeur de force relative négative signifierait que la proéminence sur la syllabe finale de N1 serait perçue comme plus forte que la proéminence sur la syllabe initiale ou finale de N2, tandis qu'une valeur de force relative positive montrerait que la proéminence sur la syllabe finale de N1 serait perçue comme plus faible que la proéminence sur la syllabe initiale ou finale de N2<sup>15</sup>.

Par ailleurs, les graphiques représentant les résultats ont été rééchelonnés pour ainsi mieux rendre compte des effets du groupement et/ou de la taille du constituant sur la perception de la force relative entre les proéminences. Nous rappelons en effet que l'échelle initiale des scores de force relative comprend 7 valeurs, entre -3 et 3. Toutefois, comme nous le signalerons tout au long des résultats, les valeurs de force relative ne s'étendent pas sur toute l'échelle des scores, et se situent davantage entre les valeurs -1,5 et 1,5, supposant par ailleurs des différences relativement ténues. Nous proposons donc de réduire l'échelle de -1,5 à 1,5 dans tous les graphiques présentés dans le cadre de cette partie.

### 2.3.3.1 Proéminence finale sur N1 – Proéminence initiale sur N1

Les résultats montrent un seul effet significatif du groupement (LR Chisq= 33,16, df=4, p<0,001). La force relative entre les deux proéminences ne varierait donc pas en fonction de la taille du constituant (p>0,05) (cf. Annexe N°16-C). Les analyses post hoc indiquent plus précisément une différence significative entre G11 d'un côté, puis G32 et G42 de l'autre. La valeur de force relative est en effet plus importante en G11 (score moyen : 0,27) qu'en G32 et G42 (score moyen : entre -0,12 et -0,22). Aucune opposition entre les autres groupements n'est par ailleurs observée (p>0,05) (cf. Tableau N°11 et Figure N°63).

Il semblerait donc, étant donné la valeur de force relative positive en G11, que la proéminence finale sur N1 soit perçue plus forte que la proéminence initiale sur N1, et étant donné la valeur de force relative négative en G32 et G42, que la proéminence initiale sur N1 soit perçue cette fois-ci plus forte que la proéminence finale sur N1. Dans les autres groupements, les deux proéminences seraient en revanche perçues avec une force relativement équivalente (cf. Figure N°63).

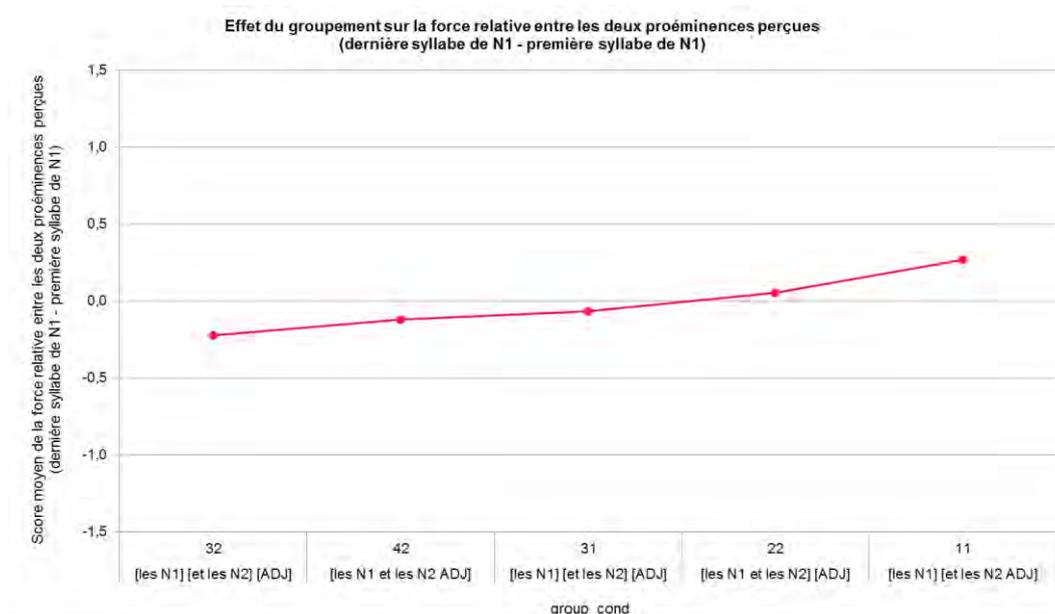
---

<sup>15</sup> Pour plus de détails sur la construction des variables de force relative, nous renvoyons le lecteur à la partie consacrée à la préparation de notre base de données (cf. 3.3 du Chapitre 4).

**Tableau 11** : Résultats des comparaisons post hoc de l'effet du groupement sur la force relative entre les deux prééminences perçues sur N1 (dernière syllabe de N1 – première syllabe de N1)

Le tableau présente les moyennes des scores de prééminences interprétées comme des valeurs numériques ordonnées de 1 à 7 pour les valeurs « mean.class » et de -3 à 3 pour les valeurs « mean.class(-4) », calculées à partir des résultats de la distribution de probabilités estimée pour chaque catégorie de groupement (cf. Annexe N°16-C). Les groupes partageant une même lettre dans la colonne ".group" ne présentent pas de différence significative.

FORCE RELATIVE ENTRE LES DEUX PROEMINENCES PERÇUES: DERNIERE SYLLABE DE N1 - PREMIERE SYLLABE DE N1			
Moyennes des scores de force relative entre les prééminences calculées à partir des résultats de la distribution de probabilités			
group_cond	mean.class	mean.class(-4)	.group
32	3,78	-0,22	a
42	3,88	-0,12	
31	3,93	-0,07	ab
22	4,05	0,05	
11	4,27	0,27	b



**Figure 63** : Illustration de l'effet du groupement sur la force relative entre les deux prééminences perçues sur N1 (dernière syllabe de N1 – première syllabe de N1)

Les points reliés par des lignes correspondent aux valeurs « mean.class (-4) » du tableau N°11 indiquant les moyennes des scores de prééminences perçues pour chaque groupement, interprétées comme des valeurs numériques ordonnées de -3 à 3. Pour faciliter la lecture des résultats, nous choisissons de réduire l'échelle des scores moyens de la force relative entre les deux prééminences aux valeurs -1,5 à 1,5.

Nous retrouvons alors ici les résultats de l'effet simple du groupement sur la perception de la proéminence finale sur N1, où le renforcement de la proéminence observé en G11 est significativement plus important que celui observé en G32 et G42. La proéminence initiale sur N1 serait alors majoritairement perçue aussi forte ou plus forte que la proéminence finale sur N1, à l'exception de G11, uniquement, où un renforcement de la proéminence finale pourrait être observé.

### 2.3.3.2 Proéminence initiale sur N2 – Proéminence finale sur N1

Les résultats indiquent un seul effet significatif du groupement (LR Chisq= 17,98, df= 4, p=0,001). La force relative entre les deux proéminences ne varierait donc pas en fonction de la taille du constituant (p>0,05) (cf. Annexe N°16-D). Il est intéressant de remarquer que l'effet seul du groupement est déjà observé dans l'analyse précédente. Les analyses post hoc montrent plus précisément une différence significative entre G11 et G32. La valeur de force relative est en effet moins importante en G11 (score moyen : -0,10) qu'en G32 (score moyen : 0,31). Aucune différence n'est par ailleurs observée entre les autres groupements (p>0,05). (cf. Tableau N°12 et Figure N°64)

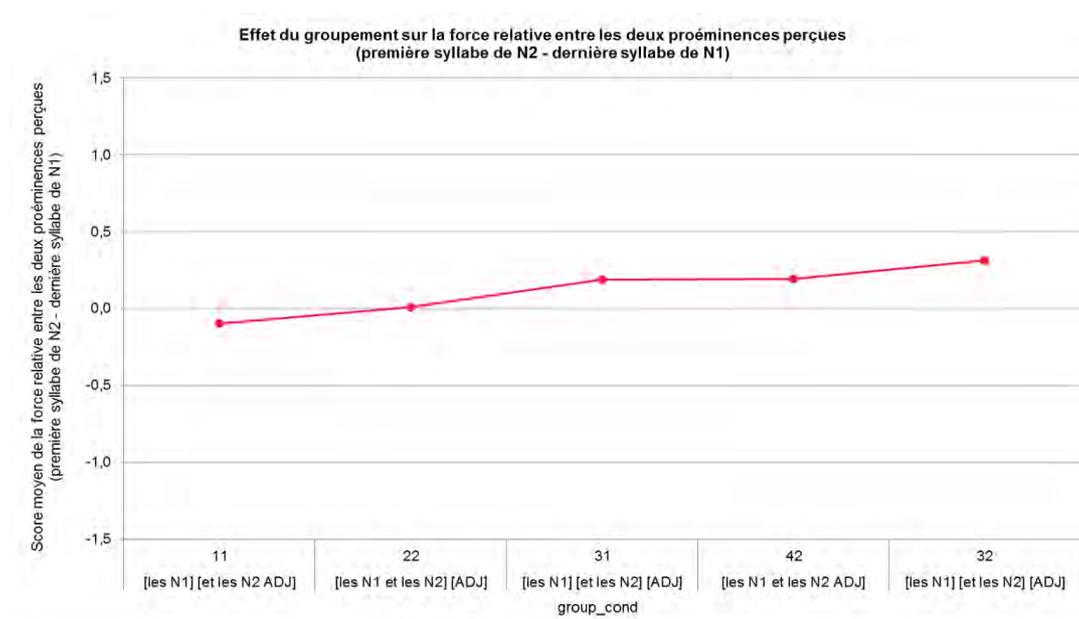
**Tableau 12 :** Résultats des comparaisons post hoc de l'effet du groupement sur la force relative entre les deux proéminences perçues sur N1 et N2 (première syllabe de N2 – dernière syllabe de N1)

Le tableau présente les moyennes des scores de proéminences interprétées comme des valeurs numériques ordonnées de 1 à 7 pour les valeurs « mean.class » et de -3 à 3 pour les valeurs « mean.class(-4) », calculées à partir des résultats de la distribution de probabilités estimée pour chaque groupement (cf. Annexe N°16-D). Les groupes partageant une même lettre dans la colonne ".group" ne présentent pas de différence significative.

FORCE RELATIVE ENTRE LES DEUX PROEMINENCES PERCUES: PREMIERE SYLLABE DE N2 - DERNIERE SYLLABE DE N1			
Moyennes des scores de force relative entre les proéminences calculées à partir des résultats de la distribution de probabilités			
group_cond	mean.class	mean.class(-4)	.group
11	3,90	-0,10	a
22	4,01	0,01	ab
31	4,19	0,19	
42	4,19	0,19	
32	4,31	0,31	b

Il semblerait donc, étant donné la valeur de force relative négative en G11, que la proéminence finale sur N1 soit perçue plus forte que la proéminence initiale sur N2, et étant

donné la valeur de force relative positive en G32, que la proéminence initiale sur N2 soit perçue plus forte que la proéminence finale sur N1. Dans les autres groupements, les deux proéminences seraient en revanche perçues avec une force relativement équivalente (cf. Figure N°64).



**Figure 64 :** Illustration de l'effet du groupement sur la force relative entre les deux proéminences perçues sur N1 et N2 (première syllabe de N2 - dernière syllabe de N1)  
 Les points reliés par des lignes correspondent aux valeurs « mean.class (-4) » du tableau N°12 indiquant les moyennes de force relative des proéminences perçues pour chaque groupement, interprétées comme des valeurs numériques ordonnées de -3 à 3. Pour faciliter la lecture des résultats, nous choisissons de réduire l'échelle des scores moyens de la force relative entre les deux proéminences aux valeurs -1,5 à 1,5.

Ce résultat semble tout d'abord confirmer les résultats de l'effet simple du groupement sur la perception de la proéminence finale sur N1, où le renforcement de la proéminence observé en G11 est significativement plus important que celui observé en G32 et G42. En revanche, dans la présente analyse, seul G32 s'oppose à G11. La proéminence initiale sur N2 serait alors majoritairement perçue aussi forte ou plus forte que la proéminence finale sur N1, à l'exception de G11, uniquement, où un renforcement de la proéminence finale pourrait être observé. D'autre part, comme pour la syllabe initiale de N1, les résultats semblent également confirmer ici le fait qu'Al soit perçu sur tous les N2, quelle que soit la catégorie de groupement ou la taille du constituant. Nous rappelons en effet que le nombre de syllabes du syntagme n'affecte pas ici la perception de la force relative entre les proéminences sur la syllabe initiale de N2 et finale de N1. De la même manière, les résultats tendent à montrer

une absence de corrélation entre le renforcement d'une proéminence ou d'une autre et la force de la frontière marquée avant ou après la syllabe accentuée (entre N1 et N2), dans la mesure où une même force relative entre la proéminence initiale de N2 et finale de N1 peut être observée pour différents niveaux de constituance. Nous constatons en effet une force relative équivalente en G31 et G42, qui sont respectivement marqués par une frontière d'IP et de PW, où la proéminence initiale sur N2 serait perçue aussi forte voire un peu plus forte que la proéminence finale sur N1. Il semblerait donc que la force relative entre les deux proéminences ne varie pas strictement en fonction du niveau de constituance.

### 2.3.3.3 Proéminence finale sur N2 – Proéminence finale sur N1

Les résultats montrent un effet significatif du groupement (LR Chisq= 16,68, df=4, p=0,002) et du nombre de syllabes de N (LR Chisq= 16,44, df= 2, p<0,001) (cf. Annexe N°16-E).

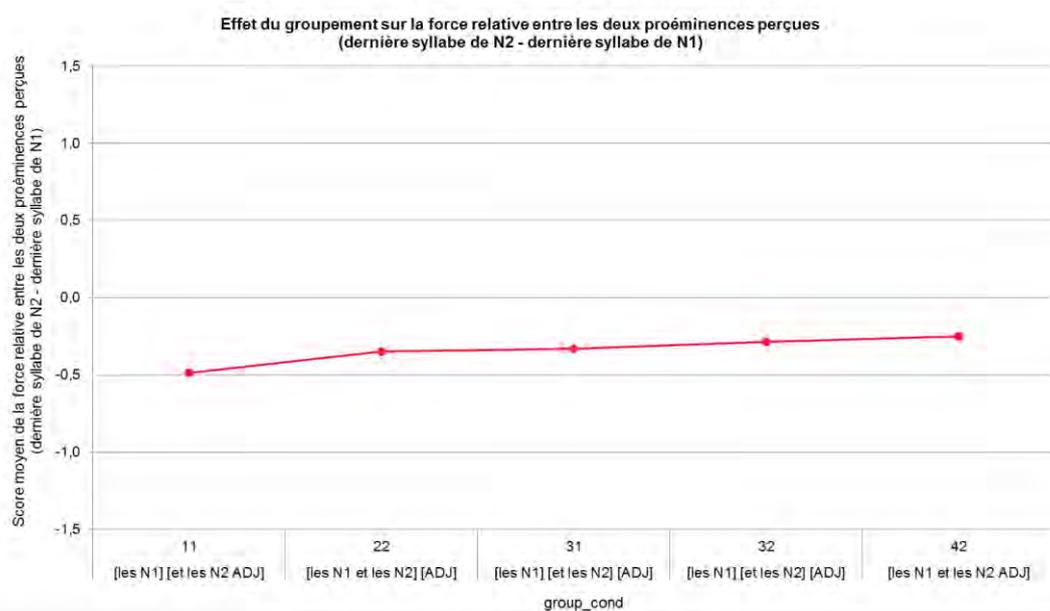
#### Effet simple du groupement :

L'analyse post hoc de l'effet du groupement indique une seule différence significative entre G11 et G42. La valeur de force relative est en effet moins importante en G11 (score moyen : -0,49) qu'en G42 (score moyen : -0,25). Aucune différence n'est par ailleurs observée entre les autres groupements (p>0,05). Nous remarquons également que toutes les valeurs de force relative sont ici négatives. Ainsi, quelle que soit la catégorie de groupement, la proéminence sur la syllabe finale de N1 est perçue plus forte que la proéminence finale sur N2. Toutefois, au vu des résultats, la proéminence finale sur N1 serait perçue encore plus forte que la proéminence finale sur N2 en G11 par rapport à G42 (cf. Tableau N°13 et Figure N°65).

**Tableau 13 :** Résultats des comparaisons post hoc de l'effet du groupement sur la force relative entre les deux proéminences perçues sur N1 et N2 (dernière syllabe de N2 – dernière syllabe de N1)

Le tableau présente les moyennes des scores de proéminences interprétées comme des valeurs numériques ordonnées de 1 à 7 pour les valeurs « mean.class » et de -3 à 3 pour les valeurs « mean.class(-4) », calculées à partir des résultats de la distribution de probabilités estimée pour chaque groupement (cf. Annexe N°16-E). Les groupes partageant une même lettre dans la colonne ".group" ne présentent pas de différence significative.

FORCE RELATIVE ENTRE LES DEUX PROEMINENCES PERCUES: DERNIERE SYLLABE DE N2 - DERNIERE SYLLABE DE N1			
Moyennes des scores de force relative entre les proéminences calculées à partir des résultats de la distribution de probabilités			
group_cond	mean.class	mean.class(-4)	.group
11	3,51	-0,49	a
22	3,65	-0,35	ab
31	3,67	-0,33	
32	3,71	-0,29	b
42	3,75	-0,25	



**Figure 65 :** Illustration de l'effet du groupement sur la force relative entre les deux proéminences perçues sur N1 et N2 (dernière syllabe de N2 – dernière syllabe de N1)

Les points reliés par des lignes correspondent aux valeurs « mean.class (-4) » du tableau N°13 indiquant les moyennes de force relative des proéminences perçues pour chaque groupement, interprétées comme des valeurs numériques ordonnées de -3 à 3. Pour faciliter la lecture des résultats, nous choisissons de réduire l'échelle des scores moyens de la force relative entre les deux proéminences aux valeurs -1,5 à 1,5.

Nous retrouvons donc là encore l'effet simple du groupement sur la perception de la proéminence finale sur N1, où le renforcement de la proéminence observé en G11 est significativement plus important que celui observé en G32 et G42. En revanche, dans la présente analyse, seul G42 s'oppose à G11. AF sur N1 serait donc systématiquement perçu plus fort qu'AF sur N2, avec cependant une différence entre les deux proéminences plus marquée en G11 qu'en G42. Il est par ailleurs intéressant de voir qu'AF sur N2 est perçu moins fort qu'AF sur N1, alors que d'un point de vue de la constituance, les niveaux de frontière après N2 sont parfois plus élevés que les niveaux de frontière après N1. En G22, par exemple, la frontière après N1 relève du niveau du PW et la frontière après N2 relève du niveau de l'ip. Cela tendrait alors à confirmer qu'AF sur N1 est perçu dès le niveau du PW.

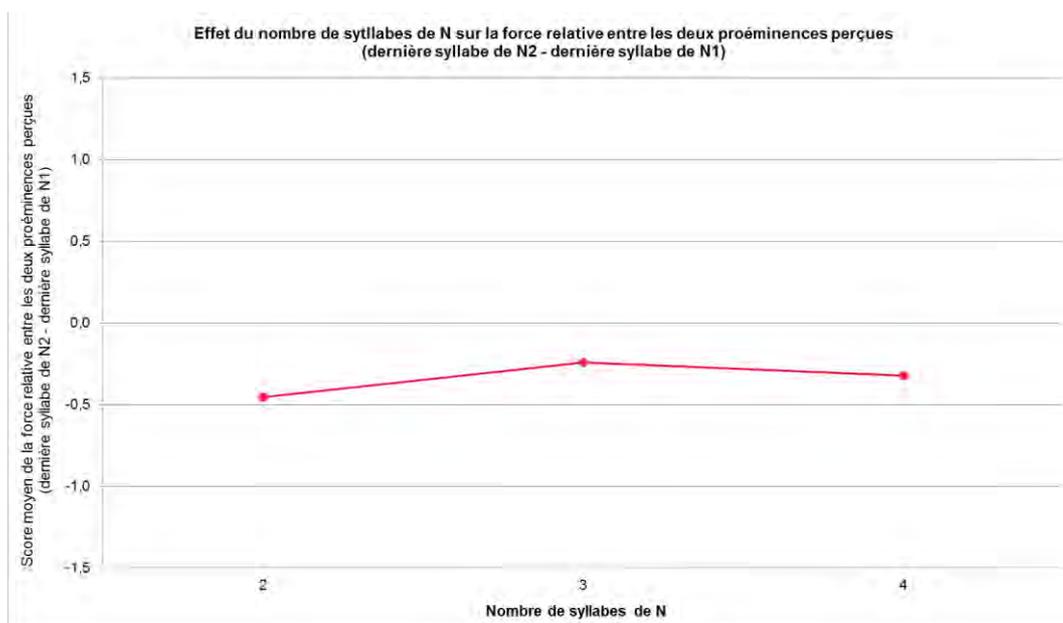
### Effet simple du nombre de syllabes de N :

Concernant l'effet du nombre de syllabes de N, les résultats de l'analyse post hoc montrent une différence entre les deux proéminences lorsque N est composé de 2 et 3 syllabes. Plus précisément, la valeur de force relative est moins importante lorsque N est composé de 2 syllabes (score moyen : -0,45) plutôt que de 3 syllabes (score moyen : -0,24). Aucune différence n'est cependant observée ni entre les N de 2 et 4 syllabes, ni entre les N de 3 et 4 syllabes ( $p > 0,05$ ). Il est par ailleurs intéressant de remarquer que toutes les valeurs de force relative sont négatives. Ainsi, quel que soit le nombre de syllabes de N, la proéminence sur la syllabe finale de N1 est perçue plus forte que la proéminence finale sur N2. Toutefois, au vu des résultats, la proéminence finale sur N1 serait perçue encore plus forte que la proéminence finale sur N2 lorsque N est composé de 2 syllabes plutôt que de 3 syllabes (cf. Tableau N°14 et Figure N°66).

**Tableau 14 :** Résultats des comparaisons post hoc de l'effet du nombre de syllabes de N sur la force relative entre les deux proéminences perçues sur N1 et N2 (dernière syllabe de N2 – dernière syllabe de N1)

Le tableau présente les moyennes des scores de proéminences interprétées comme des valeurs numériques ordonnées de 1 à 7 pour les valeurs « mean.class » et de -3 à 3 pour les valeurs « mean.class(-4) », calculées à partir des résultats de la distribution de probabilités estimée pour chaque taille de N (cf. Annexe N°16-E). Les groupes partageant une même lettre dans la colonne ".group" ne présentent pas de différence significative.

FORCE RELATIVE ENTRE LES DEUX PROEMINENCES PERCUES: DERNIERE SYLLABE DE N2 - DERNIERE SYLLABE DE N1			
Moyennes des scores de force relative entre les proéminences calculées à partir des résultats de la distribution de probabilités			
nombre de syllabes de N	mean.class	mean.class(-4)	.group
2	3,55	-0,45	a
3	3,76	-0,24	b
4	3,68	-0,32	ab



**Figure 66 :** Illustration de l'effet du nombre de syllabes de N sur la force relative entre les deux proéminences perçues sur N1 et N2 (dernière syllabe de N2 – dernière syllabe de N1)

Les points reliés par des lignes correspondent aux valeurs « mean.class (-4) » du tableau N°14 indiquant les moyennes de force relative des proéminences perçues pour chaque taille de N, interprétées comme des valeurs numériques ordonnées de -3 à 3. Pour faciliter la lecture des résultats, nous choisissons de réduire l'échelle des scores moyens de la force relative entre les deux proéminences aux valeurs -1,5 à 1,5.

Ce résultat correspondrait donc à celui obtenu dans l'analyse de l'effet simple du nombre de syllabes de N sur la perception de la proéminence finale de N1. Nous avons vu en effet que la proéminence est perçue plus forte sur les N de 2 syllabes que sur les N de 3 syllabes. Ainsi, AF sur N1 serait perçu majoritairement plus fort qu'AF sur N2, avec cependant une différence entre les deux proéminences plus marquée lorsque N est composé de 2 syllabes plutôt que de 3 syllabes. Il se pourrait donc que le renforcement du degré de proéminence lié au nombre de syllabes de N soit uniquement effectif sur la syllabe finale de N1, et par conséquent non généralisable aux syllabes finales des autres mots lexicaux. Nous rappelons en effet que N2 et ADJ ne montrent aucun effet significatif du nombre de syllabes du constituant. Dans ce contexte, AF ne serait pas directement influencé par la taille du constituant, excepté pour un site de proéminence particulier dans une taille de constituant spécifique.

Sur la base de l'ensemble des résultats, il semblerait qu'AI sur N1 ou N2 soit majoritairement perçu aussi fort ou plus fort qu'AF sur N1. Les résultats tendent également à montrer que la force relative entre la proéminence initiale et finale ne varie pas strictement

en fonction du niveau de constituance. Il apparaît par ailleurs qu'AI comme AF sont réalisés sur N1 et N2, quelles que soient la catégorie de groupement et la taille du constituant. Nous rappelons en effet que le nombre de syllabes du syntagme n'affecte pas la perception de la force relative entre les proéminences sur la syllabe initiale et finale de N1 et N2, et affecte dans une moindre mesure la force relative entre les proéminences finales sur N1 et N2. Les deux proéminences seraient alors perçues dès les constituants composés de 2 syllabes, ce qui tendrait par ailleurs à admettre un rôle plus structurel que purement rythmique de l'AI. Cette dernière interprétation pourra par ailleurs être affinée avec l'analyse de la force relative entre la proéminence initiale sur ADJ et les autres proéminences adjacentes.

### **2.3.4 Force relative perçue entre la proéminence initiale sur ADJ et les autres proéminences initiales ou finales adjacentes**

Après avoir testé la force relative entre la proéminence sur N1 et les proéminences initiales ou finales adjacentes, nous nous intéressons maintenant à la manière dont la proéminence initiale sur ADJ est perçue par rapport à la proéminence finale de ADJ, mais aussi par rapport à la proéminence initiale et finale de N2. Pour cela, plusieurs calculs ont été effectués.

Dans le cas de la force relative entre la proéminence initiale et finale de ADJ, nous avons soustrait le score de la proéminence perçue sur la syllabe finale de ADJ à celui de la proéminence perçue sur la syllabe initiale de ADJ. Nous obtenons ainsi une force relative indiquant la manière dont la proéminence sur la syllabe initiale de ADJ est perçue par rapport à la proéminence sur la syllabe finale de ADJ. Une valeur de force relative négative signifierait que la proéminence sur la syllabe initiale de ADJ serait perçue comme plus forte que la proéminence sur la syllabe finale de ADJ, tandis qu'une valeur de force relative positive montrerait que la proéminence sur la syllabe initiale de ADJ serait perçue comme plus faible que la proéminence sur la syllabe finale de ADJ. De la même manière, pour calculer la force relative entre la proéminence initiale de ADJ et la proéminence initiale ou finale de N2, nous avons soustrait le score de la proéminence perçue sur la syllabe initiale de ADJ à celui de la proéminence perçue sur la syllabe initiale ou finale de N2. Nous obtenons ainsi une force relative indiquant la manière dont la proéminence sur la syllabe initiale de ADJ est perçue par rapport à la proéminence sur la syllabe initiale ou finale de N2. Plus précisément, une valeur de force relative négative signifierait que la proéminence sur la syllabe initiale de ADJ serait perçue comme plus faible que la proéminence sur la syllabe initiale ou finale de N2, tandis qu'une valeur de force relative positive montrerait que la

proéminence sur la syllabe finale de N1 serait perçue comme plus forte que la proéminence sur la syllabe initiale ou finale de N2<sup>16</sup>.

Par ailleurs, les graphiques représentant les résultats ont été rééchelonnés pour ainsi mieux rendre compte des effets du groupement et/ou de la taille du constituant sur la perception de la force relative entre les proéminences. Nous rappelons en effet que l'échelle initiale des scores de force relative comprend 7 valeurs, entre -3 et 3. Toutefois, comme nous le signalerons tout au long des résultats, les valeurs de force relative ne s'étendent pas sur toute l'échelle des scores, et se situent davantage entre les valeurs -1,5 et 1,5, supposant par ailleurs des différences relativement ténues. Nous proposons donc de réduire l'échelle de -1,5 à 1,5 dans tous les graphiques présentés dans le cadre de cette partie.

### 2.3.4.1 Proéminence finale sur ADJ – Proéminence initiale sur ADJ

Les résultats montrent un effet du groupement (LR Chisq= 11,20, df= 4, p=0,02) ainsi qu'un effet d'interaction entre le nombre de syllabes de N et ADJ (LR Chisq= 29,91, df= 4, p<0,001) (cf. Annexe N°16-F).

#### Effet simple du groupement :

Les analyses post hoc de l'effet du groupement montrent une différence significative entre G22 et G11. Aucune différence n'est en revanche observée dans les autres groupements (p>0,05). La valeur de force relative entre les proéminences est en effet moins élevée en G22 (score moyen : -0,25) qu'en G11 (score moyen : 0,12) (cf. Tableau N°15 et Figure N°67).

Il semblerait donc, étant donné la valeur de force relative positive en G11, que la proéminence initiale sur ADJ soit perçue moins forte que la proéminence finale sur ADJ, et étant donné la valeur de force négative en G22, que la proéminence initiale sur ADJ soit perçue plus forte que la proéminence finale sur ADJ. Dans les autres groupements, les deux proéminences seraient en revanche perçues avec une force relativement équivalente.

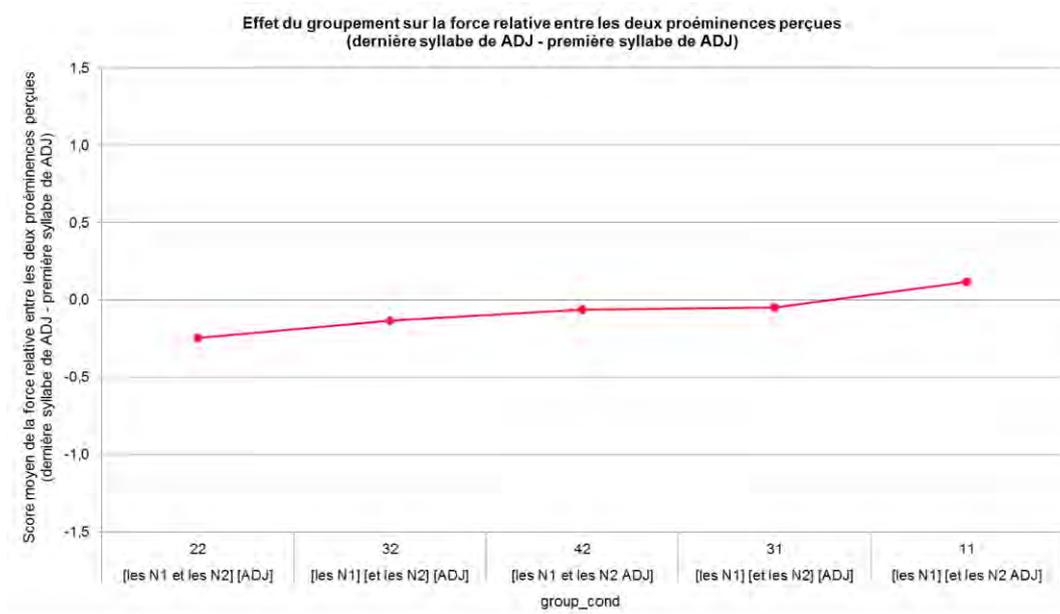
---

<sup>16</sup> Pour plus de détails sur la construction des variables de force relative, nous renvoyons le lecteur à la partie consacrée à la préparation de notre base de données (cf. 3.3 du Chapitre 4).

**Tableau 15** : Résultats des comparaisons post hoc de l'effet du groupement sur la force relative entre les deux prééminences perçues sur ADJ (dernière syllabe de ADJ – première syllabe de ADJ)

Le tableau présente les moyennes des scores de prééminences interprétées comme des valeurs numériques ordonnées de 1 à 7 pour les valeurs « mean.class » et de -3 à 3 pour les valeurs « mean.class(-4) », calculées à partir des résultats de la distribution de probabilités estimée pour chaque catégorie de groupement (cf. Annexe N°16-F). Les groupes partageant une même lettre dans la colonne ".group" ne présentent pas de différence significative.

FORCE RELATIVE ENTRE LES DEUX PROEMINENCES PERÇUES: DERNIERE SYLLABE DE ADJ - PREMIERE SYLLABE DE ADJ			
Moyennes des scores de force relative entre les prééminences calculées à partir des résultats de la distribution de probabilités			
group_cond	mean.class	mean.class(-4)	.group
22	3,75	-0,25	a
32	3,87	-0,13	ab
42	3,94	-0,06	
31	3,95	-0,05	b
11	4,12	0,12	



**Figure 67** : Illustration de l'effet du groupement sur la force relative entre les deux prééminences perçues sur ADJ (dernière syllabe de ADJ – première syllabe de ADJ)

Les points reliés par des lignes correspondent aux valeurs « mean.class (-4) » du tableau N°15 indiquant les moyennes des scores de prééminences perçues pour chaque groupement, interprétées comme des valeurs numériques ordonnées de -3 à 3. Pour faciliter la lecture des résultats, nous choisissons de réduire l'échelle des scores moyens de la force relative entre les deux prééminences aux valeurs -1,5 à 1,5.

Ce résultat semble alors correspondre au résultat de l'effet simple du groupement sur la perception de la proéminence initiale sur ADJ, où le renforcement de la proéminence observé en G22 est significativement plus important que celui observé en G11 et G42. En revanche, dans la présente analyse, seul G11 s'oppose à G22. AI sur ADJ serait donc majoritairement perçu aussi fort ou plus fort qu'AF sur ADJ, excepté en G11 où un renforcement de la proéminence finale sur ADJ pourrait être observé. Alors que nous avons supposé une certaine corrélation entre le renforcement du degré de proéminence initiale sur ADJ et le marquage du niveau de frontière située avant la syllabe accentuée, il semblerait en revanche qu'il n'y ait pas de corrélation entre la force relative et le niveau de frontière. Il est en effet possible d'observer une même force relative dans différents groupements, marqués par des niveaux plus ou moins importants. G31 et G42, par exemple, qui sont réalisés respectivement par une frontière d'ip et d'AP après N2, présentent une force relative équivalente, où AI sur ADJ serait perçu aussi fort qu'AF sur ADJ.

### Effet d'interaction entre le nombre de syllabes de N et le nombre de syllabes ADJ :

Concernant l'effet d'interaction entre le nombre de syllabes de N et ADJ, nous constatons des différences de force relative, selon la « configuration syllabique » de N2+ADJ. En regardant effectivement les contrastes post hoc (*cf.* Tableau (A) et (B) N°16), nous voyons une différence entre les forces relatives principalement sur les ADJ précédés d'un N de 4 syllabes, où une différence apparaît entre les ADJ de 2 et 3 syllabes d'un côté (scores moyens : -0,42 à -0,48) et les ADJ de 4 syllabes de l'autre (score moyen : 0,29). La valeur de la force relative est en effet positive pour les ADJ de 4 syllabes et négative pour les ADJ de 2 et 3 syllabes. Il semblerait donc que la proéminence initiale sur ADJ soit perçue moins forte que la proéminence finale sur ADJ lorsqu'un ADJ de 4 syllabes est précédé d'un N de 4 syllabes, et plus forte que la proéminence finale sur ADJ lorsqu'un ADJ de 2 et 3 syllabes est précédé d'un N de 4 syllabes (*cf.* Figure N°68). Il est donc intéressant de remarquer que lorsque le constituant ADJ augmente en nombre de syllabes, c'est la proéminence finale et non pas initiale de ADJ qui est perçue la plus forte. Nous pourrions supposer par ailleurs que le renforcement de la proéminence finale sur ADJ est lié à l'augmentation du nombre de syllabes total de N2+ADJ. C'est en effet le constituant composé du plus grand nombre de syllabes qui semble être marqué par un AF plus fort sur ADJ.

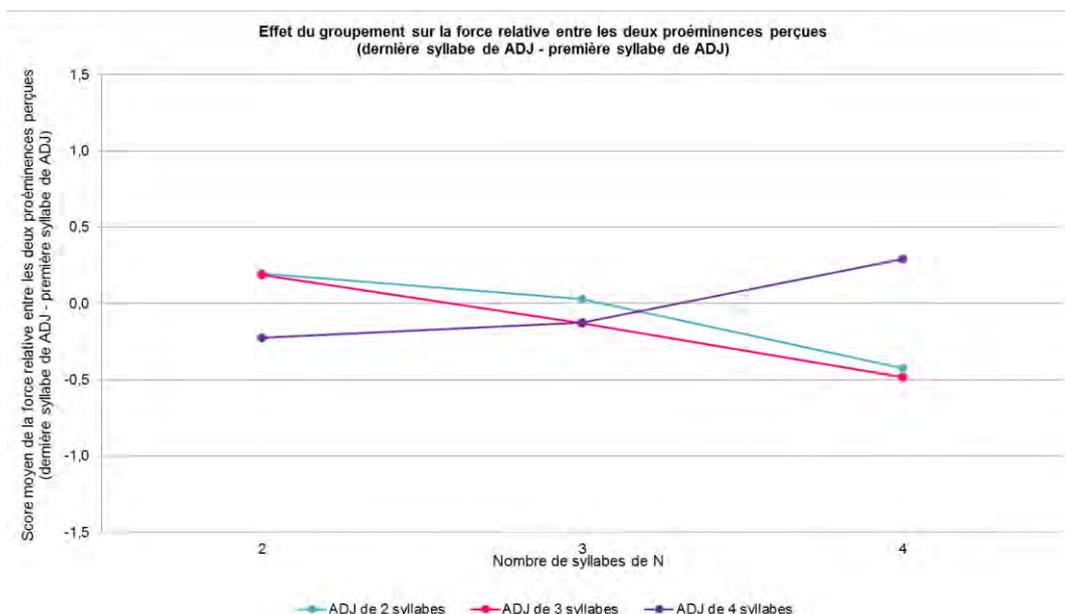
Nous observons également une différence de force relative sur les ADJ précédés d'un N de 2 syllabes. Plus précisément, une différence apparaît entre les ADJ de 2 et 4 syllabes (scores moyens respectifs : 0,20 contre -0,22). La valeur de force relative est en effet négative pour les ADJ de 4 syllabes et positive pour les ADJ de 2 syllabes. AI sur ADJ

serait donc perçu plus fort qu'AF sur ADJ lorsqu'un N de 2 syllabes est suivi d'un ADJ de 4 syllabes, et moins fort qu'AF sur ADJ lorsqu'un N de 2 syllabes est suivi d'un ADJ de 2 syllabes (cf. Figure N°68). A l'inverse de ce qu'on a pu observer précédemment, lorsque le constituant ADJ augmente en nombre de syllabes, c'est la proéminence initiale et non pas finale de ADJ qui est perçue la plus forte. De la même manière, alors que nous supposions un renforcement du degré de la proéminence finale sur ADJ lié à l'augmentation du nombre de syllabes total de N2+ADJ, nous remarquons ici l'effet inverse puisque c'est la proéminence sur la syllabe initiale de ADJ qui est renforcée lorsque le nombre de syllabes total de N2+ADJ augmente.

**Tableau 16 :** Résultats des comparaisons post hoc de l'effet d'interaction entre le nombre de syllabes de N et de ADJ sur la force relative entre les deux proéminences perçues sur ADJ (dernière syllabe de ADJ – première syllabe de ADJ)

A) Lecture des contrastes en colonne pour un même nombre de syllabes de ADJ et B) Lecture des contrastes en ligne pour un même nombre de syllabes de N. Dans chaque tableau, les moyennes des forces relatives des proéminences sont interprétées comme des valeurs numériques ordonnées de 1 à 7 pour les valeurs « mean.class » et de -3 à 3 pour les valeurs « mean.class(-4) », qui sont calculées à partir des résultats de la distribution de probabilités estimée pour chaque taille de N et de ADJ (cf. Annexe N°16-F). Les groupes partageant une même lettre dans la colonne ".group" ne présentent pas de différence significative.

FORCE RELATIVE ENTRE LES DEUX PROEMINENCES PERCUES: DERNIERE SYLLABE DE ADJ - PREMIERE SYLLABE DE ADJ						
Moyennes des scores de force relative entre les proéminences calculées à partir des résultats de la distribution de probabilités						
A) Lecture des contrastes en colonne pour un même nombre de syllabes de ADJ						
Nombre de syllabes de N	Nombre de syllabes de ADJ					
	2		3		4	
	mean.class mean.class(4)	.group	mean.class mean.class(4)	.group	mean.class mean.class(4)	.group
2	4,20	c	4,19	bc	3,78	ab
	0,20		0,19		-0,22	
3	4,03	bc	3,87	abc	3,87	abc
	0,03		-0,13		-0,13	
4	3,58	a	3,52	a	4,29	c
	-0,42		-0,48		0,29	
B) Lecture des contrastes en ligne pour un même nombre de syllabes de N						
Nombre de syllabes de N	Nombre de syllabes de ADJ					
	2		3		4	
	mean.class mean.class(4)	.group	mean.class mean.class(4)	.group	mean.class mean.class(4)	.group
2	4,20	c	4,19	bc	3,78	ab
	0,20		0,19		-0,22	
3	4,03	bc	3,87	abc	3,87	abc
	0,03		-0,13		-0,13	
4	3,58	a	3,52	a	4,29	c
	-0,42		-0,48		0,29	



**Figure 68 :** Illustration de l'effet d'interaction entre le nombre de syllabes de N et ADJ sur la force relative entre les deux proéminences perçues sur ADJ (dernière syllabe de ADJ – première syllabe de ADJ)

Les points reliés par des lignes correspondent aux valeurs « mean.class (-4) » du tableau N°16 indiquant les moyennes des scores de proéminences perçues pour chaque taille de N et ADJ, interprétées comme des valeurs numériques ordonnées de -3 à 3. Pour faciliter la lecture des résultats, nous choisissons de réduire l'échelle des scores moyens de la force relative entre les deux proéminences aux valeurs -1,5 à 1,5.

Il semble difficile dans ce contexte de déterminer précisément des règles rythmiques pouvant expliquer le renforcement d'une proéminence par rapport à une autre. La taille seule du constituant n'influencerait pas directement la perception des proéminences. En revanche, la manière dont N et ADJ se répartissent le nombre de syllabes peut venir renforcer plus ou moins la force des proéminences. Ce résultat tend alors à confirmer qu'Al, mais aussi AF, ne joue pas un rôle purement rythmique dans l'organisation du phrasé prosodique.

#### 2.3.4.2 Proéminence initiale sur ADJ – Proéminence finale sur N2

Les résultats montrent un effet significatif du groupement (LR Chisq= 11,56, df=4, p=0,02), ainsi qu'un effet d'interaction entre le nombre de syllabes de N et ADJ (LR Chisq= 18,18, df= 4, p=0,001) (cf. Annexe N°16-G).

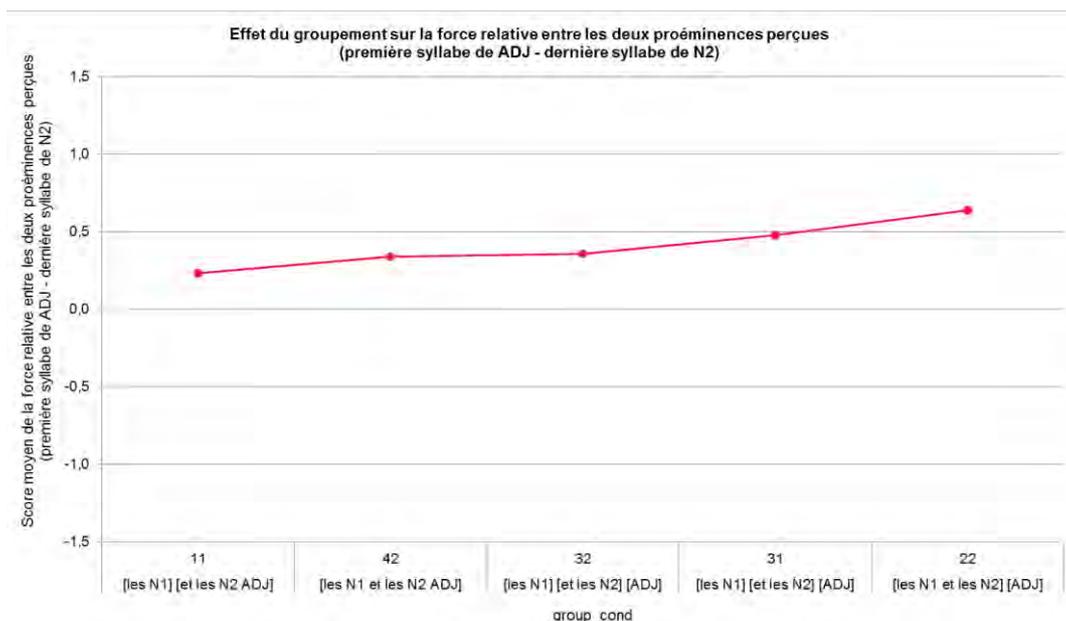
### Effet simple du groupement :

Les analyses post hoc de l'effet du groupement montrent une différence significative entre G11 et G42 d'un côté et G22 de l'autre. La valeur de force relative est en effet moins élevée en G11 et G42 (scores moyens respectifs : 0,23 et 0,34) qu'en G22 (score moyen : 0,64). Aucune différence n'est par ailleurs observée dans les autres groupements ( $p > 0,05$ ). Nous remarquons par ailleurs que toutes les valeurs de force relative sont ici positives. Ainsi, quelle que soit la catégorie de groupement, la prééminence sur la syllabe initiale sur ADJ serait perçue plus forte que la prééminence finale sur N2. Toutefois, au vu des résultats, la prééminence initiale sur ADJ serait perçue encore plus forte que la prééminence finale sur N2 en G22 par rapport à G11 et G42 (cf. Tableau N°17 et Figure N°69).

**Tableau 17 :** Résultats des comparaisons post hoc de l'effet du groupement sur la force relative entre les deux prééminences perçues sur N2 et ADJ (première syllabe de ADJ – dernière syllabe de N2)

Le tableau présente les moyennes des scores de prééminences interprétées comme des valeurs numériques ordonnées de 1 à 7 pour les valeurs « mean.class » et de -3 à 3 pour les valeurs « mean.class(-4) », calculées à partir des résultats de la distribution de probabilités estimée pour chaque groupement (cf. Annexe N°16-G). Les groupes partageant une même lettre dans la colonne ".group" ne présentent pas de différence significative.

FORCE RELATIVE ENTRE LES DEUX PROEMINENCES PERCUES: PREMIERE SYLLABE DE ADJ - DERNIERE SYLLABE DE N2			
Moyennes des scores de force relative entre les prééminences calculées à partir des résultats de la distribution de probabilités			
group_cond	mean.class	mean.class(-4)	.group
11	4,23	0,23	a
42	4,34	0,34	
32	4,36	0,36	ab
31	4,48	0,48	
22	4,64	0,64	b



**Figure 69 :** Illustration de l'effet du groupement sur la force relative entre les deux proéminences perçues sur N2 et ADJ (première syllabe de ADJ – dernière syllabe de N2) Les points reliés par des lignes correspondent aux valeurs « mean.class (-4) » du tableau N°17 indiquant les moyennes de force relative des proéminences perçues pour chaque groupement, interprétées comme des valeurs numériques ordonnées de -3 à 3. Pour faciliter la lecture des résultats, nous choisissons de réduire l'échelle des scores moyens de la force relative entre les deux proéminences aux valeurs -1,5 à 1,5.

Ce résultat confirmerait alors celui de l'effet simple du groupement sur la perception de la proéminence initiale sur ADJ, où le renforcement de la proéminence observé en G22 est significativement plus important que celui observé en G11 et G42. Ces résultats tendraient par ailleurs à confirmer la corrélation entre le renforcement du degré de proéminence initiale sur ADJ et le marquage du niveau de frontière située avant la syllabe accentuée. Nous rappelons en effet que G22 est marqué par le niveau de frontière le plus élevé (frontière d'ip), et G11 et G42 par le niveau de frontière le plus faible (frontière d'AP). G31 et G32, en revanche, se distinguent de G11 et G42 selon la « configuration syllabique » du syntagme, mais ne se différencient jamais de G22. Nous supposons dans ce contexte que G31 et G32 relèvent du même degré de frontière que G22 (frontière d'ip), avec cependant des degrés de frontière qui peuvent être perçus moins forts dans le cas de « configurations syllabiques » spécifiques. Ainsi, les résultats tendent à montrer que plus le niveau de frontière augmente, plus la différence entre AI sur ADJ et AF sur N2 serait marquée.

### Effet d'interaction entre le nombre de syllabes de N et le nombre de syllabes ADJ :

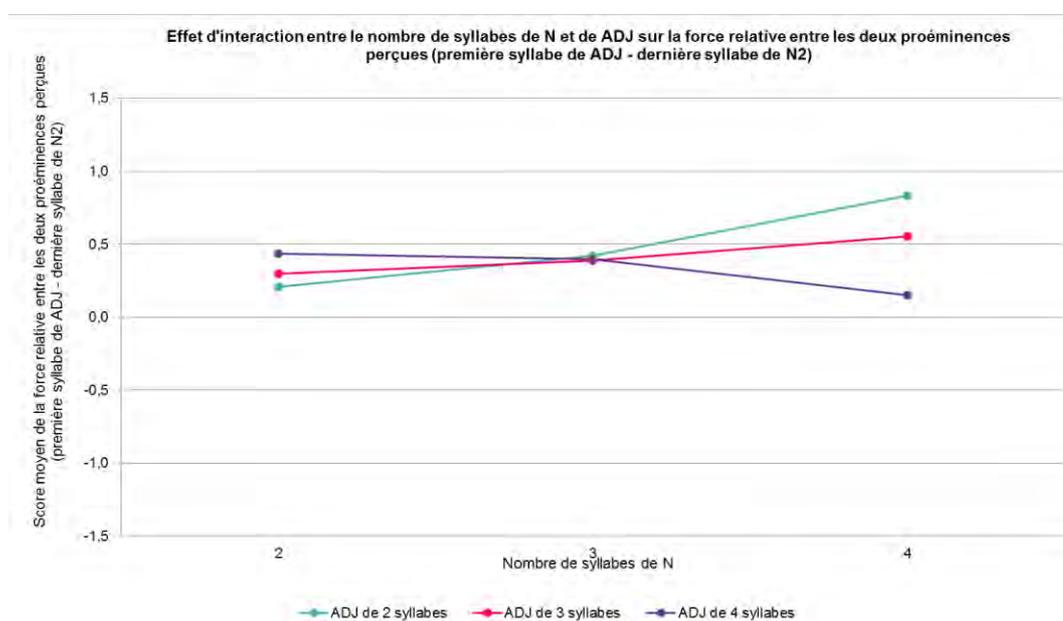
Concernant l'effet d'interaction entre le nombre de syllabes de N et ADJ, nous constatons des différences de force relative, selon la « configuration syllabique » de N2+ADJ. En regardant plus précisément les contrastes post hoc (cf. Tableau (A) et (B) N°18), nous voyons des différences entre les forces relatives principalement sur les ADJ de 2 syllabes. La force relative est en effet plus élevée sur les ADJ de 2 syllabes précédés d'un N de 4 syllabes (score moyen : 0,83) plutôt que d'un N de 2 et 3 syllabes (scores moyens respectifs : 0,21 à 0,42). Etant donné que toutes les valeurs de force relative sont ici positives, cela signifierait alors qu'AI sur ADJ serait perçu plus fort qu'AF sur N2 quelle que soit la « configuration syllabique » de N2+ADJ, avec cependant une différence entre les deux prééminences encore plus marquée lorsque ADJ de 2 syllabes est précédé d'un N composé de 4 syllabes plutôt que de 2 et 3 syllabes (cf. Figure N°70).

**Tableau 18 :** Résultats des comparaisons post hoc de l'effet d'interaction entre le nombre de syllabes de N et de ADJ sur la force relative entre les deux prééminences perçues sur N2 et ADJ (première syllabe de ADJ – dernière syllabe de N2)

A) Lecture des contrastes en colonne pour un même nombre de syllabes de ADJ et B) Lecture des contrastes en ligne pour un même nombre de syllabes de N. Dans chaque tableau, les moyennes des forces relatives des prééminences sont interprétées comme des valeurs numériques ordonnées de 1 à 7 pour les valeurs « mean.class » et de -3 à 3 pour les valeurs « mean.class(-4) », qui sont calculées à partir des résultats de la distribution de probabilités estimée pour chaque taille de N et de ADJ (cf. Annexe N°16-G). Les groupes partageant une même lettre dans la colonne ".group" ne présentent pas de différence significative.

FORCE RELATIVE ENTRE LES DEUX PROEINENCES PERCUES: PREMIERE SYLLABE DE ADJ - DERNIERE SYLLABE DE N2						
Moyennes des scores de force relative entre les prééminences calculées à partir des résultats de la distribution de probabilités						
A) Lecture des contrastes en colonne pour un même nombre de syllabes de ADJ						
Nombre de syllabes de N	Nombre de syllabes de ADJ					
	2		3		4	
	mean.class	.group	mean.class	.group	mean.class	.group
	mean.class(-4)		mean.class(-4)		mean.class(-4)	
2	4,21	a	4,30	a	4,43	ab
	0,21		0,30		0,43	
3	4,42	a	4,39	ab	4,40	a
	0,42		0,39		0,40	
4	4,83	b	4,55	ab	4,15	a
	0,83		0,55		0,15	
B) Lecture des contrastes en ligne pour un même nombre de syllabes de N						
Nombre de syllabes de N	Nombre de syllabes de ADJ					
	2		3		4	
	mean.class	.group	mean.class	.group	mean.class	.group
	mean.class(-4)		mean.class(-4)		mean.class(-4)	
2	4,21	a	4,30	a	4,43	ab
	0,21		0,30		0,43	
3	4,42	a	4,39	ab	4,40	a
	0,42		0,39		0,40	
4	4,83	b	4,55	ab	4,15	a
	0,83		0,55		0,15	

De la même manière, une différence de force relative est observée lorsque N est composé de 4 syllabes et suivi d'un ADJ de 2 ou 4 syllabes. La valeur de force relative est en effet plus élevée lorsque N de 4 syllabes est suivi d'un ADJ de 2 syllabes plutôt que de 4 syllabes (scores moyens respectifs : 0,83 contre 0,15). Etant donné que là encore toutes les valeurs de force relative sont positives, cela signifierait alors qu'AI sur ADJ serait perçu plus fort qu'AF sur N2 quelle que soit la « configuration syllabique » de N2+ADJ, avec cependant une différence entre les deux proéminences encore plus marquée lorsque N de 4 syllabes est suivi d'un ADJ composé de 2 syllabes plutôt que de 4 syllabes (cf. Figure N°70).



**Figure 70 :** Illustration de l'effet d'interaction entre le nombre de syllabes de N et de ADJ sur la force relative entre les deux proéminences perçues sur N2 et ADJ (première syllabe de ADJ – dernière syllabe de N2)

Les points reliés par des lignes correspondent aux valeurs « mean.class (-4) » du tableau N°18 indiquant les moyennes de force relative des proéminences perçues pour chaque groupement, interprétées comme des valeurs numériques ordonnées de -3 à 3. Pour faciliter la lecture des résultats, nous choisissons de réduire l'échelle des scores moyens de la force relative entre les deux proéminences aux valeurs -1,5 à 1,5.

L'ensemble de ces résultats semble correspondre à ceux obtenus dans l'effet d'interaction du nombre de syllabes de N et de ADJ sur la perception de la proéminence initiale sur ADJ, où un renforcement de la proéminence est observé sur un ADJ de 2 syllabes précédé d'un N de 4 syllabes. Nous constatons donc là encore que l'augmentation du nombre de syllabes sur ADJ ne conduirait pas au renforcement de la proéminence initiale. En effet, il semblerait

que ce renforcement soit plutôt dû à des « configurations syllabiques » très spécifiques. Comme nous l'avons déjà mentionné, la taille seule du constituant n'influencerait pas directement la perception des proéminences, tandis que la manière dont N et ADJ se répartissent le nombre de syllabes pourrait davantage impacter le renforcement ou non des proéminences.

### 2.3.4.3 *Proéminence initiale sur ADJ – Proéminence initiale sur N2*

Les résultats montrent un effet significatif du groupement (LR Chisq= 21,46, df=4,  $p < 0,001$ ) ainsi qu'un effet d'interaction en le nombre de syllabes de N et ADJ (LR Chisq= 20,88, df= 4,  $p < 0,001$ ) (cf. Annexe N°16-H).

#### Effet simple du groupement :

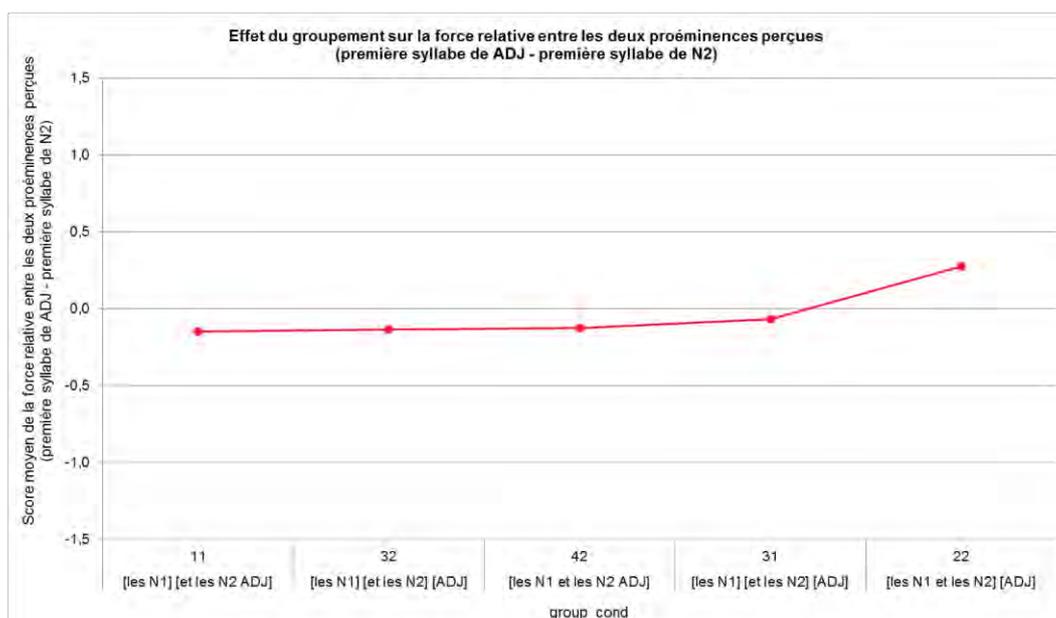
Les résultats de l'analyse post hoc de l'effet du groupement montrent une différence significative entre les deux proéminences en G22 par rapport à tous les autres groupements. La valeur de force relative est en effet plus élevée en G22 (score moyen : 0,28) qu'en G11, G32, G42 et G31 (scores moyens : entre -0,07 à -0,15). Aucune différence n'est par ailleurs observée entre les autres groupements ( $p > 0,05$ ) (cf. Tableau N°19 et Figure N°71).

**Tableau 19** : Résultats des comparaisons post hoc de l'effet du groupement sur la force relative entre les deux proéminences perçues sur N2 et ADJ (première syllabe de ADJ – première syllabe de N2)

Le tableau présente les moyennes des scores de proéminences interprétées comme des valeurs numériques ordonnées de 1 à 7 pour les valeurs « mean.class » et de -3 à 3 pour les valeurs « mean.class(-4) », calculées à partir des résultats de la distribution de probabilités estimée pour chaque groupement (cf. Annexe N°16-H). Les groupes partageant une même lettre dans la colonne ".group" ne présentent pas de différence significative.

FORCE RELATIVE ENTRE LES DEUX PROEMINENCES PERCUES: PREMIERE SYLLABE DE ADJ - PREMIERE SYLLABE DE N2			
Moyennes des scores de force relative entre les proéminences calculées à partir des résultats de la distribution de probabilités			
group_cond	mean.class	mean.class(-4)	.group
11	3,85	-0,15	a
32	3,86	-0,14	
42	3,87	-0,13	
31	3,93	-0,07	
22	4,28	0,28	b

Il semblerait donc, étant donné la valeur de force relative positive en G22, qu'Al sur ADJ soit perçu plus fort qu'Al sur N2, et étant donné la valeur de force relative négative dans tous les autres groupements, qu'Al sur ADJ soit perçu moins fort qu'Al sur N2.



**Figure 71 :** Illustration de l'effet du groupement sur la force relative entre les deux proéminences perçues sur N2 et ADJ (première syllabe de ADJ – première syllabe de N2) Les points reliés par des lignes correspondent aux valeurs « mean.class (-4) » du tableau N°19 indiquant les moyennes de force relative des proéminences perçues pour chaque groupement, interprétées comme des valeurs numériques ordonnées de -3 à 3. Pour faciliter la lecture des résultats, nous choisissons de réduire l'échelle des scores moyens de la force relative entre les deux proéminences aux valeurs -1,5 à 1,5.

Ce résultat confirmerait alors celui de l'effet simple du groupement sur la perception de la proéminence initiale sur ADJ, où le renforcement de la proéminence observé en G22 est significativement plus important que celui observé en G11 et G42. Cependant, les résultats de la présente étude montrent une différence entre G22 et tous les autres groupements. Al sur ADJ serait alors majoritairement perçu moins fort qu'Al sur N2, excepté en G22 où Al sur ADJ est perçu plus fort qu'Al sur N2. Il est par ailleurs intéressant de remarquer qu'une même force relative peut être observée entre deux groupements où les relations hiérarchiques entre les frontières après N1 et N2 sont différentes. En effet, G31 qui est marqué par une frontière d'IP après N1 et d'ip après N2, et G42 qui est marqué par une frontière de PW après N1 et d'AP après N2, présentent tous les deux la même force relative des proéminences, où Al sur N2 est perçu plus fort qu'Al sur ADJ. Il semblerait donc que la force relative entre les deux proéminences ne varie pas strictement en fonction du niveau de constituance. Par ailleurs, le fait que la proéminence initiale sur N2 après une frontière de PW soit perçue plus forte que la proéminence initiale sur ADJ après une frontière d'AP

semble indiquer qu'un AI est perçu dès le niveau le plus bas dans la hiérarchie, à savoir le PW.

### Effet d'interaction entre le nombre de syllabes de N et et le nombre de syllabes ADJ :

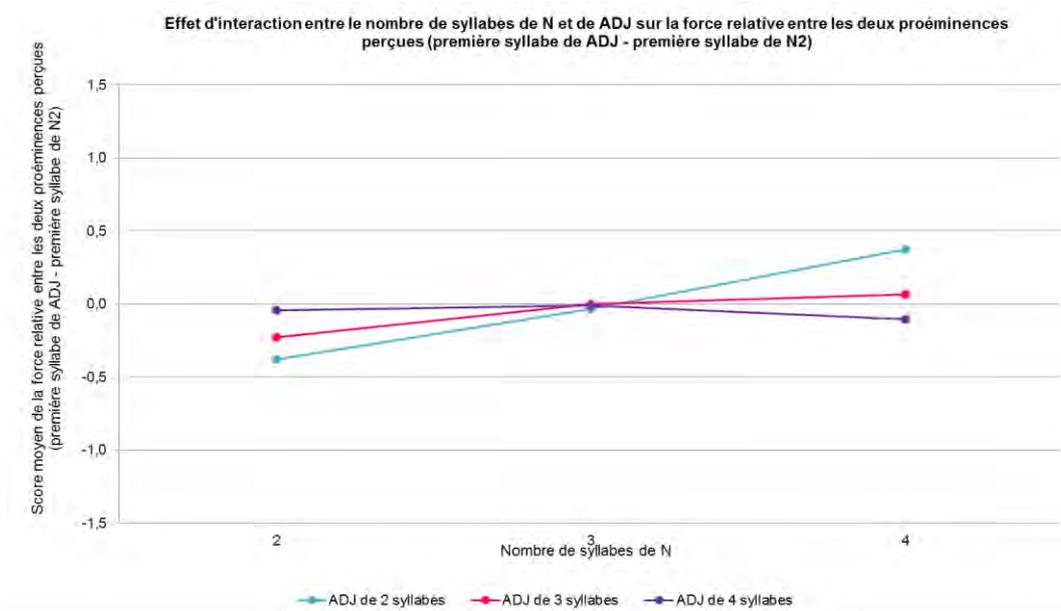
Concernant l'effet d'interaction entre le nombre de syllabes de N et ADJ, nous constatons des différences de force relative, selon la « configuration syllabique » de N2+ADJ. En regardant plus précisément les contrastes post hoc (cf. Tableau (A) et (B) N°20), nous voyons des différences entre les forces relatives principalement sur les ADJ de 2 syllabes. La force relative est en effet plus élevée sur les ADJ de 2 syllabes précédés d'un N de 4 syllabes (score moyen : 0,37) plutôt que d'un N de 2 et 3 syllabes (scores moyens respectifs : -0,38 et -0,03). Cela signifierait alors, étant donné la valeur de force relative positive lorsque N de 4 syllabes est suivi d'un ADJ de 2 syllabes, qu'AI sur ADJ serait perçu plus fort qu'AI sur N2, et étant donné la valeur de force relative négative lorsque N de 2 et 3 syllabes est suivi d'un ADJ de 2 syllabes, qu'AI sur ADJ serait perçu moins fort qu'AI sur N2.

**Tableau 20 :** Résultats des comparaisons post hoc de l'effet d'interaction entre le nombre de syllabes de N et de ADJ sur la force relative entre les deux proéminences perçues sur N2 et ADJ (première syllabe de ADJ – première syllabe de N2)

A) Lecture des contrastes en colonne pour un même nombre de syllabes de ADJ et B) Lecture des contrastes en ligne pour un même nombre de syllabes de N. Dans chaque tableau, les moyennes des forces relatives des proéminences sont interprétées comme des valeurs numériques ordonnées de 1 à 7 pour les valeurs « mean.class » et de -3 à 3 pour les valeurs « mean.class(-4) », calculées à partir des résultats de la distribution de probabilités estimée pour chaque taille de N et de ADJ (cf. Annexe N°16-H). Les groupes partageant une même lettre dans la colonne ".group" ne présentent pas de différence significative.

FORCE RELATIVE ENTRE LES DEUX PROÉMINENCES PERÇUES: PREMIERE SYLLABE DE ADJ - PREMIERE SYLLABE DE N2						
Moyennes des scores de force relative entre les proéminences calculées à partir des résultats de la distribution de probabilités						
A) Lecture des contrastes en colonne pour un même nombre de syllabes de ADJ						
Nombre de syllabes de N	Nombre de syllabes de ADJ					
	2		3		4	
	mean.class mean.class(-4)	.group	mean.class mean.class(-4)	.group	mean.class mean.class(-4)	.group
2	3,62	a	3,77	ab	3,96	ab
	-0,38		-0,23		-0,04	
3	3,97	b	4,00	abc	3,99	b
	-0,03		0,00		-0,01	
4	4,37	c	4,07	bc	3,89	ab
	0,37		0,07		-0,11	
B) Lecture des contrastes en ligne pour un même nombre de syllabes de N						
Nombre de syllabes de N	Nombre de syllabes de ADJ					
	2		3		4	
	mean.class mean.class(-4)	.group	mean.class mean.class(-4)	.group	mean.class mean.class(-4)	.group
2	3,62	a	3,77	ab	3,96	ab
	-0,38		-0,23		-0,04	
3	3,97	b	4,00	abc	3,99	b
	-0,03		0,00		-0,01	
4	4,37	c	4,07	bc	3,89	ab
	0,37		0,07		-0,11	

De la même manière, une différence de force relative est observée lorsque N est composé de 4 syllabes et suivi d'un ADJ de 2 ou 4 syllabes. La valeur de force relative est en effet plus élevée lorsque N de 4 syllabes est suivi d'un ADJ de 2 syllabes (score moyen : 0,37) plutôt que de 4 syllabes (score moyen : -0,11). Il semblerait donc, étant donné la valeur de force relative positive lorsque N de 4 syllabes est suivi d'un ADJ de 2 syllabes, qu'AI sur ADJ soit perçu plus fort qu'AI sur N2, et étant donné la valeur de force relative négative lorsque N de 4 syllabes est suivi d'un ADJ de 4 syllabes, qu'AI sur ADJ soit perçu moins fort qu'AI sur N2 (cf. Figure N°72).



**Figure 72 :** Illustration de l'effet d'interaction entre le nombre de syllabes de N et de ADJ sur la force relative entre les deux proéminences perçues sur N2 et ADJ (première syllabe de ADJ – première syllabe de N2)

Les points reliés par des lignes correspondent aux valeurs « mean.class (-4) » du tableau N°20 indiquant les moyennes de force relative des proéminences perçues pour chaque groupement, interprétées comme des valeurs numériques ordonnées de -3 à 3. Pour faciliter la lecture des résultats, nous choisissons de réduire l'échelle des scores moyens de la force relative entre les deux proéminences aux valeurs -1,5 à 1,5.

Au vu des résultats, il semble difficile de déterminer précisément des règles rythmiques pouvant expliquer qu'une proéminence initiale est perçue plus forte que l'autre proéminence initiale. Par ailleurs, nous pourrions supposer que ces règles rythmiques soient également liées à la constituance prosodique. Toutefois, aucun des résultats présentés ici ne montrent un effet d'interaction entre le groupement et le nombre de syllabes des constituants sur la perception des proéminences. Ainsi, comme pour les résultats précédents, le renforcement

de la proéminence initiale serait donc dû à des « configurations syllabiques » très spécifiques, et ne correspondrait pas des contraintes purement rythmiques.

Sur la base de l'ensemble des résultats, il semblerait qu'AI sur ADJ soit majoritairement perçu aussi fort ou plus fort qu'AF sur N2 et ADJ. AI sur N2 serait en revanche perçu majoritairement plus fort qu'AI sur ADJ. Par ailleurs, les résultats tendent à montrer que la force relative entre les proéminences ne varie pas strictement en fonction du niveau de constituance. Seule la force relative entre AI sur ADJ et AF sur N2 semble indiquer une corrélation entre la force relative et le niveau de constituance. Plus le niveau de frontière après N2 augmente, plus la différence est marquée entre les deux proéminences. Cela semble alors confirmer le rôle pertinent de l'AI sur ADJ dans le marquage de la structure prosodique. Il apparaît enfin qu'AI comme AF sont réalisés sur N2 et ADJ, quelle que soit la catégorie de groupement ou la taille du constituant. En effet, bien qu'il y ait systématiquement un effet d'interaction entre le nombre de syllabes de N et ADJ sur la perception des proéminences, il semblerait en revanche que l'influence de la taille de N et ADJ soit essentiellement due à la manière dont les mots se répartissent le nombre de syllabes qu'à la taille seule de N2+ADJ. Il apparaît donc difficile d'établir des règles rythmiques, étant donné que le renforcement d'une proéminence par rapport à une autre est lié à des « configurations syllabiques » très spécifiques. Il est par ailleurs intéressant de signaler qu'AI serait perçu dès les constituants composés de 2 syllabes, ce qui tendrait par ailleurs à confirmer le rôle plus structurel que purement rythmique de l'AI.

### 2.4 Discussion

La force des proéminences à l'initiale et à la finale des mots lexicaux semble globalement peu influencée par le groupement, et également peu influencée par la taille du constituant. Il est par ailleurs intéressant de remarquer qu'aucun effet d'interaction entre le groupement et la taille du constituant n'a été observé. Cela suppose alors que l'effet du groupement sur la perception des proéminences soit indépendant de la taille du constituant, et de la même manière, que l'effet du nombre de syllabes du constituant sur la perception des proéminences est indépendant de la catégorie de groupement.

En ce qui concerne les résultats sur l'effet de la taille du constituant, nous remarquons presque systématiquement une interaction significative entre la taille de N et ADJ. A travers les analyses post hoc de cet effet d'interaction, nous constatons des effets très disparates, auxquels il apparaît difficile de rattacher des interprétations rythmiques, pouvant expliquer pourquoi une proéminence est perçue plus ou moins forte qu'une autre

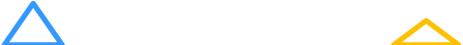
proéminence sur le même site (sur le plan paradigmatique), ou sur un autre site (sur le plan syntagmatique). Il est toutefois intéressant de constater à travers ces résultats que les proéminences initiales semblent avoir été perçues dès les constituants composés de 2 syllabes. Certains résultats tendent d'ailleurs à montrer que les AI sur les ADJ de 2 syllabes seraient perçus plus forts que les AI sur les ADJ de 3 et 4 syllabes, notamment lorsqu'un N de 4 syllabes précède ADJ. Il semblerait d'un point de vue rythmique que le renforcement de la proéminence initiale soit davantage attendu sur les ADJ de plus 2 syllabes, dans la mesure où le constituant N qui précède est déjà composé d'un grand nombre de syllabes. En admettant effectivement un rôle rythmique de l'AI, plus le constituant est long plus la probabilité d'occurrence de l'AI augmente. Nos résultats permettraient alors de remettre en doute le rôle purement rythmique de l'AI (cf. notion « *d'ictus mélodique* » de Rossi, 1980), dans la mesure où cet accent n'apparaîtrait pas seulement sur les longs constituants. Plus intéressant encore, nous rappelons que les tests du Khi deux ne montrent aucune relation de dépendance entre les groupements ou la taille du constituant et les syllabes intermédiaires des mots lexicaux. Seules les premières et dernières syllabes des mots de contenu peuvent être influencées par la catégorie de groupement et par le nombre de syllabes du constituant, ce qui tend alors à faire penser que les proéminences à l'initiale des mots sont uniquement perçues sur la première syllabe des mots lexicaux. Ce résultat contrecarrerait ainsi le postulat selon lequel la position de l'AI serait variable dans le mot, et remettrait également en doute l'idée d'un « *marqueur flou* » de frontière en français (Jun & Fougeron, 2000 ; 2002 ; Welby, 2003).

Dans le cadre de notre discussion, nous choisissons toutefois de nous intéresser davantage à l'effet du groupement qu'à l'effet de la taille du constituant. La « configuration syllabique », ou plutôt la manière dont les mots lexicaux se répartissent le nombre de syllabes dans le syntagme, semble effectivement plus pertinente dans la réalisation des degrés de proéminence que la taille seule du syntagme. Il semblerait cependant que notre matériel linguistique ne nous permette pas de tester systématiquement cette hypothèse, dans la mesure où seuls N2 et ADJ varient en nombre de syllabes. Il serait alors intéressant par la suite de proposer une étude plus spécifiquement destinée à tester l'effet des « configurations syllabiques » sur la perception des proéminences, sur la base d'un corpus qui manipule la taille de tous les mots lexicaux du syntagme, et qui propose plusieurs « configurations syllabiques » possibles pour une même taille totale du syntagme.

Afin de résumer l'ensemble des résultats de l'effet du groupement sur la force des proéminences, nous proposons la figure récapitulative ci-dessous (cf. Figure N°73). Pour constituer cette figure, nous avons tout d'abord repris les scores moyens (selon les cas, les valeurs « mean.class (-1) » ou « mean.class (-4) »), associées à chaque groupement, sur les

proéminences de toutes les syllabes révélant un effet significatif du groupement. Dans notre cas, seules les proéminences sur la syllabe finale de N1 et initiale de N2 montrent une influence du groupement. Nous avons ensuite transformé les valeurs chiffrées, associées aux proéminences perçues, en une forme triangulaire représentant la réalisation accentuelle, dont la hauteur (mesuré en centimètre « cm »), uniquement, varie proportionnellement en fonction de la valeur chiffrée (soustraction de la constante 1 à chaque valeur : 1,63 = 63 cm; 1,45 = 45 cm ; etc.). Quant aux résultats sur la force relative entre les proéminences perçues, nous n'avons effectué aucune transformation puisqu'il s'agit seulement d'indiquer si la proéminence est perçue plus ou moins forte qu'une autre proéminence (<, > ou =). Les symboles en rouge signifient que la proéminence initiale est perçue plus forte que la proéminence finale, tandis que les symboles en vert signifie que la proéminence finale est perçue plus forte que la proéminence initiale. De la même manière, pour représenter la différence de force relative entre deux proéminences du même type (entre deux AI ou entre deux AF), nous avons utilisé un code couleur pour signaler les proéminences mises en comparaison, ainsi qu'une distinction dans la taille de la police pour montrer la proéminence qui est perçue plus ou moins forte que l'autre. Dans la figure, les comparaisons sont faites entre les proéminences initiales perçues sur N2 et ADJ, et entre les proéminences finales perçues sur N1 et N2.

Nous constatons à travers la figure récapitulative que deux proéminences seulement montrent des variations significatives en fonction des groupements : la proéminence perçue sur la syllabe finale de N1 et sur la syllabe initiale de ADJ. Effectivement, la proéminence sur la syllabe finale de N1 est perçue plus forte en G11 qu'en G32 et G42, tandis qu'aucune différence n'est observée ni entre G11, G31 et G22, ni entre G32, G42, G31 et G22. Quant à la proéminence sur la syllabe initiale de ADJ, nous voyons qu'elle est perçue plus forte en G22 qu'en G11 et G42, tandis qu'aucune différence n'est observée ni entre G22, G31 et G32, ni entre G11, G42, G31 et G32. L'absence de différence entre les groupements est représentée dans le tableau via les formes triangulaires bicolores, reprenant les deux couleurs des groupements s'opposant significativement. Deux différents degrés de proéminence ont donc été perçus sur la syllabe finale de N1 (le plus élevé en G11 et le plus faible en G32 et G42) et deux autres degrés de proéminence sur la syllabe initiale de ADJ (le plus élevé en G22 et le plus faible en G11 et G42), avec néanmoins des nuances de force de proéminence qui se trouvent à l'intermédiaire des degrés de proéminence principaux: en G31 et G22 pour AF sur N1, et en G31 et G32 pour AI sur ADJ.

G11	[ les N1 ] > [ et les N2 ] ADJ AF
	Al < AF > AI AF < AI < AF
	
G31	[ les N1 ] = [ et les N2 ] [ ADJ ] AF
	Al = AF = AI AF < AI = AF
	
G32	[ les N1 ] < [ et les N2 ] [ ADJ ] AF
	Al > AF < AI AF < AI = AF
	
G42	[ les N1 ] = et les N2 ADJ AF
	Al > AF = AI AF < AI = AF
	
G22	[ les N1 ] = et les N2 ] [ ADJ ] AF
	Al = AF = AI AF < AI > AF
	

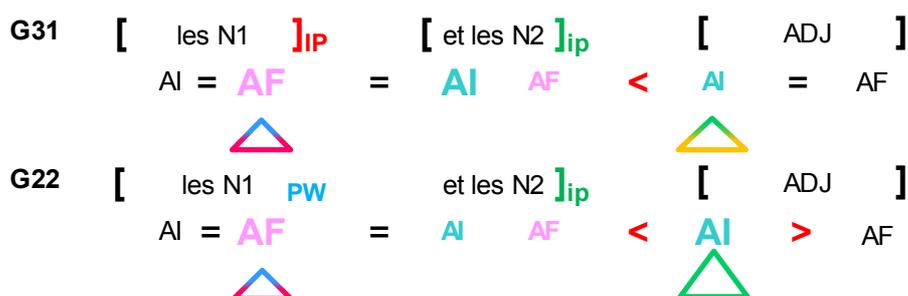
**Figure 73 :** Récapitulatif des résultats de l'effet du groupement sur les scores de prééminences perçues à l'initiale et à la finale des mots lexicaux (N1, N2 et ADJ)

Les formes triangulaires permettent de comparer la force entre les prééminences sur un même site entre les différents groupements (axe vertical). A chaque couleur de triangle correspond une force de prééminence spécifique, et quand les triangles sont bicolores, cela indique que la force de prééminence ne se distingue pas significativement des deux autres forces de prééminence. Les symboles >, < ou = permettent de comparer la force entre deux prééminences (initiale vs. finale) adjacentes (axe horizontal). Les symboles verts indiquent qu'AF est perçue plus forte qu'AI, et les symboles rouges indiquent qu'AI est perçue plus forte qu'AF. Les variations de taille de police permettent de comparer la force entre deux prééminences du même type : AI vs. AI en bleu, ou AF vs. AF en rose (axe horizontal).

Il est par ailleurs intéressant de remarquer que les principales différences entre les forces de prééminences sont observées en G11 pour AF sur N1 et en G22 pour AI sur ADJ, qui correspondent aux groupements « canoniques » représentant par défaut les deux conditions syntaxiques. Nous rappelons en effet que G1 (G11) n'a quasiment pas été associé à la Condition 2 et que G2 (G22) n'a pas été perçue pour la Condition 1. Nous avons déjà observé cette tendance dans l'analyse de la force des frontières, et nous supposons à ce moment-là que les auditeurs avaient choisi massivement ces deux groupements pour

représenter chacune des deux conditions syntaxiques, étant donné l'ambiguïté peut-être moins grande quant à la hiérarchie établie entre les deux frontières en G11 et G22, par rapport aux autres catégories de groupements. Il semblerait donc que les proéminences finales sur N1 et initiales sur N2 facilitent également le choix de ces deux groupements pour représenter les deux conditions syntaxiques.

En mettant en relation ces résultats avec nos hypothèses de départ, il semblerait que la force de la proéminence finale perçue sur N1 ne soit pas corrélée, comme prédit initialement, avec l'augmentation du niveau de frontière auquel la proéminence est associée. Les résultats tendraient plutôt à admettre que la force de l'AF sur N1 ne varierait pas strictement en fonction du niveau de constituance. Une même force de proéminence peut en effet être perçue à des niveaux de frontière différents. La proéminence est perçue aussi forte en G31 qu'en G22, alors que les niveaux de frontière marqués dans ces deux groupements relèveraient respectivement du niveau de l'IP et du PW :

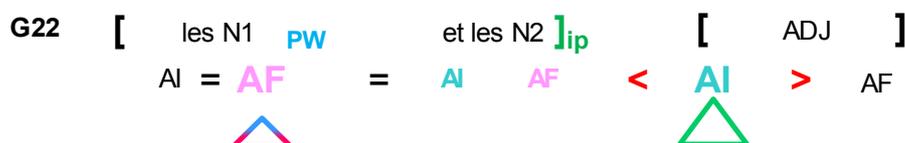


La proéminence finale sur N1 serait donc perçue toujours avec la même force dès le niveau du PW, et n'augmenterait pas strictement en fonction du niveau de frontière. Le renforcement du degré de proéminence observé en G11 pourrait plutôt être dû selon nous aux caractéristiques individuelles des locuteurs. Nous savons en effet que les syntagmes réalisés en Condition 1, donc en G11 et G31, sont caractérisés par une pause relativement longue après N1. Toutefois, les quatre locuteurs produisant les syntagmes ne réalisent pas forcément la même durée de pause. Il se pourrait donc que les syntagmes perçus en G11 soient associés à des pauses plus longues que les syntagmes perçus en G31. Cela expliquerait alors la perception d'une proéminence sur la syllabe finale de N1 qui ne soit pas significativement plus forte en G31 que dans les autres groupements, malgré la présence d'un même niveau de frontière en G11 et G31 (frontières d'IP). Comme nous l'avons déjà mentionné, cette hypothèse devra être explorée dans une future étude, en contrôlant directement les réalisations prosodiques pour chacun des locuteurs.

Ainsi, bien que ce résultat ne corresponde pas tout à fait à nos hypothèses de départ, il semblerait toutefois que le fait de percevoir une proéminence finale dès le niveau du PW,

qui ne varie pas en fonction du niveau de constituance, soit en quelque sorte la preuve d'une certaine dissociation entre les phénomènes de proéminences et de frontières. En effet, nos résultats tendent à montrer que l'accent final en français ne s'effacerait pas perceptivement en frontière prosodique majeure (Jun & Fougeron, 2000 ; 2002 ; Rossi, 1980). Il semblerait, au contraire, qu'il y ait toujours une trace de l'accent final de mot perçue, voire une « *survie de l'accent de mot* » dans le groupe (Delattre, 1966b). Ces résultats confirment par ailleurs ceux obtenus dans l'étude en perception d'Astésano et al. (2012 ; 2013), où les auteurs montrent que les syllabes marquées par un AF sont perçues comme métriquement plus fortes que les syllabes inaccentuées à l'intérieur de l'AP. De la même manière, les auteurs parlent de « *dissociation partielle* », dans la mesure où un même score de proéminence peut correspondre à différents niveaux de frontière, à l'exception de la frontière d'ip (qui selon nos résultats serait plutôt une frontière d'IP) où un renforcement de l'accent final est observé. Les auteurs considèrent cependant que ce renforcement serait davantage dû à la réalisation d'une longue pause après N1.

Nous supposons également cet effet dans nos résultats et nous affinons par ailleurs cette idée d'absence de corrélation entre le renforcement de la proéminence finale et le marquage de la frontière d'IP. En effet, en ajoutant l'information de groupement à notre analyse perceptive, nous pouvons rendre compte, à l'inverse d'Astésano et al. (2012 ; 2013), de la possibilité d'avoir un même score de proéminence en frontière de PW et d'IP. Il s'agirait donc plutôt dans nos résultats d'une « dissociation totale » entre la perception de la proéminence finale et les niveaux de frontière. Par ailleurs, à travers l'analyse de la force relative entre AF sur N2 et AF sur N1, nous remarquons qu'AF sur N2 est perçu moins fort qu'AF sur N1, alors que d'un point de vue de la constituance, les niveaux de frontière après N2 sont parfois plus élevés que les niveaux de frontière après N1. En G22, par exemple, bien que la frontière après N1 relève du niveau du PW et que la frontière après N2 relève du niveau de l'ip, AF sur N2 est perçu moins fort qu'AF sur N1. Cela tendrait alors à confirmer une fois de plus qu'AF sur N1 est perçu dès le niveau du PW.



A l'inverse, la force de la proéminence initiale sur ADJ varierait davantage en fonction du niveau de frontière. Cela confirmerait notre hypothèse de départ quant à la corrélation entre le renforcement de la proéminence initiale sur ADJ et le marquage du niveau de frontière à laquelle la proéminence est associée. Les résultats montrent effectivement que la proéminence initiale sur ADJ est perçue avec le degré de proéminence le plus élevé en G22,

qui correspond également au niveau de frontière le plus élevé (frontière d'ip), et avec le degré de proéminence le plus faible en G11 et G42, qui correspond également au niveau de frontière le plus faible (frontière d'AP). La force de la proéminence initiale sur ADJ en G31 et G32, en revanche, ne se distingue ni de celle perçue en G11 et G42, ni de celle perçue en G22. Ce résultat pourrait alors également faire écho à celui obtenu dans l'analyse de la force des frontières, où G31 et G32 se différencient de G11 et G42 dans le cas de « configurations syllabiques » spécifiques, et ne se distinguent jamais de G22. Dans ce contexte, nous avons supposé que le niveau de frontière en G31 et G32 relève du même niveau que G22 (frontière d'ip), avec cependant des degrés de frontière qui peuvent être perçus moins forts dans le cas de « configurations syllabiques » spécifiques. Ainsi, le renforcement du degré de frontière en G22 pourrait être lié à la perception d'une proéminence sur la syllabe initiale de ADJ également plus forte. De la même manière, G31 et G32 pourraient être marqués par un renforcement plus ou moins important du degré de proéminence dans le cas de « configurations syllabiques » spécifiques, impliquant alors un degré de frontière également plus ou moins important. AI pourrait alors jouer un rôle pertinent dans le marquage de la structure prosodique.

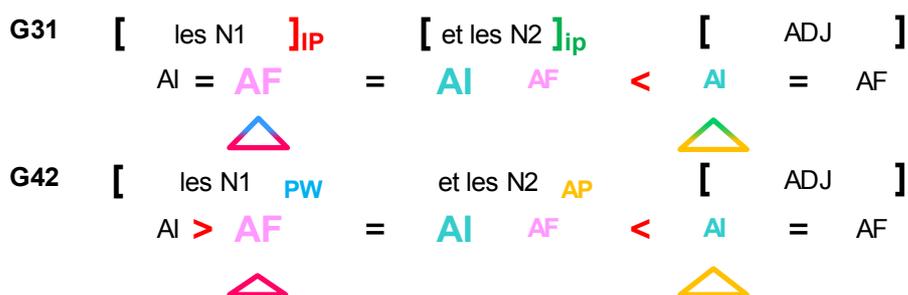
<b>G11</b>	[ les N1 ] <sub>IP</sub>	[ et les N2 ] <sub>AP</sub>	[ ADJ ]
	AI < AF	> AI AF	< AI < AF
			
<b>G31</b>	[ les N1 ] <sub>IP</sub>	[ et les N2 ] <sub>ip</sub>	[ ADJ ]
	AI = AF	= AI AF	< AI = AF
			
<b>G32</b>	[ les N1 ] <sub>AP</sub>	[ et les N2 ] <sub>ip</sub>	[ ADJ ]
	AI > AF	< AI AF	< AI = AF
			
<b>G42</b>	[ les N1 ] <sub>PW</sub>	[ et les N2 ] <sub>AP</sub>	[ ADJ ]
	AI > AF	= AI AF	< AI = AF
			
<b>G22</b>	[ les N1 ] <sub>PW</sub>	[ et les N2 ] <sub>ip</sub>	[ ADJ ]
	AI = AF	= AI AF	< AI > AF
			

Ces résultats confirmeraient ceux obtenus dans une étude en production (Astésano et *al.*, 2007) antérieure, où les auteurs montrent que la proéminence initiale marque la structure prosodique de manière plus privilégiée que la proéminence finale. Toutefois, alors qu'Astésano et *al.* (2007) rendent compte du rôle structurel de l'AI plus spécifiquement à des

niveaux inférieurs de la hiérarchie prosodique, proches du mot lexical, nos résultats ne semblent pas, en revanche, montrer de niveaux de frontière plus favorisés par l'AI que d'autres. De la même manière, Astésano et *al.* (2012 ; 2013) montrent qu'AI est perçu plus fort qu'AF lorsque le lexème se trouve en frontière d'AP, et plus faible qu'AF lorsque le lexème se trouve en frontière d'IP. En revanche, pour les frontières d'ip ou inférieures à l'AP, proche du Mot Lexical, la participation de l'AI et de l'AF semble être équivalente.

Notre analyse de la force relative des proéminences ne semble pas confirmer ces résultats. En effet, nos analyses tendent à montrer que la force relative entre les proéminences ne varie pas strictement en fonction du niveau de constituance. Seule la force relative entre AI sur ADJ et AF sur N2 semble indiquer une corrélation avec le niveau de constituance. Plus le niveau de frontière après N2 augmente, plus AI sur ADJ est perçu plus fort qu'AF sur N2. Nous constatons donc là encore le rôle pertinent de l'AI sur ADJ dans le marquage de la structure prosodique. Nous attendions également une corrélation entre le marquage des niveaux de frontière et le renforcement de la proéminence initiale sur N2. Toutefois, nos résultats n'indiquent aucun effet significatif du groupement sur la perception de cette proéminence. Nous avons supposé que l'absence de relation entre AI sur N2 et le groupement pouvait être due à la présence des mots grammaticaux précédant N2. En effet, en l'absence d'indices (les mots grammaticaux) de début potentiel de mot (Christophe et *al.*, 1997), comme pour ADJ, la proéminence initiale serait alors renforcée pour aider au marquage de la structure prosodique.

Par ailleurs, à travers l'analyse de la force relative entre AI sur ADJ et AI sur N2, il apparaît qu'une proéminence initiale est perçue dès le niveau le plus bas dans la hiérarchie, à savoir le PW. En effet, nous constatons qu'il est possible de percevoir une même force relative dans deux groupements marqués par des relations hiérarchiques différentes entre la frontière après N1 et N2. Par exemple, G31 qui est marqué par une frontière d'IP après N1 et d'ip après N2, et G42 qui est marqué par une frontière de PW après N1 et d'AP après N2, présentent tous les deux la même force relative des proéminences, où AI sur N2 est perçu plus fort qu'AI sur ADJ.



Ainsi, AI serait perçu dès le niveau du PW, et dans le cas des mots lexicaux qui ne sont pas précédés par un ou plusieurs mots grammaticaux, comme ADJ, AI aurait alors une double fonction. Il marquerait non seulement le début du mot lexical, mais la force de l'AI varierait également en fonction du niveau de constituance. Ce résultat permettrait par ailleurs de questionner le statut secondaire de l'AI en français. En effet, au-delà de percevoir un AI sur tous les mots lexicaux dès les constituants composés de 2 syllabes, il s'avère également dans nos résultats qu'AI est majoritairement perçu aussi fort ou plus fort qu'AF, quel que soit le groupement, et par conséquent quel que soit le niveau de frontière. Il semblerait donc que les PW dans nos syntagmes soient tous marqués par un AI et un AF, avec par ailleurs un marquage initial aussi fort voire plus fort métriquement que le marquage final. Ce résultat tendrait ainsi à confirmer l'idée, proposée par Astésano (2017), d'un marquage bipolaire de même force métrique au niveau du mot lexical (ou Mot Prosodique), en structure de surface. A l'inverse de Di Cristo (1999a ; 1999b ; 2000), le marquage bipolaire ne serait pas envisagé seulement au niveau sous-jacent mais serait également effectif en structure de surface, dès le niveau le plus bas dans la hiérarchie prosodique.

### **3. CORRELATS ACOUSTIQUES DES FRONTIÈRES ET DES PROÉMINENCES PERÇUES**

Sur la base de ces résultats perceptifs, nous proposons maintenant de mettre en correspondance les événements perçus avec leur réalisation physique ou acoustique. Plus précisément, nous choisissons de tester l'effet des variations acoustiques sur la perception des frontières et des proéminences. Comme nous avons pu l'observer dans la plupart des travaux sur la réalité psycho-cognitive de la constituance prosodique (cf. 1.1.1 du Chapitre 2), l'influence des indices acoustiques sur la perception des événements prosodiques est plus forte que celle des indices syntactico-sémantiques (Abney, 1991, 1992 ; Cole et *al.*, 2014 ; Gee & Grosjean, 1983 ; Simon & Christodoulides, 2016a ; 2016b). Les auteurs rendent compte notamment de la pertinence de l'allongement vocalique et de la présence d'une pause en frontière prosodique majeure (Cole et *al.*, 2014 ; Delais-Roussarie & Feldhausen, 2014 ; Simon & Christodoulides, 2016a ; 2016b ; Smith, 2009 ; Swerts, 1998). Certains auteurs s'intéressent également aux contours intonatifs et constatent que les contours tonals descendants renforcent la perception du niveau de frontière, tandis que les frontières plus faibles sont associées aux contours tonals montants (Hermès & Rump, 1994 ; Méndez, en cours ; Méndez & Astésano, 2017 ; Portes, 2002).

Dans le cadre de notre travail, nous proposons donc de tester l'influence de la durée de la syllabe précédant la frontière, et de la hauteur de  $f_0$  associée aux syllabes autour des frontières sur la perception de la force des frontières et des proéminences. Nous choisissons par ailleurs de ne pas tester l'effet de la durée de la pause, étant donné que les pauses réalisées dans notre corpus correspondent systématiquement aux frontières majeures en Condition 1 qui, comme nous l'avons vu dans nos résultats en perception, ne varient pas en termes de niveaux de frontière (cf. 2.2.4 du même chapitre) : seuls G11 et G31 correspondent aux syntagmes réalisés en Condition 1, dont la frontière après N1 semble renvoyer au même niveau de frontière dans les deux groupements, à savoir une frontière d'IP. La plupart des études montrent effectivement que la frontière d'IP est souvent corrélée à la présence d'une pause. Dans le cadre de notre travail, la pause ne nous permettrait donc pas de distinguer différents niveaux de frontière. Cependant, comme nous l'avons déjà mentionné dans les précédents résultats sur la perception, il serait intéressant par la suite de contrôler directement la production des pauses par locuteur pour pouvoir ainsi rendre compte des différences entre G11 et G31 dans la perception de la proéminence finale sur N1 (cf. 2.1.1 du Chapitre 5).

### 3.1 Démarche statistique

Concernant la modélisation statistique, il s'agit comme pour les analyses perceptives de mettre en relation une variable dépendante avec une ou plusieurs variables indépendantes (ou variables explicatives), pour ainsi tester la manière dont la variable dépendante est influencée par la ou les variables indépendantes. Alors que nous avons testé précédemment l'effet du type de groupement et de la taille du constituant sur la perception des scores de frontières et de proéminences, il s'agit ici de tester l'effet des variations acoustiques sur la perception des scores de frontières d'une part, et sur la perception des scores de proéminences d'autre part.

Dans ce contexte, nous proposons d'effectuer le même modèle que dans les analyses perceptives précédentes, à savoir un modèle de régression ordinale à effet mixte<sup>17</sup>. Plus précisément, nous considérons les variations acoustiques (durée et  $f_0$ ) comme les effets fixes de notre modèle, tandis que nous prenons en compte les participants, les

---

<sup>17</sup> Pour plus de détails sur le modèle de régression ordinale à effet mixte utilisé dans le cadre de ce travail, nous renvoyons le lecteur à la partie 5 du Chapitre 4.

locuteurs et le type de syntagme comme des effets aléatoires<sup>18</sup>. L'ensemble des modèles a été réalisé à partir du logiciel d'analyse statistique R (R Core Team, 2016), et comme pour l'analyse perceptive, nous avons utilisé le package « ordinal » (Christensen, 2015) permettant l'exécution des modèles de régression ordinale à effet mixte (clmm : « a cumulative link mixed models »).

Par ailleurs, alors que dans les analyses précédentes, les variables indépendantes sont systématiquement catégorielles, dans le présent modèle, en revanche, les variables indépendantes sont toutes de nature numérique continue. Pour interpréter le sens de nos effets, nous nous intéresserons donc au coefficient de régression. Plus précisément, un coefficient négatif suppose que l'augmentation du paramètre acoustique (durée ou  $f_0$ ) favorise les niveaux inférieurs des scores de perception (échelle des scores de 0 à 3), tandis qu'un coefficient positif suppose que l'augmentation du paramètre acoustique favorise les niveaux supérieurs des scores de perception. Nous aurions pu également proposer une analyse post-hoc plus poussée, en faisant des comparaisons entre chaque niveau de la variable dépendante ordinale, à savoir les scores de frontières et de proéminences. Toutefois, nous considérons que cette analyse n'est pas utile pour l'instant, et pourra être envisagée dans de prochains travaux. En effet, nous tenons à rendre compte dans un premier temps des tendances globales sur la manière dont les événements perçus sont corrélés avec les paramètres acoustiques.

Enfin, pour chacun de nos modèles, nous avons appliqué la même procédure : nous démarrions avec le modèle le plus complet et nous excluons progressivement du modèle final les variables non-significatives.

### **3.2 Statistiques descriptives des principales variables acoustiques**

Avant de procéder aux modèles de régression ordinale à effet mixte, nous proposons de décrire dans un premier temps les valeurs moyennes des variables acoustiques sur chaque syllabe des syntagmes. Plus précisément, nous nous intéressons au paramètre de durée, dont les valeurs brutes ont été relevées en milliseconde (noté « ms »), et au paramètre de  $f_0$ , dont les valeurs médianes ont été mesurées en Hertz (noté « Hz »). Ces relevés acoustiques<sup>19</sup> ont été effectués sur l'ensemble des 4 locuteurs de notre corpus (3

---

<sup>18</sup> Le choix et la validité des variables aléatoires utilisées dans notre modèle statistique respectent les critères établis précédemment dans la partie 5 du Chapitre 4.

<sup>19</sup> Nous tenons à remercier Rafèu Sichel-Bazin pour sa contribution dans le relevé des mesures acoustiques de notre base de données.

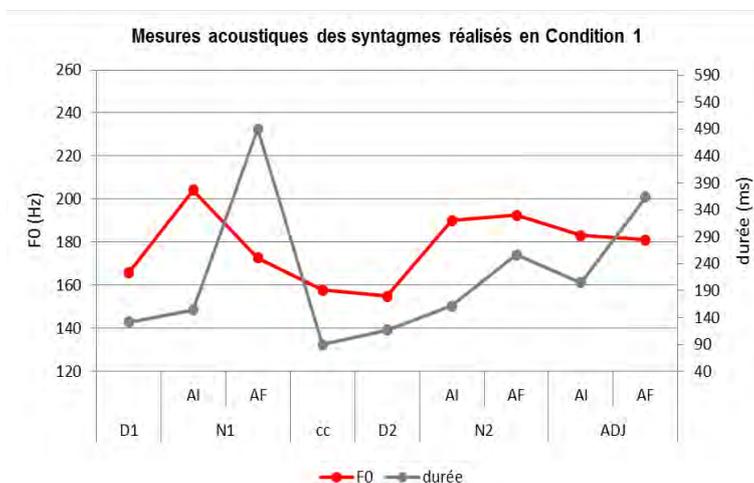
Femmes, 1 Homme). Il est en revanche important de préciser qu'aucune des ces valeurs n'ont été normalisées. Nous devons donc, dans le cadre d'une prochaine étude, prévoir une normalisation des paramètres de durée et de  $f_0$  pour pouvoir ainsi conforter nos résultats, qui représentent pour nous une première tentative, permettant de « dégrossir » notre compréhension des résultats perceptifs. Nous avons toutefois choisi de prendre en compte la médiane plutôt que la moyenne de  $f_0$  sur chaque syllabe du syntagme, pour ainsi avoir des valeurs plus représentatives.

Cette première analyse descriptive nous permettra d'avoir une idée générale de la réalisation des variations acoustiques, et facilitera également l'interprétation des résultats du modèle mixte. Nous choisissons d'effectuer cette description en séparant les deux conditions syntaxiques, pour ainsi conserver l'information sur la constituance prosodique. Nous proposons par ailleurs de décrire les variations acoustiques toutes tailles des constituants confondues dans un premier temps, pour pouvoir ensuite commenter les spécificités acoustiques dans chaque taille des constituants.

#### Description des paramètres acoustiques en Condition 1 :

La hauteur de  $f_0$  (en rouge cf. Figure N°74) semble nettement augmenter sur les syllabes accentuées, à savoir sur les premières et dernières syllabes de N1, N2 et ADJ. Nous remarquons également une  $f_0$  associée à AI sur N1 plus élevée que celle associée à AF ( $f_0$  moyennes : 204,24 Hz sur AI contre 172,76 Hz sur AF), et une valeur de  $f_0$  relativement équivalente entre AI et AF sur N2 ( $f_0$  moyennes : 190,04 Hz sur AI contre 192,49 Hz sur AF) et ADJ ( $f_0$  moyennes : 183,15 Hz sur AI contre 181,14 Hz sur AF). Par ailleurs, la valeur de  $f_0$  associée à AI sur N1 est la plus élevée parmi toutes les autres syllabes accentuées, tandis que la valeur de  $f_0$  associée à AF sur N1 est la moins élevée parmi toutes les autres syllabes accentuées. Ce phénomène de pic de  $f_0$  plus élevé sur la syllabe du premier mot lexical de l'IP est un phénomène connu dans la parole (notamment le « *early pitch accent* » chez Shattuck-Hufnagel, 1992). Enfin, les hauteurs de  $f_0$  associées aux mots grammaticaux ( $f_0$  moyennes : 165,77 Hz sur D1 ; 157,89 Hz sur cc ; 154,92 Hz sur D2) sont globalement moins importantes que celles associées aux proéminences initiales et finales des mots de contenu, à l'exception toutefois de l'AF sur N1 où la différence avec les mots grammaticaux tend à diminuer. Quant au paramètre de durée (en gris cf. Figure N°74), nous observons des « pics » de durée sur les syllabes finales des mots de contenu, et notamment un allongement plus important sur N1 (durée moyenne : 490,36 ms), moins important sur N2 (durée moyenne : 256,35 ms) et intermédiaire sur ADJ (durée moyenne : 363,64 ms). L'allongement de la durée sur les syllabes initiales semble par ailleurs nettement moins important que sur les syllabes finales, mais cependant un peu plus important que sur

les syllabes inaccentuées des mots grammaticaux (durées moyennes : entre 154,80 ms et 204,68 ms sur AI vs. entre 90,17 ms et 131,22 ms sur les mots grammaticaux).

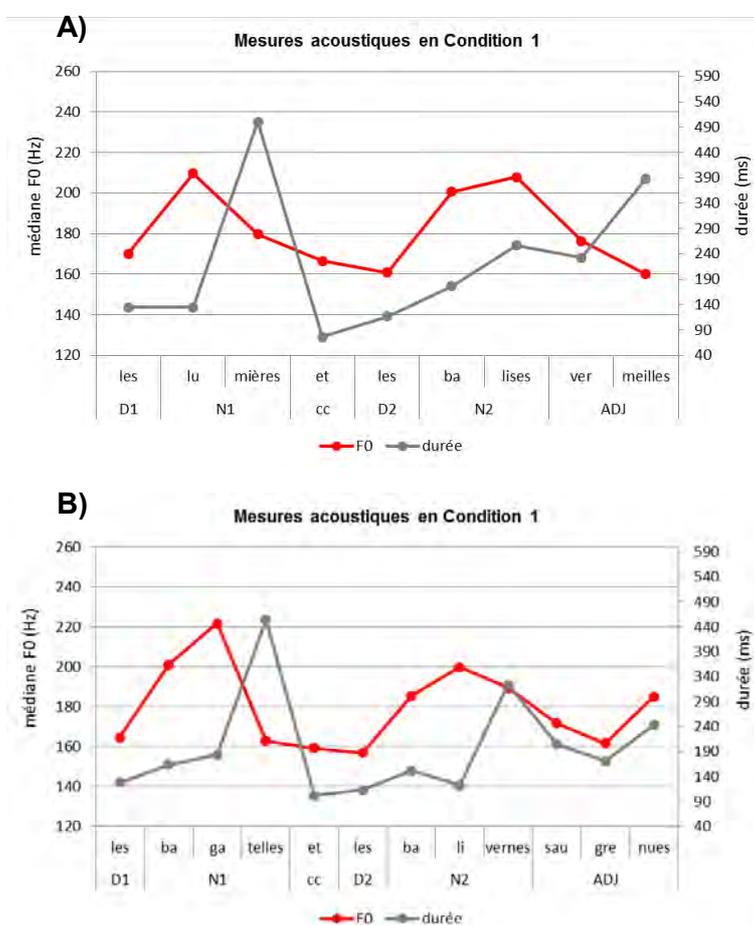


**Figure 74 :** Moyennes de  $f_0$  (Hz) et de durée (ms) de chaque syllabe des syntagmes réalisés en Condition 1, toutes tailles du constituant confondues

En mettant en lien ces mesures acoustiques avec la condition syntaxique, et par conséquent avec la constituance prosodique, il semblerait que la frontière forte (frontière d'IP) marquée après N1 soit réalisée par une association inattendue des paramètres acoustiques sur la syllabe finale du mot, à savoir une chute de  $f_0$  ( $f_0$  moyenne : 172,76 Hz) conjuguée à un allongement important de la durée syllabique (durée moyenne : 490,36 ms). Nous parlons « d'association inattendue » dans la mesure où nous aurions pu attendre en frontière forte la réalisation « canonique » d'une proéminence finale, à savoir une montée de  $f_0$  associée à un allongement de la durée. Il est également intéressant de remarquer que la proéminence initiale précédant la frontière après N1 est quant à elle réalisée par une montée de  $f_0$  ( $f_0$  moyenne : 204,24 Hz). A l'inverse, pour la frontière plus faible marquée après N2 (frontière d'ip ou d'AP), nous observons sur la première et dernière syllabe de N2 une valeur de  $f_0$  plus élevée que sur la syllabe finale de N1 ( $f_0$  moyennes : 190,04 Hz sur AI 192,49 Hz sur AF), ainsi qu'un allongement de la durée de la syllabe finale (durée moyenne : 256,35 ms) qui reste cependant nettement moins important que sur N1. Nous remarquons également sur ADJ une petite chute de  $f_0$  sur la syllabe initiale ( $f_0$  moyenne : 183,15 Hz) qui s'étend jusqu'à la dernière syllabe ( $f_0$  moyenne : 181,14 Hz), ainsi qu'un allongement de la durée de la

syllabe finale plus important que sur la dernière syllabe de N2 (durée moyenne : 363,64 ms), étant donné sans doute la présence d'une frontière d'IP après ADJ.

En prenant maintenant en compte la taille du constituant, il semblerait que la réalisation des paramètres acoustiques ne varie pas nettement à travers les différentes tailles du constituant. En comparant effectivement les mesures de  $f_0$  tout d'abord sur chaque syntagme (cf. Annexe N°17), variant en nombre de syllabes, nous observons systématiquement les mêmes réalisations : le pic de  $f_0$  en initiale puis la chute de  $f_0$  en finale de N1, ainsi que la forme de cloche plus ou moins ample sur N2 et ADJ (cf. Welby et al., 2016 pour la description de patterns acoustiques similaires).



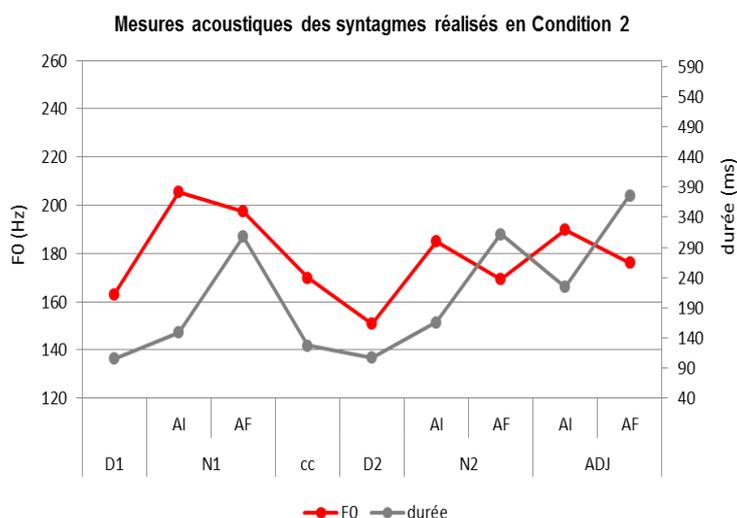
**Figure 75 :**  $f_0$  médiane (Hz) et durée (ms) de chaque syllabe des syntagmes réalisés en Condition 1 : « les lumières et les balises vermeilles » dans la Figure (A) et « les bagatelles et les balivernes saugrenues » dans la Figure (B)

Nous remarquons toutefois une différence sur l'alignement des pics de  $f_0$ . Effectivement, le pic maximal de  $f_0$  en initiale de N1 et N2 est réalisé sur la deuxième syllabe, dès lors que le mot est constitué de plus de deux syllabes (cf. Figure (A) et (B) N°75). Cette variabilité de l'alignement du pic de  $f_0$  a déjà été mentionnée dans la littérature, ce qui a d'ailleurs entraîné certains auteurs à considérer AI comme un « *marqueur de flou de frontière* » (Welby, 2003). La réalisation mélodique de la proéminence finale sur ADJ peut également varier. L'accent est réalisé soit par une montée soit par une chute de  $f_0$ , et ce indépendamment du nombre de syllabes du constituant ADJ ou N2+ADJ. En effet, pour un constituant N2+ADJ composé du même nombre de syllabes (cf. Annexe N°17-B « *les balises verticales* » et N°17-C « *les balivernes saumâtres* »), les deux types de réalisations acoustiques sont possibles. De la même manière, deux ADJ de même taille peuvent être réalisés soit par une montée de  $f_0$ , soit par une chute de  $f_0$  (cf. Annexe N°17-A « *vermeilles* » et N°17-G « *fameux* »). Il est intéressant de mettre en lien cette observation avec les résultats obtenus dans l'étude en perception de Méndez & Astésano (2017) et Méndez (en cours), qui rend compte d'une modulation de la force des frontières prosodiques par la réalisation d'un AF montant ou descendant. Plus précisément, les auteurs montrent qu'une frontière marquée par un AF descendant (une chute de  $f_0$ ) serait perçue plus forte qu'une frontière marquée par un AF montant (une montée de  $f_0$ ). Toutefois, ce résultat ne s'observerait qu'en frontière forte (frontière d'ip ou d'IP). Il se pourrait donc que la différence observée dans notre cas entre un AF marqué par une chute ou une montée de  $f_0$  relève du même phénomène.

### Description des paramètres acoustiques en Condition 2 :

Comme dans la Condition 1, la hauteur de  $f_0$  (en rouge cf. Figure N°76) semble augmenter sur les syllabes accentuées: AI et AF sur N1, N2 et ADJ. Toutefois, alors qu'en Condition 1, la valeur de  $f_0$  sur AI est majoritairement équivalente à la valeur de  $f_0$  sur AF, en Condition 2, en revanche, la hauteur de  $f_0$  est majoritairement plus élevée sur AI que sur AF ( $f_0$  moyennes : entre 184,87 Hz et 205,44 Hz sur AI vs. entre 169,46 Hz et 197,53 Hz sur AF). Seul N1 présente des valeurs de  $f_0$  relativement égales entre les deux syllabes accentuées ( $f_0$  moyennes : 205,44 Hz sur AI contre 197,53 Hz sur AF). Enfin, à l'instar de la Condition 1, la valeur de  $f_0$  associée à AI sur N1 est la plus élevée parmi toutes les autres syllabes accentuées. Nous observons également que les hauteurs de  $f_0$  sur les mots grammaticaux sont nettement moins importantes ( $f_0$  moyennes : 163,24 Hz sur D1 ; 169,94 Hz sur cc ; 150,92 Hz sur D2) que celles sur les proéminences initiales et finales des mots de contenu. Quant au paramètre de durée (en gris cf. Figure N°76), nous constatons des « pics » de durée sur les syllabes finales des mots de contenu. Il est par ailleurs intéressant de remarquer que l'allongement de la durée sur les syllabes finales de N1 et N2 est

relativement équivalent (durées moyennes : 308,13 ms sur N1 contre 311,99 ms sur N2). Enfin, nous l'observons déjà en Condition 1, l'allongement de la durée sur les syllabes initiales est nettement moins important que sur les syllabes finales, mais cependant un peu plus important que sur les syllabes inaccentuées des mots grammaticaux (durées moyennes : entre 149,44 ms et 224,59 ms sur AI vs. entre 106,19 ms et 127,77 ms sur les mots grammaticaux).

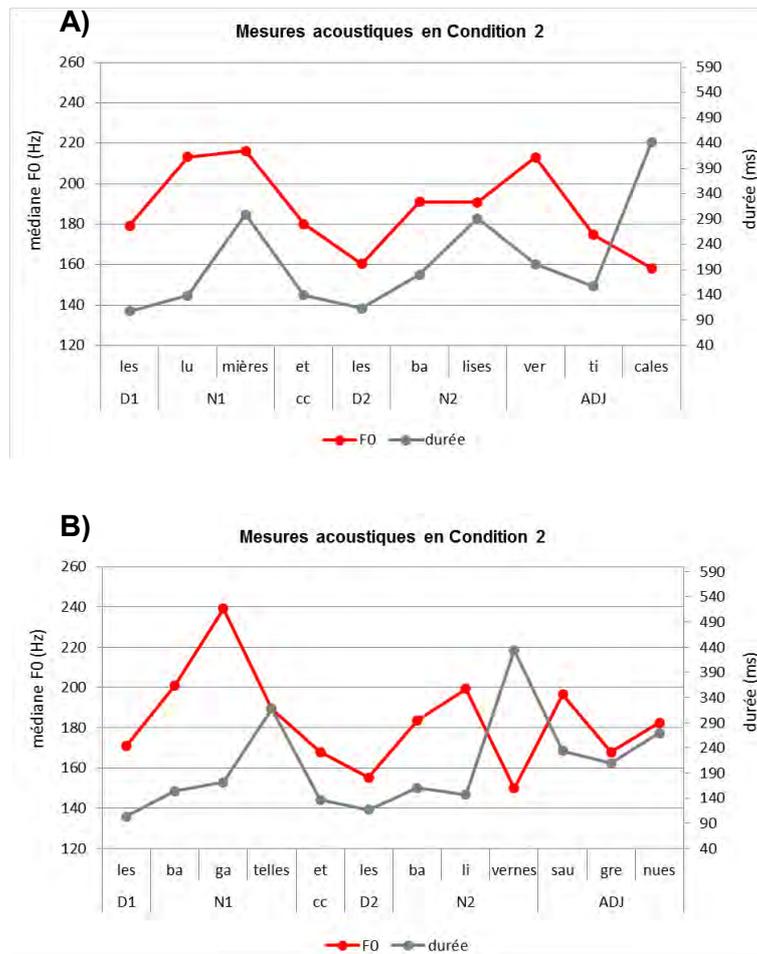


**Figure 76 :** Moyennes de  $f_0$  (Hz) et de durée (ms) de chaque syllabe des syntagmes réalisés en Condition 2, toutes tailles du constituant confondues

En mettant en lien ces mesures acoustiques avec la condition syntaxique, et par conséquent avec la constituance prosodique, il est intéressant de remarquer là encore une association inattendue des paramètres acoustiques en frontière forte, qui reste toutefois moins marquée qu'en Condition 1. La frontière la plus forte en Condition 2 est située après N2 (frontières d'ip ou d'AP), où nous constatons également une chute de  $f_0$  ( $f_0$  moyenne : 169,46 Hz) associée à un allongement de la durée de la syllabe finale (durée moyenne : 311,99 ms). Une montée de  $f_0$  est également réalisée sur la syllabe initiale de N2 ( $f_0$  moyenne : 184,97 Hz). En revanche, alors qu'en Condition 1 une chute de  $f_0$  sur la syllabe initiale de ADJ s'étend jusqu'à la dernière syllabe, nous remarquons en Condition 2 une montée de  $f_0$  sur la syllabe initiale ( $f_0$  moyenne : 189,85 Hz) suivie d'une chute de  $f_0$  sur la syllabe finale ( $f_0$  moyenne : 176,21 Hz). Cette montée de  $f_0$  sur la syllabe initiale de ADJ serait par ailleurs corrélée à la réalisation d'une frontière forte après N2, tandis que la chute de  $f_0$  sur la syllabe initiale de ADJ en Condition 1 serait corrélée, quant à elle, à la réalisation d'une frontière faible située à

l'intérieur d'un constituant plus large. Quant à la frontière faible après N1 en Condition 2, nous constatons une montée de  $f_0$  sur la syllabe initiale et finale de N1 ( $f_0$  moyennes : 205,44 Hz sur AI et 197,53 Hz sur AF), ainsi qu'un allongement de la durée de la syllabe finale seulement (durée moyenne : 308,13 ms). Cette réalisation est également observée en frontière faible après N2 en Condition 1.

En prenant maintenant en compte la taille du constituant, il apparaît là encore que la réalisation des paramètres acoustiques ne varie pas nettement à travers les différentes tailles du constituant. En comparant effectivement les mesures de  $f_0$  tout d'abord sur chaque syntagme (cf. Annexe N°17), variant en nombre de syllabes, nous observons systématiquement les mêmes variations : les pics de  $f_0$  en initiale et finale de N1 qui sont plus ou moins à hauteur égale, ainsi que la réalisation consécutive d'une montée et d'une chute de  $f_0$  sur N2 puis sur ADJ.



**Figure 77 :**  $f_0$  médiane (Hz) et durée (ms) de chaque syllabe des syntagmes réalisés en Condition 2 : « les lumières et les balises verticales » dans la Figure (A) et « les bagatelles et les balivernes saugrenues » dans la Figure (B)

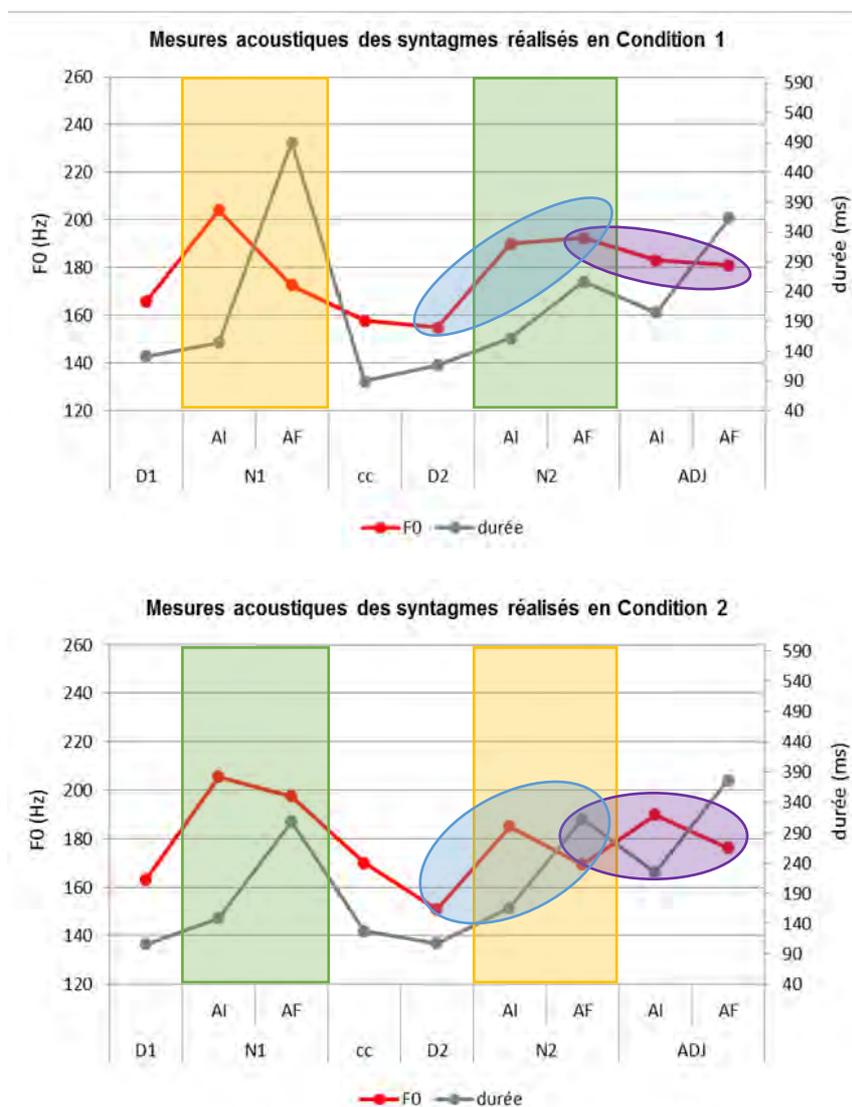
Toutefois, comme pour la Condition 1, une différence sur l'alignement des pics de  $f_0$  est observée dès lors que le nombre de syllabes augmente. Le pic maximal de  $f_0$  en initiale de N1 et N2 est en effet réalisé sur la deuxième syllabe lorsque le mot est constitué de plus de deux syllabes (cf. Figure (B) N°77). Il est également intéressant de constater que l'alignement des pics de  $f_0$  ne varie pas sur ADJ, quelle que soit la taille du constituant. En effet, lorsqu'ADJ est tri- ou quadrisyllabique le pic maximal de  $f_0$  est systématiquement aligné sur la première syllabe, tandis que la deuxième syllabe est marquée par une chute de  $f_0$  (cf. Figure (A) et (B) N°77). Ainsi, à l'instar de Welby (2003), nous observons que l'alignement du pic maximal de  $f_0$  en initiale des mots lexicaux peut être variable. En revanche, nos résultats tendent à montrer que cette variabilité n'est valable que dans le cas des mots lexicaux précédés d'un ou de plusieurs mots grammaticaux. Nous rappelons en effet que Welby montre que seul l'alignement du ton « Hi » peut varier, tandis que le ton bas « L » précédent est systématiquement ancré à la frontière entre le mot grammatical et le mot lexical. Or, nous constatons que lorsqu'il n'y a pas la possibilité d'ancrer le ton « L » au début du mot, comme dans le cas de ADJ, le ton « Hi » est alors systématiquement ancré sur la première syllabe du mot. Quant au paramètre de durée, nous observons, comme pour la Condition 1, que les variations de durée sont très stables d'un syntagme à un autre.

### Synthèse des résultats

A travers l'ensemble des analyses descriptives, il apparaît que le marquage acoustique des frontières après N1 et N2 soit spécifique au niveau de frontière, et non pas à la taille du constituant. Nous remarquons en effet que les frontières perçues plus fortes (après N1 en Condition 1 et après N2 en Condition 2) semblent être ici systématiquement associées à une chute de  $f_0$  conjuguée à un allongement important de la durée de la syllabe finale précédant la frontière, et à un pic de  $f_0$  sur la syllabe initiale précédant la frontière (indiqué dans le cadre jaune Figure N°78). Quant aux frontières perçues moins fortes (après N2 en Condition 1 et après N1 en Condition 2), nous observons systématiquement un allongement de la durée de la syllabe finale précédant la frontière, qui est globalement moins important que celui observé en frontière forte, associé à une montée de  $f_0$  sur les syllabes initiale et finale précédant la frontière (indiqué dans le cadre vert Figure N°78).

Une réalisation acoustique de  $f_0$  différente est également observée sur ADJ selon la force de la frontière précédente (indiqué dans le cadre violet Figure N°78). Après une frontière forte (en Condition 2), ADJ est marqué par une montée de  $f_0$  sur la syllabe initiale puis par une chute de  $f_0$  sur la syllabe finale, tandis qu'après une frontière faible (en Condition 1), ADJ est marqué par une chute de  $f_0$  sur la syllabe initiale qui s'étend jusqu'à la dernière syllabe. De la même manière, nous constatons que la réalisation acoustique de  $f_0$

sur N2 varie selon la force de la frontière précédente (indiqué dans le cadre bleu Figure N°78). Après une frontière forte (en Condition 1), N2 est marqué par une montée de  $f_0$  sur AI qui s'étend jusqu'à AF, et après une frontière faible (en Condition 2), N2 est marqué par une montée de  $f_0$  sur AI puis par une chute de  $f_0$  sur AF.

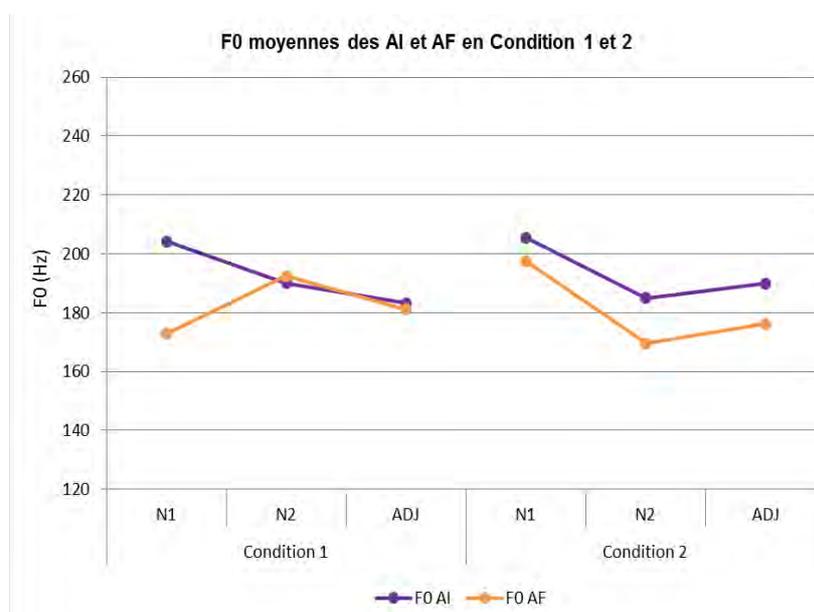


**Figure 78 :** Moyennes de  $f_0$  (Hz) et de durée (ms) de chaque syllabe des syntagmes réalisés en Condition 1 dans la figure du haut, et en Condition 2 dans la figure du bas, toutes tailles du constituant confondues

Le cadre vert indique la réalisation d'une frontière faible et le cadre jaune celle d'une frontière forte. Les cercles violets permettent de signaler la réalisation acoustique des préminences sur ADJ après une frontière faible en Condition 1 et après une frontière forte en Condition 2. Les cercles bleus permettent de signaler la réalisation acoustique des préminences sur N2 après une frontière forte en Condition 1 et après une frontière faible en Condition 2.

Il semblerait dans ce contexte que le marquage d'une frontière forte soit systématiquement associé à la réalisation d'une montée mélodique sur la syllabe initiale suivant la frontière. En revanche, après une frontière perçue plus faible, la réalisation acoustique sur la syllabe initiale suivant la frontière varie selon la position du mot lexical après la frontière. Effectivement, ADJ qui est directement positionné après la frontière prosodique, est réalisé par une chute de  $f_0$  sur la syllabe initiale (cf. Figure N°75), tandis que N2 qui est précédé de deux mots grammaticaux, est réalisé par une montée de  $f_0$  sur la syllabe initiale (cf. Figure N°77). Cette montée mélodique reste toutefois moins importante que celle réalisée sur N2 après une frontière forte (en Condition 1).

Par ailleurs, concernant plus précisément la réalisation acoustique des AI et AF, nous observons à travers les deux conditions syntaxiques que les valeurs de  $f_0$  associées à AI (en violet Figure N°79) sont majoritairement équivalentes ou supérieures aux valeurs de  $f_0$  associées à AF (en orange Figure N°79). En revanche, du point de vue de la durée, les AI sont systématiquement marqués par un allongement de la durée syllabique nettement moins important que les AF, mais malgré tout plus important que les mots grammaticaux.



**Figure 79 :** Moyennes de  $f_0$  (Hz) associées à AI et AF sur chaque mot lexical (N1, N2 et ADJ) en Condition 1 et 2

Enfin, en prenant en compte le nombre de syllabes de chaque constituant, il semblerait que les variations acoustiques de  $f_0$  et de durée restent relativement stables à travers les

différentes tailles du constituant dans les deux conditions syntaxiques. En effet, quelle que soit l'incréméntation en nombre de syllabes des mots, le même pattern de  $f_0$  et de durée est reproduit dans tous les syntagmes. Seul l'alignement des pics de  $f_0$  en initiale de N1 et N2 varie en fonction de la taille du constituant. En effet, le pic maximal de  $f_0$  est aligné sur la première syllabe dans les mots bisyllabiques, et sur la deuxième voire sur la troisième syllabe dans les mots composés de plus de deux syllabes. Il est par ailleurs intéressant de remarquer que cette variation d'alignement n'impacte pas le constituant ADJ, où le pic de  $f_0$  est systématiquement aligné avec la première syllabe du mot. Comme nous l'avons déjà mentionné, il se pourrait selon nous que la position du pic de  $f_0$  ne soit pas variable sur ADJ, dans la mesure où ADJ, à l'inverse de N1 et N2, est positionné directement après la frontière prosodique. En effet, étant donné qu'il n'y a pas la possibilité d'ancrer un ton bas « L » au début du mot lexical, le ton « Hi » est alors systématiquement ancré sur la première syllabe du mot. Ce résultat pourrait alors questionner le rôle de « *marqueur flou de frontière* » (Welby, 2003) de la proéminence initiale en français, et devra être mis en relation avec les résultats perceptifs, pour ainsi rendre compte de l'impact des variations d'alignement du pic de  $f_0$  sur la perception de la proéminence initiale.

### **3.3 Résultats du modèle de régression ordinaire à effet mixte**

Sur la base de ces statistiques descriptives, nous proposons maintenant de tester plus précisément l'effet des variations acoustiques de  $f_0$  et de durée sur la perception des scores de frontières et de proéminences, pour ainsi faire ressortir des corrélats acoustiques spécifiques à la perception des différents événements prosodiques. Avant de présenter en détails chaque résultat obtenu, nous exposons dans un premier temps le choix de nos sites d'analyse pour les corrélats acoustiques de la force des frontières tout d'abord, et ensuite pour les corrélats acoustiques de la force des proéminences. En effet, selon les résultats perceptifs obtenus dans les analyses précédentes, nous n'effectuerons pas les modèles mixtes sur les mêmes sites dans les deux types d'analyse. Enfin, les résultats obtenus dans chaque analyse seront discutés à la fin de chacune des deux prochaines parties.

### 3.3.1 Corrélats acoustiques de la force des frontières

#### 3.3.1.1 Choix des sites d'analyse

Les résultats de notre analyse perceptive montrent que les scores de frontières sont nettement influencés par le groupement, tandis que la taille des constituants ne semble pas impacter directement la perception des frontières. Plus précisément, nous avons rendu compte de 4 niveaux de frontière différents :

<b>G11</b>	[ les N1 ] <b>IP</b>	//	[ et les N2 <b>AP</b>	/	ADJ ] <b>IP</b>
<b>G31</b>	[ les N1 ] <b>IP</b>	//	[ et les N2 ] <b>ip</b>	/	[ ADJ ] <b>IP</b>
<b>G32</b>	[ les N1 ] <b>AP</b>	/	[ et les N2 ] <b>ip</b>	//	[ ADJ ] <b>IP</b>
<b>G42</b>	[ les N1 <b>PW</b>	/	et les N2 <b>AP</b>	//	ADJ ] <b>IP</b>
<b>G22</b>	[ les N1 <b>PW</b>	/	et les N2 ] <b>ip</b>	//	[ ADJ ] <b>IP</b>

Il est donc maintenant intéressant d'expérimenter les corrélats acoustiques possibles associés à ces 4 niveaux de frontière.

- **Est-ce que des variations acoustiques peuvent expliquer la perception des différentes forces de frontières ?**

Pour cela, nous choisissons de prendre en compte quatre paramètres acoustiques :

1. La durée de la syllabe finale précédant la frontière : syllabe finale de N1 pour la frontière après N1, et syllabe finale de N2 pour la frontière après N2.
2. La valeur de  $f_0$  sur la syllabe finale précédant la frontière : syllabe finale de N1 pour la frontière après N1, et syllabe finale de N2 pour la frontière après N2.
3. La valeur de  $f_0$  sur la syllabe initiale précédant la frontière : syllabe initiale de N1 pour la frontière après N1, et syllabe initiale de N2 pour la frontière après N2.
4. La valeur de  $f_0$  sur la syllabe initiale suivant la frontière : syllabe initiale de N2 pour la frontière après N1, et syllabe initiale de ADJ pour la frontière après N2.

Le choix de ces paramètres a notamment été guidé par les résultats des analyses descriptives des principales variables acoustiques. En effet, nous avons observé que le marquage d'un type de frontière est associé à des réalisations acoustiques spécifiques. Une frontière plus forte semble être marquée par un pic de  $f_0$  sur la syllabe initiale, et par une chute de  $f_0$  sur la syllabe finale du mot lexical précédant la frontière, tandis qu'une frontière plus faible semble être associée à un allongement de la durée de la syllabe finale qui peut être aussi conjugué à une montée de  $f_0$  sur la syllabe initiale et finale du mot lexical précédant la frontière. Nous avons également choisi de nous intéresser aux réalisations acoustiques sur les syllabes suivant la frontière. Après une frontière plus forte (en Condition 1), N2 est marqué par une montée de  $f_0$  sur AI qui s'étend jusqu'à AF, alors qu'après une frontière plus faible (en Condition 2), N2 est réalisé par une montée de  $f_0$  sur AI puis par une chute de  $f_0$  sur AF. De la même manière, après une frontière plus forte (en Condition 2), ADJ est marqué par une montée de  $f_0$  sur la syllabe initiale puis par une chute de  $f_0$  sur la syllabe finale de ADJ, alors qu'après une frontière plus faible (en Condition 1), ADJ est réalisé par une chute de  $f_0$  sur la syllabe initiale qui s'étend jusqu'à la dernière syllabe.

Nous proposons également d'effectuer ce modèle en séparant les deux conditions syntaxiques. Effectivement, nos résultats en perception montrent un effet important du groupement sur la perception des frontières après N1 et N2, et ces groupements sont par ailleurs étroitement liés aux deux conditions syntaxiques (*cf.* Figure N°47) :

- Groupement 1 \* Condition 1 = **G11** [les N1] [et les N2 ADJ]
- Groupement 3 \* Condition 1 = **G31** [les N1] [et les N2] [ADJ]
- Groupement 2 \* Condition 2 = **G22** [les N1 et les N2] [ADJ]
- Groupement 3 \* Condition 2 = **G32** [les N1] [et les N2] [ADJ]
- Groupement 4 \* Condition 2 = **G42** [les N1 et les N2 ADJ]

Une analyse par condition syntaxique permet alors de tester les différentes granularités des niveaux de frontière perçus dans chaque groupement, chacun associé à une condition spécifique. Une analyse par groupement, en revanche, ne permettrait pas de faire ressortir les différences potentielles entre les variations acoustiques, dans la mesure où un seul niveau de frontière est associé à chaque groupement.

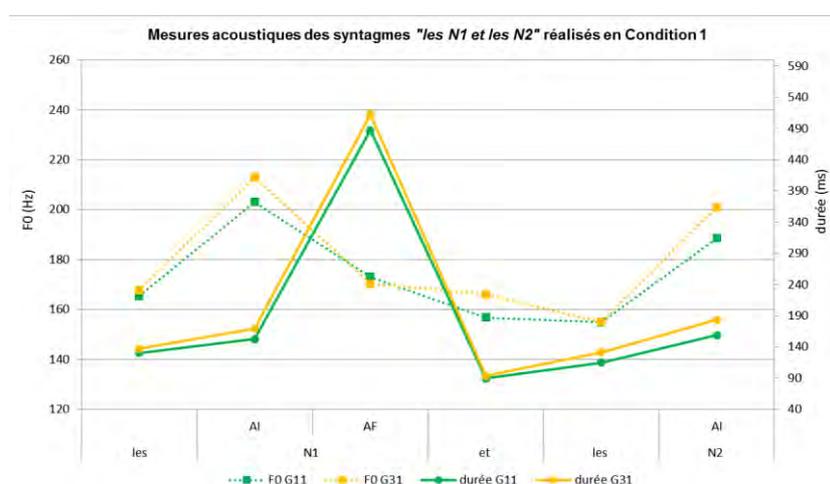
Pour chaque modèle réalisé, nous avons également analysé les composantes de la variance qui tendent à montrer une valeur de la variance liée au participant plus élevée que celle liée au locuteur ou au syntagme. Ces résultats ont également été observés dans l'analyse perceptive des frontières et des proéminences. Il s'avère en effet que les résultats

obtenus dans le cadre de notre analyse varient davantage en fonction de l'auditeur qu'en fonction du locuteur produisant les différents syntagmes.

### 3.3.1.2 Force de la frontière perçue après N1

#### En Condition 1

Les résultats ne montrent aucun effet significatif des paramètres prosodiques sur la perception de la frontière après N1 ( $p > 0,05$ ). Le score de la frontière ne varierait ni en fonction de l'allongement de la durée de la syllabe finale de N1, ni en fonction de la hauteur de  $f_0$  sur la syllabe initiale et finale de N1, ou initiale de N2. Sur le plan acoustique, aucune différence ne serait observée entre les deux groupements associés à la Condition 1 (cf. Figure N°80), ce qui confirmerait alors qu'un seul niveau de frontière ait été perçu après N1 en Condition 1 (frontière d'IP).



**Figure 80** : Moyennes de  $f_0$  (Hz) et de durée (ms) sur chaque syllabe des syntagmes « les N1 et les N2 » réalisés en G11 et G31 (Condition 1), toutes tailles du constituant confondues

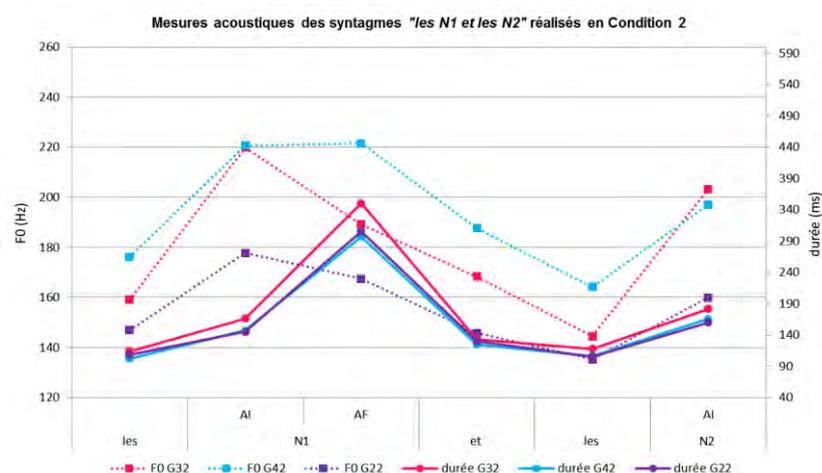
#### En Condition 2

Les résultats indiquent un effet de la durée (LR Chisq= 7,4748, df= 1,  $p=0,006$ ) et de la hauteur de  $f_0$  associées à la syllabe finale de N1 (LR Chisq= 8,2982, df= 1,  $p=0,003$ ), ainsi qu'un effet de la hauteur de  $f_0$  associée à la syllabe initiale de N2 (LR Chisq= 4,4393, df= 1,

$p=0,035$ ). En revanche, la hauteur de  $f_0$  sur la syllabe initiale de N1 n'impacterait pas la force de la frontière ( $p>0,05$ ). Plus précisément, en regardant les coefficients de régression, nous voyons que l'augmentation de la durée de la syllabe finale de N1 favorise les niveaux supérieurs du score de frontière (valeur du coefficient : 0,006718). Quant aux valeurs de  $f_0$ , alors que l'augmentation de la hauteur de  $f_0$  sur la syllabe finale de N1 favorise les niveaux inférieurs du score de frontière (valeur du coefficient : -0,008478), l'augmentation de la hauteur de  $f_0$  sur la syllabe initiale de N2, en revanche, favorise les niveaux supérieurs du score de frontière (valeur du coefficient : 0,008127) (cf. Annexe N°18-A).

- **Cela suppose alors qu'une frontière serait perçue plus forte lorsque l'allongement de la durée de la syllabe finale de N1 est plus important, et lorsque la valeur de  $f_0$  est moins élevée sur la syllabe finale de N1 et plus élevée sur la syllabe initiale de N2.**

En mettant en relation ces résultats avec ceux obtenus dans l'analyse perceptive, il semblerait que les deux forces de frontières perçues après N1 en Condition 2 (frontière d'AP en G32 et frontière de PW en G22 et G42) soient corrélées à des variations acoustiques spécifiques (cf. Figure N°81).



**Figure 81** : Moyennes de  $f_0$  (Hz) et de durée (ms) sur chaque syllabe des syntagmes « les N1 et les N2 » réalisés en G32, G42 et G22 (Condition 2), toutes tailles du constituant confondues

En regardant la Figure N°81, nous constatons effectivement que la frontière d'AP en G32 présente un allongement de la durée syllabique plus important que la frontière de PW en G42 et G22 (durées moyennes : 350,07 ms en G32 contre 305,31 ms en G22 et 293,93 ms en G42), ainsi qu'une valeur de  $f_0$  sur AI de N2 légèrement plus élevée que la frontière de PW en G42 et G22 ( $f_0$  moyennes : 203,25 Hz en G32 contre 197,04 Hz en G42 et 159,88 Hz en G22). En revanche, alors que la valeur de  $f_0$  sur AF de N1 est en effet moins élevée en G32 (AP) qu'en G42 (PW), elle apparaît cependant plus élevée en G32 (AP) qu'en G22 (PW) ( $f_0$  moyennes : 189,13 Hz en G32 ; 221,58 Hz en G42 ; 167,42 Hz en G22). Dans ce contexte, seul le paramètre de durée permettrait de distinguer les différences de force de frontière entre G32 (AP) d'un côté puis G22 et G42 (PW) de l'autre.

### 3.3.1.3 Force de la frontière perçue après N2

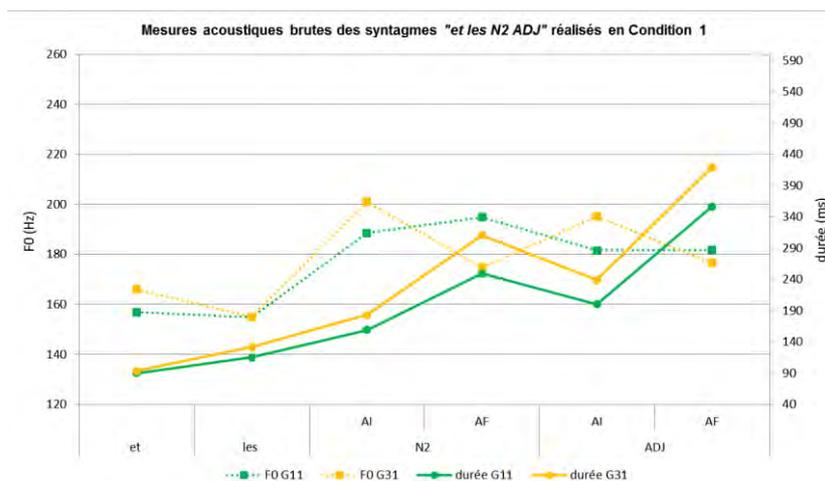
#### En Condition 1

Les résultats indiquent un effet de la durée de la syllabe finale de N2 (LR Chisq= 7,8386, df= 1, p=0,005), ainsi qu'un effet de la hauteur de  $f_0$  sur la syllabe finale de N2 (LR Chisq= 21,7966, df= 1, p<0,001) et initiale de ADJ (LR Chisq= 13,6483, df= 1, p=0,0002). La hauteur de  $f_0$  sur la syllabe initiale de N2 n'impacterait pas cependant la force de la frontière (p>0,05). Plus précisément, en regardant les coefficients de régression, il semblerait que l'augmentation de la durée de la syllabe finale de N2 favorise les niveaux supérieurs du score de frontière (valeur du coefficient : 0,004197). Quant aux valeurs de  $f_0$ , alors que l'augmentation de la hauteur de  $f_0$  sur la syllabe finale de N2 favorise les niveaux inférieurs du score de frontière (valeur du coefficient : -0,016356), l'augmentation de la hauteur de  $f_0$  sur la syllabe initiale de ADJ, en revanche, favorise les niveaux supérieurs du score de frontière (valeur du coefficient : 0,011309) (cf. Annexe N°18-B).

- **Le renforcement du niveau de frontière serait alors caractérisé par un allongement de la durée de la syllabe finale de N2 plus important, mais aussi par une valeur de  $f_0$  moins élevée sur la syllabe finale de N2 et plus élevée sur la syllabe initiale de ADJ.**

Ce résultat confirmerait donc ceux obtenus dans l'analyse perceptive, où deux niveaux de frontière après N2 ont été perçus (une frontière d'ip en G31 et d'AP en G11). En regardant la Figure N°82, nous constatons effectivement un allongement de la durée de la syllabe finale de N2 plus important en G31 qu'en G11 (durées moyennes : 310,80 ms en G31 contre

249,24 ms en G11). De la même manière, la syllabe finale de N2 est marquée par une chute de  $f_0$  en G31 et par une montée de  $f_0$  en G11 ( $f_0$  moyennes : 174,78 Hz en G31 contre 194,80 Hz en G11), et à l'inverse, la syllabe initiale de ADJ est réalisée par une montée de  $f_0$  en G31 et par une chute de  $f_0$  en G11 ( $f_0$  moyennes : 195,06 Hz en G31 contre 181,59 Hz en G11). Dans ce contexte, à la fois les paramètres de durée et de  $f_0$  permettraient de distinguer les différences de force de frontière entre G11 et G31.



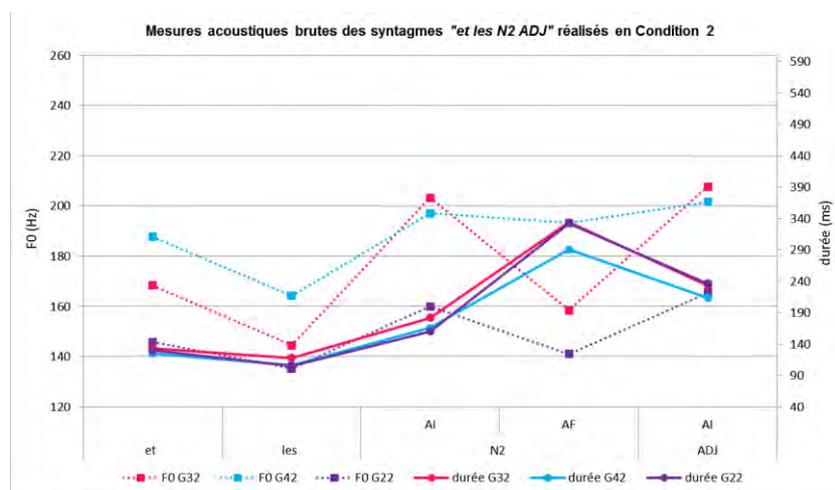
**Figure 82 :** Moyennes de  $f_0$  (Hz) et de durée (ms) sur chaque syllabe des syntagmes « et les N2 ADJ » réalisés en G11 et G31 (Condition 1), toutes tailles du constituant confondues

### En Condition 2

Les résultats indiquent un effet de la durée (LR Chisq= 11,175, df= 1, p=0,0008) et de la hauteur de  $f_0$  (LR Chisq= 17,056, df= 1, p<0,001) sur la syllabe finale de N2. Les hauteurs de  $f_0$  sur les syllabes initiales de N2 et ADJ n'impacteraient pas cependant la force de la frontière (p>0,05). A travers les coefficients de régression, nous comprenons plus précisément que l'augmentation de la durée de la syllabe finale de N2 favorise les niveaux supérieurs du score de frontière (valeur du coefficient : 0,004675), et que l'augmentation de la hauteur de  $f_0$  sur la syllabe finale de N2 favorise les niveaux inférieurs du score de frontière (valeur du coefficient : -0,013695) (cf. Annexe N°18-B).

- **Cela suppose alors que la frontière après N2 serait perçue plus forte lorsque l'allongement de la durée de la syllabe finale de N2 augmente.**

En mettant en relation ces résultats avec ceux obtenus dans l'analyse perceptive, il semblerait que la perception d'une frontière moins forte en G42 (frontière d'AP) qu'en G32 et G22 (frontières d'ip) soit corrélée à des variations acoustiques spécifiques (cf. Figure N°83).



**Figure 83 :** Moyennes de  $f_0$  (Hz) et de durée (ms) sur chaque syllabe des syntagmes « et les N2 ADJ » réalisés en G32, G42 et G22 (Condition 2), toutes tailles du constituant confondues

Nous constatons effectivement dans la Figure N°83 que l'allongement de la durée de la syllabe finale de N2 est plus important en G32 et G22 qu'en G42 (durées moyennes : 333,76 ms en G32 et 332,43 ms en G22 contre 290,60 ms en G42), et que la hauteur de  $f_0$  sur la syllabe finale de N2 est en revanche moins élevée en G32 et G22 qu'en G42 ( $f_0$  moyennes : 158,33 Hz en G32 et 140,95 Hz en G22 contre 193,23 Hz). Dans ce contexte, à la fois les paramètres de durée et de  $f_0$  permettraient de distinguer les différences de force de frontière entre G32 d'un côté puis G22 et G42 de l'autre.

### 3.3.1.4 Discussion

L'analyse des corrélats acoustiques des scores de frontières montrent que le paramètre acoustique majoritairement impliqué dans la perception de la force des frontières est l'allongement de la durée de syllabe finale précédant la frontière. Le renforcement du niveau de frontière serait en effet lié à l'augmentation de l'allongement de la durée syllabique. Les valeurs de  $f_0$  associées à la syllabe finale précédant la frontière et à la syllabe initiale suivant la frontière peuvent également influencer la perception des frontières. Une frontière est en effet perçue plus forte lorsque la syllabe finale précédant la frontière est

marquée par une chute de  $f_0$ , et lorsqu'à l'inverse, la syllabe initiale suivant la frontière est marquée par une montée de  $f_0$ . Il est en revanche intéressant de remarquer que la valeur de  $f_0$  associée à la syllabe initiale précédant la frontière n'influence jamais la perception des scores de frontières. Ce résultat est à mettre en lien avec ceux obtenus en perception, où nous n'observons aucun lien de dépendance entre le groupement et les scores de prééminences sur les syllabes initiales de N1 et N2. Selon nous, cela peut être dû au fait que N1 et N2 soient précédés d'un ou de plusieurs mots grammaticaux. En effet, en l'absence d'indices (notamment les mots grammaticaux) de début potentiel de mot (Christophe et *al.*, 1997), AI aurait une double fonction, en marquant non seulement le début du mot lexical mais également les différents niveaux de frontières. Dans ce contexte, les prééminences initiales sur N1 et N2 rempliraient uniquement le rôle de marqueurs de mot.

Etant donné que le paramètre de durée apparaît comme le plus pertinent dans le marquage des frontières, nous proposons dans un premier temps de calculer les rapports de durée entre les syllabes inaccentuées de référence (les mots grammaticaux ici) et les syllabes finales accentuées, précédant les deux frontières prosodiques (après N1 et N2). Ce rapport de durée permet notamment de rendre compte de l'importance de l'allongement de la durée syllabique à travers les différents groupements. Plus précisément, pour calculer le rapport de durée de chaque syllabe finale dans chacun des groupements, nous avons relevé la durée moyenne des syllabes associées aux mots grammaticaux, s'élevant à 113,36 ms, que nous avons ensuite soustrait à la durée de la syllabe finale accentuée. Le total de ce calcul a ensuite été divisé par la durée de la syllabe inaccentuée de référence. Nous proposons ensuite d'associer ces rapports de durée avec chaque niveau de frontière perçus (*cf.* Tableau N°21).

**Tableau 21 :** Mise en relation des valeurs « *mean.class* » des résultats perceptifs avec les rapports de durée des analyses acoustiques

Les couleurs ou les nuances de couleurs permettent d'identifier les groupements ayant le même niveau de frontière : rouge pour la frontière d'IP, vert et vert pâle pour la frontière d'ip, jaune et jaune pâle pour la frontière d'AP et bleu pour la frontière de PW.

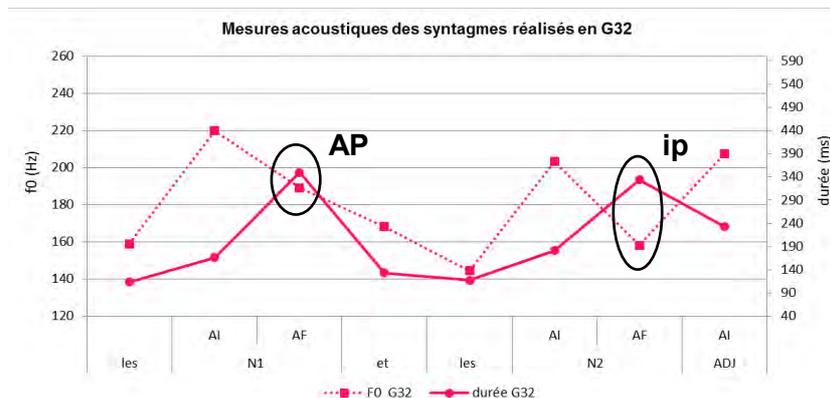
Groupement	Frontière perçue après N1	Rapport de durée	Frontière perçue après N2	Rapport de durée
G11	[ les N1 ] <sub>IP</sub> 	3,3	[ et les N2 ] <sub>AP</sub>  [ ADJ ]	1,2
G31	[ les N1 ] <sub>IP</sub> 	3,5	[ et les N2 ] <sub>ip</sub>  [ ADJ ]	1,7
G32	[ les N1 ] <sub>AP</sub> 	2,1	[ et les N2 ] <sub>ip</sub>  [ ADJ ]	1,9
G42	[ les N1 ] <sub>PW</sub> 	1,6	et les N2 ] <sub>AP</sub>  ADJ ]	1,6
G22	[ les N1 ] <sub>PW</sub> 	1,7	et les N2 ] <sub>ip</sub>  [ ADJ ]	1,9

Au vu du tableau récapitulatif, il semblerait que la différence significative entre les trois niveaux de frontière après N1 ( $G11 = G31 > G32 > G42 = G22$ ) soit corrélée à des allongements de la durée syllabique spécifiques. Le rapport de durée entre les syllabes inaccentuées de référence et les syllabes finales accentuées est en effet plus important en G11 et G31 que dans les autres groupements. Nous remarquons également que le rapport de durée est plus important en G32 qu'en G42 et G22. Ainsi, à l'instar de précédents travaux (entre autres, Astésano, 1998 ; 2001 ; Delais-Roussarie & Feldhausen, 2014 ; Delais-Roussarie et al., 2016 ; Fletcher, 1991 ; Rossi, 1985), nos résultats montrent que l'allongement de la durée varie proportionnellement en fonction de la force de la frontière. Plus précisément, dans le cas de notre corpus, la durée syllabique augmente d'environ 240% en frontière d'IP en G11 et G31, d'environ 110% en frontière d'AP en G32 et d'environ 65% en frontière de PW en G42 et G22.

De la même manière, il semble qu'il y ait une augmentation proportionnelle de l'allongement de la durée en fonction des niveaux de frontière perçus après N2. Nous constatons effectivement un rapport de durée plus important en G22, G32 et G31 qu'en G11 et G42. La durée syllabique augmente d'environ 85% en frontière d'ip en G22, G32 et G31, puis d'environ 40% en frontière d'AP en G11 et G42. Ainsi, à l'instar de Michelas & D'Imperio (2010), il semblerait que la frontière d'ip soit marquée par un allongement de la durée de la voyelle finale, qui est à la fois significativement plus important que celui observé en frontière d'AP, mais également moins important que celui observé en frontière d'IP. En comparant toutefois, sur le plan syntagmatique, les rapports de durée à la frontière après N2 avec ceux obtenus à la frontière après N1, il apparaît qu'en termes de rapport de durée, les frontières perçues après N2 dans tous les groupements relèvent à peu près du même niveau que la frontière de PW perçue après N1 en G42 et G22. Nous remarquons en effet des rapports de durée plus ou moins équivalents dans ces groupements. Or, nos résultats sur la force relative des frontières montrent que les frontières après N2 en G32, G42 et G22 sont perçues plus fortes que les frontières après N1. Il se pourrait alors dans ce contexte que le paramètre de durée ne soit pas suffisant pour expliquer la relation hiérarchique entre les niveaux de frontières perçues après N1 et N2 au sein du même groupement.

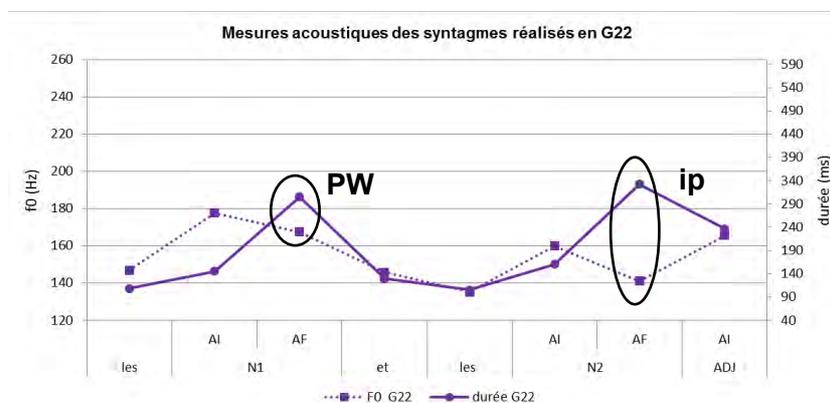
Plus précisément, nous supposons que le renforcement du niveau de frontière sur le plan syntagmatique soit dans ce cas-là compensé par les variations mélodiques. En effet, en s'intéressant dans un premier temps à G32, qui est marqué par une frontière d'AP après N1 et d'ip après N2, nous constatons que la durée de la syllabe finale est plus faible en frontière de N2 qu'en frontière de N1. En revanche, la chute de  $f_0$  sur la syllabe finale de N2 est plus importante que la chute de  $f_0$  sur la syllabe finale de N1 (cf. Figure N°84). Cela renverrait notamment au phénomène d'association inattendue des paramètres acoustiques que nous

avons observé dans les analyses descriptives, où un allongement de la durée syllabique conjugué à une chute de  $f_0$  permettrait, comme nous l'avons vu, de renforcer le niveau de frontière. Dans le cas de G32, le fait que la chute de  $f_0$  soit plus importante sur N2 que sur N1 permettrait alors de compenser l'allongement de la durée syllabique moins important sur N2 que sur N1, et ainsi renforcer le niveau de frontière d'ip après N2 en G32.



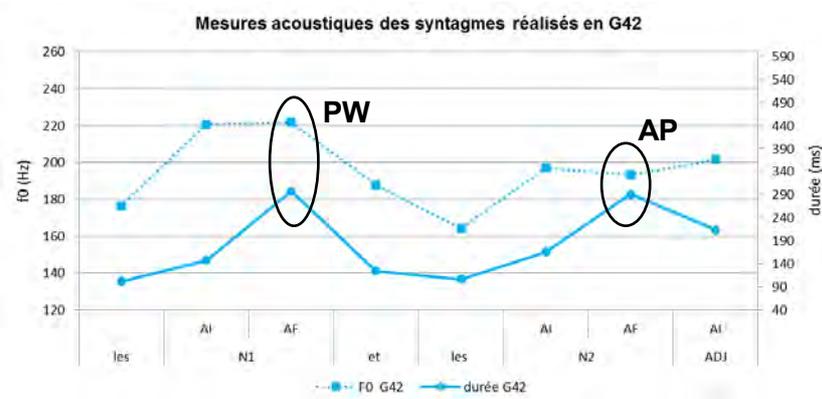
**Figure 84 :** Moyennes de  $f_0$  (Hz) et de durée (ms) sur chaque syllabe des syntagmes « les N1 et les N2 ADJ » réalisés en G32 (Condition 2), toutes tailles du constituant confondues

Ce phénomène s'observe également en G22, marqué par une frontière de PW après N1 et d'ip après N2, où cette fois-ci l'allongement de la durée est relativement équivalent entre les deux syllabes finales précédant les frontières. En revanche, la chute de  $f_0$  est là encore plus importante sur la syllabe finale de N2 que sur la syllabe finale de N1 (cf. Figure N°85). Ainsi comme dans le cas de G32, le fait que la chute de  $f_0$  soit plus importante sur N2 que sur N1 permettrait de renforcer le niveau de frontière d'ip après N2 en G22.



**Figure 85 :** Moyennes de  $f_0$  (Hz) et de durée (ms) sur chaque syllabe des syntagmes « les N1 et les N2 ADJ » réalisés en G22 (Condition 2), toutes tailles du constituant confondues

Enfin, en G42, qui est marqué par une frontière de PW après N1 et d'AP après N2, la durée des syllabes finales précédant les frontières est aussi équivalente entre les deux sites. Cependant, la valeur de  $f_0$  associée à la syllabe finale de N2 est plus faible que celle associée à la syllabe finale de N1 (cf. Figure N°86).



**Figure 86 :** Moyennes de  $f_0$  (Hz) et de durée (ms) sur chaque syllabe des syntagmes « les N1 et les N2 ADJ » réalisés en G42 (Condition 2), toutes tailles du constituant confondues

Il semblerait ainsi que le renforcement du niveau de frontière sur le plan paradigmatique soit essentiellement opéré par l’allongement de la durée, et que le renforcement du niveau de frontière sur le plan syntagmatique soit en revanche davantage lié aux variations mélodiques. Les chutes de  $f_0$  observées sur les syllabes finales de N2 permettraient en effet de compenser l’allongement de la durée syllabique moins important ou équivalent à l’allongement de la durée syllabique après N1, et de renforcer ainsi le niveau de frontière. Dans ce contexte, nous supposons que la force de la frontière après N2 augmente dès lors que la valeur de  $f_0$  sur la syllabe finale de N2 diminue. Ce résultat confirmerait ceux obtenus dans de précédentes études (notamment Hermès & Rump, 1994 ; Portes, 2002), où les auteurs constatent que les contours tonals descendants renforcent la perception d’une proéminence, et par conséquent la perception d’une frontière forte, tandis que les frontières plus faibles sont associées aux contours tonals montants. Il est également intéressant de mettre en lien nos résultats avec ceux obtenus dans l’étude en perception de Méndez & Astésano (2017) et Méndez (en cours), où les auteurs rendent compte d’une modulation de la force des frontières d’ip par l’AF descendant ou montant. Un AF marqué par une chute de  $f_0$ , plutôt que par une montée de  $f_0$ , tendrait selon eux à renforcer le degré de la frontière d’ip auquel il est associé. Toutefois, nos résultats montrent ici que la réalisation montante ou descendante des AF ne permet pas de moduler seulement le degré de frontière d’ip, mais

plutôt de distinguer, sur le plan syntagmatique, une frontière plus forte qu'une autre, à tous les niveaux de constituance. Ce phénomène pourrait également rappeler la règle « *d'inversion de pentes* » (Martin, 1981 ; 1987), selon laquelle tout constituant qui se joint à un autre constituant à sa droite pour former un groupe plus grand, doit être marqué par un contour intonatif de sens opposé à celui du constituant auquel il se joint.

Enfin, il est intéressant de préciser que la variation de  $f_0$  sur la syllabe initiale suivant la frontière n'intervient pas dans le renforcement du niveau de frontière, dans la mesure où une même valeur de  $f_0$  peut être observée pour différents niveaux de frontière. Nous observons en effet une valeur de  $f_0$  sur la syllabe initiale de N2 équivalente en G32 et G42 ( $f_0$  moyennes : 203,25 Hz en G32 et 197,04 Hz en G42) alors que la frontière après N1 est perçue plus forte en G32 (frontière d'AP) qu'en G42 (frontière de PW). Ce résultat tendrait alors à confirmer qu'un AI serait réalisé dès le niveau le plus bas dans la hiérarchie prosodique (PW). Nous avons toutefois besoin ici de l'analyse des corrélats acoustiques des proéminences perçues pour affiner davantage notre interprétation.

### **3.3.2 Corrélat acoustiques de la force des proéminences**

#### **3.3.2.1 Sites d'analyse**

Les résultats de notre analyse perceptive montrent que la force des proéminences à l'initiale et à la finale des mots lexicaux est globalement peu influencée par le groupement et par la taille du constituant. Nous avons vu en effet que l'influence de la taille du constituant sur la perception des proéminences ne répondrait pas directement à des contraintes d'ordre rythmique, dans la mesure où l'influence serait plutôt due à la « configuration syllabique » du constituant qu'à la taille seule du constituant. Certains résultats montrent d'ailleurs qu'AI est perçu plus fort sur les ADJ de 2 syllabes que sur les ADJ de 3 et 4 syllabes, alors que le N précédant est lui-même composé de 4 syllabes. Nous aurions en effet attendu d'un point de vue rythmique la perception d'un AI plus fort sur les ADJ composés de plus de 2 syllabes, pour ainsi faciliter la régulation rythmique. L'ensemble des résultats nous ont alors amenée à remettre en doute le rôle purement rythmique de l'accent initial en français. Nos résultats tendraient même à montrer un rôle plus structurel que rythmique de l'AI. Nous rappelons que seuls AF sur N1 et AI sur ADJ varient en fonction de la catégorie de groupement. Toutefois, seul AI sur ADJ présenterait une corrélation entre l'augmentation du niveau de frontière et le renforcement de la proéminence : AI en frontière d'ip est perçu plus fort qu'AI en frontière d'AP. AF sur N1, en revanche, serait perçu toujours avec la même force dès le niveau du

PW, et n'augmenterait pas strictement en fonction du niveau de frontière. Il semblerait par ailleurs que les PW dans nos syntagmes soient tous marqués par un AI et un AF, avec notamment un marquage initial aussi fort voire plus fort métriquement que le marquage final. Les résultats montrent en effet qu'AI est majoritairement perçu aussi fort ou plus fort qu'AF, quel que soit le niveau de frontière auxquels sont associées les proéminences, et quelle que soit la taille du constituant (*cf.* Figure N°73)

Il est donc intéressant de tester les corrélats acoustiques possibles associés non seulement à ces différents niveaux de proéminences, mais également aux différences de force relative entre les proéminences initiales et finales perçues.

➤ **Quelles sont les variations acoustiques pouvant favoriser la perception d'une proéminence plus forte qu'une autre?**

Pour répondre à cette question, nous proposons de tester l'effet des variations acoustiques sur la perception de la force de la proéminence finale sur N1 et initiale sur N2. Nous choisissons également de compléter ces premières analyses, en testant l'effet des variations acoustiques sur la perception de la force relative entre la proéminence initiale de N2 et finale de N1, mais aussi entre la proéminence initiale de ADJ et finale de N2. Nous rappelons que pour calculer la force relative entre la proéminence finale sur N1 et initiale sur N2, nous avons soustrait le score de la proéminence perçue sur la syllabe initiale de N2 à celui de la proéminence perçue sur la syllabe finale de N1. Nous obtenons ainsi 7 scores possibles, allant de -3 à 3. Une valeur de force relative négative signifierait que la proéminence sur la syllabe finale de N1 serait perçue comme plus forte que la proéminence sur la syllabe initiale de N2, tandis qu'une valeur de force relative positive montrerait que la proéminence sur la syllabe finale de N1 serait perçue comme plus faible que la proéminence sur la syllabe initiale de N2. Le même calcul a par ailleurs été effectué pour la force relative entre la proéminence initiale de ADJ et finale de N2. En ce qui concerne les variations acoustiques potentiellement corrélées à ces scores de proéminences, nous proposons de prendre les paramètres acoustiques suivants :

Force de la proéminence perçue sur la syllabe finale de N1 + Force relative entre la proéminence initiale sur N2 et finale sur N1 :

- La valeur de  $f_0$  sur la syllabe initiale de N1
- La durée de la syllabe finale de N1
- La valeur de  $f_0$  sur la syllabe finale de N1

- La valeur de  $f_0$  sur la syllabe initiale de N2

Force de la proéminence perçue sur la syllabe initiale de ADJ + Force relative entre la proéminence initiale sur ADJ et finale sur N2 :

- La durée de la syllabe finale de N2
- La valeur de  $f_0$  sur la syllabe finale de N2
- La valeur de  $f_0$  sur la syllabe initiale de ADJ
- La durée de la syllabe finale de ADJ
- La valeur de  $f_0$  sur la syllabe finale de ADJ

Le choix de ces paramètres a notamment été guidé par les résultats des analyses descriptives des principales variables acoustiques. En effet, nous décidons de prendre en compte les valeurs de  $f_0$  associées aux proéminences initiales et finales, étant donné que nous observons dans nos analyses une valeur de  $f_0$  qui peut être plus élevée, moins élevée ou équivalente entre les deux proéminences. Nous choisissons en revanche de prendre en compte la variable de durée uniquement dans le cas des proéminences finales, puisque nous constatons que la durée des syllabes initiales est systématiquement inférieure à celle des syllabes finales. Enfin, comme pour l'analyse des corrélats acoustiques des frontières, nous décidons d'effectuer ce modèle en séparant les deux conditions syntaxiques, pour pouvoir ainsi tester les différentes granularités des degrés de proéminence perçus dans chaque groupement. Pour rappel, la Condition 1 correspond à G11 et G31, tandis que la Condition 2 correspond à G32, G42 et G22.

Dans un deuxième temps, nous proposons également de tester l'effet de la valeur de  $f_0$  associée à la première, deuxième et troisième syllabe des mots lexicaux (N1, N2 et ADJ) sur la perception de la proéminence initiale (sur la première syllabe). Nous rappelons que nos analyses perceptives montrent que l'effet de la taille du constituant sur la perception des proéminences reste relativement faible. Nous avons d'ailleurs constaté que la force de la proéminence variait davantage en fonction de la manière dont N et ADJ se répartissent le nombre de syllabes dans le syntagme, qu'en fonction de la taille seule de ces constituants. De la même manière, nos statistiques descriptives sur les principales variables en perception montrent que les syllabes intermédiaires, ainsi que les syllabes des mots grammaticaux, ne présentent aucune relation de dépendance avec le groupement ou le nombre de syllabes du constituant.

Toutefois, en regardant sur le plan acoustique, nous remarquons que l'alignement des pics de  $f_0$  associés aux proéminences initiales varie en fonction de la taille du constituant. En effet, le pic maximal de  $f_0$  est aligné sur la première syllabe dans les mots

bisyllabiques, et sur la deuxième syllabe dans les mots composés de plus de 2 syllabes. L'ensemble de ces observations nous amènent donc à nous questionner quant à l'influence de la valeur de  $f_0$  sur la perception des proéminences initiales.

- **Est-ce que l'alignement du pic maximal de  $f_0$  sur la deuxième ou troisième syllabe du mot affaiblit le niveau de la proéminence perçue sur la première syllabe ? Et inversement, est-ce que l'alignement du pic maximal de  $f_0$  sur la première syllabe renforce le niveau de la proéminence perçue sur la première syllabe ?**

Afin de répondre à ces questions, nous choisissons de tester l'effet des variations acoustiques sur la perception des scores de proéminences sur la première syllabe de N1, de N2 puis de ADJ, en prenant en compte 3 paramètres acoustiques :

1. La valeur de  $f_0$  sur la première syllabe (de N1, N2 puis ADJ).
2. La valeur de  $f_0$  sur la deuxième syllabe (de N1, N2 puis ADJ).
3. La valeur de  $f_0$  sur la troisième syllabe (de N1, N2 puis ADJ).

Nous décidons par ailleurs d'effectuer ce modèle en prenant en compte les deux conditions syntaxiques et toutes les tailles du constituant confondues. En effet, le déplacement du pic maximal de  $f_0$  n'est pas influencé par la constituance prosodique. Quant aux tailles du constituant, étant donné que l'alignement du pic maximal de  $f_0$  dépend du nombre de syllabes, nous avons besoin de prendre en compte toutes les « configurations syllabiques » possibles pour faire ressortir des différences significatives.

Enfin, comme pour les corrélats acoustiques des scores de frontières, nous avons analysé pour chaque modèle réalisé les composantes de la variance qui tendent à montrer une valeur de la variance liée au participant plus élevée que celle liée au locuteur ou au syntagme. Ces résultats ont également été observés dans l'analyse perceptive des frontières et des proéminences, mais également dans l'analyse des corrélats acoustiques des scores de frontières. Il semblerait donc que les résultats obtenus dans le cadre de notre analyse varient davantage en fonction de l'auditeur qu'en fonction du locuteur produisant les différents syntagmes.

### 3.3.2.2 Force de la proéminence perçue sur la syllabe finale de N1

#### En Condition 1

Les résultats montrent un effet de la durée de la syllabe finale de N1 (LR Chisq= 15,8316, df= 1,  $p < 0,001$ ), ainsi qu'un effet de la valeur de  $f_0$  sur la syllabe finale de N1 (LR Chisq=12,5725, df= 1,  $p < 0,001$ ) et initiale de N2 (LR Chisq= 4,9972, df= 1,  $p = 0,02$ ). La valeur de  $f_0$  sur la syllabe initiale de N1 n'impacterait pas en revanche la force de la proéminence ( $p > 0,05$ ). Plus précisément, nous constatons à travers les coefficients de régression que l'augmentation de la durée de syllabe finale de N1 favorise les niveaux supérieurs des scores de proéminences (valeur du coefficient : 0,006316). Quant aux valeurs de  $f_0$ , l'augmentation de la  $f_0$  sur la syllabe finale de N1 favorise les niveaux supérieurs des scores de proéminences (valeur du coefficient : 0,015587), tandis que l'augmentation de la  $f_0$  sur la syllabe initiale de N2 favorise les niveaux inférieurs des scores de proéminences (valeur du coefficient : -0,015745) (cf. Annexe N°18-C).

- **Cela suppose alors que le renforcement du degré de proéminence sur la syllabe finale de N1 serait caractérisé par un allongement de la durée de la syllabe finale plus important, mais aussi par une valeur de  $f_0$  plus élevée sur la syllabe finale de N1 et moins élevée sur la syllabe initiale de N2.**

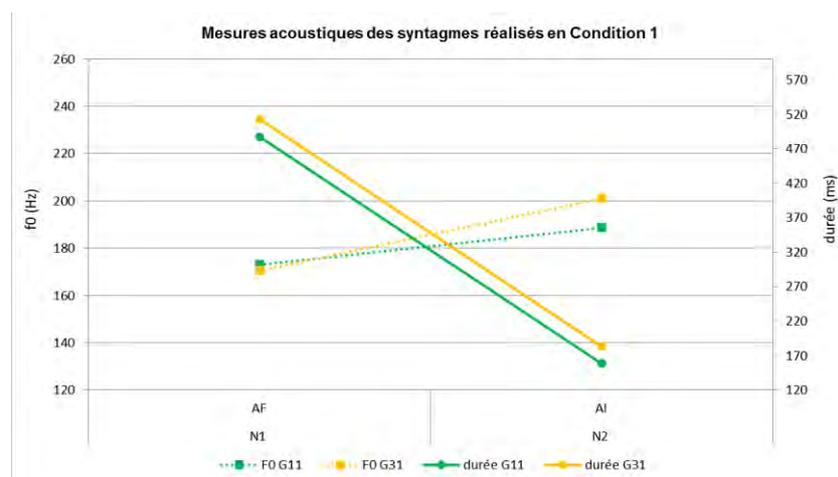
Les analyses perceptives des proéminences n'indiquent pas de différence significative entre G11 et G31 quant à la perception de la proéminence finale sur N1 (cf. Figure N°59). Nous savons cependant que le plus haut degré de proéminence est perçu en G11, et que celui perçu en G31 relève d'un niveau intermédiaire. En regardant la Figure N°87, il semblerait que la différence non significative entre G11 et G31 soit liée au fait qu'il n'y ait pas de différence forte entre les variations acoustiques sur ces deux groupements. Seule la valeur de  $f_0$  sur la syllabe initiale de N2 semble en revanche légèrement varier. La hauteur de  $f_0$  est en effet un peu plus élevée en G31 qu'en G11 ( $f_0$  moyennes : 188,60 Hz en G11 contre 201,01 en G31). Dans ce contexte, la nuance de force perçue entre les AF dans les deux groupements pourrait être uniquement liée aux variations de  $f_0$  réalisées sur la syllabe initiale de N2. A valeurs de  $f_0$  et de durée équivalentes sur la syllabe finale de N1, il semble en effet que plus la valeur de  $f_0$  associée à AI sur N2 soit élevée, plus AF sur N1 soit perçu comme faible.

Ce résultat est par ailleurs à mettre en lien avec les résultats de la force relative entre la proéminence initiale de N2 et finale de N1. Les résultats montrent un effet significatif de la

durée de la syllabe finale de N1 (LR Chisq= 4,4523, df= 1, p=0,03), ainsi qu'un effet de la valeur de  $f_0$  sur la syllabe finale de N1 (LR Chisq= 8,3345, df= 1, p=0,004) et initiale de N2 (LR Chisq= 11,4105, df= 1, p<0,001). A travers les coefficients de régression, nous comprenons plus précisément que l'augmentation de la durée de la syllabe finale de N1 favorise les niveaux inférieurs du score relatif des proéminences (valeur du coefficient : -0,003644). Quant aux valeurs de  $f_0$ , l'augmentation de la valeur de  $f_0$  sur la syllabe finale de N1 favorise les niveaux inférieurs du score relatif des proéminences (valeur du coefficient : -0,013772), tandis que l'augmentation de la  $f_0$  sur la syllabe initiale de N2 favorise les niveaux supérieurs du score relatif des proéminences (valeur du coefficient : 0,017644) (cf. Annexe N°18-D).

- **Il semblerait alors qu'AF soit perçu aussi fort ou plus fort qu'AI lorsque les valeurs de durée et de  $f_0$  augmentent sur la syllabe finale de N1. En revanche, dès lors que la valeur de  $f_0$  sur la syllabe initiale de N2 augmente, la différence entre les deux proéminences tendrait alors à s'effacer, voire à s'inverser.**

Les résultats de l'analyse perceptive de la force relative des proéminences ne montrent pas de différence significative entre G11 et G31. Nous savons cependant qu'AF est perçu plus fort qu'AI en G11, et que la différence entre les deux proéminences tend à diminuer en G31. En regardant la Figure N°87, nous constatons effectivement que les valeurs de durée et de  $f_0$  sur la syllabe finale de N1 sont à peu près égales entre les deux groupements, alors que la valeur de  $f_0$  sur la syllabe initiale de N2 est un plus élevée en G31 qu'en G11 ( $f_0$  moyennes : 201,01 Hz en G31 contre 188,61 Hz en G11). Ainsi, comme pour l'analyse de la force de la proéminence finale sur N1, seule l'augmentation de la valeur de  $f_0$  sur la syllabe initiale de N2 pourrait expliquer qu'AI et AF en G31 soient perçus avec une force relativement équivalente. A valeurs de  $f_0$  et de durée équivalentes sur la syllabe finale de N1, il semble en effet que plus la valeur de  $f_0$  associée à AI sur N2 soit élevée, plus AF sur N1 soit perçu aussi fort qu'AI sur N2 (cf. Figure N°87).



**Figure 87 :** Moyennes de  $f_0$  (Hz) et de durée (ms) associées à la syllabe finale de N1 et initiale de N2 en Condition 1

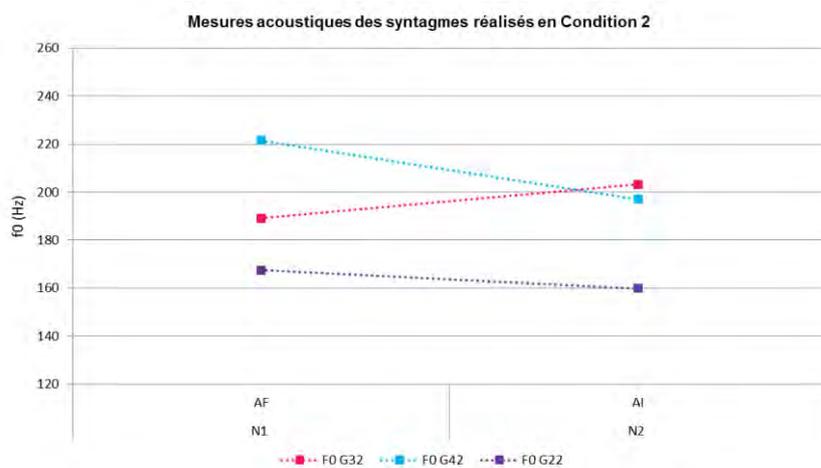
### En Condition 2

Les résultats ne montrent aucun effet significatif ( $p > 0,05$ ). La force de la proéminence finale sur N1 en Condition 2 ne varierait ni en fonction de la durée de la syllabe finale de N1, ni en fonction des valeurs de  $f_0$  associées à la syllabe initiale et finale de N1, et à la syllabe initiale de N2. En mettant en relation ce résultat avec celui obtenu dans les analyses perceptives des proéminences, nous constatons que le degré de proéminence est plus élevé en G22 qu'en G32 et G42 (cf. Figure N°59), sans qu'il y ait pour autant de différence significative entre les trois groupements. Ce résultat confirmerait qu'un seul niveau de proéminence ait été perçu sur la syllabe finale de N1 en Condition 2 dans les trois groupements.

En revanche, en testant l'effet des paramètres acoustiques sur la perception de la force relative entre la proéminence initiale de N2 et finale de N1, il semblerait que la différence de force entre les deux proéminences soit corrélée à des variations acoustiques spécifiques. Effectivement, les résultats montrent un effet significatif des valeurs de  $f_0$  sur la syllabe initiale de N2 (LR Chisq= 11,530, df= 1,  $p < 0,001$ ) et finale de N1 (LR Chisq= 16,678, df= 1,  $p < 0,001$ ). Plus précisément, les coefficients de régression indiquent que l'augmentation de la valeur de  $f_0$  sur la syllabe initiale de N2 favorise les niveaux supérieurs du score relatif des proéminences (valeur du coefficient : 0,014565), tandis que l'augmentation de la valeur de  $f_0$  sur la syllabe finale de N2 favorise les niveaux inférieurs du score relatif des proéminences (valeur du coefficient : -0,010689) (cf. Annexe N°18-D).

- Cela suppose alors qu'AF serait perçu aussi fort ou plus fort qu'AI lorsque la valeur de  $f_0$  est plus importante sur la syllabe finale de N1 que sur la syllabe initiale de N2. En revanche, dès lors que la valeur de  $f_0$  sur la syllabe initiale de N2 augmente, la différence entre les deux proéminences tendrait alors à s'effacer, voire à s'inverser.

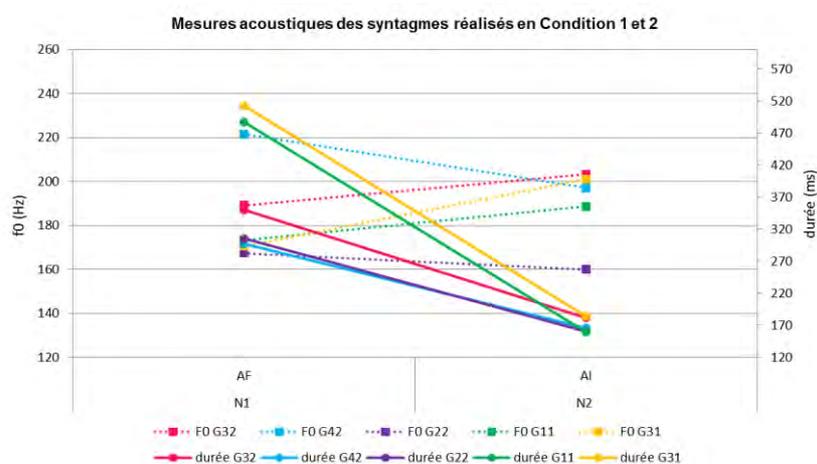
Les résultats de l'analyse perceptive ne montrent aucune différence significative entre G32, G22 et G42. Toutefois, nous savons qu'AI est perçu plus fort qu'AF en G32, et que la différence entre les deux proéminences tend à diminuer en G22 et G42. En regardant la Figure N°88, il semblerait effectivement que la valeur de  $f_0$  sur la syllabe initiale de N2 soit plus élevée que la valeur de  $f_0$  sur la syllabe finale de N1 en G32 ( $f_0$  moyennes : 203,25 Hz sur AI contre 189,13 Hz sur AF), et que la valeur de  $f_0$  sur la syllabe initiale de N2 soit moins élevée que la valeur de  $f_0$  sur la syllabe finale de N1 en G22 et G42 ( $f_0$  moyennes : G22 = 159,88 Hz sur AI contre 167,42 Hz sur AF ; G42 = 197,04 Hz sur AI contre 221,58 Hz sur AF). Ainsi, plus la valeur de  $f_0$  sur la syllabe initiale de N2 est élevée, plus AF sur N1 est perçu aussi fort qu'AI sur N2. La différence entre les deux proéminences serait donc ici principalement due à l'augmentation de la  $f_0$  sur la syllabe initiale de N2.



**Figure 88 :** Moyennes de  $f_0$  (Hz) associées à la syllabe finale de N1 et initiale de N2 en Condition 2

Les différences de force de la proéminence finale sur N1 entre les groupements semblent être principalement liées à la réalisation mélodique sur la syllabe initiale de N2. En effet, plus la valeur de  $f_0$  sur la syllabe initiale de N2 augmente, plus AF sur N1 est perçu comme

faible. De la même manière, en comparant la force relative entre AI sur N2 et AF sur N1, les résultats montrent là encore le rôle pertinent des variations mélodiques sur la syllabe initiale de N2. Toutefois, sur le plan perceptif, les différences observées entre les groupements dans chaque condition ne sont pas significatives. Dans ce contexte, nous pourrions penser que la variation mélodique sur la syllabe initiale de N2 permette de rendre compte de ces nuances de force, mais ne soit pas suffisante pour distinguer significativement les différentes forces de prééminences. En effet, lorsque nous mettons en relation tous les groupements ensemble, nous savons qu'AF sur N1 est perçu significativement plus fort en G11 qu'en G32 et G42.



**Figure 89 :** Moyennes de  $f_0$  (Hz) et de durée (ms) associées à la syllabe finale de N1 et initiale de N2 en Condition 1 et en Condition 2

En regardant la Figure N°89, nous constatons que la valeur de  $f_0$  sur la syllabe initiale de N2 n'est pas nettement différente entre les trois groupements, bien qu'elle soit un peu plus élevée en G32 et G42 qu'en G11 ( $f_0$  moyennes : G32= 203,25 Hz ; G42= 197,04 Hz ; G11= 188,61 Hz). Nous constatons en revanche une différence de durée sur la syllabe finale de N1. L'allongement de la durée syllabique est en effet plus important en G11 qu'en G32 et G42 (durées moyennes : G11= 487,42 ms ; G32= 350,07 ms ; G42= 296,93ms). Il se pourrait donc dans ce cas-là que la durée soit le paramètre le plus pertinent pour marquer la différence de force de la proéminence finale sur N1 à travers les trois groupements. Toutefois, les résultats perceptifs ne montrent aucune différence significative entre G31, G32 et G42, alors que G31 est marqué par un degré d'allongement syllabique aussi important qu'en G11 (durées moyennes : G11= 487,42 ms ; G31= 512,92 ms). De la même manière,

les résultats ne montrent aucune différence entre G11 et G22, alors que l'allongement de la durée syllabique en G22 est nettement plus faible qu'en G11 (durées moyennes : G11= 487,42 ms ; G31= 305,31 ms).

Il apparaît dans ce contexte qu'il n'y ait pas de corrélats acoustiques spécifiques pouvant expliquer la perception d'une proéminence finale sur N1 plus forte en G11 qu'en G32 et G42. Le renforcement de la proéminence finale sur N1 en G11 pourrait alors être davantage dû aux particularités individuelles des locuteurs. Comme nous l'avons déjà mentionné, les syntagmes réalisés en Condition 1, donc en G11 et G31, sont caractérisés par une pause relativement longue après N1. Toutefois, les quatre locuteurs produisant les syntagmes ne réalisent pas forcément la même durée de pause. Il se pourrait donc que les syntagmes perçus en G11 soient associés à des pauses plus longues que les syntagmes perçus en G31, ce qui expliquerait que seul G11 se différencie des autres groupements. Il serait intéressant par la suite de contrôler directement la réalisation acoustique de cet AF sur N1 par locuteur.

### 3.3.2.3 Force de la proéminence perçue sur la syllabe initiale de ADJ

#### En Condition 1

Les résultats montrent un effet de la durée de la syllabe finale de N2 (LR Chisq= 4,0528, df= 1, p=0,04), ainsi qu'un effet de la valeur de  $f_0$  sur la syllabe finale de N2 (LR Chisq= 10,9168, df= 1, p<0,001) et de ADJ (LR Chisq= 4,6837, df= 1, p=0,03). La valeur de  $f_0$  sur la syllabe initiale de ADJ n'impacte pas en revanche la force de la proéminence (p>0,05). A travers les coefficients de régression, nous constatons que l'augmentation de la durée de la syllabe finale de N2 favorise les niveaux inférieurs des scores de proéminences (valeur du coefficient : -0,0017923). Quant aux valeurs de  $f_0$ , l'augmentation de la valeur de  $f_0$  sur la syllabe finale de N2 favorise les niveaux inférieurs des scores de proéminences (valeur du coefficient : -0,0107547), tandis que l'augmentation de la valeur de  $f_0$  sur la syllabe finale de ADJ favorise les niveaux supérieurs des scores de proéminences (valeur du coefficient : 0,0033854) (cf. Annexe N°18-E).

- **Cela suppose alors que la proéminence initiale sur ADJ serait perçue plus forte lorsque les valeurs de durée et de  $f_0$  sur la syllabe finale de N2 diminuent, et que la valeur de  $f_0$  sur la syllabe finale de ADJ augmente.**

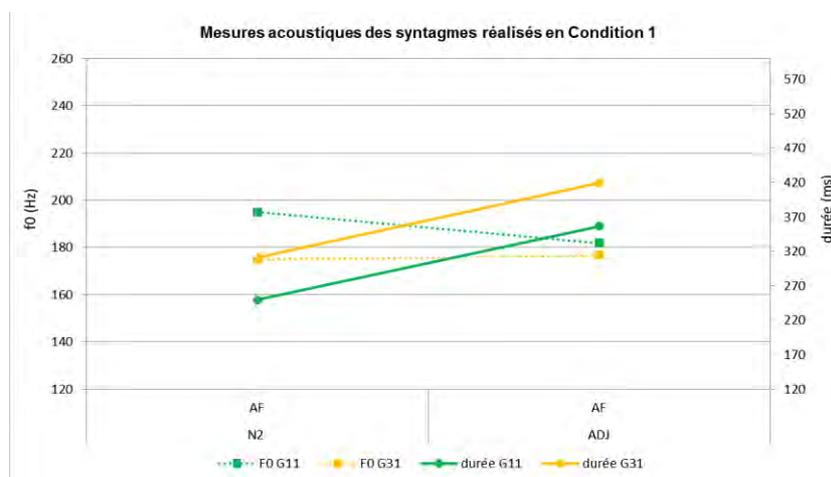
Les résultats de l'analyse perceptive ne montrent pas de différence significative entre G11 et G31 quant à la perception de la proéminence initiale sur ADJ (cf. Figure N°61). Nous savons cependant que le plus faible degré de proéminence est perçu en G11, et que celui perçu en G31 relève d'un niveau intermédiaire. En regardant la Figure N°90, il semblerait que la différence non significative entre G11 et G31 soit liée au fait qu'il n'y ait pas de différences de variation nettement marquées entre les deux groupements. La valeur de  $f_0$  sur la syllabe finale de N2 est en effet plus faible en G31 qu'en G11 ( $f_0$  moyennes : 194,80 Hz en G11 contre 174,78 Hz en G31), tandis que la valeur de  $f_0$  sur la syllabe finale de ADJ est quasi-équivalente entre les deux groupements ( $f_0$  moyennes : 181,73 Hz en G11 et 176,56 Hz en G31). Par ailleurs, l'allongement de la durée de la syllabe finale est plus important en G31 qu'en G11 (durées moyennes : 249,24 ms en G11 contre 310,80 ms en G31), ce qui semble en contradiction avec l'effet de l'augmentation de la durée de la syllabe sur l'affaiblissement de la proéminence initiale. Dans ce contexte, la perception d'une proéminence plus forte sur la syllabe initiale de ADJ en G31 serait ici principalement liée à l'abaissement de la  $f_0$  sur la syllabe finale de N2.

Ce résultat serait par ailleurs confirmé par ceux obtenus sur les corrélats acoustiques associés à la perception de la force relative entre AF sur N2 et AI sur ADJ. Les résultats montrent un seul effet de la valeur de  $f_0$  associée à la syllabe finale de N2 (LR Chisq= 5,5776, df= 1, p=0,01). Les valeurs de durée sur la syllabe finale de N2 et de  $f_0$  sur la syllabe initiale et finale de ADJ n'impacteraient pas en revanche la force de la proéminence ( $p>0,05$ ). En regardant les coefficients de régression, il semblerait que l'augmentation de la valeur de  $f_0$  sur la syllabe finale de N2 favorise les niveaux inférieurs du score relatif des proéminences (valeur du coefficient : -0,011852) (cf. Annexe N°18-F).

- **AF sur N2 serait alors perçu plus fort ou aussi fort qu'AI sur ADJ lorsque la hauteur de  $f_0$  sur la syllabe finale de N2 est plus élevée. A l'inverse, dès lors que la valeur de  $f_0$  sur la syllabe finale de N2 diminue, la différence entre les deux proéminences tendrait à s'effacer, voire à s'inverser.**

Les résultats de l'analyse perceptive de la force relative des proéminences ne montrent pas de différence significative entre G11 et G31. Nous savons cependant qu'AI sur ADJ est toujours perçu plus fort qu'AF, avec toutefois une différence entre les deux proéminences moins marquée en G11 qu'en G31. En regardant la Figure N°90, nous constatons que la valeur de  $f_0$  sur la syllabe finale de N2 est effectivement plus élevée en G11 qu'en G31 ( $f_0$

moyennes : 194,80 Hz en G11 contre 174,77 Hz en G31). Ainsi, plus la valeur de  $f_0$  sur la syllabe finale de N2 est élevée, plus AF sur N2 est perçu aussi fort qu'AI sur ADJ.



**Figure 90 :** Moyennes de  $f_0$  (Hz) et de durée (ms) associées aux syllabes finales de N2 et ADJ en Condition 1

### En Condition 2

Les résultats montrent un effet significatif de la durée des syllabes finales de N2 (LR Chisq= 4,6236, df= 1, p=0,03) et ADJ (LR Chisq= 8,8458, df= 1, p=0,002), ainsi qu'un effet des valeurs de  $f_0$  associées à la syllabe finale de N2 (LR Chisq= 16,7981, df= 1, p<0,001) et aux syllabes initiale et finale de ADJ (syllabe initiale : LR Chisq= 6,3191, df= 1, p=0,01 ; syllabe finale : LR Chisq= 12,7190, df= 1, p<0,001). Plus précisément, les coefficients de régression indiquent que l'augmentation de la durée de la syllabe finale de N2 (valeur du coefficient : -0,003777) et ADJ (valeur du coefficient : -0,005233) favorise les niveaux inférieurs des scores de proéminences. Quant aux valeurs de  $f_0$ , l'augmentation de la  $f_0$  sur la syllabe finale de N2 (valeur du coefficient : -0,013224) et ADJ (valeur du coefficient : -0,007480) favorise les niveaux inférieurs des scores de proéminences, tandis que l'augmentation de la  $f_0$  sur la syllabe initiale de ADJ favorise les niveaux supérieurs des scores de proéminences (valeur du coefficient : 0,008919) (cf. Annexe N°18-E).

- **La proéminence initiale sur ADJ devrait être perçue plus forte lorsque les valeurs de durée et de  $f_0$  sur les syllabes finales de N2 et ADJ sont faibles, mais aussi lorsque la valeur de  $f_0$  sur la syllabe initiale de ADJ est forte.**

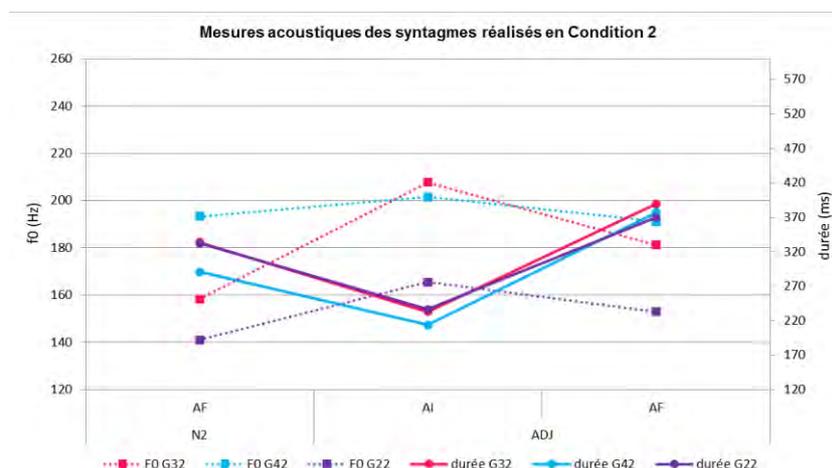
Les résultats de l'analyse perceptive des proéminences montrent que le degré de proéminence est perçu le plus fort en G22, le plus faible en G42 et avec un degré intermédiaire en G32 (cf. Figure N°61). En regardant la Figure N°91, il semblerait que les syllabes finales de N2 et ADJ soient réalisées avec des valeurs de  $f_0$  plus faibles en G22 qu'en G32 et G42 ( $f_0$  moyennes : N2 = 140,95 Hz en G22 contre 158,33 Hz en G32 et 193,23 Hz en G42 ; ADJ = 152,99 Hz en G22 contre 181,16 Hz en G32 et 191,16 Hz en G42). En revanche, l'allongement de la durée de la syllabe finale de N2 et ADJ varie peu entre les groupements (durées moyennes : N2= 332,43 ms en G22 ; 333,76 ms en G32 ; 290,60 ms en G42 ; ADJ= 369,80 ms en G22 ; 389,58 ms en G32 ; 377,04 ms en G42). De la même manière, la syllabe initiale de ADJ est marquée par une valeur de  $f_0$  nettement moins importante en G22 qu'en G32 et G42 ( $f_0$  moyennes : 165,49 Hz en G22 contre 207,66 Hz en G32 et 201,56 Hz en G42). Dans ce contexte, seules les variations de  $f_0$  sur les syllabes finales de N2 et ADJ semblent être corrélées à la perception d'une proéminence initiale plus forte en G22 que dans les autres groupements. Plus intéressant encore, AI sur ADJ est perçu plus fort en G22 alors que la valeur de  $f_0$  sur la syllabe initiale de ADJ est nettement plus faible en G22 que dans les deux autres groupements. Il se pourrait donc qu'AI sur ADJ soit toujours perçu aussi fort, quelle que soit la variation mélodique sur la première syllabe de ADJ.

De la même manière, les résultats sur les corrélats acoustiques de la force relative entre AI sur ADJ et AF sur N2 montrent un seul effet de la valeur de  $f_0$  associée à la syllabe finale de N2 (LR Chisq= 16,791, df= 1,  $p < 0,001$ ). Les valeurs de durée sur la syllabe finale de N2 et de  $f_0$  sur la syllabe initiale et finale de ADJ n'impacteraient pas en revanche la force de la proéminence ( $p > 0,05$ ). A travers les coefficients de régression, nous constatons que l'augmentation de la valeur de  $f_0$  sur la syllabe finale de N2 favorise les niveaux inférieurs du score relatif des proéminences (valeur du coefficient : -0,009794) (cf. Annexe N°18-F).

- **AF sur N2 serait perçu plus fort ou aussi fort qu'AI sur ADJ lorsque la hauteur de  $f_0$  sur la syllabe finale de N2 est plus élevée. A l'inverse, dès lors que la valeur de  $f_0$  sur la syllabe finale de N2 diminue, la différence entre les deux proéminences tendrait à s'effacer, voire à s'inverser.**

Les résultats de l'analyse perceptive de la force relative des proéminences montrent une seule différence significative entre G22 et G42. En revanche, aucune différence entre G22 et G32, ou entre G42 et G32 n'est observée. Nous savons cependant qu'AI sur ADJ est toujours perçu plus fort qu'AF sur N2, avec toutefois une différence entre les deux proéminences plus marquée en G22 qu'en G42 (cf. Figure N°69). En regardant la Figure

N°91, nous voyons effectivement que la valeur de  $f_0$  sur la syllabe finale de N2 est plus élevée en G32 et G42 qu'en G22 ( $f_0$  moyennes : 193,23 Hz en G42 ; 158,33 Hz en G32 ; 140,95 Hz). Ainsi, plus la valeur de  $f_0$  sur la syllabe finale de N2 est élevée, plus AF sur N2 est perçu aussi fort qu'AI sur ADJ.



**Figure 91** : Moyennes de  $f_0$  (Hz) et de durée (ms) associées à la syllabe finale de N2, et aux syllabes initiale et finale de ADJ en Condition 2

Il semblerait donc que le renforcement de la proéminence initiale sur ADJ soit essentiellement corrélé aux variations mélodiques sur la syllabe finale de N2. En effet, plus la valeur de  $f_0$  sur la syllabe finale de N2 est élevée, plus AI sur ADJ est perçu comme faible, et à l'inverse, plus la valeur de  $f_0$  sur la syllabe finale de N2 est faible, plus AI sur ADJ est perçu comme fort. La valeur de  $f_0$  sur la syllabe initiale de ADJ n'interviendrait donc pas dans le renforcement de la proéminence. Nous pouvons supposer dans ce contexte qu'AI sur ADJ peut être perçu comme fort y compris lorsque la valeur de  $f_0$  sur la syllabe initiale de ADJ est faible. Nous voyons en effet en G22, une valeur de  $f_0$  sur la syllabe initiale de ADJ plus faible que dans les autres groupements alors que d'un point de vue perceptif, c'est en G22 qu'AI sur ADJ est perçu le plus fort. Ce résultat pourra par ailleurs être affiné avec l'analyse suivante.

### 3.3.2.4 Force de la proéminence perçue sur les syllabes initiales des mots lexicaux

Les résultats de l'analyse ne montrent aucun effet significatif sur la perception de la proéminence initiale de N1, N2 ou ADJ ( $p > 0,05$ ). Autrement dit, la force de la proéminence initiale ne varie ni en fonction de la hauteur de  $f_0$ , ni en fonction de la position du pic de  $f_0$ . Effectivement, l'augmentation ou la diminution de la valeur de  $f_0$  sur la première, deuxième ou troisième syllabe des mots lexicaux n'impacte pas la force de la proéminence initiale. En admettant que la position de l'AI est variable dans le mot lorsque le nombre de syllabes augmente (Jun & Fougeron, 2000 ; 2002 ; Welby, 2003), nous pourrions attendre que l'augmentation de la valeur de  $f_0$  sur la deuxième ou troisième syllabe, lorsque le mot est composé de plus de 2 syllabes, favorise les niveaux inférieurs du score de proéminence sur la première syllabe des mots lexicaux. Cela supposerait en effet qu'AI soit perçu moins fort sur la première syllabe lorsque la valeur de  $f_0$  sur la deuxième ou troisième syllabe augmente. Or, nos résultats ne montrent aucun effet des variations mélodiques sur la perception de scores de proéminences.

Il semblerait dans ce contexte que la proéminence soit systématiquement perçue sur la première syllabe quel que soit l'alignement du pic maximal de  $f_0$ , mais également que la proéminence initiale soit perçue avec la même force quelle que soit la variation mélodique réalisée sur la première syllabe.

### 3.3.2.5 Discussion

Les résultats sur les corrélats acoustiques de la perception des proéminences montrent que les paramètres acoustiques majoritairement impliqués dans la perception de la force des proéminences sont les valeurs de  $f_0$  associées à la syllabe accentuée, ainsi que la valeur de  $f_0$  sur la syllabe du mot lexical précédant ou suivant directement la syllabe accentuée. Toutefois, seul AI sur ADJ pourrait être corrélé à des variations acoustiques spécifiques.

Les différentes forces de proéminence sur la syllabe finale de N1 ne seraient pas directement corrélées aux variations des paramètres acoustiques. Il se pourrait effectivement que les auditeurs se soient basés sur les variations acoustiques réalisées sur la syllabe finale de N1 uniquement pour le marquage des niveaux de frontière après N1. Nous rappelons effectivement que l'allongement de la durée de la syllabe finale de N1 est proportionnel au niveau de frontière marqué après N1. De fait, les différences d'allongement observées sur la syllabe finale de N1 seraient essentiellement corrélées à la perception des niveaux de frontière. En revanche, aucune corrélation de ce type ne peut être envisagée

pour le marquage du degré de proéminence finale, dans la mesure où AF sur N1 peut être perçu avec la même force dans deux groupements marqués ou non par un allongement de la durée syllabique. En effet, AF est perçu aussi fort en G22 qu'en G11, alors que l'allongement de la durée syllabique est uniquement réalisé en G11. Cela conduirait alors à confirmer qu'AF serait perçu dès le niveau du PW, avec une force métrique qui ne varie pas strictement en fonction du niveau de constituance. Plus intéressant encore, la force métrique de la proéminence finale ne varierait pas en fonction des variations acoustiques réalisées sur la syllabe accentuée ou sur les syllabes autour de la syllabe accentuée. AF serait donc encodé phonologiquement et serait systématiquement « attendu » par les auditeurs francophones en fin de mot lexical. Cela confirmerait par ailleurs des résultats en cours sur une étude en neuro-imagerie (te Rietmolen, en cours) montrant qu'AF n'est pas seulement perçu, mais qu'il est également anticipé par les auditeurs comme faisant partie de la représentation abstraite du mot lexical.

Quant au renforcement du degré de la proéminence initiale sur ADJ, les résultats montrent une influence importante de la valeur de  $f_0$  sur la syllabe finale sur N2. Il semblerait plus précisément que l'abaissement de la  $f_0$  sur la syllabe finale de N2 favorise la perception d'un AI sur ADJ plus fort, qui favoriserait par ailleurs la perception d'une frontière plus forte. Nous rappelons en effet que nous observons une corrélation entre le renforcement du niveau de frontière après N2 avec l'augmentation du degré de proéminence sur la syllabe initiale de ADJ. A travers nos analyses acoustiques, nous constatons en effet que la valeur de  $f_0$  sur la syllabe finale de N2 est plus faible en G22, G31 et G32 ( $f_0$  moyennes : G22= 140,95 Hz ; G31= 174,77 Hz; G32= 158,33 Hz), qui sont marqués par une frontière d'ip, qu'en G11 et G42 ( $f_0$  moyennes : G11= 194,80 Hz ; G42= 193,23 Hz), qui sont marqués par une frontière d'AP. En revanche, aucun effet de la valeur de  $f_0$  associée à la syllabe initiale de ADJ n'est observé. Plus intéressant encore, en comparant le groupement où AI est perçu le plus fort (G22) avec les groupements où AI est perçu le moins fort (G11 et G42), nous voyons que la valeur de  $f_0$  sur la syllabe initiale de ADJ est plus élevée dans les groupements où AI est perçu le moins fort, en G11 et G42, tandis que le groupement présentant le degré de proéminence le plus élevé est marqué par une valeur de  $f_0$  beaucoup plus faible, en G22 (cf. Figure N°89). Il semblerait alors que la proéminence initiale sur ADJ puisse être perçue comme forte même en l'absence d'une montée de  $f_0$  sur la syllabe initiale. De la même manière, l'analyse de la force relative entre AI sur ADJ et AF sur N2 montre uniquement une influence de la valeur de  $f_0$  associée à la syllabe finale de N2. Plus la valeur de  $f_0$  sur la syllabe finale de N2 est élevée, plus AF sur N2 est perçu aussi fort qu'AI sur ADJ. Là encore, la hauteur de  $f_0$  sur la syllabe initiale de ADJ n'impacte pas la force de la proéminence initiale perçue. Ce phénomène pourrait alors rappeler les propositions de Ladd (1996/2008),

selon lesquelles une proéminence qui est métriquement forte peut ne pas être accentuée en structure de surface. Nous ne considérons pas toutefois que la proéminence initiale n'est pas accentuée en structure de surface mais plutôt qu'elle n'est pas systématiquement corrélée à une montée de  $f_0$ . Ainsi, comme pour AF sur N1, nous pouvons supposer qu'AI sur ADJ est encodé phonologiquement et systématiquement « attendu » par les auditeurs francophones en début de mot lexical. Ce résultat confirmerait par ailleurs les résultats de précédentes études en neuro-imagerie (Aguilera et al., 2014 ; te Rietmolen et al., 2016 ), montrant que les auditeurs préfèrent les mots marqués par un pattern accentuel comprenant un AI. De la même manière, les résultats d'une seconde étude révèlent qu'AI est encodé au niveau lexical et facilite le traitement sémantique (te Rietmolen et al., 2017). Au vu de l'ensemble de ces résultats, il semblerait en effet que le mot lexical en français soit marqué par une bipolarisation accentuelle en structure de surface, où les deux accents, à l'inverse de Di Cristo (1999a ; 1999b ; 2000) et à l'instar en revanche d'Astésano (2017), relèveraient tous les deux d'une même force métrique. Il ne serait donc plus question d'accents primaires ou secondaires.

Dans le même sens, nos résultats montrent que la force des AI n'est pas influencée par l'alignement du pic maximal de  $f_0$ . AI est en effet toujours perçu sur la première syllabe même lorsque le pic mélodique maximal est aligné sur la deuxième ou troisième syllabes des mots lexicaux. Il semblerait donc que ce ne soit pas la position du pic maximal de  $f_0$  qui est importante mais plutôt le saut de  $f_0$  entre la syllabe précédant AI et la syllabe accentuée. Nous rappelons d'ailleurs que le degré de la proéminence initiale sur ADJ est renforcé lorsque la  $f_0$  sur la syllabe précédente diminue, et par conséquent lorsque la différence entre la syllabe précédente et la syllabe initiale est plus importante. Ainsi, à l'instar d'Astésano et al. (2012 ; 2013) qui observaient déjà ce résultat de manière plus préliminaire dans leur étude en perception, nous pourrions dès lors contrecarrer les propositions selon lesquelles la position de la proéminence initiale en français serait variable (Jun & Fougeron, 2000 ; 2002), et confirmer qu'AI ne serait pas un « *marqueur flou de frontière* » (Welby, 2003) puisqu'il serait systématiquement réalisé et attendu sur la première syllabe des mots lexicaux. Seule la réalisation physique, et non perceptive, du pic de  $f_0$  peut varier d'un constituant à un autre, selon notamment la possibilité ou non d'ancrer un ton bas « L » au début du mot lexical. Plus précisément, lorsqu'un ton bas ne peut être ancré, le pic maximal de  $f_0$  est alors systématiquement réalisé sur la première syllabe du mot lexical.

## Conclusion et Perspectives

---

Ce travail avait pour objectif d'explorer l'organisation du phrasé prosodique en français, à travers plus précisément le lien entre l'accentuation et la constituance prosodique. Comme nous l'avons vu, l'identité des constituants prosodiques et le rôle phonologique de l'accent dans la langue restent aujourd'hui mal établis. Sur la base de notre cadre théorique répertoriant non seulement les grands courants théoriques à l'origine de la description du phrasé prosodique d'une langue, les études sur la réalité psycho-cognitive des niveaux de constituance, mais également les modèles du phrasé prosodique consacrés plus spécifiquement au français, nous avons fait le choix d'orienter notre recherche vers une approche à la croisée de la Théorie Métrique Autosegmentale (notamment Hirst & Di Cristo, 1998 ; Post, 2000) et de la Théorie Métrique et Fonctionnaliste (Delais, 1999 ; 2000 ; Di Cristo, 1999a ; 1999b ; 2004).

En effet, les problématiques liées à l'étude du français nous ont amenée à adopter une approche avant tout métrique, testée notamment via le prisme de la perception, en complément par ailleurs d'une approche métrique autosegmentale, dont les méthodes sont davantage centrées sur la réalisation physique ou acoustique des indices prosodiques. Le lien entre l'accentuation et la constituance nécessite effectivement une analyse métrique, visant à rendre compte du rôle phonologique de l'accent en français. Sur le plan méthodologique, nous proposons de nous fonder sur la réalité perceptive des événements prosodiques, en dirigeant l'attention des auditeurs non plus sur la réalité phonétique de surface, mais davantage sur le poids métrique de ces événements. Autrement dit, « *extraire du signal physique les informations les plus pertinentes* » (Di Cristo, 2004, p. 83) intervenant dans l'élaboration du phrasé prosodique. Dans un deuxième temps, ces événements prosodiques perçus sont mis en relation avec leur réalité physique ou acoustique, permettant ainsi de faire le lien entre la réalité phonologique et la réalité de surface des différents événements prosodiques.

Sur la base d'un corpus de parole contrôlée, composé de structures D1+N1+cc+D2+N2+ADJ syntaxiquement ambiguës et variant en nombre de syllabes, nous proposons trois tâches de perception différentes, chacune effectuée par 80 auditeurs francophones natifs : une tâche de groupement, de prééminence et de frontière. A travers ce protocole, nous nous intéressons à la manière dont les auditeurs francophones segmentent et hiérarchisent le flux de parole. Comment les auditeurs découpent-ils le flux de parole en groupes de mots (tâche de groupement) ? Et comment les frontières et les prééminences

ont-elles été perçues au sein de chaque groupement ? Nous souhaitons également mettre en relation les scores de frontières et de proéminences avec la taille du constituant. Il est en effet intéressant de tester l'influence que peut avoir le nombre de syllabes du constituant sur la perception de la force des frontières et des proéminences. Nous rappelons que la taille des N et ADJ dans notre corpus est incrémentée d'une à quatre syllabes, dans toutes les combinaisons possibles, excepté pour N1 et N2 dont les longueurs covarient systématiquement. Chaque longueur de N apparaît donc avec chacune des 4 longueurs possibles de ADJ.

Par ailleurs, nous nous distinguons de précédents travaux en perception quant à la manière dont nous envisageons l'analyse des frontières et des proéminences. La plupart des études (entre autres, Astésano et *al.*, 2012 ; 2013 ; Cole et *al.*, 2010a ; 2014 ; Roux et *al.*, 2016 ; Smith, 2009) testent en effet la perception des proéminences et des frontières dans un énoncé sur le plan paradigmatique seulement, en s'intéressant à une catégorie d'événements prosodiques spécifique (frontière majeure, accent nucléaire, etc.), et sans réellement mettre en relation ces différents événements avec les autres événements prosodiques adjacents. Dans le cadre de notre travail, nous privilégions une étude testant la perception des frontières et des proéminences à la fois sur le plan paradigmatique et syntagmatique, mettant en relation, au sein d'un même énoncé ainsi qu'à travers les différents énoncés, les proéminences avec les frontières, mais également les différents niveaux de frontière ou de proéminence entre eux. A l'instar par ailleurs d'autres études en perception (Carlson et *al.*, 2005 ; de Pijper & Sanderman, 1994 ; Krivokapic & Byrd, 2012 ; Swerts, 1997), nous nous intéressons à la capacité des auditeurs à percevoir non seulement la présence et l'absence de frontières et de proéminences, mais surtout leur capacité à percevoir une granularité différentes de degrés de proéminence et de frontière.

Ce protocole expérimental vise plus précisément à questionner un certain nombre de phénomènes, notamment liées aux niveaux de constituance qui sont pertinents dans l'élaboration du phrasé prosodique en français. Nous nous interrogeons donc sur l'utilité du niveau du Mot Prosodique (noté « PW ») et du Syntagme Intermédiaire (noté « ip ») dans la hiérarchie des constituants prosodiques. Nous tenons également à rendre compte du rôle phonologique de l'accent en français, et plus spécifiquement dans le marquage de la constituance prosodique. Nous remettons par ailleurs en question l'affaiblissement perceptif de l'Accent Final (noté « AF ») en frontières prosodiques majeures, ainsi que le rôle purement rythmique, secondaire et optionnel de l'Accent Initial (noté « AI »). Afin de répondre à ces questions, nous avons procédé à un modèle de régression ordinale à effet mixte, permettant de tester la manière dont les proéminences et les frontières ont été perçues, en fonction du type de groupement choisi par les auditeurs, mais également en

fonction du nombre de syllabes du constituant. Ce modèle a également été utilisé pour tester l'influence des variations acoustiques sur la perception des scores de proéminences et de frontières. A l'issue de l'ensemble de ces analyses, plusieurs résultats ont émergé, nous entraînant à remettre en question non seulement la granularité des niveaux de frontière perçus par les auditeurs, mais également l'affaiblissement perceptif de l'AF en frontière majeure, ainsi que le statut secondaire et optionnel de l'AI en français.

➤ **Mise en évidence de la pertinence des niveaux du PW et de l'ip dans l'élaboration du phrasé prosodique en français :**

Nous avons émis l'hypothèse tout d'abord que le groupement était fortement corrélé à la condition syntaxique, et qu'il permettait notamment d'affiner les prédictions initiales faites sur les niveaux de constituance associés à chacune des deux conditions. Les tests du Khi deux confirment en effet une forte relation de dépendance entre le groupement et la condition syntaxique, ce qui nous a amenée par la suite à recoder notre base de données en associant à chaque catégorie de groupement l'information sur la condition syntaxique (G11, G31, G32, G22 et G42). Selon nous, ces différents groupements rendraient compte d'une granularité plus fine des niveaux de frontière. Nous avons supposé au départ trois niveaux de frontière différents : ip, AP et PW. Toutefois, l'analyse de la force des frontières, via le modèle de régression ordinaire à effet mixte, met en évidence quatre niveaux de frontière différents perçus par les auditeurs : IP, ip, AP et PW.

Nos résultats montrent donc dans un premier temps l'utilité d'intégrer le niveau du PW (*i.e* mot lexical : Selkirk, 1996) dans la hiérarchie des constituants prosodiques en français. En effet, les auditeurs de notre expérimentation ont eu besoin de recourir au niveau du mot pour segmenter et hiérarchiser le flux de parole, en marquant notamment un niveau de frontière hiérarchiquement inférieur au Syntagme Accentuel (noté « AP »), pourtant considéré dans la majorité des modèles du phrasé comme l'unité phonologique de base du français (entre autres, Jun & Fougeron, 2000). En effet, aucun des modèles du phrasé prosodique n'envisage de prendre en compte le mot comme une unité de planification ou de réalisation des règles accentuelles en structure de surface. De fait, le Mot Prosodique, en tant que mot lexical, n'est jamais inclus dans la hiérarchie des constituants prosodiques. Seule Astésano (2017), suite à un certain nombre d'études en production (Astésano et al., 2007), en perception (Astésano et al., 2012 ; 2013) mais également en neuro-imagerie (Aguilera et al., 2014 ; te Rietmolen et al., 2016 ; 2017), propose de prendre en compte le niveau du mot lexical, ou Mot Prosodique, comme l'unité métrique de base en français. Nos résultats tendraient alors à consolider ce postulat, en montrant notamment l'intérêt du niveau

du mot en structure de surface, dans la segmentation et la hiérarchisation du flux de parole. Plus intéressant encore, l'analyse de la force des prééminences mais également des corrélats acoustiques des scores de prééminences montre que les mots lexicaux sont systématiquement marqués par une bipolarisation accentuelle, où les deux accents relèveraient d'une même force métrique (Astésano, 2017). AI et AF seraient en effet perçus dès le niveau du PW, et encodé phonologiquement comme les marqueurs gauche et droite de la représentation abstraite du mot lexical (te Rietmolen, en cours ; Aguilera et al., 2014 ; te Rietmolen et al., 2016 ; 2017).

Le niveau de l'ip montre également son importance dans l'organisation du phrasé prosodique en français, en marquant notamment un niveau de frontière plus fort que l'AP et moins fort que l'IP (ou Syntagme Intonatif). Selon nous, l'ip ne serait pas une unité destinée à rendre compte de structures syntaxiques spécifiques (Di Cristo & Hirst, 1996 ; Jun & Fougeron, 2000), mais plutôt une unité pouvant apparaître sur tout type d'énoncé, dès lors que les structures syntaxiques et prosodiques le permettent (Michelas & D'Imperio, 2010). Par ailleurs, nous ne considérons pas l'ip comme un IP contenu dans un autre IP (Ladd, 1996 / 2008 ; Di Cristo, 2016), dans la mesure où nous observons déjà dans nos résultats un phénomène de récursivité de ce type, impliquant toutefois selon nous un niveau de frontière encore différent de celui de l'ip. Sur le plan acoustique, il apparaît également que le niveau de l'ip est marqué par un allongement de la durée de la voyelle finale, qui est à la fois significativement plus important que celui observé en frontière d'AP, mais également moins important que celui observé en frontière d'IP (Michelas & D'Imperio, 2010).

Il semblerait toutefois à travers l'analyse des corrélats acoustiques de la perception des frontières qu'il n'y ait pas de variations acoustiques spécifiques à chaque niveau de constituance, mais plutôt des variations acoustiques spécifiques au marquage de la relation hiérarchique entre les frontières. En effet, sur le plan paradigmatique, nous observons un allongement de la durée proportionnel au niveau de constituance. En revanche, en s'intéressant au plan syntagmatique, et donc à la relation hiérarchique entre les deux sites de frontières au sein d'un même groupement, il apparaît cette fois-ci possible que deux frontières associées à des niveaux différents soient marquées par un allongement de la durée relativement équivalent, voire même que le niveau de constituance le plus élevé soit marqué par un allongement de la durée syllabique moins important que le niveau de constituance le plus faible. Dans ce cas-là, la hiérarchie entre les deux frontières est davantage marquée par les variations mélodiques. Nous remarquons en effet un phénomène d'association inattendue des paramètres acoustiques, où la chute de  $f_0$  permet de compenser l'allongement de la durée syllabique, parfois moins important en frontière forte

qu'en frontière faible, et de renforcer ainsi le niveau de frontière (Hermès & Rump, 1994 ; Méndez, en cours ; Méndez & Astésano, 2017 ; Portes, 2002).

➤ **Mise en évidence d'une dissociation totale entre la perception de la proéminence finale et du niveau de frontière.**

L'augmentation de ces niveaux de frontière ne serait pas en revanche corrélée à la force de la proéminence finale sur N1 et N2. Bien que les résultats montrent un effet du groupement sur la perception de la proéminence finale sur N1, il apparaît cependant que la force de l'AF ne varierait pas strictement en fonction du niveau de constituance, puisqu'une même force de proéminence peut être perçue à des niveaux de frontière différents. La proéminence finale sur N1 serait donc perçue toujours avec la même force dès le niveau du PW. Quant à N2, étant donné que les résultats ne montrent aucune influence du groupement sur la perception de la proéminence finale, il semblerait là encore que la force de l'AF sur N2 ne varie pas directement en fonction du niveau de frontière.

Ainsi, alors qu'Astésano et *al.* (2012 ; 2013) proposent « *une dissociation partielle* » entre la perception des proéminences et des frontières, nous rendons compte dans nos résultats d'une dissociation totale. En effet, les auteurs parlent de « *dissociation partielle* » dans la mesure où un même score de proéminence peut correspondre à différents niveaux de frontières, à l'exception de la frontière d'ip (qui selon nos résultats serait plutôt une frontière d'IP) où un renforcement de l'accent final est observé. Ce renforcement serait lié selon les auteurs à la réalisation d'une longue pause après N1. Nous observons également un renforcement de la proéminence finale à la frontière d'ip (ou d'IP selon nous) en G11 uniquement, et non pas en G31 qui est pourtant marqué par une frontière du même type. Ainsi, à l'instar des auteurs, nous supposons que ce renforcement serait davantage dû à des particularités individuelles des locuteurs. Il serait donc intéressant dans le cadre d'une prochaine étude de vérifier cette hypothèse, en contrôlant directement la réalisation des pauses par locuteur. En effet, les quatre locuteurs de notre corpus ne réalisent pas forcément la même durée de pause. Il se pourrait donc que les syntagmes perçus en G11 soient associés à des pauses plus longues que les syntagmes perçus en G31. Cela pourrait expliquer le renforcement de la proéminence finale uniquement en G11. Notre analyse des scores de frontières au regard des différents types de groupements, permettent alors d'affiner les résultats d'Astésano et *al.* (2012 ; 2013) en montrant que la proéminence finale peut être perçue aussi forte en frontière d'ip ou d'IP qu'en frontière d'AP ou de PW. L'AF en français ne s'effacerait pas perceptivement en frontière prosodique majeure (*contra* Rossi,

1980 ; Jun & Fougeron, 2000 ; 2002), et garderait au contraire une trace métrique du niveau du mot lexical toujours aussi forte.

Par ailleurs, l'analyse des corrélats acoustiques montre qu'AF sur N1 n'est pas corrélé aux variations des paramètres acoustiques. Nous constatons en effet qu'AF peut être perçu avec la même force dans deux groupements marqués par des réalisations acoustiques très différentes. Il se pourrait donc que les variations acoustiques observées sur la syllabe finale de N1, telles que l'allongement de la durée, soient uniquement liées au marquage des niveaux de frontières après N1. Cela confirmerait là encore l'encodage phonologique de la proéminence finale en français. En effet, AF ne serait pas seulement perçu mais il serait également anticipé par les auditeurs comme faisant partie de la représentation abstraite du mot lexical (te Rietmolen, en cours).

➤ **Mise en évidence du rôle plus structurel que purement rythmique de la proéminence initiale :**

A l'inverse, la force de la proéminence initiale pourrait davantage varier en fonction du niveau de frontière. En effet, comme nous l'avons supposé dans nos hypothèses de départ, il y aurait une corrélation entre le renforcement de la proéminence initiale et le marquage du niveau de frontière à laquelle la proéminence est associée (similaire au renforcement articulaire démontré par Fougeron & Keating, 1997). AI jouerait alors un rôle pertinent dans le marquage de la structure prosodique. Toutefois, nous rappelons que seules les proéminences initiales réalisées sur les mots lexicaux situés directement après la frontière montrent une variation proportionnelle du degré de proéminence, en fonction du niveau de frontière. Dans le cas des mots lexicaux précédés d'un ou de plusieurs mots grammaticaux, la proéminence initiale marque le niveau du mot lexical mais ne varie pas en fonction du niveau de constituance. De plus, les résultats de l'effet de la taille du constituant sur la perception de la force des proéminences initiales ne nous permettent pas de rattacher d'interprétations rythmiques à ces effets, et tendent plutôt à montrer une influence de la « configuration syllabique » du constituant. La proéminence initiale n'apparaît donc pas seulement sur les longs constituants (*cf.* notion « *d'ictus mélodique* » de Rossi, 1980). AI est en effet perçu dès les constituants composés de 2 syllabes, y compris lorsque d'un point de vue rythmique le renforcement de la proéminence initiale est attendu sur les constituants plus grands.

Ainsi, comme nous avons vu par ailleurs que la force de la proéminence finale n'était pas corrélée aux niveaux de frontières, nous pouvons alors considérer, à l'instar de l'étude

en production d'Astésano et *al.* (2007), que la proéminence initiale marque la structure prosodique de manière plus privilégiée que la proéminence finale. Toutefois, alors que les auteurs rendent compte du rôle structurel de l'AI au niveau de l'AP, nos résultats montrent qu'AI est perçu dès le niveau du PW.

Enfin, au-delà de questionner le rôle purement rythmique de la proéminence initiale, nos résultats tendent également à mettre en doute son statut secondaire. Il s'avère en effet que la proéminence initiale est majoritairement perçue aussi forte ou plus forte que la proéminence finale, quel que soit le niveau de frontière auquel les proéminences sont associées. Par ailleurs, nous constatons qu'AI est toujours perçu sur la première syllabe même lorsque le pic mélodique maximal est aligné sur la deuxième ou troisième syllabe des mots lexicaux. D'autres résultats montrent aussi que le degré de la proéminence initiale est renforcé lorsque la  $f_0$  sur la syllabe finale du mot lexical précédant diminue. Il semblerait dans ce contexte que ce ne soit pas le pic maximal de  $f_0$  qui influence la perception de la proéminence sur la syllabe initiale, mais plutôt le saut de  $f_0$  entre la syllabe précédente et AI (Astésano, 2017 ; Astésano et *al.*, 2012 ; 2013). La position de la proéminence initiale en français ne serait donc pas variable (Jun & Fougeron, 2000 ; 2002 ; Welby, 2003), puisqu'AI serait systématiquement réalisé et attendu sur la première syllabe des mots lexicaux (Aguilera et *al.*, 2014 ; te Rietmolen et *al.*, 2016 ; 2017). Nos résultats acoustiques montrent par ailleurs que le pic maximal de  $f_0$  est systématiquement ancré sur la première syllabe des mots lexicaux lorsque le ton bas précédant ne peut s'ancrer au début du mot.

Plus intéressant encore, nos résultats sur les corrélats acoustiques montrent qu'AI peut être perçu comme fort même en l'absence d'une montée de  $f_0$  sur la syllabe accentuée. Nos résultats tendent là encore à consolider le postulat, selon lequel les mots lexicaux sont réalisés par un marquage bipolaire de même force métrique (Astésano, 2017) en structure de surface. Il ne serait donc plus question dans ce contexte d'accents primaires ou secondaires.

### ➤ **Perspectives de recherche**

L'ensemble de nos résultats questionnent donc le lien entre l'accentuation et la constituance dans l'élaboration du phrasé prosodique en français. Nous remettons par ailleurs en question la vision traditionnelle des phénomènes accentuels dans la langue, et par conséquent l'organisation de la hiérarchie des constituants prosodiques. Plus précisément, nos résultats se situent dans la lignée de précédents travaux en production et en perception (Astésano, 1998 ; 2001 ; 2017 ; Astésano et *al.* 2012 ; 2013), qui depuis un

certain temps maintenant démontrent la pertinence du niveau du mot lexical, ou du Mot Prosodique, en français à travers notamment le marquage des proéminences initiales. Dans le cadre de notre travail, nous avons choisi d'éprouver ces premiers résultats sur la base d'un plus grand corpus de parole, produit et testé sur un plus grand nombre de locuteurs et d'auditeurs. Notre recherche se distingue par ailleurs de ces précédents travaux concernant le recours à une analyse syntagmatique des événements prosodiques, via notamment notre tâche de groupement, qui a permis de confirmer l'utilité du niveau du PW et de l'ip dans la hiérarchisation du flux de parole, ainsi que la relation hiérarchique entre les proéminences initiales et finales. De la même manière, nous avons rendu compte que le paramètre de  $f_0$  pouvait venir jouer un rôle important dans le marquage des relations hiérarchiques entre les frontières, en compensant notamment le paramètre de durée. Il s'agit donc ici d'un « pas supplémentaire dans la démonstration » qui ouvre la voie maintenant à de nouvelles perspectives.

En effet, notre étude a mis en évidence des primitives prosodiques intéressantes qui restent toutefois démontrées sur de la parole contrôlée. Une des prochaines étapes de notre travail consistera alors à éprouver ces primitives prosodiques dans le cadre d'une parole plus naturelle. Comme nous l'avons déjà mentionné dans la présentation de notre matériel linguistique, deux corpus de parole semi-contrôlée et spontanée sont à notre disposition, sur lesquels un premier travail de dépouillement a été effectué. Nous pouvons donc envisager dans un premier temps de répliquer notre étude sur le corpus de la Map Task Aix (Bard et *al.*, 2013), qui est un corpus de parole semi-guidée reprenant les structures syntaxiquement ambiguës du corpus d'Edimbourg (Astésano et *al.* 2007) dans un contexte relativement plus naturel. Toutefois, comme nous l'avons vu, ce corpus nécessite des tests sémantiques préliminaires visant à vérifier que les locuteurs ont produit correctement les deux conditions syntaxiques prédites. Nous aimerions également travailler sur le corpus du CID (Bertrand et *al.*, 2008) qui est un corpus de parole spontanée. Cependant, nous n'envisageons pas de répliquer notre étude sur ce corpus, dans la mesure où nous avons vu qu'il ne se composait pas de structures syntaxiques comparables au corpus d'Edimbourg (Astésano et *al.*, 2007). En revanche, il serait intéressant de travailler sur les structures N+ADJ, nombreuses dans le corpus du CID (415 occurrences), qui pourraient nous permettre de tester plus précisément les phénomènes d'arcs accentuels. En effet, la structure en arc accentuel pourrait être un premier outil pour démontrer la trace métrique des proéminences initiales et finales au niveau du mot, à l'intérieur d'un groupe de mots en parole spontanée. Ces structures N+ADJ pourraient par ailleurs être facilement comparables aux structures N2+ADJ du corpus d'Edimbourg (Astésano, et *al.*, 2007). Le recours aux structures accentuelles nous permettrait également de mieux explorer les règles rythmiques en français, à travers

notamment les phénomènes de clash accentuels. De la même manière, ce travail sur les structures en arc accentuel permettrait de nous concentrer davantage sur les mots monosyllabiques que nous avons volontairement mis de côté dans notre étude, dans la mesure où nous voulions tester plus précisément la réalisation du marquage bipolaire AI+AF. Il semble toutefois pertinent de s'intéresser au marquage accentuel des mots monosyllabiques, notamment dans le cas d'un arc accentuel. A travers plus précisément des mesures perceptives et acoustiques, il serait intéressant de voir si l'accent réalisé sur le monosyllabique relève plus d'une proéminence initiale ou finale.

Par ailleurs, il serait également intéressant de repenser notre protocole expérimental concernant plus particulièrement la présentation orthographique des énoncés. Il est en effet possible d'imaginer que nos participants aient été influencés dans le choix des scores de frontières et de proéminences par la transcription orthographique. Leurs connaissances partagées sur la structuration syntaxique peuvent les pousser à ne pas percevoir de frontière forte entre le déterminant et le nom par exemple. Nous avons cependant fait le choix dans cette étude de proposer une échelle d'évaluation (de 0 à 3) plutôt qu'un format de réponse binaire, du type « présence ou absence », qui selon nous laisse l'opportunité aux participants de moduler suffisamment les forces de frontières et de proéminences perçues. De cette manière, même si les participants estiment qu'il n'est pas possible de percevoir une frontière entre un déterminant et un nom, mais qu'ils perçoivent malgré tout une petite force de frontière, ils peuvent l'indiquer en attribuant un très faible score de frontière après le déterminant. En revanche, il pourrait être intéressant d'utiliser, à l'instar d'autres études (entre autres, de Pijper & Sanderma, 1994 ; Simon & Christodoulides, 2016a ; 2016b), des structures délexicalisées ou des structures avec des logatomes qui conserveraient uniquement les variations prosodiques, et qui éviteraient ainsi toute influence du script orthographique. Cependant, le recueil des réponses sur une échelle d'évaluation serait dans ce contexte plus difficilement envisageable. De la même manière, nous pouvons nous interroger quant à l'intérêt de structures délexicalisées, dans la mesure où notre objectif est de tester le lien entre la force métrique avec le domaine grammatical. Il serait alors intéressant dans ce contexte d'envisager plutôt le recours à des langues artificielles (similaires aux études de Bagou et *al.*, 2002 ; Bagou & Frauenfelder, 2006).

Enfin, comme nous l'avons déjà mentionné à plusieurs reprises, nous nous intéressons également à l'effet des « configurations syllabiques » sur la perception des proéminences initiales. Nous rappelons que notre étude tend à montrer que l'AI serait davantage influencé par la manière dont les mots se répartissent le nombre de syllabes dans le constituant, que par la taille seule du constituant. Toutefois, notre matériel linguistique actuel ne permet pas de tester systématiquement cette hypothèse, et il serait donc

## Conclusion et Perspectives

---

intéressant de procéder à une étude sur la base d'un corpus manipulant la taille de tous les mots lexicaux du syntagme, et ainsi proposer plusieurs « configurations syllabiques » possibles pour une même taille totale du syntagme.

## Bibliographie

---

1. **Abercrombie, D.** (1967). *Elements of general phonetics*. London, UK: Aldine Pub. Co.
2. **Abney, S.** (1991). Parsing By Chunks. In R. Berwick, S. Abney, & C. Tenny (éd.), *Principle-Based Parsing. Studies in Linguistics and Philosophy* (Vol. 44, p. 257-278). Dordrecht, Netherlands: Springer Netherlands.
3. **Abney, S.** (1992). Prosodic structure, performance structure and phrase structure. In *Proceedings of the Workshop on Speech and Natural Language* (p. 425-428). Harriman, NY.
4. **Aguilera, M., El Yagoubi, R., & Astésano, C.** (2014). Event Related Potential investigation of Initial Accent processing in French. In *Proceedings of the 7th International Conference on Speech Prosody (SP 2014)* (p. 383-387). Dublin, Ireland.
5. **Akaike, H.** (1973). Information theory and an extension of the maximum likelihood principle. In *Proceedings of the 2nd International Symposium on Information Theory* (p. 267-281). Tsahkadsor, Armenia.
6. **Al-Mozainy, H., Bley-Vroman, R., & Mc Carthy, J. J.** (1985). Stress shift and metrical structure. *Linguistic Inquiry*, 16(1), 135-144.
7. **Anderson, A., Bader, M., Bard, E., Boyle, E., Doherty, G., Garrod, S., ... Weinert, R.** (1991). The HCRC Map Task Corpus. *Language and Speech*, 34(4), 351-366.
8. **Arvaniti, A., Ladd, D. R., & Mennen, I.** (2000). What is a Starred Tone? Evidence from Greek. In M. B. Broe & J. B. Pierrehumbert (éd.), *Acquisition and the lexicon* (p. 119-131). Cambridge, UK: Cambridge University Press.
9. **Astésano, C.** (1998). Effects of prosodic constraints on the differential lengthening of syllable constituents in French: a comparison between spontaneous and read speech. In *Actes du congrès SPoSS, les sons de la parole spontanée* (p. 143-146). La Baume-les-Aix, France.
10. **Astésano, C.** (2001). *Rythme et accentuation en français: invariance et variabilité stylistique*. Paris, France: L'Harmattan.
11. **Astésano, C.** (2017). *Le statut de l'Accent Initial dans la phonologie prosodique du français - enjeux descriptifs et psycholinguistiques* (Mémoire d'HDR). Université Toulouse 2 Jean-Jaurès.
12. **Astésano, C., Bard, E. G., & Turk, A.** (2007). Structural influences on initial accent placement in French. *Language and Speech*, 50(3), 423-446.
13. **Astésano, C., & Bertrand, R.** (2016). Accentuation et niveaux de constituance en français: enjeux phonologiques et psycholinguistiques. *Langue française*, 191(3), 11-30.

14. **Astésano, C., Bertrand, R., Espesser, R., & Nguyen, N.** (2012). Perception des frontières et des proéminences en français. In *Actes de la conférence conjointe JEP-TALN-RECITAL 2012: JEP* (Vol. 1, p. 353-360). Grenoble, France.
15. **Astésano, C., Bertrand, R., Espesser, R., & Nguyen, N.** (2013). Dissociation between prominence and boundary phenomena in French: a perception study. In *Proceedings of pS-prominenceS conference*. Viterbo, Italie.
16. **Avanzi, M.** (2007). Regards croisés sur la notion de macro-syntaxe. *Revue Tranel (Travaux neuchâtelois de linguistique)*, 47, 39-58.
17. **Avanzi, M., Lacheret-Dujour, A., & Victorri, B.** (2010). A corpus-based learning method for prominence detection in spontaneous speech. In *Proceedings of the 5th International Conference on Speech Prosody (SP 2010)* (p. 20-30). Chicago, Illinois.
18. **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 International Conference on Speech Prosody* (p. 159-162). Aix-en-Provence, France.
19. **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 XXVIèmes Journées d'Etudes sur la Parole* (p. 571-574). Dinard, France.
20. **Bally, C.** (1926). Le rythme linguistique et sa signification sociale. In A. Pfrimmer (éd.), *Compte-Rendu du Premier Congrès du Rythme* (p. 253-263). Genève, Switzerland: Institut Jacques-Dalcroze.
21. **Bard, E. G., Astésano, C., Turk, A., D'Imperio, M., Nguyen, N., Prévot, L., & Bigi, B.** (2013). Aix Map-Task: A (rather) new French resource for prosodic and discourse studies. In *Proceedings of Tools and Resources for the Analyses of Speech Prosody* (p. 15-19). Aix-en-Provence, France.
22. **Beckman, M. E.** (1986). *Stress and Non-stress Accent*. Berlin, Germany: De Gruyter Mouton.
23. **Beckman, M. E., & Edwards, J.** (1990). Lengthening and shortenings and the nature of prosodic constituency. In J. Kingston & M. E. Beckman (éd.), *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech* (p. 152-178). Cambridge, UK: Cambridge University Press.
24. **Beckman, M. E., & Hirschberg, J.** (1994). The ToBI Annotation Conventions, Appendix A. Consulté le 4 mars 2017, à l'adresse : <http://www.speech.cs.cmu.edu/tobi/ToBI.6.html>
25. **Beckman, M. E., & Jun, S.-A.** (1996). K-ToBI Labeling Conventions. Consulté le 25 juillet 2017, à l'adresse : <http://linguistics.ucla.edu/people/jun/ktobi/k-tobi.html>
26. **Beckman, M. E., & Pierrehumbert, J. B.** (1986). Intonational structure in Japanese and English. *Phonology*, 3(1), 255-309.
27. **Bertrand, R., Blache, P., Espesser, R., Ferré, G., & Meunier, C.** (2008). Le CID - Corpus of Interactional Data - Annotation et Exploitation Multimodale de Parole Conversationnelle. *Traitement automatique des langues (ATALA)*, 49(3), 105-134.
28. **Blanc, O., Constant, M., Dister, A., & Watrin, P.** (2010). Partial Parsing of

- Spontaneous Spoken French. In *Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC'10)* (6p.). Malta.
29. **Boersma, P., & Weenink, D.** (2017). Praat: doing phonetics by computer (Version 6.0.32). Consulté à l'adresse <http://www.praat.org/>
  30. **Bolinger, D.** (1951). Intonation: Levels Versus Configurations. *Word*, 7(3), 199-210.
  31. **Bolinger, D.** (1958). A Theory of Pitch Accent in English. *Word*, 14(2-3), 109-149.
  32. **Bolinger, D.** (1972). Accent Is Predictable (If You're a Mind-Reader). *Language*, 48, 633-644.
  33. **Bolinger, D.** (1989). *Intonation and its Uses*. Stanford, Cal.: Standford University Press.
  34. **Bruce, G.** (1977). *Swedish word accents in sentence perspective*. Lund, Sweden: LiberLäromedel/Gleerup.
  35. **Buhmann, J., Caspers, J., van Heuven, V. J., Hoekstra, H., Martens, J.-P., & Swerts, M.** (2002). Annotation of prominent words, prosodic boundaries and segmental lengthening by non-expert transcribers in the spoken Dutch corpus. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC'02)* (p. 779-785). Las Palmas, Canary Islands, Spain.
  36. **Carlson, R., Hirschberg, J., & Swerts, M.** (2005). Cues to upcoming Swedish prosodic boundaries: subjective judgment studies and acoustic correlates. *Speech Communication*, 46(3-4), 326-333.
  37. **Chomsky, N.** (1965). *Aspects of the theory of syntax*. Cambridge, Mass.: The MIT Press.
  38. **Chomsky, N.** (1970). Remarks on Nominalization. In R. Jacobs & P. Rosenbaum (éd.), *Readings in English Transformational Grammar* (p. 184-221). Boston, Mass.: Ginn.
  39. **Chomsky, N., & Halle, M.** (1968). *The sound pattern of English*. Cambridge, Mass.: The MIT Press.
  40. **Christensen, R. H. B.** (2015). Regression Models for Ordinal Data (Version 2015.6-28) [R package]. Consulté à l'adresse : <http://www.cran.r-project.org/package=ordinal/>.
  41. **Christophe, A., Guasti, M. T., Nespore, M., Dupoux, E., & van Ooyen, B.** (1997). Reflections on phonological bootstrapping: its role for lexical and syntactic acquisition. *Language and Cognitive Processes*, 12(5-6), 585-612.
  42. **Clements, G.** (1979). The Description of Terraced-Level Tone Languages. *Language*, 55(3), 536-558.
  43. **Cole, J., Mahrt, T., & Hualde, J. I.** (2014). Listening for sound, listening for meaning: Task effects on prosodic transcription. In *Proceedings of the 7th International Conference on Speech Prosody (SP 2014)* (p. 858-863). Dublin, Ireland.
  44. **Cole, J., Mo, Y., & Hasegawa-Johnson, M.** (2010a). Signal-based and expectation-based factors in the perception of prosodic prominence. *Laboratory Phonology*, 1(2),

425-452.

45. **Cole, J., Mo, Y., & Baek, S.** (2010b). The role of syntactic structure in guiding prosody perception with ordinary listeners and everyday speech. *Language and Cognitive Processes*, 25(7-9), 1141-1177.
46. **Cole, J., & Shattuck-Hufnagel, S.** (2011). The phonology and phonetics of perceived prosody: What do listeners imitate? In *Proceedings of the 12th Annual Conference of the International Speech Communication Association 2011 (INTERSPEECH 2011)* (p. 969-972). Florence, Italy.
47. **Cooper, W., & Paccia-Cooper, J.** (1980). *Syntax and Speech*. Cambridge, Mass.: Harvard University Press.
48. **Couper-Kuhlen, E.** (1993). *English Speech Rhythm: Form and function in everyday verbal interaction* (Vol. 25). Amsterdam, Netherlands: John Benjamins Publishing Company.
49. **Cutler, A., & Ladd, D. R.** (1983). *Prosody: Models and Measurements* (Vol. 14). Heidelberg, Germany: Springer Berlin Heidelberg.
50. **Dauer, R. M.** (1983). Stress timing and syllable timing reanalysed. *Journal of Phonetics*, 11, 51-62.
51. **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 Toulouse le Mirail - Toulouse II, Toulouse, France.
52. **Delais-Roussarie, E.** (1999). Accentuation et réalisation des clitiques en français. *Cahiers de Grammaire*, 24, 17-37.
53. **Delais-Roussarie, E.** (2000). Vers une nouvelle approche de la structure prosodique. *Langue française*, 126(1), 92-112.
54. **Delais-Roussarie, E., & Di Cristo, A.** (à paraître). L'Accentuation. In A. Abeillé & D. Godard (éd.), *Grande Grammaire du Français*. Paris, France: CNRS éditions.
55. **Delais-Roussarie, E., & Feldhausen, I.** (2014). Variation in Prosodic Boundary Strength: a study on dislocated XPs in French. In *Proceedings of the 7th International Conference on Speech Prosody (SP 2014)* (p. 1052-1056). Dublin, Ireland.
56. **Delais-Roussarie, E., & Fougeron, C.** (2004). Les notions de domaine et de constituant en phonétique et en phonologie. In *Actes des Journées d'études en linguistique* (p. 189-196). Nantes, France.
57. **Delais-Roussarie, E., Lolive, D., Yoo, H., & Guennec, D.** (2016). Patrons Rythmiques et Genres Littéraires en Synthèse de Parole. In *Actes de la conférence conjointe JEP-TALN-RECITAL 2016 : JEP* (Vol. 1, p. 446-454). Paris, France.
58. **Delais-Roussarie, E., Post, B., Avanzi, M., Buthke, C., Di Cristo, A., Feldhausen, I., ... Sichel-Bazin, R.** (2015). Intonational phonology of French: Developing a ToBI system for French. In S. Frota & P. Prieto (éd.), *Intonation in Romance* (p. 63-100). Oxford, UK: Oxford University Press.
59. **Delattre, P.** (1966a). *Studies in French and comparative phonetics: selected papers in French and English*. Berlin, Germany: Walter de Gruyter.

60. **Delattre, P.** (1966b). Les Dix Intonations de base du français. *The French Review*, 40(1), 1-14.
61. **Dell, F.** (1984). L'accentuation dans les phrases en français. In F. Dell, D. Hirst, & J.-R. Vergnaud (éd.), *Forme sonore du langage* (p. 65-122). Paris, France: Herman.
62. **de Pijper, J., & Sanderman, A.** (1994). On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues. *The Journal of the Acoustical Society of America*, 96(4), 2037-2048.
63. **Di Cristo, A.** (1976). Indices prosodiques et structure constituante. *Cahiers de Linguistique, d'Orientalisme et de Slavistique*, 7, 27-40.
64. **Di Cristo, A.** (1978). *De la microprosodie à l'intonosyntaxe* (Thèse de doctorat d'état). Université de Provence, Aix-en-Provence, France.
65. **Di Cristo, A.** (1999a). Vers une modélisation de l'accentuation du français: première partie. *Journal of French Language Studies*, 9(2), 27-44.
66. **Di Cristo, A.** (1999b). Le cadre accentuel du français contemporain: essai de modélisation. Première partie. *Langues*, 2(3), 258-267.
67. **Di Cristo, A.** (2000). La problématique de la prosodie dans l'étude de la parole dite spontanée. *Revue parole*, 15, 189-249.
68. **Di Cristo, A.** (2003). De la métrique et du rythme de la parole ordinaire: l'exemple du français. *Semen [En ligne]*, 16, consulté le 25 septembre 2017.
69. **Di Cristo, A.** (2004). La prosodie au carrefour de la phonétique, de la phonologie et de l'articulation formes-fonctions. *Travaux Interdisciplinaires du Laboratoire Parole et Langage d'Aix-en-Provence (TIPA)*, 23, 67-211.
70. **Di Cristo, A.** (2011). Une approche intégrative des relations de l'accentuation au phrasé prosodique du français. *Journal of French Language Studies*, 21(2), 143-179.
71. **Di Cristo, A.** (2013). *La prosodie de la parole*. Brussels, Belgium: De Boeck.
72. **Di Cristo, A.** (2016). *Les musiques du français parlé, Essais sur l'accentuation, la métrique, le rythme, le phrasé prosodique et l'intonation du français contemporain*. Berlin, Germany: De Gruyter Mouton.
73. **Di Cristo, A., & Hirst, D.** (1986). Modelling French micromelody: analysis and synthesis. *Phonetica*, 43(1), 11-30.
74. **Di Cristo, A., & Hirst, D.** (1993). Rythme syllabique, rythme mélodique et représentation hiérarchique de la prosodie du français. *Travaux Interdisciplinaires du Laboratoire Parole et Langage d'Aix-en-Provence (TIPA)*, (15), 9-24.
75. **Di Cristo, A., & Hirst, D.** (1996). Vers une typologie des unités intonatives du français. In *Actes des XXI<sup>e</sup> Journées d'Etude sur la Parole (JEP 1996)* (p. 219-222). Avignon, France.
76. **Di Cristo, A., & Jankowski, L.** (1999). Prosodic organisation and phrasing after focus in French. In *Proceedings of the 14th Congress of Phonetic Sciences* (p. 1565-1568). San Francisco, USA.

77. **Dubois, J.** (1969). Grammaire générative et transformationnelle. *Langue française*, 1(1), 49-57.
78. **Duez, D.** (1987). Hiérarchisation des paramètres acoustiques et identification des frontières. In *Actes des XVIe Journées d'Etude sur la Parole (JEP 1987)* (p. 221-223). Hammamet, Tunisie.
79. **Dupoux, E., Pallier, C., Sebastian, N., & Melher, J.** (1997). A Destressing « Deafness » in French? *Journal of Memory and Language*, 3(36), 406-421.
80. **Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N.** (2010). Limits on bilingualism revisited: Stress 'deafness' in simultaneous French–Spanish bilinguals. *Cognition*, 114(2), 266-275.
81. **Dupoux, E., Sebastián-Gallés, N., Navarrete, E., & Peperkamp, S.** (2008). Persistent stress 'deafness': The case of French learners of Spanish. *Cognition*, 106(2), 682-706.
82. **Fant, G., & Kruckenberg, A.** (1995). Notes on syllable duration in French and Swedish. In K. Elenius & P. Branderud (éd.), *Proceedings of the 13th International Congress of Phonetic Sciences* (p. 622-625). Stockholm, Sweden.
83. **Fant, G., Kruckenberg, A., & Nord, L.** (1991). Durational correlates of stress in Swedish, French and English. *Journal of Phonetics*, 19, 118-121.
84. **Féry, C., & Schubö, F.** (2010). Hierarchical prosodic structures in the intonation of center-embedded relative clauses. *The Linguistic Review*, 27(3), 289-313.
85. **Féry, C., & Truckenbrodt, H.** (2005). Sisterhood and tonal scaling. *Studia Linguistica*, 59(2-3), 223-243.
86. **Fletcher, J.** (1991). Rhythm and final lengthening in French. *Journal of Phonetics*, 19(2), 193-212.
87. **Fodor, J.** (1998). Learning To Parse? *Journal of Psycholinguistic Research*, 27(2), 285-319.
88. **Fónagy, I.** (1980). L'accent français : accent probabilitaire. In I. Fónagy, P. R. Léon, & F. Carton (éd.), *L'accent en français contemporain* (Vol. 15, p. 123-233). Paris, France: Didier.
89. **Fougeron, C., & Jun, S.-A.** (1998). Rate Effects on French Intonation: Phonetic Realization and Prosodic Organization. *Journal of Phonetics*, 26, 45-70.
90. **Fougeron, C., & Keating, P.** (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*, 101(6), 3726-3740.
91. **Fraisse, P.** (1956). *Les structures rythmiques, étude psychologique* (Publication de l'Université de Louvain). Namur, Belgium: Editions Erasme.
92. **Fraisse, P.** (1967). Psychologie des rythmes humains. In *Actes du Colloque sur les rythmes* (p. 23-33). Lyon, France: SIMEP-Editions.
93. **Fraisse, P.** (1974). *Psychologie du rythme*. Paris, France: Presses universitaires de France.

94. **Garde, P.** (1968). *L'accent*. Paris, France: Presses universitaires de France.
95. **Gee, J., & Grosjean, F.** (1983). Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology*, 15, 411-458.
96. **Gendrot, C., Gerdes, K., & Adda-Decker, M.** (2016). Détection automatique d'une hiérarchie prosodique dans un corpus de parole journalistique. *Langue française*, 191(3), 123-149.
97. **Ghio, A., André, C., Teston, B., & Cavé, C.** (2003). PERCEVAL: une station automatisée de tests de PERCEption et d'EVALuation auditive et visuelle. *Travaux Interdisciplinaires du Laboratoire Parole et Langage d'Aix-en-Provence (TIPA)*, 22, 115-133.
98. **Goldman-Eisler, F.** (1968). *Psycholinguistics: experiments in spontaneous speech*. London, UK: Academic Press.
99. **Goldsmith, J.** (1976). *Autosegmental Phonology* (PHD Thesis). Indiana University, Bloomington, Ind.
100. **Grice, M.** (1995). Leading tones and downstep in English. *Phonology*, 12(02), 183-233.
101. **Grice, M., & Baumann, S.** (2002). Deutsche Intonation und GToBI. *Linguistische Berichte*, 191, 267-298.
102. **Grosjean, F., & Deschamps, A.** (1972). Analyse des variables temporelles du français spontané. *Phonetica*, 26(3), 129-157.
103. **Grosjean, F., & Deschamps, A.** (1975). Analyse contrastive des variables temporelles de l'anglais et du français: vitesse de parole et variables composantes, phénomènes d'hésitation. *Phonetica*, 31, 144-184.
104. **Grosjean, F., & Dommergues, J.-Y.** (1983). Les structures de performance en psycholinguistique. *L'Année Psychologique*, 83(2), 513-536.
105. **Grosjean, F., Grosjean, L., & Lane, H.** (1979). Patterns of silence: performance structures in sentence production. *Cognitive Psychology*, 11, 58-81.
106. **Guaïtella, I.** (1986). *Considérations sur le rythme d'une structure prosodique à un seul niveau* (DEA). Aix-en-Provence, France.
107. **Gussenhoven, C.** (2005). Procliticized Phonological Phrases in English: Evidence from Rhythm. *Studia Linguistica*, 59(2-3), 174-193.
108. **Gussenhoven, C., Rietveld, T., Kerkhoff, J., & Terken, J.** (2003). Transcription of Dutch Intonation. Consulté le 25 juillet 2017, à l'adresse : [http://www.fon.hum.uva.nl/rob/Courses/InformationInSpeech/CDROM/Literature/LOT\\_winterschool2006/todi.let.kun.nl/ToDI/home.htm](http://www.fon.hum.uva.nl/rob/Courses/InformationInSpeech/CDROM/Literature/LOT_winterschool2006/todi.let.kun.nl/ToDI/home.htm)
109. **Halle, M., & Vergnaud, J.-R.** (1987). *An essay on stress*. Cambridge, Mass. : The MIT Press.
110. **Hannahs, S. J.** (1995). *Prosodic structure and French morphophonology*. Tübingen, Germany: Niemeyer.

111. **Hayes, B.** (1982). Extrametricality and English Stress. *Linguistic Inquiry*, 13(2), 227-276.
112. **Hayes, B.** (1989). The prosodic hierarchy in meter. In P. Kiparsky & G. Youmans (éd.), *Rhythm and Meter* (p. 201-260). Orlando, Florida: Academic Press.
113. **Hayes, B.** (1995). *Metrical Stress Theory: Principles and Case Studies*. Chicago, Illinois: University of Chicago Press.
114. **Hermès, D. J., & Rump, H. H.** (1994). Perception of prominence in speech intonation induced by rising and falling pitch movements. *The Journal of the Acoustical Society of America*, 96(1), 83-92.
115. **Hirst, D., & Di Cristo, A.** (1984). French Intonation: a parametric approach. *Die Neueren Sprachen*, 83(5), 554-569.
116. **Hirst, D., & Di Cristo, A.** (1996). Y-a-t-il des unités tonales en français? In *Actes des XXLe Journées d'Etude sur la Parole* (p. 223-226). Avignon, France.
117. **Hirst, D., & Di Cristo, A.** (1998). A survey of intonation systems. In D. Hirst & A. Di Cristo (éd.), *Intonation Systems: A Survey of Twenty Languages* (p. 1-44). Cambridge, UK: Cambridge University Press.
118. **Hirst, D., & Espesser, R.** (1993). Automatic Modelling of Fundamental Frequency using a quadratic spline function. *Travaux Interdisciplinaires du Laboratoire Parole et Langage d'Aix-en-Provence (TIPA)*, 15, 75-85.
119. **Huang, J.** (1985). The autosegmental and metrical nature of tone terracing. In D. L. Goyvaerts (éd.), *African Linguistics: Essays in Memory of M.W.K. Semikenke* (p. 209-238). Amsterdam, Netherlands: John Benjamins Publishing.
120. **Inkelas, S.** (1989). *Prosodic Constituency in the Lexicon* (PHD Thesis). Stanford University, Stanford, Cal.
121. **Inkelas, S., & Zec, D.** (1993). Auxiliary reduction without empty categories: A prosodic account. *Working Papers of the Cornell Phonetics Laboratory*, p. 205-253.
122. **Ito, J., & Mester, A.** (2007). Prosodic adjunction in Japanese compounds. In *Proceeding of Formal approaches to Japanese linguistics (FAJL 6)* (Vol. 4, p. 97-111). Cambridge, Mass.
123. **Ito, J., & Mester, A.** (2009). The Onset of the Prosodic Word. In S. Parker (éd.), *Phonological Argumentation: Essays on Evidence and Motivation* (p. 227-260). London, UK: Equinox.
124. **Ito, J., & Mester, R.-A.** (1992). Weak Layering and Word Binariness. In T. Honma, M. Okazaki, T. Tabata, & S.-I. Tanaka (éd.), *A New Century of Phonology and Phonological Theory. A Festschrift for Professor Shosuke Haraguchi on the Occasion of His Sixtieth Birthday* (p. 26-65). Tokyo, Japan: Kaitakusha.
125. **Jackendoff, R.** (1977). *X syntax: a study of phrase structure*. Cambridge, Mass.: The MIT Press.
126. **Jankowski, L.** (1997). *Approche linguistique, acoustique et perceptive des proéminences du français*. (Mémoire de DEA). Institut de Phonétique, Université de Provence, Aix-en-Provence, France.

127. **Jun, S.-A.** (1993). *The Phonetics and Phonology of Korean Prosody*. (PHD Thesis). The Ohio State University, Columbus, Ohio.
128. **Jun, S.-A., & Fougeron, C.** (1995). The Accentual Phrase and the Prosodic structure of French. In *Proceedings of the 13th International Congress of Phonetic Sciences was held in Stockholm (ICPHS 95)* (Vol. 2, p. 722-725). Stockholm, Sweden.
129. **Jun, S.-A., & Fougeron, C.** (2000). A Phonological Model of French Intonation. In A. Botinis (éd.), *Intonation: Analysis, Modeling and Technology* (Kluwer Academic Publishers, Vol. 15, p. 209-242). Dordrecht, Netherlands: Springer Netherlands.
130. **Jun, S.-A., & Fougeron, C.** (2002). Realizations of accentual phrase in French intonation. *Probus*, 14(1), 147-172.
131. **Kager, R.** (1993). Alternatives to the iambic-trochaic law. *Natural Language and Linguistic Theory*, 11, 381-432.
132. **Kager, R.** (1995). The Metrical Theory of Word Stress. In J. A. Glodsmith (éd.), *The Handbook of Phonological Theory* (p. 367-402). Oxford, UK: Blackwell Publishing Ltd.
133. **Kager, R.** (1999). *Optimality theory*. Cambridge, UK: Cambridge University Press.
134. **Krivokapic, J., & Byrd, D.** (2012). Prosodic boundary strength: An articulatory and perceptual study. *Journal of Phonetics*, 40(3), 430-442.
135. **Lacheret-Dujour, A.** (2003). *La prosodie des circonstants en français parlé*. Peeters Publishers.
136. **Lacheret-Dujour, A., & Beaugendre, F.** (1999). *La Prosodie du français*. Paris, France: CNRS éditions.
137. **Lacheret-Dujour, A., & Morel, M.** (2002). Génération automatique de la prosodie dans le système de synthèse vocale Kali: de la modélisation phonologique à l'implémentation des paramètres acoustiques. In *Actes des XXIVe Journées d'Etudes sur la Parole* (p. 53-57). Nancy, France.
138. **Lacheret-Dujour, A., Simon, A. C., Goldman, J.-P., & Avanzi, M.** (2014). Prominence perception and accent detection in French: from phonetic processing to grammatical analysis. *Language Sciences*, 39(Special Issues), 95-106.
139. **Ladd, D. R.** (1983). Phonological features of intonational peaks. *Language*, 59, 721-759.
140. **Ladd, D. R.** (1986). Intonational phrasing: the case for recursive prosodic structure. *Phonology Yearbook*, 3, 311-340.
141. **Ladd, D. R.** (1988). Declination « reset » and the hierarchical organization of utterances. *The Journal of the Acoustical Society of America*, 84, 530-544.
142. **Ladd, D. R.** (2008). *Intonational phonology (2nd Edition)*. Cambridge, UK: Cambridge University Press. (1st Edition: 1996).
143. **Ladd, D. R., & Campbell, N.** (1991). Theories of Prosodic Structure: Evidence from Syllable Duration. In *Proceedings of the 12th International Congress of Phonetic Sciences* (Vol. 2, p. 290-293). Aix-en-Provence, France.

144. **Lehiste, I.** (1973). Phonetic Disambiguation of Syntactic Ambiguity. *The Journal of the Acoustical Society of America*, 53(1), 380-380.
145. **Lenth, R. V.** (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, 69(1), 1-33.
146. **Léon, P.** (1992). *Phonétisme et prononciations du français*. Paris, France: Nathan Université.
147. **Liao, T. F.** (1994). *Interpreting Probability Models: Logit, Probit, and Other Generalized Linear Models*. SAGE.
148. **Liberman, M., & Pierrehumbert, J.** (1984). Intonational Invariance under Changes in Pitch Range and Length. In M. Aronoff & R. Oehrle (éd.), *Language and Sound Structure* (p. 157-233). Cambridge, Mass.: The MIT Press.
149. **Liberman, M., & Prince, A.** (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, 8(2), 249-336.
150. **Liberman, M.** (1975). *The intonational system of English*. (PHD Thesis). Massachusetts Institute of Technology, Boston, Mass.
151. **Lucci, V.** (1983). Prosodie, phonologie et variation en français contemporain. *Langue française*, 60(1), 73-84.
152. **Martin, L., Degand, L., & Simon, A.-C.** (2014). Forme et fonction de la périphérie gauche dans un corpus oral multigenres annoté. *Corpus*, 13, 243-265.
153. **Martin, P.** (1975). Analyse phonologique de la phrase française. *Linguistics*, 146, 35-68.
154. **Martin, P.** (1981). Pour une théorie de l'intonation. In M. Rossi, A. Di Cristo, D. Hirst, P. Martin, & Y. Nishinuma (éd.), *L'intonation : de l'acoustique à la sémantique* (p. 234-271). Paris, France: Klincksieck.
155. **Martin, P.** (1987). Prosodic and Rhythmic Structures in French. *Linguistics*, 5(25), 925-949.
156. **Martin, P.** (2006). Intonation du français: parole spontanée et parole lue. *Estudios de Fonética experimental*, 15, 134-162.
157. **Martin, P.** (2009). *Intonation du français*. Paris, France: Armand Colin.
158. **Maxwell, A.** (1970). Comparing the classification of subjects by two independent judges. *British Journal of Psychiatry*, 116, 651-655.
159. **McCarthy, J. J., & Prince, A.** (1993). Generalized alignment. In G. Booij & J. van Marle (éd.), *Yearbook of Morphology 1993* (p. 79-153). Dordrecht, Netherlands: Springer Netherlands.
160. **Méndez, R.** (en cours). *Perception de la structure métrique du Français: étude comparative entre des francophones et des hispanophones*. (Thèse de doctorat). Université Toulouse 2 Jean-Jaurès, Toulouse, France.
161. **Méndez, R., & Astésano, C.** (2017). Perception of the Downstepped Final Accent in French. In *Proceedings of Phonetics and Phonology in Europe (PaPE) 2017*.

Cologne, Germany.

162. **Mertens, P.** (1987). *L'intonation du français: de la description linguistique à la reconnaissance automatique* (PHD Thesis). Katholieke Universiteit Leuven, Louvain, Belgium.
163. **Mertens, P.** (1992). L'accentuation de syllabes contiguës. *ITL: Review of Applied Linguistics*, 95, 145-165.
164. **Mertens, P.** (1993a). Accentuation, intonation et morphosyntaxe. *Travaux de linguistique*, (26), 21-69.
165. **Mertens, P.** (1993b). International Grouping, Boundaries, and Syntactic Structure in French. In *Proceedings of an ESCA Workshop on Prosody* (p. 228-231). Lund, Sweden.
166. **Mertens, P.** (1997). De la chaîne linéaire à la séquence de tons. *Traitement automatique des langues*, 38(1), 27-51.
167. **Mertens, P.** (2004). Le Prosogramme: une transcription semi-automatique de la prosodie. *Cahiers de l'Institut de Linguistique de Louvain*, 30(1-3), 7-25.
168. **Mertens, P.** (2008). Syntaxe, prosodie et structure informationnelle: une approche prédictive pour l'analyse de l'intonation dans le discours. *Travaux de linguistique*, 56(1), 97-124.
169. **Michelas, A.** (2011). *Caractérisation phonétique et phonologique du syntagme intermédiaire en français: de la production à la perception* (Thèse de doctorat). Université Aix Marseille 1, Aix-en-Provence, France.
170. **Michelas, A., & D'Imperio, M.** (2010). Durational cues and prosodic phrasing in French: evidence for the intermediate phrase. In *Proceedings of the 5th International Conference on Speech Prosody (SP 2010)* (p. 11-14). Chicago, Illinois.
171. **Michelas, A., Frauenfelder, U. H., Schön, D., & Dufour, S.** (2016). How deaf are French speakers to stress? *The Journal of the Acoustical Society of America*, 139(3), 1333-1342.
172. **Millotte, S., Wales, R., & Christophe, A.** (2007). Phrasal prosody disambiguates syntax. *Language and Cognitive Processes*, 22(6), 898-909.
173. **Mo, Y., Cole, J., & Lee, E.-K.** (2008). Naïve listeners' prominence and boundary perception. In *Proceedings of the 4th International Conference on Speech Prosody (SP 2008)* (p. 735-738). Campinas, Brazil.
174. **Nespor, M., & Vogel, I.** (1986). *Prosodic phonology*. Berlin, Germany: De Gruyter Mouton.
175. **OpenCourseWare.** Pitch Accents with more than one tone: Bi-tonal accents L+H\* and L\*+H. (2006, Août). Consulté le 7 octobre 2017, à l'adresse : [https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-911-transcribing-prosodic-structure-of-spoken-utterances-with-tobi-january-iap-2006/lecture-notes/chap2\\_5.pdf](https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-911-transcribing-prosodic-structure-of-spoken-utterances-with-tobi-january-iap-2006/lecture-notes/chap2_5.pdf)
176. **Pasdeloup, V.** (1990a). *Modèle de règles rythmiques du français appliqué à la synthèse de la parole*. (Thèse de doctorat). Université de Provence, Aix-en-Provence,

France.

177. **Pasdeloup, V.** (1990b). Les contraintes linguistiques et phonotactiques dans la structuration accentuelle: analyse acoustique et perceptive d'un corpus lu. *Journal de Physique Colloques*, 51(C2), C2-511-C2-514.
178. **Pasdeloup, V.** (1991). Exploiting the secondary accent in a prosodic model for French synthesis. In *Proceedings of the 12th International Congress of Phonetic Sciences* (Vol. 2, p. 254-257). Aix-en-Provence, France.
179. **Pasdeloup, V.** (1992). A prosodic Model for French Text-to-speech synthesis: A psycholinguistic approach. In G. Bailly, C. Benoit, & T.R. Sawallis (éd.), *Talking Machines: Theories, Models, and Designs* (p. 335-348). Amsterdam, Netherlands: Elsevier Science Publishers.
180. **Pasdeloup, V.** (2004). Le rythme n'est pas élastique: étude préliminaire de l'influence du débit de parole sur la structuration temporelle. In *Actes des XXVe Journées d'Etudes sur la Parole (JEP 2004)*. Fès, Maroc.
181. **Pasdeloup, V., Espesser, R., & Faraj, M.** (2006). Sensibilité au débit et marquage accentuel des phonèmes en français. In *Actes des XXVIe Journées d'Etude sur la Parole (JEP 2006)* (p. 251-254). Dinard, France.
182. **Peperkamp, S.** (1999). Prosodic Words. *GLOT International*, 4, 15-16.
183. **Peperkamp, S., & Dupoux, E.** (2002). A typological study of stress deafness. In C. Gussenhoven & N. Warner (éd.), *Laboratory Phonology 7* (p. 203-240). Berlin, Germany: De Gruyter Mouton.
184. **Petrone, C., Truckenbrodt, H., Wellmann, C., Holzgrefe-Lang, J., Wartenburger, I., & Höhle, B.** (2017). Prosodic boundary cues in German: Evidence from the production and perception of bracketed lists. *Journal of Phonetics*, 61, 71-92.
185. **Piepho, H.P.** (2004). An algorithm for a letter-based representation of all pairwise comparisons. *Journal of Computational and Graphical Statistics*, 13(2), 456-466.
186. **Pierrehumbert, J.** (1980). *The phonology and phonetics of English intonation* (PHD Thesis). Massachusetts Institute of Technology, Boston, Mass.
187. **Pierrehumbert, J. B.** (2000). Tonal elements and their alignment. In M. Horne (éd.), *Prosody: Theory and Experiment. Studies Presented to Gosta Bruce* (p. 11-26). Dordrecht, Netherlands: Kluwer.
188. **Pierrehumbert, J. B., & Steele, S. A.** (1989). Categories of tonal alignment in English. *Phonetica*, 46(4), 181-196.
189. **Pike, K. L.** (1945). *The Intonation of American English*. East Lansing, Mich.: University of Michigan Press.
190. **Plénat, M.** (1994). L'extramétricalité des voyelles initiales. In C. Lyche (éd.), *French Generative Phonology: Retrospectives and Perspectives* (p. 237-258). Dublin, Ireland: ESRI Publication.
191. **Portes, C.** (2002). Approche instrumentale et cognitive de la prosodie du discours en français. *Travaux Interdisciplinaires du Laboratoire Parole et Langage d'Aix-en-Provence (TIPA)*, 21, 101-119.

192. **Post, B.** (2000). *Tonal and phrasal structures in French intonation* (PHD Thesis). The Hague University, The Hague, Netherlands.
193. **Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C.** (1991). The use of prosody in syntactic disambiguation. *The Journal of the Acoustical Society of America*, 90(6), 2956-2970.
194. **Prieto, P., D'Imperio, M., & Gili Fivela, B.** (2005). Pitch accent alignment in Romance: primary and secondary associations with metrical structure. *Language and Speech*, 48(4), 359-396.
195. **Prince, A.** (1990). Quantitative consequences of rhythmic organization. In *Proceedings of the 26th regional meeting of the Chicago Linguistic Society* (Vol. 2, p. 355-398). Chicago, Illinois.
196. **Prince, A., & Smolensky, P.** (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*. Cambridge, Mass. : The MIT Press.
197. **R Core Team.** (2016). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Consulté à l'adresse <http://www.R-project.org/>.
198. **Roach, P.** (1982). On the distinction between « stress-timed » and « syllable-timed » languages. In D. Crystal (éd.), *Linguistic controversies, Essays in linguistic theory and practice* (p. 73-79). London, UK: Edward Arnold.
199. **Rossi, M.** (1972). Le seuil différentiel de durée. In A. Valdman (éd.), *Papers in linguistics and phonetics to the memory of Pierre Delattre* (Mouton, p. 435-450). The Hague, Netherlands.
200. **Rossi, M.** (1980). Le français, langue sans accent? In I. Fónagy & P. Léon (éd.), *L'accent en français contemporain* (p. 13-51). Paris, France: Didier.
201. **Rossi, M.** (1981). Intonation, énonciation, syntaxe. In M. Rossi, D. Hirst, A. Di Cristo, P. Martin, & Y. Nishinuma (éd.), *L'intonation: de l'acoustique à la sémantique* (p. 184-234). Paris, France: Klincksieck.
202. **Rossi, M.** (1985). L'intonation et l'organisation de l'énoncé. *Phonetica*, 42(2-3), 135-153.
203. **Rossi, M.** (1999). *L'intonation: le système du français: description et modélisation*. Paris, France: Editions OPHRYS.
204. **Rossi, M., & Chafcouloff, M.** (1972). Les niveaux intonatifs. *Travaux Interdisciplinaires du Laboratoire Parole et Langage d'Aix-en-Provence (TIPA)*, 1, 167-176.
205. **Roux, G., Bertrand, R., Ghio, A., & Astésano, C.** (2016). Naïve listeners' perception of prominence and boundary in French spontaneous speech. In *Proceedings of the 8th International Conference on Speech Prosody (SP 2016)* (p. 912-916). Boston, Mass.
206. **Schafer, A., Carter, J., Clifton Jr, C., & Frazier, L.** (1996). Focus in Relative Clause Construal. *Language and Cognitive Processes*, 11(1-2), 135-164.
207. **Selkirk, E.** (1980). *On prosodic structure and its relation to syntactic structure*.

Bloomington, Ind. : Indiana University Linguistics Club.

208. **Selkirk, E.** (1984). *Phonology and Syntax: The Relationship between Sound and Structure*. Cambridge, Mass. : The MIT Press.
209. **Selkirk, E.** (1986). On derived domains in sentence phonology. *Phonology*, 3, 371-405.
210. **Selkirk, E.** (1995). Sentence prosody: intonation, stress and phrasing. In J. Goldsmith (éd.), *The Handbook of Phonological Theory* (p. 550-569). London, UK: Blackwell.
211. **Selkirk, E.** (1996). The Prosodic Structure of Function Words. In J. L. Morgan & K. Demuth (éd.), *Signal to Syntax: Bootstrapping From Speech to Grammar in Early Acquisition* (p. 187-211). Hove, UK: Psychology Press.
212. **Selkirk, E.** (2009). On clause and intonational phrase in Japanese: The syntactic grounding of prosodic constituent structure. *Gengo Kenkyu*, 136, 35-73.
213. **Selkirk, E.** (2011). The syntax-phonology interface. *The handbook of phonological theory*, 2, 435-483.
214. **Selkirk, E., & Tateishi, K.** (1988). Constraints on Minor Phrase formation in Japanese. In *Proceedings of the 24th Annual Meeting of the Chicago Linguistic Society* (p. 316-336). Chicago, Illinois.
215. **Shattuck-Hufnagel, S.** (1992). Early pitch accent placement within words in American English. *The Journal of the Acoustical Society of America*, 92(4), 2443-2443.
216. **Shattuck-Hufnagel, S., & Turk, A. E.** (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, 25(2), 193-247.
217. **Silverman, K., Beckman, M. E., Pitrelli, J., Ostendorf, M., Wightman, C. W., Price, P. J., ... Hirschberg, J.** (1992). ToBI: a standard for labeling English prosody. In *Proceedings of the 1992 International Conference on Spoken Language Processing* (Vol. 2, p. 867-870). Banff, AB.
218. **Simon, A. C.** (2004). *La structuration prosodique du discours en français: une approche multidimensionnelle et expérientielle*. Berne, Switzerland : P. Lang.
219. **Simon, A. C., & Christodoulides, G.** (2016a). Perception of prosodic boundaries by naïve listeners in French. In *Proceedings of the 8th International Conference on Speech Prosody (SP 2016)* (p. 1158-1162). Boston, Mass.
220. **Simon, A.-C., & Christodoulides, G.** (2016b). Frontières prosodiques perçues: corrélats acoustiques et indices syntaxiques. *Langue française*, 191(3), 83-106.
221. **Simon, A. C., & Degand, L.** (2009). Mapping Prosody and Syntax as Discourse Strategies: How Basic Discourse Units Vary Across Genres. In A. Wichmann, N. Dehé, & D. Barth-Weingarten (éds.), *Where Prosody Meets Pragmatics* (p. 79-105). Leiden, Netherlands: Brill.
222. **Smith, C.** (2009). Naïve listener's perceptions of French prosody compared to the predictions of theoretical models. In *Proceedings of the Prosody-Discourse Interface Conference (IDP-2009)* (p. 335-349). Paris, France.

223. **Steedman, M.** (2008). Information-Structural Semantics for English Intonation. In C. Lee, M. Gordon, & D. Büring (éd.), *Topic and Focus. Studies in Linguistics and Philosophy* (Vol. 82, p. 245-264). Dordrecht, Netherlands: Springer Netherlands.
224. **Streefkerk, B.** (1997). Acoustical correlates of prominence: A design for research. In *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam* (Vol. 21, p. 131-142). Amsterdam, Netherlands.
225. **Stuart, A.** (1955). A test for homogeneity of the marginal distributions in a two-way classification. *Biometrika*, 42, 412-416.
226. **Swerts, M.** (1997). Prosodic features at discourse boundaries of different strength. *The Journal of the Acoustical Society of America*, 101(1), 514-5221.
227. **Swerts, M.** (1998). Filled pauses as markers of discourse structure. *Journal of Pragmatics*, 30(4), 485-496.
228. **te Rietmolen, N.** (en cours). *Etudes neurophysiologiques du traitement de l'accentuation en français*. (Thèse de doctorat). Université Toulouse 2 Jean-Jaurès, Toulouse, France.
229. **te Rietmolen, N., El Yagoubi, R., Espesser, R., Magnen, C., & Astésano, C.** (2016). Investigating the phonological status of the initial accent in French: An Event-Related Potentials study. In *Proceedings of the 8th International Conference on Speech Prosody (SP 2016)* (p. 1181-1185). Boston, Mass.
230. **te Rietmolen, N., Yagoubi, R. E., Ghio, A., & Astésano, C.** (2017). The Phonological Status of the French Initial Accent and its Role in Semantic Processing: An Event-Related Potentials Study. In *Proceedings of 16th Annual Conference of the International Speech Communication Association 2017 (INTERSPEECH 2017)* (p. 2436-2440). Stockholm, Sweden.
231. **Trager, G. L., & Smith, H. L.** (1951). *An Outline of English Structure* (Vol. 3). Norman, OK: Battenburg Press.
232. **Tremblay, A.** (2009). Phonetic variability and the variable perception of L2 word stress by French Canadian listeners. *International Journal of Bilingualism*, 13(1), 35-62.
233. **Truckenbrodt, H.** (1999). On the Relation between Syntactic Phrases and Phonological Phrases. *Linguistic Inquiry*, 30(2), 219-255.
234. **Truckenbrodt, H.** (2002). Upstep and embedded register levels. *Phonology*, 19(01), 77-120.
235. **Turk, A. E., & Shattuck-Hufnagel, S.** (2007). Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics*, 35(4), 445-472.
236. **Vaissière, J.** (1991). Rhythm, accentuation and final lengthening in French. In J. Sundberg, L. Nord, & R. Carlson (éd.), *Music, Language, Speech and Brain* (p. 108-120). London, UK: Macmillan Press.
237. **Vaissiere, J.** (2006). *La phonétique*. Paris, France: Presses universitaires de France.
238. **van den Berg, R., Gussenhoven, C., & Rietveld, T.** (1992). Downstep in Dutch: Implications for a model. In G. Docherty & D. R. Ladd (éd.), *Papers in Laboratory*

*Phonology II: Gesture, segment, prosody* (p. 335-367). Cambridge, UK: Cambridge University Press.

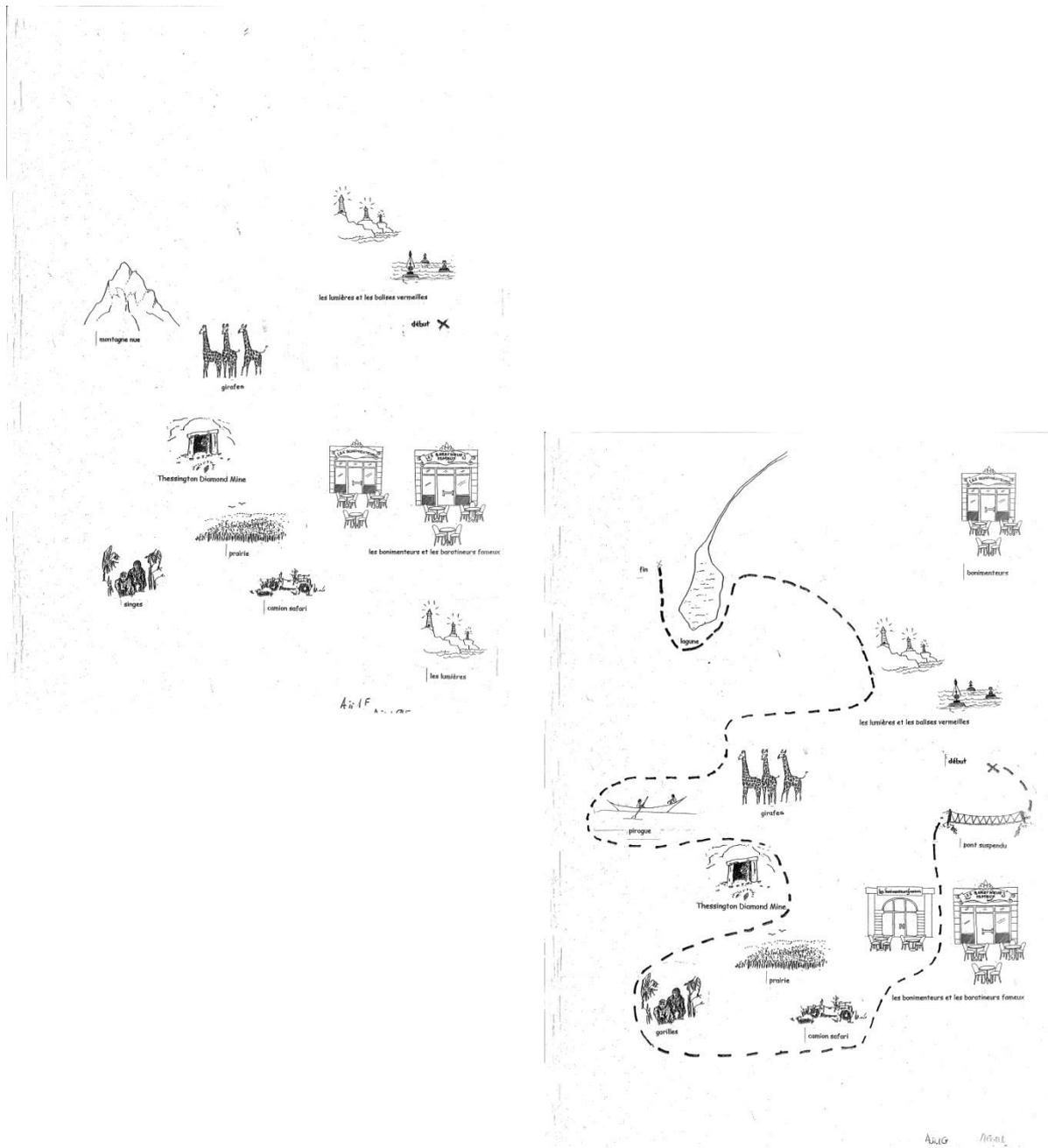
239. **Welby, P.** (2003). *The Slaying of Lady Mondegreen, being a Study of French Tonal Association and Alignment and their Role in Speech Segmentation* (PHD Thesis). The Ohio State University, Columbus, Ohio.
240. **Welby, P., Bertrand, R., Portes, C., & Astésano, C.** (2016). Realization of the French initial accent: Stability and individual differences. In *Proceedings of the 7th Conference on Tone and Intonation in Europe (TIE)*. Canterbury, UK.
241. **Wenk, B. J., & Wioland, F.** (1982). Is French really syllable-timed? *Journal of Phonetics*, 10(2), 193-216.
242. **Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J.** (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *The Journal of the Acoustical Society of America*, 91(3), 1707-1717.
243. **Zec, D.** (2005). Prosodic differences among function words. *Phonology*, 22(1), 77-112.

# **Annexes**



## Annexe 1 : Corpus de la Map Task Aix (Bard et al., 2013)

**Annexe 1 – A :** Exemple des cartes utilisées dans le corpus avec chemin (figure de gauche) ou sans chemin (figure de droite).



**Annexe 1 – B : Récapitulatif des différentes occurrences constituées de N et/ou ADJ, produites dans le corpus.**

<b>occurrences</b>	<b>Nb</b>
Adj_Cc_Adj_N	2
Adj_N	27
Adj_N_Adj	3
Adj_N_Spd_N	1
Adv_Adj_N	14
D_Adj_Adj_N	17
D_Adj_N	290
D_Adj_N_Adj	13
D_Adj_N_Spd_D_N	6
D_Adj_N_Spd_N	11
D_Adv_Adj_N	7
D_N	2945
D_N_Adj	362
D_N_Adj_Adj	18
D_N_Adj_Cc_D_N_Adj	2
D_N_Adj_Spd_D_N	9
D_N_Adv_Adj	17
D_N_Cc_D_N	96
<b>D_N_Cc_D_N_Adj</b>	<b>79</b>
D_N_Spd_Adj_N	3
D_N_Spd_D_Adj_N	7
D_N_Spd_D_N	145
D_N_Spd_D_N_Adj	6
D_N_Spd_N	169
D_N_Spd_N_Adj	14
N	683
N_Adj	64
N_Adj_Adj	4
N_Adv_Adj	4
N_Cc_D_N	6
N_Cc_D_N_Adj	1
N_Cc_N	13
N_N	34
N_Spd_Adj_N	2
N_Spd_D_Adj_N	4
N_Spd_D_N	41
N_Spd_D_N_Adj	5
N_Spd_N	19
N_Spd_N_Adj	2
<b>TOTAL</b>	<b>5145</b>

**Annexe 1 – C : Extrait du tableau Excel répertoriant chacune des structures D1+N1+cc+D2+N2+ADJ, produites dans le corpus. La cellule « couleur Map » indique le type de dialogue produit, et la cellule « Instruction F ou G » permet de signaler le rôle du locuteur (« F » pour Follower ou celui qui suit les instructions, et « G » pour Giver ou celui qui donne les instructions).**

Couleur Map	Locuteurs	Code		Instruction F ou G	contexte	Scope	
bleu	Loc 1	BAiii1	Alpha	F	faut que tu me donnes d' autres renseignements non # [les lumières et les balises vermeilles] # ouais # non # si en haut l' ai ai pas # je suis toujours à gauche # haut gauche [des des lumières et les balises vermeilles] # ouais # montagne nue # ouais # ouais # je contourne quelque chose # o.k. # donc en fait # o.k. [les lumières et les balises vermeilles] je prends tout de suite à gauche # oui ca v	B	
	Loc 2		Beta	G	alors tu vois c' est où c' est marqué en bas [les bonimenteurs et les baratineurs fameux] # non non tu vas pas tout droit # tu prends	N	
	Loc 1	BAv5	Alpha	G	et balises vermeilles # le phare qui est à gauche # [les lumières et les balises vermeilles] # non # tu vas frôler # les lumières et les balises vermeilles # non # tu vas frôler # [les lumières et les balises vermeilles] # tu t' arrêtes au niveau de # du phare qui est	B	
	Loc 2		Beta	F	as pas de dessin moi non plus # il y a [les lumières et les balises # vermeilles] # au niveau # ouais # mh mh # mh d'accord # toi c' est c' est en deux dessins [les lumières et les balises vermeilles] hein # tu as trois phares en haut et après	N	
	Loc 2	BBii6	Beta	G	# tu pars donc vers la gauche # pour passer [sous # les lumières et les balises vermeilles] # tu remontes jusqu' en dessous du mot # donc	N	
	Loc 1	BBvii2	Alpha	G	de la pyramide # est -ce que tu l' as [les lumières et les balises vermeilles] # alors c' est # il y a des arrêté en bas de la pyramide à gauche # tu vas contourner [les lumières et les balises vermeilles] # tu contournes les dessins et la phrase # sur les lumières # o.k. # et # la phrase aussi [les lumières et les balises vermeilles] tu tu passes en dessous # tu contournes # tu belle ingénue # o.k. est -ce que tu as aussi [les bonimenteurs et les baratineurs fabulateurs] # très bien # alors tu vas partir du début o.k. # tu contournes # sous # sous le titre [les bonimenteurs et les baratineurs fabulateurs] # au niveau de # tu fais un petit arrondi tu reviens	N	
	Loc 2	BCi3	Beta	G	# entre les la table # et la phrase # [les bonimenteurs et les baratineurs fameux] # c' est écrit sous les deux bars # tu	B	
	Loc 1	BCvii7	Alpha	G	le champ de riz # haut droit # tu as [les lumières et les balises vertigineuses] # tu vas descendre # et contourner # les balises tu vas descendre # et contourner # les balises # [les lumières et les balises vertigineuses] # ouais # par le bas et tu # les vertigineuses # ouais # par le bas et tu # [les lumières et les balises vertigineuses] # ouais # champ de riz bim tu descends tout droit	B	
	Loc 1	BDiv8	Alpha	F	encore je monte # je passe pas sous l' écriture [les lumières et les balises vertigineuses] # o.k. je monte # ouais # ouais # le	B	
	Loc 2		Beta	G	est à cette hauteur là # non l' écriture # [les lumières et les balises vertigineuses] # il est à ce niveau là # et il	N	
	Loc 1	BDvi4	Alpha	G	# qui est un magasin # bon # tu as [les lumières et les balises vertigineuses] # bon # alors # le cadeau promis c' est # de la croix début # à droite # et les [les lumières et les balises vertigineuses] # mais les lumières et les balises vertigineuses elles sont et les les lumières et les balises vertigineuses # mais [les lumières et les balises vertigineuses] elles sont légèrement en haut # voilà # donc début ai pas # sur la gauche # tu as quelque chose [sous les lumières et les balises vertigineuses] # o.k. elle est # en dessous en dessous ou vas contourner # jusqu'au niveau parce que moi j' ai que [les lumières et les balises vertigineuses] # jusqu'au niveau du du # l' de balises # tu je je te parle de celui qui est au-dessus # [les lumières et les balises vertigineuses] # ah voilà # ouais # à # mh # # ensuite tu as les bars là il y a [les bonimenteurs et les baratineurs fabulateurs] # tu vas descendre # entre le # le baratineur	B	
							N



**Annexe 2 :**  
**Corpus du CID (Bertrand et *al.*, 2008)**

**Annexe 2 – A :** *Récapitulatif des différentes occurrences constituées de N et/ou ADJ, produites dans le corpus.*

<b>occurrences</b>	<b>Nb</b>
Adj_ADJ_N	11
Adj_N	495
Adj_N_Adj	18
N_Cc_N	24
N_Spd_N	217
N_Spd_N_Adj	15
N_Adj	415
<b>TOTAL</b>	<b>1195</b>

**Annexe 2 – B :** Extrait des tableaux Excel répertoriant chacune des structures N+ADJ produites dans le corpus, qui sont réparties dans 5 catégories différentes. Les structures surlignées en jaune correspondent à des structures relativement figées.

Tri par nbre de syllabes Type 1 : "bons candidats"					
Ref. Stimuli	Stimuli Complet	Nombre de syllabes			
		N	ADJ	N_ADJ	Det + N_ADJ
SR gpd_131 871.59	des relations conflictuelles	3	3	6	7
SR gpd_436 2573.99	la formation doctorale	3	3	6	7
SR gpd_487 2795.12	des trous noirs	1	1	2	3
SR gpd_550 3338.17	un intervenant étranger	4	3	7	8
SR gpd_579 3530.73	une carte bleue	1	1	3	4
SR gpd_581 3539.24	la carte bleue	1	1	3	4
SR gpd_590 3570.15	le code secret	1	2	3	4
Tri par nbre de syllabes Type 2 : absence de déterminant					
Ref. Stimuli	Stimuli Complet	Nombre de syllabes			
		N	ADJ	N_ADJ	syntagme
SR gpd_124 852.28	∅ relations conflictuelles	3	3	6	6
EB gpd_234 1058.76	∅ conflit sain	2	1	3	3
EB gpd_358 1551.27	en sciences physiques	1	2	3	4
EB gpd_359 1556.05	∅ modélisation systématique	5	4	9	9
EB gpd_539 2695.77	en sciences humaines	1	2	3	4
IM gpd_570 2414.25	∅ inspecteur général	3	3	6	6
IM gpd_602 2540.46	∅ inspecteur général	3	3	6	6
Tri par nbre de syllabes Type 3 : précédé/suivi d'une particule					
Ref. Stimuli	Stimuli Complet	Nombre de syllabes			
		N	ADJ	N_ADJ	syntagme
SR gpd_574 3494.33	les cartes bleues effectivement oui	1	1	3	10
SR gpd_590 3570.15	le code secret donc	1	2	3	5
EB gpd_463 2012.32	l'été prochain quoi	2	2	4	5
EB gpd_536 2686.6	une adresse perso quoi	2	2	4	6
EB gpd_551 2776.54	des malades mentaux quoi	2	2	4	6
EB gpd_724 3516.9	les cartes bleues oui	1	1	3	5
IM gpd_213 872.25	la semaine prochaine donc	2	2	4	6
Tri par nbre de syllabes Type 4 : précédé/suivi d'une disfluece					
Ref. Stimuli	Stimuli Complet	Nombre de syllabes			
		N	ADJ	N_ADJ	syntagme
SR gpd_0 0.22	des relations conflictuelles euh	3	3	6	7
SR gpd_2 15.94	des des relations conflictuelles euh	3	3	6	8
SR gpd_156 991.419	des des des visions différentes	2	3	5	8
SR gpd_168 1056.93	le le le conflit sain	2	1	3	6
SR gpd_366 2294.43	une revue anglaise euh	2	2	4	5
SR gpd_609 3657.71	le le le code secret	1	2	3	6
SR gpd_582 3545.12	euh le code secret	1	2	3	4
Tri par nbre de syllabes Type 5 : cumul des types					
Ref. Stimuli	Stimuli Complet	Nombre de syllabes			
		N	ADJ	N_ADJ	syntagme
SR gpd_41 243.84	en position supérieure euh	3	3	6	7
EB gpd_50 240.9	en position supérieure quoi	3	3	6	8
EB gpd_233 1056.39	∅ conflit d déclaré	2	3	5	6
EB gpd_439 1932.64	euh ∅ conflit déontologique euh	2	5	7	7
EB gpd_219 983.181	une guerre larvée euh oui	1	2	3	5
IM gpd_347 1337.81	tu vois ∅ poisson pané	2	2	4	6
IM gpd_403 1632.44	en tissu blanc là	2	1	3	5

## Annexe 3 : Corpus d'Edimbourg (Astésano et al., 2007)

**Annexe 3 :** Répartition des 128 syntagmes sélectionnés pour le test de perception dans les 4 groupes expérimentaux.

Condition Syntactique	ADJ_syll	Groupe 1	Groupe 2	Groupe 3	Groupe 4
		Code Phr	Code Phr	Code Phr	Code Phr
CAS 1	1	s009_locuteur1.wav	s009_locuteur2.wav	s009_locuteur4.wav	s009_locuteur6.wav
CAS 2	1	s020_locuteur2.wav	s020_locuteur1.wav	s020_locuteur6.wav	s020_locuteur4.wav
CAS 1	2	s032_locuteur4.wav	s032_locuteur6.wav	s032_locuteur1.wav	s032_locuteur2.wav
CAS 2	2	s044_locuteur6.wav	s044_locuteur4.wav	s044_locuteur2.wav	s044_locuteur1.wav
CAS 1	3	s056_locuteur6.wav	s056_locuteur1.wav	s056_locuteur2.wav	s056_locuteur4.wav
CAS 2	3	s069_locuteur1.wav	s069_locuteur6.wav	s069_locuteur4.wav	s069_locuteur2.wav
CAS 1	4	s081_locuteur2.wav	s081_locuteur4.wav	s081_locuteur6.wav	s081_locuteur1.wav
CAS 2	4	s093_locuteur4.wav	s093_locuteur2.wav	s093_locuteur1.wav	s093_locuteur6.wav
CAS 1	1	s105_locuteur2.wav	s105_locuteur1.wav	s105_locuteur6.wav	s105_locuteur4.wav
CAS 2	1	s115_locuteur1.wav	s115_locuteur2.wav	s115_locuteur4.wav	s115_locuteur6.wav
CAS 1	2	s128_locuteur6.wav	s128_locuteur4.wav	s128_locuteur2.wav	s128_locuteur1.wav
CAS 2	2	s140_locuteur4.wav	s140_locuteur6.wav	s140_locuteur1.wav	s140_locuteur2.wav
CAS 1	3	s152_locuteur1.wav	s152_locuteur6.wav	s152_locuteur4.wav	s152_locuteur2.wav
CAS 2	3	s163_locuteur6.wav	s163_locuteur1.wav	s163_locuteur2.wav	s163_locuteur4.wav
CAS 1	4	s177_locuteur4.wav	s177_locuteur2.wav	s177_locuteur1.wav	s177_locuteur6.wav
CAS 2	4	s187_locuteur2.wav	s187_locuteur4.wav	s187_locuteur6.wav	s187_locuteur1.wav
CAS 1	1	s201_locuteur4.wav	s201_locuteur6.wav	s201_locuteur2.wav	s201_locuteur1.wav
CAS 2	1	s212_locuteur6.wav	s212_locuteur4.wav	s212_locuteur1.wav	s212_locuteur2.wav
CAS 1	2	s225_locuteur1.wav	s225_locuteur2.wav	s225_locuteur6.wav	s225_locuteur4.wav
CAS 2	2	s235_locuteur2.wav	s235_locuteur1.wav	s235_locuteur4.wav	s235_locuteur6.wav
CAS 1	3	s247_locuteur2.wav	s247_locuteur4.wav	s247_locuteur1.wav	s247_locuteur6.wav
CAS 2	3	s260_locuteur4.wav	s260_locuteur2.wav	s260_locuteur6.wav	s260_locuteur1.wav
CAS 1	4	s273_locuteur6.wav	s273_locuteur1.wav	s273_locuteur4.wav	s273_locuteur2.wav
CAS 2	4	s285_locuteur1.wav	s285_locuteur6.wav	s285_locuteur2.wav	s285_locuteur4.wav
CAS 1	1	s297_locuteur6.wav	s297_locuteur4.wav	s297_locuteur1.wav	s297_locuteur2.wav
CAS 2	1	s307_locuteur4.wav	s307_locuteur6.wav	s307_locuteur2.wav	s307_locuteur1.wav
CAS 1	2	s319_locuteur2.wav	s319_locuteur1.wav	s319_locuteur4.wav	s319_locuteur6.wav
CAS 2	2	s333_locuteur1.wav	s333_locuteur2.wav	s333_locuteur6.wav	s333_locuteur4.wav
CAS 1	3	s343_locuteur4.wav	s343_locuteur2.wav	s343_locuteur6.wav	s343_locuteur1.wav
CAS 2	3	s357_locuteur2.wav	s357_locuteur4.wav	s357_locuteur1.wav	s357_locuteur6.wav
CAS 1	4	s367_locuteur1.wav	s367_locuteur6.wav	s367_locuteur2.wav	s367_locuteur4.wav
CAS 2	4	s380_locuteur6.wav	s380_locuteur1.wav	s380_locuteur4.wav	s380_locuteur2.wav



## Annexe 4 :

### Documents remplis par les participants de l'expérimentation

**Annexe 4 – A : Formulaire de consentement éclairé.**



#### **Formulaire de Consentement libre, éclairé et exprès Expériences comportementales en psycholinguistique**

Je certifie avoir donné mon accord pour participer à une étude comportementale de psycholinguistique. J'accepte volontairement de participer à cette étude et je comprends que ma participation n'est pas obligatoire et que je peux stopper ma participation à tout moment sans avoir à me justifier ni encourir aucune responsabilité. Mon consentement ne décharge pas les organisateurs de la recherche de leurs responsabilités et je conserve tous mes droits garantis par la loi.

Au cours de cette expérience, j'accepte que soient recueillies des données de perception de la parole. Je comprends que les informations recueillies sont strictement confidentielles et à usage exclusif des investigateurs concernés.

J'ai été informé que mon identité n'apparaîtra dans aucun rapport ou publication et que toute information me concernant sera traitée de façon confidentielle. J'accepte que les données enregistrées à l'occasion de cette étude puissent être conservées dans une base de données et faire l'objet d'un traitement informatisé non nominatif par l'Unité de Recherche Interdisciplinaire OCTOGONE-Lordat (EA4156) et par l'Unité Mixte de Recherche CLLE-ERSS (5263).

Date : .....

Nom du volontaire : .....

Signature du volontaire (précédée de la mention « lu et approuvé ») :

Nom de l'expérimentateur : .....

Signature de l'expérimentateur : .....

Nom de l'expérience : .....

## FICHE DE RENSEIGNEMENTS

**SUJET N° :**

**GROUPE :**

**Nom :**

**Prénom :**

**Sexe :**

**Age :**

**Région de naissance :**

**Niveau d'études :**

**Adresse mail :**

- Droitier**
- Gaucher**

**Pratique musicale :**

Quel instrument ?

Combien d'années de pratique ?

Fréquence de la pratique ?

**Pratique des Jeux vidéo :**

Fréquence de la pratique ?

**Langue(s) parlée(s) :**

Quelle(s) langue(s) ?

Niveau dans chaque langue (scolaire ou bilinguisme) ?

-----

Avez-vous déjà participé à une expérimentation sur la perception de la parole ?

Difficultés rencontrées au cours de l'expérimentation :

Annexe 4 – C : Récapitulatif des renseignements donnés par chaque participant.

participant	groupe expé	sexe	age	région_naissance	niveau étude	gaucher	droitier	pratique musicale	pratique jeux video	anglais	allemand	italien	LSF
1	1	F	30	MIDI-PYRENEES	BAC+5		X	N	O	COURANT	X	X	X
2	2	F	33	PICARDIE	BAC+8		X	O	N	COURANT	X	COURANT	COURANT
3	3	F	26	RHONE-ALPES	BAC+8		X	N	N	SCOLAIRE	X	SCOLAIRE	X
4	4	F	21	MIDI-PYRENEES	BAC+1		X	O	O	SCOLAIRE	X	X	X
5	1	F	25	ILE DE France	BAC+8		X	O	N	SCOLAIRE	SCOLAIRE	X	X
6	2	F	21	MIDI-PYRENEES	BAC+1		X	O	N	SCOLAIRE	X	X	X
7	3	F	24	PAYS DE LA LOIRE	BAC+5		X	N	O	SCOLAIRE	X	X	X
8	4	F	22	ILE DE France	BAC+3		X	N	O	SCOLAIRE	X	X	X
9	1	F	25	RHONE-ALPES	BAC+8		X	O	O	SCOLAIRE	X	X	X
10	2	F	18	AQUITAINE	BAC+1		X	O	O	COURANT	X	X	X
11	3	F	43	ALSACE	BAC+5		X	O	N	COURANT	X	X	X
12	4	F	18	LANGUEDOC-ROUSSILLON	BAC+1		X	N	N	SCOLAIRE	X	X	X
13	1	F	28	AUVERGNE	BAC+5		X	N	N	SCOLAIRE	X	X	X
14	2	F	25	HAUTE-SAVOIE	BAC+5		X	O	O	SCOLAIRE	INITIATIONS	X	X
15	3	F	21	PROVENCE ALPES COTES D'AZU	BAC+2		X	N	O	SCOLAIRE	X	X	X
16	4	F	18	GUYANNE	BAC+1		X	O	O	SCOLAIRE	X	X	X
17	1	F	18	AQUITAINE	BAC+1		X	O	N	SCOLAIRE	SCOLAIRE	X	X
18	2	F	18	AQUITAINE	BAC+1		X	O	N	SCOLAIRE	X	X	X
19	3	F	20	MIDI-PYRENEES	BAC+1		X	N	N	SCOLAIRE	X	X	X
20	4	F	24	RHONE-ALPES	BAC+5		X	N	N	SCOLAIRE	X	X	X
21	1	F	18	MIDI-PYRENEES	BAC+1		X	N	N	SCOLAIRE	X	X	X
22	2	F	21	MIDI-PYRENEES	BAC+1	X		O	O	COURANT	X	SCOLAIRE	X
23	3	F	19	MIDI-PYRENEES	BAC+1		X	N	N	SCOLAIRE	X	X	X
24	4	F	18	MIDI-PYRENEES	BAC+1		X	O	N	SCOLAIRE	X	X	X
25	1	F	19	LANGUEDOC-ROUSSILLON	BAC+1	X		O	O	SCOLAIRE	SCOLAIRE	X	X
26	2	F	18	MIDI-PYRENEES	BAC+1	X		N	N	SCOLAIRE	SCOLAIRE	X	X
27	3	F	18	MIDI-PYRENEES	BAC+1		X	N	N	SCOLAIRE	X	X	X
28	4	F	18	MIDI-PYRENEES	BAC+1		X	N	N	SCOLAIRE	X	X	X
29	1	M	19	LANGUEDOC-ROUSSILLON	BAC+1	X		N	O	SCOLAIRE	X	X	X
30	2	F	18	ILE DE France	BAC+1	X		N	N	SCOLAIRE	X	X	INITIATION
31	3	F	20	MIDI-PYRENEES	BAC+1		X	N	N	SCOLAIRE	X	X	X
32	4	F	18	LANGUEDOC-ROUSSILLON	BAC+1		X	O	N	SCOLAIRE	X	X	X
33	1	F	19	MIDI-PYRENEES	BAC+1		X	N	O	SCOLAIRE	X	X	X
34	2	M	20	MIDI-PYRENEES	BAC+1		X	O	O	COURANT	X	INITIATIONS	X
35	3	F	19	MIDI-PYRENEES	BAC+1		X	N	N	SCOLAIRE	X	X	X
36	4	F	22	MIDI-PYRENEES	BAC+1		X	N	N	SCOLAIRE	X	X	X
37	1	F	18	ILE DE France	BAC+1		X	N	O	SCOLAIRE	X	X	X
38	2	F	20	MIDI-PYRENEES	BAC+1		X	O	N	SCOLAIRE	X	X	X
39	3	F	22	POITOU-CHARENTE	BAC+5		X	N	N	SCOLAIRE	X	X	X
40	4	F	18	MIDI-PYRENEES	BAC+1		X	O	N	SCOLAIRE	X	X	X
41	1	F	23	MIDI-PYRENEES	BAC+4		X	N	N	SCOLAIRE	SCOLAIRE	X	X
42	2	M	27	MIDI-PYRENEES	BAC+2		X	O	N	SCOLAIRE	X	X	X
43	3	F	18	MIDI-PYRENEES	BAC+1		X	O	N	SCOLAIRE	X	X	X
44	4	M	27	PICARDIE	BAC+5		X	O	N	COURANT	X	X	X
45	1	F	18	MIDI-PYRENEES	BAC+1		X	N	O	SCOLAIRE	SCOLAIRE	X	X
46	2	F	23	MIDI-PYRENEES	BAC+8		X	N	N	SCOLAIRE	X	X	X
47	3	F	27	LANGUEDOC-ROUSSILLON	BAC+8	X		O	N	COURANT	SCOLAIRE	X	X
48	4	F	27	PAYS DE LA LOIRE	BAC+4	X		N	N	SCOLAIRE	SCOLAIRE	X	X
49	1	F	23	PAYS DE LA LOIRE	BAC+5		X	O	O	SCOLAIRE	X	X	X
50	2	F	27	MIDI-PYRENEES	BAC+8		X	N	N	SCOLAIRE	X	X	X
51	3	F	30	ILE DE France	BAC+8		X	O	N	SCOLAIRE	X	X	X
52	4	F	26	BRETAGNE	BAC+8		X	O	O	COURANT	X	X	X
53	1	M	22	BRETAGNE	BAC+2		X	O	O	COURANT	X	X	X
54	2	F	18	PROVENCE ALPES COTES D'AZU	BAC+1		X	N	N	SCOLAIRE	X	X	X
55	3	M	22	AQUITAINE	BAC+4		X	O	O	COURANT	INITIATION	X	X
56	4	F	28	MIDI-PYRENEES	BAC+8		X	O	N	SCOLAIRE	X	X	X
57	1	M	38	MIDI-PYRENEES	BAC+4		X	O	O	SCOLAIRE	X	X	X
58	2	F	19	ILE DE France	BAC+1		X	N	N	SCOLAIRE	X	X	X
59	3	F	21	BOURGOGNE	BAC+4		X	N	N	SCOLAIRE	X	SCOLAIRE	X
60	4	F	19	MIDI-PYRENEES	BAC+1		X	N	N	SCOLAIRE	X	X	X
61	1	F	26	BRETAGNE	BAC+8		X	O	O	SCOLAIRE	X	X	X
62	2	F	18	AQUITAINE	BAC+1		X	O	N	SCOLAIRE	X	X	X
63	3	F	28	BRETAGNE	BAC+5		X	O	N	COURANT	X	X	X
64	4	F	43	MIDI-PYRENEES	BAC+8		X	O	N	SCOLAIRE	SCOLAIRE	SCOLAIRE	X
65	1	F	21	MIDI-PYRENEES	BAC+1		X	N	O	COURANT	X	X	X
66	2	F	24	BRETAGNE	BAC+5		X	O	N	COURANT	X	INITIATIONS	X
67	3	F	24	MIDI-PYRENEES	BAC+5		X	N	O	COURANT	X	SCOLAIRE	X
68	4	F	22	MIDI-PYRENEES	BAC+5		X	N	N	COURANT	X	X	X
69	1	F	19	MIDI-PYRENEES	BAC+1		X	N	O	SCOLAIRE	X	X	X
70	2	F	27	LIMOUSIN	BAC+5		X	N	O	SCOLAIRE	X	INITIATIONS	X
71	3	F	26	AQUITAINE	BAC+5		X	O	N	COURANT	X	X	X
72	4	F	21	MIDI-PYRENEES	BAC+5		X	N	N	SCOLAIRE	X	X	X
73	1	M	23	MIDI-PYRENEES	BAC+4		X	N	O	COURANT	X	X	X
74	2	F	18	MIDI-PYRENEES	BAC+1		X	N	N	SCOLAIRE	X	X	X
75	3	F	19	ILE DE France	BAC+1		X	N	N	COURANT	X	COURANT	X
76	4	F	23	ILE DE France	BAC+4		X	O	N	SCOLAIRE	X	X	X
77	1	F	35	AQUITAINE	BAC+8		X	O	N	SCOLAIRE	INITIATIONS	X	X
78	2	F	20	RHONE-ALPES	BAC+3		X	O	N	SCOLAIRE	SCOLAIRE	X	X
79	3	M	31	MIDI-PYRENEES	BAC+2		X	N	N	SCOLAIRE	X	X	X
80	4	F	35	LOIRET	BAC+8		X	N	O	COURANT	X	X	X

allemand	italien	LSF	espagnol	arabe	japonais	occitan	russe	roumain	chinois	coreen	malgache	polonais	croate	hongrois	portugais	Durée de l'expérimentation en min
X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75
X	COURANT	COURANT	X	X	X	X	X	X	X	X	X	X	X	X	X	40
X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	X	X	50
X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	75
SCOLAIRE	X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	SCOLAIRE	60
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	50
X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	55
X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	50
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	50
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	45
X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	60
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	40
X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	30
INITIATION	X	X	SCOLAIRE	X	INITIATION	X	X	X	X	X	X	X	X	X	X	50
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	60
X	X	X	COURANT	X	X	X	X	X	X	X	X	X	X	X	COURANT	50
SCOLAIRE	X	X	INITIATION	X	X	X	X	X	X	X	X	X	X	X	X	75
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	60
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	30
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	40
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	55
X	SCOLAIRE	X	SCOLAIRE	X	INITIATION	X	X	X	X	X	X	X	X	X	X	50
X	X	X	SCOLAIRE	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	70
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	45
SCOLAIRE	X	X	X	X	INITIATION	X	X	X	X	X	X	X	X	X	X	45
SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	40
X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	35
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	30
X	X	X	COURANT	X	X	X	X	X	X	X	X	X	X	X	X	65
X	X	INITIATION	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	40
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	45
X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	50
X	X	X	SCOLAIRE	COURANT	X	X	X	X	X	X	X	X	X	X	X	35
X	INITIATION	X	X	X	X	X	INITIATION	INITIATION	X	X	X	X	X	X	X	50
X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	45
X	X	X	X	X	X	X	X	X	INITIATION	INITIATION	X	X	X	X	X	75
X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	50
X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	70
X	X	X	COURANT	INITIATION	X	X	X	X	X	X	X	X	X	X	X	45
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	70
SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	40
X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	35
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	50
SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	60
SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	65
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	40
SCOLAIRE	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	50
SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	45
X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	40
X	X	X	COURANT	X	X	X	X	X	X	X	X	X	X	X	X	45
X	X	X	COURANT	X	X	X	X	X	X	X	X	X	X	X	X	30
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	30
X	X	X	X	X	X	X	INITIATION	X	X	X	X	X	X	X	X	40
X	X	X	COURANT	X	X	X	X	X	X	X	X	X	X	X	X	45
INITIATION	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	65
X	X	X	COURANT	X	X	X	X	X	X	X	X	X	X	X	X	30
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	55
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	INITIATION	X	X	X	X	30
X	SCOLAIRE	X	COURANT	X	X	X	X	X	X	X	X	X	X	X	X	50
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	35
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	50
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	40
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	INITIATION	X	X	X	65
SCOLAIRE	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	X	X	30
X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	50
X	INITIATION	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	60
X	SCOLAIRE	X	SCOLAIRE	X	X	X	INITIATION	X	X	X	X	X	INITIATION	SCOLAIRE	X	55
X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	40
X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	35
X	INITIATION	X	SCOLAIRE	X	INITIATION	X	X	X	X	X	X	X	X	X	X	75
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	100
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	75
X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	COURANT	45
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	60
X	COURANT	X	COURANT	X	X	X	X	X	X	X	X	X	X	X	X	70
X	X	X	COURANT	X	X	X	X	X	X	X	X	X	X	X	INITIATION	60
INITIATION	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	50
SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	45
X	X	X	SCOLAIRE	X	X	X	X	X	X	X	X	X	X	X	X	50
X	X	X	COURANT	X	X	X	X	X	X	X	X	X	X	X	X	50

**Annexe 4 – D : Reçu d'indemnisation.**



**Sujet de l'étude :**

**Perception des groupements, des prééminences et des ruptures prosodiques du français**

**Date :**

**RECU D'INDEMNISATION**

Je soussigné (nom, prénom) : .....

Demeurant à : .....

N° Sécurité Sociale : /..... /..... /..... /..... /..... /..... /..... /

Volontaire et consentant à l'étude mentionnée ci-dessus, reconnais avoir reçu ce jour, au titre de compensation de ma participation, la somme de 20€.

Signature



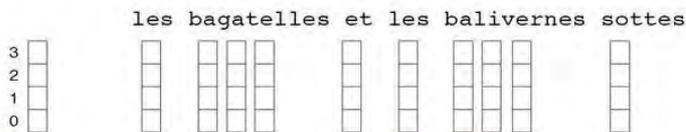
## Annexe 5 : Consignes du test de perception

**Annexe 5 – A : consignes des tâches de proéminence, frontière et groupement.**

### Tâche Proéminence

Vous allez entendre des énoncés sur lesquels vous devrez juger du **degré de mise en relief de chaque syllabe**.

Pour cela, des "barrettes" sont dessinées sous chaque syllabe. Elles sont composées de 4 cases superposées qui représentent une force graduelle de mise en relief :



**3** renvoie au degré de proéminence **le plus fort**

**2** renvoie à un degré de proéminence **un peu plus élevé**

**1** renvoie au degré de proéminence **le plus faible**

**0** indique **l'absence** de proéminence

Cochez la case qui correspond selon vous au degré de proéminence perçu.

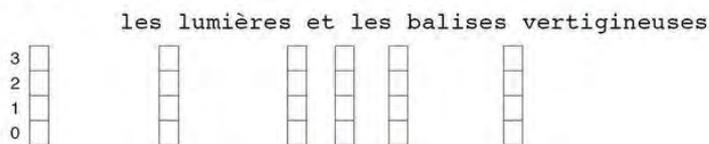
Chaque énoncé pourra être répété jusqu'à 5 fois (en cliquant sur l'image du haut-parleur) pour vous permettre de prendre le temps d'effectuer la tâche.

[Cliquer ici pour continuer](#)

### Tâche Rupture

Vous allez entendre des énoncés sur lesquels vous devrez juger du **degré de rupture entre chaque mot**.

Pour cela, des "barrettes" sont dessinées entre chaque mot. Elles sont composées de 4 cases superposées qui représentent une force graduelle de rupture :



**3** renvoie au degré de rupture **le plus fort**

**2** renvoie à un degré de rupture **un peu plus élevé**

**1** renvoie au degré de rupture **le plus faible**

**0** indique **l'absence** de rupture

Cochez la case qui correspond selon vous au degré de rupture perçu.

Chaque énoncé pourra être répété jusqu'à 5 fois (en cliquant sur l'image du haut-parleur) pour vous permettre de prendre le temps d'effectuer la tâche.

[Cliquer ici pour continuer](#)

## Tâche Groupement

Vous allez entendre des énoncés dans lesquels **certains éléments sont liés plus étroitement les uns aux autres**. Vous aurez à choisir parmi 4 possibilités de regroupement:

les gants et les bas libérateurs

**G1** : [les gants] forme **un premier groupement** et [les bas libérateurs] sont regroupés pour former **un second groupement**

**G2** : [les gants et les bas] sont regroupés pour former **un premier groupement** et [libérateurs] forme **un second groupement**

**G3** : [les gants] / [et les bas] / [libérateurs] forment **trois groupements à part entière**

**G4** : [les gants et les bas libérateurs] forment **un seul et même groupement**

Cochez la case qui correspond selon vous au type de regroupement que vous avez perçu

Chaque énoncé pourra être répété jusqu'à 5 fois (en cliquant sur l'image du haut-parleur) pour vous permettre de prendre le temps d'effectuer la tâche.

[Cliquer ici pour continuer](#)

### **Consignes du test de perception**

Merci de participer à ce test de perception de la parole.

Vous allez entendre des énoncés qui ont été extraits d'une phrase.

On vous demande d'effectuer successivement 3 tâches différentes, qui ont globalement trait à la manière dont :

- les mots sont plus ou moins regroupés les uns avec les autres.
- les mots sont plus ou moins séparés les uns des autres.
- certains éléments ressortent par rapport à d'autres.

Vous pourrez écouter chaque énoncé jusqu'à 5 fois (en cliquant sur l'image du haut-parleur) pour vous permettre de prendre le temps d'effectuer chacune des tâches. Vous avez tout le temps que vous souhaitez pour effectuer les tâches, nous ne prenons pas en compte la vitesse de réponse.

En parole, comme en musique, les éléments sont structurés et organisés, ce qui donne une certaine "musicalité" aux énoncés. Vous pourrez vous appuyer sur cette analogie entre parole et musique pour effectuer les 3 tâches qui vont vous être présentées.

Vous allez commencer par une petite séance de familiarisation avec le matériel linguistique avant d'effectuer chacune des tâches.

L'expérimentateur est là pour répondre à toute question : n'hésitez pas!



**Annexe 6 :**

**Résultats du lien de dépendance entre le groupement et la condition syntaxique.**

Khi deux de Pearson		Valeur	ddl		p-value	
		1139,84	3		0,000	
<b>TABLEAU DE CONTINGENCE</b>						
		groupement				TOTAL
		1	2	3	4	
Condition 1	Effectif	628	2	82	8	720
	Effectif théorique	323,5	124	95,5	177	720
	%	87,20%	0,30%	11,40%	1,10%	100,00%
	Résidu	304,5	-122	-13,5	-169	
Condition 2	Effectif	19	246	109	346	720
	Effectif théorique	323,5	124	95,5	177	720
	%	2,60%	34,20%	15,10%	48,10%	100,00%
	Résidu	-304,5	122	13,5	169	
TOTAL	Effectif	647	248	191	354	1440
	Effectif théorique					
	%	44,90%	17,20%	13,30%	24,60%	100,00%



## Annexe 7 :

### Résultats de la distribution des scores de frontières par groupement.

*Annexe 7 – A : scores de frontières après D1 (premier déterminant).*

Khi deux de Pearson		Valeur	ddl	p-value		
		16,358	12	0,175		
<b>TABLEAU DE CONTINGENCE</b>						
		D1				Total
		0	1	2	3	
<b>G11</b>	<b>Effectif</b>	378	196	46	8	628
	<b>Effectif théorique</b>	394,3	188,7	37,8	7,1	628,0
	<b>%</b>	60,2%	31,2%	7,3%	1,3%	100,0%
	<b>Résidu</b>	-16,3	7,3	8,2	,9	
<b>G22</b>	<b>Effectif</b>	159	72	13	2	246
	<b>Effectif théorique</b>	154,5	73,9	14,8	2,8	246,0
	<b>%</b>	64,6%	29,3%	5,3%	0,8%	100,0%
	<b>Résidu</b>	4,5	-1,9	-1,8	-,8	
<b>G31</b>	<b>Effectif</b>	44	27	9	2	82
	<b>Effectif théorique</b>	51,5	24,6	4,9	,9	82,0
	<b>%</b>	53,7%	32,9%	11,0%	2,4%	100,0%
	<b>Résidu</b>	-7,5	2,4	4,1	1,1	
<b>G32</b>	<b>Effectif</b>	76	30	2	1	109
	<b>Effectif théorique</b>	68,4	32,8	6,6	1,2	109,0
	<b>%</b>	69,7%	27,5%	1,8%	0,9%	100,0%
	<b>Résidu</b>	7,6	-2,8	-4,6	-,2	
<b>G42</b>	<b>Effectif</b>	229	99	15	3	346
	<b>Effectif théorique</b>	217,3	104,0	20,8	3,9	346,0
	<b>%</b>	66,2%	28,6%	4,3%	0,9%	100,0%
	<b>Résidu</b>	11,7	-5,0	-5,8	-,9	
<b>TOTAL</b>	<b>Effectif</b>	886	424	85	16	1411
	<b>Effectif théorique</b>	886,0	424,0	85,0	16,0	1411,0
	<b>%</b>	62,8%	30,0%	6,0%	1,1%	100,0%

Annexe 7 – B : scores de frontières après N1.

Khi deux de Pearson		Valeur	ddl	p-value		
		899,917	12	0,000		
<b>TABLEAU DE CONTINGENCE</b>						
		N1				Total
		0	1	2	3	
<b>G11</b>	<b>Effectif</b>	8	64	226	330	628
	<b>Effectif théorique</b>	166,0	138,9	142,4	180,7	628,0
	<b>%</b>	1,3%	10,2%	36,0%	52,5%	100,0%
	<b>Résidu</b>	-158,0	-74,9	83,6	149,3	
<b>G22</b>	<b>Effectif</b>	129	79	27	11	246
	<b>Effectif théorique</b>	65,0	54,4	55,8	70,8	246,0
	<b>%</b>	52,4%	32,1%	11,0%	4,5%	100,0%
	<b>Résidu</b>	64,0	24,6	-28,8	-59,8	
<b>G31</b>	<b>Effectif</b>	1	6	28	47	82
	<b>Effectif théorique</b>	21,7	18,1	18,6	23,6	82,0
	<b>%</b>	1,2%	7,3%	34,1%	57,3%	100,0%
	<b>Résidu</b>	-20,7	-12,1	9,4	23,4	
<b>G32</b>	<b>Effectif</b>	27	53	19	10	109
	<b>Effectif théorique</b>	28,8	24,1	24,7	31,4	109,0
	<b>%</b>	24,8%	48,6%	17,4%	9,2%	100,0%
	<b>Résidu</b>	-1,8	28,9	-5,7	-21,4	
<b>G42</b>	<b>Effectif</b>	208	110	20	8	346
	<b>Effectif théorique</b>	91,5	76,5	78,5	99,6	346,0
	<b>%</b>	60,1%	31,8%	5,8%	2,3%	100,0%
	<b>Résidu</b>	116,5	33,5	-58,5	-91,6	
<b>TOTAL</b>	<b>Effectif</b>	373	312	320	406	1411
	<b>Effectif théorique</b>	373,0	312,0	320,0	406,0	1411,0
	<b>%</b>	26,4%	22,1%	22,7%	28,8%	100,0%

Annexe 7 – C : scores de frontières après cc (conjonction de coordination).

Khi deux de Pearson		Valeur	ddl	p-value		
		45,933	12	0,000		
<b>TABLEAU DE CONTINGENCE</b>						
		cc				Total
		0	1	2	3	
G11	Effectif	379	173	62	14	628
	Effectif théorique	416,6	149,1	47,6	14,7	628,0
	%	60,4%	27,5%	9,9%	2,2%	100,0%
	Résidu	-37,6	23,9	14,4	-,7	
G22	Effectif	179	46	16	5	246
	Effectif théorique	163,2	58,4	18,7	5,8	246,0
	%	72,8%	18,7%	6,5%	2,0%	100,0%
	Résidu	15,8	-12,4	-2,7	-,8	
G31	Effectif	44	22	13	3	82
	Effectif théorique	54,4	19,5	6,2	1,9	82,0
	%	53,7%	26,8%	15,9%	3,7%	100,0%
	Résidu	-10,4	2,5	6,8	1,1	
G32	Effectif	87	15	5	2	109
	Effectif théorique	72,3	25,9	8,3	2,5	109,0
	%	79,8%	13,8%	4,6%	1,8%	100,0%
	Résidu	14,7	-10,9	-3,3	-,5	
G42	Effectif	247	79	11	9	346
	Effectif théorique	229,5	82,1	26,2	8,1	346,0
	%	71,4%	22,8%	3,2%	2,6%	100,0%
	Résidu	17,5	-3,1	-15,2	,9	
TOTAL	Effectif	936	335	107	33	1411
	Effectif théorique	936,0	335,0	107,0	33,0	1411,0
	%	66,3%	23,7%	7,6%	2,3%	100,0%

Annexe 7 – D : scores de frontières après D2 (deuxième déterminant).

Khi deux de Pearson		Valeur	ddl	p-value		
		33,199	12	0,001		
<b>TABLEAU DE CONTINGENCE</b>						
		D2				Total
		0	1	2	3	
<b>G11</b>	<b>Effectif</b>	341	226	44	17	628
	<b>Effectif théorique</b>	365,4	208,3	40,5	13,8	628,0
	<b>%</b>	54,3%	36,0%	7,0%	2,7%	100,0%
	<b>Résidu</b>	-24,4	17,7	3,5	3,2	
<b>G22</b>	<b>Effectif</b>	152	71	19	4	246
	<b>Effectif théorique</b>	143,1	81,6	15,9	5,4	246,0
	<b>%</b>	61,8%	28,9%	7,7%	1,6%	100,0%
	<b>Résidu</b>	8,9	-10,6	3,1	-1,4	
<b>G31</b>	<b>Effectif</b>	37	31	12	2	82
	<b>Effectif théorique</b>	47,7	27,2	5,3	1,8	82,0
	<b>%</b>	45,1%	37,8%	14,6%	2,4%	100,0%
	<b>Résidu</b>	-10,7	3,8	6,7	,2	
<b>G32</b>	<b>Effectif</b>	64	39	2	4	109
	<b>Effectif théorique</b>	63,4	36,2	7,0	2,4	109,0
	<b>%</b>	58,7%	35,8%	1,8%	3,7%	100,0%
	<b>Résidu</b>	,6	2,8	-5,0	1,6	
<b>G42</b>	<b>Effectif</b>	227	101	14	4	346
	<b>Effectif théorique</b>	201,3	114,8	22,3	7,6	346,0
	<b>%</b>	65,6%	29,2%	4,0%	1,2%	100,0%
	<b>Résidu</b>	25,7	-13,8	-8,3	-3,6	
<b>TOTAL</b>	<b>Effectif</b>	821	468	91	31	1411
	<b>Effectif théorique</b>	821,0	468,0	91,0	31,0	1411,0
	<b>%</b>	58,2%	33,2%	6,4%	2,2%	100,0%

Annexe 7 – E : scores de frontières après N2.

Khi deux de Pearson		Valeur	ddl	p-value		
		197,499	12	0,000		
<b>TABLEAU DE CONTINGENCE</b>						
		N2				Total
		0	1	2	3	
<b>G11</b>	<b>Effectif</b>	242	244	114	28	628
	<b>Effectif théorique</b>	183,8	243,5	147,3	53,4	628,0
	<b>%</b>	38,5%	38,9%	18,2%	4,5%	100,0%
	<b>Résidu</b>	58,2	,5	-33,3	-25,4	
<b>G22</b>	<b>Effectif</b>	29	77	90	50	246
	<b>Effectif théorique</b>	72,0	95,4	57,7	20,9	246,0
	<b>%</b>	11,8%	31,3%	36,6%	20,3%	100,0%
	<b>Résidu</b>	-43,0	-18,4	32,3	29,1	
<b>G31</b>	<b>Effectif</b>	9	36	21	16	82
	<b>Effectif théorique</b>	24,0	31,8	19,2	7,0	82,0
	<b>%</b>	11,0%	43,9%	25,6%	19,5%	100,0%
	<b>Résidu</b>	-15,0	4,2	1,8	9,0	
<b>G32</b>	<b>Effectif</b>	7	49	44	9	109
	<b>Effectif théorique</b>	31,9	42,3	25,6	9,3	109,0
	<b>%</b>	6,4%	45,0%	40,4%	8,3%	100,0%
	<b>Résidu</b>	-24,9	6,7	18,4	-,3	
<b>G42</b>	<b>Effectif</b>	126	141	62	17	346
	<b>Effectif théorique</b>	101,3	134,1	81,2	29,4	346,0
	<b>%</b>	36,4%	40,8%	17,9%	4,9%	100,0%
	<b>Résidu</b>	24,7	6,9	-19,2	-12,4	
<b>TOTAL</b>	<b>Effectif</b>	413	547	331	120	1411
	<b>Effectif théorique</b>	413,0	547,0	331,0	120,0	1411,0
	<b>%</b>	29,3%	38,8%	23,5%	8,5%	100,0%

*Annexe 7 – F : scores de frontières après ADJ.*

		<b>ADJ  </b>	<b>Total</b>
		<b>0</b>	
<b>G11</b>	<b>Effectif</b>	628	628
	<b>Effectif théorique</b>	628,0	628,0
	<b>%</b>	100,0%	100,0%
	<b>Résidu</b>	,0	
<b>G22</b>	<b>Effectif</b>	246	246
	<b>Effectif théorique</b>	246,0	246,0
	<b>%</b>	100,0%	100,0%
	<b>Résidu</b>	,0	
<b>G31</b>	<b>Effectif</b>	82	82
	<b>Effectif théorique</b>	82,0	82,0
	<b>%</b>	100,0%	100,0%
	<b>Résidu</b>	,0	
<b>G32</b>	<b>Effectif</b>	109	109
	<b>Effectif théorique</b>	109,0	109,0
	<b>%</b>	100,0%	100,0%
	<b>Résidu</b>	,0	
<b>G42</b>	<b>Effectif</b>	346	346
	<b>Effectif théorique</b>	346,0	346,0
	<b>%</b>	100,0%	100,0%
	<b>Résidu</b>	,0	
<b>TOTAL</b>	<b>Effectif</b>	1411	1411
	<b>Effectif théorique</b>	1411,0	1411,0
	<b>%</b>	100,0%	100,0%

## Annexe 8 :

### Résultats de la distribution des scores de frontières par taille du constituant.

*Annexe 8 – A : scores de frontières après N2 par le nombre de syllabes de N en G22.*

Khi deux de Pearson		Valeur	ddl		p-value		
		18,855	6		0,004		
<b>TABLEAU DE CONTINGENCE</b>							
			N2				TOTAL
			0	1	2	3	
en G22 en fonction du nombre de syllabes de N	2	Effectif	7	27	20	5	59
		Effectif théorique	7,0	18,5	21,6	12,0	59,0
		%	11,9%	45,8%	33,9%	8,5%	100,0%
		Résidu	,0	8,5	-1,6	-7,0	
	3	Effectif	7	26	41	30	104
		Effectif théorique	12,3	32,6	38,0	21,1	104,0
		%	6,7%	25,0%	39,4%	28,8%	100,0%
		Résidu	-5,3	-6,6	3,0	8,9	
	4	Effectif	15	24	29	15	83
		Effectif théorique	9,8	26,0	30,4	16,9	83,0
		%	18,1%	28,9%	34,9%	18,1%	100,0%
		Résidu	5,2	-2,0	-1,4	-1,9	
<b>TOTAL</b>		Effectif	29	77	90	50	246
		Effectif théorique	29,0	77,0	90,0	50,0	246,0
		%	11,8%	31,3%	36,6%	20,3%	100,0%

**Annexe 8 – B : scores de frontières après N2 par le nombre de syllabes de ADJ en G22.**

Khi deux de Pearson		Valeur	ddl	p-value			
		17,066	6	0,009			
TABLEAU DE CONTINGENCE							
			N2				TOTAL
			0	1	2	3	
en G22 en fonction du nombre de syllabes de ADJ	2	Effectif	9	20	46	20	95
		Effectif théorique	11,2	29,7	34,8	19,3	95,0
		%	9,5%	21,1%	48,4%	21,1%	100,0%
		Résidu	-2,2	-9,7	11,2	,7	
	3	Effectif	12	32	20	22	86
		Effectif théorique	10,1	26,9	31,5	17,5	86,0
		%	14,0%	37,2%	23,3%	25,6%	100,0%
		Résidu	1,9	5,1	-11,5	4,5	
	4	Effectif	8	25	24	8	65
		Effectif théorique	7,7	20,3	23,8	13,2	65,0
		%	12,3%	38,5%	36,9%	12,3%	100,0%
		Résidu	,3	4,7	,2	-5,2	
TOTAL		Effectif	29	77	90	50	246
		Effectif théorique	29,0	77,0	90,0	50,0	246,0
		%	11,8%	31,3%	36,6%	20,3%	100,0%

**Annexe 8 – C : scores de frontières après N2 par le nombre de syllabes additionné du syntagme D1+N1+cc+D2+N2.**

Khi deux de Pearson		Valeur	ddl		p-value		
		46,007	18		0,000		
<b>TABEAU DE CONTINGENCE</b>							
			N2				TOTAL
			0	1	2	3	
en G22 en fonction du nombre de syllabes additionné de : D1+N1+cc+D2+N2+ADJ	9	Effectif	1	7	7	3	18
		Effectif théorique	2,1	5,6	6,6	3,7	18,0
		%	5,6%	38,9%	38,9%	16,7%	100,0%
		Résidu	-1,1	1,4	,4	-,7	
	10	Effectif	3	13	5	1	22
		Effectif théorique	2,6	6,9	8,0	4,5	22,0
		%	13,6%	59,1%	22,7%	4,5%	100,0%
		Résidu	,4	6,1	-3,0	-3,5	
	11	Effectif	5	12	33	10	60
		Effectif théorique	7,1	18,8	22,0	12,2	60,0
		%	8,3%	20,0%	55,0%	16,7%	100,0%
		Résidu	-2,1	-6,8	11,0	-2,2	
	12	Effectif	2	15	10	16	43
		Effectif théorique	5,1	13,5	15,7	8,7	43,0
		%	4,7%	34,9%	23,3%	37,2%	100,0%
		Résidu	-3,1	1,5	-5,7	7,3	
	13	Effectif	9	14	20	13	56
		Effectif théorique	6,6	17,5	20,5	11,4	56,0
		%	16,1%	25,0%	35,7%	23,2%	100,0%
		Résidu	2,4	-3,5	-,5	1,6	
	14	Effectif	7	4	5	5	21
		Effectif théorique	2,5	6,6	7,7	4,3	21,0
		%	33,3%	19,0%	23,8%	23,8%	100,0%
		Résidu	4,5	-2,6	-2,7	,7	
	15	Effectif	2	12	10	2	26
		Effectif théorique	3,1	8,1	9,5	5,3	26,0
		%	7,7%	46,2%	38,5%	7,7%	100,0%
		Résidu	-1,1	3,9	,5	-3,3	
<b>TOTAL</b>		Effectif	29	77	90	50	246
		Effectif théorique	29,0	77,0	90,0	50,0	246,0
		%	11,8%	31,3%	36,6%	20,3%	100,0%

**Annexe 8 – D : scores de frontières après N2 par le nombre de syllabes concaténé du syntagme D1+N1+cc+D2+N2.**

Khi deux de Pearson		Valeur	ddl				p-value
		53,486	24				0,000
TABLEAU DE CONTINGENCE							
			N2				Total
			0	1	2	3	
en G22 en fonction du nombre de syllabes concaténé de : D1+N1+cc+D2+N2+ADJ	1+2+1+1+2+2	Effectif	1	7	7	3	18
		Effectif théorique	2,1	5,6	6,6	3,7	18,0
		%	5,6%	38,9%	38,9%	16,7%	100,0%
		Résidu	-1,1	1,4	,4	-,7	
	1+2+1+1+2+3	Effectif	3	13	5	1	22
		Effectif théorique	2,6	6,9	8,0	4,5	22,0
		%	13,6%	59,1%	22,7%	4,5%	100,0%
		Résidu	,4	6,1	-3,0	-3,5	
	1+2+1+1+2+4	Effectif	3	7	8	1	19
		Effectif théorique	2,2	5,9	7,0	3,9	19,0
%		15,8%	36,8%	42,1%	5,3%	100,0%	
Résidu		,8	1,1	1,0	-2,9		
1+3+1+1+3+2	Effectif	2	5	25	9	41	
	Effectif théorique	4,8	12,8	15,0	8,3	41,0	
	%	4,9%	12,2%	61,0%	22,0%	100,0%	
	Résidu	-2,8	-7,8	10,0	,7		
1+3+1+1+3+3	Effectif	2	15	10	16	43	
	Effectif théorique	5,1	13,5	15,7	8,7	43,0	
	%	4,7%	34,9%	23,3%	37,2%	100,0%	
	Résidu	-3,1	1,5	-5,7	7,3		
1+3+1+1+3+4	Effectif	3	6	6	5	20	
	Effectif théorique	2,4	6,3	7,3	4,1	20,0	
	%	15,0%	30,0%	30,0%	25,0%	100,0%	
	Résidu	,6	-,3	-1,3	,9		
1+4+1+1+4+2	Effectif	6	8	14	8	36	
	Effectif théorique	4,2	11,3	13,2	7,3	36,0	
	%	16,7%	22,2%	38,9%	22,2%	100,0%	
	Résidu	1,8	-3,3	,8	,7		
1+4+1+1+4+3	Effectif	7	4	5	5	21	
	Effectif théorique	2,5	6,6	7,7	4,3	21,0	
	%	33,3%	19,0%	23,8%	23,8%	100,0%	
	Résidu	4,5	-2,6	-2,7	,7		
1+4+1+1+4+4	Effectif	2	12	10	2	26	
	Effectif théorique	3,1	8,1	9,5	5,3	26,0	
	%	7,7%	46,2%	38,5%	7,7%	100,0%	
	Résidu	-1,1	3,9	,5	-3,3		
TOTAL		Effectif	29	77	90	50	246
		Effectif théorique	29,0	77,0	90,0	50,0	246,0
		%	11,8%	31,3%	36,6%	20,3%	100,0%

## Annexe 9 :

### Résultats de l'effet du groupement et de la taille du constituant sur la force des frontières.

Annexe 9 – A : Frontière perçue après N1.

FRONTIERE PERCUE APRES N1			
Effet principal	LR Chisq	Df	Pr(>Chisq)
group_cond	139,13	4	<2,2e-16 ***

Distribution de probabilités des scores de frontières pour chaque groupement					
score de frontière	groupement				
	42	22	32	31	11
0	60,3%	50,6%	26,1%	1,7%	2,1%
1	30,9%	36,9%	44,6%	8,8%	10,6%
2	7,2%	10,2%	22,9%	31,2%	34,1%
3	1,6%	2,3%	6,4%	58,3%	53,3%

Annexe 9 – B : Frontière perçue après N2.

FRONTIERE PERCUE APRES N2			
Effets principaux	LR Chisq	Df	Pr(>Chisq)
group_cond:N_syll	85,196	4	<2,2e-16 ***
N_syll	7,709	2	0,02119 *
Effet d'interaction	LR Chisq	Df	Pr(>Chisq)
group_cond:N_syll	17,943	8	0,02166 *

A) Distribution de probabilités des scores de frontières pour chaque groupement					
score de frontière	groupement				
	11	42	32	31	22
0	37,2%	35,9%	14,5%	13,0%	12,5%
1	41,1%	41,6%	36,5%	34,5%	33,0%
2	17,3%	17,9%	35,0%	36,5%	36,7%
3	4,4%	4,6%	14,0%	16,0%	17,9%
B) Moyennes des scores de frontières calculées à partir des résultats de la distribution de probabilités					
groupement	mean.class	mean.class(-1)	.group		
11	1,89	0,89	a		
42	1,91	0,91			
32	2,48	1,48	b		
31	2,55	1,55			
22	2,60	1,60			

**Distribution de probabilités des scores de frontières pour chaque catégorie de groupement dans chaque taille de N**

groupement	score de frontière	Nombre de syllabes de N		
		2	3	4
G11	0	40,1%	32,3%	39,2%
	1	40,4%	42,3%	40,7%
	2	15,6%	20,0%	16,1%
	3	3,9%	5,4%	4,0%
G42	0	34,3%	36,9%	36,5%
	1	42,1%	41,4%	41,3%
	2	18,7%	17,3%	17,6%
	3	4,9%	4,4%	4,6%
G32	0	15,3%	15,2%	13,0%
	1	37,1%	37,2%	35,3%
	2	34,2%	34,3%	36,6%
	3	13,5%	13,3%	15,1%
G31	0	11,4%	15,4%	12,1%
	1	32,4%	37,3%	33,9%
	2	38,1%	34,1%	37,4%
	3	18,1%	13,2%	16,6%
G22	0	16,6%	6,9%	13,9%
	1	38,8%	24,5%	35,5%
	2	32,7%	41,7%	35,5%
	3	11,8%	26,9%	15,0%

*Annexe 9 – C : Force relative entre la frontière perçue après N1 et la frontière perçue après N2.*

<b>FORCE RELATIVE ENTRE LES DEUX FRONTIÈRES PERÇUES : FRONTIÈRE APRES N2 - FRONTIÈRE APRES N1</b>			
<b>Effets principaux</b>	<b>LR Chisq</b>	<b>Df</b>	<b>Pr(&gt;Chisq)</b>
group_cond	180,942	4	<2,2e-16 ***
N_syll	9,552	2	0,008 **
<b>Effets d'interaction</b>	<b>LR Chisq</b>	<b>Df</b>	<b>Pr(&gt;Chisq)</b>
group_cond:N_syll	16,689	8	0,033 *

<b>A) Distribution de probabilités des forces relatives entre les frontières pour chaque groupement</b>					
<b>score de frontière</b>	<b>groupement</b>				
	<b>11</b>	<b>31</b>	<b>32</b>	<b>42</b>	<b>22</b>
<b>-3</b>	15,2%	5,8%	0,3%	0,3%	0,1%
<b>-2</b>	35,9%	20,5%	1,6%	1,5%	0,6%
<b>-1</b>	35,8%	42,8%	9,3%	8,9%	3,8%
<b>0</b>	12,0%	27,6%	50,3%	50,2%	32,1%
<b>1</b>	0,9%	2,7%	27,1%	27,5%	35,4%
<b>2</b>	0,2%	0,6%	9,0%	9,1%	20,7%
<b>3</b>	0,0%	0,1%	2,4%	2,4%	7,2%
<b>B) Moyennes des forces relatives entre les frontières calculées à partir des résultats de la distribution de probabilités</b>					
<b>group_cond</b>	<b>mean.class</b>	<b>mean.class(-4)</b>	<b>.group</b>		
<b>11</b>	2,48	-1,52	<b>a</b>		
<b>31</b>	3,03	-0,97	<b>b</b>		
<b>32</b>	4,39	0,39	<b>c</b>		
<b>42</b>	4,40	0,40			
<b>22</b>	4,93	0,93	<b>d</b>		

Distribution de probabilités des scores de force relative entre les frontières pour chaque catégorie de groupement dans chaque taille de N				
groupement	score de frontière	Nombre de syllabes de N		
		2	3	4
G11	-3	15,4%	12,4%	17,8%
	-2	36,3%	33,1%	38,3%
	-1	35,6%	38,8%	33,0%
	0	11,6%	14,4%	10,0%
	1	0,9%	1,1%	0,7%
	2	0,2%	0,2%	0,2%
	3	0,0%	0,1%	0,0%
G31	-3	4,2%	7,3%	5,7%
	-2	16,5%	24,4%	20,7%
	-1	41,9%	43,1%	43,3%
	0	33,0%	22,7%	27,1%
	1	3,5%	2,0%	2,6%
	2	0,8%	0,4%	0,5%
	3	0,2%	0,1%	0,1%
G32	-3	0,2%	0,4%	0,4%
	-2	1,2%	1,7%	1,9%
	-1	6,9%	10,0%	10,9%
	0	45,6%	52,0%	53,1%
	1	31,3%	25,7%	24,3%
	2	11,6%	8,1%	7,4%
	3	3,2%	2,1%	1,9%
G42	-3	0,3%	0,3%	0,4%
	-2	1,4%	1,5%	1,7%
	-1	8,5%	8,5%	9,8%
	0	49,4%	49,4%	51,7%
	1	28,3%	28,3%	26,0%
	2	9,6%	9,5%	8,3%
	3	2,5%	2,5%	2,1%
G22	-3	0,2%	0,1%	0,2%
	-2	0,8%	0,3%	0,8%
	-1	4,9%	1,7%	4,8%
	0	39,0%	18,9%	38,4%
	1	35,2%	35,7%	35,5%
	2	15,5%	30,8%	15,8%
	3	4,5%	12,6%	4,6%



## Annexe 10 :

### Résultats de la distribution des scores de prééminences par groupement.

*Annexe 10 – A : scores de prééminences sur la dernière syllabe de N1.*

Khi deux de Pearson		Valeur	ddl	p-value		
		34,480	12	0,001		
<b>TABLEAU DE CONTINGENCE</b>						
		DERNIERE SYLLABE DE N1				Total
		0	1	2	3	
<b>G11</b>	Effectif	124	137	215	152	628
	Effectif théorique	151,8	147,3	207,8	121,1	628,0
	%	19,7%	21,8%	34,2%	24,2%	100,0%
	Résidu	-27,8	-10,3	7,2	30,9	
<b>G22</b>	Effectif	68	56	79	43	246
	Effectif théorique	59,5	57,7	81,4	47,4	246,0
	%	27,6%	22,8%	32,1%	17,5%	100,0%
	Résidu	8,5	-1,7	-2,4	-4,4	
<b>G31</b>	Effectif	19	25	22	16	82
	Effectif théorique	19,8	19,2	27,1	15,8	82,0
	%	23,2%	30,5%	26,8%	19,5%	100,0%
	Résidu	-,8	5,8	-5,1	,2	
<b>G32</b>	Effectif	27	30	42	10	109
	Effectif théorique	26,3	25,6	36,1	21,0	109,0
	%	24,8%	27,5%	38,5%	9,2%	100,0%
	Résidu	,7	4,4	5,9	-11,0	
<b>G42</b>	Effectif	103	83	109	51	346
	Effectif théorique	83,6	81,2	114,5	66,7	346,0
	%	29,8%	24,0%	31,5%	14,7%	100,0%
	Résidu	19,4	1,8	-5,5	-15,7	
<b>TOTAL</b>	Effectif	341	331	467	272	1411
	Effectif théorique	341,0	331,0	467,0	272,0	1411,0
	%	24,2%	23,5%	33,1%	19,3%	100,0%

Annexe 10 – B : scores de prééminences sur la première syllabe de ADJ.

Khi deux de Pearson		Valeur	ddl	p-value		
		36,208	12	0,000		
<b>TABLEAU DE CONTINGENCE</b>						
		1ère SYLLABE DE ADJ				Total
		0	1	2	3	
<b>G11</b>	<b>Effectif</b>	155	195	172	106	628
	<b>Effectif théorique</b>	148,7	175,8	170,5	133,1	628,0
	<b>%</b>	24,7%	31,1%	27,4%	16,9%	100,0%
	<b>Résidu</b>	6,3	19,2	1,5	-27,1	
<b>G22</b>	<b>Effectif</b>	44	52	69	81	246
	<b>Effectif théorique</b>	58,2	68,9	66,8	52,1	246,0
	<b>%</b>	17,9%	21,1%	28,0%	32,9%	100,0%
	<b>Résidu</b>	-14,2	-16,9	2,2	28,9	
<b>G31</b>	<b>Effectif</b>	18	21	25	18	82
	<b>Effectif théorique</b>	19,4	23,0	22,3	17,4	82,0
	<b>%</b>	22,0%	25,6%	30,5%	22,0%	100,0%
	<b>Résidu</b>	-1,4	-2,0	2,7	,6	
<b>G32</b>	<b>Effectif</b>	22	35	29	23	109
	<b>Effectif théorique</b>	25,8	30,5	29,6	23,1	109,0
	<b>%</b>	20,2%	32,1%	26,6%	21,1%	100,0%
	<b>Résidu</b>	-3,8	4,5	-,6	-,1	
<b>G42</b>	<b>Effectif</b>	95	92	88	71	346
	<b>Effectif théorique</b>	81,9	96,9	93,9	73,3	346,0
	<b>%</b>	27,5%	26,6%	25,4%	20,5%	100,0%
	<b>Résidu</b>	13,1	-4,9	-5,9	-2,3	
<b>TOTAL</b>	<b>Effectif</b>	334	395	383	299	1411
	<b>Effectif théorique</b>	334,0	395,0	383,0	299,0	1411,0
	<b>%</b>	23,7%	28,0%	27,1%	21,2%	100,0%

## Annexe 11 :

### Résultats de la distribution des scores de prééminences par taille du constituant.

*Annexe 11 – A : scores de prééminences sur la dernière syllabe de N1 en G11 par le nombre de syllabes de N.*

Khi deux de Pearson		Valeur	ddl		p-value		
		19,09	6		0,004		
<b>TABLEAU DE CONTINGENCE</b>							
			DERNIERE SYLLABE DE N1				TOTAL
			0	1	2	3	
en G11 en fonction du nombre de syllabes de N	2	Effectif	23	43	83	60	209
		Effectif théorique	41,3	45,6	71,6	50,6	209,0
		%	11,0%	20,6%	39,7%	28,7%	100,0%
		Résidu	-18,3	-2,6	11,4	9,4	
	3	Effectif	45	49	63	46	203
		Effectif théorique	40,1	44,3	69,5	49,1	203,0
		%	22,2%	24,1%	31,0%	22,7%	100,0%
		Résidu	4,9	4,7	-6,5	-3,1	
	4	Effectif	56	45	69	46	216
		Effectif théorique	42,6	47,1	73,9	52,3	216,0
		%	25,9%	20,8%	31,9%	21,3%	100,0%
		Résidu	13,4	-2,1	-4,9	-6,3	
<b>TOTAL</b>		Effectif	124	137	215	152	628
		Effectif théorique	124,0	137,0	215,0	152,0	628,0
		%	19,7%	21,8%	34,2%	24,2%	100,0%

**Annexe 11 – B** : scores de proéminences sur la dernière syllabe de N1 en G11 par le nombre de syllabes additionné du syntagme D1+N1+cc+D2+N2+ADJ.

Khi deux de Pearson		Valeur	ddl				p-value
		30,552	18				0,032
TABLEAU DE CONTINGENCE							
			DERNIERE SYLLABE DE N1				TOTAL
			0	1	2	3	
en G11 en fonction du nombre de syllabes additionné de : D1+N1+cc+D2+N2+ADJ	9	Effectif	10	11	32	17	70
		Effectif théorique	13,8	15,3	24,0	16,9	70,0
		%	14,3%	15,7%	45,7%	24,3%	100,0%
		Résidu	-3,8	-4,3	8,0	,1	
	10	Effectif	4	17	28	22	71
		Effectif théorique	14,0	15,5	24,3	17,2	71,0
		%	5,6%	23,9%	39,4%	31,0%	100,0%
		Résidu	-10,0	1,5	3,7	4,8	
	11	Effectif	22	30	44	37	133
		Effectif théorique	26,3	29,0	45,5	32,2	133,0
		%	16,5%	22,6%	33,1%	27,8%	100,0%
		Résidu	-4,3	1,0	-1,5	4,8	
	12	Effectif	20	18	25	12	75
		Effectif théorique	14,8	16,4	25,7	18,2	75,0
		%	26,7%	24,0%	33,3%	16,0%	100,0%
		Résidu	5,2	1,6	-,7	-6,2	
	13	Effectif	32	30	34	36	132
		Effectif théorique	26,1	28,8	45,2	31,9	132,0
		%	24,2%	22,7%	25,8%	27,3%	100,0%
		Résidu	5,9	1,2	-11,2	4,1	
	14	Effectif	19	18	26	10	73
Effectif théorique		14,4	15,9	25,0	17,7	73,0	
%		26,0%	24,7%	35,6%	13,7%	100,0%	
Résidu		4,6	2,1	1,0	-7,7		
15	Effectif	17	13	26	18	74	
	Effectif théorique	14,6	16,1	25,3	17,9	74,0	
	%	23,0%	17,6%	35,1%	24,3%	100,0%	
	Résidu	2,4	-3,1	,7	,1		
TOTAL		Effectif	124	137	215	152	628
		Effectif théorique	124,0	137,0	215,0	152,0	628,0
		%	19,7%	21,8%	34,2%	24,2%	100,0%

**Annexe 11 – C : scores de prééminences sur la première syllabe de ADJ en G22 par le nombre de syllabes additionné du syntagme D1+N1+cc+D2+N2+ADJ.**

Khi deux de Pearson		Valeur	ddl		p-value		
		37,922	18		0,004		
TABLEAU DE CONTINGENCE							
			1ERE SYLLABE DE ADJ				TOTAL
			0	1	2	3	
en G22 en fonction du nombre de syllabes additionné de : D1+N1+cc+D2+N2+ADJ	9	Effectif	2	8	6	2	18
		Effectif théorique	3,2	3,8	5,0	5,9	18,0
		%	11,1%	44,4%	33,3%	11,1%	100,0%
		Résidu	-1,2	4,2	1,0	-3,9	
	10	Effectif	5	9	6	2	22
		Effectif théorique	3,9	4,7	6,2	7,2	22,0
		%	22,7%	40,9%	27,3%	9,1%	100,0%
		Résidu	1,1	4,3	-,2	-5,2	
	11	Effectif	13	8	20	19	60
		Effectif théorique	10,7	12,7	16,8	19,8	60,0
		%	21,7%	13,3%	33,3%	31,7%	100,0%
		Résidu	2,3	-4,7	3,2	-,8	
	12	Effectif	7	8	9	19	43
		Effectif théorique	7,7	9,1	12,1	14,2	43,0
		%	16,3%	18,6%	20,9%	44,2%	100,0%
		Résidu	-,7	-1,1	-3,1	4,8	
	13	Effectif	6	10	17	23	56
		Effectif théorique	10,0	11,8	15,7	18,4	56,0
		%	10,7%	17,9%	30,4%	41,1%	100,0%
		Résidu	-4,0	-1,8	1,3	4,6	
	14	Effectif	1	3	6	11	21
		Effectif théorique	3,8	4,4	5,9	6,9	21,0
		%	4,8%	14,3%	28,6%	52,4%	100,0%
		Résidu	-2,8	-1,4	,1	4,1	
15	Effectif	10	6	5	5	26	
	Effectif théorique	4,7	5,5	7,3	8,6	26,0	
	%	38,5%	23,1%	19,2%	19,2%	100,0%	
	Résidu	5,3	,5	-2,3	-3,6		
TOTAL		Effectif	44	52	69	81	246
		Effectif théorique	44,0	52,0	69,0	81,0	246,0
		%	17,9%	21,1%	28,0%	32,9%	100,0%

**Annexe 11 – D : scores de prééminences sur la première syllabe de ADJ en G22 par le nombre de syllabes concaténé du syntagme D1+N1+cc+D2+N2+ADJ.**

Khi deux de Pearson		Valeur	ddl		p-value		
		39,813	24		0,022		
TABLEAU DE CONTINGENCE							
			1ERE SYLLABE DE ADJ				Total
			0	1	2	3	
en G22 en fonction du nombre de syllabes concaténé de : D1+N1+cc+D2+N2+ADJ	1+2+1+1+2+2	Effectif	2	8	6	2	18
		Effectif théorique	3,2	3,8	5,0	5,9	18,0
		%	11,1%	44,4%	33,3%	11,1%	100,0%
		Résidu	-1,2	4,2	1,0	-3,9	
	1+2+1+1+2+3	Effectif	5	9	6	2	22
		Effectif théorique	3,9	4,7	6,2	7,2	22,0
		%	22,7%	40,9%	27,3%	9,1%	100,0%
		Résidu	1,1	4,3	-,2	-5,2	
	1+2+1+1+2+4	Effectif	5	2	5	7	19
		Effectif théorique	3,4	4,0	5,3	6,3	19,0
%		26,3%	10,5%	26,3%	36,8%	100,0%	
Résidu		1,6	-2,0	-,3	,7		
1+3+1+1+3+2	Effectif	8	6	15	12	41	
	Effectif théorique	7,3	8,7	11,5	13,5	41,0	
	%	19,5%	14,6%	36,6%	29,3%	100,0%	
	Résidu	,7	-2,7	3,5	-1,5		
1+3+1+1+3+3	Effectif	7	8	9	19	43	
	Effectif théorique	7,7	9,1	12,1	14,2	43,0	
	%	16,3%	18,6%	20,9%	44,2%	100,0%	
	Résidu	-,7	-1,1	-3,1	4,8		
1+3+1+1+3+4	Effectif	3	4	6	7	20	
	Effectif théorique	3,6	4,2	5,6	6,6	20,0	
	%	15,0%	20,0%	30,0%	35,0%	100,0%	
	Résidu	-,6	-,2	,4	,4		
1+4+1+1+4+2	Effectif	3	6	11	16	36	
	Effectif théorique	6,4	7,6	10,1	11,9	36,0	
	%	8,3%	16,7%	30,6%	44,4%	100,0%	
	Résidu	-3,4	-1,6	,9	4,1		
1+4+1+1+4+3	Effectif	1	3	6	11	21	
	Effectif théorique	3,8	4,4	5,9	6,9	21,0	
	%	4,8%	14,3%	28,6%	52,4%	100,0%	
	Résidu	-2,8	-1,4	,1	4,1		
1+4+1+1+4+4	Effectif	10	6	5	5	26	
	Effectif théorique	4,7	5,5	7,3	8,6	26,0	
	%	38,5%	23,1%	19,2%	19,2%	100,0%	
	Résidu	5,3	,5	-2,3	-3,6		
<b>TOTAL</b>		Effectif	44	52	69	81	246
		Effectif théorique	44,0	52,0	69,0	81,0	246,0
		%	17,9%	21,1%	28,0%	32,9%	100,0%

**Annexe 11 – E** : scores de prééminences sur la première syllabe de ADJ en G32 par le nombre de syllabes additionné du syntagme D1+N1+cc+D2+N2+ADJ.

Khi deux de Pearson		Valeur	ddl		p-value		
		29,45	18		0,043		
TABLEAU DE CONTINGENCE							
			1ERE SYLLABE DE ADJ				TOTAL
			0	1	2	3	
en G32 en fonction du nombre de syllabes additionné de : D1+N1+cc+D2+N2+ADJ	9	Effectif	7	6	1	1	15
		Effectif théorique	3,0	4,8	4,0	3,2	15,0
		%	46,7%	40,0%	6,7%	6,7%	100,0%
		Résidu	4,0	1,2	-3,0	-2,2	
	10	Effectif	1	3	1	3	8
		Effectif théorique	1,6	2,6	2,1	1,7	8,0
		%	12,5%	37,5%	12,5%	37,5%	100,0%
		Résidu	-,6	,4	-1,1	1,3	
	11	Effectif	5	6	10	3	24
		Effectif théorique	4,8	7,7	6,4	5,1	24,0
		%	20,8%	25,0%	41,7%	12,5%	100,0%
		Résidu	,2	-1,7	3,6	-2,1	
	12	Effectif	1	4	4	1	10
		Effectif théorique	2,0	3,2	2,7	2,1	10,0
		%	10,0%	40,0%	40,0%	10,0%	100,0%
		Résidu	-1,0	,8	1,3	-1,1	
	13	Effectif	4	10	1	6	21
		Effectif théorique	4,2	6,7	5,6	4,4	21,0
		%	19,0%	47,6%	4,8%	28,6%	100,0%
		Résidu	-,2	3,3	-4,6	1,6	
	14	Effectif	0	2	5	4	11
Effectif théorique		2,2	3,5	2,9	2,3	11,0	
%		0,0%	18,2%	45,5%	36,4%	100,0%	
Résidu		-2,2	-1,5	2,1	1,7		
15	Effectif	4	4	7	5	20	
	Effectif théorique	4,0	6,4	5,3	4,2	20,0	
	%	20,0%	20,0%	35,0%	25,0%	100,0%	
	Résidu	,0	-2,4	1,7	,8		
TOTAL		Effectif	22	35	29	23	109
		Effectif théorique	22,0	35,0	29,0	23,0	109,0
		%	20,2%	32,1%	26,6%	21,1%	100,0%

**Annexe 11 – F : scores de prééminences sur la première syllabe de ADJ en G32 par le nombre de syllabes concaténé du syntagme D1+N1+cc+D2+N2+ADJ.**

Khi deux de Pearson		Valeur	ddl		p-value		
		37,419	18		0,040		
TABLEAU DE CONTINGENCE							
			1ERE SYLLABE DE ADJ				Total
			0	1	2	3	
en G32 en fonction du nombre de syllabes concaténé de : D1+N1+cc+D2+N2+ADJ	1+2+1+1+2+2	Effectif	7	6	1	1	15
		Effectif théorique	3,0	4,8	4,0	3,2	15,0
		%	46,7%	40,0%	6,7%	6,7%	100,0%
		Résidu	4,0	1,2	-3,0	-2,2	
	1+2+1+1+2+3	Effectif	1	3	1	3	8
		Effectif théorique	1,6	2,6	2,1	1,7	8,0
		%	12,5%	37,5%	12,5%	37,5%	100,0%
		Résidu	-,6	,4	-1,1	1,3	
	1+2+1+1+2+4	Effectif	3	3	9	2	17
		Effectif théorique	3,4	5,5	4,5	3,6	17,0
%		17,6%	17,6%	52,9%	11,8%	100,0%	
Résidu		-,4	-2,5	4,5	-1,6		
1+3+1+1+3+2	Effectif	2	3	1	1	7	
	Effectif théorique	1,4	2,2	1,9	1,5	7,0	
	%	28,6%	42,9%	14,3%	14,3%	100,0%	
	Résidu	,6	,8	-,9	-,5		
1+3+1+1+3+3	Effectif	1	4	4	1	10	
	Effectif théorique	2,0	3,2	2,7	2,1	10,0	
	%	10,0%	40,0%	40,0%	10,0%	100,0%	
	Résidu	-1,0	,8	1,3	-1,1		
1+3+1+1+3+4	Effectif	3	4	0	1	8	
	Effectif théorique	1,6	2,6	2,1	1,7	8,0	
	%	,4	,5	,0	,1	1,0	
	Résidu	1,4	1,4	-2,1	-,7		
1+4+1+1+4+2	Effectif	1	6	1	5	13	
	Effectif théorique	2,6	4,2	3,5	2,7	13,0	
	%	7,7%	46,2%	7,7%	38,5%	100,0%	
	Résidu	-1,6	1,8	-2,5	2,3		
1+4+1+1+4+3	Effectif	0	2	5	4	11	
	Effectif théorique	2,2	3,5	2,9	2,3	11,0	
	%	0,0%	18,2%	45,5%	36,4%	100,0%	
	Résidu	-2,2	-1,5	2,1	1,7		
1+4+1+1+4+4	Effectif	4	4	7	5	20	
	Effectif théorique	4,0	6,4	5,3	4,2	20,0	
	%	20,0%	20,0%	35,0%	25,0%	100,0%	
	Résidu	,0	-2,4	1,7	,8		
<b>TOTAL</b>		Effectif	22	35	29	23	109
		Effectif théorique	22,0	35,0	29,0	23,0	109,0
		%	20,2%	32,1%	26,6%	21,1%	100,0%

## Annexe 12 :

### Résultats de la comparaison des degrés de proéminence perçus sur les syllabes initiales et finales de N1, N2 et ADJ.

*Annexe 12 – A : AI sur N1 vs. AF sur N1 dans chaque groupement.*

		Test de Stuart-Maxwell		SCORE DE PROEMINENCE				TOTAL
		Chisq	p-value	0	1	2	3	
G11	AI N1	22,9	< 0,001	26,6%	26,4%	30,3%	16,7%	100,0%
	AF N1			19,7%	21,8%	34,2%	24,2%	100,0%
G31	AI N1	2,02	0,568	26,8%	20,7%	31,7%	20,7%	100,0%
	AF N1			23,2%	30,5%	26,8%	19,5%	100,0%
G22	AI N1	2,8	0,423	26,4%	28,5%	26,4%	18,7%	100,0%
	AF N1			27,6%	22,8%	32,1%	17,5%	100,0%
G32	AI N1	5,73	0,056	24,8%	20,2%	37,6%	17,4%	100,0%
	AF N1			24,8%	27,5%	38,5%	9,2%	100,0%
G42	AI N1	3,59	0,309	24,3%	26,3%	32,9%	16,5%	100,0%
	AF N1			29,8%	24,0%	31,5%	14,7%	100,0%

*Annexe 12 – B : AI sur N2 vs. AF sur N2 dans chaque groupement.*

		Test de Stuart-Maxwell		SCORE DE PROEMINENCE				TOTAL
		Chisq	p-value	0	1	2	3	
G11	AI N2	91,8	0	18,3%	31,4%	29,6%	20,7%	100,0%
	AF N2			34,1%	26,1%	30,9%	8,9%	100,0%
G31	AI N2	12,2	0,006	23,2%	22,0%	28,0%	26,8%	100,0%
	AF N2			35,4%	26,8%	26,8%	11,0%	100,0%
G22	AI N2	21,8	< 0,001	26,0%	28,9%	24,4%	20,7%	100,0%
	AF N2			34,6%	30,5%	26,8%	8,1%	100,0%
G32	AI N2	24,9	< 0,001	15,6%	24,8%	39,4%	20,2%	100,0%
	AF N2			34,9%	33,0%	25,7%	6,4%	100,0%
G42	AI N2	70,2	< 0,001	17,1%	32,9%	30,1%	19,9%	100,0%
	AF N2			38,2%	28,0%	24,9%	9,0%	100,0%

*Annexe 12 – D : AI sur ADJ vs. AF sur ADJ dans chaque groupement.*

		Test de Stuart-Maxwell		SCORE DE PROEMINENCE				TOTAL
		Chisq	p-value	0	1	2	3	
G11	AI ADJ	20,8	< 0,001	24,7%	31,1%	27,4%	16,9%	100,0%
	AF ADJ			22,8%	22,9%	37,3%	17,0%	100,0%
G31	AI ADJ	1,02	0,796	22,0%	25,6%	30,5%	22,0%	100,0%
	AF ADJ			20,7%	30,5%	31,7%	17,1%	100,0%
G22	AI ADJ	20,3	< 0,001	17,9%	21,1%	28,0%	32,9%	100,0%
	AF ADJ			24,4%	22,0%	36,6%	17,1%	100,0%
G32	AI ADJ	5,58	0,134	20,2%	32,1%	26,6%	21,1%	100,0%
	AF ADJ			25,7%	25,7%	36,7%	11,9%	100,0%
G42	AI ADJ	2,98	0,225	27,5%	26,6%	25,4%	20,5%	100,0%
	AF ADJ			28,3%	26,6%	28,3%	16,8%	100,0%

*Annexe 12 – G : AFsur N2 vs. AI sur ADJ dans chaque groupement.*

		Test de Stuart-Maxwell		SCORE DE PROEMINENCE				TOTAL
		Chisq	p-value	0	1	2	3	
G11	AF N2	12,6	0,002	34,1%	26,1%	30,9%	8,9%	100,0%
	AI ADJ			24,7%	31,1%	27,4%	16,9%	100,0%
G31	AF N2	12,9	0,004	35,4%	26,8%	26,8%	11,0%	100,0%
	AI ADJ			22,0%	25,6%	30,5%	22,0%	100,0%
G22	AF N2	62,7	< 0,001	34,6%	30,5%	26,8%	8,1%	100,0%
	AI ADJ			17,9%	21,1%	28,0%	32,9%	100,0%
G32	AF N2	16	0,001	34,9%	33,0%	25,7%	6,4%	100,0%
	AI ADJ			20,2%	32,1%	26,6%	21,1%	100,0%
G42	AF N2	36,4	< 0,001	38,2%	28,0%	24,9%	9,0%	100,0%
	AI ADJ			27,5%	26,6%	25,4%	20,5%	100,0%

*Annexe 12 – F : AFsur N1 vs. AI sur N2 dans chaque groupement.*

		Test de Stuart-Maxwell		SCORE DE PROEMINENCE				TOTAL
		Chisq	p-value	0	1	2	3	
G11	AF N1	15,8	0,001	19,7%	21,8%	34,2%	24,2%	100,0%
	AI N2			18,3%	31,4%	29,6%	20,7%	100,0%
G31	AF N1	5,92	0,051	23,2%	30,5%	26,8%	19,5%	100,0%
	AI N2			23,2%	22,0%	28,0%	26,8%	100,0%
G22	AF N1	4,9	0,179	27,6%	22,8%	32,1%	17,5%	100,0%
	AI N2			26,0%	28,9%	24,4%	20,7%	100,0%
G32	AF N1	9,96	0,018	24,8%	27,5%	38,5%	9,2%	100,0%
	AI N2			15,6%	24,8%	39,4%	20,2%	100,0%
G42	AF N1	25,2	< 0,001	29,8%	24,0%	31,5%	14,7%	100,0%
	AI N2			17,1%	32,9%	30,1%	19,9%	100,0%

### Annexe 13 :

## Résultats de la comparaison des degrés de proéminence perçus sur la première, deuxième et troisième syllabe de N1, N2 et ADJ.

*Annexe 13 – A : 1<sup>ère</sup> syllabe sur N1 vs. 2<sup>ème</sup> syllabe sur N1 ; 1<sup>ère</sup> syllabe sur N1 vs. 3<sup>ème</sup> syllabe sur N1 dans chaque groupement.*

			Test de Stuart-Maxwell		SCORE DE PROEMINENCE				TOTAL
			Chisq	p-value	0	1	2	3	
G11	N1	2 <sup>ème</sup>	43,4	< 0,001	35,3%	27,4%	28,9%	8,4%	100,0%
		première			26,6%	26,4%	30,3%	16,7%	100,0%
		3 <sup>ème</sup>	36,6	< 0,001	40,7%	25,9%	27,8%	5,6%	100,0%
G31	N1	2 <sup>ème</sup>	4,46	0,108	45,3%	22,6%	20,8%	11,3%	100,0%
		première	1	0,607	26,8%	20,7%	31,7%	20,7%	100,0%
		3 <sup>ème</sup>			45,0%	25,0%	25,0%	5,0%	100,0%
G22	N1	2 <sup>ème</sup>	35,3	< 0,001	44,4%	20,9%	26,2%	8,6%	100,0%
		première			26,4%	28,5%	26,4%	18,7%	100,0%
		3 <sup>ème</sup>	14,2	0,002	39,8%	24,1%	28,9%	7,2%	100,0%
G32	N1	2 <sup>ème</sup>	20,5	< 0,001	52,2%	15,9%	21,7%	10,1%	100,0%
		première			24,8%	20,2%	37,6%	17,4%	100,0%
		3 <sup>ème</sup>	14,3	0,002	40,9%	29,5%	15,9%	13,6%	100,0%
G42	N1	2 <sup>ème</sup>	31,5	< 0,001	39,7%	23,4%	26,8%	10,0%	100,0%
		première			24,3%	26,3%	32,9%	16,5%	100,0%
		3 <sup>ème</sup>	26,5	< 0,001	44,1%	23,5%	29,4%	2,9%	100,0%

**Annexe 13 – B : 1<sup>ère</sup> syllabe sur N2 vs. 2<sup>ème</sup> syllabe sur N2 ; 1<sup>ère</sup> syllabe sur N2 vs. 3<sup>ème</sup> syllabe sur N2 dans chaque groupement.**

			Test de Stuart-Maxwell		SCORE DE PROEMINENCE				TOTAL
			Chisq	p-value	0	1	2	3	
G11	N2	2 <sup>ème</sup>	96,7	0,000	38,4%	22,7%	28,9%	10,0%	100,0%
		première			18,3%	31,4%	29,6%	20,7%	100,0%
		3 <sup>ème</sup>	53,5	< 0,001	40,7%	22,2%	28,7%	8,3%	100,0%
G31	N2	2 <sup>ème</sup>	3	0,083	45,3%	22,6%	24,5%	7,5%	100,0%
		première			23,2%	22,0%	28,0%	26,8%	100,0%
		3 <sup>ème</sup>	7,55	0,056	45,0%	20,0%	30,0%	5,0%	100,0%
G22	N2	2 <sup>ème</sup>	37,9	< 0,001	46,5%	20,3%	25,7%	7,5%	100,0%
		première			26,0%	28,9%	24,4%	20,7%	100,0%
		3 <sup>ème</sup>	9,82	0,007	43,4%	20,5%	25,3%	10,8%	100,0%
G32	N2	2 <sup>ème</sup>	25,1	< 0,001	50,7%	18,8%	18,8%	11,6%	100,0%
		première			15,6%	24,8%	39,4%	20,2%	100,0%
		3 <sup>ème</sup>	14,9	0,002	36,4%	31,8%	25,0%	6,8%	100,0%
G42	N2	2 <sup>ème</sup>	68,5	< 0,001	39,7%	29,7%	25,4%	5,3%	100,0%
		première			17,1%	32,9%	30,1%	19,9%	100,0%
		3 <sup>ème</sup>	27,2	< 0,001	40,2%	26,5%	26,5%	6,9%	100,0%

**Annexe 13 – C : 1<sup>ère</sup> syllabe sur ADJ vs. 2<sup>ème</sup> syllabe sur ADJ ; 1<sup>ère</sup> syllabe sur ADJ vs. 3<sup>ème</sup> syllabe sur ADJ dans chaque groupement.**

			Test de Stuart-Maxwell		SCORE DE PROEMINENCE				TOTAL
			Chisq	p-value	0	1	2	3	
G11	ADJ	2 <sup>ème</sup>	35,4	< 0,001	34,7%	26,4%	31,6%	7,3%	100,0%
		première			24,7%	31,1%	27,4%	16,9%	100,0%
		3 <sup>ème</sup>	20,4	0,0001	34,6%	29,8%	28,3%	7,3%	100,0%
G31	ADJ	2 <sup>ème</sup>	9,8	0,020	40,4%	12,8%	36,2%	10,6%	100,0%
		première			22,0%	25,6%	30,5%	22,0%	100,0%
		3 <sup>ème</sup>	10,1	0,017	46,9%	21,9%	18,8%	12,5%	100,0%
G22	ADJ	2 <sup>ème</sup>	36,9	< 0,001	41,1%	23,2%	23,8%	11,9%	100,0%
		première			17,9%	21,1%	28,0%	32,9%	100,0%
		3 <sup>ème</sup>	13,7	0,003	41,5%	24,6%	23,1%	10,8%	100,0%
G32	ADJ	2 <sup>ème</sup>	17,2	< 0,001	41,9%	18,9%	31,1%	8,1%	100,0%
		première			20,2%	32,1%	26,6%	21,1%	100,0%
		3 <sup>ème</sup>	9	0,029	44,4%	26,7%	22,2%	6,7%	100,0%
G42	ADJ	2 <sup>ème</sup>	29,5	< 0,001	34,6%	26,8%	31,3%	7,3%	100,0%
		première			27,5%	26,6%	25,4%	20,5%	100,0%
		3 <sup>ème</sup>	18,4	< 0,001	38,1%	31,0%	21,4%	9,5%	100,0%

## Annexe 14 :

### Résultats de la comparaison des degrés de proéminence perçus sur la deuxième, troisième et dernière syllabe de N1, N2 et ADJ.

*Annexe 14 – A : dernière syllabe sur N1 vs. 2<sup>ème</sup> syllabe sur N1 ; dernière syllabe sur N1 vs. 3<sup>ème</sup> syllabe sur N1 dans chaque groupement.*

			Test de Stuart-Maxwell		SCORE DE PROEMINENCE				TOTAL
			Chisq	p-value	0	1	2	3	
G11	N1	2 <sup>ème</sup>	46,9	< 0,001	35,3%	27,4%	28,9%	8,4%	100,0%
		dernière			19,7%	21,8%	34,2%	24,2%	100,0%
		3 <sup>ème</sup>	40,9	< 0,001	40,7%	25,9%	27,8%	5,6%	100,0%
G31	N1	2 <sup>ème</sup>	9,63	0,022	45,3%	22,6%	20,8%	11,3%	100,0%
		dernière	7,37	0,061	23,2%	30,5%	26,8%	19,5%	100,0%
		3 <sup>ème</sup>			45,0%	25,0%	25,0%	5,0%	100,0%
G22	N1	2 <sup>ème</sup>	18,1	< 0,001	44,4%	20,9%	26,2%	8,6%	100,0%
		dernière			27,6%	22,8%	32,1%	17,5%	100,0%
		3 <sup>ème</sup>	10,1	0,006	39,8%	24,1%	28,9%	7,2%	100,0%
G32	N1	2 <sup>ème</sup>	15,3	< 0,001	52,2%	15,9%	21,7%	10,1%	100,0%
		dernière			24,8%	27,5%	38,5%	9,2%	100,0%
		3 <sup>ème</sup>	9,91	0,007	40,9%	29,5%	15,9%	13,6%	100,0%
G42	N1	2 <sup>ème</sup>	5,07	0,167	39,7%	23,4%	26,8%	10,0%	100,0%
		dernière			20,3%	24,0%	31,5%	24,2%	100,0%
		3 <sup>ème</sup>	9,89	0,019	44,1%	23,5%	29,4%	2,9%	100,0%

**Annexe 14 – B : dernière syllabe sur N2 vs. 2<sup>ème</sup> syllabe sur N2 ; dernière syllabe sur N2 vs. 3<sup>ème</sup> syllabe sur N2 dans chaque groupement.**

			Test de Stuart-Maxwell		SCORE DE PROEMINENCE				TOTAL
			Chisq	p-value	0	1	2	3	
G11	N2	2 <sup>ème</sup>	4,79	0,091	38,4%	22,7%	28,9%	10,0%	100,0%
		dernière			34,1%	26,1%	30,9%	8,9%	100,0%
		3 <sup>ème</sup>	8,94	0,05	40,7%	22,2%	28,7%	8,3%	100,0%
G31	N2	2 <sup>ème</sup>	4,94	0,177	45,3%	22,6%	24,5%	7,5%	100,0%
		dernière			35,4%	26,8%	26,8%	11,0%	100,0%
		3 <sup>ème</sup>	1,78	0,62	45,0%	20,0%	30,0%	5,0%	100,0%
G22	N2	2 <sup>ème</sup>	7,3	0,06	46,5%	20,3%	25,7%	7,5%	100,0%
		dernière			34,6%	30,5%	26,8%	8,1%	100,0%
		3 <sup>ème</sup>	5,46	0,141	43,4%	20,5%	25,3%	10,8%	100,0%
G32	N2	2 <sup>ème</sup>	7,21	0,065	50,7%	18,8%	18,8%	11,6%	100,0%
		dernière			34,9%	33,0%	25,7%	6,4%	100,0%
		3 <sup>ème</sup>	3,2	0,361	36,4%	31,8%	25,0%	6,8%	100,0%
G42	N2	2 <sup>ème</sup>	2,56	0,464	39,7%	29,7%	25,4%	5,3%	100,0%
		dernière			38,2%	28,0%	24,9%	9,0%	100,0%
		3 <sup>ème</sup>	2,38	0,497	40,2%	26,5%	26,5%	6,9%	100,0%

**Annexe 14 – C : dernière syllabe sur ADJ vs. 2<sup>ème</sup> syllabe sur ADJ ; dernière syllabe sur ADJ vs. 3<sup>ème</sup> syllabe sur ADJ dans chaque groupement.**

			Test de Stuart-Maxwell		SCORE DE PROEMINENCE				TOTAL
			Chisq	p-value	0	1	2	3	
G11	ADJ	2 <sup>ème</sup>	48,7	< 0,001	34,7%	26,4%	31,6%	7,3%	100,0%
		dernière			22,8%	22,9%	37,3%	17,0%	100,0%
		3 <sup>ème</sup>	38,2	< 0,001	34,6%	29,8%	28,3%	7,3%	100,0%
G31	ADJ	2 <sup>ème</sup>	8,18	0,042	40,4%	12,8%	36,2%	10,6%	100,0%
		dernière			20,7%	30,5%	31,7%	17,1%	100,0%
		3 <sup>ème</sup>	8,94	0,03	46,9%	21,9%	18,8%	12,5%	100,0%
G22	ADJ	2 <sup>ème</sup>	14,8	0,002	41,1%	23,2%	23,8%	11,9%	100,0%
		dernière			24,4%	22,0%	36,6%	17,1%	100,0%
		3 <sup>ème</sup>	4,48	0,106	41,5%	24,6%	23,1%	10,8%	100,0%
G32	ADJ	2 <sup>ème</sup>	7,32	0,062	41,9%	18,9%	31,1%	8,1%	100,0%
		dernière			25,7%	25,7%	36,7%	11,9%	100,0%
		3 <sup>ème</sup>	9,88	0,019	44,4%	26,7%	22,2%	6,7%	100,0%
G42	ADJ	2 <sup>ème</sup>	15,4	0,001	34,6%	26,8%	31,3%	7,3%	100,0%
		dernière			28,3%	26,6%	28,3%	16,8%	100,0%
		3 <sup>ème</sup>	7,04	0,07	38,1%	31,0%	21,4%	9,5%	100,0%

## Annexe 15 :

**Résultats de la comparaison des degrés de proéminence perçus sur les syllabes intermédiaires de N1, N2 et ADJ, et sur les syllabes des mots grammaticaux.**

*Annexe 15 – A : 2<sup>ème</sup> syllabe sur N1 vs. 3<sup>ème</sup> syllabe sur N1 ; 2<sup>ème</sup> syllabe sur N2 vs. 3<sup>ème</sup> syllabe sur N2 ; 2<sup>ème</sup> syllabe sur ADJ vs. 3<sup>ème</sup> syllabe sur ADJ.*

		Test de Stuart-Maxwell		SCORE DE PROEMINENCE				TOTAL
		Chisq	p-value	0	1	2	3	
N1	2 <sup>ème</sup>	9,91	0,007	39,9%	24,1%	26,9%	9,1%	100,0%
	3 <sup>ème</sup>			41,5%	25,4%	27,1%	6,0%	100,0%
N2	2 <sup>ème</sup>	3,3	0,348	41,6%	23,5%	26,5%	8,4%	100,0%
	3 <sup>ème</sup>			40,9%	23,7%	27,3%	8,2%	100,0%
ADJ	2 <sup>ème</sup>	13,3	0,004	36,5%	24,7%	30,5%	8,3%	100,0%
	3 <sup>ème</sup>			38,3%	28,5%	24,5%	8,7%	100,0%

*Annexe 15 – B : D1 vs. cc ; D1 vs. D2 ; cc vs. D2.*

		Test de Stuart-Maxwell		SCORE DE PROEMINENCE				TOTAL
		Chisq	p-value	0	1	2	3	
D1	94	< 0,001	68,7%	25,8%	5,0%	0,4%	100,0%	
cc			61,3%	26,1%	10,1%	2,6%	100,0%	
D1	29,2	< 0,001	68,7%	25,8%	5,0%	0,4%	100,0%	
D2			65,5%	25,7%	8,4%	0,5%	100,0%	
cc	54,4	< 0,001	61,3%	26,1%	10,1%	2,6%	100,0%	
D2			65,5%	25,7%	8,4%	0,5%	100,0%	

**Annexe 15 – C :** D1 vs. 2<sup>ème</sup> syllabe sur N1 ; D1 vs. 3<sup>ème</sup> syllabe sur N1 ; D1 vs. 2<sup>ème</sup> syllabe sur N2 ; D1 vs. 3<sup>ème</sup> syllabe sur N2 ; D1 vs. 2<sup>ème</sup> syllabe sur ADJ ; D1 vs. 3<sup>ème</sup> syllabe sur ADJ ; cc vs. 2<sup>ème</sup> syllabe sur N1 ; cc vs. 3<sup>ème</sup> syllabe sur N1 ; cc vs. 2<sup>ème</sup> syllabe sur N2 ; cc vs. 3<sup>ème</sup> syllabe sur N2 ; cc vs. 2<sup>ème</sup> syllabe sur ADJ ; cc vs. 3<sup>ème</sup> syllabe sur ADJ ; D2 vs. 2<sup>ème</sup> syllabe sur N1 ; D2 vs. 3<sup>ème</sup> syllabe sur N1 ; D2 vs. 2<sup>ème</sup> syllabe sur N2 ; D2 vs. 3<sup>ème</sup> syllabe sur N2 ; D2 vs. 2<sup>ème</sup> syllabe sur ADJ ; D2 vs. 3<sup>ème</sup> syllabe sur ADJ.

		Test de Stuart-Maxwell		SCORE DE PROEMINENCE				TOTAL
		Chisq	p-value	0	1	2	3	
N1	2 <sup>ème</sup>	362	< 0,001	39,9%	24,1%	26,9%	9,1%	100,0%
D1				68,7%	25,8%	5,0%	0,4%	100,0%
N1	3 <sup>ème</sup>	166	< 0,001	41,5%	25,4%	27,1%	6,0%	100,0%
N2	2 <sup>ème</sup>	344	< 0,001	41,6%	23,5%	26,5%	8,4%	100,0%
D1				68,7%	25,8%	5,0%	0,4%	100,0%
N2	3 <sup>ème</sup>	179	< 0,001	40,9%	23,7%	27,3%	8,2%	100,0%
ADJ	2 <sup>ème</sup>	402	< 0,001	36,5%	24,7%	30,5%	8,3%	100,0%
D1				68,7%	25,8%	5,0%	0,4%	100,0%
ADJ	3 <sup>ème</sup>	185	< 0,001	38,3%	28,5%	24,5%	8,7%	100,0%
N1	2 <sup>ème</sup>	220	< 0,001	39,9%	24,1%	26,9%	9,1%	100,0%
cc				61,3%	26,1%	10,1%	2,6%	100,0%
N1	3 <sup>ème</sup>	81	< 0,001	41,5%	25,4%	27,1%	6,0%	100,0%
N2	2 <sup>ème</sup>	209	< 0,001	41,6%	23,5%	26,5%	8,4%	100,0%
cc				61,3%	26,1%	10,1%	2,6%	100,0%
N2	3 <sup>ème</sup>	96,2	< 0,001	40,9%	23,7%	27,3%	8,2%	100,0%
ADJ	2 <sup>ème</sup>	268	< 0,001	36,5%	24,7%	30,5%	8,3%	100,0%
cc				61,3%	26,1%	10,1%	2,6%	100,0%
ADJ	3 <sup>ème</sup>	120	< 0,001	38,3%	28,5%	24,5%	8,7%	100,0%
N1	2 <sup>ème</sup>	314	< 0,001	39,9%	24,1%	26,9%	9,1%	100,0%
D2				65,5%	25,7%	8,4%	0,5%	100,0%
N1	3 <sup>ème</sup>	135	< 0,001	41,5%	25,4%	27,1%	6,0%	100,0%
N2	2 <sup>ème</sup>	304	< 0,001	41,6%	23,5%	26,5%	8,4%	100,0%
D2				65,5%	25,7%	8,4%	0,5%	100,0%
N2	3 <sup>ème</sup>	150	< 0,001	40,9%	23,7%	27,3%	8,2%	100,0%
ADJ	2 <sup>ème</sup>	347	< 0,001	36,5%	24,7%	30,5%	8,3%	100,0%
D2				65,5%	25,7%	8,4%	0,5%	100,0%
ADJ	3 <sup>ème</sup>	160	< 0,001	38,3%	28,5%	24,5%	8,7%	100,0%

## Annexe 16 :

### Résultats de l'effet du groupement et de la taille du constituant sur la force des proéminences.

*Annexe 16 – A : Proéminence perçue sur la dernière syllabe de N1.*

PROÉMINENCE PERÇUE SUR LA SYLLABE FINALE DE N1			
Effets principaux	LR Chisq	Df	Pr(>Chisq)
group_cond	28,64	4	9,251e-06 ***
N_syll	19,82	2	4,972e-05 ***

Distribution de probabilités des scores de proéminences pour chaque groupement					
score de proéminence	group_cond				
	32	42	22	31	11
0	30,8%	29,8%	25,6%	24,7%	19,5%
1	25,1%	25,3%	24,2%	24,0%	21,7%
2	29,6%	30,1%	32,5%	33,0%	35,4%
3	14,5%	14,7%	17,7%	18,3%	23,4%

Distribution de probabilités des scores de proéminences pour chaque taille de N			
score de proéminence	Nombre de syllabes de N		
	2	3	4
0	31,1%	24,2%	22,9%
1	25,3%	23,8%	23,2%
2	29,3%	33,3%	33,7%
3	14,2%	18,7%	20,2%

**Annexe 16 – B : Proéminence perçue sur la première syllabe de ADJ.**

PROEMINENCE SUR LA SYLLABE INITIALE DE ADJ			
Effets principaux	LR Chisq	Df	Pr(>Chisq)
group_cond	11,41	4	0,022357*
N_syll	7,12	2	0,028425*
ADJ_syll	0,86	2	0,651263
Effets d'interaction	LR Chisq	Df	Pr(>Chisq)
N_syll:ADJ_syll	16,92	4	0,002004**

Distribution de probabilités des scores de proéminences pour chaque groupement					
score de proéminence	groupement				
	11	42	32	31	22
0	25,4%	25,1%	24,8%	25,8%	16,7%
1	29,8%	29,0%	26,4%	25,1%	23,5%
2	26,8%	26,7%	26,5%	25,4%	29,1%
3	18,0%	19,2%	22,3%	23,7%	30,7%

A) Distribution de probabilités des scores de proéminences pour chaque taille de N			
score de proéminence	Nombre de syllabes de N		
	2	3	4
0	26,4%	24,3%	18,7%
1	29,3%	27,8%	25,2%
2	26,2%	27,0%	28,9%
3	18,1%	20,9%	27,2%
B) Moyennes des scores de proéminences calculées à partir des résultats de la distribution de probabilités			
Nombre de syllabes de N	mean.class	mean.class(-1)	.group
2	2,36	1,36	a
3	2,45	1,45	ab
4	2,64	1,64	b

**Distribution de probabilités des scores de proéminences pour chaque taille de N et ADJ**

Nombre de syllabes de N	Score de proéminence	Nombre de syllabes de ADJ		
		2	3	4
2	0	31,6%	25,9%	21,6%
	1	30,3%	29,3%	28,3%
	2	23,6%	26,4%	28,7%
	3	14,5%	18,4%	21,4%
3	0	27,1%	23,7%	22,0%
	1	27,9%	27,8%	27,6%
	2	25,7%	27,1%	28,2%
	3	19,4%	21,3%	22,1%
4	0	13,8%	16,1%	26,3%
	1	22,6%	23,4%	29,6%
	2	31,1%	29,3%	26,2%
	3	32,5%	31,2%	17,8%

**Annexe 16 – C :** Force relative entre la proéminence perçue sur la dernière syllabe de N1 et la première syllabe de N1.

<b>FORCE RELATIVE ENTRE LES DEUX PROEMINENCES PERCUES : DERNIERE SYLLABE DE N1 - PREMIERE SYLLABE DE N1</b>			
<b>Effets principaux</b>	<b>LR Chisq</b>	<b>Df</b>	<b>Pr(&gt;Chisq)</b>
group_cond	33,16	4	1,109e-06 ***

<b>Distribution de probabilités des scores de force relative entre les proéminences perçues pour chaque groupement</b>					
<b>score de proéminence</b>	<b>groupement</b>				
	<b>32</b>	<b>42</b>	<b>31</b>	<b>22</b>	<b>11</b>
<b>-3</b>	2,7%	2,2%	2,3%	1,9%	1,3%
<b>-2</b>	10,9%	9,4%	9,5%	7,9%	5,9%
<b>-1</b>	32,3%	30,6%	29,3%	26,8%	22,7%
<b>0</b>	26,0%	27,0%	25,9%	26,7%	26,5%
<b>1</b>	18,7%	20,3%	21,1%	23,2%	26,4%
<b>2</b>	7,1%	7,9%	8,9%	10,0%	12,6%
<b>3</b>	2,3%	2,6%	3,1%	3,5%	4,5%

**Annexe 16 – D :** Force relative entre la proéminence perçue sur la première syllabe de N2 et la dernière syllabe de N1.

<b>FORCE RELATIVE ENTRE LES DEUX PROEMINENCES PERCUES : PREMIERE SYLLABE DE N2 - DERNIERE SYLLABE DE N1</b>			
<b>Effets principaux</b>	<b>LR Chisq</b>	<b>Df</b>	<b>Pr(&gt;Chisq)</b>
group_cond	17,98	4	0,001244 **

<b>Distribution de probabilités des scores de force relative entre les proéminences perçues pour chaque groupement</b>					
<b>score de proéminence</b>	<b>groupement</b>				
	<b>11</b>	<b>22</b>	<b>31</b>	<b>42</b>	<b>32</b>
<b>-3</b>	2,9%	2,5%	1,9%	1,9%	1,6%
<b>-2</b>	9,2%	8,0%	6,2%	6,2%	5,3%
<b>-1</b>	22,9%	20,9%	17,6%	17,5%	15,5%
<b>0</b>	34,0%	34,0%	33,2%	33,1%	32,0%
<b>1</b>	23,4%	25,8%	29,8%	29,9%	32,4%
<b>2</b>	6,2%	7,2%	9,2%	9,2%	10,8%
<b>3</b>	1,4%	1,6%	2,2%	2,2%	2,6%

**Annexe 16 – E :** Force relative entre la proéminence perçue sur la dernière syllabe de N2 et la dernière syllabe de N1.

<b>FORCE RELATIVE ENTRE LES DEUX PROEMINENCES PERCUES : DERNIERE SYLLABE DE N2 - DERNIERE SYLLABE DE N1</b>			
<b>Effets principaux</b>	<b>LR Chisq</b>	<b>Df</b>	<b>Pr(&gt;Chisq)</b>
group_cond	16,68	4	0,0022294 **
N_syll	16,44	2	0,0002695 ***

<b>Distribution de probabilités des scores de force relative entre les proéminences perçues pour chaque groupement</b>					
<b>score de proéminence</b>	<b>groupement</b>				
	<b>11</b>	<b>22</b>	<b>31</b>	<b>32</b>	<b>42</b>
<b>-3</b>	2,9%	2,2%	2,1%	1,8%	1,7%
<b>-2</b>	10,9%	8,4%	8,0%	7,3%	6,8%
<b>-1</b>	26,7%	22,9%	22,4%	21,1%	20,1%
<b>0</b>	52,1%	56,6%	57,6%	58,8%	59,6%
<b>1</b>	6,4%	8,6%	8,6%	9,4%	10,1%
<b>2</b>	0,8%	1,1%	1,1%	1,3%	1,4%
<b>3</b>	0,2%	0,2%	0,2%	0,3%	0,3%

<b>Distribution de probabilités des scores de force relative entre les proéminences perçues dans chaque taille de N</b>			
<b>score de proéminence</b>	<b>Nombre de syllabes de N</b>		
	<b>2</b>	<b>3</b>	<b>4</b>
<b>-3</b>	2,7%	1,7%	2,0%
<b>-2</b>	10,2%	6,7%	7,9%
<b>-1</b>	25,8%	19,9%	22,3%
<b>0</b>	53,3%	59,7%	57,8%
<b>1</b>	6,8%	10,3%	8,7%
<b>2</b>	0,9%	1,4%	1,1%
<b>3</b>	0,2%	0,3%	0,2%

**Annexe 16 – F :** Force relative entre la proéminence perçue sur la dernière syllabe de ADJ et la première syllabe de ADJ.

<b>FORCE RELATIVE ENTRE LES DEUX PROEMINENCES PERÇUES : DERNIERE SYLLABE DE ADJ - PREMIERE SYLLABE DE ADJ</b>			
<b>Effets principaux</b>	<b>LR Chisq</b>	<b>Df</b>	<b>Pr(&gt;Chisq)</b>
group_cond	11,20	4	0,02445 *
N_syll	3,45	2	0,17843
ADJ_syll	1,07	2	0,58601
<b>Effet d'interaction</b>	<b>LR Chisq</b>	<b>Df</b>	<b>Pr(&gt;Chisq)</b>
N_syll:ADJ_syll	29,91	4	5,114e-06 ***

<b>Distribution de probabilités des scores de force relative entre les proéminences perçues pour chaque groupement</b>					
<b>score de proéminence</b>	<b>groupement</b>				
	<b>32</b>	<b>42</b>	<b>31</b>	<b>22</b>	<b>11</b>
<b>-3</b>	3,7%	2,8%	2,7%	3,8%	2,0%
<b>-2</b>	8,7%	7,2%	7,0%	9,3%	5,5%
<b>-1</b>	24,1%	23,3%	23,4%	27,0%	19,8%
<b>0</b>	33,3%	35,7%	35,7%	34,5%	35,7%
<b>1</b>	22,3%	23,3%	23,1%	19,4%	27,1%
<b>2</b>	6,8%	6,6%	6,8%	5,1%	8,5%
<b>3</b>	1,2%	1,1%	1,2%	0,8%	1,5%

**Distribution de probabilités des scores de force relative entre les proéminences perçues pour chaque taille de N et ADJ**

Nombre de syllabes de N	score de proéminence	Nombre de syllabes de ADJ		
		2	3	4
2	-3	1,8%	1,7%	3,3%
	-2	4,9%	4,7%	8,5%
	-1	18,1%	18,1%	26,9%
	0	35,0%	35,9%	36,3%
	1	29,1%	29,0%	19,4%
	2	9,5%	9,0%	4,8%
	3	1,7%	1,5%	0,8%
3	-3	2,3%	3,0%	2,8%
	-2	6,2%	7,7%	7,5%
	-1	21,5%	24,8%	24,8%
	0	36,2%	36,3%	36,8%
	1	25,2%	21,6%	21,6%
	2	7,4%	5,7%	5,6%
	3	1,2%	0,9%	0,9%
4	-3	4,6%	5,7%	1,6%
	-2	11,3%	12,7%	4,4%
	-1	30,7%	30,3%	16,5%
	0	33,6%	31,7%	33,3%
	1	15,6%	15,2%	30,8%
	2	3,7%	3,7%	11,3%
	3	0,6%	0,6%	2,0%

**Annexe 16 – G :** Force relative entre la proéminence perçue sur la première syllabe de ADJ et la dernière syllabe de N2.

<b>FORCE RELATIVE ENTRE LES DEUX PROEMINENCES PERCUES : PREMIERE SYLLABE DE ADJ - DERNIERE SYLLABE DE N2</b>			
<b>Effets principaux</b>	<b>LR Chisq</b>	<b>Df</b>	<b>Pr(&gt;Chisq)</b>
group_cond	11,56	4	0,02095 *
N_syll	3,88	2	0,14359
ADJ_syll	2,18	2	0,33638
<b>Effets d'interaction</b>	<b>LR Chisq</b>	<b>Df</b>	<b>Pr(&gt;Chisq)</b>
N_syll:ADJ_syll	18,18	4	0,00114 **

<b>Distribution de probabilités des scores de force relative entre les proéminences perçues pour chaque groupement</b>					
<b>score de proéminence</b>	<b>groupement</b>				
	<b>11</b>	<b>42</b>	<b>32</b>	<b>31</b>	<b>22</b>
<b>-3</b>	0,6%	0,5%	0,4%	0,5%	0,3%
<b>-2</b>	2,8%	2,3%	1,9%	2,4%	1,5%
<b>-1</b>	19,5%	17,0%	14,3%	17,0%	11,3%
<b>0</b>	39,4%	38,1%	35,2%	37,2%	31,9%
<b>1</b>	27,6%	30,0%	33,0%	30,0%	36,0%
<b>2</b>	7,8%	9,2%	11,4%	9,7%	14,2%
<b>3</b>	2,3%	2,9%	3,7%	3,1%	4,8%

Distribution de probabilités des scores de force relative entre les proéminences perçues pour chaque taille de N et ADJ				
Nombre de syllabes de N	score de proéminence	Nombre de syllabes de ADJ		
		2	3	4
2	-3	0,6%	0,5%	0,4%
	-2	2,9%	2,5%	1,9%
	-1	20,2%	18,0%	14,6%
	0	39,6%	38,5%	36,9%
	1	27,0%	29,1%	32,6%
	2	7,5%	8,6%	10,3%
	3	2,2%	2,6%	3,2%
3	-3	0,4%	0,5%	0,4%
	-2	2,1%	2,3%	2,0%
	-1	15,5%	16,3%	15,4%
	0	36,3%	36,9%	37,7%
	1	31,7%	30,7%	31,7%
	2	10,6%	10,2%	9,8%
	3	3,4%	3,3%	3,0%
4	-3	0,2%	0,4%	0,7%
	-2	1,0%	1,7%	3,2%
	-1	8,0%	12,8%	21,6%
	0	27,5%	33,9%	40,2%
	1	39,3%	34,4%	25,5%
	2	17,8%	12,7%	6,8%
	3	6,3%	4,2%	2,0%

**Annexe 16 – H : Force relative entre la proéminence perçue sur la première syllabe de ADJ et la première syllabe de N2.**

<b>FORCE RELATIVE ENTRE LES DEUX PROEMINENCES PERCUES : PREMIERE SYLLABE DE ADJ - PREMIERE SYLLABE DE N2</b>			
<b>Effets principaux</b>	<b>LR Chisq</b>	<b>Df</b>	<b>Pr(&gt;Chisq)</b>
group_cond	21,46	4	0,0002565 ***
N_syll	8,01	2	0,0182308 *
ADJ_syll	0,40	2	0,8192192
<b>Effets d'interaction</b>	<b>LR Chisq</b>	<b>Df</b>	<b>Pr(&gt;Chisq)</b>
N_syll:ADJ_syll	20,89	4	0,0003335 ***

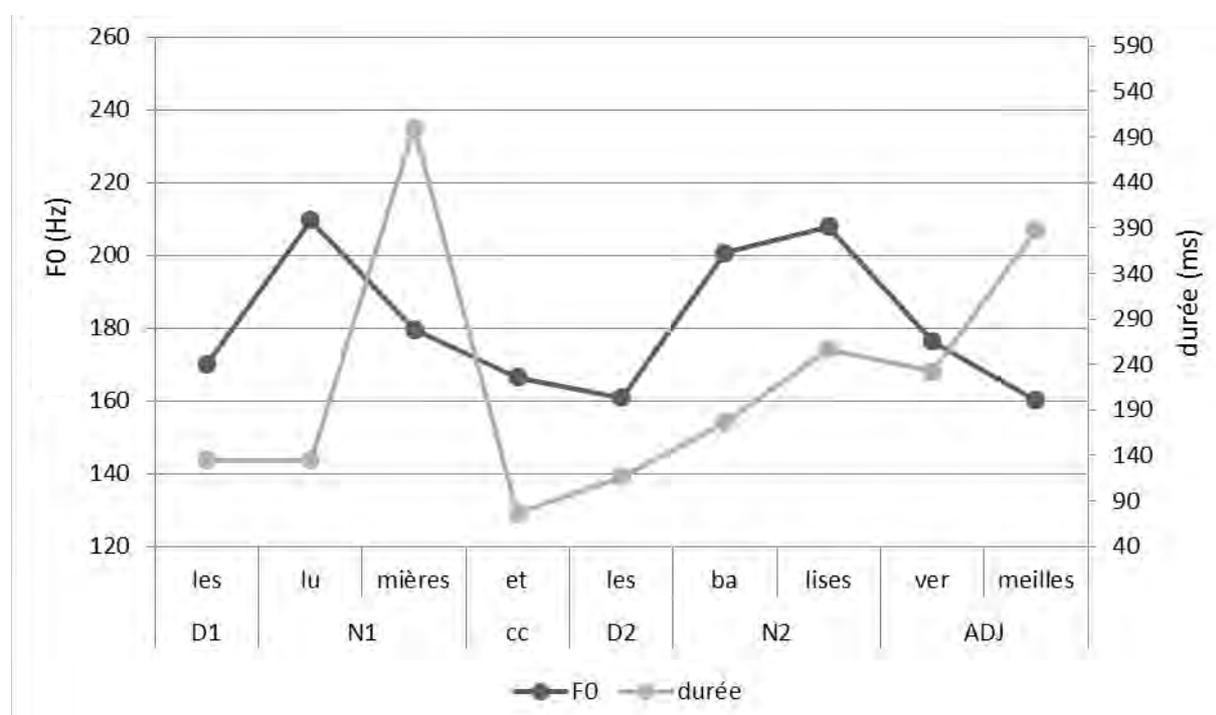
<b>Distribution de probabilités des scores de force relative entre les proéminences perçues pour chaque groupement</b>					
<b>score de proéminence</b>	<b>groupement</b>				
	<b>11</b>	<b>42</b>	<b>32</b>	<b>31</b>	<b>22</b>
<b>-3</b>	0,6%	0,5%	0,6%	0,5%	0,3%
<b>-2</b>	4,9%	4,8%	5,3%	4,2%	2,4%
<b>-1</b>	26,6%	26,0%	26,4%	24,1%	15,3%
<b>0</b>	50,1%	49,7%	48,6%	50,6%	47,2%
<b>1</b>	13,4%	14,1%	14,2%	15,2%	23,7%
<b>2</b>	3,5%	3,8%	3,8%	4,2%	8,5%
<b>3</b>	1,0%	1,1%	1,1%	1,2%	2,6%

Distribution de probabilités des scores de force relative entre les proéminences perçues pour chaque taille de N et ADJ				
Nombre de syllabes de N	score de proéminence	Nombre de syllabes de ADJ		
		2	3	4
2	-3	0,9%	0,7%	0,4%
	-2	7,8%	5,8%	3,8%
	-1	34,7%	29,6%	22,8%
	0	44,9%	48,4%	52,0%
	1	8,9%	11,7%	15,6%
	2	2,1%	3,0%	4,2%
	3	0,6%	0,8%	1,2%
3	-3	0,5%	0,5%	0,4%
	-2	4,3%	4,1%	3,6%
	-1	23,5%	22,9%	21,8%
	0	49,2%	48,7%	52,1%
	1	16,4%	17,0%	16,4%
	2	4,8%	5,3%	4,5%
	3	1,4%	1,6%	1,3%
4	-3	0,2%	0,4%	0,5%
	-2	1,7%	3,4%	4,6%
	-1	12,2%	20,3%	25,2%
	0	47,4%	50,1%	50,5%
	1	26,2%	18,4%	14,4%
	2	9,4%	5,7%	3,8%
	3	2,9%	1,7%	1,1%

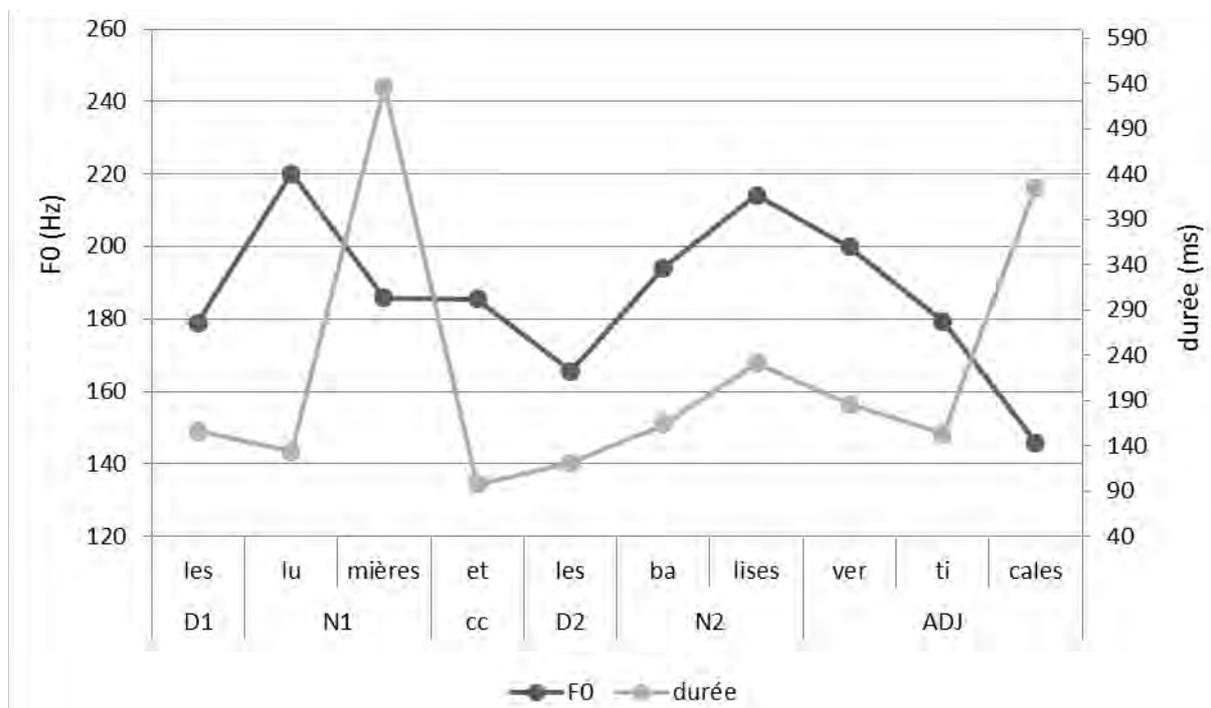
## Annexe 17 :

Moyennes de  $f_0$  et de durée sur les toutes les syllabes des syntagmes réalisés en Condition 1 et 2, tous locuteurs confondus.

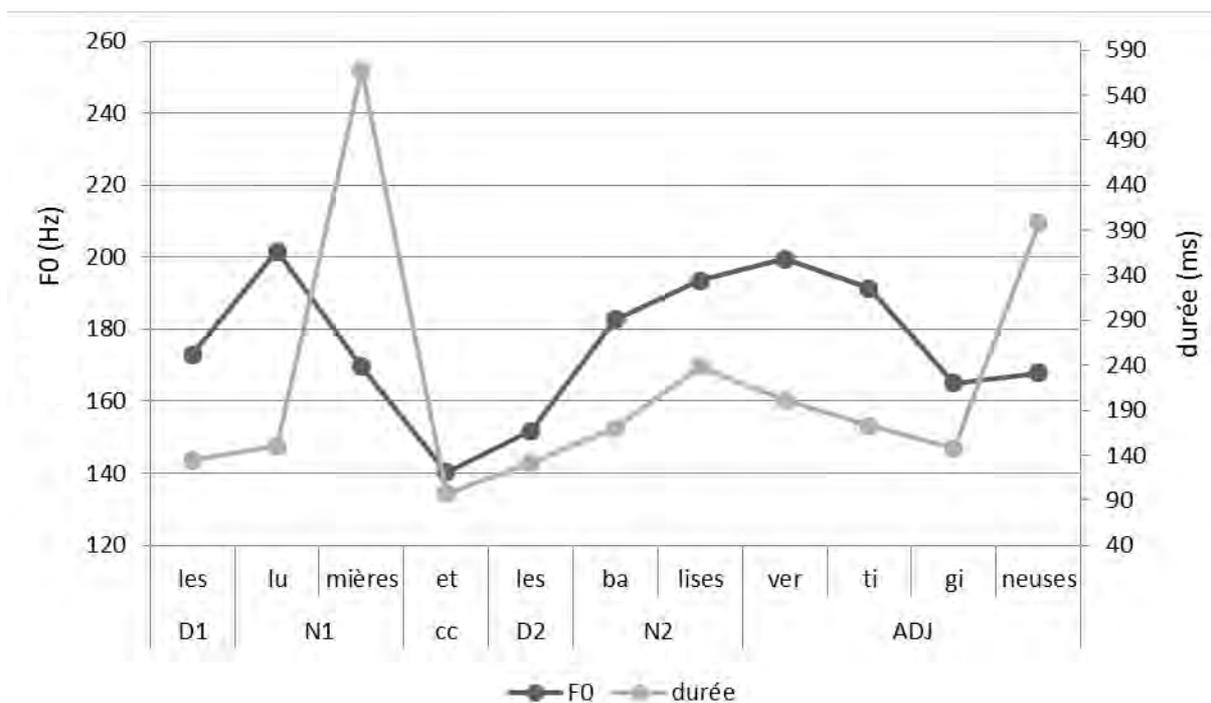
**Annexe 17 – A :** Variations de  $f_0$  et de durée sur le syntagme « les lumières et les balises vermeilles » réalisé en Condtion 1.



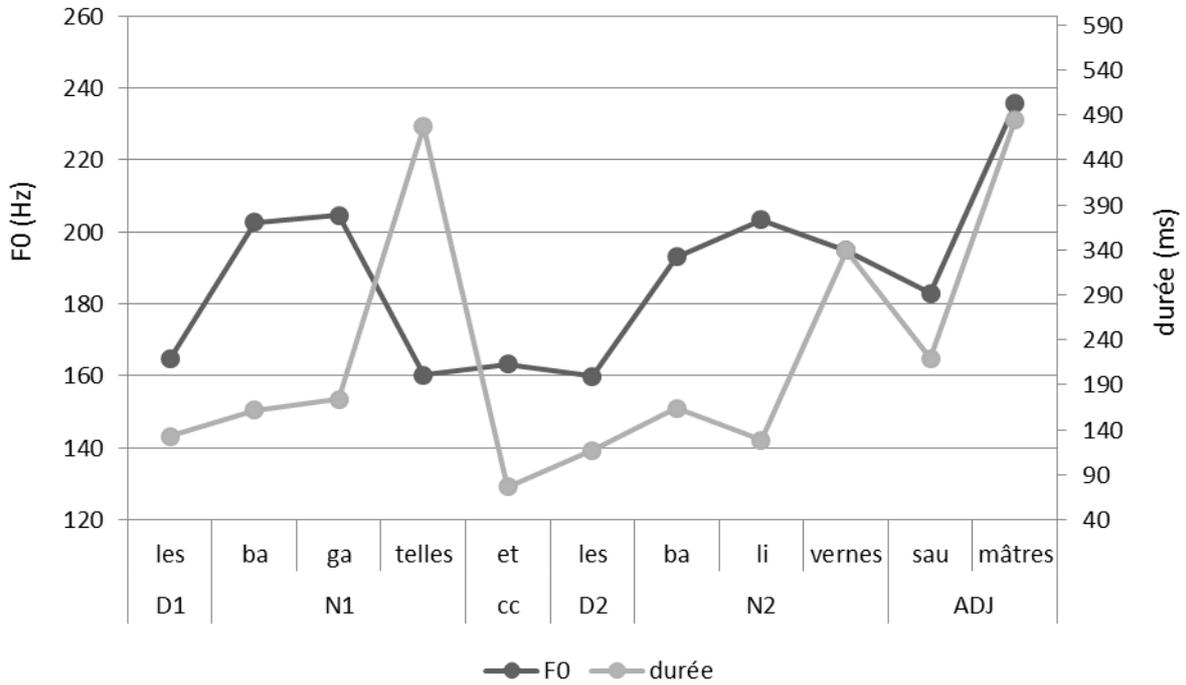
**Annexe 17 – B :** Variations de f0 et de durée sur le syntagme « les lumières et les balises verticales » réalisé en Condtion 1.



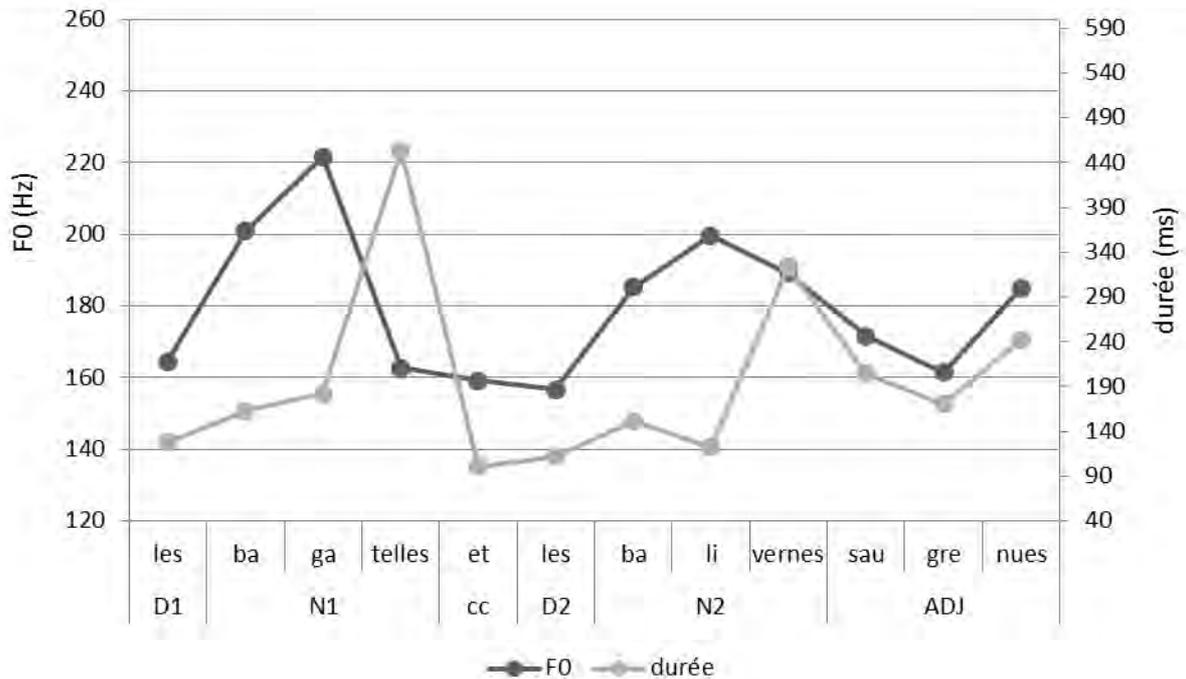
**Annexe 17 – C :** Variations de f0 et de durée sur le syntagme « les lumières et les balises vertigineuses » réalisé en Condtion 1.



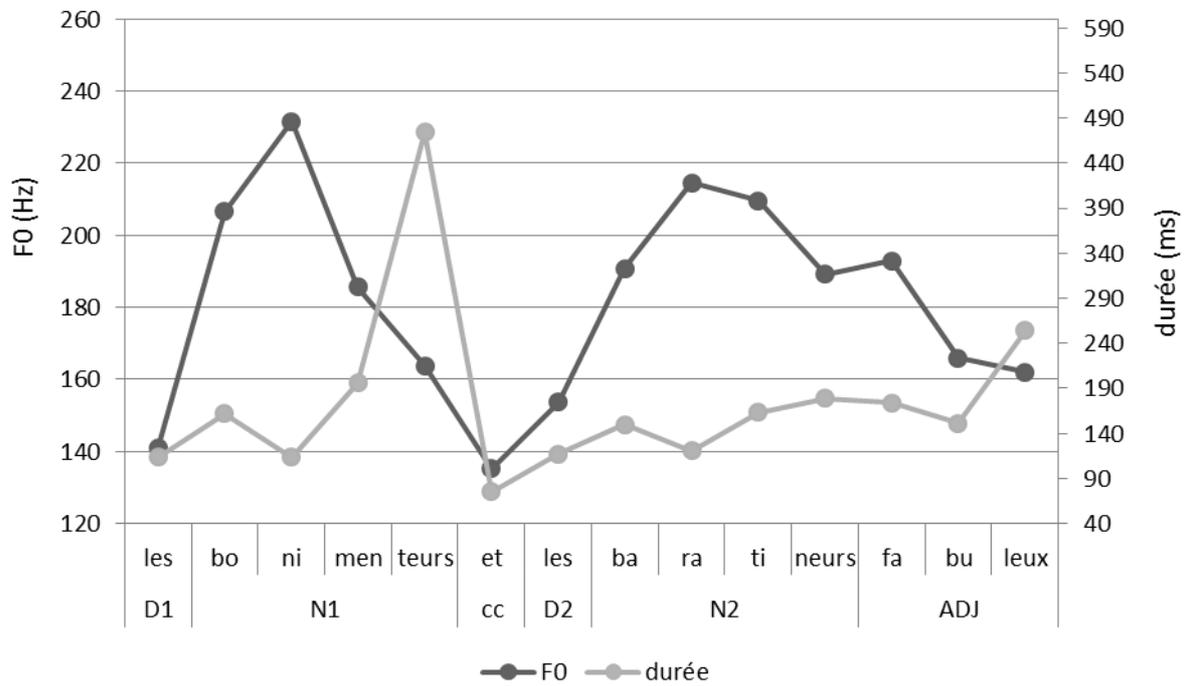
**Annexe 17 – D :** Variations de *f0* et de durée sur le syntagme « les bagatelles et les balivernes saumâtres » réalisé en Condition 1.



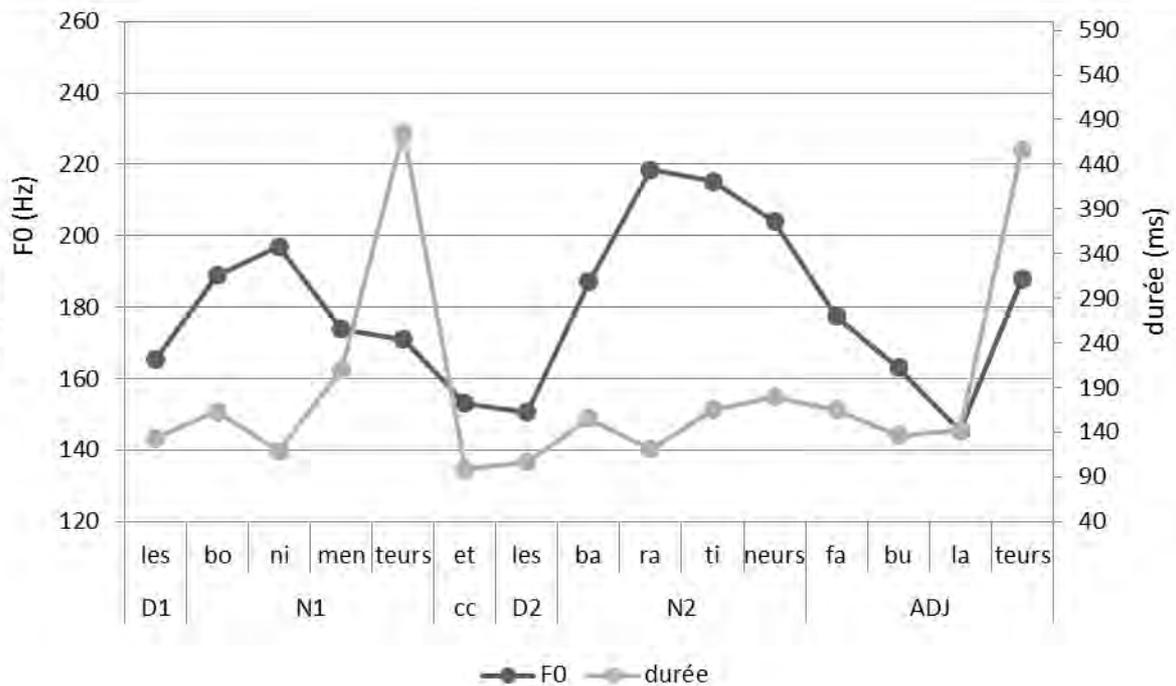
**Annexe 17 – E :** Variations de *f0* et de durée sur le syntagme « les bagatelles et les balivernes saugrenues » réalisé en Condition 1.



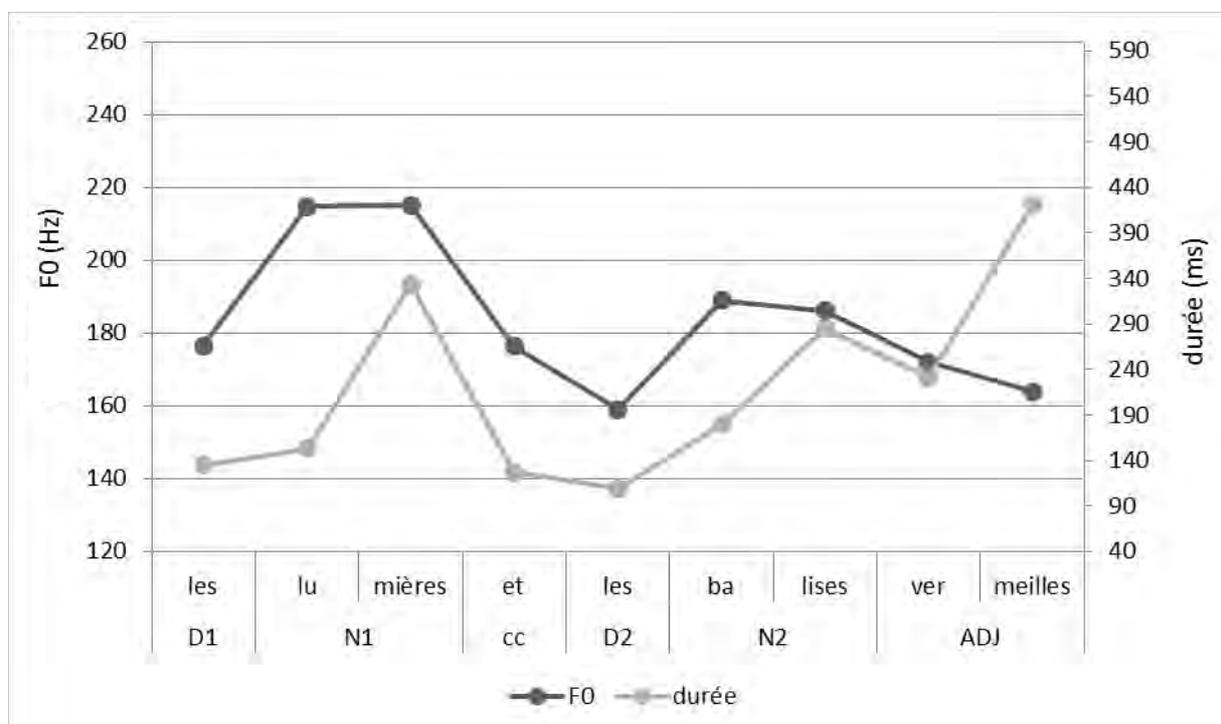
**Annexe 17 – H :** Variations de f0 et de durée sur le syntagme « les bonimenteurs et baratineurs fabuleux » réalisé en Condtion 1.



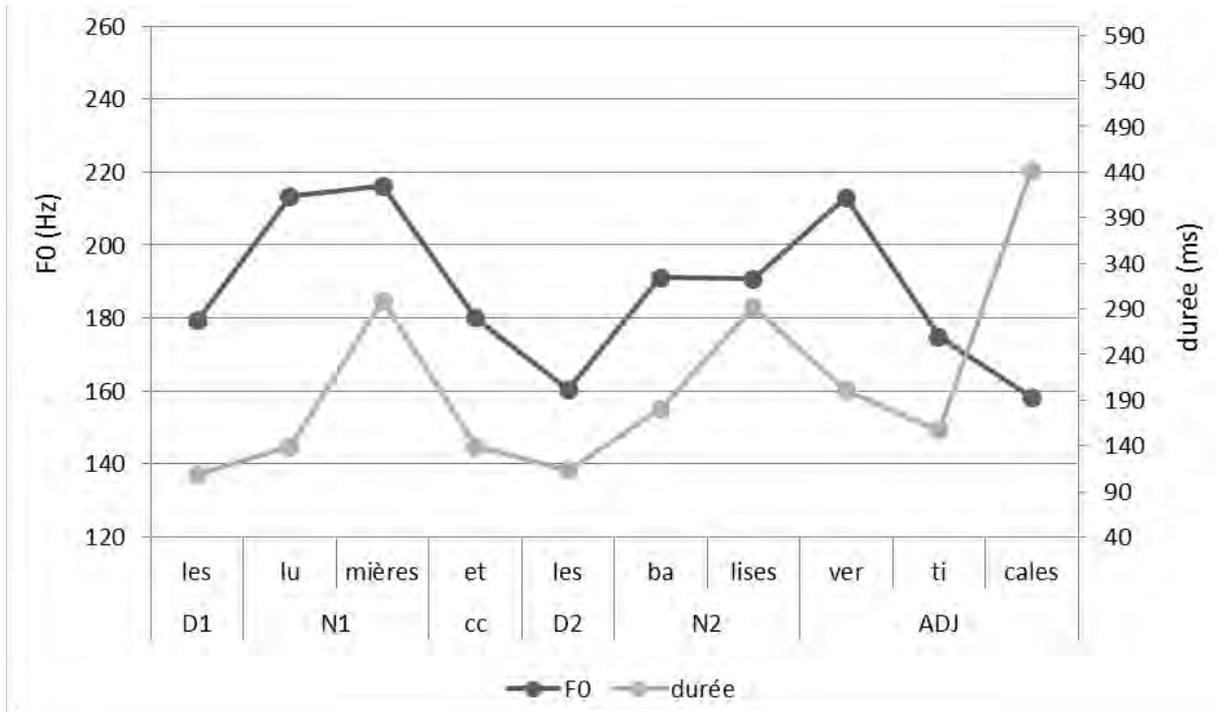
**Annexe 17 – I :** Variations de f0 et de durée sur le syntagme « les bonimenteurs et baratineurs fabulateurs » réalisé en Condtion 1.



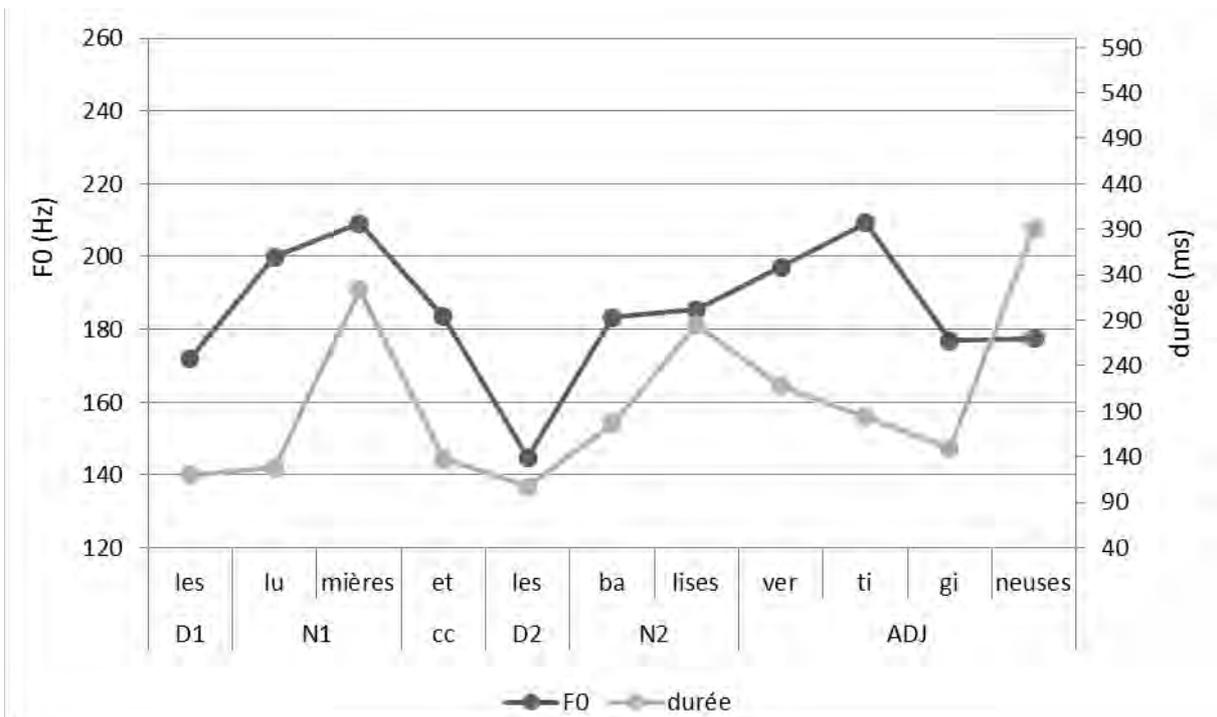
**Annexe 17 – J : Variations de f0 et de durée sur le syntagme « les lumières et les balises vertes » réalisé en Condition 2.**



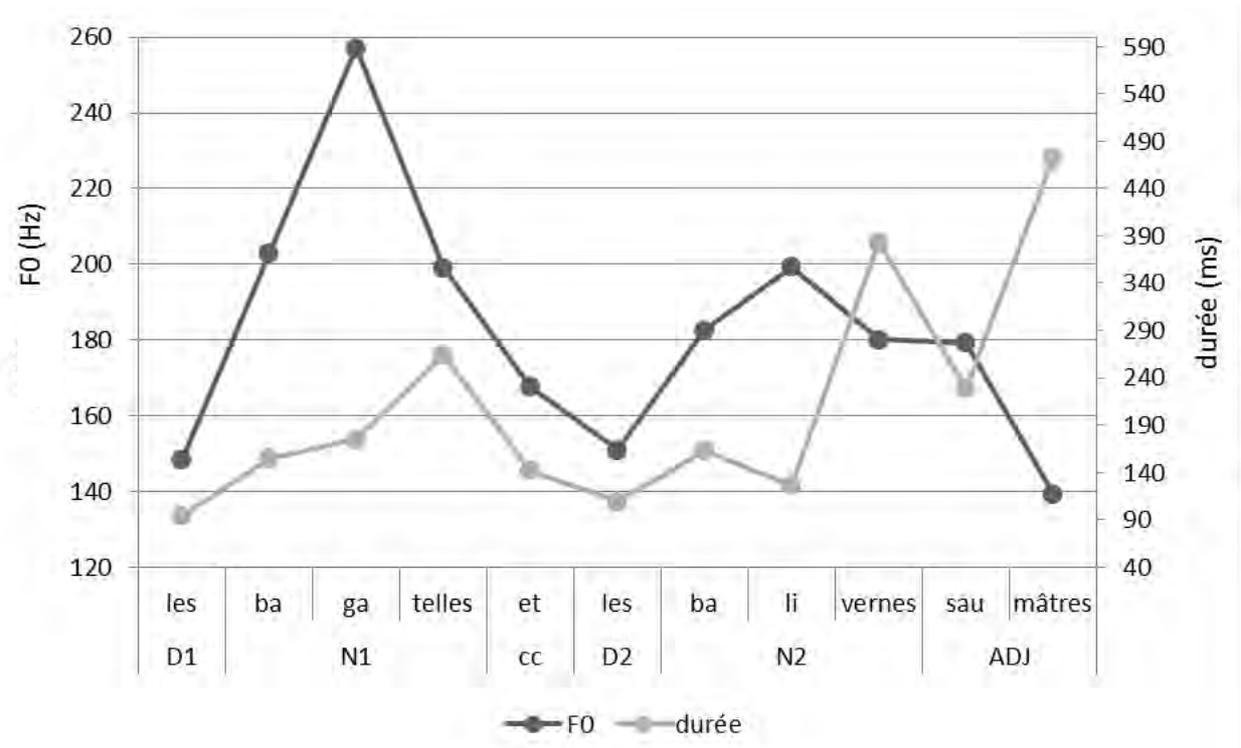
**Annexe 17 – K :** Variations de f0 et de durée sur le syntagme « les lumières et les balises verticales » réalisé en Condtion 2.



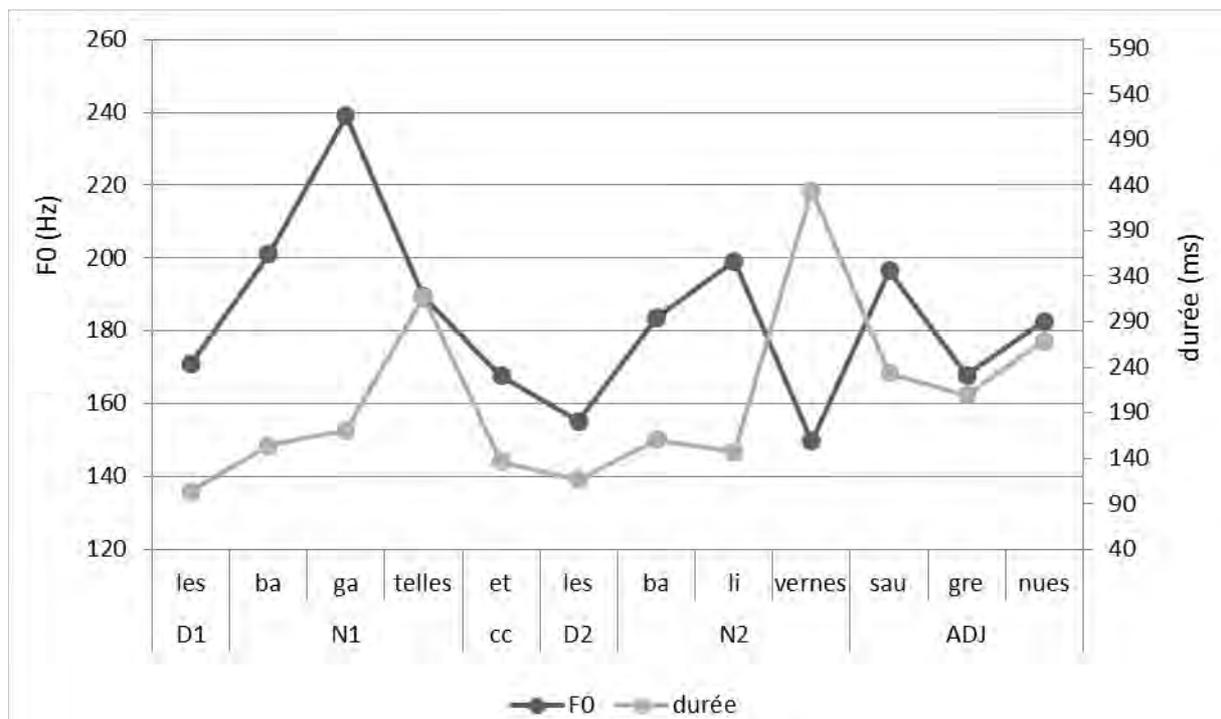
**Annexe 17 – L :** Variations de f0 et de durée sur le syntagme « les lumières et les balises vertigineuses » réalisé en Condtion 2.



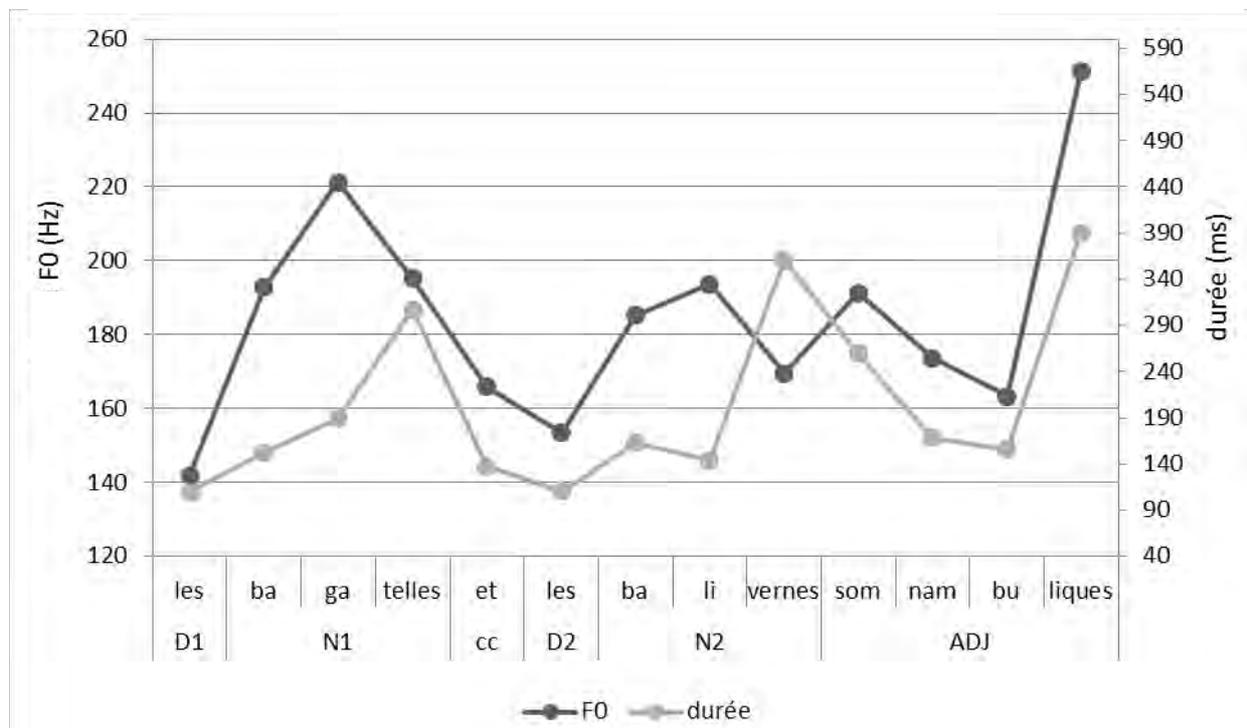
**Annexe 17 – M :** Variations de *f0* et de durée sur le syntagme « les baguettes et les balivernes saumâtres » réalisé en Condition 2.



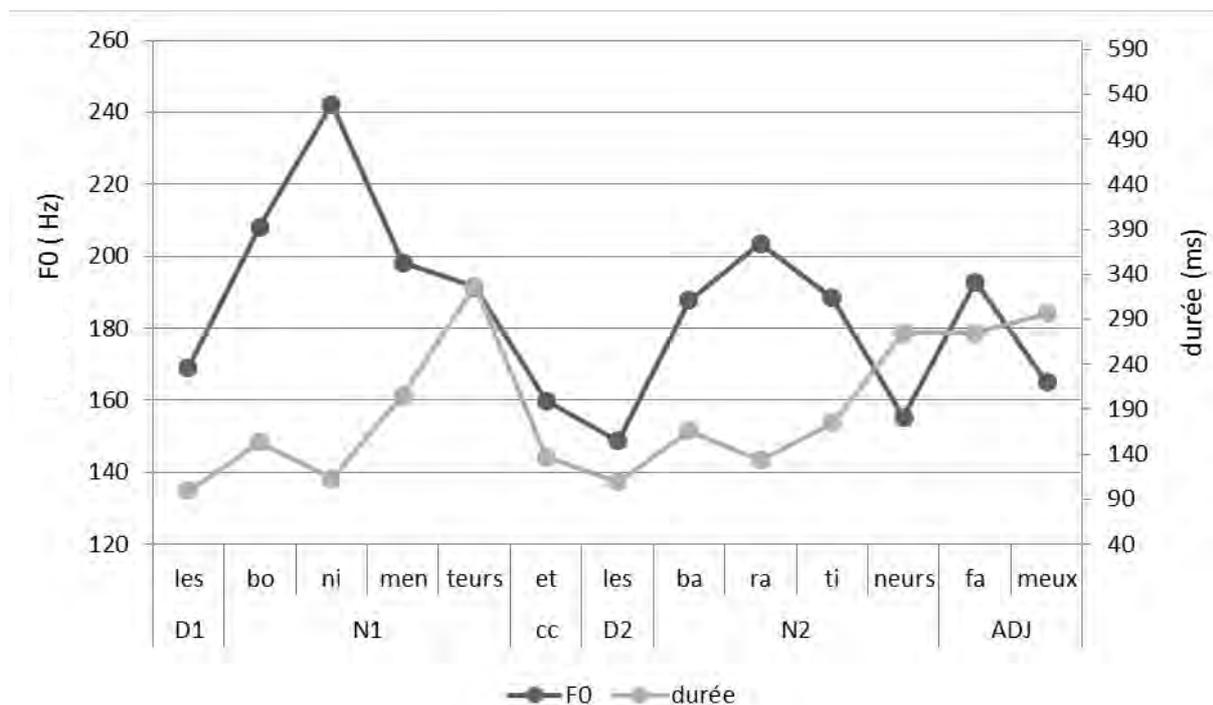
**Annexe 17 – N :** Variations de *f0* et de durée sur le syntagme « les baguettes et les balivernes saugrenues » réalisé en Condition 2.



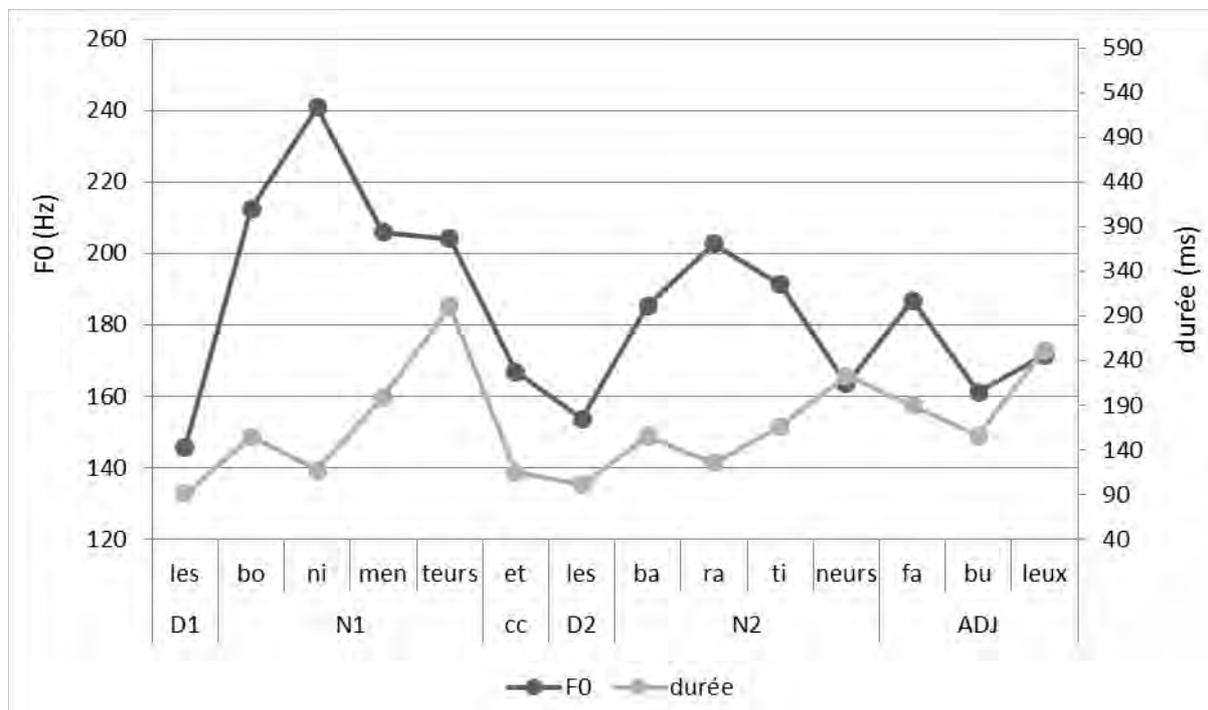
**Annexe 17 – O :** Variations de f0 et de durée sur le syntagme « les bagatelles et les balivernes sonnambuliques » réalisé en Condition 2.



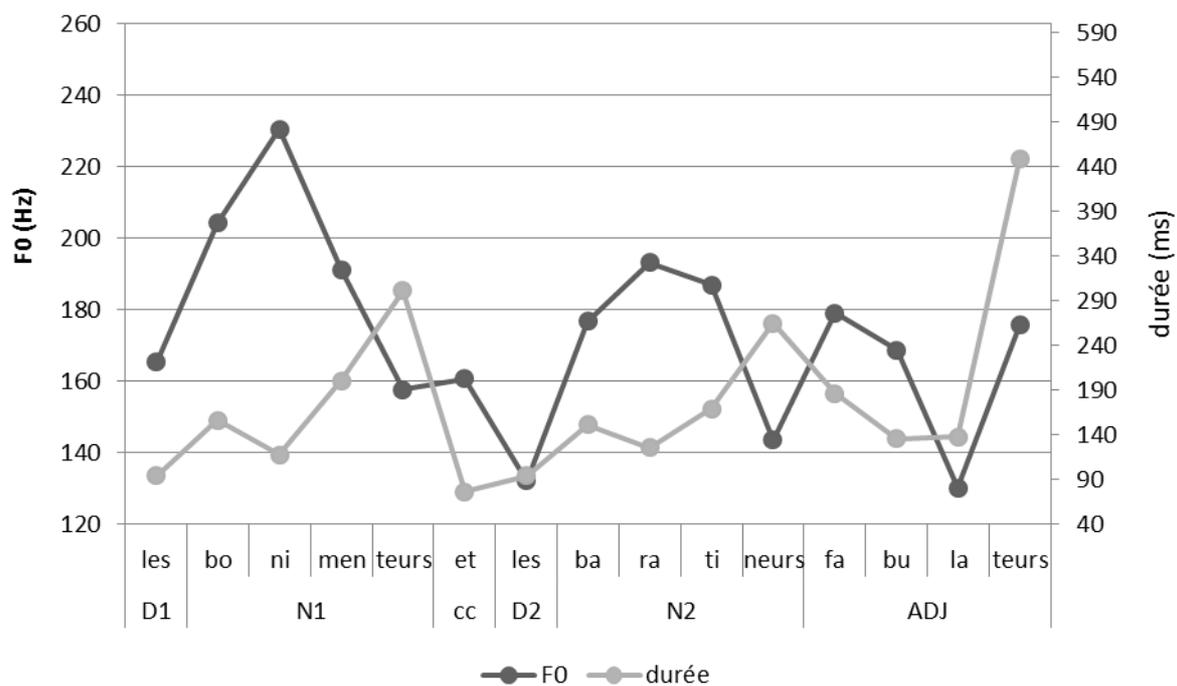
**Annexe 17 – P :** Variations de f0 et de durée sur le syntagme « les bonimenteurs et les baratineurs fameux » réalisé en Condition 2.



**Annexe 17 – Q :** Variations de  $f_0$  et de durée sur le syntagme « les bonimenteurs et les baratineurs fabuleux » réalisé en Condtion 2.



**Annexe 17 – R :** Variations de  $f_0$  et de durée sur le syntagme « les bonimenteurs et les baratineurs fabulateurs » réalisé en Condtion 2.





## Annexe 18 :

### Résultats de l'effet des variations de durée et de f0 sur la force des frontières et des proéminences.

*Annexe 18 – A : Frontière perçue après N1 en Condition 2.*

CORRELATS ACOUSTIQUES DE LA FORCE DE LA FRONTIERE PERCUE APRES N1				
EN CONDITION 2				
Effets principaux	LR Chisq	Df	Pr(>Chisq)	Estimate
duree_syll_finale_N1	7,47	1	0,006257 **	0,006718
f0_syll_finale_N1	8,30	1	0,003968 **	-0,008478
f0_syll_initiale_N2	4,44	1	0,035120 *	0,008127

*Annexe 18 – B : Frontière perçue après N2 en Condition 1 et en Condition 2.*

CORRELATS ACOUSTIQUES DE LA FORCE DE LA FRONTIERE PERCUE APRES N2				
EN CONDITION 1				
Effets principaux	LR Chisq	Df	Pr(>Chisq)	Estimate
duree_syll_finale_N2	7,84	1	0,0051143 **	0,004197
f0_syll_finale_N2	21,80	1	3,031e-06 ***	-0,016356
f0_syll_initiale_ADJ	13,65	1	0,0002204 ***	0,011309
EN CONDITION 2				
Effets principaux	LR Chisq	Df	Pr(>Chisq)	Estimate
duree_syll_finale_N2	11,18	1	0,000829 ***	0,004675
f0_syll_finale_N2	17,06	1	3,629e-05 ***	-0,013695

**Annexe 18 – C : Proéminence perçue sur la syllabe finale de N1 en Condition 1.**

CORRELATS ACOUSTIQUES DE LA FORCE DE LA PROEMINENCE PERCUE SUR LA SYLLABE FINALE DE N1				
EN CONDITION 1				
Effets principaux	LR Chisq	Df	Pr(>Chisq)	Estimate
duree_syll_finale_N1	15,83	1	6,924e-05 ***	0,006316
f0_syll_finale_N1	12,57	1	0,0003915 ***	0,015587
f0_syll_initiale_N2	5,00	1	0,0253876 *	-0,015745

**Annexe 18 – D : Force relative entre la proéminence perçue sur la syllabe initiale de N2 et la proéminence perçue sur la syllabe finale de N1 en Condition 1 et en Condition 2.**

CORRELATS ACOUSTIQUES DE LA FORCE RELATIVE ENTRE LES PROEMINENCES PERCUES SUR : PREMIERE SYLLABE DE N2 - DERNIERE SYLLABE DE N1				
EN CONDITION 1				
Effets principaux	LR Chisq	Df	Pr(>Chisq)	Estimate
duree_syll_finale_N1	4,45	1	0,0348550 *	-0,003644
f0_syll_finale_N1	8,33	1	0,0038899 **	-0,013772
f0_syll_initiale_N2	11,41	1	0,0007303 ***	0,017644
EN CONDITION 2				
Effets principaux	LR Chisq	Df	Pr(>Chisq)	Estimate
f0_syll_finale_N1	16,68	1	4,429e-05 ***	-0,010689
f0_syll_initiale_N2	11,53	1	0,0006848 ***	0,014565

**Annexe 18 – E : Proéminence perçue sur la syllabe initiale de ADJ en Condition 1 et en Condition 2.**

CORRELATS ACOUSTIQUES DE LA FORCE DE LA PROEMINENCE PERCUE SUR LA SYLLABE INITIALE DE ADJ				
EN CONDITION 1				
Effets principaux	LR Chisq	Df	Pr(>Chisq)	Estimate
duree_syll_finale_N2	4,05	1	0,044099 *	-0,0017923
f0_syll_finale_N2	10,92	1	0,000953 ***	-0,0107547
f0_syll_finale_ADJ	4,68	1	0,030451 *	0,0033854
EN CONDITION 2				
Effets principaux	LR Chisq	Df	Pr(>Chisq)	Estimate
duree_syll_finale_N2	4,62	1	0,031535 *	-0,003777
duree_syll_finale_ADJ	8,85	1	0,002938 **	-0,005233
f0_syll_finale_N2	16,80	1	4,157e-05 ***	-0,013224
f0_syll_initiale_ADJ	6,32	1	0,011944 *	0,008919
f0_syll_finale_ADJ	12,72	1	0,000362 ***	-0,00748

**Annexe 18 – F : Force relative entre la proéminence perçue sur la syllabe initiale de ADJ et la proéminence perçue sur la syllabe finale de N2 en Condition 1 et en Condition 2.**

CORRELATS ACOUSTIQUES DE LA FORCE RELATIVE ENTRE LES PROEMINENCES PERCUES SUR : PREMIERE SYLLABE DE ADJ - DERNIERE SYLLABE DE N2				
EN CONDITION 1				
Effets principaux	LR Chisq	Df	Pr(>Chisq)	Estimate
f0_syll_finale_N2	5,58	1	0,01819 *	-0,011852
EN CONDITION 2				
Effets principaux	LR Chisq	Df	Pr(>Chisq)	Estimate
f0_syll_finale_N2	16,79	1	4,173e-05 ***	-0,009794



## Annexe 19 :

# Article publié dans les actes de la conférence internationale *Speech Prosody 2016*

## Perceptual investigation of prosodic phrasing in French

Laury Garnier<sup>1,2</sup>, Lorraine Baqué<sup>3</sup>, Anne Dagnac<sup>2</sup>, Corine Astésano<sup>1</sup>

<sup>1</sup> URI Octogone- Lordat (E.A 4156), Université de Toulouse, UT2J, France

<sup>2</sup> CLLE (UMR 5263), Université de Toulouse, CNRS, UT2J, France

<sup>3</sup> Laboratori fLexSem, Universitat Autònoma de Barcelona, Spain

laury.garnier@univ-tlse2.fr, lorraine.baque@uab.cat, dagnac@univ-tlse2.fr, astesano@univ-tlse2.fr

### Abstract

The aim of this paper is to investigate prosodic phrasing and more precisely the use of prosodic cues in the marking of morphosyntactic units in French. As a first step towards this goal, a perception study was conducted on 27 listeners, who had to perform 3 distinct perceptual tasks on 32 syntactically controlled phrases read by a female speaker: a prominence strength judgment task, a boundary strength judgment task, and a task where listeners had to choose between 4 different phrase groupings intended to reflect the potential choices of prosodic phrasing. The corpus consists of syntactically ambiguous structures manipulating high and low adjective attachment on 2 coordinated nouns. It was designed to specifically test the role of prominence and boundary cues in the marking of prosodic constituency. Our results show that listeners use prosodic cues to discriminate between the two syntactic structures, with boundary cues being more readily used to capture morphosyntactic structuring. More interestingly, our results indicate that prominence and boundary cues are used to distinguish finer-grained grouping levels than those predicted by traditional descriptions on French prosodic structure.

**Index Terms:** prosodic phrasing, French, perception, prominence, boundary, grouping.

### 1. Introduction

One of the most important functions of prosody is to segment the speech flow into organized units. Prosodic cues such as prominences and boundaries are used to structure speech and partake in what is referred to as 'prosodic phrasing'. In doing so, prosodic cues help linearize constraints arising from other linguistic levels such as semantic-pragmatic constraints and syntactic rules [1; 2]. A long line of research has been specifically interested in establishing the relationship between prosody and syntax. While most prosodists do not deny some interdependencies between prosody and syntax, it is arguably difficult to find common grounds on a precise relationship between the two linguistic components. Whether one assumes a mapping between syntax and prosody, whereby prosodic cues mark morphosyntactic units, or not, it is difficult to find a clear consensus on the levels reflecting prosodic hierarchy.

Indeed, there is a wide array of propositions ranging from 2 levels (*Accentual Phrase* or *ap*, and *Intonational Phrase* or *IP*: [3]) to 5 levels (including according to models: *Prosodic/Phonological Word* or *pw*, *Clitic Group*, *Phonological Phrase* or *PP*, *Intonational Phrase* and *Utterance*: [4]-[7]). Some authors, however, propose to

envisage the possibility of recursive prosodic units in order to better capture the link between prosody and syntax [8; 9]. In these proposals, *Super-Major Phrase* (corresponding to an *IP* containing several *IPs*) can be found to posit an intermediate level between the *Utterance* and the *Major Phrase*. This recursion principle has been captured in more recent propositions revising the *Strict Layer Hypothesis* in the frame of the *Optimality Theory* [10; 11]. Also, a revision of the initial model by Selkirk [4] proposed to divide the *PP* level into 2 distinct levels: a *Minor PP* and a *Major PP* [12], equivalent to the *ap* and the *Intermediate Phrase (ip)* levels, respectively, as proposed in [13]. Discussions on the existence of an intermediate level in French have also arisen, with propositions that the *ip* exists for specific syntactic structures, such as tag questions and dislocated structures [14] (see also the '*segment d'unité intonative*' in [15]). Further evidence points towards a true intermediate level applying to longer constituents and situated between the *ap* and the *IP* in French [16], more in line with the initial proposition in [13].

For French, this question about which levels are necessary to account for prosodic phrasing is of particular interest. While stress and boundary tones are clearly dissociated in English, (final) stress is syncretic to boundary tones in French, both aligning to the right edge of prosodic domains. Stress in French is also said to be post-lexical, marking the *ap* rather than the *pw*<sup>1</sup>. Both these characteristics have led some authors to question the phonological status of the final accent (*FA*) and to characterize French as a '*language without accent*' [18] or a '*boundary language*' [19; 20]. Following up on this view, *FA* is described as a *pitch accent (H\*)* found to mark the lowest level of the hierarchy (*ap*), which however disappears in favor of the *boundary tone (H%)* at the higher (*IP*) level of the hierarchy [3]. Thus, intonation pre-empts stress at higher levels of the prosodic hierarchy. Another type of accent has also been described to mark the *ap* level in French: the initial accent (*IA*), said to be secondary and optional as opposed to the primary *FA*. *IA* has been described as a mere rhythmic device occurring on longer constituents (see among others [19; 21; 22]). Descriptions of its role in the marking of the prosodic hierarchy are scarce and its functions not well established. Its functions were specifically studied on a corpus controlling for constituents' lengths and syntactic structure [23]. This acoustic study revealed that the function of *IA* is more one of structuration than rhythmic balancing. It was also shown to

<sup>1</sup> *pw* sometimes refers to the *Clitic Group* of [5; 6], hence similar to the *ap*. Here, we refer to the sense of [5; 6] and [17], corresponding to the *Lexical Word*.

more readily mark structure than *FA* at lower levels of prosodic constituency, close to the *ap* and possibly the *pw*. A perception study on a subset of the same corpus confirmed these results [24]. They showed that French listeners can perceive *IA* and *FA* independently from boundaries, with *IA* being perceived as consistently stronger than *FA* throughout the prosodic hierarchy.

The present study is a follow up of [24]. It is not intended to address syntax/prosody mapping *per se*; rather, it is specifically designed to tackle the issue of which levels of prosodic structure are necessary to account for prosodic phrasing in French. Perception is used here as an interface between the acoustic signal and phonology, in line with propositions in [25]–[27]. It is also used to circumvent the variability or ‘flexibility’ of prosodic cues found in the signal to instantiate constituency at predicted similar levels, as described in [1; 2; 28].

## 2. Method

### 2.1. Corpus

The linguistic material used for this perception study derives from the *Edinburgh Corpus* [23]. This corpus is composed of syntactically ambiguous structures that can be disambiguated via prosodic cues (prominences, boundary tones, pauses). These structures are composed of two coordinated nouns (*N1* & *N2*) and an adjective (*A*). Syntactic ambiguity is created by manipulating the adjective scope, with low or high syntactic attachment of the adjective, yielding 2 syntactic conditions and the following phrasing (with predicted structure’s depths):

**Condition1** - The adjective qualifies only the second noun  $[(N1)_p, [(N2)_p, A]]$ , predicting a **low syntactic attachment** of the *A* with an *Intermediate Phrase (ip)* boundary between *N1* and *N2* and a *Prosodic Word (pw)* boundary between *N2* and *A*.

**Condition2** - The adjective qualifies both nouns  $[(N1)_p, (N2)_p, [A]]$ , predicting a **high syntactic attachment** of *A* with an *Accentual Phrase (ap)* boundary between *N1* and *N2* plus an *Intermediate Phrase (ip)* boundary between *N2* and *A*.

The prosodic structure proposed here is composed of 3 levels under the *IP*: *ip*, *ap* and *pw*. While *ap* is largely accepted in French, *ip* is more controversial and *pw* is hardly ever mentioned. However, in earlier acoustic/perception studies [23; 24] the question of the relevance of the *pw* unit was raised. The present study aims at further investigating this finer granularity in the prosodic hierarchy. The corpus is also manipulated with regards to constituents’ length; nouns and adjectives increase from 1 to 4 syllables, in all possible combinations. *N1* and *N2* lengths always co-vary. The original corpus is composed of 4 sets of phrases. Each phrase is extracted from a carrier sentence, the syntactic structure of which enforces a major prosodic boundary after *A*, realized as a non-terminal, high *IP* boundary. Morphology and semantics are controlled so that the adjective can apply to both nouns (see [23] for more details on the corpus).

In this perception study, one set of phrases read by one female speaker was used: 32 phrases with 4 lengths of *N* and 4 lengths of *A* in the 2 different syntactic conditions. Our data analysis however excludes phrases with monosyllabic words, because they do not allow for the distinction between initial accents (*IA*) and final accents (*FA*). Our results were thus computed on 24 phrases, namely 12 phrases by condition.

### 2.2. Participants and experimental tasks

Each sequence *N1+N2+A* was perceptually judged by 27 French native listeners. They performed 3 perception tasks:

**Boundary task**: listeners had to judge the degree of break between each word (5 potential sites) on a scale from 0 (no break) to 3 (strong break).

**Prominence task**: listeners had to judge the degree of salience of each syllable of the phrases (ranging from 6 to 15 potential sites depending of the combination of constituents’ length), on a scale from 0 (no salience) to 3 (strong salience).

**Grouping task**: listeners had to judge how the sequences *N1+N2+A* are divided into groups by selecting between 4 different phrase groupings (*G1*, *G2*, *G3*, *G4*) intended to reflect the possible choices of prosodic phrasing (see Figure1).

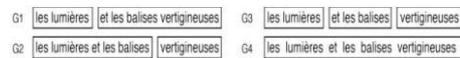


Figure 1: The 4 groups proposed in the Grouping task.

The order of presentation of the 3 tasks was counter-balanced between listeners. For each task, phrases were presented via headphones on a computer and their presentation was randomized across listeners. Participants could listen to the same phrase up to 5 times in order to perform their scoring, by pressing a ‘play’ button on the computer. A training period was undertaken before the start of the experiment, using 4 phrases from another set of *N1+N2+A* sequences than the one used for the present perception study.

### 2.3. Predictions

As exposed in the introduction, our perception study is designed to clarify the role of prosodic cues in the marking of morphosyntactic units.

In the first two perception tasks, boundaries and prominences are scored separately to capture the relative effect of these two distinct phonological events on phrasing. In the third task, groups are used to more specifically investigate prosodic phrasing strategies and to help capture the underlying factors explaining potential mismatches with syntax. *G1* and *G2* correspond to the predictions of the syntactic Conditions 1 and 2. By contrast, *G3* and *G4* were chosen to propose alternative groupings for Conditions 1 and 2, so as to help uncover potential finer-grained boundary levels. *G3* can be found in both syntactic predictions, but would reflect different relative boundary strengths in Condition1 (*N1* || *N2* | *A*) and 2 (*N1* | *N2* || *A*). *G4* may also correspond to both conditions but cannot reflect gradual boundary levels within the *IP*. *G3* and *G4* were also intended to get perceptual insight on previous acoustic results [23] on the same corpus, showing preferential marking of structure by *IA* over *FA*. *G3* will be used to perceptually test these findings, where *IA* was found to largely mark *N2* (83%) in Condition2 (after a predicted *ap* boundary), and marginally mark *A* (23%) in Condition1 (after a predicted *pw* boundary). *G4* is intended to test the potential perception of no boundary between constituents, despite the common *IA* marking of *N2* and *A*.

Ultimately, the choice of one grouping over another will be correlated to the results emerging from the boundary and prominence tasks, with the goal to clarify the respective or conjoined implication of these prosodic phenomena in speech structuration.

## 2.4. Data selection and analysis

The design of our data supposes that we first investigate listeners' interpretation of Conditions 1 and 2 in terms of the 4 grouping choices.

### 2.4.1. Syntactic condition effect on Groups

The relationship between *Conditions* and *Groups* was investigated using a chi-square test of independence. Results show a significant effect of *Condition* on *Group* ( $\chi^2(3,648)=530.89, p<.001$ ). Post-hoc analyses show that *G1* is highly associated to Condition1 (93.5% vs. *G2*: 0.6%, *G3*: 5.9% and *G4*: 0.0%; adjusted  $p$ -value<.001), while Condition2 is significantly more frequently associated to *G4* (42.9%), *G2* (32.1%) and *G3* (21.0%) than to *G1* (4.0%; adjusted  $p$ -value<.001). For lack of data, we thus decided to exclude from further analyses the cases in which Condition1 was perceived as *G2*, *G3* or *G4*, and those in which Condition2 was perceived as *G1*. Listeners are able to perceptively distinguish the two syntactic conditions: *G1* massively corresponds to Condition1, while *G2* is associated to Condition2. Finally, whereas we expected *G3* and *G4* to be distributed equally over the two conditions, they are mostly associated to Condition2.

### 2.4.2. Statistical models

On the basis of these results, we analyzed the effect of *Group* (*G1*, *G2*, *G3* and *G4*) on *Prominence* and *Boundary* scores (ranging from 0 to 3). Constituents' lengths were taken into account for *N* only, while *A* lengths were collided in order to increase statistical power. Analyses focused on the sites in the phrases reflecting structuring, i.e. between *N1|N2* and between *N2|A*. *Boundary* scores were measured after each content word, and *Prominence* scores were measured immediately before boundaries (*FA* on *N1* and *N2*) and at the beginning of content words after boundaries (*IA* on *N2* and *A*). Separate Mixed Linear Models were computed for dependent variables *Boundary* and *Prominence* respectively, and with subjects and phrases as random variables. The predictors were *Groups* and *N* length (2, 3 and 4 syllables). The non-significant variables or interactions were excluded from the final models.

## 3. Results

### 3.1. Boundaries

#### 3.1.1. *N1|N2* boundary scores

There was a main effect of *Group* ( $F(3,91.25)=494.14, p<.001$ ) and of *N* length ( $F(2, 37.54)=10.59, p<.001$ ), with an interaction between the two predictors ( $F(6,89.32)=4.04, p<.001$ ). Post-hoc analyses show significant differences ( $p<.05$ ) of boundary scores between groups, for all *N* lengths' conditions: the boundary in *G1* (mean score (*ms*) range: 3.38-3.62) is perceived as much stronger than in the other groups (*ms* ranges: *G2*= 1.44-1.83; *G3*=1.43-2.21; *G4*= 1.36-1.60). In addition, the boundary in *G3* for 2 and 4 syllables *N* is also perceived as stronger (*ms*: 2.10 and 2.21 respectively) than in *G2* (*ms*: 1.58 and 1.83 respectively) and *G4* (*ms*: 1.40 and 1.60 respectively). For 3 syllables *N*, however, no difference is observed between the boundaries in *G3* (*ms*: 1.43), *G2* (*ms*: 1.44) and *G4* (*ms*: 1.36). Altogether, these results indicate that listeners perceive 3 different boundary levels between *N1* and *N2*: the strongest in *G1*, an intermediate one in *G3* and the weakest one in *G2* and *G4*.

#### 3.1.2. *N2|A* boundary scores

Our analysis shows a main effect of *Group* ( $F(3, 111.64)=30.95, p<.001$ ). Post-hoc analyses show significant differences ( $p<.05$ ) for all comparisons except between *G2* and *G3*. The boundaries in *G3* (*ms*: 2.52) and *G2* (*ms*: 2.42) are perceived as much stronger than in *G4* (*ms*: 1.99), with the lowest boundary score for *G1* (*ms*: 1.77). Again, it seems that listeners perceived 3 different boundary levels between *N2* and *A*: the strongest in *G2* and *G3*, an intermediate one in *G4* and the weakest one in *G1*.

#### 3.1.3. Relative strength of *N1|N2* and *N2|A* boundaries

Prosodic phrasing is better accounted for when capturing the syntagmatic relationship between boundaries. This relationship was inferred in our data by subtracting *Boundary* scores perceived between *N2|A* to those perceived between *N1|N2*. Results show main effects of *Group* ( $F(3,106.96)=278.10, p<.001$ ) but no effect of *N* length. Post-hoc analyses show that the only significant contrasts ( $p<.001$ ) are between *G1* (*ms*: -1.78) and the other three groups (*ms*: *G2*= 0.96; *G3*= 0.47 and *G4*= 0.50). For *G1*, the second boundary (*N2|A*) is perceived as much weaker than the first boundary (*N1|N2*). By contrast, for *G2*, *G3* and *G4*, the second boundary (*N2|A*) is perceived as stronger than the first boundary (*N1|N2*). It seems that listeners perceive two types of syntactic/prosodic grouping with a different balance of boundary strengths: strong + weak boundaries ( $N1|N2 > N2|A$ ) in Condition1 associated to *G1*, and weak + strong boundaries ( $N1|N2 < N2|A$ ) in Condition2 associated to *G2*, *G3* and *G4*.

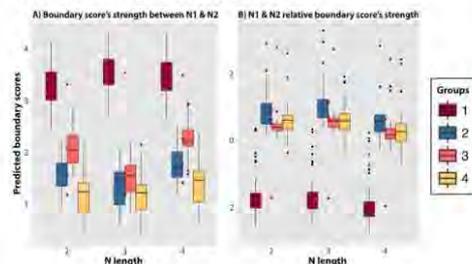


Figure 2: Perception results as a function of *N* length and *Groups* for A) *N1|N2* boundary scores' strength, and B) Relative strength of *N1|N2* and *N2|A* boundaries.

### 3.2. Prominences

#### 3.2.1. Final and Initial Accents

For *FA* on *N1*, our analysis shows main effects of *Group* ( $F(3,568.71)=6.24, p<.001$ ) and of *N* length ( $F(2,556.49)=8.14, p<.001$ ). Post-hoc analyses show that *FA* on *N1* in *G1* (*ms*: 2.39) is perceived as stronger ( $p<.035$ ) than in *G2* (*ms*: 2.09) and *G4* (*ms*: 2.17), while the difference with *G3* (*ms*: 2.13) doesn't reach significance ( $p=.087$ ). The other contrasts are not significant. As far as *N* length effect is concerned, *FA* on *N1* is globally perceived as significantly weaker ( $p<.014$ ) when *N* is composed of 3 syllables (*ms*: 2.03) as compared to nouns of 2 or 4 syllables (*ms*: 2.25 and 2.31 respectively). No difference is observed between *N* of 2 and 4 syllables. For *FA* on *N2*, our analysis shows no effect of *Group*. and a main

effect of *N* length ( $F(2,42.64)=3.87, p=.029$ ). Post-hoc analyses show that *FA* on *N2* is perceived as weaker when *N* is composed of 3 rather than 4 syllables (*ms*: 1.92 vs 2.14 respectively,  $p<.022$ ). For *IA* on *N2*, there are no significant main effects of *Group* and *N* length ( $p>.05$ ; *ms* range: 2.26-2.80). The same goes for *IA* on *A* (*ms* range: 2.00-2.72).

### 3.2.2. *IA* and *FA* relative prominence strengths

The question remains as to the relative weight of prominences in the marking of the structure. We computed a paired two-tailed t-test with *IA* and *FA* scores around the two boundary sites, for each group and independently from *N* length. Results show a significant difference between *FA* on *N1* and *IA* on *N2*: *IA* is perceived as stronger than *FA* in *G3* (*ms*: 1.63 vs. 1.07;  $t(67)=3.30, p=.002$ ) and in *G4* (*ms*: 1.56 vs. 1.07,  $t(138)=5.55, p<.001$ ). *IA* is also marginally stronger than *FA* in *G2* (*ms*: 1.34 vs. 1.11;  $t(103)=1.96, p=.053$ ). There is however no effect for *G1* ( $p=.338$ ). Our results also show a significant difference between *IA* on *A* and *FA* on *N2* for all groups: *IA* is perceived as stronger than *FA* in *G1* (*ms*: 1.23 vs. 0.92;  $t(302)=4.83, p<.001$ ), in *G2* (*ms*: 1.38 vs. 0.99;  $t(103)=3.30, p=.001$ ), in *G3* (*ms*: 1.45 vs. 0.82;  $t(67)=3.80, p<.001$ ) and in *G4* (*ms*: 1.42 vs. 1.04;  $t(138)=4.06, p<.001$ ).

## 4. Discussion & Conclusion

The aim of this study was to investigate prosodic phrasing in French, and more precisely the use of prosodic cues in the marking of morphosyntactic units. Indeed, as exposed in the introduction, there is no clear consensus on how many levels are necessary to reflect prosodic hierarchy, especially for French descriptions. Investigating prosodic phrasing is best done manipulating syntactic ambiguity, as shown in the seminal work by Lehiste [29; see also 28]. Our corpus uses low and high adjective attachment to two conjoined nouns in order to elicit prosodic disambiguation in French. It also manipulates constituents' length. If structure depth can be uncovered by means of acoustical analyses, some authors have however demonstrated the difficulty of matching acoustic features to syntactic structures [28]. Perceptual investigations of such a link might thus be a better gateway towards this end [25]–[27], a useful interface between acoustic cues and phonology. Our participants were asked to perform three separate perceptual judgments: two tasks (*Boundary* and *Prominence strengths*' scoring) were designed to investigate the role of these prosodic cues in structure marking. Our statistical design aimed at interpreting these results with regard to the *Grouping* task, in order to more specifically test their relationship to prosodic phrasing.

Taken together, our results indicate that listeners are able to distinguish the two intended syntactic conditions: while Condition1 was exclusively associated with *G1*, Condition2, however, was distributed across *G2*, *G3* and *G4*. Prosodic phrasing is thus perceived with different degrees of granularity in Condition2. This granularity could be investigated by testing the effect of each *Grouping* choices on *Boundary* and *Prominence* scores independently. Altogether, our results indicate that morphosyntactic structuring is better accounted for by *Boundary* than *Prominence* perception. Indeed, while *Grouping* choices very marginally affect *Prominence* perception, they largely explain *Boundary* scores. Indeed, *Grouping* choices indicate that listeners perceive various *Boundary* levels on each site (*N1* | *N2*; *N2* | *A*). However, a syntagmatic investigation indicates that listeners adjust

relative boundary strengths in order to maintain the intended syntactic structure: strong + weak boundary (*N1*/*N2* > *N2*/*A*) in Condition1 and weak + strong boundary (*N1*/*N2* < *N2*/*A*) in Condition2. Confronting the syntagmatic analysis with the analysis by site allows for finer interpretation of the granularity of these two boundary strengths: listeners perceive up to three boundary levels on each site below the *IP* level. Between *N1* and *N2*, the strongest boundary perceived in *G1* (*N1*/*N2* > *N2*/*A*) could reflect an *ip* boundary. The weaker one perceived in *G3* (*N1*/*N2* < *N2*/*A*) would correspond to an *ap* boundary. The lowest boundary is perceived in *G2* and *G4* (*N1*/*N2* < *N2*/*A*), implying potentially a *pw* boundary. Between *N2* and *A*, the strongest boundary perceived in *G2* and *G3* (*N1*/*N2* < *N2*/*A*) could reflect an *ip* boundary, and the weaker one perceived in *G4* (*N1*/*N2* < *N2*/*A*) could correspond to an *ap* boundary. The lowest boundary perceived in *G1* (*N1*/*N2* > *N2*/*A*) could also reflect a *pw* boundary. These results are particularly interesting when considering the alleged propositions for French prosodic constituency. In most theoretical descriptions, indeed, French accent is said to be post-lexical, which makes it impossible to recognize a boundary level lower than the *ap*, namely the *pw* level close to the lexical word [5; 6; 17]. Our results however indicate that such a constituency level is actually perceived: between *N2*/*A* in *G1* and between *N1*/*N2* in *G2* and *G4*.

Our investigations on *Prominence* also confirm these assumptions. Two main results emerge: first, *FA* was clearly used to mark structures perceived as *G1* at the *ip* boundary, which can be interpreted as an independent and "supporting role" [30] to the strong boundary scores found at this site. This result questions the proposition in [3] according to which *boundary tones* (*H%*) supersede *FA pitch accent* (*H\**) at higher prosodic boundaries. Namely, the syncretic occurrence of *FA* prominence and intonation boundaries does not block the independent perception of both these prosodic events. The second important result concerns *IA*. The perception of *IA* does not depend on *Grouping* choices. More surprisingly with regards to the literature in French, it is not linked to constituents' length, contradicting the hypothesis that *IA* essentially plays a rhythmic role in French (see among others [19; 21; 22]). Rather, our results indicate that *IA* is perceived as stronger than *FA*, with consistent high scores throughout the prosodic structure. Previous acoustic results suggested that *IA* preferentially marks the lowest *ap* level over *FA* [23]. The present perceptual results rather point towards preferential marking of the *pw*. Indeed, results show that *IA* after *pw* boundaries in *G1* (*N2*/*A*), and in *G2* and *G4* (*N1*/*N2*) was perceived as more salient than *FA* (situated before the *pw* boundary): here, *IA* is a left boundary marker of constituency, starting as early as the *pw* level.

We wish to elaborate on these results, which are particularly interesting to address the question of prosodic phrasing in French. New analyses are currently under way on a much larger database, using the same corpus (now 4 speakers) and the same tasks run on 80 listeners.

## 5. Acknowledgements

This study is supported by the Agence Nationale de la Recherche grant ANR-12-BSH2-0001 (PI: Corine Astésano) and by the Spanish Ministry of Economy and Competitiveness grant FFI2013-40419-P (PI: Lorraine Baqué). The authors also wish to thank Rafeu Sichel-Bazin for useful comments on an earlier version of this text.

## 6. References

- [1] M. Rossi, "Is Syntactic Structure Prosodically Retrievable," in *Proceedings of 5th European Conference of Speech Communication and Technology*, Rhodes, Greece, 1997, vol. 1, pp. 1–8.
- [2] M. Rossi, *L'intonation, le système du français: description et modélisation*. Paris, France: Editions OPHRYS, 1999.
- [3] S.-A. Jun and C. Fougeron, "Realizations of accentual phrase in French intonation," *Probus*, vol. 14, no. 1, pp. 147–172, 2002.
- [4] E. Selkirk, "On derived domains in sentence phonology," *Phonology*, vol. 3, pp. 371–405, May 1986.
- [5] M. Nespoulet and I. Vogel, *Prosodic phonology*. Foris: Dordrecht, 1986.
- [6] B. Hayes, "The prosodic hierarchy in meter," in *Rhythm and Meter*, Academic Press., P. Kiparsky and G. Youmans, Eds. Orlando, Florida, 1989, pp. 201–260.
- [7] B. Post, "Tonal and phrasal structures in French intonation," Thesis, The Hague, 2000.
- [8] D. R. Ladd and N. Campbell, "Theories of Prosodic Structure: Evidence from Syllable Duration," in *Proceedings of the 12th Congress of Phonetic Sciences*, 1991, vol. 2, pp. 290–293.
- [9] D. R. Ladd, *Intonational Phonology*. Cambridge, U.K: Cambridge University Press, 1996.
- [10] E. Selkirk, "Sentence prosody: intonation, stress and phrasing," in *The Handbook of Phonological Theory*. Blackwell., J. Goldsmith, Ed. London, UK, 1995, pp. 550–569.
- [11] R. Kager, *Optimality theory*. Cambridge, U.K: Cambridge University Press, 1999.
- [12] E. Selkirk and K. Tateishi, "Constraints on Minor Phrase formation in Japanese," in *Proceedings of the 24th Annual Meeting of the Chicago Linguistic Society*, 1988, pp. 316–336.
- [13] M. E. Beckman and J. B. Pierrehumbert, "Intonational structure in Japanese and English," *Phonology*, vol. 3, no. 01, pp. 255–309, May 1986.
- [14] S.-A. Jun and C. Fougeron, "A Phonological Model of French Intonation," in *Intonation: Analysis, Modeling and Technology*, Kluwer Academic Publishers, vol. 15, A. Botinis, Ed. Dordrecht: Springer Netherlands, 2000, pp. 209–242.
- [15] A. Di Cristo and D. Hirst, "Rythme syllabique, rythme mélodique et représentation hiérarchique de la prosodie du français," *Trav. Interdiscip. Lab. Parole Lang. Aix-En-Provence TIPPA*, no. 15, pp. 9–24, 1996.
- [16] A. Michelas and M. D'Imperio, "Durational cues and prosodic phrasing in French: evidence for the intermediate phrase," in *Speech Prosody 2010 100881:1-4*, Chicago, Illinois, 2010.
- [17] E. Selkirk, "The prosodic structure of function words," in *Signal to syntax: Prosodic bootstrapping from speech to grammar in early acquisition*, John J. Mc Carthy, Mahwah, NJ: Blackwell Publishing Ltd, 1996, pp. 187–213.
- [18] M. Rossi, "Le français, langue sans accent?," in *L'accent en français contemporain*, vol. 15, I. Fonagy and P. Léon, Eds. 1980, pp. 13–51.
- [19] J. Vaissière, "Rhythm, accentuation and final lengthening in French," in *Music, Language, Speech and Brain*, Macmillan Press, J. Sundberg, L. Nord, and R. Carlson, Eds. London, UK, 1991, pp. 108–120.
- [20] M. E. Beckman, "Evidence for Speech Rhythms across Languages," in *Speech Perception, Production and Linguistic Structure*, Y. Tohkura, E. vatikiotis-Bateson, and Y. Sagisaka, Eds. Tokyo, Japan, 1992, pp. 457–463.
- [21] V. Lucci, "Prosodie, phonologie et variation en français contemporain," *Lang. Fr.*, vol. 60, no. 1, pp. 73–84, 1983.
- [22] M. Rossi, "L'intonation et l'organisation de l'énoncé," *Phonetica*, vol. 42, no. 2–3, pp. 135–153, 1985.
- [23] C. Astésano, E. G. Bard, and A. Turk, "Structural influences on initial accent placement in French," *Lang. Speech*, vol. 50, no. Pt 3, pp. 423–446, 2007.
- [24] C. Astésano, R. Bertrand, R. Espesser, and N. Nguyen, "Dissociation between prominence and boundary phenomena in French: a perception study," in *Proceedings of pS-prominenceS conference*, Viterbo, Italie, 2013.
- [25] A. Di Cristo, "La prosodie au carrefour de la phonétique, de la phonologie et de l'articulation formes-fonctions," *Trav. Interdiscip. Lab. Parole Lang. Aix-En-Provence TIPPA*, vol. 23, pp. 67–211, 2004.
- [26] J. Cole, Y. Mo, and S. Baek, "The role of syntactic structure in guiding prosody perception with ordinary listeners and everyday speech," *Lang. Cogn. Process.*, vol. 25, no. 7–9, pp. 1141–1177, Sep. 2010.
- [27] J. Cole, T. Mahrt, and J. I. Hualde, "Listening for sound, listening for meaning: Task effects on prosodic transcription," in *Speech Prosody 2014*, Dublin, Ireland, 2014, pp. 858–863.
- [28] E. Delais-Roussarie and I. Feldhausen, "Variation in Prosodic Boundary Strength: a study on dislocated XPs in French," in *Speech Prosody 2014*, Dublin, Ireland, 2014, pp. 1052–1056.
- [29] I. Lehiste, "Phonetic Disambiguation of Syntactic Ambiguity," *J. Acoust. Soc. Am.*, vol. 53, no. 1, p. 380, 1973.
- [30] P. J. Price, M. Ostendorf, S. Shattuck-Hufnagel, and C. Fong, "The use of prosody in syntactic disambiguation," *J. Acoust. Soc. Am.*, vol. 90, no. 6, pp. 2956–2970, Dec. 1991.



## Annexe 20 :

### Article publié dans les actes de la conférence internationale *JEP-TALN-RECITAL 2016*.

#### **Constituance et phrasé prosodique en français : une étude perceptive.**

Laury Garnier<sup>1,2</sup> Corine Astésano<sup>1</sup> Lorraine Baqué<sup>3</sup> Anne Dagnac<sup>2</sup>

(1) URI Octogone-Lordat (E.A 4156), Université Toulouse 2 Jean-Jaurès, France

(2) Laboratoire CLLE-ERSS (UMR 5263), Université Toulouse 2 Jean-Jaurès, France

(3) Laboratori fLexSem, Universitat Autònoma de Barcelona, Espagne

[laury.garnier@univ-tlse2.fr](mailto:laury.garnier@univ-tlse2.fr), [astesano@univ-tlse2.fr](mailto:astesano@univ-tlse2.fr),  
[lorraine.baque@uab.cat](mailto:lorraine.baque@uab.cat), [dagnac@univ-tlse2.fr](mailto:dagnac@univ-tlse2.fr)

#### RESUME

---

L'objectif de cette étude est d'explorer l'organisation du phrasé prosodique en français. Il n'existe pas de consensus clair sur le nombre de niveaux nécessaires pour refléter la hiérarchie prosodique de la langue. Dans ce cadre, nous proposons une étude perceptive, via un corpus de parole contrôlée manipulant des structures syntaxiques ambiguës, où 27 participants ont effectué 3 tâches de perception : proéminence, frontière et groupement. Nos résultats montrent une utilisation privilégiée des indices de frontières dans le marquage des groupes prosodiques. Plus précisément, on observe que les auditeurs sont capables de percevoir des niveaux de granularité de frontières plus fins que ce que les descriptions traditionnelles du français prédisent. Par ailleurs, les résultats de la tâche de proéminence montrent que l'accent initial est toujours perçu plus fort que l'accent final, et ce dès les niveaux les plus bas de la hiérarchie.

#### ABSTRACT

---

##### **Prosodic constituency and phrasing in French: a perception study**

The aim of the present study is to investigate the organization of prosodic phrasing in French. There is no clear consensus on how many levels are necessary to reflect the prosodic hierarchy in this language. In this context, we propose a perception study on a corpus manipulating syntactically ambiguous structures, where 27 participants had to perform 3 distinct perceptual tasks: prominence, boundary and grouping tasks. Our results show a preferential use of boundary cues in prosodic groups' marking. More precisely, we observe that listeners are able to distinguish finer-grained grouping levels than those predicted in traditional French descriptions. Moreover, the results of the prominence task show that initial accents are always perceived stronger than final accents, even at the lowest levels of the prosodic hierarchy.

**MOTS-CLES :** phrasé prosodique, perception, proéminence, frontière, groupement, français.

**KEYWORDS:** prosodic phrasing, perception, prominence, boundary, grouping, French.

---

## **1 Introduction**

Les indices prosodiques, tels que les proéminences et les frontières, vont segmenter le flux de parole en groupes de mots pour ainsi faciliter la compréhension du message ; c'est ce qu'on appelle le phrasé prosodique. Là intervient alors le lien entre la prosodie et d'autres niveaux linguistiques,

notamment le niveau syntaxique. Un certain nombre de travaux se sont intéressés au lien à établir entre la structure prosodique et la structure syntaxique. Au-delà de la question du mapping entre les deux domaines, il apparaît difficile de trouver un consensus sur le choix des niveaux qui reflètent la hiérarchie prosodique. En effet, différentes hiérarchies prosodiques ont été proposées dans la littérature, variant de 2 niveaux (*syntagme accentuel* (ci-après *ap*) et *syntagme intonatif* (ci-après *IP*): Pierrehumbert, 1980, et Jun & Fougeron, 2002 pour le français) à 5 niveaux (incluant, selon les modèles : *mot prosodique* (ci-après *pw*), *groupe clitique*, *syntagme phonologique* (ci-après *PP*), *syntagme intonatif* et *énoncé* ; Selkirk, 1986 ; Nespor & Vogel, 1986 ; Post, 2000, notamment). Certains auteurs envisagent également la possibilité d'unités prosodiques récursives, dans le but de mieux rendre compte du lien entre prosodie et syntaxe. Parmi ces propositions, on retrouve le « *Super-Major Phrase* » (*super-syntagme majeur* : une *IP* pouvant contenir elle-même plusieurs *IP*) servant de niveau intermédiaire entre l'*énoncé* et le *syntagme majeur* (Ladd, 1996). De la même manière, Selkirk & Tateishi (1988) proposent de diviser le niveau du *PP* en 2 niveaux distincts : le *PP mineur* (équivalent à l'*ap*) et le *PP majeur* (équivalent au *syntagme intermédiaire* (ci-après *ip*)) proposé par Beckman & Pierrehumbert, 1986). Plus récemment, ce principe de récursivité a été observé via la révision de la *Strict Layer Hypothesis* dans le cadre de la *Théorie de l'Optimalité* (Selkirk, 1995 ; Kager, 1999). L'existence d'un niveau intermédiaire a également été discutée pour le français. Quand certains auteurs considèrent l'*ip* comme une unité propre à des structures syntaxiques spécifiques, telles que les « *question tags* » ou les structures disloquées (Jun & Fougeron, 2000 ; voir aussi le *segment d'unité intonative* de Di Cristo & Hirst, 1996), d'autres, en revanche, proposent l'existence d'un « vrai » niveau intermédiaire entre le niveau de l'*ap* et de l'*IP*, qui permet de rendre compte non plus d'un type de structure syntaxique particulier mais davantage d'une unité de niveau supérieur à l'*ap*, fonction de la taille des constituants (Michelas & d'Imperio, 2010, en lien avec les propositions de Beckman & Pierrehumbert, 1986).

Pour le français, cette question autour des niveaux hiérarchiques est particulièrement intéressante. Alors que dans les contours intonatifs de l'anglais les prééminences se distinguent clairement des mouvements tonaux associés aux frontières, le français, au contraire, est caractérisé par un syncrétisme entre les tons de frontière et les prééminences finales (ci-après *AF*), tous deux se manifestant sur la dernière syllabe des constituants prosodiques. Également, à l'inverse d'autres langues, l'accent en français est dit post-lexical et marquerait de manière privilégiée le niveau de l'*ap* plutôt que le niveau du *pw*<sup>1</sup>. Ces caractéristiques du français ont d'ailleurs poussé certains auteurs à parler de *langue sans accent* (Rossi, 1980) ou de *langue de frontière* (Vaissière, 1991). Plus récemment, dans le cadre de l'approche métrique autosegmentale, *AF* en français est décrit comme un *accent mélodique* (*pitch accent* : *H\**), marquant le plus petit niveau de la hiérarchie prévu en français (*ap*). Dans cette conception, *AF* s'effacerait au profit du seul contour intonatif (*boundary tone* : *H%*) à un haut niveau de la hiérarchie prosodique : l'*IP* (Jun & Fougeron, 2002). Il existe également un autre type d'accent en français marquant le bord gauche de l'*ap* : il s'agit de l'accent initial (ci-après *AI*). Cet accent est dit secondaire et optionnel, *AF* étant primaire. Traditionnellement, *AI* est décrit comme un marqueur rythmique dont l'apparition dépend de la taille des constituants (Vaissière, 1991 ; Rossi, 1985). Toutefois, la description du rôle et des fonctions de cet accent dans le marquage même de la hiérarchie prosodique reste encore aujourd'hui mal établie. En effet, Astésano *et al.* (2007) se sont intéressés aux fonctions de *AI* via l'analyse acoustique d'un corpus contrôlé, manipulant à la fois la taille des constituants et la structure syntaxique. Dans les résultats de cette étude, *AI* se révèle être un marqueur plus structurel que purement rythmique ; les auteurs montrent que *AI* marque la structure prosodique de manière privilégiée par rapport à *AF*,

<sup>1</sup> *pw* est parfois compris comme l'équivalent du *groupe clitique* de Nespor & Vogel (1986), donc similaire à l'*ap*. Nous entendons ici *pw* dans l'acception de Selkirk (1996) qui correspond au *mot lexical*, les *clitiques* étant rattachés au niveau immédiatement supérieur.

spécifiquement à des niveaux inférieurs de la hiérarchie prosodique : le niveau de l'*ap*, voire le niveau du *pw*. Également, une étude en perception sur le même corpus confirme ces résultats (Astésano *et al.*, 2012). Dans cette dernière, les auteurs montrent que les auditeurs francophones sont capables de percevoir les proéminences finales (*AF*) indépendamment des frontières intonatives, et que *AI* est perçu systématiquement plus fort que *AF* à tous les niveaux de la hiérarchie prosodique.

Dans le prolongement de ces précédents travaux (Astésano *et al.* 2007 et 2012), nous proposons une étude perceptive, plus spécifiquement destinée à aborder la question des niveaux hiérarchiques de la structure prosodique en français. À travers la perception des francophones natifs, nous nous intéressons à la manière dont les indices de proéminences et de frontières sont utilisés pour instancier la structure morphosyntaxique. Plus précisément, cette étude vise à éprouver la hiérarchie à 3 niveaux (*ap*, *ip*, *IP*) proposée pour le français, sur la base de structures syntaxiques très contrôlées, afin de tester si ce fonctionnement à 3 niveaux est suffisant pour décrire le phrasé prosodique en français. Par ailleurs, la perception est utilisée ici comme une interface entre le niveau acoustique et le niveau phonologique permettant de mettre en évidence les indices pertinents dans le traitement du phrasé prosodique (comme proposé par Di Cristo, 2004 ; Cole *et al.*, 2010). Enfin, le prisme de la perception nous permet de contourner le problème de la variabilité ou « flexibilité » des indices acoustiques instanciant les constituants de même niveau dans la hiérarchie prosodique (Rossi, 1997 et 1999 ; Delais-Roussarie & Feldhausen, 2014).

## 2 Matériel linguistique et procédure expérimentale

Le matériel linguistique utilisé pour cette étude de perception est issu du *Corpus d'Edimbourg*, à la base de l'étude d'Astésano *et al.* (2007). Le corpus se constitue de structures syntaxiquement ambiguës, composées de deux noms coordonnés (*N1* et *N2*) et d'un adjectif (*A*), que les indices prosodiques (frontières et proéminences) aident à désambiguïser. Les deux structures syntaxiques possibles sont créées en manipulant la portée de l'adjectif : alors qu'en Condition 1, l'adjectif qualifie uniquement *N2* [*les lumières*] [*et les balises vertes*], en Condition 2, l'adjectif qualifie à la fois *N1* et *N2* [*les lumières et les balises*] [*vertes*]. Ces séquences *N1+N2+A* ont été placées dans un contexte phrastique plus large forçant ainsi la réalisation d'une frontière prosodique forte (frontière d'*IP*) après le syntagme cible. Selon la condition syntaxique, une structure prosodique différente est prédite, comprenant 3 niveaux au-dessous du niveau de l'*IP* : *ip*, *ap* et *pw*.

- **Condition 1** : [ { ( les / lumières /<sub>pw</sub> )<sub>ap</sub> }<sub>ip</sub> { ( et les / balises /<sub>pw</sub> / vertes /<sub>pw</sub> )<sub>ap</sub> }<sub>ip</sub> ]<sub>IP</sub>
- **Condition 2** : [ { ( les / lumières /<sub>pw</sub> )<sub>ap</sub> ( et les / balises /<sub>pw</sub> )<sub>ap</sub> }<sub>ip</sub> { ( / vertes /<sub>pw</sub> )<sub>ap</sub> }<sub>ip</sub> ]<sub>IP</sub>

Alors que le niveau de l'*ap* est communément accepté pour le français, celui de l'*ip* est controversé et le *pw* n'est même presque jamais mentionné. Cependant, la question de la pertinence de ce dernier niveau d'analyse avait déjà été soulevée dans les études acoustiques et perceptives précédentes d'Astésano *et al.* (2007 et 2012). La présente étude vise à approfondir l'investigation de cette granularité plus fine de la hiérarchie prosodique en français. Le corpus manipule également la taille des constituants (mots de 1 à 4 syllabes) et se compose de 4 sets de séquences *N1+N2+A*, lues par 8 locuteurs différents. Un test de jugement sémantique a été réalisé sur toutes les structures afin de ne garder que les locuteurs parvenant à instancier les deux conditions syntaxiques (cf. Astésano *et al.*, 2007 pour plus de détails sur la constitution du corpus). Pour cette étude perceptive, un seul des 4 sets lu par 1 locutrice a été utilisé : 32 syntagmes (4 longueurs de *N* \* 4 longueurs de *A* \* 2 conditions syntaxiques). Toutefois, nous avons exclu de notre analyse de données les séquences composées de *N* monosyllabiques puisqu'elles ne permettaient pas de distinguer *AI* et *AF*. Nos

résultats concernent donc 24 structures syntaxiquement ambiguës, soit 12 séquences par condition. Chacun de ces syntagmes a été jugé perceptivement par 27 auditeurs francophones natifs. Les auditeurs effectuaient 3 tâches de perception (cf. Figure 1) : une *tâche de proéminence* où ils devaient juger le degré de mise en relief de chaque syllabe du syntagme (variant de 6 à 15 sites potentiels selon la combinaison du nombre de syllabes des constituants), sur une échelle d'évaluation allant de 0 à 3 ; une *tâche de frontière* où les auditeurs devaient juger le degré de rupture entre chaque mot (5 sites potentiels) sur une échelle d'évaluation allant de 0 à 3 ; une *tâche de groupement* où ils devaient juger la manière dont les mots se regroupaient les uns avec les autres, parmi 4 propositions de groupements.

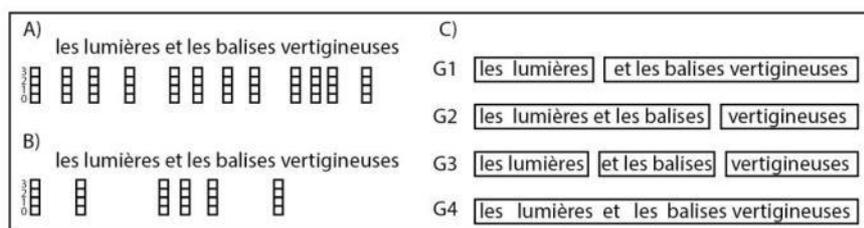


FIGURE 1 : A) Tâche de proéminence ; B) Tâche de frontière ; C) Tâche de groupement.

Chaque participant était muni d'un casque audio, confortablement assis devant un ordinateur, et pouvait écouter jusqu'à 5 fois le même syntagme. Avant chaque tâche, les participants effectuaient une phase d'entraînement, comprenant 4 séquences issues d'un set différent de celui utilisé dans cette étude. L'ordre de présentation des 3 tâches était contrebalancé entre tous les participants.

Notre étude de perception vise à préciser le rôle des indices prosodiques dans le marquage des unités morphosyntaxiques. Dans les deux premières tâches, proéminences et frontières sont volontairement jugées séparément, afin de rendre compte du rôle distinct de ces deux événements phonologiques dans le marquage de la structure. La troisième tâche de groupement sert plus spécifiquement à questionner les stratégies du phrasé prosodique, mais aussi à mettre en évidence les facteurs sous-jacents expliquant, notamment, les variations de correspondance potentielles avec la syntaxe. Alors que *G1* et *G2* correspondent aux deux conditions syntaxiques de départ (respectivement Condition 1 et Condition 2), *G3* et *G4* ont été choisis pour proposer une alternative à ces deux groupements, en permettant de faire ressortir des degrés de granularité de niveaux de frontières plus ou moins fins. *G3* satisfait à la fois les deux prédictions syntaxiques, bien qu'on attende néanmoins différentes forces relatives de frontières en Condition 1 (*NI || N2 | A*) et en Condition 2 (*NI | N2 || A*). *G4* peut également refléter les deux conditions syntaxiques, mais ne permet pas de mettre en évidence des frontières internes graduées à l'intérieur de l'*IP*.

L'analyse des scores de perception issus de ces 3 tâches suit 2 grandes étapes : une première phase consiste à tester l'adéquation des prédictions associées aux deux conditions syntaxiques avec les résultats de perception des groupements, en observant la manière dont les deux conditions sont distribuées à travers *G1*, *G2*, *G3* et *G4*. Dans une seconde étape, les résultats de groupements sont mis en relation avec les scores de frontières et de proéminences séparément, afin de faire émerger l'implication respective, mais aussi conjointe, de ces phénomènes prosodiques dans la structuration de la parole. Pour cette seconde analyse, la taille des constituants est prise en compte uniquement pour les noms, alors que pour les adjectifs elles sont regroupées afin d'augmenter le pouvoir statistique. Nous choisissons d'effectuer des analyses sur deux sites particuliers, points clés de la structure (cf. Figure 2) : entre *N1* et *N2* (*NI|N2*) et entre *N2* et *A* (*N2|A*). Pour chacun de ces sites,

nous relevons le score de frontière perçue entre chaque mot, ainsi que les scores de proéminences perçues sur les syllabes positionnées directement avant (*AF* potentiels sur *N1* et *N2*) et après (*AI* potentiels sur *N2* et *A*) la frontière (hors mots grammaticaux).



FIGURE 2 : Sites d'intérêt : *N1|N2* et *N2|A*.

Sur la base de ces données, nous construisons un modèle linéaire mixte avec le score de groupement et le nombre de syllabes du nom comme prédicteurs, et les participants et les items lexicaux comme variables aléatoires. Nous effectuons une analyse séparée pour chaque variable dépendante : le score de proéminence d'une part, et le score de frontière d'autre part. Les variables ou interactions non-significatives sont progressivement exclues du modèle final.

### 3 Résultats et discussions

#### 3.1 Effet de la condition syntaxique sur la perception du groupement

Un test  $\chi^2$  d'indépendance montre un effet significatif de la condition syntaxique sur le groupement ( $\chi^2(3,648)=530.89, p<.001$ ). Les analyses post-hoc indiquent que la Condition 1 est fortement associée à *G1* (93.5% vs. *G2* : 0.6%, *G3* : 5.9% et *G4* : 0.0% ; valeur de *p* ajustée <.001), alors que la Condition 2 est plus fréquemment associée à *G4* (42.9%), *G2* (32.1%) et *G3* (21.0%) qu'à *G1* (4.0% ; valeur de *p* ajustée <.001). À l'issue de ce test, nous décidons d'exclure de nos futures analyses les cas trop marginaux où la Condition 1 a été perçue comme *G2*, *G3* et *G4*, ainsi que les cas où la Condition 2 a été perçue comme *G1*. Nos résultats indiquent que les auditeurs sont capables de distinguer les deux conditions syntaxiques comme prédit : *G1* n'est quasiment pas associé à la Condition 2 et *G2* n'est pas perçue pour la Condition 1. En revanche, alors qu'on pouvait s'attendre à une répartition relativement équivalente de *G3* et *G4* sur les deux conditions, nous observons que ces deux groupements sont massivement associés à la Condition 2. Différents degrés de granularité dans le phrasé prosodique ont donc été perçus dans le cadre de cette dernière condition. La mise en relation de la perception des groupements avec les scores de frontières et de proéminences nous permet d'éclairer ce premier résultat (cf. infra).

#### 3.2 Effets du groupement et de la longueur des *N* sur les scores de frontières

##### A) Force des frontières entre *N1|N2* (cf. Figure 3-A) :

Les résultats montrent un effet significatif du groupement ( $F(3,91.25)= 494.14, p<.001$ ) et du nombre de syllabes du nom ( $F(2,37.54)= 10.59, p<.001$ ), ainsi qu'une interaction entre ces deux prédicteurs ( $F(6,89.32)= 4.04, p<.001$ ). Les tests post-hoc montrent des différences significatives ( $p<.05$ ) des scores de frontières entre les différents types de groupements, quel que soit le nombre de syllabes du nom : la frontière en *G1* (étendue du score moyen (*sm*) : 3.38–3.62) est perçue beaucoup plus forte que dans les autres groupements (étendue du *sm* pour *G2* : 1.44–1.83 ; *G3* : 1.43–2.21 ; *G4* : 1.36–1.60). Également, la frontière en *G3* pour les noms de 2 et 4 syllabes est perçue plus forte

(*sm* respectifs : 2.10 et 2.21) qu'en *G2* (*sm* respectifs : 1.58 et 1.83) et *G4* (*sm* respectifs : 1.40 et 1.60). Cependant, pour les noms de 3 syllabes, nous n'observons pas de différences significatives (*sm* pour *G3* : 1.43 ; *G2* : 1.44 ; *G4* : 1.36). À l'issue de cette analyse, il semble donc que les auditeurs perçoivent 3 niveaux de frontière différents entre *NI|N2* : la frontière la plus forte en *G1*, intermédiaire en *G3*, et la plus faible en *G2* et *G4*.

**B) Force des frontières entre *N2|A* (cf. Figure 3-B):**

Nous observons un effet significatif du groupement ( $F(3,111.64) = 30.95, p < .001$ ) seulement. Les tests post-hoc montrent une différence significative ( $p < .05$ ) entre tous les groupements, excepté entre *G2* et *G3*. La frontière en *G1* (*sm* : 1.77) est perçue plus faible qu'en *G4* (*sm* : 1.99), qui, elle-même, est perçue plus faible qu'en *G2* (*sm* : 2.42) et *G3* (*sm* : 2.52). Les auditeurs semblent donc avoir perçu 3 niveaux de frontières différents entre *N2|A* : la frontière la plus forte en *G2* et *G3*, intermédiaire en *G4*, et la plus faible en *G1*.

Alors que la taille des constituants n'a pas d'effet sur le score des frontières, le type de groupement, en revanche, explique le degré de force de ces frontières. Néanmoins, ces analyses par sites (*NI|N2* et *N2|A*) ne permettent pas de rendre compte du rapport hiérarchique entre les frontières des deux sites. Nous proposons alors une analyse syntagmatique de la force relative des frontières en soustrayant le score de frontière perçue entre *NI|N2* à celui perçue entre *N2|A*.

**C) Force relative des frontières entre les 2 sites (*N2|A* - *NI|N2*) (cf. Figure 3-C):**

Les résultats montrent un effet significatif du groupement ( $F(3,106.96) = 278.10, p < .001$ ) seulement. Les tests post-hoc indiquent que les seuls contrastes significatifs ( $p < .001$ ) sont ceux entre *G1* (*sm* : -1.78) et les 3 autres groupements (*sm* pour *G2* : 0.96 ; *G3* : 0.47 ; *G4* : 0.50). En effet, pour *G1* seulement, la frontière sur le second site (*N2|A*) est perçue plus faible que celle sur le premier site (*NI|N2*). À l'inverse, pour *G2*, *G3* et *G4*, la frontière sur le second site (*N2|A*) est perçue plus forte que celle sur le premier site (*NI|N2*). Les résultats montrent donc que les auditeurs perçoivent deux types de hiérarchies prosodiques, avec une force relative des frontières différente, correspondant aux prédictions des deux conditions syntaxiques : rapport fort + faible pour la Condition 1 représentée ici par *G1* (*NI|N2* > *N2|A*) et rapport faible + fort pour la Condition 2 associée ici à *G2*, *G3* et *G4* (*NI|N2* < *N2|A*).

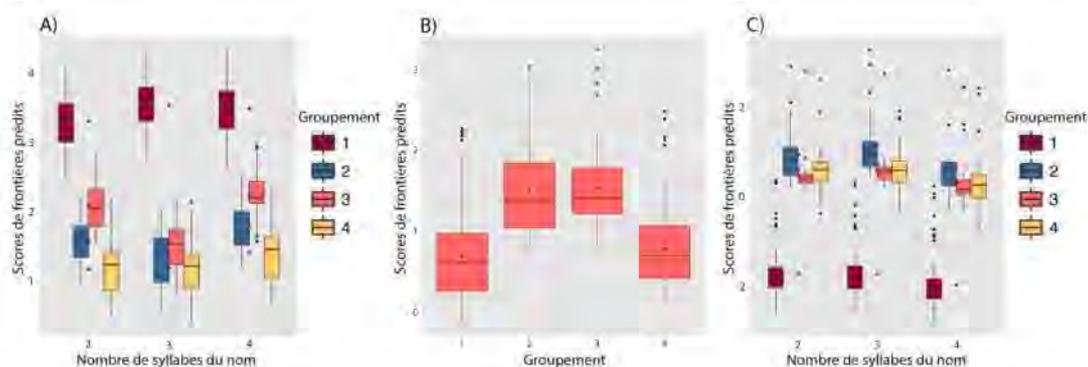


FIGURE 3: A) scores de la force des frontières entre *NI|N2* en fonction du groupement et du nombre de syllabes du nom ; B) scores de la force des frontières entre *N2|A* en fonction du groupement ; C) scores relatifs des frontières obtenus après soustraction des scores de frontières entre *N2|A* et *NI|N2*, en fonction du groupement et du nombre de syllabes du nom.

Nous observons une flexibilité dans la perception des frontières intra-site, expliquée par le type de groupement (cf. résultats A et B). Néanmoins, à travers l'analyse syntagmatique (cf. résultats C), on voit que la force relative des frontières inter-sites perçue par les auditeurs s'ajuste à la hiérarchie structurelle prédite par la syntaxe. Cette dernière analyse met également en évidence la nécessité de reconsidérer le nombre de niveaux hiérarchiques proposés pour la description du phrasé prosodique en français. En effet, on observe que selon la condition syntaxique, nous n'avons pas la même force relative entre les deux sites de frontières : rapport fort + faible en Condition 1 et rapport faible + fort en Condition 2. Nos résultats par sites nous permettent ensuite d'affiner l'interprétation de la granularité de ces 2 niveaux de frontières. En effet, il apparaît que les auditeurs perçoivent, sur chaque site, jusqu'à 3 niveaux de frontières inférieurs à l'*IP*. Entre *NI* et *N2*, la frontière perçue la plus forte en *G1* ( $NI|N2 > N2|A$ ) pourrait correspondre à une frontière d'*ip*, alors que la frontière perçue un peu plus faible en *G3* ( $NI|N2 < N2|A$ ) correspondrait à une frontière d'*ap*. Une frontière encore plus faible est perçue en *G2* et *G4* ( $NI|N2 < N2|A$ ), impliquant potentiellement une frontière de *p<sub>w</sub>*. Entre *N2* et *A*, la frontière la plus forte perçue en *G2* et *G3* ( $NI|N2 < N2|A$ ) pourrait correspondre à une frontière d'*ip*, la frontière perçue un peu plus faible en *G4* ( $NI|N2 < N2|A$ ) correspondrait à une frontière d'*ap*. La frontière perçue encore plus faible en *G1* ( $NI|N2 > N2|A$ ) impliquerait potentiellement, comme pour les résultats précédents, une frontière de *p<sub>w</sub>*. Ce résultat est particulièrement intéressant au regard des propositions faites sur les niveaux de constituance prosodique du français. Dans la plupart des approches théoriques, en effet, l'accent en français serait post-lexical, ce qui ne permet pas d'envisager un marquage de frontière à un niveau inférieur à l'*ap*, i.e. au niveau du *p<sub>w</sub>* proche du mot lexical chez Selkirk (1996). Or, il semble qu'il y ait bien un tel degré de niveau de frontière perçu par les auditeurs : entre *N2|A* en *G1* et entre *NI|N2* en *G2* et *G4*.

### 3.3 Effets du groupement et de la longueur des *N* sur les scores de proéminences

Pour les *AF* (sur *NI* et *N2*), nous notons un effet significatif du groupement ( $F(3,568.71) = 6.24$ ,  $p < .001$ ) seulement (effet trop marginal de la longueur des *N*) et ce, uniquement pour *AF* sur *NI*. Les tests post hoc indiquent que *AF* sur *NI* est perçu plus fort en *G1* ( $sm : 2.39$ ,  $p < .035$ ) qu'en *G2* et *G4* ( $sm$  respectifs : 2.09 et 2.17). Ce résultat montre donc qu'en *G1*, *AF* sur *NI* est clairement perçu sur le plan métrique, indépendamment de l'indice de frontière dans le marquage de la structure, et notamment ici dans le marquage de la frontière d'*ip*. Ce dernier point remet en question certaines propositions (Jun & Fougeron, 2002) selon lesquelles l'accent mélodique (*pitch accent* :  $H^*$ ) disparaîtrait au profit du ton de frontière (*boundary tone* :  $H^\%$ ) à un haut niveau de la hiérarchie prosodique. Concernant les *AI*, on n'observe ni d'effet du groupement, ni de la longueur des *N* ( $p > .05$ ). Ce résultat questionne notamment le rôle purement rythmique de *AI* que les descriptions traditionnelles s'accordent à lui attribuer (Vaissière, 1991 ; Rossi, 1985). Pour aller plus loin, on propose alors de tester la contribution relative des proéminences initiales et finales dans le marquage des structures. Nous avons effectué un test de Student avec les scores d'*AI* et d'*AF* autour des deux sites *NI|N2* (*AF* sur *NI* et *AI* sur *N2*) et *N2|A* (*AF* sur *N2* et *AI* sur *A*), pour chaque groupement, sans tenir compte du nombre de syllabes du nom. Les résultats montrent une différence significative entre *AF* sur *NI* et *AI* sur *N2* : *AI* est perçu plus fort que *AF* en *G3* ( $sm : 1.63$  vs. 1.07,  $t(67) = 3.30$ ,  $p = .002$ ) et en *G4* ( $sm : 1.56$  vs. 1.07,  $t(138) = 5.55$ ,  $p < .001$ ). Également, *AI* est marginalement plus fort que *AF* en *G2* ( $sm : 1.34$  vs. 1.11,  $t(103) = 1.96$ ,  $p = .053$ ). En revanche, il n'y a pas d'effet pour *G1* ( $p = .338$ ). De la même manière, *AI* sur *A* est perçu significativement plus fort que *AF* sur *N2* pour tous les groupements ( $p < .001$ ). Il semble donc que *AI* soit plus saillant perceptivement que *AF*, quel que soit le nombre de syllabes, et quel que soit le niveau dans la hiérarchie prosodique. Il est intéressant de noter que *AI* marque un niveau inférieur à l'*ap*. En effet, *AI* réalisé après la frontière du *p<sub>w</sub>* (entre *N2|A* en *G1*, entre *NI|N2* en *G2* et *G4*) est perçu plus fort que *AF* situé avant cette même frontière : *AI* sert donc ici à marquer la frontière gauche des *p<sub>w</sub>*.

## 4 Conclusion

L'objectif de ce travail était d'explorer, à travers une étude en perception, l'organisation du phrasé prosodique en français, et plus précisément l'utilisation des indices prosodiques dans le marquage de la structure. En effet, comme nous l'avons déjà exposé dans l'introduction, il n'existe pas de consensus clair sur le nombre de niveaux hiérarchiques nécessaires pour la description de la structure prosodique, et tout particulièrement pour la description du français. Deux résultats majeurs émergent : les participants utilisent de manière privilégiée les indices de frontières dans le marquage de la structure. Plus intéressant encore, les auditeurs sont capables de percevoir des niveaux de granularité de frontières plus fins que ce que les descriptions traditionnelles du français prédisent. En effet, nos résultats mettent en évidence la nécessité de prendre en compte un niveau supplémentaire dans la hiérarchie prosodique du français, inférieur à l'*ap* : le niveau du *piv*. Nos analyses sur les scores de proéminences vont également dans ce sens, en montrant que les *AI* sont plus saillants perceptivement que les *AF* et ce, dès le niveau du *piv*. Ces résultats, particulièrement intéressants sur la question des niveaux hiérarchiques nécessaires dans la structure prosodique du français, ouvrent la voie à de nouvelles analyses qui seront effectuées sur une plus large base de données, comprenant les mêmes syntagmes lus par 4 locuteurs différents (au total : 128 syntagmes) et entendus par 80 sujets. Avec un pouvoir statistique plus important, nous pourrons alors tester ces premiers résultats, en complétant notamment l'analyse des scores de proéminences avec d'autres sites accentuels (*AI* sur *NI* ; *AF* sur *A*), mais également envisager des analyses acoustiques afin de comparer la réalité acoustique et perceptive des niveaux de constituance prosodique du français.

## Remerciements

Cette étude a été réalisée dans le cadre du projet ANR-12-BSH2-0001 (IP: Corine Astésano) et soutenue par le ministère espagnol de l'économie et de la compétitivité FFI2013-40419-P (IP: Lorraine Baqué). Nous tenons également à remercier Rafèu Sichel-Bazin, avec qui nous avons pu partager de fructueuses discussions sur la constituance en français.

## Références

- ASTESANO C., BARD E.G., TURK A. (2007). Structural influences on initial accent placement in French. *Language and Speech*, 50(3), 423–446.
- ASTESANO C., BERTRAND R., ESPESSER R., NGUYEN N. (2012). Perception des frontières et des proéminences en français. Actes des *JEP-TALN-RECITAL*, Grenoble, 353-360.
- BECKMAN M.E., PIERREHUMBERT J.B. (1986). Intonational structure in Japanese and English. *Phonology*, 3(01), 255–309.
- COLE J., MO Y., BAEK S. (2010). The role of syntactic structure in guiding prosody perception with ordinary listeners and everyday speech. *Lang. Cogn. Process.*, 25(7–9), 1141–1177.
- DELAIS-ROUSSARIE E., FELDHAUSEN I. (2014). Variation in Prosodic Boundary Strength: a study on dislocated XPs in French. *Speech Prosody 2014*, Dublin, 1052–1056.
- DI CRISTO A. (2004). La prosodie au carrefour de la phonétique, de la phonologie et de l'articulation formes-fonctions. *TIPA*, 23, 67–211.

- DI CRISTO A., HIRST D. (1996). Rythme syllabique, rythme mélodique et représentation hiérarchique de la prosodie du français. *TIPA*, 15, 9–24.
- JUN S-A., FOUGERON C. (2000). A Phonological Model of French Intonation. *Intonation: Analysis, Modeling and Technology*, Kluwer Academic Publishers, 15, 209–242.
- JUN S-A., FOUGERON C. (2002). Realizations of accentual phrase in French intonation. *Probus*, 14(1), 147–172.
- KAGER R. (1999). *Optimality theory*. Cambridge : Cambridge University Press.
- LADD D.R. (1996). *Intonational Phonology*. Cambridge : Cambridge University Press.
- MICHELAS A., D'IMPERIO M. (2010). Durational cues and prosodic phrasing in French: evidence for the intermediate phrase. *Speech Prosody 2010*, Chicago, 100881:1-4.
- NESPOR M., VOGEL I. (1986). *Prosodic phonology*. Foris: Dordrecht.
- PIERREHUMBERT J.B. (1980). *The phonology and phonetics of English intonation*. Thesis, Massachusetts Institute of Technology, United States.
- POST B. (2000). *Tonal and phrasal structures in French intonation*. Thesis, The Hague.
- ROSSI M. (1980). Le français, langue sans accent? In *L'accent en français contemporain*, 15, I. Fonagy and P. Léon, Eds., 13–51.
- ROSSI M. (1985). L'intonation et l'organisation de l'énoncé. *Phonetica*, 42 (2–3), 135–153.
- ROSSI M. (1997). Is Syntactic Structure Prosodically Retrievable. *Proceedings of 5th European Conference of Speech Communication and Technology*, Greece, 1, 1–8.
- ROSSI M. (1999). *L'intonation: le système du français : description et modélisation*. Paris, France: Editions OPHRYS.
- SELKIRK E. (1986). On derived domains in sentence phonology. *Phonology*, 3, 371–405.
- SELKIRK E. (1995). Sentence prosody: intonation, stress and phrasing. In *The Handbook of Phonological Theory*, Blackwell., J. Goldsmith, Ed. London, 550–569.
- SELKIRK E. (1996). The prosodic structure of function words. In *Signal to syntax: Prosodic bootstrapping from speech to grammar in early acquisition*, J-L. Morgan and K. Demuth, Eds. Mahwah, NJ: Lawrence Erlbaum, pp. 187-214.
- SELKIRK E., TATEISHI K. (1988). Constraints on Minor Phrase formation in Japanese. *Proceedings of Chicago Linguistic Society*, 24.
- VAISSIÈRE J. (1991). Rhythm, accentuation and final lengthening in French. In *Music, Language, Speech and Brain*, Macmillan Press., J. Sundberg, L. Nord, and R. Carlson, Eds. London, 108–120.



## Quels liens entre accentuation et niveaux de constituance en français ? Une analyse perceptive et acoustique

L'accent en français est considéré comme un accent post-lexical marquant le niveau du groupe de mots plutôt que le niveau du mot lui-même. L'Accent Final (AF) primaire, cooccurrent à la frontière prosodique, s'effacerait perceptivement en frontière prosodique majeure. L'Accent Initial (AI), dit secondaire et optionnel, serait un accent rythmique apparaissant sur les longs constituants. Dans ce contexte, seuls deux niveaux de constituance sont communément admis en français : le Syntagme Intonatif (IP) et le Syntagme Accentuel (AP). L'existence d'un Syntagme Intermédiaire (ip) est en revanche controversée. Enfin, la prise en compte du Mot Prosodique (PW) (*i.e.* mot lexical) comme unité de planification, ou de réalisation des règles accentuelles, en structure de surface ne semble pas envisagé. L'objectif de cette étude est d'explorer l'organisation du phrasé prosodique en français. Dans ce cadre, nous proposons une étude perceptive, via un corpus de parole contrôlée manipulant des structures syntaxiques ambiguës, où 80 participants ont effectué 3 tâches de perception : proéminence, frontière et groupement. Les événements prosodiques perçus ont ensuite été mis en relation avec leurs réalités acoustiques. Les résultats montrent que les auditeurs sont capables de percevoir des niveaux de granularité de frontières plus fins que ce que les descriptions traditionnelles du français prédisent. Par ailleurs, les mots lexicaux sont systématiquement réalisés par un marquage bipolaire (AI+AF) de même force métrique. AI joue également un rôle plus structurel que rythmique, en marquant la structure prosodique de manière plus privilégiée qu'AF. Enfin, AF ne s'efface pas perceptivement en frontière prosodique majeure, et garde au contraire une trace métrique au niveau du mot lexical, qui ne varie pas strictement en fonction du niveau de constituance.

### **MOTS-CLES :**

phrasé prosodique, perception, acoustique, proéminences, frontières, groupements.

In French, accentuation is said to be post-lexical, marking the phrase rather than the word. That is, the primary final accent (FA) is considered to be perceptively weakened when co-occurring with a major prosodic boundary, while the Initial Accent (IA), regarded as a secondary and optional accent, is thought to hold merely a rhythmic function in balancing longer constituents. Consequently, only two levels of prosodic constituency are accounted for in French: the Intonational Phrase (IP), and the Accentual Phrase (AP). The existence of a third level, the Intermediate Phrase (ip), while advanced by some authors, remains controversial. Moreover, the Prosodic Word (PW) (*i.e.* lexical word) as a phonological unit, or as the domain of accentual rules, is disregarded altogether. The aim of our study is to investigate the organization of prosodic phrasing in French. We propose a perception study on a corpus in which syntactically ambiguous structures were manipulated, and asked 80 participants to perform 3 distinct perception tasks: a prominence, boundary and grouping task. The perceived prosodic events were then related to their acoustic realization. Taken together, our results indicate that listeners are able to distinguish finer-grained grouping levels than those predicted in traditional French descriptions. Moreover, lexical words are systematically realized by an accentual bipolarization (IA+FA), with each accent carrying the same metrical weight. The function of IA is shown to be more one of structuration than rhythmic balancing, with IA even marking structure more readily than FA. Finally, our results indicate that FA is not perceptively weakened when co-occurring with major prosodic boundaries, but instead remains a metrical mark at the level of the lexical word, in a manner independent from the level of constituency.

### **KEYWORDS:**

Prosodic Phrasing, Perception, Acoustics, Prominence, Boundary, Grouping.